

# Distribuição Marshall-Olkin Normal e Aplicações

Gabriela Oliveira

Orientador: Prof. Fredy Walther Castellares Cáceres

Coorientador: Prof<sup>a</sup>. Lourdes Coral Contreras Montenegro

Área de Concentração: Estatística Matemática

Dissertação apresentada para a obtenção do grau de mestre  
em estatística pela Universidade Federal de Minas Gerais

Belo Horizonte, fevereiro de 2014

# Distribuição Marshall-Olkin Normal e Aplicações

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 27/02/2014. Uma cópia da versão original está disponível no Departamento de Estatística da Universidade Federal de Minas Gerais.

Comissão Julgadora:

- Prof. Fredy Walther Castellares Cáceres (orientador) - UFMG
- Prof<sup>a</sup>. Lourdes Coral Contreras Montenegro (coorientadora)- UFMG
- Prof. Marcos Antonio da Cunha Santos - UFMG
- Prof. Artur José Lemonte - UFPE

# Agradecimentos

Ao longo desses sete anos que se passaram desde a graduação na matemática até a conclusão dessa dissertação, diversas pessoas me guiaram direta e indiretamente nessa jornada pela UFMG.

Agradeço a toda a minha família, em especial à tia Lilian, minha maior influência pela escolha da matemática.

Aos meus amigos e aos colegas da matemática, pelas conversas, cafés, risadas e pelos brindes à passagem. Em especial, quero agradecer de coração ao meu namorado Bruno e às minhas queridas amigas Denise, Giovanna, Lorena, Luciana e Marina por estarem sempre comigo, na alegria e na tristeza, escutando minhas lamentações e é claro por toda amizade e companheirismo. Com vocês tudo foi mais feliz!

Ao meu orientador Fredy Walther Castellares Cáceres e à minha coorientadora Lourdes Coral Contreras Montenegro, por confiarem no meu trabalho e pelo aprendizado ao longo do mestrado.

Aos membros da banca examinadora, professores Marcos Antonio da Cunha Santos (UFMG) e Artur José Lemonte (UFPE), pela participação e sugestões.

Aos professores do departamento de matemática da UFMG, pelos conhecimentos transmitidos, em especial ao professor mais bem humorado da universidade, Fabio Enrique Brochero Martinez, pelas inúmeras dúvidas esclarecidas, conversas e risadas.

Aos professores, colegas e funcionários do departamento de estatística da UFMG, pelo apoio e conhecimentos compartilhados.

Por fim, quero agradecer à FAPEMIG pela bolsa de mestrado.

À todos vocês, meu MUITO OBRIGADA!



# Resumo

OLIVEIRA, G. **Distribuição Marshall-Olkin Normal**. 2014. Dissertação (Mestrado) - Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

Marshall e Olkin em 1997 introduziram uma nova família de distribuições adicionando um novo parâmetro em uma distribuição primitiva já existente. Entretanto, os autores não estudaram as propriedades gerais desta nova família de distribuições. Neste trabalho, foi estudada uma generalização da família normal utilizando o mecanismo de Marshall-Olkin, denominada de distribuição Marshall-Olkin normal (MON). Para esta distribuição obteve-se as principais propriedades matemáticas, como expansões gerais para as funções de densidade de probabilidade e de distribuição acumulada. Expressões dos estimadores de máxima verossimilhança dos parâmetros, matriz de informação esperada e identificabilidade do modelo foram apresentadas. A utilidade dessa distribuição é ilustrada através de um conjunto de dados, mostrando que a distribuição MON é mais flexível do que outras distribuições existentes na literatura como as distribuições logística, t-Student e Cauchy.



# Abstract

OLIVEIRA, G. **Marshall-Olkin Normal Distribution**. 2014. Dissertação (Mestrado) - Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

Marshall and Olkin in 1997 introduced a new family of distributions by adding a new parameter in a primitive distribution  $G(x)$ . However, the authors did not study the general properties of this new family of distributions. In this work, a generalization of the normal family using the mechanism of Marshall-Olkin called normal Marshall-Olkin (MON) distribution was studied. For this distribution we obtained the mathematical properties, such as general expansions for the probability density functions and cumulative distribution, expressions of the maximum likelihood estimators of the parameters, the expected information matrix and identifiability. The usefulness of this distribution is illustrated by a set of data showing that the MON distribution is more flexible than other existing distributions in the literature as the logistic, Student  $t$  and Cauchy distributions.





# Sumário

<b>Lista de Figuras</b>	<b>ix</b>
<b>Lista de Tabelas</b>	<b>xi</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Distribuição Marshall-Olkin</b>	<b>3</b>
2.1 Definição . . . . .	3
2.2 Função densidade de probabilidade . . . . .	4
2.3 Expansões . . . . .	4
2.4 Sobrevivência e risco . . . . .	10
2.5 Momentos . . . . .	11
2.6 Função Quantílica . . . . .	12
<b>3 Distribuição Marshall-Olkin Normal</b>	<b>13</b>
3.1 Funções de distribuição acumulada e de densidade de probabilidade . . . . .	13
3.2 Função de sobrevivência e função de risco . . . . .	15
3.3 Propriedades da distribuição MON . . . . .	17
3.3.1 Moda . . . . .	17
3.3.2 Mediana . . . . .	17
3.4 Momentos . . . . .	19
3.5 Variância . . . . .	29
3.6 Assimetria . . . . .	30
3.7 Curtose . . . . .	31
3.8 Entropia . . . . .	33
3.9 Identificabilidade . . . . .	35
<b>4 Inferência</b>	<b>41</b>
4.1 Estimação . . . . .	41

4.1.1	Estimação pontual . . . . .	41
4.1.2	Algoritmo de Newton-Raphson . . . . .	42
4.1.3	Função escore e informação . . . . .	44
4.1.4	Estimação intervalar . . . . .	45
4.2	Testes de hipóteses . . . . .	47
4.2.1	Teste da razão verossimilhança . . . . .	47
4.2.2	Teste da razão de verossimilhança generalizada . . . . .	48
4.2.3	Teste bondade do ajuste . . . . .	48
4.3	Critérios de seleção de modelos . . . . .	50
4.3.1	Critério de informação de Akaike - AIC . . . . .	50
4.3.2	Critério de informação de Akaike corrigido - AICc . . . . .	50
4.3.3	Critério Bayesiano de Schwarz - BIC . . . . .	50
<b>5</b>	<b>Simulação e Aplicação</b>	<b>51</b>
5.1	Simulação . . . . .	51
5.2	Aplicação . . . . .	52
<b>6</b>	<b>Conclusões e discussões</b>	<b>57</b>
<b>A</b>		<b>63</b>
<b>Apêndice</b>		<b>63</b>
A.1	Terceiro momento . . . . .	63
A.2	Quarto momento . . . . .	65
<b>B</b>		<b>69</b>
<b>Apêndice</b>		<b>69</b>
<b>C</b>		<b>71</b>
<b>Apêndice</b>		<b>71</b>

# Lista de Figuras

3.1	Gráfico da função de distribuição acumulada da distribuição MON para alguns valores de $p$ , $\mu$ e $\sigma$ . . . . .	14
3.2	Gráfico da função densidade de probabilidade da distribuição MON para alguns valores de $p$ , $\mu$ e $\sigma$ . . . . .	15
3.3	Gráfico da função de risco da distribuição MON para alguns valores de $p$ , $\mu$ e $\sigma$ . . . . .	16
3.4	Derivada segunda da função de distribuição acumulada de uma variável aleatória com distribuição MON para alguns valores dos parâmetros. . . . .	18
3.5	Gráfico da moda e mediana de $X \sim MON(p, 0, 1)$ para diferentes valores de $p$ . . . . .	19
3.6	Variância de uma variável aleatória com distribuição MON padrão para diferentes valores de $p$ . . . . .	30
3.7	Medidas de assimetria da distribuição MON: (a) terceiro momento padronizado; (b) medida de assimetria de Pearson; (c) medida de assimetria AG . . . . .	32
3.8	Curtose da distribuição MON padrão para determinados valores de $p$ . . . . .	32
4.1	Superfície da função de verossimilhança fixando $p = 0, 5$ e $p = 2$ . . . . .	43
4.2	Superfície da função de verossimilhança fixando $\mu = -2$ e $\mu = 2$ . . . . .	43
4.3	Superfície da função de verossimilhança fixando $\sigma = 2$ . . . . .	43
5.1	Gráficos da função densidade exata da distribuição MON com histogramas para dados simulados. . . . .	52
5.2	Histograma do conjunto de dados de fibras de carbono com as funções de densidades de probabilidades dos modelos MON, normal, logística, t-Student e Cauchy. . . . .	54
C.1	Histograma dos dados de fibras de carbono com as funções de densidades de probabilidades substituindo-se as EMVs dos modelos MON, normal, skew-normal e normal-potência. . . . .	72



# Lista de Tabelas

3.1	Primeiro momento da distribuição MON padrão para alguns valores de $p$ . . . . .	26
5.1	EMVs dos parâmetros dos modelos MON, normal, logística, t-Student e Cauchy para dados de fibras de carbono, correspondentes erros-padrão (entre parênteses), log-verossimilhança e valores dos AIC, BIC, AICc. . . . .	53
5.2	Teste bondade do ajuste . . . . .	55
5.3	Teste de Kolmogorov-Smirnov para as distribuições testadas . . . . .	55
C.1	EMVs dos parâmetros dos modelos MON, normal, NP, SN e SNB para dados de IQ Scores OITS, correspondentes erros-padrão (entre parênteses) e valores dos AIC, BIC e CAIC. . . . .	71
C.2	Teste de Hipóteses . . . . .	71
C.3	Teste de Kolmogorov-Smirnov para as distribuições testadas. . . . .	72



# Capítulo 1

## Introdução

Modelos probabilísticos utilizando distribuições generalizadas de probabilidade surgiram na literatura pela primeira vez em 1895 com Karl Pearson quando ele propôs uma família de distribuições a qual incluía diversas distribuições de probabilidade. No século passado, após a formalização das técnicas da estatística matemática foram definidas e estudadas várias famílias de distribuições de probabilidade. Entre as principais referências podem ser consultados Ord (1972), Mardia (1972, 2000), Patil (1975, 1981), Johnson et al. (1994, 2000, 2005). Atualmente tem sido publicados artigos em revistas especializadas propondo famílias de distribuições de probabilidade generalizadas e que ainda não foram publicadas em livros texto ou manuais de distribuições. Este contínuo interesse em definir novas distribuições de probabilidade mostra que este é um tema atual e que o estudo sistemático das propriedades e inferência destas famílias merece atenção.

