



TESTES ESTATÍSTICOS DE CLASSIFICAÇÃO DE DOMÍNIOS DE PROTEÍNAS EM FAMÍLIAS E CLÃS

Edgar Augusto Gonçalves Gurgel do Amaral

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Rubem Pinto Mondaini

Rio de Janeiro

Junho de 2018

TESTES ESTATÍSTICOS DE CLASSIFICAÇÃO DE DOMÍNIOS DE PROTEÍNAS
EM FAMÍLIAS E CLÃS

Edgar Augusto Gonçalves Gurgel do Amaral

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Rubem Pinto Mondaini, D.Sc.

Prof. Cláudio Luís de Amorim, Ph.D.

Prof. José Abdalla Helayel Neto, Ph.D.

Prof. Jules Ghislain Slama, D.Sc.

Prof. Marco Antônio von Kruger, Ph.D.

RIO DE JANEIRO, RJ - BRASIL
JUNHO DE 2018

Amaral, Edgar Augusto Gonçalves Gurgel do

Testes Estatísticos de Classificação de Domínios de Proteínas em Famílias e Clãs / Edgar Augusto Gonçalves Gurgel do Amaral – Rio de Janeiro: UFRJ/COPPE, 2018.

XII, 249 p.: il.; 29,7 cm.

Orientador: Rubem Pinto Mondaini

Tese (doutorado) – UFRJ/ COPPE/ Programa de Engenharia de Sistemas e Computação, 2018.

Referências Bibliográficas: p. 246-249.

1. Domínio de Proteínas. 2. Medidas Entropia. 3. Testes Estatísticos. I. Mondaini, Rubem Pinto. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

DEDICATÓRIA

Às mulheres da minha vida: minha esposa Katia e as minhas filhas Luana e Luciana.

AGRADECIMENTOS

A Deus.

A meus pais Jandrisson e Eliana, pela educação e pelo apoio durante toda a minha vida.

Ao Professores da Coppe pelos ensinamentos.

Aos Professores que participaram da banca de Doutorado.

Ao Professor José Abdalla Helayel Neto pelo apoio na finalização do trabalho.

Ao Professor Adilson Elias Xavier pelo incentivo recebido.

Aos meus colegas de doutorado, Sandro, Simão e Wellington pelo companheirismo.

À funcionária Josefina Solange Silva Santos pela atenção recebida durante toda esta jornada.

Aos Funcionários Adilson Jorge Barros Magalhães e Nerivaldo pelo apoio recebido.

Ao funcionário Francisco Antônio Coelho de Souza pela grande ajuda em ceder uma máquina para realizar os cálculos estatísticos para o meu trabalho.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

TESTES ESTATÍSTICOS DE CLASSIFICAÇÃO DE DOMÍNIOS DE PROTEÍNAS EM FAMÍLIAS E CLÃS

Edgar Augusto Gonçalves Gurgel do Amaral

Junho/2018

Orientador: Rubem Pinto Mondaini

Programa: Engenharia de Sistemas e Computação

As decisões de biólogos especialistas acerca da identificação de Famílias e Clãs de Domínios de Proteínas são confirmadas usualmente pela aplicação de algoritmos de Programação Dinâmica e por técnicas de Cadeias Ocultas de Markov na literatura de Biologia Computacional. Este trabalho é uma introdução a métodos estatísticos alternativos associados à distribuição de valores de funções (Medidas de Entropia) de variáveis aleatórias (probabilidades de ocorrência de aminoácidos nos domínios de proteínas). O teste de hipóteses de uma estatística ANOVA (one way) é então adequado para mostrar a eficiência desses métodos em uma nova confirmação da existência das Famílias e Clãs introduzidos previamente.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

STATISTICAL CLASSIFICATION OF PROTEIN DOMAINS IN FAMILIES AND
CLANS

Edgar Augusto Gonçalves Gurgel do Amaral

June/2018

Advisor: Rubem Pinto Mondaini

Department: Systems and Computing Engineering

The decisions of expert biologists on the recognition of Families and Clans of Protein Domains are usually confirmed by the application of algorithms of Dynamical Programming and techniques of Hidden Markov Chains in the literature of Computational Biology. The present work is an introduction to alternative statistical methods associated with the distribution of values of functions (Entropy Measures) of random variables (probabilities of occurrence of amino acids on the protein domains). The Hypothesis testing of ANOVA Statistics (one way) is then able to display the efficiency of these methods on a new confirmation of the existence of Families and Clans previously introduced.

SUMÁRIO

Capítulo I – Introdução.....	1
1.1 Motivação.....	1
1.2 Objetivos	2
1.3 Organização do Trabalho.....	3
Capítulo II – Classificação de Proteínas em Clãs e Famílias	5
2.1 Domínios de Proteínas.....	5
2.2 Probabilidades Simples e Compostas	8
2.3 Tratamentos prévios	12
Capítulo III – Construção do espaço de Amostra – “janelas m x n”.....	16
3.1 Estruturas de vetores e matrizes construídas das probabilidades.....	18
3.2 Primeiras noções de medidas de Entropias. Por que Entropia?.....	23
Capítulo IV – Entropias Sharma-Mittal e Jaccard	24
4.1 Entropias Havrda-Charvat, Rényi, Landsberg-Vedral e limite Gibbs-Shannon.....	25
4.2 Entropias Jaccard associadas ao item anterior	27
Capítulo V – Banco de Dados Pfam - Generalidades.....	30
5.1 Classificação dos dados da versão Pfam 27.0	30
5.2 Variações médias estatísticas com parâmetro das Entropias Havrda-Charvat	53
5.3 Histogramas correspondentes	62
Capítulo VI – Comparação entre Sistemas Computacionais, entre Sistemas Operacionais e entre Estruturas no desempenho dos cálculos.....	73
6.1 Sistemas computacionais e operacionais	73
6.1 Estruturas.....	75
Capítulo VII – Análise Estatística ANOVA	75
7.1 ANOVA para medidas de Entropia advindas de probabilidades simples	77
Figura 7.1 – Entropia de distribuição de aminoácidos em uma coluna.....	78
7.2 ANOVA para medidas de Entropia advindas de probabilidades compostas ...	80
7.3 ANOVA para medidas de Entropia Jaccard.....	82
7.4 ANOVA para medidas de Entropia - janelas 100 x 100	81
7.4.1 Teste de Homogeneidade de Variância.....	160
7.4.1.1 Levene Teste.....	160
7.4.1.2 Barlett Teste.....	162
7.4.2 Anova paramétrica.....	164
7.4.2.1 Teste Turkey HSD.....	165
7.4.3 ANOVA não paramétrica.....	166

7.4.3.1	Kruskal-Wallis rank sum test	166
7.4.3.2	Dunn's Kruskal-Wallis Multiple Comparisons.....	167
7.4.4	Resultados das aplicações dos Testes.....	168
Capítulo VIII	– Conclusões e Trabalhos Futuros	245
8.1	Introdução	245
Referências Bibliográficas	246

LISTA DE FIGURAS

Figura 2.1: Tela inicial do Banco de Dados Pfam.....	13
Figura 2.2: Tela do Banco de Dados Pfam – ftp.....	14
Figura 2.3: Sequência Fasta.....	15
Figura 3.1: Bloco m x n de aminoácidos.....	19
Figura 5.1: Valores médios das medidas de entropia Havrda-Charvat.....	55
Figura 5.2: Desvio padrão das medidas de entropia Havrda-Charvat.....	57
Figura 5.3: Valores médios da Entropia de Havrda-Charvat para um conjunto de 8 Clãs. Desvio padrão das médias para a Entropia de Havrda-Charvat de um conjunto de 8 Clãs.....	59
Figura 5.4: Valores médios da Entropia de Havrda-Charvat para um conjunto de 8 Pseudo-Clãs. Desvio padrão das médias para a Entropia de Havrda-Charvat de um conjunto de 8 Pseudo-Clãs	61
Figura 5.5: Histogramas da família PF00005, clã CL0023.....	64
Figura 5.6: Histogramas da família PF00135, clã CL0028.....	67
Figura 5.7: Histogramas da família PF06850, clã CL0028.....	70
Figura 7.1: Entropia de distribuição de aminoácidos em uma coluna.....	75
Figura 7.2: Entropia de distribuição de aminoácidos em um par de colunas.....	77
Figura 7.3 – Histogramas da família PF0005, clã CL0023 (s = 0.1, 0.3, 0.5, 0.7, 0.9 e 1.0) probabilidade simples.....	83
Figura 7.4 – Histogramas da família PF00135, clã CL0028 (s = 0.1, 0.3, 0.5, 0.7, 0.9 e 1.0) – probabilidade simples.....	85
Figura 7.5 – Histogramas da família PF06850, clã CL0028 (s = 0.1, 0.3, 0.5, 0.7, 0.9 e 1.0) – probabilidade simples.....	87
Figura 7.6 – Histogramas da família PF0005, clã CL0023 (s = 0.1, 0.3, 0.5, 0.7, 0.9 e 1.0) – probabilidade composta.....	89
Figura 7.7 – Histogramas da família PF00135, clã CL0028 (s = 0.1, 0.3, 0.5, 0.7, 0.9 e 1.0) – probabilidade composta.....	91
Figura 7.8 – Histogramas da família PF06850, clã CL0028 (s = 0.1, 0.3, 0.5, 0.7, 0.9 e 1.0) – probabilidade composta.....	93
Figura 7.9 – Boxplots para as famílias do clã CL0004 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	99
Figura 7.10 – Boxplots para as famílias do clã CL0013 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	105

Figura 7.11 – Boxplots para as famílias do clã CL0014 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	111
Figura 7.12 – Boxplots para as famílias do clã CL0015 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	117
Figura 7.13 – Boxplots para as famílias do clã CL0016 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	123
Figura 7.14 – Boxplots para as famílias do clã CL0020 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	129
Figura 7.15 – Boxplots para as famílias do clã CL0021 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	134
Figura 7.16 – Boxplots para as famílias do clã CL0023 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	142
Figura 7.17 – Boxplots para as famílias do clã CL0025 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	148
Figura 7.18 – Boxplots para as famílias do clã CL0026 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	154
Figura 7.19 – Boxplots para as famílias do clã CL0027 (s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3) - probabilidade composta.....	160
Figura 7.20 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade simples para 8 clãs, com s = 0.1, s = 0.5, s = 1.0 (Shannon) e s = 1.3.....	235
Figura 7.21 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade simples para 23 clãs, com s = 0.1, s = 0.5, s = 1.0 (Shannon) e s = 1.3.....	237
Figura 7.22 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade simples para 26 clãs, com s = 0.1, s = 0.5, s = 1.0 (Shannon) e s = 1.3.....	238
Figura 7.23 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade composta para 8 clãs, com s = 0.1, s = 0.5, s = 1.0 (Shannon) e s = 1.3.....	240
Figura 7.24 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade composta para 23 clãs, com s = 0.1, s = 0.5, s = 1.0 (Shannon) e s = 1.3.....	241
Figura 7.25 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade composta para 26 clãs, com s = 0.1, s = 0.5, s = 1.0 (Shannon) e s = 1.3.....	243

LISTA DE TABELAS

Tabela 2.1: Sequência fictícia de nucleotídeo.....	06
Tabela 2.2: Bases complementares DNA/RNA.....	08
Tabela 3.1: Quantidade de clãs e famílias analisadas no Pfam v. 27.0.....	17
Tabela 5.1: Evolução do Banco de Dados Pfam.....	31
Tabela 5.2: Clãs e Famílias extraídas do Pfam v. 27.0.....	33
Tabela 5.3: Valores médios e desvio padrão entropia Havrda-Charvat(Clãs).....	58
Tabela 5.4: Valores médios e desvio padrão entropia Havrda-Charvat(Pseudo-Clãs).....	60

Capítulo I – Introdução

1.1 Motivação

Com a descrição da estrutura do DNA por Watson e Crick (1953), grandes avanços têm sido presenciados na Biologia. Um grande avanço nas pesquisas genéticas foi obtido com o Projeto do Genoma Humano,

Uma nova etapa nas pesquisas genéticas iniciou-se com a identificação das informações presentes no genoma humano, a qual só foi possível com o auxílio da computação.

Essa união da computação com a da Biologia, deu origem ao surgimento da Biologia Molecular. Alguns autores organizam esta área em duas subáreas: a Bioinformática e a Biologia Computacional [LANCIA, 2004]. A primeira, analisa os problemas relacionados a produção de grandes massas de dados, aos seu tratamento e armazenamento. A segunda, discute a interpretação e a análise dos dados genômicos gerados.

O maior desafio atualmente na Biologia Molecular é analisar e entender a grande quantidade de dados de sequências e de estruturas de DNAs e de proteínas que são geradas pelos projetos atualmente em desenvolvimento na área.

Dada a necessidade de um melhor gerenciamento e armazenamento de domínios de proteínas, que até o momento eram armazenados em sistemas de arquivos, foram desenvolvidos sistemas especializados para manipular essas informações em conjunto com o uso de Banco de Dados para armazenamento e gerenciamento de proteínas, dentre outras informações.

Atualmente há dezenas de banco de dados biológicos disponíveis na Web e muitas interfaces para consultar os seus conjuntos de dados, como o GenBank mantido pela National Center for Biotechnology Information (NCBI), EMBL, PDB e

PFAM, dentre outros, em sua maioria disponíveis na internet para pesquisa, possibilitando o acesso de pesquisadores das mais diversas áreas de atuação, auxiliando portanto, no diagnóstico e tratamento de doenças, projetos de novas drogas farmacêuticas, prevenção de pragas, dentre outros.

1.2 Objetivos

O grande aumento na quantidade de dados disponíveis não requer apenas novas tecnologias para consultar e manipular os dados gerados pela Biologia, mas, o desenvolvimento de novos métodos e técnicas de análise que sejam eficientes e eficazes na manipulação de grandes massas de dados. Assim, o desafio para lidar com esse volume de dados não pode ser resolvido apenas com computadores mais rápidos e com maior capacidade, mas através do desenvolvimento de estruturas de dados e de algoritmos que explorem a complexa estrutura dos bancos de dados de proteínas disponíveis na internet. Bancos de Dados são utilizados para armazenar essas grandes massas de dados gerados, como sequências de proteínas.

e as classificações definidas para a definição de famílias, superfamílias e clãs.

Existem critérios, descritos na literatura, utilizados pelos Biólogos para classificar as sequências de aminoácidos em famílias e clãs, além da verificação de existência de novas proteínas e nas regras de construção de novas sequências, famílias e clãs. Atualmente, os Banco de Dados de proteínas são construídos e mantidos de acordo com essa classificação utilizada.

Este trabalho tem como objetivo analisar regiões de famílias extraídas do Banco de Dados PFAM, versão 27.0, com o propósito de discutir essa classificação do Banco de Dados de Proteínas. Cálculos de distribuições de probabilidade simples e conjunta, e funções de variáveis aleatórias denominadas de Medidas de Entropia, foram aplicados ao Banco de Dados PFAM. Existem medidas estatísticas utilizadas

na literatura que melhor descrevem o comportamento dos Bancos de Dados, como por exemplo, a de Jensen Shannon (TOPSE, 2007) e a de Jaccard que foram utilizadas neste trabalho. Assim, é possível melhorar a precisão das atuais famílias no banco de dados analisado, bem como a avaliação de novos agrupamentos no domínio de famílias.

Para o cálculo das medidas estatísticas, foram utilizados dois conjuntos de famílias de proteínas:

1. um conjunto de 68 clãs, com 1069 famílias, cada clã com um mínimo de 5 famílias, onde cada família foi selecionada de uma área disponível de 100 linhas x 200 colunas de aminoácidos, extraídas do Banco de Dados PFAM, versão 27.0;
2. um conjunto de 146 clãs com 2180 famílias, cada qual também obedecendo a restrição de ter pelo menos 5 famílias, selecionadas de uma área disponível de 100 linhas x 100 colunas de aminoácidos, também extraídas do Banco de Dados PFAM, versão 27.0.

1.3 Organização do Trabalho

Este trabalho está organizado em sete capítulos. Este primeiro capítulo, apresenta a motivação, os objetivos e a organização do trabalho. Os demais capítulos desta tese estão estruturados conforme a descrição abaixo.

O segundo capítulo faz uma revisão dos conceitos básicos sobre proteínas, probabilidades simples e compostas, e os tratamentos iniciais utilizados para a extração de famílias do banco de dados Pfam.

No terceiro capítulo, o espaço amostral adequado para uso na abordagem estatística deste trabalho é descrito em detalhes. Começa pela definição de probabilidades de ocorrência e as restrições impostas aos números de famílias

possíveis pela estrutura do espaço amostral. Além disso, introduz a fundamentação teórica sobre medidas de entropia.

O quarto capítulo, introduz o conjunto de Medidas de Entropias de Havrda-Charvat, Reny, Landsberg-Vedral, Shannon e Sharma-Mittal e a medida de Informação Mútua.

O quinto capítulo discute os conceitos de banco de dados de proteínas e apresenta os principais bancos atualmente utilizados. Além disso, introduz o banco de dados Pfam, do qual as famílias de proteínas foram extraídas, e os resultados estatísticos obtidos com a medida de entropia Havrda-Charvat.

O sexto capítulo descreve a não adequação do sistema computacional Maple para os cálculos estatísticos realizados.

O sétimo capítulo discute a análise estatística ANOVA e a aplica aos resultados obtidos no capítulo anterior, com o mesmo espaço amostral utilizado.

E no oitavo capítulo, são apresentadas as considerações finais deste trabalho, as suas contribuições e as propostas para futuros trabalhos.

Capítulo II – Classificação de Proteínas em Clãs e Famílias

2.1 Domínios de Proteínas

Desde a antiguidade, sabe-se que as características como cor dos olhos, cor dos cabelos, formato das orelhas, tamanho do nariz, são transmitidas de pais para filhos, de gerações para gerações, ou seja, são características hereditárias. Pensava-se, antigamente, que essas características eram de alguma forma mescladas de geração em geração.

Somente no século XIX é que o conhecimento humano sobre a hereditariedade passou por um grande avanço. Gregor Mendel foi o pesquisador responsável, que realizou uma série de experimentos com uma determinada espécie de ervilha, denominada de *Pisum sativum*. Essa planta produz frutos, as vagens, dentro dos quais existem sementes, os grãos de ervilha.

Os trabalhos de Mendel levaram à conclusão de que as características hereditárias dos pais não são simplesmente mescladas pela natureza para originar descendentes como filhos com características de seus ascendentes. A hereditariedade está relacionada a características fornecidas por ambos os pais a cada um de seus descendentes.

As características transferidas dos pais para os filhos são denominadas de genes. Os genes compõem o material genético existente nas células, relacionado às chamadas características hereditárias, ou genéticas. Os genes são responsáveis por fornecer as condições necessárias para o funcionamento do metabolismo celular e para o desenvolvimento das características de um organismo. O metabolismo celular é o conjunto de todos os processos que ocorrem em uma determinada célula. As informações genéticas estão registradas nos cromossomos, os quais são constituídos de estruturas de DNA. (ALBERTS *et al*, 2014).

Os ácidos nucleicos estão associados à transmissão de características de pais para filhos, e também comandam e gerenciam todas as atividades das células. Eles são encontrados em todos os seres vivos e são substâncias complexas, formados pela repetição de um conjunto de moléculas menores, denominados de nucleotídeos. (LEWIS, 2003).

Bases nitrogenadas são classificadas de acordo com a sua estrutura em púricas ou pirimídicas, da seguinte forma:

- Bases púricas – Adenina (A) e guanina (G);
- Bases pirimídicas – citosina (C), timina (T) e uracila (U).

São dois os tipos de ácidos nucleicos: ácido desoxirribonucleico (DNA) e o ácido ribonucleico (RNA). O DNA e o RNA podem ser diferenciados, em sua composição química, pelas bases nitrogenadas. Citosina, guanina e adenina ocorrem em todos os ácidos nucléicos. Enquanto que a timina ocorre no DNA e a uracila ocorre no RNA.

Tabela 2.1 – Bases complementares DNA/RNA

DNA	RNA
C	G
G	C
T	A
A	U

O modelo de DNA foi estabelecido pelos cientistas James Watson e Francis Crick, em 1953. Os dois cientistas perceberam que a largura da molécula do DNA é constante. Além disso, verificaram que numa molécula de DNA as proporções entre adenina e timina eram iguais, o mesmo ocorrendo nas proporções entre citosina e guanina. Esses estudos indicaram que as cadeias de DNA estão relacionadas de acordo com um determinado emparelhamento das bases nitrogenadas, com a

adenina complementando a timina e a guanina complementando a citosina (RASHIDI e BUEHLER, 2000).

O DNA comanda a síntese de proteínas nas células e é assim, portanto que atua na determinação das características de um indivíduo. Nas regiões do DNA que correspondem aos genes, as sequências de nucleotídeos funcionam como um código: determinadas sequências correspondem a um determinado aminoácido. Assim, as letras do DNA codificam uma sequência de aminoácidos, que constituirão uma proteína. (ALBERTS *et al*, 2014).

Proteínas são macromoléculas que têm funções específicas dentro de um organismo. Elas podem ter um caráter estrutural ou podem estar ligadas a determinadas atividades, como é caso dos anticorpos, de hormônios e de enzimas. Uma proteína é formada por unidades conhecidas por aminoácidos (TAMARIN, 1996).

Proteínas podem ser compostas de uma ou mais regiões funcionais, denominadas de domínios. A identificação de domínios que ocorrem em uma proteína pode fornecer informações importantes sobre suas funções. Por esta razão biólogos classificam domínios de proteínas em famílias e importam-se sobre a confiabilidade de sua classificação. (BUSCH *et al*, 2009).

As proteínas são formadas por uma sequência linear de pequenas moléculas denominadas de aminoácidos, derivados dos processos de transcrição e tradução do DNA. Segundo Hunter(1993), proteínas podem conter sequências de 4.500 aminoácidos, desta forma o espaço de possíveis proteínas é bastante significativo, cerca de 204.500. Os aminoácidos são representados por um conjunto de vinte letras, conforme Tabela 2.4, abaixo:

Tabela 2.2 –Aminoácidos representados por códigos (letras)

Nome	Abreviação	Código
Alanina	Ala	A
Cisteína	Cys	C
Ácido aspártico	Asp	D
Ácido glutâmico	Glu	E
Fenilalanina	Phe	F
Glicenina	Gly	G
Histidina	His	H
Isoleucina	Ile	I
Lisina	Lys	K
Leucina	Leu	L
Metionina	Met	M
Asparagina	Asn	N
Prolina	Pro	P
Glutamina	Gln	Q
Arginina	Arg	R
Serina	Ser	S
Treonina	Thr	T
Valina	Val	V
Triptofano	Trp	W
Tirosina	Tyr	Y

2.2 Probabilidades Simples e Compostas

Conhecer os fundamentos do cálculo de probabilidades é essencial para descrever grande parte dos fenômenos que ocorrem na natureza e para a aplicação da teoria da probabilidade às mais diversas áreas de conhecimento. Nos dias atuais, ele tem um amplo emprego na sociedade em geral, especialmente dando apoio a tomada de decisões que envolvem riscos. Cada área faz uso dela de forma distinta, como no governo, na saúde, educação, finanças, agricultura, dentre outras.

O termo aleatoriedade sugere uma imprevisibilidade. Um bom e simples exemplo pode ser o lançamento de uma moeda. Entretanto, outros exemplos aleatórios podem ser bem mais complexos e difíceis de prever, como por exemplo: os preços flutuantes do estoque, o preço das ações, dentre outros.

Assim, pode-se dizer que fenômenos cujos resultados em cada tentativa são imprevisíveis e distintos, que não seguem leis causais, são ditos como experimentos aleatórios ou probabilísticos. O modelo adequado ao estudo de regularidades quando um determinado experimento é repetido muitas vezes é o probabilístico. Ele é indicado ao uso em experimentos com muitos eventos aleatórios, com muitas incertezas quanto a determinada ocorrência de um fato ou a experimentos cujos resultados futuros são desconhecidos e imprevisíveis (GRAVETTER e WALLNAU, 2013).

Um experimento aleatório (E) é um fato ou fenômeno, cujos resultados possíveis só podem ser avaliados quando as ocorrências são coletivamente consideradas. O espaço amostral representa o conjunto de todos os possíveis resultados de um experimento. O espaço amostral é representado por S e o seu número de elementos por N. Talvez o que interesse em um experimento seja uma parte do espaço amostral, um subconjunto do espaço amostral do experimento, denominado de evento. Um evento é indicado por uma letra maiúscula do alfabeto e seus números de elementos por n. (MEYER, 1970).

A teoria das probabilidades objetiva mensurar as chances de ocorrência dos diversos resultados que um experimento pode apresentar. Dentre os métodos existentes, há o método clássico que considera que todos os eventos simples em um espaço amostral tem a mesma chance de ocorrer. (EWENS E GRANT, 2005).

Então a probabilidade de um evento A ($A \subset S$), o número real $P(A)$, é dada por:

$$P(A) = \frac{n(A)}{n(S)}$$

onde:

$n(A)$ é o número de elementos de A;

$n(S)$ é o número de elementos de S.

Onde $P(A)$ indica a chance de ocorrência do evento indicado por A. E $n(A)$ representa também o número de casos favoráveis e $n(S)$ o número de casos possíveis. Como A é um subconjunto de S, espaço amostral do experimento, então pode-se afirmar: $0 \leq n \leq N$

A probabilidade condicional consiste em associar a ocorrência de dado evento ao acontecimento de outro, por exemplo, o evento A acontece após o evento B ter ocorrido. A chance de um evento A ocorrer sob condições estabelecidas previamente pelo evento B é definido da seguinte forma:

$$p(A|B) = \frac{p(A, B)}{p(B)} \quad (2.1)$$

E ainda pode ser escrito, considerando a chance de um evento B ocorrer sob determinadas condições definidas previamente pelo evento A, da seguinte forma:

$$p(B|A) = \frac{p(B, A)}{p(A)} \quad (2.2)$$

Existem três usos importantes para a probabilidade condicional: a determinação da ocorrência simultânea de eventos, o cálculo da probabilidade total e o teorema de Bayes.

O teorema de Bayes calcula a probabilidade de um determinado evento ocorrer, considerando a conhecimento prévio de um outro evento.

Considerando a equação (2.1) da probabilidade condicional, resulta em:

$$p(A,B) = p(A|B) p(B)$$

E de acordo com a equação (2.2) da probabilidade condicional, resulta em:

$$p(B,A) = p(B|A) p(A)$$

Como,

$$p(A,B) = p(B,A)$$

então:

$$p(A|B) p(B) = p(B|A) p(A)$$

obtendo-se:

$$p(A|B) = \frac{p(B|A) p(A)}{p(B)}$$

Em muitos cenários, como na biologia e na física, duas quantidades podem variar probabilisticamente de tal forma que a distribuição de uma depende da outra. Neste caso, não é aceitável trabalhar com distribuições de probabilidade de uma única variável aleatória independentemente. Deve-se usar a distribuição de probabilidade conjunta para variáveis aleatórias contínuas, a qual considera como a distribuição de uma pode mudar quando o valor de outra variável mudar. A probabilidade conjunta pode ser definida como a probabilidade de um evento ocorrer em combinação com outro evento (GRAVETTER e WALLNAU, 2013).

As variáveis aleatórias são classificadas em duas categorias, que são (GAIL *et al*, 2005):

- variável aleatória discreta – que assume valores em um conjunto discreto de valores possíveis.
- variável aleatória contínua – as que assumem quaisquer valores em um intervalo contínuo de valores;

Para um intervalo discreto, considerando que X e Y são duas variáveis aleatórias, então para cada par de possíveis resultados $X = x$, $Y = y$, pode ser atribuída a probabilidade $P_{XY}(x,y)$.

A distribuição de probabilidade conjunta para as variáveis aleatórias discretas, é dada por:

$$f_{XY}(x,y) = p(X=x, Y=y), \text{ para todos os números reais } x \text{ e } y.$$

Então, a função $f_{XY}(x,y)$ é chamada de função de probabilidade conjunta de X e Y, se o seguinte item for satisfeito:

$$\sum_x \sum_y f_{xy}(x, y) = 1$$

2.3 Tratamentos prévios

A complexidade na Biologia evoluiu através da modificação e da recombinação de blocos de proteínas. Esses blocos são denominados de domínios e a identificação e caracterização de domínios e de novos domínios de proteínas, atualmente é um dos grandes desafios da ciência.

Agrupar domínios em famílias é muito importante por dois motivos: primeiro, leva ao reconhecimento de novos membros de uma forma mais aprimorada e melhora a eficiência contra falsos membros. As características essenciais

conservadas em uma família são normalmente expressas por Cadeias de Markov Ocultas ou ainda por padrões (HEGER e HOLM, 2003).

Neste trabalho a verificação da classificação das proteínas em famílias e clãs será realizada através da análise de domínios de proteínas (no formato de “janelas”), extraídas especificamente de um banco de dados biológico.

Essas janelas foram recortadas utilizando um programa desenvolvido (script), para leitura de um arquivo baixado do sitio do Pfam. O arquivo foi obtido utilizando o protocolo ftp. A Figura 2.1, abaixo, mostra a tela inicial com o acesso ao link do ftp do sitio (PFAM, 2018).

EMBL-EBI

HOME | SEARCH | BROWSE | FTP | HELP | ABOUT

Pfam
keyword search Go

Pfam 31.0 (March 2017, 16712 entries)

The Pfam database is a large collection of protein families, each represented by **multiple sequence alignments** and **hidden Markov models (HMMs)**. [More...](#)

QUICK LINKS YOU CAN FIND DATA IN PFAM IN VARIOUS WAYS...

- SEQUENCE SEARCH** Analyze your protein sequence for Pfam matches
- VIEW A PFAM ENTRY** View Pfam annotation and alignments
- VIEW A CLAN** See groups of related entries
- VIEW A SEQUENCE** Look at the domain organisation of a protein sequence
- VIEW A STRUCTURE** Find the domains on a PDB structure
- KEYWORD SEARCH** Query Pfam by keywords

JUMP TO

Enter any type of accession or ID to jump to the page for a Pfam entry or clan, UniProt sequence, PDB structure, etc.

Or view the [help](#) pages for more information

Recent Pfam [blog](#) posts Hide this

[Pfam 31.0 is released](#) (posted 8 March 2017)

Pfam 31.0 contains a total of 16712 families and 604 clans. Since the last release, we have built 415 new families, killed 9 families and created 11 new clans. We have also been working on expanding our clan classification; in Pfam 31.0, over 36% of Pfam entries are placed within a clan. The new "stuff" [...]


[Pfam train online](#) (posted 8 November 2016)

Figura 2.1 – Tela inicial do Banco de Dados Pfam

Entrando na opção do ftp, é possível navegar pelos diretórios de arquivos e encontrar o arquivo que contém as famílias de proteínas na sua forma bruta, antes

do tratamento que foi realizado pelo script desenvolvido para recortar as famílias de proteínas originais do banco de dados. A Figura 2.2, abaixo, mostra o arquivo Pfam-A.fasta.gz utilizado para o recorte da janela (PFAM, 2018).

Index of /pub/databases/Pfam/current_release/

 [parent directory]




























Name	Size	Date Modified
 Pfam-A.clans.tsv.gz	298 kB	2/23/17, 9:00:00 PM
 Pfam-A.dead.gz	21.2 kB	2/23/17, 9:00:00 PM
 Pfam-A.fasta.gz	2.0 GB	2/23/17, 9:00:00 PM
 Pfam-A.full.gz	5.6 GB	2/23/17, 9:00:00 PM
 Pfam-A.full.metagenomics.gz	742 MB	2/23/17, 9:00:00 PM
 Pfam-A.full.ncbi.gz	27.7 GB	2/23/17, 9:00:00 PM
 Pfam-A.full.uniprot.gz	9.1 GB	2/23/17, 9:00:00 PM
 Pfam-A.hmm.dat.gz	434 kB	2/23/17, 9:00:00 PM
 Pfam-A.hmm.gz	246 MB	2/23/17, 9:00:00 PM
 Pfam-A.regions.tsv.gz	1.1 GB	2/23/17, 9:00:00 PM
 Pfam-A.regions.uniprot.tsv.gz	2.8 GB	2/23/17, 9:00:00 PM
 Pfam-A.rp15.gz	828 MB	2/23/17, 9:00:00 PM
 Pfam-A.rp35.gz	2.2 GB	2/23/17, 9:00:00 PM
 Pfam-A.rp55.gz	3.8 GB	2/23/17, 9:00:00 PM
 Pfam-A.rp75.gz	5.4 GB	2/23/17, 9:00:00 PM
 Pfam-A.seed.gz	150 MB	2/23/17, 9:00:00 PM
 Pfam-C.gz	140 kB	2/23/17, 9:00:00 PM
 Pfam.version.gz	115 B	2/23/17, 9:00:00 PM
 active_site.dat.gz	12.2 kB	2/23/17, 9:00:00 PM
 database_files/		2/23/17, 9:00:00 PM
 diff.gz	210 kB	2/23/17, 9:00:00 PM
 md5_checksums	1.6 kB	2/23/17, 9:00:00 PM
 metaseq.gz	803 MB	2/23/17, 9:00:00 PM
 ncbi.gz	24.0 GB	2/23/17, 9:00:00 PM
 pdbmap.gz	2.5 MB	2/23/17, 9:00:00 PM
 pfamseq.gz	5.8 GB	2/23/17, 9:00:00 PM
 proteomes/		2/23/17, 9:00:00 PM

Figura 2.2 – Tela do Banco de Dados Pfam - ftp

É importante ressaltar que o arquivo Pfam.fasta.gz foi convertido previamente no arquivo Pfam-A.txt, com o objetivo de transformar um arquivo no formato Fasta em um arquivo no formato .txt. O formato Fasta é muito utilizado em Bioinformática para representar sequências de nucleotídeos ou de peptídeos, nas quais nucleotídeos ou aminoácidos são representados utilizando códigos de uma única letra. Este formato origina-se do pacote de software Fasta, utilizado para o alinhamento de sequências de proteínas e de DNA. Foi inicialmente descrito com

FastaP (Fasta Pearson), desenvolvido por David J. Lipman e William R. Pearson (LIPMAN e PEARSON, 1985).

Um arquivo do tipo Fasta inicia com o símbolo de maior (>) ou ainda, de forma menos frequente, com o símbolo de ponto e vírgula (;). As linhas que se iniciam com o ponto e vírgula (;) são consideradas somente como comentários. Seguindo a linha inicial, está descrita a sequência representada por um conjunto de códigos de letras. Comumente, a sequência é finalizada com o uso de asterisco (*). A Figura 2.3 abaixo, representa um arquivo com a sequência Fasta.

```
LCBO - Prolactin precursor - Bovine ; a sample sequence in FASTA
format
MDSKGSSQKGSRLLLLLLVSNLLLCQGVVSTPVCPNGPGNCQVSLRDLFDRAVMVSHYIHDLSS
EMFNEFDKRYAQKGFITMALNSCHTSSLPTPEDKEQAQQTTHHEVLMSLILGLLRSWNDPLYHL
VTEVRGMKGAPDAILSRAIEIEEENKRLLEGMEMIFGQVIPGAKETEPYPVWSGLPSLQTKDED
ARYSAFYNLLHCLRRDSSKIDTYLKLLNCRIIYNNNC* >MCHU - Calmodulin -
Human, rabbit, bovine, rat, and chicken
ADQLTEEQIAEFKEAFSLFDKDGDTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTID
FPEFLTMMARKMKDSDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREA
DIDGDQVNYEEFVQMMTAK* >gi|5524211|gb|AAD44166.1| cytochrome b
[Elephas maximus maximus]
LCLYTHIGRNIYYGSYLYSETWNTGIMLLLITMATAFMGYVLPWGQMSFWGATVITNLFSAIPYI
GTNLV
EWIWGGFVSVDKATLNRFFAFHFILPFTMVALAGVHLTFLHETGSNNPLGLTSDSDKIPFHPYYTI
KDFLG
LLILILLLLLLLALLSPDMLGDPDNHMPADPLNTPHLIKPEWYFLFAYAILRSVPNKLGGVLALFL
SIVIL
GLMPFLHTSKHRSMMLRPLSQALFWTLTMDLLTLTWIGSQPVEYPYTIIGQMASILYFSIILAFI
PIAGX IENY
```

Figura 2.3 – Sequência Fasta

Capítulo III – Construção do espaço de Amostra – “janelas m x n”

Com o uso da tecnologia para a análise de sequências genéticas, além de informações obtidas com o andamento de diversos projetos, uma grande quantidade de sequências de proteínas está sendo produzida e não se sabe a quantidade de dados que ainda serão produzidos. Para armazenar essa grande massa de dados é necessário o uso de bancos de dados.

Bancos de dados são utilizados como uma boa ferramenta para gerenciar melhor esses dados, além de permitir a otimização de consultas. Eles agregam vários arquivos em uma única base de dados, facilitando tarefas de uso (inclusão, extração, alteração, exclusão) e comparação dos dados armazenados.

Dentre esses banco de dados, o Pfam é largamente utilizado pelos biólogos como fonte de informações sobre proteínas, além de disponibilizar ferramentas de análise dessas proteínas. A população a ser considerada neste trabalho é representada pelo conjunto de proteínas que foram extraídas do Banco de Dados Pfam.

Foi utilizada a versão 27.0, datada de 22/03/2013, a qual possui 14.831 famílias de proteínas, sendo que desse total somente 4.563 famílias estão classificadas em 515 clãs. Alguns clãs, foram extintos, como é o caso do clã CL0002, o CL0008, o CL0017, dentre outros. Para uma melhor compreensão sobre os clãs que possuem famílias, eles foram listados e informada para cada um a quantidade de famílias existentes. No trabalho que foi realizado aqui, uma família é composta por um grande conjunto de proteínas disponibilizadas na forma de uma

matriz de n linhas e m colunas, em um arquivo .txt. Foram selecionadas duas regiões (janelas) de cada uma das famílias de domínios de proteínas, representadas por blocos de aminoácidos do banco de dados Pfam, com as seguintes dimensões:

1. 100 linhas x 200 colunas;
2. 100 linhas x 100 colunas.

Foram utilizados somente os clãs que possuem pelo menos 05 famílias com as restrições do tamanho da região, acima citadas. A restrição foi utilizada para que fosse possível utilizar somente os clãs que tivessem uma amostra representativa de famílias para realização de cálculos estatísticos.

Para a restrição de 100 linhas x 200 colunas, do total de 515 clãs encontrados no Pfam, foram selecionados 68 clãs com um total de 1.441 famílias que atendem à restrição definida de terem no mínimo 05 famílias.

Enquanto que para a restrição de 100 linhas x 100 linhas, do total de 515 clãs encontrados no Pfam, foram selecionados 190 clãs com um total de 2.180 famílias que atendem à restrição de apresentarem pelo menos 05 famílias.

A Tabela 3.1, abaixo, mostra as quantidades de clãs e de famílias que foram analisadas no banco de dados Pfam, versão 27.0:

Tabela 3.1 – Quantidade de clãs e famílias analisadas no Pfam 27.0

Total	sem restrição	com restrição (linhas x colunas)	
		100 x 200	100 x 100
clãs	515	68	190
famílias	4.563	1.069	2.180

Esses conjuntos de clãs foram utilizados para o cálculo de probabilidades simples e conjunta, além das medidas de entropia.

3.1 Estruturas de vetores e matrizes construídas das probabilidades

Observações e experiências necessitam de um planejamento e precisam ser analisadas através da aplicação de um método estatístico. Para um determinado conjunto de dados, experimentos aleatórios podem ser realizados um conjunto de vezes sob condições similares e vão apresentar resultados não previsíveis. Esses experimentos podem ser organizados considerando todo o conjunto de elementos que serão analisados, definido como população do experimento, ou ainda, considerando uma parte desse conjunto de elementos que será analisada, definida como amostra do experimento.

Conclusões obtidas a partir de dados empíricos são vistas como incertezas, as quais serão tratadas em termos de probabilidade. Estas são utilizadas para prever que tipo de conclusões são prováveis de serem obtidas com o uso de uma determinada amostra, estabelecendo uma relação direta entre população e amostra, de tal forma que conclusões possam ser obtidas com o uso de amostras e consideradas para a população (BATSCHELET, 1978).

Inicialmente cálculos de probabilidade apropriados necessitam ser realizados, para que em seguida operações estatísticas possam ser executadas, permitindo um melhor entendimento dos dados em tratamento.

Uma determinada proteína pode ser vista como uma sequência de aminoácidos ou ainda computacionalmente pode ser tratada como uma sequência de vinte letras .

Para organizar o espaço amostral para o teste de hipóteses, o domínio de uma proteína será representado por um bloco de aminoácidos, organizado em dois procedimentos básicos, da seguinte forma:

1. seleciona-se uma determinada família de proteínas do banco de dados Pfam com m domínios (linhas). Em cada domínio existem n_1, n_2, \dots, n_m aminoácidos. Então, considera-se um bloco de $m \times n$ aminoácidos, onde n é dado por:

$$n = \min(n_1, n_2, \dots, n_m)$$

2. o tamanho do bloco a ser “recortado” deve ser especificado previamente. Em seguida serão removidos todos os domínios de proteínas onde $n_j < n$ e também serão removidos $(n_j - n)$ aminoácidos nos domínios remanescentes.

O segundo procedimento é utilizado para organizar o espaço amostral de blocos ($m \times n$) de aminoácidos, conforme Figura 3.1, abaixo:

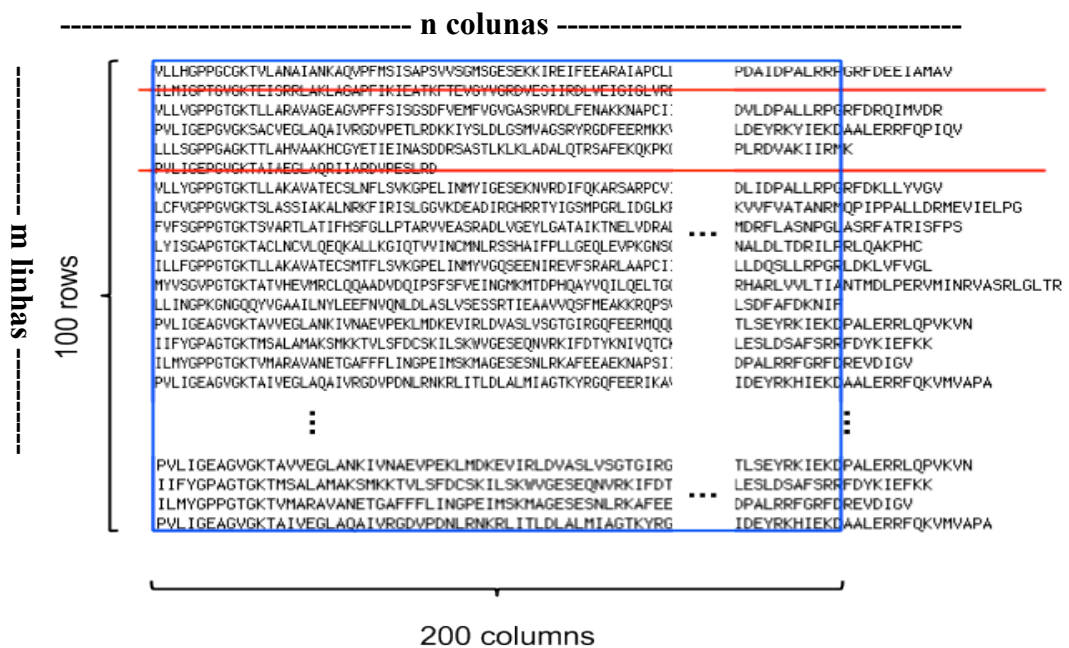


Figura 3.1 – Bloco $m \times n$ de aminoácidos

Para cada coluna de m linhas, um vetor p_j de 20 linhas é associado, conforme descrito abaixo:

$$p_j^T = (p_j(a_1), p_j(a_2), \dots, p_j(a_{20})),$$

$$j = 1, 2, \dots, n;$$

$a_1, a_2, \dots, a_{20} = A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y$; onde cada letra representa um aminoácido, e

$$p_j(a) = \frac{n_j(a)}{m}, \quad \dots \quad (3.1)$$

Onde $n_j(a)$ é o número de ocorrências do aminoácido a na j -ésima coluna.

A equação (3. 1) pode ainda ser interpretada como as componentes de n vetores de 20 componentes cada, da seguinte forma: (3.2)

$$\begin{bmatrix} p_1(A) \\ \vdots \\ p_1(Y) \end{bmatrix}, \quad \begin{bmatrix} p_j(A) \\ \vdots \\ p_j(Y) \end{bmatrix}, \quad \begin{bmatrix} p_n(A) \\ \vdots \\ p_n(Y) \end{bmatrix}$$

E assim, tem-se que:

$$\sum_a n_j(a) = m, \forall_j \rightarrow \sum_a p_j(a) = 1, \forall_j \quad (3.3)$$

De forma semelhante a equação (3.1), pode-se escrever a probabilidade conjunta de ocorrência de um par de aminoácidos **a**, **b** nas colunas **j**, **k**, respectivamente, da seguinte forma:

$$p_{jk}(a,b) = \frac{n_{jk}(a,b)}{m} \quad (3.4)$$

Onde $n_{jk}(a,b)$ é o número de ocorrências do par de aminoácidos **a**, **b** nas colunas **j**, **k**, respectivamente.

Uma interpretação conveniente dessa probabilidade conjunta podem ser os elementos de $\frac{n(n-1)}{2}$ matrizes quadradas de 20 x 20 elementos, que podem ser escritas da seguinte forma:

$$P_{jk} = \begin{pmatrix} P_{jk}(A,A) & \dots & P_{jk}(A,Y) \\ \vdots & \ddots & \vdots \\ P_{jk}(Y,A) & \dots & P_{jk}(Y,Y) \end{pmatrix} \quad (3.5)$$

Onde $j = 1, 2, \dots, (n-1)$;

$k = j + 1, \dots, n$;

$a, b = A, C, D, \dots, W, Y$.

Pode-se escrever também,

$$P_{jk}(a,b) = P_{jk}(a|b)p_k(b) \quad (3.6)$$

Esta equação também pode também ser entendida como outra definição de probabilidade conjunta. $P_{jk}(a|b)$ é a probabilidade condicional de ocorrência do aminoácido **a** na coluna **j**, se o aminoácido **b** já foi encontrado na coluna **k**.

Então, tem-se:

$$\sum_a P_{jk}(a|b) = 1 \quad (3.7)$$

Das equações (3.6) e (3.7), tem-se:

$$\sum_a P_{jk}(a, b) = p_k(b) \quad (3.8)$$

E a partir da equação (3.8),

$$\sum_a \sum_b P_{jk}(a, b) = 1 \quad (3.9)$$

o que é uma identidade, desde que $P_{jk}(a, b)$ também é uma probabilidade.

Equações (3.8) e (3.9) podem também serem derivadas de:

$$\begin{aligned} \sum_a n_{jk}(a, b) &= n_k(b) \\ \sum_a \sum_b n_{jk}(a, b) &= m \end{aligned} \quad (3.10)$$

e das definições das equações (3.1) e (3.4).

E do teorema de Bayes:

$$\begin{aligned} P(A|B) &= \frac{P(A, B)}{P(B)} \\ P_{jk}(a, b) &= P_{kj}(a|b) p_j(b) \\ P(A, B) &= P(B, A) \\ P_{jk}(a, b) &= P_{kj}(b, a) \\ P_{kj}(b, a) &= P_{kj}(b|a) p_j(a) \\ P_{jk}(a|b) P_k(b) &= P_{kj}(b|a) p_j(a) \end{aligned} \quad (3.11)$$

e da equação (3.11), a propriedade da simetria:

$$P_{jk}(a, b) = P_{kj}(b, a) \quad (3.12)$$

3.2 Primeiras noções de medidas de Entropias. Por que Entropia?

É importante estimar a diversidade de proteínas de uma determinada família ou de superfamílias, denominadas de clãs, para assegurar que a classificação dessas famílias foi realizada de forma condizente com as suas funções biológicas. Uma forma de medir a diversidade das famílias de proteínas com a mesma função é a de medir a desordem de um sistema.

A grandeza usada para medir esse tipo de desordem é a entropia, pois quanto maior a desordem, maior a entropia. Assim, uma maneira de expressar a segunda lei da termodinâmica é dizer que o sistema mudará espontaneamente para o estado de organização que tiver a maior entropia.(ALBERTS *et al*, 2010).

Portanto, a análise das distribuições de probabilidades dos blocos que representam as famílias de proteínas é feita através da utilização de medidas de entropia.

A primeira medida de entropia a ser apresentada no próximo capítulo é a Sharma-Mittal, pois as demais são obtidas a partir dela.

Capítulo IV – Entropias Sharma-Mittal e Jaccard

A informação que está armazenada em bancos de dados de Proteínas nas últimas décadas, está aguardando pela sua primeira descrição a ser feita por uma teoria de sucesso de enovelamento e desenovelamento de proteínas. Entende-se que este estudo deve ser iniciado a partir da evolução de famílias de proteínas e da sua associação em clãs, como um passo necessário ao seu desenvolvimento.

A primeira e fundamental ideia a ser desenvolvida é que as proteínas não evoluem independentemente. Foram apresentados vários argumentos e muitos cálculos foram realizados acerca dos fatos sobre a evolução de proteínas, para enfatizar a existência de um processo de formação de proteínas (Protein Family Formation Process – PFFP). Proteínas, ou suas partes inteligentes que são os domínios de proteínas, evoluem juntos como uma família de domínios de proteínas.

A descoberta dos estágios intermediários da dinâmica de enovelamento e desenovelamento da proteína, com o propósito de contornar o paradoxo de Levinthal (LEVINTHAL, 1968) (KARPLUS, 1997), bem como a determinação das condições iniciais, deve ser encontrada a partir de um estudo detalhado desse processo PFFP. Um subproduto dessa abordagem é a possibilidade de testar a hipótese de junção de famílias de proteínas em clãs utilizando métodos estatísticos rigorosos como ANOVA (DE GROOT e SCHERVISH, 2012).

Muitos exemplos de Medidas de Entropia foram utilizados nesta modelagem, como funções genéricas de variáveis aleatórias (probabilidades). Estas são as

probabilidades de ocorrência de aminoácidos em blocos retangulares, os quais são escolhidos de domínios de proteínas.

A medida de Entropia Sharma-Mittal possui dois parâmetros (r, s), sendo definida da seguinte forma (DE GROOT, 2012) :

$$(SM)_{jk}(r, s) = -\frac{1}{1-r} \left(1 - \left(\sum_a \sum_b (P_{jk}(a, b))^s \right)^{\frac{1-r}{1-s}} \right) \quad (4.1)$$

$$(SM)_j(r, s) = -\frac{1}{1-r} \left(1 - \left(\sum_a (p_j(a))^s \right)^{\frac{1-r}{1-s}} \right) \quad (4.2)$$

onde $p_j(a)$ e $P_{jk}(a,b)$ são as probabilidades simples e conjunta de ocorrência de aminoácidos como definido nas equações (3.1) e (3.4), respectivamente.

4.1 Entropias Havrda-Charvat, Rényi, Landsberg-Vedral e limite Gibbs-Shannon

Tomando-se o limite do parâmetro r tendendo a s , obtém-se a entropia Havrda-Charvat (HAVRDA e CHARVAT, 1967):

$$H_{jk}(s) = \lim_{r \rightarrow s} (SM)_{jk}(r, s) = -\frac{1}{1-s} \left(1 - \sum_a \sum_b (P_{jk}(a, b))^s \right) \quad (4.3)$$

$$H_j(s) = \lim_{r \rightarrow s} (SM)_j(r, s) = -\frac{1}{1-s} \left(1 - \sum_a (p_j(a))^s \right) \quad (4.4)$$

As medidas de entropia de Havrda-Charvat serão especialmente destacadas neste trabalho.

A entropia Rényi (RÉNYI, 1961) é obtida tomando o limite de r tendendo a 1:

$$R_{jk}(s) = \lim_{r \rightarrow 1} (SM)_{jk}(r, s) = \frac{1}{1-s} \log \left(\sum_a \sum_b (P_{jk}(a, b))^s \right) \quad (4.5)$$

$$R_j(s) = \lim_{r \rightarrow 1} (SM)_j(r, s) = \frac{1}{1-s} \log \left(\sum_a (P_j(a))^s \right) \quad (4.6)$$

A entropia Landsberg-Vedral (LANDSBERG e VEDRAL, 1997), uma normalização da Havrda-Charvat, é obtida da seguinte forma:

$$L_{jk}(s) = \lim_{r \rightarrow 2-1} (SM)_{jk}(r, s) = \frac{1}{1-s} \left(1 - \left(\sum_a \sum_b (P_{jk}(a, b))^s \right)^{-1} \right) = \frac{H_{jk}(s)}{\sum_a \sum_b (P_{jk}(a, b))^s} \quad (4.7)$$

$$L_j(s) = \lim_{r \rightarrow 2-1} (SM)_j(r, s) = \frac{1}{1-s} \left(1 - \left(\sum_a (p_j(a))^s \right)^{-1} \right) = \frac{H_j(s)}{\sum_a (p_j(a))^s} \quad (4.8)$$

A entropia Shannon (SHANNON, 1948) é obtida através do limite do parâmetro s tendendo a 1 de qualquer uma das três entropias (Havrda-Charvat, Landsberg-Vedral e Rényi):

$$\lim_{s \rightarrow 1} H_{jk}(s) = \lim_{s \rightarrow 1} R_{jk}(s) = \lim_{s \rightarrow 1} L_{jk}(s) = S_{jk} \quad (4.9)$$

$$\lim_{s \rightarrow 1} H_j(s) = \lim_{s \rightarrow 1} R_j(s) = \lim_{s \rightarrow 1} L_j(s) = S_j \quad (4.10)$$

onde,

$$S_{jk} = - \sum_a \sum_b P_{jk}(a,b) \log P_{jk}(a,b) \quad (4.11)$$

$$S_j = - \sum_a p_j(a) \log P_j(a) \quad (4.12)$$

4.2 Entropias Jaccard associadas ao item anterior

Existe um tipo de relação na ocorrência dos aminoácidos nos pares de colunas. Essa relação pode ser medida através do cálculo da Informação Mútua.

Para cada uma das entropias apresentadas previamente, tem-se a Informação Mútua da seguinte forma:

1. Sharma-Mittal

$$M_{jk}^{(SM)}(r,s) = \frac{1}{1-r} \left(1 - \frac{\sum_a \sum_b (p_{jk}(a,b))^s \frac{1-r}{1-s}}{\sum_a \sum_b (p_j(a)p_k(b))^s} \right) \quad (4.13)$$

2. Havrda-Charvat

$$M_{jk}^{(HC)}(s) = \lim_{r \rightarrow s} M_{jk}^{(SM)}(r,s) = \frac{1}{1-s} \left(1 - \frac{\sum_a \sum_b (p_{jk}(a,b))^s}{\sum_a \sum_b (p_j(a)p_k(b))^s} \right) \quad (4.14)$$

3. Landsberg-Vedral

$$M_{jk}^{(LV)}(s) = \lim_{r \rightarrow 2 \rightarrow s} M_{jk}^{(SM)}(r, s) = -\frac{1}{1-s} \left(1 - \frac{\sum_a \sum_b (p_j(a) p_k(b))^s}{\sum_a \sum_b (p_{jk}(a, b))^s} \right) \quad (4.15)$$

4. Rényi

$$M_{jk}^{(R)}(s) = \lim_{r \rightarrow 1} M_{jk}^{(SM)}(r, s) = -\frac{1}{1-s} \ln \left(1 - \frac{\sum_a \sum_b (p_{jk}(a, b))^s}{\sum_a \sum_b (p_j(a) p_k(b))^s} \right) \quad (4.16)$$

5. Shannon

$$\begin{aligned} M_{jk}^{(S)} &= \lim_{s \rightarrow 1} M_{jk}^{(HC)}(s) = \lim_{s \rightarrow 1} M_{jk}^{(LV)}(s) = \lim_{s \rightarrow 1} M_{jk}^{(R)}(s) \\ &= \sum_a \sum_b p_{jk}(a, b) \ln(p_{jk}(a, b)) - \sum_a \sum_b p_j(a) p_k(b) \ln(p_j(a) p_k(b)) \\ &= -S_{jk} + S_j + S_k \end{aligned} \quad (4.17)$$

O cálculo da distância da informação é dado pela diferença entre a medida de entropia e a informação mútua:

$$d_{jk}(r, s) = (SM)_{jk}(r, s) - M_{jk}^{(SM)}(r, s), \quad (4.18)$$

onde,

$$(SM)_{jk}(r, s) \geq 0;$$

$$M_{jk}^{(SM)}(r, s) \geq 0; \quad (4.19)$$

$$(SM)_{jk}(r, s) - M_{jk}^{(SM)}(r, s) \geq 0.$$

A entropia da Jaccard é obtida normalizando a distância da informação, obtendo-se:

$$J_{jk}(r,s) = 1 - \frac{M_{jk}^{(SM)}(r,s)}{(SM)_{jk}(r,s)}, \quad (4.20)$$

assim,

$$0 \leq J_{jk}(r,s) \leq 1.$$

A média da entropia de Jaccard para todos os pares de colunas é calculado da seguinte forma:

$$\begin{aligned} J(r,s) &= \frac{2}{n(n-1)} \sum_j \sum_k J_{jk}(r,s) & (4.21) \\ &= \frac{2}{200(200-1)} \sum_{j=1}^{199} \sum_{k=(j+1)}^{200} J_{jk}(r,s) \\ &= \frac{2}{19900} \sum_{j=1}^{199} \sum_{k=(j+1)}^{200} J_{jk}(r,s) \end{aligned}$$

Capítulo V – Banco de Dados Pfam - Generalidades

5.1 Classificação dos dados da versão Pfam 27.0

Devido ao rápido crescimento de dados à partir do Projeto Genoma e do alinhamento de proteínas, houve uma forte preocupação por parte da comunidade acadêmica em armazenar essas informações.

Banco de Dados Biológicos estão disponíveis na web para consulta de suas bases, de tal forma que pesquisadores de diferentes centros de pesquisas podem ter acesso a essas bases de dados. O fato de banco de dados estarem disponíveis é importante para a pesquisa, pois permite a integração de diversas Instituições, facilitando a troca de informações e a geração de novos conhecimentos.

Banco de Dados Biológicos podem ser divididos em bancos de sequências e bancos de estruturas. Os bancos de sequências são compostos de ácidos nucleicos e de proteínas enquanto que os bancos de dados de estruturas somente de proteínas.

O surgimento do primeiro banco de dados biológico ocorreu em 1956 com o sequenciamento da insulina. Na década de 60 outras sequências de proteínas foram estabelecidas. Em seguida, outros banco de dados surgiram, como o Protein Data Bank (PDB) em 1971 pelo Laboratório Nacional de Brookhaven. No início foram armazenadas sete estruturas de macromoléculas e nos anos seguintes mais estruturas foram depositadas. No início da década de 80 houve um grande crescimento da quantidade de estruturas armazenadas no banco de dados.

Outros banco de dados surgiram, dentre ele surgiu o Swiss-Prot em 1986 com aproximadamente 70.000 sequências de proteínas.

A primeira atividade para obtenção das sequências de aminoácidos consiste em definir um Banco de Dados de proteínas para a extração de dados.

Desta forma, informações sobre proteínas foram extraídas à partir do banco de dados Pfam, que possui alinhamentos de domínios de milhares de proteínas e perfis baseados no modelo *Hidden Markov*, denominado de *profile HMM*. A definição de limites de domínios, membros de famílias, bem como alinhamento é baseada na habilidade do *profile HMM* para identificar corretamente e alinhar membros.

O Banco de Dados Pfam armazena proteínas, famílias de proteínas (conjuntos de sequências de proteínas que possuem um determinado grau de similaridade de sequências) e superfamílias, também referenciadas como clãs.

Atualmente o Pfam está versão 31.0. A tabela 5.1, abaixo, descreve as versões do Pfam desde a 18.0, apresentando o ano de lançamento da versão do banco, a quantidade de famílias existentes, o número de famílias organizadas em clãs e a quantidade de clãs existentes em cada versão.

Tabela 5.1 – Evolução do Banco de Dados Pfam

versão	ano	quantidade de famílias	quantidade de famílias em clãs	quantidade de clãs
18.0	2005	7973	1181	172
19.0	2005	8183	1399	205
20.0	2006	8296	1560	239
21.0	2006	8957	1683	262
22.0	2007	9318	1815	283
23.0	2008	10340	2016	303
24.0	2009	11912	3132	423
25.0	2011	12273	3439	458
26.0	2011	13672	4243	499

27.0	2013	14831	4563	515
28.0	2015	16230	4939	541
29.0	2015	16295	5282	559
30.0	2016	16306	5423	595
31.0	2017	16712	5996	604

O Pfam, até a versão 27.0, era composto por dois tipos de famílias (PUNTA et al, 2012) :

- **Pfam-A**: composto de famílias de proteínas de alta qualidade, organizadas de forma manual e com um critério para inserção.
- **Pfam-B**: composto por famílias, automaticamente adicionadas.

Foi utilizada a versão 27.0 do banco de dados Pfam, com as famílias extraídas do arquivo Pfam-A. Os dados do arquivo Pfam-A podem ser baixados na sítio do Instituto de Bioinformática Europeu (European Bioinformatics Institute – EMBL-EBI) em <https://www.ebi.ac.uk/>, que disponibiliza publicamente dados biológicos para a comunidade científica através de um conjunto de serviços e ferramentas, executa pesquisa básica e provê treinamento profissional em bioinformática. Ele faz parte do Laboratório de Biologia Molecular Europeu (European Molecular Biology Laboratory - EMBL), uma organização de pesquisa internacional interdisciplinar fundada por vinte e dois estados membros e dois estados membros associados.

O EMBL-EBI disponibiliza uma grande quantidade de ferramentas e serviços, como por exemplo: consulta por semelhanças de sequências, visualização de sequencias de proteínas, consulta a famílias de proteínas, dentre outros.

Para baixar o arquivo Pfam-A.fasta.gz que contém os dados das famílias que serão utilizados neste trabalho, será utilizado um protocolo de rede padrão usado para a transferência de arquivos de um computador servidor para um computador

cliente denominado de FTP (file transfer protocol), disponível no seguinte link:
<ftp://ftp.ebi.ac.uk/pub/databases/Pfam/releases/Pfam27.0/>.

Após o arquivo Pfam-A.fasta.gz ser baixado, ele precisa ser convertido para o formato .txt para ser manipulado para a extração de famílias, gerando um arquivo Pfam-A.txt. Foi desenvolvido um script com comandos básicos do sistema operacional Linux necessários para gerar todos os arquivos e a estrutura de diretórios que serão utilizados como entrada, posteriormente, nos cálculos de entropia.

Para extrair um clã do arquivo Pfam-A.txt é necessário informar o nome do clã, o script aplicará uma restrição de que as famílias devam ter pelo menos uma família contendo um bloco de 100 linhas x 200 colunas. Desta forma somente foram gerados clãs contendo famílias na restrição acima descrita (100 linhas x 200 colunas).

A versão 27.0 do Pfam possui 14.831 famílias de proteínas, com 4.563 famílias classificadas em 515 clãs. Alguns clãs não existem na sequência de numeração, como o clã CL0002. Desta forma, os números dos clãs foram colocados de forma sequencial e foram informados os clãs inexistentes.

A tabela 5.2, abaixo, relaciona todos os clãs extraídos do arquivo Pfam-A.txt, informando a quantidade de famílias em cada clã e a quantidade de famílias que obedecem a restrição de terem famílias com blocos de 100 linhas x 200 colunas e blocos de 100 linhas x 100 colunas.

Tabela 5.2 – Clãs e Famílias extraídas do Pfam 27.0

número do Clã	quantidade de Famílias	quantidade de Famílias 100 linhas x 100 colunas	quantidade Famílias 100 linhas x 200 colunas
CL0001	13	0	0

CL0002	-	-	-
CL0003	5	0	0
CL0004	16	14	7
CL0005	2	0	0
CL0006	5	0	0
CL0007	5	0	0
CL0008	-	-	-
CL0009	3	3	1
CL0010	9	0	0
CL0011	24	3	0
CL0012	11	1	0
CL0013	7	6	6
CL0014	13	13	8
CL0015	25	22	21
CL0016	19	13	10
CL0017	-	-	-
CL0018	3	0	0
CL0019	-	-	-
CL0020	117	63	38
CL0021	45	15	1
CL0022	11	3	0
CL0023	198	174	119
CL0024	-	-	-
CL0025	9	6	0
CL0026	9	7	0
CL0027	8	7	7
CL0028	67	54	41
CL0029	53	40	13

CL0030	8	7	6
CL0031	9	9	3
CL0032	18	9	2
CL0033	5	3	0
CL0034	16	14	10
CL0035	12	10	9
CL0036	57	53	48
CL0037	12	7	2
CL0038	-	-	-
CL0039	26	26	20
CL0040	9	9	7
CL0041	5	0	0
CL0042	7	7	1
CL0043	3	3	2
CL0044	19	18	5
CL0045	4	1	0
CL0046	13	11	7
CL0047	2	1	0
CL0048	2	2	0
CL0049	14	4	0
CL0050	13	11	3
CL0051	24	16	1
CL0052	14	12	9
CL0053	22	4	1
CL0054	11	0	0
CL0055	17	4	2
CL0056	7	2	0
CL0057	21	0	0

CL0058	50	36	34
CL0059	28	25	19
CL0060	10	8	3
CL0061	15	14	14
CL0062	18	16	16
CL0063	180	156	92
CL0064	13	13	11
CL0065	5	6	1
CL0066	15	11	1
CL0067	4	2	1
CL0068	2	0	0
CL0069	2	1	1
CL0070	9	0	0
CL0071	2	2	2
CL0072	41	8	1
CL0073	7	5	1
CL0074	5	1	0
CL0075	7	0	0
CL0076	5	4	1
CL0077	2	2	1
CL0078	4	3	2
CL0079	9	3	0
CL0080	3	1	0
CL0081	6	0	0
CL0082	6	2	0
CL0083	19	0	0
CL0084	6	3	2
CL0085	7	5	3

CL0086	2	2	2
CL0087	4	2	1
CL0088	9	9	7
CL0089	9	4	0
CL0090	4	1	0
CL0091	2	1	1
CL0092	2	1	0
CL0093	9	7	6
CL0094	4	2	1
CL0095	3	2	1
CL0096	2	1	0
CL0097	3	2	0
CL0098	11	8	3
CL0099	4	3	3
CL0100	2	1	0
CL0101	5	3	0
CL0102	-	-	-
CL0103	11	9	9
CL0104	8	7	1
CL0105	17	12	5
CL0106	3	3	2
CL0107	4	2	0
CL0108	29	25	23
CL0109	16	7	2
CL0110	44	38	29
CL0111	21	19	16
CL0112	6	5	3
CL0113	35	32	20

CL0114	7	2	0
CL0115	6	5	3
CL0116	16	9	0
CL0117	5	0	0
CL0118	6	6	5
CL0119	-	-	-
CL0120	-	-	-
CL0121	4	2	0
CL0122	4	2	0
CL0123	192	42	6
CL0124	24	11	0
CL0125	60	32	15
CL0126	54	43	26
CL0127	10	10	5
CL0128	16	14	8
CL0129	14	9	1
CL0130	4	2	2
CL0131	9	7	2
CL0132	3	1	0
CL0133	3	2	1
CL0134	-	-	-
CL0135	6	5	3
CL0136	7	1	0
CL0137	21	19	11
CL0138	-	-	-
CL0139	2	2	1
CL0140	3	2	0
CL0141	5	2	0

CL0142	6	5	5
CL0143	5	5	4
CL0144	9	9	8
CL0145	9	7	4
CL0146	2	-	0
CL0147	3	2	0
CL0148	2	1	1
CL0149	7	7	7
CL0150	-	-	-
CL0151	10	9	9
CL0152	-	-	-
CL0153	7	2	0
CL0154	9	6	0
CL0155	2	1	0
CL0156	2	2	1
CL0157	2	2	1
CL0158	6	5	5
CL0159	63	30	3
CL0160	3	3	3
CL0161	10	9	1
CL0162	2	1	0
CL0163	8	7	7
CL0164	2	2	0
CL0165	6	3	1
CL0166	2	1	0
CL0167	50	2	1
CL0168	4	2	0
CL0169	9	6	4

CL0170	7	5	2
CL0171	3	3	3
CL0172	45	27	4
CL0173	4	3	0
CL0174	9	6	0
CL0175	11	0	0
CL0176	4	3	3
CL0177	23	19	17
CL0178	10	6	1
CL0179	20	17	11
CL0180	-	-	-
CL0181	10	10	8
CL0182	17	13	12
CL0183	13	11	0
CL0184	20	17	7
CL0185	-	-	-
CL0186	50	31	29
CL0187	3	1	0
CL0188	4	2	0
CL0189	2	2	1
CL0190	4	3	1
CL0191	4	2	1
CL0192	35	29	26
CL0193	54	40	33
CL0194	4	3	3
CL0195	1	1	0
CL0196	3	2	0
CL0197	4	3	4

CL0198	20	6	1
CL0199	5	3	0
CL0200	3	1	0
CL0201	4	2	0
CL0202	27	21	3
CL0203	5	3	0
CL0204	7	3	0
CL0205	2	2	2
CL0206	3	1	0
CL0207	3	2	0
CL0208	09	4	0
CL0209	14	10	3
CL0210	3	1	0
CL0211	-	-	-
CL0212	7	6	2
CL0213	2	-	0
CL0214	9	0	0
CL0215	-	-	-
CL0216	-	-	-
CL0217	2	1	0
CL0218	2	0	0
CL0219	46	33	16
CL0220	16	5	0
CL0221	15	5	0
CL0222	6	6	4
CL0223	3	2	0
CL0224	3	3	3
CL0225	2	0	0

CL0226	5	4	1
CL0227	3	1	0
CL0228	4	3	3
CL0229	24	1	1
CL0230	5	4	3
CL0231	5	2	0
CL0232	2	0	0
CL0233	2	0	0
CL0234	2	1	1
CL0235	2	2	1
CL0236	123	46	21
CL0237	10	8	4
CL0238	3	3	3
CL0239	2	0	0
CL0240	3	2	2
CL0241	4	3	2
CL0242	2	2	1
CL0243	4	2	1
CL0244	3	1	0
CL0245	4	4	3
CL0246	11	11	8
CL0247	9	4	2
CL0248	3	2	1
CL0249	2	1	1
CL0250	2	0	0
CL0251	2	1	0
CL0252	2	2	0
CL0253	-	-	-

CL0254	9	9	6
CL0255	12	6	1
CL0256	5	3	2
CL0257	31	21	5
CL0258	2	1	0
CL0259	3	2	0
CL0260	22	18	10
CL0261	5	4	1
CL0262	4	3	0
CL0263	19	7	2
CL0264	08	6	5
CL0265	5	5	2
CL0266	33	16	2
CL0267	3	1	0
CL0268	22	13	8
CL0269	2	1	1
CL0270	5	5	5
CL0271	4	2	0
CL0272	2	0	0
CL0273	2	1	2
CL0274	6	0	0
CL0275	2	1	0
CL0276	3	2	1
CL0277	6	4	4
CL0278	3	1	0
CL0279	2	1	0
CL0280	16	13	1
CL0281	2	1	0

CL0282	2	0	0
CL0283	2	2	2
CL0284	2	1	0
CL0285	2	2	2
CL0286	9	8	8
CL0287	13	4	0
CL0288	4	4	3
CL0289	3	2	1
CL0290	2	2	2
CL0291	8	6	0
CL0292	10	8	6
CL0293	2	2	1
CL0294	3	3	3
CL0295	16	14	9
CL0296	2	0	0
CL0297	2	1	1
CL0298	3	2	0
CL0299	2	0	0
CL0300	2	0	0
CL0301	2	1	0
CL0302	3	2	2
CL0303	3	2	1
CL0304	3	2	0
CL0305	2	0	0
CL0306	5	0	0
CL0307	9	9	4
CL0308	3	2	2
CL0309	-	-	-

CL0310	8	7	0
CL0311	3	-	0
CL0312	2	-	0
CL0313	-	-	-
CL0314	4	2	0
CL0315	12	11	2
CL0316	6	6	5
CL0317	11	6	1
CL0318	10	5	4
CL0319	7	4	0
CL0320	8	1	1
CL0321	2	1	0
CL0322	6	6	4
CL0323	3	3	3
CL0324	7	2	0
CL0325	2	2	2
CL0326	4	3	0
CL0327	10	3	0
CL0328	10	9	3
CL0329	14	11	2
CL0330	5	5	2
CL0331	8	6	1
CL0332	3	2	1
CL0333	5	0	0
CL0334	6	5	0
CL0335	5	3	0
CL0336	8	7	0
CL0337	2	0	0

CL0338	-	-	-
CL0339	6	5	4
CL0340	2	2	1
CL0341	2	1	1
CL0342	6	6	1
CL0343	4	3	1
CL0344	24	6	0
CL0345	2	0	0
CL0346	2	1	0
CL0347	3	2	1
CL0348	2	1	0
CL0349	4	2	1
CL0350	3	1	0
CL0351	6	0	0
CL0352	6	3	1
CL0353	2	0	0
CL0354	3	2	0
CL0355	4	2	0
CL0356	2	0	0
CL0357	3	1	0
CL0358	-	-	-
CL0359	3	1	0
CL0360	2	0	0
CL0361	19	2	0
CL0362	7	6	4
CL0363	7	4	2
CL0364	3	1	1
CL0365	2	0	0

CL0366	6	5	1
CL0367	2	1	0
CL0368	5	3	0
CL0369	8	1	0
CL0370	2	1	0
CL0371	3	1	0
CL0372	2	1	1
CL0373	6	5	5
CL0374	3	3	2
CL0375	10	5	2
CL0376	2	1	1
CL0377	2	2	1
CL0378	3	3	3
CL0379	2	2	1
CL0380	3	2	2
CL0381	7	5	4
CL0382	7	5	3
CL0383	2	0	0
CL0384	4	2	1
CL0385	2	2	0
CL0386	2	1	0
CL0387	2	2	1
CL0388	2	1	0
CL0389	3	1	0
CL0390	9	3	0
CL0391	2	0	0
CL0392	2	1	0
CL0393	2	2	0

CL0394	3	2	0
CL0395	3	1	2
CL0396	3	1	0
CL0397	2	0	0
CL0398	2	1	0
CL0399	3	2	1
CL0400	5	0	0
CL0401	6	5	4
CL0402	3	1	0
CL0403	2	2	2
CL0404	4	3	2
CL0405	3	1	0
CL0406	2	0	0
CL0407	3	1	0
CL0408	2	2	2
CL0409	3	3	1
CL0410	2	2	1
CL0411	2	2	2
CL0412	2	1	1
CL0413	7	6	0
CL0414	2	1	0
CL0415	-	-	-
CL0416	3	2	2
CL0417	3	1	0
CL0418	4	2	0
CL0419	2	0	0
CL0420	2	1	0
CL0421	4	3	0

CL0422	2	2	2
CL0423	2	2	1
CL0424	3	1	0
CL0425	2	1	1
CL0426	3	2	0
CL0427	-	-	-
CL0428	3	1	0
CL0429	2	0	0
CL0430	4	3	0
CL0431	3	1	0
CL0432	-	-	-
CL0433	3	3	2
CL0434	7	4	3
CL0435	2	2	2
CL0436	4	1	0
CL0437	4	3	0
CL0438	3	1	0
CL0439	2	1	0
CL0440	-	-	-
CL0441	2	0	0
CL0442	6	5	1
CL0443	-	-	-
CL0444	3	1	0
CL0445	4	1	0
CL0446	2	0	0
CL0447	2	0	0
CL0448	2	2	2
CL0449	2	1	0

CL0450	2	2	2
CL0451	3	2	0
CL0452	2	1	1
CL0453	2	2	2
CL0454	2	1	0
CL0455	2	0	0
CL0456	2	2	2
CL0457	3	2	0
CL0458	2	0	0
CL0459	2	1	0
CL0460	-	-	-
CL0461	5	0	0
CL0462	2	1	0
CL0463	-	-	-
CL0464	4	2	1
CL0465	7	1	1
CL0466	5	2	0
CL0467	-	-	-
CL0468	2	1	1
CL0469	5	2	0
CL0470	2	1	1
CL0471	2	2	2
CL0472	3	3	3
CL0473	-	-	-
CL0474	2	0	0
CL0475	6	4	3
CL0476	2	1	0
CL0477	2	1	0

CL0478	2	0	0
CL0479	8	6	3
CL0480	2	1	1
CL0481	2	0	0
CL0482	5	2	0
CL0483	2	1	0
CL0484	2	1	1
CL0485	-	-	-
CL0486	5	1	0
CL0487	5	3	0
CL0488	2	0	0
CL0489	2	1	1
CL0490	5	1	0
CL0491	3	1	0
CL0492	2	0	0
CL0493	2	2	2
CL0494	2	0	0
CL0495	-	-	-
CL0496	3	1	0
CL0497	5	3	0
CL0498	4	1	0
CL0499	2	1	1
CL0500	5	2	0
CL0501	3	0	0
CL0502	2	1	1
CL0503	2	0	0
CL0504	8	3	1
CL0505	3	1	0

CL0506	2	1	0
CL0507	2	2	2
CL0508	2	1	1
CL0509	2	1	1
CL0510	-	-	-
CL0511	6	0	0
CL0512	2	0	0
CL0513	2	1	0
CL0514	-	-	-
CL0515	2	0	0
CL0516	2	0	0
CL0517	2	1	1
CL0518	-	-	-
CL0519	-	-	-
CL0520	3	1	0
CL0521	2	1	0
CL0522	2	0	0
CL0523	7	6	0
CL0524	2	0	0
CL0525	3	1	1
CL0526	7	6	0
CL0527	3	1	0
CL0528	2	0	0
CL0529	2	1	1
CL0530	2	2	1
CL0531	2	1	0
CL0532	3	1	0
CL0533	6	5	2

CL0534	3	1	0
CL0535	2	0	0
CL0536	2	0	0
CL0537	2	0	0
CL0538	2	1	0
CL0539	3	1	0
CL0540	2	1	1
CL0541	2	1	1
CL0542	2	0	0
CL0543	2	0	0
CL0544	2	1	0
CL0545	2	1	0
CL0546	3	1	0
CL0547	2	0	0
CL0548	3	1	0
CL0549	6	6	4
CL0550	3	1	0
CL0551	3	2	1
CL0552	2	2	2
CL0553	2	2	2

5.2 Variações médias estatísticas com parâmetro das Entropias Havrda-Charvat

Aplicando a equação (4.3) da seção 4.1 para o cálculo de médias e de desvios padrões de medidas de entropias, considere um clã com F famílias. A medida de entropia Havrda-Charvat associada a um par de colunas em um *array* $m \times n$ de cada família e com um valor específico de parâmetro s é dado por:

$$H_{jk}(s;f) = -\frac{1}{1-s} \left(1 - \sum_a \sum_b (P_{jk}(a,b;f))^s\right) \quad (5.1)$$

Pode-se definir uma média dessas medidas de entropia para cada família, da seguinte forma:

$$\langle H(s;f) \rangle = \frac{2}{n(n-1)} \sum_j \sum_k H_{jk}(s;f) \quad (5.2)$$

E pode-se definir também o valor da média de todas as médias do conjunto de famílias F:

$$\langle H(s) \rangle_F = \frac{1}{F} \sum_{f=1}^F \langle H(s;f) \rangle \quad (5.3)$$

O desvio padrão das medidas de entropia $H_{jk}(s;f)$, relacionadas à média dada na equação (5.2), pode ser escrita da seguinte forma:

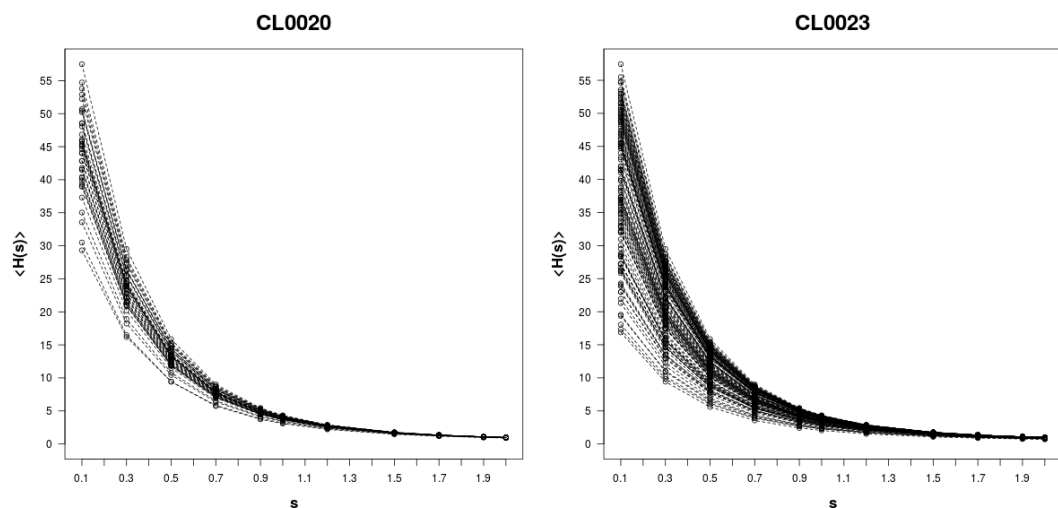
$$\sigma(s;f) = \left(\frac{1}{\frac{n(n-1)}{2} - 1} \sum_j \sum_k (H_{jk}(s;f) - \langle H(s;f) \rangle)^2 \right)^{\frac{1}{2}} \quad (5.4)$$

E finalmente, o desvio padrão da média $\langle H(s;f) \rangle$ relacionada à média $\langle H(s) \rangle_F$ é dada por:

$$\sigma_F(s) = \left(\frac{1}{(F-1)} \sum_{f=1}^F (\langle H(s; f) \rangle - \langle H(s) \rangle_F)^2 \right)^{\frac{1}{2}} \quad (5.5)$$

As figuras 5.1 e 5.2, apresentam, respectivamente, os valores da média e do desvio padrão das medidas de entropia de Havrda-Charvat para as famílias de um conjunto de oito clãs selecionados e para onze valores de parâmetro s , de acordo com as equações (5.2) e (5.4).

A região selecionada (janela) de cada família do domínio de proteína, corresponde a 100 linhas x 200 colunas, com a restrição de clãs selecionados com o mínimo de 5 famílias.



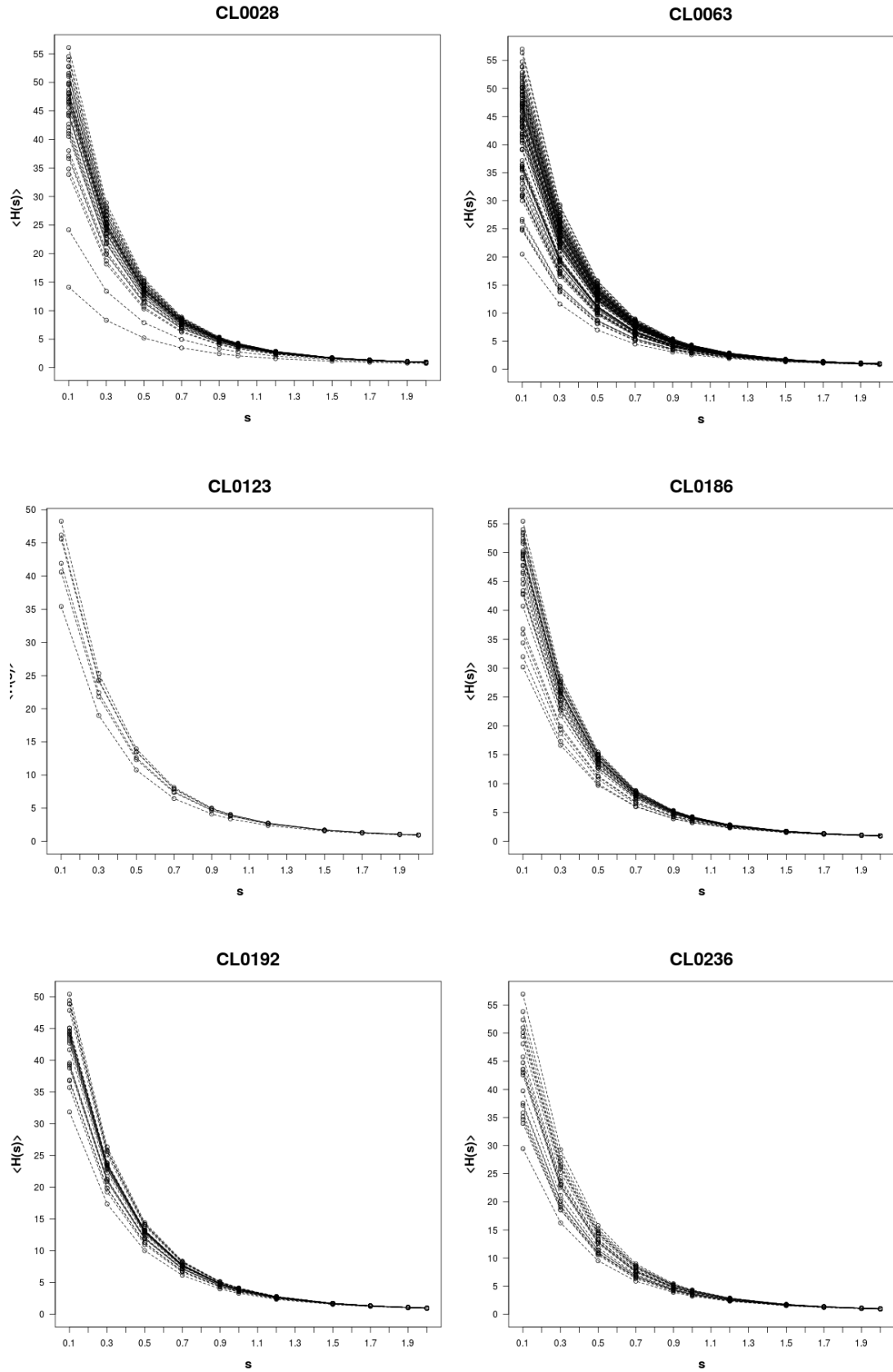
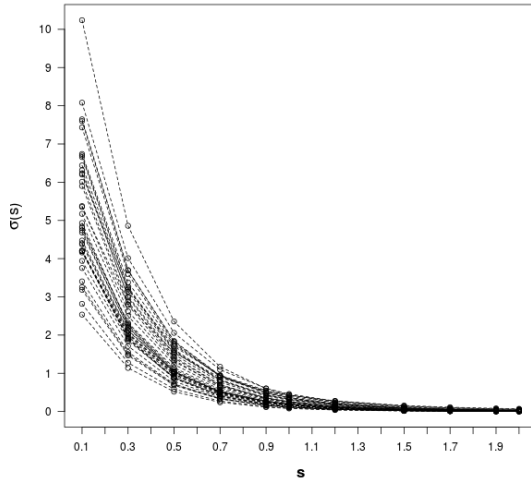
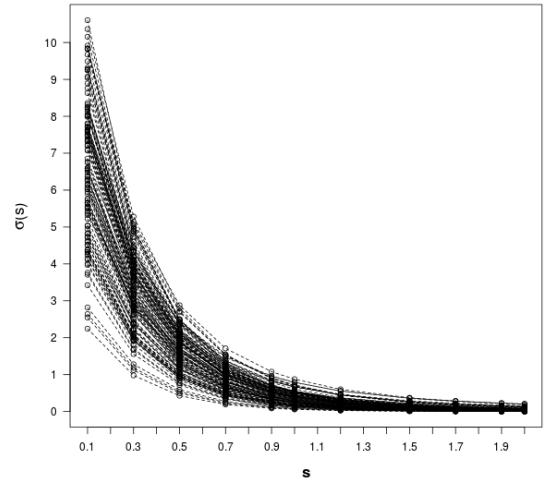


Figura 5.1 – Valores médios das medidas de entropia de Havrda-Charvat

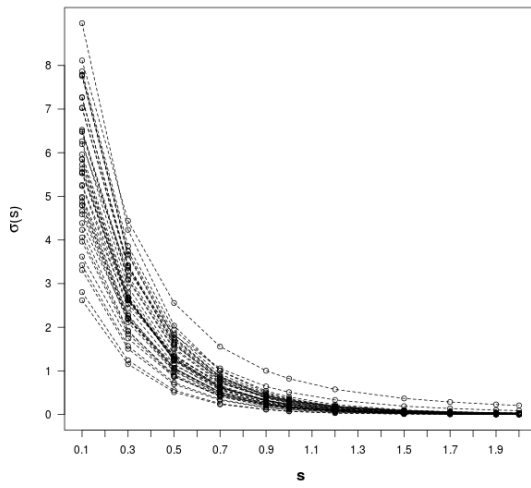
CL0020



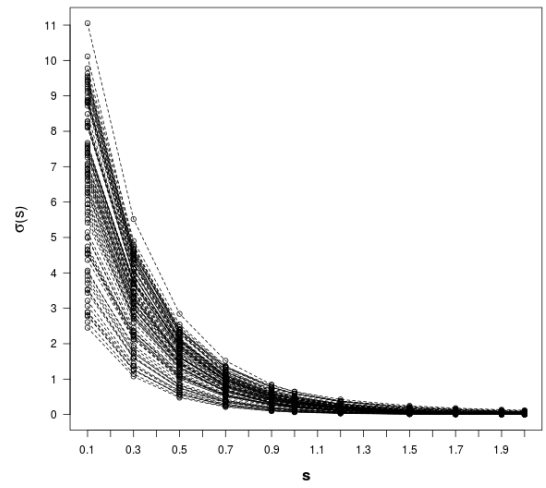
CL0023



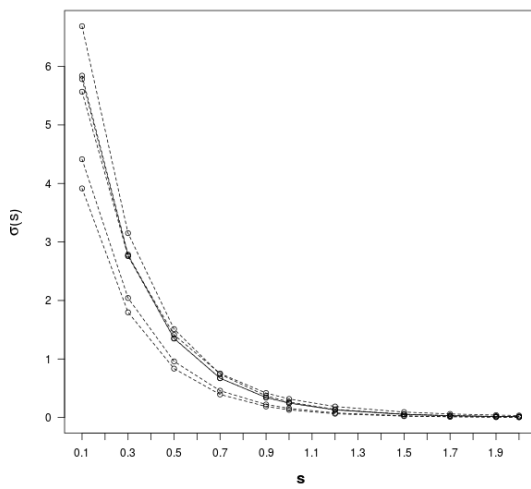
CL0028



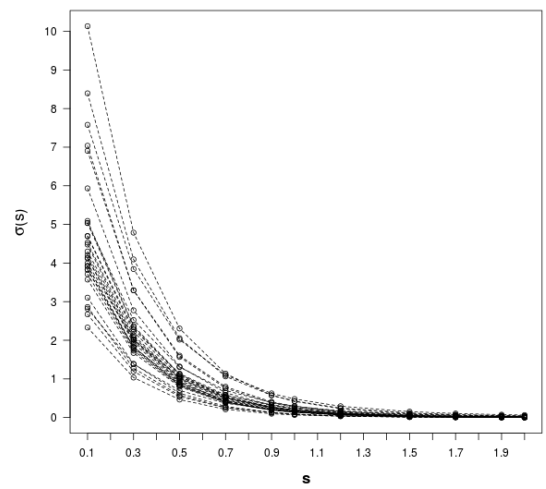
CL0063



CL0123



CL0186



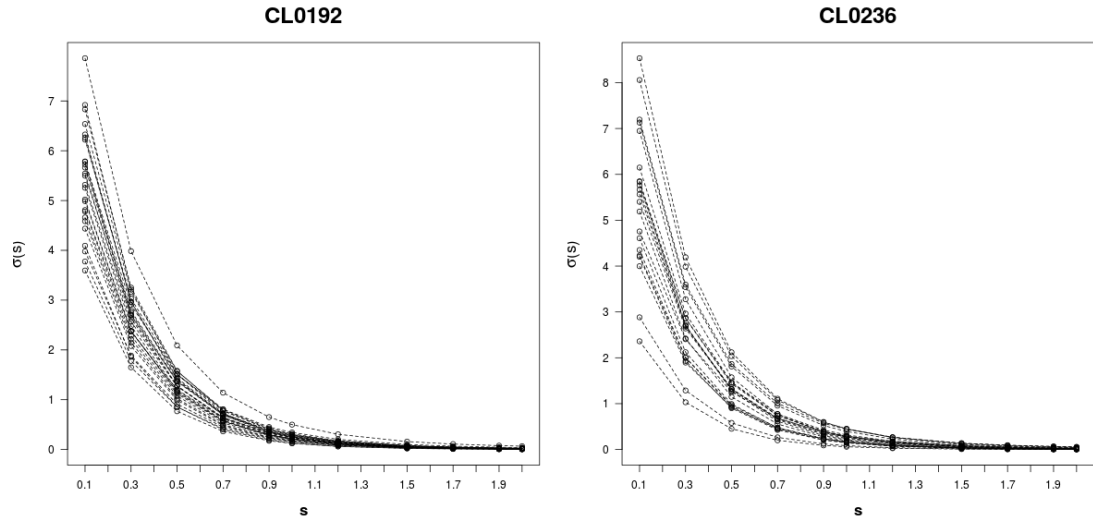


Figura 5.2 – Desvio padrão das medidas de entropia de Havrda-Charvat

A Tabela 5.3, abaixo, apresenta os valores de $\langle H(s) \rangle_F$ e de $\sigma(s)_F$ para um número selecionado de Clãs e onze valores de parâmetro s , de acordo com as equações (5.3) e (5.5).

Tabela 5.3 – Valores médios e desvio padrão Entropia Havrda-Charvat (Clãs)

Clãs do Pfam 27.0 - Entropia Havrda -Chavat							
Número Clã	s	$\langle H(s) \rangle_F$	$\sigma(s)_F$	Número Clã	s	$\langle H(s) \rangle_F$	$\sigma(s)_F$
CL0020 (38 famílias)	0.1	44.21	6.39	CL0123 (06 famílias)	0.1	43.00	4.66
	0.3	23.38	3.05		0.3	22.85	2.29
	0.5	12.99	1.49		0.5	12.77	1.16
	0.7	7.64	0.74		0.7	7.55	0.60
	0.9	4.78	0.38		0.9	4.74	0.32
	1.0	3.88	0.27		1.0	3.84	0.24
	1.2	2.66	0.15		1.2	2.65	0.13
	1.5	1.68	0.06		1.5	1.68	0.06
	1.7	1.31	0.03		1.7	1.31	0.03
	1.9	1.06	0.02		1.9	1.06	0.02
2.0	0.97	0.01	2.0	0.97	0.01		
CL0023 (119 famílias)	0.1	39.23	10.17	CL0186 (29 famílias)	0.1	46.08	6.79
	0.3	20.91	4.98		0.3	24.26	3.22
	0.5	11.73	2.51		0.5	13.42	1.56
	0.7	6.98	1.30		0.7	7.85	0.77
	0.9	4.42	0.70		0.9	4.89	0.39
	1.0	3.60	0.52		1.0	3.95	0.28
	1.2	2.50	0.30		1.2	2.70	0.14
	1.5	1.61	0.14		1.5	1.70	0.06
	1.7	1.26	0.09		1.7	1.32	0.03
	1.9	1.03	0.06		1.9	1.07	0.02
2.0	0.94	0.05	2.0	0.97	0.01		
CL0028 (41 famílias)	0.1	44.91	7.99	CL0192 (26 famílias)	0.1	42.86	4.56
	0.3	23.67	3.90		0.3	22.79	2.19
	0.5	13.11	1.96		0.5	12.74	1.07
	0.7	7.69	1.01		0.7	7.54	0.53
	0.9	4.80	0.54		0.9	4.74	0.27
	1.0	3.88	0.40		1.0	3.85	0.19
	1.2	2.66	0.23		1.2	2.65	0.10
	1.5	1.68	0.10		1.5	1.68	0.04
	1.7	1.31	0.07		1.7	1.31	0.02
	1.9	1.06	0.04		1.9	1.06	0.01
2.0	0.96	0.03	2.0	0.97	0.01		
CL0063 (92 famílias)	0.1	42.31	8.02	CL0236 (21 famílias)	0.1	43.25	7.46
	0.3	22.45	3.85		0.3	22.91	3.56
	0.5	12.53	1.89		0.5	12.76	1.73
	0.7	7.41	0.95		0.7	7.52	0.86
	0.9	4.66	0.49		0.9	4.72	0.44
	1.0	3.78	0.36		1.0	3.83	0.31
	1.2	2.61	0.19		1.2	2.63	0.16
	1.5	1.66	0.08		1.5	1.67	0.06
	1.7	1.30	0.05		1.7	1.30	0.04
	1.9	1.05	0.03		1.9	1.06	0.02
2.0	0.96	0.02	2.0	0.96	0.01		

A Figura 5.3, abaixo, apresenta os gráficos correspondentes a Tabela 5.3.

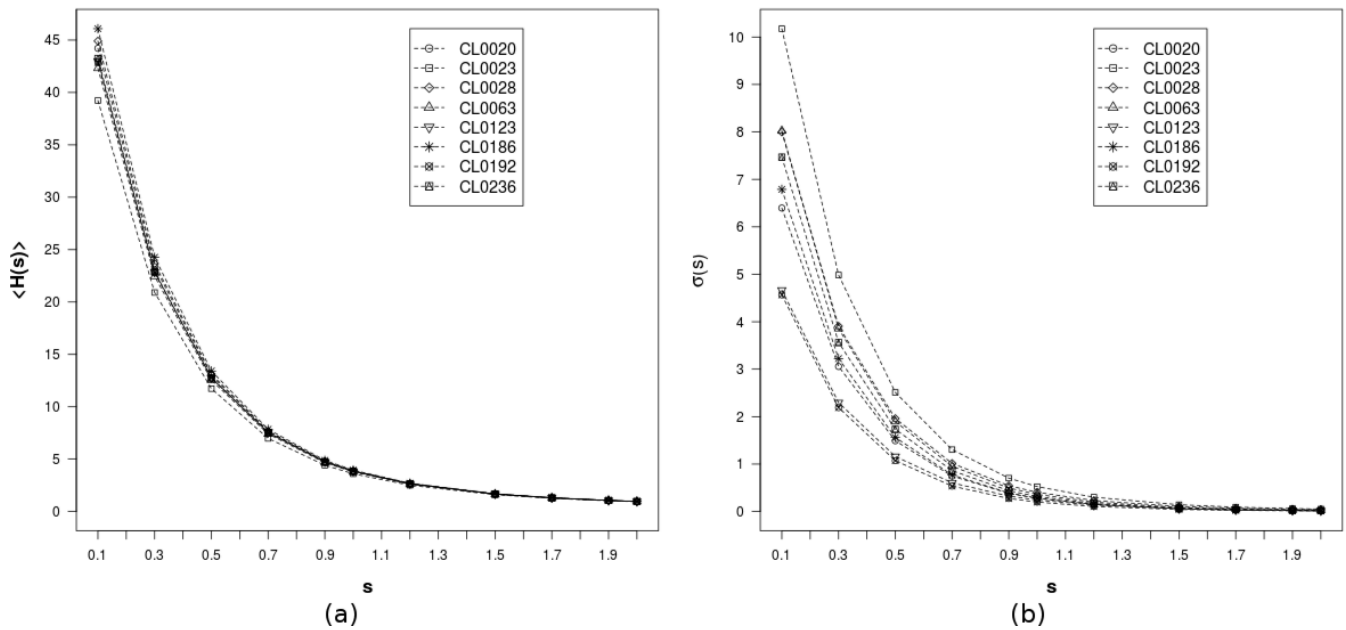


Figura 5.3 – (a) Valores médios da Entropia de Havrda-Charvat para um conjunto de 8 Clãs. (b) Desvio padrão das médias para a Entropia de Havrda-Charvat de um conjunto de 8 Clãs.

Para testar a robustez do conceito de Clã, foi criado o conceito de Pseudo-Clãs. A proposta foi verificar se os Pseudo-Clãs, com o mesmo número de famílias (com um mínimo de 05 famílias), dos Clãs correspondentes teriam diferentes valores de $\langle H(s) \rangle_F$ e $\sigma(s)_F$. As famílias associadas aos Pseudo-Clãs foram selecionadas do conjunto de 1069 famílias e pela retirada das famílias já escolhidas.

A Tabela 5.4 apresenta os valores de $\langle H(s) \rangle_F$ e $\sigma(s)_F$ para os Pseudo-Clãs obtidos conforme procedimento acima descrito.

A Figura 5.4 apresenta a comparação dos dados da Tabela 5.3 (Clãs) com os da Tabela 5.4 (Pseudo-Clãs). Os Clãs estão representados na cor vermelha, enquanto que os Pseudo-Clãs estão em azul.

A partir das Figuras (5.3 e 5.4) e das Tabelas (5.3 e 5.4), pode-se perceber que a região com $s \leq 1$ conduz a uma melhor caracterização das distribuições de medidas de entropia em bancos de dados de proteínas.

