Fortalecimentos de representações de algumas sequências permutáveis

Luís Gustavo Esteves

Tese apresentada ao
Instituto de Matemática e Estatística da
Universidade de São Paulo
para a obtenção do grau de
Doutor em Estatística

Área de concentração: Estatística
Orientador: Prof. Dr. Sergio Wechsler

Durante a elaboração deste trabalho, o autor recebeu apoio financeiro do CNPq

- São Paulo, Dezembro de 2001 -

Este exemplar corresponde à redação final

da dissertação devidamente corrigida e

defendida por Luís Gustavo Esteves

e aprovada pela comissão julgadora.

São Paulo, 11 de Dezembro de 2001.

Banca examinadora:

- Prof. Dr. Sergio Wechsler (Orientador) - IME - USP
- Profa. Dr. Pilar Loreto Iglesias Zuazola - PUC - Chile
- Prof. Dr. José Galvão Leite - UFSCar
- Prof. Dr. Antônio Luiz Pereira - IME - USP
- Prof. Dr. Dani Gammerman - IM - UFRJ

Agradecimentos

Ao longo destes quatro anos do programa de doutorado, não faltaram dificuldades a serem superadas. No entanto, acredito que alguns destes obstáculos foram satisfatoriamente vencidos, em decorrência de muito esforço e dedicação e, em grande parte, da generosidade, paciência e dedicação de muitas pessoas com as quais convivi neste período. A estas pessoas, expresso nesta página meu modesto agradecimento pelas suas valorosas contribuições.

Inicialmente, agradeço a minha família: a meus pais pela generosidade, carinho, dedicação,..., enfim, meus pais; à Dê e ao Má pela amizade e acolhimento, e ao Lucca, meu "chapinha", e sua doce inocência.

Agradeço também aos meus amigos de sempre, Marcão, P. A., Adriana, Patrícia e Elis, pelo convívio de tantos anos e pela sincera torcida pelo meu êxito.

Agradeço aos vários amigos do IME-USP: à galera da graduação e do mestrado pelo incentivo e à turma de 1998 do doutorado, a qual foi fundamental para que tivesse uma jornada mais amena. Em especial, agradeço ao Celso (vascaíno), à Verónica, ao Simon, à Denise e ao Osvaldo pelos bons papos, ao Esteban, à Daisy e à Neide, sempre dispostos a oferecer um ombro amigo nos momentos de maior dificuldade, e ao André, Matoso e Paulo Maranhão pelas boas (boas?!?) raquetadas.

Agradeço aos professores do IME-USP em geral pela formação que me propiciaram. Em especial, meu muito obrigado aos professores Adilson, Caio, Carlinhos, Cláudia, Fábio e Nélson pela constante motivação, ao Luiz Renato pelo trabalho de orientação inicial e aos Corinthianíssimos Toninho e Galvão pela minuciosa leitura de toda a tese e pela ajuda essencial na compreensão de muitos resultados que figuram no trabalho. Agradeço ainda às secretárias do departamento de estatística do IME e ao pessoal da CPG pela prontidão em diversos momentos. Muito obrigado também ao Válter por todo o trabalho de digitação da tese.

Por fim, um agradecimento muito especial à Pilar pela sugestão dos problemas abordados neste trabalho e sua infindável disposição para estudá-los, e ao "guru", professor Sérgio Wechsler, pela amizade, generosidade e paciência ao longo de todos estes anos, pela orientação do trabalho e pelas agradáveis e intermináveis apreciações sobre o alvi-negro.

Resumo

Neste trabalho, consideramos três problemas de caracterização de sequências de variáveis aleatórias (ou vetores) permutáveis. Apresentamos, primeiramente, teoremas do tipo de De Finetti (formas finita e infinita) para algumas distribuições uniformes bivariadas dependentes de um único parâmetro. Estes resultados fornecem uma representação preditivista, no sentido de De Finetti, para os modelos paramétricos bivariados. Na sequência, caracterizamos vetores aleatórios cujas densidades de probabilidade podem ser expressas como função do mínimo e do máximo ou do máximo das somas das componentes. Provamos que independência e específicas condições de simetria caracterizam modelos uniformes conhecidos (discretos e contínuos, univariados e bivariados). Finalmente, introduzimos uma nova versão do modelo de urna de Pólya-Eggenberger, considerando uma distribuição de probabilidade para a composição inicial da urna, isto é, para os números de bolas brancas e pretas inicialmente na urna. Determinamos também quando um processo permutável com valores em $\{0,1\}^{\infty}$ pode ser bem aproximado por um adequado processo de Pólya com configuração inicial desconhecida.

Abstract

In this work, we consider three problems on characterizations of exchangeable sequences of random variables (vectors). We first present De Finetti's type theorems (both finite and infinite forms) for certain bivariate uniform distributions which depend on a single parameter. These results provide a predictivistic representation, in De Finetti's sense, for such bivariate parametric models. Extensions of these results in some directions are also discussed. In the sequel, we characterize random vectors the densities of which can be expressed as a function of the minimum and maximum or of the maximum of sums of components. We show that independence and suitable conditions of symmetry characterize well-known uniform models (discrete and continuous, univariate and bivariate). Finally, we introduce a new version of Pólya's urn scheme via the consideration of a probability distribution for the initial composition, that is, for the numbers of black and white balls initially in the urn. We also determine when an exchangeable process taking values in $\{0,1\}^\infty$ may be well aproximated by a suitable Pólya process with unknown initial configuration.

Sumário

Introdução 1

1 Teoremas do tipo de De Finetti para modelos uniformes bivariados 3
1.1 Introdução 3
1.2 Distribuição Uniforme no Triângulo isósceles 5
1.2.1 Caso discreto 5
1.2.2 Caso contínuo 19
1.3 Distribuição Uniforme no Círculo 34
1.4 Conclusões 47

2 Caracterização de variáveis aleatórias uniformes independentes e identicamente distribuídas 50
2.1 Introdução 50
2.2 Distribuições uniformes discretas 52
2.2.1 Caso univariado 52
2.2.2 Caso bivariado 64
2.3 Distribuições uniformes contínuas 73
2.3.1 Caso univariado 73
2.3.2 Caso bivariado 83
2.4 Conclusões 89

3 Uma variante do modelo de urna de Pólya-Eggenberger 91
3.1 Introdução 91
3.2 O modelo de Pólya-Eggenberger e suas variações 92
3.3 Resultado principal 99
3.4 Conclusões 112

Conclusão 115

Introdução

Neste trabalho, abordamos três problemas de caracterização de sequências (finitas e infinitas) de variáveis aleatórias (ou vetores) permutáveis. Mais precisamente, procedemos à derivação de representações preditivistas de modelos uniformes bivariados dependendo de um único parâmetro, à caracterização de modelos uniformes (univariados e bivariados) dentro de classes de distribuições multivariadas satisfazendo específicas condições de simetria e à determinação de condições suficientes para que um processo permutável com valores em $\{0,1\}^\infty$ possa ser bem aproximado, em um sentido preciso, por uma mistura de processos de Pólya-Eggenberger na composição inicial da urna, respectivamente nos capítulos 1, 2 e 3. Apresentamos agora uma breve descrição do conteúdo de cada capítulo.

No capítulo 1, apresentamos representações preditivistas para modelos uniformes bivariados dependendo de um único parâmetro. Conforme veremos, a representação preditivista de um modelo paramétrico, uma caracterização deste modelo baseada exclusivamente em julgamentos probabilísticos sobre quantidades observáveis, possui um caráter fundamental para a escola Bayesiana preditivista de inferência estatística, pois, entre outras razões, fornece uma justificativa subjetivista para a adoção de tal modelo paramétrico como distribuição amostral (condicional) dos dados gerados em experimentos aleatórios. Após 1937, quando Bruno de Finetti estabeleceu o famoso Teorema da Representação para variáveis aleatórias permutáveis e obteve, em particular, uma representação preditivista para o modelo Bernoulli, vários modelos paramétricos tiveram suas representações preditivistas determinadas. No entanto, modelos multivariados mais gerais não pertencentes a famílias exponenciais foram pouco contemplados na literatura e, deste modo, optamos por focar alguns destes modelos no capítulo 1. Especificamente, consideramos os modelos uniformes sobre $\{(x,y) \in \mathbb{N}^2 : x+y \leq \theta\}$, $\theta \in \mathbb{N}$, $\{(x,y) \in \mathbb{R}_+^2 : x+y \leq \theta\}$, $\theta > 0$, e $\{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq \theta^2\}$, $\theta > 0$ (verificamos que os resultados obtidos podem ser estendidos, de maneira bem simples, para os correspondentes modelos d-variados, $d > 2$). Finalmente, apresentamos alguns problemas decorrentes da introdução de novo(s) parâmetro(s) nestes modelos para a obtenção de suas respectivas representações preditivistas.

No capítulo 2, identificamos modelos uniformes conhecidos dentro de específicas classes de distribuições simétricas multivariadas. Caracterizações de distribuições simétricas

multivariadas têm sido desenvolvidas por diversos autores nas últimas décadas tendo em vista não apenas a extensão de distribuições univariadas conhecidas para as suas respectivas versões multivariadas, mas também a consideração de novos modelos, que não o normal, para fins de inferência. Em particular, as distribuições esféricas multivariadas e generalizações em norma l_p, $p \neq 2$, têm sido constantemente investigadas. No entanto, distribuições simétricas multivariadas cujas funções de densidade de probabilidade dependem, segundo Iglesias et al. (1998), da norma l_∞ apenas recentemente foram contempladas na literatura. Deste modo, no capítulo 2, fazemos uma apreciação das classes de distribuições multivariadas cujas funções de densidade de probabilidade dependem do máximo das coordenadas ou do mínimo e do máximo das coordenadas ou da máxima soma das coordenadas em cada componente do vetor, nas versões discreta e contínua. Para cada uma das classes de distribuições consideradas, a suposição adicional de independência fornece uma caracterização de um modelo uniforme conhecido.

Finalmente, no capítulo 3, apresentamos uma nova versão para o modelo de urna de Pólya-Eggenberger, com a introdução de uma distribuição de probabilidade para a composição inicial da urna (o modelo de Pólya-Eggenberger ocupa uma posição central dentre os diversos modelos de urna desenvolvidos para representar formas de contágio). Conforme veremos, este novo processo é permutável, diferentemente de outras variações do modelo de Pólya-Eggenberger já consideradas na literatura. Exibimos também uma condição suficiente para que um processo permutável com valores em $\{0,1\}^\infty$ possa ser bem aproximado por um conveniente modelo de urna de Pólya-Eggenberger com configuração inicial aleatória, isto é, por uma mistura de processos de Pólya na composição inicial da urna.

Capítulo 1

Teoremas do tipo de De Finetti para modelos uniformes bivariados

1.1 Introdução

Em 1937, Bruno de Finetti estabeleceu o famoso Teorema da Representação para variáveis aleatórias permutáveis assumindo valores no conjunto $\{0, 1\}$. Este resultado, não obstante sua formulação matemática relativamente simples, incorpora boa parte do ideário sustentado pela escola bayesiana preditivista em inferência estatística (De Finetti [1937]). Dentre as várias idéias que podemos abstrair do Teorema da Representação de De Finetti, destacamos a justificativa, sob o ponto de vista bayesiano, da adoção do modelo Bernoulli, por parte de um estatístico, como distribuição amostral (condicional) para dados binários gerados a partir de um experimento aleatório para fins de inferência. Esta nova forma de caracterização de modelos paramétricos através da expressão de julgamentos probabilísticos sobre quantidades, segundo De Finetti, consideradas observáveis recebe o nome de Representação Preditivista.

A partir de Bruno de Finetti, muitos autores dedicaram-se ao estudo de representações preditivistas, uma vez que tais representações fundamentariam a adoção de modelos paramétricos mais gerais em procedimentos de inferência estatística. Dentre outros, des-

tacamos os trabalhos de Freedman e Diaconis (1980) que obtiveram uma representação preditivista para Cadeias de Markov recorrentes com espaço de estados discreto através da introdução de um conceito de permutabilidade parcial, de Dawid (1977) e Lauritzen (1988) que desenvolveram teoremas do tipo de De Finetti para modelos lineares multivariados considerando processos invariantes em grupos de transformações ortogonais, de Freedman e Diaconis (1990) que derivaram representações preditivistas para modelos gerais pertencentes às famílias exponenciais multivariadas, e de Gnedin (1995) e Iglesias, Pereira e Tanaka (1998) que obtiveram resultados para famílias não-exponenciais (em particular, para distribuições uniformes univariadas). Diaconis e Freedman (1984) e Fortini, Ladelli e Regazzini (2000) também apresentaram resultados gerais relacionados com a noção de suficiência preditiva, que fornece uma forma infinita de representação para caracterizar a distribuição uniforme em $(0, \theta)$, $\theta > 0$. Uma descrição bastante detalhada do trabalho desenvolvido na área de representações preditivistas para modelos paramétricos é apresentada em Iglesias (1993).

Apesar do grande avanço alcançado nesta área de pesquisa nas últimas décadas, vários modelos paramétricos multivariados que não pertencem à família exponencial ainda não tiveram suas representações preditivistas determinadas, ao menos utilizando-se estritamente condições probabilísticas sobre quantidades observáveis (apenas recentemente Arellano et al. (1994) obtiveram uma caracterização preditivista para a distribuição t multivariada, Bolfarine et al. (1999) estabeleceram resultados similares para os modelos com erro tipo II de Pearson nas variáveis e Iglesias et al. (1998) caracterizaram distribuições esféricas multivariadas na norma l_∞).

Dentro deste contexto, propomos, neste capítulo, teoremas do tipo de De Finetti para alguns modelos paramétricos bivariados, bem como suas respectivas extensões para dimensão $d > 2$. Mais precisamente, derivamos as representações preditivistas para os

seguintes modelos: (A) distribuição uniforme no triângulo isósceles (versões discreta e contínua) e (B) distribuição uniforme no círculo (as descrições destes modelos são encontradas nas seções 2 e 3, respectivamente).

O restante deste capítulo está organizado em três seções: na seção 2 são apresentadas as representações preditivistas para as versões discreta e contínua do modelo (A); na seção 3, teoremas do tipo de De Finetti são derivados para o modelo (B); finalmente, na seção 4, apontamos os problemas originados para a obtenção de representações preditivistas para os modelos (A) e (B) quando novos parâmetros são incorporados a estes modelos.

Ao longo deste capítulo, adotaremos as seguintes notações. Para um conjunto A não-vazio, denotaremos por $\mathcal{P}(A)$ o conjunto das partes de A e por A^N o N-ésimo produto cartesiano de A, $N \in \mathbb{N}$. Para $A \subset \mathbb{R}$, A_+ denotará o conjunto $\{a \in A : a \geq 0\}$. Representaremos a σ-álgebra de Borel em $\mathbb{R}^N$ por $\mathcal{B}(\mathbb{R}^N)$, a medida de Lebesgue em $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ por λ_N, $N \in \mathbb{N}$ e a medida de Lebesgue em $((\mathbb{R}^k)^N, \mathcal{B}((\mathbb{R}^k)^N))$ por $\lambda_k^{(N)}$. Finalmente, se P e Q são duas medidas definidas num mesmo espaço mensurável, então $P \ll Q$ indicará que P é absolutamente contínua em relação a Q e $\|P - Q\|$ denotará a distância em variação total entre P e Q.

1.2 Distribuição Uniforme no Triângulo isósceles

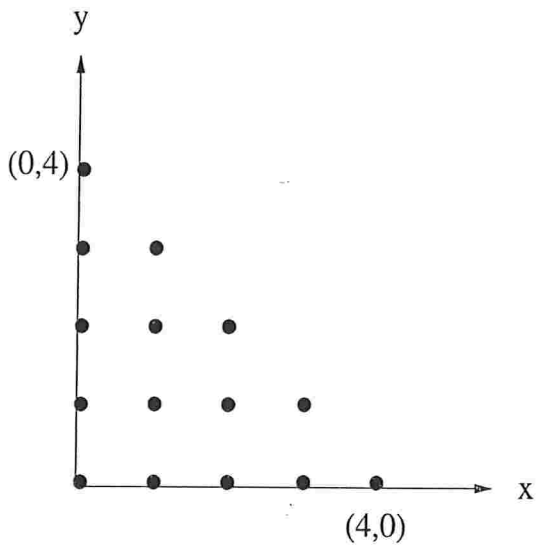
Nesta seção, derivamos teoremas do tipo de De Finetti para a distribuição uniforme sobre o triângulo isósceles (versões discreta e contínua).

1.2.1 Caso discreto

Vamos considerar a distribuição uniforme sobre os conjuntos $\{(x, y) \in \mathbb{N}^2 : x + y \leq \theta\}$, $\theta \in \mathbb{N}$. Em termos de distribuição de probabilidade, este modelo tem a seguinte caracterização:

$$\text{Prob}\{X = x, Y = y \mid \theta\} = \frac{2}{(\theta+1)(\theta+2)} \mathbb{I}_{\{(a,b)\in\mathbb{N}^2 : a+b\leq\theta\}}(x,y) \tag{1.1}$$

O suporte desta distribuição é esboçado abaixo para $\theta = 4$.



Para a obtenção da representação preditivista deste modelo, procederemos aos seguintes passos: primeiro, consideraremos uma distribuição uniforme sobre uma específica região de $(\mathbb{N}^2)^N$ determinada por uma estatística suficiente para o modelo em (1.1). Mais precisamente, esta região corresponderá à imagem inversa de $\{m\}$ gerada pela estatística suficiente mencionada acima, conforme veremos na sequência (salientamos que a construção de tal distribuição uniforme é simples neste caso, uma vez que estamos considerando um modelo discreto; para a versão contínua deste modelo, será necessária uma construção mais sofisticada). Em seguida, obteremos a marginal n-dimensional desta distribuição uniforme, $n < N$, e estimaremos sua distância em variação total relativamente à distribuição de n vetores aleatórios bidimensionais independentes e identicamente distribuídos com lei dada por (1.1). A forma finita do teorema do tipo de De Finetti seguirá diretamente desta estimação. Para obter sua forma infinita, estudaremos a convergência

de medidas de probabilidade específicas em um espaço métrico conveniente (a distância em variação total sendo sua métrica) utilizando a estimação supracitada. Mais precisamente, avaliaremos a convergência da sequência de medidas relativa a sequência de estatísticas suficientes consideradas.

Consideremos, então, o seguinte subconjunto de $(\mathbb{N}^2)^N$, $\chi_m^N = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{N}^2)^N : \max_{1 \leq i \leq N}\{x_i + y_i\} = m\}$, $m \in \mathbb{N}$. Denotemos por Q_{Nm} a distribuição uniforme sobre χ_m^N, isto é

$$Q_{Nm}(\{((x_1, y_1), \ldots, (x_N, y_N))\}) = \frac{\mathbb{I}_{\{m\}}\left(\max_{1 \leq i \leq N}\{x_i + y_i\}\right)}{\nu(\chi_m^N)}, \tag{1.2}$$

onde ν é a medida de contagem em $((\mathbb{N}^2)^N, \mathcal{P}((\mathbb{N}^2)^N))$.

Notemos que considerando N vetores aleatórios (X_1, Y_1), ..., (X_N, Y_N) que, dado θ, são condicionalmente independentes e com distribuição comum dada por (1.1), temos que $\max_{1 \leq i \leq N}\{X_i + Y_i\}$ é uma estatística suficiente mínima para este modelo. Neste sentido, dizemos que Q_{Nm} é uma distribuição uniforme sobre a imagem inversa de $\{m\}$ gerada pela estatística suficiente para o modelo (A).

Verifiquemos que $\nu(\chi_m^N) = \binom{m+2}{2}^N - \binom{m+1}{2}^N$.

Seja $M_K^{Nm} = \left\{((x_1, y_1), \ldots, (x_N, y_N)) \in \chi_m^N : \sum_{i=1}^N \mathbb{I}_{\{m\}}(x_i + y_i) = K\right\}$, $K = 1, \ldots, N$.

É claro que $\chi_m^N = \dot{\cup}_{1 \leq K \leq N} M_K^{Nm}$ e que $\nu(M_K^{Nm}) = \binom{N}{K}(m+1)^K \left[\frac{m(m+1)}{2}\right]^{N-K}$. (Para obter $\nu(M_K^{Nm})$, devemos escolher K pares (x_i, y_i) que somem m. Para os demais $N - K$ pares, existem $m(m+1)/2$ combinações cujas somas são estritamente menores que m.)

Assim,

$$\nu(\chi_m^N) = \nu\left(\dot{\cup}_{1 \leq K \leq N} M_K^{Nm}\right) = \sum_{K=1}^N \nu(M_K^{Nm}) = \sum_{K=1}^N \binom{N}{K}(m+1)^K \left[\frac{m(m+1)}{2}\right]^{N-K} =$$

$$= \left[\frac{(m+1)(m+2)}{2}\right]^N - \left[\frac{m(m+1)}{2}\right]^N = \binom{m+2}{2}^N - \binom{m+1}{2}^N.$$

Substituindo a expressão acima em (1.2), resulta que

$$Q_{Nm}(\{((x_1,y_1),\ldots,(x_N,y_N))\}) = \frac{\mathbb{I}_{\{m\}}\left(\max_{1\leq K\leq N}\{x_i+y_i\}\right)}{\binom{m+2}{2}^N - \binom{m+1}{2}^N} \tag{1.3}$$

Notemos que se $((X_1,Y_1),\ldots,(X_N,Y_N))$ tem distribuição dada por (1.3), então a sequência é permutável.

Destacamos ainda que se $((X_1,Y_1),\ldots,(X_N,Y_N))$ é uma sequência de vetores aleatórios bivariados satisfazendo

$$P(((X_1=x_1,Y_1=y_1),\ldots,(X_n=x_n,Y_n=y_n))) = \phi(\max_{1\leq i\leq n}\{x_i+y_i\})$$

para alguma função não-negativa apropriada ϕ, então a lei de probabilidade de $((X_1,Y_1),\ldots,(X_N,Y_N))$ pode ser expressa como uma mistura, no sentido usual, dos elementos da família $\{Q_{Nm} : m\in\mathbb{N}\}$, conforme veremos no próximo capítulo.

Derivamos agora a distribuição de $((X_1,Y_1),\ldots,(X_n,Y_n))$, $n<N$, a marginal n-dimensional de $((X_1,Y_1),\ldots,(X_N,Y_N))$, a qual denotamos por Q_{Nmn}. Para $((x_1,y_1),\ldots,(x_n,y_n))\in(\mathbb{N}^2)^n$, temos que:

$$Q_{Nmn}(\{((x_1,y_1),\ldots,(x_n,y_n))\}) = \sum_{((u_1,v_1),\ldots,(u_N,v_N))\in C_x} Q_{Nm}(\{((u_1,v_1),\ldots,(u_n,v_n))\}),$$

onde $C_x = \{((u_1,v_1),\ldots,(u_N,v_N)) \in (\mathbb{N}^2)^N : (u_i,v_i) = (x_i,y_i), i=1,\ldots,n,$ e $\max_{1\leq i\leq N}\{u_i+v_i\}=m\}$.

Então,

$$Q_{Nmn}(((x_1,y_1),\ldots,(x_n,n_n))) =$$

$$= \sum_{((u_{n+1},v_{n+1}),\ldots,(u_N,v_N))\in C} Q_{Nm}(((x_1,y_1),\ldots,(x_n,y_n),(u_{n+1},v_{n+1}),\ldots,(u_N,v_N))),$$

onde $C = \left\{((u_{n+1}, v_{n+1}), \ldots, (u_N, v_N)) \in (\mathbb{N}^2)^{N-n} : \max_{1 \leq i \leq N}\{u_i + v_i\} = m\right\}$. Temos, então, duas situações a considerar:

(i) se $\max_{1 \leq i \leq n}\{x_i + y_i\} = m$, então temos $C = \left\{((u_{n+1}, v_{n+1}), \ldots, (u_N, v_N)) \in (\mathbb{N}^2)^{N-n} : \max_{n+1 \leq i \leq N}\{u_i + v_i\} \leq m\right\}$, e, consequentemente,

$$Q_{Nmn}(\{((x_1, y_1), \ldots, (x_n, y_n))\}) = \sum_{((u_{n+1}, v_{n+1}), \ldots, (u_N, v_N)) \in C} \frac{1}{\binom{m+2}{2}^N - \binom{m+1}{2}^N} =$$

$$= \frac{\binom{m+2}{2}^{N-n}}{\binom{m+2}{2}^N - \binom{m+1}{2}^N}.$$

(ii) se $\max_{1 \leq i \leq n}\{x_i + y_i\} < m$, então $C = \left\{((u_{n+1}, v_{n+1}), \ldots, (u_N, v_N)) \in (\mathbb{N}^2)^{N-n} : \max_{n+1 \leq i \leq N}\{u_i + v_i\} = m\right\} = \chi_m^{N-n}$. Neste último caso, temos que

$$Q_{Nmn}((x_1, y_1), \ldots, (x_n, y_n)) = \sum_{((u_{n+1}, v_{n+1}), \ldots, (u_N, v_N)) \in \chi_m^{N-n}} \frac{1}{\binom{m+2}{2}^N - \binom{m+1}{2}^N} =$$

$$= \frac{\binom{m+2}{2}^{N-n} - \binom{m+1}{2}^{N-n}}{\binom{m+2}{2}^N - \binom{m+1}{2}^N}$$

De (i) e (ii), obtemos, finalmente, que

$$Q_{Nmn}((x_1, y_1), ..., (x_n, y_n)) = \frac{\binom{m+2}{2}^{N-n} - \binom{m+1}{2}^{N-n}(1 - \mathbb{I}_{\{m\}}(\max_{1 \leq i \leq n}\{x_i + y_i\}))}{\binom{m+2}{2}^N - \binom{m+1}{2}^N} \tag{1.4}$$

Continuando, estimamos a distância em variação total entre Q_{Nmn} e a distribuição de n vetores aleatórios bivariados independentes e com idêntica distribuição uniforme sobre $\{(x, y) \in \mathbb{N}^2 : x + y \leq m\}$, $m \in \mathbb{N}$, a qual denotamos por P_m^n. Notemos que, fixando-se

valores de $m, n \in \mathbb{N}$ e tomando-se limites em (1.4) quando $N \to \infty$, a sequência de medidas de probabilidade $\{Q_{Nmn} : N \in \mathbb{N}\}$ converge para P_m^n em quase toda parte.

Teorema 1.2.1 *Sejam Q_{Nmn} e P_m^n as medidas de probabilidade em $((\mathbb{N}^2)^n, \mathcal{P}((\mathbb{N}^2)^n))$ definidas anteriormente. Então, para $1 \leq n < N$,*

$$\|Q_{Nmn} - P_m^n\| \leq \frac{2n}{N}, \tag{1.5}$$

onde $\|P - Q\|$ denota a distância em variação total entre P e Q.

Prova Denotando por ν a medida de contagem em $((\mathbb{N}^2)^n, \mathcal{P}((\mathbb{N}^2)^n))$, temos que $Q_{Nmn} \ll \nu$ e $P_m^n \ll \nu$. Assim,

$$\|Q_{Nmn} - P_m^n\| = \int_{(\mathbb{N}^2)^n} |Q_{Nmn} - P_m^n| \mathrm{d}\nu =$$

$$= \int_{C_1} \left| \frac{\binom{m+2}{2}^{N-n} - \binom{m+1}{2}^{N-n}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}} - \frac{1}{\binom{m+2}{2}^{n}} \right| \mathrm{d}\nu + \int_{C_2} \left| \frac{\binom{m+2}{2}^{N-n}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}} - \frac{1}{\binom{m+2}{2}^{n}} \right| \mathrm{d}\nu,$$

onde $C_1 = \{((x_1, y_1), \ldots, (x_n, y_n)) \in (\mathbb{N}^2)^n : \max_{1 \leq i \leq n}\{x_i + y_i\} < m\}$ e $C_2 = \{((x_1, y_1), \ldots, (x_n, y_n)) \in (\mathbb{N}^2)^n : \max_{1 \leq i \leq n}\{x_i + y_i\} = m\}$. Logo,

$$\|Q_{Nmn} - P_m^n\| = \left| \frac{\binom{m+2}{2}^{N-n} - \binom{m+1}{2}^{N-n}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}} - \frac{1}{\binom{m+2}{2}^{n}} \right| \cdot \binom{m+1}{2}^{n} +$$

$$+ \left| \frac{\binom{m+2}{2}^{N-n}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}} - \frac{1}{\binom{m+2}{2}^{n}} \right| \left[\binom{m+2}{2}^{n} - \binom{m+1}{2}^{n} \right] =$$

$$= \left| \frac{1 - \left[\frac{\binom{m+1}{2}}{\binom{m+2}{2}}\right]^{N-n}}{1 - \left[\frac{\binom{m+1}{2}}{\binom{m+2}{2}}\right]^{N}} - 1 \right| \cdot \left[\frac{\binom{m+1}{2}}{\binom{m+2}{2}}\right]^{n} + \left| \frac{1}{1 - \left[\frac{\binom{m+1}{2}}{\binom{m+2}{2}}\right]^{N}} - 1 \right| \left\{ 1 - \left[\frac{\binom{m+1}{2}}{\binom{m+2}{2}}\right]^{n} \right\}$$

Como $\binom{m+1}{2}/\binom{m+2}{2} = \frac{m}{m+2} < 1$, $\forall m \in \mathbb{N}$, segue que

$$\|Q_{Nmn} - P_m^n\| =$$

$$= \left\{1 - \frac{1 - \left(\frac{m}{m+2}\right)^{N-n}}{1 - \left(\frac{m}{m+2}\right)^{N}}\right\}\left(\frac{m}{m+2}\right)^n + \left\{\frac{1}{1 - \left(\frac{m}{m+2}\right)^{N}} - 1\right\}\left[1 - \left(\frac{m}{m+2}\right)^n\right] \leq$$

$$\leq \sup_{x\in(0,1)} \left\{\left(1 - \frac{1 - x^{N-n}}{1 - x^N}\right)x^n + \left(\frac{1}{1 - x^N} - 1\right)(1 - x^n)\right\},$$

isto é,

$$\|Q_{Nmn} - P_m^n\| \leq 2 \sup_{x\in(0,1)} \frac{x^N - x^{N+n}}{1 - x^N}.$$

Mas, é fácil verificar que a função $f(x) = \frac{x^N - x^{N+n}}{1-x^N}$ é não-decrescente no intervalo $(0,1)$ e que $\lim_{x\to 0^+} f(x) = 0$ e $\lim_{x\to 1^-} f(x) = \frac{n}{N}$. Logo, $\sup_{x\in(0,1)} \frac{x^N - x^{N+n}}{1-x^N} = \frac{n}{N}$ e, consequentemente,

$$\|Q_{Nmn} - P_m^n\| \leq \frac{2n}{N} \qquad \blacksquare$$

Do Teorema 1.2.1, podemos derivar a forma finita do Teorema do tipo de De Finetti para a versão discreta do modelo (A). Assim, suponhamos que $((X_1, Y_1), \ldots, (X_N, Y_N))$ sejam vetores aleatórios assumindo valores em $\mathbb{N}^2$ com lei P_N. Seja $T^{(N)} = \max_{1\leq i\leq N}\{X_i + Y_i\}$ e suponha que, dado $T^{(N)} = m$, $m \in \mathbb{N}$, $((X_1, Y_1), \ldots, (X_N, Y_N))$ tenha distribuição Q_{Nm} (distribuição uniforme sobre χ_m^N definida no começo desta seção). Consideremos P_m^n como antes e seja $P_{\mu,n} = \int_{\mathbb{N}} P_m^n \mathrm{d}\mu(m)$ uma medida de probabilidade em $((\mathbb{N}^2)^n, \mathcal{P}((\mathbb{N}^2)^n))$ para uma medida de probabilidade qualquer μ em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ (é fácil verificar que $P_{\mu,n}$

é, de fato, uma medida de probabilidade). Finalmente, denotando por P_n e μ_N a lei da marginal n-dimensional de $((X_1, Y_1), \ldots, (X_N, Y_N))$, $n < N$, e a lei de $T^{(N)}$ (obtida a partir de P_N), respectivamente, podemos estabelecer o seguinte teorema.

Teorema 1.2.2 *Sejam P_N e $P_{\mu,n}$ as medidas de probabilidade definidas acima. Então, existe uma medida de probabilidade μ_0 em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ tal que, para $1 \leq n < N$*

$$\|P_n - P_{\mu_0,n}\| \leq \frac{2n}{N}. \tag{1.6}$$

Prova Da definição de P_n, temos que

$$P_n = \int_{\mathbb{N}} Q_{Nmn} \mathrm{d}\mu_N(m),$$

uma vez que, condicionalmente a $T^{(N)} = m$, $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído sobre χ_m^N (com lei Q_{Nm}) e, consequentemente, sua marginal n-dimensional tem lei Q_{Nmn}. Assim, para $\mu : \mathcal{P}(\mathbb{N}) \to [0, 1]$, temos que

$$\|P_n - P_{\mu,n}\| = 2 \sup_{A \in \mathcal{P}((\mathbb{N}^2)^n)} |P_n(A) - P_{\mu,n}(A)| =$$

$$= 2 \sup_{A \in \mathcal{P}((\mathbb{N}^2)^n)} \left| \int_{\mathbb{N}} Q_{Nmn}(A) \mathrm{d}\mu_N(m) - \int_{\mathbb{N}} P_m^n(A) \mathrm{d}\mu(m) \right|.$$

Para $\mu = \mu_N$, segue que

$$\|P_n - P_{\mu,n}\| = 2 \sup_{A \in \mathcal{P}((\mathbb{N}^2)^n)} \left| \int_{\mathbb{N}} \{Q_{Nmn}(A) - P_m^n(A)\} \mathrm{d}\mu_N(m) \right| \leq$$

$$\leq \sup_{A \in \mathcal{P}((\mathbb{N}^2)^n)} \int 2|Q_{Nmn}(A) - P_m^n(A)| \mathrm{d}\mu_N(m)$$

Mas, do teorema 1.2.1, $2|Q_{Nmn}(A) - P_m^n(A)| \leq \frac{2n}{N}$, $\forall A \in \mathcal{P}((\mathbb{N}^2)^n)$, pois

$$\|Q_{Nmn} - P_m^n\| = 2 \sup_{A \in \mathcal{P}((\mathbb{N}^2)^n)} |Q_{Nmn}(A) - P_m^n(A)| \leq \frac{2n}{N}.$$

Então,

$$\|P_n - P_{\mu_N,n}\| \leq \int_{\mathbb{N}} \frac{2n}{N} \mathrm{d}\mu_N(m) \Rightarrow \|P_n - P_{\mu_n,n}\| \leq \frac{2n}{N}.$$

Portanto, existe uma medida de probabilidade $\mu_0 = \mu_N$ em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ tal que $\|P_n - P_{\mu_0,n}\| \leq \frac{2n}{N}$, para $1 \leq n < N$. ■

O Teorema 1.2.2 estabelece que a lei da marginal n-dimensional de $((X_1, Y_1), \ldots, (X_N, Y_N))$, a saber, P_n, pode ser bem aproximada, em distância em variação total, por uma mistura da distribuição de n vetores aleatórios independentes e identicamente distribuídos segundo (1.1), sempre que $((X_1, Y_1), \ldots, (X_N, Y_N))$, dado $\max_{1 \leq i \leq N}\{X_i + Y_i\} = m$, tenha distribuição uniforme sobre χ_m^N.

No que segue, obtemos a versão infinita do Teorema do tipo de De Finetti para a versão discreta do modelo (A).