Marshall e Olkin (1997) propuseram um método para adicionar um novo parâmetro com valor positivo em uma distribuição primitiva  $G(x)$  já existente cuja distribuição resultante será denominada de distribuição Marshall-Olkin generalizada (MO). Alguns casos particulares desta distribuição têm sido discutidos na literatura tomando  $G(x)$  como as distribuições Pareto (Alice et al. 2003-2004, Ghitany 2005), Weibull (Jose et al. 2001, Alice et al. 2005, Ghitany et al. 2005, Jose et al. 2009-2010), gamma (Ristić et al. 2007), Lomax (Ghitany et al. 2007), exponencial (Alice et al. 2004) e a distribuição de taxa de falha linear (Ghitany e Kotz 2007). Outras extensões e generalizações da família de distribuições Marshall-Olkin foram apresentadas por Lam e Leung (2001), Economou e Caroni (2007), Gupta e Peng (2009), Gómez- Déniz (2010), Caroni (2010), Gupta et al. (2010), Nanda e Das (2012). Mais recentemente Barreto-Souza et al. (2013) derivam propriedades gerais desta distribuição.

O objetivo do presente trabalho é estudar a distribuição Marshall-Olkin normal (denotada por MON), a qual é definida aplicando-se a transformação Marshall-Olkin com  $G(x)$  sendo a distribuição normal. Esta distribuição foi proposta inicialmente por García et al. (2010) onde eles investigaram o comportamento do parâmetro adicional como um parâmetro de assimetria. García et al. (2010) afirmaram que a distribuição MON compete com distribuições assimétricas tais como skew-normal e skew-normal Balakrishnan. Posteriormente, Maiti e Dey (2012) propuseram a mesma distribuição que García et al. (2010) (curiosamente sem citá-los) e a chamaram de distribuição *tilted* normal com o objetivo de modelar dados de sobrevivência. No mesmo ano, Rubio e Stell (2012) analisaram o uso da transformação Marshall-Olkin como um mecanismo de introduzir assimetria. Estes autores mostraram que se esse mecanismo é aplicado em distri-

buições primitivas simétricas a distribuição resultante não é flexível para modelar dados com assimetria moderada ou alta, contradizendo assim García et al. (2010) e Maiti e Dey (2012).

A principal motivação para escrever esta dissertação deve-se as afirmações contraditórias sobre o parâmetro adicional da nova distribuição Marshall-Olkin normal como mecanismo de produzir assimetria. Uma outra motivação foi encontrar de forma rigorosa algumas propriedades matemáticas da nova distribuição incluindo a inferência e a identificabilidade do modelo que não foram apresentadas por García et al. (2010), Maiti e Dey (2012) e Rubio e Stell (2012).

Neste trabalho, algumas contribuições para a distribuição Marshall-Olkin normal foram feitas. A distribuição MON foi caracterizada como uma mistura de máximos e mínimos aleatórios em que os coeficientes dessa mistura são as probabilidades de uma variável aleatória com distribuição geométrica. Os quatro primeiros momentos da distribuição MON foram representados como combinações infinitas dos momentos do máximo e mínimo da distribuição normal. Expressões para as entropias de Rényi e de Shannon foram apresentadas. Provou-se também que o modelo MON é identificável. Como ilustração, foram feitas simulações e uma análise de um conjunto de dados real que mostrou que o modelo MON compete com as distribuições logística, Cauchy e t-Student. Desta forma, a distribuição em estudo compete com distribuições com caudas pesadas e curtose maior do que a distribuição normal.

O conteúdo da dissertação está organizado em 6 capítulos e 3 apêndices. No Capítulo 2 é apresentado a distribuição Marshall-Olkin generalizada destacando-se as expansões gerais para as funções de distribuição acumulada e de densidade de probabilidade. O Capítulo 3 introduz a distribuição Marshall-Olkin normal exibindo-se suas principais propriedades, entre elas os momentos, entropia e a identificabilidade do modelo. O Capítulo 4 aborda resultados de inferência tais como os estimadores de máxima verossimilhança, matriz de informação esperada, testes de hipóteses e alguns critérios de seleção de modelos. O Capítulo 5 mostra simulações e a análise de um conjunto de dados reais para verificar a flexibilidade do modelo. As conclusões são apresentadas no Capítulo 7. Nos apêndices A e B encontram-se as demonstrações de alguns resultados mostrados neste trabalho. Por fim, no Apêndice C uma análise do banco de dados apresentado no artigo de García et al. (2010) é realizada.



# Capítulo 2

## Distribuição Marshall-Olkin

Existem diversos métodos de introduzir um novo parâmetro para expandir famílias de distribuições para ter mais flexibilidade em um determinado modelo. Este capítulo mostra o método usado por Marshall e Olkin (1997) que acrescenta um novo parâmetro em uma distribuição primitiva  $G$  definindo assim a distribuição Marshall-Olkin (MO). Algumas propriedades dessa nova família de distribuições serão apresentadas baseando-se nos artigos de Marshall e Olkin (1997), Rubio e Stell (2012) e Castellares e Lemonte (2013). Dentre elas destaca-se a caracterização da distribuição Marshall-Olkin como uma mistura de máximos e mínimos de uma quantidade aleatória da distribuição primitiva  $G$ .

### 2.1 Definição

Sejam  $G(x)$  a função de distribuição acumulada (fda) de uma variável aleatória contínua  $X$  e  $S(x) = 1 - G(x)$  sua função de sobrevivência. Adicionando um novo parâmetro  $p > 0$ , Marshall e Olkin (1997) introduziram uma nova família de distribuições generalizadas com função de distribuição acumulada expressa por

$$F(x) = \frac{G(x)}{1 - (1-p)S(x)} = \frac{p^{-1}G(x)}{1 - (1-p^{-1})G(x)}, \quad x \in \mathbb{R}, \quad p > 0. \quad (2.1)$$

A distribuição com função de distribuição acumulada  $G(x)$  será nomeada de distribuição primitiva. A nova família de distribuições Marshall-Olkin, denotada por MO, generaliza a distribuição primitiva  $G$ . De fato, a função de distribuição acumulada da distribuição MO coincide com  $G(x)$  quando  $p = 1$ .

Para expressões futuras será utilizada a reparametrização  $q = p^{-1}$ , assim a função de distribuição acumulada  $F(x)$  pode ser reescrita como

$$F(x) = \frac{qG(x)}{1 - (1-q)G(x)}, \quad x \in \mathbb{R}, \quad q > 0. \quad (2.2)$$

## 2.2 Função densidade de probabilidade

As funções de densidades de probabilidades correspondentes as funções de distribuições acumuladas  $F(x)$  definidas em (2.1) e (2.2) são representadas, respectivamente, por

$$f(x) = \frac{pg(x)}{[1 - (1-p)S(x)]^2}, \quad x \in \mathbb{R}, \quad p > 0, \quad (2.3)$$

$$f(x) = \frac{qg(x)}{[1 - (1-q)G(x)]^2}, \quad x \in \mathbb{R}, \quad q > 0, \quad (2.4)$$

em que  $g(x)$  é a função densidade de probabilidade (fdp) primitiva, isto é,  $g(x) = dG(x)/dx$ .

## 2.3 Expansões para as funções de distribuição acumulada e de densidade de probabilidade

Castellares e Lemonte (2013) propuseram expansões para as funções de distribuição acumulada e de densidade de probabilidade da distribuição MO caracterizando-a como uma mistura de máximos e mínimos aleatórios de uma quantidade aleatória da distribuição primitiva  $G$ . Para essa caracterização eles utilizaram lemas, proposições e teoremas que serão apresentados nesta seção com demonstrações detalhadas.

Seja  $\{X_n, n \in \mathbb{N}\}$  uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d) em que cada variável aleatória  $X_n$  tem como função de distribuição acumulada primitiva  $G$ . Denotando por  $N$  e  $M$  variáveis aleatórias independentes de  $\{X_n, n \in \mathbb{N}\}$ , ambas com distribuições geométricas de parâmetros  $p \in (0, 1]$  e  $q = p^{-1}$  ( $p > 1$ ), respectivamente, tem-se

$$P(N = n) = p(1-p)^{n-1}, \quad P(M = n) = q(1-q)^{n-1},$$

com  $n \in \mathbb{N}$ .

Para a caracterização proposta por eles é necessário definir  $Y_N$  e  $Z_M$  da forma

$$Y_N = \min\{X_1, X_2, \dots, X_N\} \quad \text{e} \quad Z_M = \max\{X_1, X_2, \dots, X_M\}.$$

O Lema 1 mostra uma expressão para a função de distribuição acumulada do mínimo de  $n$  variáveis aleatórias com distribuição geométrica de parâmetro  $p$ , em que  $0 < p \leq 1$ .