Tabela 5.4 – Valores médios e desvio padrão Entropia Havrda-Charvat (Pseudo-Clãs)

Pseudo-Clãs do Pfam 27.0 - Entropia Havrda -Chavat								
Número Clã	s	$\langle H(s) \rangle_F$	$\sigma(s)_F$		Número Clã	s	$\langle H(s) \rangle_F$	$\sigma(s)_F$
PCL0020 (38 famílias)	0.1	42.38	8.22		PCL0123 (06 famílias)	0.1	42.04	4.84
	0.3	22.48	3.97			0.3	22.35	2.17
	0.5	12.54	1.97			0.5	12.50	1.08
	0.7	7.41	1.00			0.7	7.40	0.55
	0.9	4.65	0.52			0.9	4.66	0.29
	1.0	3.78	0.38			1.0	3.79	0.21
	1.2	2.60	0.21			1.2	2.61	0.12
	1.5	1.65	0.09			1.5	1.66	0.05
	1.7	1.29	0.05			1.7	1.30	0.03
	1.9	1.05	0.03			1.9	1.06	0.02
2.0	0.95	0.03	2.0	0.97	0.01			
PCL0023 (119 famílias)	0.1	41.02	9.00		PCL0186 (29 famílias)	0.1	40.89	7.72
	0.3	21.82	4.36			0.3	21.79	3.76
	0.5	12.21	2.16			0.5	12.22	1.89
	0.7	7.24	1.10			0.7	7.25	0.98
	0.9	4.57	0.58			0.9	4.58	0.52
	1.0	3.71	0.42			1.0	3.73	0.39
	1.2	2.57	0.24			1.2	2.58	0.22
	1.5	1.64	0.10			1.5	1.64	0.10
	1.7	1.28	0.06			1.7	1.29	0.07
	1.9	1.04	0.04			1.9	1.04	0.04
2.0	0.95	0.03	2.0	0.95	0.04			
PCL0028 (41 famílias)	0.1	42.71	7.43		PCL0192 (26 famílias)	0.1	42.75	9.36
	0.3	23.65	3.57			0.3	22.67	4.53
	0.5	12.64	1.75			0.5	12.63	2.25
	0.7	7.46	0.88			0.7	7.45	1.51
	0.9	4.69	0.460			0.9	4.68	0.60
	1.0	3.80	0.33			1.0	3.79	0.44
	1.2	2.62	0.18			1.2	2.61	0.25
	1.5	1.66	0.07			1.5	1.65	0.11
	1.7	1.30	0.04			1.7	1.29	0.07
	1.9	1.05	0.03			1.9	1.05	0.04
2.0	0.96	0.02	2.0	0.95	0.03			
PCL0063 (92 famílias)	0.1	41.87	8.13		PCL0236 (21 famílias)	0.1	41.55	7.47
	0.3	22.25	3.91			0.3	22.09	3.59
	0.5	14.44	1.92			0.5	12.35	1.76
	0.7	7.36	0.97			0.7	7.32	0.88
	0.9	4.63	0.510			0.9	4.61	0.46
	1.0	3.71	0.37			1.0	3.75	0.33
	1.2	2.60	0.20			1.2	2.60	0.18
	1.5	1.65	0.09			1.5	1.65	0.08
	1.7	1.29	0.05			1.7	1.29	0.04
	1.9	1.05	0.03			1.9	1.05	0.03
2.0	0.95	0.03	2.0	0.96	0.02			

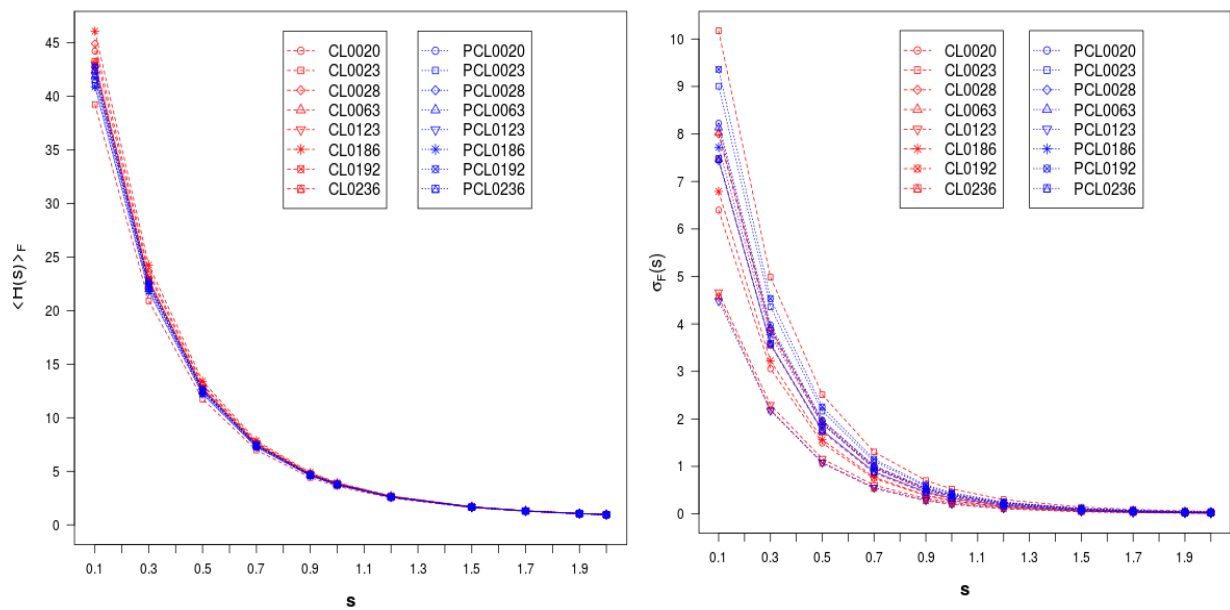
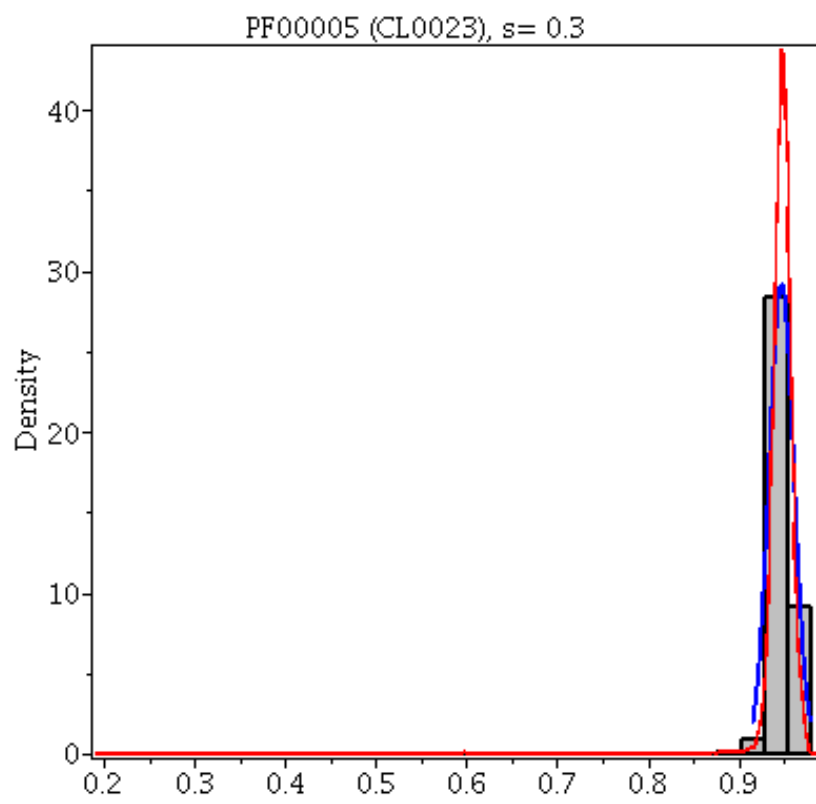
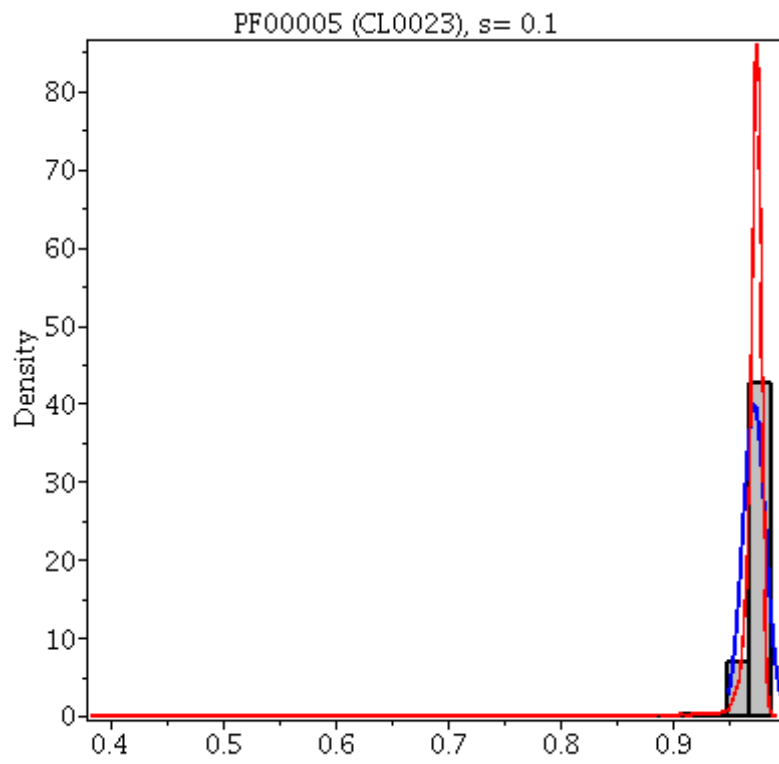
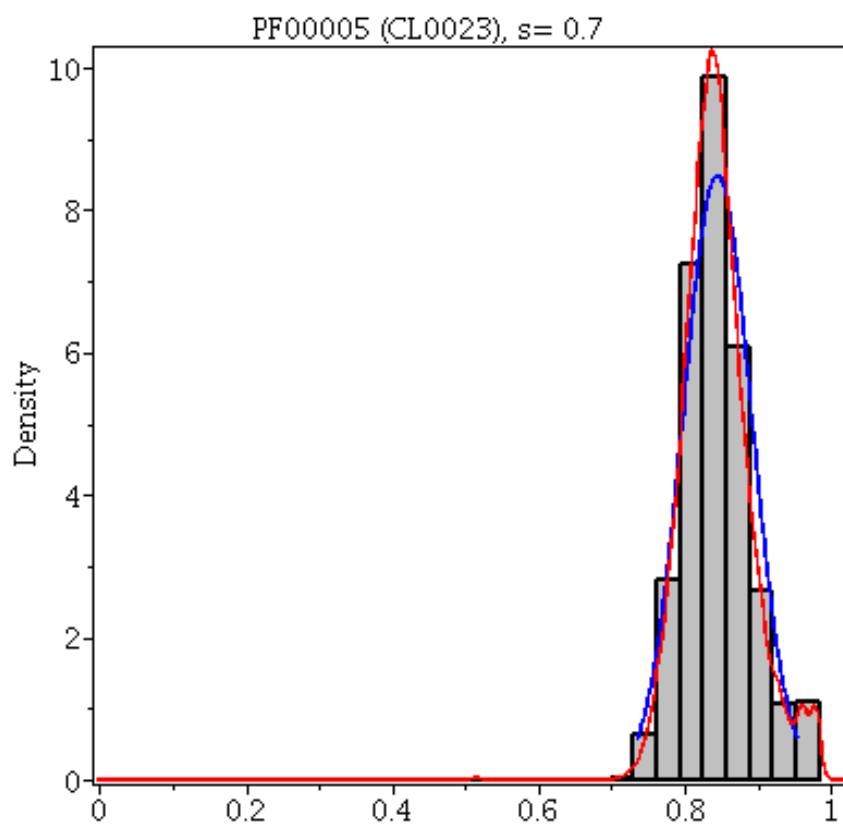
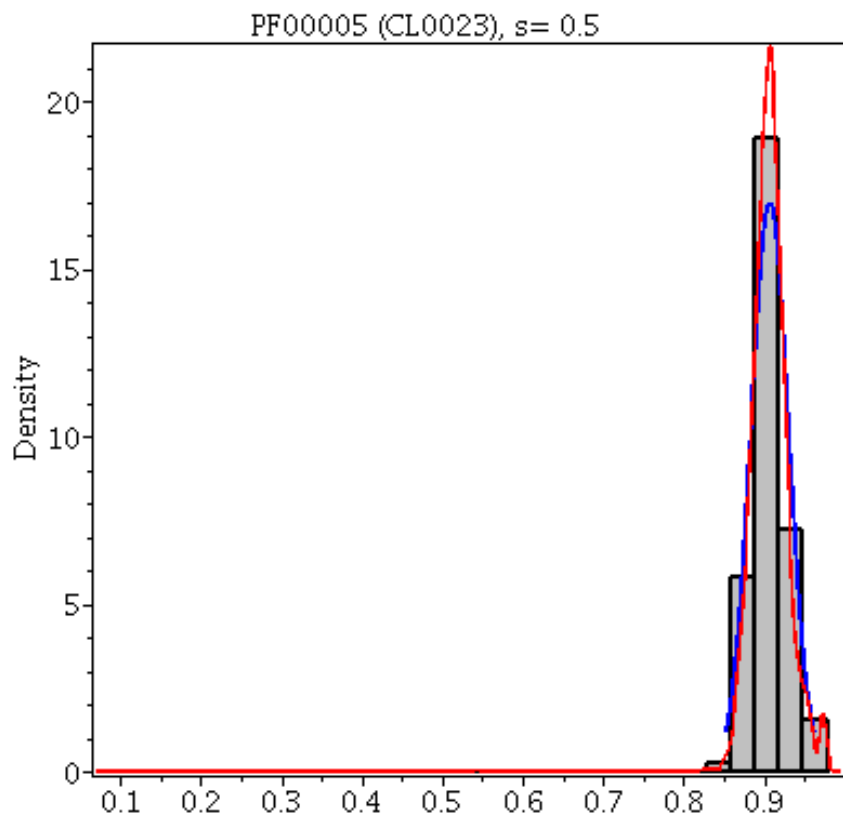


Figura 5.4 – (a) Valores médios da Entropia de Havrda-Charvat para um conjunto de 8 Pseudo-Clãs. (b) Desvio padrão das médias para a Entropia de Havrda-Charvat de um conjunto de 8 Pseudo-Clãs.

5.3 Histogramas correspondentes

Foram feitos os histogramas para três famílias (PF00005, PF00135, PF06850) de dois clãs (CL0023 e CL0028) e para diferentes valores do parâmetro s (0.1, 0.3, 0.5, 0.7, 0.9 e 1.0), selecionados a partir dos Clãs utilizados na Tabela 5.3.





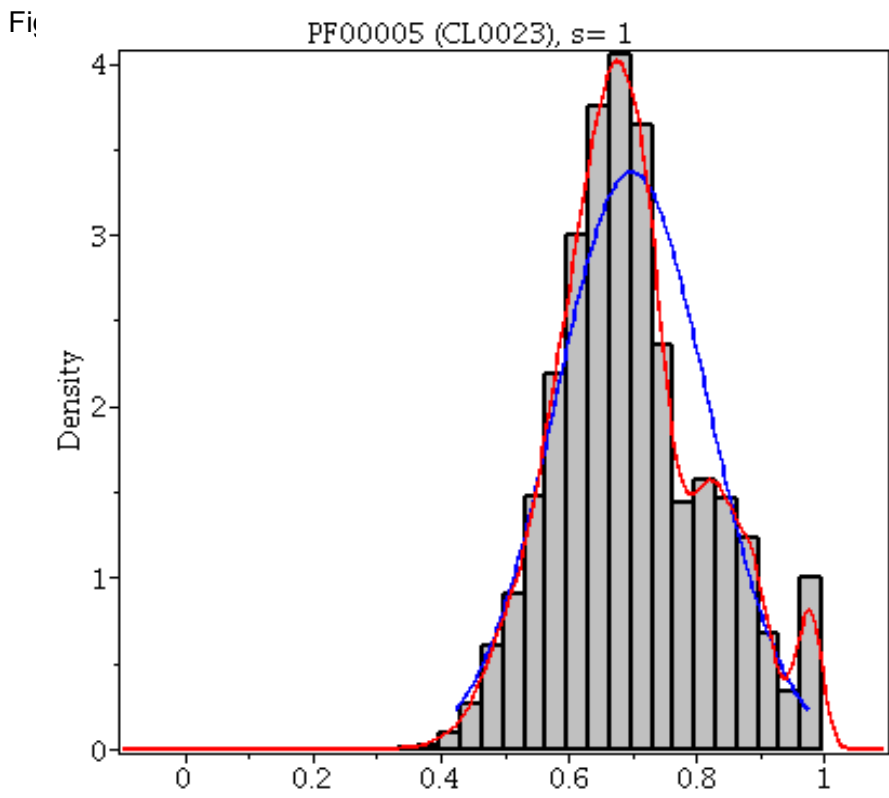
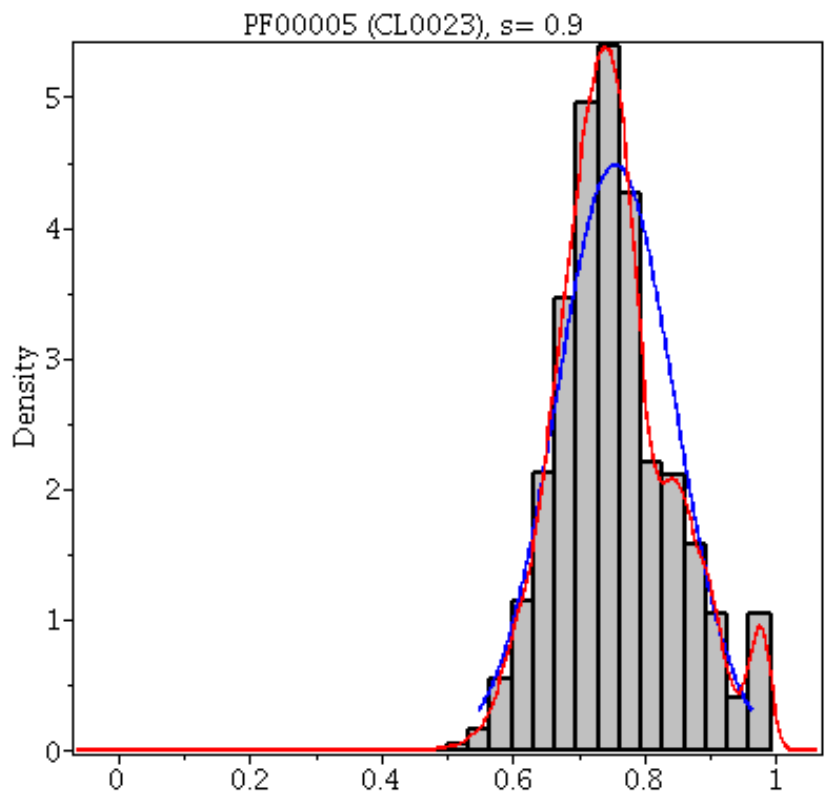
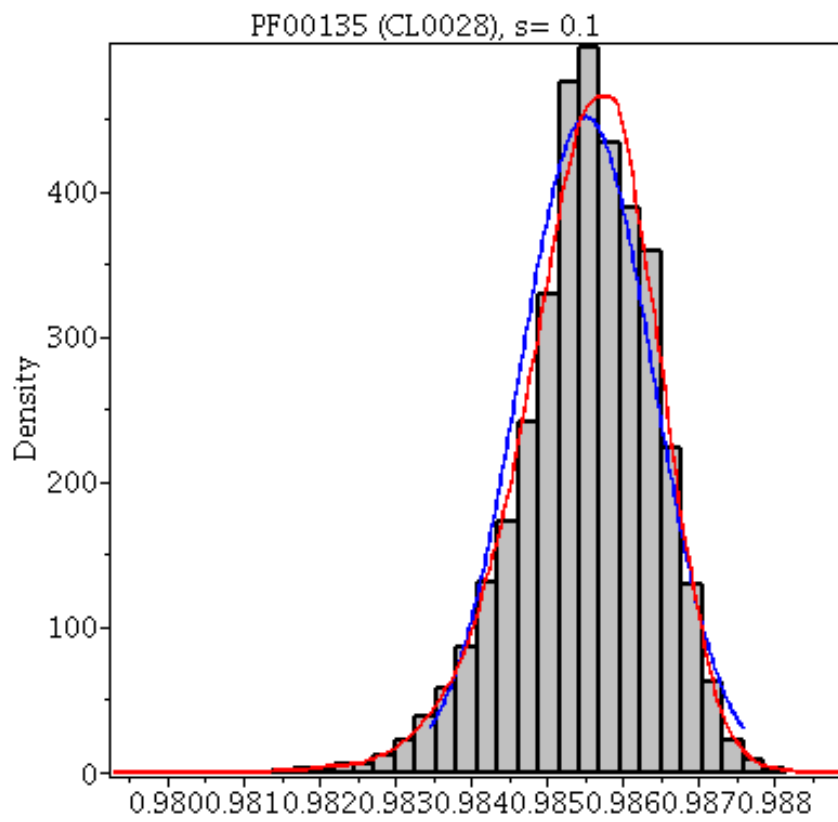
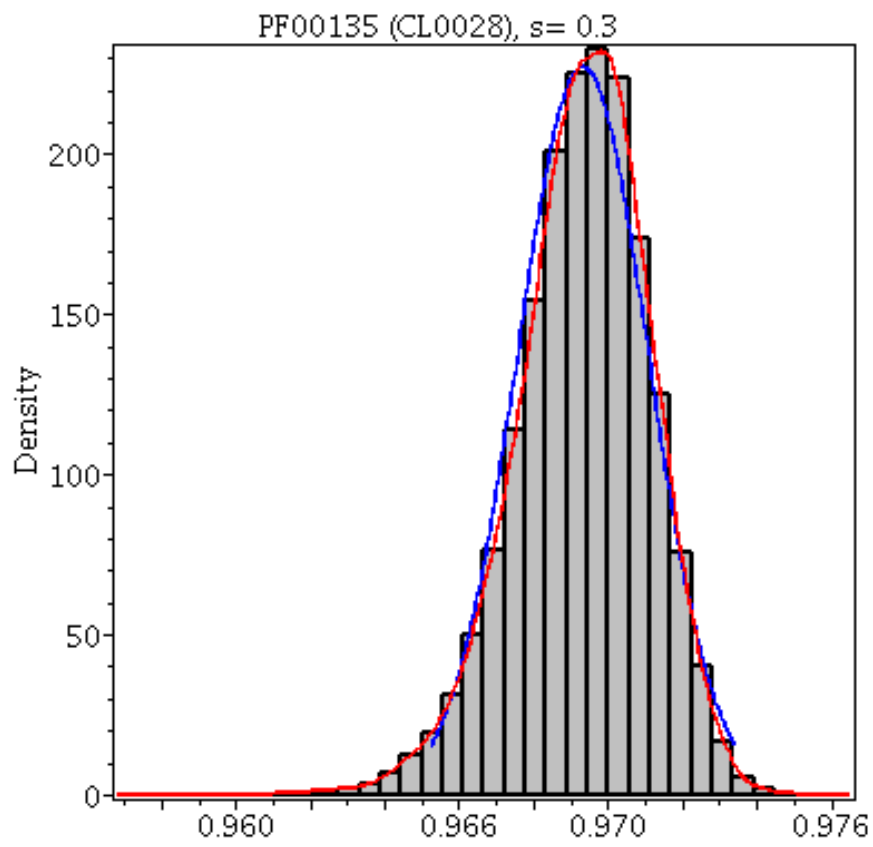
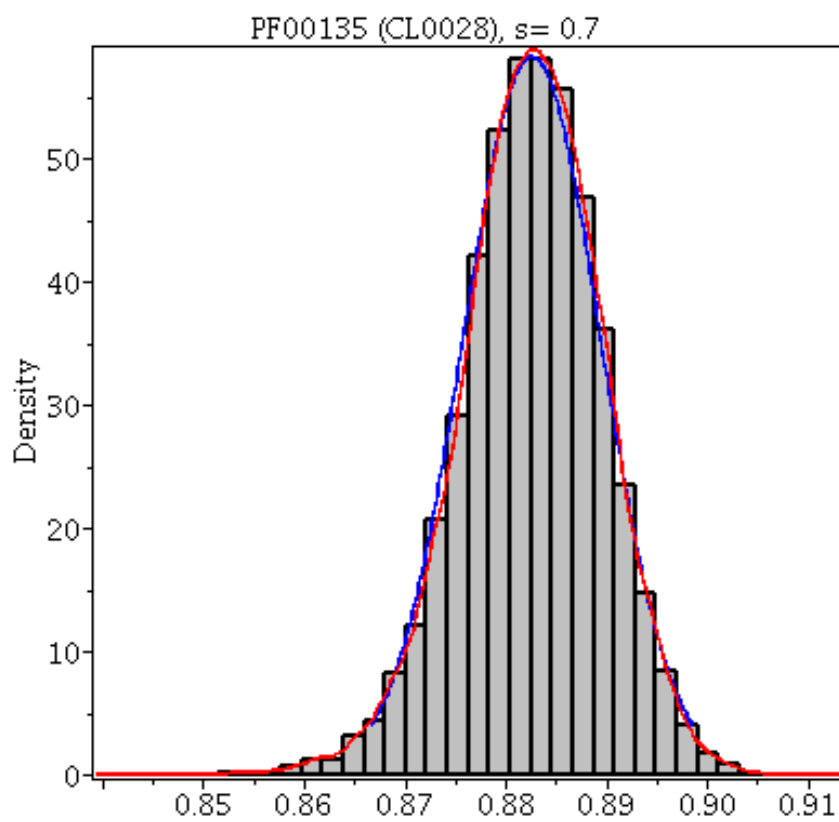
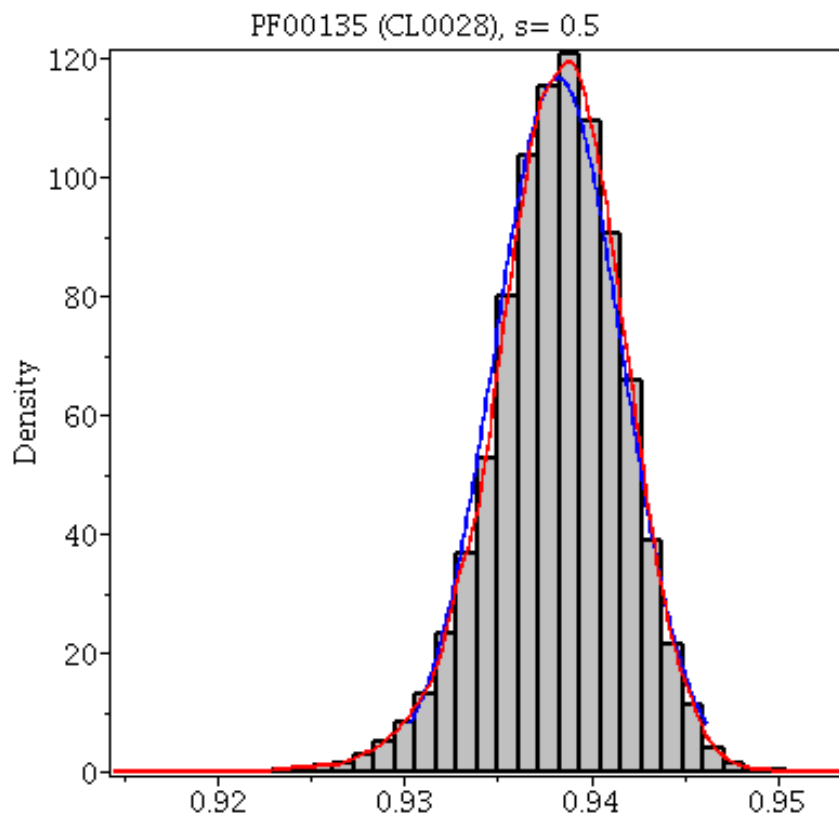


Fig. Figura 5.5 – Histogramas da família PF0005, clă CL0023







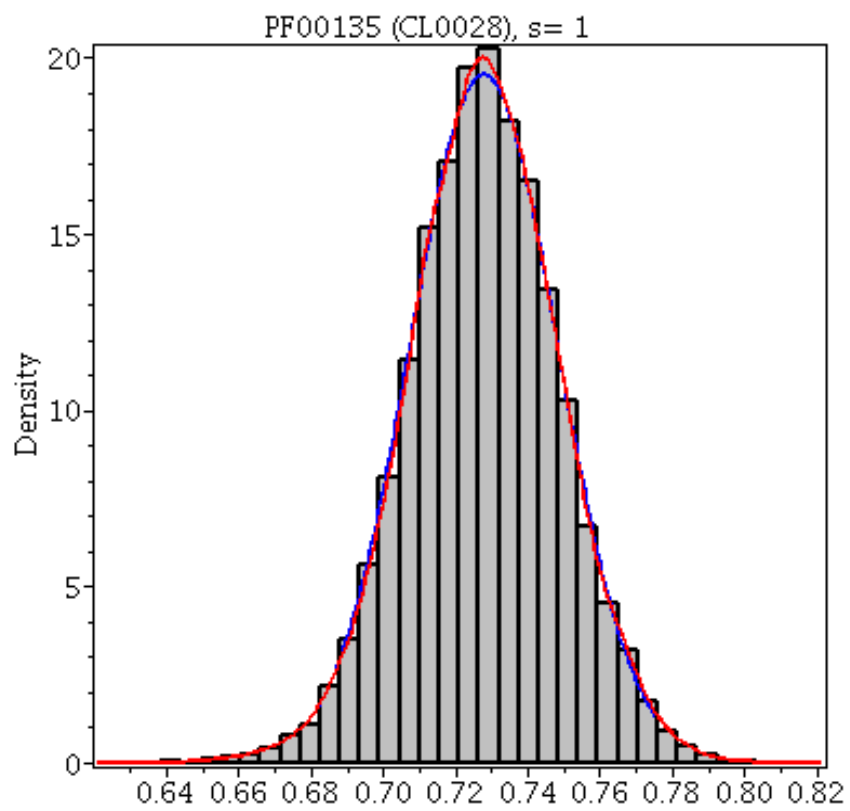
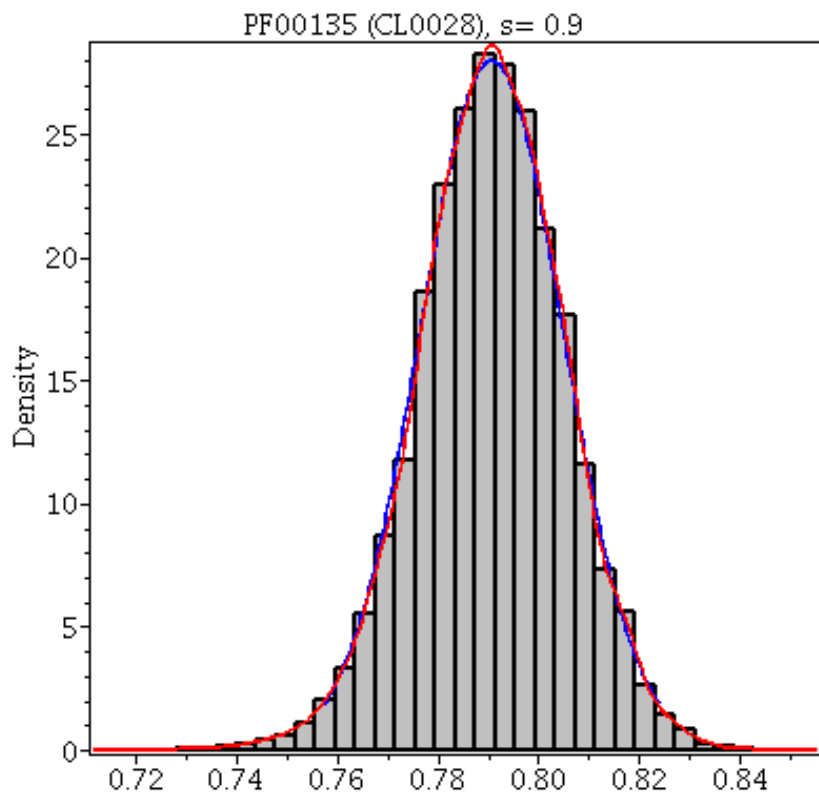
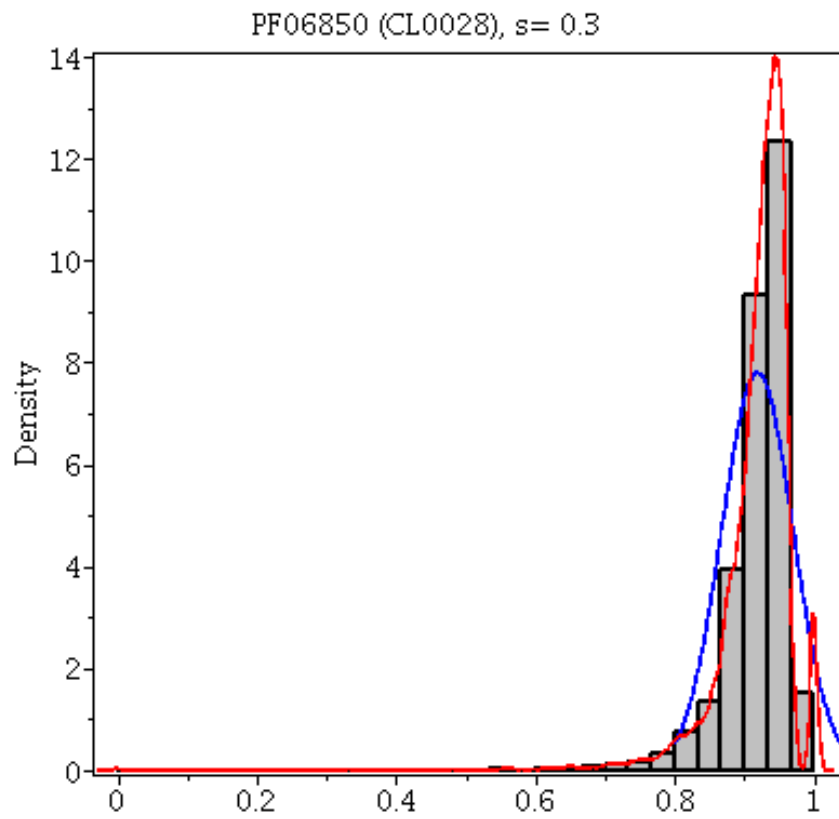
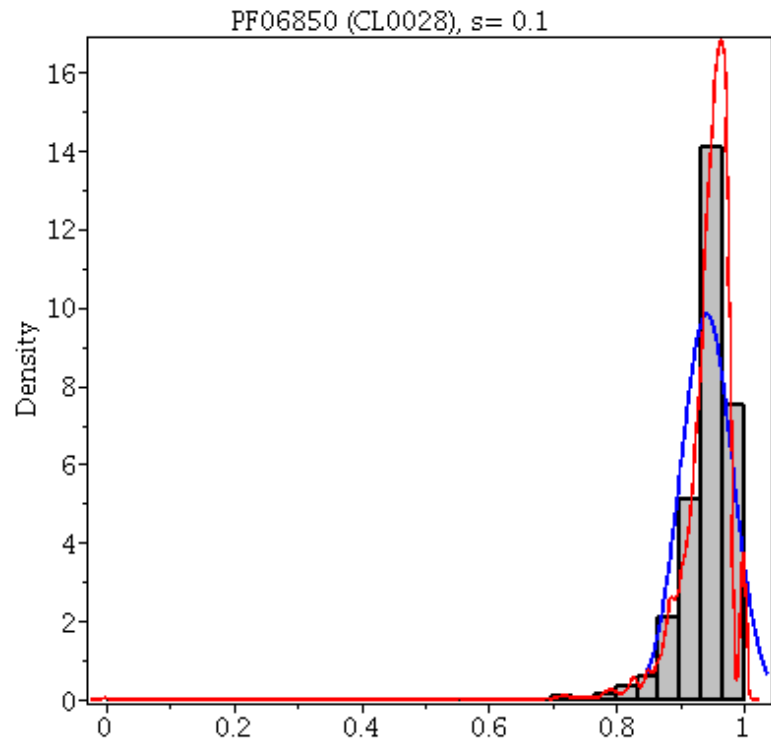
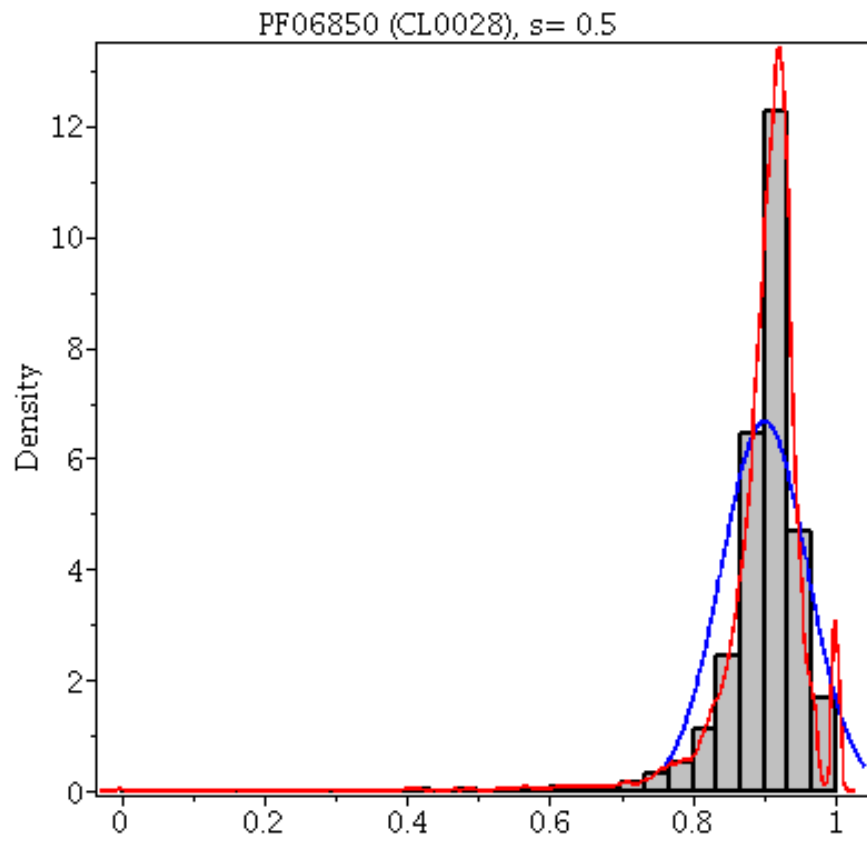
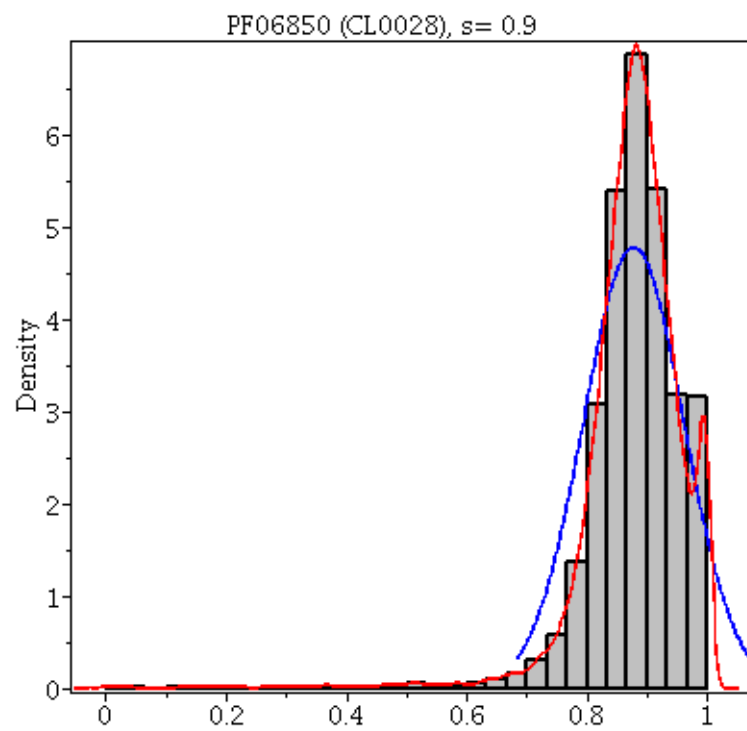
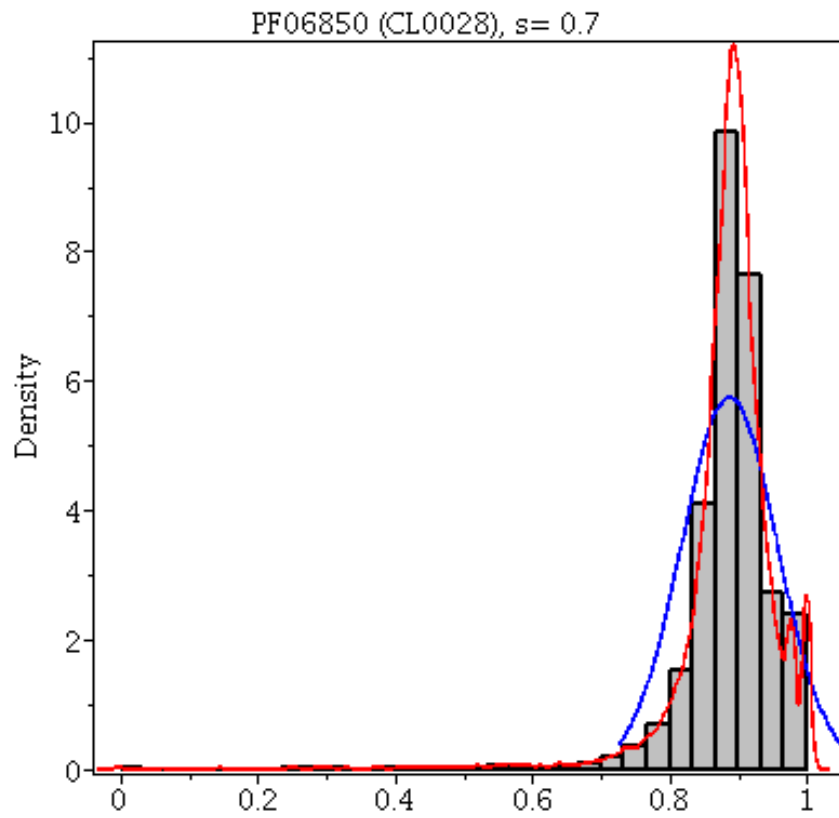


Figura 5.6 – Histogramas da família PF00135, clã CL0028







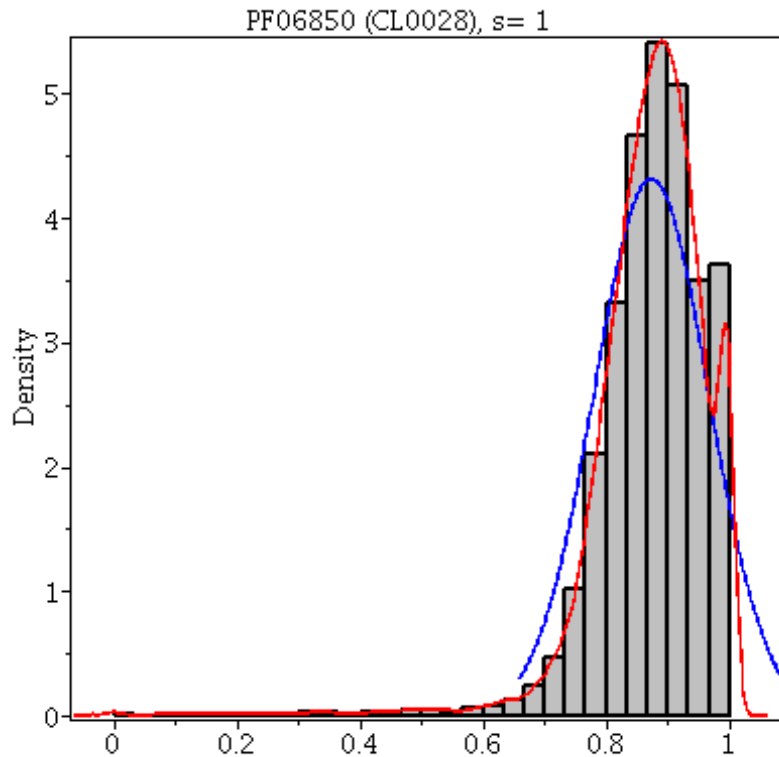


Figura 5.7 – Histogramas da família PF06850, clã CL0028

Capítulo VI – Comparação entre Sistemas Computacionais, entre Sistemas Operacionais e entre Estruturas no desempenho dos cálculos.

Atualmente parece ser inevitável a utilização de software para a resolução de problemas matemáticos sem a utilização de computadores e de produtos de software, pois muito tempo deveria ser investido para a realização de cálculos repetitivos, sem complexidade e que poderiam levar a difusão de erros na solução do problema e conseqüentemente na análise dos dados obtidos.

6.1 Sistemas computacionais e operacionais

Para o problema do cálculo da probabilidade simples e conjunta das famílias de entropia, inicialmente, foi utilizado o software matemático Maple. As principais

razões para utilizar o Maple são: a sua simplicidade de uso, facilidade de ter um ambiente de programação com instruções simples e uma interface intuitiva e amigável para o usuário. Além disso, a facilidade de aprendizado do software, permite que iniciantes aprendam técnicas de programação e do uso de funções em um curto período de tempo.

O Maple é uma solução computacional interessante por combinar em um ambiente integrado facilidades para descrever, visualizar, analisar e resolver problemas matemáticos. Oferece também a possibilidade de desenvolvimento de gráficos em duas e três dimensões para obter uma melhor compreensão e percepção dos dados analisados, sendo muito utilizados nas áreas de matemática, física, engenharias, dentre outras. O uso das características computacionais do software matemático *Maple*, problemas de difícil solução podem ser solucionados de forma mais fácil e com maior rapidez, auxiliando na evolução da pesquisa em desenvolvimento. Além das características e vantagens descritas com o uso do *Maple*, ele ainda oferece a possibilidade de ser executado em diferentes plataformas. Dessa forma é possível instalar e configurar o *Maple* em computadores que utilizam os seguintes sistemas operacionais: Windows da Microsoft, los da Apple e nas distribuições do Linux.

Entretanto o software *Maple* não apresentou resultados satisfatórios quando necessitou realizar cálculos estatísticos com longas *strings* de aminoácidos extraídas de bancos de dados de proteínas. Para a manipulação de arquivos de texto muito grande, uma das linguagens de programação mais adequada é o Perl. Esta linguagem é bastante utilizada pelos biólogos e por programadores que necessitam realizar tratamento de strings.

A linguagem Perl foi utilizada para avaliar o banco de dados Pfam, realizando cálculos de probabilidade de ocorrência de aminoácidos e de medidas de entropia.

Um grande número de operações elementares foram realizadas em uma estrutura de array multidimensional.

6.2 Estruturas

Algumas adaptações dos métodos utilizados foram feitas para serem usadas com o sistema computacional Perl e os novos tempos de cálculos são comparados com aqueles utilizados com Maple. Alguns comentários são incluídos com o uso do Perl, especialmente em sua forma para calcular com dados de entrada em forma de “arrays” e a maneira de contornar isso. Entretanto o cálculo de medidas de entropia é favorecido com o uso da estrutura intrínseca Hash do sistema computacional Perl. Em seguida, uma estrutura “array” modificada é introduzida para cálculo no Perl.

Capítulo VII – Análise Estatística ANOVA

Um teste de hipótese descreve teorias inicialmente consideradas como verdadeiras até serem rejeitadas pela referência aos fatos. Ela utiliza médias, variâncias, por exemplo, para afirmar algo sobre uma população. É muito utilizada para a avaliação de afirmativas ou hipóteses relativas a populações. Entretanto, pode também comparar os resultados obtidos com os esperados em um experimento ou ainda verificar se uma determinada distribuição de probabilidade é adequada a uma amostra em particular.

Uma hipótese estatística é uma opinião sobre um parâmetro de uma população. A opinião pode ou não ser verdadeira. As hipóteses são testadas

comparando-se estatísticas amostrais com o esperado, considerando que as amostras foram aleatoriamente retiradas de populações normalmente distribuídas.

O princípio que rege o teste de hipóteses diz que quanto maior o desvio entre o valor suposto e o amostral obtido, menor a chance de veracidade da hipótese. Rejeita-se a hipótese quando a diferença entre o proposto e o observado nas amostras é relevante. E aceita-se a hipótese quando a diferença entre o valor alegado e o amostral observado é estatisticamente insignificante.

Os testes iniciam com a formulação de duas proposições mutuamente excludentes: uma afirma uma dada característica da amostra estudada e a outra nega-a. A veracidade de uma implica na falsidade da outra e a ocorrência de uma obriga a ausência da outra.

Existem dois tipos de hipóteses estatísticas: a hipótese nula (H_0) e a hipótese alternativa (H_a). A hipótese nula (H_0) declara a proposição como verdadeira em relação a característica a avaliar, enquanto que a hipótese alternativa (H_a) nega a proposição. Em princípio, o teste objetiva rejeitar a H_0 , mas também pode-se considerar outras ocorrências que impactem na observação final do experimento.

Em alguns experimentos é fundamental compreender se há ou não uma diferença significativa entre as médias das populações. Por exemplo, pode ocorrer que a variação entre a média das amostras seja grande em relação à variação encontrada na média das amostras individuais.

Caso exista uma diferença é importante verificar se a diferença observada é devido a variação da amostra ou variações nos dados da amostra.

Existe uma técnica para determinar o quão provável que as diferenças observadas entre as médias das amostras, sejam devidas a erros nas amostras. Essa técnica é conhecida como Análise de Variância (ANOVA). Ela possibilita

determinar se qualquer diferença observada entre as entre as médias das amostras é devido a erro na amostra.

Assim ANOVA é um método estatístico para determinar a existência de diferenças entre várias médias de uma população. Ela é projetada para detectar diferenças entre médias de populações sujeitas a diferentes tratamentos.

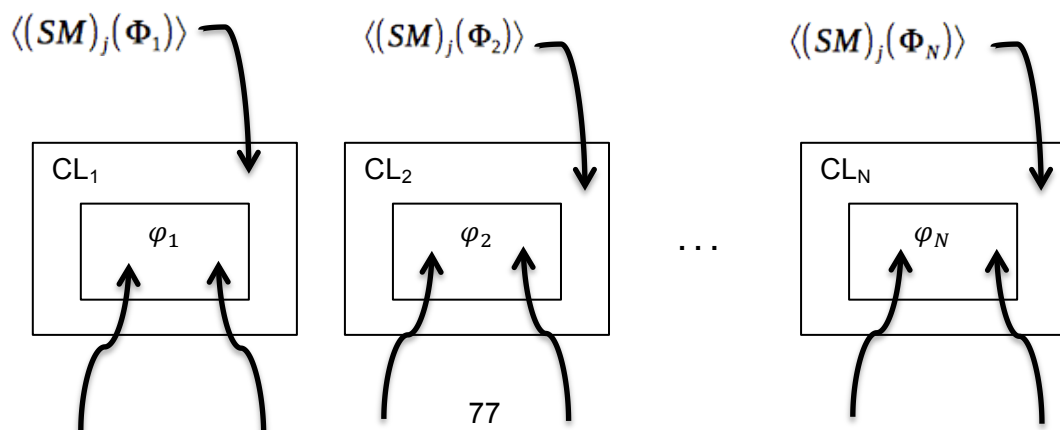
7.1 ANOVA para medidas de Entropia advindas de probabilidades simples

ANOVA foi utilizada para validar a classificação das famílias em clãs, através da análise e comparação das variâncias das distribuições de entropia dos clãs e entre os clãs que foram comparadas a um valor teórico da distribuição de Fisher.

Considere um conjunto de clãs CL_1, CL_2, \dots, CL_N com $\Phi_1, \Phi_2, \dots, \Phi_n$ famílias e que são reduzidas a $\varphi_1, \dots, \varphi_n$ famílias em virtude da restrição de que contenham blocos $m \times n$. Cálculos de médias e de variâncias das entropias por coluna ou por par de colunas foram realizados com os clãs que após a restrição do bloco $m \times n$, contenham um mínimo de 5 famílias.

A Figura 7.1 apresenta o l -ésimo clã com média de entropia por coluna

$\langle (SM)_j(\Phi_l) \rangle$ e variância $\sigma_{j\Phi_l}$, é representado pela média $\langle (SM)_j(\varphi_l) \rangle$ e variância $\sigma_{j\varphi_l}$.



$$\langle (SM)_j(\varphi_1) \rangle \sigma_{j\varphi_1} \quad \langle (SM)_j(\varphi_2) \rangle \sigma_{j\varphi_2} \quad \langle (SM)_j(\varphi_N) \rangle \sigma_{j\varphi_N}$$

Figura 7.1 – Entropia de distribuição de aminoácidos em uma coluna

A equação abaixo representa a média das famílias de um clã l da entropia de uma coluna:

$$\langle (SM)_j(\varphi_l) \rangle = \frac{1}{\varphi_l} \sum_{p=1}^{\varphi_l} (SM)_j^p(\varphi_l) \quad (7.1)$$

onde $p = 1, 2, \dots, \varphi_l$ são as famílias que compõem o l -ésimo clã. A j -ésima coluna pertence ao intervalo $j = 1, 2, \dots, n$ e o l -ésimo clã ao intervalo $l = 1, 2, \dots, N$.

O desvio padrão da média de entropias da probabilidade simples ($\sigma_{j\varphi_l}$) é dado pela equação:

$$(\varphi - 1)(\sigma_{j\varphi_l}^2) = \sum_{p=1}^{\varphi_l} ((SM)_j^p(\varphi_l) - \langle (SM)_j(\varphi_l) \rangle)^2 \quad (7.2)$$

A média de todas as famílias de todos os clãs, média total, para a probabilidade simples (de uma coluna j) é dada por:

$$\langle (SM)_j \rangle = \frac{1}{\sum_{l=1}^N \varphi_l} \sum_{l=1}^N \sum_{p=1}^{\varphi_l} (SM)_j^p(\varphi_l) \quad (7.3)$$

O desvio padrão das médias totais de entropias de probabilidade simples (σ_j) é dado pela equação:

$$\left(\sum_{l=1}^N \varphi_l - 1 \right) \sigma_j^2 = \sum_{l=1}^N \sum_{p=1}^{\varphi_l} ((SM)_j^p(\varphi_l) - \langle (SM)_j \rangle)^2 \quad (7.4)$$

E então tem-se que a variação entre as médias dos clãs em relação a média total da soma dos quadrados das amostras entre grupos (*Sum of Squares of samples between Groups* - SSG), é dada por:

$$SSG = \sum_{l=1}^N \varphi_l (\langle (SM)_j(\varphi_l) \rangle - \langle (SM)_j \rangle)^2 \quad (7.5)$$

E assim a seguinte relação é definida:

$$\underbrace{\left(\sum_{l=1}^N \varphi_l - 1 \right) \sigma_j^2}_{SST} = \underbrace{\sum_{l=1}^N (\varphi_l - 1) \sigma_{j\varphi_l}^2}_{SSE} + \underbrace{\sum_{l=1}^N \varphi_l (\langle (SM)_j(\varphi_l) \rangle - \langle (SM)_j \rangle)^2}_{SSG} \quad (7.6)$$

onde, a soma dos quadrados de todas as amostras (*Sum of Squares of Total samples* - SST) é a variabilidade entre o valor da média da entropia de cada família em relação à média total (7.3).

Enquanto que a soma dos quadrados das amostras pertencentes aos grupos (*Sum of Squares of samples within Groups* – SSE) é a variabilidade entre as entropias de cada família em relação à média dos clãs aos quais elas pertencem (7.1).

Para verificar o número de termos independentes, tem-se que:

$$\sum_{l=1}^N \varphi_l - 1 = \sum_{l=1}^N (\varphi_l - 1) + N - 1 = \sum_{l=1}^N \varphi_l - N + N - 1$$

O coeficiente de correlação de Fisher para a entropia de probabilidade simples é dado da seguinte forma:

$$F_j = \frac{\frac{SSG}{N-1}}{\frac{SSE}{\sum_{l=1}^N \varphi_l - N}} = \left(\frac{\sum_{l=1}^N \varphi_l - N}{N-1} \right) \cdot \left(\frac{\left(\sum_{l=1}^N \varphi_l - 1 \right) \sigma_j^2}{\sum_{l=1}^N (\varphi_l - 1) \sigma_{j\varphi_l}^2} - 1 \right), j=1, 2, \dots, n \quad (7.7)$$

Serão realizados n testes ANOVA para o cálculo individualmente em cada coluna, por valor de parâmetro.

7.2 ANOVA para medidas de Entropia advindas de probabilidades compostas

O mesmo tratamento é dado para a média e variância de entropia de pares de colunas, conforme a Figura 7.2.

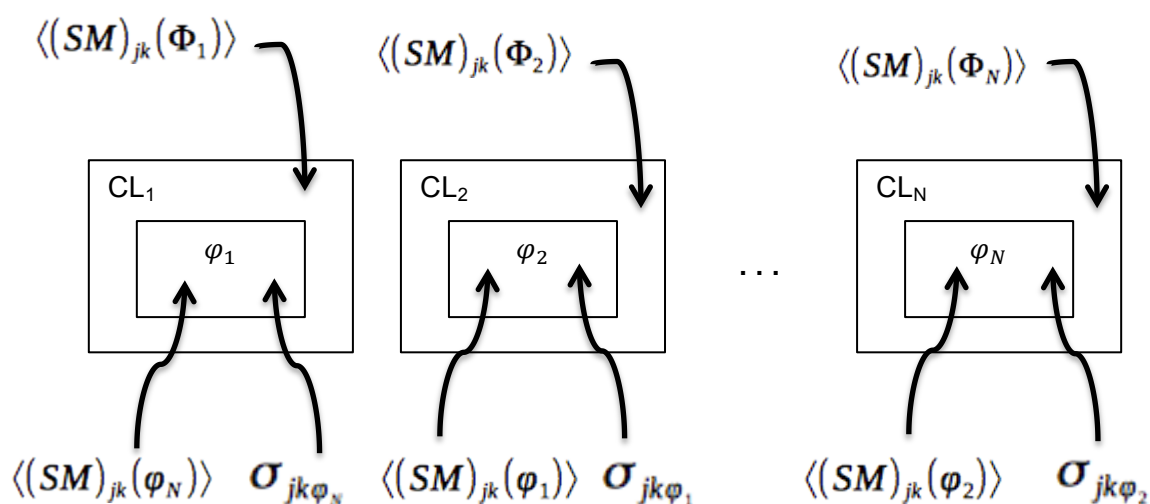


Figura 7.2 – Entropia de distribuição de aminoácidos em um par de colunas

Da mesma forma que foi apresentado para uma coluna, a média de um clã da entropia para um par de colunas jk é dado por:

$$\langle (SM)_{jk}(\varphi_l) \rangle = \frac{1}{\varphi_l} \sum_{p=1}^{\varphi_l} (SM)_{jk}^p(\varphi_l) \quad (7.8)$$

com $j = 1, 2, \dots, (n-1)$ e $k = (j + 1), (j + 2), \dots, n$.

Enquanto que o desvio padrão da média de entropia da probabilidade conjunta $\sigma_{jk\varphi_l}$ é dado por:

$$(\varphi - 1)(\sigma_{jk\varphi_l}^2) = \sum_{p=1}^{\varphi_l} ((SM)_{jk}^p(\varphi_l) - \langle (SM)_{jk}(\varphi_l) \rangle)^2 \quad (7.9)$$

E a média total para a probabilidade conjunta (par de colunas) é dada por:

$$\langle (SM)_{jk} \rangle = \frac{1}{\sum_{l=1}^N \varphi_l} \sum_{l=1}^N \sum_{p=1}^{\varphi_l} (SM)_{jk}^p(\varphi_l) \quad (7.10)$$

E o desvio padrão das médias totais de entropias de probabilidade conjunta (σ_{jk}) é dado pela equação:

$$\left(\sum_{l=1}^N \varphi_l - 1 \right) \sigma_{jk}^2 = \sum_{l=1}^N \sum_{p=1}^{\varphi_l} ((SM)_{jk}^p(\varphi_l) - \langle (SM)_{jk} \rangle)^2 \quad (7.11)$$

De forma equivalente, para entropias de pares de colunas, as equações (7.5) e (7.6) podem ser escritas da seguinte forma:

$$SSG = \sum_{l=1}^N \varphi_l (\langle (SM)_{jk}(\varphi_l) \rangle - \langle (SM)_{jk} \rangle)^2 \quad (7.12)$$

$$\underbrace{\left(\sum_{l=1}^N \varphi_l - 1 \right)}_{SST} \sigma_{jk}^2 = \underbrace{\sum_{l=1}^N (\varphi_l - 1) \sigma_{jk\varphi_l}^2}_{SSE} + \underbrace{\sum_{l=1}^N \varphi_l (\langle (SM)_{jk}(\varphi_l) \rangle - \langle (SM)_{jk} \rangle)^2}_{SSG} \quad (7.13)$$

E o coeficiente de correlação de Fisher para a entropia de probabilidade composta é dado da seguinte forma:

$$F_{jk} = \frac{\frac{SSG}{N-1}}{\frac{SSE}{\sum_{l=1}^N \varphi_l - N}} = \left(\frac{\sum_{l=1}^N \varphi_l - N}{N-1} \right) \cdot \left(\frac{\sum_{l=1}^N \varphi_l - 1}{\sum_{l=1}^N (\varphi_l - 1) \sigma_{jk\varphi_l}^2} - 1 \right), \quad j=1, 2, \dots, n-1 \quad k=(j+1), (j+2), \dots, n \quad (7.14)$$

Serão realizados $n \frac{(n-1)}{2}$ testes ANOVA para o cálculo com pares de colunas, por valor de parâmetro.

7.3 ANOVA para medidas de Entropia Jaccard

O valor teórico para ser comparado com os valores calculados, $F_{\mu\nu\alpha}$ é obtido através de uma função de variável aleatória denominada distribuição F de Fisher-Snedecor.

A função de distribuição acumulada (*cumulative distribution function* – cdf) é dada por:

$$g(\mu, \nu; F_{\mu\nu\alpha}) = \int_0^{F_{\mu\nu\alpha}} f(\mu, \nu; t) dt = 1 - \alpha \quad (7.15)$$

sendo $f(\mu, \nu; t)$ a função densidade de probabilidade (*probability density function* – pdf).

O teste de Fisher consiste na realização de um teste de hipótese para verificar a relação dos valores de F (experimental), calculados a partir de um banco de dados, com o F teórico da distribuição. O objetivo inicial é rejeitar a H_0 (Hipótese Nula), considerada inicialmente como verdadeira. Caso ela não possa ser comprovada, a H_a (Hipótese Alternativa) será considerada.

Neste trabalho, as H_0 e H_a são, respectivamente:

- $H_0: \langle (SM)_{jk}(\Phi_1) \rangle = \langle (SM)_{jk}(\Phi_2) \rangle = \dots = \langle (SM)_{jk}(\Phi_N) \rangle \Rightarrow$ não existência de clãs;

- $H_a : \langle (SM)_{jk}(\Phi_1) \rangle \neq \langle (SM)_{jk}(\Phi_2) \rangle \neq \dots \neq \langle (SM)_{jk}(\Phi_N) \rangle \Rightarrow$
existência de clãs. Porém, não sendo todos, necessariamente,
diferentes.

Para rejeitar a H_0 e reconhecer a existência de clãs, o valor experimental deve ser maior que o valor teórico: $F_j > F_{\mu \nu \alpha}$

E para a realização do teste, algumas suposições devem ser consideradas em relação aos dados a serem analisados (DEGROOT e SCHERVISH, 2012).

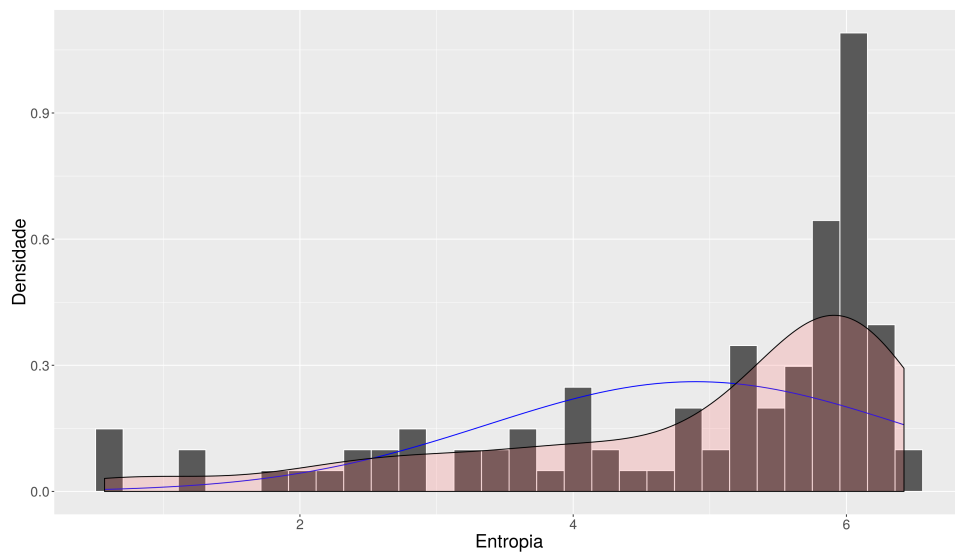
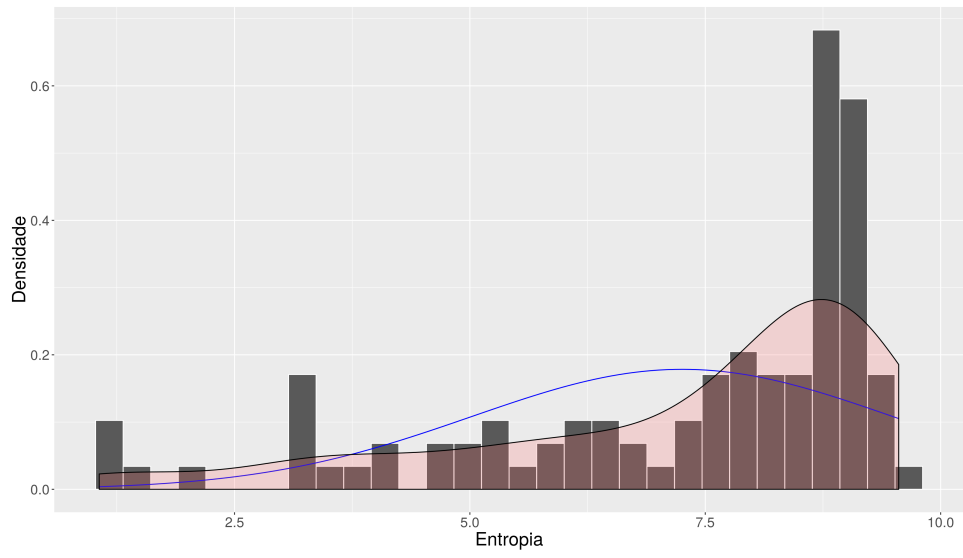
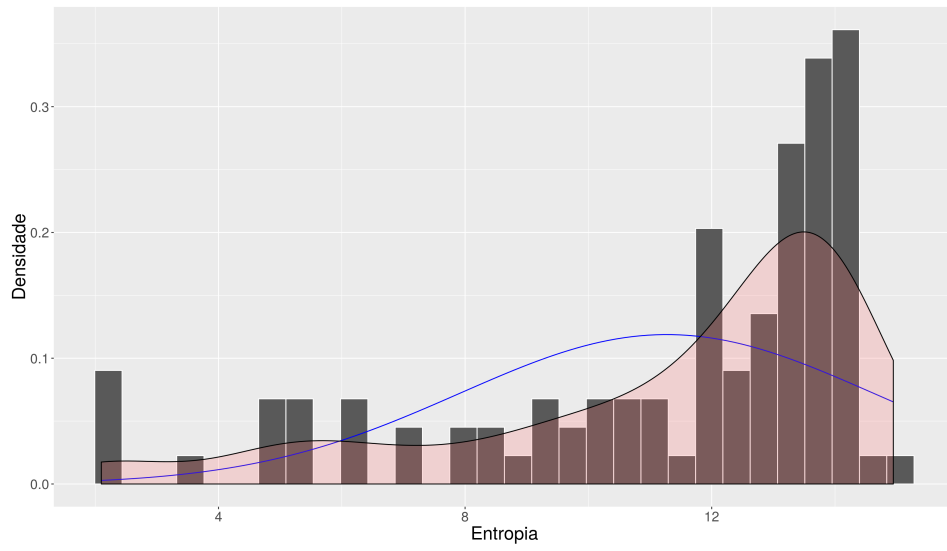
1. Os blocos $m \times n$ das N populações (clãs) são independentes;
2. Os blocos $m \times n$ devem ser normalmente distribuídos;
3. Os blocos $m \times n$ devem ser selecionados de populações com $\sigma_{j\varphi}^2$ variância iguais.

A versão 27.0 do Pfam, utilizada neste trabalho, possui um total de 14.831 famílias, entretanto somente 4.563 famílias estão organizadas em 515 clãs. Com o uso do bloco de 100×200 , foram obtidas 1.441 famílias organizadas em 267 clãs. E com a restrição de clãs formados com um mínimo de 5 famílias, obteve-se um espaço amostral de 1.069 famílias organizadas em 68 clãs.

7.4 ANOVA para medidas de Entropia - janelas 100×100

Foram feitos os histogramas de probabilidade simples e conjunta para três famílias (PF00005, PF00135, PF06850) de dois clãs (CL0023 e CL0028) e para diferentes valores do parâmetro s (0.1, 0.3, 0.5, 0.7, 0.9 e 1.0), selecionados a partir dos Clãs obtidos nas janelas 100×100 .

Inicialmente são apresentados os histogramas para probabilidades simples:



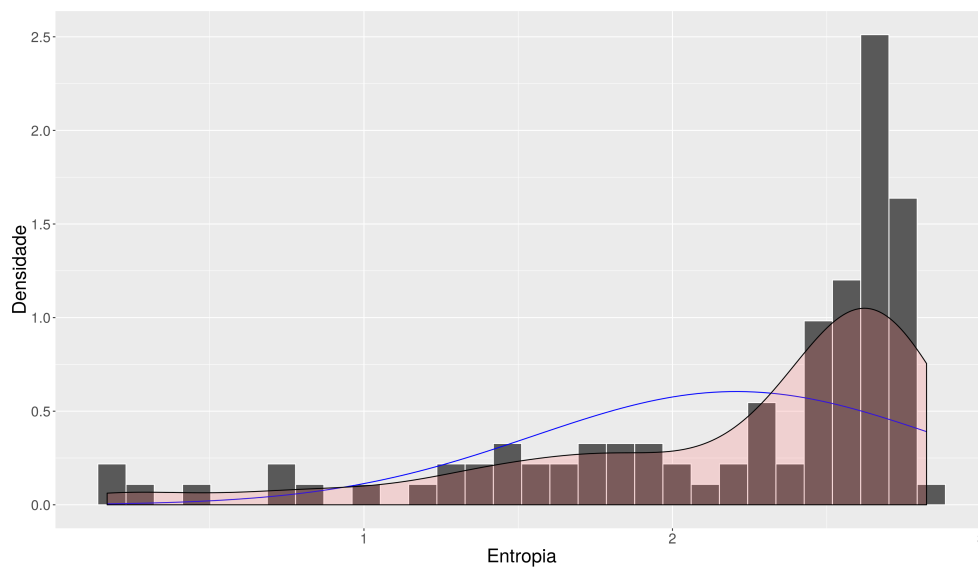
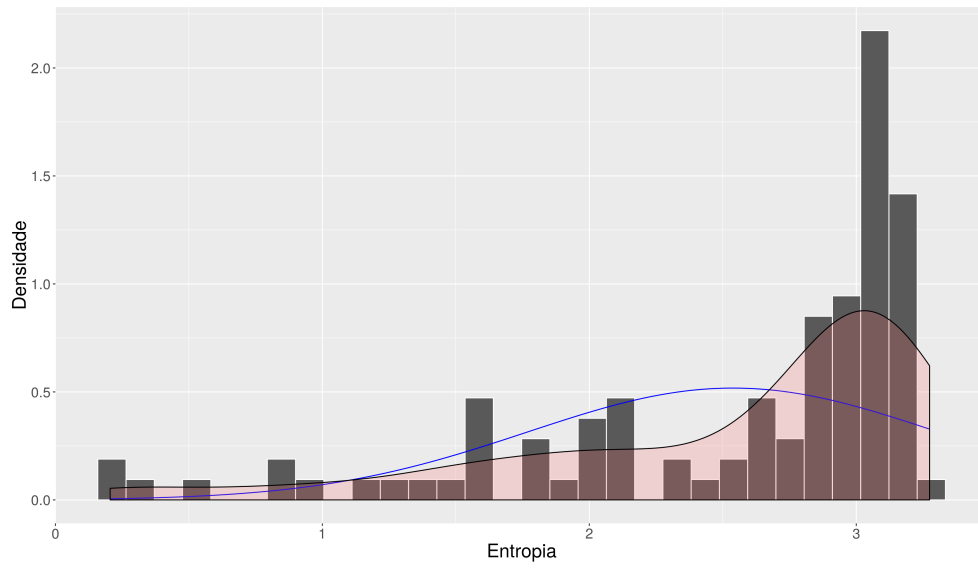
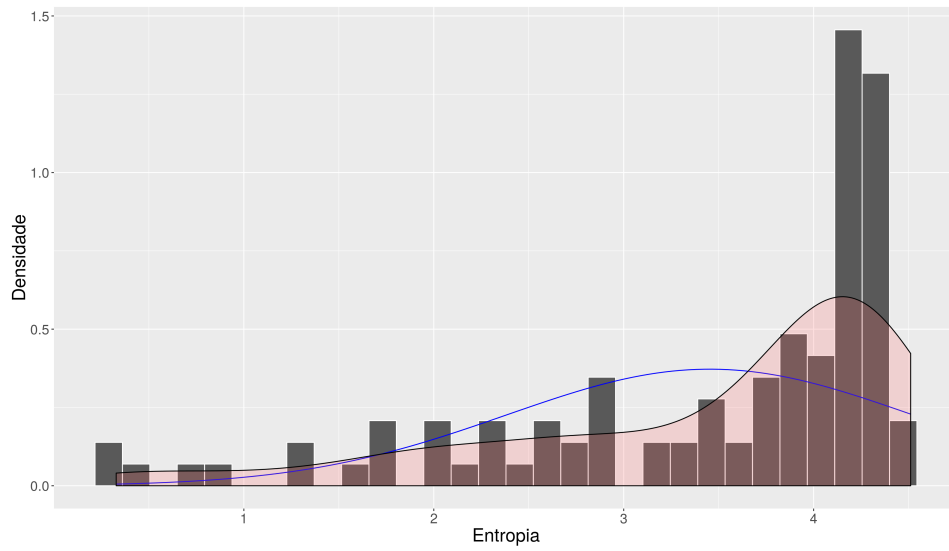
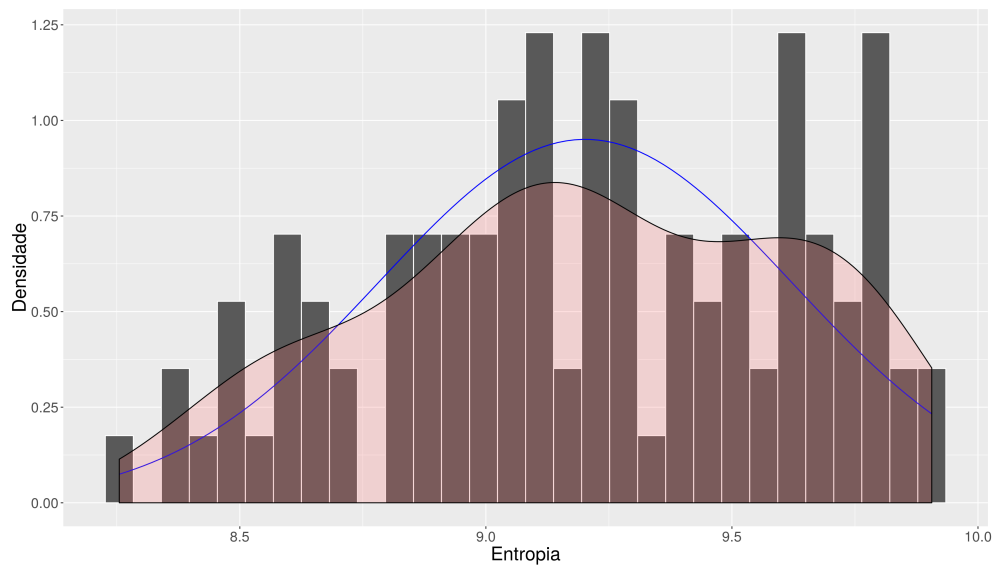
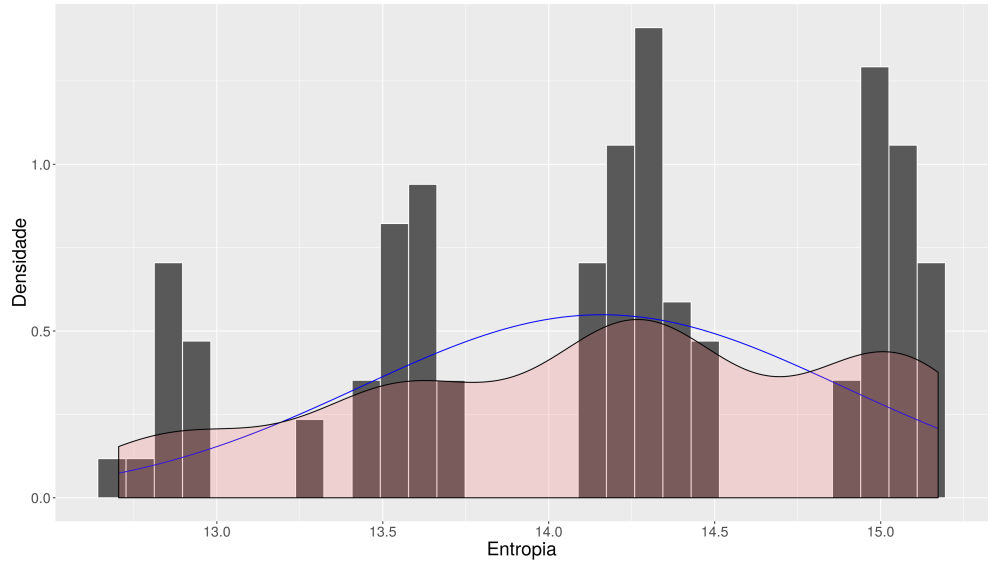
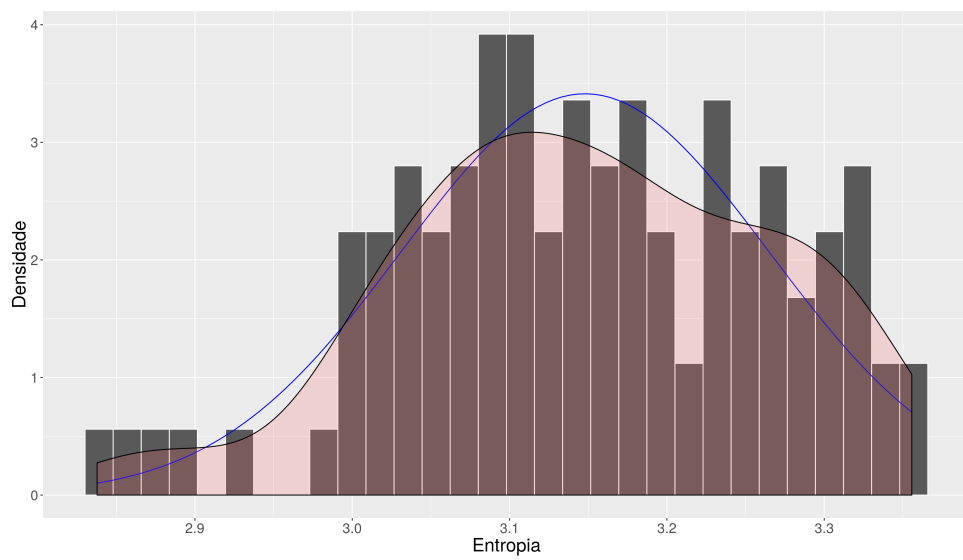
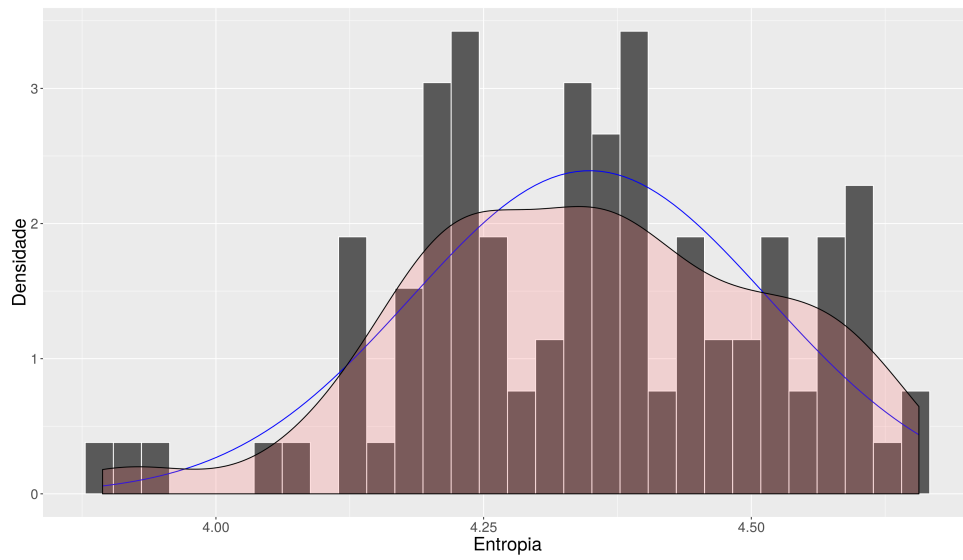
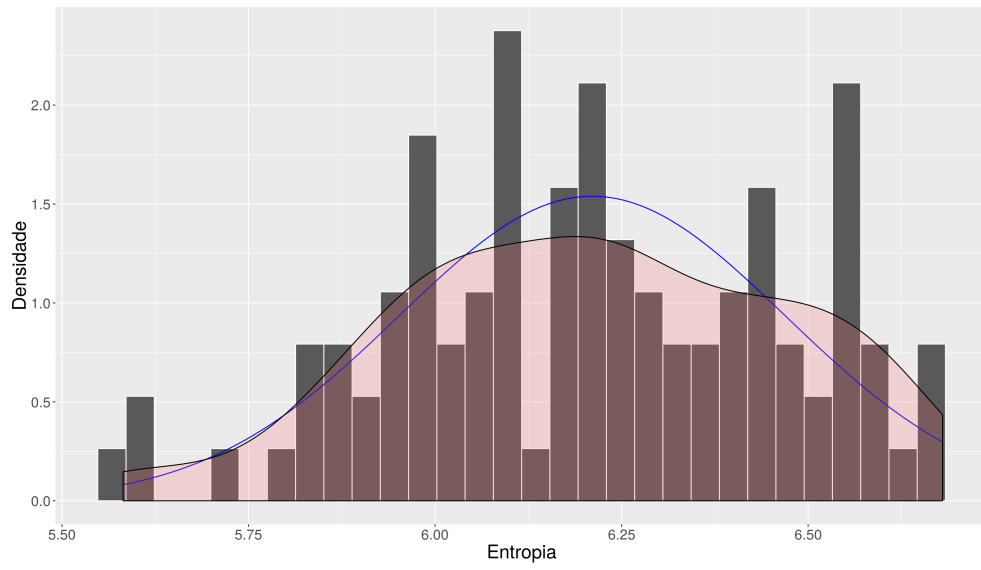


Figura 7.3 – Histogramas da família PF0005, clã CL0023 ($s = 0.1, 0.3, 0.5, 0.7, 0.9$ e 1.0) – probabilidade simples





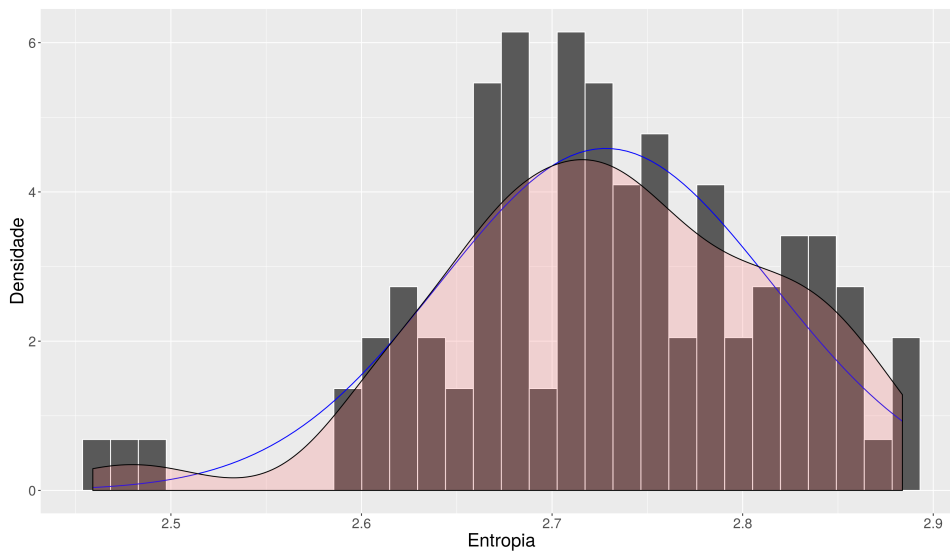
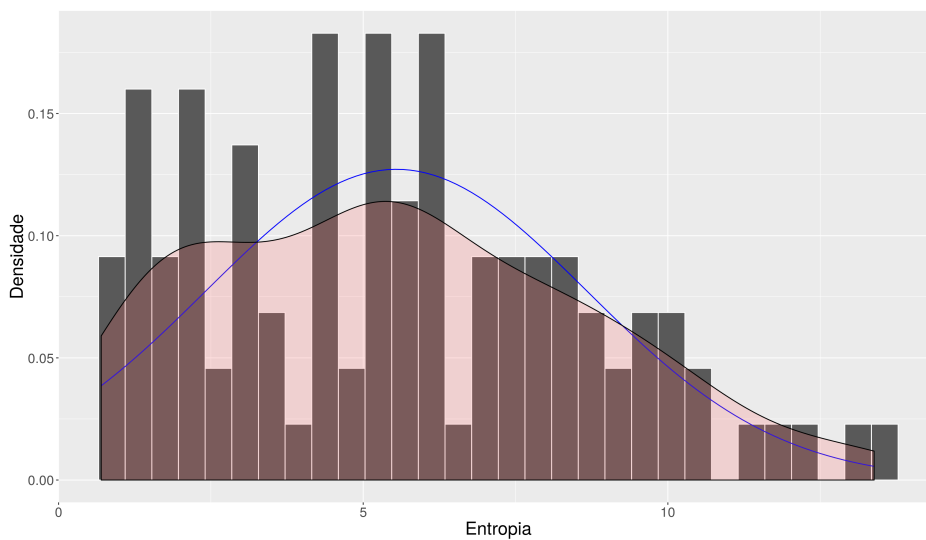
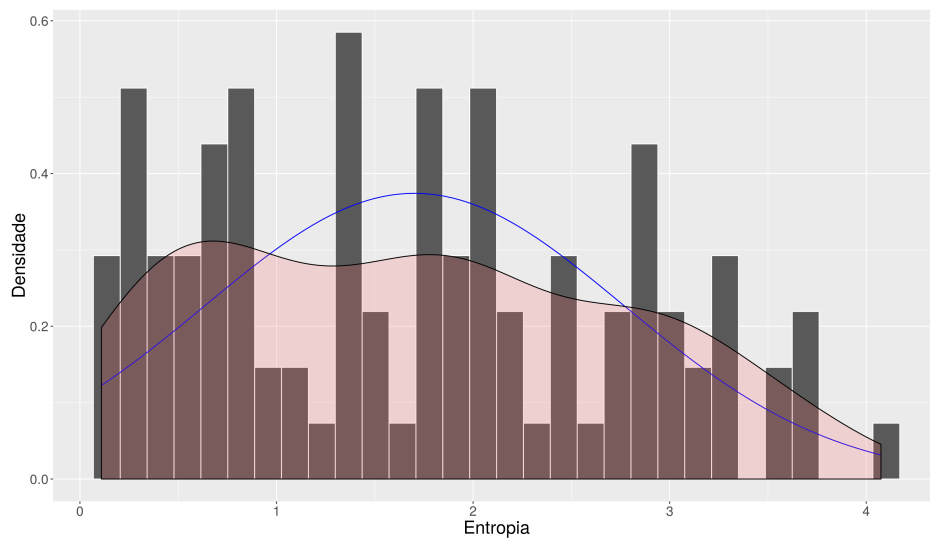
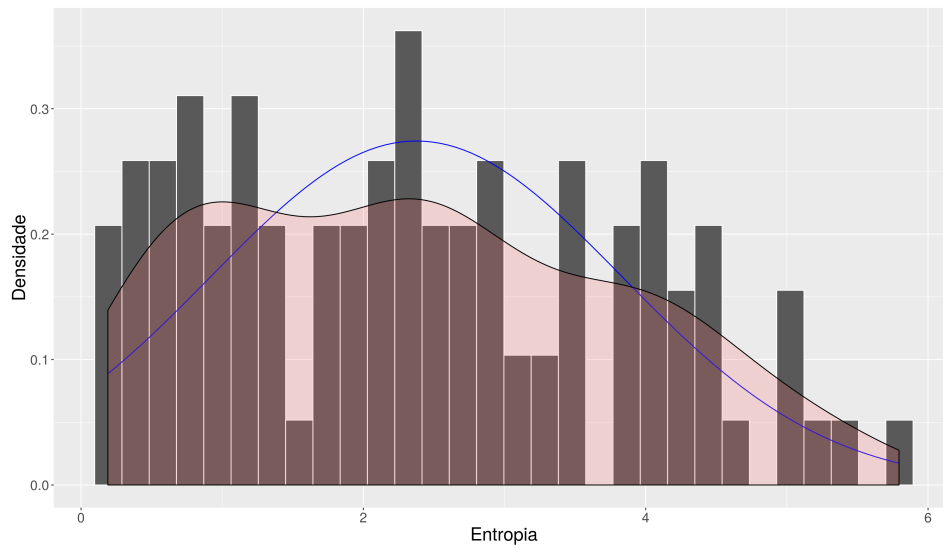
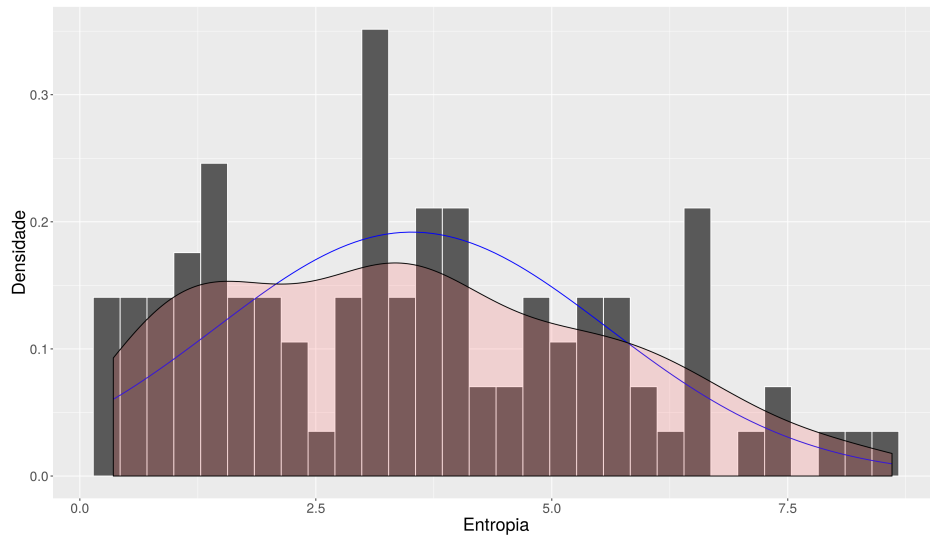


Figura 7.4 – Histogramas da família PF00135, clã CL0028 ($s = 0.1, 0.3, 0.5, 0.7, 0.9$ e 1.0) – probabilidade simples





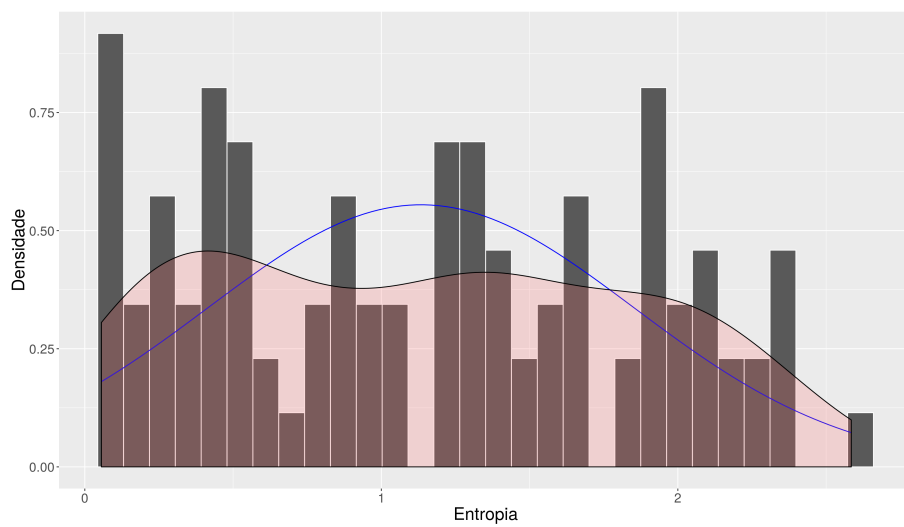
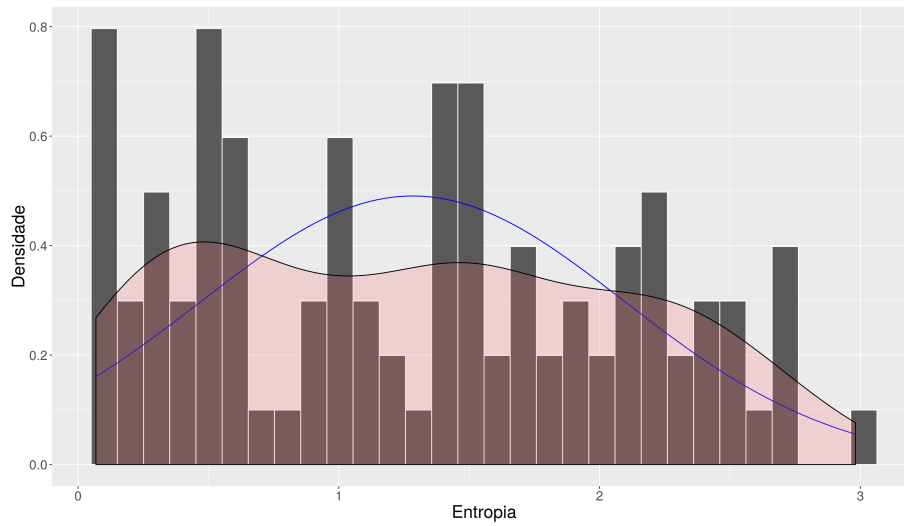
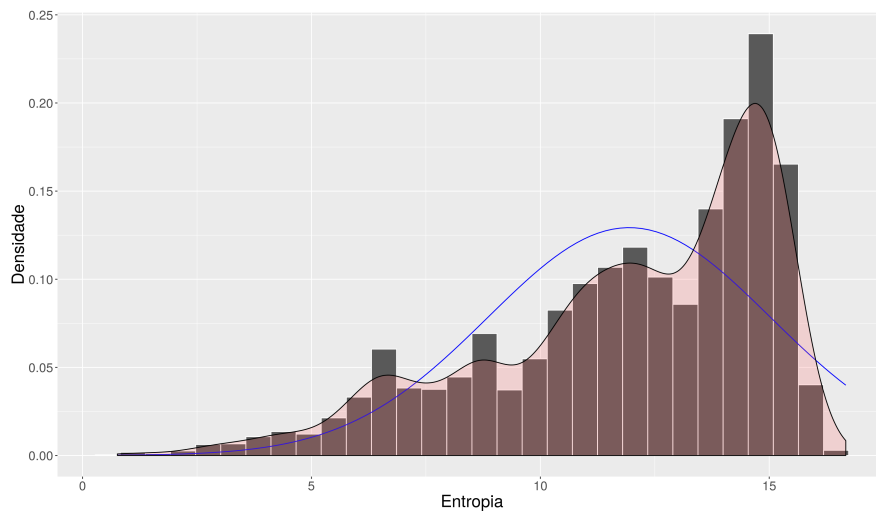
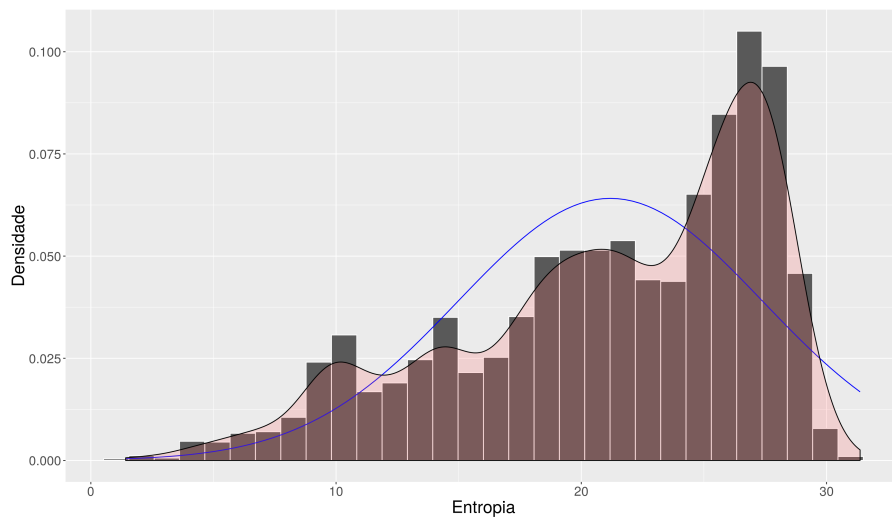
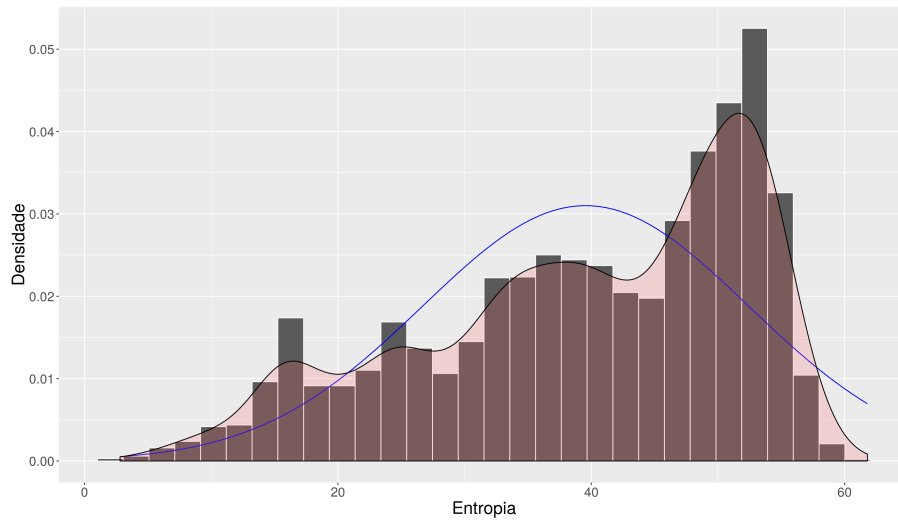


Figura 7.5 – Histogramas da família PF06850, clã CL0028 ($s = 0.1, 0.3, 0.5, 0.7, 0.9$ e 1.0) – probabilidade simples

Em seguida são apresentados os histogramas para probabilidades conjuntas:



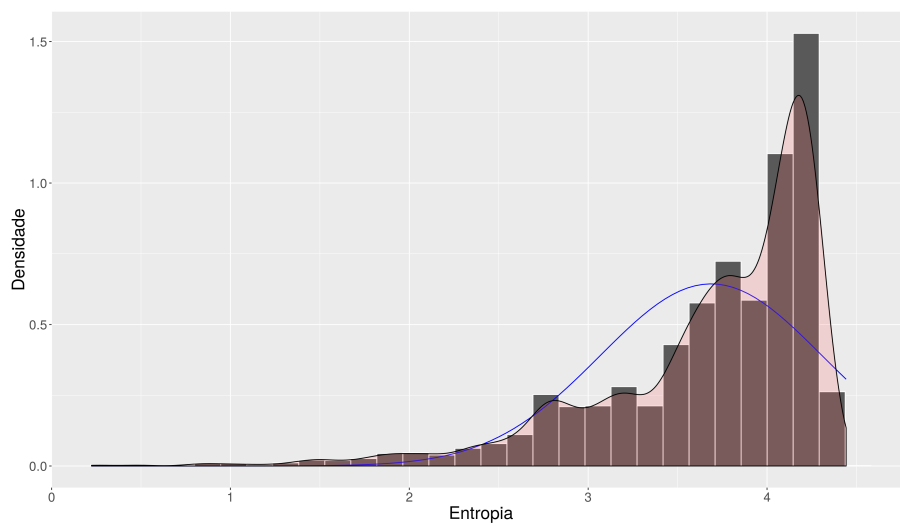
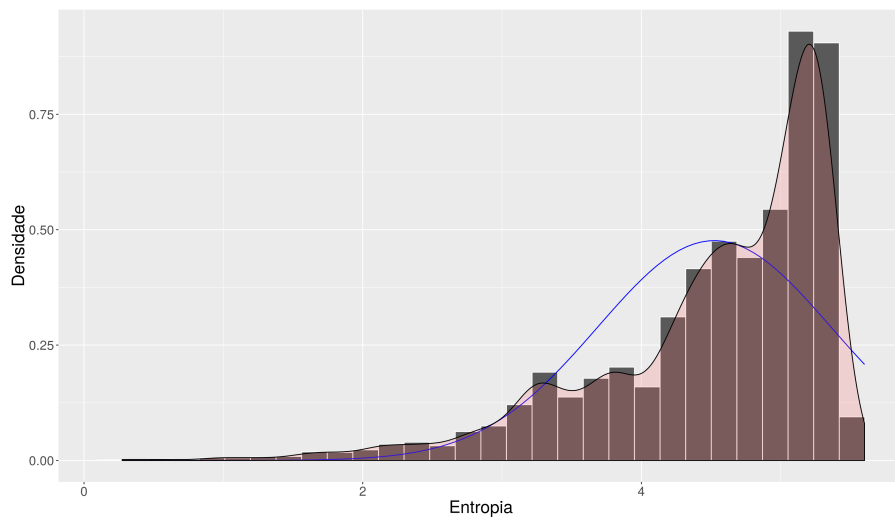
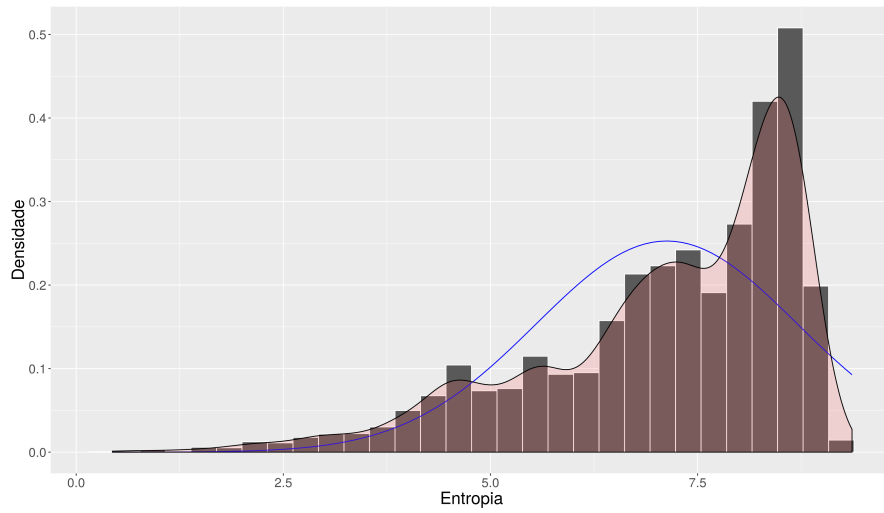
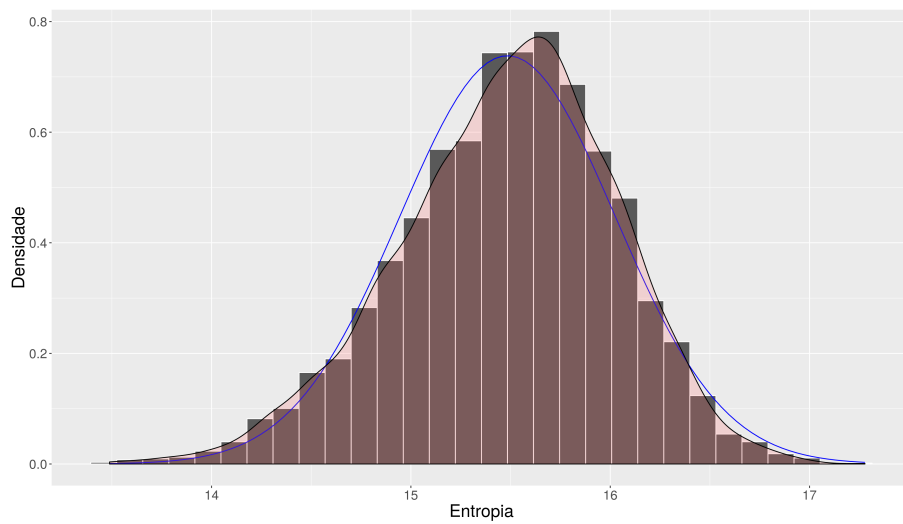
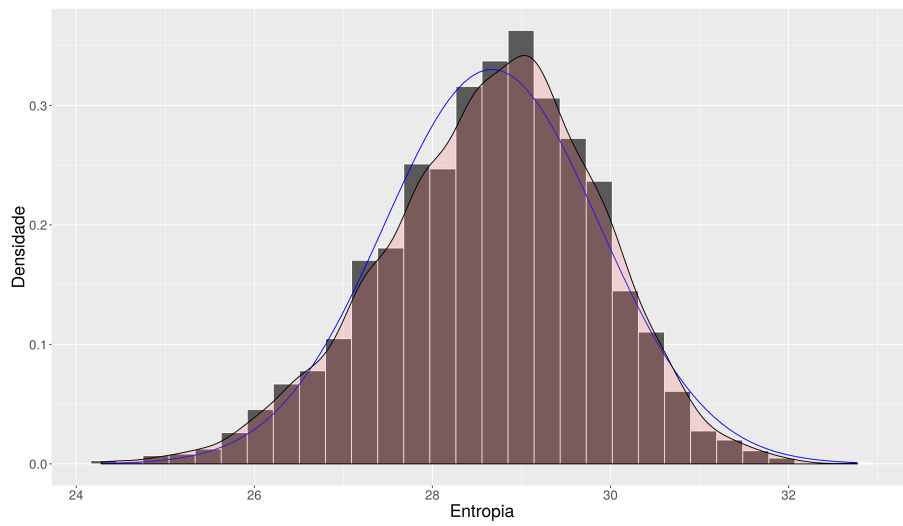
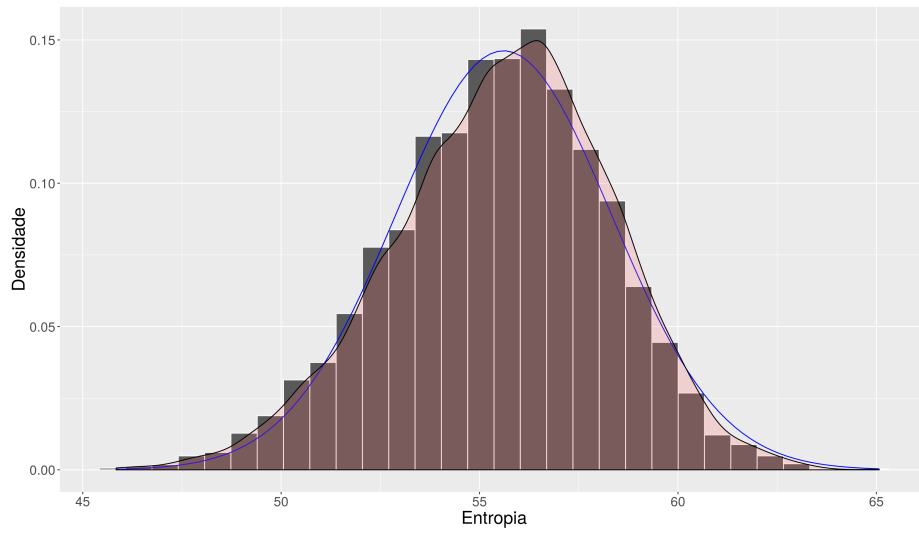


Figura 7.6 – Histogramas da família PF0005, clã CL0023 ($s = 0.1, 0.3, 0.5, 0.7, 0.9$ e 1.0) – probabilidade composta