Teorema 1.2.3 *Seja $\{(X_n, Y_n)\}_{n \geq 1}$ uma sequência de vetores aleatórios assumindo valores em $\mathbb{N}^2$. Suponha que para cada $N \in \mathbb{N}$ e $m \in \mathbb{N}$, a distribuição condicional de $((X_1, Y_1), \ldots, (X_N, Y_N))$, dado $\max_{1 \leq i \leq N}\{X_i + Y_i\} = m$, seja uniforme sobre χ_m^N. Então, existe uma única medida de probabilidade μ^* em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ tal que $\forall n \in \mathbb{N}$ e $\forall((x_1, y_1), \ldots, (x_n, y_n)) \in (\mathbb{N}^2)^n$*

$$P((X_1 = x_1, Y_1 = y_1), \ldots, (X_n = x_n, Y_n = y_n)) =$$

$$= \int_{\mathbb{N}} \left[\frac{2}{(\theta+1)(\theta+2)}\right]^n \mathbb{I}_{\{0,\ldots,\theta\}} \left(\max_{1 \leq i \leq n}\{x_i + y_i\}\right) \mathrm{d}\mu^*(\theta) \tag{1.7}$$

Prova Seja $\mathcal{M} = \{\mathbb{P} : ((\mathbb{N}^2)^n, \mathcal{P}((\mathbb{N}^2)^n)) \to [0,1]\}$, isto é, o conjunto de medidas de probabilidade em $((\mathbb{N}^2)^n, \mathcal{P}((\mathbb{N}^2)^n))$. Considerando a distância em variação total em $\mathcal{M}^2$, que denotamos aqui por $\| \ \|$, o par $(\mathcal{M}, \| \ \|)$ constitui um espaço métrico com distância $\| \ \|$. Vamos verificar que os dois lados da igualdade acima correspondem ao limite de uma particular sequência de medidas em $\mathcal{M}$. Consequentemente, pela propriedade da unicidade do limite em $(\mathcal{M}, \| \ \|)$, concluiremos o resultado do Teorema 1.2.3.

Da hipótese do Teorema e do resultado estabelecido no Teorema 1.2.2, tomando $\mu_0 = \mu_N$, temos que o lado esquerdo de (1.7) corresponde a lei P_n e que

$$\|P_n - P_{\mu_N,n}\| \leq \frac{2n}{N}.$$

Tomando limites nesta desigualdade ($N \to \infty$), temos

$$\lim_{N\to\infty} \|P_n - P_{\mu_N,n}\| = 0.$$

Logo, $\{P_{\mu_N,n}\}_{N\geq 1}$ converge uniformemente a P_n quando $N \to \infty$, isto é,

$$P_{\mu_N,n} \xrightarrow{\mu} P_n, \ N \to \infty. \tag{1.8}$$

Agora, verificaremos que a sequência $\{P_{\mu_N,n}\}_{N\geq 1}$ possui uma subsequência convergente à medida apresentada no lado direito de (1.7). Para tanto, provaremos que a sequência $\{\mu_N\}_{N\geq 1}$ é "tight", o que implicatá o resultado limite procurado.

Para $x \in \mathbb{N}$, temos, para todo $N \in \mathbb{N}$, que

$$P(X_1 \leq x, Y_1 \leq x) = \int_{\mathbb{N}} P\left(X_1 \leq x, Y_1 \leq x \mid \max_{1\leq i\leq N}\{X_1 + Y_1\} = m\right) \mathrm{d}\mu_N(m),$$

μ_N sendo a lei de $\max_{1\leq i\leq N}\{X_i + Y_i\}$ como antes. Então,

$$P(X_1 \leq x, Y_1 \leq x) = \int_{\mathbb{N}} Q_{Nm1}(\{0,\ldots,x\}^2)\mathrm{d}\mu_N(m) =$$

$$= \int_{\{0,\ldots,2x\}} Q_{Nm1}(\{0,\ldots,x\}^2)\mathrm{d}\mu_N(m) + \int_{\{2x+1,\ldots\}} Q_{Nm1}(\{0,\ldots,x\})\mathrm{d}\mu_N(m) \leq$$

$$\leq \int_{\{0,\ldots,2x\}} 1 \mathrm{d}\mu_N(m) + \int_{\{2x+1,\ldots\}} Q_{Nm1}(\{0,\ldots,x\}^2) \mathrm{d}\mu_N(m).$$

Para $m \geq 2x+1$, $Q_{Nm1}(\{0,\ldots,x\}^2)$ pode ser obtida pela primeira sentença de (1.4). Assim,

$$P(X_1 \leq x, Y_1 \leq x) \leq \mu_N(\{0,\ldots,2x\}) + \int_{\{2x+1,\ldots\}} \left[\sum_{x_1=0}^{x} \sum_{y_1=0}^{x} \frac{\binom{m+2}{2}^{N-1} - \binom{m+1}{2}^{N-1}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}} \right] \mathrm{d}\mu_N(m)$$

$$= \mu_N(\{0,\ldots,2x\}) + \int_{\{2x+1,\ldots\}} 2(x+1)^2 \frac{1}{(m+1)(m+2)} \underbrace{\frac{1-\left(\frac{m}{m+2}\right)^{N-1}}{1-\left(\frac{m}{m+2}\right)^{N}}}_{\leq 1} \mathrm{d}\mu_N(m)$$

$$\leq \mu_N(\{0,\ldots,2x\}) + \int_{\{2x+1,\ldots\}} \frac{2(x+1)^2}{(m+1)(m+2)} \mathrm{d}\mu_N(m).$$

Mas, para $m \geq 2x+1$, $(m+1)(m+2) \geq 4(x+1)^2$. Então,

$$\Pr(X_1 \leq x, Y_1 \leq x) \leq \mu_N(\{0,\ldots,2x\}) + \int_{\{2x+1,\ldots\}} \frac{1}{2} \mathrm{d}\mu_N(m) \Rightarrow$$

$$\Rightarrow \quad \Pr(X_1 \leq x, Y_1 \leq x) \leq \frac{1}{2} + \frac{1}{2}\mu_N(\{0,\ldots,2x\}), \ \forall x \in \mathbb{N}, \ \forall N \in \mathbb{N}. \tag{1.9}$$

Como $\Pr(X_1 \leq x, Y_1 \leq x) \to 1$ quando $x \to \infty$ (é uma função de distribuição bivariada), temos que $\forall \varepsilon > 0$, $\exists x_0 = x_0(\varepsilon) \in \mathbb{N}$ tal que $\forall x \geq x_0$ $\Pr(X_1 \leq x, X_2 \leq x) > 1 - \frac{\varepsilon}{2}$. Assim, tomando $x = x_0$, juntamente com (1.9), segue que

$$1 - \frac{\varepsilon}{2} < \Pr(X_1 \leq x_0, Y_1 \leq x_0) \leq \frac{1}{2} + \frac{1}{2}\mu_N(\{0,\ldots,2x_0\}) \Rightarrow$$

$$\Rightarrow \mu_N(\{0,\ldots,2x_0\}) > 1 - \varepsilon, \quad \forall N \in \mathbb{N}.$$

Como $\forall \varepsilon > 0$, $\exists x_0 \in \mathbb{N}$ tal que $\mu_N(\{0,\ldots,2x_0\}) > 1-\varepsilon$ ($\{0,\ldots,2x_0\}$ sendo um conjunto compacto em $\mathbb{N}$), $\forall N \in \mathbb{N}$, segue que $\{\mu_N\}_{N\geq 1}$ é "tight". Nestas condições, existe uma subsequência $\{\mu_{N_k}\}_{k\geq 1}$ e uma medida de probabilidade μ^* em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ tais que $\mu_{N_k} \to \mu^*$ fracamente quando $k \to \infty$, ou, equivalentemente,

$$\int_{\mathbb{N}} f \mathrm{d}\mu_{N_k} \xrightarrow{w} \int_{\mathbb{N}} f \mathrm{d}\mu^*, \text{ quando } k \to \infty,$$

para toda função contínua e limitada (em quase toda parte) $f : \mathbb{N} \to \mathbb{R}$. Em particular, para $f : \mathbb{N} \to \mathbb{R}$ dada por $f(\theta) = \left[\frac{2}{(\theta+1)(\theta+2)}\right]^n \mathbb{I}_{\{0,\ldots,\theta\}}(\max_{1\leq i\leq n}\{x_i+y_i\})$, que é limitada, temos que

$$\int_{\mathbb{N}} f(\theta)\mathrm{d}\mu_{N_k}(\theta) \xrightarrow{w} \int_{\mathbb{N}} f(\theta)\mathrm{d}\mu^*(\theta), \text{ quando } k \to \infty,$$

ou,

$$P_{\mu_{N_k}}((x_1,y_1),\ldots,(x_n,y_n)) \xrightarrow{w} \int_{\mathbb{N}} \left[\frac{2}{(\theta+1)(\theta+2)}\right]^n \mathbb{I}_{\{0,\ldots,\theta\}}(\max_{1\leq i\leq n}\{x_i+y_i\})\mathrm{d}\mu^*(\theta). \tag{1.10}$$

De (1.8), (1.10) e da unicidade do limite de $\{P_{\mu_N,n}\}_{N\geq 1}$, concluimos que

$$P_n((x_1,y_1),\ldots,(x_n,y_n)) = \int_{\mathbb{N}} \left[\frac{2}{(\theta+1)(\theta+2)}\right]^n \mathbb{I}_{\{0,\ldots,\theta\}}(\max_{1\leq i\leq n}\{x_i+y_i\})\mathrm{d}\mu^*(\theta).$$

Portanto, $\forall n \in \mathbb{N}$ e $\forall((x_1,y_1),\ldots,(x_n,y_n)) \in (\mathbb{N}^2)^n$,

$$P_n((x_1,y_1),\ldots,(x_n,y_n)) = \int_{\mathbb{N}} \left[\frac{2}{(\theta+1)(\theta+2)}\right]^n \mathbb{I}_{\{0,\ldots,\theta\}}(\max_{1\leq i\leq n}\{x_i+y_i\})\mathrm{d}\mu^*(\theta),$$

Verificamos, finalmente, a unicidade de μ^*. Para N, $t \in \mathbb{N}$, temos

$$\mu_N(\{0,\ldots,t\}) = P\left(\max_{1\leq i\leq N}\{X_i+Y_i\} \leq t\right) = P\left(\bigcap_{i=1}^{N}\{X_i+Y_i \leq t\}\right) =$$

$$= P\left(\underbrace{\{((x_1,y_1),\ldots,(x_N,y_N)) \in (\mathbb{N}^2)^N : x_i+y_i \leq t,\ i=1,\ldots,N\}}_{H_{Nt}}\right) =$$

$$= \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in H_{Nt}} \int_{\mathbb{N}} \left[\frac{2}{(\theta+1)(\theta+2)}\right]^N \mathbb{I}_{\{0,\ldots,\theta\}}\left(\max_{1\le i\le N}\{x_i+y_i\}\right) \mathrm{d}\mu^*(\theta) =$$

$$= \int_{\mathbb{N}} \left(\sum_{((x_1,y_1),\ldots,(x_N,y_N))\in H_{Nt}} \left[\frac{2}{(\theta+1)(\theta+2)}\right]^N \mathbb{I}_{\{0,\ldots,\theta\}}\left(\max_{1\le i\le N}\{x_i+y_i\}\right)\right) \mathrm{d}\mu^*(\theta)$$

Como a soma interna da integral acima é 1 para $\theta \le t$ e $\left[\frac{(t+1)(t+2)}{(\theta+1)(\theta+2)}\right]^N$ para $\theta > t$, segue que

$$\mu_N(\{0,\ldots,t\}) = \int_{\{0,\ldots,t\}} 1\mathrm{d}\mu^*(\theta) + \int_{\{t+1,\ldots\}} \left[\frac{(t+1)(t+2)}{(\theta+1)(\theta+2)}\right]^N \mathrm{d}\mu^*(\theta) =$$

$$= \mu^*(\{0,\ldots,t\}) + \int_{\{t+1,\ldots\}} \left[\frac{(t+1)(t+2)}{(\theta+1)(\theta+2)}\right]^N \mathrm{d}\mu^*(\theta)\ .$$

Pelo Teorema da Convergência Dominada, a segunda parcela do lado direito da igualdade acima tende a zero quando $N \to \infty$ e, conseqüentemente,

$$\lim_{N\to\infty} \mu_N(\{0,\ldots,t\}) = \mu^*(\{0,\ldots,t\}).$$

Assim, $\{\mu_N\}_{N\ge1}$ possui limite, completando a prova do teorema 1.2.3. ■

A partir do Teorema 1.2.3 podemos derivar a representação preditivista para outro modelo discreto cuja distribuição de probabilidade é dada por

$$\mathrm{Prob}\{Z = z|\ \theta\} = \frac{2(z+1)}{(\theta+1)(\theta+2)} I_{\{0,\ldots,\theta\}}(z).$$

A obtenção deste resultado consiste basicamente na construção de um novo processo estocástico $\{Z_n\}_{n\ge1}$ a partir de $\{(X_n,Y_n)\}_{n\ge1}$ definindo $Z_n = X_n + Y_n$, $\forall n \in \mathbb{N}$, e na especificação de convenientes condições sobre $\{Z_n\}_{n\ge1}$ resultantes da uniformidade sobre χ_m^N, $\forall N, m \in \mathbb{N}$, para o processo original $\{(X_n,Y_n)\}_{n\ge1}$.

Destacamos que se $(X_1,Y_1),\ldots,(X_N,Y_N)$, $N \ge 2$, são vetores aleatórios independentes e identicamente distribuídos assumindo valores em $\mathbb{N}^2$, então $((X_1,Y_1),\ldots,(X_N,Y_N))$ é

uniformemente distribuído sobre χ_m^N, dado que $\max_{1\leq i\leq N}\{X_i+Y_i\}=m$, se, e somente se, (X_i,Y_i), $i=1,\ldots,N$, tem distribuição uniforme em $\{(a,b)\in\mathbb{N}^2 : a+b\leq k\}$ para algum $k\in\mathbb{N}$ (a demonstração deste resultado é apresentada no capítulo 2).

As formas finita e infinita do Teorema do tipo de De Finetti podem ser generalizadas para o caso de um vetor aleatório d-variado, $d>2$, uniformemente distribuído sobre $\{(x_1,\ldots,x_d)\in\mathbb{N}^d : x_1+\cdots+x_d\leq\theta\}$, $\theta\in\mathbb{N}$. Isto é, podemos derivar resultados análogos aos Teoremas 1.2.2 e 1.2.3 para o modelo cuja distribuição de probabilidade é dada por

$$\text{Prob}\{X_1=x_1,\ldots,X_d=x_d\mid\theta\}=\frac{1}{\binom{\theta+d}{d}}\mathbb{I}_{\{(a_1,\ldots,a_d)\in\mathbb{N}^d:a_1+\cdots+a_d\leq\theta\}}(x_1,\ldots,x_d) \qquad (1.11)$$

Os resultados para o modelo em (1.11) podem ser obtidos de uma maneira similar ao caso $d=2$. À título de ilustração, enunciamos (sem demonstração) a forma infinita do Teorema do tipo de De Finetti para este modelo.

Teorema 1.2.4 *Seja $\{(X_n^1,\ldots,X_n^d)\}_{n\geq1}$ uma sequência de vetores aleatórios d-variados, $d>2$, assumindo valores em $\mathbb{N}^d$. Suponha que para cada $N\in\mathbb{N}$ e $m\in\mathbb{N}$, a distribuição condicional de $((X_1^1,\ldots,X_1^d),\ldots,(X_N^1,\ldots,X_N^d))$, dado $\max_{1\leq i\leq N}\{X_i^1+\cdots+X_i^d\}=m$, seja uniforme sobre $\{((x_1^1,\ldots,x_1^d),\ldots,(x_N^1,\ldots,x_N^d))\in(\mathbb{N}^d)^N : \max_{1\leq i\leq N}\{x_i^1+\cdots+x_i^d\}=m\}$. Então, existe uma medida de probabilidade μ^* em $(\mathbb{N},\mathcal{P}(\mathbb{N}))$ tal que $\forall n\in\mathbb{N}$ e $\forall((x_1^1,\ldots,x_1^d),\ldots,(x_n^1,\ldots,x_n^d))\in(\mathbb{N}^d)^n$*

$$\begin{aligned}&P((X_1^1=x_1^1,\ldots,X_1^d=x_1^d),\ldots,(X_n^1=x_n^1,\ldots,X_n^d=x_n^d))\\ =&\int_{\mathbb{N}}\left[\frac{1}{\binom{\theta+d}{d}}\right]^n\mathbb{I}_{\{0,\ldots,\theta\}}(\max_{1\leq i\leq N}\{x_i^1+\cdots+x_i^d\})\mathrm{d}\mu^*(\theta).\end{aligned}$$

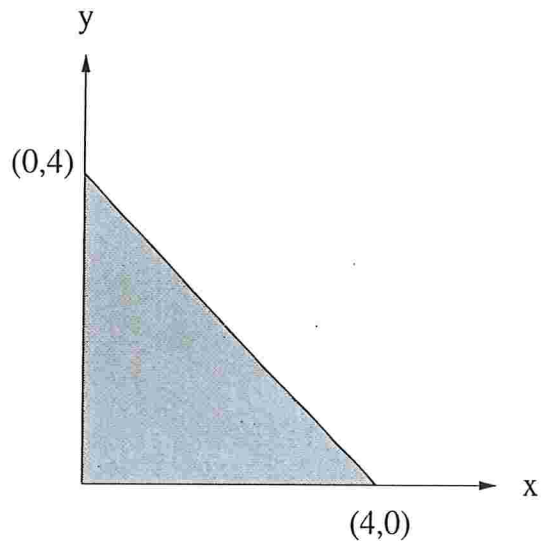
No que segue, derivamos as formas finita e infinita do Teorema do tipo de De Finetti para a versão contínua do modelo (A).

1.2.2 Caso contínuo

Consideramos, a partir de agora, a versão contínua do modelo (A). Sua função densidade de probabilidade é dada por

$$f(x, y|\theta) = \frac{2}{\theta^2} \mathbb{I}_{\{(a,b)\in\mathbb{R}_+^2 : a+b\leq\theta\}}(x, y). \tag{1.12}$$

O suporte de (1.12) é esboçado abaixo para $\theta = 4$.



Para obtermos a representação preditivista deste modelo, procederemos como na subseção 1.2.1, primeiro derivando a forma finita do Teorema do tipo de De Finetti e, então, sua forma infinita. Assim, consideremos o conjunto $\chi_m^N = \{((x_1, y_1)\ldots, (x_N, y_N)) \in (\mathbb{R}_+^2)^N : \max_{1\leq i\leq N}\{x_i + y_i\} = m\}$, $m > 0$. Construiremos, a seguir, uma medida de probabilidade uniforme sobre χ_m^N.

Inicialmente, destacamos, visando a construção de uma medida de probabilidade uniforme sobre χ_m^N, que $\lambda_2^{(N)}(\chi_m^N) = 0$. Isto implica que a maneira mais natural de construirmos uma medida de probabilidade uniforme μ sobre um subconjunto compacto B de

$(\mathbb{R}^2)^N$, a saber, $\mu(A) = \frac{\lambda_2^{(N)}(A\cap B)}{\lambda_2^{(N)}(B)}$, $A \in \mathcal{B}((\mathbb{R}^2)^N)$, não é adequada no presente caso. Portanto, vamos introduzir um enfoque um pouco diferente, baseado no trabalho de Iglesias et al. (1998).

Formalmente, seja χ_m^N definido como antes. Consideremos os conjuntos $M_i(m) = \{((x_1, y_1)\ldots, (x_N, y_N)) \in (\mathbb{R}_+^2)^N : x_i + y_i = m, x_j + y_j \leq m, j \neq i\}$, $i = 1, \ldots, N$. É claro que $\chi_m^N = \bigcup_{i=1}^N M_i(m)$. Para cada $i = 1, \ldots, N$, consideremos a "projeção" $\varphi^i = (\mathbb{R}^2)^N \to \mathbb{R}^{2N-2}$ dada por

$$\varphi^i(((x_1, y_1)\ldots, (x_N, y_N))) = (x_1, y_1, \ldots, x_{i-1}, y_{i-1}, x_{i+1}, y_{i+1}, \ldots, x_N, y_N)$$

Seja $\mathcal{A} = \chi_m^N \cap \mathcal{B}((\mathbb{R}^2)^N)$ uma σ-álgebra de subconjuntos de χ_m^N. Considerando os espaços de medida $(\chi_m^N, \mathcal{A}, \mu_i)$, $i = 1, \ldots, N$, onde $\mu_i : \mathcal{A} \to \mathbb{R}_+$ é dada por

$$\mu_i(A) = \lambda_{2N-2}(\varphi^i(A \cap M_i(m))), \ A \in \mathcal{A}, \tag{1.13}$$

podemos definir $\tilde{Q}_{Nm} : \mathcal{A} \to [0, 1]$, uma medida de probabilidade em $(\chi_m^N, \mathcal{A})$, tal que

$$\tilde{Q}_{Nm}(A) = \frac{\sum_{i=1}^N \mu_i(A)}{N\left(\frac{m^2}{2}\right)^{N-1}}, \quad A \in \mathcal{A}.$$

Finalmente, definimos $Q_{Nm} : \mathcal{B}((\mathbb{R}^2)^N) \to [0, 1]$ por

$$Q_{Nm}(B) = \tilde{Q}_{Nm}(B \cap \chi_m^N), \quad B \in \mathcal{B}((\mathbb{R}^2)^N). \tag{1.14}$$

A medida de probabilidade Q_{Nm} em $((\mathbb{R}^2)^N, \mathcal{B}((\mathbb{R}^2)^N))$ chamamos de distribuição uniforme sobre χ_m^N. Notemos que se $((X_1, Y_1), \ldots, (X_N, Y_N))$ tem lei Q_{Nm}, então suas marginais bivariadas $(X_1, Y_1), \ldots, (X_N, Y_N)$ são permutáveis.

Analogamente à subseção 1.2.1, continuamos estimando a distância em variação total entre a lei da marginal n-dimensional de Q_{Nm}, $n < N$, que denotamos por Q_{Nmn}, e a

lei de n vetores aleatórios bivariados independentes e com mesma distribuição uniforme em $\{(a,b) \in \mathbb{R}_+^2 : a+b \leq m\}$, $m > 0$. Para isto, estabelecemos uma caracterização de uniformidade que será de grande importância na sequência.

Proposição 1.2.1 *Sejam $(U_1, V_1), \ldots, (U_N, V_N)$ vetores aleatórios independentes assumindo valores em $\mathbb{R}_+^2$, cada um uniformemente distribuído sobre $E = \{(a,b) \in \mathbb{R}_+^2 : a+b \leq 1\}$. Seja $S^{(N)} = \max_{1 \leq i \leq N}\{U_i + V_i\}$, $X_i = \frac{mU_i}{S^{(N)}}$ e $Y_i = \frac{mV_i}{S^{(N)}}$, $i = 1, \ldots, N$, com $m > 0$. Então, $((X_1, Y_1), \ldots, (X_N, Y_N))$ é o único vetor Q_{Nm}-uniformemente distribuído em χ_m^N.*

Prova Seja Q a lei de $((X_1, Y_1), \ldots, (X_N, Y_N))$. Claramente, $Q(\chi_m^N) = 1$, pois $X_i + Y_i = \frac{m(U_i+V_i)}{S^{(N)}} \leq m, \forall i = 1, \ldots, N$, e, para $j \in \{1, \ldots, N\}$ tal que $S^{(N)} = U_j + V_j$, temos que $X_j + Y_j = m$ e, portanto, $((X_1, Y_1), \ldots, (X_N, Y_N)) \in \bigcup_{j=1}^N M_j(m) = \chi_m^N$. Para $A \in \mathcal{B}((\mathbb{R}^2)^N)$,

$$Q(A) = P(((X_1, Y_1), \ldots, (X_N, Y_N)) \in A) = P\left(\bigcup_{i=1}^N A_i\right),$$

onde A_i é o evento $((X_1, Y_1), \ldots, (X_N, Y_N)) \in A \cap M_i(m)$. Então,

$$Q(A) = P\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N P(A_i), \tag{1.15}$$

onde a última igualdade segue do fato que $\lambda_{2N-2}(\varphi^i(M_i(m) \cap M_j(m))) = 0$, para $i \neq j$. Calculemos, então, $P(A_i)$.

$$P(A_i) = P(((X_1, Y_1), \ldots, (X_N, Y_N)) \in A \cap M_i(m))$$

$$= P(X_i + Y_i = m, \varphi^i(((X_1, Y_1), \ldots, (X_N, Y_N))) \in \varphi^i(A \cap M_i(m))) =$$

$$= P\left(\frac{m}{S^{(N)}}(U_i + V_i) = m, \varphi^i\left(\frac{m}{S^{(N)}}((U_1, V_1), \ldots, (U_N, V_N))\right)\right) \in \varphi^i(A \cap M_i(m))\Big)$$

$$= P\left(\underbrace{\bigcap_{j\neq i}\{U_j+V_j \leq U_i+V_i\}}_{B}, \underbrace{\varphi^i\left(\frac{m}{U_i+V_i}((U_1,V_1),\ldots,(U_N,V_N))\right) \in \varphi^i(A\cap M_i(m))}_{C}\right)$$

Mas, se $((r_1,s_1)\ldots,(r_N,s_N)) \in C$, então $\frac{m(r_k+s_k)}{U_i+V_i} \leq m$, $\forall k = 1,\ldots,N$, $k \neq i$, isto é, $r_k + s_k \leq U_i + V_i$, $\forall k \neq i$. Logo, $((r_1,s_1),\ldots,(r_N,s_N)) \in \bigcap_{j\neq i}\{U_j + V_j \leq U_i + V_i\} = B$, implicando que $C \subseteq B$. Portanto,

$$P(A_i) = P\left(\varphi^i\left(\frac{m}{U_i+V_i}((U_1,V_1),\ldots,(U_N,V_N))\right) \in \varphi^i(A\cap M_i(m))\right) =$$

$$= \int_E P\left(\varphi^i\left(\frac{m}{a+b}((U_1,V_1),\ldots,(U_N,V_N))\right) \in \varphi^i(A\cap M_i(m))|U_i=a, V_i=b\right) 2\mathrm{d}a\mathrm{d}b,$$

lembrando que (U_i, V_i) tem função densidade de probabilidade dada por (1.12) com $\theta = 1$. Como os vetores aleatórios $(U_1,V_1),\ldots,(U_N,V_N)$ são independentes, segue que

$$P(A_i) = \int_E P\left(\varphi^i\left(\frac{m}{a+b}((U_1,V_1),\ldots,(U_N,V_N))\right) \in \varphi^i(A\cap M_i(m))|U_i=a, V_i=b\right) 2\mathrm{d}a\mathrm{d}b$$

$$= \int_E \left[\int_{\varphi^i(A\cap M_i(m))} f_{\mathcal{Z}_{(a,b)}} \mathrm{d}\lambda_{2N-2}\right] 2\mathrm{d}a\mathrm{d}b, \tag{1.16}$$

onde $f_{\mathcal{Z}_{(a,b)}}$ é a função densidade de probabilidade do vetor $(2N-2)$-dimensional

$$\mathcal{Z}_{(a,b)} = \left(\varphi^i\left(\frac{m}{a+b}((U_1,V_1),\ldots,(U_N,V_N))\right)\right) =$$

$$= \frac{m}{a+b}(U_1,V_1,\ldots,U_{i-1},V_{i-1},U_{i+1},V_{i+1},\ldots,U_N,V_N).$$

Pela hipótese da proposição e aplicando o método do Jacobiano para transformações de vetores aleatórios, obtemos que

$$f_{\mathcal{Z}_{(a,b)}} = \left(\frac{a+b}{m}\right)^{2N-2} 2^{N-1}.$$

Então,

$$P(A_i) = \int_E \left[\int_{\varphi^i(A\cap M_i(m))} \left(\frac{a+b}{m}\right)^{2N-2} 2^{N-1} \mathrm{d}\lambda_{2N-2}\right] 2\mathrm{d}a\mathrm{d}b =$$

$$= \int_E \frac{2^N(a+b)^{2N-2}}{m^{2N-2}} \lambda_{2N-2}(\varphi^i(A\cap M_i(m)))\mathrm{d}a\mathrm{d}b =$$

$$= \frac{2^N \lambda_{2N-2}(\varphi^i(A\cap M_i(m)))}{m^{2N-2}} \int_E (a+b)^{2N-2}\mathrm{d}a\mathrm{d}b \Rightarrow$$

$$\Rightarrow P(A_i) = \frac{\lambda_{2N-2}(\varphi^i(A\cap M_i(m)))}{N\left(\frac{m^2}{2}\right)^{N-1}}$$

Substituindo este resultado em (1.15), obtemos

$$Q(A) = \frac{\sum_{i=1}^N \lambda_{2N-2}(\varphi^i(A\cap \chi_m^N \cap M_i(m)))}{N\left(\frac{m^2}{2}\right)^{N-1}} = \frac{\sum_{i=1}^N \mu_i(A\cap \chi_m^N)}{N\left(\frac{m^2}{2}\right)^{N-1}} = \tilde{Q}_{Nm}(A\cap\chi_m^N) \Rightarrow$$

$$\Rightarrow Q(A) = Q_{Nm}(A).$$

Assim, $((X_1,Y_1),\ldots,(X_N,Y_N))$ é Q_{Nm}-uniformemente distribuído em χ_m^N. ■

A Proposição 1.2.1 estabelece que um vetor aleatório N-dimensional com lei Q_{Nm} pode ser caracterizado como uma função de N vetores aleatórios bivariados independentes e identicamente distribuídos com distribução uniforme em E.

A seguir, apresentamos a distribuição da marginal unidimensional de Q_{Nm}, a qual denotaremos por Q_{Nm1} (para $1 < n < N$, denotamos a marginal n- dimensional por Q_{Nmn}). Mais precisamente, exibimos Q_{Nm1} em termos de sua função de distribuição. Este resultado figurará na demonstração do Teorema 1.2.7 adiante.

Proposição 1.2.2 *Seja $((X_1, Y_1), \ldots, (X_N, Y_N))$ um vetor aleatório uniformemente distribuído sobre χ_m^N. Então, a função de distribuição de sua marginal unidimensional, (X_1, Y_1), é dada por*

$$P(X_1 \leq x, Y_1 \leq y) = \begin{cases} 0, & \min\{x,y\} < 0 \\ \dfrac{(N-1)}{N} \cdot \dfrac{2xy}{m^2}, & \min\{x,y\} \geq 0 \ e \ x+y < m \\ \dfrac{1}{N} + \dfrac{(N-1)}{N} \cdot \dfrac{2xy - (x+y-m)^2}{m^2}, & 0 \leq x,y \leq m \ e \ x+y \geq m \\ \dfrac{1}{N} + \dfrac{(N-1)}{N} \cdot \dfrac{\min\{x,y\}(2m - \min\{x,y\})}{m^2}, & \min\{x,y\} < m \leq \max\{x,y\} \\ 1, & \min\{x,y\} \geq m. \end{cases}$$

Prova Provamos o resultado para o caso $x, y \in \mathbb{R}_+$ tais que $x + y < m$ (os demais casos podem ser provados de maneira similar).

$$P(X_1 \leq x, Y_1 \leq y) = Q_{Nm1}((-\infty, x] \times (-\infty, y]) =$$
$$= Q_{Nm}\left(\underbrace{((-\infty, x] \times (-\infty, y]) \times (\mathbb{R}^2)^{N-1}}_{D}\right)$$

Da definição de Q_{Nm}, temos

$$Q_{Nm}(D) = \frac{\sum_{i=2}^{N} \lambda_{2N-2}(\varphi^i(D \cap \chi_m^N \cap M_i(m)))}{N\left(\frac{m^2}{2}\right)^{N-1}},$$

pois $D \cap M_1(m) = \emptyset$. Como $D \cap \chi_m^N \cap M_i(m) = ([0,x] \times [0,y]) \times E_m^{N-1}$, onde $E_m = \{(a,b) \in \mathbb{R}_+^2 : a + b \leq m\}$, para $i \neq 1$, segue que

$$Q_{Nm}(D) = \frac{\sum_{i=2}^{N} \lambda_{2N-2}(([0,x] \times [0,y]) \times E_m^{N-2})}{N\left(\frac{m^2}{2}\right)^{N-1}} =$$

$$= \frac{\sum_{i=2}^{N} \lambda_2([0,x]\times[0,y])\,[\lambda_2(E_m)]^{N-2}}{N\left(\frac{m^2}{2}\right)^{N-1}} = \frac{\sum_{i=2}^{N} xy\left(\frac{m^2}{2}\right)^{N-2}}{N\left(\frac{m^2}{2}\right)^{N-1}} \Rightarrow$$

$$\Rightarrow\ Q_{Nm}(D) = \frac{(N-1)}{N}\frac{2xy}{m^2},$$

concluindo a prova. ■

No que segue, derivamos a função de distribuição de $T^{(N)} = \max_{1\le i\le n}\{X_i + Y_i\}$, um resultado que terá um papel relevante na obtenção da forma finita do Teorema do tipo de De Finetti para a versão contínua do modelo (A).

Proposição 1.2.3 *Seja $((X_1,Y_1),\ldots,(X_n,Y_n))$ um vetor aleatório com lei Q_{Nmn}. Então, a função de distribuição de $T^{(n)} = \max_{1\le i\le n}\{X_i + Y_i\}$ é dada por*

$$P(T^{(n)} \le t) = \begin{cases} 0 & , t < 0, \\ \dfrac{N-n}{N}\left(\dfrac{t}{m}\right)^{2n} & , 0 \le t < m, \\ 1 & , t \ge m. \end{cases}$$

Prova Para $t \in (0,m)$ (as outras duas situações são triviais), temos que

$$P(T^{(n)} \le t) \;=\; P\left(\bigcap_{i=1}^{n}\{X_i + Y_i \le t\}\right) =$$

$$= P\left(\bigcup_{k=1}^{N}\left[\bigcap_{i=1}^{n}\{X_i + Y_i \le t\} \cap \{\max_{1\le i\le N}\{X_i + Y_i\} = X_k + Y_k\}\right]\right) =$$

$$= \sum_{k=1}^{N} P\left(\{\max_{1\le i\le N}\{X_i + Y_i\} = X_k + Y_k\}\bigcap_{i=1}^{n}\{X_i + Y_i \le t\}\right) =$$

$$= \sum_{k=1}^{N} P\left(\bigcap_{i\ne k}\{X_i + Y_i \le X_k + Y_k\}\bigcap_{i=1}^{n}\{X_i + Y_i \le t\}\right).$$

Da Proposição 1.2.1., segue que

$$P(T^{(n)} \le t) \;=\; \sum_{k=1}^{N} P\left(\bigcap_{i\ne k}\{U_i + V_i \le U_k + V_k\}\bigcap_{i=1}^{n}\left\{U_i + V_i \le \frac{t(U_k + V_k)}{m}\right\}\right) =$$

Como $t < m$, segue que as n primeiras parcelas da soma acima são todas iguais a zero. Portanto,

$$P(T^{(n)} \leq t) = \sum_{k=n+1}^{N} P\left(\bigcap_{i=1}^{n}\left\{U_i + V_i \leq \frac{t(U_k+V_k)}{m}\right\} \cap \bigcap_{i\geq n+1, i\neq k}\{U_i + V_i \leq U_k + V_k\}\right) =$$

$$\sum_{k=n+1}^{N} \int_E P\left(\bigcap_{i=1}^{n}\left\{U_i + V_i \leq \frac{t(x+y)}{m}\right\} \cap \bigcap_{i\geq n+1, i\neq k}\{U_i + V_i \leq x + y\}\right| U_k = x, V_k = y 2\mathrm{d}x\mathrm{d}y$$

Como $\{(U_i, V_i)\}_{i=1}^{N}$ são independentes com mesma distribuição uniforme em E, segue que

$$P(T^{(n)} \leq t) =$$

$$= \sum_{k=n+1}^{N} \int_E \left[P\left(U_1 + V_1 \leq \frac{t(x+y)}{m}\right)\right]^n [P(U_1 + V_1 \leq x + y)]^{N-n-1} 2\mathrm{d}x\mathrm{d}y =$$

$$= \sum_{k=n+1}^{N} \int_E \left[\frac{t(x+y)}{m}\right]^{2n} (x+y)^{2N-2n-2} 2\mathrm{d}x\mathrm{d}y =$$

$$= \sum_{k=n+1}^{N} 2\left(\frac{t}{m}\right)^{2n} \int_E (x+y)^{2N-2}\mathrm{d}x\mathrm{d}y = \sum_{k=n+1}^{N} \frac{1}{N}\left(\frac{t}{m}\right)^{2n} \Rightarrow$$

$$\Rightarrow P(T^{(n)} \leq t) = \frac{N-n}{N}\left(\frac{t}{m}\right)^{2n},$$

concluindo a prova. ■

Proposição 1.2.4 *Seja* $((X_1, Y_1), \ldots, (X_N, Y_N))$ *um vetor aleatório com lei* Q_{Nmn}, *dado* $m > 0$. *Então,* $T^{(n)} = \max_{1 \leq i \leq n}\{X_i + Y_i\}$ *é uma estatística suficiente para a família* $\{Q_{Nmn} : m > 0\}$.