**Lema 1.** *A função de distribuição acumulada de  $Y_N$  é dada por*

$$F_N(x) = \frac{G(x)}{1 - (1-p)S(x)}, \quad x \in \mathbb{R}, \quad 0 < p \leq 1.$$

*Demonstração.*

A função de distribuição acumulada de  $Y_N(x)$  é definida por

$$\begin{aligned}
F_N(x) = P(Y_N \leq x) &= 1 - P(Y_N > x) \\
&= 1 - \sum_{n=1}^{\infty} P(N = n)P(Y_N > x|N = n) \\
&= 1 - \sum_{n=1}^{\infty} P(N = n)P(Y_n > x) \\
&= 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1}P(\min\{X_1, X_2, \dots, X_n\} > x) \\
&= 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1}P(X_1 > x, X_2 > x, \dots, X_n > x) \\
&= 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1}P(X_1 > x)P(X_2 > x) \cdots P(X_n > x) \\
&= 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1}P(X_1 > x)^n \\
&= 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1}S(x)^n,
\end{aligned}$$

pois as variáveis aleatórias  $X_i$ 's são independentes e a função de sobrevivência  $S(x)$  é igual a  $1 - G(x)$ . Como  $0 < p \leq 1$  e  $0 \leq S(x) \leq 1$ , a série geométrica converge e, desta forma,

$$\begin{aligned}
F_N(x) &= 1 - pS(x) \sum_{n=1}^{\infty} (1-p)^{n-1}S(x)^{n-1} \\
&= 1 - pS(x) \sum_{n=0}^{\infty} (1-p)^n S(x)^n \\
&= 1 - pS(x) \frac{1}{1 - (1-p)S(x)} \\
&= \frac{1 - S(x)}{1 - (1-p)S(x)} = \frac{G(x)}{1 - (1-p)S(x)}.
\end{aligned}$$

□

Analogamente, o Lema 2 mostra uma expressão para a distribuição acumulada do máximo de  $n$  variáveis aleatórias com distribuição geométrica com parâmetro  $0 < q < 1$ , em que foi utilizada a reparametrização  $q = p^{-1}$  (quando  $p > 1$ ).

**Lema 2.** *A função de distribuição acumulada de  $Z_M$  é dada por*

$$F_M(x) = \frac{qG(x)}{1 - (1-q)G(x)}, \quad x \in \mathbb{R}, \quad 0 < q < 1.$$

*Demonstração.*

A função de distribuição acumulada de  $Z_M(x)$  é definida por

$$\begin{aligned}
 F_M(x) = P(Y_M \leq x) &= \sum_{n=1}^{\infty} P(M = n)P(Y_M \leq x|M = n) \\
 &= \sum_{n=1}^{\infty} P(M = n)P(Y_n \leq x) \\
 &= \sum_{n=1}^{\infty} q(1 - q)^{n-1}P(\max\{X_1, X_2, \dots, X_n\} \leq x) \\
 &= \sum_{n=1}^{\infty} q(1 - q)^{n-1}P(X_1 \leq x)P(X_2 \leq x) \cdots P(X_n \leq x) \\
 &= \sum_{n=1}^{\infty} q(1 - q)^{n-1}P(X_1 \leq x)^n \\
 &= \sum_{n=1}^{\infty} q(1 - q)^{n-1}G(x)^n,
 \end{aligned}$$

pois as variáveis aleatórias  $X_i$ 's são idênticas e independentes. Como  $0 < q < 1$  e  $0 \leq G(x) \leq 1$ , a série geométrica é convergente e, portanto,

$$\begin{aligned}
 F_M(x) &= qG(x) \sum_{n=1}^{\infty} (1 - q)^{n-1}G(x)^{n-1} \\
 &= qG(x) \sum_{n=0}^{\infty} (1 - q)^n G(x)^n \\
 &= qG(x) \frac{1}{1 - (1 - q)G(x)} \\
 &= \frac{qG(x)}{1 - (1 - q)G(x)}.
 \end{aligned}$$

□

Utilizando o Lema 1 e o Lema 2, pode ser definida uma nova função de distribuição acumulada na qual é possível observar uma semelhança com a distribuição MO, como mostra o Lema 3 a seguir.

**Lema 3.** *Sob as definições anteriores, existe a função de distribuição acumulada representada por*

$$F(x; \alpha) = \begin{cases} \frac{G(x)}{1 - (1 - \alpha)S(x)}, & \text{se } 0 < \alpha \leq 1, \\ \frac{\alpha^{-1}G(x)}{1 - (1 - \alpha^{-1})G(x)} = \frac{G(x)}{1 - (1 - \alpha)S(x)}, & \text{se } \alpha > 1. \end{cases} \quad (2.5)$$

*Demonstração.*

Escrevendo  $\alpha = p$  e  $F(x; \alpha) = F_N(x)$  quando  $0 < p \leq 1$ , pelo Lema 1 tem-se

$$F(x; \alpha) = F_N(x) = \frac{G(x)}{1 - (1 - \alpha)S(x)}.$$

Se  $p > 1$  considere a reparametrização  $q = p^{-1}$ . Desta forma, escrevendo  $\alpha = q^{-1} > 1$  e substituindo  $F_M(x)$  por  $F(x; \alpha)$  no Lema 2, a função de distribuição acumulada de  $M$  pode ser expressa como

$$F(x; \alpha) = F_M(x) = \frac{\alpha^{-1}G(x)}{1 - (1 - \alpha^{-1})G(x)} = \frac{G(x)}{1 - (1 - \alpha)S(x)}.$$

□

A proposição a seguir mostra que se uma variável aleatória tem função de distribuição acumulada representada por (2.5) então  $X$  tem distribuição MO.

**Proposição 1.** *Se  $X$  é uma variável aleatória com função de distribuição acumulada da forma (2.5), então,  $X$  segue uma distribuição MO.*

*Demonstração.*

Para  $0 < \alpha \leq 1$  defina  $\alpha = p$  e quando  $\alpha > 1$ ,  $\alpha = q^{-1}$  ( $0 < q < 1$ ). Pelo Lema (3), com essas mudanças de parâmetros obtêm-se exatamente a função de distribuição acumulada definida por Marshall e Olkin (1997). Assim, a função de distribuição acumulada representada em (2.5) pode ser reescrita como

$$F(x; \alpha) = \frac{G(x)}{1 - (1 - \alpha)S(x)}, \quad x \in \mathbb{R}, \quad \alpha > 0.$$

□

Castellares e Lemonte (2013) caracterizaram a distribuição MO como uma mistura de máximos e mínimos aleatórios de uma quantidade aleatória da distribuição primitiva  $G$ , em que os coeficientes dessa mistura são as probabilidades de variáveis aleatórias independentes com distribuição geométrica, como será apresentado a seguir. O Teorema 1 mostra uma expansão em série para a função de distribuição acumulada da distribuição MO, quando o parâmetro  $p$  assume valores entre 0 e 1.

**Teorema 1.** *Se  $X$  tem distribuição MO, então a função de distribuição acumulada pode ser representada como a mistura*

$$F(x) = \sum_{n=1}^{\infty} p(1-p)^{n-1} \{1 - [1 - G(x)]^n\}, \quad x \in \mathbb{R}, \quad 0 < p \leq 1,$$

em que os seus coeficientes são as probabilidades de  $N$ , ou seja,  $X$  é uma mistura de distribuições do mínimo.

*Demonstração.*

Utilizando o Lema 1 e a Proposição 1 tem-se

$$\begin{aligned}
F(x) = F(x, \alpha) &= \frac{G(x)}{1 - (1 - \alpha)S(x)} \\
&= 1 - pS(x) \sum_{n=1}^{\infty} (1 - p)^{n-1} S(x)^{n-1} \\
&= \sum_{n=1}^{\infty} p(1 - p)^{n-1} \{1 - [1 - G(x)]^n\}.
\end{aligned}$$

□

Analogamente quando  $p > 1$ , utilizando a reparametrização  $q = p^{-1}$ , a função de distribuição acumulada da distribuição MO pode ser representada como uma série absolutamente convergente como pode ser constatado no Teorema 2.

**Teorema 2.** *Se  $X$  tem distribuição MO, então a função de distribuição acumulada pode ser representada como a mistura*

$$F(x) = \sum_{n=1}^{\infty} q(1 - q)^{n-1} G(x)^n, \quad x \in \mathbb{R}, \quad 0 < q < 1,$$

em que os seus coeficientes são as probabilidades de  $M$ , ou seja,  $X$  é uma mistura de distribuições do máximo.

*Demonstração.*

Utilizando o Lema 2 e a Proposição 1 tem-se

$$F(x) = F(x; \alpha) = F_M(x) = \sum_{n=1}^{\infty} q(1 - q)^{n-1} G(x)^n.$$

□

O Teorema 1 e o Teorema 2 podem ser resumidos no Corolário 1, desta forma, tem-se uma representação para a função de distribuição acumulada da distribuição MO como expansões em séries infinitas.

**Corolário 1.** *Se  $X$  tem uma distribuição MO, então sua função de distribuição acumulada pode ser representada pelas séries*

$$F(x) = \begin{cases} 1 - \sum_{n=1}^{\infty} p(1 - p)^{n-1} S(x)^n, & \text{se } 0 < p \leq 1, \\ \sum_{n=1}^{\infty} q(1 - q)^{n-1} G(x)^n, & \text{se } q^{-1} = p > 1. \end{cases} \quad (2.6)$$

Derivando a expansão em série da função de distribuição acumulada tem-se uma expansão para a função densidade de probabilidade de uma variável aleatória com distribuição MO, como mostra o Teorema 3.