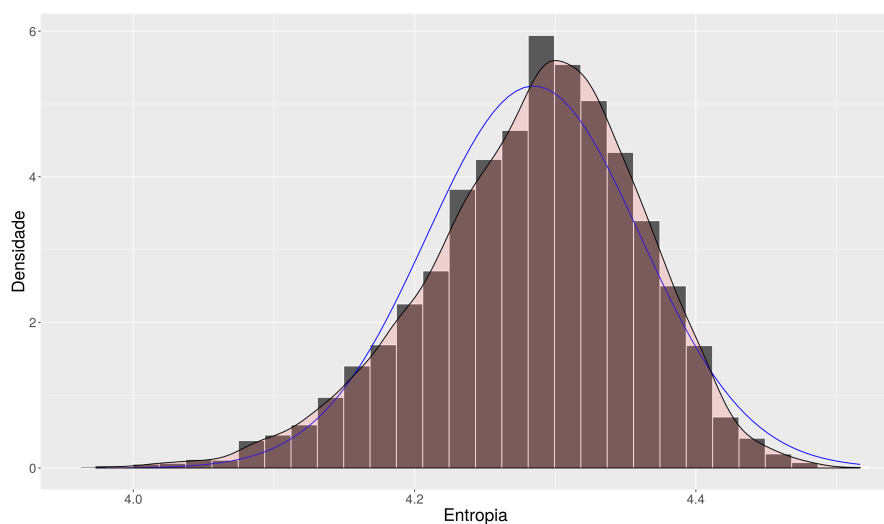
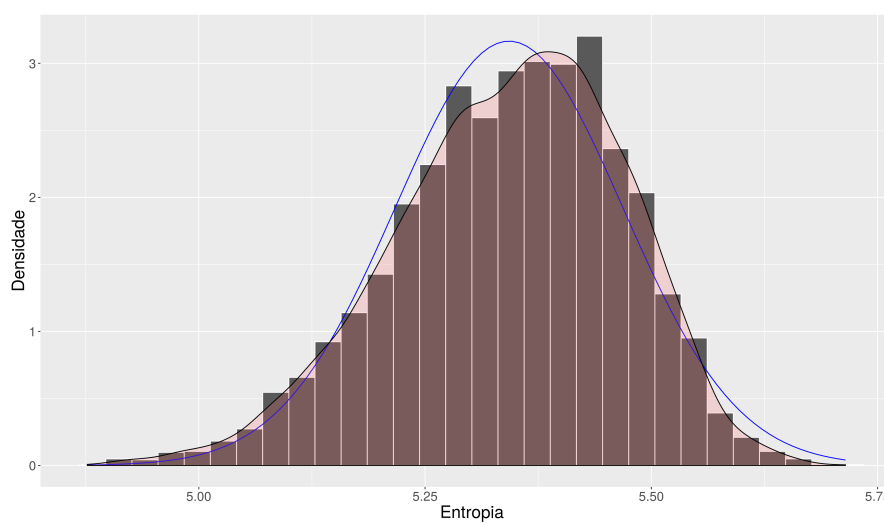
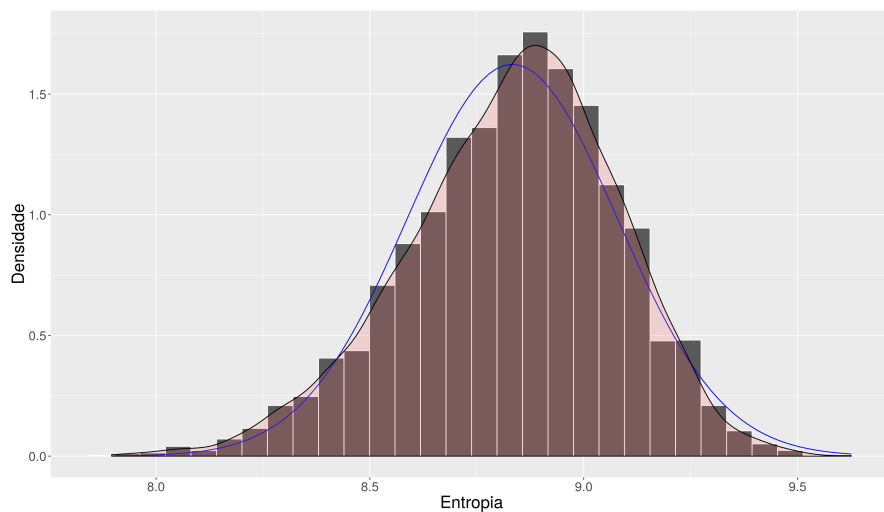
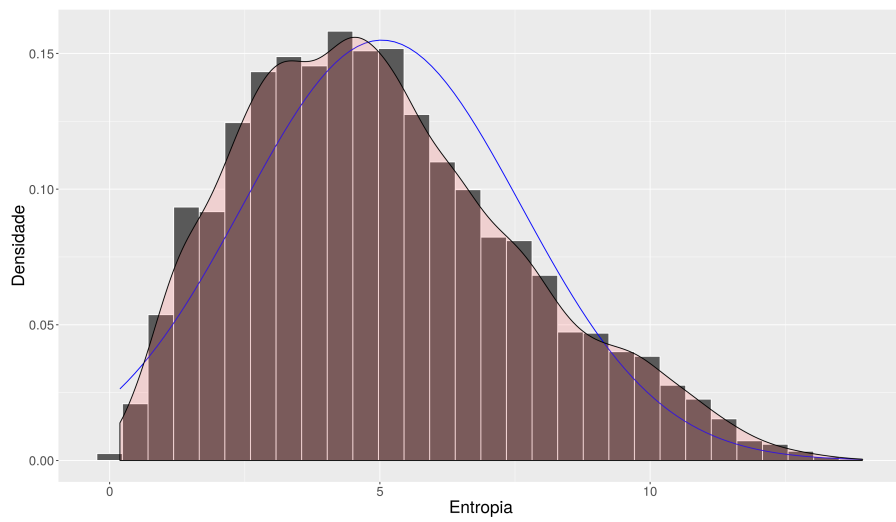
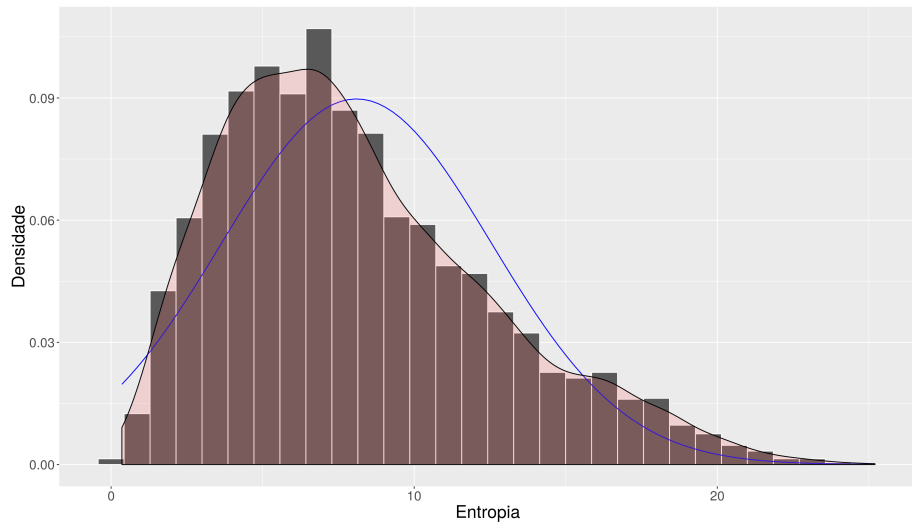
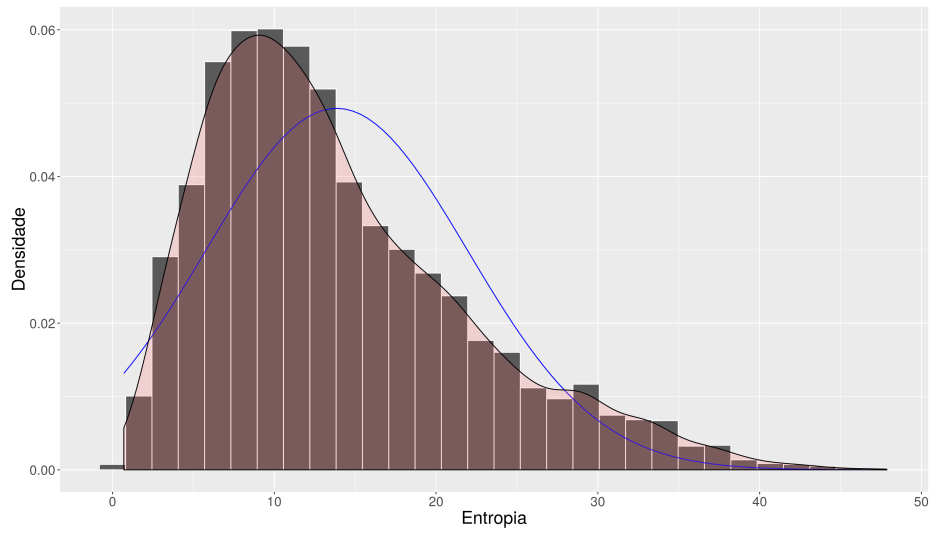


Figura 7.7 – Histogramas da família PF00135, clã CL0028 ($s = 0.1, 0.3, 0.5, 0.7, 0.9$ e 1.0) – probabilidade composta



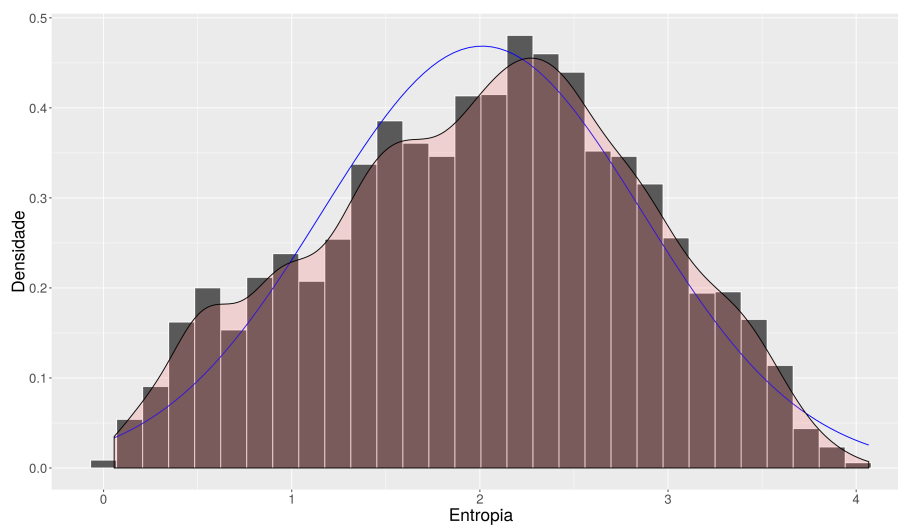
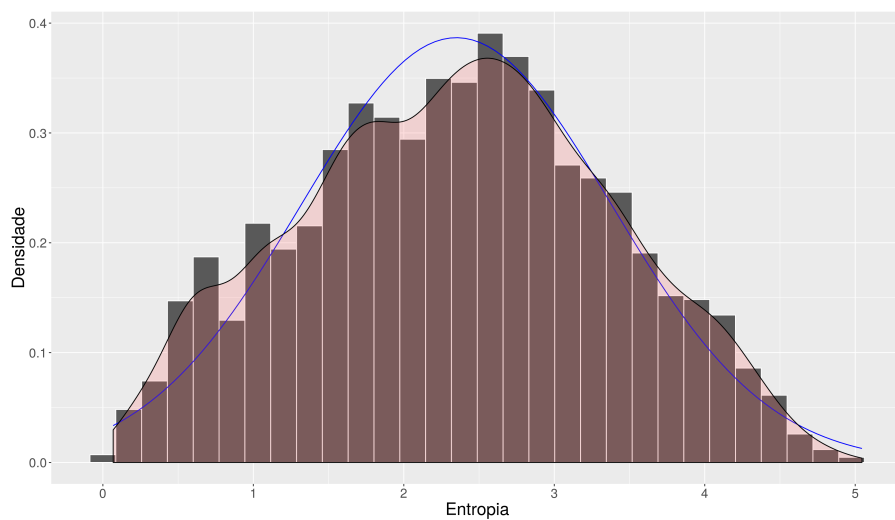
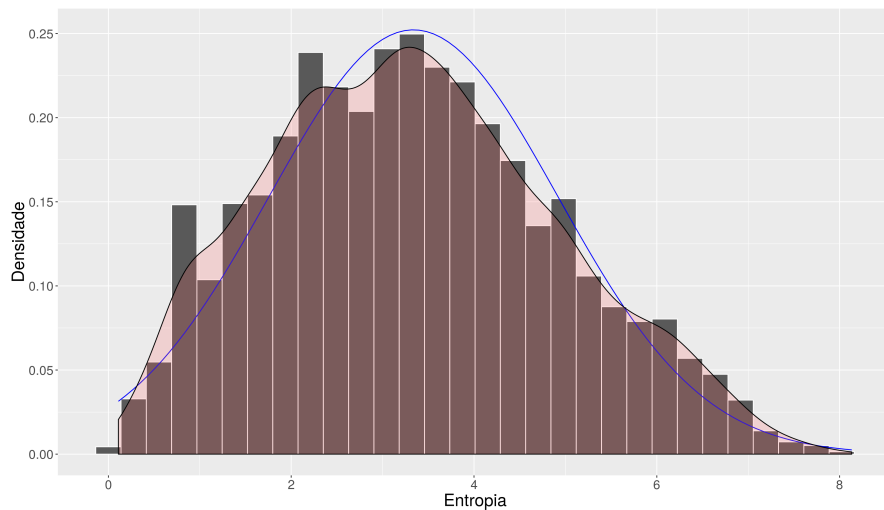
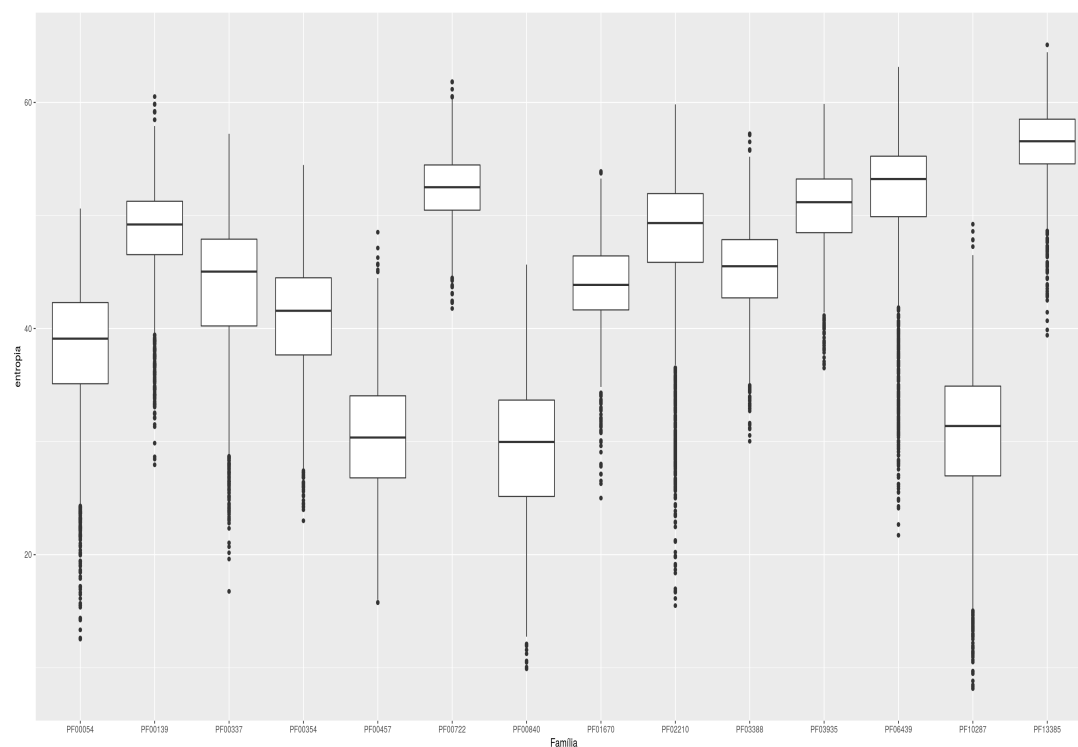


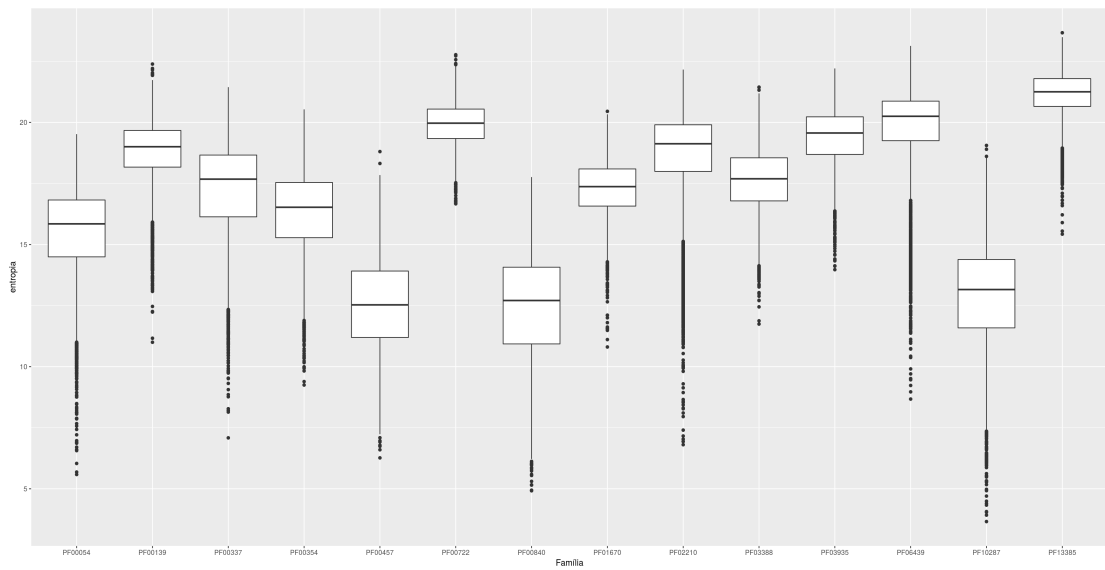
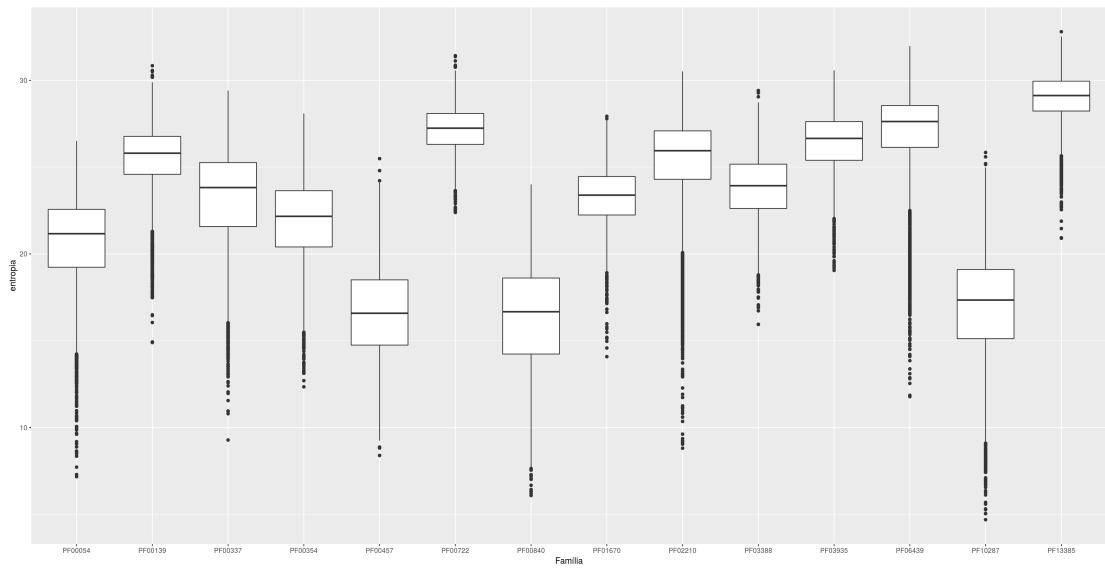
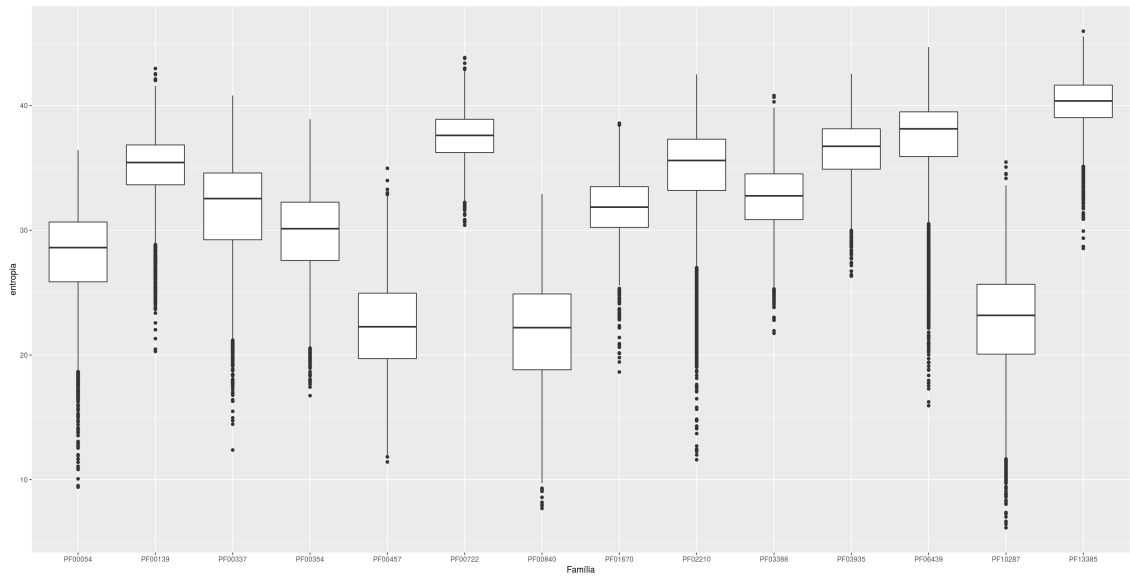
Figura 7.8 – Histogramas da família PF06850, clã CL0028 ($s = 0.1, 0.3, 0.5, 0.7, 0.9$ e 1.0) – probabilidade composta

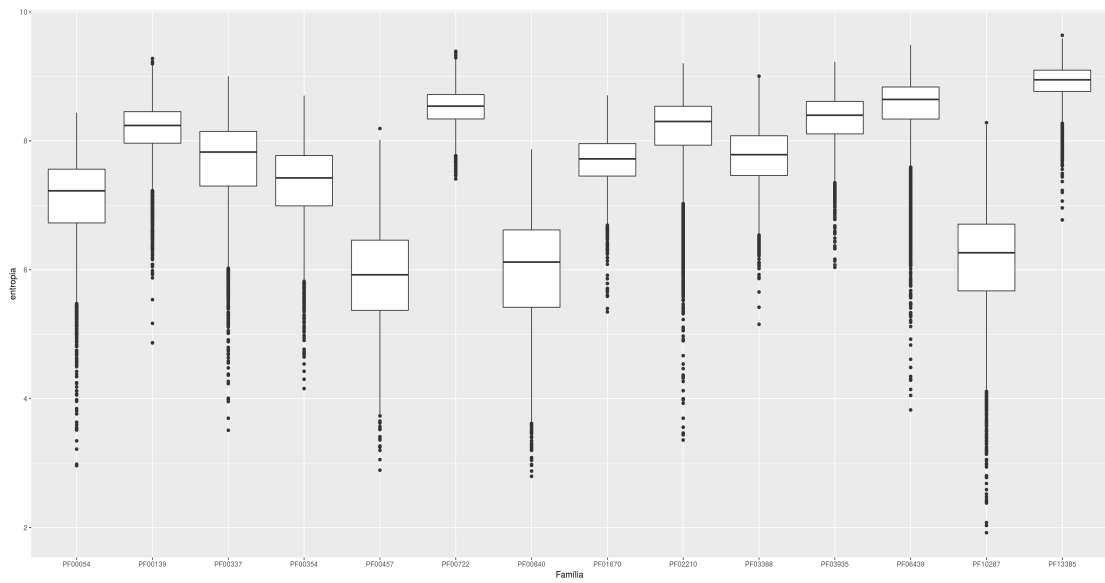
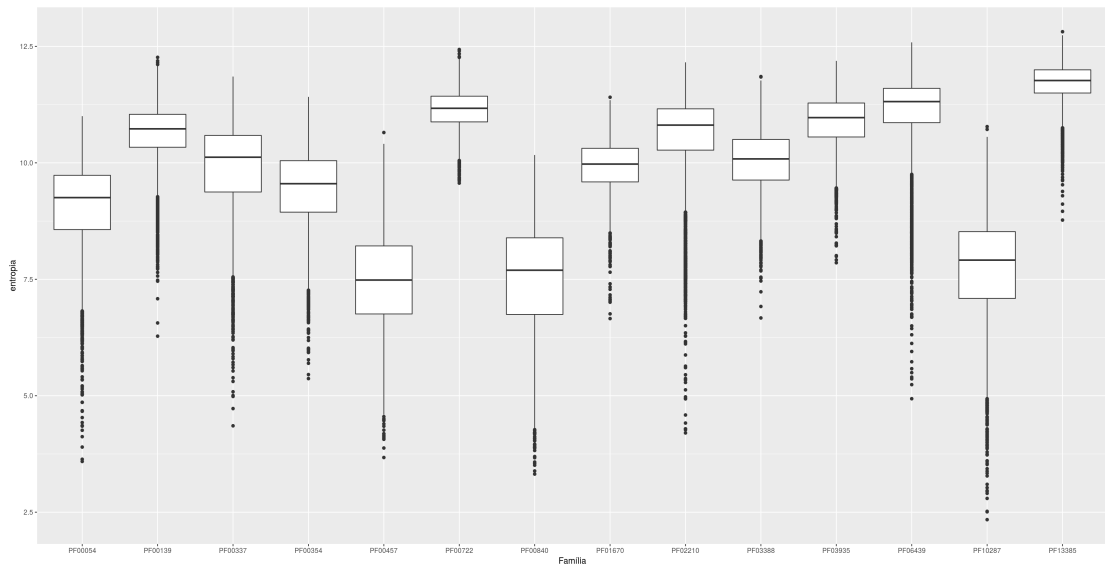
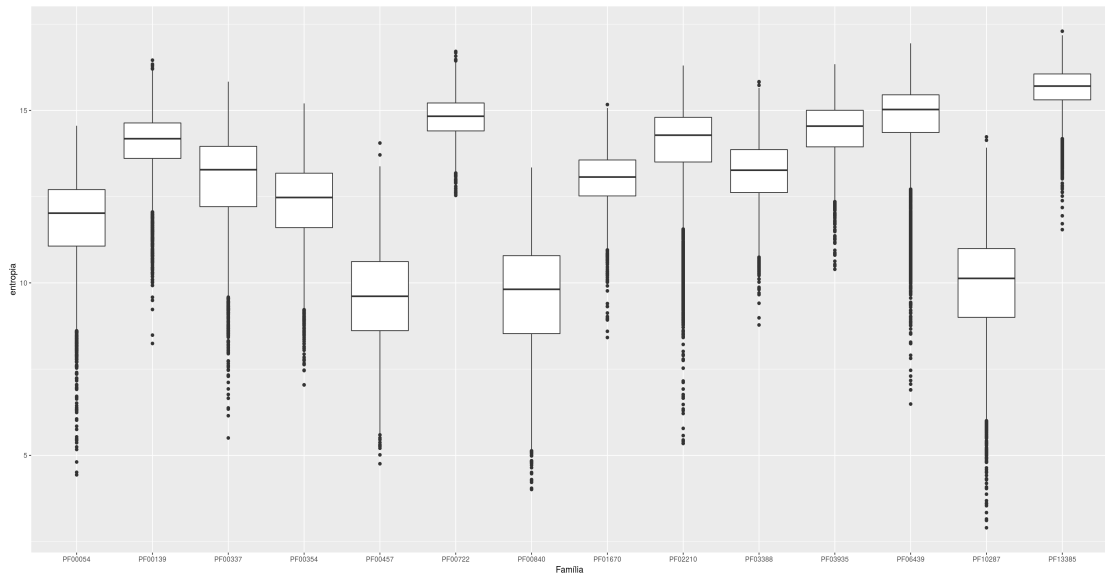
O boxplot foi utilizado neste trabalho para explorar e comparar as características das distribuições de entropia das famílias em um clã, como o nível de posição central, simetria da distribuição, amplitude de variação da distribuição e presença de dados extremos.

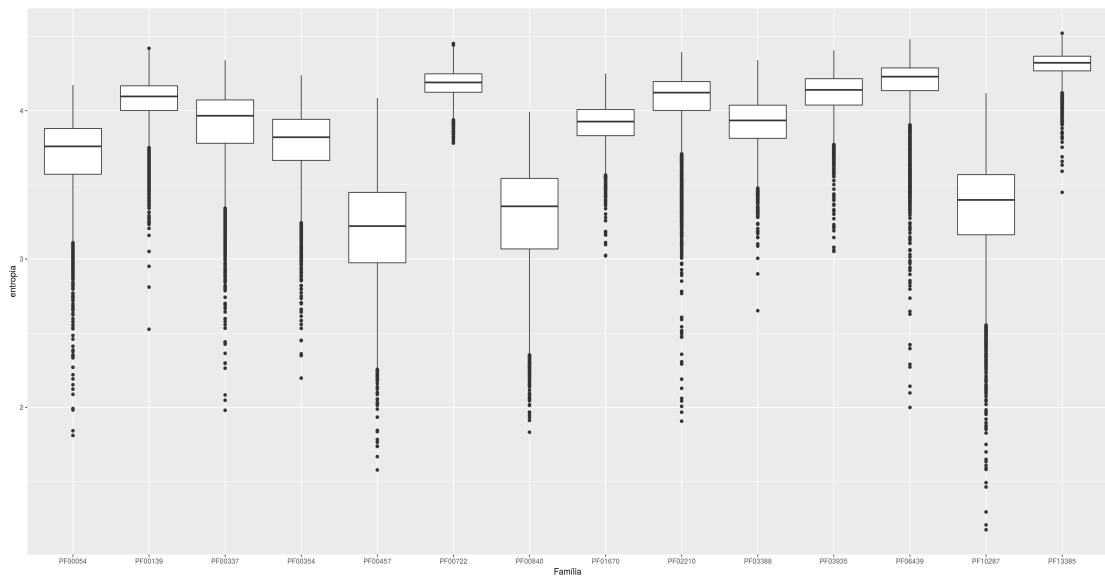
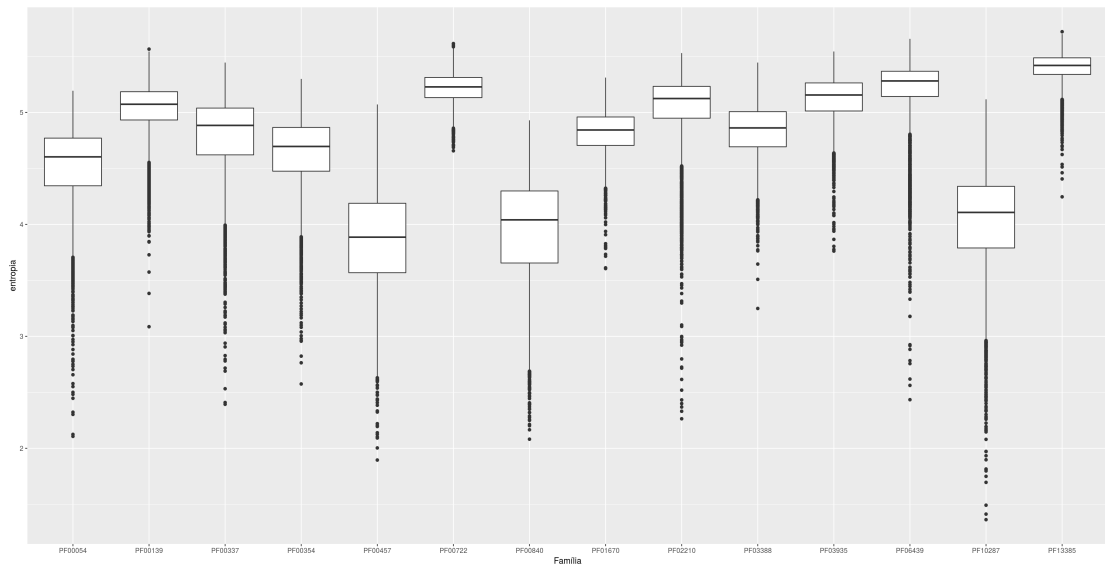
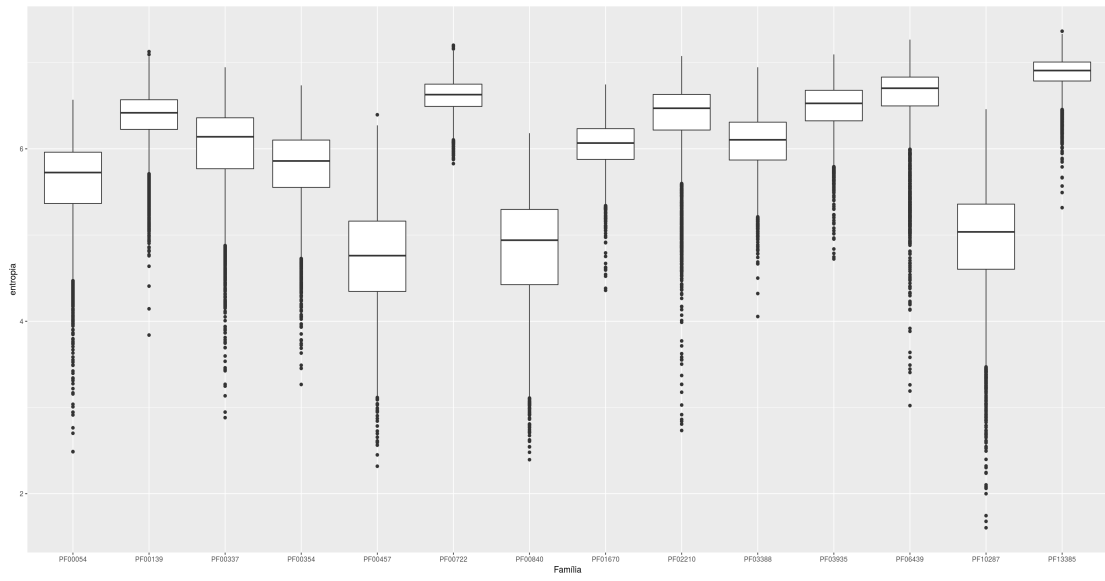
Foram feitos os boxplots de probabilidade conjunta para todas as famílias de onze clãs (CL0004, CL0013, CL0014, CL0015, CL0016, CL0020, CL0021, CL0023, CL0025, CL0026 e CL0027) e para diferentes valores do parâmetro s (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3), selecionados a partir dos Clãs obtidos nas janelas 100 x100.

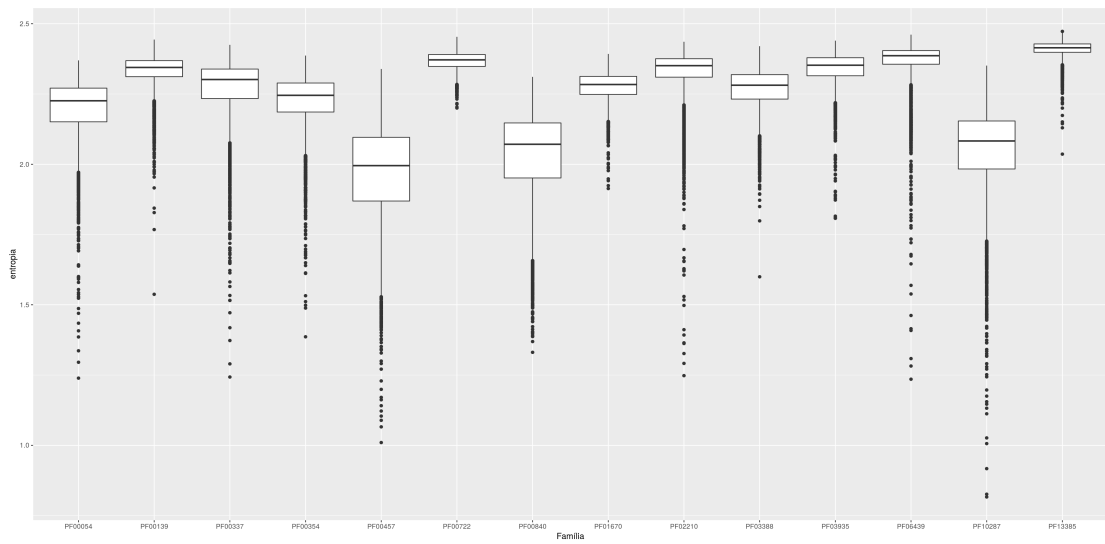
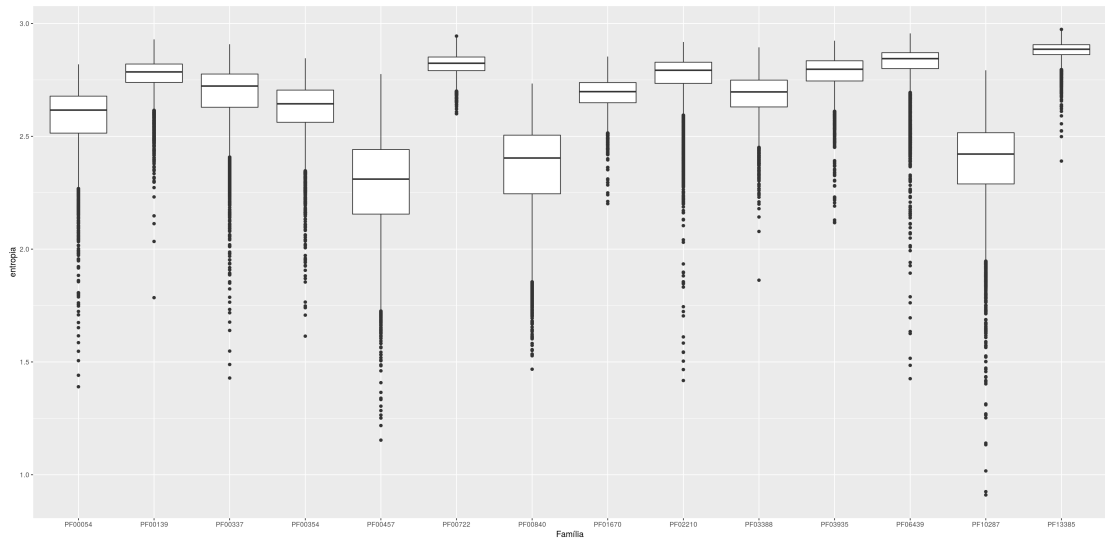
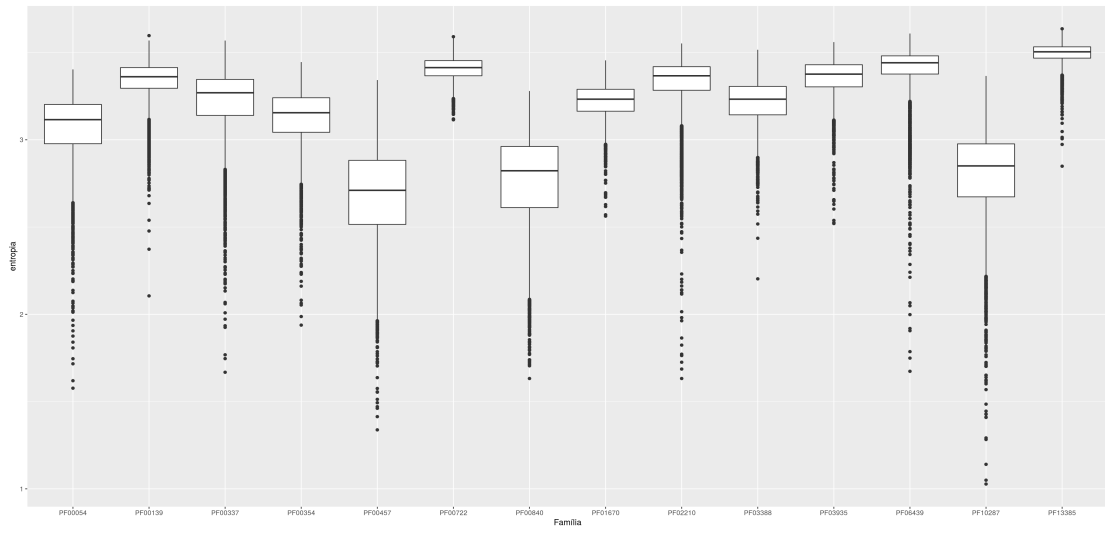
Em seguida serão apresentados os boxplots para as distribuições de entropias das famílias de cada um dos clãs e parâmetros acima citados:











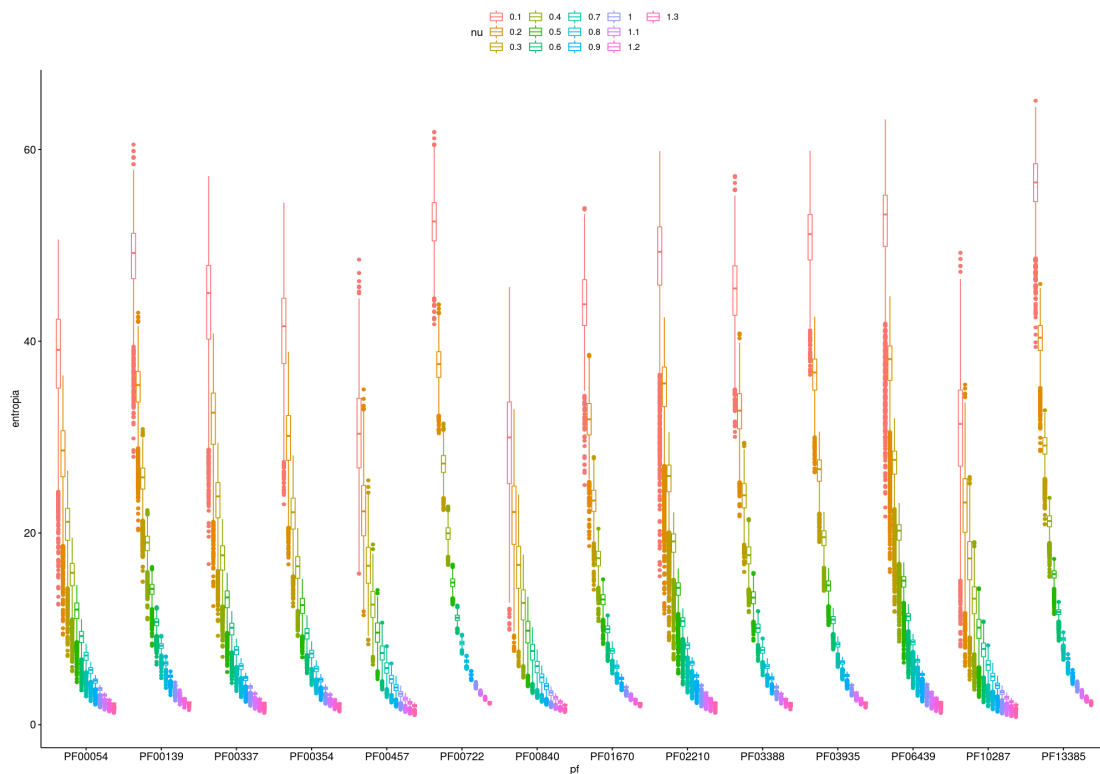
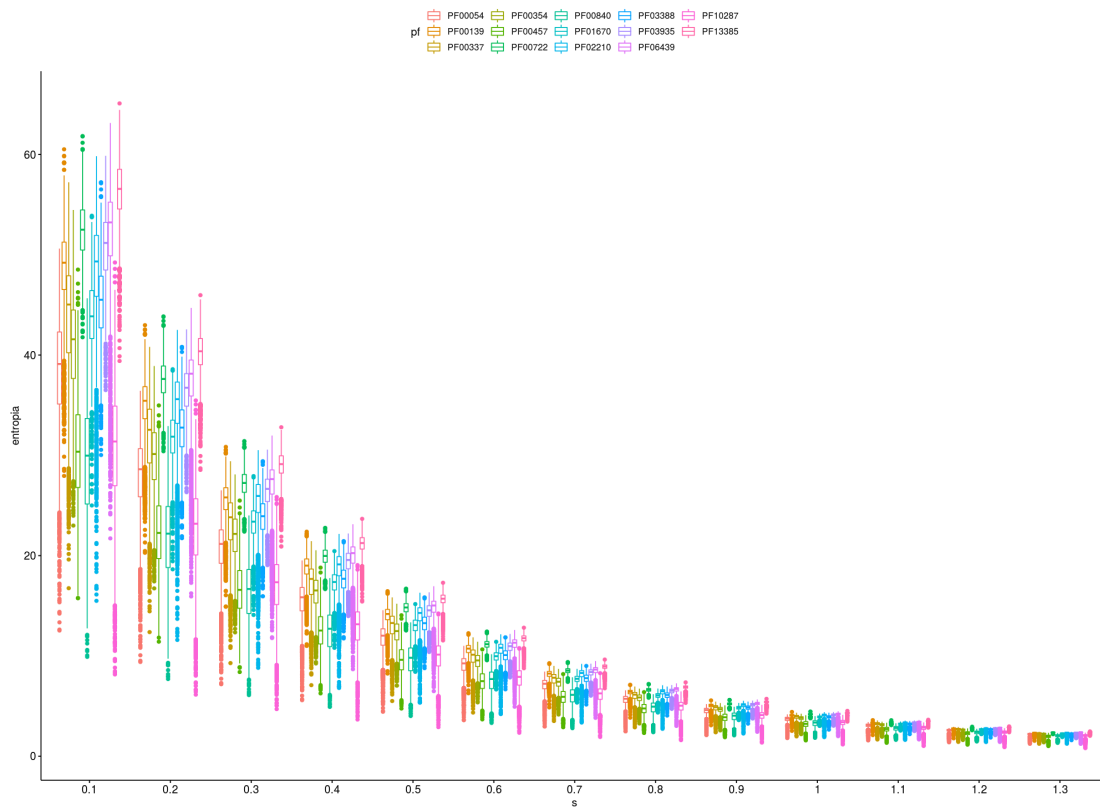
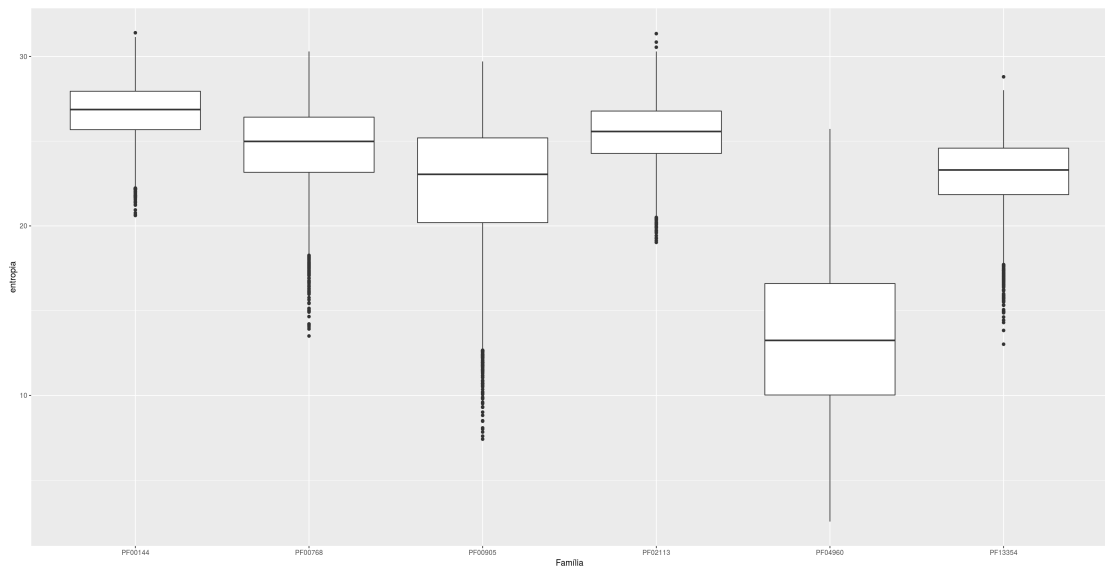
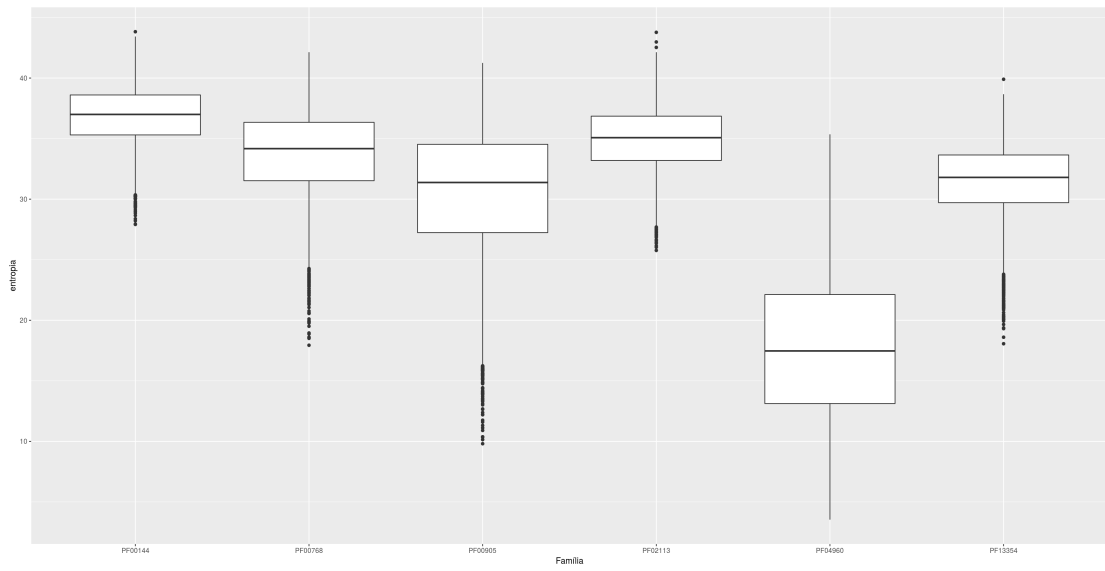
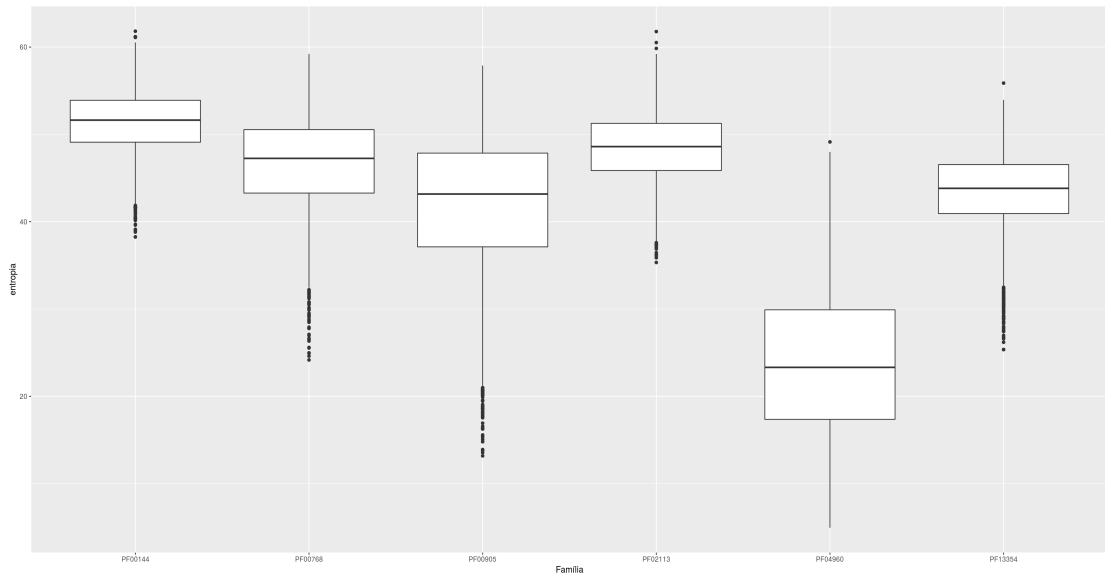
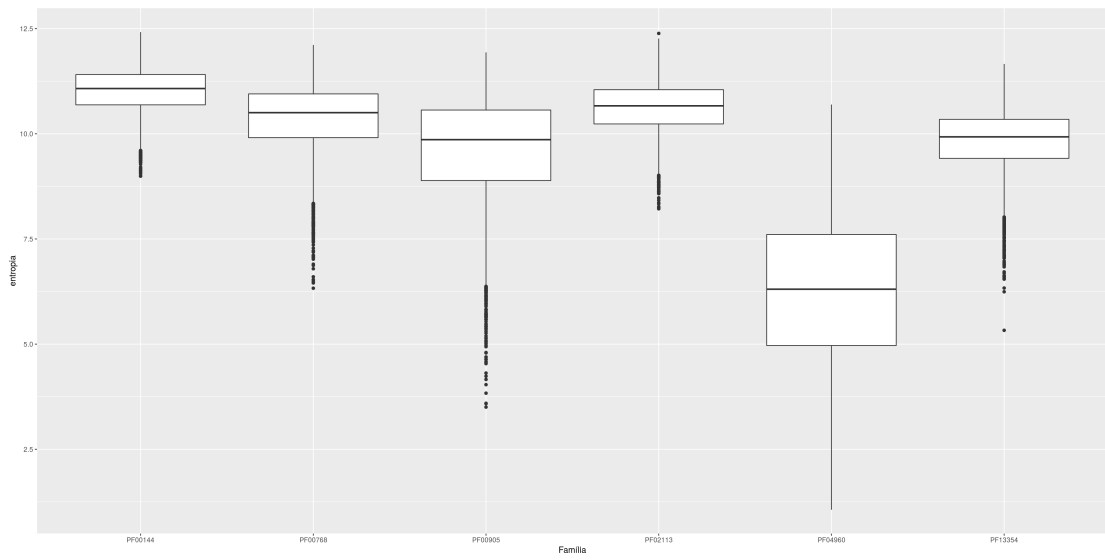
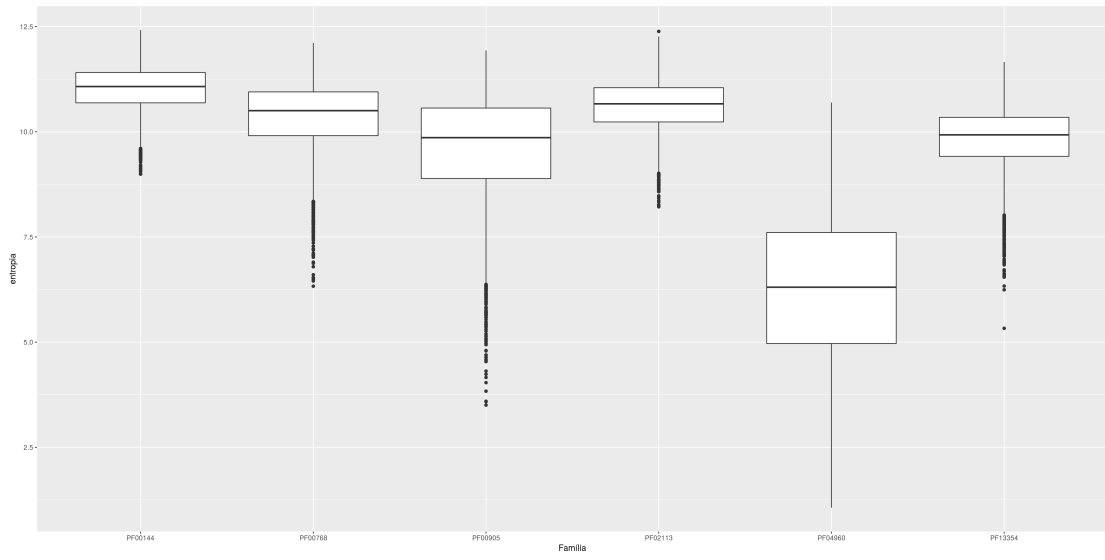
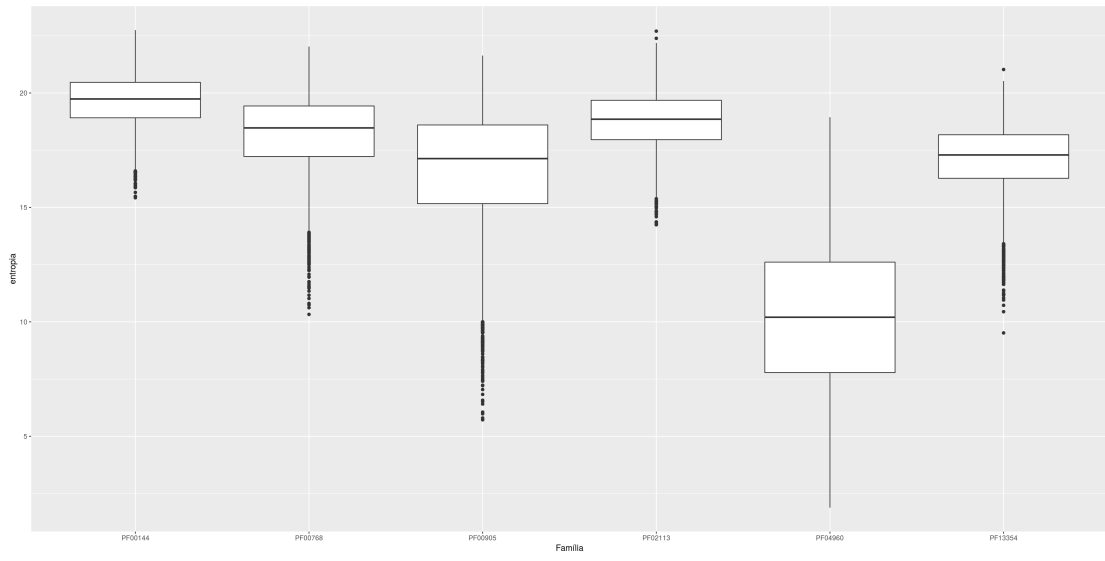
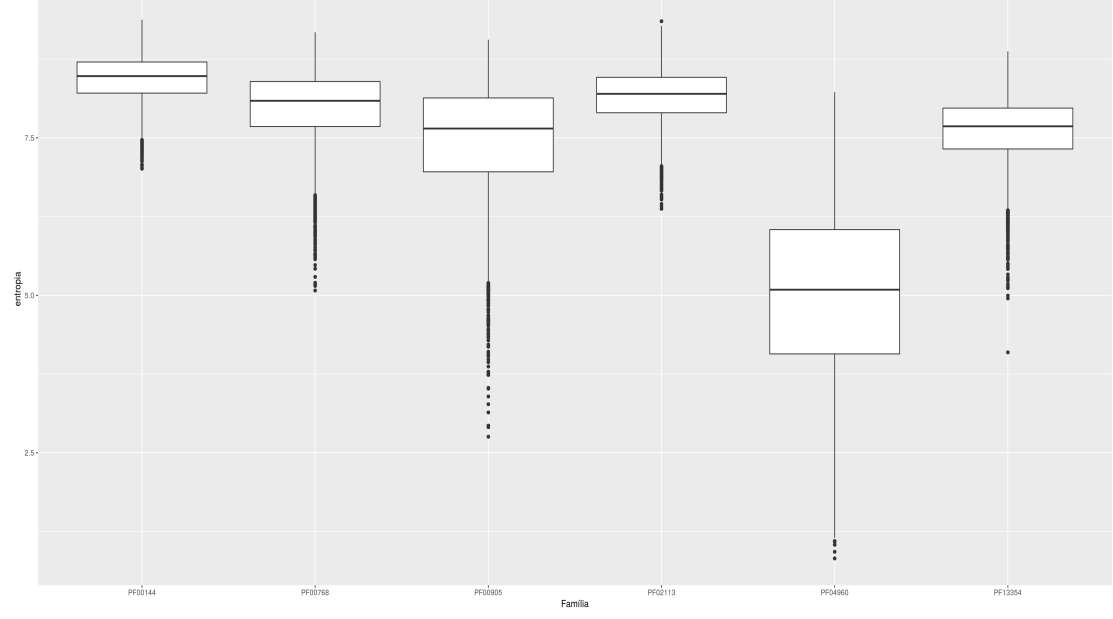
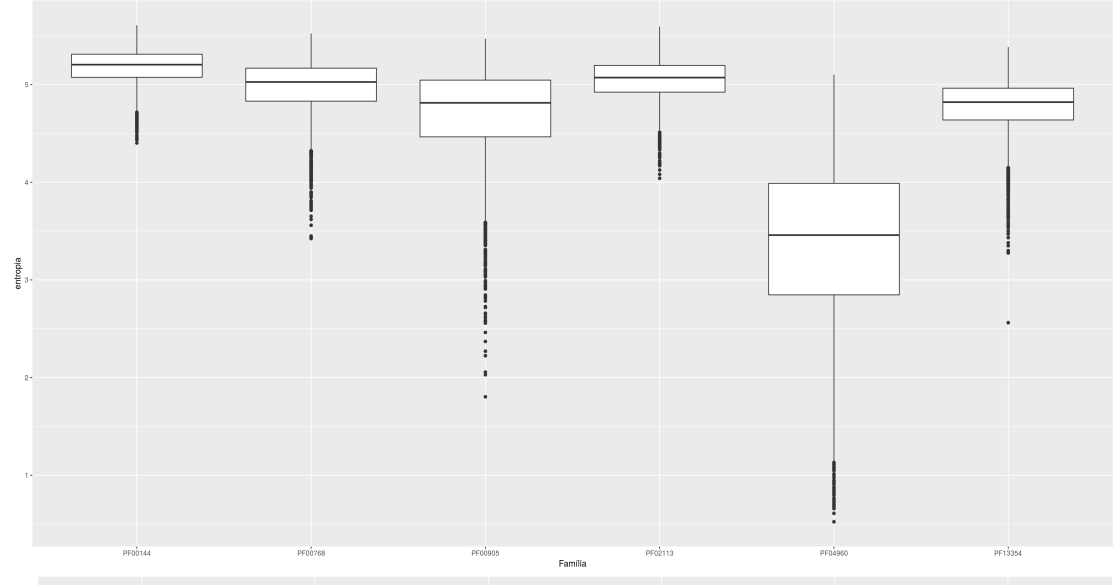
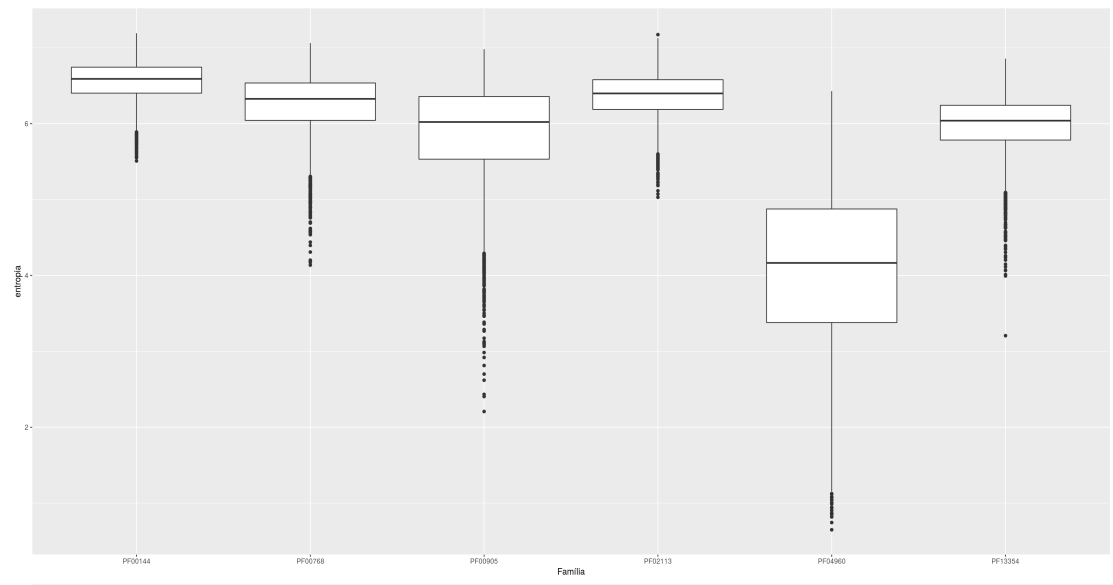
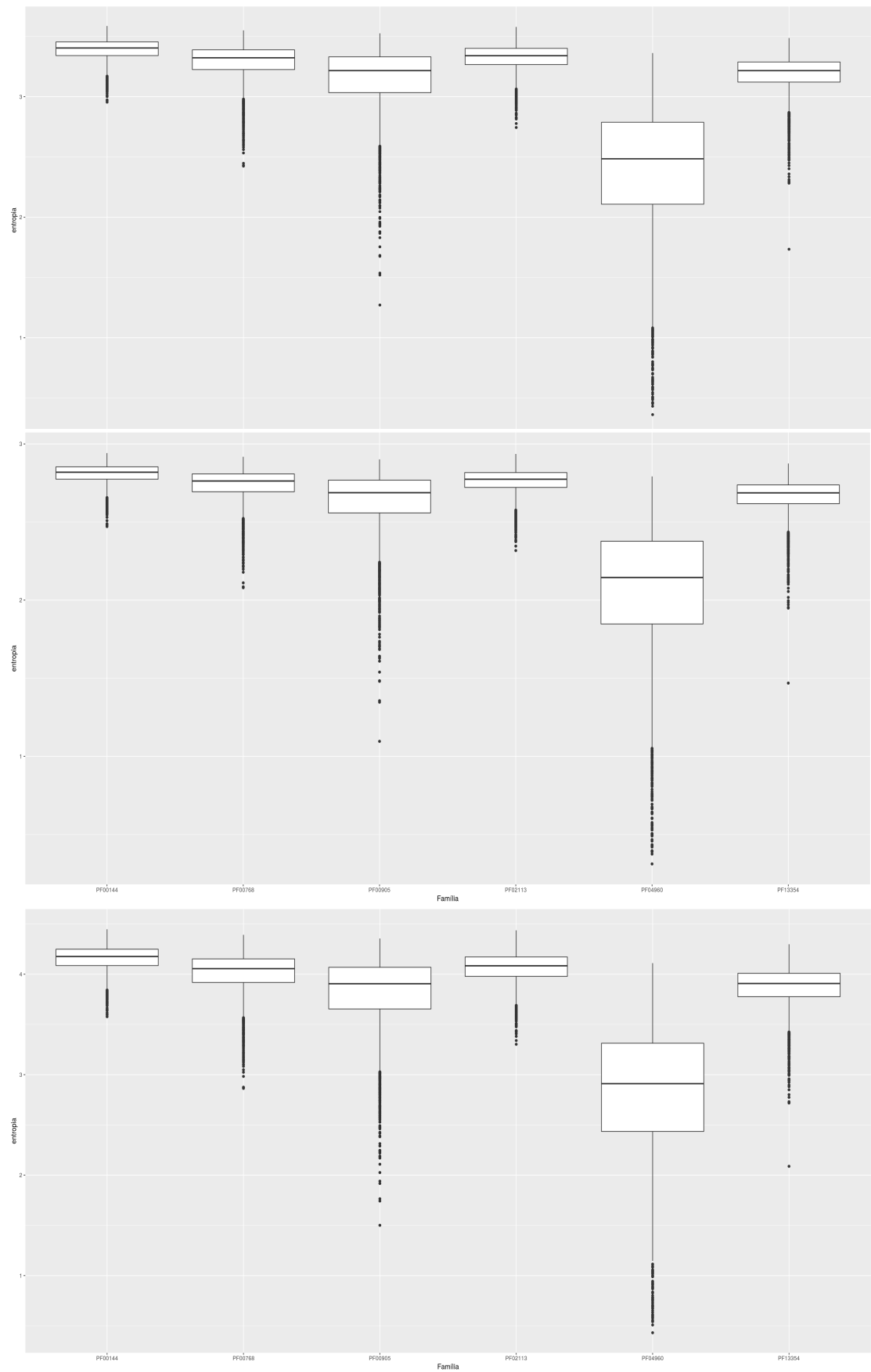


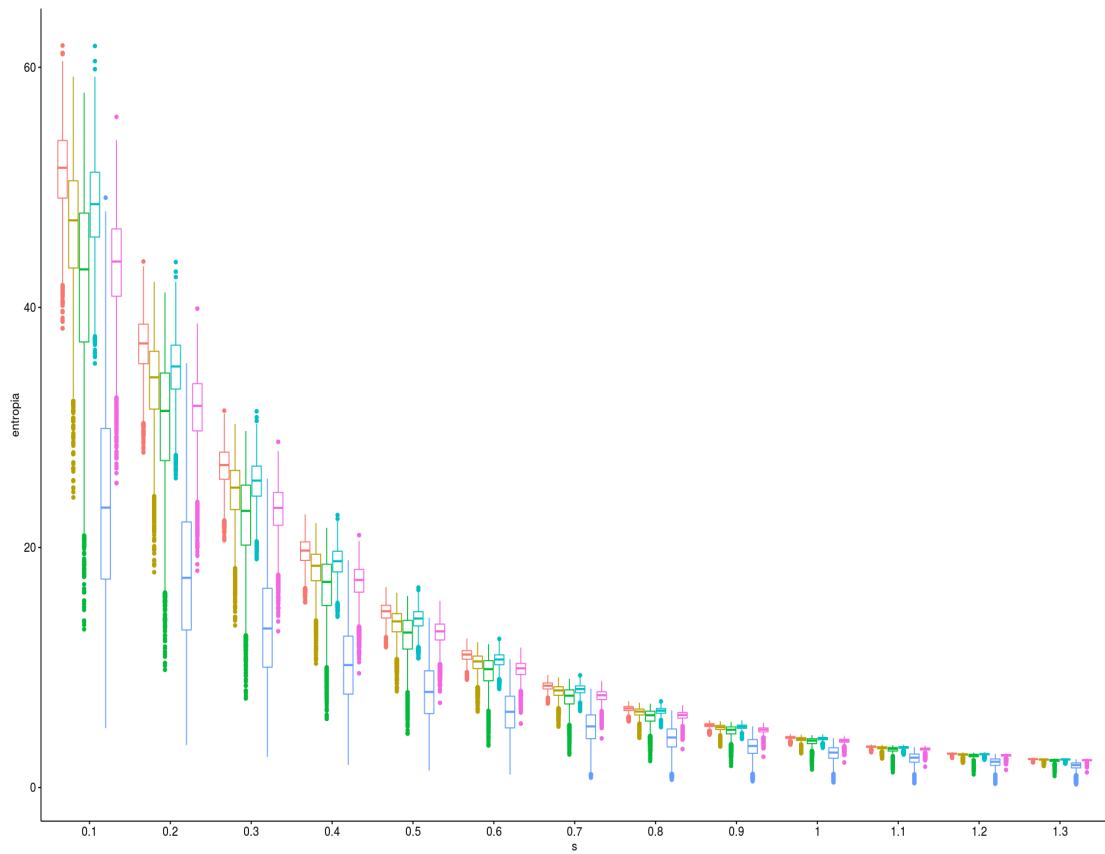
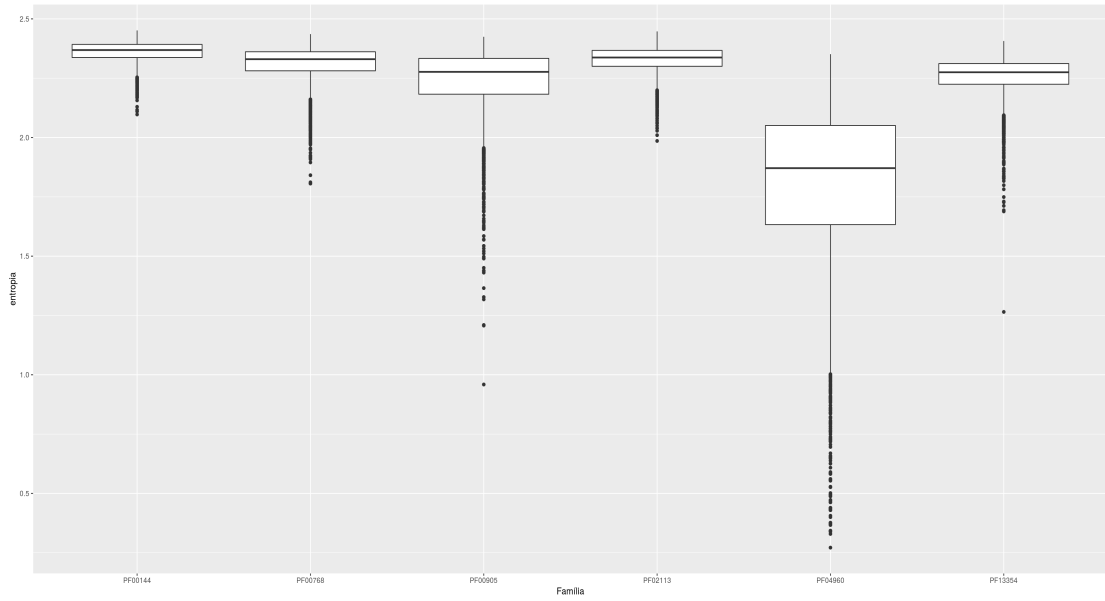
Figura 7.9 – Boxplots para as famílias do clã CL0004 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta











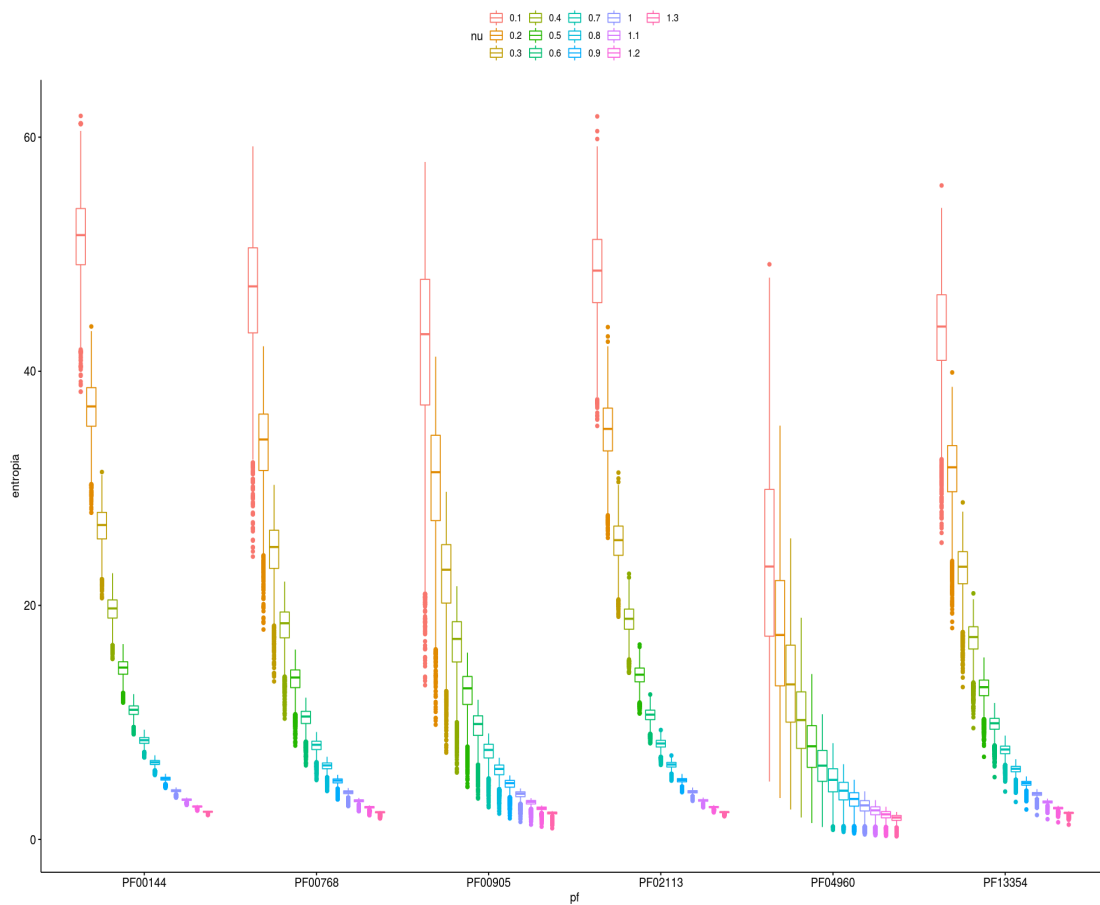
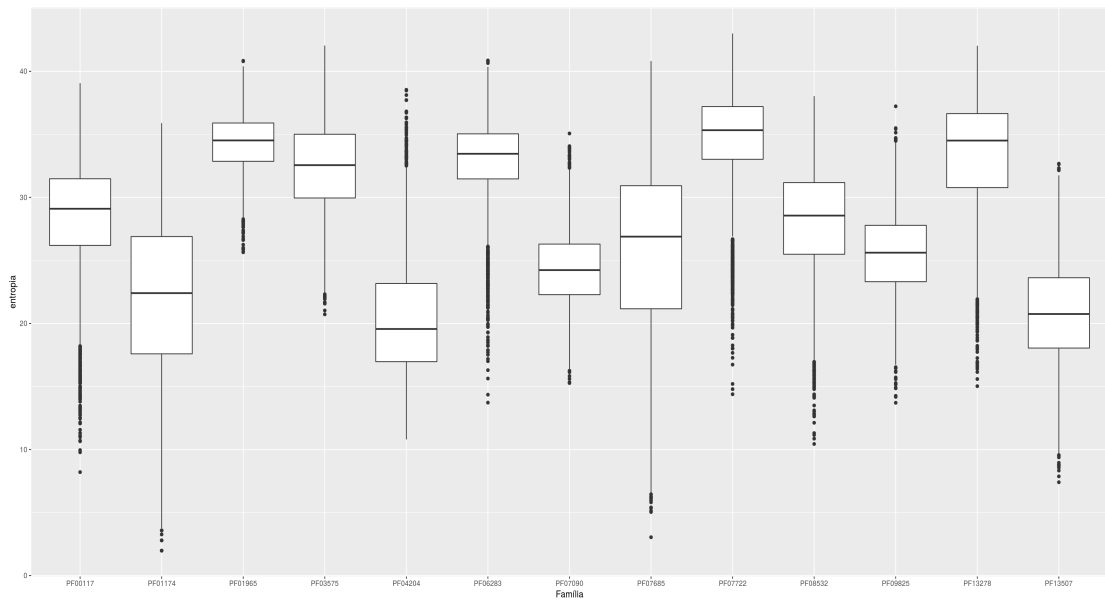
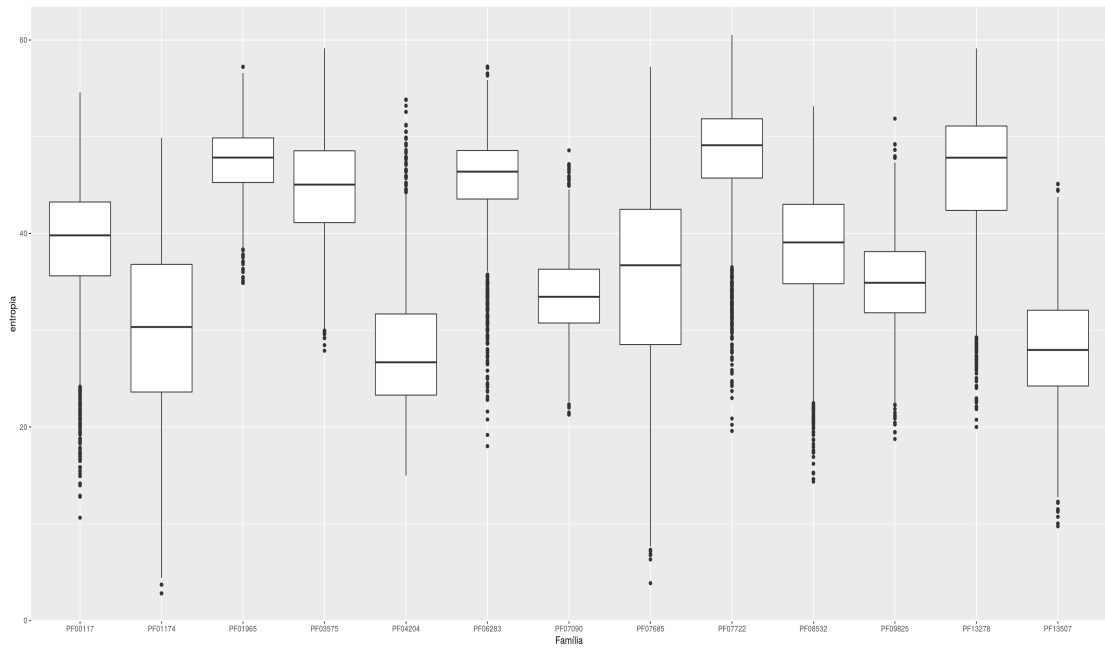
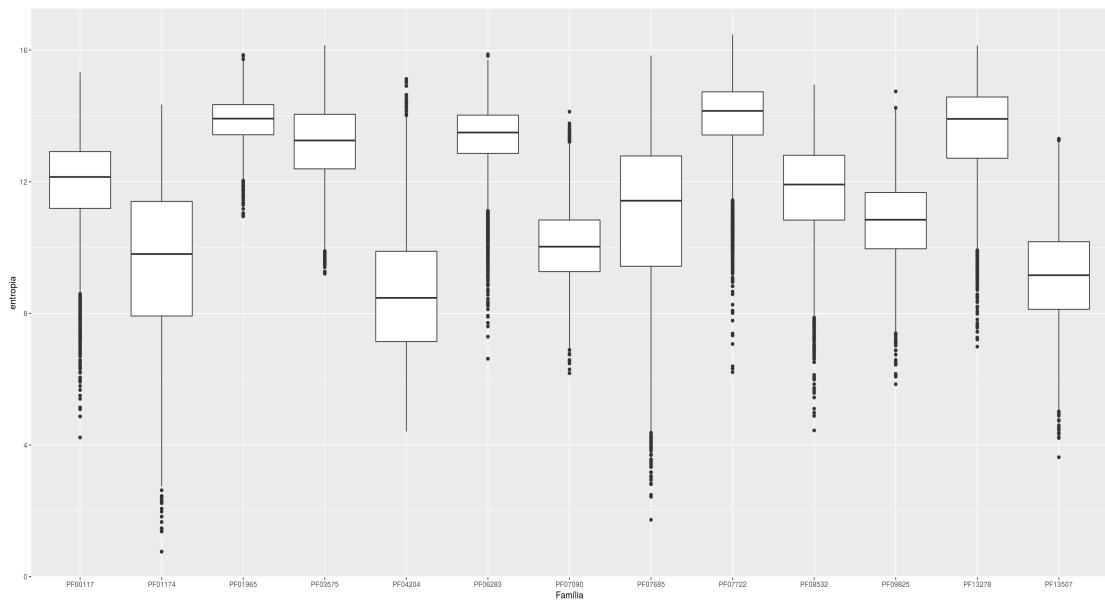
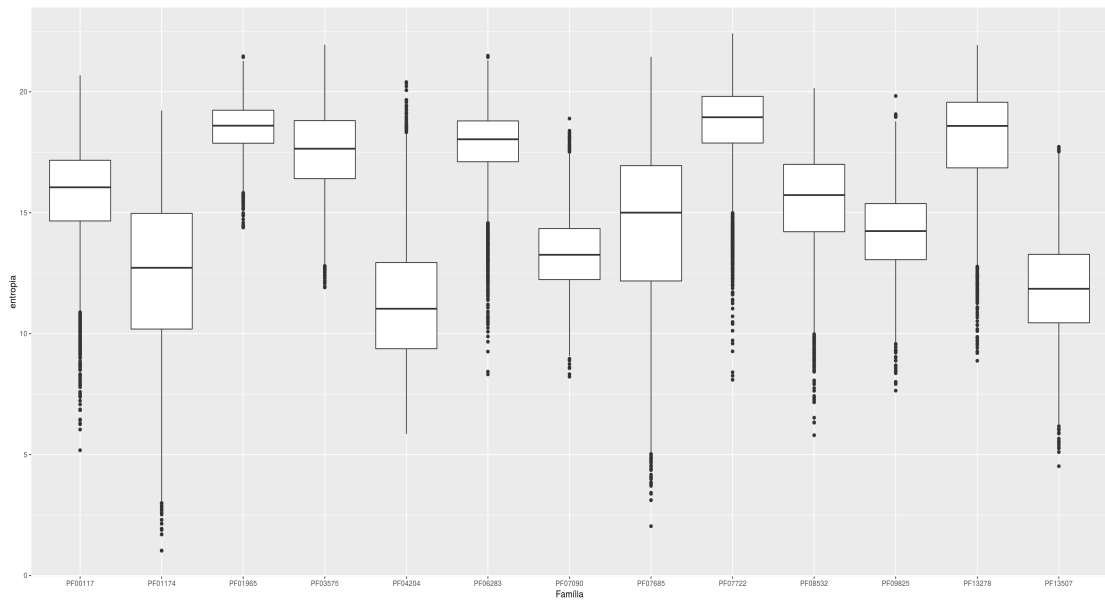
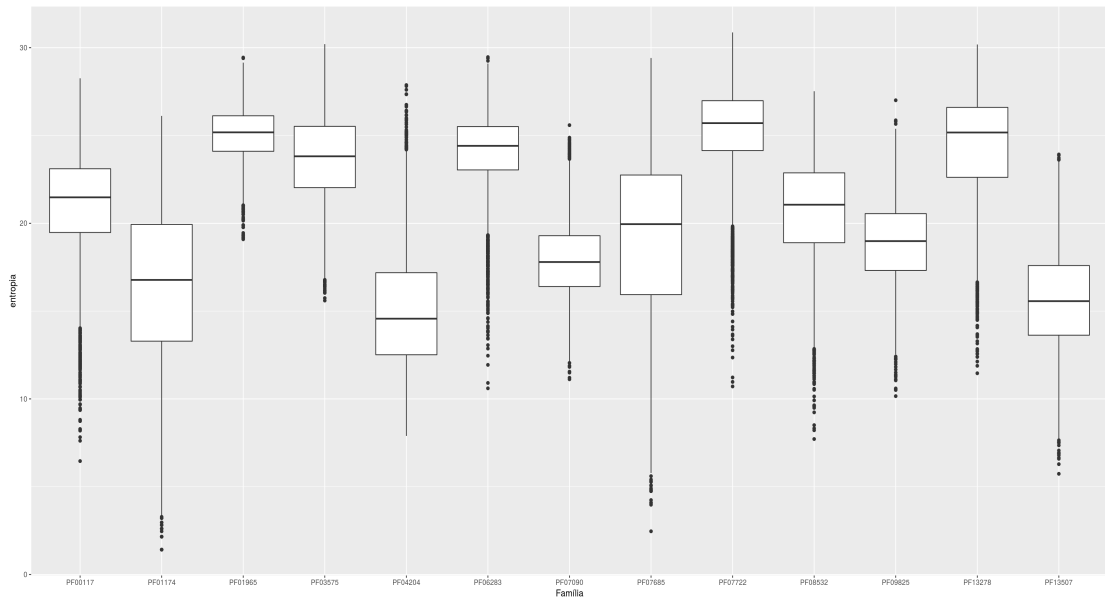
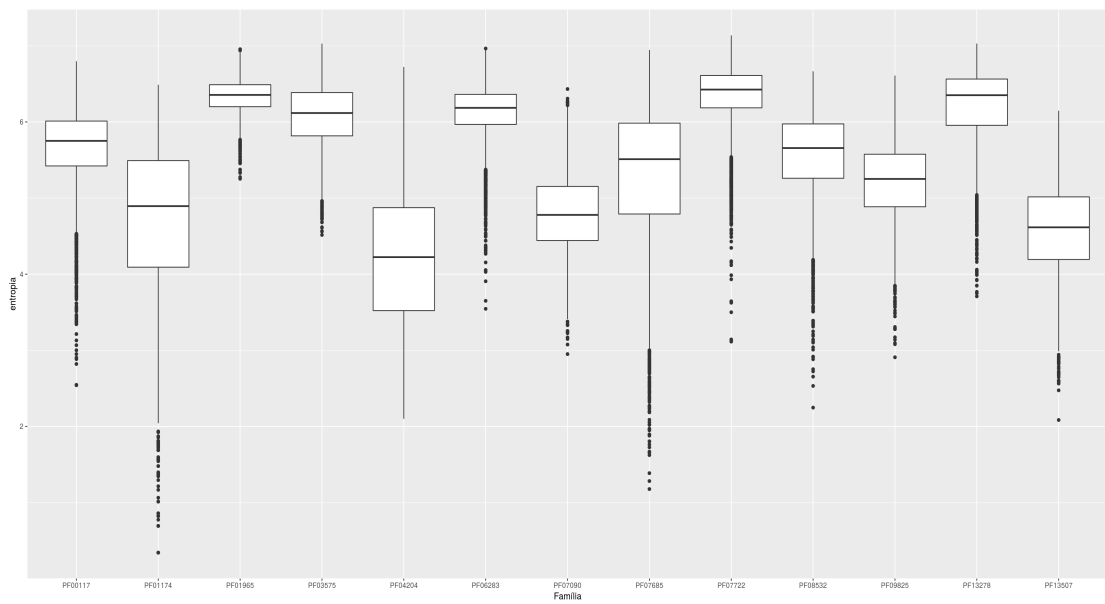
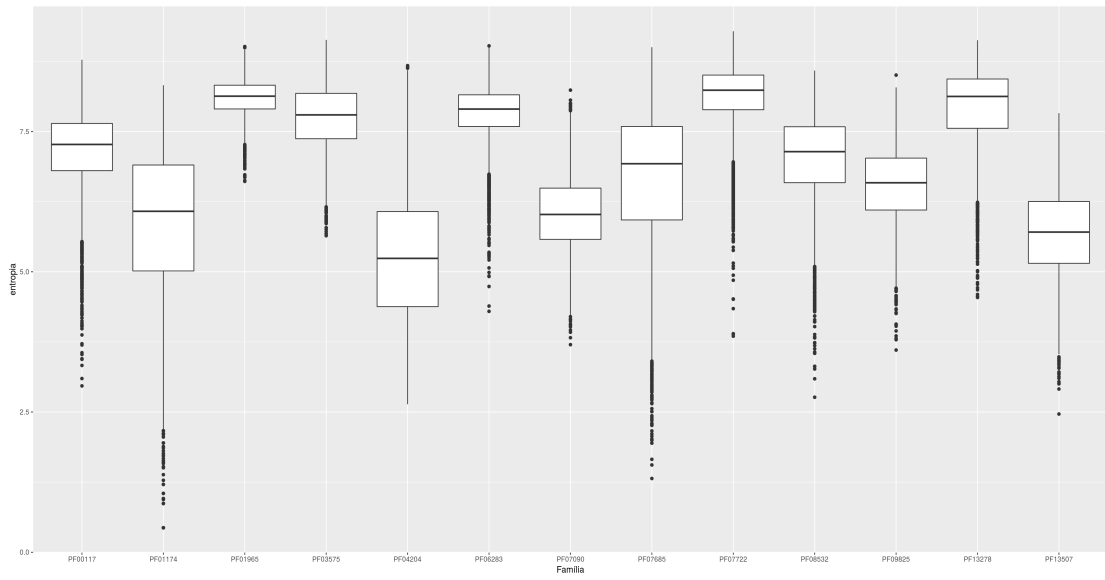
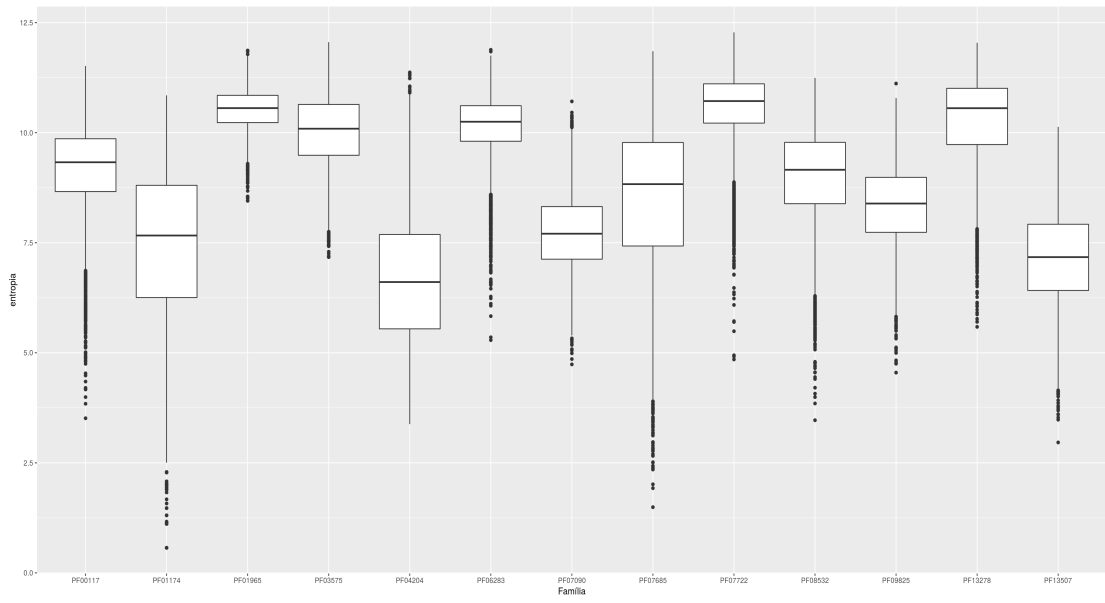
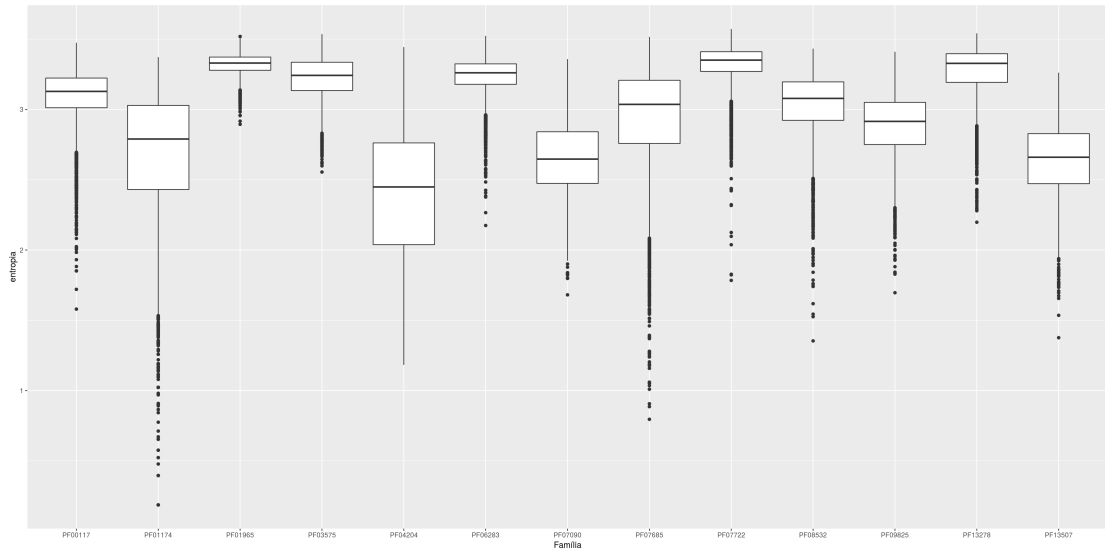
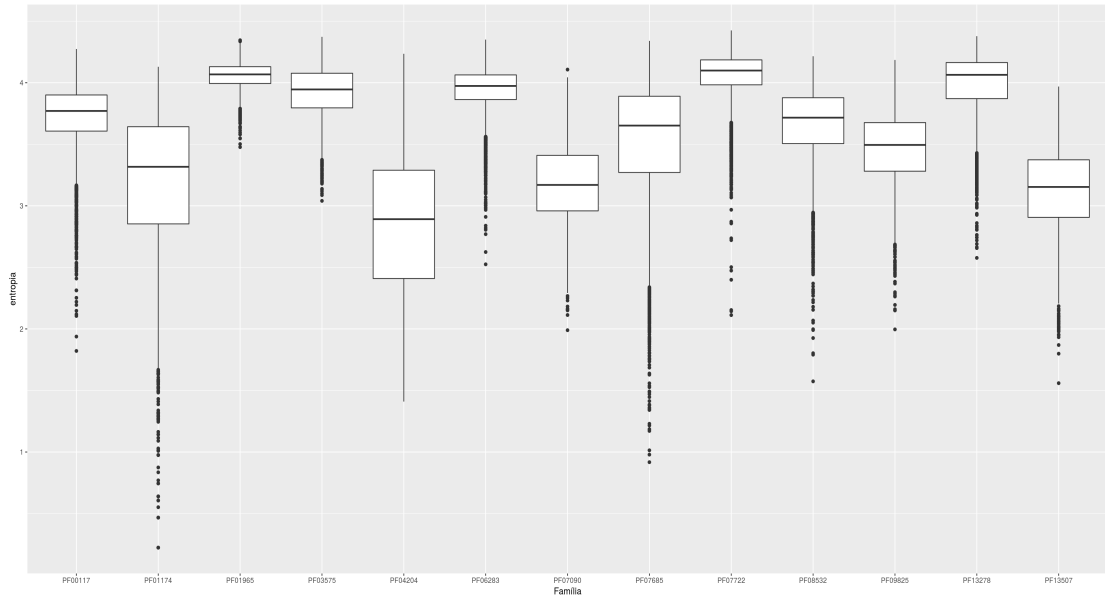
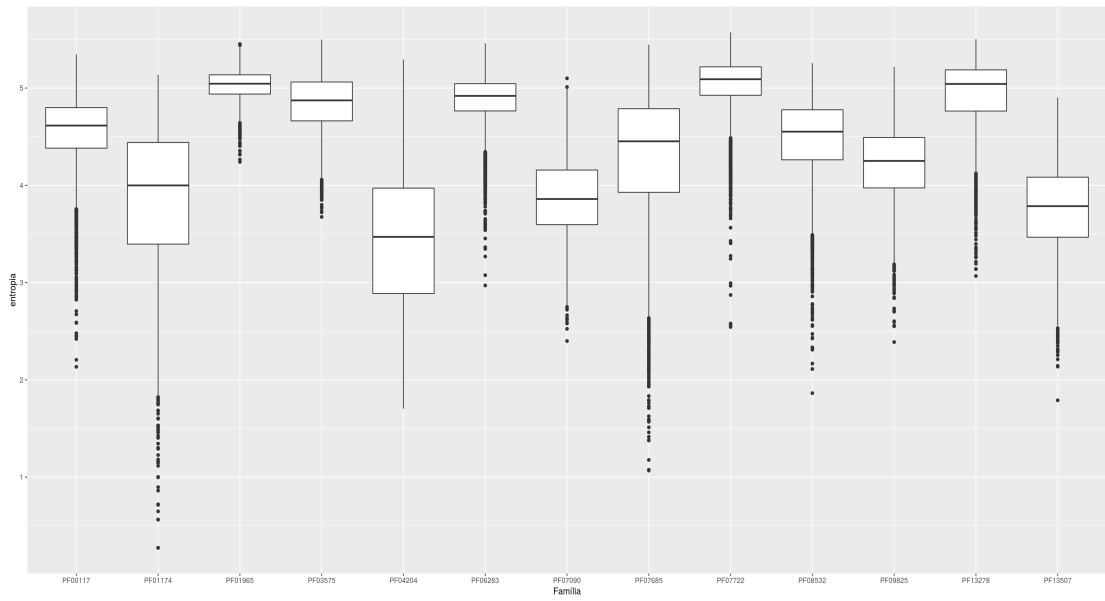


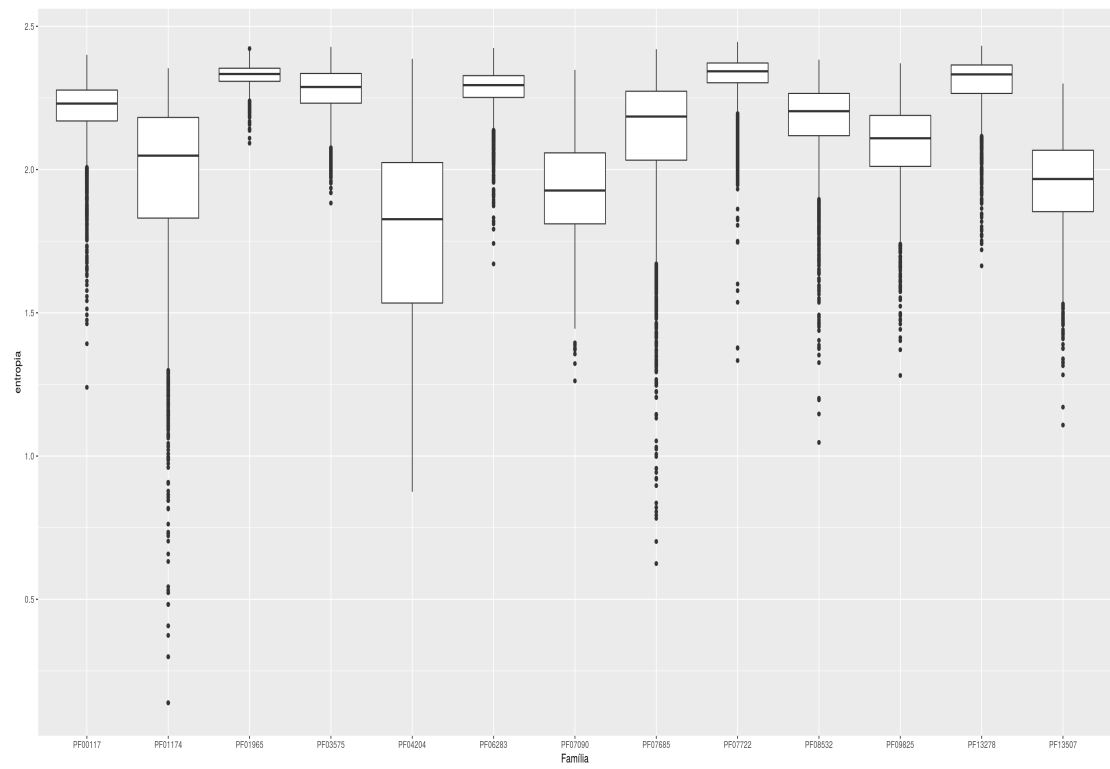
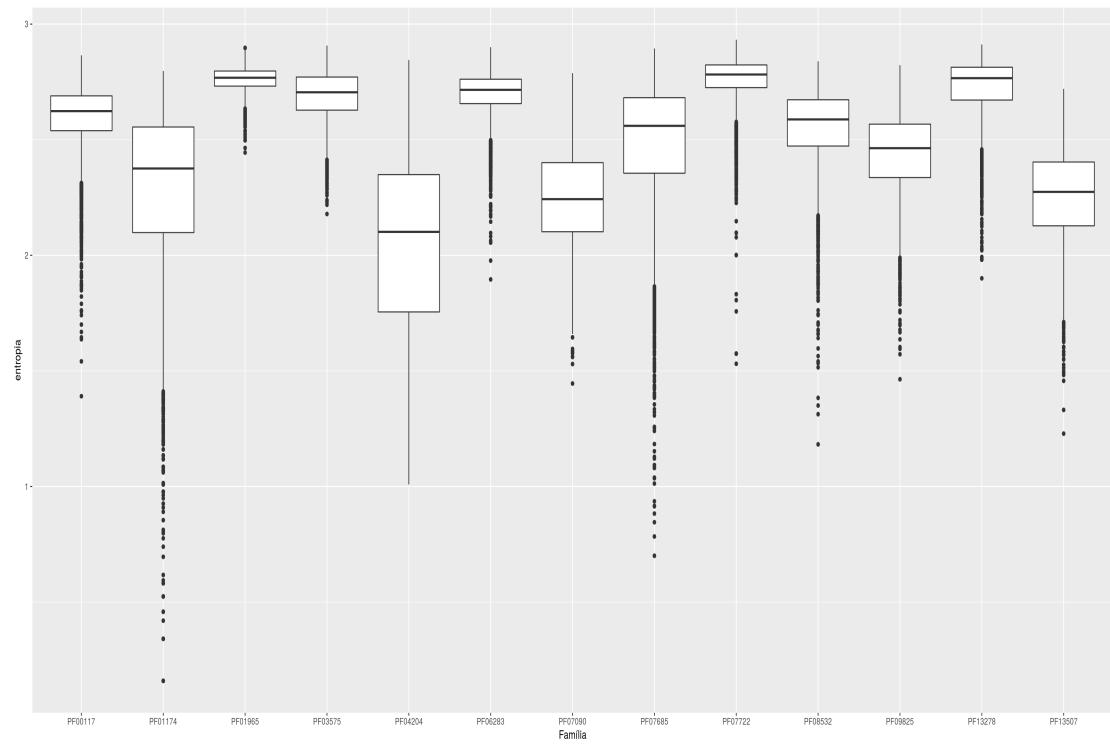
Figura 7.10 – Boxplots para as famílias do clã CL0013 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta











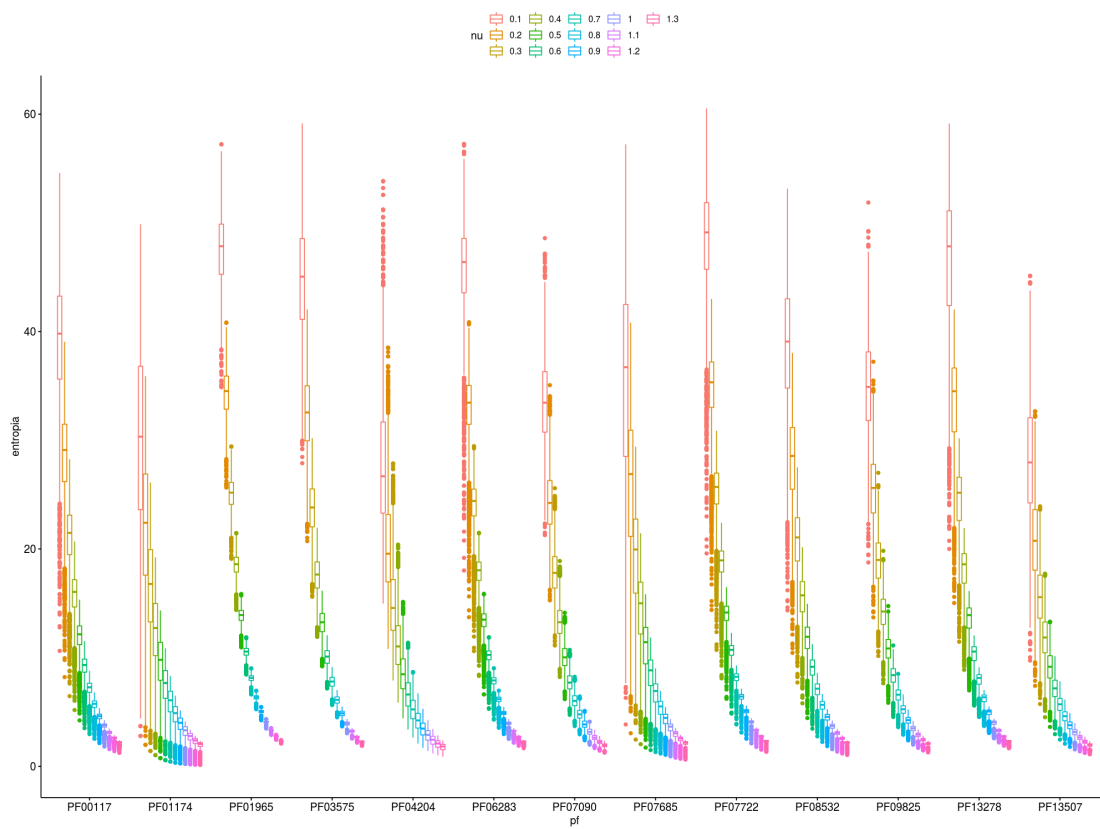
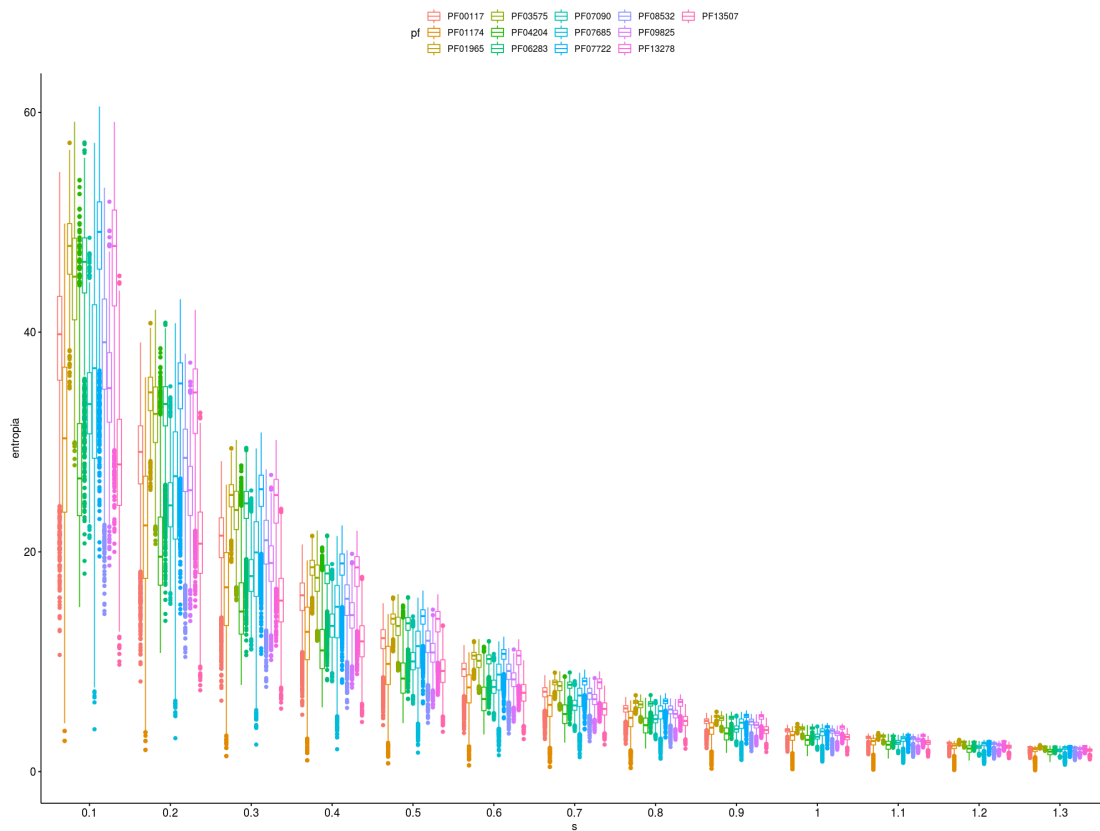
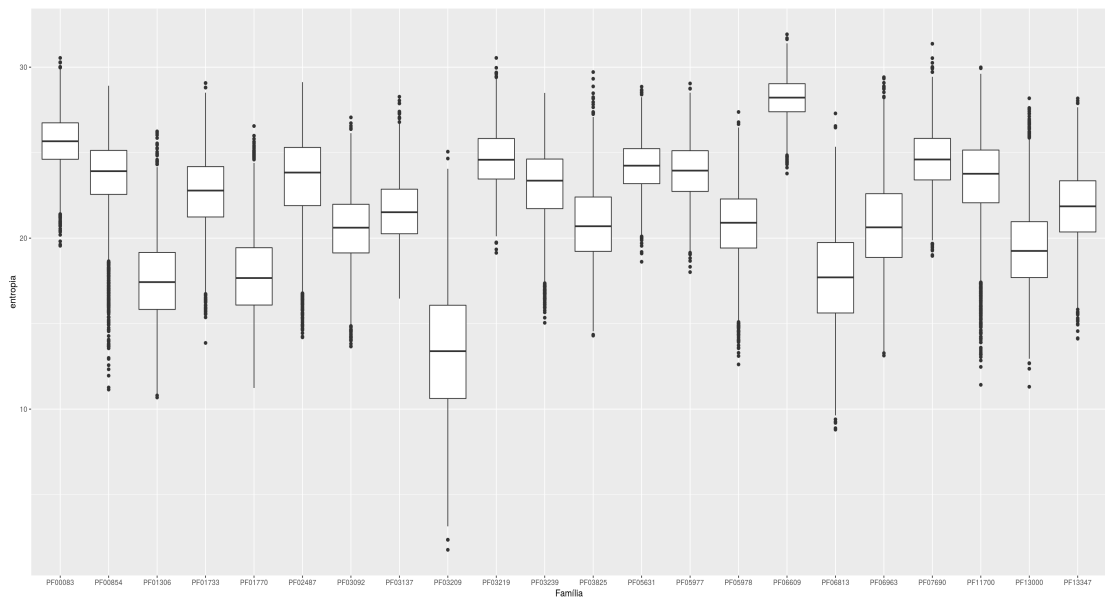
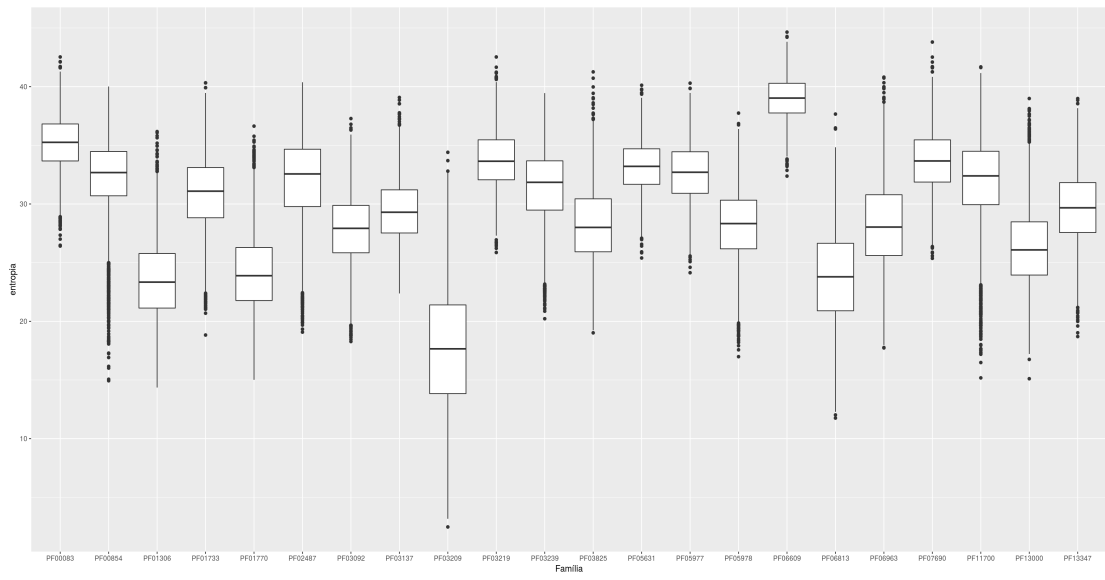
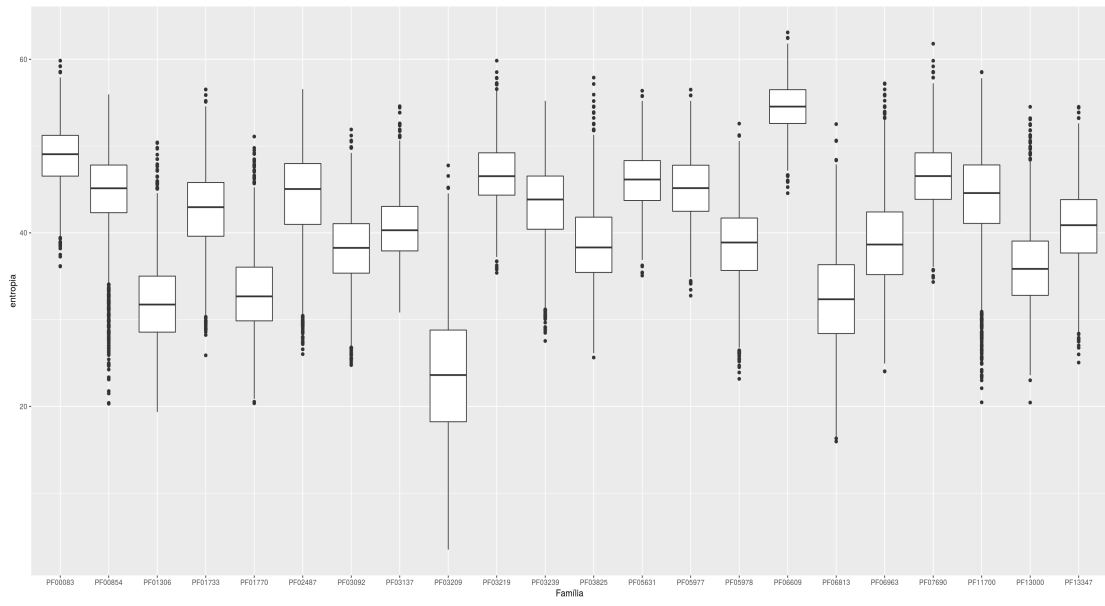
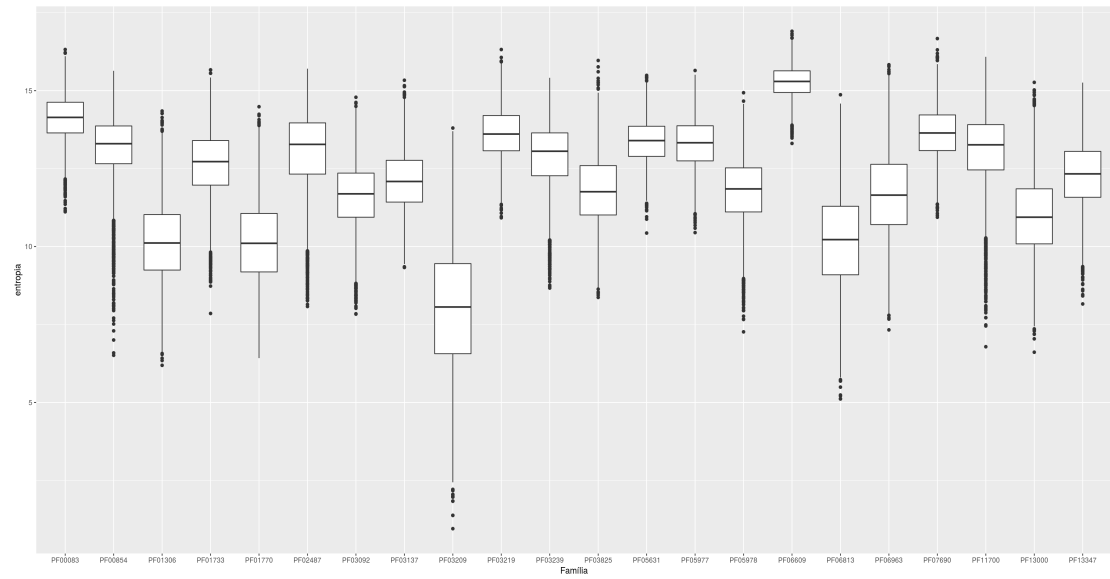
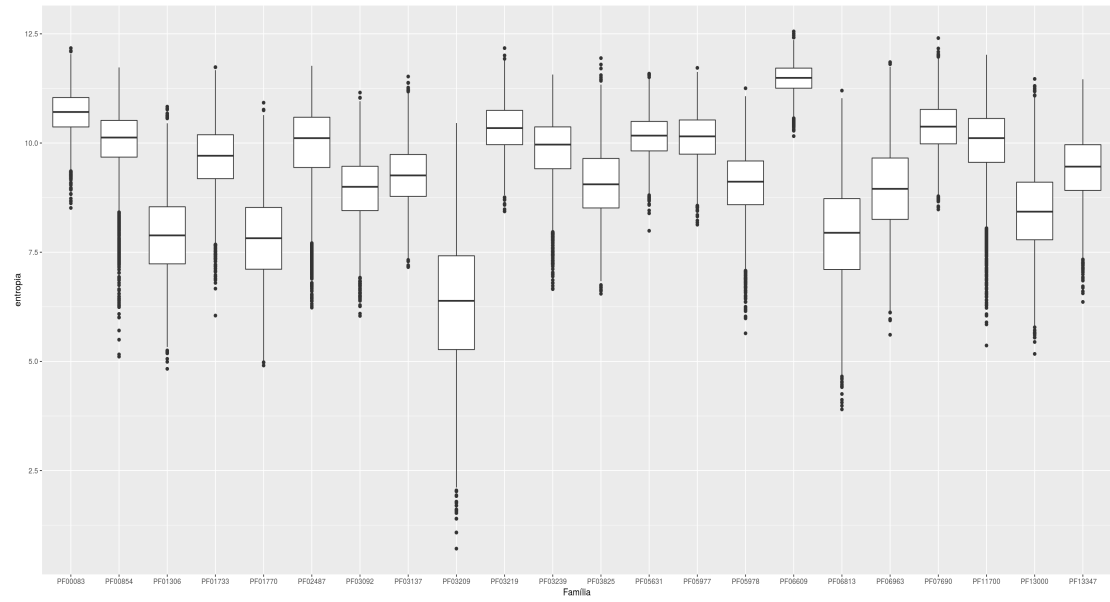
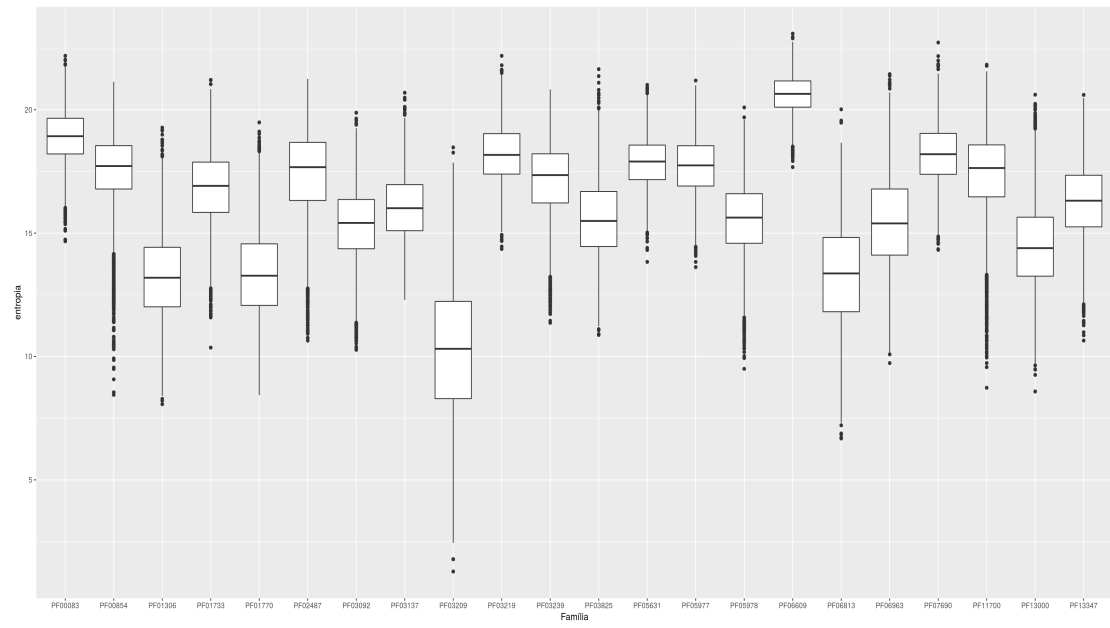
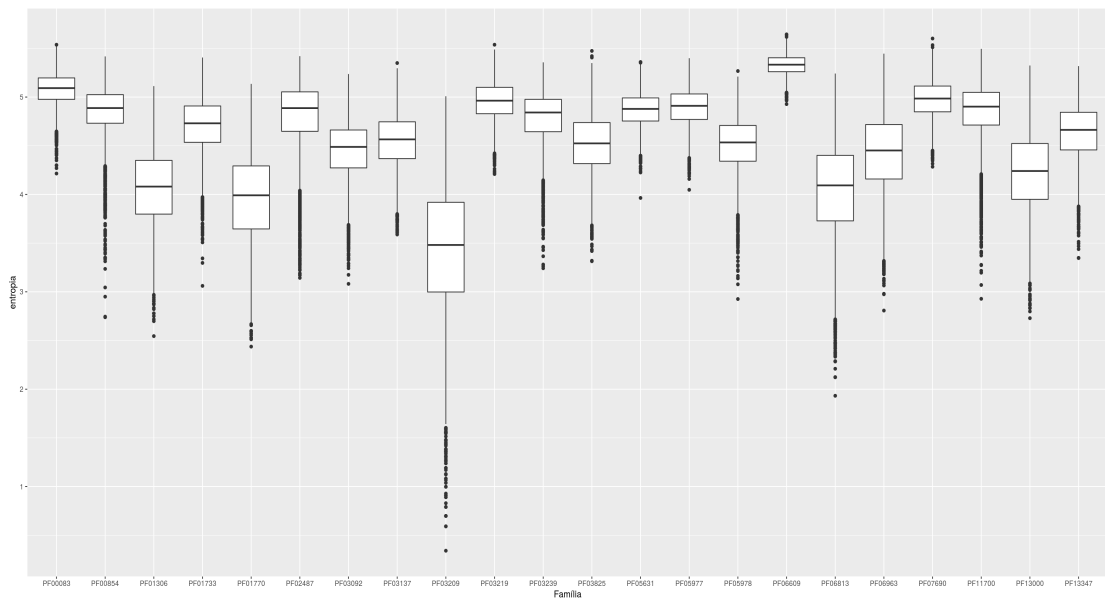
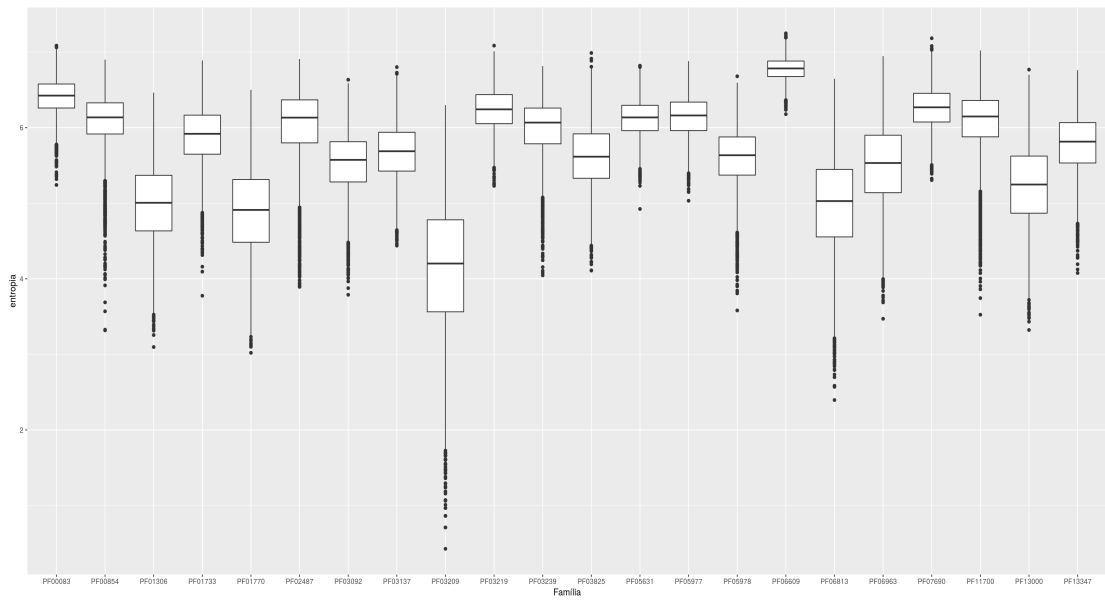
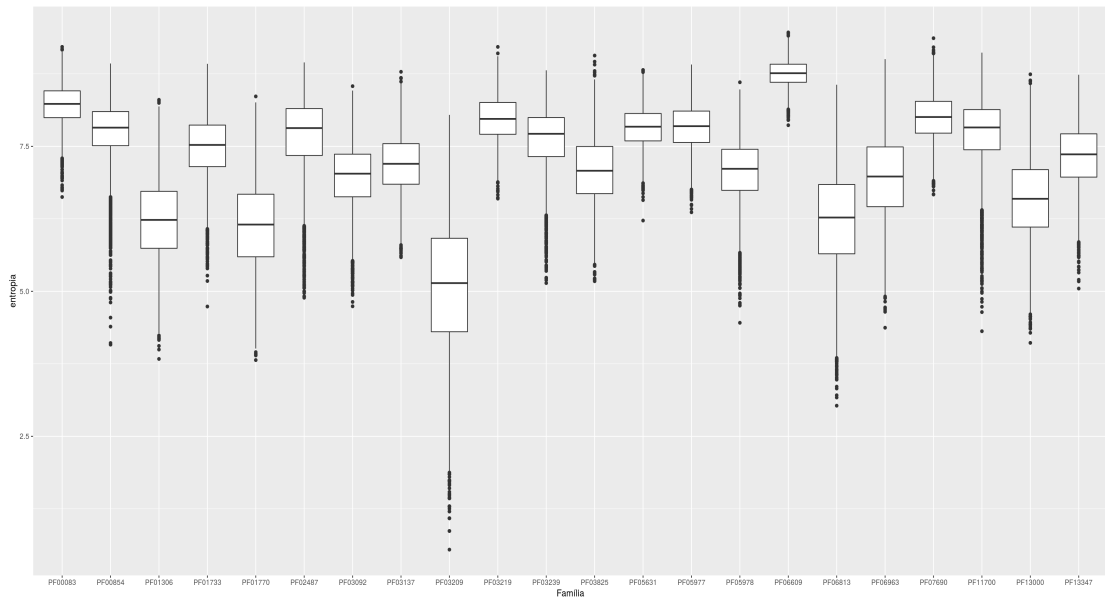
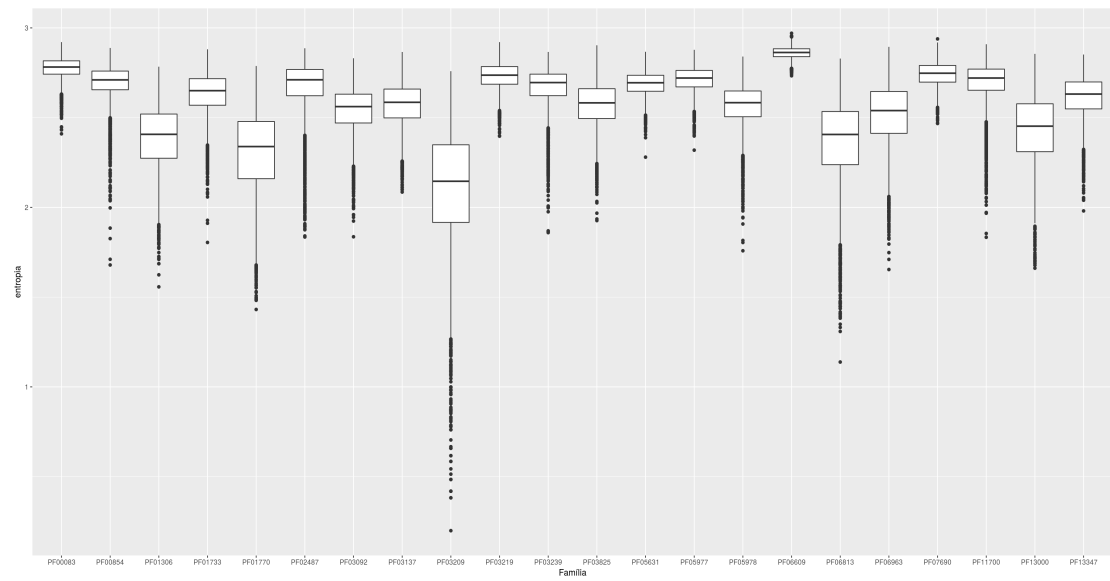
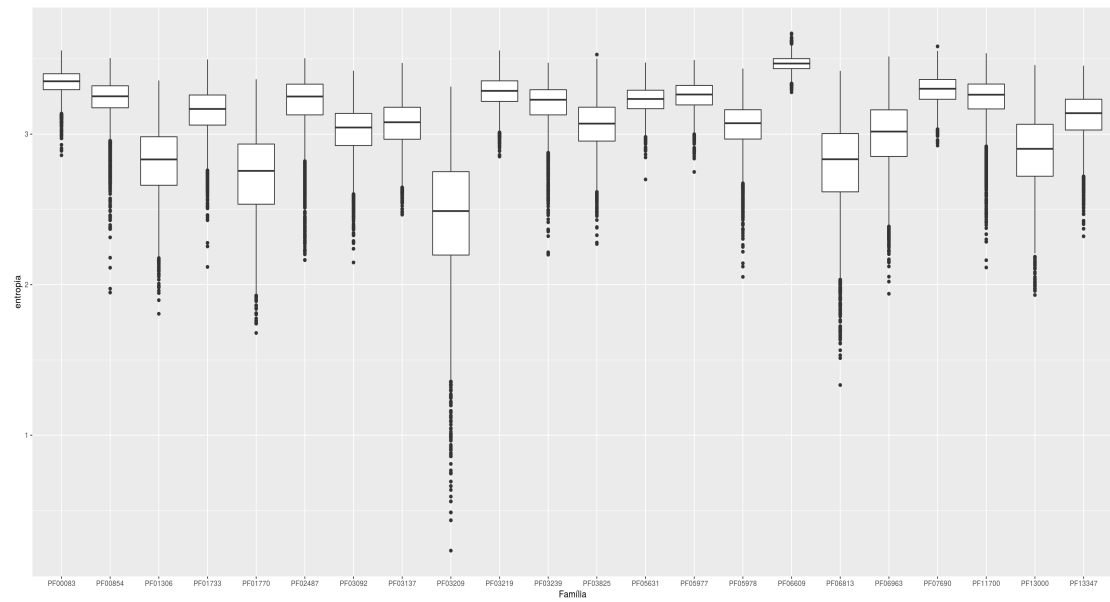
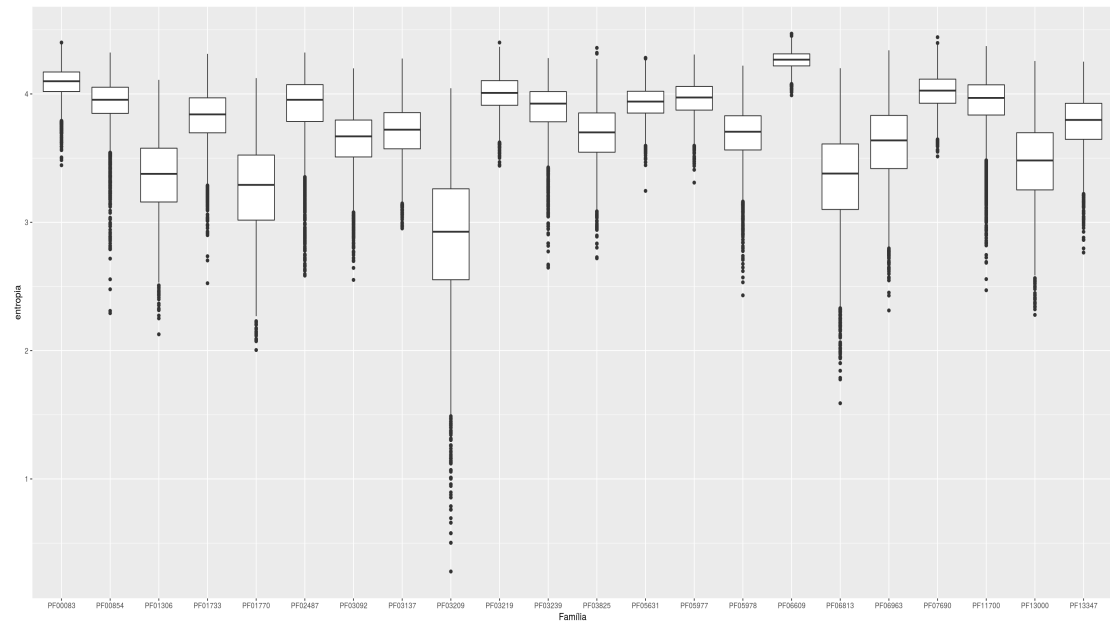


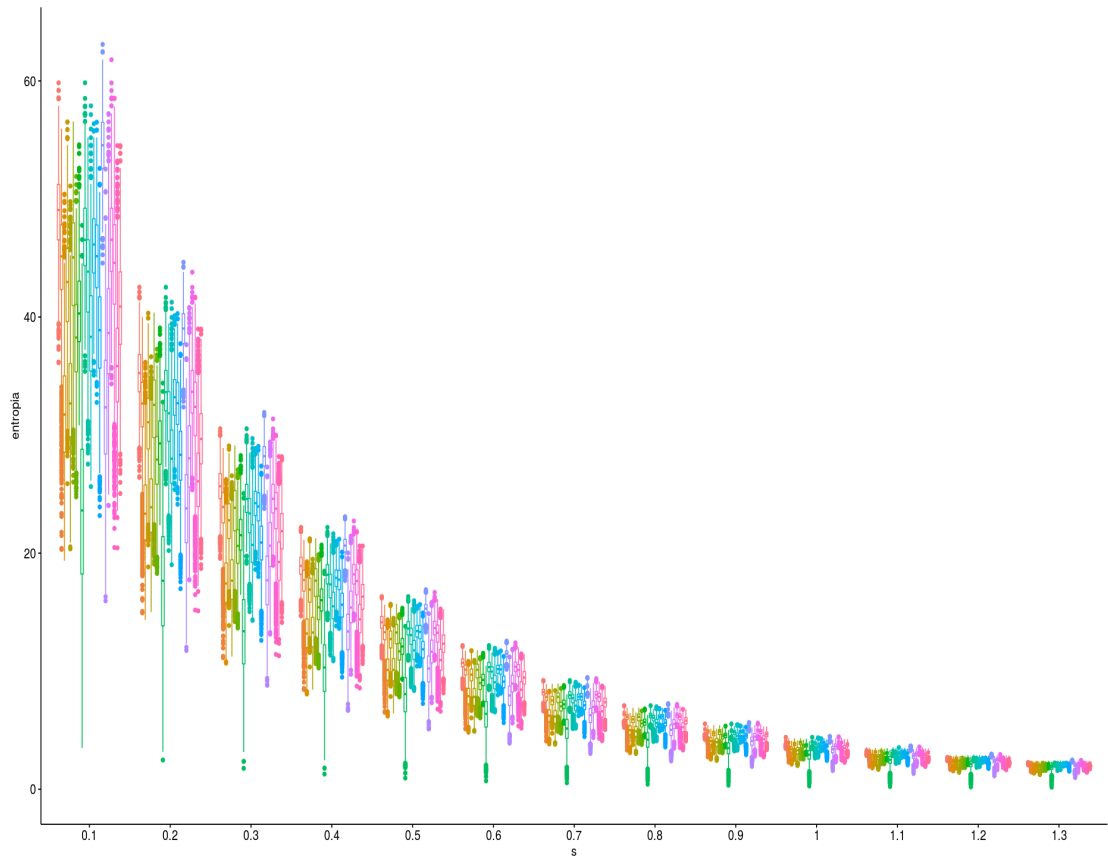
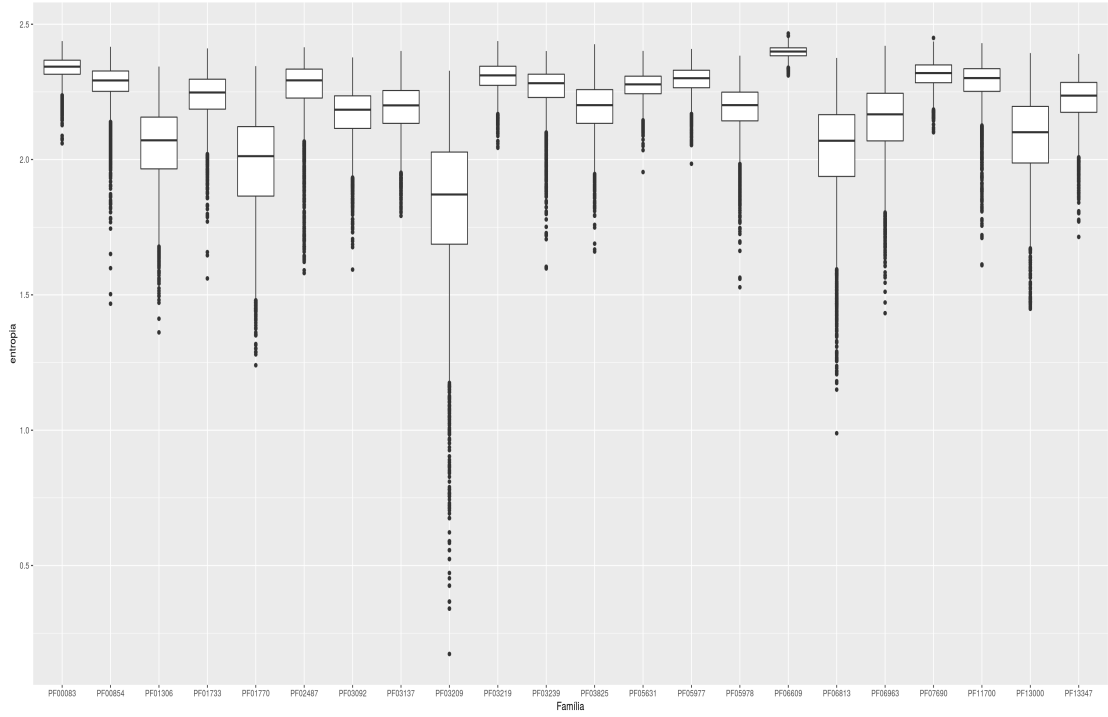
Figura 7.11 – Boxplots para as famílias do clã CL0014 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta











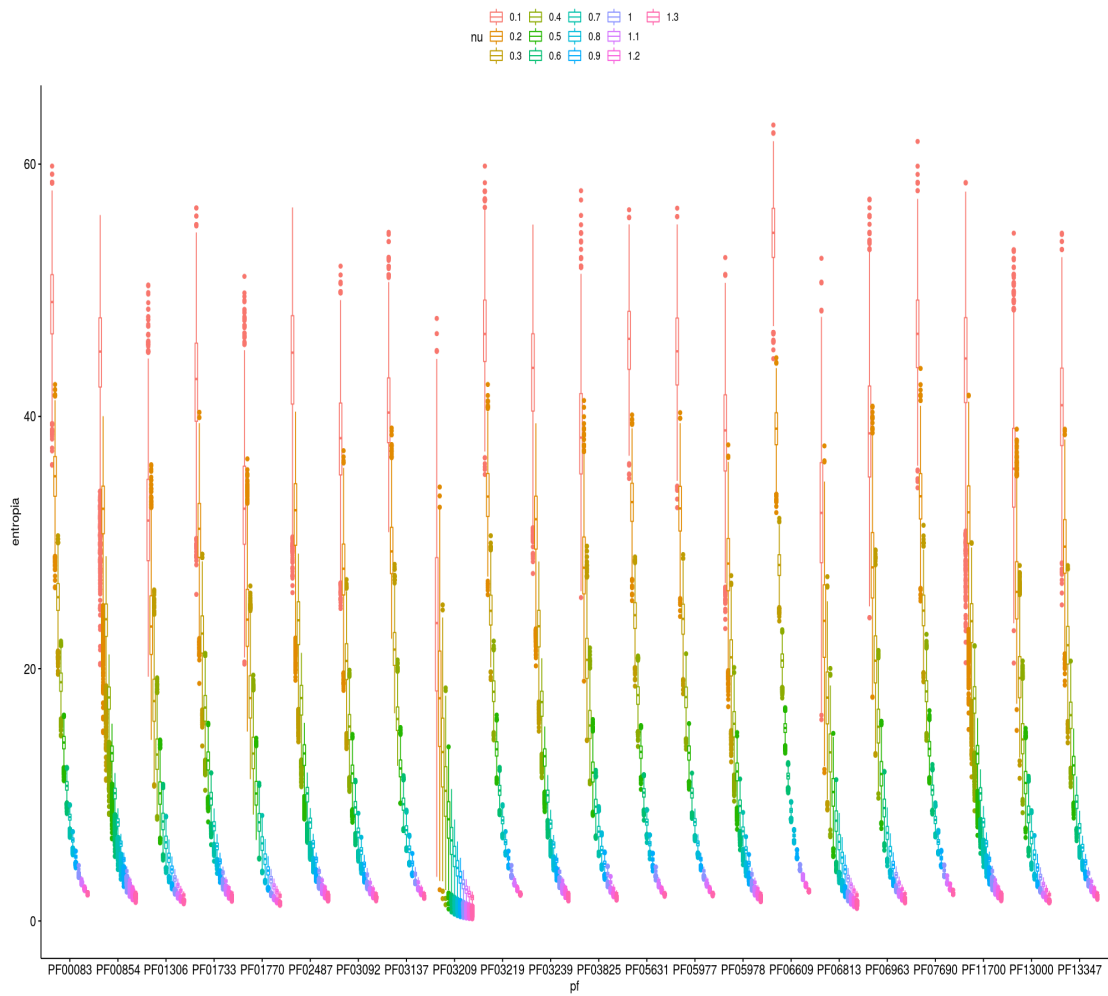
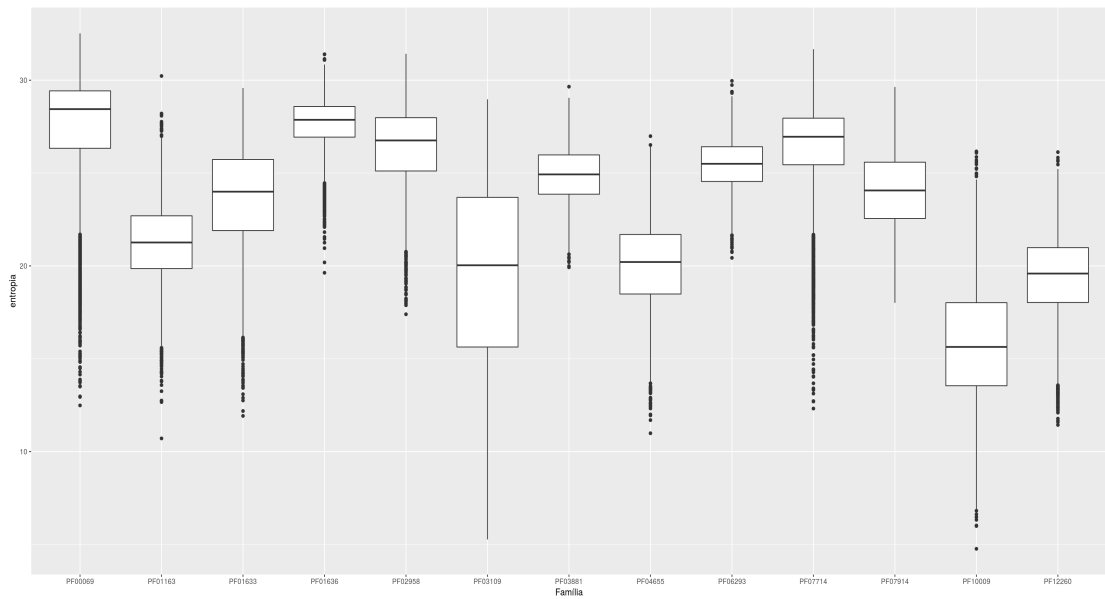
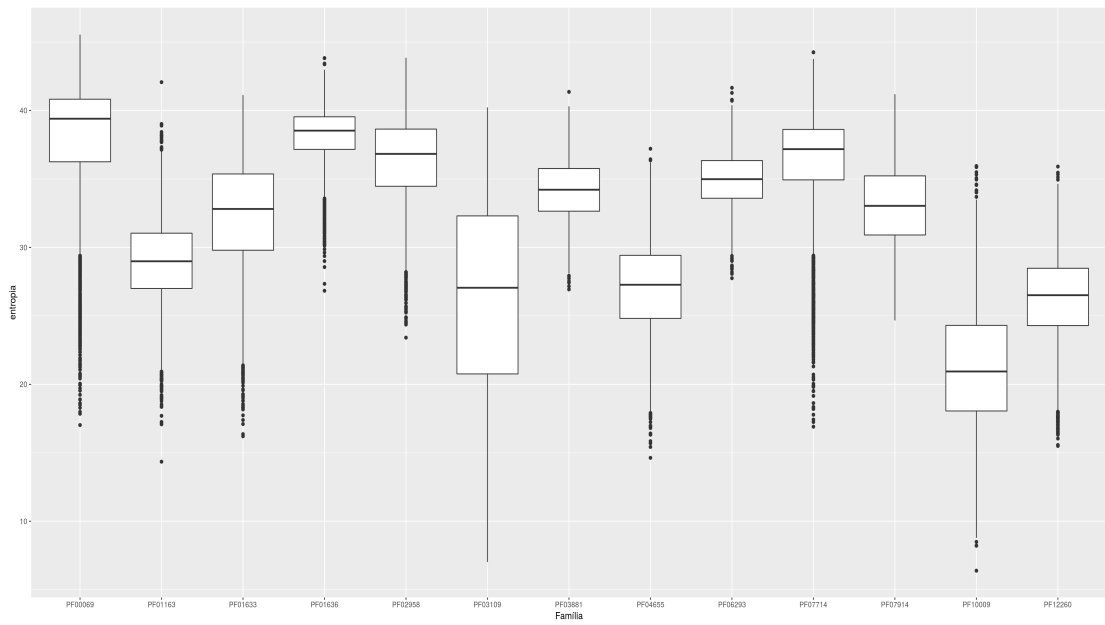
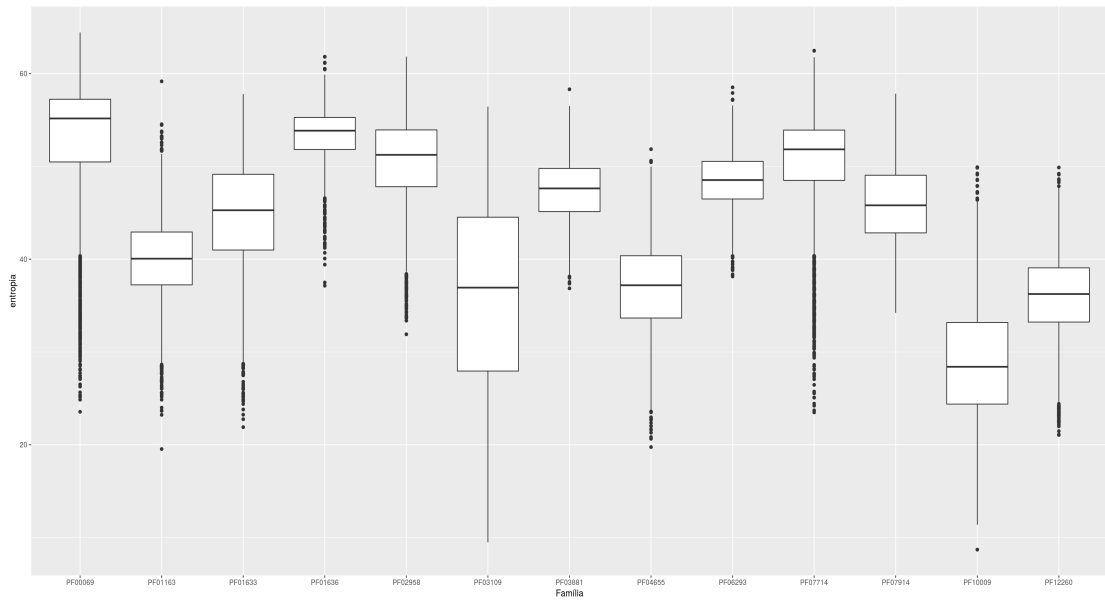
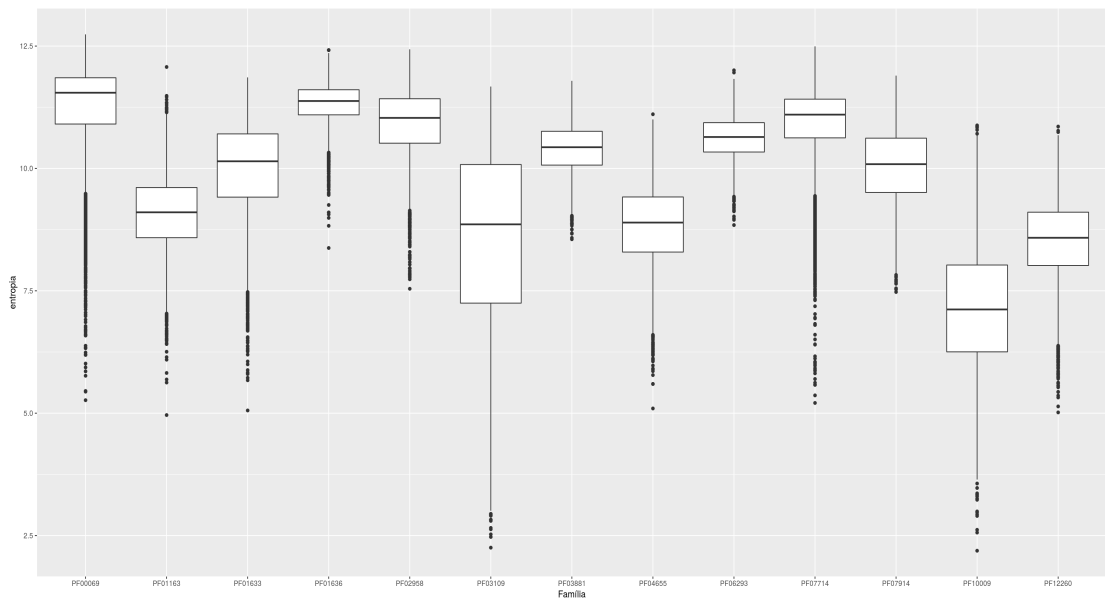
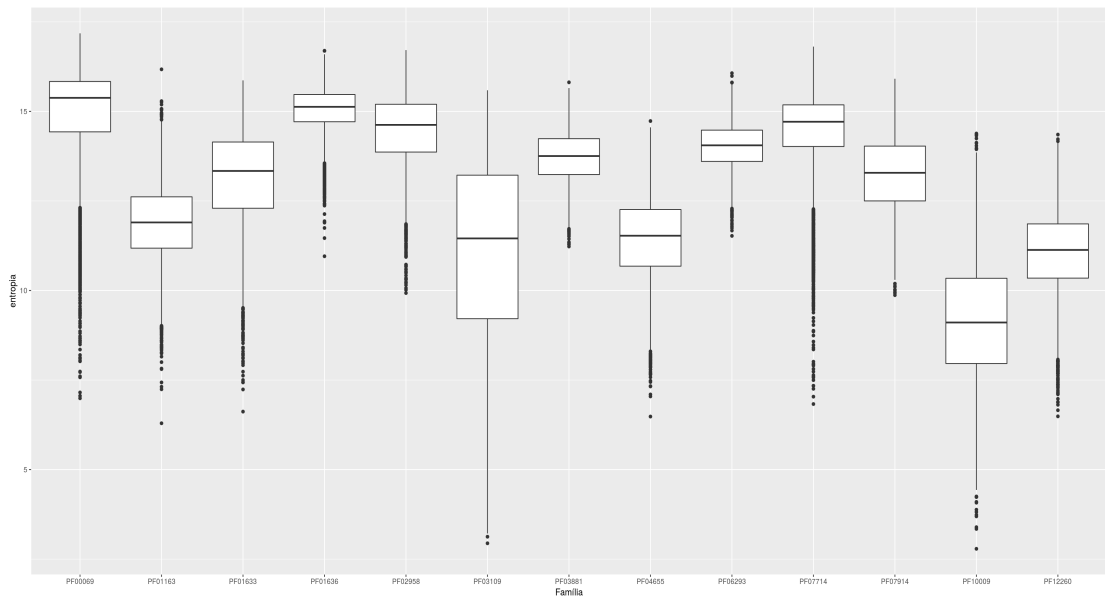
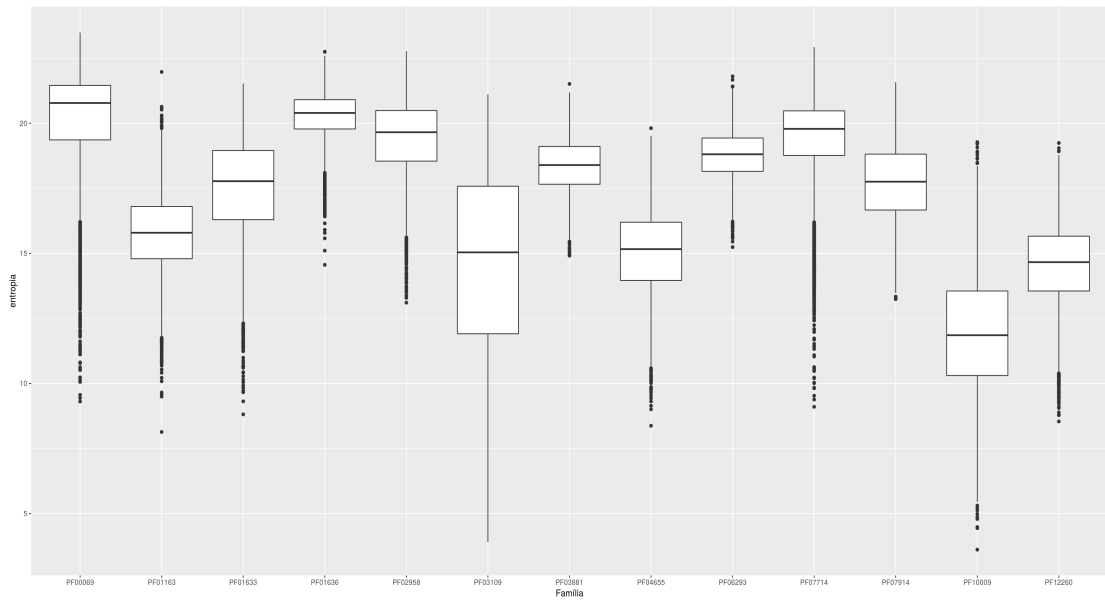
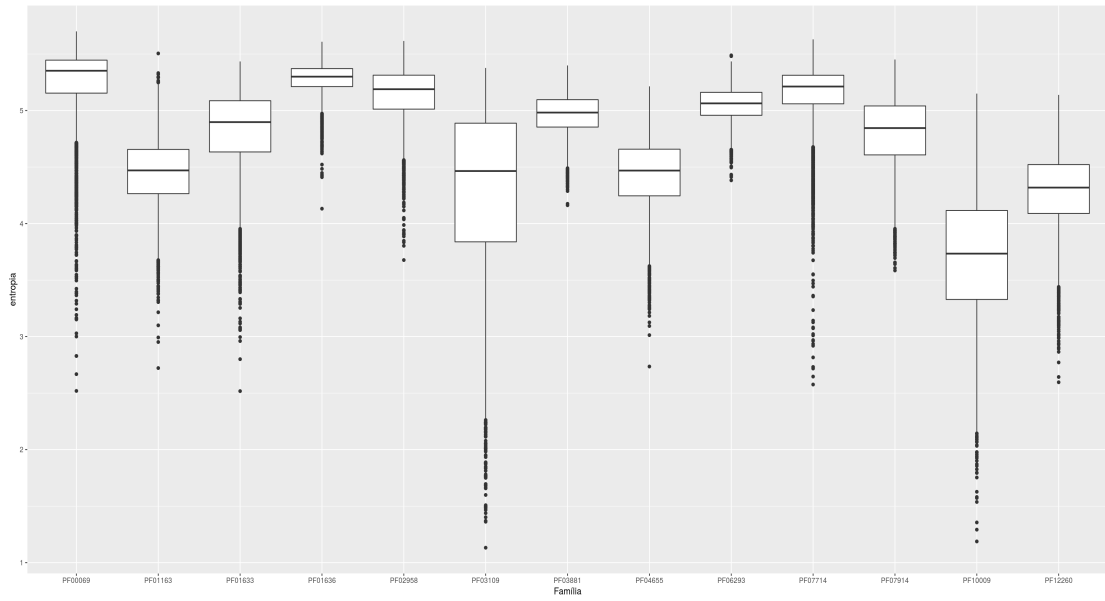
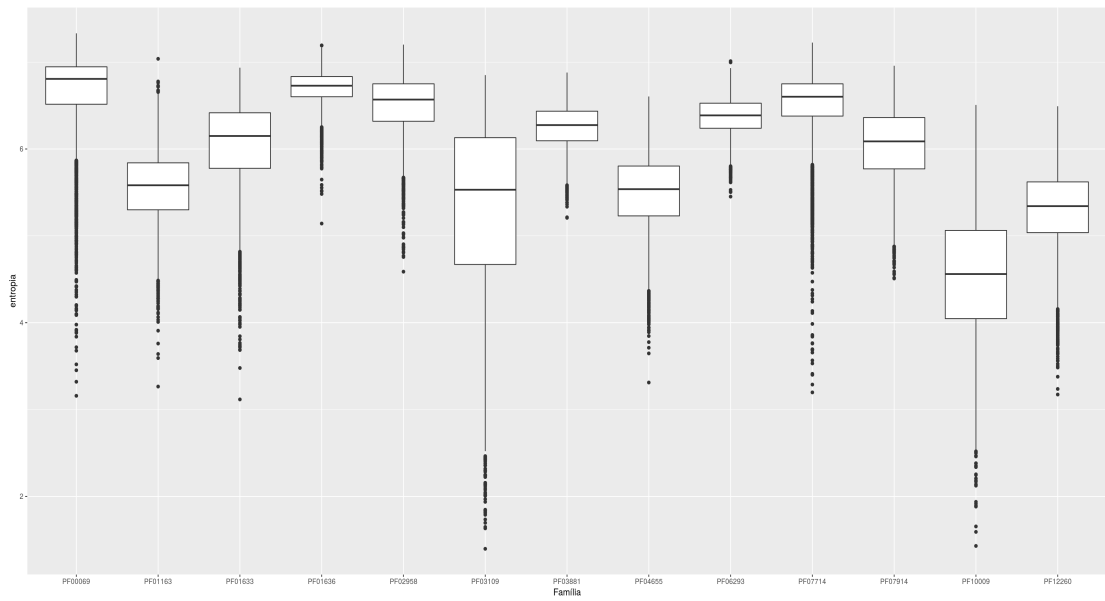
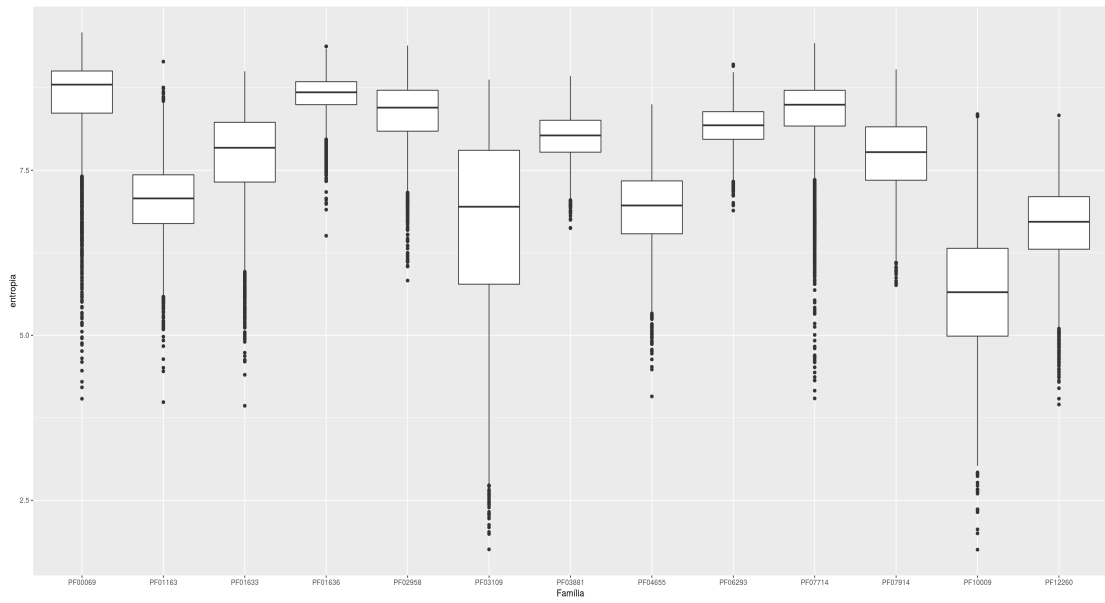
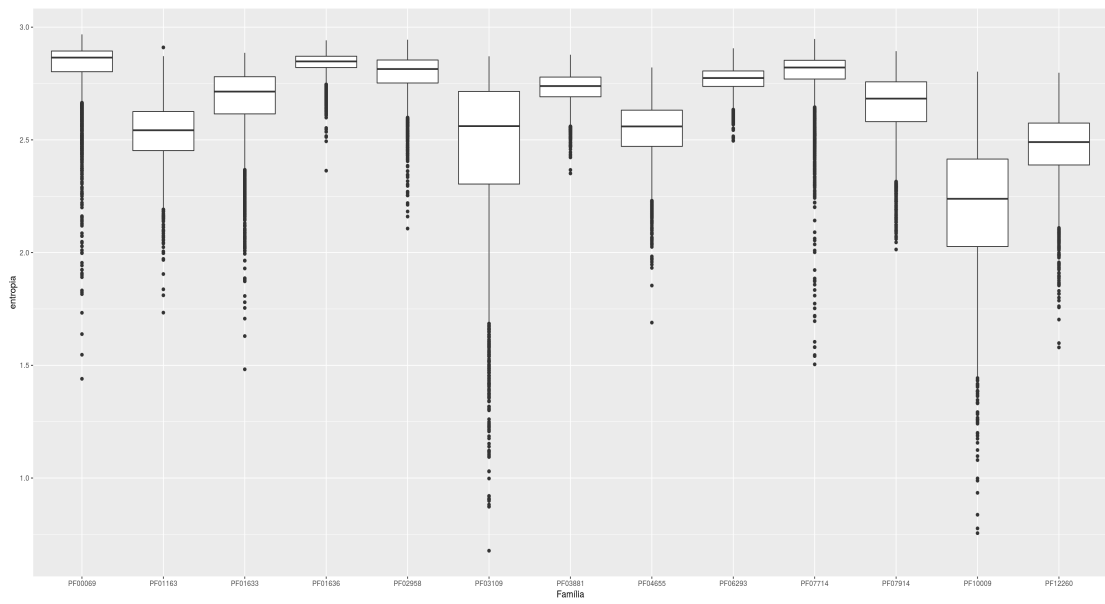
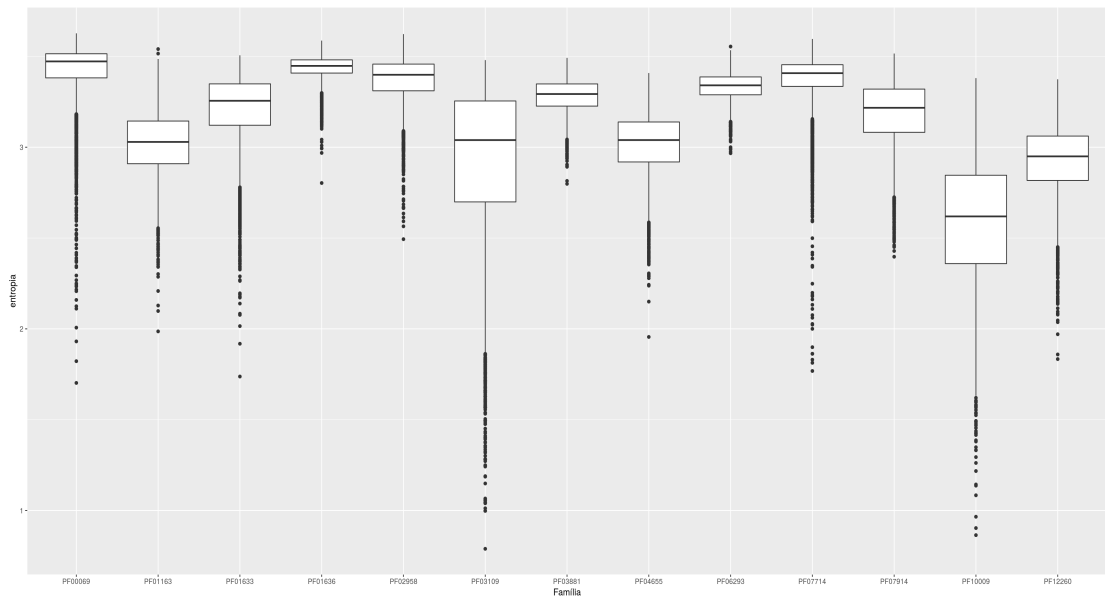
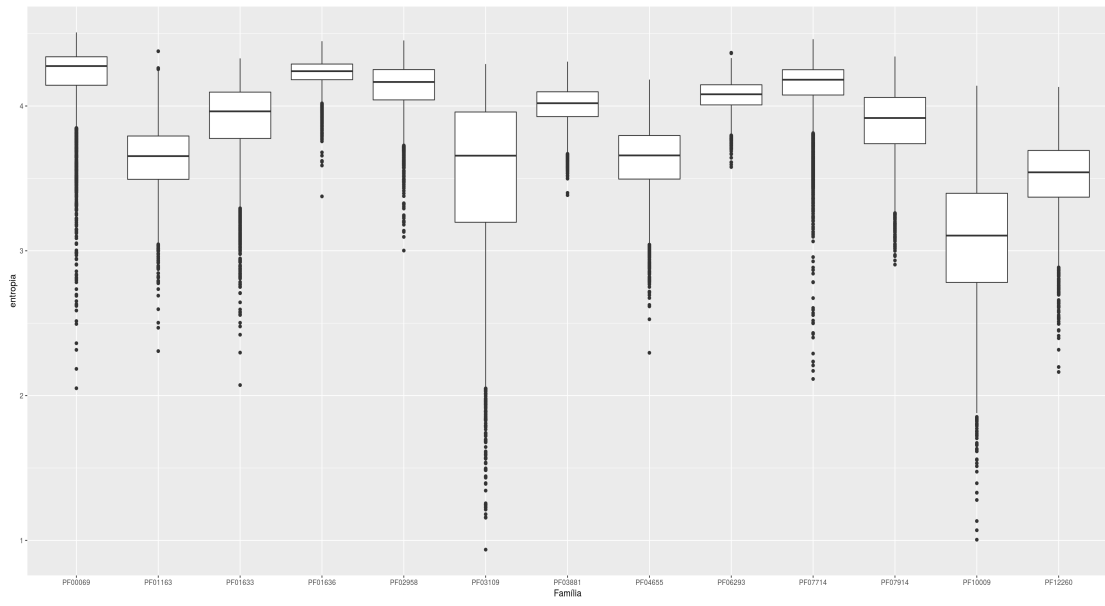


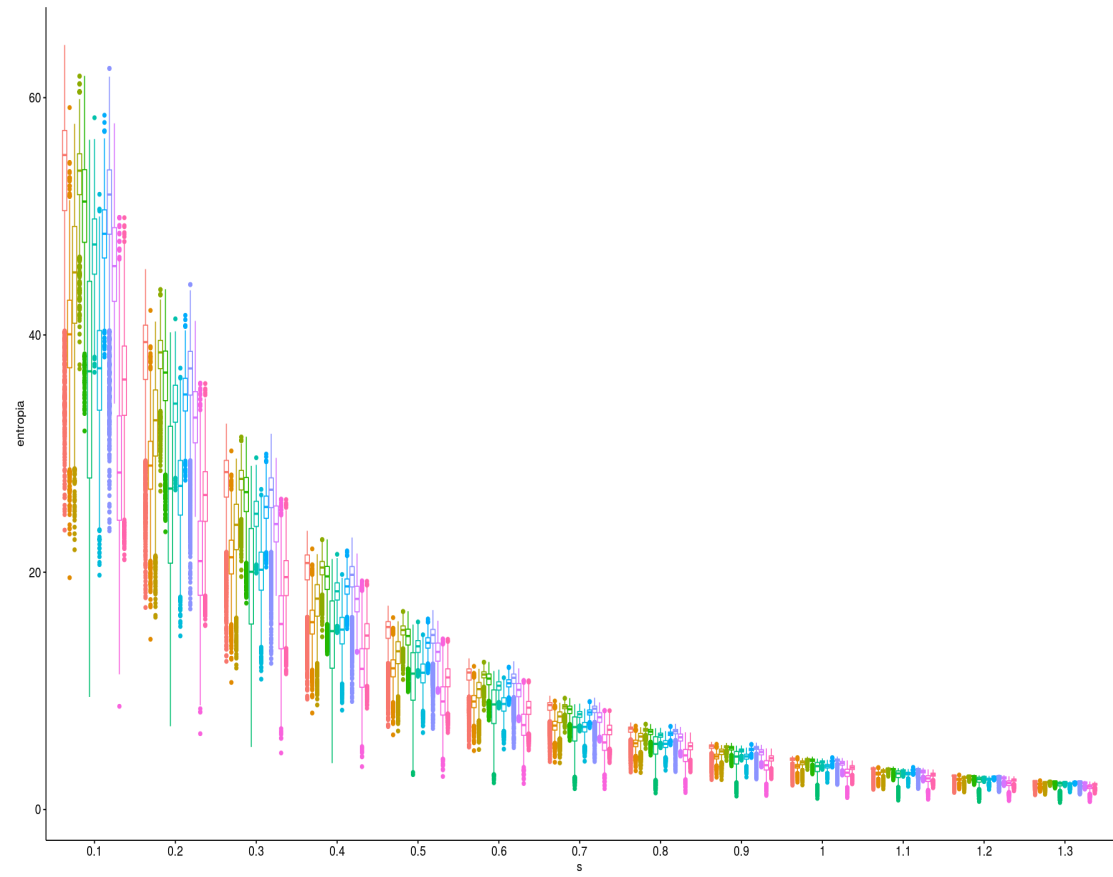
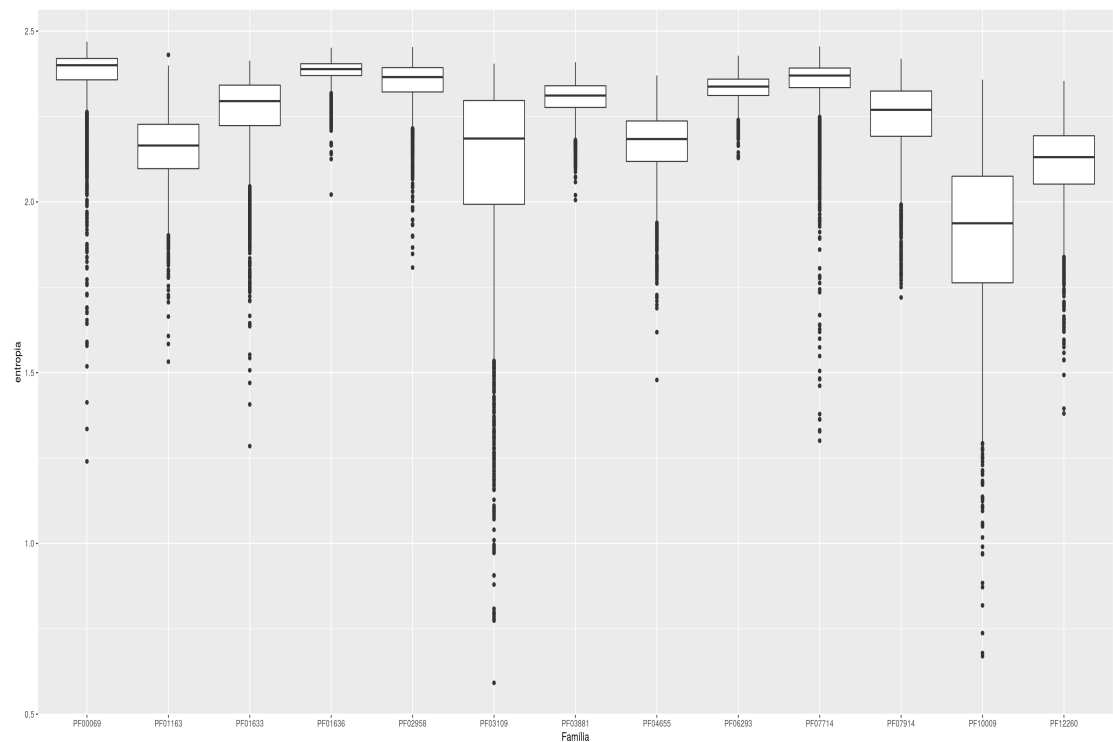
Figura 7.12 – Boxplots para as famílias do clã CL0015 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta











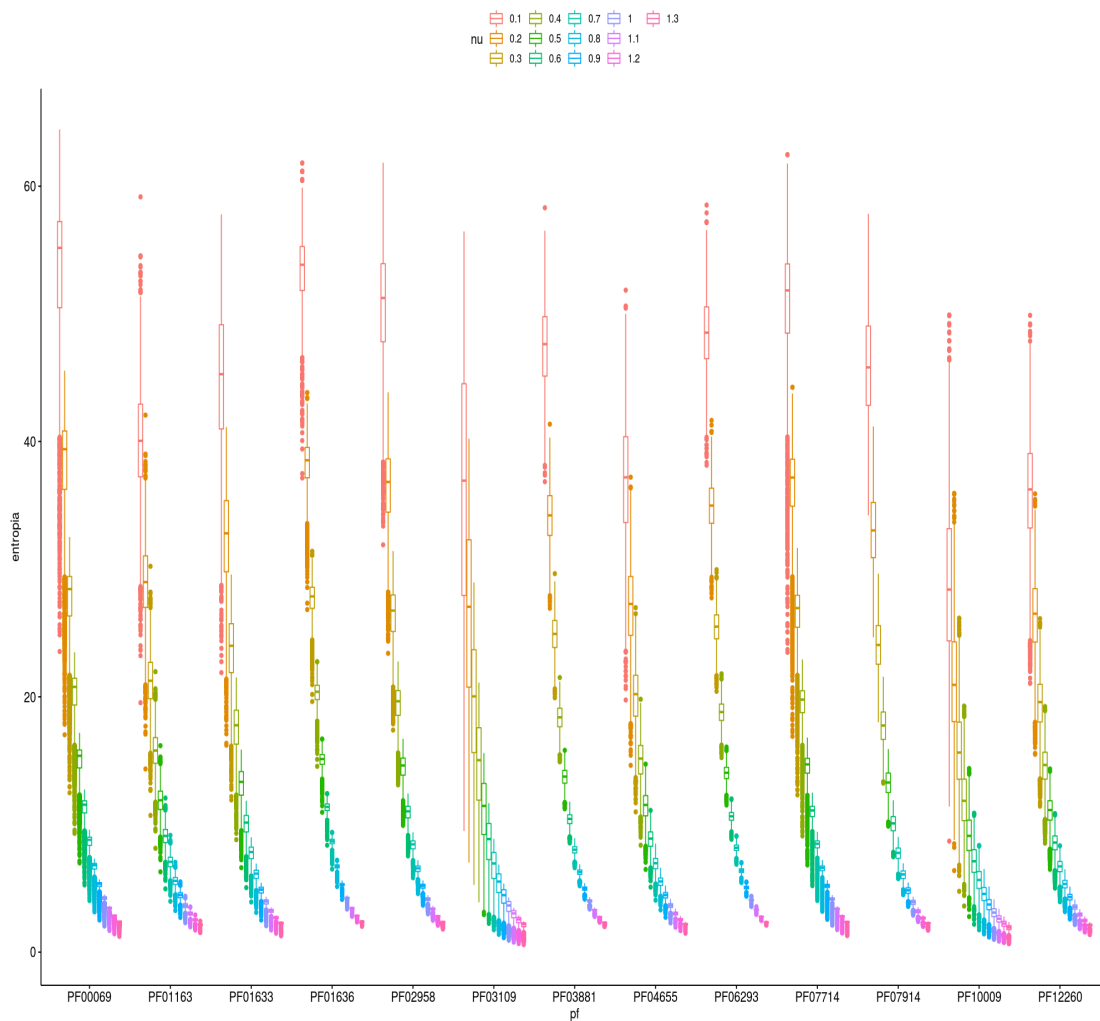
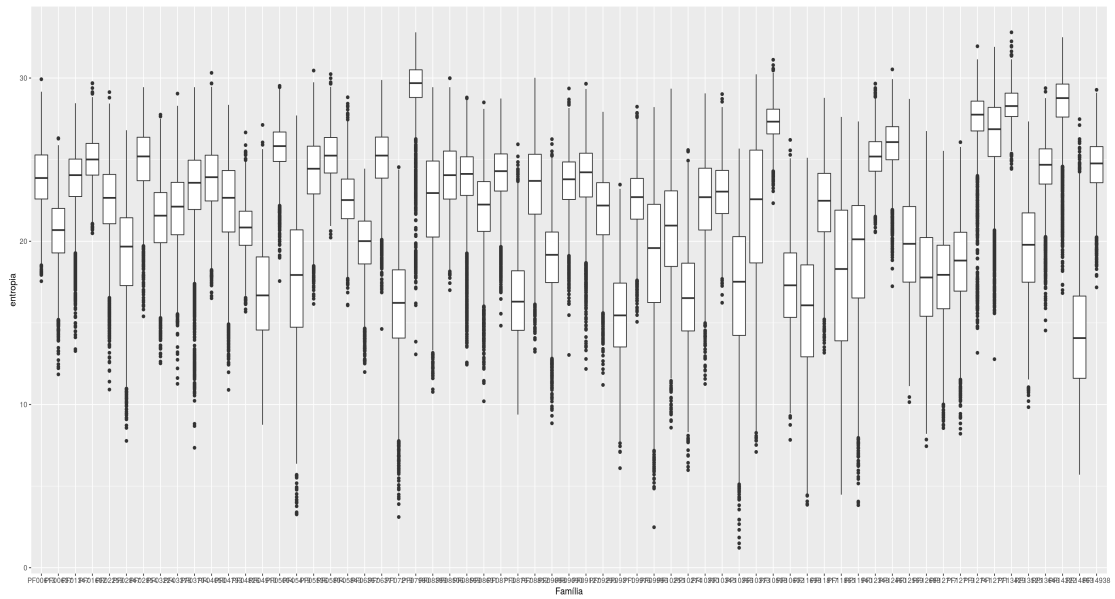
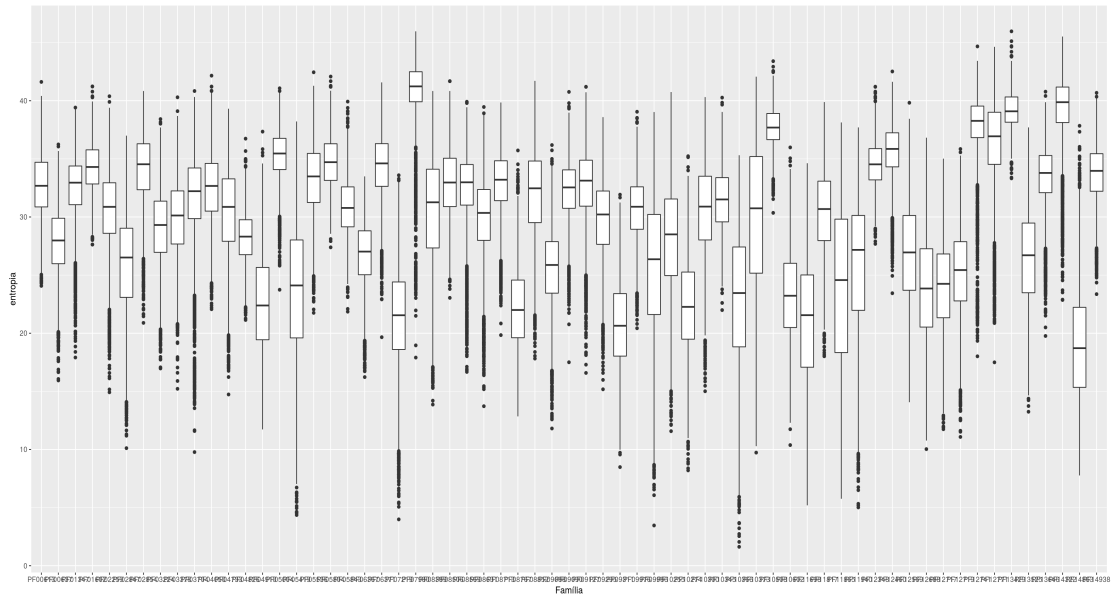
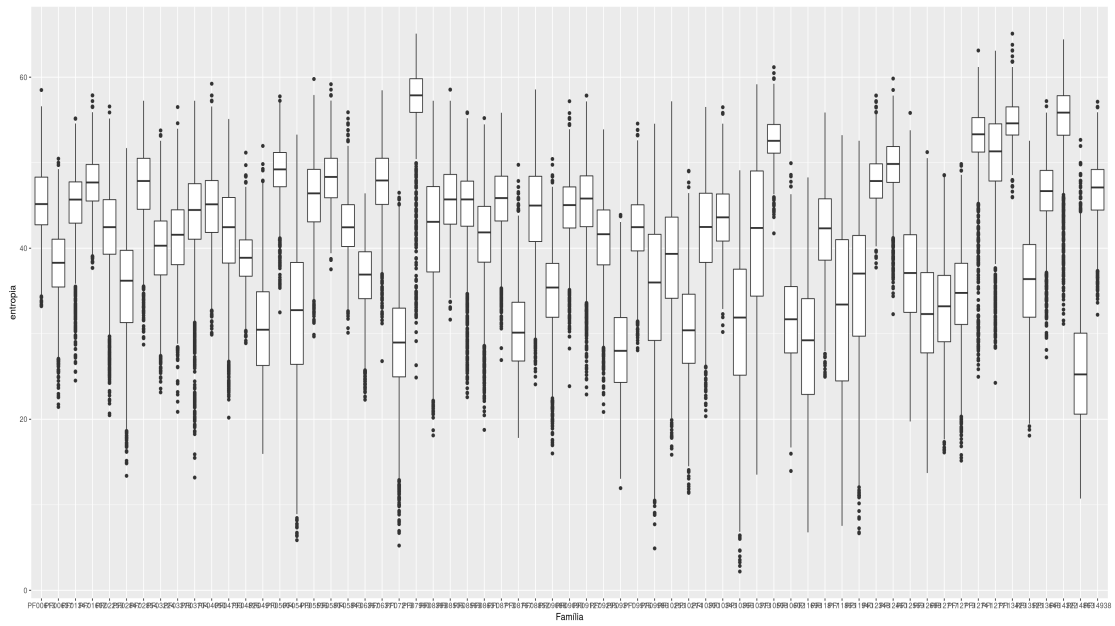
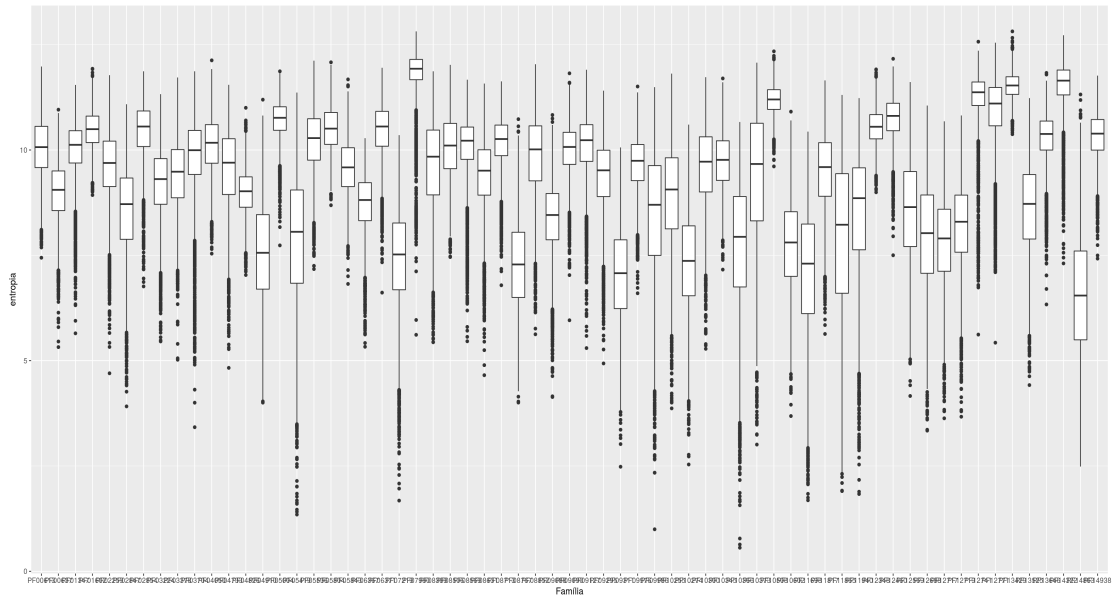
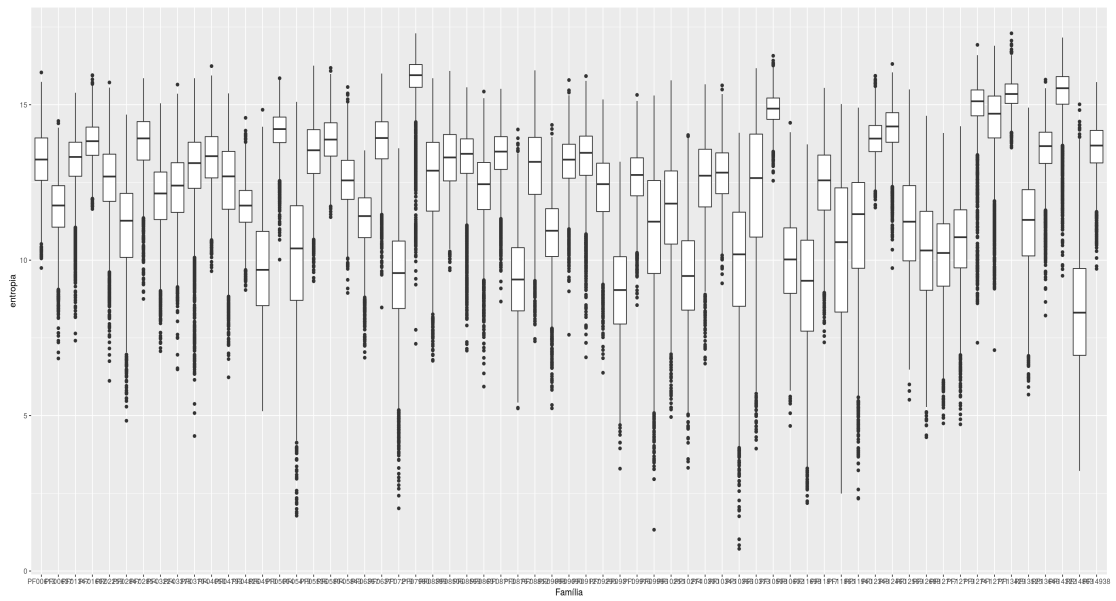
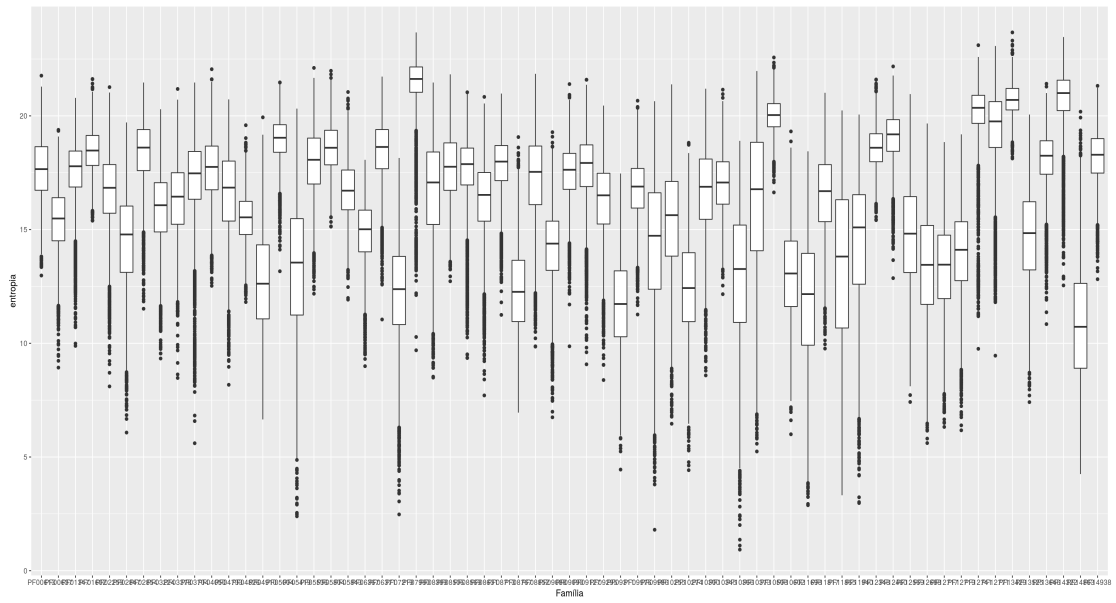
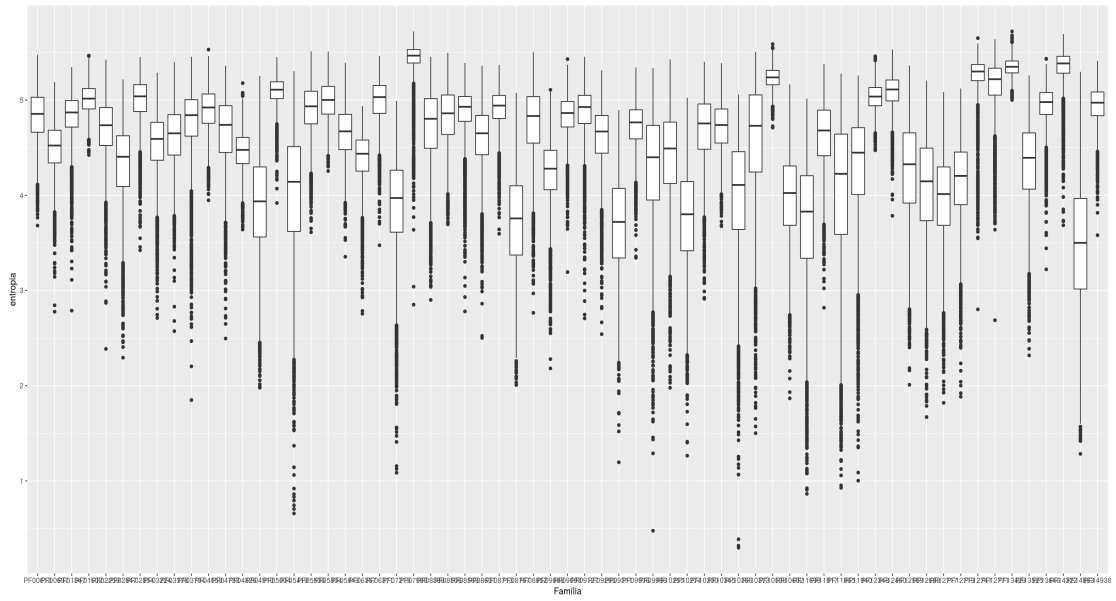
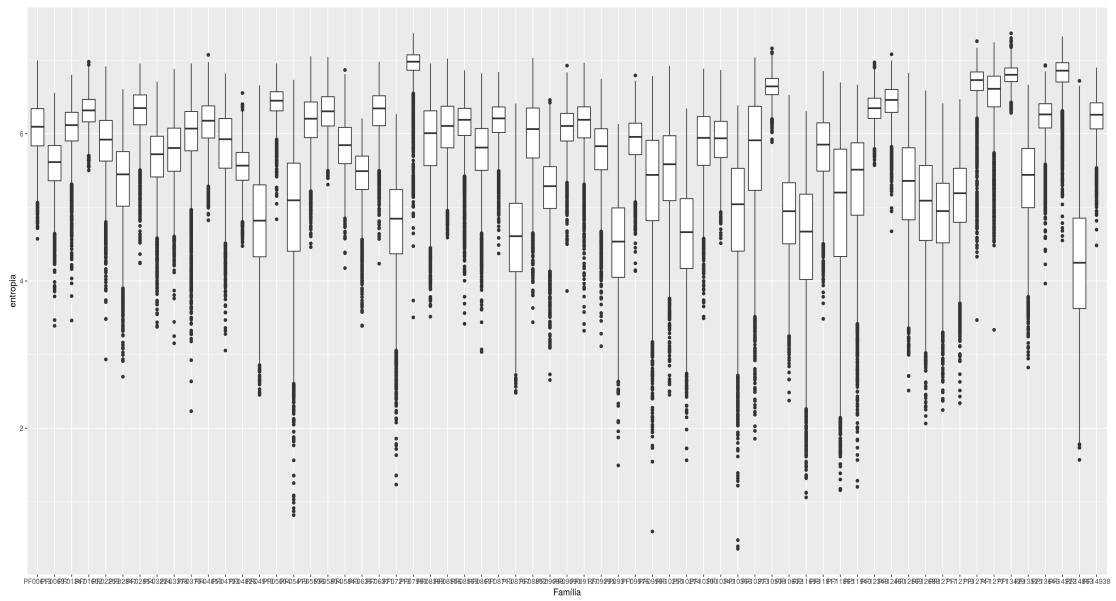
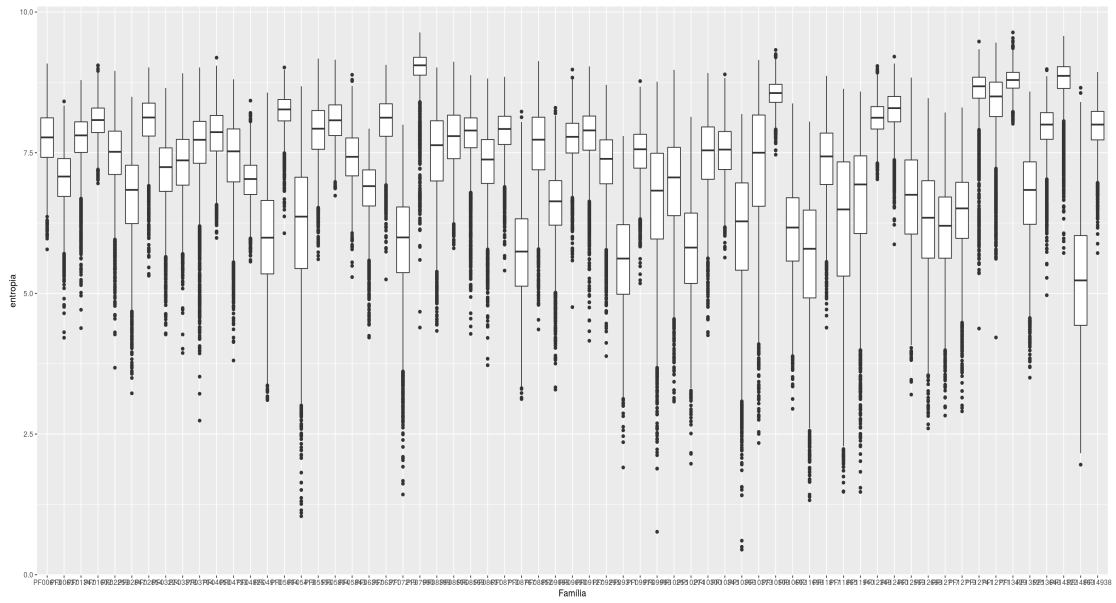
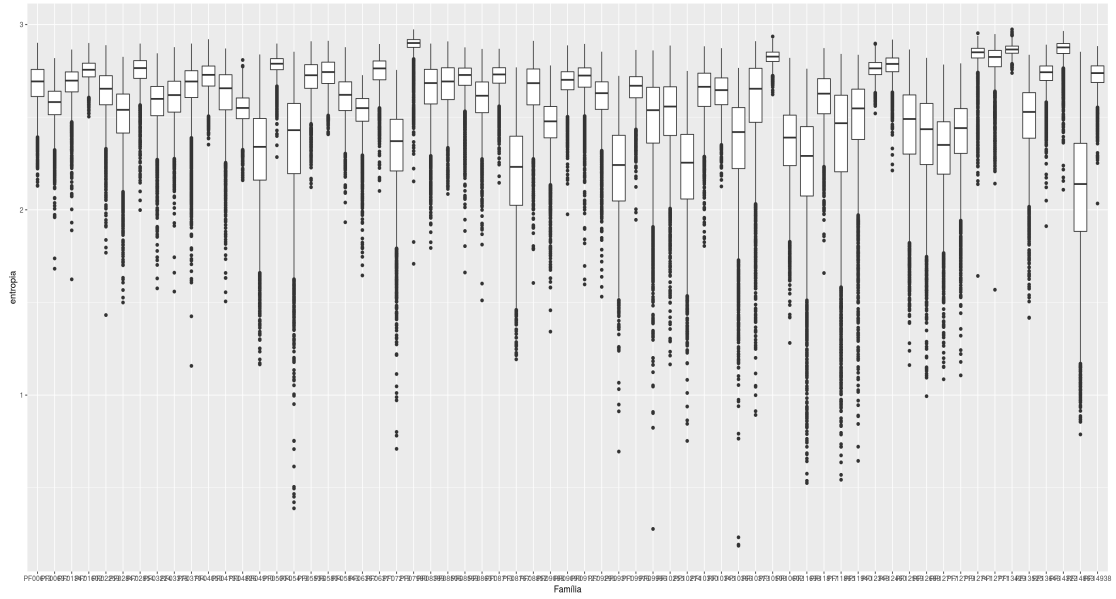
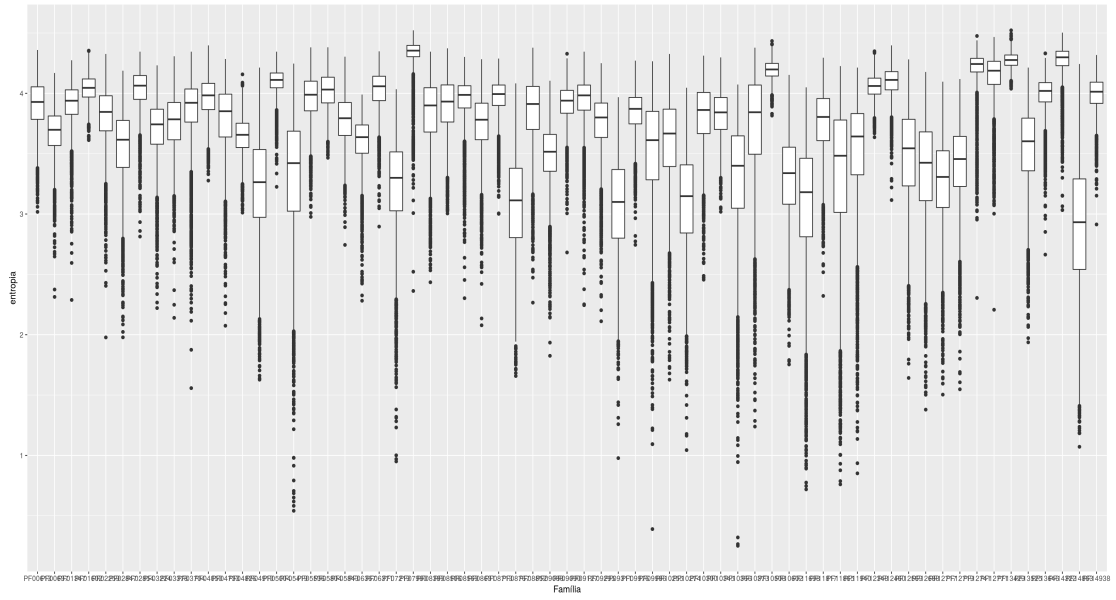
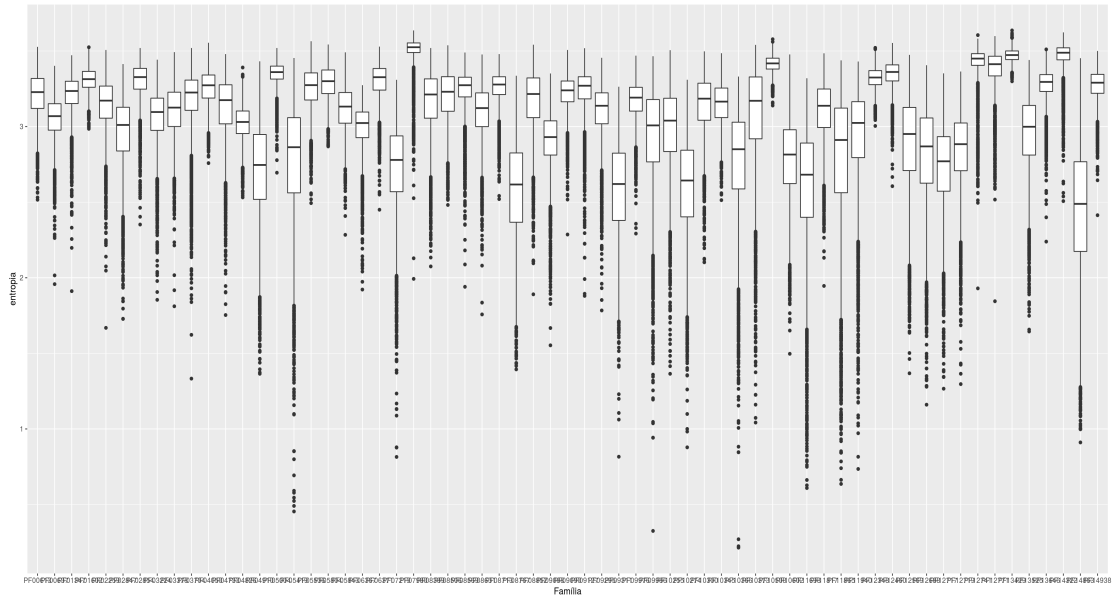


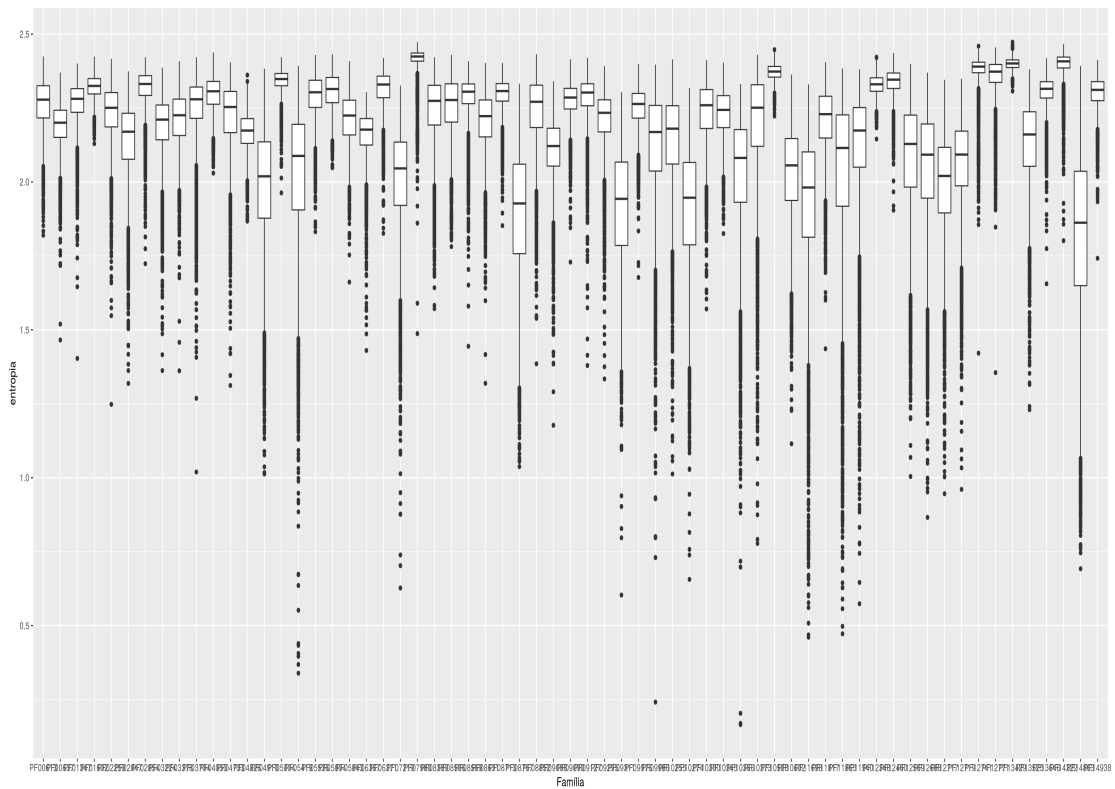
Figura 7.13 – Boxplots para as famílias do clã CL0016 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta



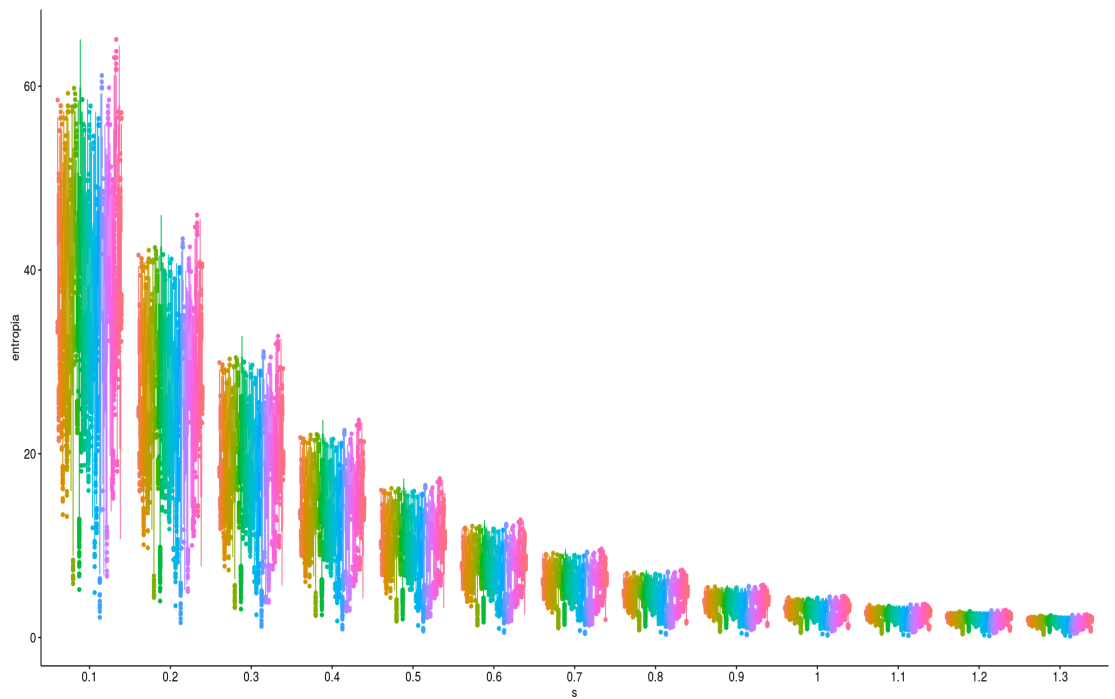








- PF00613 PF04910 PF08631 PF10300 PF12688
- PF00637 PF05004 PF08713 PF10345 PF12717
- PF01347 PF05419 PF08767 PF10366 PF12719
- PF01602 PF05536 PF08852 PF10373 PF12741
- PF02259 PF05804 PF09088 PF10508 PF12771
- PF02847 PF05843 PF09090 PF10602 PF13429
- PF02854 PF06367 PF09127 PF11898 PF13525
- PF03224 PF06371 PF09295 PF11817 PF13846
- PF03378 PF07219 PF09311 PF11865 PF14322
- PF03704 PF07980 PF09976 PF11940 PF14863
- PF04050 PF08389 PF09986 PF12348 PF14938
- PF04733 PF08508 PF10255 PF12460
- PF04826 PF08569 PF10274 PF12569