A fácil demonstração da Proposição 1.2.4 é aqui omitida. Continuando, estimamos agora a distância em variação total entre Q_{Nmn} e a lei de n vetores aleatórios bidimensionais independentes e com mesma distribuição uniforme sobre $\{(x, y) \in \mathbb{R}_+^2 : x + y \leq m\}$, a qual denotamos por P_m^n.

Teorema 1.2.5 *Sejam* Q_{Nmn} *e* P_m^n *as medidas de probabilidade em* $((\mathbb{R}^2)^n, \mathcal{B}((\mathbb{R}^2)^n))$ *definidas anteriormente. Então, para* $1 \leq n < N$,

$$\|Q_{Nmn} - P_m^n\| \leq \frac{4n}{N}.$$

Prova É fácil verificar que $T^{(n)} = \max_{1 \leq i \leq n}\{X_i + Y_i\}$ é suficiente para a família $\{P_m^n : m > 0\}$. Assim, $T^{(n)}$ é suficiente para ambas as famílias $\{Q_{Nmn} : m > 0\}$ e $\{P_m^n : m > 0\}$, o que implica, de acordo com uma propriedade da distância em variação total, que

$$\|Q_{Nmn} - P_m^n\| = \|Q_{Nmn,T} - P_{m,T}^n\|,$$

onde $Q_{Nmn,T}(P_{m,T}^n)$ é a $Q_{Nmn}(P_m^n)$-lei de $T^{(n)}$. Assim, vamos avaliar $\|Q_{Nmn,T} - P_{m,T}^n\| = 2\sup_{B \in \mathcal{B}(\mathbb{R})} |Q_{Nmn,T}(B) - P_{m,T}^n(B)|$. Como

$$Q_{Nmn,T}(B) = \int_{B \cap (0,m)} \frac{N-n}{N} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a + Q_{Nmn,T}(B \cap \{m\}) \quad \text{e}$$
$$P_{m,T}^n(B) = \int_{B \cap (0,m)} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a,$$

segue que

$$\|Q_{Nmn,T} - P_{m,T}^n\| = 2 \sup_{B \in \mathcal{B}(\mathbb{R})} |Q_{Nmn,T}(B) - P_{m,T}^n(B)| =$$

$$= 2 \sup_{B \in \mathcal{B}(\mathbb{R})} \left| \int_{B \cap (0,m)} \frac{N-n}{N} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a + Q_{Nmn,T}(B \cap \{m\}) - \int_{B \cap (0,m)} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a \right| \leq$$

$$\leq 2 \sup_{B \in \mathcal{B}(\mathbb{R})} \left| \int_{B \cap (0,m)} \frac{N-n}{N} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a - \int_{B \cap (0,m)} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a \right| + 2 \sup_{B \in \mathcal{B}(\mathbb{R})} |Q_{Nmn,T}(B \cap \{m\})|$$

$$\leq 2 \sup_{B \in \mathcal{B}(\mathbb{R})} \int_{B \cap (0,m)} \frac{n}{N} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a + 2Q_{Nmn,T}(\{m\}) \leq$$

$$\leq 2 \int_{(0,m)} \frac{n}{N} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a + \frac{2n}{N} = \frac{4n}{N} \Rightarrow \|Q_{Nmn} - P_m^n\| \leq \frac{4n}{N}.$$

■

Podemos, finalmente, estabelecer a versão finita do Teorema do tipo de De Finetti para a versão contínua do modelo (A). Como na subseção 1.2.1, P_N denota a lei do vetor aleatório N-dimensional $((X_1, Y_1), \ldots, (X_N, Y_N))$, cuja distribuição condicional, dado $\max_{1 \leq i \leq N}\{X_i + Y_i\} = m$, $m > 0$, é uniforme sobre χ_m^N. Consideremos P_m^n como antes e seja $P_{\mu,n} = \int_{\mathbb{R}_+} P_m^n \mathrm{d}\mu(m)$ uma medida de probabilidade em $((\mathbb{R}^2)^n, \mathcal{B}((\mathbb{R}^2)^n))$ para alguma medida de probabilidade μ em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. Finalmente, denotando por P_n a lei da marginal n-dimensional de P_N e por μ_N a lei de $T^{(N)} = \max_{1 \leq i \leq N}\{X_i + Y_i\}$, estabelecemos o seguinte teorema.

Teorema 1.2.6 *Sejam P_N e $P_{\mu,n}$ as medidas de probabilidade definidas acima. Então, existe uma medida de probabilidade μ_0 em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ tal que, para $1 \leq n < N$*

$$\|P_n - P_{\mu_0,n}\| \leq \frac{4n}{N}.$$

Prova Da definição de P_n, temos que

$$P_n = \int_{\mathbb{R}_+} Q_{Nmn} \mathrm{d}\mu_N(m),$$

pois, condicionalmente a $T^{(N)} = m$, $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído em χ_m^N. Assim, para $\mu : \mathcal{B}(\mathbb{R}_+) \to [0, 1]$, temos

$$\|P_n - P_{\mu,n}\| = 2 \sup_{B \in \mathcal{B}((\mathbb{R}^2)^n)} |P_n(B) - P_{\mu,n}(B)| =$$

$$= 2 \sup_{B \in \mathcal{B}((\mathbb{R}^2)^n)} \left| \int_{R_+} Q_{Nmn}(B) \mathrm{d}\mu_N(m) - \int_{\mathbb{R}_+} P_m^n(B) \mathrm{d}\mu(m) \right|.$$

Para $\mu = \mu_N$, segue que

$$\|P_n - P_{\mu_N,n}\| = 2 \sup_{B \in \mathcal{B}((\mathbb{R}^2)^n)} \left| \int_{\mathbb{R}_+} [Q_{Nmn}(B) - P_m^n(B)] \mathrm{d}\mu_N(m) \right| \leq$$

$$\leq 2 \sup_{B \in \mathcal{B}((\mathbb{R}^2)^n)} \int_{\mathbb{R}_+} |Q_{Nmn}(B) - P_m^n(B)| \mathrm{d}\mu_N(m) \leq \frac{4n}{N},$$

onde a última desigualdade segue do Teorema 1.2.5. Portanto, existe uma medida de probabilidade $\mu_0 = \mu_N$ em $(R_+, \mathcal{B}(\mathbb{R}_+))$ tal que $\|P_n - P_{\mu_0,n}\| \leq \frac{4n}{N}$, para $1 \leq n < N$. ■

O último teorema estabelece que a lei da marginal n-dimensional de $((X_1, Y_1), \ldots, (X_N, Y_N))$, a saber, P_n, pode ser bem aproximada por uma mistura da distribuição de n vetores aleatórios bivariados independentes e identicamente distribuídos uniformes em $\{(x, y) \in \mathbb{R}_+^2 : x + y \leq m\}$, $m > 0$, sempre que $((X_1, Y_1), \ldots, (X_N, Y_N))$, dado $\max_{1 \leq i \leq N}\{X_i + Y_i\} = m$, tem distribuição uniforme sobre χ_m^N.

Na sequência, obtemos a versão infinita do Teorema do tipo de De Finetti para a versão contínua do modelo (A).

Teorema 1.2.7 *Seja $\{(X_n, Y_n)\}_{n\geq 1}$ uma sequência de vetores aleatórios bivariados tomando valores em $\mathbb{R}_+^2$. Suponha que para cada $N \in \mathbb{N}$ e $m > 0$, a distribuição condicional de $((X_1, Y_1), \ldots, (X_N, Y_N))$, dado $\max_{1\leq i\leq N}\{X_i + Y_i\} = m$, seja uniforme sobre χ_m^N. Então, existe uma única medida de probabilidade μ^* em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ tal que $\forall n \in \mathbb{N}$ e $B \in \mathcal{B}((\mathbb{R}^2)^n)$*

$$P(((X_1, Y_1), \ldots, (X_n, Y_n)) \in B) = \int_{R_+} P_\theta^n(B)\mathrm{d}\mu^*(\theta) \tag{1.17}$$

Prova A demonstração é análoga à prova do Teorema 1.2.3, na subseção 1.2.1. Neste caso, seja $\mathcal{M} = \{\mathbb{P} : ((\mathbb{R}^2)^n, \mathcal{B}((\mathbb{R}^2)^n)) \to [0,1]\}$ e denotemos por $\| \ \|$ a distância em variação total em $\mathcal{M}^2$. Vamos verificar que os dois lados da igualdade em (1.17) correspondem ao limite de $\{P_{\mu_N,n}\}_{N\geq 1}$ em $\mathcal{M}$. Primeiro, da hipótese do teorema e do resultado do Teorema 1.2.6, tomando $\mu_0 = \mu_N$, obtemos que

$$\lim_{N\to\infty} \|P_n - P_{\mu_N,n}\| = 0$$

ou, de maneira equivalente,

$$P_{\mu_N,n} \xrightarrow{\mu} P_n, \text{ quando } N \to \infty. \tag{1.18}$$

Em seguida, verificamos que $\{\mu_N\}_{N\geq 1}$ é "tight", o que implicará a convergência fraca de $\{P_{\mu_N,n}\}_{N\geq 1}$. Para $x > 0$ e $\forall N \in \mathbb{N}$, temos que

$$P(X_1 \leq x, Y_1 \leq x) = \int_{\mathbb{R}_+} P\left(X_1 \leq x, Y_1 \leq x \mid \max_{1\leq i\leq N}\{X_i + Y_i\} = m\right) \mathrm{d}\mu_N(m),$$

onde μ_N é a lei de $\max_{1\leq i\leq N}\{X_i + Y_i\}$. Então,

$$P(X_1 \leq x, Y_1 \leq x) = \int_{\mathbb{R}_+} Q_{Nm1}([0,x]^2)\mathrm{d}\mu_N(m) =$$

$$= \int_{[0,2x]} Q_{Nm1}([0,x]^2)\mathrm{d}\mu_N(m) + \int_{(2x,\infty)} Q_{Nm1}([0,x]^2)\mathrm{d}\mu_N(m) \Rightarrow$$

$$\Rightarrow P(X_1 \leq x, Y_1 \leq x) \leq \int_{[0,2x]} 1\mathrm{d}\mu_N(m) + \int_{(2x,\infty)} Q_{Nm1}([0,x]^2)\mathrm{d}\mu_N(m). \qquad (1.19)$$

Para $m > 2x$, $Q_{Nm1}([0,x]^2)$ pode ser obtida da Proposição 1.2.2:

$$Q_{Nm1}([0,x]^2) = \frac{(N-1)}{N}\frac{2x^2}{m^2}.$$

Substituindo a expressão acima em (1.19), obtemos

$$P(X_1 \leq x, Y_1 \leq x) \leq \mu_N([0,2x]) + \int_{(2x,\infty)} \frac{(N-1)}{N}\frac{2x^2}{m^2}\mathrm{d}\mu_N(m) \leq$$

$$\leq \mu_N([0,2x]) + \int_{(2x,\infty)} \frac{2x^2}{m^2}\mathrm{d}\mu_N(m).$$

Mas, para $m > 2x$, $\frac{2x^2}{m^2} \leq \frac{2x^2}{4x^2} = \frac{1}{2}$. Então,

$$P(X_1 \leq x, Y_1 \leq x) \leq \mu_N([0,2x]) + \int_{(2x,\infty)} \frac{1}{2}\mathrm{d}\mu_N(m) \Rightarrow$$

$$\Rightarrow P(X_1 \leq x, Y_1 \leq x) \leq \frac{1}{2} + \frac{1}{2}\mu_N([0,2x]), \quad \forall N \in \mathbb{N}, \ \forall x > 0. \qquad (1.20)$$

Como $P(X_1 \leq x, Y_1 \leq x) \to 1$, quando $x \to \infty$, temos que $\forall \varepsilon > 0$, $\exists x_0 \in \mathbb{R}_+$ tal que $\forall x \geq x_0$, $P(X_1 \leq x, Y_1 \leq x) > 1 - \frac{\varepsilon}{2}$. Assim, de (1.20) e tomando $x = x_0$, obtemos

$$1 - \frac{\varepsilon}{2} < P(X_1 \leq x_0, Y_1 \leq x_0) \leq \frac{1}{2} + \frac{1}{2}\mu_N([0,2x_0]) \Rightarrow$$

$$\Rightarrow \mu_N([0,2x_0]) > 1 - \varepsilon$$

e, portanto, $\{\mu_N\}_{N\geq 1}$ é "tight", implicando que existem uma subsequência $\{\mu_{N_k}\}_{k\geq 1}$ e uma medida de probabilidade μ^* tais que

$$\int_{\mathbb{R}_+} f\mathrm{d}\mu_{N_k} \xrightarrow{w} \int_{\mathbb{R}_+} f\mathrm{d}\mu^*, \text{ quando } k \to \infty,$$

para toda função $f : \mathbb{R}_+ \to \mathbb{R}_+$ limitada e contínua (em quase toda parte). Tomando $f(\theta) = P_\theta^n$, segue que

$$\int_{\mathbb{R}_+} P_\theta^n \mathrm{d}\mu_{N_k}(\theta) \xrightarrow{w} \int_{\mathbb{R}_+} P_\theta^n \mathrm{d}\mu^*(\theta), \text{ quando } k \to \infty,$$

isto é,

$$P_{\mu_{N_k},n} \xrightarrow{w} \int_{\mathbb{R}_+} P_\theta^n \mathrm{d}\mu^*(\theta), \text{ quando } k \to \infty. \tag{1.21}$$

De (1.18), (1.21) e da unicidade do limite de $\{P_{\mu_N,n}\}_{N\geq 1}$, resulta que

$$P_n = \int_{\mathbb{R}_+} P_\theta^n \mathrm{d}\mu^*(\theta).$$

Portanto, $\forall n \in \mathbb{N}$ e $\forall B \in \mathcal{B}((\mathbb{R}^2)^n)$, existe $\mu^* : (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)) \to [0,1]$ tal que

$$P(((X_1,Y_1),\ldots,(X_n,Y_n)) \in B) = \int_{\mathbb{R}_+} P_\theta^n(B)\mathrm{d}\mu^*(\theta)$$

Verifiquemos, finalmente, a unicidade de μ^*. Para $N \in \mathbb{N}$ e $t > 0$, temos

$$\mu_N([0,t]) = P\left(\max_{1\leq i\leq N}\{X_i + Y_i\} \leq t\right) = P\left(\bigcap_{i=1}^{N}\{X_i + Y_i \leq t\}\right) =$$

$$= P\left(\underbrace{\{((x_1,y_2),\ldots,(x_N,y_N)) \in (\mathbb{R}_+^2)^N : x_i + y_i \leq t,\ i = 1,\ldots,N\}}_{H_{Nt}}\right) =$$

$$= \int_{\mathbb{R}_+} P_\theta^N(H_{Nt})\mathrm{d}\mu^*(\theta) = \int_{\mathbb{R}_+} \prod_{i=1}^{N} P_\theta(\{X_i + Y_i \leq t\})\mathrm{d}\mu^*(\theta).$$

Mas, para $\theta \leq t$, $P_\theta(\{X_i + Y_i \leq t\}) = 1$ e, para $\theta > t$, $P_\theta(\{X_i + Y_i \leq t\}) = \left(\frac{t}{\theta}\right)^2$. Logo,

$$\mu_N([0,t]) = \int_{[0,t]} 1\mathrm{d}\mu^*(\theta) + \int_{(t,\infty)} \left(\frac{t}{\theta}\right)^{2N} \mathrm{d}\mu^*(\theta).$$

Pelo Teorema da Convergência Dominada, segue que a segunda parcela do lado direito da última igualdade tende a zero quando $N \to \infty$ e, conseqüentemente,

$$\lim_{N\to\infty} \mu_N([0,t]) = \int_{[0,1]} 1 \mathrm{d}\mu^*(\theta) = \mu^*([0,t]).$$

Assim, $\{\mu_N\}_{N\geq 1}$ possui limite, concluindo a prova. ■

A partir do Teorema 1.2.7 podemos derivar a representação preditivista para a distribuição Beta (2,1) com parâmetro de escala θ (dizemos que X tem distribuição Beta (2.1) com parâmetro de escala θ se $\frac{X}{\theta}$ tem distribuição Beta (2,1)), cuja função densidade de probabilidade é dada por

$$f(z|\theta) \doteq \frac{2z}{\theta^2} \mathbb{I}_{(0,\theta)}(z)$$

A obtenção deste resultado consiste basicamente na construção de um novo processo estocástico $\{Z_n\}_{n\geq 1}$ a partir de $\{(X_n, Y_n)\}_{n\geq 1}$ definindo $Z_n = X_n + Y_n$, $\forall n \in \mathbb{N}$, e na especificação de convenientes condições sobre $\{Z_n\}_{n\geq 1}$ resultantes da uniformidade sobre χ_m^N, $\forall N, m \in \mathbb{N}$, para o processo original $\{(X_n, Y_n)\}_{n\geq 1}$.

As formas finita e infinita do Teorema do tipo de De Finetti podem ser estendidas para o caso de um vetor aleatório d-variado, $d > 2$, do mesmo modo que na subseção 1.2.1. Podemos derivar resultados similares aos Teoremas 1.2.6 e 1.2.7 para o modelo cuja função densidade de probabilidade é dada por

$$f(x_1, \ldots, x_d|\theta) = \frac{d!}{\theta^d} \mathbb{I}_{\{(a_1,\ldots,a_d)\in\mathbb{R}_+^d : a_1+\cdots+a_d\leq\theta\}}(x_1, \ldots, x_d). \tag{1.22}$$

Os resultados para o modelo em (1.22) podem ser obtidos sem dificuldades procedendo-se como na subseção 1.2.1. Para encerrar esta seção, enunciamos a forma infinita do Teorema do tipo de De Finetti (sem prova) para o modelo acima.

Teorema 1.2.8 *Seja* $\{(X_n^1, \ldots, X_n^d)\}_{n\geq 1}$ *uma sequência de vetores aleatórios d-variados,* $d > 2$, *assumindo valores em* $\mathbb{R}_+^d$. *Suponha que para cada* $N \in \mathbb{N}$ *e* $m > 0$, *a distribuição condicional de* $((X_1^1, \ldots, X_1^d), \ldots, (X_N^1, \ldots, X_N^d))$, *dado* $\max_{1\leq i\leq N}\{X_i^1 + \cdots + X_i^d\} = m$, *seja uniforme sobre* $\{((x_1^1, \ldots, x_1^d), \ldots, (x_N^1, \ldots, x_N^d)) \in (\mathbb{R}_+^d)^N : \max_{1\leq i\leq N}\{x_i^1 + \cdots + x_i^d\} = m\}$. *Então, existe uma medida de probabilidade* μ^* *em* $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ *tal que* $\forall n \in \mathbb{N}$ *e* $\forall B \in \mathcal{B}((\mathbb{R}^d)^N)$

$$P(((X_1^1, \ldots, X_1^d), \ldots, (X_n^1, \ldots, X_n^d)) \in B) = \int_{\mathbb{R}_+} P_\theta^{n,d}(B)\mathrm{d}\mu^*(\theta),$$

onde $P_\theta^{n,d}$ *é a lei de n vetores aleatórios independentes e com distribuição comum uniforme em* $\{(a_1, \ldots, a_d) \in \mathbb{R}_+^d : a_1 + \cdots + a_d \leq \theta\}$.

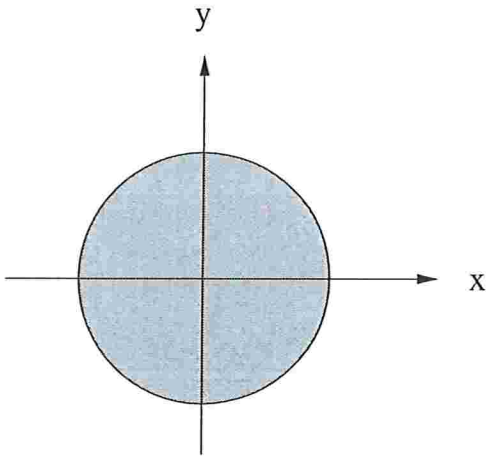
1.3 Distribuição Uniforme no Círculo

Nesta seção, obteremos a representação preditivista do modelo (B). Sua função densidade de probabilidade é dada por

$$f(x, y|\theta) = \frac{1}{\pi\theta^2}\mathbb{I}_{\{(a,b)\in\mathbb{R}^2 : a^2+b^2\leq\theta^2\}}(x, y). \tag{1.23}$$

O suporte de (1.23) é esboçado na sequência.

Procederemos do mesmo modo que na seção 1.2. Assim, vamos considerar o subconjunto de $(\mathbb{R}^2)^N$ $\chi_m^N = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{R}^2)^N : \max_{1\leq i\leq N}\{x_i^2 + y_i^2\} = m^2\}$, $m > 0$. Vamos definir uma distribuição uniforme sobre χ_m^N analogamente àquela especificada em (1.14). Para isto, consideremos os conjuntos $M_i(m) = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{R}^2)^N : x_i^2 + y_i^2 = m^2, x_j^2 + y_j^2 \leq m^2, j \neq i\}$, $i = 1, \ldots, N$, e as "projeções" φ^i e medidas μ_i, $i = 1, \ldots, N$, introduzidas na seção 1.2. Podemos, então, definir a distribuição uniforme sobre χ_m^N, $Q_{Nm} : \mathcal{B}((\mathbb{R}^2)^N) \to [0, 1]$, por



$$Q_{Nm}(B) = \frac{\sum_{i=1}^{N} \mu_i(B \cap \chi_m^N)}{N(\pi m^2)^{N-1}}, \quad B \in \mathcal{B}((\mathbb{R}^2)^N). \tag{1.24}$$

Pode-se facilmente checar que se um vetor aleatório $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído em χ_m^N, então suas marginais bivariadas $(X_1, Y_1), \ldots, (X_N, Y_N)$ são permutáveis.

Continuando, enunciamos agora três proposições que serão utilizadas na estimação da distância em variação total entre a lei da marginal n-dimensional de Q_{Nm}, $n < N$, a qual denotaremos por Q_{Nmn}, e a lei de n vetores aleatórios bidimensionais independentes com distribuição comum uniforme sobre $\{(a, b) \in \mathbb{R}^2 : a^2 + b^2 \leq m^2\}$.

Proposição 1.3.1 *Sejam $(U_1, V_1), \ldots, (U_N, V_N)$ vetores aleatórios independentes com distribuição comum uniforme sobre o conjunto $E = \{(a, b) \in \mathbb{R}^2 : a^2 + b^2 \leq 1\}$. Seja $S^{(N)} = \sqrt{\max_{1 \leq i \leq N}\{U_i^2 + V_i^2\}}$, $X_i = \frac{mU_i}{S^{(N)}}$ e $Y_i = \frac{mV_i}{S^{(N)}}$, $i = 1, \ldots, N$, com $m > 0$. Então, $((X_1, Y_1), \ldots, (X_N, Y_N))$ é o único vetor Q_{Nm}-uniformemente distribuído sobre χ_m^N.*

Prova Seja Q a lei de $((X_1, Y_1), \ldots, (X_N, Y_N))$. Claramente, $Q(\chi_m^N) = 1$. Para $A \in$

$\mathcal{B}((\mathbb{R}^2)^n)$,

$$Q(A) = P(((X_1, Y_1), \ldots, (X_N, Y_N)) \in A) = P\left(\bigcup_{i=1}^{N} A_i\right),$$

A_i denotando o evento $((X_1, Y_1), \ldots, (X_N, Y_N)) \in A \cap M_i(m)$, $i = 1, \ldots, N$. Então,

$$Q(A) = P\left(\bigcup_{i=1}^{N} A_i\right) = \sum_{i=1}^{N} P(A_i), \tag{1.25}$$

onde a última igualdade decorre do fato que $\lambda_{2N-2}(\varphi^i(M_i(m) \cap M_j(m))) = 0$, para $i \neq j$. Calculemos, então, $P(A_i)$.

$$P(A_i) = P(((X_1, Y_1), \ldots, (X_N, Y_N)) \in A \cap M_i(m)) =$$

$$= P(X_i^2 + Y_i^2 = m^2, \varphi^i(((X_1, Y_1), \ldots, (X_N, Y_N))) \in \varphi^i(A \cap M_i(m))) =$$

$$= P\left(\frac{m}{S^{(N)}}\sqrt{U_i^2 + V_i^2} = m, \varphi^i\left(\frac{m}{S^{(N)}}((U_1, V_1), \ldots, (U_N, V_N))\right) \in \varphi^i(A \cap M_i(m))\right) =$$

$$= P\left(S^{(N)} = \sqrt{U_i^2 + V_i^2}, \varphi^i\left(\frac{m}{\sqrt{U_i^2 + V_i^2}}((U_1, V_1), \ldots, (U_N, V_N))\right) \in \varphi^i(A \cap M_i(m))\right) =$$

$$= P\left(\underbrace{\bigcap_{j \neq i}\{\sqrt{U_j^2 + V_j^2} \leq \sqrt{U_i^2 + V_i^2}\}}_{B}, \underbrace{\varphi^i\left(\frac{m}{\sqrt{U_i^2 + V_i^2}}((U_1, V_1), \ldots, (U_N, V_N))\right) \in \varphi^i(A \cap M_i(m))}_{C}\right)$$

Como na subseção 1.2.2, não é difícil verificar que $C \subseteq B$. Portanto,

$$P(A_i) = P\left(\varphi^i\left(\frac{m}{\sqrt{U_i^2 + V_i^2}}((U_1, V_1), \ldots, (U_N, V_N))\right) \in \varphi^i(A \cap M_i(m))\right) =$$

$$= \int_E P\left(\varphi^i\left(\frac{m}{\sqrt{U_i^2 + V_i^2}}((U_1, V_1), \ldots, (U_N, V_N))\right) \in \varphi^i(A \cap M_i(m)) | U_i = a, V_i = b\right) \frac{1}{\pi} da db,$$

pois (U_i, V_i) tem função densidade de probabilidade dada por (1.23) com $\theta = 1$. Pela independência dos vetores aleatórios $(U_1, V_1), \ldots, (U_N, V_N)$ segue que

$$P(A_i) = \int_E P\left(\varphi^i\left(\frac{m}{\sqrt{a^2 + b^2}}((U_1, V_1), \ldots, (U_N, V_N))\right) \in \varphi^i(A \cap M_i(m))\right) \frac{1}{\pi} da db$$

$$= \int_E \left[\int_{\varphi^i(A \cap M_i(m))} f_{\mathcal{Z}_{(a,b)}} \, \mathrm{d}\lambda_{2N-2} \right] \frac{1}{\pi} \mathrm{d}a\mathrm{d}b,$$

onde $f_{\mathcal{Z}_{(a,b)}}$ é a função densidade de probabilidade do vetor aleatório $(2N-2)$-dimensional

$$\mathcal{Z}_{(a,b)} = \varphi^i \left(\frac{m}{\sqrt{a^2+b^2}} ((U_1, V_1), \ldots, (U_N, V_N)) \right) =$$

$$= \frac{m}{\sqrt{a^2+b^2}} (U_1, V_1, \ldots, U_{i-1}, V_{i-1}, U_{i+1} V_{i+1}, \ldots, U_N, V_N).$$

Pela hipótese da proposição e aplicando o método do Jacobiano para transformações de vetores aleatórios, obtemos que

$$f_{\mathcal{Z}_{(a,b)}} = \frac{(a^2+b^2)^{N-1}}{m^{2N-2}} \left(\frac{1}{\pi} \right)^{N-1}.$$

Então,

$$P(A_i) = \int_E \left[\int_{\varphi^i(A \cap M_i(m))} \frac{(a^2+b^2)^{N-1}}{m^{2N-2}} \left(\frac{1}{\pi} \right)^{N-1} \mathrm{d}\lambda_{2N-2} \right] \frac{1}{\pi} \mathrm{d}a\mathrm{d}b =$$

$$= \int_E \frac{(a^2+b^2)^{N-1}}{m^{2N-2}\pi^N} \lambda_{2N-2}(\varphi^i(A \cap M_i(m)))\mathrm{d}a\mathrm{d}b \Rightarrow$$

$$\Rightarrow P(A_i) = \frac{\lambda_{2N-2}(\varphi^i(A \cap \chi_m^N \cap M_i(m)))}{N(\pi m^2)^{N-1}} = \frac{\mu_i(A \cap \chi_m^N)}{N(\pi m^2)^{N-1}}.$$

Substituindo a expressão acima em (1.25), resulta que

$$Q(A) = \frac{\sum_{i=1}^N \mu_i(A \cap \chi_m^N)}{N(\pi m^2)^{N-1}} = \tilde{Q}_{Nm}(A \cap \chi_m^N) = Q_{Nm}(A).$$

Assim, $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído sobre χ_m^N. ■

Como podemos ver, a proposição 1.3.1 fornece uma caracterização da lei Q_{Nm} em termos de vetores aleatórios independentes e identicamente distribuídos.

Apresentamos, na sequência, a função de distribuição da marginal unidimensional de Q_{Nm} (como antes, para $1 < n < N$, denotamos a marginal n- dimensional de Q_{Nm} por Q_{Nmn}).

Proposição 1.3.2 *Seja $((X_1, Y_1), \ldots, (X_N, Y_N))$ um vetor aleatório uniformemente distribuído sobre χ_m^N. Então, a função de distribuição de sua marginal unidimensional, (X_1, Y_1), é dada por*

$$P(X_1 \leq x, Y_1 \leq y) = \begin{cases} 0, se \ \min\{x, y\} < -m \ ou \ x, y \leq 0 \ com \ x^2 + y^2 > m^2 \\ \dfrac{1}{N} + \dfrac{(N-1)\lambda_2(((-\infty, x] \times (-\infty, y]) \cap B_m)}{N\pi m^2}, \ c.c.. \end{cases}$$

Prova Tomemos $x, y \in \mathbb{R}$.

$$P(X_1 \leq x, Y_1 \leq y) = Q_{Nm1}((-\infty, x] \times (-\infty, y]) =$$

$$= Q_{Nm}(\underbrace{((-\infty, x] \times (-\infty, y]) \times (\mathbb{R}^2)^{N-1}}_{D})$$

Da definição de Q_{Nm}, temos

$$Q_{Nm}(D) = \frac{\sum_{i=1}^{N} \lambda_{2N-2}(\varphi^i(D \cap \chi_m^N \cap M_i(m)))}{N(\pi m^2)^{N-1}} =$$

$$= \frac{\lambda_{2N-2}(\varphi^1(D \cap M_1(m))) + \sum_{i=2}^{N} \lambda_{2N-2}(\varphi^i(D \cap M_i(m)))}{N(\pi m^2)^{N-1}}$$

Se $D \cap M_1(m) \neq 0$, então $\varphi^1(D \cap \chi_m^N \cap M_1(m)) = B_m^{N-1}$, onde $B_m = \{(a, b) \in \mathbb{R}^2 : a^2 + b^2 \leq m^2\}$. Além disso, $\varphi^i(D \cap \chi_m^N \cap M_i(m)) = ((((-\infty, x] \times (-\infty, y]) \cap B_m)) \times B_m^{N-2}$, para $i \neq 1$. Notando que $D \cap M_1(m) = \phi$ se, e somente se, $\min\{x, y\} < -m$ ou $x, y \leq 0$ com $x^2 + y^2 > m^2$, segue que

$$Q_{Nm}(D) = \frac{\lambda_{2N-2}(B_m^{N-1}) + \sum_{i=2}^{N} \lambda_{2N-2}(((((-\infty, x] \times (-\infty, y]) \cap B_m)) \times B_m^{N-2})}{N(\pi m^2)^{N-1}}$$

$$= \frac{[\lambda_2(B_m)]^{N-1} + \sum_{i=2}^{N} \lambda_2(((-\infty, x] \times (-\infty, y]) \cap B_m)[\lambda_2(B_m)]^{N-2}}{N(\pi m^2)^{N-1}} =$$

$$= \frac{(\pi m^2)^{N-1} + \sum_{i=2}^{N} \lambda_2(((-\infty, x] \times (-\infty, y]) \cap B_m)(\pi m^2)^{N-2}}{N(\pi m^2)^{N-1}} \Rightarrow$$

$$\Rightarrow \; Q_{Nm}(D) = \frac{1}{N} + \frac{(N-1)\lambda_2(((-\infty, x] \times (-\infty, y]) \cap B_m)}{N\pi m^2},$$

se $D \cap M_1(m) \neq \emptyset$ (caso contrário, a parcela $\frac{1}{N}$ não figura na expressão de $Q_{Nm}(D)$).

■

Em particular, se $x \in \left(0, \frac{m\sqrt{2}}{2}\right)$, temos

$$P(X_1 \leq x, Y_1 \leq x) = \frac{1}{N} + \left(\frac{N-1}{N}\right)\left(\frac{3}{4} + \frac{x^2}{\pi m^2} - \frac{\operatorname{arcos}\left(\frac{x}{m}\right)}{\pi} + \frac{x}{\pi}\frac{\sqrt{m^2 - x^2}}{m^2}\right).$$

Proposição 1.3.3 *Seja $((X_1, Y_1), \ldots, (X_n, Y_n))$ um vetor aleatório com lei Q_{Nmn}. Então, a função de distribuição de $T^{(n)} = \max_{1 \leq i \leq n}\{\sqrt{X_i^2 + Y_i^2}\}$ é dada por*

$$P(T^{(n)} \leq t) = \begin{cases} 0 & , t < 0, \\ \dfrac{N-n}{N}\left(\dfrac{t}{m}\right)^{2n} & , 0 \leq t < m, \\ 1 & , t \geq m. \end{cases}$$

A demonstração desta proposição é omitida pois é similar à prova da Proposição 1.2.3.

Proposição 1.3.4 *Seja $((X_1, Y_1), \ldots, (X_N, Y_N))$ um vetor aleatório com lei Q_{Nmn}, dado $m > 0$. Então, $T^{(n)} = \max_{1 \leq i \leq n}\{\sqrt{X_i^2 + Y_i^2}\}$ é uma estatística suficiente para a família $\{Q_{Nmn} : m > 0\}$.*

A demonstração da Proposição 1.3.4 é aqui omitida. Estimamos agora a distância em variação total entre Q_{Nmn} e a lei de n vetores aleatórios bivariados independentes e identicamente distribuídos uniformemente sobre $\{(a,b) \in \mathbb{R}^2 : a^2 + b^2 \leq m^2\}$, a qual denotamos por P_m^n.