**Teorema 3.** *Se  $X$  segue uma distribuição MO com função de distribuição acumulada (2.6), então sua função densidade de probabilidade é representada por uma série absolutamente convergente da forma*

$$f(x) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1)g(x)S(x)^n, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1)g(x)G(x)^n, & \text{se } q^{-1} = p > 1. \end{cases} \quad (2.7)$$

*Demonstração.*

Por (2.6) a função de distribuição acumulada de  $X$  é representada por uma série absolutamente convergente, uma vez que a função de sobrevivência  $S(x)$  é limitada (menor ou igual a um) e a série geométrica converge (a série da função de distribuição acumulada é cotada por um número real e desta forma é absolutamente convergente). Para obter a função densidade de probabilidade é suficiente derivar (2.6) com respeito a  $x$ . Para  $0 < p \leq 1$ , fazendo-se uma mudança no índice do somatório, a função de distribuição acumulada é expressa como

$$F(x; p) = 1 - \sum_{n=0}^{\infty} p(1-p)^n S(x)^{n+1}.$$

Calculando a derivada em relação a  $x$  tem-se que

$$f(x; p) = \sum_{n=0}^{\infty} p(1-p)^n(n+1)g(x)S(x)^n.$$

Para  $q^{-1} = p > 1$  a função de distribuição acumulada de  $X$  é representada por

$$F(x; q) = \sum_{n=0}^{\infty} q(1-q)^n G(x)^{n+1}.$$

Derivando tem-se

$$f(x; q) = \sum_{n=0}^{\infty} q(1-q)^n(n+1)g(x)G(x)^n.$$

□

As expansões (2.7) para a função densidade de probabilidade, apresentadas por Castellares e Lemonte (2013), são fundamentais para se obter algumas propriedades matemáticas da família de distribuições MO. No Capítulo 3 essas expansões serão utilizadas para caracterizar a função de distribuição Marshall-Olkin normal como uma mistura de máximos e mínimos aleatórios de uma quantidade aleatória da distribuição normal.

## 2.4 Função de Sobrevivência e função de risco

A função de sobrevivência da família de distribuição MO com função de distribuição acumulada (2.1) é definida como

$$1 - F(x) = \bar{F}(x) = \frac{p[1 - G(x)]}{1 - (1 - p)[1 - G(x)]} = \frac{pS(x)}{1 - (1 - p)S(x)}, \quad x \in \mathbb{R}, \quad p > 0, \quad (2.8)$$

em que  $S(x) = 1 - G(x)$  denota a função de sobrevivência da distribuição primitiva  $G$ .

A função de risco da família de distribuição MO para  $-\infty < x < \infty$  é expressa como

$$r(x; p) = \frac{f(x)}{\bar{F}(x)} = \frac{1}{[1 - (1 - p)S(x)]} \frac{g(x)}{S(x)} = \frac{r_G(x)}{1 - (1 - p)S(x)}, \quad (2.9)$$

em que  $r_G(x) = \frac{g(x)}{S(x)}$  é a função de risco da distribuição primitiva  $G$ .

A função de risco da nova distribuição possui algumas propriedades que foram apresentadas por Marshall e Olkin (1997) que serão dadas a seguir.

Seja  $R(x)$  a razão  $r(x; p)/r_G(x)$ . Calculando a derivada  $R'(x)$  tem-se que

$$R'(x) = \frac{(p - 1)g(x)}{[p + (1 - p)G(x)]^2},$$

em que  $g(x)$  é a função densidade de probabilidade correspondente à função de distribuição acumulada primitiva  $G(x)$ . Portanto, a razão  $r(x; p)/r_G(x)$  é crescente em  $x$  para  $p > 1$  e decrescente quando  $0 < p < 1$ , e para  $p = 1$  é igual a zero.

Os limites infinitos da função de sobrevivência  $S(x)$  são dados por

$$\begin{aligned} \lim_{x \rightarrow -\infty} S(x) &= 1 - \lim_{x \rightarrow -\infty} G(x) = 1 \\ \lim_{x \rightarrow \infty} S(x) &= 1 - \lim_{x \rightarrow \infty} G(x) = 0. \end{aligned}$$

Pelas propriedades de limites quando  $n$  vai para menos infinito, o limite para a função de risco  $r(x; p)$  da distribuição MO é dado por

$$\begin{aligned} \lim_{x \rightarrow -\infty} r(x; p) &= \frac{\lim_{x \rightarrow -\infty} r_G(x)}{1 - (1 - p) \lim_{x \rightarrow -\infty} S(x)} \\ &= \lim_{x \rightarrow -\infty} \frac{r_G(x)}{p}. \end{aligned}$$

Analogamente,

$$\begin{aligned} \lim_{x \rightarrow +\infty} r(x; p) &= \frac{\lim_{x \rightarrow +\infty} r_G(x)}{1 - (1 - p) \lim_{x \rightarrow +\infty} S(x)} \\ &= \lim_{x \rightarrow +\infty} r_G(x). \end{aligned}$$



No artigo de Marshall e Olkin (1997) são apresentadas as seguintes desigualdades que comparam as funções de risco e de sobrevivência da distribuição primitiva  $G$  com as funções de risco e de sobrevivência da distribuição MO.

$$r_G(x)/p \leq r(x;p) \leq r_G(x), \quad (-\infty < x < \infty, p \geq 1), \quad (2.10)$$

$$r_G(x) \leq r(x;p) \leq r_G(x)/p, \quad (-\infty < x < \infty, 0 < p \leq 1). \quad (2.11)$$

## 2.5 Momentos

Utilizando a representação em série da função densidade de probabilidade mostrada na expressão (2.7), tem-se que o  $s$ -ésimo momento da distribuição MO pode ser expresso como

$$E(X^s) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x^s g(x) S(x)^n dx, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^s g(x) G(x)^n dx, & \text{se } q^{-1} = p > 1. \end{cases} \quad (2.12)$$

Rubio e Stell (2012) mostraram que a transformação na função de distribuição primitiva preserva a existência dos momentos, como é apresentado no Teorema 4. A demonstração é completada com detalhes.

**Teorema 4.** *Os momentos da função de distribuição acumulada da distribuição MO existem exatamente para as mesmas ordens que os momentos da distribuição primitiva  $G$ .*

*Demonstração.* Se  $p < 1$  segue que  $p[1 - G(x)] < [1 - G(x)]$  e então  $G(x) + p[1 - G(x)] < 1$ . Como  $G(x) + p[1 - G(x)]$  é maior que zero, pode-se elevar ambos os lados da desigualdade ao quadrado obtendo-se

$$\{G(x) + p[1 - G(x)]\}^2 < 1.$$

Multiplicando a desigualdade anterior por  $p$  e reorganizando os termos, tem-se que

$$p < \frac{p}{\{G(x) + p[1 - G(x)]\}^2}. \quad (2.13)$$

Com efeito,  $G(x) > pG(x)$  e desta forma  $p < p + G(x) - pG(x)$ . Sendo assim,  $p^2 < \{G(x) + p[1 - G(x)]\}^2$  e portanto

$$\frac{p}{\{G(x) + p[1 - G(x)]\}^2} < \frac{1}{p}. \quad (2.14)$$

Reunindo as desigualdades (2.13) e (2.14) tem-se que

$$p < \frac{p}{\{G(x) + p[1 - G(x)]\}^2} < \frac{1}{p}.$$

Por outro lado, quando  $p > 1$  segue que  $G(x)[1 - p] + p < p$ . Elevando ao quadrado os dois membros desta última desigualdade o sinal se mantém e desta forma

$$\frac{1}{p} < \frac{p}{\{G(x) + p[1 - G(x)]\}^2}. \quad (2.15)$$

Como  $p > 1$  a desigualdade  $p[1 - G(x)] > [1 - G(x)]$  é verdadeira. Portanto,  $G(x) + p[1 - G(x)] > 1$  e a desigualdade é mantida quando eleva-se os dois lados da inequação ao quadrado, ou seja,  $\{G(x) + p[1 - G(x)]\}^2 > 1$ . Multiplicando ambos os lados por  $p$  resulta que

$$\frac{p}{\{G(x) + p[1 - G(x)]\}^2} < p. \quad (2.16)$$

Pelas desigualdades (2.15) e (2.16) segue que

$$\frac{1}{p} < \frac{p}{\{G(x) + p[1 - G(x)]\}^2} < p.$$

Relembrando que a função densidade de probabilidade da distribuição MO é representada por

$$f(x) = \frac{pg(x)}{\{G(x) + p[1 - G(x)]\}^2},$$

tem-se que

$$f(x) = K(x)g(x),$$

em que  $K(x)$  toma valores entre  $\min\{p, 1/p\}$  e  $\max\{p, 1/p\}$ . Desta forma, a transformação MO preserva a existência dos momentos em relação a distribuição primitiva  $G$ , ou seja, possui momentos de mesma ordem que os da distribuição  $G$ .  $\square$

## 2.6 Função Quantílica

Castellares e Lemonte (2013) apresentam também a função quantílica para a distribuição MO, representada pelas expressões:

$$x = Q(u) = G^{-1}\left(\frac{up}{1 - (1 - p)u}\right), \quad \text{se } u \in (0, 1), \quad p > 0, \quad (2.17)$$

ou

$$x = Q(u) = G^{-1}\left(\frac{u}{q + (1 - q)u}\right), \quad \text{se } u \in (0, 1), \quad q > 0. \quad (2.18)$$

No próximo capítulo será definida a distribuição MO normal em que os resultados apresentados neste capítulo serão utilizados para encontrar suas principais propriedades tais como fórmulas explícitas para os momentos, entropia e identificabilidade, até então não presentes na literatura.