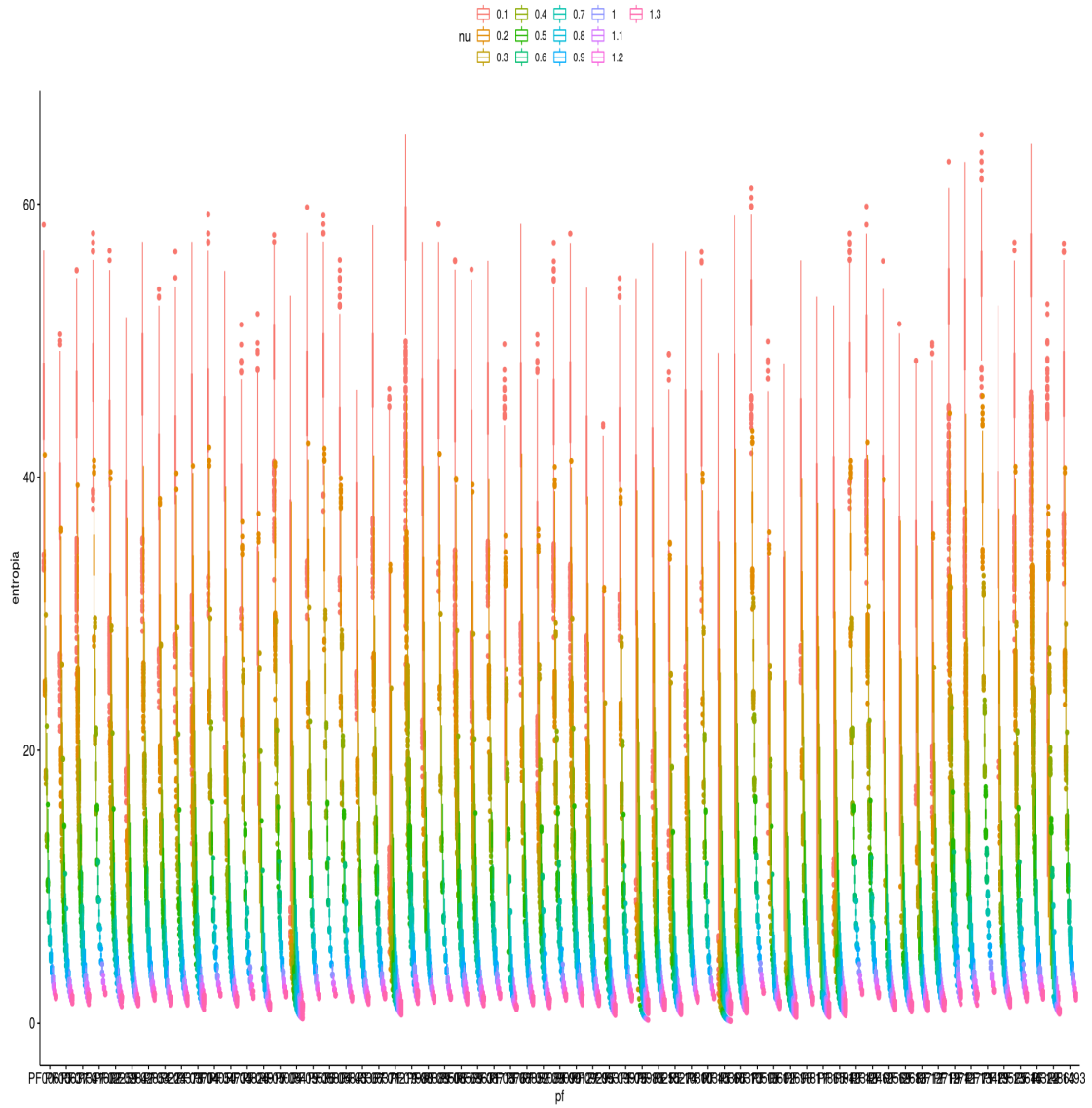
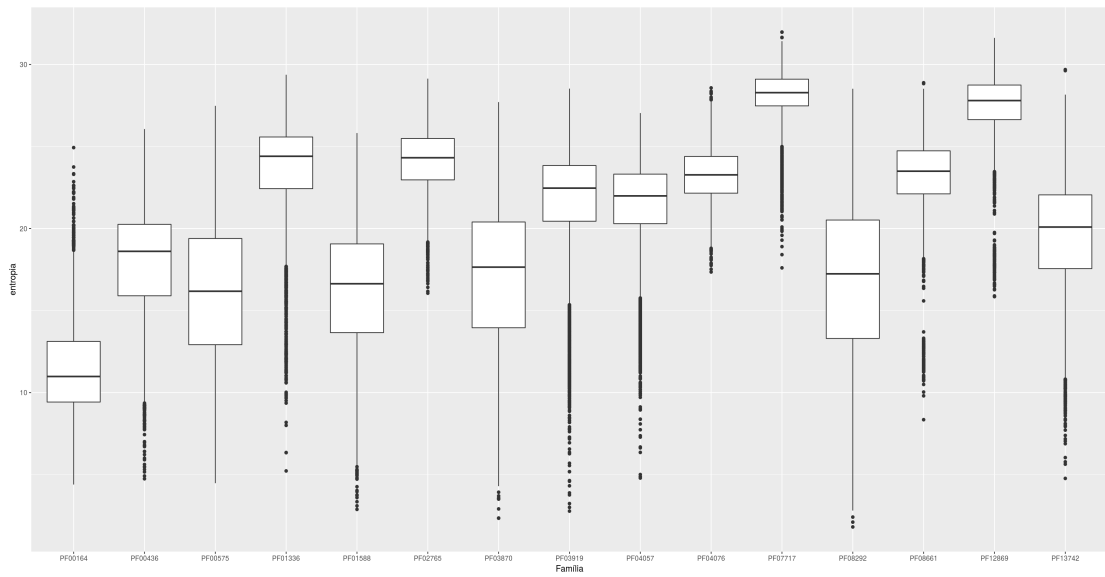
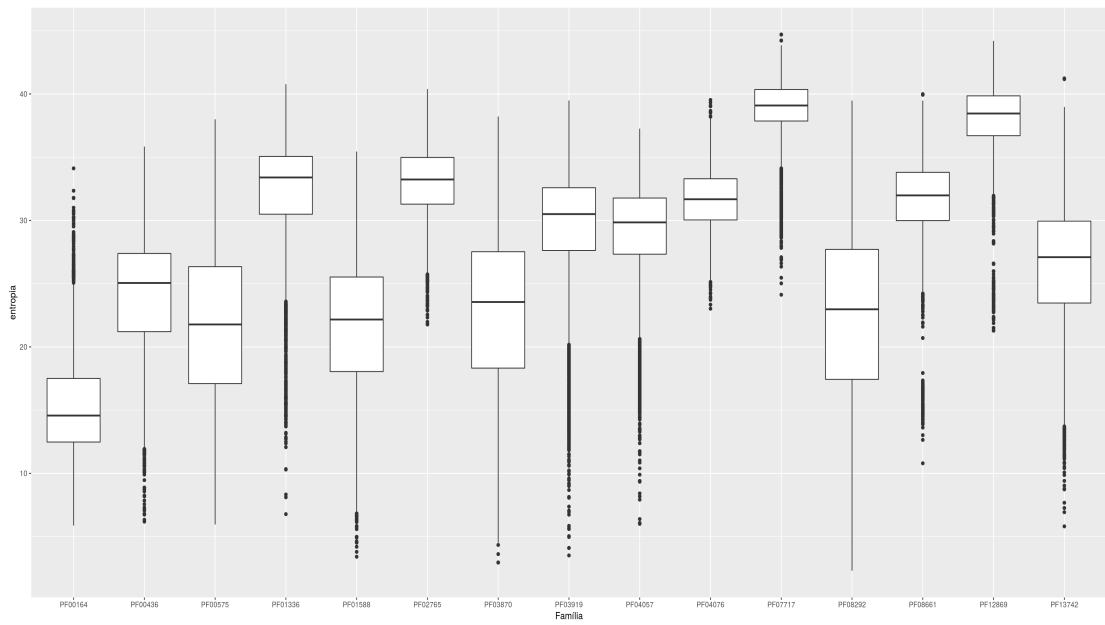
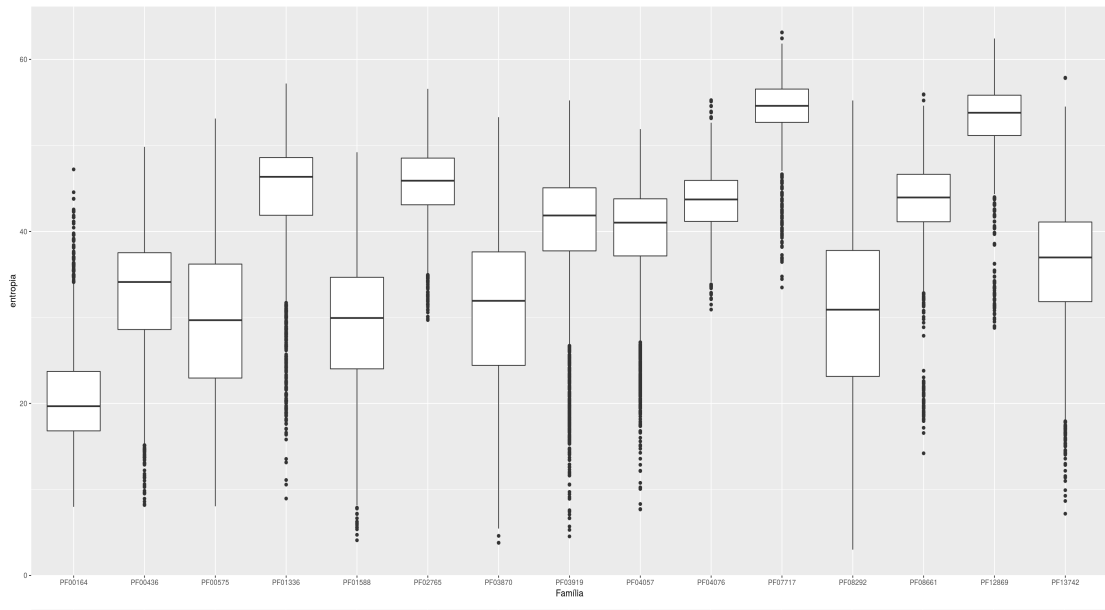
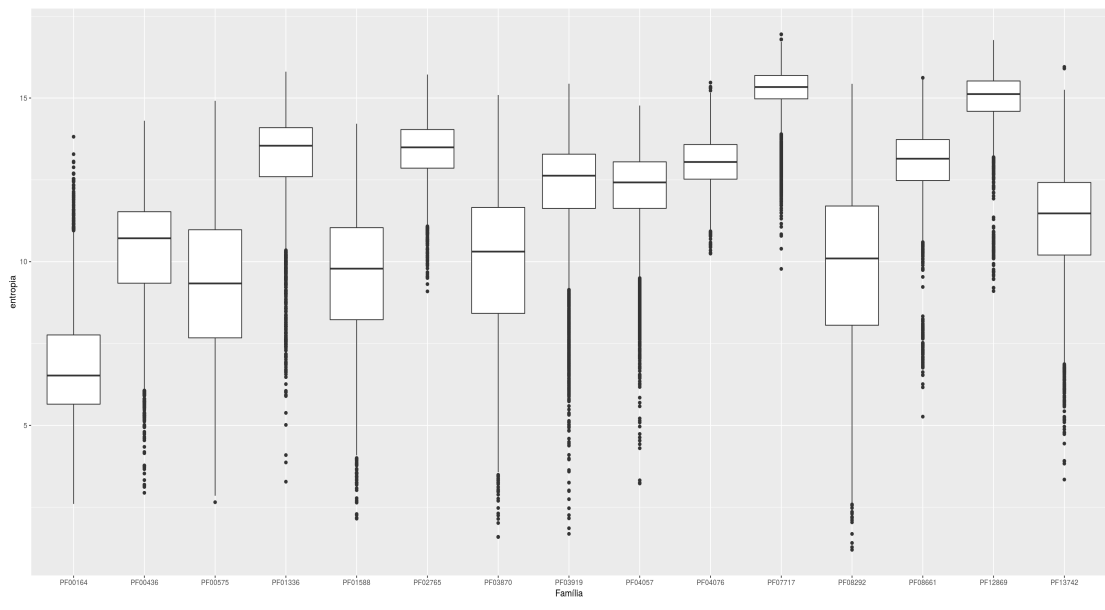
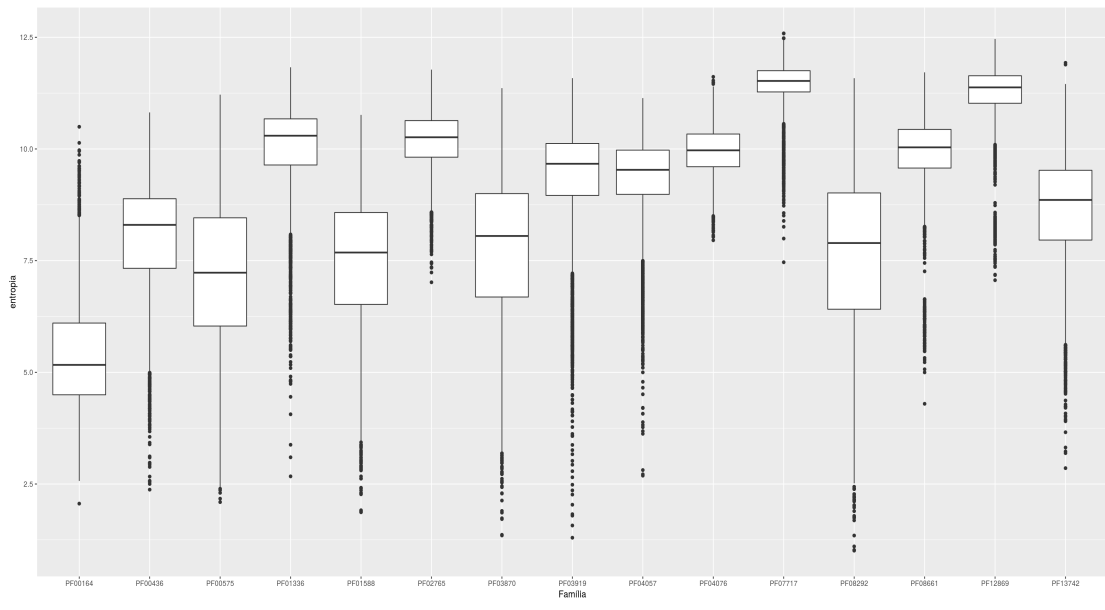
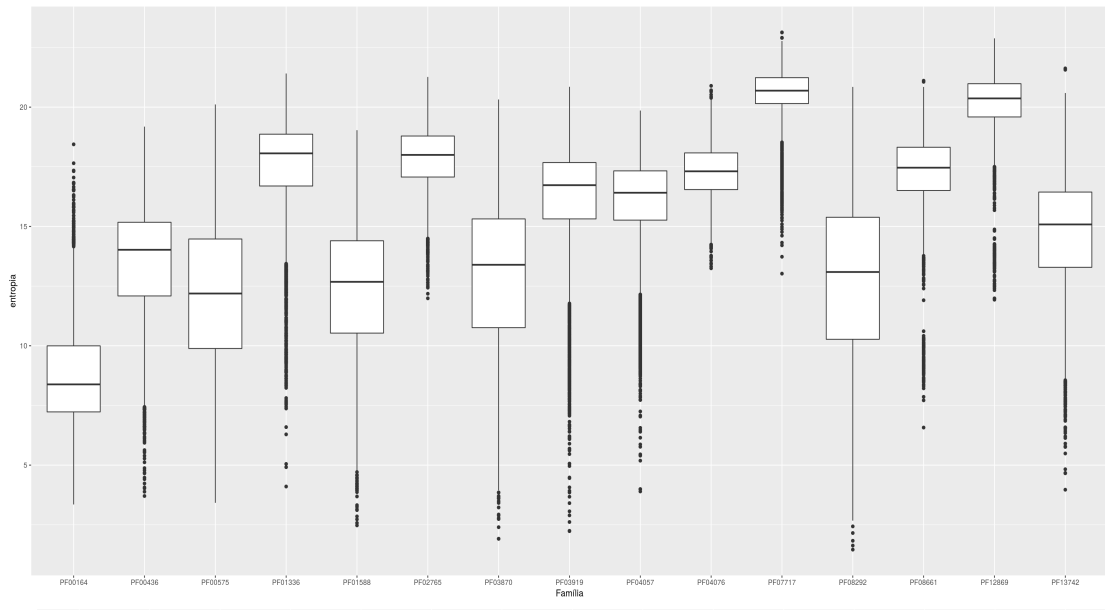
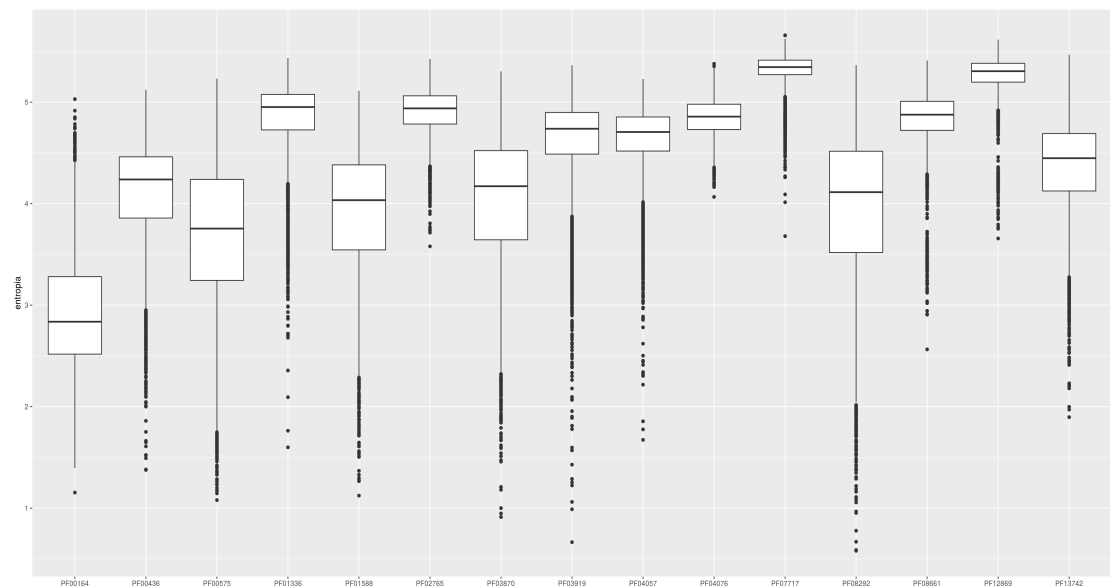
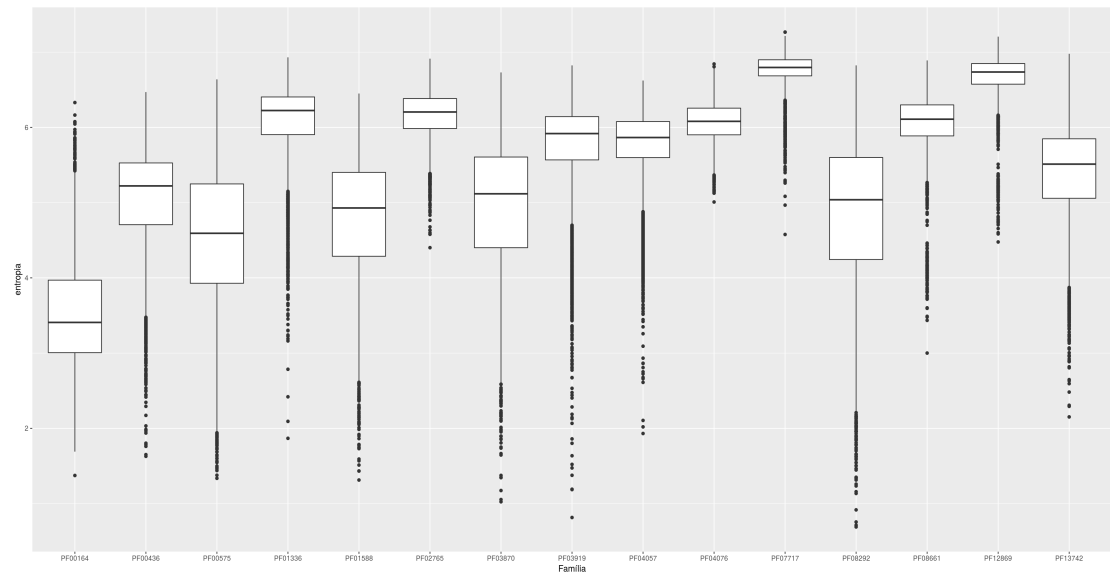
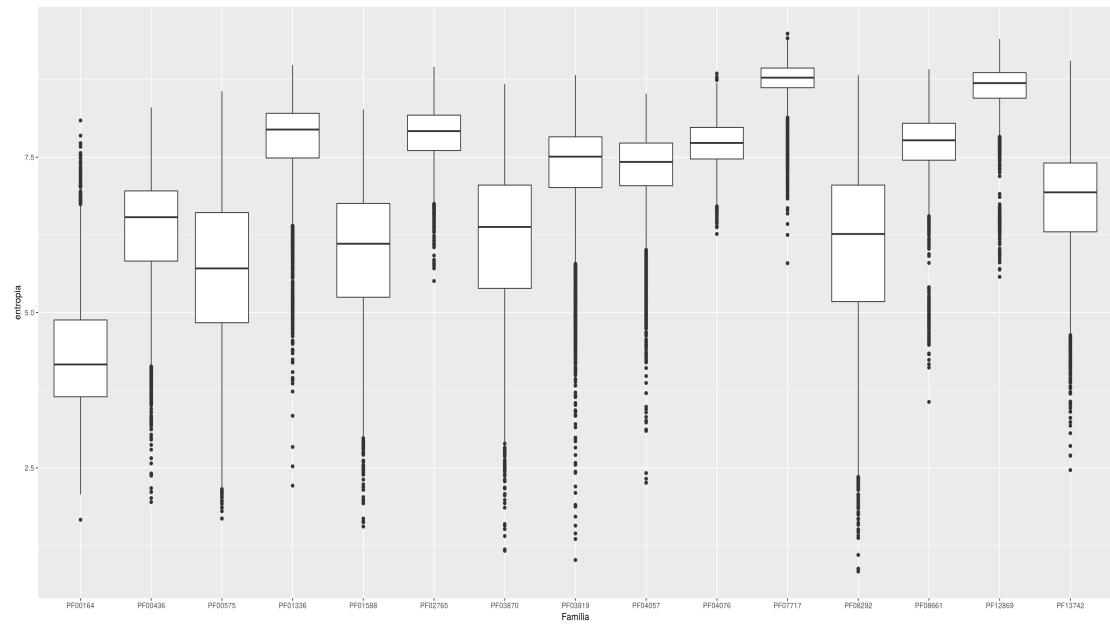
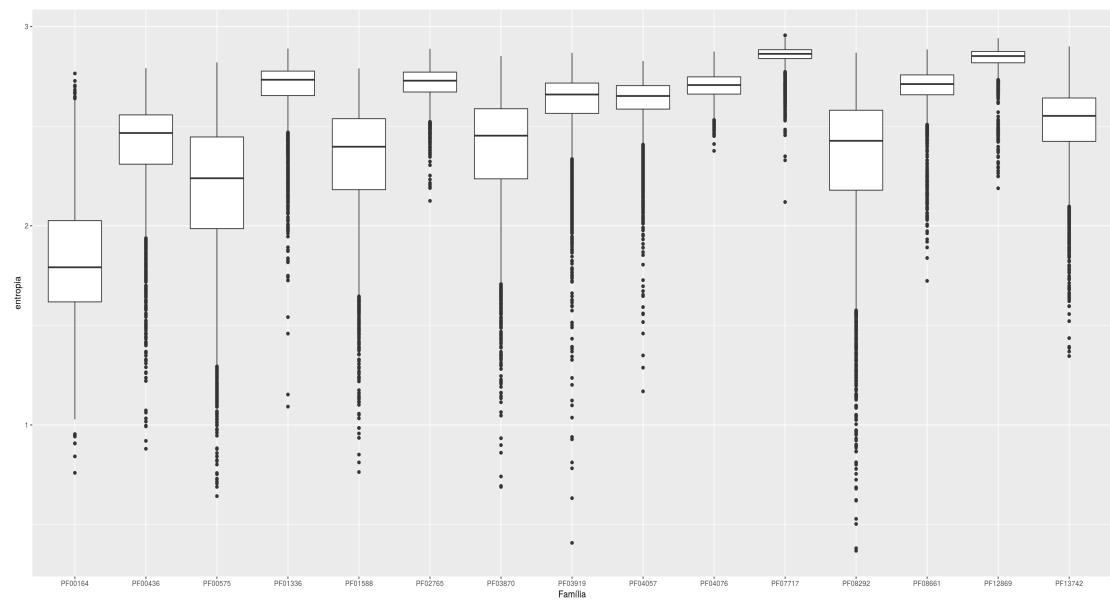
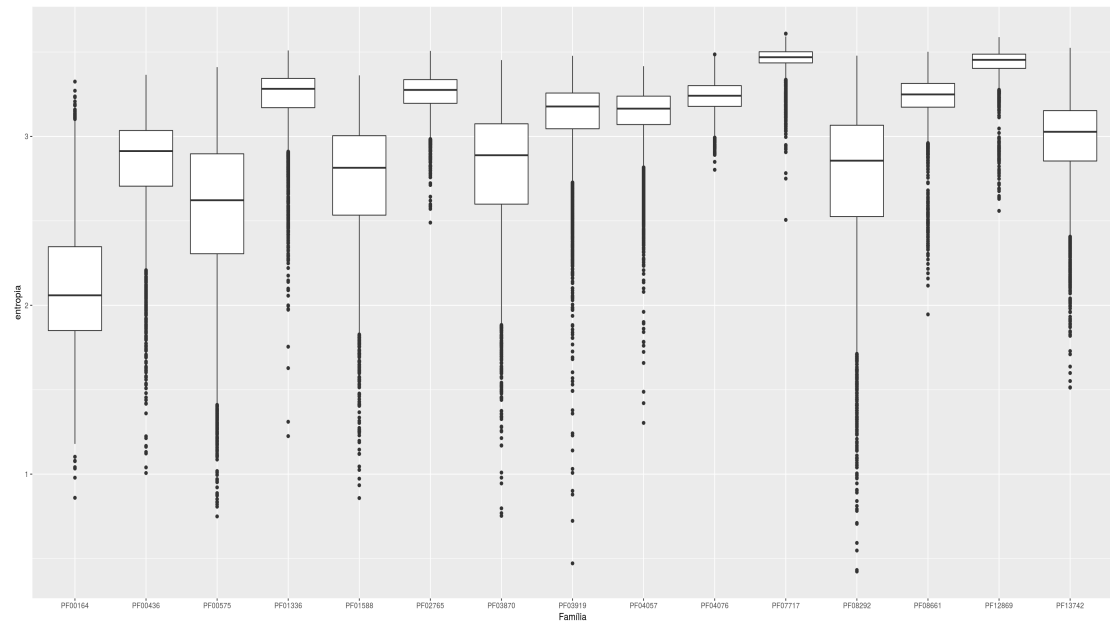
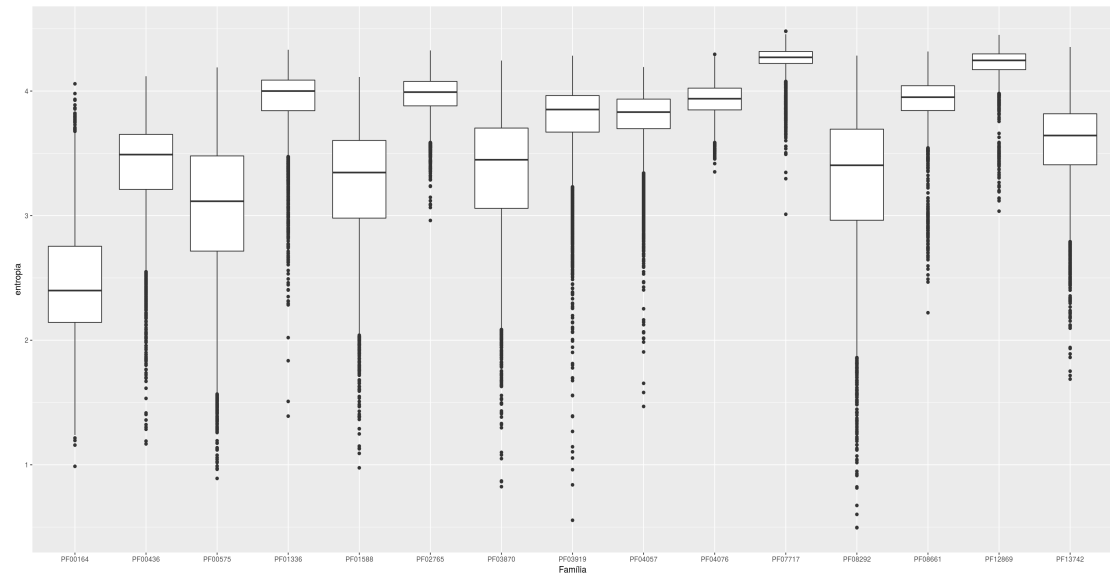


Figura 7.14 – Boxplots para as famílias do clã CL0020 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta









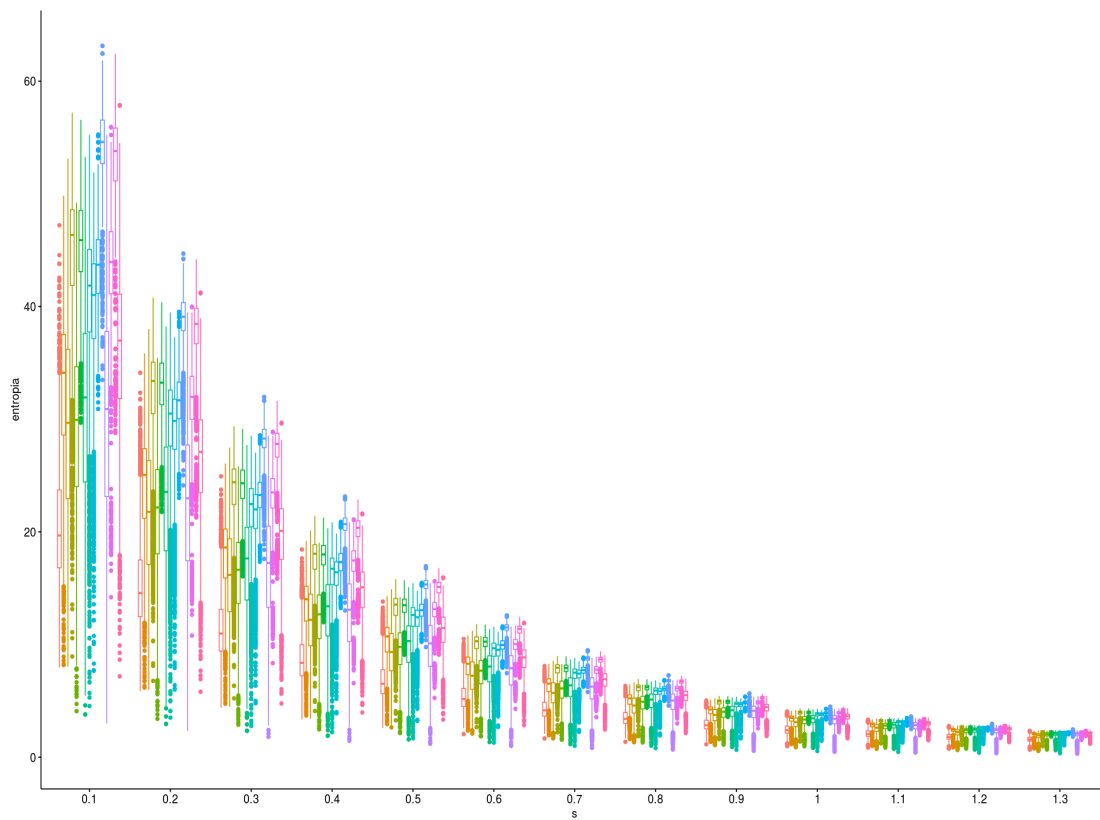
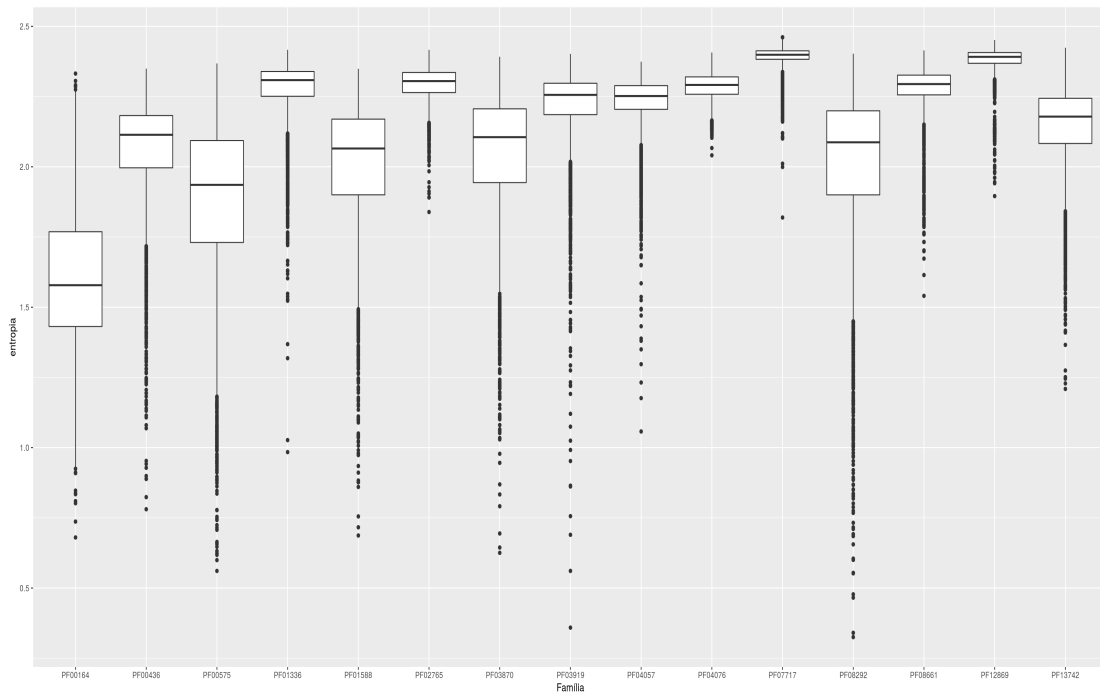
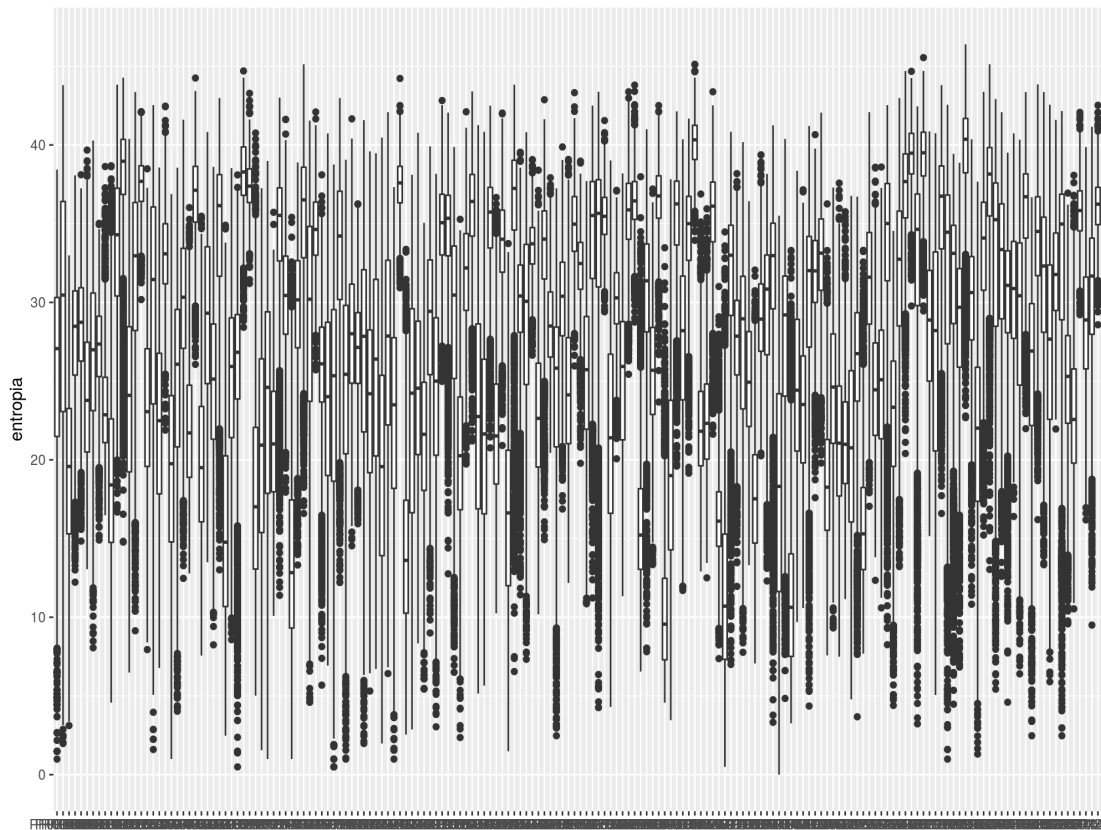
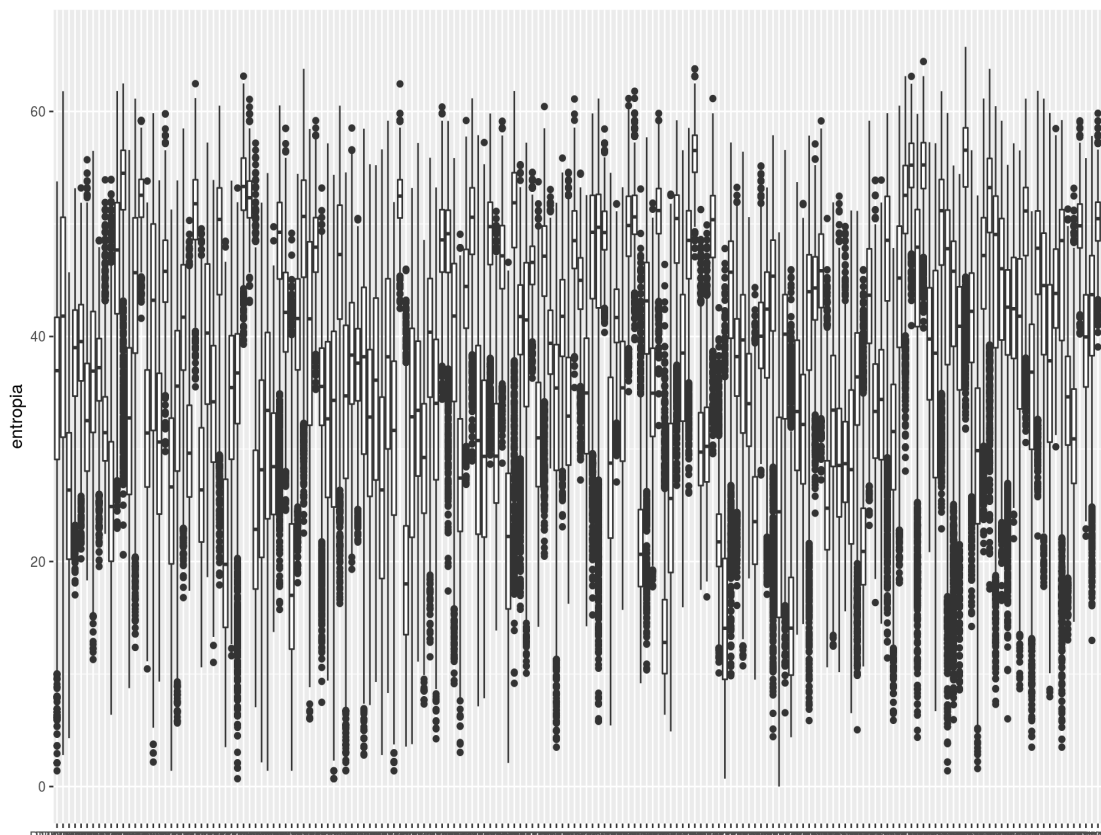


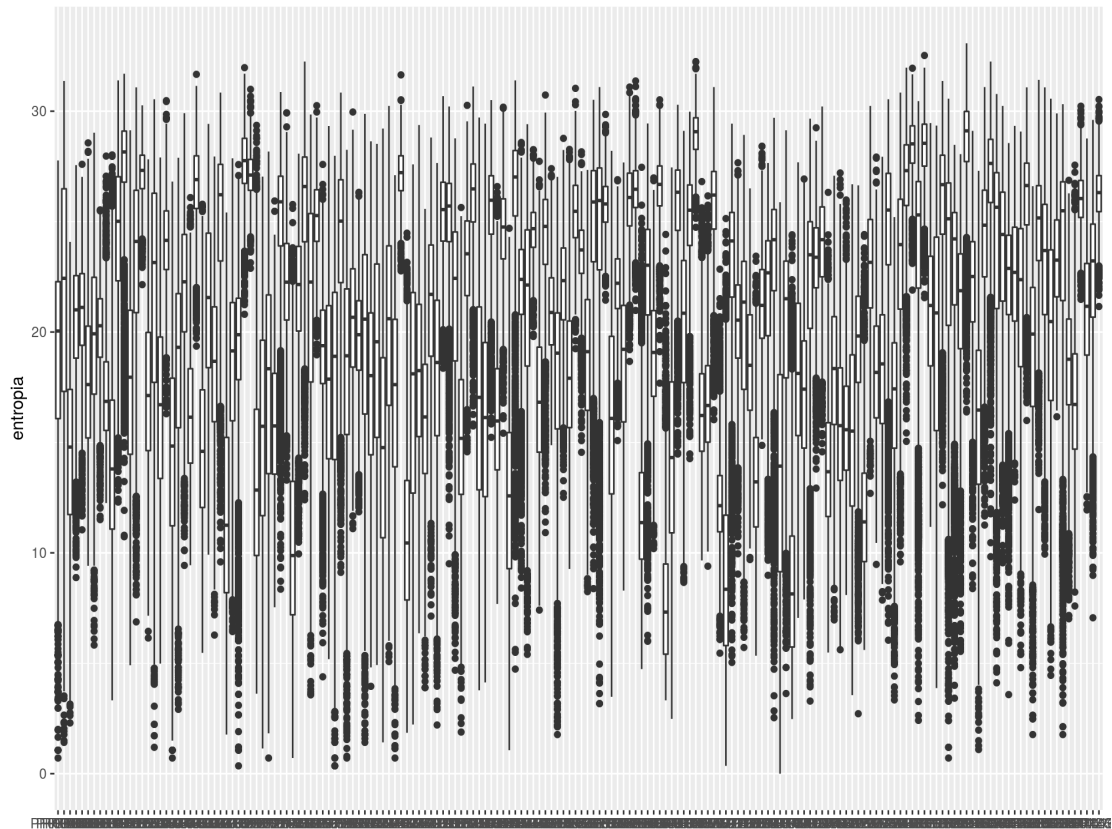
Figura 7.15 – Boxplots para as famílias do clã CL0021 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta



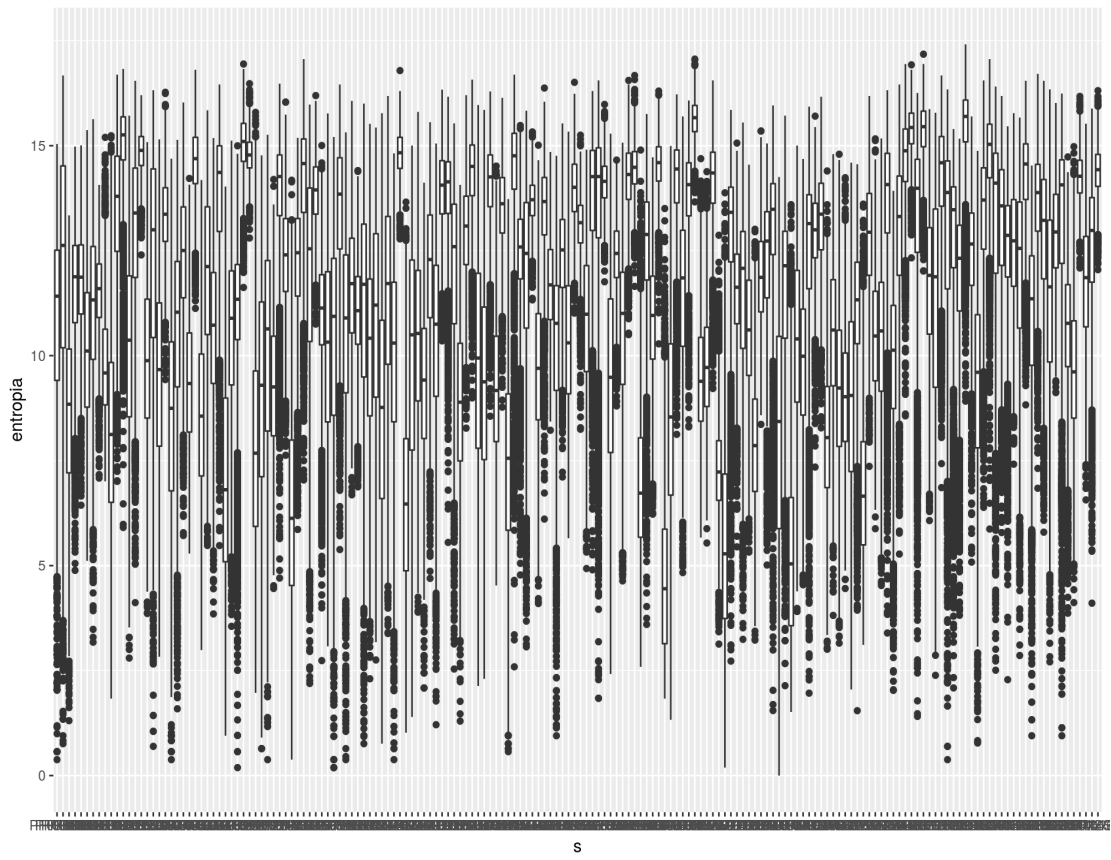
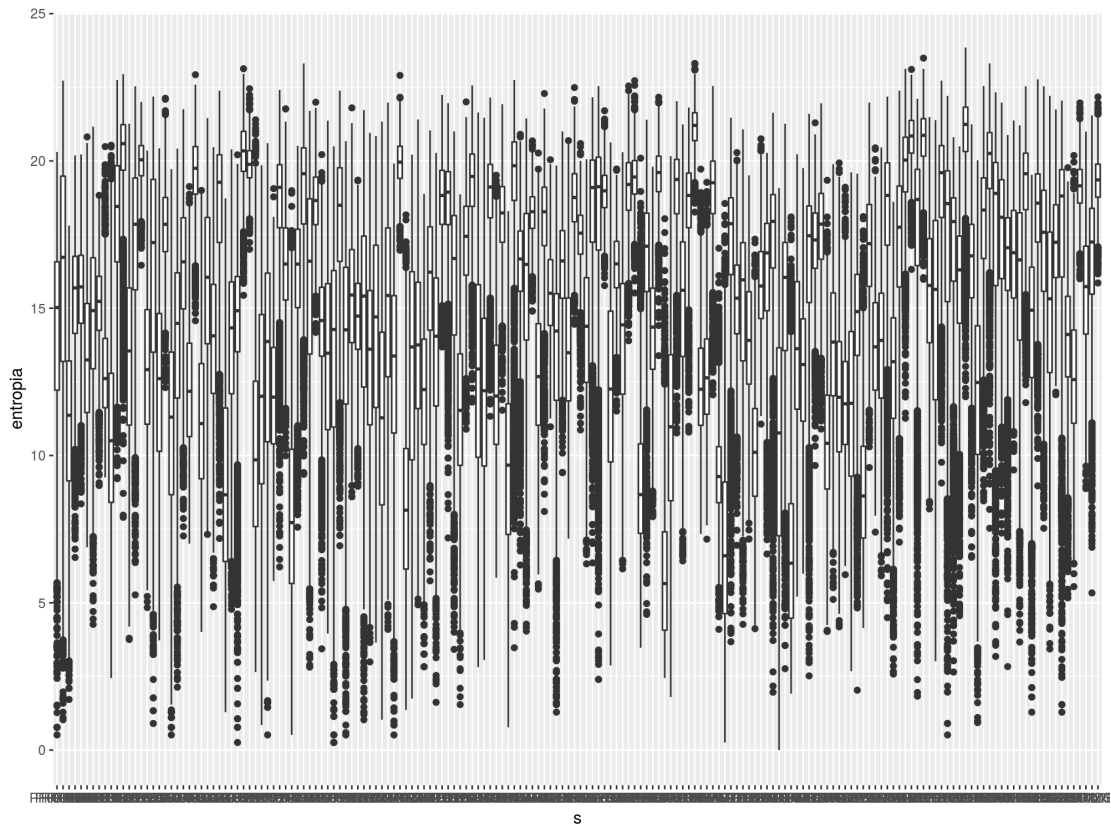
s

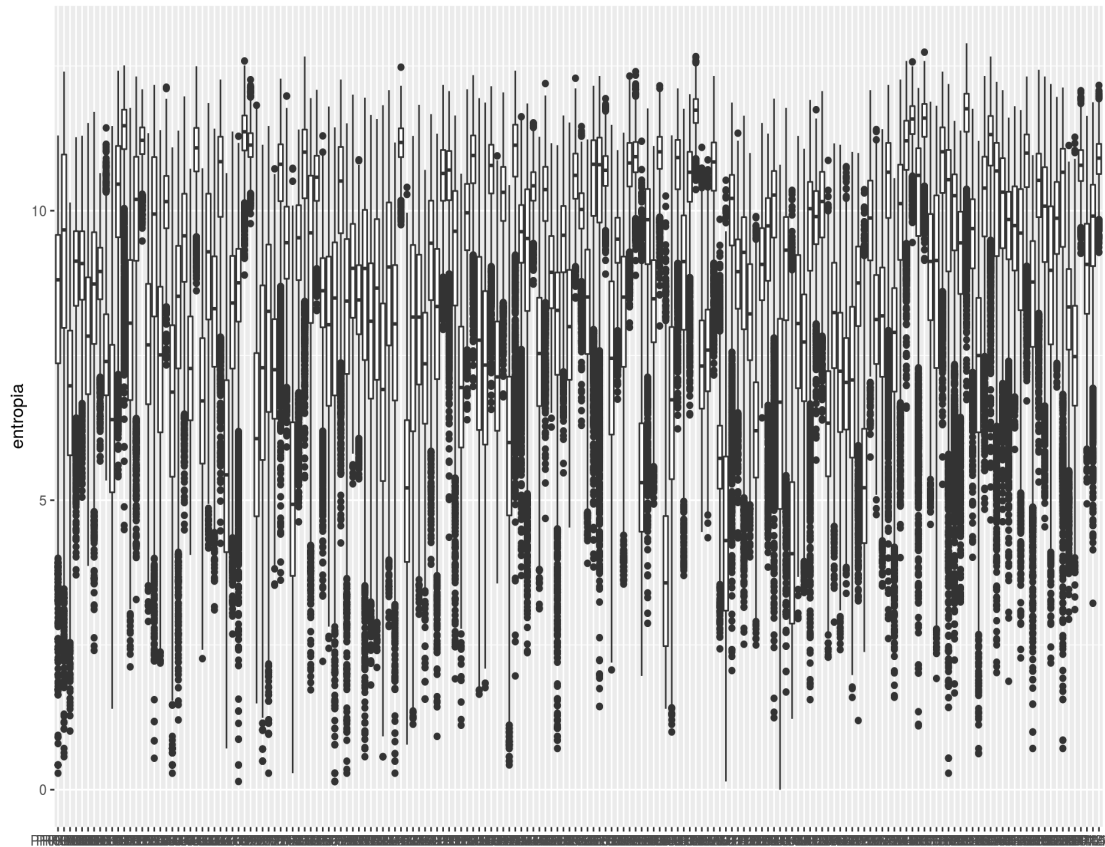


s

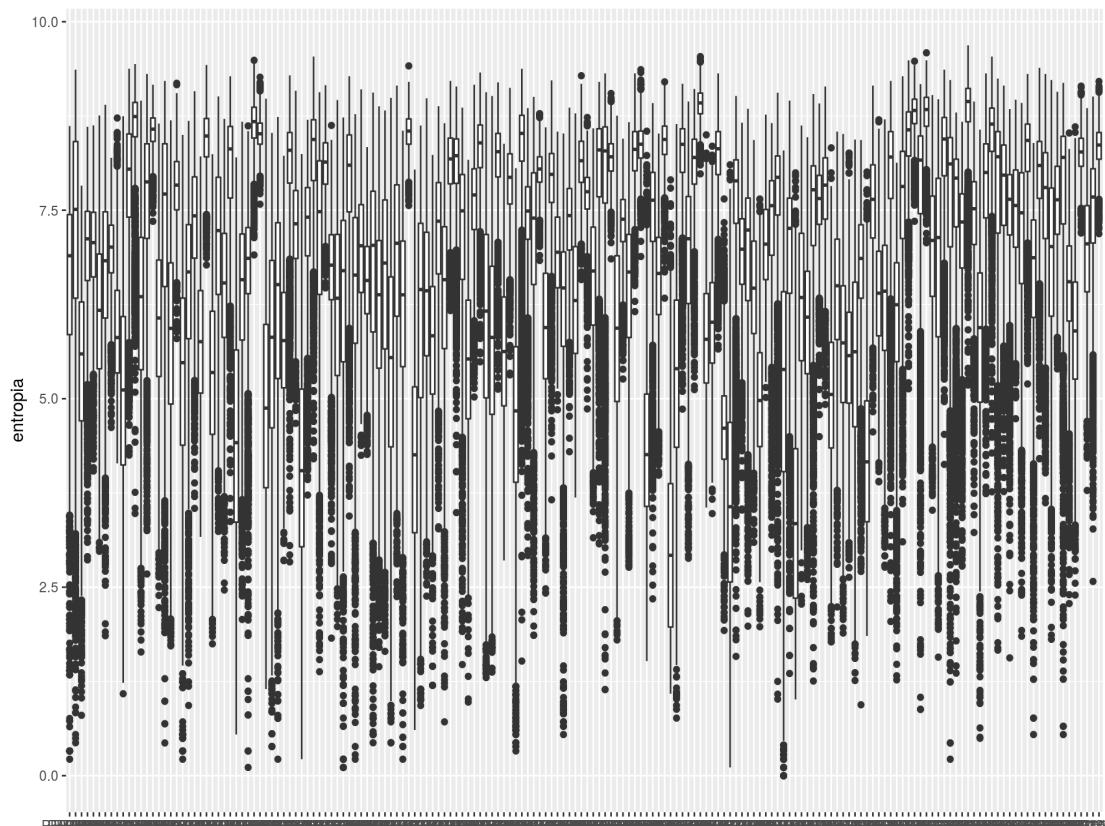


s

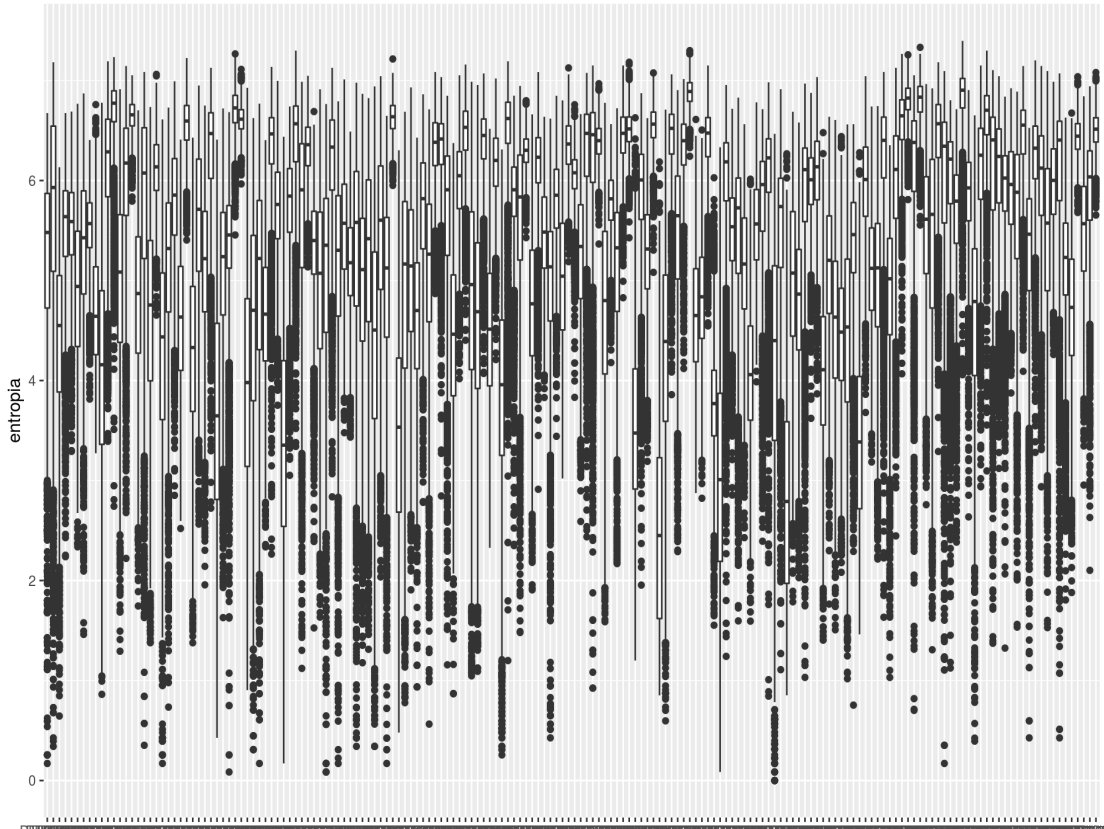




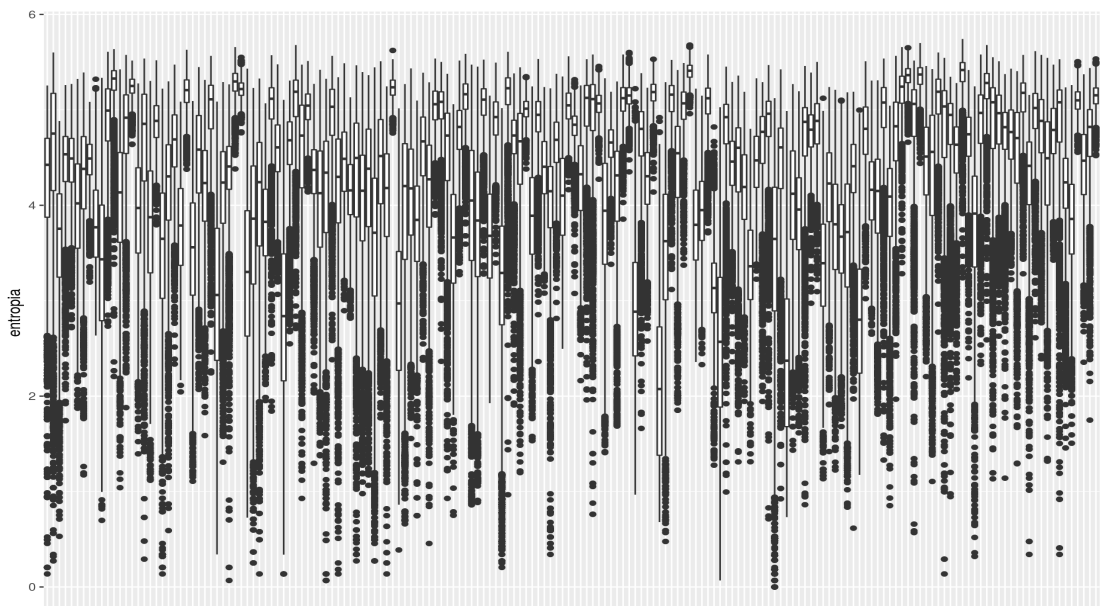
s



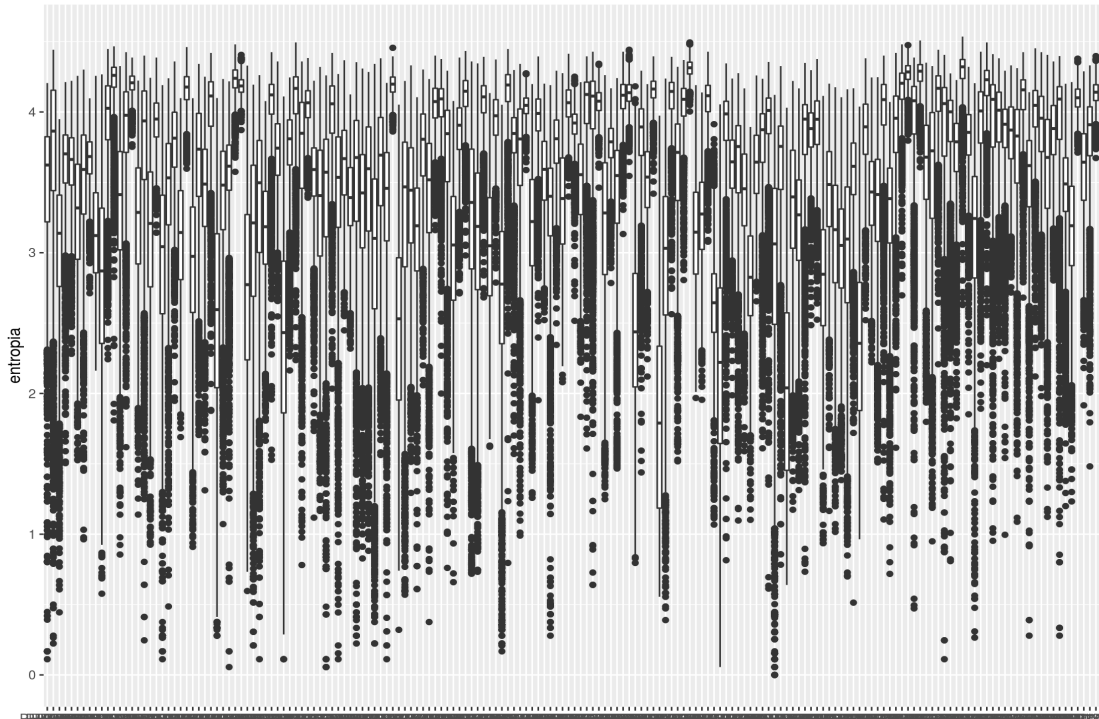
s



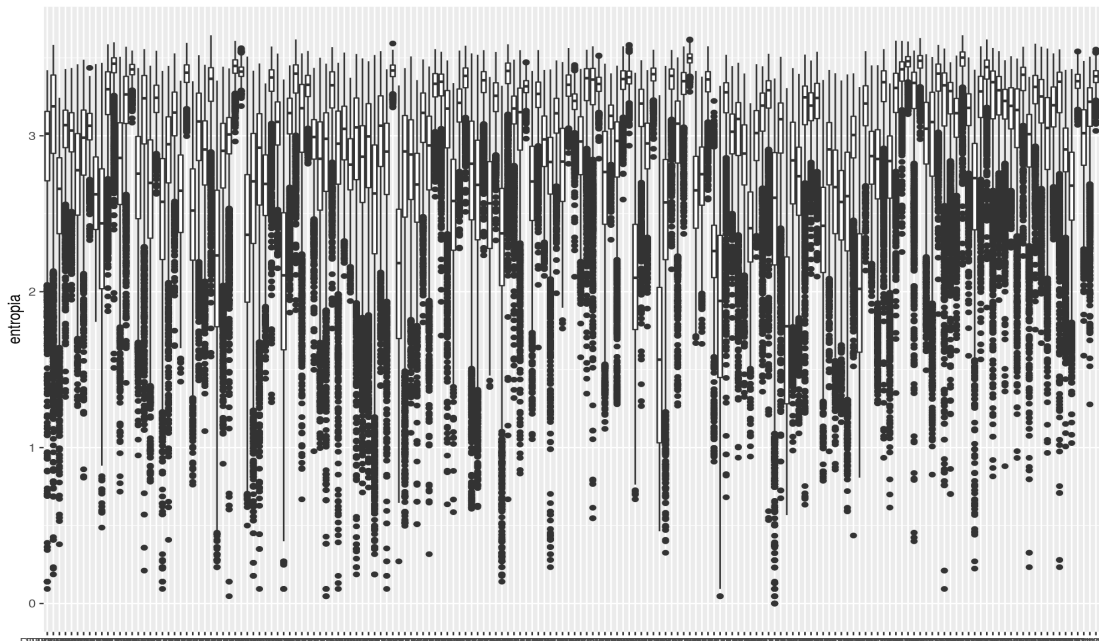
S



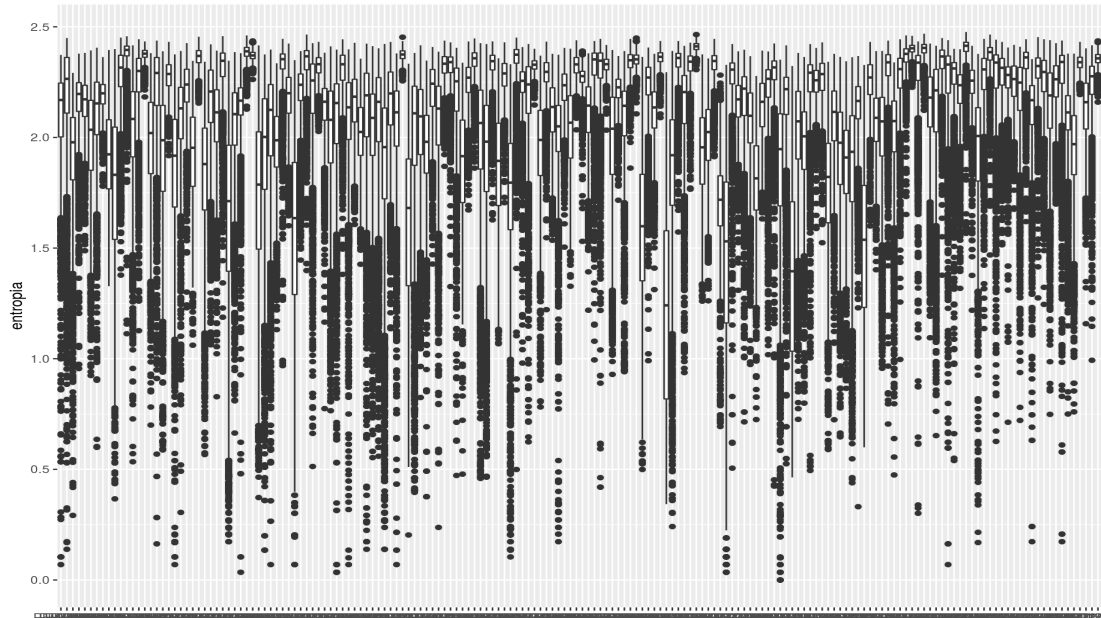
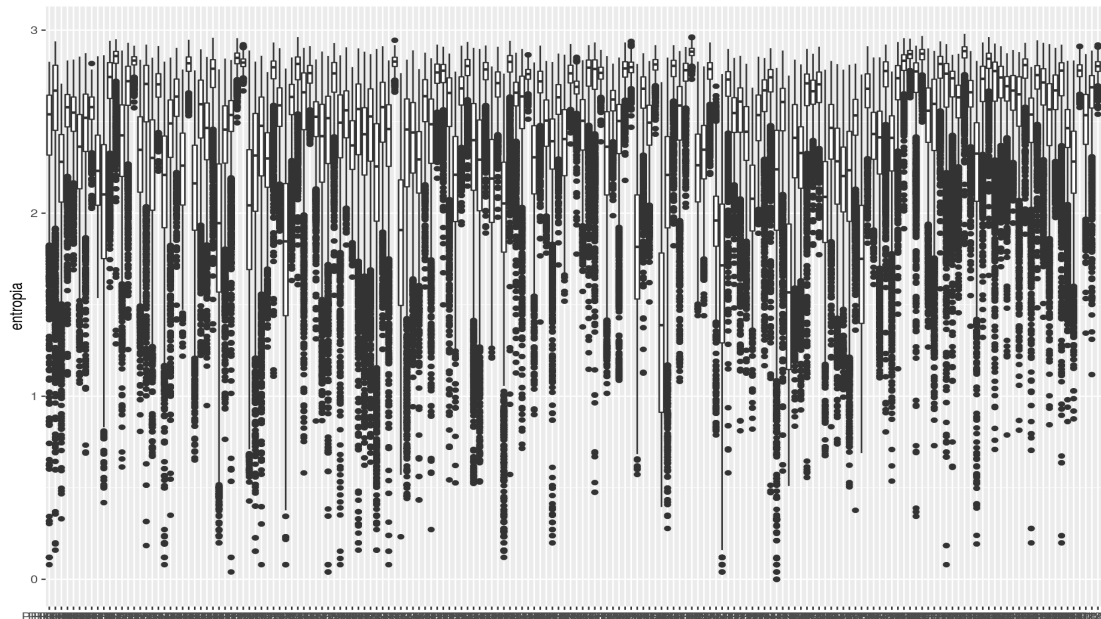
S



S



S



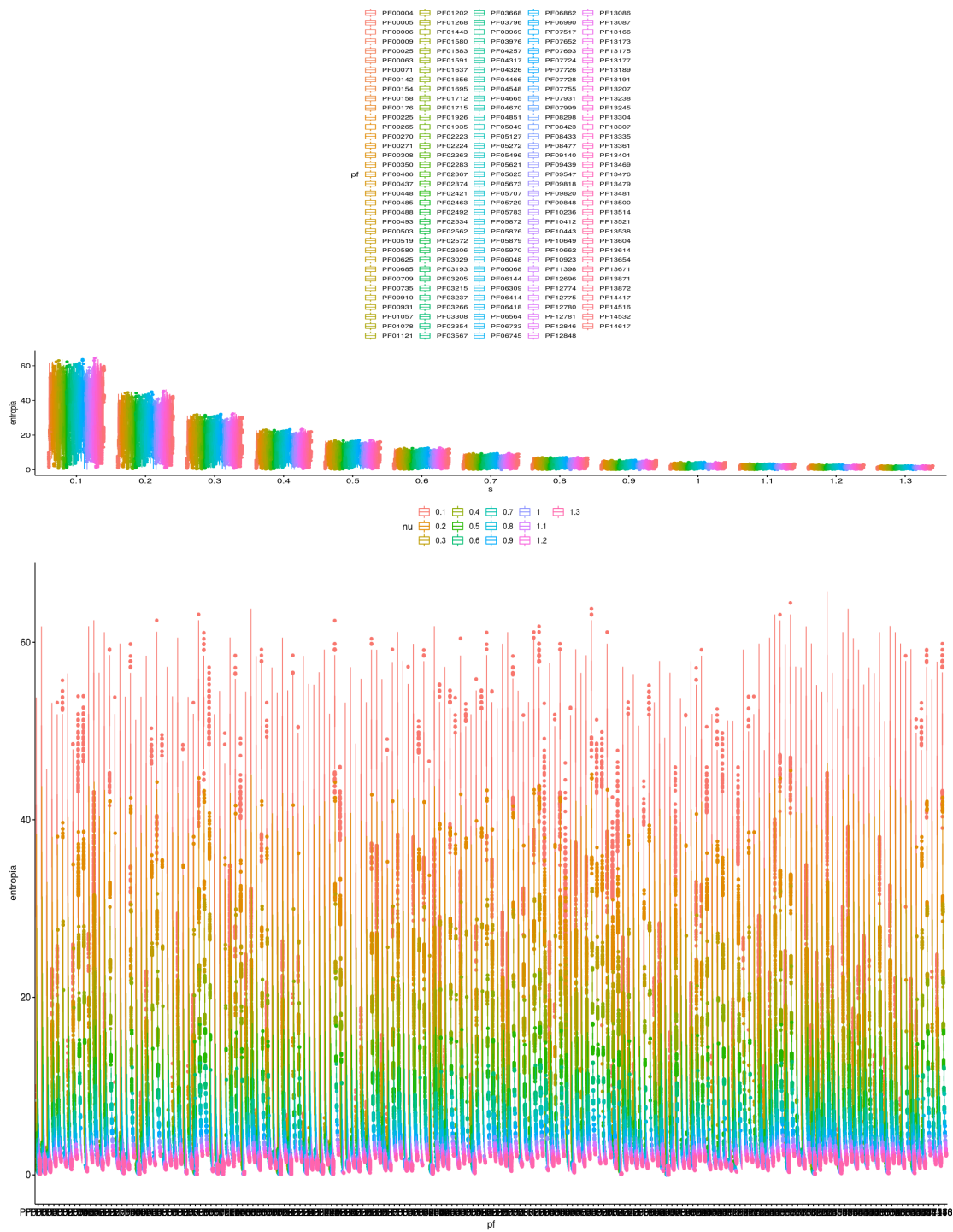
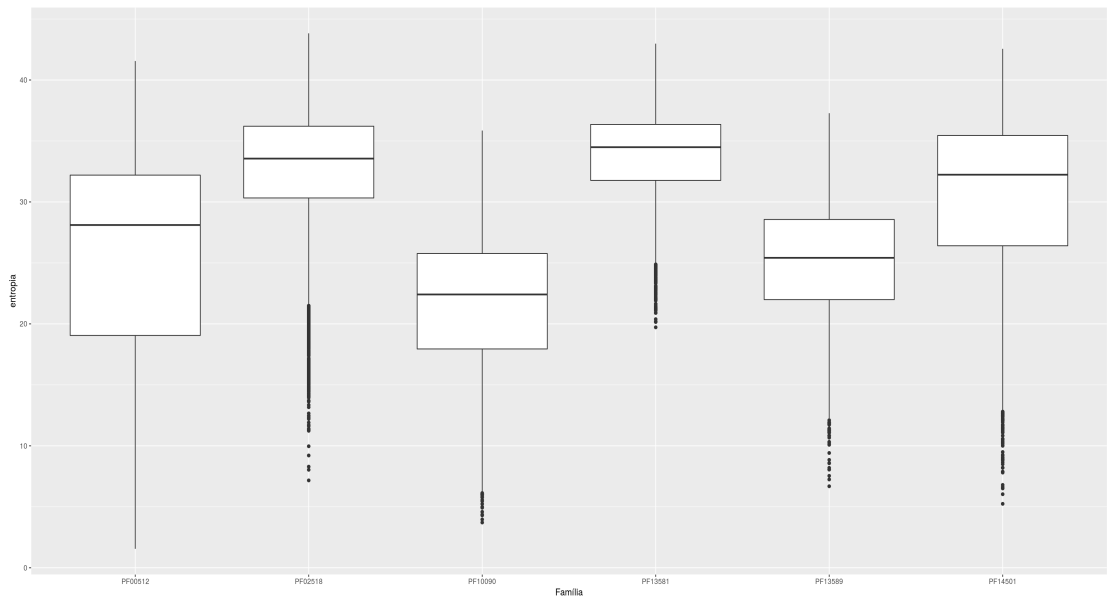
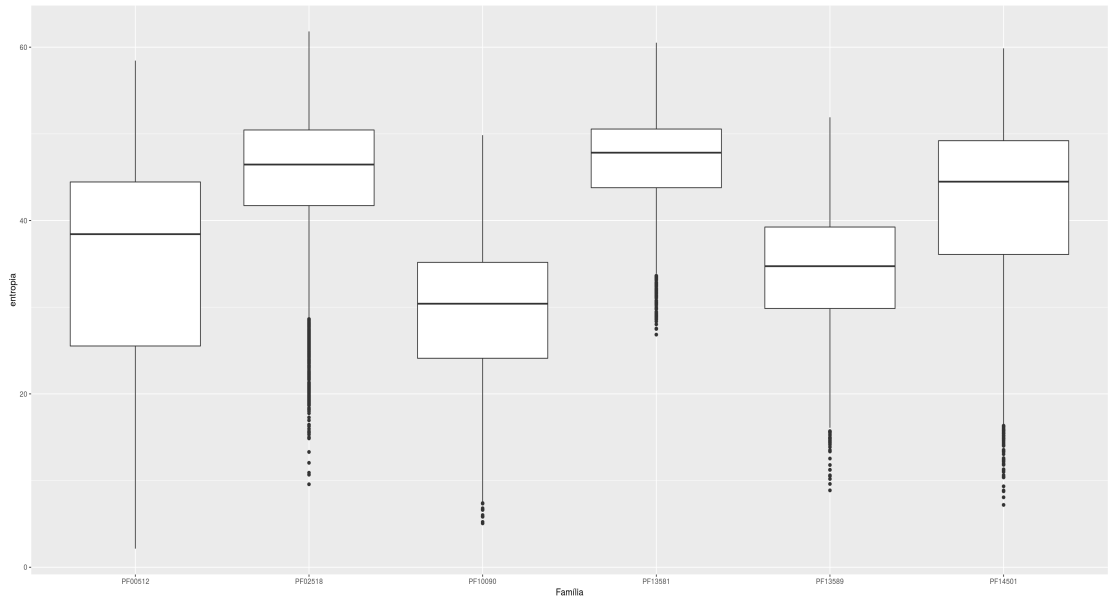
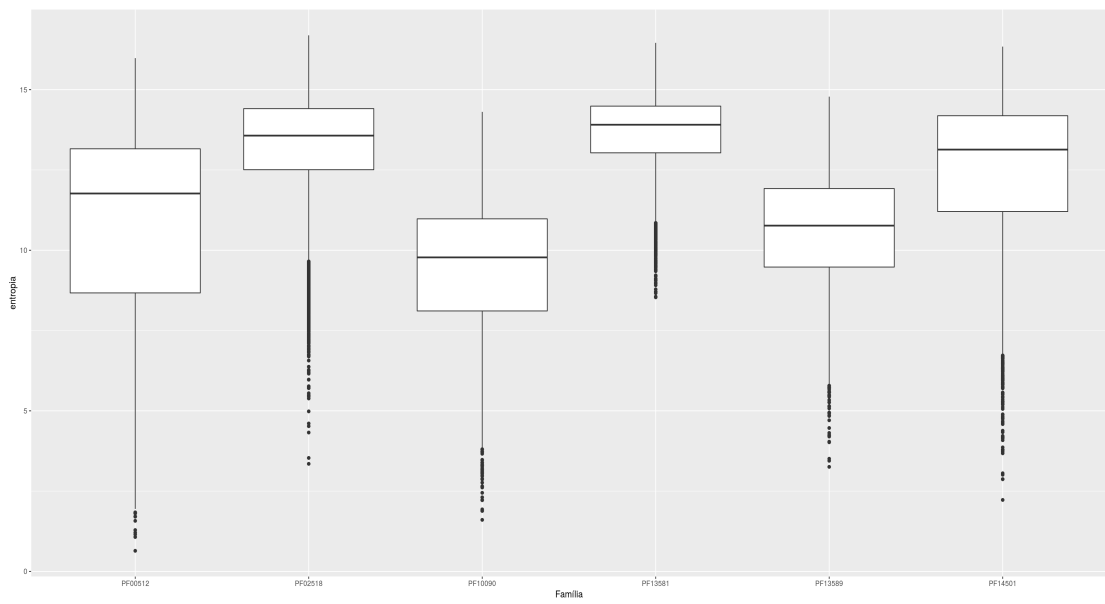
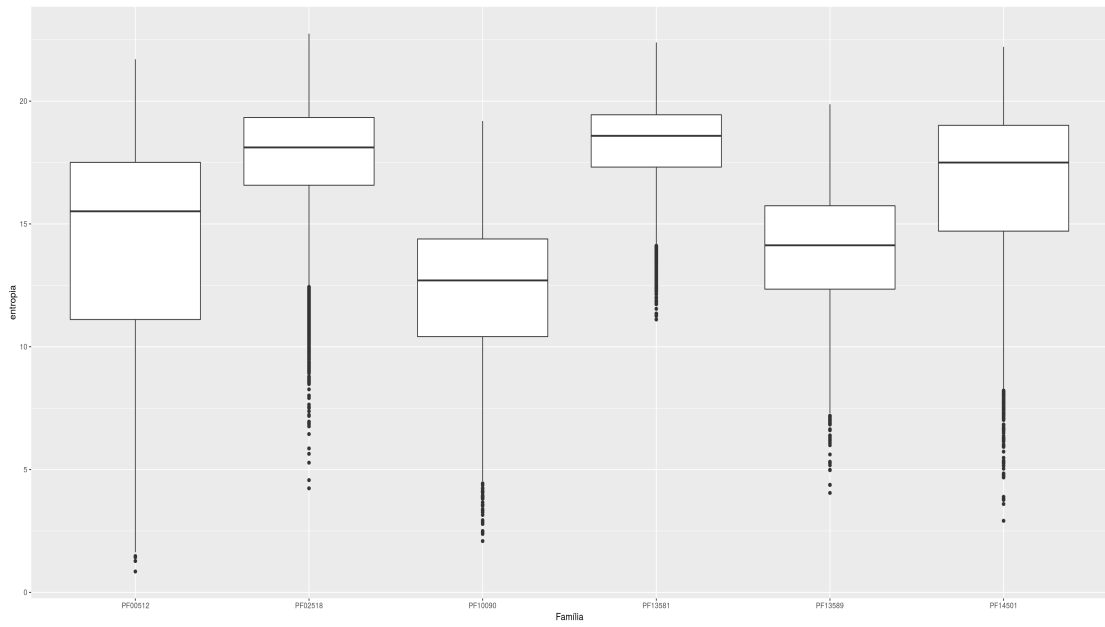
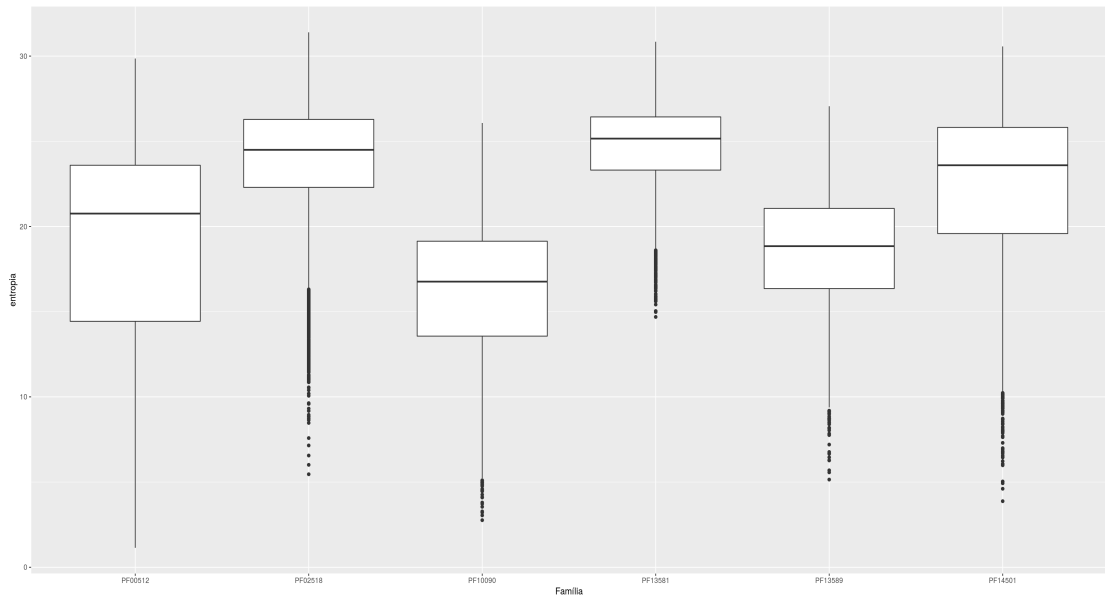
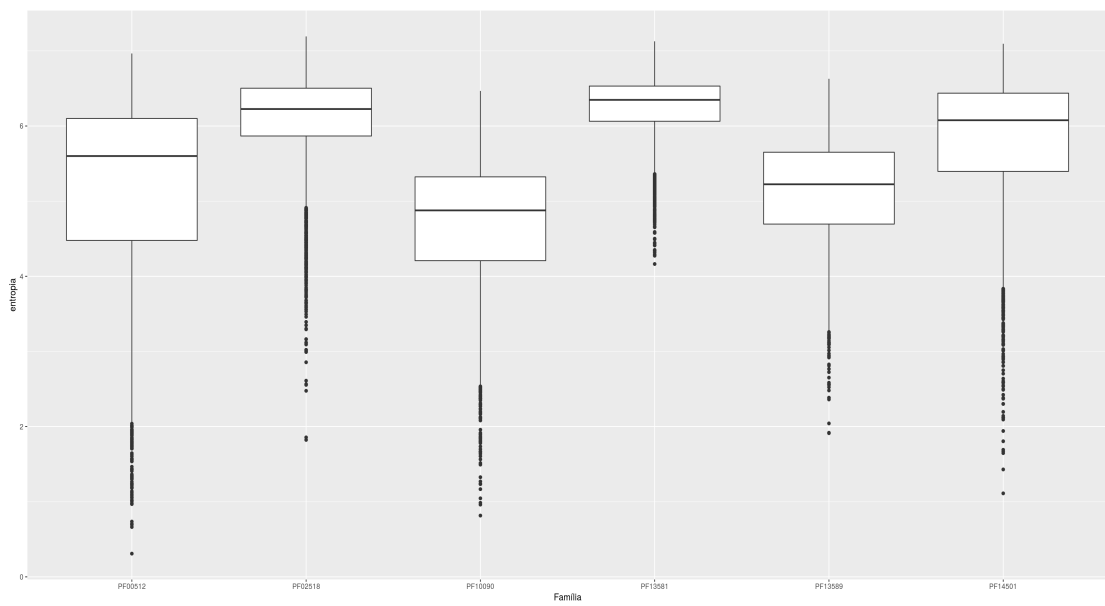
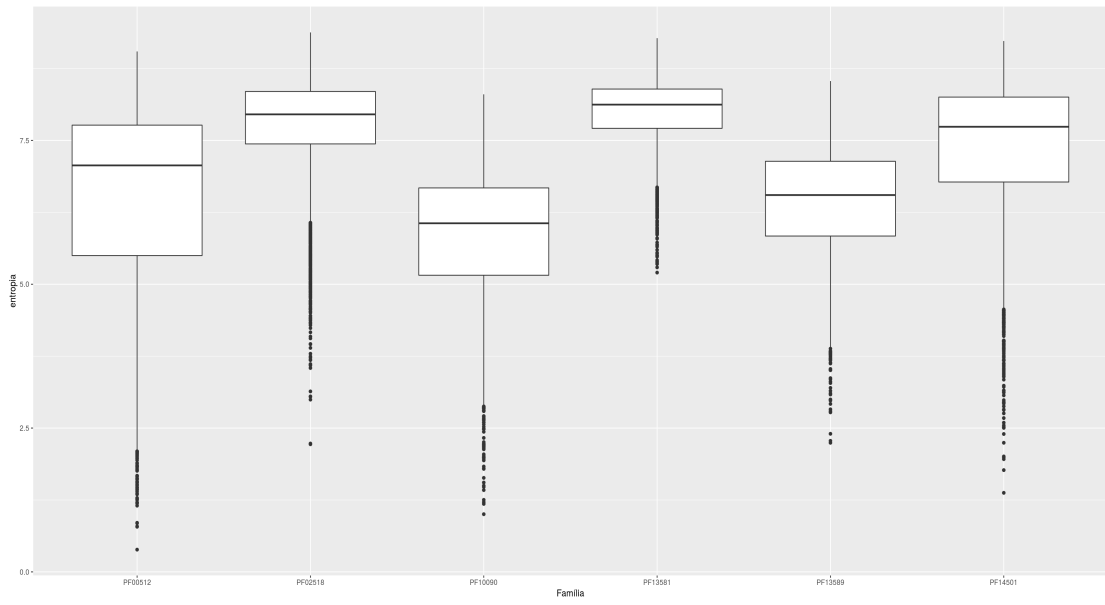
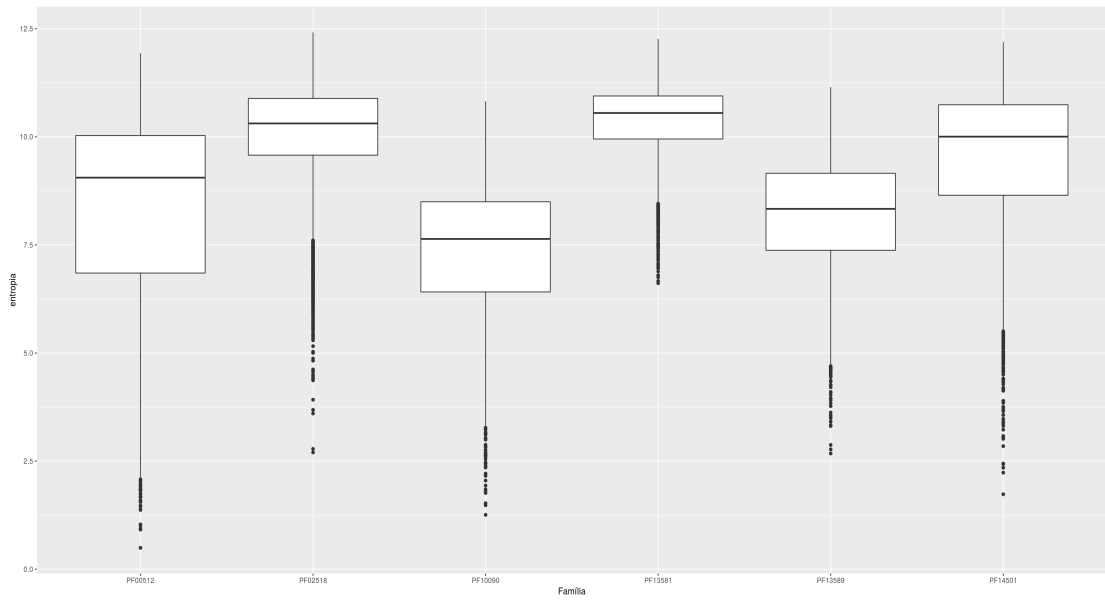
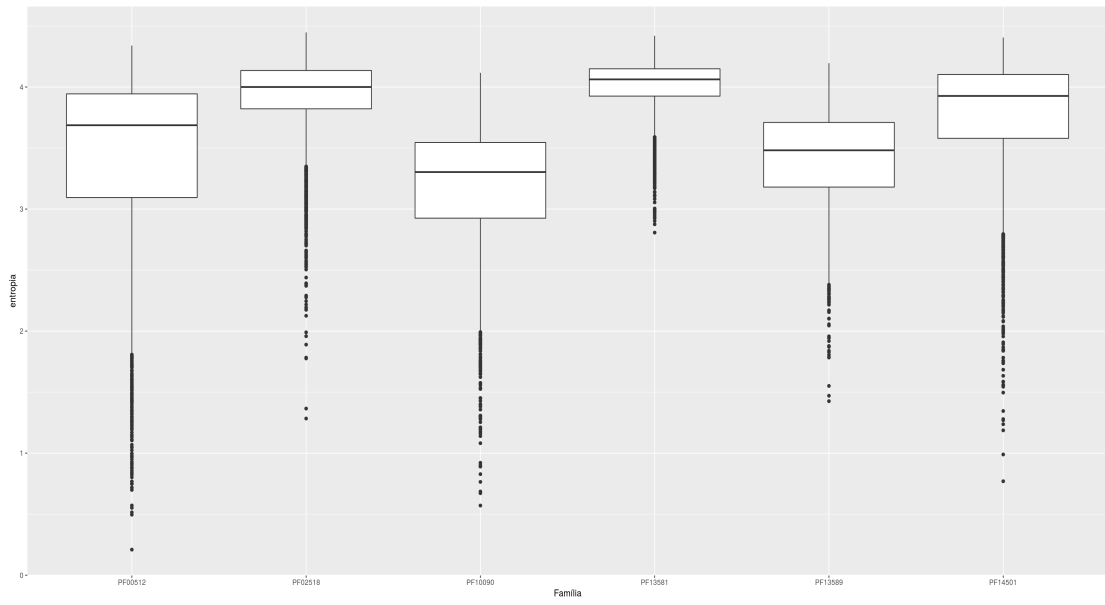
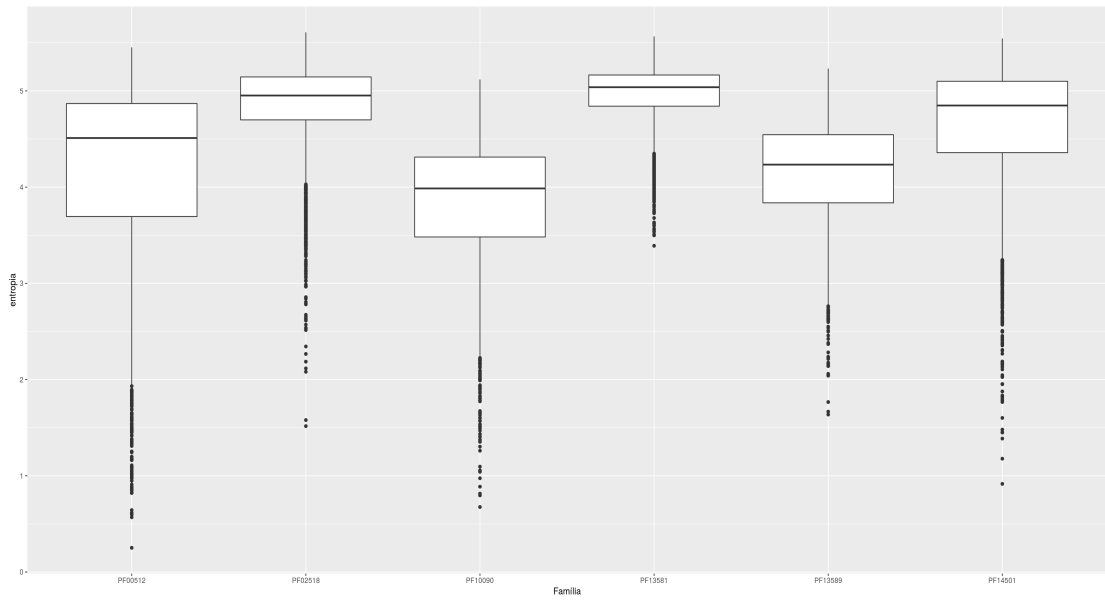


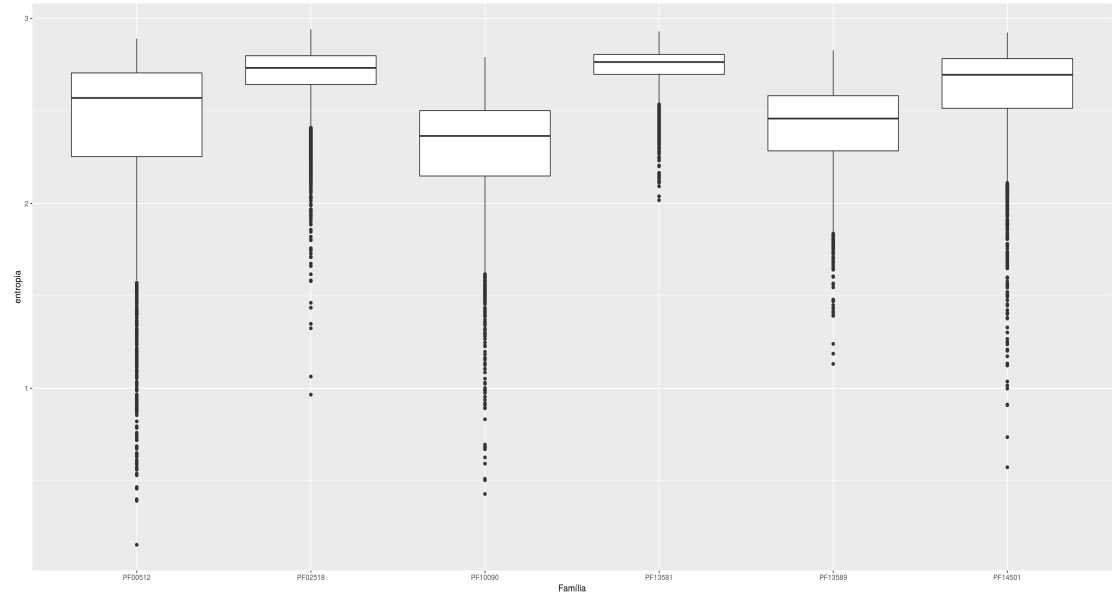
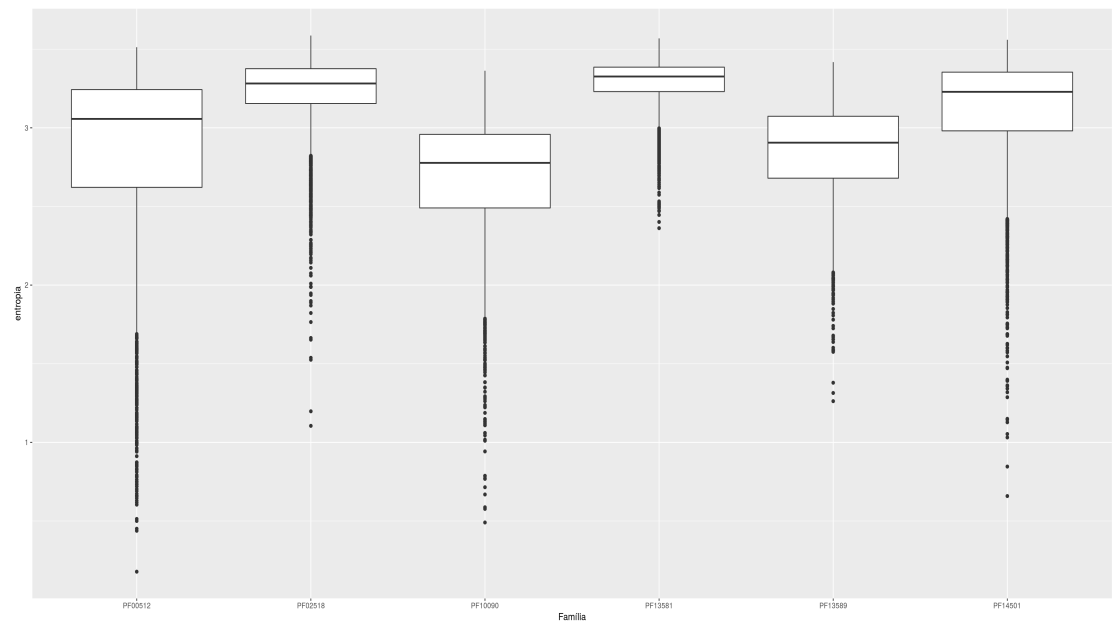
Figura 7.16 – Boxplots para as famílias do clã CL0023 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta

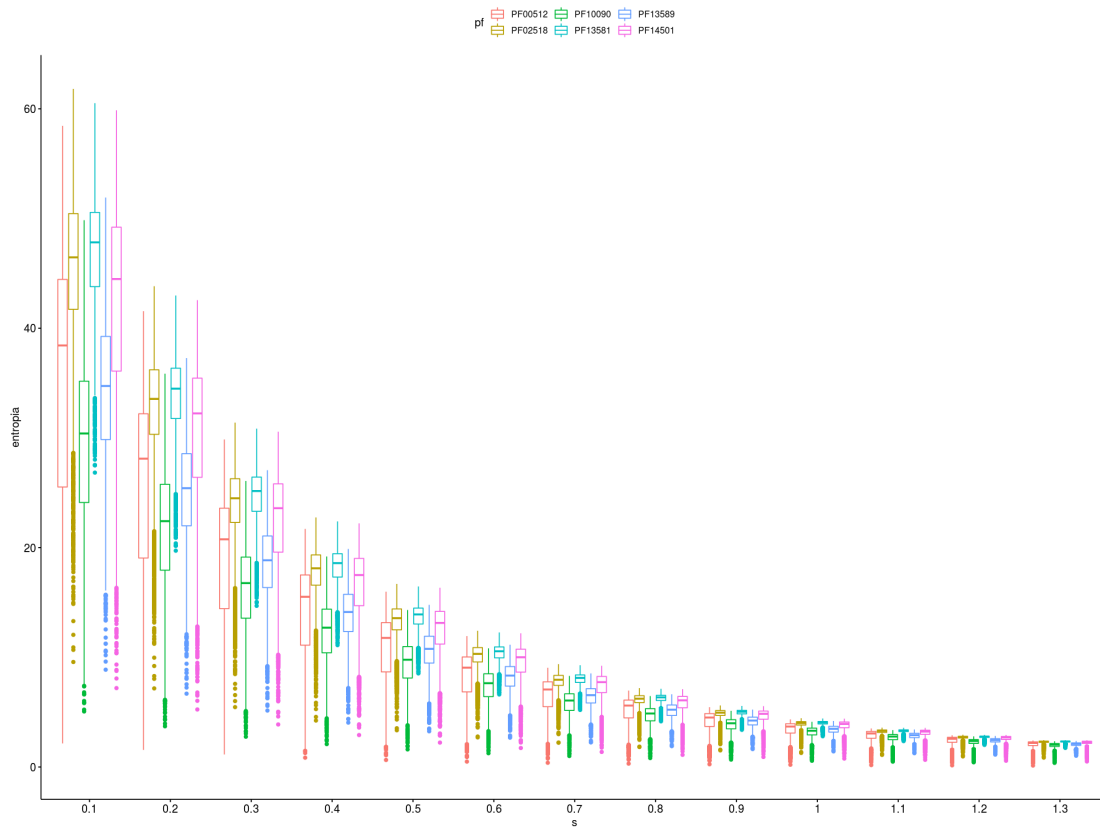
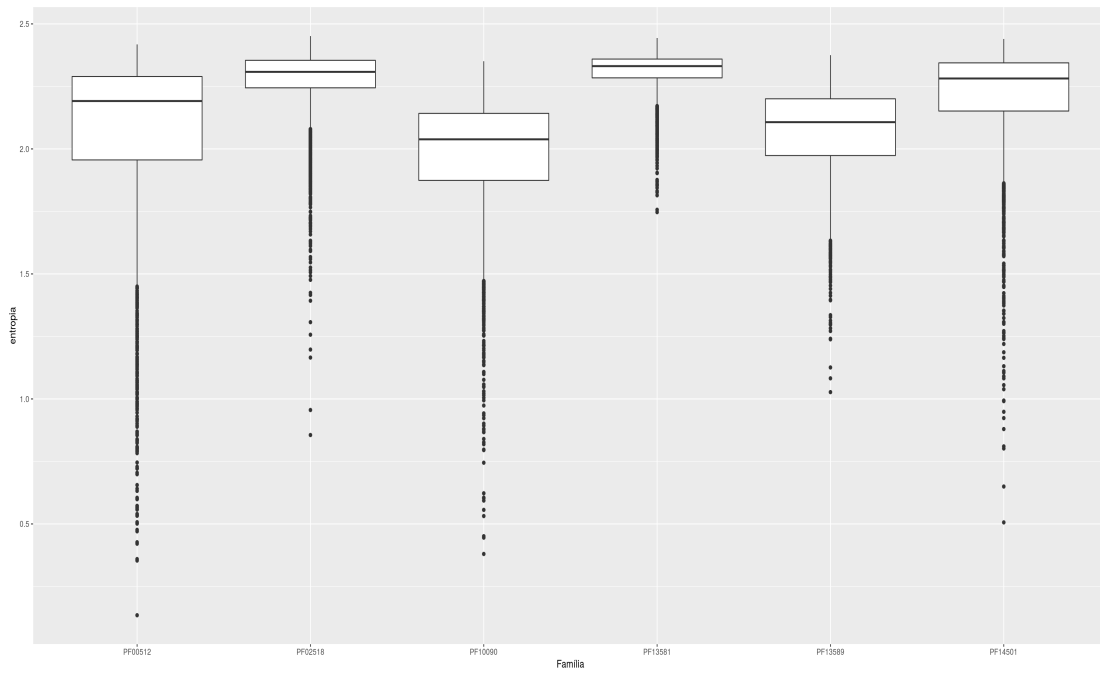