Teorema 1.3.1 *Sejam Q_{Nmn} e P_m^n as medidas de probabilidade em $((\mathbb{R}^2)^n, \mathcal{B}((\mathbb{R}^2)^n))$ definidas anteriormente. Então, para $1 \leq n < N$,*

$$\|Q_{Nmn} - P_m^n\| \leq \frac{4n}{N}.$$

Prova É fácil verificar que $T^{(n)} = \max_{1 \leq i \leq n}\{\sqrt{X_i^2 + Y_i^2}\}$ é suficiente para a família $\{P_m^n : m > 0\}$. Assim, $T^{(n)}$ é suficiente para $\{Q_{Nmn} : m > 0\}$ e $\{P_m^n : n > 0\}$, o que implica que

$$\|Q_{Nmn} - P_m^n\| = \|Q_{Nmn,T} - P_{m,T}^n\|,$$

onde $Q_{Nmn,T}(P_{m,T}^n)$ é a $Q_{Nmn}(P_m^n)$-lei de $T^{(n)}$. Assim, vamos avaliar $\|Q_{Nmn,T} - P_{m,T}^n\|$. Para $B \in \mathcal{B}(\mathbb{R})$,

$$Q_{Nmn,T}(B) = \int_{B\cap(0,m)} \frac{N-n}{N}\frac{2na^{2n-1}}{m^{2n}}\mathrm{d}a + Q_{Nmn,T}(B \cap \{m\}) \text{ e}$$

$$P_{m,T}^n(B) = \int_{B\cap(0,m)} \frac{2na^{2n-1}}{m^{2n}}\mathrm{d}a.$$

Logo,

$$\|Q_{Nmn,T} - P_{m,T}^n\| = 2 \sup_{B\in\mathcal{B}(\mathbb{R})} |Q_{Nmn,T}(B) - P_{m,T}^n(B)| =$$

$$= 2 \sup_{B\in\mathcal{B}(\mathbb{R})} \left| \int_{B\cap(0,m)} \frac{N-n}{N}\frac{2na^{2n-1}}{m^{2n}}\mathrm{d}a + Q_{Nmn,T}(B \cap \{m\}) - \int_{B\cap(0,m)} \frac{2na^{2n-1}}{m^{2n}}\mathrm{d}a \right|$$

$$\leq 2 \sup_{B\in\mathcal{B}(\mathbb{R})} \left| \int_{B\cap(0,m)} \left(-\frac{n}{N}\right) \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a \right| 2 \sup_{B\in\mathcal{B}(\mathbb{R})} |Q_{Nmn,T}(B\cap\{m\})| \leq$$

$$\leq \frac{2n}{N} \int_{(0,m)} \frac{2na^{2n-1}}{m^{2n}} \mathrm{d}a + 2Q_{Nmn,T}(\{m\}) \leq \frac{4n}{N} \Rightarrow$$

$$\Rightarrow \|Q_{Nmn} - P_m^n\| = \|Q_{Nmn,T} - P_{m,T}^n\| \leq \frac{4n}{N}.$$

■

Finalmente, obtemos a versão finita do Teorema do tipo de De Finetti para o modelo (B). Denotemos por P_N a lei de um vetor aleatório N-dimensional $((X_1, Y_1), \ldots, (X_N, Y_N))$, cuja distribuição condicional, dado $\max_{1\leq i\leq N}\{\sqrt{X_i^2 + Y_i^2}\} = m$, $m > 0$, é uniforme sobre χ_m^N. Seja $P_{\mu,n} = \int_{\mathbb{R}} P_m^n \mathrm{d}\mu(m)$ uma medida de probabilidade em $((\mathbb{R}^2)^n, \mathcal{B}((\mathbb{R}^2)^n))$ para alguma medida de probabilidade μ em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. Além disso, sejam P_n a lei da marginal n-dimensional de P_n e μ_N a lei de $T^{(N)} = \max_{1\leq i\leq N}\{\sqrt{X_i^2 + Y_i^2}\}$. Enunciamos o seguinte teorema.

Teorema 1.3.2 *Sejam P_N e $P_{\mu,n}$ as medidas de probabilidades definidas acima. Então, existe uma medida de probabilidade μ_0 em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ tal que, para $1 \leq n < N$,*

$$\|P_n - P_{\mu_0,n}\| \leq \frac{4n}{N}.$$

Prova Da definição de P_n, segue que

$$P_n = \int_{\mathbb{R}_+} Q_{Nmn} \mathrm{d}\mu_N(m),$$

uma vez que $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído sobre χ_m^N, dado $T^{(N)} = m$. Assim, para $\mu : \mathcal{B}(\mathbb{R}) \to [0, 1]$, temos que

$$\|P_n - P_{\mu,n}\| = 2 \sup_{B\in\mathcal{B}((\mathbb{R}^2)^n)} |P_n(B) - P_{\mu,n}(B)| =$$

$$= 2 \sup_{B\in\mathcal{B}((\mathbb{R}^2)^n)} \left| \int_{\mathbb{R}_+} Q_{Nmn}(B)\mathrm{d}\mu_N - \int_{\mathbb{R}_+} P_m^n(B)\mathrm{d}\mu \right|.$$

Tomando $\mu = \mu_N$, obtemos que

$$\|P_n - P_{\mu,n}\| = 2 \sup_{B\in\mathcal{B}((\mathbb{R}^2)^n)} \left| \int_{\mathbb{R}_+} [Q_{Nmn}(B) - P_m^n(B)]\mathrm{d}\mu_N(m) \right| \leq$$

$$\leq 2 \sup_{B\in\mathcal{B}((\mathbb{R}^2)^n)} \int_{\mathbb{R}_+} |Q_{Nmn}(B) - P_m^n(B)|\mathrm{d}\mu_N(m) \leq \frac{4n}{N},$$

onde a última desigualdade segue do resultado do Teorema 1.3.1. Portanto, existe uma medida de probabilidade $\mu_0 = \mu_N$ em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ tal que $\|P_n - P_{\mu_0,n}\| \leq \frac{4n}{N}$, para $1 \leq n < N$. ∎

Na sequência, derivamos a forma infinita do Teorema do tipo de De Finetti para o modelo (B).

Teorema 1.3.3 *Seja $\{(X_n, Y_n)\}_{n\geq 1}$ uma sequência de vetores aleatórios bivariados assumindo valores em $\mathbb{R}^2$. Suponha que para cada $N \in \mathbb{N}$ e $m > 0$, a distribuição condicional de $((X_1, Y_1), \ldots, (X_N, Y_N))$, dado $\max_{1\leq i\leq N}\{\sqrt{X_i^2 + Y_i^2}\} = m$, seja uniforme em χ_m^N. Então, existe uma única medida de probabilidade μ^* em $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ tal que $\forall n \in \mathbb{N}$ e $B \in \mathcal{B}((\mathbb{R}^2)^n)$*

$$P(((X_1, Y_1), \ldots, (X_N, Y_N)) \in B) = \int_{\mathbb{R}_+} P_\theta^n(B)\mathrm{d}\mu^*(\theta). \tag{1.26}$$

Prova Seja $\mathcal{M} = \{P : ((\mathbb{R}^2)^n, \mathcal{B}((\mathbb{R}^2)^n)) \to [0, 1]\}$ e denotemos por $\| \ \|$ a distância em variação total em $\mathcal{M}^2$. Provaremos que os dois lados da igualdade em (1.26) correspondem

ao limite da sequência $\{P_{\mu_N,n}\}_{N\geq 1}$, no espaço métrico $(\mathcal{M}, \| \ \|)$. Primeiro, tomando limites no resultado do Teorema 1.3.2, com $\mu_0 = \mu_N$, obtemos que

$$\lim_{N\to\infty} \|P_n - P_{\mu_N,n}\| = 0,$$

ou,

$$P_{\mu_N,n} \xrightarrow{\mu} P_n, \text{ quando } N \to \infty, \tag{1.27}$$

onde P_n é a distribuição de $((X_1, Y_1), \ldots, (X_N, Y_N))$. Em seguida, verificamos que $\{\mu_N\}_{N\geq 2}$ é "tight", o que implicará a convergência fraca de $\{P_{\mu_N,n}\}_{N\geq 1}$. Para $x > 0$ e $\forall N \geq 2$, temos que

$$P(X_1 \leq x, Y_1 \leq x) = \int_{\mathbb{R}_+} \underbrace{P\left(X_1 \leq x, Y_1 \leq x \mid \max_{1\leq i\leq N}\{\sqrt{X_i^2 + Y_i^2}\} = m\right)}_{Q_{Nm1}((-\infty,x]^2)} \mathrm{d}\mu_N(m) =$$

$$= \int_{[0,x\sqrt{2}]} Q_{Nm1}((-\infty, x]^2)\mathrm{d}\mu_N(m) + \int_{(x\sqrt{2},\infty)} Q_{Nm1}((-\infty, x]^2)\mathrm{d}\mu_N(m) \Rightarrow$$

$$\Rightarrow P(X_1 \leq x, Y_1 \leq x) \leq \int_{[0,x\sqrt{2}]} 1\mathrm{d}\mu_N(m) + \int_{(x\sqrt{2},\infty)} Q_{Nm1}((-\infty, x]^2)\mathrm{d}\mu_N(m). \tag{1.28}$$

Para $m > x\sqrt{2}$, temos

$$Q_{Nm1}((-\infty, x]^2) = \frac{(N-1)}{N}\left[\frac{3}{4} + \frac{x^2}{\pi m^2} - \frac{\arccos(\frac{x}{m})}{\pi} + \frac{x}{\pi}\frac{\sqrt{m^2 - x^2}}{m^2}\right] + \frac{1}{N},$$

conforme a proposição 1.3.2. Voltando para (1.28), obtemos que

$$P(X_1 \leq x, Y_1 \leq x) \leq$$

$$\leq \mu_N([0, x\sqrt{2}]) + \int_{(x\sqrt{2},\infty)} \frac{(N-1)}{N}\left[\frac{3}{4} + \frac{1}{2\pi} - \frac{\arccos(\frac{x}{m})}{\pi} + \frac{x}{\pi}\frac{\sqrt{m^2 - x^2}}{m^2}\right] + \frac{1}{N}\mathrm{d}\mu_N(m)$$

$$\leq \mu_N([0, x\sqrt{2}]) + \int_{(x\sqrt{2},\infty)} \frac{(N-1)}{N}\left[\frac{3}{4} + \frac{1}{2\pi} - \frac{1}{4} + \frac{x}{\pi}\frac{\sqrt{m^2-x^2}}{m^2}\right] + \frac{1}{N}\mathrm{d}\mu_N(m),$$

uma vez que para $m > x\sqrt{2}$, $\arccos\left(\frac{x}{m}\right) > \frac{\pi}{4}$. Então,

$$P(X_1 \leq x, Y_1 \leq x) \leq \mu_N([0, x\sqrt{2}]) + \int_{(x\sqrt{2},\infty)} \frac{(N-1)}{N}\left[\frac{1}{2} + \frac{1}{2\pi} + \frac{x}{\pi}\frac{\sqrt{m^2-x^2}}{m^2}\right] + \frac{1}{N}\mathrm{d}\mu_N(m)$$

$$\leq \mu_N([0, x\sqrt{2}]) + \int_{(x\sqrt{2},\infty)} \frac{(N-1)}{N}\left[\frac{1}{2} + \frac{1}{2\pi} + \frac{1}{2\pi}\right] + \frac{1}{N}\mathrm{d}\mu_N(m),$$

pois $\sup\left\{\frac{\sqrt{m^2-x^2}}{m^2} : m > x\sqrt{2}\right\} = \frac{1}{2x}$. Assim,

$$P(X_1 \leq x, Y_1 \leq x) \leq \mu_N([0, x\sqrt{2}]) + \int_{(x\sqrt{2},\infty)} \left(\frac{5N+1}{6N}\right)\mathrm{d}\mu_N(m) \leq$$

$$\leq \mu_N([0, x\sqrt{2}]) + \frac{11}{12}\mu_N((x\sqrt{2}, \infty)) \Rightarrow$$

$$\Rightarrow P(X_1 \leq x, Y_1 \leq x) \leq \frac{1}{12}\mu_N([0, x\sqrt{2}]) + \frac{11}{12}$$

Mas, $P(X_1 \leq x, Y_1 \leq x) \to 1$, quando $x \to \infty$. Portanto, $\forall \varepsilon > 0$, $\exists x_0 \in \mathbb{R}_+$ tal que para $x \geq x_0$, $P(X_1 \leq x, Y_1 \leq x) > 1 - \frac{\varepsilon}{12}$. Assim, tomando $x = x_0$, obtemos que

$$1 - \frac{\varepsilon}{12} < P(X_1 \leq x_0, Y_1 \leq x_0) \leq \frac{1}{12}\mu_N([0, x_0\sqrt{2}]) + \frac{11}{12} \Rightarrow$$

$$\Rightarrow \mu_N([0, x_0\sqrt{2}]) \geq 1 - \varepsilon, \quad \forall N \geq 2.$$

Assim, $\{\mu_N\}_{N\geq 2}$ é "tight". Portanto, existem uma medida de probabilidade $\mu^* : \mathcal{B}(\mathbb{R}_+) \to [0,1]$ e uma subsequência $\{\mu_{N_k}\}_{k\geq 1}$ tais que

$$\int_{\mathbb{R}_+} f\mathrm{d}\mu_{N_k} \xrightarrow{w} \int_{\mathbb{R}_+} f\mathrm{d}\mu^*, \text{ quando } k \to \infty,$$

para toda função $f : \mathbb{R}_+ \to \mathbb{R}_+$ limitada e contínua. Tomando $f = P_\theta^n$, segue que

$$\int_{R_+} P_\theta^n \mathrm{d}\mu_{N_k}(\theta) \xrightarrow{w} \int_{\mathbb{R}_+} P_\theta^n \mathrm{d}\mu^*(\theta), \text{ quando } k \to \infty,$$

isto é,

$$P_{\mu_{N_k},n} \xrightarrow{w} \int_{R_+} P_\theta^n \mathrm{d}\mu^*(\theta), \text{ quando } k \to \infty. \tag{1.29}$$

De (1.27), (1.29) e da propriedade de unicidade do limite de $\{P_{\mu_N,n}\}_{N\geq 1}$, temos, finalmente, que

$$P_n = \int_{\mathbb{R}_+} P_\theta^n \mathrm{d}\mu^*(\theta)$$

Por fim, verificamos a unicidade de μ^*. Para $N \in \mathbb{N}$ e $t > 0$, temos

$$\mu_N([0,t]) = P\left(\max_{1\leq i\leq N}\{\sqrt{X_i^2+Y_i^2}\} \leq t\right) =$$

$$= P\left(\bigcap_{i=1}^N \{\sqrt{X_i^2+Y_i^2} \leq t\}\right) = P\left(\bigcap_{i=1}^N \{X_i^2+Y_i^2 \leq t^2\}\right) =$$

$$= P\left(\underbrace{\{((x_1,y_1),\ldots,(x_n,y_n)) \in (\mathbb{R}^2)^N : x_i^2+y_i^2 \leq t^2,\ i=1,\ldots,N\}}_{H_{Nt}}\right) =$$

$$= \int_{\mathbb{R}_+} P_\theta^N(H_{Nt})\mathrm{d}\mu^*(\theta) = \int_{R_+} \prod_{i=1}^N P_\theta(\{X_i^2+Y_i^2 \leq t^2\})\mathrm{d}\mu^*(\theta).$$

Para $\theta \leq t$, $P_\theta(\{X_i^2+Y_i^2 \leq t^2\}) = 1$ e, para $\theta > t$, $P_\theta(\{X_i^2+Y_i^2 \leq t^2\}) = \left(\frac{t}{\theta}\right)^2$. Logo,

$$\mu_N([0,t]) = \int_{[0,t]} 1\ \mathrm{d}\mu^*(\theta) + \int_{(t,\infty)} \left(\frac{t}{\theta}\right)^{2N} \mathrm{d}\mu^*(\theta).$$

Mas, pelo Teorema da Convergência Dominada, temos que

$$\lim_{N\to\infty} \int_{(t,\infty)} \left(\frac{t}{\theta}\right)^{2N} \mathrm{d}\mu^*(\theta) = 0, \text{ quando } N \to \infty,$$

e, portanto,

$$\lim_{N\to\infty} \mu_N([0,t]) = \mu^*([0,t]),$$

concluindo a demonstração. ■

Os resultados dos Teoremas 1.3.2 e 1.3.3 podem ser estendidos para o caso de um vetor aleatório d-variado, $d > 2$, como nas subseções 1.2.1 e 1.2.2. Neste caso, o modelo paramétrico tem a seguinte forma (função densidade de probabilidade)

$$f(x_1,\ldots,x_d|\theta) = \frac{d\Gamma\left(\frac{d}{2}\right)}{2(\pi)^{d/2}} \mathbb{I}_{\{(a_1,\ldots,a_d)\in\mathbb{R}^d : a_1^2+\cdots+a_d^2\leq\theta^2\}}(x_1,\ldots,x_d), \tag{1.30}$$

onde $\Gamma(t) = \int_0^\infty e^{-s}s^{t-1}\mathrm{d}s$ é a função gama usual.

Para concluir esta seção, enunciamos (sem demonstração) a forma infinita do Teorema do tipo de De Finetti para o modelo em (1.30).

Teorema 1.3.4 *Seja $\{(X_n^1,\ldots,X_n^d)\}_{n\geq 1}$ uma sequência de vetores aleatórios d-variados, $d > 2$, assumindo valores em $\mathbb{R}^d$. Suponha que para cada $N \in \mathbb{N}$ e $m > 0$, a distribuição de $((X_1^1,\ldots,X_1^d),\ldots,(X_N^1,\ldots,X_N^d))$ condicional a $\max_{1\leq i\leq N}\{\sqrt{(X_i^1)^2+\cdots+(X_i^d)^2}\} = m$, seja uniforme em $\{((x_1^1,\ldots,x_1^d),\ldots,(x_N^1,\ldots,x_N^d)) \in (\mathbb{R}^d)^N : \max_{1\leq i\leq N}\{\sqrt{(x_i^1)^2+\cdots+(x_i^d)^2}\} = m\}$. Então, existe uma medida de probabilidade μ^* em $(\mathbb{R}_+,\mathcal{B}(\mathbb{R}_+))$ tal que $\forall n \in \mathbb{N}$ e $\forall B \in \mathcal{B}((\mathbb{R}^d)^n)$*

$$P(((X_1^1,\ldots,X_1^d),\ldots,(X_n^1,\ldots,X_n^d)) \in B) = \int_{\mathbb{R}_+} P_\theta^{n,d}(B)\mathrm{d}\mu^*(\theta),$$

onde $P_\theta^{n,d}$ é a lei de n vetores aleatórios independentes com lei comum uniforme sobre $\{(a_1,\ldots,a_d) \in \mathbb{R}^d : a_1^2+\cdots+a_d^2 \leq \theta^2\}$.

1.4 Conclusões

Neste capítulo, apresentamos a representação preditivista para modelos uniformes bivariados dependendo de um único parâmetro. Conforme mencionado anteriormente, a representação preditivista de um modelo paramétrico tem um caráter relevante para a escola Bayesiana preditivista de inferência estatística, pois torna plausível, sob o ponto de vista Bayesiano, a adoção de tal modelo paramétrico como a distribuição amostral (condicional) para dados gerados em certos experimentos aleatórios.

Os resultados obtidos para o caso bivariado podem ser estendidos, de maneira bem simples, para os correspondentes modelos d-variados, $d > 2$ (independentemente de d, estes modelos, cujos espaços paramétricos são unidimensionais, apresentam estatísticas suficientes mínimas também unidimensionais). No entanto, a introdução de outro(s) parâmetro(s) nestes modelos dificulta a obtenção de suas respectivas representações preditivistas, gerando problemas interessantes que descrevemos na sequência.

A versão contínua do modelo (A), cujo suporte $\{(x,y) \in \mathbb{R}_+^2 : \frac{x}{\theta} + \frac{y}{\theta} \leq 1\}$, $\theta > 0$ corresponde a um triângulo isósceles com vértices $(0,0), (\theta,0)$ e $(0,\theta)$ em $\mathbb{R}^2$, pode ser modificado de modo a ter como novo suporte $\{(x,y) \in \mathbb{R}_+^2 : \frac{x}{\theta_1} + \frac{y}{\theta_2} \leq 1\}$, $\theta_1, \theta_2 > 0$, um triângulo escaleno de vértices $(0,0), (\theta_1,0)$ e $(0,\theta_2)$. Neste caso, o novo modelo, cujo espaço paramétrico é $\mathbb{R}_+^2$, não pode ter sua representação preditivista determinada através da abordagem utilizada ao longo deste capítulo, uma vez que possui estatística suficiente mínima de dimensão maior que 2. Acreditamos que outras técnicas, porém, possam ser desenvolvidas de modo a obter tal resultado.

Ao modelo (B), cujo suporte $\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq \theta^2\}$, $\theta > 0$, corresponde a um círculo de centro $(0,0)$ em $\mathbb{R}^2$ e raio $\theta > 0$, podem ser incorporados parâmetros de modo que seu novo suporte seja um círculo de centro agora desconhecido em $\mathbb{R}^2$ e raio

(desconhecido) $\theta > 0$, a saber, $\{(x,y) \in \mathbb{R}^2 : (x-\alpha)^2 + (y-\beta)^2 \leq \theta^2\}$, $\alpha, \beta \in \mathbb{R}$. Assim, o modelo obtido tem como estatística suficiente mínima o envoltório convexo (ou os pontos extremais) da amostra, que não possui a mesma dimensão do novo espaço paramétrico $\mathbb{R}_+ \times \mathbb{R}^2$, inviabilizando, pois, o emprego da técnica aqui utilizada (vários modelos discretos não contemplados neste trabalho como, por exemplo, a distribuição Poisson bivariada de Holgate, também apresentam esta característica). Novamente, como no caso da versão contínua do modelo (A), um enfoque diferente precisa ser desenvolvido para a obtenção de sua representação preditivista.

Referências

[1] Arellano-Valle, R.; Bolfarine, H.; Iglesias, P. (1994). *A predictivistic interpretation to the multivariate t distribution.* Test, 3, 1–16.

[2] Bolfarine, H.; Gasco, L.; Iglesias, P. (1999). *Pearson Type II Errors-in-Variables Models.* Em Bayesian Statistics, 6, 713–722. Oxford University Press.

[3] Dawid, A.P. (1977). *Invariant distributions and analysis of variance models.* Biometrika, 64(2), 291–297.

[4] de Finetti, B. (1937). *Le prévision: ses lois logiques, ses sources subjectives.* Ann. Inst. H. Poincaré. Prob. Statist., 7, 1–68.

[5] Diaconis, P.; Freedman, D. (1984). *Partial exchangeability and sufficient statistics: Applications and new directions.* (J. K. Ghosh and J. Roy, eds.) 205–274.

[6] Fortini, S.; Ladelli, L.; Regazzini, E. (2000). *Exchangeability, predictive distributions and parametric models.* Sankhyã, 62(A), 86–109.

[7] Freedman, D.; Diaconis, P. (1980). *De Finetti's theorem for Markov chains.* The Annals of Probability, 8(1), 115–130.

[8] Freedman, D.; Diaconis, P. (1990). *Cauchy's equation and de Finetti's theorem.* Scandinavian Journal of Statistics, 17, 235–250.

[9] Gnedin, A.V. (1995). *On a class of exchangeable sequences.* Statist. Probab. Letters, 25, 351–355

[10] Iglesias, P. (1993). *Formas finitas do Teorema de De Finetti: A visão preditivista da Inferência Estatística em populações finitas.* Tese de doutorado, Instituto de Matemática e Estatística - Univ. de São Paulo.

[11] Iglesias, P.; Pereira, C.A.B.; Tanaka, N.I. (1998). *Characterizations of multivariate spherical distributions in l_∞-norm.* Test, 7(2), 307–324.

[12] Lauritzen, S.L. (1988). *Extremal families and systems of sufficient statistics.* Springer Lecture Notes in Statistics, vol. 49, Springer, Berlin.

Capítulo 2

Caracterização de variáveis aleatórias uniformes independentes e identicamente distribuídas

2.1 Introdução

Caracterizações de distribuições simétricas multivariadas têm sido desenvolvidas por diversos autores nas últimas décadas (segundo Fang et al. (1990), os primeiros trabalhos nesta área foram publicados na segunda metade do século XIX), visando não apenas a extensão de distribuições univariadas conhecidas as suas respectivas versões multivariadas, mas também a consideração de novos modelos multivariados, que não o normal, para fins de inferência. Fang et al. (1990) afirmam que "o aspecto multivariado caracteriza o modelo de realizações dependentes de experimentos aleatórios, enquanto que simetria (e o conceito estritamente relacionado de invariância) passa pelo coração da teoria e filosofia das leis da natureza e matemática". Efron (1969) reforça a importância de distribuições simétricas citando que "a condição mais fraca de simetria pode substituir, com sucesso, a distribuição normal ao menos no que diz respeito ao clássico teste t de Student".

Por outro lado, a especificação do tipo de simetria (em adição à permutabilidade) de uma distribuição multivariada permite situá-la numa certa classe de distribuições. As-

sim, por exemplo, a simetria descrita por invariância sob um grupo de transformações ortogonais caracteriza as distribuições esféricas (em particular, com a condição adicional de independência, obtemos uma caracterização de normalidade, isto é, a única medida produto dentro da classe de distribuições esféricas n-variadas, $n \in \mathbb{N}$, corresponde à distribuição de n variáveis aleatórias independentes com mesma distribuição normal $(0, \sigma^2)$, para algum $\sigma^2 > 0$). Em geral, a imposição de simetria e condições adicionais a uma distribuição simétrica multivariada permite a caracterização de suas marginais, como no exemplo anterior. Devemos destacar que, de um modo geral, um particular tipo de simetria pode ser especificado através da função de distribuição, da função característica, da função densidade de probabilidade, ou da estatística suficiente do modelo sob consideração, dentre outras. Uma detalhada revisão sobre distribuições simétricas multivariadas em geral é apresentada em Fang et al. (1990).

Neste capítulo, apresentamos caracterizações de conhecidas distribuições uniformes (univariadas e bivariadas) dentro de classes de distribuições multivariadas satisfazendo específicas condições de simetria. Mais precisamente, consideramos as classes de distribuições multivariadas cujas distribuições de probabilidade dependem apenas do máximo, do mínimo e do máximo, ou da máxima soma das coordenadas em cada componente do vetor aleatório. Com a suposição adicional de independência, conseguimos caracterizar certos modelos uniformes. Na seção 2.2, tratamos dos casos discretos. A seção 2.3 é dedicada às respectivas versões contínuas.

Ao longo deste capítulo adotaremos a notação que segue. Para um conjunto não-vazio $A \subset \mathbb{R}$, A^N denotará o N-ésimo produto cartesiano de A e A_+ será o conjunto $\{a \in A : a > 0\}$. Além disso, $\mathcal{P}(A)$ indicará o conjunto das partes de A. Para o conjunto $\{a_1, a_2, \ldots, a_n\}$, $a_{(1)}$ e $a_{(n)}$ denotarão, respectivamente, o mínimo e o máximo de tal conjunto. Por fim, representaremos a σ-álgebra de Borel em $\mathbb{R}^N$ por $\mathcal{B}(\mathbb{R}^N)$, a medida

de Lebesgue em $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ por λ_N, $N \in \mathbb{N} = \{0, 1, \ldots\}$, e a medida de Lebesgue em $((\mathbb{R}^k)^N, \mathcal{B}((\mathbb{R}^k)^N))$ por $\lambda_k^{(N)}$.

2.2 Distribuições uniformes discretas

Nesta seção, vamos apresentar teoremas de caracterização para variáveis aleatórias uniformes independentes e identicamente distribuídas assumindo valores em Z. Primeiro, exibimos os resultados relativos à distribuição uniforme no conjunto $\{0, \ldots, \theta\}$, para $\theta \in \mathbb{N}$, e, em seguida, para a distribuição uniforme em $\{\theta_1, \theta_1 + 1, \ldots, \theta_2\}$, para $\theta_1, \theta_2 \in Z$, com $\theta_1 < \theta_2$. Por fim, é derivada uma caracterização para vetores aleatórios bivariados independentes e com distribuição comum uniforme sobre o triângulo isósceles discreto $\{(a, b) \in \mathbb{N}^2 : a + b \leq \theta\}$, $\theta \in \mathbb{N}$.

Como cada um destes resultados (teoremas 2.2.1, 2.2.2 e 2.2.3 na sequência) incorpora, em suas hipóteses, condições de uniformidade, descrevemos, no início de cada subseção, a distribuição uniforme condicional a ser considerada.

2.2.1 Caso univariado

Consideremos o seguinte subconjunto de $\mathbb{N}^N$, para $N \geq 2$ e $m \in \mathbb{N}$

$$\chi_m^N = \{(x_1, \ldots, x_N) \in \mathbb{N}^N : \max_{1 \leq i \leq N} \{x_i\} = m\}.$$

Dizemos que $(X_1, \ldots, X_N)$ é uniformemente distribuído sobre χ_m^N (Iglesias, Pereira e Tanaka (1998)), com lei de probabilidade denotada por Q_{Nm}, se

$$\text{Prob}(X_1 = x_1, \ldots, X_N = x_N) = Q_{Nm}((x_1, \ldots, x_N)) = \frac{\mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i\})}{(m+1)^N - m^N}, \qquad (2.1)$$

onde o denominador do lado direito da igualdade acima é simplesmente a medida de contagem (densidade) de χ_m^N.

Iglesias, Pereira e Tanaka (1998) demonstram que se $(X_1, \ldots, X_N)$ é uniformemente distribuído em χ_m^N, então sua marginal n-dimensional $(X_1, \ldots, X_n)$, $n < N$, tem lei Q_{Nm}^n dada por

$$Q_{Nm}^n(x_1, \ldots, x_n) = \begin{cases} \dfrac{1}{(m+1)^n} \left\{ \dfrac{1 - \left(\frac{m}{m+1}\right)^{N-n}}{1 - \left(\frac{m}{m+1}\right)^N} \right\}, & \text{se } \max_{1 \leq i \leq n}\{x_i\} < m, \\ \dfrac{1}{(m+1)^n} \left\{ \dfrac{1}{1 - \left(\frac{m}{m+1}\right)^N} \right\}, & \text{se } \max_{1 \leq i \leq n}\{x_i\} = m. \end{cases}$$

A partir deste último resultado e das propriedades da distância em variação total, as formas finita e infinita do Teorema do tipo de De Finetti para o modelo uniforme em $\{0, \ldots, k\}$, $k \in \mathbb{N}$, podem ser obtidas (Iglesias, Pereira e Tanaka (1998)), a segunda sendo derivada da primeira fixando-se n e tomando-se limites quando $N \to \infty$.

O primeiro objetivo desta seção é fornecer uma caracterização das medidas de probabilidade que correspondem a misturas, no sentido usual, dos elementos da família $\{Q_{Nm} : m \in \mathbb{N}\}$. Consideremos, então, a classe $\mathcal{C}_N$ de medidas de probabilidade em $(\mathbb{N}^N, \mathcal{P}(\mathbb{N}^N))$ obtidas por mistura dos elementos de $\{Q_{Nm} : m \in \mathbb{N}\}$ na variável "radial" (Fang et al. (1990)).

Proposição 2.2.1 *Uma medida de probabilidade P pertence a $\mathcal{C}_N$ se, e somente se, existe uma função real não-negativa φ tal que*

$$P((x_1, \ldots, x_N)) = \varphi(\max_{1 \leq i \leq N}\{x_i\}) \ ,$$

para cada $(x_1, \ldots, x_N) \in \mathbb{N}^N$.

Prova ($\Rightarrow$) Se P pertence a $\mathcal{C}_N$, então, para $(x_1, \ldots, x_N) \in \mathbb{N}^N$, temos que

$$P((x_1, \ldots, x_N)) = \int_{\mathbb{N}} Q_{Nm}((x_1, \ldots, x_N)) \mathrm{d}\mu(m) = \int_{\mathbb{N}} \frac{\mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i\})}{(m+1)^N - m^N} \mathrm{d}\mu(m).$$

Como μ é uma medida de probabilidade em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, podemos reescrever a última integral acima como uma série. Então,

$$P((x_1, \ldots, x_N)) = \sum_{m=1}^{\infty} \frac{\mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i\})}{(m+1)^N - m^N} \mu(\{m\}). \tag{2.2}$$

Considerando a função não-negativa $\varphi : \mathbb{N} \to \mathbb{R}_+$ definida por

$$\varphi(t) = \sum_{m=1}^{\infty} \frac{\mathbb{I}_{\{m\}}(t)}{(m+1)^N - m^N} \mu(\{m\}), \quad t \in \mathbb{N},$$

obtemos de (2.2) que

$$P((x_1, \ldots, x_N)) = \varphi(\max_{1 \leq i \leq N}\{x_i\}).$$

($\Leftarrow$) Suponhamos agora que $\forall (x_1, \ldots, x_N) \in \mathbb{N}^N$,

$$P((x_1, \ldots, x_N)) = \varphi(\max_{1 \leq i \leq N}\{x_i\}).$$

Então, para $m \in \mathbb{N}$, temos que

$$P(X_{(N)} = m) = \sum_{\substack{(x_1, \ldots, x_N) \in \mathbb{N}^N: \\ \max\limits_{1 \leq i \leq N}\{x_i\} = m}} P((x_1, \ldots, x_N)) =$$

$$= \sum_{\substack{(x_1, \ldots, x_N) \in \mathbb{N}^N: \\ \max\limits_{1 \leq i \leq N}\{x_i\} = m}} \varphi(\max_{1 \leq i \leq N}\{x_i\}) = \sum_{\substack{(x_1, \ldots, x_N) \in \mathbb{N}^N: \\ \max\limits_{1 \leq i \leq N}\{x_i\} = m}} \varphi(m) \Rightarrow$$

$$\Rightarrow P(X_{(N)} = m) = \varphi(m) |\chi_m^N| , \tag{2.3}$$

Agora, vamos provar que $P(A) = \int_{\mathbb{N}} Q_{Nm}(A) \mathrm{d}\mu(m)$, $\forall A \in \mathcal{P}(\mathbb{N}^N)$, para uma específica μ em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

$$P(A) = \sum_{(x_1, \ldots, x_N) \in A} P(x_1, \ldots, x_N) = \sum_{(x_1, \ldots, x_N) \in A} \sum_{m=1}^{\infty} P((x_1, \ldots, x_N), X_{(N)} = m) =$$

$$= \sum_{(x_1,\ldots,x_N)\in A} \sum_{m=1}^{\infty} P(x_1,\ldots,x_N)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i\}) =$$

$$= \sum_{(x_1,\ldots,x_N)\in A} \sum_{m=1}^{\infty} \varphi(\max_{1\leq i\leq N}\{x_i\})\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i\}) =$$

$$= \sum_{(x_1,\ldots,x_N)\in A} \sum_{m=1}^{\infty} \varphi(m)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i\}) = \sum_{m=1}^{\infty} \sum_{(x_1,\ldots,x_N)\in A} \varphi(m)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i\}),$$

onde a última igualdade segue de uma aplicação do Teorema de Tonelli. Substituindo (2.3) na expressão acima obtemos que

$$P(A) = \sum_{m=1}^{\infty} \sum_{(x_1,\ldots,x_N)\in A} \frac{\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i\})}{|\chi_m^N|} P(X_{(N)} = m) =$$

$$= \sum_{m=1}^{\infty} \sum_{(x_1,\ldots,x_N)\in A} Q_{Nm}((x_1,\ldots,x_N))P(X_{(N)} = m) = \sum_{m=1}^{\infty} Q_{Nm}(A)P(X_{(N)} = m) \Rightarrow$$

$$\Rightarrow P(A) = \int_{\mathbb{N}} Q_{Nm}(A)\mathrm{d}\mu(m),$$

onde μ é a P-lei de $X_{(N)} = \max\{X_1,\ldots,X_N\}$. Portanto,

$$P = \int_{\mathbb{N}} Q_{Nm}\mathrm{d}\mu(m)$$

e, consequentemente, $P \in \mathcal{C}_N$. ■

A proposição 2.2.1 estabelece que uma medida de probabilidade em $(\mathbb{N}^N, \mathcal{P}(\mathbb{N}^N))$ é mistura dos elementos de $\{Q_{Nm} : m \in \mathbb{N}\}$ se, e somente se, corresponde à distribuição

de um vetor aleatório $(X_1, \ldots, X_N)$ a valores em $\mathbb{N}^N$ com distribuição dependendo de $(x_1, \ldots, x_N) \in \mathbb{N}^N$ somente através de $\max_{1 \leq i \leq N}\{x_i\}$.

No que segue, mostramos que a única medida produto em $\mathcal{C}_N$ corresponde à distribuição de n variáveis aleatórias independentes e identicamente distribuídas uniformes em $\{0, \ldots, \theta\}$, para algum $\theta \in \mathbb{N}$.