# Capítulo 3

## Distribuição Marshall-Olkin Normal

Neste capítulo apresenta-se o modelo Marshall-Olkin normal (MON), em que sua função de distribuição acumulada será caracterizada como uma mistura de máximos e mínimos aleatórios, em que os coeficientes dessa mistura são as probabilidades de uma variável aleatória com distribuição geométrica. Destaca-se que esta representação é mais simples para encontrar propriedades analíticas para esta distribuição do que as expansões que foram obtidas por García et al. (2010). Os quatro primeiros momentos da distribuição MON são representados como combinações infinitas dos momentos do máximo e mínimo da distribuição normal e expressões para as entropias de Rényi e de Shannon também são obtidas. Vale a pena ressaltar a prova de identificabilidade do modelo em estudo, o que não é natural encontrar em diversas bibliografias que trabalham com novas distribuições.

### 3.1 Funções de distribuição acumulada e de densidade de probabilidade

Uma variável aleatória  $X$  tem distribuição Marshall-Olkin normal (MON) quando, nas definições do capítulo anterior, tem-se  $G(x) = \Phi(x; \mu, \sigma)$ , ou seja, a função de distribuição acumulada primitiva é a função de distribuição acumulada da distribuição normal. Se  $G(x) = \Phi(x; 0, 1)$ , a distribuição será nomeada de distribuição MON padrão.

A função de distribuição acumulada e a função densidade de probabilidade da distribuição MON são dadas, respectivamente, por

$$F(x; p, \mu, \sigma) = \frac{\Phi\left(\frac{x-\mu}{\sigma}\right)}{\{1 - (1-p)[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)]\}} \quad (3.1)$$

$$f(x; p, \mu, \sigma) = \frac{p\phi\left(\frac{x-\mu}{\sigma}\right)}{\sigma\{1 - (1-p)[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)]\}^2}, \quad (3.2)$$

com  $x \in \mathbb{R}$ ,  $\mu \in \mathbb{R}$ ,  $p > 0$  e  $\sigma > 0$ .

Utilizando as expansões para as funções de distribuição acumulada (2.6) e de densidade de probabilidade (2.7) dadas no Capítulo 2 tem-se para a distribuição MON padrão as repre-

representações:

$$F(x; p, 0, 1) = \begin{cases} 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1}[1 - \Phi(x)]^n, & \text{se } 0 < p \leq 1, \\ \sum_{n=1}^{\infty} q(1-q)^{n-1}\Phi(x)^n, & \text{se } q^{-1} = p > 1, \end{cases} \quad (3.3)$$

$$f(x; p, 0, 1) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1)\phi(x)[1 - \Phi(x)]^n, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1)\phi(x)\Phi(x)^n, & \text{se } q^{-1} = p > 1. \end{cases} \quad (3.4)$$

Essas representações para as funções de distribuição acumulada (fda) e de densidade de probabilidade (fdp) serão utilizadas para o desenvolvimento dos momentos da distribuição MON padrão em seções posteriores.

As Figuras 3.1 e 3.2 ilustram, respectivamente, algumas possíveis formas da fda e da fdp de uma variável aleatória  $X$  com distribuição MON para determinados valores dos parâmetros, incluindo a distribuição normal que é um caso particular da distribuição MON quando  $p = 1$ . Observa-se que o gráfico da função de distribuição acumulada da distribuição MON é semelhante à curva da função de distribuição acumulada da normal, porém, para  $0 < p < 1$ , tem-se um deslocamento para a esquerda, e, para  $p > 1$ , o deslocamento é para direita. Outra característica da distribuição MON é que ela é mais alta (afunilada) e concentrada que a distribuição normal, ou seja, uma simulação da distribuição MON pode gerar valores maiores que uma simulação da distribuição normal.

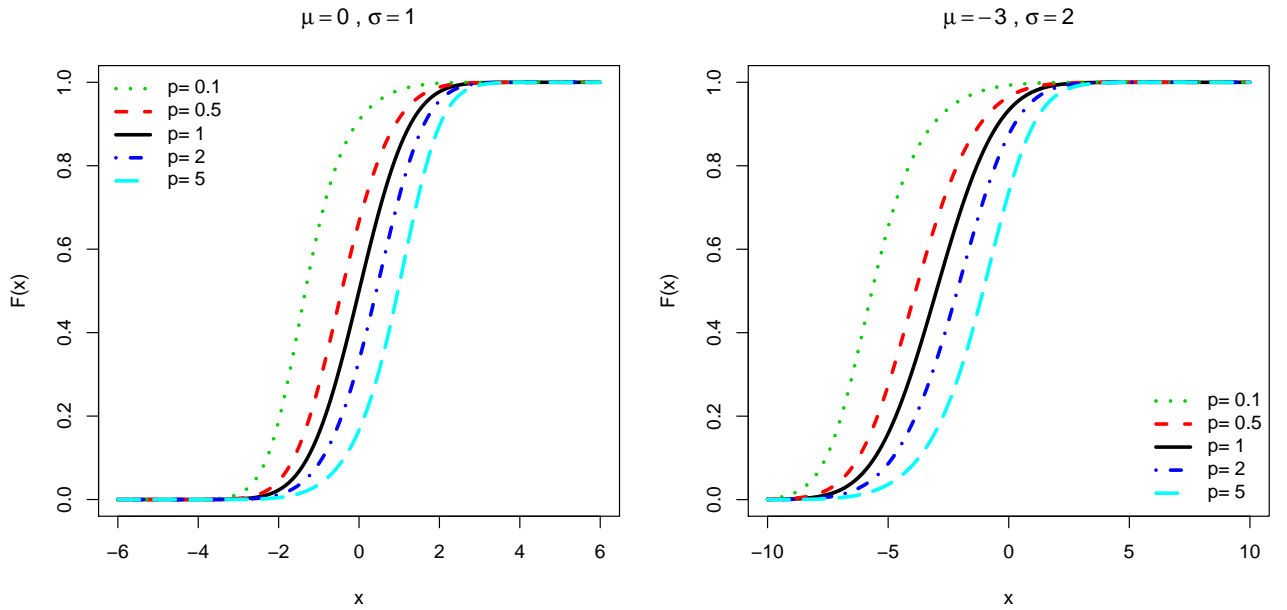


Figura 3.1: Gráfico da função de distribuição acumulada da distribuição MON para alguns valores de  $p$ ,  $\mu$  e  $\sigma$ .

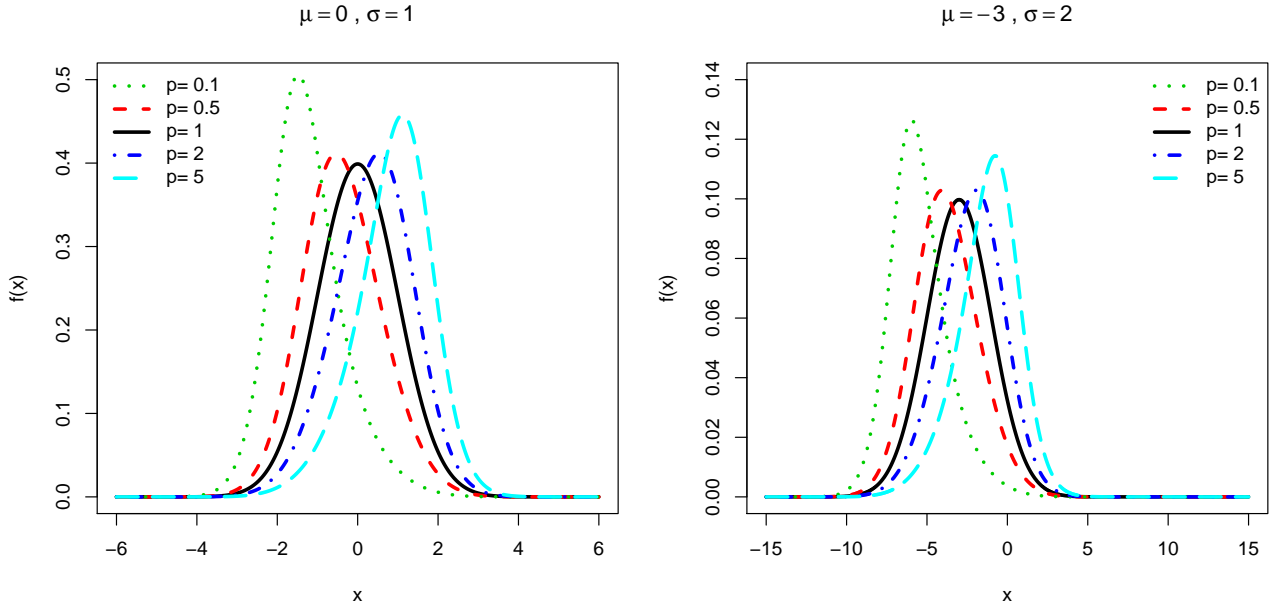


Figura 3.2: Gráfico da função densidade de probabilidade da distribuição MON para alguns valores de  $p$ ,  $\mu$  e  $\sigma$ .

### 3.2 Função de sobrevivência e função de risco

As funções de sobrevivência e de risco são importantes funções utilizadas para descrever estudos em análise de sobrevivência e em confiabilidade. As funções de sobrevivência e de risco para variável aleatória  $X$  com distribuição MON podem ser escritas, respectivamente, como

$$\begin{aligned}\bar{F}(x; p, \mu, \sigma) &= \frac{p[1 - \Phi(\frac{x-\mu}{\sigma})]}{1 - (1-p)[1 - \Phi(\frac{x-\mu}{\sigma})]}, \\ r(x; p, \mu, \sigma) &= \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma [1 - \Phi(\frac{x-\mu}{\sigma})] \{1 - (1-p)(1 - \Phi(\frac{x-\mu}{\sigma}))\}}.\end{aligned}\quad (3.5)$$

A Figura 3.3 ilustra a função de risco para uma variável aleatória  $X$  com distribuição MON para alguns valores de  $p$ ,  $\mu$  e  $\sigma$ .