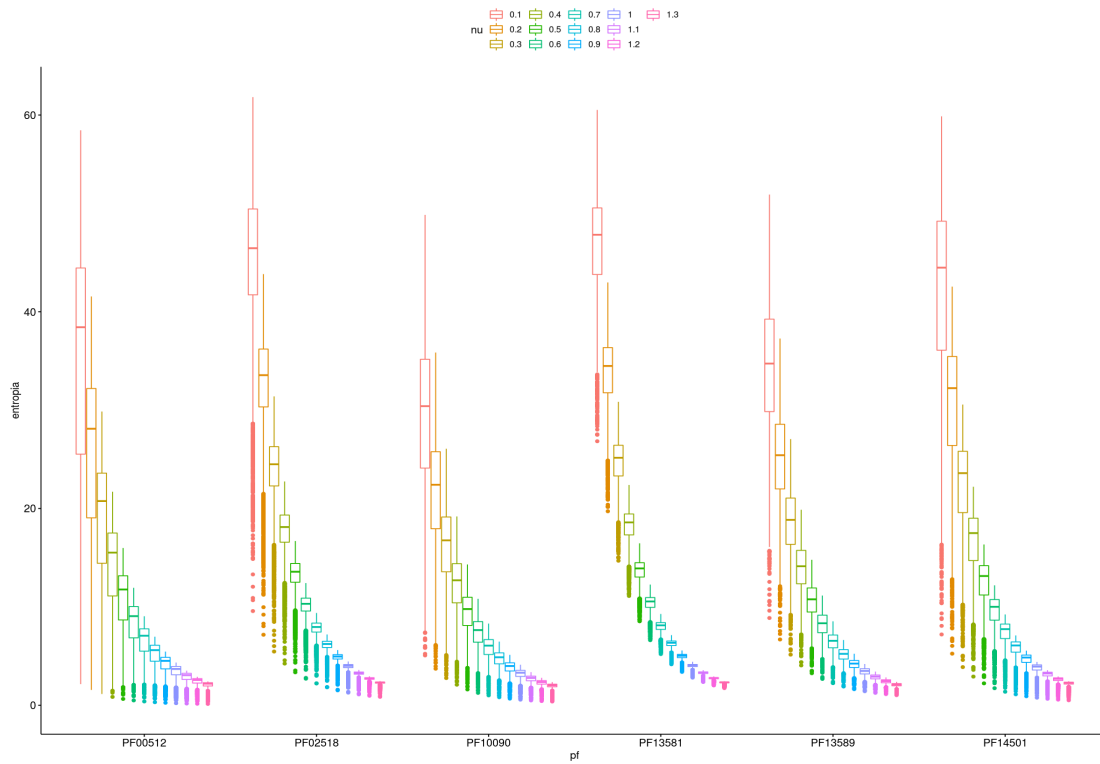
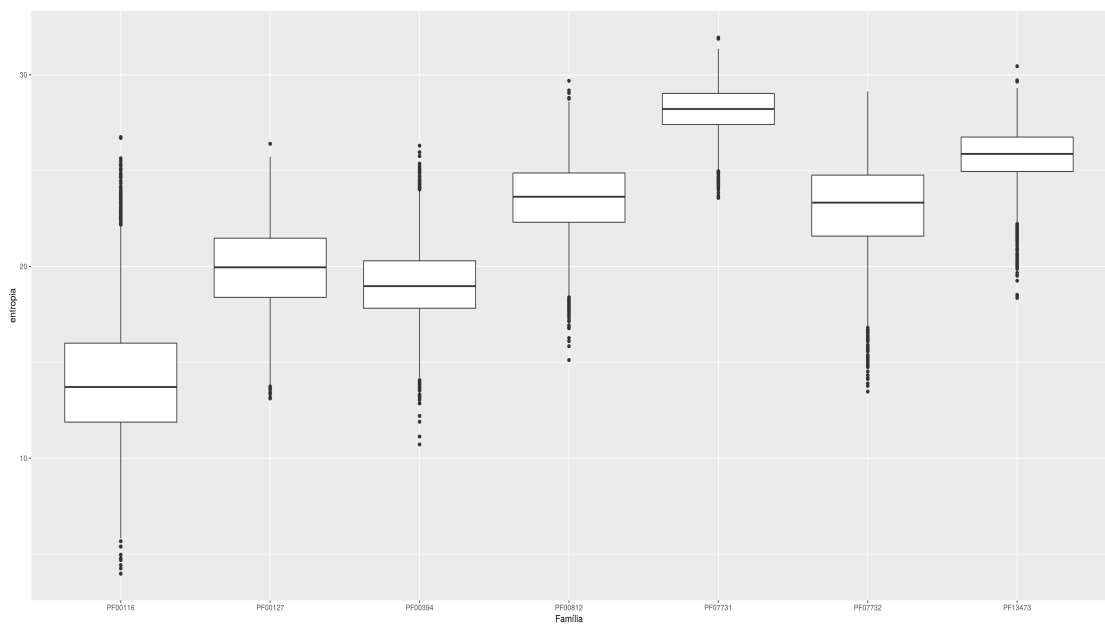
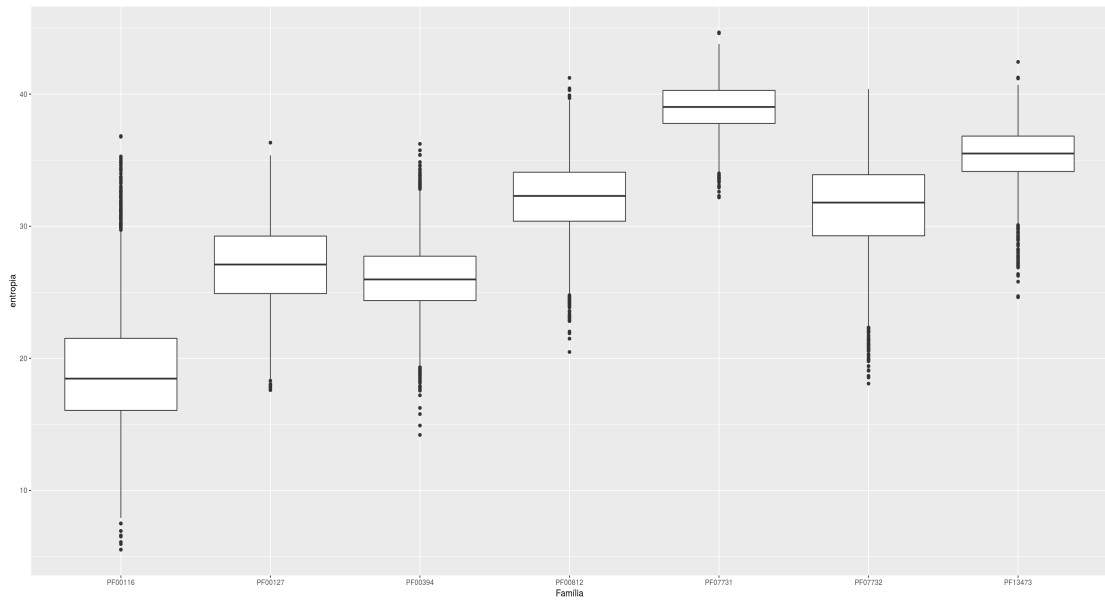
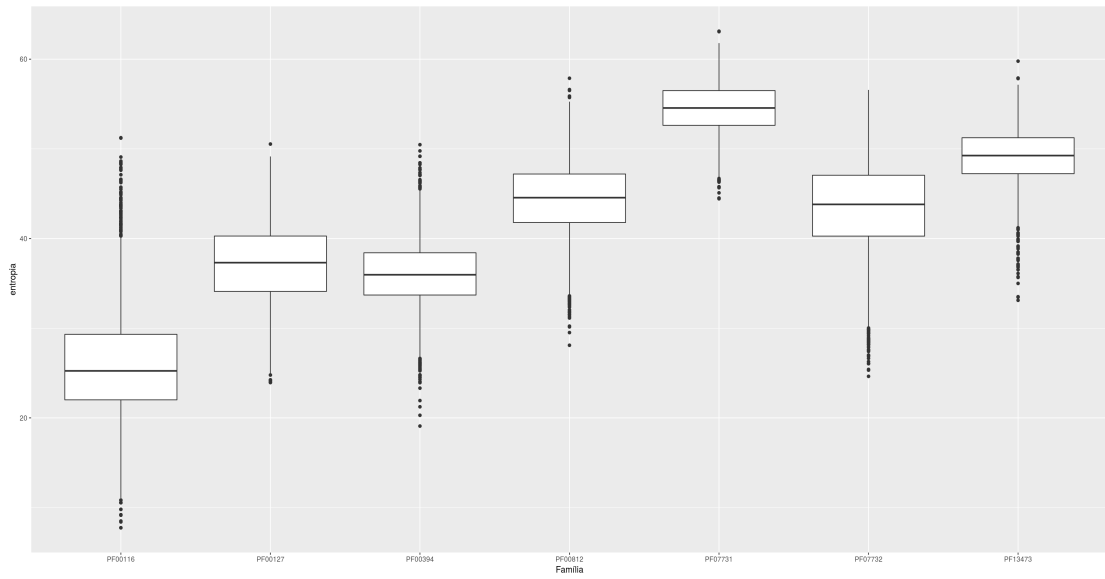
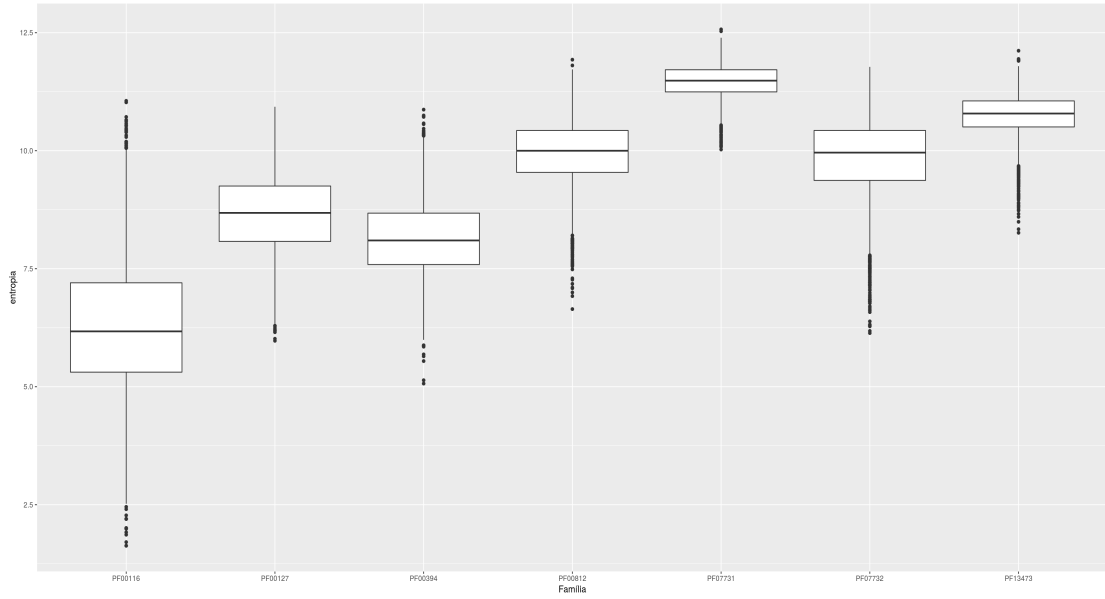
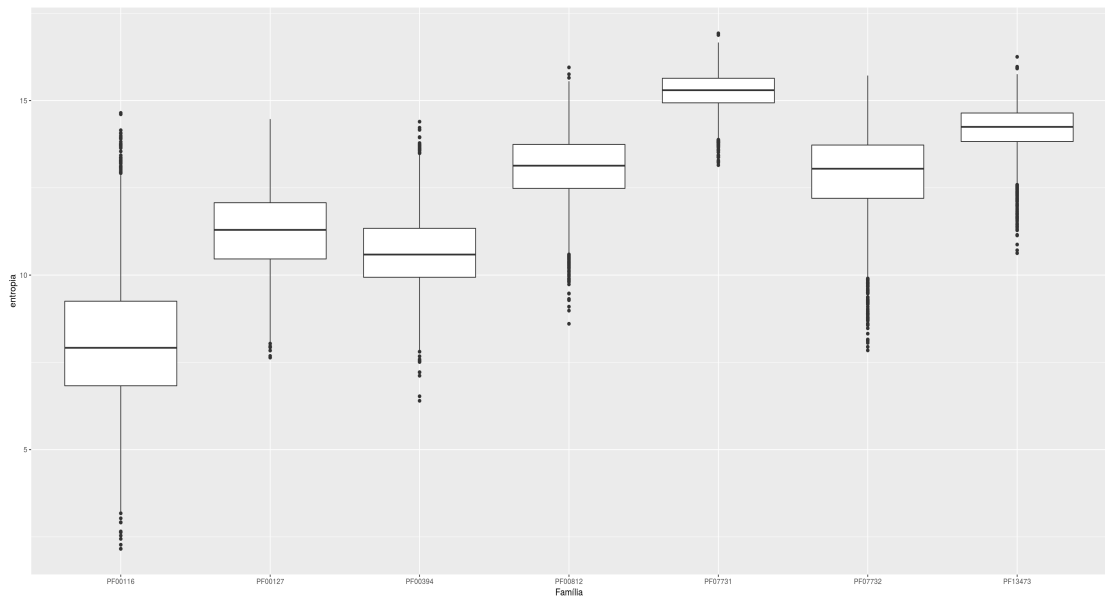
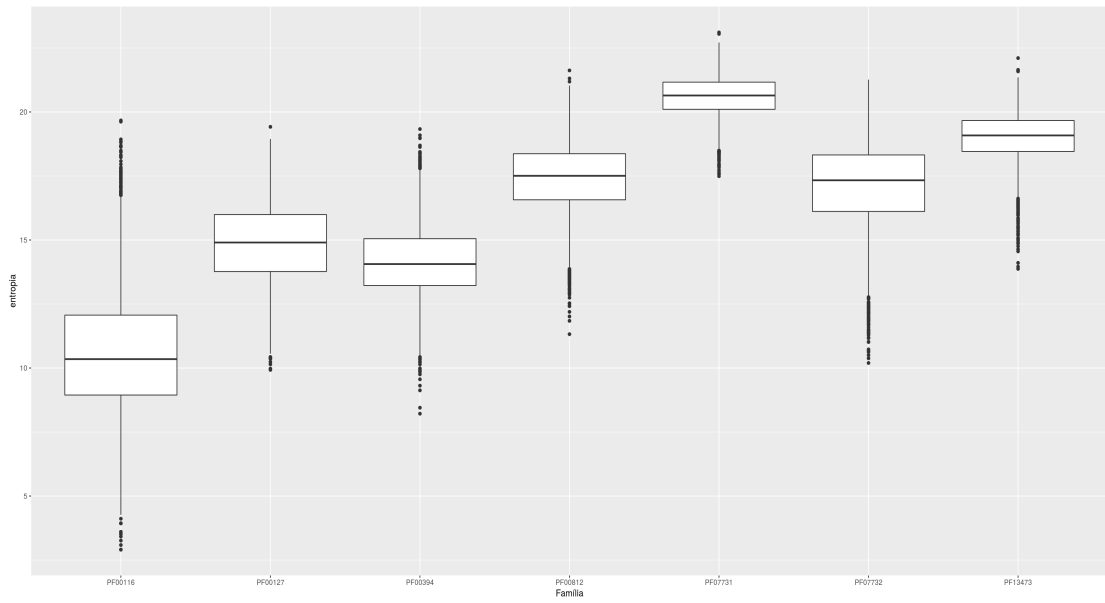
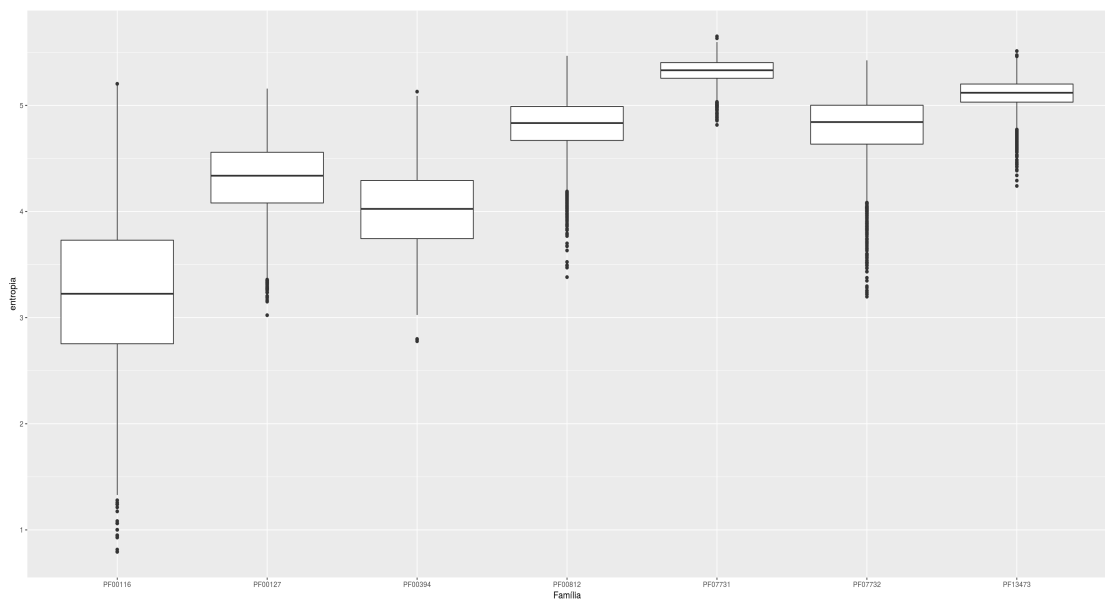
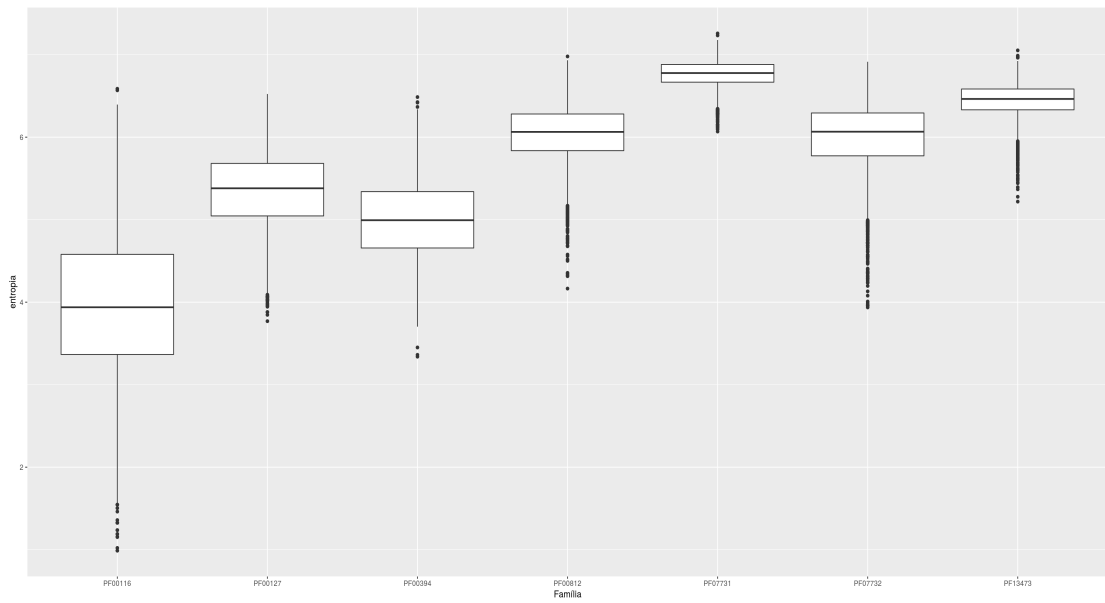
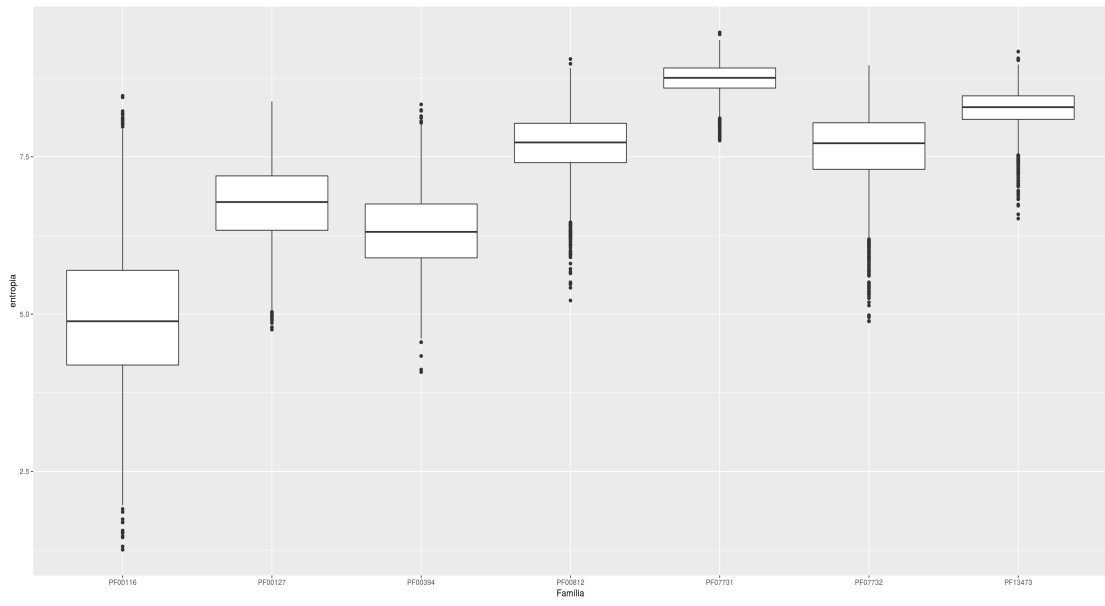
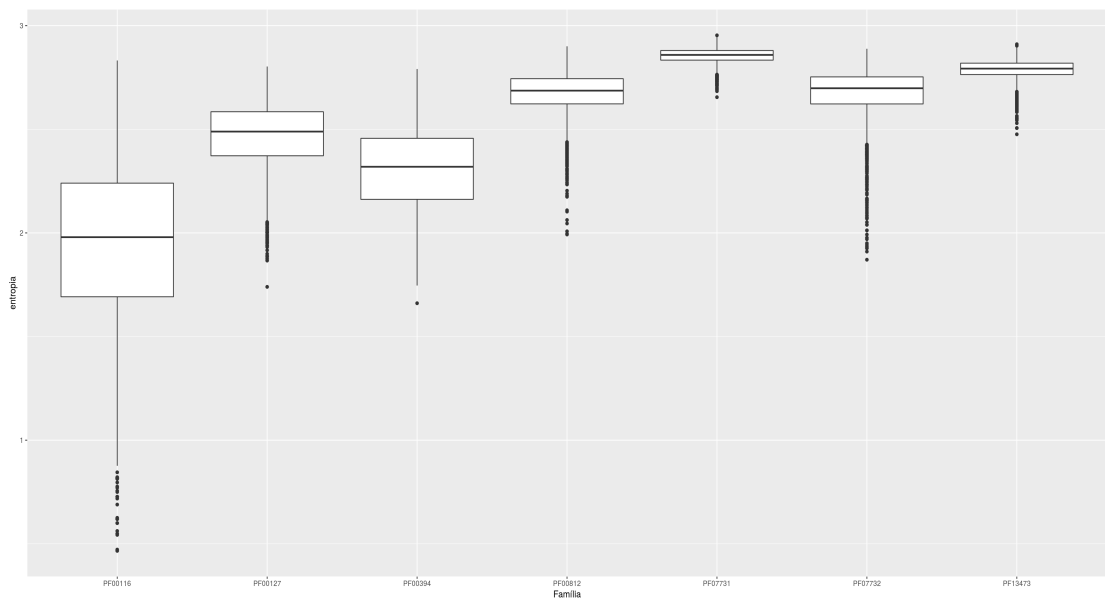
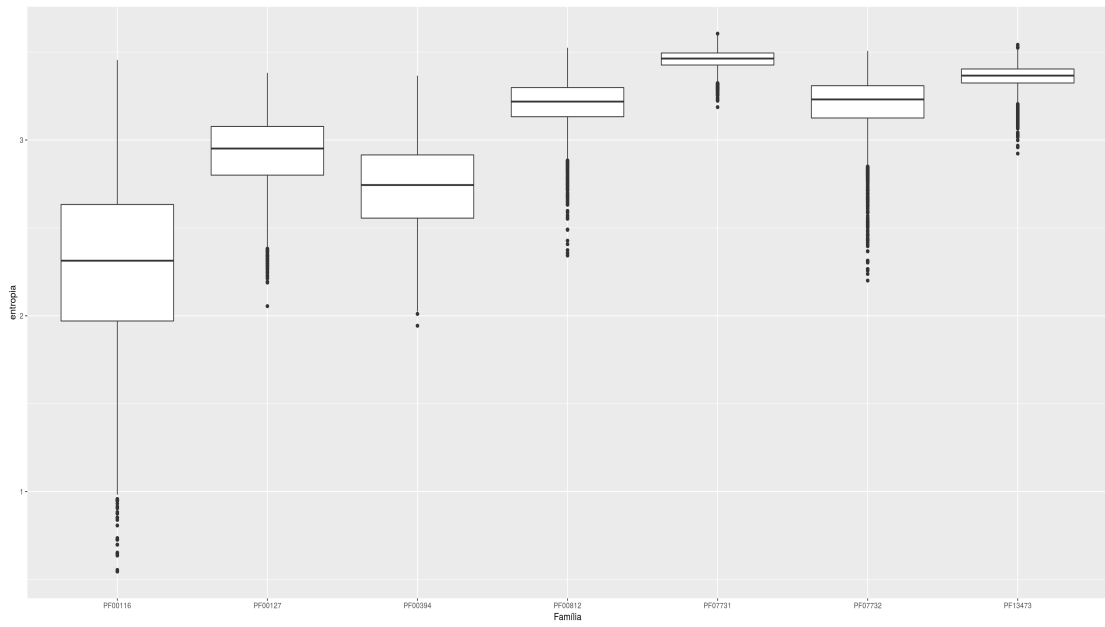
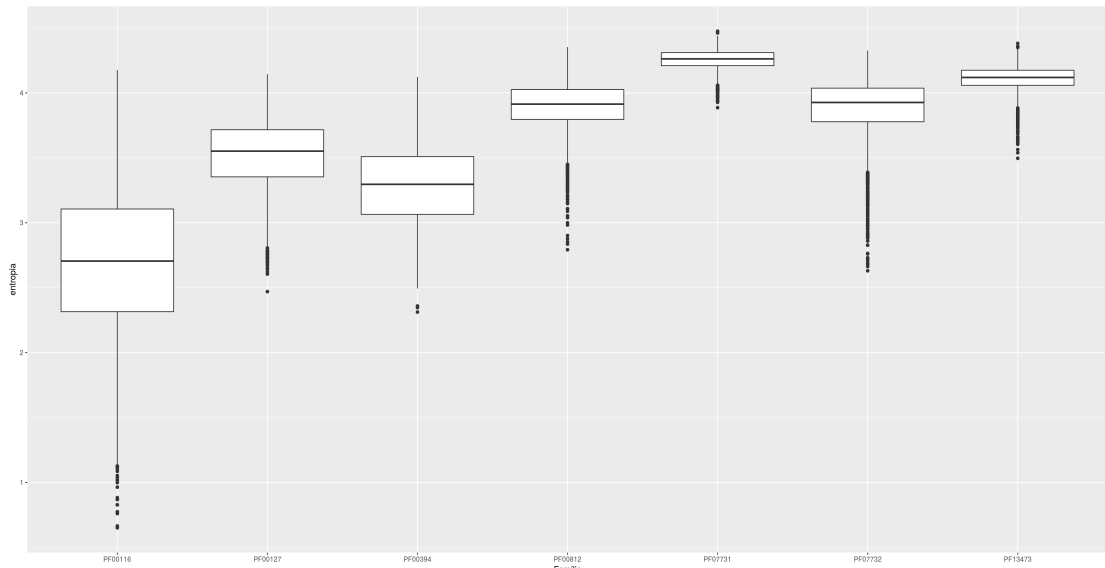


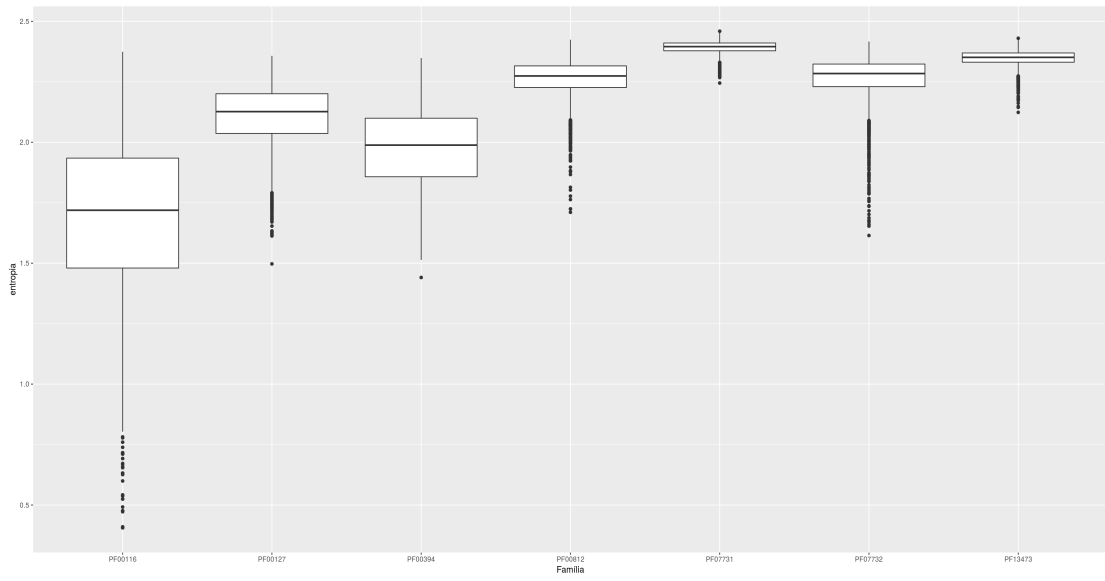
Figura 7.17 – Boxplots para as famílias do clã CL0025 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta



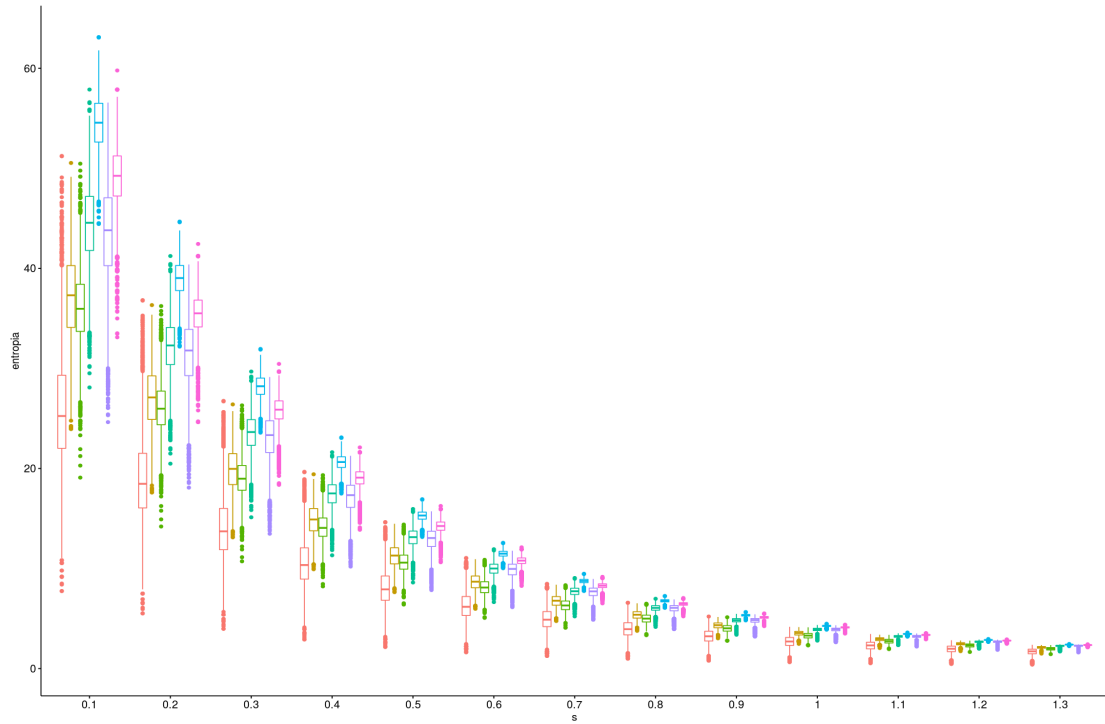








pf PF00116 PF00394 PF07731 PF13473
 PF00127 PF00812 PF07732



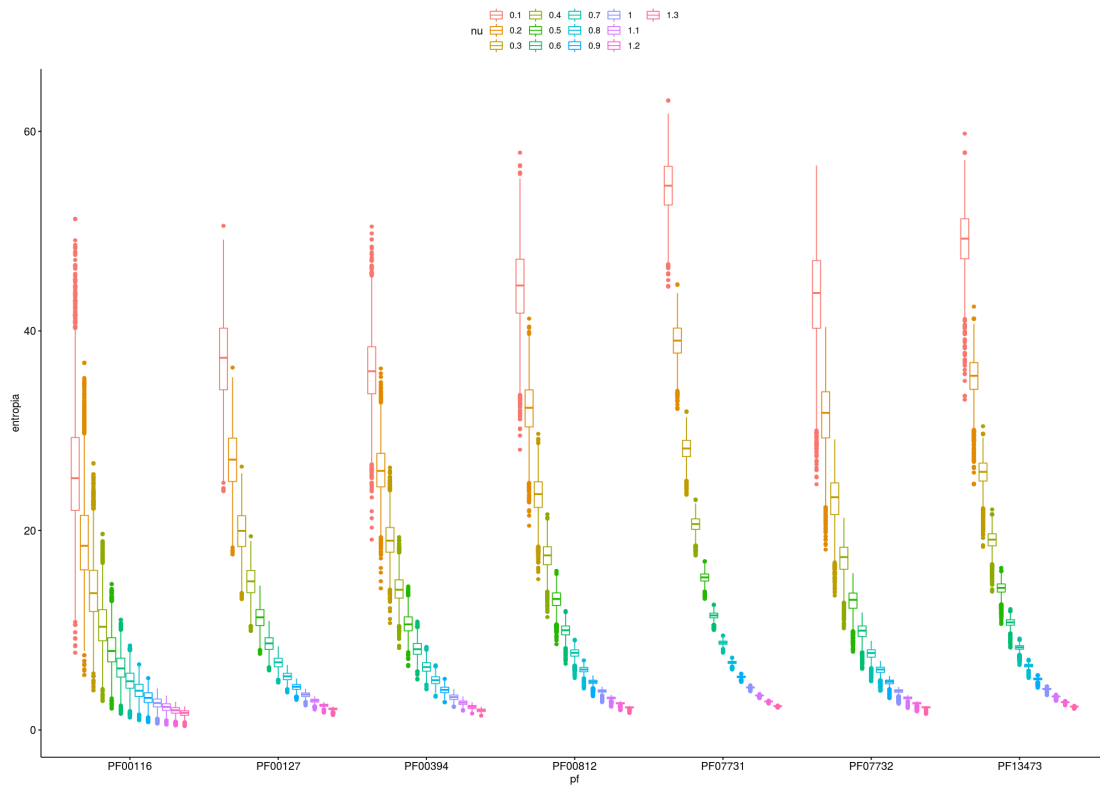
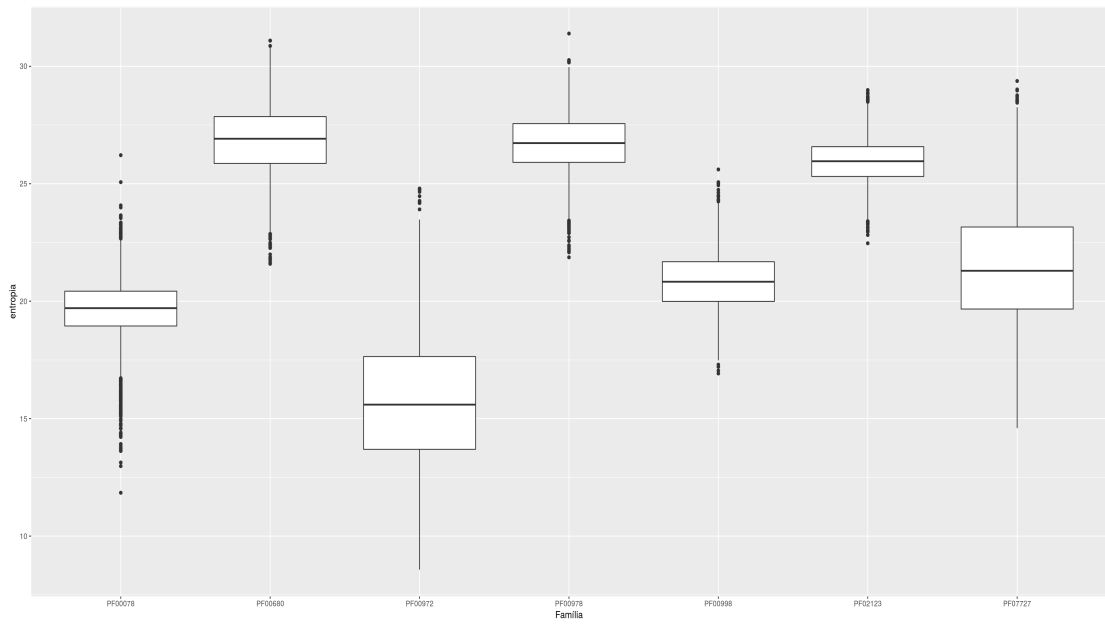
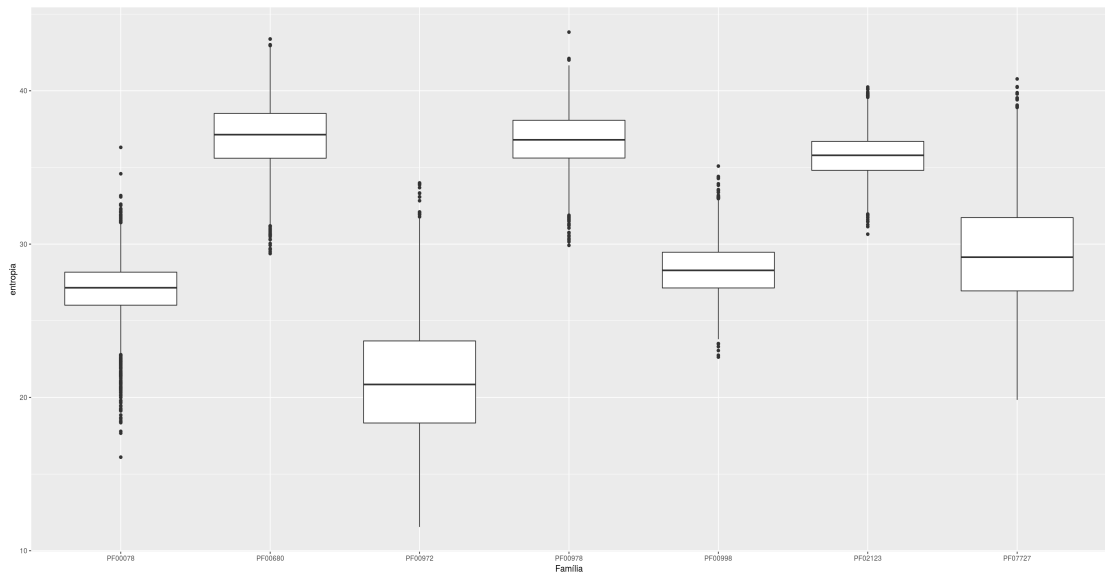
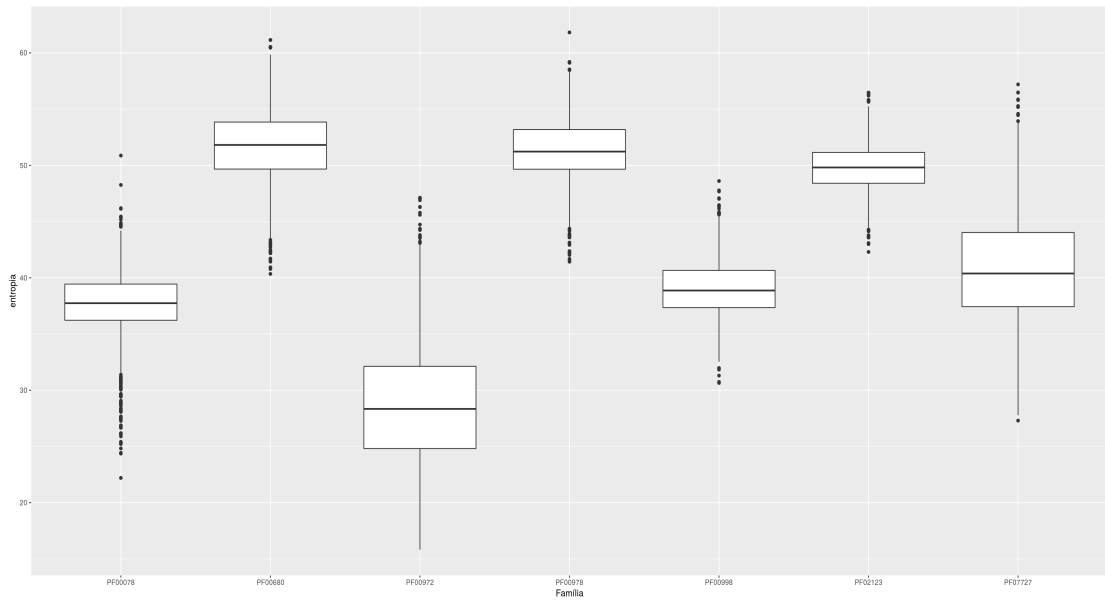
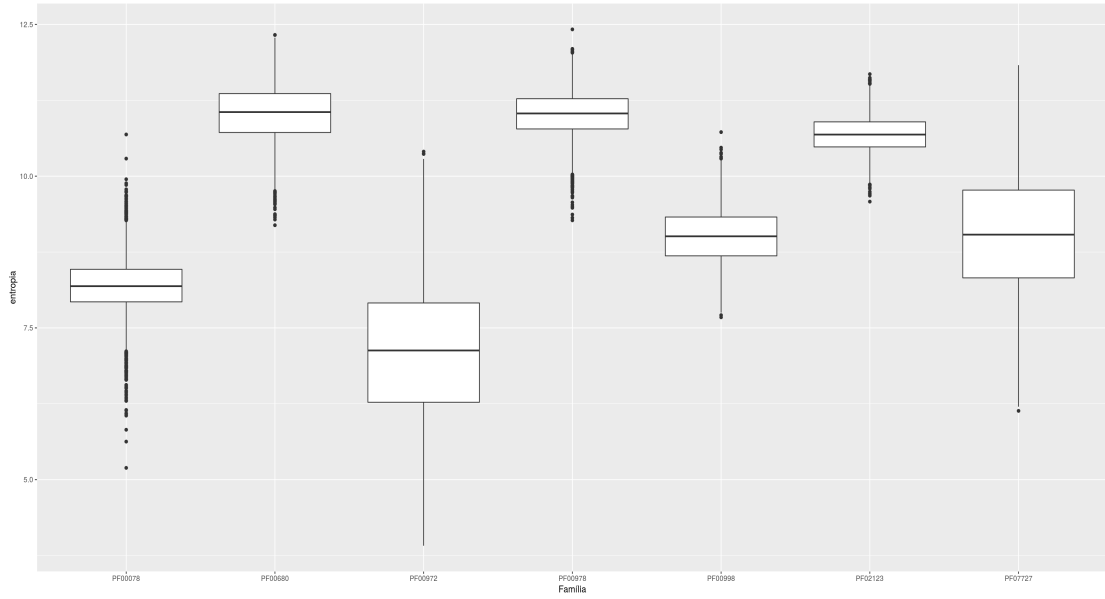
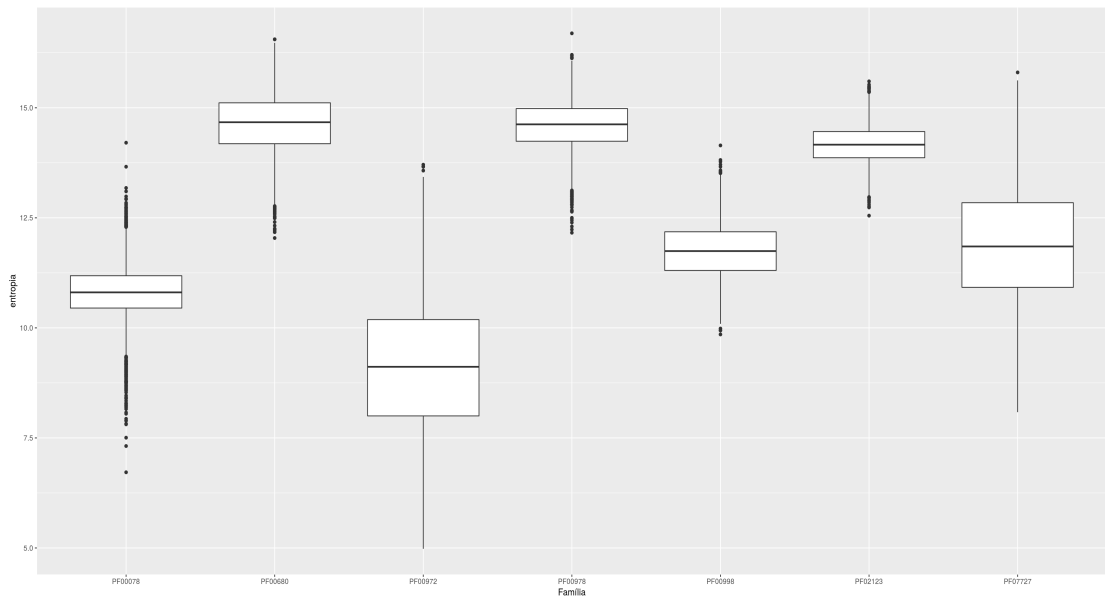
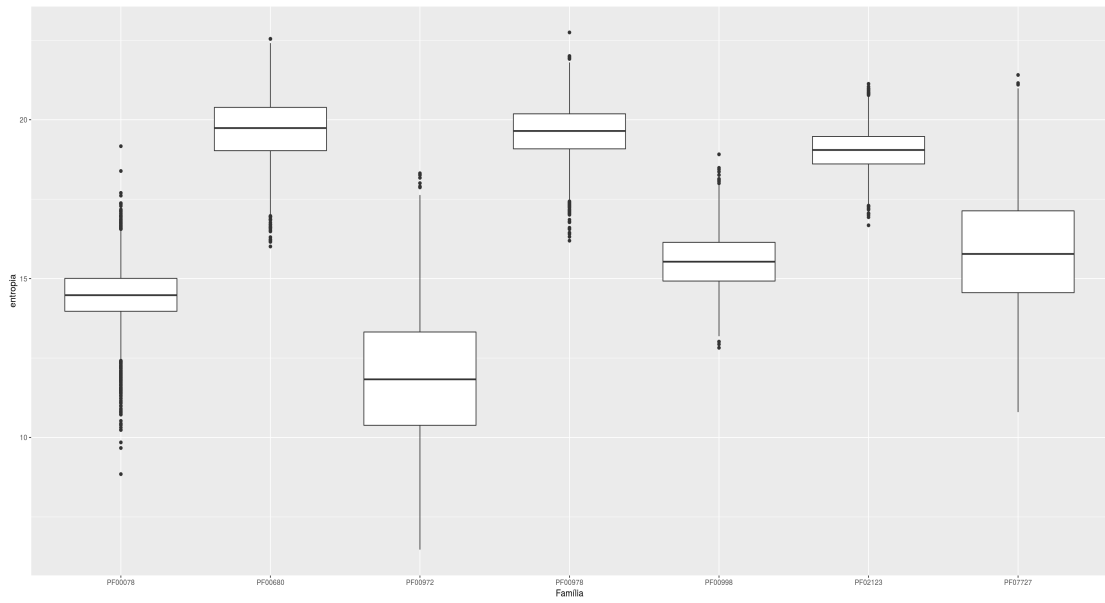
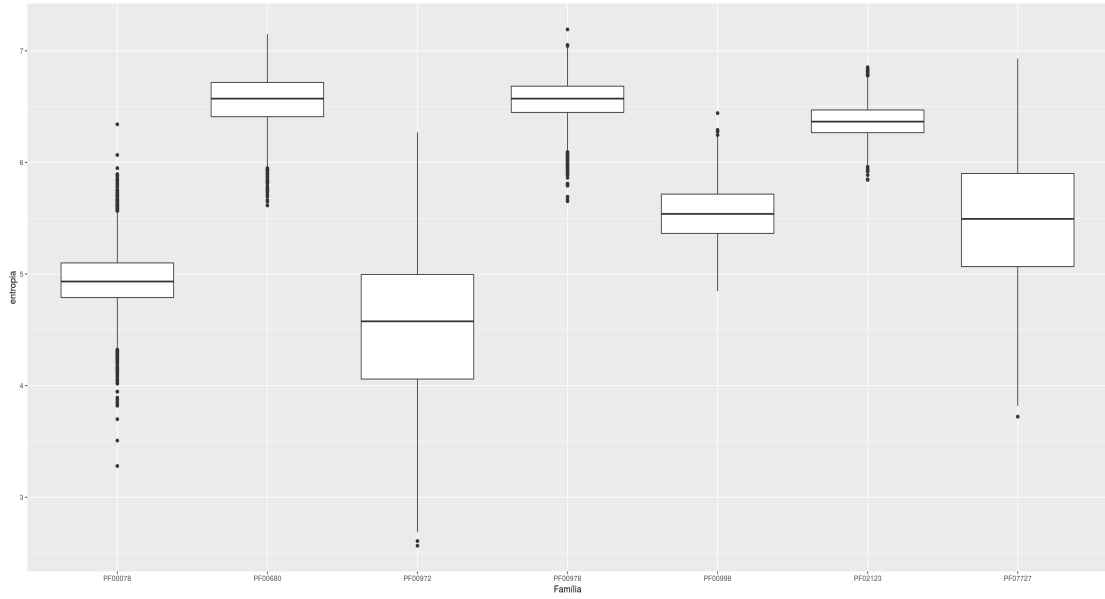
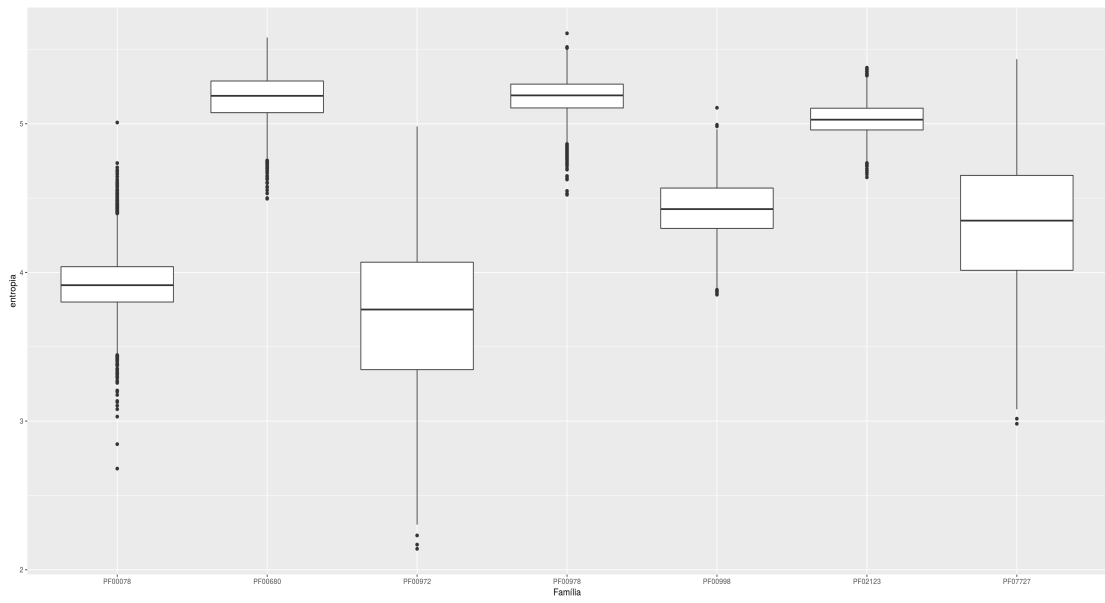
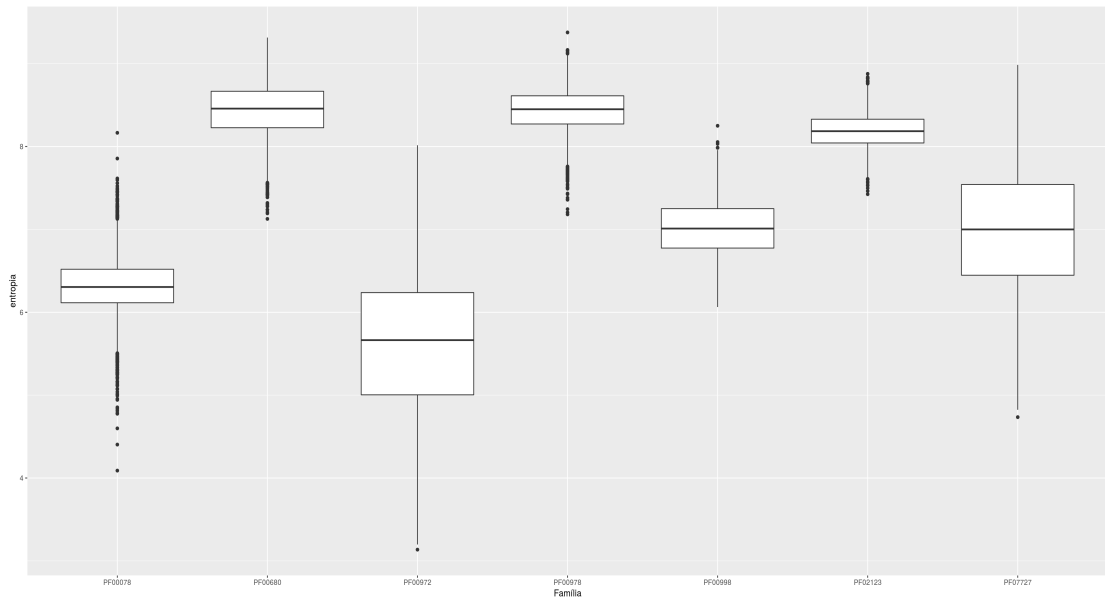
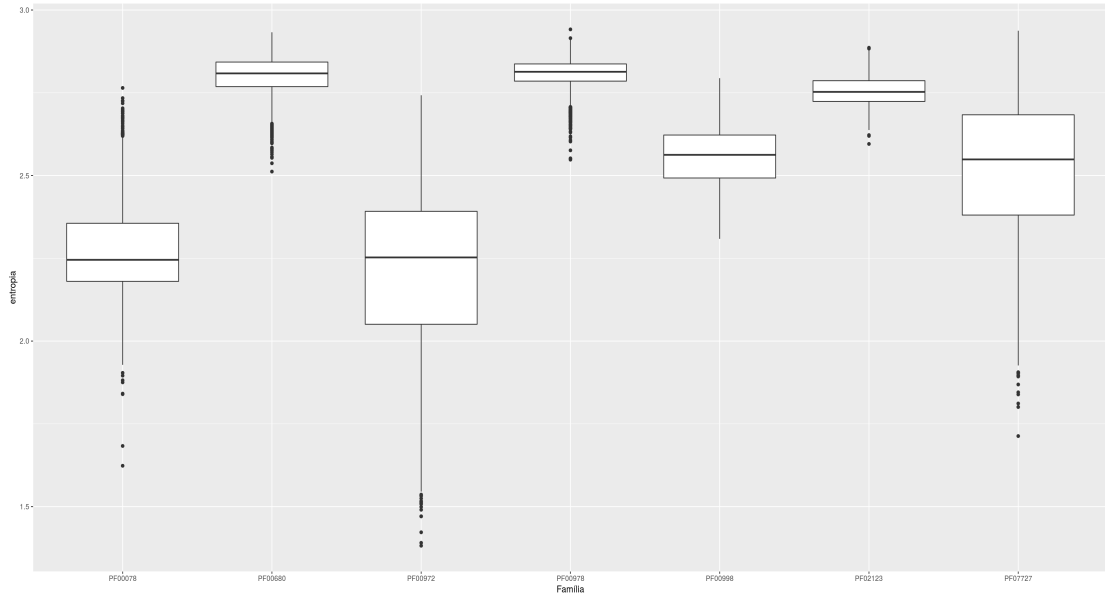
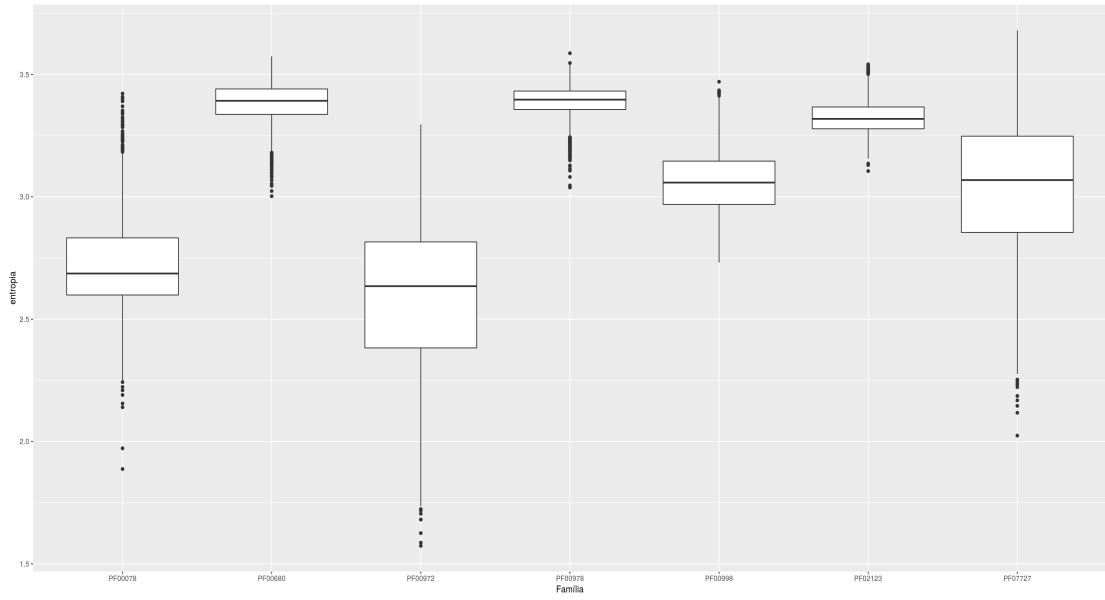
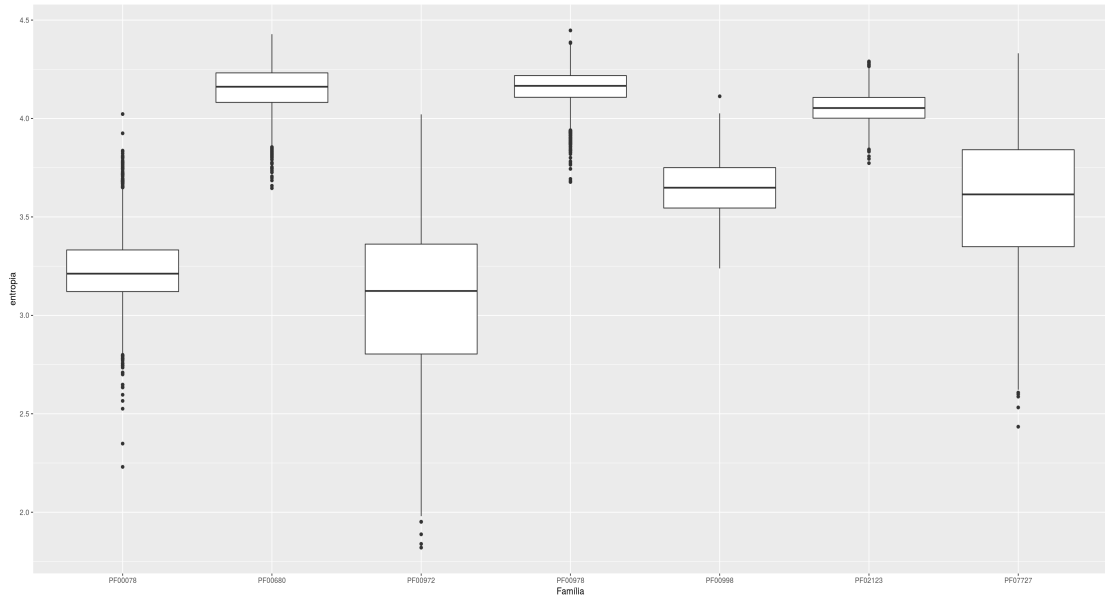


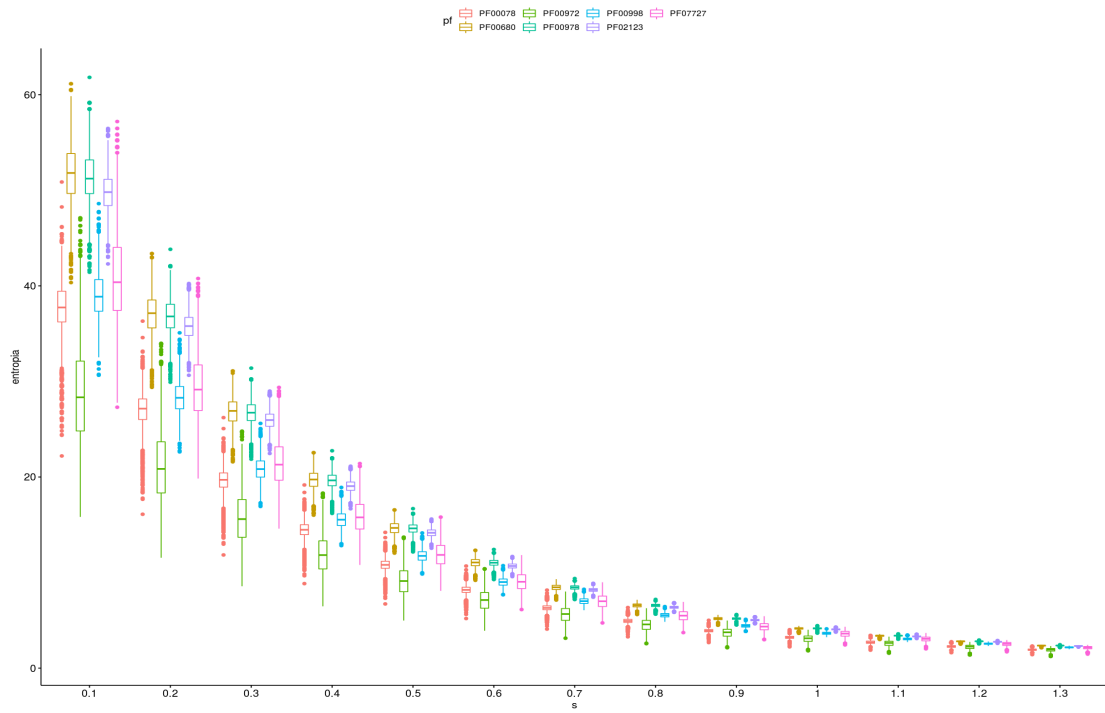
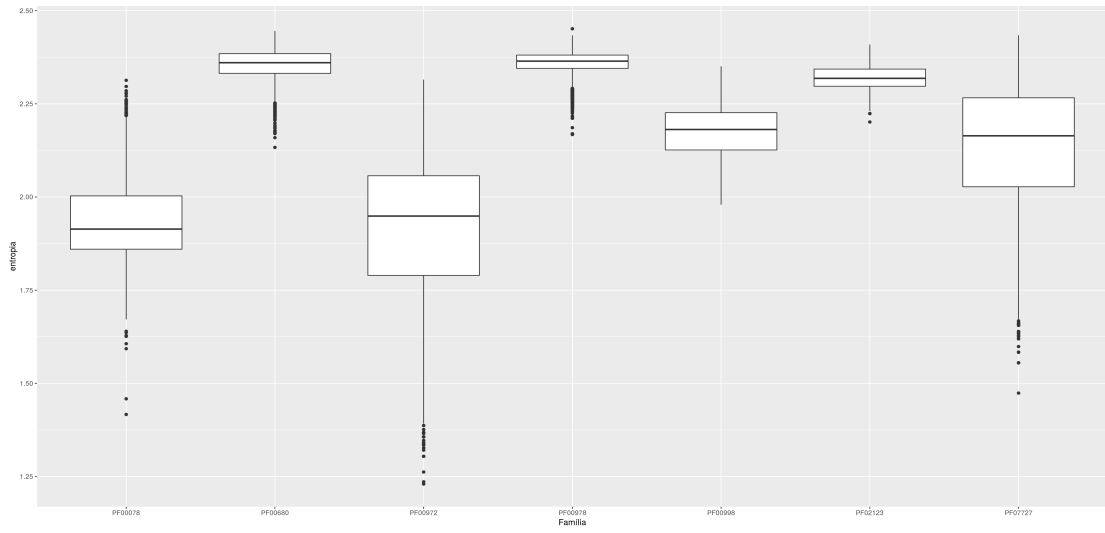
Figura 7.18 – Boxplots para as famílias do clã CL0026 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta











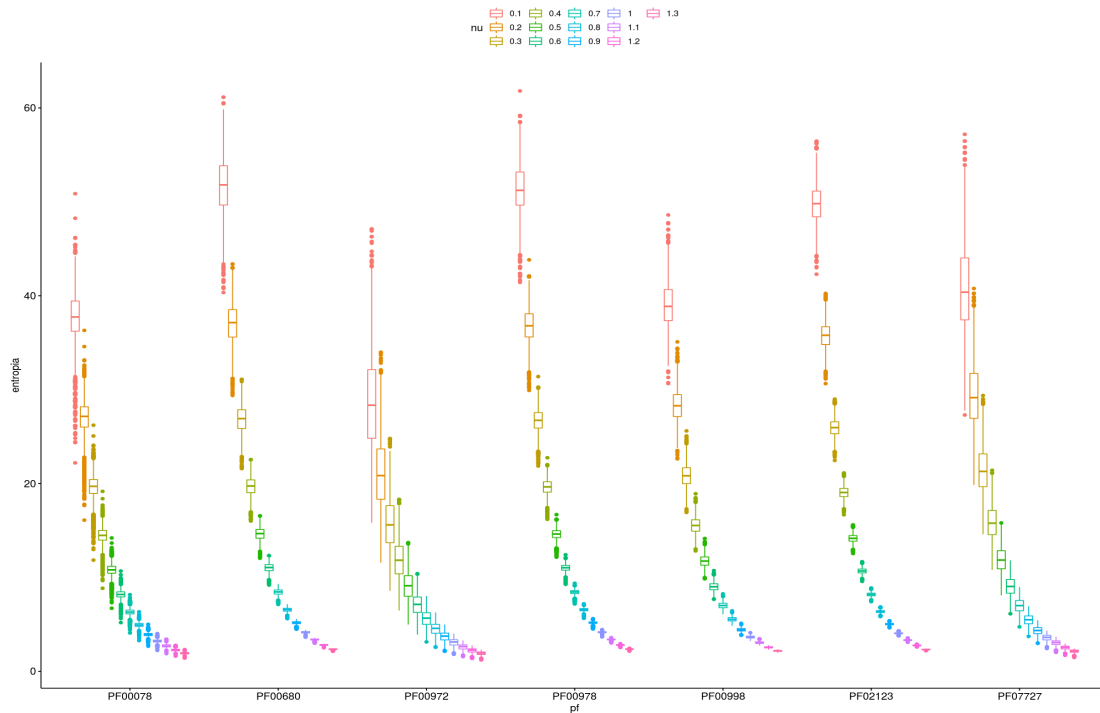


Figura 7.19 – Boxplots para as famílias do clã CL0027 ($s = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$ e 1.3) - probabilidade composta

Para as análises estatísticas ANOVA, foram realizados os seguintes testes:

7.4.1 Teste de Homogeneidade de Variância

7.4.1.1 Levene Teste

O teste de Levene (MONTGOMERY, 2012) é uma estatística inferencial usada para avaliar a igualdade de variâncias para uma variável calculada para dois ou mais grupos. Alguns procedimentos estatísticos comuns assumem que as variações das populações das quais diferentes amostras são extraídas são iguais.

Ele avalia essa suposição. Ele testa a hipótese nula de que as variâncias populacionais são iguais, chamadas de homogeneidade de variância ou homocedasticidade. Se o valor p resultante do teste de Levene for menor do que algum nível de significância, tipicamente 0.05, as diferenças obtidas nas variâncias

da amostra provavelmente não terão ocorrido com base na amostragem aleatória de uma população com variâncias iguais. Assim, a hipótese nula de variâncias iguais é rejeitada e conclui-se que existe uma diferença entre as variâncias na população.

Alguns dos procedimentos tipicamente assumindo homocedasticidade, para os quais se pode usar os testes de Levene, incluem análise de variância e testes-t. O teste de Levene é frequentemente usado antes de uma comparação de médias. Quando o teste de Levene mostra significância, deve-se passar para testes mais generalizados, livres de suposições de homocedasticidade, às vezes até testes não-paramétricos.

O teste de Levene também pode ser usado como um teste principal para responder a uma pergunta independente sobre se duas sub amostras em uma dada população têm variâncias iguais ou diferentes.

Para testar a homogeneidade da variância da entropia entre as famílias, foi utilizada a função `leveneTest` do pacote `car` do ambiente estatístico R (FARAWAY, 2015).

Exemplo:

Levene's Test for Homogeneity of Variance (center = mean)

```
Df F value Pr(>F)
group 13 1011.6 < 2.2e-16 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

O resultado da análise mostra um P-valor < 2.2e-16 significativo considerando nível de significância de 0.05, então rejeita-se a hipótese nula que os grupos tem variâncias iguais. Isto é, as entropias das famílias consideradas apresentam variabilidades diferentes.

Este resultado viola uma condição da ANOVA que supõe que os grupos tenham igual variância. Neste caso podemos explorar visualmente, utilizando

boxplots, para verificar o nível de discrepância da variabilidade entre os grupos. Uma alternativa para superar este casos é usar ANOVA não paramétrico ou teste Kruskal-Wallis.

7.4.1.2 Barlett Teste

O teste de Bartlett é útil para verificar as suposições de uma análise de variância, permitindo comparar a variância de duas ou mais amostras para determinar se são extraídas de populações com variações iguais. É adequado para dados normalmente distribuídos. O teste tem a hipótese nula de que as variâncias são iguais e a hipótese alternativa de que elas não são iguais.

Para testar a homogeneidade da variância da entropia entre as famílias, foi utilizada a função `bartlett.test` do pacote `stats` do ambiente estatístico R (FARAWAY, 2015).

Ambos os testes, Levene e Barlett, assumem que os dados tem distribuição normal.

Exemplo:

```
bartlett.test( datb.x$entropia, datb.x$pf )  
Bartlett test of homogeneity of variances  
data: datb.x$entropia and datb.x$pf  
Bartlett's K-squared = 22307, df = 13, p-value < 2.2e-16
```

O resultado teve o P valor < 2.2e-16, ou seja significativo considerando nível de significância de 0.05, então rejeita-se a hipótese nula e as variâncias das entropias entre os grupos não são iguais.

Este resultado coincide com o resultado anterior usando o teste de Levene. Ou seja, a distribuição das entropias das famílias tem variabilidade não homogênea. Isto, pode ser gerado pela presença de dados *outliers*.

7.4.2 ANOVA paramétrica - supõe normalidade

A análise de variância(ANOVA) pode determinar se as médias de dois ou mais grupos são diferentes. A ANOVA usa testes F para testar estatisticamente a igualdade de médias. Existem dois tipos de ANOVA que são comumente usados, o One-Way ANOVA e o Two-Way ANOVA (MONTGOMERY, 2012).

Em uma ANOVA unidirecional, há duas hipóteses possíveis. A hipótese nula (H0) é que não há diferença entre os grupos e igualdade entre médias. A hipótese alternativa (H1) é que existe uma diferença entre as médias e os grupos.

As suposições da One-Way ANOVA são:

- Normalidade - que cada amostra é retirada de uma população normalmente distribuída;
- Independência da amostra - que cada amostra foi retirada independentemente das outras amostras;
- Igualdade de Variação - que a variação de dados nos diferentes grupos deve ser a mesma.

Para testar a hipótese nula de igualdade de médias entre as entropias de um grupo de famílias foi utilizada a função `aov` do pacote `stats` do ambiente estatístico R (HORTON and KLEINMAN, 2011).

Exemplo:

```
anova.b <- aov( entropia ~ pf , datb.x )  
print( summary( anova.b ) )
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
pf	13	8056	619.7	10156	<2e-16 ***
Residuals	69286	4228	0.1		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Resultado o P valor=<2.2e-16 significativo, então rejeita hipótese nula e as médias não são iguais.

Ou seja, o resultado da ANOVA indica que os valores das médias das entropias do conjunto de famílias são diferentes.

7.4.2.1 Teste Tukey HSD

Quando a análise ANOVA indica que há diferenças entre as médias, não sabemos quais pares são diferentes. Nesse caso usamos o teste de Tukey HSD para descobrir os pares com médias diferentes. Ele pode ser usado em dados brutos ou em conjunto com uma ANOVA (análise post-hoc) para encontrar médias que são significativamente diferentes umas das outras. (MONTGOMERY, 2012)

Para testar dois a dois pares de médias de entropias de duas famílias se usou a função TukeyHSD do pacote stats do R (FARAWAY, 2015).

Exemplo:

Usando nível de confiança de 90% temos

```
TukeyHSD( anova.b , conf.level = 0.90 )
```

Tukey multiple comparisons of means

90% family-wise confidence level

```
Fit: aov(formula = entropia ~ pf, data = datb.x)
```

\$pf	diff	lwr	upr	p adj
------	------	-----	-----	-------

PF00139-PF00054 0.369173583 0.353683591 0.38466357 0.0000000

Compara par de médias para duas famílias, onde:

diff = diferença das médias

lwr = limite inferior da diferença

upr = limite superior da diferença

p adj = p valor ajustado

Se for menor que 0.05 rejeita-se a hipótese nula (diferença = 0) ou as médias são diferentes.

O resultado mostra que a média da entropia da família PF00139 e a média da entropia da família PF00054 são diferentes.

7.4.3 ANOVA não paramétrica - não supõe normalidade dos dados

7.4.3.1 Kruskal-Wallis rank sum test

Equivalente a ANOVA clássica. É um teste não paramétrico. Em situações em que a suposição de normalidade é injustificada, o experimentador pode querer usar um procedimento alternativo para a análise de variância do teste F que não depende desta condição.

Tal procedimento foi desenvolvido por Kruskal e Wallis(1952). O Kruskal-Wallis teste é usado para testar a hipótese nula de que os tratamentos são idênticos contra a alternativa hipótese de que alguns dos tratamentos geram observações maiores que outras (MONTGOMERY, 2012).

Para comparar a entropia entre um grupo de famílias, foi utilizado o teste `kruskal.test` do pacote `stats` do ambiente estatístico R(FARAWAY, 2015).

Exemplo:

```
kruskal.test( entropia ~ pf, data = datb.x )
```

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 49538, df = 13, p-value < 2.2e-16

É significativo se p-value for menor que 5% ou 1%.

O resultado P valor= $<2.2e-16$ é significativo, então rejeita-se a hipótese nula que as médias das entropias das famílias são iguais. Ou seja pelo menos um par de famílias têm médias de entropia diferentes.

Para descobrir quais pares são diferentes foi utilizado o Dunn's Kruskal-Wallis Multiple Comparisons.

7.4.3.2 Dunn's Kruskal-Wallis Multiple Comparisons

O teste de Dunn é a comparação múltipla emparelhada não paramétrica apropriada quando o teste de Kruskal-Wallis é rejeitado. Se o teste de Kruskal-Wallis for significativo, uma análise post-hoc pode ser realizada para determinar quais níveis da variável independente diferem um do outro. Provavelmente, o teste mais popular para isso é o teste de Dunn, que é executado com a função `dunnTest` no pacote FSA (MONTGOMERY, 2012).

Foi utilizada a função `dunnTest` do pacote FSA do ambiente estatístico R para comparar pares de médias de entropia de famílias (FARAWAY, 2015).

Exemplo

```
PT = dunnTest( entropia ~ pf, data = datb.x, method = "bh" )
```

```
print( PT )
```

Comparison	Z	P .unadj	P.adj
1 PF00054 - PF00139	-54.8170052	0.000000e+00	0.000000e+00

A primeira linha da saída do teste informa que o p valor ajustado (P.adj) é o valor quase zero ou seja significativo, então rejeita-se a hipótese nula que as médias são iguais. Ou seja as famílias PF00054 e PF00139 têm médias de entropia diferentes.

7.4.4 Resultados das aplicações dos Testes

Os testes acima descritos foram aplicados aos valores de entropia Harvard-Charvat das famílias de onze clãs (CL0004, CL0013, CL0014, CL0015, CL0016, CL0020, CL0021, CL0023, CL0025, CL0026 e CL0027), com os seguintes valores para o parâmetro s : 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 e 1.3.

Os resultados obtidos dos testes realizados para o clã CL0004 são apresentados abaixo:

```

=====
ANOVA da entropia entre familias controlando s = 0.1

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)
  Df F value Pr(>F)
group 13 474.6 < 2.2e-16 ***
 69286
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x$entropia and datb.x$pf
Bartlett's K-squared = 9897, df = 13, p-value < 2.2e-16
=====
ANOVA

      Df Sum Sq Mean Sq F value Pr(>F)
pf      13 4813168 370244 13828 <2e-16 ***
Residuals 69286 1855065 27

```

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste Tukey HSD

Tukey multiple comparisons of means
 90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	10.6144472	10.28997227	10.9389222	0.0000000
PF00337-PF00054	5.5961698	5.27169483	5.9206447	0.0000000
PF00354-PF00054	2.9230667	2.59859179	3.2475417	0.0000000
PF00457-PF00054	-7.6866668	-8.01114172	-7.3621918	0.0000000
PF00722-PF00054	14.2353460	13.91087107	14.5598210	0.0000000
PF00840-PF00054	-8.7147629	-9.03923788	-8.3902880	0.0000000
PF01670-PF00054	5.7085865	5.38411160	6.0330615	0.0000000
PF02210-PF00054	9.4912210	9.16674608	9.8156960	0.0000000
PF03388-PF00054	7.1312597	6.80678472	7.4557346	0.0000000
PF03935-PF00054	12.7453308	12.42085582	13.0698057	0.0000000
PF06439-PF00054	13.1694402	12.84496524	13.4939151	0.0000000
PF10287-PF00054	-7.5299419	-7.85441689	-7.2054670	0.0000000
PF13385-PF00054	18.0471898	17.72271484	18.3716647	0.0000000
PF00337-PF00139	-5.0182774	-5.34275239	-4.6938025	0.0000000
PF00354-PF00139	-7.6913805	-8.01585543	-7.3669055	0.0000000
PF00457-PF00139	-18.3011140	-18.62558894	-17.9766391	0.0000000
PF00722-PF00139	3.6208988	3.29642385	3.9453737	0.0000000
PF00840-PF00139	-19.3292102	-19.65368510	-19.0047352	0.0000000
PF01670-PF00139	-4.9058607	-5.23033562	-4.5813857	0.0000000
PF02210-PF00139	-1.1232262	-1.44770113	-0.7987512	0.0000000
PF03388-PF00139	-3.4831876	-3.80766250	-3.1587126	0.0000000
PF03935-PF00139	2.1308836	1.80640861	2.4553585	0.0000000
PF06439-PF00139	2.5549930	2.23051802	2.8794679	0.0000000
PF10287-PF00139	-18.1443892	-18.46886411	-17.8199142	0.0000000
PF13385-PF00139	7.4327426	7.10826762	7.7572175	0.0000000
PF00354-PF00337	-2.6731030	-2.99757798	-2.3486281	0.0000000
PF00457-PF00337	-13.2828366	-13.60731150	-12.9583616	0.0000000
PF00722-PF00337	8.6391762	8.31470130	8.9636512	0.0000000
PF00840-PF00337	-14.3109327	-14.63540766	-13.9864578	0.0000000
PF01670-PF00337	0.1124168	-0.21205818	0.4368917	0.9985024
PF02210-PF00337	3.8950513	3.57057631	4.2195262	0.0000000
PF03388-PF00337	1.5350899	1.21061495	1.8595648	0.0000000
PF03935-PF00337	7.1491610	6.82468605	7.4736359	0.0000000
PF06439-PF00337	7.5732704	7.24879547	7.8977454	0.0000000
PF10287-PF00337	-13.1261117	-13.45058666	-12.8016368	0.0000000
PF13385-PF00337	12.4510200	12.12654507	12.7754950	0.0000000

PF00457-PF00354 -10.6097335 -10.93420846 -10.2852586 0.0000000
PF00722-PF00354 11.3122793 10.98780433 11.6367542 0.0000000
PF00840-PF00354 -11.6378297 -11.96230462 -11.3133547 0.0000000
PF01670-PF00354 2.7855198 2.46104486 3.1099947 0.0000000
PF02210-PF00354 6.5681543 6.24367935 6.8926292 0.0000000
PF03388-PF00354 4.2081929 3.88371798 4.5326679 0.0000000
PF03935-PF00354 9.8222640 9.49778909 10.1467390 0.0000000
PF06439-PF00354 10.2463734 9.92189850 10.5708484 0.0000000
PF10287-PF00354 -10.4530087 -10.77748363 -10.1285337 0.0000000
PF13385-PF00354 15.1241230 14.79964810 15.4485980 0.0000000
PF00722-PF00457 21.9220128 21.59753785 22.2464877 0.0000000
PF00840-PF00457 -1.0280962 -1.35257110 -0.7036212 0.0000000
PF01670-PF00457 13.3952533 13.07077838 13.7197283 0.0000000
PF02210-PF00457 17.1778878 16.85341286 17.5023628 0.0000000
PF03388-PF00457 14.8179264 14.49345150 15.1424014 0.0000000
PF03935-PF00457 20.4319975 20.10752260 20.7564725 0.0000000
PF06439-PF00457 20.8561070 20.53163202 21.1805819 0.0000000
PF10287-PF00457 0.1567248 -0.16775011 0.4811998 0.9667675
PF13385-PF00457 25.7338566 25.40938162 26.0583315 0.0000000
PF00840-PF00722 -22.9501090 -23.27458390 -22.6256340 0.0000000
PF01670-PF00722 -8.5267595 -8.85123442 -8.2022845 0.0000000
PF02210-PF00722 -4.7441250 -5.06859993 -4.4196500 0.0000000
PF03388-PF00722 -7.1040863 -7.42856129 -6.7796114 0.0000000
PF03935-PF00722 -1.4900152 -1.81449019 -1.1655403 0.0000000
PF06439-PF00722 -1.0659058 -1.39038077 -0.7414309 0.0000000
PF10287-PF00722 -21.7652880 -22.08976290 -21.4408130 0.0000000
PF13385-PF00722 3.8118438 3.48736883 4.1363187 0.0000000
PF01670-PF00840 14.4233495 14.09887454 14.7478244 0.0000000
PF02210-PF00840 18.2059840 17.88150902 18.5304589 0.0000000
PF03388-PF00840 15.8460226 15.52154766 16.1704976 0.0000000
PF03935-PF00840 21.4600937 21.13561876 21.7845687 0.0000000
PF06439-PF00840 21.8842031 21.55972818 22.2086781 0.0000000
PF10287-PF00840 1.1848210 0.86034605 1.5092959 0.0000000
PF13385-PF00840 26.7619527 26.43747778 27.0864277 0.0000000
PF02210-PF01670 3.7826345 3.45815954 4.1071094 0.0000000
PF03388-PF01670 1.4226731 1.09819818 1.7471481 0.0000000
PF03935-PF01670 7.0367442 6.71226928 7.3612192 0.0000000
PF06439-PF01670 7.4608536 7.13637870 7.7853286 0.0000000
PF10287-PF01670 -13.2385285 -13.56300343 -12.9140535 0.0000000
PF13385-PF01670 12.3386032 12.01412830 12.6630782 0.0000000
PF03388-PF02210 -2.3599614 -2.68443631 -2.0354864 0.0000000
PF03935-PF02210 3.2541097 2.92963479 3.5785847 0.0000000
PF06439-PF02210 3.6782192 3.35374421 4.0026941 0.0000000
PF10287-PF02210 -17.0211630 -17.34563792 -16.6966880 0.0000000
PF13385-PF02210 8.5559688 8.23149381 8.8804437 0.0000000
PF03935-PF03388 5.6140711 5.28959616 5.9385460 0.0000000
PF06439-PF03388 6.0381805 5.71370557 6.3626555 0.0000000
PF10287-PF03388 -14.6612016 -14.98567656 -14.3367267 0.0000000
PF13385-PF03388 10.9159301 10.59145517 11.2404051 0.0000000

PF06439-PF03935	0.4241094	0.09963447	0.7485844	0.0036047
PF10287-PF03935	-20.2752727	-20.59974766	-19.9507978	0.0000000
PF13385-PF03935	5.3018590	4.97738407	5.6263340	0.0000000
PF10287-PF06439	-20.6993821	-21.02385708	-20.3749072	0.0000000
PF13385-PF06439	4.8777496	4.55327465	5.2022245	0.0000000
PF13385-PF10287	25.5771317	25.25265678	25.9016067	0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF01670-PF00337	0.1124168	-0.21205818	0.4368917	0.99850237
PF10287-PF00457	0.1567248	-0.16775011	0.4811998	0.96676750
PF06439-PF03935	0.4241094	0.09963447	0.7485844	0.00360473

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50987, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

	Comparison	Z	P.unadj	P.adj
1	PF00054 - PF00139	-59.223666	0.000000e+00	0.000000e+00
2	PF00054 - PF00337	-29.156889	6.832330e-187	1.002810e-186
3	PF00139 - PF00337	30.066777	1.317859e-198	1.965986e-198
4	PF00054 - PF00354	-12.936591	2.797706e-38	3.143101e-38
5	PF00139 - PF00354	46.287075	0.000000e+00	0.000000e+00
6	PF00337 - PF00354	16.220298	3.624317e-59	4.228370e-59
7	PF00054 - PF00457	28.931918	4.738817e-184	6.844959e-184
8	PF00139 - PF00457	88.155584	0.000000e+00	0.000000e+00
9	PF00337 - PF00457	58.088807	0.000000e+00	0.000000e+00
10	PF00354 - PF00457	41.868509	0.000000e+00	0.000000e+00
11	PF00054 - PF00722	-83.517502	0.000000e+00	0.000000e+00
12	PF00139 - PF00722	-24.293836	2.277596e-130	2.960875e-130
13	PF00337 - PF00722	-54.360613	0.000000e+00	0.000000e+00
14	PF00354 - PF00722	-70.580911	0.000000e+00	0.000000e+00
15	PF00457 - PF00722	-112.449420	0.000000e+00	0.000000e+00
16	PF00054 - PF00840	30.372431	1.270609e-202	1.927090e-202
17	PF00139 - PF00840	89.596097	0.000000e+00	0.000000e+00
18	PF00337 - PF00840	59.529320	0.000000e+00	0.000000e+00
19	PF00354 - PF00840	43.309022	0.000000e+00	0.000000e+00
20	PF00457 - PF00840	1.440513	1.497223e-01	1.497223e-01
21	PF00722 - PF00840	113.889933	0.000000e+00	0.000000e+00
22	PF00054 - PF01670	-26.668768	1.084377e-156	1.472810e-156
23	PF00139 - PF01670	32.554898	1.784949e-232	2.800523e-232
24	PF00337 - PF01670	2.488121	1.284199e-02	1.313057e-02

25 PF00354 - PF01670 -13.732177 6.514659e-43 7.410424e-43
 26 PF00457 - PF01670 -55.600686 0.000000e+00 0.000000e+00
 27 PF00722 - PF01670 56.848734 0.000000e+00 0.000000e+00
 28 PF00840 - PF01670 -57.041199 0.000000e+00 0.000000e+00
 29 PF00054 - PF02210 -55.423427 0.000000e+00 0.000000e+00
 30 PF00139 - PF02210 3.800239 1.445566e-04 1.512029e-04
 31 PF00337 - PF02210 -26.266538 4.626983e-152 6.102253e-152
 32 PF00354 - PF02210 -42.486836 0.000000e+00 0.000000e+00
 33 PF00457 - PF02210 -84.355345 0.000000e+00 0.000000e+00
 34 PF00722 - PF02210 28.094075 1.157286e-173 1.620200e-173
 35 PF00840 - PF02210 -85.795858 0.000000e+00 0.000000e+00
 36 PF01670 - PF02210 -28.754659 7.920716e-182 1.126227e-181
 37 PF00054 - PF03388 -35.309254 4.234521e-273 6.760376e-273
 38 PF00139 - PF03388 23.914411 2.168805e-126 2.779736e-126
 39 PF00337 - PF03388 -6.152365 7.633580e-10 8.269712e-10
 40 PF00354 - PF03388 -22.372663 7.266638e-111 9.058412e-111
 41 PF00457 - PF03388 -64.241172 0.000000e+00 0.000000e+00
 42 PF00722 - PF03388 48.208247 0.000000e+00 0.000000e+00
 43 PF00840 - PF03388 -65.681685 0.000000e+00 0.000000e+00
 44 PF01670 - PF03388 -8.640487 5.597402e-18 6.136911e-18
 45 PF02210 - PF03388 20.114172 5.545699e-90 6.728782e-90
 46 PF00054 - PF03935 -73.067964 0.000000e+00 0.000000e+00
 47 PF00139 - PF03935 -13.844299 1.377180e-43 1.586371e-43
 48 PF00337 - PF03935 -43.911075 0.000000e+00 0.000000e+00
 49 PF00354 - PF03935 -60.131373 0.000000e+00 0.000000e+00
 50 PF00457 - PF03935 -101.999883 0.000000e+00 0.000000e+00
 51 PF00722 - PF03935 10.449537 1.472417e-25 1.634023e-25
 52 PF00840 - PF03935 -103.440396 0.000000e+00 0.000000e+00
 53 PF01670 - PF03935 -46.399197 0.000000e+00 0.000000e+00
 54 PF02210 - PF03935 -17.644538 1.120875e-69 1.324671e-69
 55 PF03388 - PF03935 -37.758710 0.000000e+00 0.000000e+00
 56 PF00054 - PF06439 -78.192700 0.000000e+00 0.000000e+00
 57 PF00139 - PF06439 -18.969035 3.075024e-80 3.681936e-80
 58 PF00337 - PF06439 -49.035811 0.000000e+00 0.000000e+00
 59 PF00354 - PF06439 -65.256109 0.000000e+00 0.000000e+00
 60 PF00457 - PF06439 -107.124618 0.000000e+00 0.000000e+00
 61 PF00722 - PF06439 5.324802 1.010632e-07 1.081971e-07
 62 PF00840 - PF06439 -108.565131 0.000000e+00 0.000000e+00
 63 PF01670 - PF06439 -51.523933 0.000000e+00 0.000000e+00
 64 PF02210 - PF06439 -22.769273 9.246595e-115 1.168667e-114
 65 PF03388 - PF06439 -42.883446 0.000000e+00 0.000000e+00
 66 PF03935 - PF06439 -5.124736 2.979554e-07 3.152784e-07
 67 PF00054 - PF10287 26.880131 3.749973e-159 5.170418e-159
 68 PF00139 - PF10287 86.103797 0.000000e+00 0.000000e+00
 69 PF00337 - PF10287 56.037020 0.000000e+00 0.000000e+00
 70 PF00354 - PF10287 39.816722 0.000000e+00 0.000000e+00
 71 PF00457 - PF10287 -2.051787 4.019037e-02 4.063693e-02
 72 PF00722 - PF10287 110.397633 0.000000e+00 0.000000e+00
 73 PF00840 - PF10287 -3.492300 4.788800e-04 4.952055e-04

74	PF01670 - PF10287	53.548899	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.303558	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	62.189385	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	99.948096	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	105.072831	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-104.561282	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-45.337617	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-75.404393	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-91.624691	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-133.493200	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-21.043780	2.607398e-98	3.206395e-98
85	PF00840 - PF13385	-134.933713	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-77.892515	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-49.137855	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-69.252028	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-31.493318	1.072362e-217	1.653981e-217
90	PF06439 - PF13385	-26.368582	3.142639e-153	4.205590e-153
91	PF10287 - PF13385	-131.441413	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
20	PF00457 - PF00840	1.440513	0.14972231	0.14972231
24	PF00337 - PF01670	2.488121	0.01284199	0.01313057
71	PF00457 - PF10287	-2.051787	0.04019037	0.04063693

ANOVA da entropia entre familias controlando **s = 0.2**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	506.1	< 2.2e-16 ***
		69286	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 10516, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 2270533 174656 13596 <2e-16 ***

Residuals 69286 890031 13

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	7.2039394	6.97918713	7.4286918	0.0000000
PF00337-PF00054	3.8072705	3.58251821	4.0320228	0.0000000
PF00354-PF00054	1.9466840	1.72193167	2.1714363	0.0000000
PF00457-PF00054	-5.5244571	-5.74920942	-5.2997048	0.0000000
PF00722-PF00054	9.6491447	9.42439237	9.8738970	0.0000000
PF00840-PF00054	-6.0943174	-6.31906972	-5.8695651	0.0000000
PF01670-PF00054	3.9037041	3.67895181	4.1284564	0.0000000
PF02210-PF00054	6.4591172	6.23436492	6.6838696	0.0000000
PF03388-PF00054	4.7798440	4.55509173	5.0045964	0.0000000
PF03935-PF00054	8.6044535	8.37970120	8.8292058	0.0000000
PF06439-PF00054	8.9408153	8.71606298	9.1655676	0.0000000
PF10287-PF00054	-5.2629333	-5.48768564	-5.0381810	0.0000000
PF13385-PF00054	12.2235656	11.99881333	12.4483180	0.0000000
PF00337-PF00139	-3.3966689	-3.62142124	-3.1719166	0.0000000
PF00354-PF00139	-5.2572555	-5.48200777	-5.0325031	0.0000000
PF00457-PF00139	-12.7283965	-12.95314886	-12.5036442	0.0000000
PF00722-PF00139	2.4452052	2.22045292	2.6699576	0.0000000
PF00840-PF00139	-13.2982568	-13.52300916	-13.0735045	0.0000000
PF01670-PF00139	-3.3002353	-3.52498763	-3.0754830	0.0000000
PF02210-PF00139	-0.7448222	-0.96957452	-0.5200699	0.0000000
PF03388-PF00139	-2.4240954	-2.64884771	-2.1993431	0.0000000
PF03935-PF00139	1.4005141	1.17576176	1.6252664	0.0000000

PF06439-PF00139	1.7368759	1.51212354	1.9616282	0.0000000
PF10287-PF00139	-12.4668728	-12.69162509	-12.2421205	0.0000000
PF13385-PF00139	5.0196262	4.79487388	5.2443785	0.0000000
PF00354-PF00337	-1.8605865	-2.08533885	-1.6358342	0.0000000
PF00457-PF00337	-9.3317276	-9.55647994	-9.1069753	0.0000000
PF00722-PF00337	5.8418742	5.61712184	6.0666265	0.0000000
PF00840-PF00337	-9.9015879	-10.12634025	-9.6768356	0.0000000
PF01670-PF00337	0.0964336	-0.12831872	0.3211859	0.9880245
PF02210-PF00337	2.6518467	2.42709440	2.8765990	0.0000000
PF03388-PF00337	0.9725735	0.74782121	1.1973258	0.0000000
PF03935-PF00337	4.7971830	4.57243068	5.0219353	0.0000000
PF06439-PF00337	5.1335448	4.90879246	5.3582971	0.0000000
PF10287-PF00337	-9.0702038	-9.29495617	-8.8454515	0.0000000
PF13385-PF00337	8.4162951	8.19154280	8.6410474	0.0000000
PF00457-PF00354	-7.4711411	-7.69589341	-7.2463888	0.0000000
PF00722-PF00354	7.7024607	7.47770837	7.9272130	0.0000000
PF00840-PF00354	-8.0410014	-8.26575371	-7.8162491	0.0000000
PF01670-PF00354	1.9570201	1.73226782	2.1817725	0.0000000
PF02210-PF00354	4.5124332	4.28768093	4.7371856	0.0000000
PF03388-PF00354	2.8331601	2.60840774	3.0579124	0.0000000
PF03935-PF00354	6.6577695	6.43301721	6.8825218	0.0000000
PF06439-PF00354	6.9941313	6.76937899	7.2188836	0.0000000
PF10287-PF00354	-7.2096173	-7.43436964	-6.9848650	0.0000000
PF13385-PF00354	10.2768817	10.05212934	10.5016340	0.0000000
PF00722-PF00457	15.1736018	14.94884946	15.3983541	0.0000000
PF00840-PF00457	-0.5698603	-0.79461262	-0.3451080	0.0000000
PF01670-PF00457	9.4281612	9.20340891	9.6529135	0.0000000
PF02210-PF00457	11.9835743	11.75882202	12.2083267	0.0000000
PF03388-PF00457	10.3043011	10.07954883	10.5290535	0.0000000
PF03935-PF00457	14.1289106	13.90415830	14.3536629	0.0000000
PF06439-PF00457	14.4652724	14.24052008	14.6900247	0.0000000
PF10287-PF00457	0.2615238	0.03677145	0.4862761	0.0198887
PF13385-PF00457	17.7480227	17.52327042	17.9727751	0.0000000
PF00840-PF00722	-15.7434621	-15.96821440	-15.5187098	0.0000000
PF01670-PF00722	-5.7454406	-5.97019287	-5.5206882	0.0000000
PF02210-PF00722	-3.1900274	-3.41477976	-2.9652751	0.0000000
PF03388-PF00722	-4.8693006	-5.09405295	-4.6445483	0.0000000
PF03935-PF00722	-1.0446912	-1.26944348	-0.8199388	0.0000000
PF06439-PF00722	-0.7083294	-0.93308170	-0.4835771	0.0000000
PF10287-PF00722	-14.9120780	-15.13683033	-14.6873257	0.0000000
PF13385-PF00722	2.5744210	2.34966864	2.7991733	0.0000000
PF01670-PF00840	9.9980215	9.77326921	10.2227738	0.0000000
PF02210-PF00840	12.5534346	12.32868232	12.7781870	0.0000000
PF03388-PF00840	10.8741615	10.64940913	11.0989138	0.0000000
PF03935-PF00840	14.6987709	14.47401860	14.9235232	0.0000000
PF06439-PF00840	15.0351327	14.81038039	15.2598850	0.0000000
PF10287-PF00840	0.8313841	0.60663176	1.0561364	0.0000000
PF13385-PF00840	18.3178830	18.09313073	18.5426354	0.0000000
PF02210-PF01670	2.5554131	2.33066079	2.7801654	0.0000000
PF03388-PF01670	0.8761399	0.65138760	1.1008922	0.0000000
PF03935-PF01670	4.7007494	4.47599707	4.9255017	0.0000000
PF06439-PF01670	5.0371112	4.81235886	5.2618635	0.0000000
PF10287-PF01670	-9.1666375	-9.39138977	-8.9418851	0.0000000
PF13385-PF01670	8.3198615	8.09510920	8.5446138	0.0000000

PF03388-PF02210 -1.6792732 -1.90402551 -1.4545209 0.0000000
 PF03935-PF02210 2.1453363 1.92058396 2.3700886 0.0000000
 PF06439-PF02210 2.4816981 2.25694574 2.7064504 0.0000000
 PF10287-PF02210 -11.7220506 -11.94680288 -11.4972982 0.0000000
 PF13385-PF02210 5.7644484 5.53969609 5.9892007 0.0000000
 PF03935-PF03388 3.8246095 3.59985715 4.0493618 0.0000000
 PF06439-PF03388 4.1609713 3.93621894 4.3857236 0.0000000
 PF10287-PF03388 -10.0427774 -10.26752969 -9.8180251 0.0000000
 PF13385-PF03388 7.4437216 7.21896928 7.6684739 0.0000000
 PF06439-PF03935 0.3363618 0.11160947 0.5611141 0.0002608
 PF10287-PF03935 -13.8673868 -14.09213916 -13.6426345 0.0000000
 PF13385-PF03935 3.6191121 3.39435981 3.8438644 0.0000000
 PF10287-PF06439 -14.2037486 -14.42850095 -13.9789963 0.0000000
 PF13385-PF06439 3.2827503 3.05799802 3.5075027 0.0000000
 PF13385-PF10287 17.4864990 17.26174665 17.7112513 0.0000000
 Pares iguais

diff lwr upr p adj

PF01670-PF00337 0.0964336 -0.12831872 0.3211859 0.98802450

PF10287-PF00457 0.2615238 0.03677145 0.4862761 0.01988871

=====

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50926, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-59.0021356	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-29.0255458	3.132832e-185	4.525202e-185
3 PF00139 - PF00337	29.9765898	1.981745e-197	2.956374e-197
4 PF00054 - PF00354	-12.7121898	5.059740e-37	5.684400e-37
5 PF00139 - PF00354	46.2899457	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	16.3133559	7.931296e-60	9.253179e-60
7 PF00054 - PF00457	29.4671299	7.596329e-191	1.114945e-190
8 PF00139 - PF00457	88.4692655	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	58.4926757	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	42.1793198	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-83.3156045	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-24.3134690	1.412229e-130	1.835897e-130
13 PF00337 - PF00722	-54.2900588	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-70.6034147	0.000000e+00	0.000000e+00
15 PF00457 - PF00722	-112.7827344	0.000000e+00	0.000000e+00
16 PF00054 - PF00840	30.3599553	1.856602e-202	2.815847e-202
17 PF00139 - PF00840	89.3620909	0.000000e+00	0.000000e+00

18	PF00337 - PF00840	59.3855011	0.000000e+00	0.000000e+00
19	PF00354 - PF00840	43.0721451	0.000000e+00	0.000000e+00
20	PF00457 - PF00840	0.8928254	3.719507e-01	3.719507e-01
21	PF00722 - PF00840	113.6755598	0.000000e+00	0.000000e+00
22	PF00054 - PF01670	-26.5732567	1.383472e-155	1.879044e-155
23	PF00139 - PF01670	32.4288789	1.075396e-230	1.687260e-230
24	PF00337 - PF01670	2.4522891	1.419506e-02	1.435278e-02
25	PF00354 - PF01670	-13.8610668	1.090415e-43	1.256047e-43
26	PF00457 - PF01670	-56.0403866	0.000000e+00	0.000000e+00
27	PF00722 - PF01670	56.7423479	0.000000e+00	0.000000e+00
28	PF00840 - PF01670	-56.9332120	0.000000e+00	0.000000e+00
29	PF00054 - PF02210	-55.3136047	0.000000e+00	0.000000e+00
30	PF00139 - PF02210	3.6885308	2.255527e-04	2.359229e-04
31	PF00337 - PF02210	-26.2880590	2.626269e-152	3.463630e-152
32	PF00354 - PF02210	-42.6014149	0.000000e+00	0.000000e+00
33	PF00457 - PF02210	-84.7807346	0.000000e+00	0.000000e+00
34	PF00722 - PF02210	28.0019998	1.536184e-172	2.150657e-172
35	PF00840 - PF02210	-85.6735600	0.000000e+00	0.000000e+00
36	PF01670 - PF02210	-28.7403481	1.195779e-181	1.700249e-181
37	PF00054 - PF03388	-34.8284688	9.020199e-266	1.440067e-265
38	PF00139 - PF03388	24.1736668	4.210727e-129	5.396847e-129
39	PF00337 - PF03388	-5.8029230	6.516871e-09	7.059944e-09
40	PF00354 - PF03388	-22.1162789	2.203735e-108	2.747121e-108
41	PF00457 - PF03388	-64.2955987	0.000000e+00	0.000000e+00
42	PF00722 - PF03388	48.4871357	0.000000e+00	0.000000e+00
43	PF00840 - PF03388	-65.1884241	0.000000e+00	0.000000e+00
44	PF01670 - PF03388	-8.2552121	1.516327e-16	1.662479e-16
45	PF02210 - PF03388	20.4851359	2.921673e-93	3.544964e-93
46	PF00054 - PF03935	-72.6665749	0.000000e+00	0.000000e+00
47	PF00139 - PF03935	-13.6644393	1.655741e-42	1.883405e-42
48	PF00337 - PF03935	-43.6410291	0.000000e+00	0.000000e+00
49	PF00354 - PF03935	-59.9543851	0.000000e+00	0.000000e+00
50	PF00457 - PF03935	-102.1337048	0.000000e+00	0.000000e+00
51	PF00722 - PF03935	10.6490296	1.761921e-26	1.955302e-26
52	PF00840 - PF03935	-103.0265302	0.000000e+00	0.000000e+00
53	PF01670 - PF03935	-46.0933182	0.000000e+00	0.000000e+00
54	PF02210 - PF03935	-17.3529702	1.873045e-67	2.213599e-67
55	PF03388 - PF03935	-37.8381061	0.000000e+00	0.000000e+00
56	PF00054 - PF06439	-78.0218658	0.000000e+00	0.000000e+00
57	PF00139 - PF06439	-19.0197303	1.170843e-80	1.401930e-80
58	PF00337 - PF06439	-48.9963201	0.000000e+00	0.000000e+00
59	PF00354 - PF06439	-65.3096760	0.000000e+00	0.000000e+00
60	PF00457 - PF06439	-107.4889957	0.000000e+00	0.000000e+00
61	PF00722 - PF06439	5.2937387	1.198406e-07	1.268081e-07
62	PF00840 - PF06439	-108.3818211	0.000000e+00	0.000000e+00
63	PF01670 - PF06439	-51.4486091	0.000000e+00	0.000000e+00
64	PF02210 - PF06439	-22.7082611	3.712441e-114	4.692113e-114
65	PF03388 - PF06439	-43.1933970	0.000000e+00	0.000000e+00
66	PF03935 - PF06439	-5.3552909	8.541887e-08	9.144844e-08
67	PF00054 - PF10287	26.9158323	1.433522e-159	1.976523e-159
68	PF00139 - PF10287	85.9179678	0.000000e+00	0.000000e+00
69	PF00337 - PF10287	55.9413780	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	39.6280221	0.000000e+00	0.000000e+00
71	PF00457 - PF10287	-2.5512976	1.073226e-02	1.097344e-02

72 PF00722 - PF10287 110.2314368 0.000000e+00 0.000000e+00
73 PF00840 - PF10287 -3.4441230 5.729150e-04 5.924462e-04
74 PF01670 - PF10287 53.4890890 0.000000e+00 0.000000e+00
75 PF02210 - PF10287 82.2294370 0.000000e+00 0.000000e+00
76 PF03388 - PF10287 61.7443011 0.000000e+00 0.000000e+00
77 PF03935 - PF10287 99.5824072 0.000000e+00 0.000000e+00
78 PF06439 - PF10287 104.9376981 0.000000e+00 0.000000e+00
79 PF00054 - PF13385 -104.3654195 0.000000e+00 0.000000e+00
80 PF00139 - PF13385 -45.3632839 0.000000e+00 0.000000e+00
81 PF00337 - PF13385 -75.3398737 0.000000e+00 0.000000e+00
82 PF00354 - PF13385 -91.6532296 0.000000e+00 0.000000e+00
83 PF00457 - PF13385 -133.8325494 0.000000e+00 0.000000e+00
84 PF00722 - PF13385 -21.0498150 2.295743e-98 2.823144e-98
85 PF00840 - PF13385 -134.7253748 0.000000e+00 0.000000e+00
86 PF01670 - PF13385 -77.7921628 0.000000e+00 0.000000e+00
87 PF02210 - PF13385 -49.0518148 0.000000e+00 0.000000e+00
88 PF03388 - PF13385 -69.5369507 0.000000e+00 0.000000e+00
89 PF03935 - PF13385 -31.6988446 1.611688e-220 2.485824e-220
90 PF06439 - PF13385 -26.3435537 6.083976e-153 8.141791e-153
91 PF10287 - PF13385 -131.2812518 0.000000e+00 0.000000e+00

Pares iguais

Comparison	Z	P.unadj	P.adj
20 PF00457 - PF00840	0.8928254	0.37195069	0.37195069
24 PF00337 - PF01670	2.4522891	0.01419506	0.01435278
71 PF00457 - PF10287	-2.5512976	0.01073226	0.01097344

=====

ANOVA da entropia entre familias controlando **s = 0.3**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	543.26	< 2.2e-16 ***
	69286		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 11258, df = 13, p-value < 2.2e-16

=====

ANOVA

Df	Sum Sq	Mean Sq	F value	Pr(>F)
pf	13 1080630	83125	13325	<2e-16 ***
Residuals	69286	432238	6	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	4.90466462	4.74803895	5.0612903	0.0000000
PF00337-PF00054	2.59874840	2.44212273	2.7553741	0.0000000
PF00354-PF00054	1.29957045	1.14294478	1.4561961	0.0000000
PF00457-PF00054	-3.99534507	-4.15197074	-3.8387194	0.0000000
PF00722-PF00054	6.56109096	6.40446529	6.7177166	0.0000000
PF00840-PF00054	-4.28271712	-4.43934279	-4.1260914	0.0000000
PF01670-PF00054	2.67996469	2.52333902	2.8365904	0.0000000
PF02210-PF00054	4.41072784	4.25410217	4.5673535	0.0000000
PF03388-PF00054	3.21186744	3.05524177	3.3684931	0.0000000
PF03935-PF00054	5.82558628	5.66896061	5.9822120	0.0000000
PF06439-PF00054	6.08936288	5.93273721	6.2459885	0.0000000
PF10287-PF00054	-3.69687105	-3.85349672	-3.5402454	0.0000000
PF13385-PF00054	8.30423892	8.14761325	8.4608646	0.0000000
PF00337-PF00139	-2.30591622	-2.46254189	-2.1492905	0.0000000
PF00354-PF00139	-3.60509416	-3.76171983	-3.4484685	0.0000000
PF00457-PF00139	-8.90000969	-9.05663536	-8.7433840	0.0000000
PF00722-PF00139	1.65642634	1.49980067	1.8130520	0.0000000
PF00840-PF00139	-9.18738173	-9.34400740	-9.0307561	0.0000000
PF01670-PF00139	-2.22469992	-2.38132559	-2.0680743	0.0000000
PF02210-PF00139	-0.49393677	-0.65056244	-0.3373111	0.0000000
PF03388-PF00139	-1.69279717	-1.84942284	-1.5361715	0.0000000
PF03935-PF00139	0.92092167	0.76429600	1.0775473	0.0000000

PF06439-PF00139	1.18469826	1.02807259	1.3413239	0.0000000
PF10287-PF00139	-8.60153567	-8.75816134	-8.4449100	0.0000000
PF13385-PF00139	3.39957430	3.24294863	3.5562000	0.0000000
PF00354-PF00337	-1.29917794	-1.45580361	-1.1425523	0.0000000
PF00457-PF00337	-6.59409347	-6.75071914	-6.4374678	0.0000000
PF00722-PF00337	3.96234256	3.80571689	4.1189682	0.0000000
PF00840-PF00337	-6.88146551	-7.03809118	-6.7248398	0.0000000
PF01670-PF00337	0.08121629	-0.07540938	0.2378420	0.9419500
PF02210-PF00337	1.81197945	1.65535378	1.9686051	0.0000000
PF03388-PF00337	0.61311904	0.45649337	0.7697447	0.0000000
PF03935-PF00337	3.22683789	3.07021222	3.3834636	0.0000000
PF06439-PF00337	3.49061448	3.33398881	3.6472401	0.0000000
PF10287-PF00337	-6.29561945	-6.45224512	-6.1389938	0.0000000
PF13385-PF00337	5.70549052	5.54886485	5.8621162	0.0000000
PF00457-PF00354	-5.29491553	-5.45154120	-5.1382899	0.0000000
PF00722-PF00354	5.26152050	5.10489483	5.4181462	0.0000000
PF00840-PF00354	-5.58228757	-5.73891324	-5.4256619	0.0000000
PF01670-PF00354	1.38039424	1.22376857	1.5370199	0.0000000
PF02210-PF00354	3.11115739	2.95453172	3.2677831	0.0000000
PF03388-PF00354	1.91229699	1.75567132	2.0689227	0.0000000
PF03935-PF00354	4.52601583	4.36939016	4.6826415	0.0000000
PF06439-PF00354	4.78979242	4.63316675	4.9464181	0.0000000
PF10287-PF00354	-4.99644151	-5.15306718	-4.8398158	0.0000000
PF13385-PF00354	7.00466846	6.84804280	7.1612941	0.0000000
PF00722-PF00457	10.55643603	10.39981036	10.7130617	0.0000000
PF00840-PF00457	-0.28737204	-0.44399771	-0.1307464	0.0000009
PF01670-PF00457	6.67530976	6.51868409	6.8319354	0.0000000
PF02210-PF00457	8.40607292	8.24944725	8.5626986	0.0000000
PF03388-PF00457	7.20721251	7.05058684	7.3638382	0.0000000
PF03935-PF00457	9.82093136	9.66430569	9.9775570	0.0000000
PF06439-PF00457	10.08470795	9.92808228	10.2413336	0.0000000
PF10287-PF00457	0.29847402	0.14184835	0.4550997	0.0000003
PF13385-PF00457	12.29958399	12.14295832	12.4562097	0.0000000
PF00840-PF00722	-10.84380807	-11.00043374	-10.6871824	0.0000000
PF01670-PF00722	-3.88112627	-4.03775194	-3.7245006	0.0000000
PF02210-PF00722	-2.15036311	-2.30698878	-1.9937374	0.0000000
PF03388-PF00722	-3.34922352	-3.50584919	-3.1925978	0.0000000
PF03935-PF00722	-0.73550467	-0.89213034	-0.5788790	0.0000000
PF06439-PF00722	-0.47172808	-0.62835375	-0.3151024	0.0000000
PF10287-PF00722	-10.25796201	-10.41458768	-10.1013363	0.0000000
PF13385-PF00722	1.74314796	1.58652229	1.8997736	0.0000000
PF01670-PF00840	6.96268181	6.80605614	7.1193075	0.0000000
PF02210-PF00840	8.69344496	8.53681929	8.8500706	0.0000000
PF03388-PF00840	7.49458456	7.33795889	7.6512102	0.0000000
PF03935-PF00840	10.10830340	9.95167773	10.2649291	0.0000000
PF06439-PF00840	10.37207999	10.21545432	10.5287057	0.0000000
PF10287-PF00840	0.58584606	0.42922039	0.7424717	0.0000000
PF13385-PF00840	12.58695603	12.43033036	12.7435817	0.0000000
PF02210-PF01670	1.73076315	1.57413748	1.8873888	0.0000000
PF03388-PF01670	0.53190275	0.37527708	0.6885284	0.0000000
PF03935-PF01670	3.14562159	2.98899592	3.3022473	0.0000000
PF06439-PF01670	3.40939819	3.25277252	3.5660239	0.0000000
PF10287-PF01670	-6.37683575	-6.53346141	-6.2202101	0.0000000
PF13385-PF01670	5.62427423	5.46764856	5.7808999	0.0000000

PF03388-PF02210	-1.19886040	-1.35548607	-1.0422347	0.0000000
PF03935-PF02210	1.41485844	1.25823277	1.5714841	0.0000000
PF06439-PF02210	1.67863503	1.52200936	1.8352607	0.0000000
PF10287-PF02210	-8.10759890	-8.26422457	-7.9509732	0.0000000
PF13385-PF02210	3.89351107	3.73688541	4.0501367	0.0000000
PF03935-PF03388	2.61371884	2.45709317	2.7703445	0.0000000
PF06439-PF03388	2.87749544	2.72086977	3.0341211	0.0000000
PF10287-PF03388	-6.90873850	-7.06536416	-6.7521128	0.0000000
PF13385-PF03388	5.09237148	4.93574581	5.2489971	0.0000000
PF06439-PF03935	0.26377659	0.10715092	0.4204023	0.0000133
PF10287-PF03935	-9.52245734	-9.67908301	-9.3658317	0.0000000
PF13385-PF03935	2.47865263	2.32202696	2.6352783	0.0000000
PF10287-PF06439	-9.78623393	-9.94285960	-9.6296083	0.0000000
PF13385-PF06439	2.21487604	2.05825037	2.3715017	0.0000000
PF13385-PF10287	12.00110997	11.84448430	12.1577356	0.0000000