Teorema 2.2.1 *Sejam $X_1, X_2, \ldots, X_N$ variáveis aleatórias independentes assumindo valores em $\mathbb{N}$. Então, $(X_1, \ldots, X_N)$ é uniformemente distribuído sobre χ_m^N, dado $X_{(N)} = m$, $m \in \mathbb{N}$ tal que $P(X_{(N)} = m) > 0$, se, e somente se, X_i é uniformemente distribuído em $\{0, \ldots, \theta_0\}$ para algum $\theta_0 \in \mathbb{N}$, $i = 1, \ldots, N$.*

Prova ($\Rightarrow$) Notemos, inicialmente, que $(X_1, \ldots, X_N)$ é permutável, pois sua distribuição pertence a $\mathcal{C}_N$. Assim, como $X_1, \ldots, X_N$ são variáveis aleatórias independentes e identicamente distribuídas, temos, $\forall n \leq N$, que

$$\Pr(X_1 = x_1, \ldots, X_n = x_n \mid X_{(n)} = m) = \frac{\prod_{i=1}^n \Pr(X_1 = x_i) \mathbb{I}_{\{m\}}(\max_{1 \leq i \leq n}\{x_i\})}{[\Pr(X_1 \leq m)]^n - [\Pr(X_1 \leq m-1)]^n} \tag{2.4}$$

Mas, por hipótese, $(X_1, \ldots, X_N)$, dado $X_{(N)} = m$, $m \in \mathbb{N}$, é uniformemente distribuído sobre χ_m^N e, então, $(X_1, \ldots, X_n)$, dado $X_{(n)} = m$, $m \in \mathbb{N}$, é uniformemente distribuído sobre χ_m^n. Segue, $\forall n \leq N$, que

$$\frac{\prod_{i=1}^n \Pr(X_1 = x_i) \mathbb{I}_{\{m\}}(\max_{1 \leq i \leq n}\{x_i\})}{[\Pr(X_1 \leq m)]^n - [\Pr(X_1 \leq m-1)]^n} = \frac{\mathbb{I}_{\{m\}}(\max_{1 \leq i \leq n}\{x_i\})}{(m+1)^n - m^n}.$$

Para $n = 2$ e tomando $x_1 = x_2 = m$, segue que

$$\frac{\Pr(X_1 = m)]^2}{[\Pr(X_1 \leq m)]^2 - [\Pr(X_1 \leq m-1)]^2} = \frac{1}{(m+1)^2 - m^2} \Rightarrow$$

$$\Rightarrow \frac{\Pr(X_1 = m)]^2}{\Pr(X_1 = m)[2\Pr(X_1 \leq m) - \Pr(X_1 = m)]} = \frac{1}{2m+1} \Rightarrow$$

$$\Rightarrow \frac{\Pr(X_1 = m)}{\Pr(X_1 \leq m)} = \frac{1}{m+1},$$

$\forall m \in \mathbb{N}$ tal que $P(X_1 = m) > 0$. Isto implica que $P(X_1 = 0) > 0$. Além disso, $\{x \in \mathbb{N} : P(X_1 = x) = 0\} \neq \emptyset$, pois, caso contrário, teríamos, desenvolvendo a última igualdade acima, que

$$P(X_1 = x) = P(X_1 = 0) > 0 \ , \ \forall x \in \mathbb{N} \ ,$$

o que seria absurdo, uma vez que $P(X_1 = \cdot)$ é uma distribuição de probabilidade. Seja, então, $\theta_0 + 1 = \min\{x \in \mathbb{N} : P(X_1 = x) = 0\}$. Afirmamos que $\{x \in \mathbb{N} : P(X_1 = x) = 0\} = \{\theta_0 + 1, \theta_0 + 2, \ldots\}$. Se não, existiria $b = \min\{x > \theta + 1 : P(X_1 = x) > 0\}$ tal que $P(X_1 = b) = \frac{\theta+1}{b} P(X_1 = 0)$, produzindo

$$P((0, \ldots, 0, b) | \ X_{(N)} = b) = \frac{(\frac{b}{\theta+1})^{N-1}}{(b+1)^N - b^N} \ ,$$

o que contradiria o fato de que a distribuição de $(X_1, \ldots, X_N)$ pertence a $\mathcal{C}_N$. Logo, $P(X_1 = x) = P(X_1 = 0) > 0, \forall x \in \{0, \ldots, \theta_0\}$, e $P(X_1 = x) = 0, \forall x \geq \theta_0 + 1$, e, portanto,

$$P(X_1 = x) = \frac{1}{\theta_0 + 1} \mathbb{I}_{\{0, \ldots, \theta_0\}}(x).$$

($\Leftarrow$) Por outro lado, se $X_1, \ldots, X_N$ são variáveis aleatórias independentes com distribuição comum uniforme em $\{0, 1, \ldots, \theta_0\}$, temos, de (2.4), que

$$P(X_1 = x_1, \ldots, X_N = x_n \mid X_{(N)} = m) = \frac{\frac{1}{(\theta_0+1)^N} \mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i\})}{\left(\frac{m+1}{\theta_0+1}\right)^N - \left(\frac{m}{\theta_0+1}\right)^N} \Rightarrow$$

$$\Rightarrow \quad \Pr(X_1 = x_1, \ldots, X_N = x_N \mid X_{(N)} = n) = \frac{\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i\})}{(m+1)^N - m^N},$$

isto é, dado $X_{(N)} = m$, $(X_1, \ldots, X_N)$ é uniformemente distribuído sobre χ_m^N. ■

Consideremos agora a distribuição uniforme no seguinte subconjunto de Z^N, para $N > 3$, $m_1, m_2 \in Z$, com $m_1 < m_2$,

$$\chi_{m_1,m_2}^N = \left\{(x_1, \ldots, x_N) \in Z^N : \min_{1\leq i\leq N}\{x_i\} = m_1, \ \max_{1\leq i\leq N}\{x_i\} = m_2\right\}.$$

$(X_1, \ldots, X_N)$ é dito uniformemente distribuído sobre χ_{m_1,m_2}^N, com lei de probabilidade denotada por $Q_N^{m_1,m_2}$, se

$$\text{Prob}(X_1 = x_1, \ldots, X_N = x_N) = Q_N^{m_1,m_2}((x_1, \ldots, x_N)) =$$

$$= \frac{\mathbb{I}_{\{m_1\}}(\min_{1\leq i\leq N}\{x_i\})\mathbb{I}_{\{m_2\}}(\max_{1\leq i\leq N}\{x_i\})}{(m_2 - m_1 + 1)^N - 2(m_2 - m_1)^N + (m_2 - m_1 - 1)^N}, \tag{2.5}$$

onde o denominador do lado direito da igualdade acima corresponde à medida de contagem de χ_{m_1,m_2}^N.

Na sequência, apresentamos uma caracterização das medidas de probabilidade que são misturas dos elementos da família $\{Q_N^{m_1,m_2} : (m_1, m_2) \in Z^2, m_1 < m_2\}$. Assim, consideremos a classe $\mathcal{C}_N$ de medidas de probabilidade em $(Z^N, \mathcal{P}(Z^N))$ obtidas por mistura dos elementos de $\{Q_N^{m_1,m_2} : (m_1, m_2) \in Z^2, m_1 < m_2\}$.

Proposição 2.2.2 *Uma medida de probabilidade P pertence a $\mathcal{C}_N$ se, e somente se, existe uma função não-negativa φ tal que*

$$P((x_1, \ldots, x_N)) = \varphi(\min_{1\leq i\leq N}\{x_i\}, \ \max_{1\leq i\leq N}\{x_i\})$$

para cada $(x_1, \ldots, x_N) \in Z^N$.

Prova ($\Rightarrow$) Se P pertence a $\mathcal{C}_N$, então, para $(x_1, \ldots, x_N) \in Z^N$, temos que

$$P((x_1, \ldots, x_N)) = \int_{Z^2} Q_N^{a,b}((x_1, \ldots, x_N))\mathrm{d}\mu(a,b) =$$

$$= \int_{Z^2} \frac{\mathbb{I}_{\{a\}}(\min_{1\leq i\leq N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\leq i\leq N}\{x_i\})}{(b-a+1)^N - 2(b-a)^N + (b-a-1)^N}\mathrm{d}\mu(a,b) \ .$$

Como μ é uma medida de probabilidade em $(Z^2, \mathcal{P}(Z^2))$, reescrevemos a última integral acima como uma série. Então,

$$P((x_1, \ldots, x_N)) = \sum_{(a,b)\in Z^2} \frac{\mathbb{I}_{\{a\}}(\min_{1\leq i\leq N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\leq i\leq N}\{x_i\})}{(b-a+1)^N - 2(b-a)^N + (b-a-1)^N}\mu(\{(a,b)\}) \ . \qquad (2.6)$$

Considerando a função não-negativa $\varphi : Z^2 \to \mathbb{R}_+$ definida por

$$\varphi(u,v) = \sum_{(a,b)\in Z^2} \frac{\mathbb{I}_{\{a\}}(u)\mathbb{I}_{\{b\}}(v)}{(b-a+1)^N - 2(b-a)^N + (b-a-1)^N}\mu(\{(a,b)\}), \ (u,v) \in Z^2,$$

obtemos de (2.6) que

$$P((x_1, \ldots, x_N)) = \varphi\left(\min_{1\leq i\leq N}\{x_i\}, \ \max_{1\leq i\leq N}\{x_i\}\right).$$

($\Leftarrow$) Suponha agora que $\forall (x_1, \ldots, x_N) \in Z^N$,

$$P((x_1, \ldots, x_N)) = \varphi\left(\min_{1\leq i\leq N}\{x_i\}, \ \max_{1\leq i\leq N}\{x_i\}\right).$$

Então, para $(a,b) \in Z^2$, com $a < b$, temos que

$$P(X_{(1)} = a, \ X_{(N)} = b) = \sum_{\substack{(x_1,\ldots,x_N)\in Z^N: \\ \min_{1\leq i\leq N}\{x_i\}=a, \max_{1\leq i\leq N}\{x_i\}=b}} P((x_1, \ldots, x_N)) =$$

$$= \sum_{\substack{(x_1,\ldots,x_N)\in Z^N:\\ \min_{1\le i\le N}\{x_i\}=a,\max_{1\le i\le N}\{x_i\}=b}} \varphi\left(\min_{1\le i\le N}\{x_i\}, \max_{1\le i\le N}\{x_i\}\right) =$$

$$= \sum_{\substack{(x_1,\ldots,x_N)\in Z^N:\\ \min_{1\le i\le N}\{x_i\}=a,\max_{1\le i\le N}\{x_i\}=b}} \varphi(a,b) \Rightarrow$$

$$\Rightarrow P(X_{(1)} = a,\ X_{(N)} = b) = \varphi(a,b)|\chi^N_{a,b}| \,. \tag{2.7}$$

Verifiquemos que $P(A) = \int_{Z^2} Q_N^{a,b}(A)\mathrm{d}\mu(a,b)$, $\forall A \in \mathcal{P}(Z^N)$, para específica μ em $(Z^2, \mathcal{P}(Z^2))$.

$$P(A) = \sum_{(x_1,\ldots,x_N)\in A} P((x_1,\ldots,x_N)) =$$

$$= \sum_{(x_1,\ldots,x_N)\in A} \sum_{(a,b)\in Z^2} P((x_1,\ldots,x_N), X_{(1)} = a, X_{(N)} = b) =$$

$$= \sum_{(x_1,\ldots,x_N))\in A} \sum_{(a,b)\in Z^2} P((x_1,\ldots,x_N))\mathbb{I}_{\{a\}}(\min_{1\le i\le N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\le i\le N}\{x_i\}) =$$

$$= \sum_{(x_1,\ldots,x_N))\in A} \sum_{(a,b)\in Z^2} \varphi(\min_{1\le i\le N}\{x_i\}, \max_{1\le i\le N}\{x_i\})\mathbb{I}_{\{a\}}(\min_{1\le i\le N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\le i\le N}\{x_i\}) =$$

$$= \sum_{(x_1,\ldots,x_N))\in A} \sum_{(a,b)\in Z^2} \varphi(a,b)\mathbb{I}_{\{a\}}(\min_{1\le i\le N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\le i\le N}\{x_i\}) =$$

$$= \sum_{(a,b)\in Z^2} \sum_{(x_1,\ldots,x_N)\in A} \varphi(a,b)\mathbb{I}_{\{a\}}(\min_{1\le i\le N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\le i\le N}\{x_i\}),$$

onde a úlitma igualdade segue do Teorema de Tonelli. Substituindo (2.7) na expressão para $P(A)$, resulta que

$$P(A) = \sum_{(a,b)\in Z^2} \sum_{(x_1,\ldots,x_N)\in A} \frac{\mathbb{I}_{\{a\}}(\min_{1\le i\le N}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\le i\le N}\{x_i\})}{|\chi^N_{a,b}|} P(X_{(1)}=a, X_{(N)}=b)$$

$$= \sum_{(a,b)\in Z^2} \sum_{(x_1,\ldots,x_N)\in A} Q_N^{a,b}((x_1,\ldots,x_N))P(X_{(1)}=a, X_{(N)}=b) =$$

$$= \sum_{(a,b)\in Z^2} Q_N^{a,b}(A)P(X_{(1)}=a, X_{(N)}=b) \Rightarrow P(A) = \int_{Z^2} Q_N^{a,b}(A)\mathrm{d}\mu(a,b),$$

onde μ denota a P-lei de $(X_{(1)}, X_{(N)})$. Portanto,

$$P = \int_{Z^2} Q_N^{a,b}\mathrm{d}\mu(a,b)$$

e, consequentemente, $P \in \mathcal{C}_N$, concluindo a prova. ■

A proposição 2.2.2 estabelece que uma medida de probabilidade em $(Z^N, \mathcal{P}(Z^N))$ é mistura dos elementos de $\{Q_N^{m_1,m_2} : (m_1,m_2) \in Z^2, m_1 < m_2\}$ se, e somente se, corresponde à distribuição de um vetor aleatório $(X_1,\ldots,X_N)$ a valores em Z^N com distribuição dependendo de $(x_1,\ldots,x_N) \in Z^N$ somente através de $(\min_{1\le i\le N}\{x_i\}, \max_{1\le i\le N}\{x_i\})$.

Teorema 2.2.2 *Sejam $X_1, X_2, X_3, \ldots, X_N$ variáveis aleatórias independentes assumindo valores em Z. Então, $X_1,\ldots,X_N$ é uniformemente distribuído sobre $\chi^N_{a,b}$, dado $X_{(1)} = a$ e $X_{(N)} = b$, $a, b \in Z$, $a < b$, tais que $P(X_{(1)} = a, X_{(N)} = b) > 0$, se, e somente se, X_i tem distribuição uniforme em $\{\theta_1, \theta_1+1, \ldots, \theta_2\}$ para algum par $\theta_1, \theta_2 \in Z$ com $\theta_1 < \theta_2$, $i = 1,\ldots,N$.*

Prova ($\Rightarrow$) Destacamos, inicialmente, que $(X_1,\ldots,X_N)$ é permutável, pois sua distribuição pertence a $\mathcal{C}_N$. Assim, por definição e levando em conta que $X_1,\ldots,X_N$ são variáveis aleatórias independentes e identicamente distribuídas, temos, para $a, b \in Z$, $a < b$, e $\forall n \le N$, que

$$P\left(X = x_1, \ldots, X_n = x_n \mid \min_{1\leq i\leq n}\{X_i\} = a,\ \max_{1\leq i\leq n}\{X_i\} = b\right) = \tag{2.8}$$

$$= \frac{\prod_{i=1}^{n} P(X_1 = x_i)\mathbb{I}_{\{a\}}(\min_{1\leq i\leq n}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\leq i\leq n}\{x_i\})}{[P(a\leq X_1\leq b)]^n - [P(a+1\leq X_1\leq b)]^n - [P(a\leq X_1\leq b-1)]^n + [P(a+1\leq X_1\leq b-1)]^n}$$

Mas, condicionalmente a $\max_{1\leq i\leq N}\{X_i\} = b$ e $\min_{1\leq i\leq N}\{X_i\} = a$, $(X_1, \ldots, X_N)$ é uniformemente distribuído sobre $\chi_{a,b}^N$. Então, $(X_1, \ldots, X_n)$, dado $X_{(1)} = a$ e $X_{(n)} = b$, $a, b \in Z$, com $a < b$, é uniformemente distribuído sobre $\chi_{a,b}^n$, de modo que, para $2 \leq n \leq N$, temos

$$\frac{\prod_{i=1}^{n} P(X_i = x_i)\mathbb{I}_{\{a\}}(\min_{1\leq i\leq n}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\leq i\leq n}\{x_i\})}{[P(a\leq X_1\leq b)]^n - [P(a+1\leq X_1\leq b)]^n - [P(a\leq X_1\leq b-1)]^n + [P(a+1\leq X_1\leq b-1)]^n} =$$

$$= \frac{\mathbb{I}_{\{a\}}(\min_{1\leq i\leq n}\{x_i\})\mathbb{I}_{\{b\}}(\max_{1\leq i\leq n}\{x_i\})}{(b-a+1)^n - 2(b-a)^n + (b-a-1)^n} \tag{2.9}$$

Para $n = 3$, $x_1 = a$, $x_2 = b$ e $x_3 = x \in \{a, \ldots, b\}$, obtemos, de (2.9), que

$$\frac{P(X_1 = a)P(X_1 = b)P(X_1 = x)}{[P(a\leq X_1\leq b)]^3 - [P(a+1\leq X_1\leq b)]^3 - [P(a\leq X_1\leq b-1)]^3 + [P(a+1\leq X_1\leq b-1)]^3} =$$

$$= \frac{1}{(b-a+1)^3 - 2(b-a)^3 + (b-a-1)^3}$$

Desenvolvendo os denominadores na igualdade acima, resulta, $\forall a, b \in Z$ tais que $P(X_1 = a) > 0$ e $P(X_1 = b) > 0$, que

$$\frac{P(X_1 = x)}{6P(a \leq X_1 \leq b) - 3P(X_1 = a) - 3P(X_1 = b)} = \frac{1}{6(b-a)} \Rightarrow$$

$$\Rightarrow \; P(X_1 = x) \; = \; \frac{6P(a \leq X_1 \leq b) - 3P(X_1 = a) - 3P(X_1 = b)}{6(b-a)} \; \Rightarrow$$

$$\Rightarrow \; P(X_1 = x) = \; \text{constante} \; > 0 \; ,$$

$\forall x \in \{a, \ldots, b\}$. Como $P(X_1 = \cdot)$ é uma distribuição de probabilidade, $\exists \theta_1, \theta_2 \in Z$, $\theta_1 < \theta_2$, tais que $P(X_1 = x) = 0$ para $x < \theta_1$ ou $x > \theta_2$ (caso contrário, a série $\sum_{x \in Z} P(X_1 = x)$ não seria convergente). Logo, $P(X_1 = \theta_1) = P(X_1 = \theta_1 + 1) = \cdots =$ $= P(X_1 = \theta_2) > 0$ e $P(X_1 = x) = 0$, para $x < \theta_1$ ou $x > \theta_2$, e, portanto,

$$P(X_1 = x) = \frac{\mathbb{I}_{\{\theta_1, \theta_1+1, \ldots, \theta_2\}}(x)}{\theta_2 - \theta_1 + 1} \; ,$$

isto é, X_1 é uniformemente distribuído em $\{\theta_1, \theta_1 + 1, \ldots, \theta_2\}$ para $\theta_1, \theta_2 \in Z$, com $\theta_1 < \theta_2$.

($\Leftarrow$) Se $X_1, \ldots, X_N$ são variáveis aleatórias independentes e identicamente distribuídas uniformes em $\{\theta_1, \theta_1 + 1, \ldots, \theta_2\}$, com $\theta_1 < \theta_2$, então, para $a, b \in Z$, $a < b$,

$$P(X_1 = x_1, \ldots, X_N = x_n \mid \min_{1 \leq i \leq N}\{X_i\} = a, \; \max_{1 \leq i \leq N}\{X_i\} = b) \; =$$

$$= \frac{\prod_{i=1}^{N} P(X_i = x_i) \mathbb{I}_{\{a\}}(\min_{1 \leq i \leq N}\{x_i\}) \mathbb{I}_{\{b\}}(\max_{1 \leq i \leq N}\{x_i\})}{P(\min_{1 \leq i \leq N}\{X_i\} = a, \; \max_{1 \leq i \leq N}\{X_i\} = b)} \; =$$

$$= \frac{\prod_{i=1}^{N} \frac{1}{k_2 - k_1 + 1} \mathbb{I}_{\{a\}}(\min_{1 \leq i \leq N}\{x_i\}) \mathbb{I}_{\{b\}}(\max_{1 \leq i \leq N}\{x_i\})}{[P(a \leq X_1 \leq b)]^N - [P(a+1 \leq X_1 \leq b)]^N - [P(a \leq X_1 \leq b-1)]^N + [P(a+1 \leq X_1 \leq b-1)]^N}$$

$$= \frac{\frac{1}{(k_2 - k_1 + 1)^N} \mathbb{I}_{\{a\}}(\min_{1 \leq i \leq N}\{x_i\}) \mathbb{I}_{\{b\}}(\max_{1 \leq i \leq N}\{x_i\})}{(\frac{b-1+1}{k_2 - k_1 + 1})^N - (\frac{b-(a+1)+1}{k_2 - k_1 + 1})^N - (\frac{(b-1)-a+1}{k_2 - k_2 + 1})^N + (\frac{(b-1)-(a+1)+1}{k_2 - k_1 + 1})^N} \; \Rightarrow$$

$$\Rightarrow P\left(X_1 = x_1, \ldots, X_N = x_N \mid \min_{1\leq i\leq N}\{X_i\} = a,\ \max_{1\leq i\leq N}\{X_i\} = b\right) =$$

$$= \frac{\mathbb{I}_{\{b\}}(\max_{1\leq i\leq N}\{x_i\})\mathbb{I}_{\{a\}}(\min_{1\leq i\leq N}\{x_i\})}{(b-a+1)^N - 2(b-a)^N + (b-a-1)^N},$$

isto é, dados $\min_{1\leq i\leq N}\{X_i\} = a$ e $\max_{1\leq i\leq N}\{X_i\} = b$, $a, b \in Z$, $a < b$, $(X_1, \ldots, X_N)$ é uniformemente distribuído sobre $\chi^N_{a,b}$. ■

2.2.2 Caso bivariado

Como na subseção 2.2.1, primeiramente vamos descrever a distribuição uniforme sobre um específico subconjunto de $(\mathbb{N}^2)^N$, $N \geq 2$ (esta uniformidade figurará nas hipóteses do Teorema 2.2.3), e, então, enunciar os resultados de caracterização.

Consideremos o seguinte subconjunto de $(\mathbb{N}^2)^N$, para $N \geq 2$ e $m \in \mathbb{N}$

$$\chi^N_m = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{N}^2)^N : \max_{1\leq i\leq N}\{x_i + y_i\} = m\}.$$

$((X_1, Y_1), \ldots, (X_N, Y_N))$ é dito uniformemente distribuído sobre χ^N_m (conforme (1.3), capítulo 1), com lei de probabilidade denotada por Q_{Nm}, se

$$Pr((X_1, Y_1) = (x_1, y_1), \ldots, (X_N, Y_N) = (x_N, y_N)) = Q_{Nm}(((x_1 y_1), \ldots, (x_N, y_N))) =$$

$$= \frac{\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i + y_i\})}{\binom{m+2}{2}^N - \binom{m+1}{2}^N}, \tag{2.10}$$

onde o denominador no lado direito da igualdade acima é a cardinalidade de χ^N_m.

Verificamos, no capítulo 1, que se $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído em χ^N_m, então a marginal n-dimensional $((X_1, Y_1), \ldots, (X_n, Y_n))$, $n < N$, tem lei Q^n_{Nm} dada por

$$Q^n_{Nm}((x_1,y_1),\ldots,(x_n,y_n)) = \begin{cases} \dfrac{\binom{m+2}{2}^{N-n} - \binom{m+1}{2}^{N-n}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}}, & \text{se } \max_{1\le i\le n}\{x_i+y_i\} < m, \\ \dfrac{\binom{m+2}{2}^{N-m}}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}}, & \text{se } \max_{1\le i\le n}\{x_i+y_i\} = m. \end{cases}$$

Este último resultado, juntamente com propriedades da distância em variação total, fornecem as desigualdades básicas (conforme (1.5) no capítulo 1) para a obtenção das formas finita e infinita do Teorema do tipo de De Finetti para o modelo uniforme discreto no triângulo isósceles $\{(x,y) \in \mathbb{N}^2 : x+y \le \theta\}$, $\theta \in \mathbb{N}$ (a segunda sendo derivada a partir da primeira fixando-se um valor de n e tomando-se limites quando $N \to \infty$).

Como na subseção 2.2.1, apresentamos, no que segue, uma caracterização das medidas de probabilidade que correspondem a misturas dos elementos da família $\{Q_{Nm} : m \in \mathbb{N}\}$. Consideremos, então, a classe $\mathcal{C}_N$ de medidas de probabilidade em $((\mathbb{N}^2)^N, \mathcal{P}((\mathbb{N}^2)^N))$ geradas por misturas dos elementos de $\{Q_{Nm} : m \in \mathbb{N}\}$ na variável "radial".

Proposição 2.2.3 *Uma medida de probabilidade P pertence a $\mathcal{C}_N$ se, e somente se, existe uma função não-negativa φ tal que*

$$P(((x_1,y_1),\ldots,(x_N,y_N))) = \varphi\left(\max_{1\le i\le N}\{x_i+y_i\}\right)$$

para cada $((x_1,\, y_1),\, \ldots,\, (x_N,y_N)) \in (\mathbb{N}^2)^N$.

Prova ($\Rightarrow$) Se P pertence a $\mathcal{C}_N$, então, para $((x_1,y_1),\ldots,(x_N,y_N)) \in (\mathbb{N}^2)^N$, temos

$$P(((x_1,y_1),\ldots,(x_N,y_N))) = \int_{\mathbb{N}} Q_{Nm}(((x_1,y_1),\ldots,(x_N,y_N)))\mathrm{d}\mu(m) =$$

$$= \int_{\mathbb{N}} \frac{\mathbb{I}_{\{m\}}(\max_{1\le i\le N}\{x_i+y_i\})}{\binom{m+2}{2}^{N} - \binom{m+1}{2}^{N}}\mathrm{d}\mu(m)\ .$$

Como μ é uma medida de probabilidade em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, podemos reescrever a última integral acima como uma série, isto é,

$$P(((x_1, y_1), \ldots, (x_N, y_N))) = \sum_{m=1}^{\infty} \frac{\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i + y_i\})}{\binom{m+2}{2}^N - \binom{m+1}{2}^N} \mu(\{m\}) \ . \tag{2.11}$$

Considerando a função não-negativa $\varphi : \mathbb{N} \to \mathbb{R}_+$ definida por

$$\varphi(t) = \sum_{m=1}^{\infty} \frac{\mathbb{I}_{\{m\}}(t)}{\binom{m+2}{2}^N - \binom{m+1}{2}^N} \mu(\{m\}), \ t \in \mathbb{N},$$

temos de (2.11) que

$$P(((x_1, y_1), \ldots, (x_N, y_N))) = \varphi\left(\max_{1\leq i\leq N}\{x_i\}\right).$$

($\Leftarrow$) Suponhamos agora que $\forall((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{N}^2)^N$

$$P(((x_1, y_1), \ldots, (x_N, y_N))) = \varphi\left(\max_{1\leq i\leq N}\{x_i + y_i\}\right).$$

Então, para $m \in \mathbb{N}$, temos que

$$P\left(\max_{1\leq i\leq N}\{X_i + Y_i\} = m\right) = \sum_{\substack{((x_1,y_1),\ldots,(x_N,y_N))\in(\mathbb{N}^2)^N: \\ \max_{1\leq i\leq N}\{x_i+y_i\}=m}} P(((x_1, y_1), \ldots, (x_N, y_N))) =$$

$$= \sum_{\substack{((x_1,y_1),\ldots,(x_N,y_N))\in(\mathbb{N}^2)^N: \\ \max_{1\leq i\leq N}\{x_i+y_i\}=m}} \varphi(\max_{1\leq i\leq N}\{x_i + y_i\}) = \sum_{\substack{((x_1,y_1),\ldots,(x_N,y_N))\in(\mathbb{N}^2)^N: \\ \max_{1\leq i\leq N}\{x_i+y_i\}=m}} \varphi(m) \ \Rightarrow$$

$$\Rightarrow \ P\left(\max_{1\leq i\leq N}\{X_i + Y_i\} = m\right) = \varphi(m)|\chi_m^N| \ . \tag{2.12}$$

Demonstramos agora que $P(A) = \int_{\mathbb{N}} Q_{Nm}(A)\mathrm{d}\mu(m)$, $\forall A \in \mathcal{P}((\mathbb{N}^2)^N)$, para uma específica μ em $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

$$P(A) = \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} P(((x_1,y_1),\ldots,(x_N,y_N))) =$$

$$= \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} \sum_{m=1}^{\infty} P\left(((x_1,y_1),\ldots,(x_N,y_N)), \max_{1\leq i\leq N}\{X_i+Y_i\} = m\right) =$$

$$= \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} \sum_{m=1}^{\infty} P(((x_1,y_1),\ldots,(x_N,y_N)))\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i+y_i\}) =$$

$$= \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} \sum_{m=1}^{\infty} \varphi(\max_{1\leq i\leq N}\{x_i+y_i\})\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i+y_i\}) =$$

$$= \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} \sum_{m=1}^{\infty} \varphi(m)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i+y_i\}) \Rightarrow$$

$$\Rightarrow \quad P(A) = \sum_{m=1}^{\infty} \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} \varphi(m)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i+y_i\}),$$

onde a última igualdade segue do Teorema de Tonelli. Substituindo (2.12) na expressão para $P(A)$, obtemos que

$$P(A) = \sum_{m=1}^{\infty} \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} \frac{\mathbb{I}_{\{m\}}(\max_{1\leq i\leq N}\{x_i+y_i\})}{|\chi_m^N|} P(\max_{1\leq i\leq N}\{X_i+Y_i\} = m)$$

$$= \sum_{m=1}^{\infty} \sum_{((x_1,y_1),\ldots,(x_N,y_N))\in A} Q_{Nm}(((x_1,y_1),\ldots,(x_N,y_N)))P(\max_{1\leq i\leq N}\{X_i+Y_i\} = m) =$$

$$= \sum_{m=1}^{\infty} Q_{Nm}(A) P(\max_{1\leq i\leq N}\{X_i + Y_i\} = m) \quad \Rightarrow \quad P(A) = \int_{\mathbb{N}} Q_{Nm}(A) \mathrm{d}\mu(m),$$

onde μ é a P-lei de $\max_{1\leq i\leq N}\{X_i + Y_i\}$. Portanto,

$$P = \int_{\mathbb{N}} Q_{Nm} \mathrm{d}\mu(m)$$

e, consequentemente, $P \in \mathcal{C}_N$, concluindo a demonstração. ■

A proposição 2.2.3 estabelece que uma medida de probabilidade em $((\mathbb{N}^2)^N, \mathcal{P}((\mathbb{N}^2)^N))$ é mistura dos elementos de $\{Q_{Nm} : m \in \mathbb{N}\}$ se, e somente se, corresponde à distribuição de um vetor aleatório $((X_1, Y_1), \ldots, (X_N, Y_N))$ a valores em $(\mathbb{N}^2)^N$ com distribuição dependendo de $((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{N}^2)^N$ somente através de $\max_{1\leq i\leq N}\{x_i + y_i\}$.

Para concluirmos esta seção, demonstramos que a única medida produto em $\mathcal{C}_N$ corresponde à distribuição de n vetores aleatórios independentes e com mesma distribuição uniforme em $\{(x, y) \in \mathbb{N}^2 : x + y \leq \theta\}$, para algum $\theta \in \mathbb{N}$.