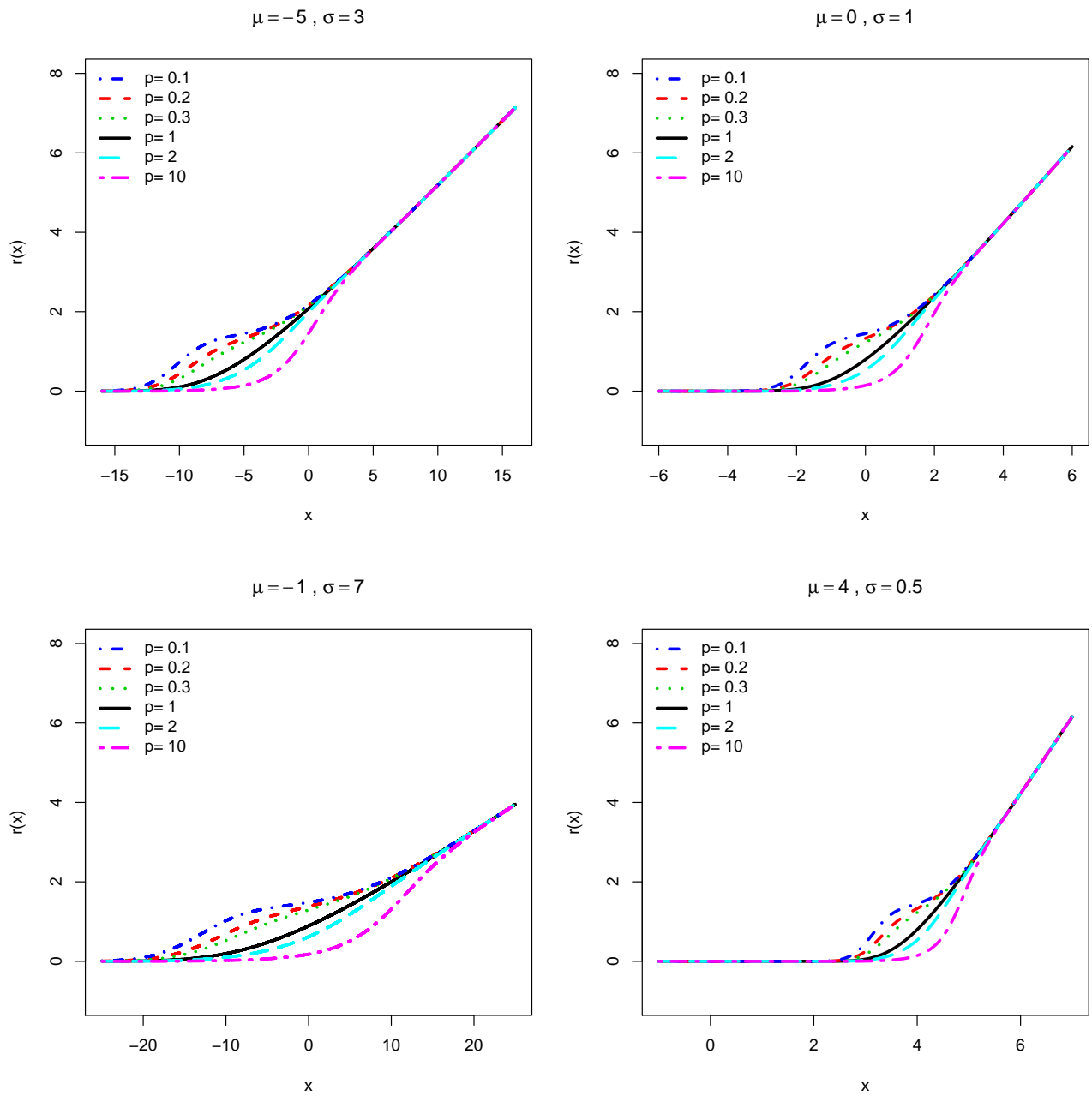


Figura 3.3: Gráfico da função de risco da distribuição MON para alguns valores de  $p, \mu$  e  $\sigma$ .

### 3.3 Propriedades da distribuição MON

#### 3.3.1 Moda

A moda da distribuição MON pode ser obtida por meio da maximização do logaritmo da função densidade de probabilidade dada na expressão (3.2), isto é,

$$\frac{d}{dx}[\log f(x; p, \mu, \sigma)] = -\frac{x - \mu}{\sigma^2} - \frac{2(1 - p)\phi\left(\frac{x - \mu}{\sigma}\right)}{1 - (1 - p)[1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]}. \quad (3.6)$$

Igualando a equação (3.6) a zero, pode-se encontrar o maior valor ou os valores mais frequentes da variável aleatória que representa a moda. Portanto, a moda da distribuição MON é dada pela solução da equação não linear

$$x = -\frac{2\sigma(1 - p)\phi\left(\frac{x - \mu}{\sigma}\right)}{1 - (1 - p)[1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]} + \mu.$$

Em distribuições contínuas a unimodalidade pode ser definida pelo comportamento da função de distribuição acumulada. Se existe um  $M$  tal que a função de distribuição acumulada é convexa para  $x < M$  e côncava se  $x > M$  então a distribuição é unimodal sendo  $M$  a moda. Para determinar a concavidade de uma função é necessário fazer a derivada segunda da função de distribuição acumulada que será denotada por  $F''(x; p, \mu, \sigma)$ . A derivada segunda da função de distribuição acumulada de um variável aleatória  $X$  com distribuição MON é dada por

$$F''(x; p, \mu, \sigma) = \frac{-p\phi\left(\frac{x - \mu}{\sigma}\right) \left[ (x - \mu) \{1 - (1 - p)[1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]\} + 2\sigma(1 - p)\phi\left(\frac{x - \mu}{\sigma}\right) \right]}{\sigma^3 \{1 - (1 - p)[1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]\}^3}. \quad (3.7)$$

Nota-se que não é possível resolver analiticamente para que valores de  $x$  a derivada segunda (3.7) é positiva ou negativa. Desta forma, alguns gráficos para  $F''(x; p, \mu, \sigma)$  foram construídos pelos quais tem-se evidências numéricas que a distribuição MON é unimodal uma vez que  $F''(x; p, \mu, \sigma)$  muda de sinal (é côncava e convexa) apenas uma vez (um único ponto crítico) como mostra a Figura 3.4.

#### 3.3.2 Mediana

A mediana é uma medida de tendência central que separa a metade inferior de uma amostra, de uma população, ou de uma distribuição de probabilidade, da outra metade superior. Denota-se  $m$  como a mediana de uma distribuição de probabilidade se ela satisfaz

$$P(X \leq m) \geq \frac{1}{2} \quad \text{e} \quad P(X \geq m) \geq \frac{1}{2}. \quad (3.8)$$

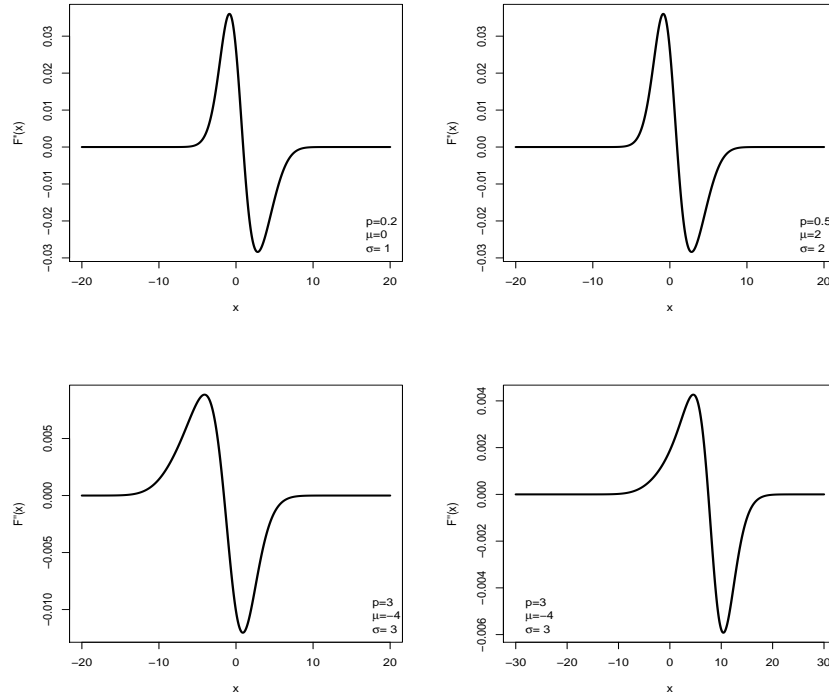


Figura 3.4: Derivada segunda da função de distribuição acumulada de uma variável aleatória com distribuição MON para alguns valores dos parâmetros.