Pares iguais

diff	lwr	upr	p adj
0.08121629	-0.07540938	0.23784196	0.94195004

=====

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50817, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-58.6129106	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-28.8857721	1.801864e-183	2.602692e-183
3 PF00139 - PF00337	29.7271384	3.425158e-194	5.027249e-194
4 PF00054 - PF00354	-12.3094298	8.059049e-35	9.053994e-35
5 PF00139 - PF00354	46.3034807	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	16.5763423	1.033297e-61	1.205513e-61
7 PF00054 - PF00457	30.2041387	2.089985e-200	3.117846e-200
8 PF00139 - PF00457	88.8170493	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	59.0899108	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	42.5135686	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-82.9144754	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-24.3015648	1.887054e-130	2.418619e-130
13 PF00337 - PF00722	-54.0287033	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-70.6050456	0.000000e+00	0.000000e+00
15 PF00457 - PF00722	-113.1186141	0.000000e+00	0.000000e+00
16 PF00054 - PF00840	30.4146122	3.520714e-203	5.339750e-203
17 PF00139 - PF00840	89.0275227	0.000000e+00	0.000000e+00

18 PF00337 - PF00840 59.3003843 0.000000e+00 0.000000e+00
 19 PF00354 - PF00840 42.7240420 0.000000e+00 0.000000e+00
 20 PF00457 - PF00840 0.2104735 8.332982e-01 8.332982e-01
 21 PF00722 - PF00840 113.3290876 0.000000e+00 0.000000e+00
 22 PF00054 - PF01670 -26.4324584 5.805849e-154 7.885556e-154
 23 PF00139 - PF01670 32.1804522 3.313322e-227 5.110378e-227
 24 PF00337 - PF01670 2.4533137 1.415468e-02 1.431196e-02
 25 PF00354 - PF01670 -14.1230285 2.739598e-45 3.155740e-45
 26 PF00457 - PF01670 -56.6365971 0.000000e+00 0.000000e+00
 27 PF00722 - PF01670 56.4820170 0.000000e+00 0.000000e+00
 28 PF00840 - PF01670 -56.8470706 0.000000e+00 0.000000e+00
 29 PF00054 - PF02210 -55.2123506 0.000000e+00 0.000000e+00
 30 PF00139 - PF02210 3.4005600 6.724798e-04 7.033984e-04
 31 PF00337 - PF02210 -26.3265785 9.519516e-153 1.273935e-152
 32 PF00354 - PF02210 -42.9029207 0.000000e+00 0.000000e+00
 33 PF00457 - PF02210 -85.4164893 0.000000e+00 0.000000e+00
 34 PF00722 - PF02210 27.7021248 6.582548e-169 9.215567e-169
 35 PF00840 - PF02210 -85.6269628 0.000000e+00 0.000000e+00
 36 PF01670 - PF02210 -28.7798922 3.829436e-182 5.444979e-182
 37 PF00054 - PF03388 -33.9473821 1.333224e-252 2.128480e-252
 38 PF00139 - PF03388 24.6655285 2.507835e-134 3.260185e-134
 39 PF00337 - PF03388 -5.0616100 4.157308e-07 4.399012e-07
 40 PF00354 - PF03388 -21.6379522 7.893160e-104 9.839418e-104
 41 PF00457 - PF03388 -64.1515208 0.000000e+00 0.000000e+00
 42 PF00722 - PF03388 48.9670933 0.000000e+00 0.000000e+00
 43 PF00840 - PF03388 -64.3619943 0.000000e+00 0.000000e+00
 44 PF01670 - PF03388 -7.5149237 5.694414e-14 6.243273e-14
 45 PF02210 - PF03388 21.2649685 2.396513e-100 2.947063e-100
 46 PF00054 - PF03935 -71.8551272 0.000000e+00 0.000000e+00
 47 PF00139 - PF03935 -13.2422167 5.004950e-40 5.693130e-40
 48 PF00337 - PF03935 -42.9693551 0.000000e+00 0.000000e+00
 49 PF00354 - PF03935 -59.5456974 0.000000e+00 0.000000e+00
 50 PF00457 - PF03935 -102.0592660 0.000000e+00 0.000000e+00
 51 PF00722 - PF03935 11.0593481 1.975278e-28 2.192077e-28
 52 PF00840 - PF03935 -102.2697394 0.000000e+00 0.000000e+00
 53 PF01670 - PF03935 -45.4226689 0.000000e+00 0.000000e+00
 54 PF02210 - PF03935 -16.6427767 3.414129e-62 4.034879e-62
 55 PF03388 - PF03935 -37.9077452 0.000000e+00 0.000000e+00
 56 PF00054 - PF06439 -77.8098475 0.000000e+00 0.000000e+00
 57 PF00139 - PF06439 -19.1969369 3.925709e-82 4.700520e-82
 58 PF00337 - PF06439 -48.9240753 0.000000e+00 0.000000e+00
 59 PF00354 - PF06439 -65.5004176 0.000000e+00 0.000000e+00
 60 PF00457 - PF06439 -108.0139862 0.000000e+00 0.000000e+00
 61 PF00722 - PF06439 5.1046279 3.314460e-07 3.548422e-07
 62 PF00840 - PF06439 -108.2244596 0.000000e+00 0.000000e+00
 63 PF01670 - PF06439 -51.3773891 0.000000e+00 0.000000e+00
 64 PF02210 - PF06439 -22.5974969 4.586439e-113 5.796749e-113
 65 PF03388 - PF06439 -43.8624654 0.000000e+00 0.000000e+00
 66 PF03935 - PF06439 -5.9547202 2.605171e-09 2.822269e-09
 67 PF00054 - PF10287 27.0510798 3.709369e-161 5.114433e-161
 68 PF00139 - PF10287 85.6639903 0.000000e+00 0.000000e+00
 69 PF00337 - PF10287 55.9368519 0.000000e+00 0.000000e+00
 70 PF00354 - PF10287 39.3605096 0.000000e+00 0.000000e+00
 71 PF00457 - PF10287 -3.1530590 1.615692e-03 1.651999e-03

72 PF00722 - PF10287 109.9655552 0.000000e+00 0.000000e+00
73 PF00840 - PF10287 -3.3635324 7.695179e-04 7.957515e-04
74 PF01670 - PF10287 53.4835381 0.000000e+00 0.000000e+00
75 PF02210 - PF10287 82.2634303 0.000000e+00 0.000000e+00
76 PF03388 - PF10287 60.9984618 0.000000e+00 0.000000e+00
77 PF03935 - PF10287 98.9062070 0.000000e+00 0.000000e+00
78 PF06439 - PF10287 104.8609272 0.000000e+00 0.000000e+00
79 PF00054 - PF13385 -104.0635816 0.000000e+00 0.000000e+00
80 PF00139 - PF13385 -45.4506711 0.000000e+00 0.000000e+00
81 PF00337 - PF13385 -75.1778095 0.000000e+00 0.000000e+00
82 PF00354 - PF13385 -91.7541518 0.000000e+00 0.000000e+00
83 PF00457 - PF13385 -134.2677203 0.000000e+00 0.000000e+00
84 PF00722 - PF13385 -21.1491062 2.812194e-99 3.412129e-99
85 PF00840 - PF13385 -134.4781938 0.000000e+00 0.000000e+00
86 PF01670 - PF13385 -77.6311232 0.000000e+00 0.000000e+00
87 PF02210 - PF13385 -48.8512310 0.000000e+00 0.000000e+00
88 PF03388 - PF13385 -70.1161995 0.000000e+00 0.000000e+00
89 PF03935 - PF13385 -32.2084544 1.343890e-227 2.108516e-227
90 PF06439 - PF13385 -26.2537342 6.479287e-152 8.545146e-152
91 PF10287 - PF13385 -131.1146614 0.000000e+00 0.000000e+00

Pares iguais

Comparison	Z	P.unadj	P.adj
20 PF00457 - PF00840	0.2104735	0.833298163	0.833298163
24 PF00337 - PF01670	2.4533137	0.014154683	0.014311957
71 PF00457 - PF10287	-3.1530590	0.001615692	0.001651999

=====

ANOVA da entropia entre familias controlando **s = 0.4**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	587.26	< 2.2e-16 ***
	69286		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 12149, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 519186 39937 13010 <2e-16 ***

Residuals 69286 212694 3

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf	diff	lwr	upr	p adj
PF00139-PF00054	3.35011986	3.24024985	3.459989880	0.0000000
PF00337-PF00054	1.77976513	1.66989512	1.889635148	0.0000000
PF00354-PF00054	0.86983110	0.75996108	0.979701114	0.0000000
PF00457-PF00054	-2.90805257	-3.01792258	-2.798182551	0.0000000
PF00722-PF00054	4.47616420	4.36629418	4.586034215	0.0000000
PF00840-PF00054	-3.02491191	-3.13478193	-2.915041895	0.0000000
PF01670-PF00054	1.84748225	1.73761223	1.957352266	0.0000000
PF02210-PF00054	3.02280363	2.91293362	3.132673651	0.0000000
PF03388-PF00054	2.16408814	2.05421812	2.273958155	0.0000000
PF03935-PF00054	3.95612680	3.84625678	4.065996814	0.0000000
PF06439-PF00054	4.16122933	4.05135932	4.271099349	0.0000000
PF10287-PF00054	-2.61026075	-2.72013077	-2.500390735	0.0000000
PF13385-PF00054	5.65958199	5.54971198	5.769452009	0.0000000
PF00337-PF00139	-1.57035473	-1.68022475	-1.460484716	0.0000000
PF00354-PF00139	-2.48028877	-2.59015878	-2.370418750	0.0000000
PF00457-PF00139	-6.25817243	-6.36804245	-6.148302415	0.0000000
PF00722-PF00139	1.12604434	1.01617432	1.235914352	0.0000000
PF00840-PF00139	-6.37503177	-6.48490179	-6.265161759	0.0000000
PF01670-PF00139	-1.50263761	-1.61250763	-1.392767598	0.0000000
PF02210-PF00139	-0.32731623	-0.43718625	-0.217446213	0.0000000
PF03388-PF00139	-1.18603172	-1.29590174	-1.076161709	0.0000000
PF03935-PF00139	0.60600693	0.49613692	0.715876950	0.0000000

PF06439-PF00139 0.81110947 0.70123945 0.920979485 0.0000000
PF10287-PF00139 -5.96038061 -6.07025063 -5.850510599 0.0000000
PF13385-PF00139 2.30946213 2.19959211 2.419332145 0.0000000
PF00354-PF00337 -0.90993403 -1.01980405 -0.800064017 0.0000000
PF00457-PF00337 -4.68781770 -4.79768771 -4.577947682 0.0000000
PF00722-PF00337 2.69639907 2.58652905 2.806269084 0.0000000
PF00840-PF00337 -4.80467704 -4.91454706 -4.694807026 0.0000000
PF01670-PF00337 0.06771712 -0.04215290 0.177587135 0.8143674
PF02210-PF00337 1.24303850 1.13316849 1.352908519 0.0000000
PF03388-PF00337 0.38432301 0.27445299 0.494193023 0.0000000
PF03935-PF00337 2.17636167 2.06649165 2.286231682 0.0000000
PF06439-PF00337 2.38146420 2.27159419 2.491334218 0.0000000
PF10287-PF00337 -4.39002588 -4.49989590 -4.280155867 0.0000000
PF13385-PF00337 3.87981686 3.76994685 3.989686877 0.0000000
PF00457-PF00354 -3.77788367 -3.88775368 -3.668013649 0.0000000
PF00722-PF00354 3.60633310 3.49646309 3.716203117 0.0000000
PF00840-PF00354 -3.89474301 -4.00461302 -3.784872993 0.0000000
PF01670-PF00354 0.97765115 0.86778114 1.087521168 0.0000000
PF02210-PF00354 2.15297254 2.04310252 2.262842552 0.0000000
PF03388-PF00354 1.29425704 1.18438702 1.404127057 0.0000000
PF03935-PF00354 3.08629570 2.97642568 3.196165716 0.0000000
PF06439-PF00354 3.29139823 3.18152822 3.401268251 0.0000000
PF10287-PF00354 -3.48009185 -3.58996187 -3.370221833 0.0000000
PF13385-PF00354 4.78975089 4.67988088 4.899620911 0.0000000
PF00722-PF00457 7.38421677 7.27434675 7.494086782 0.0000000
PF00840-PF00457 -0.11685934 -0.22672936 -0.006989328 0.0558345
PF01670-PF00457 4.75553482 4.64566480 4.865404833 0.0000000
PF02210-PF00457 5.93085620 5.82098619 6.040726217 0.0000000
PF03388-PF00457 5.07214071 4.96227069 5.182010722 0.0000000
PF03935-PF00457 6.86417936 6.75430935 6.974049381 0.0000000
PF06439-PF00457 7.06928190 6.95941188 7.179151916 0.0000000
PF10287-PF00457 0.29779182 0.18792180 0.407661832 0.0000000
PF13385-PF00457 8.56763456 8.45776454 8.677504576 0.0000000
PF00840-PF00722 -7.50107611 -7.61094613 -7.391206094 0.0000000
PF01670-PF00722 -2.62868195 -2.73855197 -2.518811933 0.0000000
PF02210-PF00722 -1.45336056 -1.56323058 -1.343490549 0.0000000
PF03388-PF00722 -2.31207606 -2.42194608 -2.202206044 0.0000000
PF03935-PF00722 -0.52003740 -0.62990742 -0.410167385 0.0000000
PF06439-PF00722 -0.31493487 -0.42480488 -0.205064850 0.0000000
PF10287-PF00722 -7.08642495 -7.19629497 -6.976554935 0.0000000
PF13385-PF00722 1.18341779 1.07354778 1.293287809 0.0000000
PF01670-PF00840 4.87239416 4.76252414 4.982264177 0.0000000
PF02210-PF00840 6.04771555 5.93784553 6.157585561 0.0000000
PF03388-PF00840 5.18900005 5.07913003 5.298870066 0.0000000
PF03935-PF00840 6.98103871 6.87116869 7.090908725 0.0000000
PF06439-PF00840 7.18614124 7.07627123 7.296011260 0.0000000
PF10287-PF00840 0.41465116 0.30478114 0.524521175 0.0000000
PF13385-PF00840 8.68449390 8.57462389 8.794363919 0.0000000
PF02210-PF01670 1.17532138 1.06545137 1.285191400 0.0000000
PF03388-PF01670 0.31660589 0.20673587 0.426475905 0.0000000
PF03935-PF01670 2.10864455 1.99877453 2.218514564 0.0000000
PF06439-PF01670 2.31374708 2.20387707 2.423617099 0.0000000
PF10287-PF01670 -4.45774300 -4.56761302 -4.347872985 0.0000000
PF13385-PF01670 3.81209974 3.70222973 3.921969759 0.0000000

PF03388-PF02210	-0.85871550	-0.96858551	-0.748845480	0.0000000
PF03935-PF02210	0.93332316	0.82345315	1.043193179	0.0000000
PF06439-PF02210	1.13842570	1.02855568	1.248295714	0.0000000
PF10287-PF02210	-5.63306439	-5.74293440	-5.523194370	0.0000000
PF13385-PF02210	2.63677836	2.52690834	2.746648374	0.0000000
PF03935-PF03388	1.79203866	1.68216864	1.901908675	0.0000000
PF06439-PF03388	1.99714119	1.88727118	2.107011210	0.0000000
PF10287-PF03388	-4.77434889	-4.88421891	-4.664478874	0.0000000
PF13385-PF03388	3.49549385	3.38562384	3.605363870	0.0000000
PF06439-PF03935	0.20510254	0.09523252	0.314972551	0.0000005
PF10287-PF03935	-6.56638755	-6.67625757	-6.456517533	0.0000000
PF13385-PF03935	1.70345519	1.59358518	1.813325211	0.0000000
PF10287-PF06439	-6.77149008	-6.88136010	-6.661620068	0.0000000
PF13385-PF06439	1.49835266	1.38848264	1.608222676	0.0000000
PF13385-PF10287	8.26984274	8.15997273	8.379712760	0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF01670-PF00337	0.06771712	-0.0421529	0.177587135	0.81436736
PF00840-PF00457	-0.11685934	-0.2267294	-0.006989328	0.05583448
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50691, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-58.1589029	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-28.7499054	9.082263e-182	1.291384e-181
3 PF00139 - PF00337	29.4089974	4.213768e-190	6.184724e-190
4 PF00054 - PF00354	-11.8678176	1.739462e-32	1.954210e-32
5 PF00139 - PF00354	46.2910853	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	16.8820878	6.095057e-64	7.203249e-64
7 PF00054 - PF00457	30.9708608	1.330935e-210	2.018585e-210
8 PF00139 - PF00457	89.1297637	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	59.7207663	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	42.8386784	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-82.4152145	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-24.2563116	5.672191e-130	7.269992e-130
13 PF00337 - PF00722	-53.6653090	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-70.5473969	0.000000e+00	0.000000e+00
15 PF00457 - PF00722	-113.3860753	0.000000e+00	0.000000e+00

16 PF00054 - PF00840 30.4764102 5.353508e-204 7.986381e-204
 17 PF00139 - PF00840 88.6353131 0.000000e+00 0.000000e+00
 18 PF00337 - PF00840 59.2263156 0.000000e+00 0.000000e+00
 19 PF00354 - PF00840 42.3442278 0.000000e+00 0.000000e+00
 20 PF00457 - PF00840 -0.4944506 6.209880e-01 6.209880e-01
 21 PF00722 - PF00840 112.8916247 0.000000e+00 0.000000e+00
 22 PF00054 - PF01670 -26.2750841 3.695279e-152 4.945153e-152
 23 PF00139 - PF01670 31.8838188 4.475662e-223 6.903140e-223
 24 PF00337 - PF01670 2.4748213 1.333028e-02 1.347840e-02
 25 PF00354 - PF01670 -14.4072665 4.657833e-47 5.365352e-47
 26 PF00457 - PF01670 -57.2459449 0.000000e+00 0.000000e+00
 27 PF00722 - PF01670 56.1401304 0.000000e+00 0.000000e+00
 28 PF00840 - PF01670 -56.7514943 0.000000e+00 0.000000e+00
 29 PF00054 - PF02210 -55.1380043 0.000000e+00 0.000000e+00
 30 PF00139 - PF02210 3.0208986 2.520258e-03 2.576893e-03
 31 PF00337 - PF02210 -26.3880989 1.876676e-153 2.548919e-153
 32 PF00354 - PF02210 -43.2701867 0.000000e+00 0.000000e+00
 33 PF00457 - PF02210 -86.1088651 0.000000e+00 0.000000e+00
 34 PF00722 - PF02210 27.2772101 7.906428e-164 1.106900e-163
 35 PF00840 - PF02210 -85.6144145 0.000000e+00 0.000000e+00
 36 PF01670 - PF02210 -28.8629202 3.488404e-183 5.038805e-183
 37 PF00054 - PF03388 -32.9777052 1.695862e-238 2.707429e-238
 38 PF00139 - PF03388 25.1811977 6.437207e-140 8.368369e-140
 39 PF00337 - PF03388 -4.2277998 2.359877e-05 2.497079e-05
 40 PF00354 - PF03388 -21.1098876 6.452596e-99 7.829149e-99
 41 PF00457 - PF03388 -63.9485660 0.000000e+00 0.000000e+00
 42 PF00722 - PF03388 49.4375093 0.000000e+00 0.000000e+00
 43 PF00840 - PF03388 -63.4541154 0.000000e+00 0.000000e+00
 44 PF01670 - PF03388 -6.7026211 2.047138e-11 2.244452e-11
 45 PF02210 - PF03388 22.1602991 8.299761e-109 1.034628e-108
 46 PF00054 - PF03935 -70.9276180 0.000000e+00 0.000000e+00
 47 PF00139 - PF03935 -12.7687151 2.451683e-37 2.788789e-37
 48 PF00337 - PF03935 -42.1777125 0.000000e+00 0.000000e+00
 49 PF00354 - PF03935 -59.0598004 0.000000e+00 0.000000e+00
 50 PF00457 - PF03935 -101.8984788 0.000000e+00 0.000000e+00
 51 PF00722 - PF03935 11.4875965 1.522901e-30 1.690049e-30
 52 PF00840 - PF03935 -101.4040281 0.000000e+00 0.000000e+00
 53 PF01670 - PF03935 -44.6525339 0.000000e+00 0.000000e+00
 54 PF02210 - PF03935 -15.7896136 3.668199e-56 4.279565e-56
 55 PF03388 - PF03935 -37.9499128 0.000000e+00 0.000000e+00
 56 PF00054 - PF06439 -77.6302373 0.000000e+00 0.000000e+00
 57 PF00139 - PF06439 -19.4713344 1.921990e-84 2.301330e-84
 58 PF00337 - PF06439 -48.8803319 0.000000e+00 0.000000e+00
 59 PF00354 - PF06439 -65.7624197 0.000000e+00 0.000000e+00
 60 PF00457 - PF06439 -108.6010981 0.000000e+00 0.000000e+00
 61 PF00722 - PF06439 4.7849772 1.710068e-06 1.830778e-06
 62 PF00840 - PF06439 -108.1066475 0.000000e+00 0.000000e+00
 63 PF01670 - PF06439 -51.3551532 0.000000e+00 0.000000e+00
 64 PF02210 - PF06439 -22.4922330 4.944768e-112 6.249638e-112
 65 PF03388 - PF06439 -44.6525321 0.000000e+00 0.000000e+00
 66 PF03935 - PF06439 -6.7026193 2.047163e-11 2.217759e-11
 67 PF00054 - PF10287 27.2020579 6.140850e-163 8.466930e-163
 68 PF00139 - PF10287 85.3609608 0.000000e+00 0.000000e+00
 69 PF00337 - PF10287 55.9519633 0.000000e+00 0.000000e+00

70	PF00354 - PF10287	39.0698755	0.000000e+00	0.000000e+00
71	PF00457 - PF10287	-3.7688029	1.640323e-04	1.715740e-04
72	PF00722 - PF10287	109.6172723	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-3.2743523	1.059045e-03	1.095149e-03
74	PF01670 - PF10287	53.4771420	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.3400622	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	60.1797631	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	98.1296758	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.8322952	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-103.7551377	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-45.5962348	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-75.0052322	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-91.8873201	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-134.7259985	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-21.3399232	4.837406e-101	5.948702e-101
85	PF00840 - PF13385	-134.2315478	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-77.4800536	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-48.6171333	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-70.7774325	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-32.8275197	2.384698e-236	3.741509e-236
90	PF06439 - PF13385	-26.1249004	1.901065e-150	2.507202e-150
91	PF10287 - PF13385	-130.9571955	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
20	PF00457 - PF00840	-0.4944506	0.620987966	0.620987966
24	PF00337 - PF01670	2.4748213	0.013330284	0.013478399
30	PF00139 - PF02210	3.0208986	0.002520258	0.002576893
73	PF00840 - PF10287	-3.2743523	0.001059045	0.001095149

ANOVA da entropia entre familias controlando **s = 0.5**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	638.38	< 2.2e-16 ***
		69286	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 13216, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 251952 19381 12649 <2e-16 ***

Residuals 69286 106163 2

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	2.2958614	2.21823898	2.37348386	0.0000000
PF00337-PF00054	1.2228923	1.14526986	1.30051474	0.0000000
PF00354-PF00054	0.5838353	0.50621290	0.66145778	0.0000000
PF00457-PF00054	-2.1305508	-2.20817320	-2.05292832	0.0000000
PF00722-PF00054	3.0645060	2.98688351	3.14212839	0.0000000
PF00840-PF00054	-2.1477005	-2.22532290	-2.07007802	0.0000000
PF01670-PF00054	1.2791684	1.20154596	1.35679084	0.0000000
PF02210-PF00054	2.0794737	2.00185126	2.15709614	0.0000000
PF03388-PF00054	1.4623615	1.38473910	1.53998398	0.0000000
PF03935-PF00054	2.6952013	2.61757890	2.77282378	0.0000000
PF06439-PF00054	2.8536503	2.77602783	2.93127272	0.0000000
PF10287-PF00054	-1.8528493	-1.93047173	-1.77522685	0.0000000
PF13385-PF00054	3.8701240	3.79250159	3.94774647	0.0000000
PF00337-PF00139	-1.0729691	-1.15059156	-0.99534668	0.0000000
PF00354-PF00139	-1.7120261	-1.78964851	-1.63440363	0.0000000
PF00457-PF00139	-4.4264122	-4.50403462	-4.34878974	0.0000000
PF00722-PF00139	0.7686445	0.69102210	0.84626698	0.0000000
PF00840-PF00139	-4.4435619	-4.52118432	-4.36593944	0.0000000
PF01670-PF00139	-1.0166930	-1.09431546	-0.93907058	0.0000000
PF02210-PF00139	-0.2163877	-0.29401016	-0.13876528	0.0000000

PF03388-PF00139 -0.8334999 -0.91112231 -0.75587743 0.0000000
 PF03935-PF00139 0.3993399 0.32171748 0.47696237 0.0000000
 PF06439-PF00139 0.5577889 0.48016642 0.63541130 0.0000000
 PF10287-PF00139 -4.1487107 -4.22633315 -4.07108827 0.0000000
 PF13385-PF00139 1.5742626 1.49664018 1.65188506 0.0000000
 PF00354-PF00337 -0.6390570 -0.71667939 -0.56143451 0.0000000
 PF00457-PF00337 -3.3534431 -3.43106550 -3.27582062 0.0000000
 PF00722-PF00337 1.8416137 1.76399122 1.91923610 0.0000000
 PF00840-PF00337 -3.3705928 -3.44821520 -3.29297032 0.0000000
 PF01670-PF00337 0.0562761 -0.02134634 0.13389854 0.5836166
 PF02210-PF00337 0.8565814 0.77895896 0.93420384 0.0000000
 PF03388-PF00337 0.2394692 0.16184681 0.31709169 0.0000000
 PF03935-PF00337 1.4723090 1.39468660 1.54993148 0.0000000
 PF06439-PF00337 1.6307580 1.55313554 1.70838042 0.0000000
 PF10287-PF00337 -3.0757416 -3.15336403 -2.99811915 0.0000000
 PF13385-PF00337 2.6472317 2.56960929 2.72485418 0.0000000
 PF00457-PF00354 -2.7143861 -2.79200855 -2.63676366 0.0000000
 PF00722-PF00354 2.4806706 2.40304817 2.55829305 0.0000000
 PF00840-PF00354 -2.7315358 -2.80915825 -2.65391337 0.0000000
 PF01670-PF00354 0.6953331 0.61771061 0.77295549 0.0000000
 PF02210-PF00354 1.4956384 1.41801591 1.57326079 0.0000000
 PF03388-PF00354 0.8785262 0.80090376 0.95614864 0.0000000
 PF03935-PF00354 2.1113660 2.03374356 2.18898844 0.0000000
 PF06439-PF00354 2.2698149 2.19219249 2.34743737 0.0000000
 PF10287-PF00354 -2.4366846 -2.51430708 -2.35906220 0.0000000
 PF13385-PF00354 3.2862887 3.20866625 3.36391113 0.0000000
 PF00722-PF00457 5.1950567 5.11743427 5.27267915 0.0000000
 PF00840-PF00457 -0.0171497 -0.09477214 0.06047274 0.9999903
 PF01670-PF00457 3.4097192 3.33209672 3.48734160 0.0000000
 PF02210-PF00457 4.2100245 4.13240202 4.28764690 0.0000000
 PF03388-PF00457 3.5929123 3.51528986 3.67053474 0.0000000
 PF03935-PF00457 4.8257521 4.74812966 4.90337454 0.0000000
 PF06439-PF00457 4.9842010 4.90657860 5.06182348 0.0000000
 PF10287-PF00457 0.2777015 0.20007903 0.35532391 0.0000000
 PF13385-PF00457 6.0006748 5.92305235 6.07829723 0.0000000
 PF00840-PF00722 -5.2122064 -5.28982886 -5.13458397 0.0000000
 PF01670-PF00722 -1.7853376 -1.86296000 -1.70771512 0.0000000
 PF02210-PF00722 -0.9850323 -1.06265470 -0.90740982 0.0000000
 PF03388-PF00722 -1.6021444 -1.67976685 -1.52452197 0.0000000
 PF03935-PF00722 -0.3693046 -0.44692705 -0.29168217 0.0000000
 PF06439-PF00722 -0.2108557 -0.28847812 -0.13323324 0.0000000
 PF10287-PF00722 -4.9173552 -4.99497769 -4.83973281 0.0000000
 PF13385-PF00722 0.8056181 0.72799564 0.88324052 0.0000000
 PF01670-PF00840 3.4268689 3.34924642 3.50449130 0.0000000
 PF02210-PF00840 4.2271742 4.14955172 4.30479660 0.0000000
 PF03388-PF00840 3.6100620 3.53243956 3.68768445 0.0000000
 PF03935-PF00840 4.8429018 4.76527936 4.92052424 0.0000000
 PF06439-PF00840 5.0013507 4.92372830 5.07897318 0.0000000
 PF10287-PF00840 0.2948512 0.21722873 0.37247361 0.0000000
 PF13385-PF00840 6.0178245 5.94020205 6.09544694 0.0000000
 PF02210-PF01670 0.8003053 0.72268286 0.87792774 0.0000000
 PF03388-PF01670 0.1831931 0.10557071 0.26081559 0.0000000
 PF03935-PF01670 1.4160329 1.33841050 1.49365538 0.0000000
 PF06439-PF01670 1.5744819 1.49685944 1.65210432 0.0000000

PF10287-PF01670	-3.1320177	-3.20964013	-3.05439525	0.0000000
PF13385-PF01670	2.5909556	2.51333319	2.66857808	0.0000000
PF03388-PF02210	-0.6171122	-0.69473460	-0.53948971	0.0000000
PF03935-PF02210	0.6157276	0.53810520	0.69335008	0.0000000
PF06439-PF02210	0.7741766	0.69655414	0.85179902	0.0000000
PF10287-PF02210	-3.9323230	-4.00994543	-3.85470055	0.0000000
PF13385-PF02210	1.7906503	1.71302789	1.86827278	0.0000000
PF03935-PF03388	1.2328398	1.15521736	1.31046224	0.0000000
PF06439-PF03388	1.3912887	1.31366629	1.46891117	0.0000000
PF10287-PF03388	-3.3152108	-3.39283328	-3.23758840	0.0000000
PF13385-PF03388	2.4077625	2.33014005	2.48538493	0.0000000
PF06439-PF03935	0.1584489	0.08082649	0.23607138	0.0000000
PF10287-PF03935	-4.5480506	-4.62567307	-4.47042819	0.0000000
PF13385-PF03935	1.1749227	1.09730025	1.25254513	0.0000000
PF10287-PF06439	-4.7064996	-4.78412201	-4.62887713	0.0000000
PF13385-PF06439	1.0164738	0.93885132	1.09409620	0.0000000
PF13385-PF10287	5.7229733	5.64535089	5.80059577	0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF01670-PF00337	0.0562761	-0.02134634	0.13389854	0.5836166
PF00840-PF00457	-0.0171497	-0.09477214	0.06047274	0.9999903
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50549, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-57.651588	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-28.602171	6.313719e-180	8.977320e-180
3 PF00139 - PF00337	29.049417	1.565111e-185	2.297180e-185
4 PF00054 - PF00354	-11.421714	3.257452e-30	3.614978e-30
5 PF00139 - PF00354	46.229874	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	17.180457	3.719683e-66	4.395989e-66
7 PF00054 - PF00457	31.750300	3.145074e-221	4.850877e-221
8 PF00139 - PF00457	89.401889	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	60.352472	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	43.172014	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-81.848228	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-24.196640	2.413489e-129	3.093345e-129
13 PF00337 - PF00722	-53.246057	0.000000e+00	0.000000e+00

14 PF00354 - PF00722 -70.426514 0.000000e+00 0.000000e+00
15 PF00457 - PF00722 -113.598529 0.000000e+00 0.000000e+00
16 PF00054 - PF00840 30.534976 8.951553e-205 1.335396e-204
17 PF00139 - PF00840 88.186564 0.000000e+00 0.000000e+00
18 PF00337 - PF00840 59.137147 0.000000e+00 0.000000e+00
19 PF00354 - PF00840 41.956690 0.000000e+00 0.000000e+00
20 PF00457 - PF00840 -1.215324 2.242425e-01 2.242425e-01
21 PF00722 - PF00840 112.383204 0.000000e+00 0.000000e+00
22 PF00054 - PF01670 -26.108363 2.929836e-150 3.920810e-150
23 PF00139 - PF01670 31.543225 2.220802e-218 3.368217e-218
24 PF00337 - PF01670 2.493808 1.263808e-02 1.277850e-02
25 PF00354 - PF01670 -14.686649 7.850109e-49 9.042531e-49
26 PF00457 - PF01670 -57.858663 0.000000e+00 0.000000e+00
27 PF00722 - PF01670 55.739865 0.000000e+00 0.000000e+00
28 PF00840 - PF01670 -56.643339 0.000000e+00 0.000000e+00
29 PF00054 - PF02210 -55.068637 0.000000e+00 0.000000e+00
30 PF00139 - PF02210 2.582952 9.795904e-03 1.001604e-02
31 PF00337 - PF02210 -26.466465 2.358720e-154 3.203635e-154
32 PF00354 - PF02210 -43.646922 0.000000e+00 0.000000e+00
33 PF00457 - PF02210 -86.818937 0.000000e+00 0.000000e+00
34 PF00722 - PF02210 26.779592 5.586504e-158 7.702604e-158
35 PF00840 - PF02210 -85.603613 0.000000e+00 0.000000e+00
36 PF01670 - PF02210 -28.960274 2.083460e-184 3.009443e-184
37 PF00054 - PF03388 -31.970675 2.788332e-224 4.374796e-224
38 PF00139 - PF03388 25.680913 1.910066e-145 2.483085e-145
39 PF00337 - PF03388 -3.368504 7.557735e-04 7.905217e-04
40 PF00354 - PF03388 -20.548961 7.862810e-94 9.540209e-94
41 PF00457 - PF03388 -63.720975 0.000000e+00 0.000000e+00
42 PF00722 - PF03388 49.877553 0.000000e+00 0.000000e+00
43 PF00840 - PF03388 -62.505651 0.000000e+00 0.000000e+00
44 PF01670 - PF03388 -5.862312 4.564662e-09 4.945050e-09
45 PF02210 - PF03388 23.097961 4.853631e-118 6.134450e-118
46 PF00054 - PF03935 -69.926865 0.000000e+00 0.000000e+00
47 PF00139 - PF03935 -12.275277 1.229701e-34 1.398785e-34
48 PF00337 - PF03935 -41.324694 0.000000e+00 0.000000e+00
49 PF00354 - PF03935 -58.505151 0.000000e+00 0.000000e+00
50 PF00457 - PF03935 -101.677166 0.000000e+00 0.000000e+00
51 PF00722 - PF03935 11.921363 9.159749e-33 1.029058e-32
52 PF00840 - PF03935 -100.461841 0.000000e+00 0.000000e+00
53 PF01670 - PF03935 -43.818502 0.000000e+00 0.000000e+00
54 PF02210 - PF03935 -14.858229 6.153058e-50 7.178568e-50
55 PF03388 - PF03935 -37.956190 0.000000e+00 0.000000e+00
56 PF00054 - PF06439 -77.467091 0.000000e+00 0.000000e+00
57 PF00139 - PF06439 -19.815503 2.188110e-87 2.619974e-87
58 PF00337 - PF06439 -48.864920 0.000000e+00 0.000000e+00
59 PF00354 - PF06439 -66.045377 0.000000e+00 0.000000e+00
60 PF00457 - PF06439 -109.217391 0.000000e+00 0.000000e+00
61 PF00722 - PF06439 4.381137 1.180615e-05 1.249255e-05
62 PF00840 - PF06439 -108.002067 0.000000e+00 0.000000e+00
63 PF01670 - PF06439 -51.358728 0.000000e+00 0.000000e+00
64 PF02210 - PF06439 -22.398455 4.074687e-111 5.079405e-111
65 PF03388 - PF06439 -45.496416 0.000000e+00 0.000000e+00
66 PF03935 - PF06439 -7.540226 4.691581e-14 5.143782e-14
67 PF00054 - PF10287 27.357548 8.781935e-165 1.229471e-164

68	PF00139 - PF10287	85.009137	0.000000e+00	0.000000e+00
69	PF00337 - PF10287	55.959720	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	38.779262	0.000000e+00	0.000000e+00
71	PF00457 - PF10287	-4.392752	1.119248e-05	1.198254e-05
72	PF00722 - PF10287	109.205777	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-3.177428	1.485877e-03	1.536532e-03
74	PF01670 - PF10287	53.465911	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.426185	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	59.328223	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	97.284414	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.824640	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-103.446023	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-45.794435	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-74.843852	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-92.024309	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-135.196323	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-21.597795	1.883968e-103	2.316771e-103
85	PF00840 - PF13385	-133.980999	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-77.337660	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-48.377387	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-71.475348	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-33.519158	2.534924e-246	4.046984e-246
90	PF06439 - PF13385	-25.978932	8.569172e-149	1.130137e-148
91	PF10287 - PF13385	-130.803572	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
20	PF00457 - PF00840	-1.215324	0.224242468	0.224242468
24	PF00337 - PF01670	2.493808	0.012638080	0.012778503
30	PF00139 - PF02210	2.582952	0.009795904	0.010016036
73	PF00840 - PF10287	-3.177428	0.001485877	0.001536532

=====

=====

ANOVA da entropia entre familias controlando **s = 0.6**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	696.72	< 2.2e-16 ***

69286

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 14487, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 123571 9505 12239 <2e-16 ***

Residuals 69286 53810 1

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	1.57845916	1.523196240	1.63372208	0.0000000
PF00337-PF00054	0.84280559	0.787542669	0.89806851	0.0000000
PF00354-PF00054	0.39306315	0.337800230	0.44832607	0.0000000
PF00457-PF00054	-1.57132228	-1.626585200	-1.51605936	0.0000000
PF00722-PF00054	2.10586301	2.050600086	2.16112593	0.0000000
PF00840-PF00054	-1.53307762	-1.588340543	-1.47781470	0.0000000
PF01670-PF00054	0.88975405	0.834491133	0.94501698	0.0000000
PF02210-PF00054	1.43625382	1.380990894	1.49151674	0.0000000
PF03388-PF00054	0.99130335	0.936040429	1.04656627	0.0000000
PF03935-PF00054	1.84242230	1.787159377	1.89768522	0.0000000
PF06439-PF00054	1.96422326	1.908960335	2.01948618	0.0000000
PF10287-PF00054	-1.32236413	-1.377627049	-1.26710120	0.0000000
PF13385-PF00054	2.65584003	2.600577107	2.71110295	0.0000000
PF00337-PF00139	-0.73565357	-0.790916493	-0.68039065	0.0000000
PF00354-PF00139	-1.18539601	-1.240658932	-1.13013309	0.0000000
PF00457-PF00139	-3.14978144	-3.205044362	-3.09451852	0.0000000

PF00722-PF00139 0.52740385 0.472140924 0.58266677 0.0000000
PF00840-PF00139 -3.11153678 -3.166799705 -3.05627386 0.0000000
PF01670-PF00139 -0.68870511 -0.743968029 -0.63344219 0.0000000
PF02210-PF00139 -0.14220535 -0.197468268 -0.08694242 0.0000000
PF03388-PF00139 -0.58715581 -0.642418733 -0.53189289 0.0000000
PF03935-PF00139 0.26396314 0.208700215 0.31922606 0.0000000
PF06439-PF00139 0.38576409 0.330501173 0.44102702 0.0000000
PF10287-PF00139 -2.90082329 -2.956086211 -2.84556037 0.0000000
PF13385-PF00139 1.07738087 1.022117944 1.13264379 0.0000000
PF00354-PF00337 -0.44974244 -0.505005361 -0.39447952 0.0000000
PF00457-PF00337 -2.41412787 -2.469390791 -2.35886495 0.0000000
PF00722-PF00337 1.26305742 1.207794495 1.31832034 0.0000000
PF00840-PF00337 -2.37588321 -2.431146134 -2.32062029 0.0000000
PF01670-PF00337 0.04694846 -0.008314458 0.10221139 0.3082979
PF02210-PF00337 0.59344822 0.538185303 0.64871115 0.0000000
PF03388-PF00337 0.14849776 0.093234838 0.20376068 0.0000000
PF03935-PF00337 0.99961671 0.944353786 1.05487963 0.0000000
PF06439-PF00337 1.12141767 1.066154744 1.17668059 0.0000000
PF10287-PF00337 -2.16516972 -2.220432640 -2.10990680 0.0000000
PF13385-PF00337 1.81303444 1.757771515 1.86829736 0.0000000
PF00457-PF00354 -1.96438543 -2.019648352 -1.90912251 0.0000000
PF00722-PF00354 1.71279986 1.657536934 1.76806278 0.0000000
PF00840-PF00354 -1.92614077 -1.981403695 -1.87087785 0.0000000
PF01670-PF00354 0.49669090 0.441427981 0.55195382 0.0000000
PF02210-PF00354 1.04319066 0.987927742 1.09845359 0.0000000
PF03388-PF00354 0.59824020 0.542977277 0.65350312 0.0000000
PF03935-PF00354 1.44935915 1.394096225 1.50462207 0.0000000
PF06439-PF00354 1.57116011 1.515897183 1.62642303 0.0000000
PF10287-PF00354 -1.71542728 -1.770690201 -1.66016436 0.0000000
PF13385-PF00354 2.26277688 2.207513955 2.31803980 0.0000000
PF00722-PF00457 3.67718529 3.621922364 3.73244821 0.0000000
PF00840-PF00457 0.03824466 -0.017018265 0.09350758 0.6596422
PF01670-PF00457 2.46107633 2.405813411 2.51633925 0.0000000
PF02210-PF00457 3.00757609 2.952313172 3.06283902 0.0000000
PF03388-PF00457 2.56262563 2.507362707 2.61788855 0.0000000
PF03935-PF00457 3.41374458 3.358481655 3.46900750 0.0000000
PF06439-PF00457 3.53554554 3.480282613 3.59080846 0.0000000
PF10287-PF00457 0.24895815 0.193695229 0.30422107 0.0000000
PF13385-PF00457 4.22716231 4.171899385 4.28242523 0.0000000
PF00840-PF00722 -3.63894063 -3.694203551 -3.58367771 0.0000000
PF01670-PF00722 -1.21610895 -1.271371875 -1.16084603 0.0000000
PF02210-PF00722 -0.66960919 -0.724872114 -0.61434627 0.0000000
PF03388-PF00722 -1.11455966 -1.169822579 -1.05929673 0.0000000
PF03935-PF00722 -0.26344071 -0.318703631 -0.20817779 0.0000000
PF06439-PF00722 -0.14163975 -0.196902673 -0.08637683 0.0000000
PF10287-PF00722 -3.42822713 -3.483490057 -3.37296421 0.0000000
PF13385-PF00722 0.54997702 0.494714099 0.60523994 0.0000000
PF01670-PF00840 2.42283168 2.367568754 2.47809460 0.0000000
PF02210-PF00840 2.96933144 2.914068515 3.02459436 0.0000000
PF03388-PF00840 2.52438097 2.469118050 2.57964389 0.0000000
PF03935-PF00840 3.37549992 3.320236997 3.43076284 0.0000000
PF06439-PF00840 3.49730088 3.442037956 3.55256380 0.0000000
PF10287-PF00840 0.21071349 0.155450572 0.26597642 0.0000000
PF13385-PF00840 4.18891765 4.133654727 4.24418057 0.0000000

PF02210-PF01670	0.54649976	0.491236839	0.60176268	0.0000000
PF03388-PF01670	0.10154930	0.046286374	0.15681222	0.0000009
PF03935-PF01670	0.95266824	0.897405322	1.00793117	0.0000000
PF06439-PF01670	1.07446920	1.019206280	1.12973212	0.0000000
PF10287-PF01670	-2.21211818	-2.267381104	-2.15685526	0.0000000
PF13385-PF01670	1.76608597	1.710823052	1.82134890	0.0000000
PF03388-PF02210	-0.44495047	-0.500213387	-0.38968754	0.0000000
PF03935-PF02210	0.40616848	0.350905561	0.46143140	0.0000000
PF06439-PF02210	0.52796944	0.472706519	0.58323236	0.0000000
PF10287-PF02210	-2.75861794	-2.813880865	-2.70335502	0.0000000
PF13385-PF02210	1.21958621	1.164323290	1.27484913	0.0000000
PF03935-PF03388	0.85111895	0.795856026	0.90638187	0.0000000
PF06439-PF03388	0.97291991	0.917656984	1.02818283	0.0000000
PF10287-PF03388	-2.31366748	-2.368930400	-2.25840456	0.0000000
PF13385-PF03388	1.66453668	1.609273756	1.71979960	0.0000000
PF06439-PF03935	0.12180096	0.066538037	0.17706388	0.0000000
PF10287-PF03935	-3.16478643	-3.220049348	-3.10952350	0.0000000
PF13385-PF03935	0.81341773	0.758154808	0.86868065	0.0000000
PF10287-PF06439	-3.28658738	-3.341850306	-3.23132446	0.0000000
PF13385-PF06439	0.69161677	0.636353850	0.74687969	0.0000000
PF13385-PF10287	3.97820416	3.922941234	4.03346708	0.0000000

Pares iguais

	diff	lwr	upr	p adj
--	------	-----	-----	-------

PF01670-PF00337	0.04694846	-0.008314458	0.10221139	0.3082979
PF00840-PF00457	0.03824466	-0.017018265	0.09350758	0.6596422
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50388, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

	Comparison	Z	P.unadj	P.adj
1	PF00054 - PF00139	-57.077780	0.000000e+00	0.000000e+00
2	PF00054 - PF00337	-28.425682	9.738972e-178	1.384760e-177
3	PF00139 - PF00337	28.652099	1.509388e-180	2.180227e-180
4	PF00054 - PF00354	-10.977974	4.877402e-28	5.412727e-28
5	PF00139 - PF00354	46.099807	0.000000e+00	0.000000e+00
6	PF00337 - PF00354	17.447708	3.583260e-68	4.234761e-68
7	PF00054 - PF00457	32.535662	3.340194e-232	5.240650e-232
8	PF00139 - PF00457	89.613442	0.000000e+00	0.000000e+00
9	PF00337 - PF00457	60.961344	0.000000e+00	0.000000e+00

10 PF00354 - PF00457 43.513636 0.000000e+00 0.000000e+00
11 PF00054 - PF00722 -81.239248 0.000000e+00 0.000000e+00
12 PF00139 - PF00722 -24.161467 5.657417e-129 7.251056e-129
13 PF00337 - PF00722 -52.813566 0.000000e+00 0.000000e+00
14 PF00354 - PF00722 -70.261274 0.000000e+00 0.000000e+00
15 PF00457 - PF00722 -113.774910 0.000000e+00 0.000000e+00
16 PF00054 - PF00840 30.591435 1.591085e-205 2.373585e-205
17 PF00139 - PF00840 87.669216 0.000000e+00 0.000000e+00
18 PF00337 - PF00840 59.017117 0.000000e+00 0.000000e+00
19 PF00354 - PF00840 41.569409 0.000000e+00 0.000000e+00
20 PF00457 - PF00840 -1.944227 5.186813e-02 5.186813e-02
21 PF00722 - PF00840 111.830683 0.000000e+00 0.000000e+00
22 PF00054 - PF01670 -25.935877 2.624387e-148 3.461148e-148
23 PF00139 - PF01670 31.141904 6.528391e-213 1.006921e-212
24 PF00337 - PF01670 2.489805 1.278131e-02 1.306853e-02
25 PF00354 - PF01670 -14.957903 1.383168e-50 1.613696e-50
26 PF00457 - PF01670 -58.471538 0.000000e+00 0.000000e+00
27 PF00722 - PF01670 55.303371 0.000000e+00 0.000000e+00
28 PF00840 - PF01670 -56.527312 0.000000e+00 0.000000e+00
29 PF00054 - PF02210 -54.995232 0.000000e+00 0.000000e+00
30 PF00139 - PF02210 2.082548 3.729244e-02 3.770680e-02
31 PF00337 - PF02210 -26.569551 1.526854e-155 2.105208e-155
32 PF00354 - PF02210 -44.017259 0.000000e+00 0.000000e+00
33 PF00457 - PF02210 -87.530894 0.000000e+00 0.000000e+00
34 PF00722 - PF02210 26.244015 8.365239e-152 1.136174e-151
35 PF00840 - PF02210 -85.586668 0.000000e+00 0.000000e+00
36 PF01670 - PF02210 -29.059356 1.172159e-185 1.720427e-185
37 PF00054 - PF03388 -30.943072 3.148869e-210 4.775784e-210
38 PF00139 - PF03388 26.134709 1.470715e-150 1.968162e-150
39 PF00337 - PF03388 -2.517390 1.182279e-02 1.222584e-02
40 PF00354 - PF03388 -19.965098 1.108106e-88 1.326811e-88
41 PF00457 - PF03388 -63.478733 0.000000e+00 0.000000e+00
42 PF00722 - PF03388 50.296176 0.000000e+00 0.000000e+00
43 PF00840 - PF03388 -61.534507 0.000000e+00 0.000000e+00
44 PF01670 - PF03388 -5.007195 5.522892e-07 5.912743e-07
45 PF02210 - PF03388 24.052161 7.924323e-128 1.001546e-127
46 PF00054 - PF03935 -68.873132 0.000000e+00 0.000000e+00
47 PF00139 - PF03935 -11.795351 4.124755e-32 4.633984e-32
48 PF00337 - PF03935 -40.447450 0.000000e+00 0.000000e+00
49 PF00354 - PF03935 -57.895158 0.000000e+00 0.000000e+00
50 PF00457 - PF03935 -101.408794 0.000000e+00 0.000000e+00
51 PF00722 - PF03935 12.366116 3.986358e-35 4.534483e-35
52 PF00840 - PF03935 -99.464567 0.000000e+00 0.000000e+00
53 PF01670 - PF03935 -42.937255 0.000000e+00 0.000000e+00
54 PF02210 - PF03935 -13.877899 8.623428e-44 9.933316e-44
55 PF03388 - PF03935 -37.930060 0.000000e+00 0.000000e+00
56 PF00054 - PF06439 -77.297898 0.000000e+00 0.000000e+00
57 PF00139 - PF06439 -20.220117 6.512847e-91 7.902255e-91
58 PF00337 - PF06439 -48.872216 0.000000e+00 0.000000e+00
59 PF00354 - PF06439 -66.319924 0.000000e+00 0.000000e+00
60 PF00457 - PF06439 -109.833559 0.000000e+00 0.000000e+00
61 PF00722 - PF06439 3.941350 8.102417e-05 8.573488e-05
62 PF00840 - PF06439 -107.889333 0.000000e+00 0.000000e+00
63 PF01670 - PF06439 -51.362021 0.000000e+00 0.000000e+00

64	PF02210 - PF06439	-22.302665	3.481376e-110	4.339798e-110
65	PF03388 - PF06439	-46.354826	0.000000e+00	0.000000e+00
66	PF03935 - PF06439	-8.424766	3.614729e-17	3.963137e-17
67	PF00054 - PF10287	27.517558	1.082414e-166	1.515380e-166
68	PF00139 - PF10287	84.595339	0.000000e+00	0.000000e+00
69	PF00337 - PF10287	55.943240	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	38.495532	0.000000e+00	0.000000e+00
71	PF00457 - PF10287	-5.018104	5.218401e-07	5.653267e-07
72	PF00722 - PF10287	108.756806	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-3.073877	2.112965e-03	2.210113e-03
74	PF01670 - PF10287	53.453435	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.512790	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	58.460630	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	96.390690	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.815456	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-103.125094	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-46.047313	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-74.699412	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-92.147120	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-135.660756	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-21.885846	3.543745e-106	4.357848e-106
85	PF00840 - PF13385	-133.716529	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-77.189217	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-48.129861	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-72.182022	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-34.251962	4.077694e-257	6.510003e-257
90	PF06439 - PF13385	-25.827196	4.389695e-147	5.706603e-147
91	PF10287 - PF13385	-130.642652	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
20	PF00457 - PF00840	-1.944227	0.051868129	0.051868129
24	PF00337 - PF01670	2.489805	0.012781312	0.013068532
30	PF00139 - PF02210	2.082548	0.037292439	0.037706800
39	PF00337 - PF03388	-2.517390	0.011822790	0.012225840
73	PF00840 - PF10287	-3.073877	0.002112965	0.002210113

=====

=====

ANOVA da entropia entre familias controlando **s = 0.7**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

Df	F value	Pr(>F)
----	---------	--------

group 13 762.73 < 2.2e-16 ***

69286

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 15993, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 61287 4714 11779 <2e-16 ***

Residuals 69286 27731 0

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	1.08828263	1.048611e+00	1.12795447	0.0000000
PF00337-PF00054	0.58213186	5.424600e-01	0.62180370	0.0000000
PF00354-PF00054	0.26549312	2.258213e-01	0.30516496	0.0000000
PF00457-PF00054	-1.16663969	-1.206312e+00	-1.12696785	0.0000000
PF00722-PF00054	1.45286357	1.413192e+00	1.49253541	0.0000000
PF00840-PF00054	-1.10036785	-1.140040e+00	-1.06069601	0.0000000
PF01670-PF00054	0.62187928	5.822074e-01	0.66155112	0.0000000
PF02210-PF00054	0.99622504	9.565532e-01	1.03589688	0.0000000
PF03388-PF00054	0.67434915	6.346773e-01	0.71402100	0.0000000
PF03935-PF00054	1.26407608	1.224404e+00	1.30374793	0.0000000

PF06439-PF00054 1.35735031 1.317678e+00 1.39702216 0.0000000
PF10287-PF00054 -0.94893055 -9.886024e-01 -0.90925870 0.0000000
PF13385-PF00054 1.82941091 1.789739e+00 1.86908276 0.0000000
PF00337-PF00139 -0.50615077 -5.458226e-01 -0.46647892 0.0000000
PF00354-PF00139 -0.82278951 -8.624613e-01 -0.78311766 0.0000000
PF00457-PF00139 -2.25492231 -2.294594e+00 -2.21525047 0.0000000
PF00722-PF00139 0.36458094 3.249091e-01 0.40425278 0.0000000
PF00840-PF00139 -2.18865048 -2.228322e+00 -2.14897864 0.0000000
PF01670-PF00139 -0.46640335 -5.060752e-01 -0.42673151 0.0000000
PF02210-PF00139 -0.09205759 -1.317294e-01 -0.05238574 0.0000000
PF03388-PF00139 -0.41393347 -4.536053e-01 -0.37426163 0.0000000
PF03935-PF00139 0.17579346 1.361216e-01 0.21546530 0.0000000
PF06439-PF00139 0.26906769 2.293958e-01 0.30873953 0.0000000
PF10287-PF00139 -2.03721317 -2.076885e+00 -1.99754133 0.0000000
PF13385-PF00139 0.74112829 7.014564e-01 0.78080013 0.0000000
PF00354-PF00337 -0.31663874 -3.563106e-01 -0.27696690 0.0000000
PF00457-PF00337 -1.74877155 -1.788443e+00 -1.70909970 0.0000000
PF00722-PF00337 0.87073171 8.310599e-01 0.91040355 0.0000000
PF00840-PF00337 -1.68249971 -1.722172e+00 -1.64282787 0.0000000
PF01670-PF00337 0.03974742 7.557517e-05 0.07941926 0.0983562
PF02210-PF00337 0.41409318 3.744213e-01 0.45376502 0.0000000
PF03388-PF00337 0.09221729 5.254545e-02 0.13188914 0.0000000
PF03935-PF00337 0.68194423 6.422724e-01 0.72161607 0.0000000
PF06439-PF00337 0.77521845 7.355466e-01 0.81489030 0.0000000
PF10287-PF00337 -1.53106241 -1.570734e+00 -1.49139056 0.0000000
PF13385-PF00337 1.24727905 1.207607e+00 1.28695090 0.0000000
PF00457-PF00354 -1.43213281 -1.471805e+00 -1.39246097 0.0000000
PF00722-PF00354 1.18737045 1.147699e+00 1.22704229 0.0000000
PF00840-PF00354 -1.36586097 -1.405533e+00 -1.32618913 0.0000000
PF01670-PF00354 0.35638616 3.167143e-01 0.39605800 0.0000000
PF02210-PF00354 0.73073192 6.910601e-01 0.77040376 0.0000000
PF03388-PF00354 0.40885603 3.691842e-01 0.44852787 0.0000000
PF03935-PF00354 0.99858296 9.589111e-01 1.03825481 0.0000000
PF06439-PF00354 1.09185719 1.052185e+00 1.13152904 0.0000000
PF10287-PF00354 -1.21442367 -1.254096e+00 -1.17475182 0.0000000
PF13385-PF00354 1.56391779 1.524246e+00 1.60358964 0.0000000
PF00722-PF00457 2.61950325 2.579831e+00 2.65917510 0.0000000
PF00840-PF00457 0.06627183 2.659999e-02 0.10594368 0.0000167
PF01670-PF00457 1.78851897 1.748847e+00 1.82819081 0.0000000
PF02210-PF00457 2.16286473 2.123193e+00 2.20253657 0.0000000
PF03388-PF00457 1.84098884 1.801317e+00 1.88066068 0.0000000
PF03935-PF00457 2.43071577 2.391044e+00 2.47038762 0.0000000
PF06439-PF00457 2.52399000 2.484318e+00 2.56366184 0.0000000
PF10287-PF00457 0.21770914 1.780373e-01 0.25738098 0.0000000
PF13385-PF00457 2.99605060 2.956379e+00 3.03572245 0.0000000
PF00840-PF00722 -2.55323142 -2.592903e+00 -2.51355958 0.0000000
PF01670-PF00722 -0.83098429 -8.706561e-01 -0.79131245 0.0000000
PF02210-PF00722 -0.45663853 -4.963104e-01 -0.41696669 0.0000000
PF03388-PF00722 -0.77851441 -8.181863e-01 -0.73884257 0.0000000
PF03935-PF00722 -0.18878748 -2.284593e-01 -0.14911564 0.0000000
PF06439-PF00722 -0.09551325 -1.351851e-01 -0.05584141 0.0000000
PF10287-PF00722 -2.40179411 -2.441466e+00 -2.36212227 0.0000000
PF13385-PF00722 0.37654735 3.368755e-01 0.41621919 0.0000000
PF01670-PF00840 1.72224713 1.682575e+00 1.76191897 0.0000000

PF02210-PF00840	2.09659289	2.056921e+00	2.13626473	0.0000000
PF03388-PF00840	1.77471701	1.735045e+00	1.81438885	0.0000000
PF03935-PF00840	2.36444394	2.324772e+00	2.40411578	0.0000000
PF06439-PF00840	2.45771817	2.418046e+00	2.49739001	0.0000000
PF10287-PF00840	0.15143731	1.117655e-01	0.19110915	0.0000000
PF13385-PF00840	2.92977877	2.890107e+00	2.96945061	0.0000000
PF02210-PF01670	0.37434576	3.346739e-01	0.41401760	0.0000000
PF03388-PF01670	0.05246987	1.279803e-02	0.09214172	0.0029515
PF03935-PF01670	0.64219681	6.025250e-01	0.68186865	0.0000000
PF06439-PF01670	0.73547104	6.957992e-01	0.77514288	0.0000000
PF10287-PF01670	-1.57080982	-1.610482e+00	-1.53113798	0.0000000
PF13385-PF01670	1.20753164	1.167860e+00	1.24720348	0.0000000
PF03388-PF02210	-0.32187589	-3.615477e-01	-0.28220404	0.0000000
PF03935-PF02210	0.26785105	2.281792e-01	0.30752289	0.0000000
PF06439-PF02210	0.36112528	3.214534e-01	0.40079712	0.0000000
PF10287-PF02210	-1.94515558	-1.984827e+00	-1.90548374	0.0000000
PF13385-PF02210	0.83318588	7.935140e-01	0.87285772	0.0000000
PF03935-PF03388	0.58972693	5.500551e-01	0.62939878	0.0000000
PF06439-PF03388	0.68300116	6.433293e-01	0.72267300	0.0000000
PF10287-PF03388	-1.62327970	-1.662952e+00	-1.58360786	0.0000000
PF13385-PF03388	1.15506176	1.115390e+00	1.19473360	0.0000000
PF06439-PF03935	0.09327423	5.360239e-02	0.13294607	0.0000000
PF10287-PF03935	-2.21300663	-2.252678e+00	-2.17333479	0.0000000
PF13385-PF03935	0.56533483	5.256630e-01	0.60500667	0.0000000
PF10287-PF06439	-2.30628086	-2.345953e+00	-2.26660902	0.0000000
PF13385-PF06439	0.47206060	4.323888e-01	0.51173244	0.0000000
PF13385-PF10287	2.77834146	2.738670e+00	2.81801330	0.0000000

Pares iguais

	diff	lwr	upr	p adj
--	------	-----	-----	-------

PF01670-PF00337	0.03974742	7.557517e-05	0.07941926	0.098356243
PF03388-PF01670	0.05246987	1.279803e-02	0.09214172	0.002951456
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 50203, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

	Comparison	Z	P.unadj	P.adj
1	PF00054 - PF00139	-56.391775	0.000000e+00	0.000000e+00

2 PF00054 - PF00337 -28.188382 8.117369e-175 1.154188e-174
 3 PF00139 - PF00337 28.203393 5.313256e-175 7.674703e-175
 4 PF00054 - PF00354 -10.546699 5.261342e-26 5.838806e-26
 5 PF00139 - PF00354 45.845076 0.000000e+00 0.000000e+00
 6 PF00337 - PF00354 17.641683 1.178966e-69 1.393323e-69
 7 PF00054 - PF00457 33.319961 1.984658e-243 3.113860e-243
 8 PF00139 - PF00457 89.711736 0.000000e+00 0.000000e+00
 9 PF00337 - PF00457 61.508342 0.000000e+00 0.000000e+00
 10 PF00354 - PF00457 43.866659 0.000000e+00 0.000000e+00
 11 PF00054 - PF00722 -80.602765 0.000000e+00 0.000000e+00
 12 PF00139 - PF00722 -24.210990 1.704318e-129 2.154068e-129
 13 PF00337 - PF00722 -52.414383 0.000000e+00 0.000000e+00
 14 PF00354 - PF00722 -70.056066 0.000000e+00 0.000000e+00
 15 PF00457 - PF00722 -113.922725 0.000000e+00 0.000000e+00
 16 PF00054 - PF00840 30.649070 2.719132e-206 4.193916e-206
 17 PF00139 - PF00840 87.040846 0.000000e+00 0.000000e+00
 18 PF00337 - PF00840 58.837452 0.000000e+00 0.000000e+00
 19 PF00354 - PF00840 41.195769 0.000000e+00 0.000000e+00
 20 PF00457 - PF00840 -2.670890 7.565043e-03 7.822942e-03
 21 PF00722 - PF00840 111.251835 0.000000e+00 0.000000e+00
 22 PF00054 - PF01670 -25.767638 2.045041e-146 2.736746e-146
 23 PF00139 - PF01670 30.624137 5.841654e-206 8.859843e-206
 24 PF00337 - PF01670 2.420743 1.548881e-02 1.583688e-02
 25 PF00354 - PF01670 -15.220940 2.568151e-52 2.996176e-52
 26 PF00457 - PF01670 -59.087599 0.000000e+00 0.000000e+00
 27 PF00722 - PF01670 54.835126 0.000000e+00 0.000000e+00
 28 PF00840 - PF01670 -56.416709 0.000000e+00 0.000000e+00
 29 PF00054 - PF02210 -54.924286 0.000000e+00 0.000000e+00
 30 PF00139 - PF02210 1.467489 1.422431e-01 1.422431e-01
 31 PF00337 - PF02210 -26.735904 1.801059e-157 2.483278e-157
 32 PF00354 - PF02210 -44.377587 0.000000e+00 0.000000e+00
 33 PF00457 - PF02210 -88.244247 0.000000e+00 0.000000e+00
 34 PF00722 - PF02210 25.678479 2.033471e-145 2.681824e-145
 35 PF00840 - PF02210 -85.573357 0.000000e+00 0.000000e+00
 36 PF01670 - PF02210 -29.156648 6.880663e-187 1.009904e-186
 37 PF00054 - PF03388 -29.914021 1.293192e-196 1.929189e-196
 38 PF00139 - PF03388 26.477754 1.748682e-154 2.375075e-154
 39 PF00337 - PF03388 -1.725640 8.441228e-02 8.535020e-02
 40 PF00354 - PF03388 -19.367323 1.456365e-83 1.743806e-83
 41 PF00457 - PF03388 -63.233982 0.000000e+00 0.000000e+00
 42 PF00722 - PF03388 50.688744 0.000000e+00 0.000000e+00
 43 PF00840 - PF03388 -60.563092 0.000000e+00 0.000000e+00
 44 PF01670 - PF03388 -4.146383 3.377690e-05 3.616115e-05
 45 PF02210 - PF03388 25.010265 4.727511e-138 6.059204e-138
 46 PF00054 - PF03935 -67.783967 0.000000e+00 0.000000e+00
 47 PF00139 - PF03935 -11.392191 4.573404e-30 5.138022e-30
 48 PF00337 - PF03935 -39.595585 0.000000e+00 0.000000e+00
 49 PF00354 - PF03935 -57.237268 0.000000e+00 0.000000e+00
 50 PF00457 - PF03935 -101.103927 0.000000e+00 0.000000e+00
 51 PF00722 - PF03935 12.818798 1.286792e-37 1.463726e-37
 52 PF00840 - PF03935 -98.433037 0.000000e+00 0.000000e+00
 53 PF01670 - PF03935 -42.016328 0.000000e+00 0.000000e+00
 54 PF02210 - PF03935 -12.859680 7.588972e-38 8.741727e-38
 55 PF03388 - PF03935 -37.869945 0.000000e+00 0.000000e+00

56 PF00054 - PF06439 -77.114862 0.000000e+00 0.000000e+00
 57 PF00139 - PF06439 -20.723087 2.144836e-95 2.602401e-95
 58 PF00337 - PF06439 -48.926480 0.000000e+00 0.000000e+00
 59 PF00354 - PF06439 -66.568163 0.000000e+00 0.000000e+00
 60 PF00457 - PF06439 -110.434822 0.000000e+00 0.000000e+00
 61 PF00722 - PF06439 3.487903 4.868247e-04 5.151285e-04
 62 PF00840 - PF06439 -107.763932 0.000000e+00 0.000000e+00
 63 PF01670 - PF06439 -51.347224 0.000000e+00 0.000000e+00
 64 PF02210 - PF06439 -22.190576 4.235348e-109 5.279681e-109
 65 PF03388 - PF06439 -47.200841 0.000000e+00 0.000000e+00
 66 PF03935 - PF06439 -9.330895 1.049802e-20 1.150988e-20
 67 PF00054 - PF10287 27.681911 1.152946e-168 1.614124e-168
 68 PF00139 - PF10287 84.073686 0.000000e+00 0.000000e+00
 69 PF00337 - PF10287 55.870293 0.000000e+00 0.000000e+00
 70 PF00354 - PF10287 38.228610 0.000000e+00 0.000000e+00
 71 PF00457 - PF10287 -5.638049 1.719873e-08 1.863196e-08
 72 PF00722 - PF10287 108.284676 0.000000e+00 0.000000e+00
 73 PF00840 - PF10287 -2.967159 3.005652e-03 3.143843e-03
 74 PF01670 - PF10287 53.449550 0.000000e+00 0.000000e+00
 75 PF02210 - PF10287 82.606197 0.000000e+00 0.000000e+00
 76 PF03388 - PF10287 57.595932 0.000000e+00 0.000000e+00
 77 PF03935 - PF10287 95.465878 0.000000e+00 0.000000e+00
 78 PF06439 - PF10287 104.796773 0.000000e+00 0.000000e+00
 79 PF00054 - PF13385 -102.780564 0.000000e+00 0.000000e+00
 80 PF00139 - PF13385 -46.388789 0.000000e+00 0.000000e+00
 81 PF00337 - PF13385 -74.592183 0.000000e+00 0.000000e+00
 82 PF00354 - PF13385 -92.233866 0.000000e+00 0.000000e+00
 83 PF00457 - PF13385 -136.100525 0.000000e+00 0.000000e+00
 84 PF00722 - PF13385 -22.177799 5.626451e-109 6.919014e-109
 85 PF00840 - PF13385 -133.429635 0.000000e+00 0.000000e+00
 86 PF01670 - PF13385 -77.012926 0.000000e+00 0.000000e+00
 87 PF02210 - PF13385 -47.856278 0.000000e+00 0.000000e+00
 88 PF03388 - PF13385 -72.866543 0.000000e+00 0.000000e+00
 89 PF03935 - PF13385 -34.996598 2.534569e-268 4.046418e-268
 90 PF06439 - PF13385 -25.665702 2.824253e-145 3.671529e-145
 91 PF10287 - PF13385 -130.462475 0.000000e+00 0.000000e+00
 Pares iguais

Comparison	Z	P.unadj	P.adj
20 PF00457 - PF00840	-2.670890	0.007565043	0.007822942
24 PF00337 - PF01670	2.420743	0.015488814	0.015836877
30 PF00139 - PF02210	1.467489	0.142243080	0.142243080
39 PF00337 - PF03388	-1.725640	0.084412281	0.085350195
73 PF00840 - PF10287	-2.967159	0.003005652	0.003143843

=====

ANOVA da entropia entre familias controlando **s = 0.8**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

```

Df F value Pr(>F)
group 13 835.94 < 2.2e-16 ***
69286
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 17752, df = 13, p-value < 2.2e-16

=====

ANOVA

```

Df Sum Sq Mean Sq F value Pr(>F)

```

```

pf      13 30752 2365.6 11264 <2e-16 ***

```

```

Residuals 69286 14551 0.2

```

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	0.75120148	0.722463814	0.77993915	0.0000000
PF00337-PF00054	0.40183638	0.373098711	0.43057405	0.0000000
PF00354-PF00054	0.17995820	0.151220527	0.20869587	0.0000000
PF00457-PF00054	-0.87186760	-0.900605271	-0.84312993	0.0000000
PF00722-PF00054	1.00673611	0.977998442	1.03547378	0.0000000
PF00840-PF00054	-0.79421814	-0.822955812	-0.76548047	0.0000000
PF01670-PF00054	0.43685226	0.408114587	0.46558993	0.0000000
PF02210-PF00054	0.69427272	0.665535054	0.72301039	0.0000000

PF03388-PF00054 0.46066431 0.431926639 0.48940198 0.0000000
PF03935-PF00054 0.87081200 0.842074331 0.89954967 0.0000000
PF06439-PF00054 0.94203458 0.913296905 0.97077225 0.0000000
PF10287-PF00054 -0.68457262 -0.713310294 -0.65583495 0.0000000
PF13385-PF00054 1.26531198 1.236574312 1.29404965 0.0000000
PF00337-PF00139 -0.34936510 -0.378102774 -0.32062743 0.0000000
PF00354-PF00139 -0.57124329 -0.599980957 -0.54250562 0.0000000
PF00457-PF00139 -1.62306909 -1.651806756 -1.59433142 0.0000000
PF00722-PF00139 0.25553463 0.226796957 0.28427230 0.0000000
PF00840-PF00139 -1.54541963 -1.574157297 -1.51668196 0.0000000
PF01670-PF00139 -0.31434923 -0.343086897 -0.28561156 0.0000000
PF02210-PF00139 -0.05692876 -0.085666431 -0.02819109 0.0000001
PF03388-PF00139 -0.29053717 -0.319274845 -0.26179950 0.0000000
PF03935-PF00139 0.11961052 0.090872846 0.14834819 0.0000000
PF06439-PF00139 0.19083309 0.162095421 0.21957076 0.0000000
PF10287-PF00139 -1.43577411 -1.464511779 -1.40703644 0.0000000
PF13385-PF00139 0.51411050 0.485372828 0.54284817 0.0000000
PF00354-PF00337 -0.22187818 -0.250615854 -0.19314051 0.0000000
PF00457-PF00337 -1.27370398 -1.302441652 -1.24496631 0.0000000
PF00722-PF00337 0.60489973 0.576162061 0.63363740 0.0000000
PF00840-PF00337 -1.19605452 -1.224792193 -1.16731685 0.0000000
PF01670-PF00337 0.03501588 0.006278207 0.06375355 0.0106689
PF02210-PF00337 0.29243634 0.263698673 0.32117401 0.0000000
PF03388-PF00337 0.05882793 0.030090259 0.08756560 0.0000000
PF03935-PF00337 0.46897562 0.440237950 0.49771329 0.0000000
PF06439-PF00337 0.54019819 0.511460524 0.56893586 0.0000000
PF10287-PF00337 -1.08640901 -1.115146675 -1.05767133 0.0000000
PF13385-PF00337 0.86347560 0.834737931 0.89221327 0.0000000
PF00457-PF00354 -1.05182580 -1.080563469 -1.02308813 0.0000000
PF00722-PF00354 0.82677791 0.798040244 0.85551558 0.0000000
PF00840-PF00354 -0.97417634 -1.002914010 -0.94543867 0.0000000
PF01670-PF00354 0.25689406 0.228156390 0.28563173 0.0000000
PF02210-PF00354 0.51431453 0.485576856 0.54305220 0.0000000
PF03388-PF00354 0.28070611 0.251968442 0.30944378 0.0000000
PF03935-PF00354 0.69085380 0.662116133 0.71959147 0.0000000
PF06439-PF00354 0.76207638 0.733338708 0.79081405 0.0000000
PF10287-PF00354 -0.86453082 -0.893268492 -0.83579315 0.0000000
PF13385-PF00354 1.08535378 1.056616115 1.11409146 0.0000000
PF00722-PF00457 1.87860371 1.849866042 1.90734138 0.0000000
PF00840-PF00457 0.07764946 0.048911789 0.10638713 0.0000000
PF01670-PF00457 1.30871986 1.279982188 1.33745753 0.0000000
PF02210-PF00457 1.56614032 1.537402654 1.59487799 0.0000000
PF03388-PF00457 1.33253191 1.303794240 1.36126958 0.0000000
PF03935-PF00457 1.74267960 1.713941932 1.77141727 0.0000000
PF06439-PF00457 1.81390218 1.785164506 1.84263985 0.0000000
PF10287-PF00457 0.18729498 0.158557306 0.21603265 0.0000000
PF13385-PF00457 2.13717958 2.108441913 2.16591725 0.0000000
PF00840-PF00722 -1.80095425 -1.829691924 -1.77221658 0.0000000
PF01670-PF00722 -0.56988385 -0.598621524 -0.54114618 0.0000000
PF02210-PF00722 -0.31246339 -0.341201058 -0.28372572 0.0000000
PF03388-PF00722 -0.54607180 -0.574809472 -0.51733413 0.0000000
PF03935-PF00722 -0.13592411 -0.164661781 -0.10718644 0.0000000
PF06439-PF00722 -0.06470154 -0.093439207 -0.03596387 0.0000000
PF10287-PF00722 -1.69130874 -1.720046406 -1.66257107 0.0000000

PF13385-PF00722	0.25857587	0.229838200	0.28731354	0.0000000
PF01670-PF00840	1.23107040	1.202332730	1.25980807	0.0000000
PF02210-PF00840	1.48849087	1.459753196	1.51722854	0.0000000
PF03388-PF00840	1.25488245	1.226144781	1.28362012	0.0000000
PF03935-PF00840	1.66503014	1.636292473	1.69376781	0.0000000
PF06439-PF00840	1.73625272	1.707515047	1.76499039	0.0000000
PF10287-PF00840	0.10964552	0.080907848	0.13838319	0.0000000
PF13385-PF00840	2.05953012	2.030792454	2.08826779	0.0000000
PF02210-PF01670	0.25742047	0.228682796	0.28615814	0.0000000
PF03388-PF01670	0.02381205	-0.004925618	0.05254972	0.3499103
PF03935-PF01670	0.43395974	0.405222073	0.46269741	0.0000000
PF06439-PF01670	0.50518232	0.476444647	0.53391999	0.0000000
PF10287-PF01670	-1.12142488	-1.150162552	-1.09268721	0.0000000
PF13385-PF01670	0.82845972	0.799722054	0.85719739	0.0000000
PF03388-PF02210	-0.23360841	-0.262346084	-0.20487074	0.0000000
PF03935-PF02210	0.17653928	0.147801607	0.20527695	0.0000000
PF06439-PF02210	0.24776185	0.219024181	0.27649952	0.0000000
PF10287-PF02210	-1.37884535	-1.407583018	-1.35010768	0.0000000
PF13385-PF02210	0.57103926	0.542301588	0.59977693	0.0000000
PF03935-PF03388	0.41014769	0.381410021	0.43888536	0.0000000
PF06439-PF03388	0.48137027	0.452632596	0.51010794	0.0000000
PF10287-PF03388	-1.14523693	-1.173974604	-1.11649926	0.0000000
PF13385-PF03388	0.80464767	0.775910003	0.83338534	0.0000000
PF06439-PF03935	0.07122257	0.042484904	0.09996024	0.0000000
PF10287-PF03935	-1.55538463	-1.584122296	-1.52664696	0.0000000
PF13385-PF03935	0.39449998	0.365762311	0.42323765	0.0000000
PF10287-PF06439	-1.62660720	-1.655344870	-1.59786953	0.0000000
PF13385-PF06439	0.32327741	0.294539737	0.35201508	0.0000000
PF13385-PF10287	1.94988461	1.921146936	1.97862228	0.0000000

Pares iguais

	diff	lwr	upr	p adj
--	------	-----	-----	-------

PF01670-PF00337	0.03501588	0.006278207	0.06375355	0.01066894
PF03388-PF01670	0.02381205	-0.004925618	0.05254972	0.34991026

=====

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 49991, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-55.479708	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-27.809188	3.358732e-170	4.775697e-170
3 PF00139 - PF00337	27.670520	1.580883e-168	2.213236e-168
4 PF00054 - PF00354	-10.141480	3.615798e-24	3.964309e-24
5 PF00139 - PF00354	45.338228	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	17.667708	7.435766e-70	8.787724e-70
7 PF00054 - PF00457	34.101123	7.098885e-255	1.113791e-254
8 PF00139 - PF00457	89.580831	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	61.910310	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	44.242602	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-79.939493	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-24.459785	3.959861e-132	5.004825e-132
13 PF00337 - PF00722	-52.130305	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-69.798013	0.000000e+00	0.000000e+00
15 PF00457 - PF00722	-114.040615	0.000000e+00	0.000000e+00
16 PF00054 - PF00840	30.714386	3.657542e-207	5.641293e-207
17 PF00139 - PF00840	86.194094	0.000000e+00	0.000000e+00
18 PF00337 - PF00840	58.523573	0.000000e+00	0.000000e+00
19 PF00354 - PF00840	40.855866	0.000000e+00	0.000000e+00
20 PF00457 - PF00840	-3.386737	7.072917e-04	7.572182e-04
21 PF00722 - PF00840	110.653878	0.000000e+00	0.000000e+00
22 PF00054 - PF01670	-25.614075	1.063293e-144	1.402314e-144
23 PF00139 - PF01670	29.865633	5.501699e-196	8.344243e-196
24 PF00337 - PF01670	2.195112	2.815556e-02	2.878827e-02
25 PF00354 - PF01670	-15.472596	5.312126e-54	6.197480e-54
26 PF00457 - PF01670	-59.715198	0.000000e+00	0.000000e+00
27 PF00722 - PF01670	54.325417	0.000000e+00	0.000000e+00
28 PF00840 - PF01670	-56.328461	0.000000e+00	0.000000e+00
29 PF00054 - PF02210	-54.867486	0.000000e+00	0.000000e+00
30 PF00139 - PF02210	0.612222	5.403909e-01	5.403909e-01
31 PF00337 - PF02210	-27.058299	3.050466e-161	4.205946e-161
32 PF00354 - PF02210	-44.726006	0.000000e+00	0.000000e+00
33 PF00457 - PF02210	-88.968609	0.000000e+00	0.000000e+00
34 PF00722 - PF02210	25.072007	1.004864e-138	1.287925e-138
35 PF00840 - PF02210	-85.581872	0.000000e+00	0.000000e+00
36 PF01670 - PF02210	-29.253411	4.063438e-188	6.061851e-188
37 PF00054 - PF03388	-28.911683	8.514192e-184	1.249664e-183
38 PF00139 - PF03388	26.568025	1.590087e-155	2.159670e-155
39 PF00337 - PF03388	-1.102495	2.702465e-01	2.732493e-01
40 PF00354 - PF03388	-18.770203	1.323796e-78	1.585072e-78
41 PF00457 - PF03388	-63.012805	0.000000e+00	0.000000e+00
42 PF00722 - PF03388	51.027810	0.000000e+00	0.000000e+00
43 PF00840 - PF03388	-59.626068	0.000000e+00	0.000000e+00
44 PF01670 - PF03388	-3.297607	9.751249e-04	1.031818e-03
45 PF02210 - PF03388	25.955804	1.563708e-148	2.092609e-148
46 PF00054 - PF03935	-66.677812	0.000000e+00	0.000000e+00
47 PF00139 - PF03935	-11.198104	4.165538e-29	4.679801e-29
48 PF00337 - PF03935	-38.868624	0.000000e+00	0.000000e+00
49 PF00354 - PF03935	-56.536332	0.000000e+00	0.000000e+00
50 PF00457 - PF03935	-100.778934	0.000000e+00	0.000000e+00
51 PF00722 - PF03935	13.261681	3.861421e-40	4.447966e-40
52 PF00840 - PF03935	-97.392198	0.000000e+00	0.000000e+00

53	PF01670 - PF03935	-41.063736	0.000000e+00	0.000000e+00
54	PF02210 - PF03935	-11.810326	3.452209e-32	3.926888e-32
55	PF03388 - PF03935	-37.766129	0.000000e+00	0.000000e+00
56	PF00054 - PF06439	-76.922208	0.000000e+00	0.000000e+00
57	PF00139 - PF06439	-21.442500	5.365319e-102	6.509921e-102
58	PF00337 - PF06439	-49.113020	0.000000e+00	0.000000e+00
59	PF00354 - PF06439	-66.780728	0.000000e+00	0.000000e+00
60	PF00457 - PF06439	-111.023330	0.000000e+00	0.000000e+00
61	PF00722 - PF06439	3.017285	2.550497e-03	2.667761e-03
62	PF00840 - PF06439	-107.636593	0.000000e+00	0.000000e+00
63	PF01670 - PF06439	-51.308132	0.000000e+00	0.000000e+00
64	PF02210 - PF06439	-22.054722	8.605964e-108	1.058301e-107
65	PF03388 - PF06439	-48.010525	0.000000e+00	0.000000e+00
66	PF03935 - PF06439	-10.244396	1.254060e-24	1.391701e-24
67	PF00054 - PF10287	27.850656	1.057600e-170	1.527644e-170
68	PF00139 - PF10287	83.330364	0.000000e+00	0.000000e+00
69	PF00337 - PF10287	55.659844	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	37.992136	0.000000e+00	0.000000e+00
71	PF00457 - PF10287	-6.250466	4.092293e-10	4.433317e-10
72	PF00722 - PF10287	107.790149	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-2.863729	4.186857e-03	4.329591e-03
74	PF01670 - PF10287	53.464732	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.718142	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	56.762339	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	94.528468	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.772864	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-102.417046	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-46.937338	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-74.607859	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-92.275566	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-136.518169	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-22.477553	6.882962e-112	8.580131e-112
85	PF00840 - PF13385	-133.131432	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-76.802971	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-47.549560	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-73.505364	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-35.739234	9.723714e-280	1.552382e-279
90	PF06439 - PF13385	-25.494839	2.248964e-143	2.923653e-143
91	PF10287 - PF13385	-130.267703	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
24	PF00337 - PF01670	2.195112	0.0281555596	0.028788269
30	PF00139 - PF02210	0.612222	0.5403909134	0.540390913
39	PF00337 - PF03388	-1.102495	0.2702465387	0.273249278
44	PF01670 - PF03388	-3.297607	0.0009751249	0.001031818
61	PF00722 - PF06439	3.017285	0.0025504968	0.002667761
73	PF00840 - PF10287	-2.863729	0.0041868570	0.004329591
=====				

```

=====
ANOVA da entropia entre familias controlando s = 0.9
Teste de homogeneidade de variância: Levene Teste
Levene's Test for Homogeneity of Variance (center = mean)

      Df F value  Pr(>F)
group  13  910.51 < 2.2e-16 ***
      69286

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Teste de homogeneidade de variância: Barlett Teste

      Bartlett test of homogeneity of variances
data:  datb.x$entropia and datb.x$pf
Bartlett's K-squared = 19711, df = 13, p-value < 2.2e-16

```

```

=====
ANOVA

Df Sum Sq Mean Sq F value Pr(>F)
pf      13  15610  1200.8  10662 <2e-16 ***
Residuals 69286  7803   0.1