Teorema 2.2.3 *Sejam $(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)$ vetores aleatórios independentes assumindo valores em $\mathbb{N}^2$. Então, $((X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N))$ é uniformemente distribuído em χ_m^N, dado $\max_{1\leq i\leq N}\{X_i+Y_i\} = m$, $m \in \mathbb{N}$ tal que $P(\max_{1\leq i\leq N}\{X_i+Y_i\} = m) > 0$, se, e somente se, (X_i, Y_i) é distribuído uniformemente sobre $\{(a, b) \in \mathbb{N}^2 : a+b \leq \theta\}$ para algum $\theta \in \mathbb{N}$, i = 1,...,N.*

Prova ($\Rightarrow$) Neste caso, $(X_1, Y_1), \ldots, (X_N, Y_N)$ são permutáveis. Assim, pela definição e considerando a independência dos vetores aleatórios bivariados, temos, $\forall n \leq N$, que

$$P\left((X_1 = x_1, Y_1 = y_1), \ldots, (X_n = x_n, Y_n = y_n) \mid \max_{1\leq i\leq n}\{X_i + Y_i\} = m\right) =$$

$$= \frac{\prod_{i=1}^{n} P(X_1 = x_i, Y_1 = y_i)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq n}\{x_i + y_i\})}{[\Pr(X_1 + Y_1 \leq m)]^n - [\Pr(X_1 + Y_1 \leq m-1)]^n}. \tag{2.13}$$

Como $((X_1, Y_1), \ldots, (X_N, Y_N))$, dado $\max_{1\leq i\leq N}\{X_i + Y_i\} = m$, $m \in \mathbb{N}$, é uniformemente distribuído sobre χ_m^N, então $((X_1, Y_1), \ldots, (X_n, Y_n))$, dado $\max_{1\leq i\leq n}\{X_i + Y_i\} = m$, é uniformemente distribuído em χ_m^n, $n \leq N$. Assim, $\forall n \leq N$

$$\frac{\prod_{i=1}^{n} \Pr(X_1 = x_i, Y_1 = y_i)\mathbb{I}_{\{m\}}(\max_{1\leq i\leq n}\{x_i + y_i\})}{[\Pr(X_1 + Y_1 \leq m)]^n - [\Pr(X_1 + Y_1 \leq m-1)]^n} = \frac{\mathbb{I}_{\{m\}}(\max_{1\leq i\leq n}\{x_i + y_i\})}{\binom{m+2}{2}^n - \binom{m+1}{2}^n}, \tag{2.14}$$

Para $m \in \mathbb{N}$ tal que $P(X_1 + Y_1 = m) > 0$, consideremos em (2.14), $n = 1$ e $(x_1, y_1) \in \mathbb{N}^2$ tal que $x_1 + y_1 = m$. Assim,

$$\frac{P(X_1 = x_1, Y_1 = y_1)}{P(X_1 + Y_1 = m)} = \frac{1}{m+1}. \tag{2.15}$$

De (2.15), podemos ver que é suficiente obtermos a distribuição de $X_1 + Y_1$ para determinarmos a lei de (X_1, Y_1). Em (2.14), tomando $n = 2$ e $(x_1, y_1) \in \mathbb{N}^2$ tal que $x_1 + y_1 = m$, obtemos

$$\frac{[P(X_1 = x_1, Y_1 = y_1)]^2}{P(X_1 + Y_1 = m)[P(X_1 + Y_1 \leq m) + P(X_1 + Y_1 \leq m-1)]} = \frac{1}{\binom{m+2}{2}^2 - \binom{m+1}{2}^2} \Rightarrow$$

$$\Rightarrow \frac{[P(X_1 = x_1, Y_1 = y_1)]^2}{P(X_1 + Y_1 = m)[P(X_1 + Y_1 \leq m) + P(X_1 + Y_1 \leq m-1)]} = \frac{1}{(m+1)^3}.$$

Lembrando de (2.15) que $P(X_1 = x_1, Y_1 = y_1) = \frac{P(X_1+Y_1=m)}{m+1}$, obtemos, na igualdade acima, que

$$\frac{P(X_1 + Y_1 \leq m) - P(X_1 + Y_1 \leq m-1)}{P(X_1 + Y_1 \leq m) + P(X_1 + Y_1 \leq m-1)} = \frac{1}{m+1},$$

$\forall m \in \mathbb{N}$ tal que $P(X_1 + Y_1 = m) > 0$. Logo, $P(X_1 + Y_1 = 0) > 0$. Além disso, $\{x \in \mathbb{N} : P(X_1 + Y_1 = x) = 0\} \neq \emptyset$, pois, caso contrário, teríamos

$$F(m) = \frac{m+2}{m} F(m-1) \ , \tag{2.16}$$

onde $F(m) = P(X_1 + Y_1 \leq m)$. Através da relação de recorrência para $F(m)$ em (2.16), obteríamos

$$F(m) = \binom{m+2}{2} F(0), \ \forall m \geq 1$$

Consequentemente,

$$P(X_1 + Y_1 = m) \ = \ P(X_1 + Y_1 \leq m) - P(X_1 + Y_1 \leq m-1) \ =$$

$$= \ F(m) - F(m-1) \ = \ \binom{m+2}{2} F(0) - \binom{m+1}{2} F(0) \ = \ \binom{m+1}{1} F(0) \ \Rightarrow$$

$$\Rightarrow \ P(X_1 + Y_1 = m) \ = \ \binom{m+1}{1} P(X_1 + Y_1 = 0) \ , \ \forall m \geq 1 \ , \tag{2.17}$$

e, portanto, $P(X_1 + Y_1 = \cdot)$ seria crescente em $\mathbb{N}$, absurdo pois $P(X_1 + Y_1 = \cdot)$ é uma distribuição de probabilidade. Seja, então, $\theta + 1 = \min\{x \in \mathbb{N} : P(X_1 + Y_1 = x) = 0\}$. Afirmamos que $\{x \in \mathbb{N} : P(X_1 + Y_1 = x) = 0\} = \{\theta + 1, \theta + 2, \ldots\}$. Se não, existiria $b = \min\{x > \theta + 1 : P(X_1 + Y_1 = x) > 0\}$ tal que

$$P(X_1 + Y_1 = b) \ = \ P(X_1 + Y_1 \leq b) - P(X_1 + Y_1 \leq \theta) \ =$$

$$= \ \frac{b+2}{b} \binom{\theta+2}{2} P(X_1 + Y_1 = 0) - \binom{\theta+2}{2} P(X_1 + Y_1 = 0) \ \Rightarrow$$

$$\Rightarrow \; P(X_1 + Y_1 = b) = \frac{2}{b}\binom{\theta+2}{2} P(X_1 + Y_1 = 0) \; .$$

Deste modo, teríamos

$$P(((0,0),\ldots,(0,0),(0,b))|\max_{1\leq i\leq N}\{X_i + Y_i\} = b) \; =$$

$$= \frac{[P(X_1 + Y_1 = 0)]^{N-1} P(X_1 = 0, Y_1 = b)}{[\frac{b+2}{b}\binom{\theta+2}{2} P(X_1 + Y_1 = 0)]^N - [\binom{\theta+2}{2} P(X_1 + Y_1 = 0)]^N} \; .$$

Desenvolvendo a expressão acima, resultaria que

$$P(((0,0),\ldots,(0,0),(0,b))|\max_{1\leq i\leq N}\{X_i + Y_i\} = b) \; =$$

$$= \frac{[\frac{\binom{b+1}{2}}{\binom{\theta+1}{2}}]^{N-1}}{\binom{b+2}{2}^N - \binom{b+1}{2}^N} \; ,$$

isto é, $((X_1, Y_1), \ldots, (X_N, Y_N))$, condicionalmente a $\max_{1\leq i\leq N}\{X_i + Y_i\} = b$, não seria uniforme em χ_b^N. Logo,

$$P(X_1 + Y_1 = m) > 0 \; , \; \forall m \in \{0, \ldots, \theta\} \;\; \text{e} \;\; P(X_1 + Y_1 = m) = 0 \; , \; \forall m > \theta \; .$$

Assim,

$$\sum_{m=0}^{\theta} P(X_1 + Y_1 = m) = P(X_1 + Y_1 = 0) \sum_{m=0}^{\theta} \binom{m+1}{1} \Rightarrow P(X_1 + Y_1 = 0) = \frac{1}{\binom{\theta+2}{2}}.$$

Substituindo este valor em (2.16), obtemos que

$$P(X_1 + Y_1 = m) = \frac{m+1}{\binom{\theta+2}{2}}, \; m \in \{0, \ldots, \theta\}.$$

Finalmente, substituindo a expressão acima em (2.15), temos que

$$P(X_1 = x, Y_1 = y) = \frac{1}{\binom{\theta+2}{2}},$$

$\forall (x, y) \in \mathbb{N}^2$ tal que $x + y \in \{0, \ldots, \theta\}$, isto é, (X_1, Y_1) é uniformemente distribuído em $\{(a, b) \in \mathbb{N}^2 : a + b \leq \theta\}$, $\theta \in \mathbb{N}$.

($\Leftarrow$) Supondo que $(X_1, Y_1), \ldots, (X_N, Y_N)$ são independentes tendo distribuição comum uniforme sobre $\{(a, b) \in \mathbb{N}^2 : a + b \leq \theta\}$, $\theta \in \mathbb{N}$, temos, de (2.13), que

$$P\left((X_1 = x_1, Y_1 = y_1), \ldots, (X_N = x_N, Y_N = y_N) \mid \max_{1 \leq i \leq N} \{X_i + Y_i\} = m\right) =$$

$$= \frac{\prod_{i=1}^{N} \Pr(X_1 = x_i, Y_1 = y_i) \mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i + y_i\})}{[\Pr(X_1 + Y_1 \leq m)]^N - [\Pr(X_1 + Y_1 \leq m - 1)]^N} =$$

$$= \frac{\left[\frac{1}{\binom{\theta+2}{2}}\right]^N \mathbb{I}_{\{0,\ldots,k\}}(\max_{1 \leq i \leq N}\{x_i + y_i\}) \mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i + y_i\})}{\left[\frac{\binom{m+2}{2}}{\binom{\theta+2}{2}}\right]^N - \left[\frac{\binom{m+1}{2}}{\binom{\theta+2}{2}}\right]^N} \Rightarrow$$

$$\Rightarrow P\left((X_1 = x_1, Y_1 = y_1), \ldots, (X_N = x_N, Y_N = y_N) \mid \max_{1 \leq i \leq N} \{X_i + Y_i\} = m\right) =$$

$$= \frac{\mathbb{I}_{\{m\}}(\max_{1 \leq i \leq N}\{x_i + y_i\})}{\binom{m+2}{2}^N - \binom{m+1}{2}^N},$$

isto é, dado $\max_{1 \leq i \leq N}\{X_i + Y_i\} = m$, $((X_1, Y_1), \ldots, (X_N, Y_N))$ é uniformemente distribuído sobre χ_m^N. ■

2.3 Distribuições uniformes contínuas

Nesta seção, derivamos teoremas de caracterização para as versões contínuas dos modelos considerados na seção anterior.

2.3.1 Caso univariado

Primeiro, consideramos variáveis aleatórias $X_1, \ldots, X_N$, $N \geq 2$, absolutamente contínuas com função densidade de probabilidade conjunta f dada por

$$f(x_1, \ldots, x_N) = \psi_N \left(\max_{1 \leq i \leq N} \{|X_i|\} \right), \tag{2.18}$$

para alguma função não-negativa ψ que torne f definida acima uma função densidade de probabilidade em $\mathbb{R}^N$, isto é, $\psi : \mathbb{R}^N \to \mathbb{R}_+$ tal que

$$\int_0^\infty u^{N-1} \psi_N(u) \mathrm{d}u = \frac{1}{N2^N} \tag{2.19}$$

Fang et al. (1990) denominam ψ_N de função geradora de densidade da distribuição esférica l_∞ e a variável $R = \max_{1 \leq i \leq N}\{|X_i|\}$ de variável "radial". Neste caso, se g é a função densidade de probabilidade de R, então

$$g(r) = N2^N r^{N-1} \psi_N(r) \mathbb{I}_{(0,\infty)}(r) \ . \tag{2.20}$$

(a demonstração deste fato é apresentada no teorema que segue). O próximo resultado fornece uma caracterização de variáveis aleatórias independentes com função densidade de probabilidade conjunta dada por (2.18).

Teorema 2.3.1 *Sejam $X_1, \ldots, X_N$, $N \geq 2$, variáveis aleatórias absolutamente contínuas independentes a valores em $\mathbb{R}$ com função densidade de probabilidade conjunta f. Se f satisfaz (2.18), então X_i é uniformemente distribuído em $(-\theta, \theta)$, para algum $\theta > 0$, $i = 1, \ldots, N$.*

Prova Para caracterizarmos X_i, $i = 1, \ldots, N$, derivaremos duas expressões para a função densidade de probabilidade de $T_N = \max_{1 \leq i \leq N}\{|X_i|\}$ e igualaremos tais expressões. O resultado é obtido então desta última igualdade.

Inicialmente, vamos obter $\mathbb{P}_N$, a distribuição de probabilidade de $(X_1, \ldots, X_N)$, em termos de ψ e da distribuição uniforme sobre $H_m^N = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : \max_{1 \leq i \leq N}\{|x_i|\} = m\}$, $m > 0$, Q_{Nm}, dada por (Iglesias et al. (1998))

$$Q_{Nm}(A) = \frac{\sum_{i=1}^{N} [\lambda_{N-1}(\varphi^i(A \cap M_i(m))) + \lambda_{N-1}(\varphi^i(A \cap M_i(-m)))]}{2N(2m)^{N-1}}, \quad A \in \mathcal{B}(\mathbb{R}^N), \tag{2.21}$$

onde $M_i(m) = \{(x_1, \ldots, x_N) \in H_m^N : x_i = m\}$ e $\varphi^i : \mathbb{R}^N \to \mathbb{R}^{N-1}$ é dada por

$$\varphi^i(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_N) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N).$$

Assim, para $A \in \mathcal{B}(\mathbb{R}^N)$, temos

$$\mathbb{P}_N(A) = \int_A f(x_1, \ldots, x_N) \mathrm{d}\lambda_N \ .$$

Definindo

$$B_i = \left\{(x_1, \ldots, x_N) \in \mathbb{R}^N : \max_{1 \leq j \leq N}\{|x_j|\} = |x_i|\right\} = \bigcup_{t \in \mathbb{R}} M_i(|t|), \ i = 1, \ldots, N,$$

segue que

$$\mathbb{P}_N(A) = \sum_{i=1}^{N} \int_{A \cap B_i} f(x_1, \ldots, x_N) \mathrm{d}\lambda_N \ ,$$

pois $\mathbb{R}^N = \bigcup_{i=1}^N B_i$ e $\lambda_N(B_i \cap B_j) = 0$, $i \neq j$. Assim,

$$\mathbb{P}_N(A) = \sum_{i=1}^N \int_{A\cap B_i} \psi\left(\max_{1\leq j\leq N}\{|x_j|\}\right) \mathrm{d}\lambda_N =$$

$$= \sum_{i=1}^N \int_{A\cap B_i} \psi(|x_i|)\mathrm{d}\lambda_N = \sum_{i=1}^N \int_{\mathbb{R}^N} \psi(|x_i|)\mathbb{I}_{A\cap B_i}((x_1,\ldots,x_N))\mathrm{d}\lambda_N .$$

Pelo teorema de Tonelli, segue que

$$\mathbb{P}_N(A) = \sum_{i=1}^N \Big[\int_{\mathbb{R}} \Big[\int_{\mathbb{R}^{N-1}} \psi(|m|)\mathbb{I}_{\varphi^i(A\cap M_i(m))}(\varphi^i(x_1,\ldots,x_N))\mathrm{d}\lambda_{N-1}\Big]\mathrm{d}\lambda_1(m)\Big]$$

$$= \sum_{i=1}^N \Big[\int_{\mathbb{R}} \psi(|m|)\Big[\int_{\mathbb{R}^{N-1}} \mathbb{I}_{\varphi^i(A\cap M_i(m))}(\varphi^i(x_1,\ldots,x_N))\mathrm{d}\lambda_{N-1}\Big]\lambda_1(m)\Big] =$$

$$= \sum_{i=1}^N \int_{\mathbb{R}} \psi(|m|)\lambda_{N-1}(\varphi^i(A\cap M_i(m)))\mathrm{d}m =$$

$$= \sum_{i=1}^N \Big[\int_{-\infty}^0 \psi(-m)\lambda_{N-1}(\varphi^i(A\cap M_i(m)))\mathrm{d}m + \int_0^\infty \psi(m)\lambda_{N-1}(\varphi^i(A\cap M_i(m)))\mathrm{d}m\Big] =$$

$$= \sum_{i=1}^N \Big[\int_0^\infty \psi(m)\lambda_{N-1}(\varphi^i(A\cap M_i(-m)))\mathrm{d}m + \int_0^\infty \psi(m)\lambda_{N-1}(\varphi^i(A\cap M_i(m)))\mathrm{d}m\Big] =$$

$$= \sum_{i=1}^N \int_0^\infty \psi(m)\left[\lambda_{N-1}(\varphi^i(A\cap M_i(-m))) + \lambda_{N-1}(\varphi^i(A\cap M_i(m)))\right]\mathrm{d}m =$$

$$= \int_0^\infty \psi(m)\sum_{i=1}^N \left[\lambda_{N-1}(\varphi^i(A\cap M_i(m))) + \lambda_{N-1}(\varphi^i(A\cap M_i(-m)))\right]\mathrm{d}m .$$

Da última igualdade e de (2.21), segue que

$$\mathbb{P}_N(A) = \int_0^\infty \varphi(m) 2N(2m)^{N-1} Q_{Nm}(A) \mathrm{d}m, \quad A \in \mathcal{B}(\mathbb{R}^N). \tag{2.22}$$

Para $A = [-t, t]^N$, temos

$$Q_{Nm}(A) = \mathbb{I}_{(0,t]}(m),$$

uma vez que é uma medida concentrada em $H_m^N = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : \max_{1 \leq i \leq N}\{|x_i|\} = m\}$. Portanto,

$$\mathbb{P}_N([-t, t]^N) = \int_0^t \psi(m) 2N(2m)^{N-1} \mathrm{d}m \ .$$

Mas, $\mathbb{P}_N([-t, t]^N)$ é o valor da função de distribuição de $T_N = \max_{1 \leq i \leq N}\{|X_i|\}$, F_T, no ponto $t > 0$. Assim,

$$F_T(t) = \int_0^t \psi(m) 2N(2m)^{N-1} \mathrm{d}m \ .$$

Da última igualdade, segue que

$$f_T(t) = \psi(t) 2N(2t)^{N-1} \mathbb{I}_{\mathbb{R}_+}(t) \tag{2.23}$$

é a função densidade de probabilidade de T_N (provamos, assim, o resultado em (2.20)).

De (2.20) e (2.22), notamos que se $(X_1, \ldots, X_N)$ são variáveis aleatórias absolutamente contínuas com função densidade de probabilidade conjunta satisfazendo (2.18), então a distribuição de $(X_1, \ldots, X_N)$ corresponde a uma mistura dos elementos da família $\{Q_{Nm} : m > 0\}$ na variável T_N.

Mas, por hipótese, $X_1, \ldots, X_N$ são independentes e permutáveis (notemos que $f(x_1, \ldots, x_N) = f(x_{\pi_1}, \ldots, x_{\pi_N}) = \psi(\max_{1 \leq i \leq N}\{|x_i|\})$ para qualquer permutação π dos elementos

de $\{1, \ldots, N\}$). Assim, para todo $t > 0$,

$$f_T(t) = N[G(t) - G(-t)]^{N-1}[g(t) + g(-t)], \tag{2.24}$$

onde $G(g)$ é a função de distribuição (função densidade de probabilidade) de X_1. De (2.23) e (2.24), temos que

$$\psi(t)2N(2t)^{N-1} = N[G(t) - G(-t)]^{N-1}[g(t) + g(-t)]. \tag{2.25}$$

Mas, $\psi(t) = f(t, \ldots, t) = [g(t)]^N$ e g é simétrica (notemos que $f(x_1, \ldots, x_N) = f(-x_1, \ldots, x_N)$), de modo que $g(t) = g(-t)$, $t \in \mathbb{R}$, e $G(t) - G(-t) = 2G(t) - 1$. Voltando em (2.25), obtemos que

$$[g(t)]^N 2N(2t)^{N-1} = N2^{N-1}[G(t) - \frac{1}{2}]^{N-1}2g(t) \Rightarrow$$

$$\Rightarrow [tg(t)]^{N-1} = [G(t) - \frac{1}{2}]^{N-1} \Rightarrow tg(t) = G(t) - \frac{1}{2},$$

$\forall t > 0$ tal que $g(t) > 0$. Da última relação, segue que se $c > 0$ é ponto da imagem de g, então $A_c = \{x > 0 : g(x) = c\}$ é limitado superiormente (caso contrário, existiria $x' > \frac{1}{2c}$ pertencente a A_c com $G(x') > 1$). Por outro lado, afirmamos que g assume apenas dois valores, a saber, 0 e c, $c > 0$. De fato, sejam $c_1, c_2 > 0$ pontos da imagem de g. Vamos verificar que devemos ter, necessariamente, $c_1 = c_2$. Se $c_1, c_2 > 0$ são pontos da imagem de g, então tomemos $x_1, x_2 \in \mathbb{R}_+$ tais que $g(x_i) = c_i$, i=1,2. Sem perda de generalidade, suponhamos $x_1 < x_2$. Assim, por hipótese, temos

$$f(x_1, x_2, \ldots, x_2) = g(x_1)[g(x_2)]^{N-1} = c_1(c_2)^{N-1} = \psi(x_2) \text{ e}$$

$$f(x_2, x_2, \ldots, x_2) = [g(x_2)]^N = (c_2)^N = \psi(x_2),$$

donde segue que $c_1 = c_2$. Assim, como g assume apenas os valores 0 e c e A_c é limitado superiormente, segue que o suporte de g é limitado superiormente. Seja, então, θ o seu máximo. Logo, $G(\theta) = 1$, o que implica que $c = \frac{1}{2\theta}$. Por fim, verificamos que g é igual a $\frac{1}{2\theta}$ no intervalo $(0, \theta)$ λ_1-q.c.. Com efeito,

$$G(\theta) - G(0) = \frac{1}{2} = \int_{\{g=0\}\cap(0,\theta)} g \, \mathrm{d}\lambda_1 + \int_{\{g=\frac{1}{2\theta}\}\cap(0,\theta)} g \, \mathrm{d}\lambda_1 =$$

$$= \frac{1}{2\theta}\lambda_1(\{g = \frac{1}{2\theta}\} \cap (0, \theta)) \quad \Rightarrow \quad \lambda_1((0, \theta) \cap \{g = 0\}) = 0.$$

Finalmente, da simetria de g e dos resultados do último parágrafo, segue que

$$G(t) = \begin{cases} 0, & t < -\theta, \\ \frac{1}{2} + ct, & -\theta \leq t < \theta, \\ 1, & t \geq \theta, \end{cases}$$

isto é, X_1 tem distribuição uniforme em $(-\theta, \theta)$, $\theta > 0$, concluindo a prova. ■

Pode-se verificar facilmente que a recíproca do teorema 2.3.1 é verdadeira. Neste caso, se $X_1, \ldots, X_N$ são variáveis aleatórias independentes uniformes em $(-\theta, \theta)$, $\theta > 0$, então

$$f(x_1, \ldots, x_N) = \left(\frac{1}{2\theta}\right)^N \mathbb{I}_{(-\theta,\theta)} \left(\max_{1 \leq i \leq N} \{|x_i|\}\right).$$

Assim, considerando-se a função não-negativa $\psi : \mathbb{R} \to \mathbb{R}_+$ dada por

$$\psi(t) = \left(\frac{1}{2\theta}\right)^N \mathbb{I}_{(-\theta,\theta)}(t),$$

obtemos o resultado.

No que segue, consideramos a distribuição uniforme dependendo de dois parâmetros $\theta_1, \theta_2 \in \mathbb{R}$, $\theta_1 < \theta_2$.

Teorema 2.3.2 *Sejam $X_1, \ldots, X_N$, $N \geq 3$, variáveis aleatórias absolutamente contínuas independentes a valores em $\mathbb{R}$ com função densidade de probabilidade conjunta f. Se $f(x_1, \ldots, x_N) = \psi(x_{(1)}, x_{(N)})$, $\forall (x_1, \ldots, x_N) \in \mathbb{R}^N$, para alguma função não-negativa ψ que torna f uma densidade de probabilidade em $\mathbb{R}^N$, então X_i tem distribuição uniforme em (θ_1, θ_2), para $\theta_1, \theta_2 \in \mathbb{R}$, com $\theta_1 < \theta_2$, $i = 1, \ldots, N$.*

Prova Para obtermos a distribuição de X_i, $i = 1, \ldots, N$, derivaremos duas formas funcionais para a função densidade de probabilidade conjunta de $(X_{(1)}, X_{(N)})$. Da igualdade de tais funções seguirá o resultado.

Inicialmente, vamos obter $\mathbb{P}_N$, a distribuição de probabilidade de $(X_1, \ldots, X_N)$, em termos de ψ e da distribuição uniforme sobre $\chi^N_{m_1,m_2} = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : x_{(1)} = m_1, x_{(N)} = m_2\}$, $m_1, m_2 \in \mathbb{R}$, $m_1 < m_2$, Q_{Nm_1,m_2}, dada por (Iglesias et al. (1998))

$$Q_{Nm_1,m_2}(B) = \frac{\sum_{i=1}^N \sum_{j \neq i} \lambda_{N-2}(\varphi^{ij}(B \cap M_{ij}(m_1, m_2)))}{N(N-1)(m_2 - m_1)^{N-2}}, \quad B \in \mathcal{B}(\mathbb{R}^N), \tag{2.26}$$

onde $M_{ij}(m_1, m_2) = \{(x_1, \ldots, x_N) \in \chi^N_{m_1,m_2} : x_i = m_1, x_j = m_2, i \neq j\}$ e $\varphi^{ij} : \mathbb{R}^N \to \mathbb{R}^{N-2}$ é dada por $\varphi^{ij}(x_1, \ldots, x_N) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{j-1}, x_{j+1}, \ldots, x_N)$.

Assim, para $A \in \mathcal{B}(\mathbb{R}^N)$, temos

$$\mathbb{P}_N(A) = \int_A f(x_1, \ldots, x_N) \mathrm{d}\lambda_N \ .$$

Definindo $B_{ij} = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : x_{(1)} = x_i, x_{(N)} = x_j\}$, $i = 1, \ldots, N$, $j \neq i$, segue que

$$\mathbb{P}_N(A) = \sum_{i=1}^N \sum_{j \neq i} \int_{A \cap B_{ij}} f(x_1, \ldots, x_N) \mathrm{d}\lambda_N \ ,$$

uma vez que $\mathbb{R}^N = \bigcup_{\substack{i \in \{1, \ldots, N\} \\ j \neq i}} B_{ij}$ e $\lambda_N(B_{i_1 j_1} \cap B_{i_2 j_2}) = 0$, se $(i_1, j_1) \neq (i_2, j_2)$, $i_k \neq j_k$, $k = 1, 2$. Assim,

$$\mathbb{P}_N(A) = \sum_{i=1}^{N}\sum_{j\neq i}\int_{A\cap B_{ij}} \psi(x_{(1)}, x_{(N)})\mathrm{d}\lambda_N =$$

$$= \sum_{i=1}^{N}\sum_{j\neq i}\int_{A\cap B_{ij}} \psi(x_i, x_j)\mathrm{d}\lambda_N = \sum_{i=1}^{N}\sum_{j\neq i}\int_{\mathbb{R}^N} \psi(x_i, x_j)\mathbb{I}_{A\cap B_{ij}}((x_1,\ldots,x_N))\mathrm{d}\lambda_N .$$

Pelo teorema de Tonelli, segue que

$$\mathbb{P}_N(A) = \sum_{i=1}^{N}\sum_{j\neq i}\Big[\int_{\mathbb{R}^2}\Big[\int_{\mathbb{R}^{N-2}} \psi(u,v)\mathbb{I}_{\varphi^{ij}(A\cap M_{ij}(u,v))}(\varphi^{ij}(x_1,\ldots,x_N))\mathrm{d}\lambda_{N-2}\Big]\mathrm{d}\lambda_2(u,v)$$

$$= \sum_{i=1}^{N}\sum_{j\neq i}\Big[\int_{\mathbb{R}^2} \psi(u,v)\Big[\int_{\mathbb{R}^{N-2}} \mathbb{I}_{\varphi^{ij}(A\cap M_{ij}(u,v))}(\varphi^{ij}(x_1,\ldots,x_N))\mathrm{d}\lambda_{N-2}\Big]\mathrm{d}\lambda_2(u,v)\Big] =$$

$$= \sum_{i=1}^{N}\sum_{j\neq i}\int_{\mathbb{R}^2} \psi(u,v)\lambda_{N-2}(\varphi^{ij}(A\cap M_{ij}(u,v)))\mathrm{d}\lambda_2(u,v) =$$

$$= \int_{\mathbb{R}^2} \psi(u,v)\sum_{i=1}^{N}\sum_{j\neq i}\lambda_{N-2}(\varphi^{ij}(A\cap M_{ij}(u,v)))\mathrm{d}\lambda_2(u,v) =$$

$$= \int_{-\infty}^{\infty}\int_{u}^{\infty} \psi(u,v)\sum_{i=1}^{N}\sum_{j\neq i}\lambda_{N-2}(\varphi^{ij}(A\cap M_{ij}(u,v)))\mathrm{d}u\mathrm{d}v .$$

Da última igualdade e de (2.26), segue que

$$\mathbb{P}_N(A) = \int_{-\infty}^{\infty}\int_{u}^{\infty} \psi(u,v)N(N-1)(v-u)^{N-2}Q_{Nu,v}(A)\mathrm{d}v\mathrm{d}u, \quad A\in\mathcal{B}(\mathbb{R}^N) . \quad (2.27)$$

Considerando $A = \{(x_1,\ldots,x_N)\in\mathbb{R}^N : x_{(1)}\leq a,\ x_{(n)}\leq b\}$, temos

$$Q_{Nu,v}(A) = \mathbb{I}_{(-\infty,a]}(u)\mathbb{I}_{(-\infty,b]}(v),$$

uma vez que $Q_{Nu,v}$ é uma medida concentrada em $\chi_{u,v}^N = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : x_{(1)} = u,\ x_{(N)} = v\}$. Logo,

$$\mathbb{P}_N(A) = \int_{-\infty}^{a} \int_{u}^{b} \psi(u,v) N(N-1)(v-u)^{N-2} \mathrm{d}v\mathrm{d}u \ .$$

Mas, $\mathbb{P}_N(A)$ é o valor da função de distribuição de $(X_{(1)}, X_{(N)})$, $F_{X_{(1)},X_{(N)}}$, no ponto (a,b). Assim,

$$F_{X_{(1)},X_{(N)}}(a,b) = \int_{-\infty}^{a} \int_{u}^{b} \psi(u,v) N(N-1)(v-u)^{N-2} \mathrm{d}v\mathrm{d}u \ .$$

Da última igualdade, segue que

$$f_{X_{(1)},X_{(N)}}(a,b) = \psi(a,b) N(N-1)(b-a)^{N-2} \quad , a < b, \tag{2.28}$$

é a função densidade de probabilidade conjunta de $(X_{(1)}, X_{(N)})$.

De (2.28) e (2.27), é fácil ver que se $(X_1, \ldots, X_N)$ são variáveis aleatórias absolutamente contínuas com função densidade de probabilidade conjunta satisfazendo $f(x_1, \ldots, x_N) = \psi(x_{(1)}, x_{(N)})$, $\forall (x_1, \ldots, x_N) \in \mathbb{R}^N$, então a distribuição de $(X_1, \ldots, X_N)$ corresponde a uma mistura dos elementos da família $\{Q_{Nm_1,m_2} : m_1, m_2 \in \mathbb{R}, m_1 < m_2\}$ no vetor $(X_{(1)}, X_{(N)})$.

Por outro lado, $X_1, \ldots, X_N$ são independentes e permutáveis ($f(x_1, \ldots, x_N) = f(x_{\pi_1}, \ldots, x_{\pi_N}) = \psi(x_{(1)}, x_{(N)})$ para qualquer permutação π dos elementos de $\{1, \ldots, N\}$). Assim, para todo $a, b \in \mathbb{R}$, $a < b$,

$$f_{X_{(1)},X_{(N)}}(a,b) = N(N-1)[G(b) - G(a)]^{N-2} g(b) g(a), \tag{2.29}$$

onde $G(g)$ é a função de distribuição (função densidade de probabilidade) de X_1. De (2.28) e (2.29), temos que

$$\psi(a,b) N(N-1)(b-a)^{N-2} = N(N-1)[G(b) - G(a)]^{N-2} g(b) g(a). \tag{2.30}$$

Mas, $\psi(a,b) = f(a,t,\ldots,t,b) = g(a)g(b)[g(t)]^{N-2}$, $\forall t \in [a,b]$. Assim,

$$g(a)g(b)[g(t)]^{N-2}N(N-1)(b-a)^{N-2} = N(N-1)[G(b)-G(a)]^{N-2}g(a)g(b) \Rightarrow$$

$$\Rightarrow \; [g(t)]^{N-2} = \left[\frac{G(b)-G(a)}{b-a}\right]^{N-2} \Rightarrow$$

$$\Rightarrow \; g(t) = \frac{G(b)-G(a)}{b-a} = \text{ constante, } \forall t \in [a,b], \tag{2.31}$$

para todo $a, b \in \mathbb{R}$, $a < b$, tais que $g(a), g(b) > 0$.

De (2.31), concluímos que $\exists\ \theta_1, \theta_2 \in \mathbb{R}$, $\theta_1 < \theta_2$, tais que $g(t) = 0$, $\forall t < \theta_1$, e $g(t) = 0$, $\forall t > \theta_2$ (caso contrário, g seria constante em um intervalo ilimitado de $\mathbb{R}$ e, conseqüentemente, não seria uma função densidade de probabilidade). Deste modo, temos que g deve satisfazer

$$g(t) = \begin{cases} 0, & t < \theta_1, \\ \dfrac{1}{\theta_2 - \theta_1}, & \theta_1 \leq t \leq \theta_2, \\ 0, & t > \theta_2, \end{cases}$$

isto é, X_1 é uniformemente distribuído em (θ_1, θ_2), $\theta_1, \theta_2 \in \mathbb{R}$, com $\theta_1 < \theta_2$. ■

Notemos ainda que a recíproca do teorema 2.3.2 é válida. Se $X_1, \ldots, X_N$ são variáveis aleatórias independentes uniformes em (θ_1, θ_2), então

$$f(x_1, \ldots, x_N) = \left(\frac{1}{\theta_2 - \theta_1}\right)^N \mathbb{I}_{(\theta_1,\infty)}(x_{(1)})\mathbb{I}_{(x_{(1)},\theta_2)}(x_{(N)}) \; .$$

Assim, basta considerarmos a função não-negativa $\psi : \mathbb{R}^2 \to \mathbb{R}_+$ dada por

$$\psi(t_1, t_2) = \left(\frac{1}{\theta_2 - \theta_1}\right)^N \mathbb{I}_{(\theta_1,\infty)}(t_1)\mathbb{I}_{(t_1,\theta_2)}(t_2) \; .$$

2.3.2 Caso bivariado

Nesta subseção, consideramos uma sequência $(X_1, Y_1), \ldots, (X_N, Y_N)$, $N \geq 2$, de vetores aleatórios bivariados absolutamente contínuos assumindo valores em $\mathbb{R}_+^2$, cuja função densidade de probabilidade conjunta f satisfaz

$$f((x_1, y_1), \ldots, (x_N, y_N)) = \psi_N(\max_{1 \leq i \leq N}\{x_i + y_i\}), \tag{2.32}$$

para uma apropriada função não-negativa ψ_N em $\mathbb{R}_+$ que torna f uma função densidade de probabilidade em $(\mathbb{R}^2)^N$.

Teorema 2.3.3 *Sejam $X_1, \ldots, X_N$, $N \geq 2$, vetores aleatórios absolutamente contínuos independentes assumindo valores em $\mathbb{R}_+^2$ com função densidade de probabilidade conjunta f. Nestas condições, $f((x_1, y_1), \ldots, (x_N, y_N)) = \psi_N(\max_{1 \leq i \leq N}\{x_i + y_i\})$ para alguma função não-negativa ψ, $\forall((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{R}_2^+)^N$, se, e somente se, (X_i, Y_i) é uniformemente distribuído sobre $\{(x, y) \in \mathbb{R}_+^2 : x + y \leq \theta\})$, para algum $\theta > 0$, $i = 1, \ldots, N$.*

Prova ($\Leftarrow$) Neste sentido, a demonstração é imediata. Se $((X_1, Y_1), \ldots, (X_N, Y_N))$ são vetores aleatórios independentes uniformes no triângulo $\{(x, y) \in \mathbb{R}_+^2 : x + y \leq \theta\}$, $\theta > 0$, então

$$f((x_1, y_1), \ldots, (x_N, y_N)) = \left(\frac{2}{\theta^2}\right)^N \mathbb{I}_{(0,\theta)}(\max_{1 \leq i \leq N}\{x_i + y_i\}) \ .$$

Considerando-se então a função não-negativa $\psi : \mathbb{R} \to \mathbb{R}_+$ dada por

$$\psi(t) = \left(\frac{2}{\theta^2}\right)^N \mathbb{I}_{(0,\theta)}(t),$$

obtemos o resultado desejado.

($\Rightarrow$) Para caracterizarmos (X_i, Y_i), $i = 1, \ldots, N$, obteremos duas expressões para a função densidade de probabilidade de $T_N = \max_{1 \leq i \leq N}\{X_i + Y_i\}$. A distribuição de (X_1, Y_1) decorrerá da igualdade de tais expressões.