Considerando  $X$  uma variável aleatória com distribuição MON padrão e resolvendo o sistema de inequações (3.8), tem-se

$$\begin{aligned}
 P(X \leq m) &= \frac{\Phi(x)}{\Phi(x) + p[1 - \Phi(x)]} \geq \frac{1}{2} \\
 2\Phi(x) &\geq \Phi(x) + p[1 - \Phi(x)] \\
 \Phi(x) &\geq \frac{p}{p+1},
 \end{aligned} \tag{3.9}$$

pois  $\Phi(x) + p[1 - \Phi(x)]$  é positivo para todo  $p$ . Por outro lado,

$$\begin{aligned}
 P(X \geq m) &= 1 - \frac{\Phi(x)}{\Phi(x) + p[1 - \Phi(x)]} \geq \frac{1}{2} \\
 2[\Phi(x) + p[1 - \Phi(x)] - \Phi(x)] &\geq \Phi(x) + p[1 - \Phi(x)] \\
 \Phi(x) &\leq \frac{p}{p+1}.
 \end{aligned} \tag{3.10}$$

Então, pelas desigualdades (3.9) e (3.10), a mediana da distribuição MON padrão é dada por

$$\Phi(m) = \frac{p}{1+p},$$



ou seja,

$$m = \Phi^{-1} \left( \frac{p}{1+p} \right).$$

Em outras palavras, a mediana  $m$  de uma distribuição MON padrão é o  $\frac{p}{1+p}$ -ésimo quantil de uma distribuição normal padrão. A Figura 3.5 mostra o gráfico para a moda e mediana de  $X$  com distribuição MON padrão para alguns valores de  $p$ . Pode-se observar que a medida que o valor de  $p$  aumenta a distância entre a moda e a mediana fica maior, indicando que a distribuição se afasta um pouco da simetria.

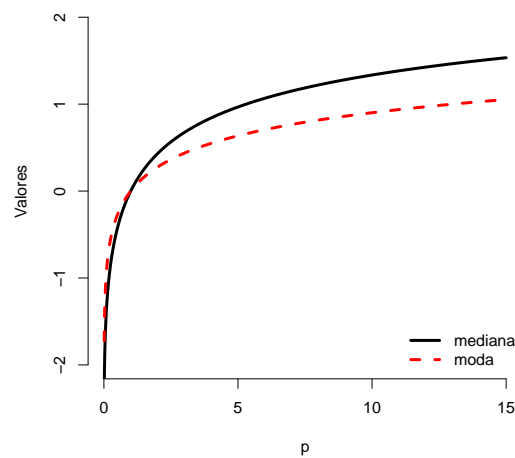


Figura 3.5: Gráfico da moda e mediana de  $X \sim MON(p, 0, 1)$  para diferentes valores de  $p$ .

### 3.4 Momentos

Nesta seção uma expressão geral para calcular os momentos de primeira, segunda, terceira e quarta ordens da distribuição MON padrão, como uma combinação infinita dos momentos da distribuição normal, é apresentada. Além disso, os primeiros termos dessas expansões foram calculados explicitamente. Estes resultados, até o presente trabalho, não estavam disponíveis na literatura. Para o desenvolvimento desses momentos utilizou-se alguns resultados apresentados no artigo de Bose e Gupta (1959). Para o caso geral, os cálculos podem ser feitos de maneira análoga fazendo-se a transformação  $X = \sigma Z + \mu$ , com  $X$  sendo uma variável aleatória com distribuição MON.

### Primeiro momento

O primeiro momento de uma variável aleatória com distribuição MON padrão utilizando a expressão (2.12) é dado por

$$E(X) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x\phi(x)[1-\Phi(x)]^n dx, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x\phi(x)\Phi(x)^n dx, & \text{se } q^{-1} = p > 1. \end{cases} \quad (3.11)$$

Para  $q^{-1} = p > 1$  tem-se que

$$\begin{aligned} E(X) &= \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x\phi(x)[\Phi(x)]^n dx \\ &= q \int_{-\infty}^{\infty} x\phi(x) dx + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x\phi(x)[\Phi(x)]^n dx \\ &= 0 + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x\phi(x)[\Phi(x)]^n dx \\ &= \sum_{n=1}^{\infty} q(1-q)^n \mu'_1(n+1, n+1), \end{aligned} \quad (3.12)$$

com  $\mu'_1(n+1, n+1) = (n+1) \int_{-\infty}^{\infty} x\phi(x)[\Phi(x)]^n dx$  representando o primeiro momento do máximo de  $n+1$  variáveis com distribuição normal padrão.

Bose e Gupta (1959) mostraram que  $\mu'_1(n+1, n+1)$  pode ser representado como

$$\begin{aligned} \mu'_1(n+1, n+1) &= \frac{\binom{n+1}{2}}{\pi} \int_{-\infty}^{\infty} [\Phi(x)]^{n-1} e^{-x^2} dx \\ &= \frac{n(n+1)}{2\pi} \int_{-\infty}^{\infty} [\Phi(x)]^{n-1} e^{-x^2} dx, \end{aligned} \quad (3.13)$$

e denotaram

$$I_n(a) = \int_{-\infty}^{\infty} [\Phi(ax)]^n e^{-x^2} dx. \quad (3.14)$$

Desta forma, o primeiro momento representado na expressão (3.12) pode ser escrito como

$$E(X) = \sum_{n=1}^{\infty} \frac{q(1-q)^n n(n+1)}{2\pi} I_{n-1}(1). \quad (3.15)$$

Para encontrar explicitamente os valores do primeiro momento do máximo de  $n+1$  variáveis com distribuição normal padrão, os mesmos autores apresentaram uma fórmula de recorrência para os  $I_n$ 's dada na expressão (3.16), a qual será provada com detalhes.

Sabe-se que a função  $h(x) = \Phi(ax) - \frac{1}{2}$  é uma função ímpar, enquanto que  $q(x) = e^{-x^2}$  é uma função par. Por sua vez, o produto dessas funções é uma função ímpar e portanto sua

integral é zero. Desta forma,

$$\int_{-\infty}^{\infty} \left[ \Phi(ax) - \frac{1}{2} \right]^{2m+1} e^{-x^2} dx = 0$$

Utilizando a expansão do binômio de Newton para  $[\Phi(ax) - \frac{1}{2}]^{2m+1}$ , tem-se que

$$\begin{aligned} & \int_{-\infty}^{\infty} \sum_{r=0}^{2m+1} \binom{2m+1}{r} (-1)^r \left(\frac{1}{2}\right)^r \Phi(ax)^{2m+1-r} e^{-x^2} dx = 0 \\ & \int_{-\infty}^{\infty} \Phi(ax)^{2m+1} e^{-x^2} dx + \int_{-\infty}^{\infty} \sum_{r=1}^{2m+1} \binom{2m+1}{r} (-1)^r \left(\frac{1}{2}\right)^r \Phi(ax)^{2m+1-r} e^{-x^2} dx = 0 \\ & I_{2m+1}(a) + \int_{-\infty}^{\infty} \sum_{r=1}^{2m+1} \binom{2m+1}{r} (-1)^r \left(\frac{1}{2}\right)^r \Phi(ax)^{2m+1-r} e^{-x^2} dx = 0 \\ & I_{2m+1}(a) = - \sum_{r=1}^{2m+1} \binom{2m+1}{r} (-1)^r \left(\frac{1}{2}\right)^r \int_{-\infty}^{\infty} \Phi(ax)^{2m+1-r} e^{-x^2} dx \\ & I_{2m+1}(a) = \sum_{r=1}^{2m+1} \binom{2m+1}{r} (-1)^{r+1} \left(\frac{1}{2}\right)^r I_{2m+1-r}(a). \end{aligned} \quad (3.16)$$

Fazendo  $n = 0$  na expressão (3.14) e substituindo  $m = 0$  em (3.16), respectivamente, encontra-se

$$\begin{aligned} I_0(a) &= \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \int_{-\infty}^{\infty} \frac{e^{-x^2}}{\sqrt{\pi}} dx = \sqrt{\pi}, \\ I_1(a) &= \frac{1}{2} I_0(a) = \frac{\sqrt{\pi}}{2}. \end{aligned}$$

pois  $\frac{e^{-x^2}}{\sqrt{\pi}}$  é a função densidade de probabilidade de uma variável aleatória com distribuição  $N(0, 1/2)$ .

Derivando a expressão (3.14) com respeito a  $a$  pode-se obter  $I_2(a)$ . Sabendo que a integral (3.14) converge uniformemente e o integrando é contínuo, a troca da integral com a derivada pode ser realizada. Assim,

$$\begin{aligned} \frac{d}{da} I_2(a) &= \int_{-\infty}^{\infty} \frac{d}{da} [\Phi(ax)]^2 e^{-x^2} dx \\ &= \int_{-\infty}^{\infty} 2\Phi(ax)\phi(ax)xe^{-x^2} dx \\ &= 2 \int_{-\infty}^{\infty} \Phi(ax) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}(a^2+2)} x dx. \end{aligned} \quad (3.17)$$

Utilizando a integração por partes com a mudança de variáveis  $u = \Phi(ax)$  e  $dv = e^{-\frac{x^2}{2}(a^2+2)}xdx$ , a expressão (3.17) pode ser escrita por

$$\begin{aligned} \frac{d}{da}I_2(a) &= \frac{2}{\sqrt{2\pi}} \left( \frac{a}{a^2+2} \right) \frac{1}{\sqrt{2a^2+1}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \sqrt{2(a^2+1)} e^{-\frac{x^2}{2}(a^2+2)} dx \\ &= \frac{2}{\sqrt{2\pi}} \left( \frac{a}{a^2+2} \right) \frac{1}{\sqrt{2(a^2+1)}} \\ &= \frac{1}{\sqrt{\pi}} \frac{a}{(a^2+2)\sqrt{(a^2+1)}}. \end{aligned}$$

Para recuperar  $I_2(a)$  é suficiente integrar com respeito a  $a$  ambos os lados da igualdade anterior obtendo

$$I_2(a) = \frac{1}{\sqrt{\pi}} \arctan(\sqrt{a^2+1}).$$

Novamente pela fórmula de recorrência (3.16), agora com  $m = 1$ , tem-se que  $I_3(a)$  é dado por

$$I_3(a) = \frac{3}{2\sqrt{\pi}} \arctan(\sqrt{a^2+1}) - \frac{\sqrt{\pi}}{4}.$$

Bose e Gupta (1959) afirmaram que os termos de ordem ímpar  $I_{2m+1}$  podem ser expressos como combinação linear de  $I_{2m}(a), I_{2m-2}(a), \dots, I_0(a)$  e para encontrar os termos de ordem par é suficiente derivar a equação (3.14) com respeito a  $a$ , como foi mostrado para o cálculo de  $I_2(a)$ . Desta forma, os termos da expansão do primeiro momento para a distribuição MON padrão podem ser encontrados.