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

=====
Teste Tukey HSD

Tukey multiple comparisons of means

      90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

$pf

```


	diff	lwr	upr	p adj
PF00139-PF00054	0.514411705	0.49336748	0.535455932	0.0000000
PF00337-PF00054	0.273091733	0.25204751	0.294135960	0.0000000
PF00354-PF00054	0.122442200	0.10139797	0.143486427	0.0000000
PF00457-PF00054	-0.655222054	-0.67626628	-0.634177827	0.0000000
PF00722-PF00054	0.701502992	0.68045877	0.722547220	0.0000000
PF00840-PF00054	-0.576507249	-0.59755148	-0.555463022	0.0000000
PF01670-PF00054	0.308498017	0.28745379	0.329542244	0.0000000
PF02210-PF00054	0.486914520	0.46587029	0.507958747	0.0000000
PF03388-PF00054	0.316814953	0.29577073	0.337859181	0.0000000
PF03935-PF00054	0.603172112	0.58212789	0.624216340	0.0000000
PF06439-PF00054	0.657438494	0.63639427	0.678482722	0.0000000
PF10287-PF00054	-0.495845853	-0.51689008	-0.474801626	0.0000000
PF13385-PF00054	0.879610083	0.85856586	0.900654310	0.0000000
PF00337-PF00139	-0.241319972	-0.26236420	-0.220275745	0.0000000
PF00354-PF00139	-0.391969505	-0.41301373	-0.370925278	0.0000000
PF00457-PF00139	-1.169633759	-1.19067799	-1.148589532	0.0000000
PF00722-PF00139	0.187091288	0.16604706	0.208135515	0.0000000
PF00840-PF00139	-1.090918954	-1.11196318	-1.069874727	0.0000000
PF01670-PF00139	-0.205913688	-0.22695792	-0.184869461	0.0000000
PF02210-PF00139	-0.027497185	-0.04854141	-0.006452958	0.0036244
PF03388-PF00139	-0.197596751	-0.21864098	-0.176552524	0.0000000
PF03935-PF00139	0.088760408	0.06771618	0.109804635	0.0000000
PF06439-PF00139	0.143026790	0.12198256	0.164071017	0.0000000
PF10287-PF00139	-1.010257558	-1.03130179	-0.989213331	0.0000000
PF13385-PF00139	0.365198378	0.34415415	0.386242606	0.0000000
PF00354-PF00337	-0.150649533	-0.17169376	-0.129605306	0.0000000
PF00457-PF00337	-0.928313787	-0.94935801	-0.907269560	0.0000000
PF00722-PF00337	0.428411260	0.40736703	0.449455487	0.0000000
PF00840-PF00337	-0.849598982	-0.87064321	-0.828554755	0.0000000
PF01670-PF00337	0.035406284	0.01436206	0.056450511	0.0000137
PF02210-PF00337	0.213822787	0.19277856	0.234867014	0.0000000
PF03388-PF00337	0.043723220	0.02267899	0.064767448	0.0000000
PF03935-PF00337	0.330080379	0.30903615	0.351124607	0.0000000
PF06439-PF00337	0.384346761	0.36330253	0.405390989	0.0000000
PF10287-PF00337	-0.768937586	-0.78998181	-0.747893359	0.0000000
PF13385-PF00337	0.606518350	0.58547412	0.627562577	0.0000000
PF00457-PF00354	-0.777664254	-0.79870848	-0.756620027	0.0000000
PF00722-PF00354	0.579060792	0.55801657	0.600105020	0.0000000
PF00840-PF00354	-0.698949449	-0.71999368	-0.677905222	0.0000000
PF01670-PF00354	0.186055817	0.16501159	0.207100044	0.0000000
PF02210-PF00354	0.364472320	0.34342809	0.385516547	0.0000000
PF03388-PF00354	0.194372753	0.17332853	0.215416981	0.0000000
PF03935-PF00354	0.480729912	0.45968569	0.501774140	0.0000000
PF06439-PF00354	0.534996294	0.51395207	0.556040522	0.0000000
PF10287-PF00354	-0.618288053	-0.63933228	-0.597243826	0.0000000
PF13385-PF00354	0.757167883	0.73612366	0.778212110	0.0000000
PF00722-PF00457	1.356725046	1.33568082	1.377769274	0.0000000
PF00840-PF00457	0.078714805	0.05767058	0.099759032	0.0000000
PF01670-PF00457	0.963720071	0.94267584	0.984764298	0.0000000
PF02210-PF00457	1.142136574	1.12109235	1.163180801	0.0000000
PF03388-PF00457	0.972037007	0.95099278	0.993081235	0.0000000
PF03935-PF00457	1.258394166	1.23734994	1.279438394	0.0000000

PF06439-PF00457	1.312660548	1.29161632	1.333704776	0.0000000
PF10287-PF00457	0.159376201	0.13833197	0.180420428	0.0000000
PF13385-PF00457	1.534832137	1.51378791	1.555876364	0.0000000
PF00840-PF00722	-1.278010242	-1.29905447	-1.256966014	0.0000000
PF01670-PF00722	-0.393004976	-0.41404920	-0.371960749	0.0000000
PF02210-PF00722	-0.214588472	-0.23563270	-0.193544245	0.0000000
PF03388-PF00722	-0.384688039	-0.40573227	-0.363643812	0.0000000
PF03935-PF00722	-0.098330880	-0.11937511	-0.077286653	0.0000000
PF06439-PF00722	-0.044064498	-0.06510873	-0.023020271	0.0000000
PF10287-PF00722	-1.197348846	-1.21839307	-1.176304618	0.0000000
PF13385-PF00722	0.178107091	0.15706286	0.199151318	0.0000000
PF01670-PF00840	0.885005266	0.86396104	0.906049493	0.0000000
PF02210-PF00840	1.063421769	1.04237754	1.084465996	0.0000000
PF03388-PF00840	0.893322203	0.87227798	0.914366430	0.0000000
PF03935-PF00840	1.179679362	1.15863513	1.200723589	0.0000000
PF06439-PF00840	1.233945744	1.21290152	1.254989971	0.0000000
PF10287-PF00840	0.080661396	0.05961717	0.101705623	0.0000000
PF13385-PF00840	1.456117332	1.43507311	1.477161560	0.0000000
PF02210-PF01670	0.178416503	0.15737228	0.199460731	0.0000000
PF03388-PF01670	0.008316937	-0.01272729	0.029361164	0.9944365
PF03935-PF01670	0.294674096	0.27362987	0.315718323	0.0000000
PF06439-PF01670	0.348940478	0.32789625	0.369984705	0.0000000
PF10287-PF01670	-0.804343870	-0.82538810	-0.783299643	0.0000000
PF13385-PF01670	0.571112067	0.55006784	0.592156294	0.0000000
PF03388-PF02210	-0.170099567	-0.19114379	-0.149055339	0.0000000
PF03935-PF02210	0.116257592	0.09521337	0.137301820	0.0000000
PF06439-PF02210	0.170523974	0.14947975	0.191568202	0.0000000
PF10287-PF02210	-0.982760373	-1.00380460	-0.961716146	0.0000000
PF13385-PF02210	0.392695563	0.37165134	0.413739790	0.0000000
PF03935-PF03388	0.286357159	0.26531293	0.307401386	0.0000000
PF06439-PF03388	0.340623541	0.31957931	0.361667768	0.0000000
PF10287-PF03388	-0.812660807	-0.83370503	-0.791616579	0.0000000
PF13385-PF03388	0.562795130	0.54175090	0.583839357	0.0000000
PF06439-PF03935	0.054266382	0.03322215	0.075310609	0.0000000
PF10287-PF03935	-1.099017966	-1.12006219	-1.077973738	0.0000000
PF13385-PF03935	0.276437971	0.25539374	0.297482198	0.0000000
PF10287-PF06439	-1.153284347	-1.17432857	-1.132240120	0.0000000
PF13385-PF06439	0.222171589	0.20112736	0.243215816	0.0000000
PF13385-PF10287	1.375455936	1.35441171	1.396500164	0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF02210-PF00139	-0.027497185	-0.04854141	-0.006452958	0.003624395
PF03388-PF01670	0.008316937	-0.01272729	0.029361164	0.994436535
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 49697, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-53.684555	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-26.863484	5.869113e-159	8.216758e-159
3 PF00139 - PF00337	26.821071	1.835209e-158	2.530364e-158
4 PF00054 - PF00354	-9.818286	9.392680e-23	1.029800e-22
5 PF00139 - PF00354	43.866269	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	17.045198	3.794578e-65	4.484501e-65
7 PF00054 - PF00457	34.920830	3.590662e-267	5.633625e-267
8 PF00139 - PF00457	88.605385	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	61.784314	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	44.739116	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-79.229792	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-25.545237	6.202332e-144	8.179887e-144
13 PF00337 - PF00722	-52.366308	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-69.411505	0.000000e+00	0.000000e+00
15 PF00457 - PF00722	-114.150622	0.000000e+00	0.000000e+00
16 PF00054 - PF00840	30.850794	5.465953e-209	8.430538e-209
17 PF00139 - PF00840	84.535349	0.000000e+00	0.000000e+00
18 PF00337 - PF00840	57.714278	0.000000e+00	0.000000e+00
19 PF00354 - PF00840	40.669080	0.000000e+00	0.000000e+00
20 PF00457 - PF00840	-4.070036	4.700583e-05	5.032389e-05
21 PF00722 - PF00840	110.080586	0.000000e+00	0.000000e+00
22 PF00054 - PF01670	-25.539098	7.256900e-144	9.433970e-144
23 PF00139 - PF01670	28.145457	2.723745e-174	4.063292e-174
24 PF00337 - PF01670	1.324386	1.853750e-01	1.895407e-01
25 PF00354 - PF01670	-15.720812	1.089184e-55	1.270715e-55
26 PF00457 - PF01670	-60.459928	0.000000e+00	0.000000e+00
27 PF00722 - PF01670	53.690693	0.000000e+00	0.000000e+00
28 PF00840 - PF01670	-56.389892	0.000000e+00	0.000000e+00
29 PF00054 - PF02210	-54.855700	0.000000e+00	0.000000e+00
30 PF00139 - PF02210	-1.171145	2.415406e-01	2.415406e-01
31 PF00337 - PF02210	-27.992216	2.020975e-172	2.873573e-172
32 PF00354 - PF02210	-45.037413	0.000000e+00	0.000000e+00
33 PF00457 - PF02210	-89.776530	0.000000e+00	0.000000e+00
34 PF00722 - PF02210	24.374092	3.220310e-131	4.070114e-131
35 PF00840 - PF02210	-85.706494	0.000000e+00	0.000000e+00
36 PF01670 - PF02210	-29.316601	6.372060e-189	9.664291e-189
37 PF00054 - PF03388	-28.038549	5.509415e-173	7.958044e-173
38 PF00139 - PF03388	25.646006	4.684893e-145	6.269489e-145
39 PF00337 - PF03388	-1.175065	2.399686e-01	2.426349e-01
40 PF00354 - PF03388	-18.220263	3.564367e-74	4.267860e-74
41 PF00457 - PF03388	-62.959379	0.000000e+00	0.000000e+00
42 PF00722 - PF03388	51.191242	0.000000e+00	0.000000e+00
43 PF00840 - PF03388	-58.889343	0.000000e+00	0.000000e+00
44 PF01670 - PF03388	-2.499451	1.243859e-02	1.286263e-02
45 PF02210 - PF03388	26.817150	2.038984e-158	2.769366e-158
46 PF00054 - PF03935	-65.570556	0.000000e+00	0.000000e+00

47	PF00139 - PF03935	-11.886001	1.399484e-32	1.591913e-32
48	PF00337 - PF03935	-38.707072	0.000000e+00	0.000000e+00
49	PF00354 - PF03935	-55.752270	0.000000e+00	0.000000e+00
50	PF00457 - PF03935	-100.491386	0.000000e+00	0.000000e+00
51	PF00722 - PF03935	13.659236	1.778404e-42	2.048541e-42
52	PF00840 - PF03935	-96.421350	0.000000e+00	0.000000e+00
53	PF01670 - PF03935	-40.031458	0.000000e+00	0.000000e+00
54	PF02210 - PF03935	-10.714856	8.669181e-27	9.620676e-27
55	PF03388 - PF03935	-37.532007	0.000000e+00	0.000000e+00
56	PF00054 - PF06439	-76.686523	0.000000e+00	0.000000e+00
57	PF00139 - PF06439	-23.001968	4.454605e-117	5.553000e-117
58	PF00337 - PF06439	-49.823039	0.000000e+00	0.000000e+00
59	PF00354 - PF06439	-66.868237	0.000000e+00	0.000000e+00
60	PF00457 - PF06439	-111.607353	0.000000e+00	0.000000e+00
61	PF00722 - PF06439	2.543269	1.098207e-02	1.148700e-02
62	PF00840 - PF06439	-107.537317	0.000000e+00	0.000000e+00
63	PF01670 - PF06439	-51.147425	0.000000e+00	0.000000e+00
64	PF02210 - PF06439	-21.830823	1.182712e-105	1.435024e-105
65	PF03388 - PF06439	-48.647974	0.000000e+00	0.000000e+00
66	PF03935 - PF06439	-11.115967	1.049078e-28	1.178594e-28
67	PF00054 - PF10287	28.057413	3.243662e-173	4.760859e-173
68	PF00139 - PF10287	81.741968	0.000000e+00	0.000000e+00
69	PF00337 - PF10287	54.920897	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	37.875699	0.000000e+00	0.000000e+00
71	PF00457 - PF10287	-6.863417	6.723244e-12	7.283514e-12
72	PF00722 - PF10287	107.287204	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-2.793381	5.216020e-03	5.519277e-03
74	PF01670 - PF10287	53.596511	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.913112	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	56.095962	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	93.627969	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.743936	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-101.952645	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-48.268090	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-75.089161	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-92.134359	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-136.873475	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-22.722853	2.663331e-114	3.275177e-114
85	PF00840 - PF13385	-132.803439	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-76.413547	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-47.096945	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-73.914096	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-36.382089	8.173110e-290	1.304830e-289
90	PF06439 - PF13385	-25.266122	7.532433e-141	9.654246e-141
91	PF10287 - PF13385	-130.010058	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
24	PF00337 - PF01670	1.324386	0.18537499	0.189540719
30	PF00139 - PF02210	-1.171145	0.24154063	0.241540628
39	PF00337 - PF03388	-1.175065	0.23996860	0.242634914

44 PF01670 - PF03388 -2.499451 0.01243859 0.012862634
61 PF00722 - PF06439 2.543269 0.01098207 0.011486996
73 PF00840 - PF10287 -2.793381 0.00521602 0.005519277
=====

=====

ANOVA da entropia entre familias controlando **s = 1.0**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	1011.6	< 2.2e-16 ***
		69286	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 22307, df = 13, p-value < 2.2e-16

=====

ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
pf	13	8056	619.7	10156	<2e-16 ***
Residuals	69286	4228	0.1		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD (ANOVA Clássico - paramétrico)

Supõe a normalidade

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	0.369173583	0.353683591	0.38466357	0.0000000
PF00337-PF00054	0.199581496	0.184091504	0.21507149	0.0000000
PF00354-PF00054	0.083646881	0.068156889	0.09913687	0.0000000
PF00457-PF00054	-0.498288973	-0.513778965	-0.48279898	0.0000000
PF00722-PF00054	0.488687287	0.473197295	0.50417728	0.0000000
PF00840-PF00054	-0.420878780	-0.436368772	-0.40538879	0.0000000
PF01670-PF00054	0.219055170	0.203565178	0.23454516	0.0000000
PF02210-PF00054	0.340738788	0.325248797	0.35622878	0.0000000
PF03388-PF00054	0.216481713	0.200991721	0.23197171	0.0000000
PF03935-PF00054	0.417206976	0.401716984	0.43269697	0.0000000
PF06439-PF00054	0.458487852	0.442997860	0.47397784	0.0000000
PF10287-PF00054	-0.363656469	-0.379146461	-0.34816648	0.0000000
PF13385-PF00054	0.611750477	0.596260485	0.62724047	0.0000000
PF00337-PF00139	-0.169592087	-0.185082079	-0.15410210	0.0000000
PF00354-PF00139	-0.285526702	-0.301016694	-0.27003671	0.0000000
PF00457-PF00139	-0.867462556	-0.882952548	-0.85197256	0.0000000
PF00722-PF00139	0.119513704	0.104023712	0.13500370	0.0000000
PF00840-PF00139	-0.790052363	-0.805542355	-0.77456237	0.0000000
PF01670-PF00139	-0.150118413	-0.165608405	-0.13462842	0.0000000
PF02210-PF00139	-0.028434795	-0.043924786	-0.01294480	0.0000009
PF03388-PF00139	-0.152691870	-0.168181862	-0.13720188	0.0000000
PF03935-PF00139	0.048033393	0.032543401	0.06352338	0.0000000
PF06439-PF00139	0.089314269	0.073824277	0.10480426	0.0000000
PF10287-PF00139	-0.732830052	-0.748320044	-0.71734006	0.0000000
PF13385-PF00139	0.242576894	0.227086902	0.25806689	0.0000000
PF00354-PF00337	-0.115934615	-0.131424607	-0.10044462	0.0000000
PF00457-PF00337	-0.697870469	-0.713360461	-0.68238048	0.0000000
PF00722-PF00337	0.289105791	0.273615799	0.30459578	0.0000000
PF00840-PF00337	-0.620460276	-0.635950268	-0.60497028	0.0000000
PF01670-PF00337	0.019473674	0.003983682	0.03496367	0.0067192
PF02210-PF00337	0.141157292	0.125667301	0.15664728	0.0000000
PF03388-PF00337	0.016900217	0.001410225	0.03239021	0.0426492
PF03935-PF00337	0.217625480	0.202135488	0.23311547	0.0000000
PF06439-PF00337	0.258906356	0.243416364	0.27439635	0.0000000
PF10287-PF00337	-0.563237966	-0.578727957	-0.54774797	0.0000000
PF13385-PF00337	0.412168981	0.396678989	0.42765897	0.0000000
PF00457-PF00354	-0.581935854	-0.597425845	-0.56644586	0.0000000
PF00722-PF00354	0.405040406	0.389550414	0.42053040	0.0000000
PF00840-PF00354	-0.504525661	-0.520015653	-0.48903567	0.0000000
PF01670-PF00354	0.135408289	0.119918297	0.15089828	0.0000000
PF02210-PF00354	0.257091908	0.241601916	0.27258190	0.0000000
PF03388-PF00354	0.132834833	0.117344841	0.14832482	0.0000000
PF03935-PF00354	0.333560095	0.318070103	0.34905009	0.0000000
PF06439-PF00354	0.374840971	0.359350979	0.39033096	0.0000000

PF10287-PF00354	-0.447303350	-0.462793342	-0.43181336	0.0000000
PF13385-PF00354	0.528103596	0.512613605	0.54359359	0.0000000
PF00722-PF00457	0.986976260	0.971486268	1.00246625	0.0000000
PF00840-PF00457	0.077410193	0.061920201	0.09290018	0.0000000
PF01670-PF00457	0.717344143	0.701854151	0.73283413	0.0000000
PF02210-PF00457	0.839027761	0.823537770	0.85451775	0.0000000
PF03388-PF00457	0.714770686	0.699280694	0.73026068	0.0000000
PF03935-PF00457	0.915495949	0.900005957	0.93098594	0.0000000
PF06439-PF00457	0.956776825	0.941286833	0.97226682	0.0000000
PF10287-PF00457	0.134632503	0.119142512	0.15012250	0.0000000
PF13385-PF00457	1.110039450	1.094549458	1.12552944	0.0000000
PF00840-PF00722	-0.909566067	-0.925056059	-0.89407608	0.0000000
PF01670-PF00722	-0.269632117	-0.285122109	-0.25414213	0.0000000
PF02210-PF00722	-0.147948498	-0.163438490	-0.13245851	0.0000000
PF03388-PF00722	-0.272205574	-0.287695565	-0.25671558	0.0000000
PF03935-PF00722	-0.071480311	-0.086970303	-0.05599032	0.0000000
PF06439-PF00722	-0.030199435	-0.045689427	-0.01470944	0.0000001
PF10287-PF00722	-0.852343756	-0.867833748	-0.83685376	0.0000000
PF13385-PF00722	0.123063190	0.107573198	0.13855318	0.0000000
PF01670-PF00840	0.639933950	0.624443958	0.65542394	0.0000000
PF02210-PF00840	0.761617568	0.746127577	0.77710756	0.0000000
PF03388-PF00840	0.637360493	0.621870501	0.65285049	0.0000000
PF03935-PF00840	0.838085756	0.822595764	0.85357575	0.0000000
PF06439-PF00840	0.879366632	0.863876640	0.89485662	0.0000000
PF10287-PF00840	0.057222311	0.041732319	0.07271230	0.0000000
PF13385-PF00840	1.032629257	1.017139265	1.04811925	0.0000000
PF02210-PF01670	0.121683618	0.106193627	0.13717361	0.0000000
PF03388-PF01670	-0.002573457	-0.018063449	0.01291654	0.9999997
PF03935-PF01670	0.198151806	0.182661814	0.21364180	0.0000000
PF06439-PF01670	0.239432682	0.223942690	0.25492267	0.0000000
PF10287-PF01670	-0.582711640	-0.598201631	-0.56722165	0.0000000
PF13385-PF01670	0.392695307	0.377205315	0.40818530	0.0000000
PF03388-PF02210	-0.124257075	-0.139747067	-0.10876708	0.0000000
PF03935-PF02210	0.076468187	0.060978196	0.09195818	0.0000000
PF06439-PF02210	0.117749063	0.102259072	0.13323906	0.0000000
PF10287-PF02210	-0.704395258	-0.719885250	-0.68890527	0.0000000
PF13385-PF02210	0.271011689	0.255521697	0.28650168	0.0000000
PF03935-PF03388	0.200725263	0.185235271	0.21621525	0.0000000
PF06439-PF03388	0.242006139	0.226516147	0.25749613	0.0000000
PF10287-PF03388	-0.580138183	-0.595628175	-0.56464819	0.0000000
PF13385-PF03388	0.395268764	0.379778772	0.41075876	0.0000000
PF06439-PF03935	0.041280876	0.025790884	0.05677087	0.0000000
PF10287-PF03935	-0.780863445	-0.796353437	-0.76537345	0.0000000
PF13385-PF03935	0.194543501	0.179053509	0.21003349	0.0000000
PF10287-PF06439	-0.822144321	-0.837634313	-0.80665433	0.0000000
PF13385-PF06439	0.153262625	0.137772633	0.16875262	0.0000000
PF13385-PF10287	0.975406947	0.959916955	0.99089694	0.0000000

Pares iguais

diff lwr upr p adj

PF01670-PF00337 0.019473674 0.003983682 0.03496367 0.00671925
 PF03388-PF00337 0.016900217 0.001410225 0.03239021 0.04264915
 PF03388-PF01670 -0.002573457 -0.018063449 0.01291654 0.99999969
 =====

ANOVA Kruskal-Wallis Rank Sum Test (ANOVA não paramétrico)

Não supõe a normalidade

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 49538, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-54.8170052	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-27.9079951	2.133943e-171	2.987520e-171
3 PF00139 - PF00337	26.9090101	1.722867e-159	2.375468e-159
4 PF00054 - PF00354	-9.2350514	2.581620e-20	2.830450e-20
5 PF00139 - PF00354	45.5819537	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	18.6729437	8.219505e-78	9.841776e-78
7 PF00054 - PF00457	35.5762828	3.260447e-277	5.115529e-277
8 PF00139 - PF00457	90.3932879	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	63.4842779	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	44.8113342	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-78.3789978	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-23.5619927	9.459296e-123	1.195550e-122
13 PF00337 - PF00722	-50.4710027	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-69.1439464	0.000000e+00	0.000000e+00
15 PF00457 - PF00722	-113.9552806	0.000000e+00	0.000000e+00
16 PF00054 - PF00840	30.7673187	7.173903e-208	1.106483e-207
17 PF00139 - PF00840	85.5843239	0.000000e+00	0.000000e+00
18 PF00337 - PF00840	58.6753138	0.000000e+00	0.000000e+00
19 PF00354 - PF00840	40.0023702	0.000000e+00	0.000000e+00
20 PF00457 - PF00840	-4.8089640	1.517145e-06	1.624238e-06
21 PF00722 - PF00840	109.1463166	0.000000e+00	0.000000e+00
22 PF00054 - PF01670	-25.1544785	1.262420e-139	1.641147e-139
23 PF00139 - PF01670	29.6625266	2.338177e-193	3.546235e-193
24 PF00337 - PF01670	2.7535165	5.895879e-03	6.238663e-03
25 PF00354 - PF01670	-15.9194271	4.646114e-57	5.420466e-57
26 PF00457 - PF01670	-60.7307613	0.000000e+00	0.000000e+00
27 PF00722 - PF01670	53.2245193	0.000000e+00	0.000000e+00
28 PF00840 - PF01670	-55.9217973	0.000000e+00	0.000000e+00
29 PF00054 - PF02210	-54.5710132	0.000000e+00	0.000000e+00
30 PF00139 - PF02210	0.2459919	8.056885e-01	8.056885e-01
31 PF00337 - PF02210	-26.6630181	1.264322e-156	1.717214e-156
32 PF00354 - PF02210	-45.3359618	0.000000e+00	0.000000e+00
33 PF00457 - PF02210	-90.1472960	0.000000e+00	0.000000e+00

34 PF00722 - PF02210 23.8079846 2.760751e-125 3.538427e-125
35 PF00840 - PF02210 -85.3383320 0.000000e+00 0.000000e+00
36 PF01670 - PF02210 -29.4165347 3.375053e-190 5.034915e-190
37 PF00054 - PF03388 -26.6303701 3.021397e-156 4.043340e-156
38 PF00139 - PF03388 28.1866350 8.527531e-175 1.251622e-174
39 PF00337 - PF03388 1.2776250 2.013817e-01 2.036193e-01
40 PF00354 - PF03388 -17.3953187 8.952742e-68 1.058051e-67
41 PF00457 - PF03388 -62.2066529 0.000000e+00 0.000000e+00
42 PF00722 - PF03388 51.7486277 0.000000e+00 0.000000e+00
43 PF00840 - PF03388 -57.3976889 0.000000e+00 0.000000e+00
44 PF01670 - PF03388 -1.4758916 1.399730e-01 1.431185e-01
45 PF02210 - PF03388 27.9406431 8.565281e-172 1.217876e-171
46 PF00054 - PF03935 -64.1522694 0.000000e+00 0.000000e+00
47 PF00139 - PF03935 -9.3352642 1.007397e-20 1.117965e-20
48 PF00337 - PF03935 -36.2442743 1.223089e-287 1.952651e-287
49 PF00354 - PF03935 -54.9172180 0.000000e+00 0.000000e+00
50 PF00457 - PF03935 -99.7285521 0.000000e+00 0.000000e+00
51 PF00722 - PF03935 14.2267284 6.253982e-46 7.203954e-46
52 PF00840 - PF03935 -94.9195881 0.000000e+00 0.000000e+00
53 PF01670 - PF03935 -38.9977908 0.000000e+00 0.000000e+00
54 PF02210 - PF03935 -9.5812562 9.587171e-22 1.077077e-21
55 PF03388 - PF03935 -37.5218993 0.000000e+00 0.000000e+00
56 PF00054 - PF06439 -76.4652880 0.000000e+00 0.000000e+00
57 PF00139 - PF06439 -21.6482829 6.308725e-104 7.654586e-104
58 PF00337 - PF06439 -48.5572930 0.000000e+00 0.000000e+00
59 PF00354 - PF06439 -67.2302366 0.000000e+00 0.000000e+00
60 PF00457 - PF06439 -112.0415708 0.000000e+00 0.000000e+00
61 PF00722 - PF06439 1.9137098 5.565725e-02 5.755466e-02
62 PF00840 - PF06439 -107.2326068 0.000000e+00 0.000000e+00
63 PF01670 - PF06439 -51.3108095 0.000000e+00 0.000000e+00
64 PF02210 - PF06439 -21.8942748 2.945542e-106 3.622221e-106
65 PF03388 - PF06439 -49.8349179 0.000000e+00 0.000000e+00
66 PF03935 - PF06439 -12.3130187 7.708510e-35 8.768430e-35
67 PF00054 - PF10287 28.1764153 1.137787e-174 1.643470e-174
68 PF00139 - PF10287 82.9934204 0.000000e+00 0.000000e+00
69 PF00337 - PF10287 56.0844104 0.000000e+00 0.000000e+00
70 PF00354 - PF10287 37.4114667 2.543743e-306 4.133582e-306
71 PF00457 - PF10287 -7.3998675 1.363205e-13 1.476805e-13
72 PF00722 - PF10287 106.5554131 0.000000e+00 0.000000e+00
73 PF00840 - PF10287 -2.5909035 9.572433e-03 1.001255e-02
74 PF01670 - PF10287 53.3308938 0.000000e+00 0.000000e+00
75 PF02210 - PF10287 82.7474285 0.000000e+00 0.000000e+00
76 PF03388 - PF10287 54.8067854 0.000000e+00 0.000000e+00
77 PF03935 - PF10287 92.3286846 0.000000e+00 0.000000e+00
78 PF06439 - PF10287 104.6417033 0.000000e+00 0.000000e+00
79 PF00054 - PF13385 -101.6732201 0.000000e+00 0.000000e+00
80 PF00139 - PF13385 -46.8562150 0.000000e+00 0.000000e+00
81 PF00337 - PF13385 -73.7652250 0.000000e+00 0.000000e+00
82 PF00354 - PF13385 -92.4381687 0.000000e+00 0.000000e+00
83 PF00457 - PF13385 -137.2495029 0.000000e+00 0.000000e+00
84 PF00722 - PF13385 -23.2942223 5.073257e-120 6.324198e-120
85 PF00840 - PF13385 -132.4405389 0.000000e+00 0.000000e+00
86 PF01670 - PF13385 -76.5187416 0.000000e+00 0.000000e+00
87 PF02210 - PF13385 -47.1022069 0.000000e+00 0.000000e+00

88 PF03388 - PF13385 -75.0428500 0.000000e+00 0.000000e+00
 89 PF03935 - PF13385 -37.5209508 0.000000e+00 0.000000e+00
 90 PF06439 - PF13385 -25.2079321 3.278793e-140 4.324205e-140
 91 PF10287 - PF13385 -129.8496354 0.000000e+00 0.000000e+00

Pares iguais

Comparison	Z	P.unadj	P.adj
24 PF00337 - PF01670	2.7535165	0.005895879	0.006238663
30 PF00139 - PF02210	0.2459919	0.805688477	0.805688477
39 PF00337 - PF03388	1.2776250	0.201381689	0.203619264
44 PF01670 - PF03388	-1.4758916	0.139973003	0.143118464
61 PF00722 - PF06439	1.9137098	0.055657251	0.057554657
73 PF00840 - PF10287	-2.5909035	0.009572433	0.010012545

=====

=====

ANOVA da entropia entre familias controlando **s = 1.1**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	1107.1	< 2.2e-16 ***
		69286	

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 25121, df = 13, p-value < 2.2e-16

=====

ANOVA

Df	Sum Sq	Mean Sq	F value	Pr(>F)
----	--------	---------	---------	--------

pf 13 4227 325.1 9585 <2e-16 ***

Residuals 69286 2350 0.0

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	0.272753668	0.261203910	0.284303427	0.0000000
PF00337-PF00054	0.152045322	0.140495564	0.163595081	0.0000000
PF00354-PF00054	0.057391623	0.045841865	0.068941382	0.0000000
PF00457-PF00054	-0.382406521	-0.393956279	-0.370856762	0.0000000
PF00722-PF00054	0.341003857	0.329454099	0.352553616	0.0000000
PF00840-PF00054	-0.309036387	-0.320586146	-0.297486629	0.0000000
PF01670-PF00054	0.156432817	0.144883058	0.167982575	0.0000000
PF02210-PF00054	0.238570836	0.227021078	0.250120595	0.0000000
PF03388-PF00054	0.147510867	0.135961108	0.159060625	0.0000000
PF03935-PF00054	0.288786552	0.277236793	0.300336311	0.0000000
PF06439-PF00054	0.320153177	0.308603419	0.331702936	0.0000000
PF10287-PF00054	-0.269214616	-0.280764374	-0.257664857	0.0000000
PF13385-PF00054	0.426305020	0.414755261	0.437854778	0.0000000
PF00337-PF00139	-0.120708346	-0.132258104	-0.109158587	0.0000000
PF00354-PF00139	-0.215362045	-0.226911803	-0.203812286	0.0000000
PF00457-PF00139	-0.655160189	-0.666709947	-0.643610430	0.0000000
PF00722-PF00139	0.068250189	0.056700431	0.079799948	0.0000000
PF00840-PF00139	-0.581790055	-0.593339814	-0.570240297	0.0000000
PF01670-PF00139	-0.116320851	-0.127870610	-0.104771093	0.0000000
PF02210-PF00139	-0.034182832	-0.045732590	-0.022633073	0.0000000
PF03388-PF00139	-0.125242801	-0.136792560	-0.113693043	0.0000000
PF03935-PF00139	0.016032884	0.004483125	0.027582642	0.0012308
PF06439-PF00139	0.047399509	0.035849751	0.058949268	0.0000000
PF10287-PF00139	-0.541968284	-0.553518042	-0.530418525	0.0000000
PF13385-PF00139	0.153551351	0.142001593	0.165101110	0.0000000
PF00354-PF00337	-0.094653699	-0.106203458	-0.083103940	0.0000000
PF00457-PF00337	-0.534451843	-0.546001602	-0.522902084	0.0000000
PF00722-PF00337	0.188958535	0.177408776	0.200508293	0.0000000
PF00840-PF00337	-0.461081710	-0.472631468	-0.449531951	0.0000000
PF01670-PF00337	0.004387494	-0.007162264	0.015937253	0.9962115
PF02210-PF00337	0.086525514	0.074975755	0.098075272	0.0000000
PF03388-PF00337	-0.004534456	-0.016084214	0.007015303	0.9947798

PF03935-PF00337 0.136741230 0.125191471 0.148290988 0.0000000
PF06439-PF00337 0.168107855 0.156558096 0.179657613 0.0000000
PF10287-PF00337 -0.421259938 -0.432809697 -0.409710180 0.0000000
PF13385-PF00337 0.274259697 0.262709939 0.285809456 0.0000000
PF00457-PF00354 -0.439798144 -0.451347903 -0.428248385 0.0000000
PF00722-PF00354 0.283612234 0.272062475 0.295161992 0.0000000
PF00840-PF00354 -0.366428011 -0.377977769 -0.354878252 0.0000000
PF01670-PF00354 0.099041193 0.087491435 0.110590952 0.0000000
PF02210-PF00354 0.181179213 0.169629454 0.192728971 0.0000000
PF03388-PF00354 0.090119243 0.078569485 0.101669002 0.0000000
PF03935-PF00354 0.231394929 0.219845170 0.242944687 0.0000000
PF06439-PF00354 0.262761554 0.251211795 0.274311312 0.0000000
PF10287-PF00354 -0.326606239 -0.338155998 -0.315056481 0.0000000
PF13385-PF00354 0.368913396 0.357363638 0.380463155 0.0000000
PF00722-PF00457 0.723410378 0.711860619 0.734960136 0.0000000
PF00840-PF00457 0.073370133 0.061820375 0.084919892 0.0000000
PF01670-PF00457 0.538839337 0.527289579 0.550389096 0.0000000
PF02210-PF00457 0.620977357 0.609427598 0.632527115 0.0000000
PF03388-PF00457 0.529917387 0.518367629 0.541467146 0.0000000
PF03935-PF00457 0.671193073 0.659643314 0.682742831 0.0000000
PF06439-PF00457 0.702559698 0.691009939 0.714109456 0.0000000
PF10287-PF00457 0.113191905 0.101642146 0.124741663 0.0000000
PF13385-PF00457 0.808711540 0.797161782 0.820261299 0.0000000
PF00840-PF00722 -0.650040245 -0.661590003 -0.638490486 0.0000000
PF01670-PF00722 -0.184571040 -0.196120799 -0.173021282 0.0000000
PF02210-PF00722 -0.102433021 -0.113982779 -0.090883262 0.0000000
PF03388-PF00722 -0.193492991 -0.205042749 -0.181943232 0.0000000
PF03935-PF00722 -0.052217305 -0.063767064 -0.040667547 0.0000000
PF06439-PF00722 -0.020850680 -0.032400439 -0.009300922 0.0000016
PF10287-PF00722 -0.610218473 -0.621768231 -0.598668714 0.0000000
PF13385-PF00722 0.085301162 0.073751404 0.096850921 0.0000000
PF01670-PF00840 0.465469204 0.453919446 0.477018963 0.0000000
PF02210-PF00840 0.547607224 0.536057465 0.559156982 0.0000000
PF03388-PF00840 0.456547254 0.444997496 0.468097013 0.0000000
PF03935-PF00840 0.597822939 0.586273181 0.609372698 0.0000000
PF06439-PF00840 0.629189565 0.617639806 0.640739323 0.0000000
PF10287-PF00840 0.039821772 0.028272013 0.051371530 0.0000000
PF13385-PF00840 0.735341407 0.723791648 0.746891165 0.0000000
PF02210-PF01670 0.082138020 0.070588261 0.093687778 0.0000000
PF03388-PF01670 -0.008921950 -0.020471709 0.002627808 0.4727927
PF03935-PF01670 0.132353735 0.120803977 0.143903494 0.0000000
PF06439-PF01670 0.163720360 0.152170602 0.175270119 0.0000000
PF10287-PF01670 -0.425647432 -0.437197191 -0.414097674 0.0000000
PF13385-PF01670 0.269872203 0.258322444 0.281421961 0.0000000
PF03388-PF02210 -0.091059970 -0.102609728 -0.079510211 0.0000000
PF03935-PF02210 0.050215716 0.038665957 0.061765474 0.0000000
PF06439-PF02210 0.081582341 0.070032582 0.093132099 0.0000000
PF10287-PF02210 -0.507785452 -0.519335211 -0.496235693 0.0000000
PF13385-PF02210 0.187734183 0.176184425 0.199283942 0.0000000
PF03935-PF03388 0.141275685 0.129725927 0.152825444 0.0000000
PF06439-PF03388 0.172642310 0.161092552 0.184192069 0.0000000
PF10287-PF03388 -0.416725482 -0.428275241 -0.405175724 0.0000000
PF13385-PF03388 0.278794153 0.267244394 0.290343911 0.0000000
PF06439-PF03935 0.031366625 0.019816867 0.042916384 0.0000000

PF10287-PF03935 -0.558001168 -0.569550926 -0.546451409 0.0000000
 PF13385-PF03935 0.137518468 0.125968709 0.149068226 0.0000000
 PF10287-PF06439 -0.589367793 -0.600917551 -0.577818034 0.0000000
 PF13385-PF06439 0.106151842 0.094602084 0.117701601 0.0000000
 PF13385-PF10287 0.695519635 0.683969877 0.707069394 0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF03935-PF00139	0.016032884	0.004483125	0.027582642	0.001230821
PF01670-PF00337	0.004387494	-0.007162264	0.015937253	0.996211474
PF03388-PF00337	-0.004534456	-0.016084214	0.007015303	0.994779846
PF03388-PF01670	-0.008921950	-0.020471709	0.002627808	0.472792693

=====

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 49121, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

	Comparison	Z	P.unadj	P.adj
1	PF00054 - PF00139	-57.6643441	0.000000e+00	0.000000e+00
2	PF00054 - PF00337	-30.4108956	3.942516e-203	5.786596e-203
3	PF00139 - PF00337	27.2534486	1.512596e-163	2.085549e-163
4	PF00054 - PF00354	-8.4362037	3.278113e-17	3.682818e-17
5	PF00139 - PF00354	49.2281404	0.000000e+00	0.000000e+00
6	PF00337 - PF00354	21.9746918	5.029502e-107	6.184928e-107
7	PF00054 - PF00457	36.3301050	5.417486e-289	8.648969e-289
8	PF00139 - PF00457	93.9944492	0.000000e+00	0.000000e+00
9	PF00337 - PF00457	66.7410006	0.000000e+00	0.000000e+00
10	PF00354 - PF00457	44.7663087	0.000000e+00	0.000000e+00
11	PF00054 - PF00722	-76.6553190	0.000000e+00	0.000000e+00
12	PF00139 - PF00722	-18.9909749	2.025334e-80	2.457406e-80
13	PF00337 - PF00722	-46.2444235	0.000000e+00	0.000000e+00
14	PF00354 - PF00722	-68.2191153	0.000000e+00	0.000000e+00
15	PF00457 - PF00722	-112.9854240	0.000000e+00	0.000000e+00
16	PF00054 - PF00840	30.7446704	1.440763e-207	2.149335e-207
17	PF00139 - PF00840	88.4090145	0.000000e+00	0.000000e+00
18	PF00337 - PF00840	61.1555659	0.000000e+00	0.000000e+00
19	PF00354 - PF00840	39.1808741	0.000000e+00	0.000000e+00
20	PF00457 - PF00840	-5.5854347	2.331164e-08	2.466697e-08
21	PF00722 - PF00840	107.3999894	0.000000e+00	0.000000e+00
22	PF00054 - PF01670	-24.5987154	1.303803e-133	1.719508e-133
23	PF00139 - PF01670	33.0656287	9.275011e-240	1.455217e-239

24 PF00337 - PF01670 5.8121801 6.166444e-09 6.680314e-09
 25 PF00354 - PF01670 -16.1625117 9.270615e-59 1.081572e-58
 26 PF00457 - PF01670 -60.9288205 0.000000e+00 0.000000e+00
 27 PF00722 - PF01670 52.0566036 0.000000e+00 0.000000e+00
 28 PF00840 - PF01670 -55.3433858 0.000000e+00 0.000000e+00
 29 PF00054 - PF02210 -53.5275885 0.000000e+00 0.000000e+00
 30 PF00139 - PF02210 4.1367557 3.522508e-05 3.642594e-05
 31 PF00337 - PF02210 -23.1166929 3.145857e-118 3.976014e-118
 32 PF00354 - PF02210 -45.0913848 0.000000e+00 0.000000e+00
 33 PF00457 - PF02210 -89.8576935 0.000000e+00 0.000000e+00
 34 PF00722 - PF02210 23.1277306 2.436094e-118 3.122318e-118
 35 PF00840 - PF02210 -84.2722588 0.000000e+00 0.000000e+00
 36 PF01670 - PF02210 -28.9288730 5.175765e-184 7.476105e-184
 37 PF00054 - PF03388 -24.7929053 1.069232e-135 1.430884e-135
 38 PF00139 - PF03388 32.8714388 5.627157e-237 8.679174e-237
 39 PF00337 - PF03388 5.6179902 1.931914e-08 2.068285e-08
 40 PF00354 - PF03388 -16.3567016 3.896695e-60 4.605185e-60
 41 PF00457 - PF03388 -61.1230104 0.000000e+00 0.000000e+00
 42 PF00722 - PF03388 51.8624137 0.000000e+00 0.000000e+00
 43 PF00840 - PF03388 -55.5375757 0.000000e+00 0.000000e+00
 44 PF01670 - PF03388 -0.1941899 8.460272e-01 8.460272e-01
 45 PF02210 - PF03388 28.7346831 1.407466e-181 2.001241e-181
 46 PF00054 - PF03935 -61.9519296 0.000000e+00 0.000000e+00
 47 PF00139 - PF03935 -4.2875855 1.806258e-05 1.889304e-05
 48 PF00337 - PF03935 -31.5410340 2.379886e-218 3.609494e-218
 49 PF00354 - PF03935 -53.5157259 0.000000e+00 0.000000e+00
 50 PF00457 - PF03935 -98.2820346 0.000000e+00 0.000000e+00
 51 PF00722 - PF03935 14.7033894 6.131238e-49 7.062565e-49
 52 PF00840 - PF03935 -92.6966000 0.000000e+00 0.000000e+00
 53 PF01670 - PF03935 -37.3532142 2.248411e-305 3.788989e-305
 54 PF02210 - PF03935 -8.4243411 3.627859e-17 4.026038e-17
 55 PF03388 - PF03935 -37.1590243 3.134392e-302 5.185995e-302
 56 PF00054 - PF06439 -75.5987081 0.000000e+00 0.000000e+00
 57 PF00139 - PF06439 -17.9343640 6.358713e-72 7.613723e-72
 58 PF00337 - PF06439 -45.1878125 0.000000e+00 0.000000e+00
 59 PF00354 - PF06439 -67.1625044 0.000000e+00 0.000000e+00
 60 PF00457 - PF06439 -111.9288131 0.000000e+00 0.000000e+00
 61 PF00722 - PF06439 1.0566109 2.906892e-01 2.939191e-01
 62 PF00840 - PF06439 -106.3433785 0.000000e+00 0.000000e+00
 63 PF01670 - PF06439 -50.9999927 0.000000e+00 0.000000e+00
 64 PF02210 - PF06439 -22.0711196 5.989006e-108 7.465747e-108
 65 PF03388 - PF06439 -50.8058027 0.000000e+00 0.000000e+00
 66 PF03935 - PF06439 -13.6467785 2.110021e-42 2.400149e-42
 67 PF00054 - PF10287 28.4403995 6.405451e-178 8.967631e-178
 68 PF00139 - PF10287 86.1047437 0.000000e+00 0.000000e+00
 69 PF00337 - PF10287 58.8512951 0.000000e+00 0.000000e+00
 70 PF00354 - PF10287 36.8766032 1.096076e-297 1.781124e-297
 71 PF00457 - PF10287 -7.8897055 3.029005e-15 3.320957e-15
 72 PF00722 - PF10287 105.0957185 0.000000e+00 0.000000e+00
 73 PF00840 - PF10287 -2.3042708 2.120744e-02 2.168402e-02
 74 PF01670 - PF10287 53.0391150 0.000000e+00 0.000000e+00
 75 PF02210 - PF10287 81.9679880 0.000000e+00 0.000000e+00
 76 PF03388 - PF10287 53.2333049 0.000000e+00 0.000000e+00
 77 PF03935 - PF10287 90.3923291 0.000000e+00 0.000000e+00

```

78 PF06439 - PF10287 104.0391076 0.000000e+00 0.000000e+00
79 PF00054 - PF13385 -100.9810362 0.000000e+00 0.000000e+00
80 PF00139 - PF13385 -43.3166920 0.000000e+00 0.000000e+00
81 PF00337 - PF13385 -70.5701406 0.000000e+00 0.000000e+00
82 PF00354 - PF13385 -92.5448325 0.000000e+00 0.000000e+00
83 PF00457 - PF13385 -137.3111412 0.000000e+00 0.000000e+00
84 PF00722 - PF13385 -24.3257172 1.047908e-130 1.362281e-130
85 PF00840 - PF13385 -131.7257065 0.000000e+00 0.000000e+00
86 PF01670 - PF13385 -76.3823207 0.000000e+00 0.000000e+00
87 PF02210 - PF13385 -47.4534477 0.000000e+00 0.000000e+00
88 PF03388 - PF13385 -76.1881308 0.000000e+00 0.000000e+00
89 PF03935 - PF13385 -39.0291066 0.000000e+00 0.000000e+00
90 PF06439 - PF13385 -25.3823281 3.952717e-142 5.368615e-142
91 PF10287 - PF13385 -129.4214357 0.000000e+00 0.000000e+00

```

Pares iguais

Comparison	Z	P.unadj	P.adj
44 PF01670 - PF03388	-0.1941899	0.84602719	0.84602719
61 PF00722 - PF06439	1.0566109	0.29068919	0.29391907
73 PF00840 - PF10287	-2.3042708	0.02120744	0.02168402

=====

=====

ANOVA da entropia entre familias controlando **s = 1.2**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	1207.2	< 2.2e-16 ***
		69286	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 28314, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 2235 171.94 8944 <2e-16 ***

Residuals 69286 1332 0.02

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	0.189674637	0.180979980	0.1983692950	0.0000000
PF00337-PF00054	0.105024667	0.096330009	0.1137193248	0.0000000
PF00354-PF00054	0.039559403	0.030864745	0.0482540604	0.0000000
PF00457-PF00054	-0.293409491	-0.302104149	-0.2847148336	0.0000000
PF00722-PF00054	0.240927526	0.232232868	0.2496221837	0.0000000
PF00840-PF00054	-0.228225481	-0.236920138	-0.2195308229	0.0000000
PF01670-PF00054	0.112372445	0.103677787	0.1210671024	0.0000000
PF02210-PF00054	0.169685549	0.160990891	0.1783802063	0.0000000
PF03388-PF00054	0.102742736	0.094048079	0.1114373943	0.0000000
PF03935-PF00054	0.202592339	0.193897682	0.2112869972	0.0000000
PF06439-PF00054	0.226406807	0.217712149	0.2351014650	0.0000000
PF10287-PF00054	-0.198519029	-0.207213687	-0.1898243712	0.0000000
PF13385-PF00054	0.300245247	0.291550589	0.3089399050	0.0000000
PF00337-PF00139	-0.084649970	-0.093344628	-0.0759553125	0.0000000
PF00354-PF00139	-0.150115235	-0.158809892	-0.1414205769	0.0000000
PF00457-PF00139	-0.483084129	-0.491778786	-0.4743894709	0.0000000
PF00722-PF00139	0.051252889	0.042558231	0.0599475465	0.0000000
PF00840-PF00139	-0.417900118	-0.426594776	-0.4092054602	0.0000000
PF01670-PF00139	-0.077302193	-0.085996850	-0.0686075349	0.0000000
PF02210-PF00139	-0.019989089	-0.028683747	-0.0112944310	0.0000000
PF03388-PF00139	-0.086931901	-0.095626559	-0.0782372430	0.0000000
PF03935-PF00139	0.012917702	0.004223044	0.0216123599	0.0003065
PF06439-PF00139	0.036732170	0.028037512	0.0454268277	0.0000000
PF10287-PF00139	-0.388193666	-0.396888324	-0.3794990085	0.0000000

PF13385-PF00139	0.110570610	0.101875952	0.1192652677	0.0000000
PF00354-PF00337	-0.065465264	-0.074159922	-0.0567706066	0.0000000
PF00457-PF00337	-0.398434158	-0.407128816	-0.3897395006	0.0000000
PF00722-PF00337	0.135902859	0.127208201	0.1445975167	0.0000000
PF00840-PF00337	-0.333250148	-0.341944805	-0.3245554900	0.0000000
PF01670-PF00337	0.007347778	-0.001346880	0.0160424353	0.3169181
PF02210-PF00337	0.064660881	0.055966224	0.0733555392	0.0000000
PF03388-PF00337	-0.002281931	-0.010976588	0.0064127272	0.9999276
PF03935-PF00337	0.097567672	0.088873015	0.1062623301	0.0000000
PF06439-PF00337	0.121382140	0.112687482	0.1300767979	0.0000000
PF10287-PF00337	-0.303543696	-0.312238354	-0.2948490382	0.0000000
PF13385-PF00337	0.195220580	0.186525922	0.2039152379	0.0000000
PF00457-PF00354	-0.332968894	-0.341663552	-0.3242742362	0.0000000
PF00722-PF00354	0.201368123	0.192673466	0.2100627811	0.0000000
PF00840-PF00354	-0.267784883	-0.276479541	-0.2590902256	0.0000000
PF01670-PF00354	0.072813042	0.064118384	0.0815076997	0.0000000
PF02210-PF00354	0.130126146	0.121431488	0.1388208036	0.0000000
PF03388-PF00354	0.063183334	0.054488676	0.0718779916	0.0000000
PF03935-PF00354	0.163032937	0.154338279	0.1717275945	0.0000000
PF06439-PF00354	0.186847405	0.178152747	0.1955420623	0.0000000
PF10287-PF00354	-0.238078432	-0.246773089	-0.2293837739	0.0000000
PF13385-PF00354	0.260685845	0.251991187	0.2693805023	0.0000000
PF00722-PF00457	0.534337017	0.525642360	0.5430316751	0.0000000
PF00840-PF00457	0.065184011	0.056489353	0.0738786684	0.0000000
PF01670-PF00457	0.405781936	0.397087278	0.4144765937	0.0000000
PF02210-PF00457	0.463095040	0.454400382	0.4717896976	0.0000000
PF03388-PF00457	0.396152228	0.387457570	0.4048468856	0.0000000
PF03935-PF00457	0.496001831	0.487307173	0.5046964885	0.0000000
PF06439-PF00457	0.519816299	0.511121641	0.5285109563	0.0000000
PF10287-PF00457	0.094890462	0.086195805	0.1035851201	0.0000000
PF13385-PF00457	0.593654739	0.584960081	0.6023493963	0.0000000
PF00840-PF00722	-0.469153007	-0.477847664	-0.4604583489	0.0000000
PF01670-PF00722	-0.128555081	-0.137249739	-0.1198604236	0.0000000
PF02210-PF00722	-0.071241977	-0.079936635	-0.0625473197	0.0000000
PF03388-PF00722	-0.138184789	-0.146879447	-0.1294901317	0.0000000
PF03935-PF00722	-0.038335187	-0.047029844	-0.0296405288	0.0000000
PF06439-PF00722	-0.014520719	-0.023215377	-0.0058260610	0.0000168
PF10287-PF00722	-0.439446555	-0.448141213	-0.4307518972	0.0000000
PF13385-PF00722	0.059317721	0.050623063	0.0680123790	0.0000000
PF01670-PF00840	0.340597925	0.331903268	0.3492925830	0.0000000
PF02210-PF00840	0.397911029	0.389216371	0.4066056870	0.0000000
PF03388-PF00840	0.330968217	0.322273559	0.3396628749	0.0000000
PF03935-PF00840	0.430817820	0.422123162	0.4395124778	0.0000000
PF06439-PF00840	0.454632288	0.445937630	0.4633269456	0.0000000
PF10287-PF00840	0.029706452	0.021011794	0.0384011095	0.0000000
PF13385-PF00840	0.528470728	0.519776070	0.5371653857	0.0000000
PF02210-PF01670	0.057313104	0.048618446	0.0660077617	0.0000000
PF03388-PF01670	-0.009629708	-0.018324366	-0.0009350504	0.0360835
PF03935-PF01670	0.090219895	0.081525237	0.0989145525	0.0000000
PF06439-PF01670	0.114034363	0.105339705	0.1227290204	0.0000000
PF10287-PF01670	-0.310891474	-0.319586131	-0.3021968158	0.0000000
PF13385-PF01670	0.187872803	0.179178145	0.1965674604	0.0000000
PF03388-PF02210	-0.066942812	-0.075637470	-0.0582481543	0.0000000
PF03935-PF02210	0.032906791	0.024212133	0.0416014486	0.0000000

PF06439-PF02210	0.056721259	0.048026601	0.0654159164	0.0000000
PF10287-PF02210	-0.368204577	-0.376899235	-0.3595099197	0.0000000
PF13385-PF02210	0.130559699	0.121865041	0.1392543564	0.0000000
PF03935-PF03388	0.099849603	0.091154945	0.1085442607	0.0000000
PF06439-PF03388	0.123664071	0.114969413	0.1323587285	0.0000000
PF10287-PF03388	-0.301261765	-0.309956423	-0.2925671077	0.0000000
PF13385-PF03388	0.197502511	0.188807853	0.2061971685	0.0000000
PF06439-PF03935	0.023814468	0.015119810	0.0325091256	0.0000000
PF10287-PF03935	-0.401111368	-0.409806026	-0.3924167106	0.0000000
PF13385-PF03935	0.097652908	0.088958250	0.1063475656	0.0000000
PF10287-PF06439	-0.424925836	-0.433620494	-0.4162311784	0.0000000
PF13385-PF06439	0.073838440	0.065143782	0.0825330978	0.0000000
PF13385-PF10287	0.498764276	0.490069618	0.5074589339	0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF01670-PF00337	0.007347778	-0.00134688	0.0160424353	0.31691806
PF03388-PF00337	-0.002281931	-0.01097659	0.0064127272	0.99992758
PF03388-PF01670	-0.009629708	-0.01832437	-0.0009350504	0.03608349
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 48956, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

Comparison	Z	P.unadj	P.adj
1 PF00054 - PF00139	-55.9041146	0.000000e+00	0.000000e+00
2 PF00054 - PF00337	-29.3549246	2.067570e-189	2.986490e-189
3 PF00139 - PF00337	26.5491900	2.624104e-155	3.618083e-155
4 PF00054 - PF00354	-8.1696144	3.093768e-16	3.433328e-16
5 PF00139 - PF00354	47.7345003	0.000000e+00	0.000000e+00
6 PF00337 - PF00354	21.1853103	1.304635e-99	1.604349e-99
7 PF00054 - PF00457	36.9524023	6.664351e-299	1.102647e-298
8 PF00139 - PF00457	92.8565169	0.000000e+00	0.000000e+00
9 PF00337 - PF00457	66.3073269	0.000000e+00	0.000000e+00
10 PF00354 - PF00457	45.1220166	0.000000e+00	0.000000e+00
11 PF00054 - PF00722	-76.2514828	0.000000e+00	0.000000e+00
12 PF00139 - PF00722	-20.3473682	4.898794e-92	5.943870e-92
13 PF00337 - PF00722	-46.8965582	0.000000e+00	0.000000e+00
14 PF00354 - PF00722	-68.0818685	0.000000e+00	0.000000e+00

15 PF00457 - PF00722 -113.2038851 0.000000e+00 0.000000e+00
16 PF00054 - PF00840 30.7732194 5.981634e-208 8.923421e-208
17 PF00139 - PF00840 86.6773340 0.000000e+00 0.000000e+00
18 PF00337 - PF00840 60.1281440 0.000000e+00 0.000000e+00
19 PF00354 - PF00840 38.9428338 0.000000e+00 0.000000e+00
20 PF00457 - PF00840 -6.1791829 6.443422e-10 6.980374e-10
21 PF00722 - PF00840 107.0247022 0.000000e+00 0.000000e+00
22 PF00054 - PF01670 -24.4784139 2.508324e-132 3.308079e-132
23 PF00139 - PF01670 31.4257008 9.017824e-217 1.367703e-216
24 PF00337 - PF01670 4.8765108 1.079788e-06 1.129433e-06
25 PF00354 - PF01670 -16.3087995 8.545576e-60 1.009932e-59
26 PF00457 - PF01670 -61.4308161 0.000000e+00 0.000000e+00
27 PF00722 - PF01670 51.7730689 0.000000e+00 0.000000e+00
28 PF00840 - PF01670 -55.2516333 0.000000e+00 0.000000e+00
29 PF00054 - PF02210 -53.7646793 0.000000e+00 0.000000e+00
30 PF00139 - PF02210 2.1394354 3.240043e-02 3.312852e-02
31 PF00337 - PF02210 -24.4097547 1.347351e-131 1.751556e-131
32 PF00354 - PF02210 -45.5950649 0.000000e+00 0.000000e+00
33 PF00457 - PF02210 -90.7170816 0.000000e+00 0.000000e+00
34 PF00722 - PF02210 22.4868035 5.588304e-112 7.062995e-112
35 PF00840 - PF02210 -84.5378987 0.000000e+00 0.000000e+00
36 PF01670 - PF02210 -29.2862654 1.551549e-188 2.206109e-188
37 PF00054 - PF03388 -23.9926186 3.320670e-127 4.256070e-127
38 PF00139 - PF03388 31.9114961 1.849560e-223 2.901896e-223
39 PF00337 - PF03388 5.3623060 8.216617e-08 8.796614e-08
40 PF00354 - PF03388 -15.8230042 2.159384e-56 2.519281e-56
41 PF00457 - PF03388 -60.9450209 0.000000e+00 0.000000e+00
42 PF00722 - PF03388 52.2588642 0.000000e+00 0.000000e+00
43 PF00840 - PF03388 -54.7658380 0.000000e+00 0.000000e+00
44 PF01670 - PF03388 0.4857953 6.271123e-01 6.271123e-01
45 PF02210 - PF03388 29.7720607 8.987252e-195 1.319097e-194
46 PF00054 - PF03935 -61.0611062 0.000000e+00 0.000000e+00
47 PF00139 - PF03935 -5.1569916 2.509491e-07 2.655392e-07
48 PF00337 - PF03935 -31.7061816 1.276918e-220 1.969484e-220
49 PF00354 - PF03935 -52.8914919 0.000000e+00 0.000000e+00
50 PF00457 - PF03935 -98.0135085 0.000000e+00 0.000000e+00
51 PF00722 - PF03935 15.1903766 4.095599e-52 4.717716e-52
52 PF00840 - PF03935 -91.8343256 0.000000e+00 0.000000e+00
53 PF01670 - PF03935 -36.5826924 5.390456e-293 8.605815e-293
54 PF02210 - PF03935 -7.2964269 2.955102e-13 3.239931e-13
55 PF03388 - PF03935 -37.0684876 9.047061e-301 1.524597e-300
56 PF00054 - PF06439 -75.6852642 0.000000e+00 0.000000e+00
57 PF00139 - PF06439 -19.7811496 4.327126e-87 5.181164e-87
58 PF00337 - PF06439 -46.3303396 0.000000e+00 0.000000e+00
59 PF00354 - PF06439 -67.5156499 0.000000e+00 0.000000e+00
60 PF00457 - PF06439 -112.6376665 0.000000e+00 0.000000e+00
61 PF00722 - PF06439 0.5662186 5.712452e-01 5.775924e-01
62 PF00840 - PF06439 -106.4584836 0.000000e+00 0.000000e+00
63 PF01670 - PF06439 -51.2068504 0.000000e+00 0.000000e+00
64 PF02210 - PF06439 -21.9205849 1.653195e-106 2.060832e-106
65 PF03388 - PF06439 -51.6926456 0.000000e+00 0.000000e+00
66 PF03935 - PF06439 -14.6241580 1.969929e-48 2.240794e-48
67 PF00054 - PF10287 28.5342992 4.399507e-179 6.159310e-179
68 PF00139 - PF10287 84.4384139 0.000000e+00 0.000000e+00

69	PF00337 - PF10287	57.8892239	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	36.7039136	6.325254e-295	1.027854e-294
71	PF00457 - PF10287	-8.4181030	3.826292e-17	4.298673e-17
72	PF00722 - PF10287	104.7857820	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-2.2389202	2.516111e-02	2.601888e-02
74	PF01670 - PF10287	53.0127131	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.2989785	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	52.5269178	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	89.5954055	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.2195635	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-100.7721073	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-44.8679927	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-71.4171827	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-92.6024930	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-137.7245096	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-24.5206245	8.902588e-133	1.191376e-132
85	PF00840 - PF13385	-131.5453267	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-76.2936935	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-47.0074280	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-76.7794887	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-39.7110011	0.000000e+00	0.000000e+00
90	PF06439 - PF13385	-25.0868431	6.922396e-139	9.402060e-139
91	PF10287 - PF13385	-129.3064066	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
30	PF00139 - PF02210	2.1394354	0.03240043	0.03312852
44	PF01670 - PF03388	0.4857953	0.62711231	0.62711231
61	PF00722 - PF06439	0.5662186	0.57124519	0.57759236
73	PF00840 - PF10287	-2.2389202	0.02516111	0.02601888

=====

ANOVA da entropia entre familias controlando **s = 1.3**

Teste de homogeneidade de variância: Levene Teste

Levene's Test for Homogeneity of Variance (center = mean)

	Df	F value	Pr(>F)
group	13	1303.2	< 2.2e-16 ***
		69286	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Teste de homogeneidade de variância: Barlett Teste

Bartlett test of homogeneity of variances

data: datb.x\$entropia and datb.x\$pf

Bartlett's K-squared = 31828, df = 13, p-value < 2.2e-16

=====

ANOVA

Df Sum Sq Mean Sq F value Pr(>F)

pf 13 1201.1 92.39 8287 <2e-16 ***

Residuals 69286 772.5 0.01

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Teste Tukey HSD

Tukey multiple comparisons of means

90% family-wise confidence level

Fit: aov(formula = entropia ~ pf, data = datb.x)

\$pf

	diff	lwr	upr	p adj
PF00139-PF00054	0.134133543	0.1275123125	0.140754774	0.0000000
PF00337-PF00054	0.074215777	0.0675945462	0.080837007	0.0000000
PF00354-PF00054	0.027401542	0.0207803117	0.034022773	0.0000000
PF00457-PF00054	-0.226963649	-0.2335848792	-0.220342418	0.0000000
PF00722-PF00054	0.170840605	0.1642193750	0.177461836	0.0000000
PF00840-PF00054	-0.169516003	-0.1761372339	-0.162894773	0.0000000
PF01670-PF00054	0.081213255	0.0745920247	0.087834486	0.0000000
PF02210-PF00054	0.121049295	0.1144280640	0.127670525	0.0000000
PF03388-PF00054	0.071683286	0.0650620556	0.078304517	0.0000000
PF03935-PF00054	0.142555729	0.1359344988	0.149176960	0.0000000
PF06439-PF00054	0.160626845	0.1540056142	0.167248075	0.0000000
PF10287-PF00054	-0.147537339	-0.1541585698	-0.140916109	0.0000000
PF13385-PF00054	0.212224958	0.2056037274	0.218846188	0.0000000
PF00337-PF00139	-0.059917766	-0.0665389968	-0.053296536	0.0000000
PF00354-PF00139	-0.106732001	-0.1133532314	-0.100110770	0.0000000
PF00457-PF00139	-0.361097192	-0.3677184223	-0.354475961	0.0000000

PF00722-PF00139	0.036707062	0.0300858319	0.043328293	0.0000000
PF00840-PF00139	-0.303649546	-0.3102707770	-0.297028316	0.0000000
PF01670-PF00139	-0.052920288	-0.0595415184	-0.046299057	0.0000000
PF02210-PF00139	-0.013084249	-0.0197054791	-0.006463018	0.0000001
PF03388-PF00139	-0.062450257	-0.0690714874	-0.055829026	0.0000000
PF03935-PF00139	0.008422186	0.0018009558	0.015043417	0.0056019
PF06439-PF00139	0.026493302	0.0198720711	0.033114532	0.0000000
PF10287-PF00139	-0.281670882	-0.2882921129	-0.275049652	0.0000000
PF13385-PF00139	0.078091415	0.0714701844	0.084712645	0.0000000
PF00354-PF00337	-0.046814235	-0.0534354651	-0.040193004	0.0000000
PF00457-PF00337	-0.301179425	-0.3078006560	-0.294558195	0.0000000
PF00722-PF00337	0.096624829	0.0900035982	0.103246059	0.0000000
PF00840-PF00337	-0.243731780	-0.2503530107	-0.237110550	0.0000000
PF01670-PF00337	0.006997478	0.0003762479	0.013618709	0.0595864
PF02210-PF00337	0.046833518	0.0402122872	0.053454748	0.0000000
PF03388-PF00337	-0.002532491	-0.0091537211	0.004088740	0.9959491
PF03935-PF00337	0.068339953	0.0617187221	0.074961183	0.0000000
PF06439-PF00337	0.086411068	0.0797898374	0.093032298	0.0000000
PF10287-PF00337	-0.221753116	-0.2283743466	-0.215131886	0.0000000
PF13385-PF00337	0.138009181	0.1313879507	0.144630412	0.0000000
PF00457-PF00354	-0.254365191	-0.2609864214	-0.247743960	0.0000000
PF00722-PF00354	0.143439063	0.1368178327	0.150060294	0.0000000
PF00840-PF00354	-0.196917546	-0.2035387762	-0.190296315	0.0000000
PF01670-PF00354	0.053811713	0.0471904824	0.060432944	0.0000000
PF02210-PF00354	0.093647752	0.0870265218	0.100268983	0.0000000
PF03388-PF00354	0.044281744	0.0376605134	0.050902974	0.0000000
PF03935-PF00354	0.115154187	0.1085329566	0.121775418	0.0000000
PF06439-PF00354	0.133225302	0.1266040720	0.139846533	0.0000000
PF10287-PF00354	-0.174938882	-0.1815601121	-0.168317651	0.0000000
PF13385-PF00354	0.184823416	0.1782021852	0.191444646	0.0000000
PF00722-PF00457	0.397804254	0.3911830236	0.404425485	0.0000000
PF00840-PF00457	0.057447645	0.0508264147	0.064068876	0.0000000
PF01670-PF00457	0.308176904	0.3015556734	0.314798134	0.0000000
PF02210-PF00457	0.348012943	0.3413917127	0.354634174	0.0000000
PF03388-PF00457	0.298646935	0.2920257043	0.305268165	0.0000000
PF03935-PF00457	0.369519378	0.3628981475	0.376140609	0.0000000
PF06439-PF00457	0.387590493	0.3809692629	0.394211724	0.0000000
PF10287-PF00457	0.079426309	0.0728050788	0.086047540	0.0000000
PF13385-PF00457	0.439188607	0.4325673761	0.445809837	0.0000000
PF00840-PF00722	-0.340356609	-0.3469778394	-0.333735378	0.0000000
PF01670-PF00722	-0.089627350	-0.0962485808	-0.083006120	0.0000000
PF02210-PF00722	-0.049791311	-0.0564125415	-0.043170080	0.0000000
PF03388-PF00722	-0.099157319	-0.1057785499	-0.092536089	0.0000000
PF03935-PF00722	-0.028284876	-0.0349061067	-0.021663646	0.0000000
PF06439-PF00722	-0.010213761	-0.0168349913	-0.003592530	0.0001300
PF10287-PF00722	-0.318377945	-0.3249991753	-0.311756714	0.0000000
PF13385-PF00722	0.041384352	0.0347631219	0.048005583	0.0000000
PF01670-PF00840	0.250729259	0.2441080281	0.257350489	0.0000000
PF02210-PF00840	0.290565298	0.2839440674	0.297186528	0.0000000
PF03388-PF00840	0.241199290	0.2345780590	0.247820520	0.0000000
PF03935-PF00840	0.312071733	0.3054505022	0.318692963	0.0000000
PF06439-PF00840	0.330142848	0.3235216176	0.336764079	0.0000000
PF10287-PF00840	0.021978664	0.0153574335	0.028599895	0.0000000
PF13385-PF00840	0.381740961	0.3751197308	0.388362192	0.0000000

PF02210-PF01670	0.039836039	0.0332148088	0.046457270	0.0000000
PF03388-PF01670	-0.009529969	-0.0161511996	-0.002908739	0.0006015
PF03935-PF01670	0.061342474	0.0547212436	0.067963705	0.0000000
PF06439-PF01670	0.079413590	0.0727923590	0.086034820	0.0000000
PF10287-PF01670	-0.228750595	-0.2353718251	-0.222129364	0.0000000
PF13385-PF01670	0.131011703	0.1243904722	0.137632933	0.0000000
PF03388-PF02210	-0.049366008	-0.0559872389	-0.042744778	0.0000000
PF03935-PF02210	0.021506435	0.0148852043	0.028127665	0.0000000
PF06439-PF02210	0.039577550	0.0329563197	0.046198781	0.0000000
PF10287-PF02210	-0.268586634	-0.2752078644	-0.261965403	0.0000000
PF13385-PF02210	0.091175663	0.0845544329	0.097796894	0.0000000
PF03935-PF03388	0.070872443	0.0642512127	0.077493674	0.0000000
PF06439-PF03388	0.088943559	0.0823223280	0.095564789	0.0000000
PF10287-PF03388	-0.219220625	-0.2258418560	-0.212599395	0.0000000
PF13385-PF03388	0.140541672	0.1339204413	0.147162902	0.0000000
PF06439-PF03935	0.018071115	0.0114498848	0.024692346	0.0000000
PF10287-PF03935	-0.290093069	-0.2967142992	-0.283471838	0.0000000
PF13385-PF03935	0.069669229	0.0630479980	0.076290459	0.0000000
PF10287-PF06439	-0.308164184	-0.3147854146	-0.301542953	0.0000000
PF13385-PF06439	0.051598113	0.0449768827	0.058219344	0.0000000
PF13385-PF10287	0.359762297	0.3531410667	0.366383528	0.0000000

Pares iguais

	diff	lwr	upr	p adj
PF03935-PF00139	0.008422186	0.0018009558	0.01504342	0.00560192
PF01670-PF00337	0.006997478	0.0003762479	0.01361871	0.05958638
PF03388-PF00337	-0.002532491	-0.0091537211	0.00408874	0.99594911
=====				

ANOVA Kruskal-Wallis Rank Sum Test

Kruskal-Wallis rank sum test

data: entropia by pf

Kruskal-Wallis chi-squared = 48689, df = 13, p-value < 2.2e-16

Dunn's Kruskal-Wallis Multiple Comparisons

	Comparison	Z	P.unadj	P.adj
1	PF00054 - PF00139	-54.92501754	0.000000e+00	0.000000e+00
2	PF00054 - PF00337	-28.96124884	2.025368e-184	2.879819e-184
3	PF00139 - PF00337	25.96376870	1.271223e-148	1.752747e-148
4	PF00054 - PF00354	-7.81871471	5.336540e-15	5.922258e-15
5	PF00139 - PF00354	47.10630283	0.000000e+00	0.000000e+00
6	PF00337 - PF00354	21.14253413	3.232470e-99	3.975064e-99
7	PF00054 - PF00457	37.59106697	0.000000e+00	0.000000e+00
8	PF00139 - PF00457	92.51608452	0.000000e+00	0.000000e+00

9 PF00337 - PF00457 66.55231581 0.000000e+00 0.000000e+00
10 PF00354 - PF00457 45.40978168 0.000000e+00 0.000000e+00
11 PF00054 - PF00722 -75.48616205 0.000000e+00 0.000000e+00
12 PF00139 - PF00722 -20.56114451 6.117300e-94 7.324662e-94
13 PF00337 - PF00722 -46.52491321 0.000000e+00 0.000000e+00
14 PF00354 - PF00722 -67.66744734 0.000000e+00 0.000000e+00
15 PF00457 - PF00722 -113.07722903 0.000000e+00 0.000000e+00
16 PF00054 - PF00840 30.81484319 1.657948e-208 2.514555e-208
17 PF00139 - PF00840 85.73986073 0.000000e+00 0.000000e+00
18 PF00337 - PF00840 59.77609203 0.000000e+00 0.000000e+00
19 PF00354 - PF00840 38.63355790 0.000000e+00 0.000000e+00
20 PF00457 - PF00840 -6.77622379 1.233576e-11 1.352474e-11
21 PF00722 - PF00840 106.30100524 0.000000e+00 0.000000e+00
22 PF00054 - PF01670 -24.28178482 3.053591e-130 3.969669e-130
23 PF00139 - PF01670 30.64323273 3.252448e-206 4.773754e-206
24 PF00337 - PF01670 4.67946402 2.876258e-06 3.008500e-06
25 PF00354 - PF01670 -16.46307011 6.758391e-61 7.987189e-61
26 PF00457 - PF01670 -61.87285179 0.000000e+00 0.000000e+00
27 PF00722 - PF01670 51.20437724 0.000000e+00 0.000000e+00
28 PF00840 - PF01670 -55.09662800 0.000000e+00 0.000000e+00
29 PF00054 - PF02210 -53.71765063 0.000000e+00 0.000000e+00
30 PF00139 - PF02210 1.20736691 2.272909e-01 2.298163e-01
31 PF00337 - PF02210 -24.75640179 2.645268e-135 3.488687e-135
32 PF00354 - PF02210 -45.89893592 0.000000e+00 0.000000e+00
33 PF00457 - PF02210 -91.30871760 0.000000e+00 0.000000e+00
34 PF00722 - PF02210 21.76851143 4.613721e-105 5.751351e-105
35 PF00840 - PF02210 -84.53249381 0.000000e+00 0.000000e+00
36 PF01670 - PF02210 -29.43586581 1.909638e-190 2.758366e-190
37 PF00054 - PF03388 -23.01463207 3.326799e-117 4.263925e-117
38 PF00139 - PF03388 31.91038547 1.916351e-223 3.006688e-223
39 PF00337 - PF03388 5.94661677 2.737413e-09 2.930642e-09
40 PF00354 - PF03388 -15.19591736 3.763577e-52 4.281069e-52
41 PF00457 - PF03388 -60.60569905 0.000000e+00 0.000000e+00
42 PF00722 - PF03388 52.47152998 0.000000e+00 0.000000e+00
43 PF00840 - PF03388 -53.82947526 0.000000e+00 0.000000e+00
44 PF01670 - PF03388 1.26715274 2.051007e-01 2.097097e-01
45 PF02210 - PF03388 30.70301856 5.187385e-207 7.738558e-207
46 PF00054 - PF03935 -59.86125609 0.000000e+00 0.000000e+00
47 PF00139 - PF03935 -4.93623855 7.964368e-07 8.427413e-07
48 PF00337 - PF03935 -30.90000725 1.194190e-209 1.841886e-209
49 PF00354 - PF03935 -52.04254138 0.000000e+00 0.000000e+00
50 PF00457 - PF03935 -97.45232306 0.000000e+00 0.000000e+00
51 PF00722 - PF03935 15.62490596 4.926479e-55 5.674804e-55
52 PF00840 - PF03935 -90.67609928 0.000000e+00 0.000000e+00
53 PF01670 - PF03935 -35.57947127 2.910527e-277 4.646631e-277
54 PF02210 - PF03935 -6.14360546 8.066907e-10 8.739149e-10
55 PF03388 - PF03935 -36.84662402 3.312262e-297 5.480288e-297
56 PF00054 - PF06439 -75.51159479 0.000000e+00 0.000000e+00
57 PF00139 - PF06439 -20.58657725 3.620604e-94 4.392999e-94
58 PF00337 - PF06439 -46.55034595 0.000000e+00 0.000000e+00
59 PF00354 - PF06439 -67.69288009 0.000000e+00 0.000000e+00
60 PF00457 - PF06439 -113.10266177 0.000000e+00 0.000000e+00
61 PF00722 - PF06439 -0.02543274 9.797098e-01 9.797098e-01
62 PF00840 - PF06439 -106.32643798 0.000000e+00 0.000000e+00

63	PF01670 - PF06439	-51.22980998	0.000000e+00	0.000000e+00
64	PF02210 - PF06439	-21.79394417	2.648294e-105	3.347149e-105
65	PF03388 - PF06439	-52.49696272	0.000000e+00	0.000000e+00
66	PF03935 - PF06439	-15.65033871	3.304561e-55	3.855321e-55
67	PF00054 - PF10287	28.68337716	6.150407e-181	8.610570e-181
68	PF00139 - PF10287	83.60839470	0.000000e+00	0.000000e+00
69	PF00337 - PF10287	57.64462600	0.000000e+00	0.000000e+00
70	PF00354 - PF10287	36.50209187	1.027364e-291	1.669467e-291
71	PF00457 - PF10287	-8.90768981	5.210694e-19	5.853990e-19
72	PF00722 - PF10287	104.16953921	0.000000e+00	0.000000e+00
73	PF00840 - PF10287	-2.13146603	3.305077e-02	3.417750e-02
74	PF01670 - PF10287	52.96516198	0.000000e+00	0.000000e+00
75	PF02210 - PF10287	82.40102779	0.000000e+00	0.000000e+00
76	PF03388 - PF10287	51.69800923	0.000000e+00	0.000000e+00
77	PF03935 - PF10287	88.54463325	0.000000e+00	0.000000e+00
78	PF06439 - PF10287	104.19497195	0.000000e+00	0.000000e+00
79	PF00054 - PF13385	-100.39692391	0.000000e+00	0.000000e+00
80	PF00139 - PF13385	-45.47190637	0.000000e+00	0.000000e+00
81	PF00337 - PF13385	-71.43567507	0.000000e+00	0.000000e+00
82	PF00354 - PF13385	-92.57820920	0.000000e+00	0.000000e+00
83	PF00457 - PF13385	-137.98799088	0.000000e+00	0.000000e+00
84	PF00722 - PF13385	-24.91076185	5.688425e-137	7.726069e-137
85	PF00840 - PF13385	-131.21176710	0.000000e+00	0.000000e+00
86	PF01670 - PF13385	-76.11513909	0.000000e+00	0.000000e+00
87	PF02210 - PF13385	-46.67927328	0.000000e+00	0.000000e+00
88	PF03388 - PF13385	-77.38229184	0.000000e+00	0.000000e+00
89	PF03935 - PF13385	-40.53566782	0.000000e+00	0.000000e+00
90	PF06439 - PF13385	-24.88532911	1.072607e-136	1.435400e-136
91	PF10287 - PF13385	-129.08030107	0.000000e+00	0.000000e+00

Pares iguais

	Comparison	Z	P.unadj	P.adj
30	PF00139 - PF02210	1.20736691	0.22729087	0.2298163
44	PF01670 - PF03388	1.26715274	0.20510068	0.2097097
61	PF00722 - PF06439	-0.02543274	0.97970980	0.9797098
73	PF00840 - PF10287	-2.13146603	0.03305077	0.0341775
=====				

O teste de Tukey foi aplicado depois de verificar que a ANOVA foi significativa na distribuição de entropia Havrda-Charvat para as probabilidades simples e compostas. Em todos os experimentos realizados, o nível de significância utilizado foi de $\alpha = 0.05$.

Os resultados obtidos para o teste nas entropias de probabilidades simples são apresentados em gráficos bidimensionais, onde o eixo horizontal relaciona as

100 colunas e o eixo vertical os valores (representados como pontos) dos coeficientes p value de cada ANOVA, e representado pela reta horizontal o valor correspondente do nível de significância de 0.05.

Para este experimento a quantidade de clãs utilizados foi de: 8, 23 e 26; com parâmetros $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$, conforme as figuras 7.20, 7.21 e 7.22.

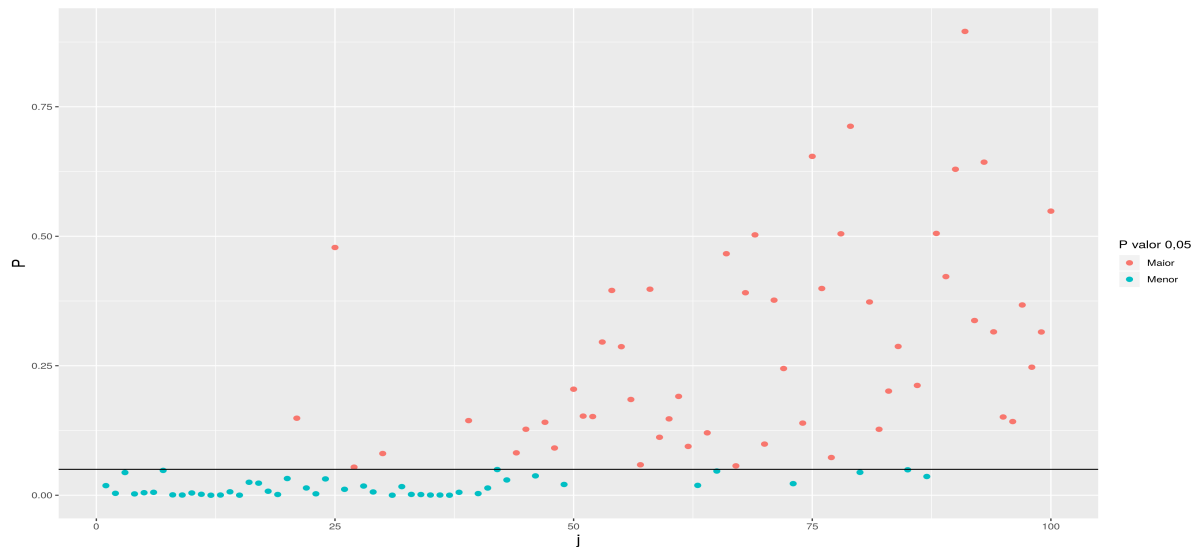
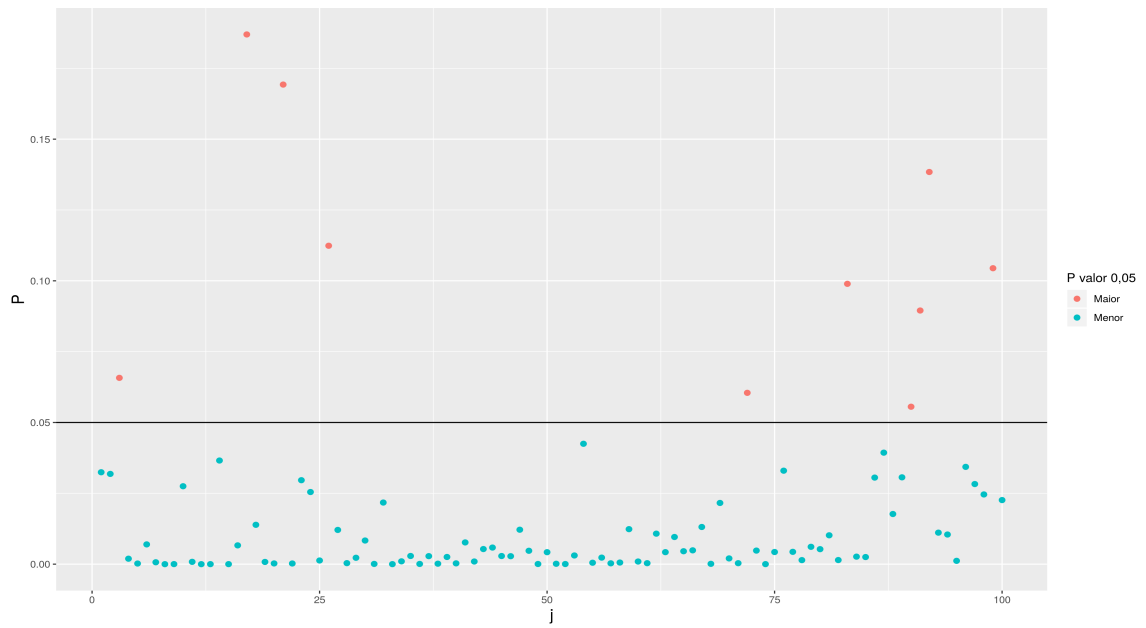
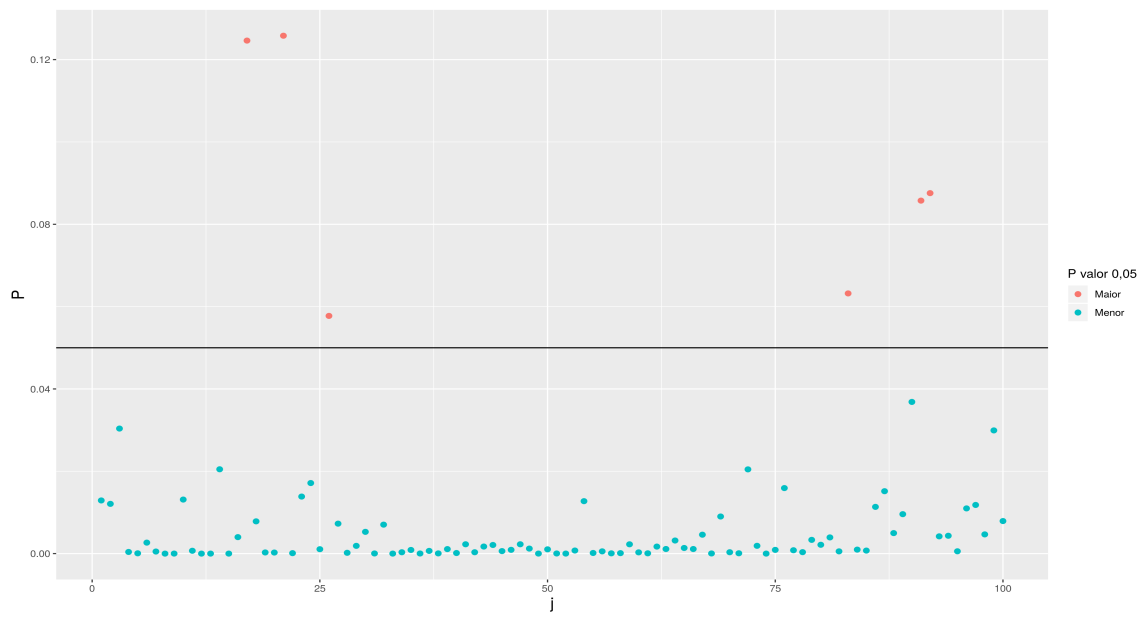
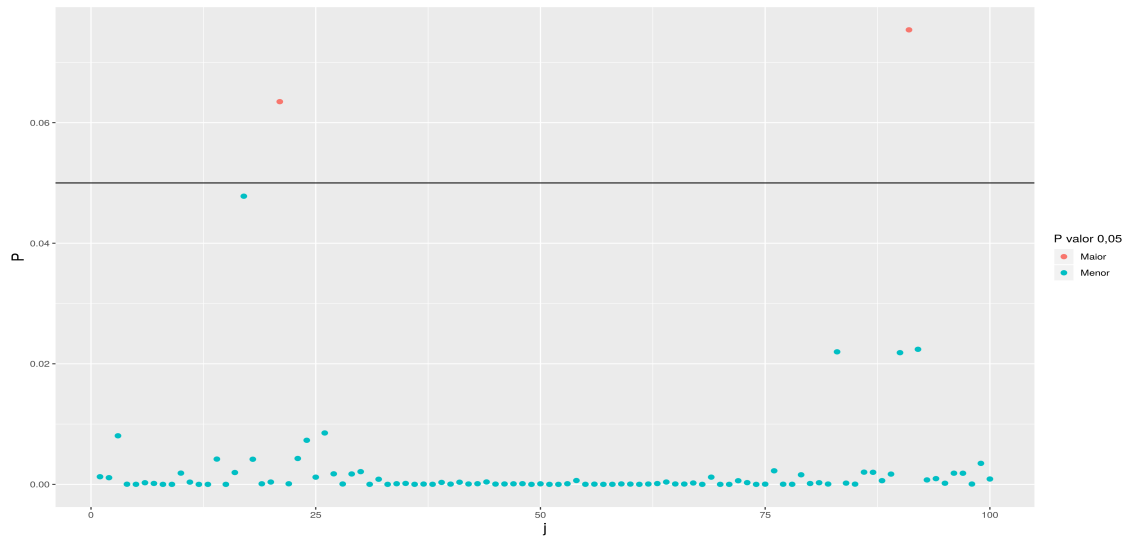




Figura 7.20 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade simples para 8 clãs, com $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$.



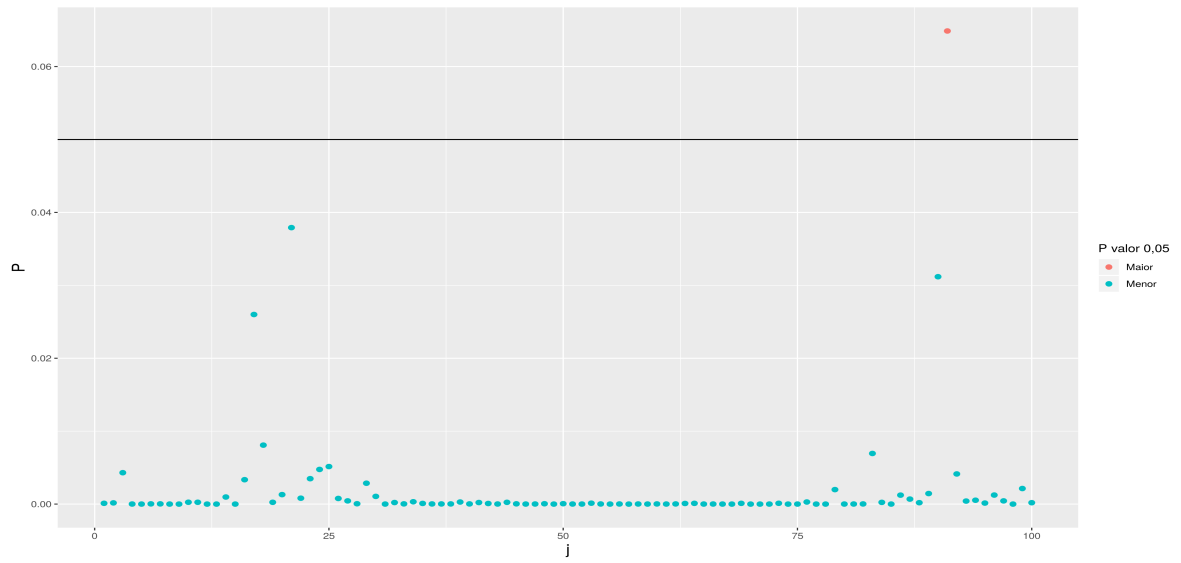
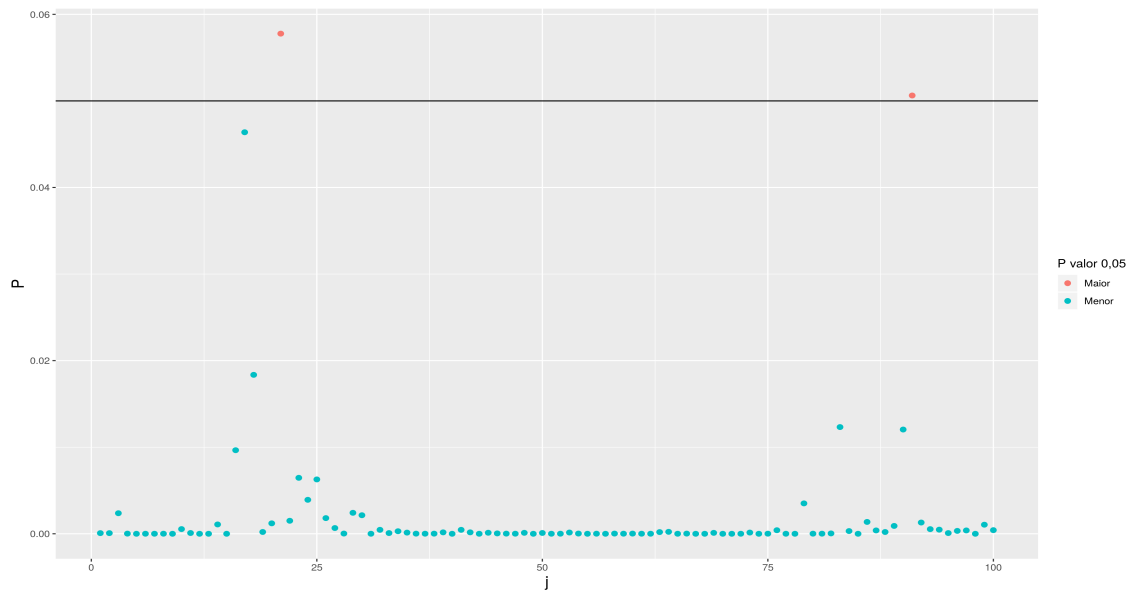


Figura 7.21 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade simples para 23 clãs, com $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$.



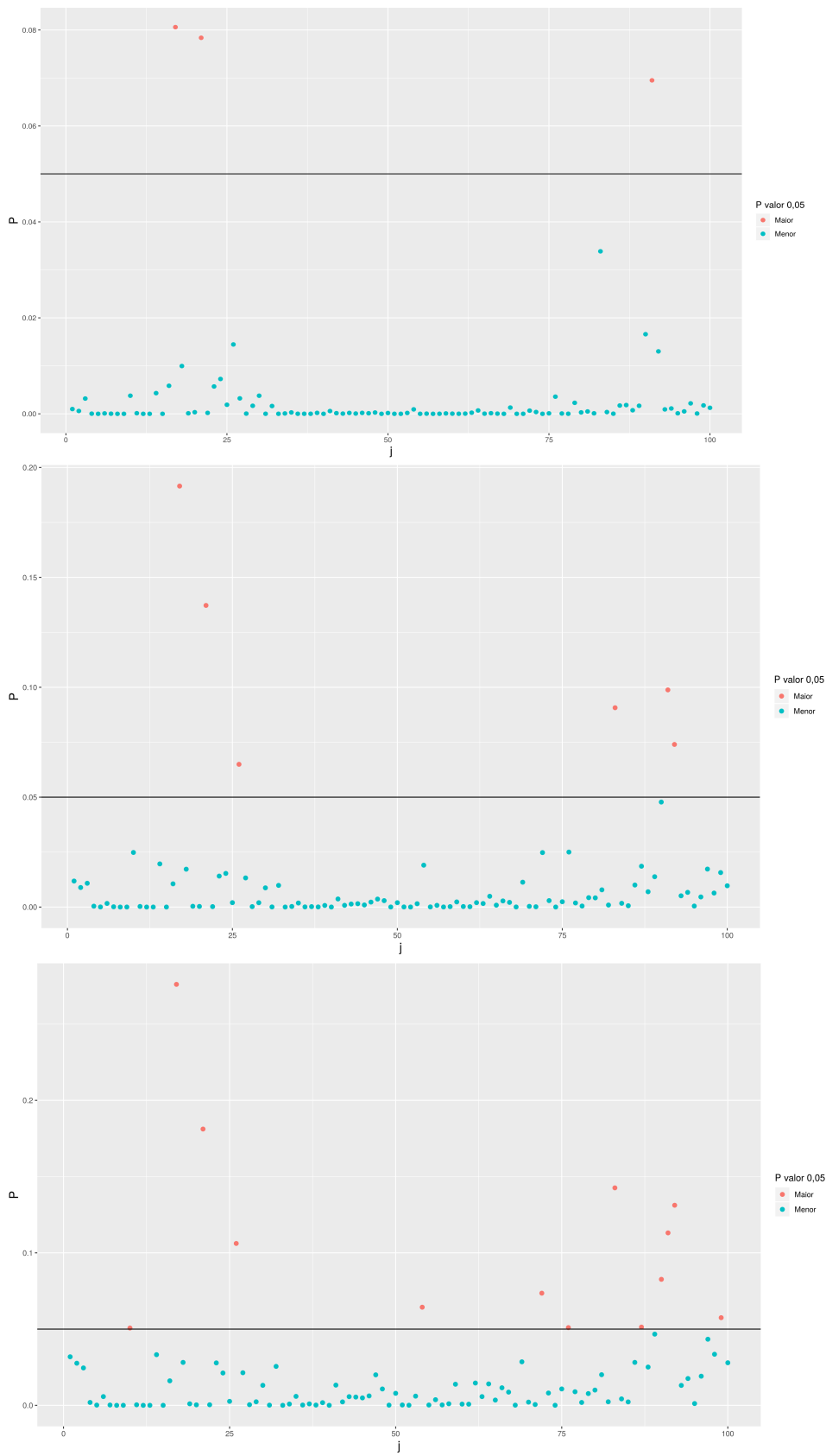
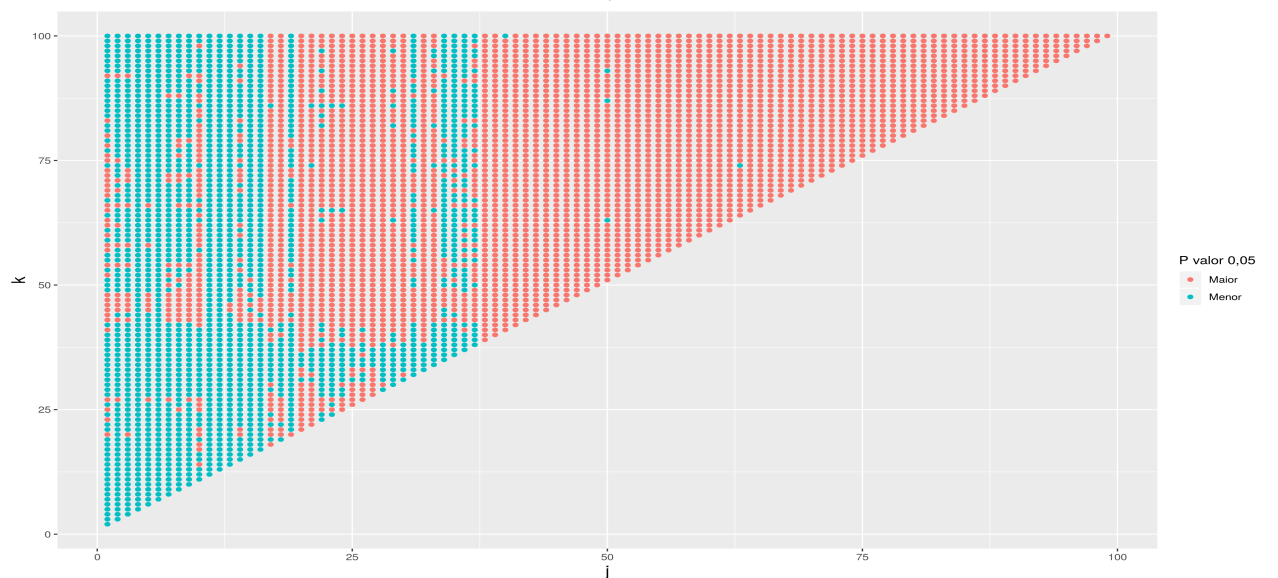
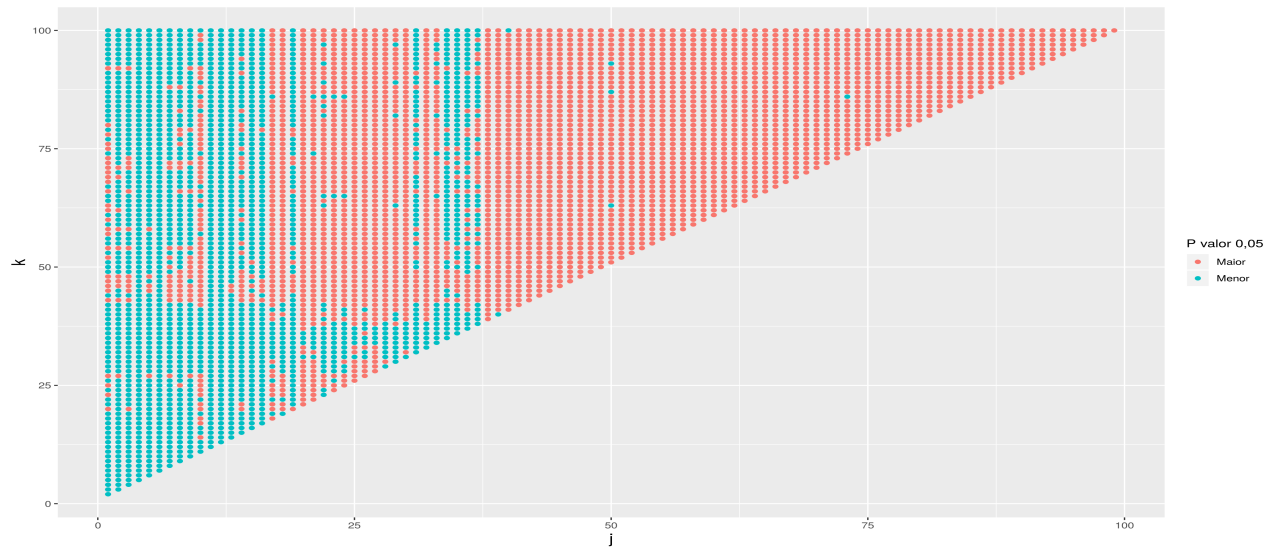


Figura 7.22 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade simples para 26 clãs, com $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$.

Enquanto que os resultados obtidos para o teste de ANOVA nas entropias de probabilidade conjunta são apresentados em gráficos bidimensionais, onde os pontos representam os valores dos coeficientes de p value da ANOVA e os eixos representam os pares de colunas ($j = 1$ a 99) e ($k = 2$ a 100), respeitando sempre a ordenação jk ($k \geq j + 1$). Os pontos (p value da ANOVA) são classificados em duas cores diferentes em função ao seu valor ser maior o menor ao nível de significância de 0.05 .

Para este experimento a quantidade de clãs utilizados também foi de: 8 , 23 e 26 ; com parâmetros $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$, conforme as figuras 7.23, 7.24 e 7.25.



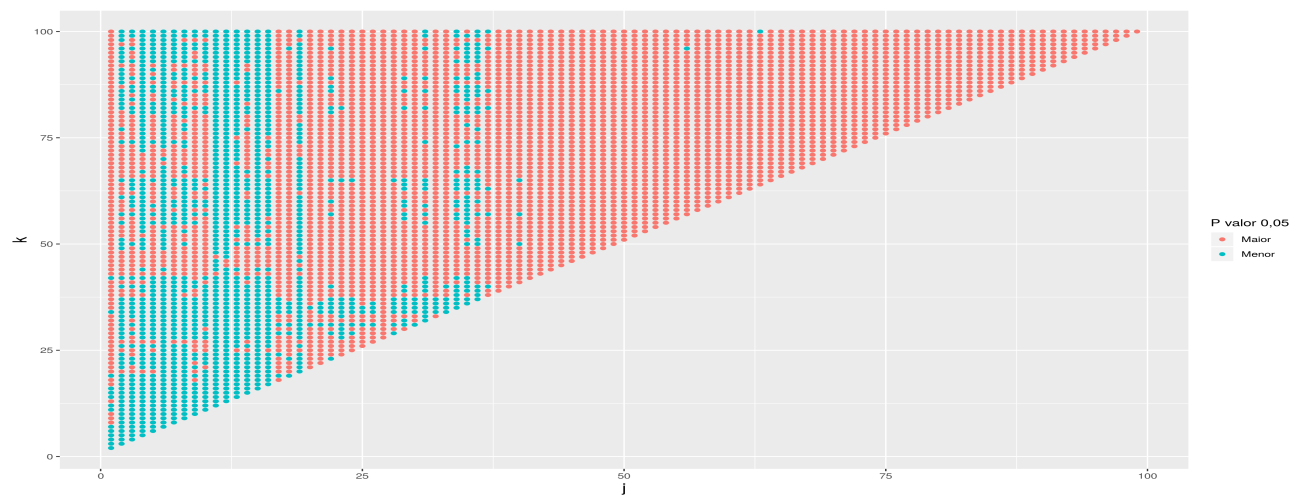
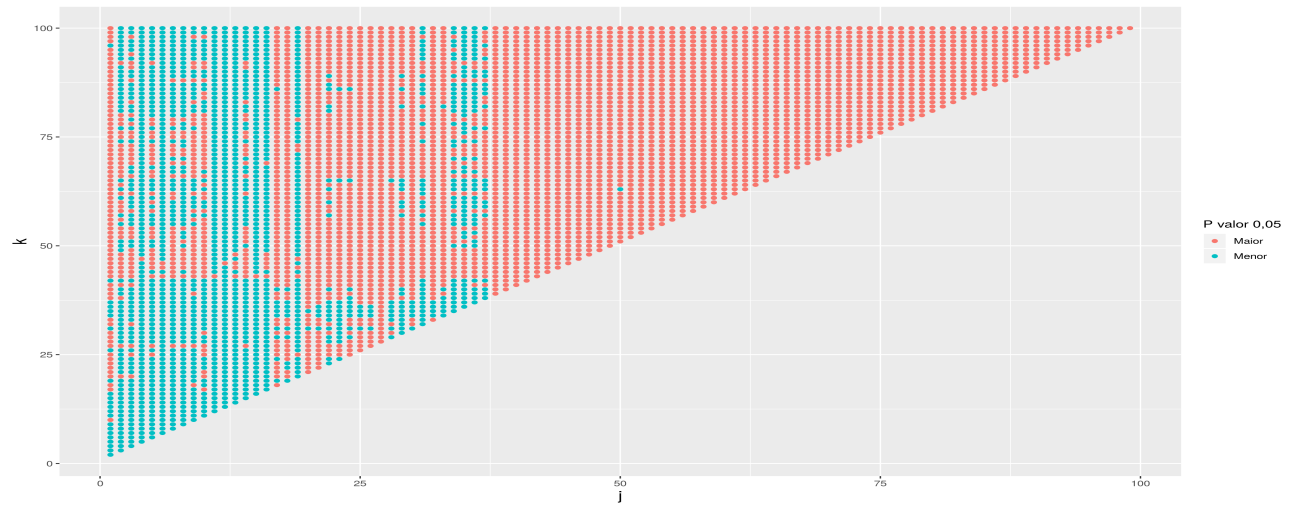
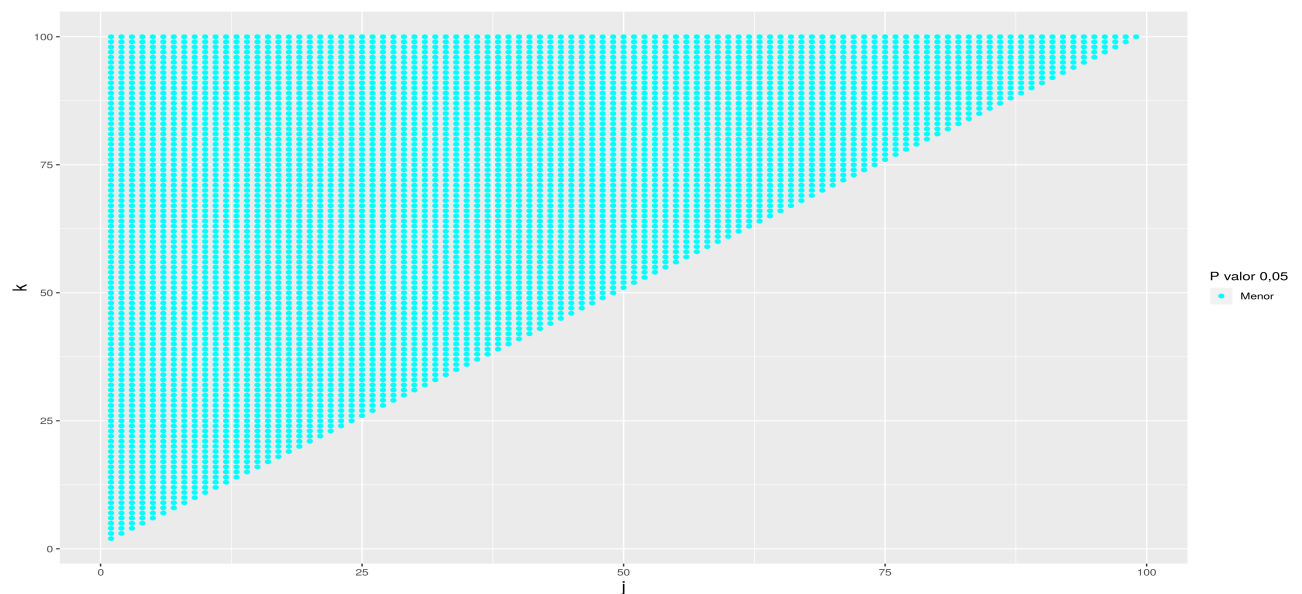


Figura 7.23 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade composta para 8 clãs, com $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$.



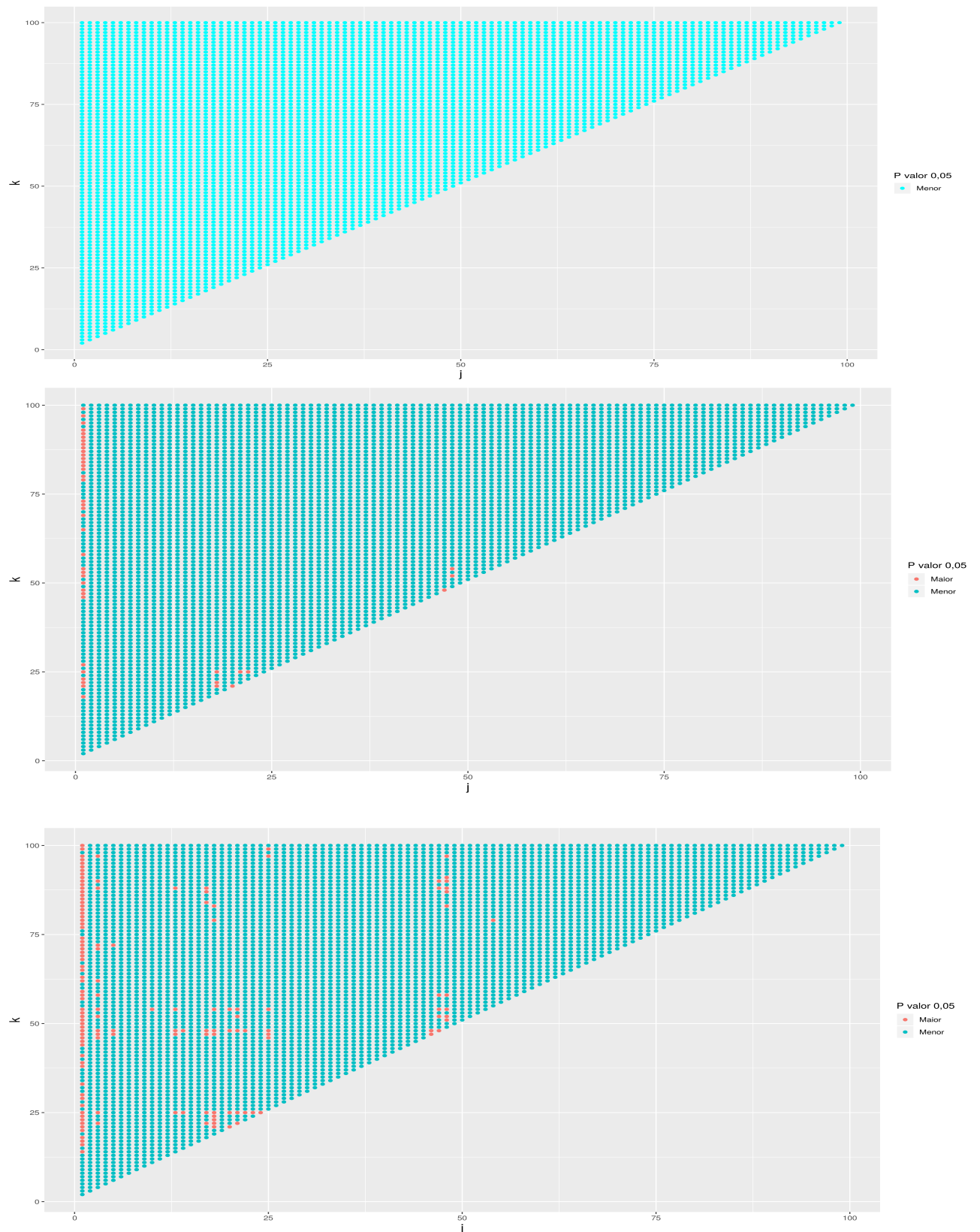
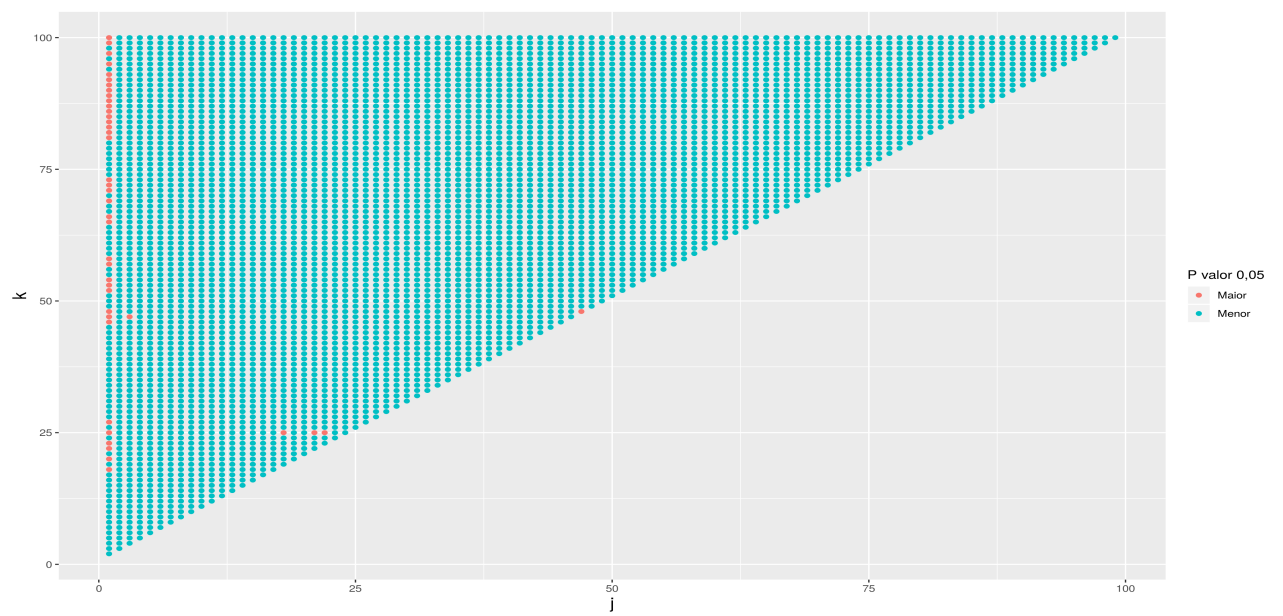
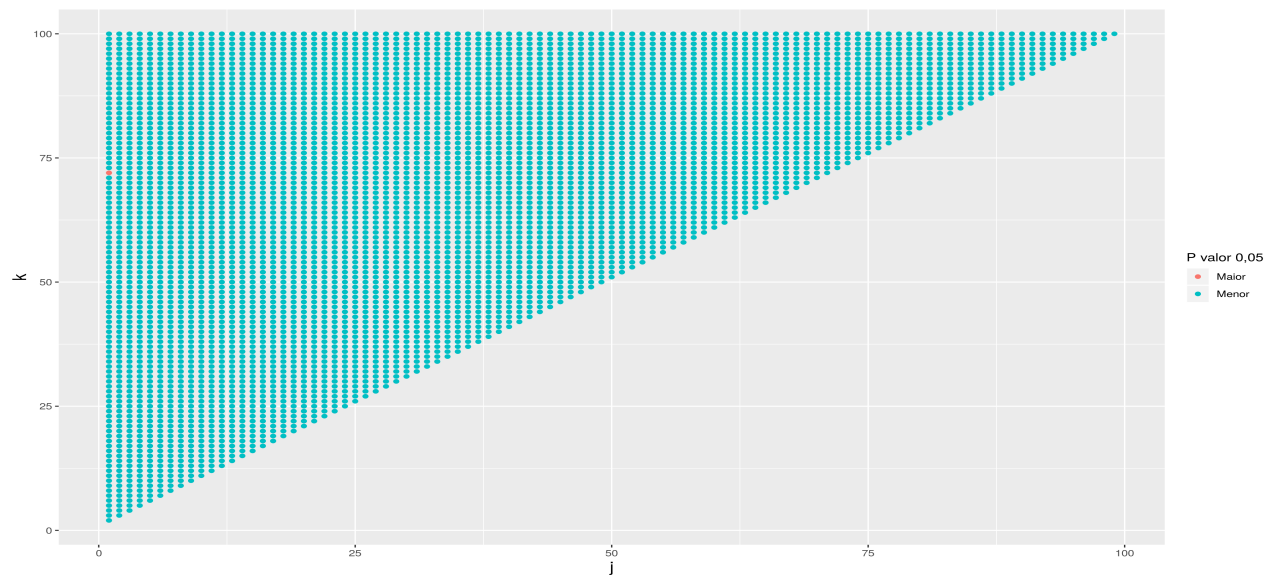
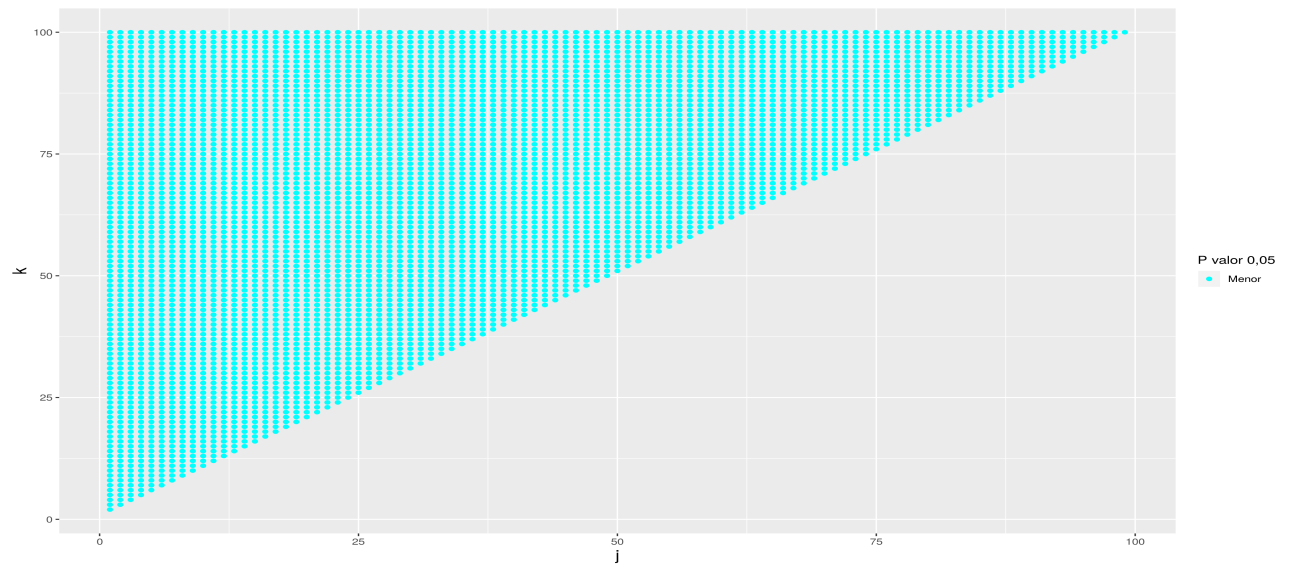


Figura 7.24 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade composta para 23 clãs, com $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$.



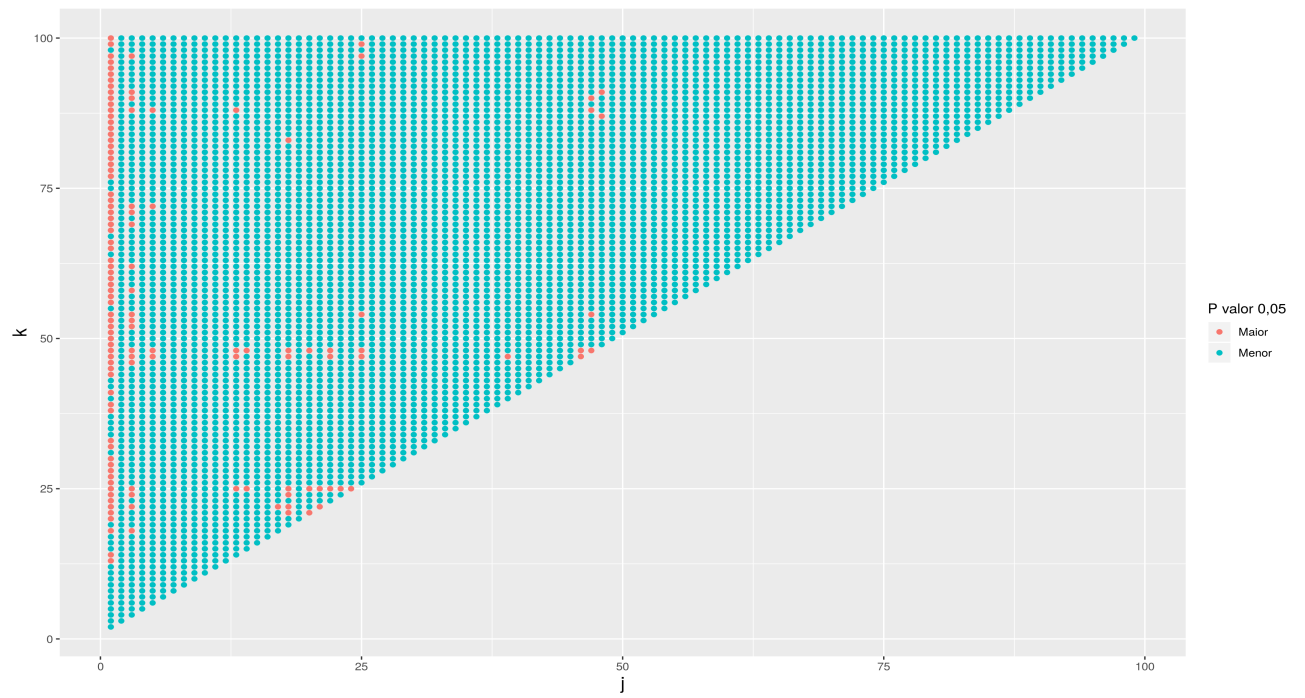


Figura 7.25 – Teste ANOVA com a entropia Harvda-Charvat de probabilidade composta para 26 clãs, com $s = 0.1$, $s = 0.5$, $s = 1.0$ (Shannon) e $s = 1.3$.

Capítulo VIII – Conclusões e Trabalhos Futuros

8.1 Introdução

O trabalho foi conclusivo em relação a existência de clãs. Porém é necessário que mais blocos de famílias sejam utilizados, além dos blocos 100 x 200 e 100 x 100. Experimentos com “recortes” de 80 x 80 e 100 x 80, por exemplo, devem ser utilizados e os resultados comparados com os obtidos pelos blocos calculados (100 x 200 e 100 x 100).

Outras versões do Pfam devem ser utilizadas e comparadas com a versão 27.0 utilizada neste trabalho.

É importante também avaliar o uso de outras medidas de entropia e de métodos estatísticos, que talvez possam ser mais adequados para lidar com o espaço amostral utilizado.

Além disso, os cálculos devem ser otimizados com o uso de técnicas de programação paralela e uso de uma infraestrutura de um supercomputador.

Referências Bibliográficas

ALBERTS, Bruce; JOHNSON, Alexander D.; LEWIS, Julian; MORGAN, David; RAFF, Martin; ROBERTS, Keith; WALTER, Peter. Molecular Biology of the Cell. W. W. Norton & Company. Sixth Edition, November, 2014.

ALURU, Srinivas. Handbook of Computacional Molecular Biology. Chapman and Hall , 2006.

BATSCHULET, Edward. Introdução à matemática para Biocientistas. São Paulo: Ed. da Universidade de São Paulo, 1978.

BAXEVANIS, Andrea; OUELLETTE, Francis. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Wiley-Interscience, 1998.

BROWN, T.A. Genomes. Wiley-Liss Bros., 1999.

BUSCH, Jorge R.; FERRARI, Pablo A.; FLESIA, Ana G.; FRAIMAN, Ricardo; GRYNBERG, Sebastian P.; LEONARDI, Florencia. Testing Statistical Hypothesis on Random Trees and Applications to the Protein Classification Problem. Annals of Applied Statistics 2009, Vol. 3, No. 2, 542-563

CANTOR, C. and Smith, C.L. Genomics: The Science and Technology Behind the Human Genome Project. Wiley-Interscience, New York, 2000.

DEGROOT, M. H; SCHERVISH, M. J. Probability and Statistics. 4th edition. Addison-Wesley, 2012.

EWENS, Warren J.; GRANT, Gregory. Statistical Methods in Bioinformatics: An Introduction. Springer-Verlag, New York, 2005.

FARAWAY, Julian J. Linear Models with R. Second Edition, 2015.

GAIL, M. *et al.* Statistics Methods in Bioinformatics: An Introduction. Springer, Second Edition, 2005.

GRAVETTER, Frederick J.; WALLNAU, Larry B. Statistics for the Behavioral Sciences. Wadsworth, Cengage Learning, Ninth Edition, 2013.

GUSFIELD, D. Algorithms on Strings, Trees, and Sequences. Cambridge University Press, 1997.

HAVRDA, J.; CHARVAT, F. Quantification Method of Classification Process. Concept of Structural α -Entropy. Kybernetika, v. 3, n. 1, pp. 30-35, 1967,

HEGER, Andreas; HOLM, Liisa. Exhaustive Enumeration of Protein Domain Families. Elsevier Science Ltd., J. Mol. Biol., 2003.

HORTON, Nicholas J.; KLEINMAN, Ken. Using R for data management, statistical analysis, and graphics, 2011.

HUNTER, Lawrence E. Artificial Intelligence and Molecular Biology. AAAI Press; First edition. March, 1993.

JONES, Neil C.; PEVZNER, Pavel. An Introduction to Bioinformatics Algorithms. The MIT Press, 2004.

LANCIA, G. Applications to Computational Molecular Biology, Kluwer International Series in Operations Research and Management Science, Volume on Modeling for Discrete Optimization, (G. Appa and P. Williams eds), 2004.

LANDSBERG, P. T.; VEDRAL, V. Distributions and Channel Capacities in Generalized Statistical Mechanics. Phys. Lett. A, v. 224, pp. 326-330, 1997.

LEWIN, Benjamin. Genes VIII. Prentice Hall, 2003.

LIPMAN, David J.; PEARSON, William R. "Rapid and sensitive protein similarity searches". Science, 1985. 227 (4693): 1435–41.

LODISH, Harvey; BERK, Arnold; ZIPURSKY, S. Lawrence; PAUL, Matsudaira; BALTIMORE, David; DARNELL, James. Biologia Celular e Molecular. Rio de Janeiro: Revinter, 2002.

MEYER, Paul L. Introductory Probability and Statistical Applications. Addison-Wesley Publishing Company, Second Edition, 1970.

MONTGOMERY, Douglas C. Design and analysis of experiments. Eighth edition, 2012.

PEVZNER, P. Computational Molecular Biology - An Algorithmic Approach. MIT Press, 2000.

Protein Family Database (PFAM). <https://pfam.xfam.org/>. Acessado em: 05/03/2018.

PUNTA, Marco *et al.* The Pfam Protein Families Database. Nucleic Acids Research, Vol. 40, 2012.

SHANNON, C. E. The Mathematical Theory of Communication. Bell Syst, Tech. Journal, v. 27, pp. 379-423; 623-656, 1948.

RASHIDI, Hooman, H; BUEHLER, Lukas, K. Applications in Biological Science and Medicine. CRC Press LLC, 2000.

RÉNYI, A. On Measures of Entropy and Information. In: Proc. Fourth Berkely Symp. Math. Statist. and Probability, v. 1, pp 547-561. University of California Press Berkely, 1961.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

RStudio (2018). RStudio: Integrated development environment for R (Version 1.1.453) [Computer software]. Boston, MA.

SILVA, Fabricio A. B. da. Big Data e Nuvens Computacionais: Aplicações em Saúde Pública e Genômica. JHI, Journal of Health Informatics, 2016.

TAMARIN, Robert H. Principles of Genetics. Brown Publishers, fifth edition, 1996.

TOPSE, Flemming. Jensen-Shannon Divergence and norm-based measures of Discrimination and Variation. Departmente of Mathematics, University of Copenhagen, 2007.

WATERMAN, M.S. Introduction to Computational Biology, Chapman and Hall, 1995.

WATSON, J.D. and F.H.C. Crick. Molecular structure of nucleic acid. A structure for deoxyribose nucleic acid. Nature 171:737-738, 1953.

ZHANG, M.Q. Statistical features of human exons and their flanking regions. Hum. Mol. Genet., 7, 919–932, 1998.