Primeiramente, vamos calcular a distribuição de probabilidade de $((X_1, Y_1), \ldots, (X_N, Y_N))$, $\mathbb{P}_N$, em termos de ψ e da distribuição uniforme sobre $\chi_m^N = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{R}_+^2)^N : \max_{1 \leq i \leq N}\{x_i + y_i\} = m\}$, $m > 0$, Q_{Nm}, dada por (Esteves et al. (2001))

$$Q_{Nm}(B) = \frac{\sum_{i=1}^N \lambda_{2N-2}(\varphi^i(B \cap M_i(m)))}{N(\frac{m^2}{2})^{N-1}}, \quad B \in \mathcal{B}((\mathbb{R}_+^2)^N), \tag{2.33}$$

onde $M_i(m) = \{((x_1, y_1), \ldots, (x_N, y_N)) \in \chi_m^N : x_i + y_i = m\}$ e $\varphi^i : (\mathbb{R}^2)^N \to \mathbb{R}^{2N-2}$ é dada por $\varphi^i(((x_1, y_1), \ldots, (x_N, y_N))) = (x_1, y_1, \ldots, x_{i-1}, y_{i-1}, x_{i+1}, y_{i+1}, \ldots, x_N, y_N)$. Assim, para $A \in \mathcal{B}((\mathbb{R}_+^2)^N)$,

$$\mathbb{P}_N(A) = \int_A f(((x_1, y_1), \ldots, (x_N, y_N)))\mathrm{d}\lambda_2^{(N)} .$$

Definindo $B_i = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{R}_+^2)^N : \max_{1 \leq j \leq N}\{x_j + y_j\} = x_i + y_i\}$, temos que

$$\mathbb{P}_N(A) = \sum_{i=1}^N \int_{A \cap B_i} f(((x_1, y_1), \ldots, (x_N, y_N)))\mathrm{d}\lambda_2^{(N)} ,$$

pois $(\mathbb{R}_+^2)^N = \bigcup_{i=1}^N B_i$ e $\lambda_2^{(N)}(B_i \cap B_j) = 0$, $i \neq j$. Logo,

$$\mathbb{P}_N(A) = \sum_{i=1}^N \int_{A \cap B_i} \psi(\max_{1 \leq j \leq N}\{x_j + y_j\})\mathrm{d}\lambda_2^{(N)} =$$

$$= \sum_{i=1}^N \int_{A \cap B_i} \psi(x_i + y_i)\mathrm{d}\lambda_2^{(N)} =$$

$$= \sum_{i=1}^{N} \int_{(\mathbb{R}_+^2)^N} \psi(x_i + y_i) \mathbb{I}_{A \cap B_i}(((x_1, y_1), \ldots, (x_N, y_N))) \mathrm{d}\lambda_2^{(N)} .$$

Pelo Teorema de Tonelli, segue que

$$\mathbb{P}_N(A) \quad =$$

$$= \sum_{i=1}^{N} \Big[\int_{\mathbb{R}_+^2} \Big[\int_{(\mathbb{R}_+^2)^{N-1}} \psi(u+v) \mathbb{I}_{\varphi^{-i}(A \cap M_i(u+v))}(\varphi^{-i}(((x_1, y_1), \ldots, (x_N, y_N))) \mathrm{d}\lambda_2^{(N-1)} \Big] \mathrm{d}\lambda_2(u, v) \Big],$$

onde $\varphi^{-i} : (\mathbb{R}^2)^N \to (\mathbb{R}^2)^{N-1}$ é dada por

$$\varphi^{-i}(((x_1, y_1), \ldots, (x_N, y_N))) = ((x_1, y_1), \ldots, (x_{i-1}, y_{i-1}), (x_{i+1}, y_{i+1}), \ldots, (x_N, y_N)).$$

Então,

$$\mathbb{P}_N(A) = \sum_{i=1}^{N} \Big[\int_{\mathbb{R}_+^2} \psi(u+v) \Big[\int_{\mathbb{R}_+^{2N-2}} \mathbb{I}_{\varphi^i(A \cap M_i(u+v))}(\varphi^i(((x_1, y_1), \ldots, (x_N, y_N))) \mathrm{d}\lambda_{2N-2} \Big] \mathrm{d}\lambda_2(u, v) \Big]$$

$$= \sum_{i=1}^{N} \int_{\mathbb{R}_+^2} \psi(u+v) \lambda_{2N-2}(\varphi^i(A \cap M_i(u+v))) \mathrm{d}u\mathrm{d}v \quad =$$

$$= \int_{\mathbb{R}_+^2} \psi(u+v) \sum_{i=1}^{N} \lambda_{2N-2}(\varphi^i(A \cap M_i(u+v))) \mathrm{d}u\mathrm{d}v .$$

De (2.33), segue que

$$\mathbb{P}_N(A) = \int_{\mathbb{R}_+^2} \psi(u+v) N \left(\frac{(u+v)^2}{2} \right)^{N-1} Q_{N(u+v)}(A) \mathrm{d}u\mathrm{d}v, \quad A \in \mathcal{B}((\mathbb{R}_+^2)^N). \qquad (2.34)$$

Tomando, em particular, $A = [\{(x, y) \in \mathbb{R}_+^2 : x + y \leq t\}]^N$, temos

$$Q_{N(u+v)}(A) = \mathbb{I}_{(0,t)}(\mu + v) = \mathbb{I}_{(0,t)}(u) \mathbb{I}_{(0,t-u)}(v),$$

pois $Q_{N(u+v)}$ é uma medida concentrada em $\chi^N_{u+v} = \{((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathbb{R}^2_+)^N : \max_{1\leq i\leq N}\{x_i + y_i\} = u + v\}$. Portanto,

$$\mathbb{P}_N(A) = \int_0^t \int_0^{t-u} \psi(u+v) N \left(\frac{(u+v)^2}{2}\right)^{N-1} \mathrm{d}v\mathrm{d}u \ .$$

Mas, $\mathbb{P}_N(A)$ é o valor da função de distribuição de $T_N = \max_{1\leq i\leq N}\{X_i + Y_i\}$, F_{T_N}, calculada no ponto $t > 0$. Assim,

$$F_{T_N}(t) = \int_0^t \int_0^{t-u} \psi(u+v) N \left(\frac{(u+v)^2}{2}\right)^{N-1} \mathrm{d}v\mathrm{d}u \ .$$

Fazendo a transformação $a = u + v$ e $b = v$, segue que

$$F_{T_N}(t) = \int_0^t \psi(a) \frac{N}{2^{N-1}} a^{2N-2} \mathrm{d}a \ .$$

Logo,

$$f_{T_N}(t) = \psi(t) \frac{N}{2^{N-1}} t^{2N-2} \mathbb{I}_{\mathbb{R}_+}(t) \tag{2.35}$$

é a função densidade de probabilidade de T_N.

De (2.35) e (2.34), verificamos que se $((X_1, Y_1), \ldots, (X_N, Y_N))$ são vetores aleatórias absolutamente contínuos com função densidade de probabilidade conjunta satisfazendo (2.32), então a distribuição de $((X_1, Y_1), \ldots, (X_N, Y_N))$ corresponde a uma mistura dos elementos da família $\{Q_{Nm} : m > 0\}$ na variável T_N.

Mas, por outro lado, $(X_1, Y_1), \ldots, (X_N, Y_N)$ são independentes e permutáveis ($f(((x_1, y_1, \ldots, x_N, y_N))) = f(((x_{\pi_1}, y_{\pi_1}), \ldots, (x_{\pi_N}, y_{\pi_N}))) = \psi(\max_{1\leq i\leq N}\{x_i + y_i\})$ para toda permutação π dos elementos de $\{1, \ldots, N\}$). Assim, temos, para $t > 0$, que

$$f_{T_N}(t) = N[F_{X_1+Y_1}(t)]^{N-1} f_{X_1+Y_1}(t), \tag{2.36}$$

onde $F_{X_1+Y_1}(f_{X_1+Y_1})$ é a função de distribuição (função densidade de probabilidade) de $X_1 + Y_1$. De (2.35) e (2.36), temos que

$$\frac{N}{2^{N-1}}\psi(t)t^{2N-1} = N[F_{X_1+Y_1}(t)]^{N-1}f_{X_1+Y_1}(t). \tag{2.37}$$

Mas, $\psi(t) = f((0,t),\ldots,(0,t)) = [g(0,t)]^N$, onde g é a função densidade de probabilidade de (X_1, Y_1). Por outro lado, $g(a, t-a) = g(0,t)$, $\forall a \in [0,t]$, $\forall t > 0$, o que implica que

$$f_{X_1+Y_1}(t) = \int_{-\infty}^{\infty} g(a,t-a)\mathrm{d}a = \int_0^t g(a,t-a)\mathrm{d}a = \int_0^t g(0,t)\mathrm{d}a = tg(0,t) .$$

Da última relação e da expressão para $\psi(t)$ em função de $g(0,t)$, obtemos que

$$\psi(t) = \left[\frac{f_{X_1+Y_1}(t)}{t}\right]^N .$$

Substituindo em (2.37), resulta que

$$\frac{N}{2^{N-1}}\frac{[f_{X_1+Y_1}(t)]^N}{t^N}t^{2N-1} = N[F_{X_1+Y_1}(t)]^{N-1}f_{X_1+Y_1}(t) \Rightarrow$$

$$\Rightarrow [f_{X_1+Y_1}(t)]^{N-1}t^{N-1} = [2F_{X_1+Y_1}(t)]^{N-1} \Rightarrow$$

$$F_{X_1+Y_1}(t) = \frac{t^2}{2}g(0,t) ,$$

$\forall t > 0$ tal que $g(0,t) > 0$. Da relação acima, segue que se $c > 0$ pertence à imagem de g, então $A_c = \{t > 0 : g(0,t) = c\}$ é limitado superiormente (caso contrário, existiria $t' > \sqrt{\frac{2}{c}}$ pertencente a A_c com $F_{X_1+Y_1}(t') > 1$). Por outro lado, afirmamos que g assume apenas dois valores, a saber, 0 e c, $c > 0$. De fato, sejam $c_1, c_2 > 0$ pontos da imagem de g. Vamos verificar que, necessariamente, $c_1 = c_2$. Se $c_1, c_2 > 0$ são pontos da imagem

de g, então tomemos $(x_1, y_1), (x_2, y_2) \in \mathbb{R}_+^2$ tais que $g(x_i, y_i) = c_i$, i=1,2. Sem perda de generalidade, suponhamos que $x_1 + y_1 < x_2 + y_2$. Assim, por hipótese, temos

$$f((x_1, y_1), (x_2, y_2), \ldots, (x_2, y_2)) = g(x_1, y_1)[g(x_2, y_2)]^{N-1} = c_1(c_2)^{N-1} = \psi(x_2 + y_2) \quad \text{e}$$

$$f((x_2, y_2), (x_2, y_2), \ldots, (x_2, y_2)) = [g(x_2, y_2)]^N = (c_2)^N = \psi(x_2 + y_2) \ ,$$

donde segue que $c_1 = c_2$. Assim, como g assume apenas os valores 0 e c e A_c é limitado superiormente, segue que o suporte de $f_{X_1+Y_1}$ é limitado superiormente e, portanto, existe $\theta > 0$ tal que $\theta = \max\{t > 0 : f_{X_1+Y_1}(t) > 0\}$, com $g(0, \theta) = c$. Logo, $F_{X_1+Y_1}(\theta) = 1$, o que implica que $c = \frac{2}{\theta^2}$. Por fim, verificamos que g é igual a $\frac{2}{\theta^2}$ em $B_\theta = \{(x, y) \in \mathbb{R}_+^2 : x + y \leq \theta\}$ λ_2-q.c.. Com efeito,

$$P((X_1, Y_1) \in B_\theta) \ = \ F_{X_1+Y_1}(\theta) \ = \ 1 \ \Rightarrow$$

$$\Rightarrow \int_{B_\theta \cap \{g=0\}} g \ \mathrm{d}\lambda_2 \ + \int_{B_\theta \cap \{g=\frac{2}{\theta^2}\}} g \ \mathrm{d}\lambda_2 \ = \int_{B_\theta} \frac{2}{\theta^2} \ \mathrm{d}\lambda_2 \ = \ 1 \ \Rightarrow$$

$$\Rightarrow \ \frac{2}{\theta^2}\lambda_2(B_\theta \cap \{g = \frac{2}{\theta^2}\}) \ = \ 1 \ \Rightarrow \ \lambda_2(B_\theta \cap \{g = \frac{2}{\theta^2}\}) \ = \ \frac{\theta^2}{2} \ \Rightarrow$$

$$\Rightarrow \ \lambda_2(B_\theta \cap \{g = 0\}) \ = \ 0.$$

Logo, temos, finalmente, que

$$g(x, y) = \begin{cases} \frac{2}{\theta^2}, & (x, y) \in B_\theta, \\ 0, & \text{caso contrário}, \end{cases}$$

isto é, (X_1, Y_1) tem distribuição uniforme sobre $\{(x, y) \in \mathbb{R}_+^2 : x + y \leq \theta\}$, $\theta > 0$. ■

2.4 Conclusões

Neste capítulo, apresentamos caracterizações de distribuições uniformes, discretas e contínuas, univariadas e bivariadas, dentro de classes de distribuições multivariadas satisfazendo certas condições de simetria. No caso discreto, a especificação de distribuições de probabilidade multivariadas dependendo apenas do máximo, do mínimo e do máximo, e, ou do máximo das somas das coordenadas e a suposição adicional de independência fornecem uma caracterização dos modelos uniformes em $\{0, \ldots, \theta\}$, $\theta \in \mathbb{N}$, em $\{\theta_1, \ldots, \theta_2\}$, $\theta_1, \theta_2 \in Z$, $\theta_1 < \theta_2$, e em $\{(x, y) \in \mathbb{N}^2 : x + y \leq \theta\}$, $\theta \in \mathbb{N}$, respectivamente. No caso contínuo, a imposição de condições similares sobre as funções densidade de probabilidade de distribuições simétricas multivariadas mais independência permitem uma caracterização dos modelos uniformes em $(-\theta, \theta)$, $\theta > 0$, em (θ_1, θ_2), $\theta_1, \theta_2 \in \mathbb{R}$, $\theta_1 < \theta_2$, e em $\{(x, y) \in \mathbb{R}_+^2 : x + y \leq \theta\}$, $\theta > 0$. Algumas questões, no entanto, não foram contempladas neste capítulo e merecem futura apreciação. Dentre outras, destacamos a possibilidade de caracterizar os vetores aleatórios $\mathbf{X}$ com distribuições pertencentes às classes de distribuições simétricas multivariadas aqui consideradas através de uma representação estocástica da forma $\mathbf{X} = R\mathbf{U}$, para alguma variável aleatória $R \geq 0$ independente de um vetor aleatório $\mathbf{U}$ uniformemente distribuído sobre uma específica região do $\mathbb{R}^N$ $((\mathbb{R}^2)^N)$ (no caso das distribuições esféricas n-variadas, $n \in \mathbb{N}$, sabemos que existe uma caracterização desta forma, com $\mathbf{U}$ sendo uniformemente distribuído sobre a superfície da esfera n-dimensional de raio 1).

Referências

[1] Efron, B. (1969). *Student's t-test under symmetry conditions.* J. Am. Statist. Assoc., 64, 1278–1302.

[2] Fang, K.T.; Kotz, S.; Ng, K.W. (1990). *Symmetric multivariate and related distributions.* Chapman and Hall.

[3] Fortini, S.; Ladelli, L.; Regazzini, E. (2000). *Exchangeability, predictive distributions and parametric models.* Sankhyã, 62(A), 86–109.

[4] Gnedin, A.V. (1995). *On a class of exchangeable sequences.* Statist. Probab. Letters, 25, 351–355.

[5] Iglesias, P.; Pereira, C.A.B.; Tanaka, N.I. (1998). *Characterizations of multivariate spherical distributions in l_∞-norm.* Test, vol.7, 2, 307-324.

[6] Kagan, A.M.; Linnik, Y.V.; Rao, C.R. (1973). *Characterization problems in mathematical statistics.* John Wiley & Sons, New York.

Capítulo 3

Uma variante do modelo de urna de Pólya-Eggenberger

3.1 Introdução

O estudo de modelos de urnas, cujos primórdios remontam ao século XVII (segundo Johnson & Kotz, 1977, p. 22, "a primeira referência a modelos de urnas em problemas de probabilidade aparece nos trabalhos de Huygen (1629–1695)"), tem sido, desde então, alvo de trabalhos de muitos autores devido a sua importância, que, resumidamente, reside em pelo menos dois aspectos: primeiro, a possibilidade de derivação de muitos resultados da Teoria de Probabilidade (ao menos no que diz respeito à probabilidade discreta) via modelos de urnas, o que os tornam um vigoroso instrumento didático (Johnson & Kotz, 1977, por exemplo, obtêm as distribuições de probabilidade discretas comumente utilizadas em métodos estatísticos a partir de um esquema de urnas mais geral); segundo, a possibilidade de modelagem de diversos fenômenos da natureza bem como de problemas reais em várias áreas do conhecimento através de esquemas de urnas (Heitele, 1975, afirma que é possível associar modelos de urnas a grande parte dos experimentos que envolvem incerteza). Finalmente, Pólya (1954) ratifica a importância de modelos de urnas: "Qualquer problema de probabilidade parece comparável a um adequado problema de urnas conten-

do bolas e qualquer fenômeno aleatório parece similar, em certos aspectos essenciais, a sucessivas retiradas de bolas de um sistema de urnas combinadas convenientemente".

Johnson & Kotz (1977) distinguem os principais modelos de urnas em duas categorias: (i) os modelos de urnas com reposição de bolas, dentre os quais podemos destacar os modelos estocásticos de Pólya-Eggenberger para fenômenos envolvendo algumas formas de "contágio" e de Ehrenfest para transferência de calor entre dois corpos isolados e (ii) os modelos de urnas para problemas de ocupação (sem reposição de bolas) tais como os modelos de Bose-Einstein e de Maxwell-Boltzmann. Neste trabalho, no entanto, vamos nos ater apenas a um particular modelo da categoria (i) citada acima: o modelo de Pólya-Enggenberger e suas variações. Uma descrição bastante detalhada sobre o estudo de modelos de urnas em geral é encontrado em Johnson & Kotz (1977).

No que segue, recordamos o modelo de urna de Pólya-Eggenberger bem como apresentamos uma variação deste modelo que acreditamos não ter sido ainda contemplada na literatura.

3.2 O modelo de Pólya-Eggenberger e suas variações

Dentre os diversos modelos de urnas desenvolvidos para representar formas de contágio, o modelo de Pólya-Eggenberger ocupa, indubitavelmente, uma posição central, não só pelo seu cunho pioneiro nesta área de estudo, mas também pela generalidade e riqueza de propriedades que encerra. Na seqüência, faremos uma breve descrição deste modelo e de algumas de suas propriedades.

Podemos descrever o modelo de urna de Pólya-Eggenberger de modo bastante simples. Imaginemos uma urna contendo, inicialmente, a bolas brancas e b bolas pretas. Pólya e Eggenberger (1923) consideram o seguinte procedimento de retiradas sucessivas de bolas da urna: retira-se uma bola da urna de maneira equiprovável e, em seguida, retorna-se

esta bola à urna juntamente com c bolas desta mesma cor (o caso particular $c = 1$ foi estudado anteriormente por Markov (1906)). Notemos que, segundo esta construção, a retirada de uma bola de uma determinada cor, digamos branca, em uma certa etapa do processo aumenta a probabilidade de nova ocorrência deste evento (retirada de uma bola branca) na etapa seguinte (Feller, 1957, dá o nome de "aftereffect" a este fenômeno em que a ocorrência de um evento aumenta (ou diminui) a probabilidade de nova ocorrência deste evento). Esta característica do processo, em concordância com seu conceito de "influence globale", teria levado Pólya a considerá-lo um protótipo bastante razoável para descrever específicas formas de "contágio". A seguir, apresentamos uma formalização do modelo de Pólya-Eggenberger.

Seja $\{X_n\}_{n\geq 1}$ um processo estocástico assumindo valores em $\{0,1\}^\infty$ com medida $\mathbb{P}$ tal que, $\forall n \in \mathbb{N}$, $\forall (x_1, \ldots, x_n) \in \{0,1\}^n$,

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \frac{a(a+c)\cdots(a+(t_n-1)c)b(b+c)\cdots(b+(n-t_n-1)c)}{(a+b)(a+b+c)\cdots(a+b+(n-1)c)}$$

ou

$$P(X_1 = x_1, \ldots, X_n = x_n) = \frac{\Gamma\left(\frac{a}{c}+t_n\right)\Gamma\left(\frac{b}{c}+n-t_n\right)\Gamma\left(\frac{a+b}{c}\right)}{\Gamma\left(\frac{a}{c}\right)\Gamma\left(\frac{b}{c}\right)\Gamma\left(\frac{a+b}{c}+n\right)}, \tag{3.1}$$

onde $t_n = \sum_{i=1}^n x_i$, $a, b, c \in \mathbb{N}$. O processo $\{X_n\}_{n\geq 1}$ descreve a evolução do modelo de urna de Pólya-Eggenberger com configuração inicial (a, b), isto é, com a bolas brancas e b bolas pretas inicialmente na urna, e c bolas acrescidas em cada etapa. X_n corresponde então à variável indicadora de retirada de uma bola branca na n-ésima etapa do processo, $n \in \mathbb{N}$. De (3.1), é fácil ver que o processo de Pólya-Eggenberger é permutável. Deste modo, o processo de Pólya-Eggenberger satisfaz as condições do Teorema da Representação de De Finetti e, portanto, qualquer marginal n-dimensional deste processo pode ser escrita como uma mistura de n variáveis aleatórias de Bernoulli independentes

e identicamente distribuídas (neste caso, a medida de De Finetti, ou misturadora, possui função densidade de probabilidade Beta com parâmetros $\frac{a}{c}$ e $\frac{b}{c}$). A partir do processo de Pólya-Eggenberger, são derivadas as distribuições de Pólya-Eggenberger e Pólya inversa, bem como suas versões multivariadas para as situações em que existem mais de duas cores para as bolas da urna. Um estudo bastante detalhado dessas distribuições e outras caracterizações do modelo de Pólya-Eggenberger são encontrados em Johnson & Kotz (1977).

A partir do trabalho de Pólya e Eggenberger (1923), modificações (ou generalizações) em várias direções para o modelo de Pólya-Eggenberger têm sido desenvolvidas, desde a consideração de mais de duas cores para as bolas na urna até a adoção de diferentes mecanismos de reposição de bolas à urna. Dentre estas alterações destacamos: (1) reposição não apenas de c bolas da mesma cor da bola retirada, mas também de d bolas da outra cor; (2) reposição de bolas das duas cores, como em (1), dependendo da etapa do processo, isto é, $c = c_n$ e $d = d_n$, ou da cor da bola retirada, $c = c_n(x_n)$ e $d = d_n(x_n)$ e (3) reposição de números C e D aleatórios de bolas à urna. Em qualquer uma das três situações acima, o processo obtido não é, em geral, permutável. Consideremos, a título de ilustração, os exemplos.

Exemplo 3.2.1 No caso (1) acima, sejam $a = b = 1$, $c = 3$ e $d = 1$. É fácil ver que

$$P(X_1 = 1, X_2 = 1, X_3 = 0) = \frac{1}{10} \neq P(X_1 = 0, X_2 = 1, X_3 = 1) = \frac{1}{12}.$$

Exemplo 3.2.2 No caso (2), tomemos $a = b = 1$, $c_n(1) = 4$, $c_n(0) = 1$ e $d_n(1) = d_n(0) = 1$, $\forall n \in \mathbb{N}$. Assim,

$$P(X_1 = 1, X_2 = 1, X_3 = 0) = \frac{5}{56} \neq P(X_1 = 0, X_2 = 1, X_3 = 1) = \frac{1}{6}.$$

Exemplo 3.2.3 No caso (3), consideremos $a = b = 1$, $d = 0$ e C distribuído segundo

$P(C=1)=P(C=3)=\frac{1}{2}$. Neste caso,

$$P(X_1=1,X_2=1,X_3=0)=\frac{23}{360}\neq P(X_1=0,X_2=1,X_3=1)=\frac{25}{360}.$$

Apresentamos agora uma nova versão para o modelo de urna de Pólya-Eggenberger.

Consideremos uma urna contendo bolas brancas e pretas. Suponhamos, no entanto, que sejam desconhecidos os números iniciais A e B de bolas brancas e pretas, respectivamente. Como o par (A,B) é desconhecido, é natural, do ponto de vista subjetivista, que atribuamos uma medida de probabilidade sobre $\mathbb{N}^2$ que expresse nossa incerteza a respeito da composição inicial da urna. Assim, a configuração inicial da urna para o modelo que propomos é um vetor aleatório em $(\mathbb{N}^2,\mathcal{P}(\mathbb{N}^2))$. Quanto à evolução do processo, admitimos que, em cada etapa, seja escolhida uma bola da urna uniformemente e que, em seguida, esta seja reposta à urna juntamente com outra bola da mesma cor (notemos que este é o procedimento de retiradas sucessivas do modelo de urna de Pólya-Eggenberger com $c=1$). Ao processo descrito acima, damos o nome de Modelo de Urna de Pólya-Eggenberger com composição (ou configuração) inicial aleatória. Neste caso, se $\{((a_k,b_k),p_k)\}_{k\in\mathbb{N}}$, $(a_k,b_k)\in\mathbb{N}^2$, $p_k\geq 0$, $k\in\mathbb{N}$, com $\sum_{k\in\mathbb{N}}p_k=1$, exprime nossa incerteza a respeito da composição inicial da urna, temos, utilizando a notação do começo desta seção, que $\forall n\in\mathbb{N}$, $\forall(x_1,\ldots,x_n)\in\{0,1\}^n$,

$$\mathbb{P}(X_1=x_1,\ldots,X_n=x_n)=\sum_{k=1}^{\infty}p_k\cdot\text{Pólya}\left((x_1,\ldots,x_n)|(a_k,b_k)\right),$$

ou

$$\mathbb{P}(X_1=x_1,\ldots,X_n=x_n)=\int_{\mathbb{N}^2}\text{Pólya}\left((x_1,\ldots,x_n)|(a,b)\right)\mathrm{d}\mu(a,b), \tag{3.2}$$

onde $\mu:\mathcal{P}(\mathbb{N}^2)\to[0,1]$ é tal que $\mu(A)=\sum_{k:(a_k,b_k)\in A}p_k$, $A\in\mathcal{P}(\mathbb{N}^2)$, e

$$\text{Pólya}\left((x_1,\ldots,x_n)|(a,b)\right) = \frac{\Gamma\left(a+\sum_{i=1}^n x_i\right)\Gamma\left(b+n-\sum_{i=1}^n x_i\right)\Gamma(a+b)}{\Gamma(a)\Gamma(b)\Gamma(a+b+n)}.$$

A expressão (3.2) acima é facilmente obtida condicionando-se a evolução do processo à composição inicial da urna, isto é,

$$P(X_1 = x_1,\ldots,X_n = x_n) = \sum_{k\in\mathbb{N}} P((A,B)=(a_k,b_k)) \prod_{j=1}^n P(X_j = x_j \mid x_1,\ldots,x_{j-1},(a_k,b_k)).$$

Notemos que em (3.2) poderíamos ter formalmente considerado medidas misturadoras μ mais gerais sobre $(\mathbb{R}_+^2, \mathcal{B}(\mathbb{R}_+^2))$. No entanto, como tal medida deve refletir nossa incerteza sobre a composição inicial da urna, (A, B), sabidamente assumindo valores em $\mathbb{N}^2$, não contemplamos estas situações neste trabalho. No que segue, apresentamos algumas propriedades do Modelo de Urna de Pólya-Eggenberger com configuração inicial aleatória.

Proposição 3.2.1 *Qualquer que seja a medida misturadora μ, o modelo de Urna de Pólya-Eggenberger com composição inicial aleatória é permutável, isto é, $\forall n \in \mathbb{N}$, $\forall (x_1,\ldots,x_n) \in \{0,1\}^n$ e $\forall \pi : \{1,\ldots,n\} \to \{1,\ldots,n\}$ bijetora (permutação dos índices $1,\ldots,n$)*

$$P(x_{\pi(1)},\ldots,x_{\pi(n)}) = P(x_1,\ldots,x_n).$$

Prova

$$P(x_{\pi(1)},\ldots,x_{\pi(n)}) = \sum_{k\in\mathbb{N}} p_k \cdot \text{Pólya}\left((x_{\pi(1)},\ldots,x_{\pi(n)})|(a_k,b_k)\right) =$$

$$= \sum_{k\in\mathbb{N}} p_k \cdot \text{Pólya}\left((x_1,\ldots,x_n)|(a_k,b_k)\right) = P(x_1,\ldots,x_n)\ ,$$

onde a penúltima igualdade decorre da permutabilidade do processo de Pólya-Eggenberger usual (isto é, com composição inicial não-aleatória). ■

O resultado da Proposição 3.2.1 é, na verdade, conseqüência imediata da convexidade do conjunto das medidas permutáveis em $\{0,1\}^n$ com a usual σ-álgebra de conjuntos cilíndricos. A Proposição 3.2.1 revela que o Modelo de Urna de Pólya-Eggenberger com composição inical aleatória é permutável, ao contrário das outras variações do Modelo de Pólya-Eggenberger mencionadas anteriormente, que não gozam de tal propriedade. Também por isso, vemos que o modelo que propomos não corresponde a generalização alguma do modelo de urna de Pólya-Eggenberger, dentre as mencionadas. Ademais, o modelo de urna de Pólya-Eggenberger com configuração inicial aleatória, sendo um processo permutável em $\{0,1\}^\infty$, satisfaz as condições do Teorema da Representação de De Finetti (1937), de modo que qualquer marginal n-dimensional deste processo equivale a uma mistura de n variáveis de Bernoulli independentes e identicamente distribuídas. A seguir, caracterizamos então a medida de De Finetti deste processo.

Proposição 3.2.2 *Seja $\{X_n\}_{n\geq 1}$ um modelo de urna de Pólya-Eggenberger com configuração inicial aleatória dada por $\{((a_k,b_k);p_k)\}_{k\in\mathbb{N}}$, $(a_k,b_k)\in\mathbb{N}^2$, $p_k\geq 0$, $k\in\mathbb{N}$, com $\sum_{k\in\mathbb{N}} p_k = 1$. Então a medida de De Finetti deste processo possui função densidade de probabilidade dada por*

$$g(\theta) = \sum_{k\in\mathbb{N}} p_k \mathrm{Beta}\,(\theta;(a_k,b_k)), \tag{3.3}$$

onde $\mathrm{Beta}\,(\theta;(a,b)) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\mathbb{I}_{[0,1]}(\theta)$.

Prova Seja $\mathbb{P}$ a medida do processo $\{X_n\}_{n\geq 1}$. Assim, $\forall n\in\mathbb{N}$, $\forall (x_1,\ldots,x_n)\in\{0,1\}^n$, temos

$$\mathbb{P}(x_1,\ldots,x_n) = \sum_{k\in\mathbb{N}} p_k \cdot \text{Pólya}\,((x_1,\ldots,x_n)|(a_k,b_k)).$$

Mas,

$$\text{Pólya}\left((x_1,\ldots,x_n)|(a_k,b_k)\right) = \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\text{Beta}\left(\theta;(a_k,b_k)\right)\mathrm{d}\theta, \qquad (3.4)$$

uma vez que o modelo de urna de Pólya-Eggenberger com configuração inicial fixa (a_k, b_k) é permutável com medida de De Finetti possuindo função densidade de probabilidade Beta com parâmetros a_k e b_k. Logo,

$$\mathbb{P}(x_1,\ldots,x_n) = \sum_{k\in\mathbb{N}} p_k \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\text{Beta}\left(\theta;(a_k,b_k)\right)\mathrm{d}\theta =$$

$$= \sum_{k\in\mathbb{N}} \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} p_k \text{Beta}\left(\theta;(a_k,b_k)\right)\mathrm{d}\theta =$$

$$= \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\left(\sum_{k\in\mathbb{N}} p_k \cdot \text{Beta}\left(\theta;(a_k,b_k)\right)\right)\mathrm{d}\theta,$$

onde a última igualdade decorre do Teorema de Tonelli. Assim,

$$\mathbb{P}(x_1,\ldots,x_n) = \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\left(\sum_{k\in\mathbb{N}} p_k \cdot \text{Beta}\left(\theta;(a_k,b_k)\right)\right)\mathrm{d}\theta\ .$$

Da unicidade da medida de De Finetti no Teorema da Representação segue o resultado.

■

A Proposição 3.2.2 mostra que o modelo de urna de Pólya-Eggenberger com configuração inicial dada por $\{((a_k,b_k);p_k)\}_{k\in\mathbb{N}}$ tem medida de De Finetti possuindo função densidade de probabilidade igual a uma mistura enumerável de funções densidade Betas, a saber, Betas com parâmetros a_k e b_k, $k \in \mathbb{N}$, ponderadas pelos respectivos pesos p_k.

Na próxima seção, apresentamos o principal resultado deste trabalho, que afirma que, sob certas condições, um processo permutável em $\{0,1\}^\infty$ pode ser "bem aproximado" por um modelo de urna de Pólya-Eggenberger com configuração inicial aleatória.

3.3 Resultado principal

Estabelecemos, a partir de agora, condições para que a medida de probabilidade de um processo permutável a valores em $\{0,1\}^\infty$ possa ser bem aproximada, em distância de variação total, pela medida de um conveniente modelo de urna de Pólya-Eggenberger com composição inicial aleatória. Antes, porém, relembremos um teorema devido a Bernstein que terá papel fundamental nos resultados subseqüentes.

Teorema 3.3.1 (S. Bernstein) *Seja $f : [0,1] \to \mathbb{R}$ contínua. Então, a seqüência de polinômios de Bernstein*

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k} \ , \ n = 1, 2, \ldots$$

converge uniformemente a f quando $n \to \infty$.

Prova Notemos, inicialmente, que se $r_k(x) = \binom{n}{k} x^k (1-x)^{n-k}$, então

$$\sum_{k=0}^{n} r_k(x) = 1, \quad \sum_{k=0}^{n} k r_k(x) = nx \ \sum_{k=0}^{n} (k - nx)^2 \cdot r_k(x) = nx(1-x).$$

Vamos avaliar

$$|f(x) - P_n(x)| = \left| f(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) r_k(x) \right| =$$

$$= \left| \sum_{k=0}^{n} f(x) r_k(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) r_k(x) \right| = \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) r_k(x) \right| \leq$$

$$\leq \sum_{k=0}^{n} \left| f(x) - f\left(\frac{k}{n}\right) \right| r_k(x) \tag{3.1}$$

Como f é contínua em $[0,1]$, temos que f é absolutamente contínua e, portanto, $\forall \varepsilon > 0$, $\exists = (\varepsilon) > 0$ tal que $|x - y| <\Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{2}$. Além disso, $\exists M > 0$ tal que $|f(x)| \leq M$, $\forall x \in [0,1]$.

Dividimos a soma em (3.1) em duas partes. Para $\varepsilon > 0$ e $= (\varepsilon)$, temos

$$|f(x) - P_n(x)| \leq \sum_{k:|k-nx|<n} \left| f(x) - f\left(\frac{k}{n}\right) \right| r_k(x) + \sum_{k:|k-nx|\geq n} \left| f(x) - f\left(\frac{k}{n}\right) \right| r_k(x).$$

Mas,

$$|k - nx| < n \Rightarrow \left| \frac{k}{n} - x \right| <\Rightarrow \left| f(x) - f\left(\frac{k}{n}\right) \right| < \frac{\varepsilon}{2}.$$

Logo,

$$|f(x) - P_n(x)| \leq \frac{\varepsilon}{2} \sum_{k:|k-nx|<n} r_k(x) + \sum_{k:|k-nx|\geq n} \left(|f(x)| + \left| f\left(\frac{k}{n}\right) \right| \right) r_k(x) \leq$$

$$\leq \frac{\varepsilon}{2} + 2M \sum_{k:|k-nx|\geq n} r_k(x) \leq \frac{\varepsilon}{2} + \frac{2Mnx(1-x)}{n^{22}} \leq \frac{\varepsilon}{2} + \frac{M}{2^2 n},$$

uma vez que

$$\sum_{k:|k-nx|\geq n} r_k(x) \leq \frac{nx(1-x)}{n^{22}} \quad x(1-x) \leq \frac{1}{4}.$$

Assim,

$$|f(x) - P_n(x)| < \frac{\varepsilon}{2} + \frac{M}{2^2 n}, \quad \forall x \in [0,1], \ \forall n \in \mathbb{N}.$$

Tomando $n_0 \in \mathbb{N}$ tal que $\frac{M}{2^2 n_0} < \frac{\varepsilon}{2}$, segue que

$$|f(x) - P_n(x)| < \varepsilon, \quad \forall n \geq n_0.$$

Assim, $\forall \varepsilon > 0$, $\exists n_0 = n_0(\varepsilon) \in \mathbb{N}$ tal que

$$n \geq n_0 \Rightarrow |f(x) - P_n(x)| < \varepsilon, \quad \forall x \in [0,1], \text{ ou seja,}.$$

P_n converge uniformemente a f quando $n \to \infty$. ■

Marsden (1974) argumenta que o conhecimento de Teoria da Probabilidade possibilitou a Bernstein melhor compreender e demonstrar este teorema. De fato, várias passagens da prova acima correspondem a resultados bem conhecidos em probabilidade. Além disso, a consideração de um cenário envolvendo um jogo com premiação fornece uma explicação intuitiva, ainda que sob uma perspectiva frequentista, da validade do teorema 3.3.1. Com efeito, considerando-se n lançamentos independentes de uma moeda, nos quais a probabilidade de resultar cara é x, e admitindo-se que $f(\frac{k}{n})$ unidades monetárias são pagas a um jogador quando k caras são obtidas nestes n lançamentos da moeda, $P_n(x)$ é o ganho esperado deste jogador. Para valores grandes de n, a fração de lançamentos que resultam em cara aproxima-se de x e, consequentemente, o ganho médio de tal jogador aproxima-se de $f(x)$, isto é, $P_n(x)$ deve aproximar $f(x)$.

Na seqüência, exibimos as condições para que a solução do problema dos momentos de Hausdorff (Shohat e Tamarkin (1943)) seja absolutamente contínua possuindo função densidade limitada.

Teorema 3.3.2 *Sejam $\{\alpha_n\}_{n\geq 1}$, $\alpha_n \in [0,1]$, uma sequência de números e $\psi(t)$ uma função de distribuição com $\psi(0) = 0$ e $\psi(1) = 1$ tais que*

$$\alpha_n = \int_0^1 t^n \mathrm{d}\psi(t), \quad \forall\, n \in \mathbb{N}, \tag{3.2}$$

Então, $\psi(t) = \int_0^t \varphi(u)\mathrm{d}u$, com $\varphi \geq 0$ limitada, se, e somente se, existe $M > 0$ tal que, $\forall n \in \mathbb{N}$, tem-se

$$\max_{v\in\{0,\ldots,n\}} \left\{ \binom{n}{v} \int_0^1 t^v (1-t)^{n-v} \mathrm{d}\psi(t) \right\} \leq \frac{M}{n+1}.$$

Prova A demonstração que apresentamos aqui é devida a Shohat e Tamarkin (1943), onde são abordadas várias versões do problema clássico dos momentos proposto por Stieltjes

(1894-95).

($\Rightarrow$) Se $\psi(t) = \int_0^t \varphi(u)\mathrm{d}u$, φ limitada, então $\forall n \in \mathbb{N}$, $\forall v \in \{0, \ldots, n\}$, temos

$$\binom{n}{v} \int_0^1 t^v (1-t)^{n-v} \mathrm{d}\psi(t) = \frac{\Gamma(n+1)}{\Gamma(v+1)\Gamma(n-v+1)} \int_0^1 t^v (1-t)^{n-v} \varphi(t) \mathrm{d}t \leq$$

$$\leq \frac{\Gamma(n+1)}{\Gamma(v+1)\Gamma(n-v+1)} M \int_0^1 t^v (1-t)^{n-v} \mathrm{d}t \ ,$$

sendo $M = \sup_{[0,1]} \varphi(t)$. Logo,

$$\binom{n}{v} \int_0^1 t^v (1-t)^{n-v} \mathrm{d}\psi \leq M \frac{\Gamma(n+1)}{\Gamma(v+1)\Gamma(n-v+1)} \frac{\Gamma(v+1)\Gamma(n-v+1)}{\Gamma(n+2)} \leq \frac{M}{n+1}$$

$$\Rightarrow \max_{v \in \{0,\ldots,n\}} \left\{ \binom{n}{v} \int_0^1 t^v (1-t)^{n-v} \mathrm{d}\psi \right\} \leq \frac{M}{n+1}.$$

($\Leftarrow$) Neste caso, consideremos os conjuntos $\mathcal{O}([0,1])$ dos polinômios definidos em $[0,1]$ e $\mathcal{L}_1([0,1], \mathcal{B}([0,1], \lambda_1)$ das classes de equivalência de funções integráveis, isto é, das classes de funções $[f]$ tais que $\int_0^1 |f| \mathrm{d}\lambda_1 < \infty$, munidos da norma $\|\cdot\|_1 = \int_{[0,1]} |\cdot| \mathrm{d}\lambda_1$.

Claramente, $\mathcal{O}([0,1]) \subset \mathcal{L}_1([0,1], \mathcal{B}([0,1]), \lambda_1)$. Definiremos, no que segue, um funcional linear em $\mathcal{O}([0,1])$ e verificaremos que este funcional, limitado, pode ser estendido a $\mathcal{L}_1([0,1], \mathcal{B}([0,1], \lambda_1)$. Seja $\nu : \mathcal{O}([0,1]) \to \mathbb{R}$ o funcional linear que associa ao polinômio $P(t) = \sum_{j=0}^n x_j t^j$ o valor

$$\nu(P) = \sum_{j=0}^n \alpha_j x_j.$$

Verificamos que este funcional é limitado, considerando, para $P \in \mathcal{O}([0,1])$, a seqüência de Polinômios de Bernstein $\{B_n\}_{n \geq 1}$ tal que $B_n \xrightarrow{u} P$, quando $n \to \infty$.

$$|\nu(B_n)| = \left|\sum_{j=0}^{n} \alpha_{n-j} \sum_{k=0}^{n-j} (-1)^{n-k-j} \binom{n}{k} P\left(\frac{k}{n}\right)\right|, \quad \text{pois}$$

$$B_n(t) = \sum_{j=0}^{n} \underbrace{\left(\sum_{k=0}^{n-j} (-1)^{n-j-k} \binom{n}{k} \binom{n-k}{j} P\left(\frac{k}{n}\right)\right)}_{x_{n-j}} t^{n-j}. \text{ Logo,}$$

$$|\nu(B_n)| = \left|\sum_{j=0}^{n} \sum_{k=0}^{n-j} \binom{n}{k} \binom{n-k}{j} P\left(\frac{k}{n}\right) (-1)^{n-j-k} \int_0^1 t^{n-j} \mathrm{d}\psi\right| =$$

$$= \left|\sum_{j=0}^{n} \sum_{k=0}^{n-j} \binom{n}{k} \binom{n-k}{j} P\left(\frac{k}{n}\right) t^k \int_0^1 (-1)^{n-j-k} \mathrm{d}\psi\right| =$$

$$= \left|\sum_{j=0}^{n} \sum_{k=0}^{n-j} \binom{n}{k} \binom{n-k}{j} P\left(\frac{k}{n}\right) t^k \int_0^1 (-1)^{n-k-j} \mathrm{d}\psi\right| =$$

$$= \left|\sum_{k=0}^{n} \binom{n}{k} P\left(\frac{k}{n}\right) t^k \sum_{j=0}^{n-k} \int_0^1 \binom{n-k}{j} (-t)^{n-k-j} \mathrm{d}\psi\right| =$$

$$= \left|\sum_{k=0}^{n} P\left(\frac{k}{n}\right) \binom{n}{k} t^k \int_0^1 \sum_{j=0}^{n-k} \binom{n-k}{j} (-t)^{n-k-j} \mathrm{d}\psi\right| =$$

$$= \left|\sum_{k=0}^{n} P\left(\frac{k}{n}\right) \binom{n}{k} t^k \int_0^1 (1-t)^{n-k} \mathrm{d}\psi\right| = \left|\sum_{k=0}^{n} P\left(\frac{k}{n}\right) \binom{n}{k} \int_0^1 t^k (1-t)^{n-k} \mathrm{d}\psi\right| \leq$$

$$\leq \sum_{k=0}^{n} \left|P\left(\frac{k}{n}\right) \binom{n}{k} \int_0^1 t^k (1-t)^{n-k} \mathrm{d}\psi\right| .$$

Mas, por hipótese, $\binom{n}{k} \int_0^1 t^k (1-t)^{n-k} \mathrm{d}\psi \leq \frac{M}{n+1}$, $\forall k = 0, \ldots, n$. Assim,

$$|\nu(B_n)| \leq M \sum_{k=0}^{n} \frac{|P(\frac{k}{n})|}{n+1} .$$

Da última desigualdade segue que

$$\limsup |\nu(B_n)| \;\leq\; M\|P\|_1 \;, \tag{3.3}$$

pois $\sum_{k=0}^{n} |P(\frac{k}{n})|(n+1)^{-1} \to \int_0^1 |P| \mathrm{d}\lambda_1$, quando $n \to \infty$. Por outro lado, temos que (Shohat e Tamarkin (1943), págs. 8 e 99)

$$|\nu(P)| \;\leq\; |\nu(B_n)| \;+\; |\nu(\varepsilon_n)| \;, \tag{3.4}$$

onde ε_n é um elemento de $\mathcal{O}([0,1])$ tal que $|\nu(\varepsilon_n)| \to 0$, quando $n \to \infty$. De (3.3) e (3.4) e tomando limites quando $n \to \infty$, resulta que

$$|\nu(P)| \;\leq\; M\|P\|_1 \;+\; \lim_{n\to\infty} |\nu(\varepsilon_n)| \;=\; M\|P\|_1 \;.$$

Portanto, ν é um funcional linear contínuo em $\mathcal{O}([0,1])$ e, pelo Teorema de Hahn-Banach para extensão de funcionais lineares, segue que ν em $\mathcal{O}([0,1])$ pode ser estendido a ν^* em $\mathcal{L}_1([0,1], \mathcal{B}[0,1]), \lambda_1)$ de modo a preservar linearidade e limitação, isto é, $\nu^* : \mathcal{L}_1([0,1], \mathcal{B}([0,1]), \lambda_1) \to \mathbb{R}$ é linear e tal que $\nu^*(P) = \nu(P)$, $\forall\, P \in \mathcal{O}([0,1])$ e

$$|\nu^*(f)| \leq M \cdot \|f\|_1, \quad \forall f \in \mathcal{L}_1([0,1], \mathcal{B}([0,1]), \lambda_1).$$

Mas, pelo Teorema da Representação de Riesz, temos que $\exists\, \varphi \in \mathcal{L}_\infty([0,1], \mathcal{B}([0,1]), \lambda_1)$, $\varphi \geq 0$, tal que $\forall f \in \mathcal{L}_1([0,1], \mathcal{B}([0,1]), \lambda_1)$,

$$\nu^*(f) = \int_{[0,1]} f\varphi \mathrm{d}\lambda_1,$$

isto é, $\exists\, \varphi$ limitada tal que

$$\nu^*(f) = \int_{[0,1]} f(t)\varphi(t)\mathrm{d}t \;.$$

Tomando, $\forall n \in \mathbb{N}$, $P_n(t) = t^n$, segue que

$$\nu^*(P_n) = \alpha_n = \int_0^1 t^n \mathrm{d}\psi \ .$$

Por outro lado,

$$\nu^*(P_n) = \int_0^1 P_n(t)\varphi(t)\mathrm{d}t = \int_0^1 t^n\varphi(t)\mathrm{d}t,$$

e, conseqüentemente,

$$\int_0^1 t^n \mathrm{d}\psi = \int_0^1 t^n\varphi(t)\mathrm{d}t, \quad \forall n \in \mathbb{N},$$

implicando que os funcionais lineares $\mu_1 : \mathcal{L}_1([0,1], \mathcal{B}([0,1]), \psi) \to \mathbb{R}$ e $\mu_2 : \mathcal{L}_1([0,1], \mathcal{B}([0,1]), \lambda_1) \to \mathbb{R}$, dados, respectivamente, por

$$\mu_1(f) = \int_0^1 f\mathrm{d}\psi \quad \text{e} \quad \mu_2(f) = \int_0^1 f(t)\varphi(t)\mathrm{d}t \ ,$$

coincidem em $\mathcal{O}([0,1])$ e são limitados na norma $\mathcal{L}_1$. Portanto, temos, em particular, que $\mu_1(\mathbb{I}_{[0,t]}) = \mu_2(\mathbb{I}_{[0,t]})$, $\forall t \in [0,1]$, isto é,

$$\psi(t) = \int_0^t \varphi(u)\mathrm{d}u, \quad 0 \leq t \leq 1 \ ,$$

concluindo a demonstração. ■

Enunciamos a seguir o principal resultado deste capítulo.

Teorema 3.3.3 *Seja $\{X_n\}_{n\geq 1}$ um processo permutável com valores em $\{0,1\}^\infty$. Suponha que exista $M > 0$ tal que, $\forall n \in \mathbb{N}$, $\max_{v\in\{0,\ldots,n\}} P\left(\sum_{i=1}^n X_i = v\right) \leq \frac{M}{n+1}$. Então, $\forall \varepsilon > 0$, $\exists m_0 = m_0(\varepsilon) \in \mathbb{N}$ e $\{p_0, \ldots, p_{m_0}\}$ com $p_k \geq 0$, $k = 0, \ldots, m_0$, e $\sum_{k=0}^{m_0} p_k = 1$, tais que $\forall n \in \mathbb{N}$ e $\forall (x_1, \ldots, x_n) \in \{0,1\}^n$,*

$$\left| P(x_1, \ldots, x_n) - \sum_{k=0}^{m_0} p_k \cdot \textit{Pólya}\left((x_1, \ldots, x_n)|(k+1, m_0 - k + 1)\right) \right| < \varepsilon, \tag{3.5}$$

onde P denota a medida do processo $\{X_n\}_{n\geq 1}$.

Prova Como $\{X_n\}_{n\geq 1}$ é permutável assumindo valores em $\{0,1\}^\infty$, segue, pelo Teorema da Representação de De Finetti, que, $\forall n \in \mathbb{N}$, $\forall (x_1,\ldots,x_n) \in \{0,1\}^n$,

$$P(x_1,\ldots,x_n) = \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \mathrm{d}\psi(\theta),$$

onde $\psi : \mathcal{B}([0,1]) \to [0,1]$ é a medida de De Finetti. Em particular, para o vetor $1_n = (1,\ldots,1)_{1\mathrm{X}n}$, temos

$$\underbrace{P(1_n)}_{\alpha_n} = \int_0^1 \theta^n \mathrm{d}\psi(\theta), \quad \forall n \in \mathbb{N}.$$

Como por hipótese existe $M > 0$ tal que

$$\max_{v\in\{0,\ldots,n\}} P\left(\sum_{i=1}^n X_i = v\right) = \max_{v\in\{0,\ldots,n\}} \binom{n}{v} \int_0^1 \theta^v (1-\theta)^{n-v} \mathrm{d}\psi \leq \frac{M}{n+1}, \forall n \in \mathbb{N},$$

segue, pelo Teorema 3.3.2, que ψ é absolutamente contínua e que $\exists\ \varphi : [0,1] \to \mathbb{R}_+$, limitada, tal que

$$\psi(t) = \varphi(t)\mathrm{d}t\ .$$

Neste caso, existe $\{\varphi_n\}_{n\geq 1}$, $\varphi_n \geq 0$, simples, $\forall n \in \mathbb{N}$, tal que $\varphi_n \xrightarrow{u} \varphi$. Assim, $\forall \varepsilon > 0$, $\exists\ \varphi_{n_0} \geq 0$, simples, tal que

$$\int_0^1 |\varphi - \varphi_{n_0}| \mathrm{d}\lambda_1 \ < \ \frac{\varepsilon}{4}\ . \tag{3.6}$$

Além disso, existem uma "step function" $S : [0,1] \to \mathbb{R}_+$ tal que

$$\int_0^1 |\varphi_{n_0} - S| \mathrm{d}\lambda_1 \ < \ \frac{\varepsilon}{4}\ , \tag{3.7}$$

uma função contínua limitada $C : [0,1] \to \mathbb{R}_+$ tal que

$$\int_0^1 |S - C| \mathrm{d}\lambda_1 \ < \ \frac{\varepsilon}{4} \tag{3.8}$$

e uma sequência de polinômios de Bernstein $\{B_m\}_{m\geq 1}$ tal que $B_m \xrightarrow{u} C$. Avaliemos, então,

$$\left| P(x_1,\ldots,x_n) - \sum_{k=0}^{m} p_k^{(m)} \cdot \text{Pólya}\left((x_1,\ldots,x_n)|(k+1,m-k+1)\right) \right|,$$

com $p_k^{(m)} = \dfrac{C(\frac{k}{m})}{\sum_{j=0}^{m} C(\frac{j}{m})}$, $k = 0,\ldots,m$. Temos

$$\left| P(x_1,\ldots,x_n) - \sum_{k=0}^{m} p_k^{(m)} \cdot \text{Pólya}\left((x_1,\ldots,x_n)|(k+1,m-k+1)\right) \right| =$$

$$= \left| \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\varphi(\theta)\mathrm{d}\theta \;\; - \right.$$

$$\left. - \sum_{k=0}^{m} p_k^{(m)} \cdot \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \text{Beta}\left(\theta;(k+1,m-k+1)\right)\mathrm{d}\theta \right|,$$

onde a última igualdade segue do Teorema da Representação de De Finetti e (3.4). Pelo Teorema de Tonelli, segue que

$$= \left| P(x_1,\ldots,x_n) - \sum_{k=0}^{m} p_k^{(m)} \cdot \text{Pólya}\left((x_1,\ldots,x_n);(k+1,m-k+1)\right) \right| =$$

$$= \left| \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\varphi(\theta)\mathrm{d}\theta \;\; - \right.$$

$$\left. - \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \left(\sum_{k=0}^{m} p_k^{(m)} \cdot \text{Beta}\left(\theta;(k+1,m-k+1)\right) \right) \mathrm{d}\theta \right| =$$

$$= \left| \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \left[\varphi(\theta) - \sum_{k=0}^{m} \frac{C(\frac{k}{m})}{\sum_{j=0}^{m} C(\frac{j}{m})} \frac{\Gamma(m+2)}{\Gamma(k+1)\Gamma(m-k+1)} \theta^k (1-\theta)^{m-k} \right] \mathrm{d}\theta \right|$$

$$= \left| \int_0^1 \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \left[\varphi(\theta) - \frac{1}{\sum_{j=0}^m \frac{C(\frac{j}{m})}{m+1}} \sum_{k=0}^m \binom{m}{k} C\left(\frac{k}{m}\right) \theta^k (1-\theta)^{m-k} \right] \mathrm{d}\theta \right| =$$

$$= \left| \int_0^1 \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \left[\varphi(\theta) - \frac{P_m(\theta)}{b_m} \right] \mathrm{d}\theta \right| ,$$

onde $P_m(\theta) = \sum_{k=0}^m \binom{m}{k} C\left(\frac{k}{m}\right) \theta^k (1-\theta)^{m-k}$ e $b_m = \dfrac{\sum_{j=0}^m C(\frac{j}{m})}{m+1}$. Por fim,

$$\left| P(x_1, \ldots, x_n) - \sum_{k=0}^m p_k^{(m)} \cdot \text{Pólya}\left((x_1, \ldots, x_n) | (k+1, m-k+1)\right) \right| =$$

$$= \left| \int_0^1 \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \left[\varphi(\theta) - \frac{P_m(\theta)}{b_m} \right] \mathrm{d}\theta \right| \leq$$

$$\leq \int_0^1 \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \left| \varphi(\theta) - \frac{P_m(\theta)}{b_m} \right| \mathrm{d}\theta .$$

Mas, como $b_m \to 1$, quando $m \to \infty$, temos que

$$\frac{P_m(\theta)}{b_m} \xrightarrow{u} C(\theta), \quad \text{quando } m \to \infty$$

e, portanto, $\forall \varepsilon > 0$, $\exists m_0 = m_0(\varepsilon) \in \mathbb{N}$ tal que se $m \geq m_0$, então

$$\int_0^1 \left| \frac{P_m(\theta)}{b_m} - C(\theta) \right| \mathrm{d}\lambda_1 < \frac{\varepsilon}{4} \tag{3.9}$$

Assim, temos, finalmente, que

$$\left| P(x_1, \ldots, x_n) - \sum_{k=0}^{m_0} p_k^{(m_0)} \cdot \text{Pólya}\left((x_1, \ldots, x_n) | (k+1, m-k+1)\right) \right| \leq$$

$$\leq \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\left|\varphi(\theta) - \frac{P_m(\theta)}{b_m}\right| \mathrm{d}\theta \leq$$

$$\leq \int_0^1 \left|\varphi(\theta) - \frac{P_m(\theta)}{b_m}\right| \mathrm{d}\theta \leq$$

$$\leq \int_0^1 |\varphi - \varphi_{n_0}|\,\mathrm{d}\lambda_1 + \int_0^1 |\varphi_{n_0} - S|\,\mathrm{d}\lambda_1 + \int_0^1 |S - C|\,\mathrm{d}\lambda_1 + \int_0^1 \left|C - \frac{P_{m_0}}{b_{m_0}}\right| \mathrm{d}\lambda_1 \Rightarrow$$

$$\Rightarrow \left|P(x_1,\ldots,x_n) - \sum_{k=0}^{m_0} p_k^{(m_0)} \cdot \text{Pólya}((x_1,\ldots,x_n)|(k+1, m-k+1))\right| < \varepsilon\ ,$$

onde a última desigualdade segue de (3.6), (3.7), (3.8) e (3.9). Portanto, $\forall \varepsilon > 0$, $\exists m_0 = m_0(\varepsilon) \in \mathbb{N}$ e $\{p_0, \ldots, p_{m_0}\}$ com $p_k \geq 0$, $k = 0, \ldots, m_0$, e $\sum_{k=0}^{m_0} p_k = 1$, tais que $\forall n \in \mathbb{N}$ e $\forall (x_1, \ldots, x_n) \in \{0,1\}^n$,

$$\left|P(x_1,\ldots,x_n) - \sum_{k=0}^{m_0} p_k \cdot \text{Pólya}((x_1,\ldots,x_n)|(k+1, m_0-k+1))\right| < \varepsilon\ ,$$

concluindo a demonstração. ■

O Teorema 3.3.3 estabelece que qualquer processo permutável com valores em $\{0,1\}^\infty$ satisfazendo a condição, $\forall n \in \mathbb{N}$, $\max_{v\in\{0,\ldots,n\}} P\left(\sum_{i=1}^n X_i = v\right) \leq \frac{M}{n+1}$, para algum $M > 0$, pode ser bem aproximado por um adequado modelo de urna de Pólya-Eggenberger com configuração inicial aleatória. Mais precisamente, a probabilidade de qualquer evento relativo a uma marginal n-dimensional deste processo, $n \in \mathbb{N}$, pode ser avaliada, com erro máximo $\varepsilon > 0$ fixado, pela probabilidade do correspondente evento num modelo de urna de Pólya-Eggenberger com composição inicial aleatória contendo, inicialmente, k (k desconhecido) bolas brancas e m_0+2-k bolas pretas com probabilidade $p_{k-1}^{(m_0)}$ especificada no teorema, $k = 1, \ldots, m_0 + 1$.

A condição citada acima indica que tal aproximação é possível sempre que a distribuição de $\sum_{i=1}^{n} X_i$ não apresente, em seu suporte, pontos com probabilidade de ordem superior a $\frac{1}{n}$, isto é, sempre que a sequência $\{(n+1)P\left(\sum_{i=1}^{n} X_i = k\right)\}_{n\geq 1}$ seja convergente (finita), $\forall k \in \mathbb{N}$. Esta condição é suficiente para que a medida de De Finetti do processo $\{X_n\}_{n\geq 1}$ sob consideração seja absolutamente contínua possuindo função densidade de probabilidade limitada (na verdade, é também condição necessária). Deste modo, o resultado do Teorema 3.3.3 não se aplica, por exemplo, a um processo estocástico permutável em $\{0,1\}^{\infty}$ cuja medida de De Finetti possui função densidade de probabilidade Beta (1/2,1), que não é limitada (com efeito, temos, neste caso, $\lim_{n\to\infty}(n+1)P\left(\sum_{i=1}^{n} X_i = 0\right) = \infty$).

Ainda nas condições do Teorema 3.3.3, com a hipótese adicional da medida de De Finetti do processo possuir função densidade de probabilidade não apenas limitada, mas também contínua em $[0,1]$, obtemos uma majoração bem mais "fina" em (3.5). Neste caso, temos, $\forall n \in \mathbb{N}$ e $\forall (x_1,\ldots,x_n) \in \{0,1\}^n$,

$$\left| P(x_1,\ldots,x_n) - \sum_{k=0}^{m_0} p_k^{(m_0)} \cdot \text{Pólya}\left((x_1,\ldots,x_n)|(k+1, m-k+1)\right) \right| <$$

$$\leq \frac{\varepsilon\, \Gamma\left(\sum_{i=1}^{n} x_i + 1\right) \Gamma\left(n - \sum_{i=1}^{n} x_i + 1\right)}{\Gamma(n+2)} \leq \frac{\varepsilon}{n+1} << \varepsilon \quad \text{e,}$$

além disso, a distribuição do número de "sucessos" ($\{X_i = 1\}$) nas n primeiras etapas do processo, $\sum_{i=1}^{n} X_i$, $n \geq 1$, é bem aproximada por uma mistura de distribuições de Pólya-Eggenberger, isto é,

$$\left| P\left(\sum_{i=1}^{n} X_i = t\right) - \sum_{k=0}^{m_0} p_k \cdot P - E(t|(k+1, m_0 - k + 1)) \right| < \varepsilon,$$

$\forall n \in \mathbb{N}$, $\forall t = 0,1,\ldots,n$, $P - E(\ \cdot\ |(k+1, m_0 - k + 1))$ denotando a distribuição de Pólya-Eggenberger para o número de retiradas de bola branca nas n primeiras etapas

do processo de urna de Pólya-Eggenberger com configuração inicial $(k+1, m_0-k+1)$:

$$P - E(t|(k+1, m_0-k+1)) = \binom{n}{t} \text{Pólya}((x_1, \ldots, x_n)|(k+1, m_0-k+1)),$$

com $(x_1, \ldots, x_n) \in \{0,1\}^n$ tal que $\sum_{i=1}^n x_i = t$.

Devemos salientar que o resultado do Teorema 3.3.3 é essencialmente qualitativo, no sentido em que apenas aponta a existência de um modelo de urna de Pólya-Eggenberger com configuração inicial aleatória conveniente para "aproximar" o processo estocástico permutável original. Assim, neste trabalho, não contemplamos, por exemplo, a questão da possível otimalidade do modelo de urna de Pólya-Eggenberger com configuração inicial aleatória considerado no Teorema 3.3.3 (Diaconis e Ylvisaker (1985) sugerem uma modificação nos coeficientes do Polinômio de Bernstein de modo a obter uma majoração mais "fina" no Teorema 3.3.1).

Destacamos ainda que o Teorema 3.3.3, embora estabeleça condições somente sobre as quantidades observáveis $\{X_n\}_{n\geq 1}$ para que (3.5) se verifique, não corresponde a uma representação preditivista exata para o modelo de urna de Pólya-Eggenberger com configuração inicial aleatória com $c = 1$. Afinal, o resultado em (3.5) fornece apenas aproximações para as distribuições marginais do processo $\{X_n\}_{n\geq 1}$, não garantindo a existência de uma medida $\mu^* : \mathcal{P}(\mathbb{N}^2) \to [0,1]$ tal que, $\forall n \in \mathbb{N}$ e $\forall (x_1, \ldots, x_n) \in \{0,1\}^n$,

$$P(x_1, \ldots, x_n) = \int_{\mathbb{N}^2} \text{Pólya}((x_1, \ldots, x_n)|(a,b)) \mathrm{d}\mu^*(a,b) .$$

De fato, como a sequência de medidas em $(\mathbb{N}^2, \mathcal{P}(\mathbb{N}^2))$ $\{\{((k+1, m-k+1), p_k^m)\}_{k=0}^m\}_{m\geq 1}$ não é "tight" (pois, seus respectivos suportes são $\{(a,b) \in \mathbb{N}^2 : a+b = m+2\}$), não podemos assegurar a existência de tal medida μ^*. Ainda assim, a hipótese do teorema 3.3.3 é totalmente preditivista, uma vez que leva em conta somente condições sobre as quantidades observáveis $X_1, X_2, \ldots$.

Devemos salientar ainda que o resultado do Teorema 3.3.3 possui forte apelo intuitivo, uma vez que envolve a incerteza que possuímos a respeito da composição inicial de uma urna contendo bolas de duas cores; os modelos de urnas são objetos de concepção bem mais simples, se não mais natural, que limites de freqüências relativas ou probabilidades de cara em um lançamento de uma moeda (usuais interpretações para a quantidade aleatória θ que figura no Teorema da Representação de De Finetti), além de serem passíveis de experimentação e de possuirem grande flexibilidade para modelagem probabilística (Freudenthal (1960) afirma que "Lançamentos de moedas ou dados ou jogos de cartas não são suficientemente flexíveis. O instrumento de aleatoriedade mais geral é a urna preenchida com bolas de diferentes cores").

3.4 Conclusões

Neste capítulo, apresentamos uma nova versão para o modelo de Pólya-Eggenberger, introduzindo uma distribuição de probabilidade para a composição inicial da urna. Este novo processo, diferentemente de outras variações do modelo de Pólya-Eggenberger, é permutável, com medida de De Finetti possuindo função densidade de probabilidade igual a uma mistura de funções densidade Betas. Exibimos, também, em nosso principal resultado, uma condição suficiente e preditivista para que um processo permutável a valores em $\{0,1\}^\infty$ possa ser bem aproximado por um conveniente modelo de urna de Pólya-Eggenberger com configuração inicial aleatória. Além disso, mostramos que o fortalecimento desta condição conduz a uma aproximação bem melhor, de ordem $\frac{1}{n}$. Devemos destacar que o resultado do teorema 3.3.3 possui forte apelo intuitivo, uma vez que envolve a incerteza que possuímos a respeito da composição inicial de uma urna contendo bolas de duas cores, ao invés de objetos de concepção mais sofisticada como limites de frequencias relativas ou probabilidades de cara em um lançamento de uma moeda.

Algumas questões, no entanto, não foram contempladas neste trabalho e, indubitavelmente, constituem objeto de futura investigação. Dentre outras, destacamos a obtenção de condições necessárias para a validade do teorema 3.3.3 (do qual foi caracterizada apenas uma condição suficiente), a possível derivação de uma representação preditivista exata para o modelo de urna de Pólya-Eggenberger, o que corresponde à especificação de condições adicionais sobre um processo permutável a valores em $\{0,1\}^\infty$ de modo a restringir a sua medida de De Finetti à classe das medidas em $([0,1], \mathcal{B}([0,1]))$ com função densidade igual a mistura de Betas (na verdade, desconhecemos a existência de uma representação preditivista neste caso), a introdução de uma medida de probabilidade conjunta para a configuração inicial da urna (A,B) e para o número de bolas repostas à urna em cada etapa (C) e, também, a determinação de condições para que as distribuições preditivas de um processo estocástico em $\{0,1\}^\infty$ possam ser bem aproximadas por respectivas distribuições preditivas de um processo de urna de Pólya-Eggenberger com configuração inicial aleatória (Dalal e Hall (1983) demonstram resultados desta natureza para famílias de distribuições conjugadas em Inferência Bayesiana).

Referências

[1] Dalal, S.R.; Hall, W.J. (1983). *Approximating Priors by Mixtures of Natural Conjugate Priors.* Journal of Royal Statistical Society, 45(B), 2, 278–286.

[2] de Finetti, B. (1937). *Le prévision: ses lois logiques, ses sources subjectives.* Ann. Inst. H. Poincaré. Prob. Statist., 7, 1–68.

[3] Diaconis, P.; Ylvisaker, D. (1985). *Quantifying prior opinion.* Em Bayesian Statistics, 2, Elsevier Science Publishers B. V. and Valencia University Press.

[4] Feller, W. (1957). *An Introduction to Probability Theory and its Applications.* Vol. I. Wiley, New York.

[5] Freudenthal, H. (1960). *Models in applied probability.* Synthèse, 12, 202–212.

[6] Heitele, D. (1975). *An epistemological view on fundamental stochastic ideas.* Educational Studies in Mathematics, Vol. 6, 187–205. Reidel, Dordrecht, Netherlands.

[7] Johnson, N.L.; Kotz, S. (1977). *Urn Models and Their Applications.* Wiley, New York.

[8] Markov, A.A. (1906). *Extension of the law of large numbers to dependent variables.* Izv. Fiz-Mat. Obschch. Kazan Univ., Ser. 2, 15(4), 135–256. (in Russian)

[9] Marsden, J.E. (1974). *Elementary Classical Analysis.* W. H. Freeman and Company, San Francisco.

[10] Pólya, G.; Eggenberger, F. (1923). *Uber die Statistik verketteter Vorgange.* Z. Angew. Math. Mech., 3, 279–289.

[11] Pólya, G. (1954). *Patterns of Plausible Inference.* Princeton University Press, Princeton, N.J..

[12] Shohat, J.A.; Tamarkin, J.D. (1943). *The problem of moments.* American Mathematical Society, New York.

[13] Stieltjes, T.J. (1894 - 1895). *Recherches sur les fractions continues.* Annales de la Faculté des Sciences de Toulouse, (1) 8 (1894), T 1–122, (1) 9 (1895), A 5–47.

Conclusão

Neste trabalho, abordamos três problemas de caracterização de sequências (finitas e infinitas) de variáveis aleatórias (ou vetores) permutáveis. No capítulo 1, apresentamos representações preditivistas para modelos uniformes bivariados dependendo de um único parâmetro. Mais precisamente, consideramos os modelos uniformes sobre $\{(x,y) \in \mathbb{N}^2 : x+y \leq \theta\}$, $\theta \in \mathbb{N}$, $\{(x,y) \in \mathbb{R}_+^2 : x+y \leq \theta\}$, $\theta > 0$, e $\{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq \theta^2\}$, $\theta > 0$. Lembramos que a representação preditivista de um modelo paramétrico, uma caracterização deste modelo baseada exclusivamente em julgamentos probabilísticos sobre quantidades observáveis, possui um caráter fundamental para a escola Bayesiana preditivista de inferência estatística, pois, entre outras razões, fornece uma justificativa subjetivista para a adoção de tal modelo paramétrico como distribuição amostral (condicional) dos dados gerados em experimentos aleatórios. Verificamos também que os resultados obtidos podem ser estendidos, de maneira bem simples, para os correspondentes modelos d-variados, $d > 2$. Finalmente, apresentamos alguns problemas decorrentes da introdução de novo(s) parâmetro(s) nestes modelos para a obtenção de suas respectivas representações preditivistas. Em geral, a introdução de novos parâmetros torna inviável o emprego da técnica utilizada neste trabalho baseada em suficiência minimal.

No capítulo 2, identificamos modelos uniformes conhecidos, discretos e contínuos, dentro de específicas classes de distribuições simétricas multivariadas. No caso discreto, a especificação de distribuições de probabilidade multivariadas que dependem apenas do máximo das coordenadas ou do mínimo e do máximo das coordenadas ou da máxima soma das coordenadas em cada componente do vetor, e a suposição adicional de independência fornecem uma caracterização dos modelos uniformes em $\{0,\ldots,\theta\}$, $\theta \in \mathbb{N}$, em $\{\theta_1,\ldots,\theta_2\}$, $\theta_1,\theta_2 \in Z$, $\theta_1 < \theta_2$, e em $\{(x,y) \in \mathbb{N}^2 : x+y \leq \theta\}$, $\theta \in \mathbb{N}$, respectivamente. No caso contínuo, a imposição de condições similares sobre as funções de densidade de probabilidade de distribuições simétricas multivariadas, além de independência, permite uma caracterização dos modelos uniformes em $(-\theta,\theta)$, $\theta > 0$, em (θ_1,θ_2), $\theta_1,\theta_2 \in \mathbb{R}$,

$\theta_1 < \theta_2$, e em $\{(x,y) \in \mathbb{R}_+^2 : x + y \leq \theta\}$, $\theta > 0$. Outras caracterizações das distribuições simétricas multivariadas consideradas neste trabalho não foram apreciadas e constituem objeto de futura investigação.

Finalmente, no capítulo 3, apresentamos uma nova versão para o modelo de Pólya-Eggenberger, introduzindo uma distribuição de probabilidade para a composição inicial da urna. Este novo processo, diferentemente de outras variações do modelo de Pólya-Eggenberger, é permutável, com medida de De Finetti possuindo função densidade de probabilidade igual a uma mistura de funções densidade Betas. Exibimos, também, em nosso principal resultado, uma condição suficiente para que um processo permutável com valores em $\{0,1\}^\infty$ possa ser bem aproximado por um conveniente modelo de urna de Pólya-Eggenberger com configuração inicial aleatória. Algumas questões de interesse, no entanto, não foram contempladas neste trabalho. Dentre outras, destacamos a obtenção de condições necessárias para a validade do resultado principal, a derivação de um teorema do tipo de De Finetti para o modelo de urna de Pólya-Eggenberger e a introdução de uma medida de probabilidade conjunta para a configuração inicial da urna (A,B) e para o número de bolas repostas à urna em cada etapa (C).