Atribuindo alguns valores para  $n$  na série (3.12) tem-se que o primeiro momento,  $E(X)$ , da distribuição MON padrão para  $p > 1$ , utilizando a reparametrização  $q = p^{-1}$  pode ser representado como

$$\begin{aligned} E(X) &= \sum_{n=1}^{\infty} \frac{q(1-q)^n n(n+1)}{2\pi} I_{n-1}(1) \\ &= \frac{q(1-q)}{\pi} I_0(1) + \frac{3q(1-q)^2}{\pi} I_1(1) + \frac{6q(1-q)^3}{\pi} I_2(1) + \frac{10q(1-q)^4}{\pi} I_3(1) \\ &\quad + \sum_{n=5}^{\infty} \frac{q(1-q)^n n(n+1)}{2\pi} I_{n-1}(1) \\ &= \frac{q(1-q)}{\sqrt{\pi}} + \frac{3q(1-q)^2}{2\sqrt{\pi}} + 6q(1-q)^3 \frac{\arctan \sqrt{2}}{\sqrt{\pi^3}} + \\ &\quad \frac{5q(1-q)^4}{\pi} \left( \frac{3}{\sqrt{\pi}} \arctan \sqrt{2} - \frac{\sqrt{\pi}}{2} \right) + \sum_{n=5}^{\infty} \frac{q(1-q)^n n(n+1)}{2\pi} I_{n-1}(1). \end{aligned}$$

Para se obter a expansão do primeiro momento da distribuição MON padrão quando  $0 <$

$p \leq 1$ , pode-se definir

$$J_n(a) = \int_{-\infty}^{+\infty} [1 - \Phi(ax)]^n e^{-x^2} dx, \quad (3.18)$$

em que a relação de recorrência entre os  $J'_n(a)$ s será encontrada de forma similar ao desenvolvimento dos  $I'_n$ s apresentado anteriormente. Sabe-se que

$$\begin{aligned} \int_{-\infty}^{+\infty} \left[ \Phi(ax) - \frac{1}{2} \right]^{2m+1} e^{-x^2} dx &= 0 \\ \int_{-\infty}^{+\infty} \left[ 1 - \Phi(ax) - \frac{1}{2} \right]^{2m+1} e^{-x^2} dx &= 0 \\ \int_{-\infty}^{+\infty} \sum_{r=0}^{2m+1} \binom{2m+1}{r} \left( -\frac{1}{2} \right)^r [1 - \Phi(ax)]^{2m-r+1} e^{-x^2} dx &= 0 \\ \int_{-\infty}^{+\infty} \sum_{r=0}^{2m+1} \binom{2m+1}{r} \left( \frac{1}{2^r} \right) [1 - \Phi(ax)]^{2m-r+1} (-1)^r e^{-x^2} dx &= 0 \\ \sum_{r=0}^{2m+1} \binom{2m+1}{r} \left( \frac{1}{2^r} \right) (-1)^r \int_{-\infty}^{+\infty} [1 - \Phi(ax)]^{2m-r+1} e^{-x^2} dx &= 0 \\ \sum_{r=0}^{2m+1} \binom{2m+1}{r} \left( \frac{1}{2^r} \right) (-1)^r J_{2m-r+1}(a) &= 0 \\ J_{2m+1}(a) + \sum_{r=1}^{2m+1} \binom{2m+1}{r} \left( \frac{1}{2^r} \right) (-1)^r J_{2m-r+1}(a) &= 0 \end{aligned}$$

Portanto,

$$J_{2m+1}(a) = \sum_{r=1}^{2m+1} \binom{2m+1}{r} \left( \frac{1}{2^r} \right) (-1)^{r+1} J_{2m-r+1}(a). \quad (3.19)$$

Substituindo, respectivamente,  $n = 0$  em (3.18) e  $m = 0$  na equação (3.19), obtém-se

$$\begin{aligned} J_0(a) &= \sqrt{\pi}, \\ J_1(a) &= \frac{1}{2} J_0(a) = \frac{\sqrt{\pi}}{2}. \end{aligned}$$

De forma similar, como foram calculados anteriormente para os  $I'_n$ s, os termos ímpares  $J_{2m+1}(a)$  podem ser expressos como função linear dos  $J_{2m}(a)$ ,  $J_{2m-2}(a)$ ,  $\dots$ ,  $J_0(a)$ . Para encontrar os termos de ordem par, deriva-se a equação (3.18) com respeito a  $a$ . Neste caso, a troca da integral com a derivada é possível pois a integral converge uniformemente e o integrando é

contínuo. Derivando  $J_2(a)$  com respeito a  $a$  tem-se que

$$\begin{aligned}
\frac{d}{da} J_2(a) &= \int_{-\infty}^{+\infty} \frac{d}{da} [1 - \Phi(ax)]^2 e^{-x^2} dx \\
&= \int_{-\infty}^{+\infty} 2xe^{-x^2} \Phi(ax) \phi(ax) dx - \int_{-\infty}^{+\infty} 2xe^{-x^2} \phi(ax) dx \\
&= \frac{d}{da} I_2(a) - 2 \int_{-\infty}^{+\infty} xe^{-x^2} \phi(ax) dx \\
&= \frac{1}{\sqrt{\pi}} \frac{a}{(a^2 + 2)(\sqrt{a^2 + 1})} - 2 \int_{-\infty}^{+\infty} x \frac{e^{-\frac{x^2}{2}(a^2+2)}}{\sqrt{2\pi}} dx \\
&= \frac{1}{\sqrt{\pi}} \frac{a}{(a^2 + 2)(\sqrt{a^2 + 1})} - \frac{2}{\sqrt{a^2 + 2}} E(Y) \quad \text{onde } Y \sim N\left(0, \frac{1}{a^2 + 2}\right) \\
&= \frac{1}{\sqrt{\pi}} \frac{a}{(a^2 + 2)(\sqrt{a^2 + 1})}. \tag{3.20}
\end{aligned}$$

Integrando a equação (3.20) com respeito a  $a$ , pode-se obter o termo  $J_2(a)$ , resultando em

$$J_2(a) = \frac{\arctan(\sqrt{a^2 + 1})}{\sqrt{\pi}}.$$

Fazendo  $m = 1$  em (3.19), o termo  $J_3(a)$  é dado por

$$\begin{aligned}
J_3(a) &= \frac{3}{2} J_2(a) - \frac{3}{4} J_1(a) + \frac{1}{8} J_0(a) \\
&= \frac{3}{2} J_2(a) - \frac{1}{4} J_0(a) \\
&= \frac{3 \arctan(\sqrt{a^2 + 1})}{2\sqrt{\pi}} - \frac{\sqrt{\pi}}{4}.
\end{aligned}$$

Considerando  $0 < p \leq 1$ , o primeiro momento da distribuição MON padrão pode ser escrito como

$$\begin{aligned}
E(X) &= \sum_{n=0}^{\infty} (n+1)p(1-p)^n \int_{-\infty}^{+\infty} x\phi(x)[1-\Phi(x)]^n dx \\
&= p \int_{-\infty}^{+\infty} x\phi(x) dx + \sum_{n=1}^{\infty} (n+1)p(1-p)^n \int_{-\infty}^{+\infty} x\phi(x)[1-\Phi(x)]^n dx \\
&= 0 + \sum_{n=1}^{\infty} (n+1)p(1-p)^n \int_{-\infty}^{+\infty} x\phi(x)[1-\Phi(x)]^n dx
\end{aligned}$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} p(1-p)^n \int_{-\infty}^{+\infty} (n+1)x\phi(x)[1-\Phi(x)]^n dx \\
&= \sum_{n=1}^{\infty} p(1-p)^n \mu'_1(n+1, 1),
\end{aligned}$$

com  $\mu'_1(n+1, 1)$  representando o primeiro momento do mínimo de  $n+1$  variáveis com distribuição normal padrão. Por Bose e Gupta (1959),  $\mu'_1(n+1, 1)$  pode ser escrito como

$$\begin{aligned}
\mu'_1(n+1, 1) &= -\frac{1}{2\pi} \int_{-\infty}^{+\infty} n(n+1)[1-\Phi(x)]^{n-1} e^{-x^2} dx \\
&= -\frac{n(n+1)}{2\pi} J_{n-1}(1).
\end{aligned}$$

Portanto, o primeiro momento da distribuição MON padrão, quando  $0 < p \leq 1$ , é dado por

$$\begin{aligned}
E(X) &= -\sum_{n=1}^{\infty} \frac{p(1-p)^n n(n+1)}{2\pi} J_{n-1}(1) \\
&= -\frac{p(1-p)J_0(1)}{\pi} - \frac{3p(1-p)^2 J_1(1)}{\pi} - \frac{6p(1-p)^3 J_2(1)}{\pi} - \frac{10p(1-p)^4 J_3(1)}{\pi} \\
&\quad - \sum_{n=5}^{\infty} \frac{p(1-p)^n n(n+1)}{2\pi} J_{n-1}(1) \\
&= -\frac{p(1-p)}{\sqrt{\pi}} - \frac{3p(1-p)^2}{2\sqrt{\pi}} - \frac{6p(1-p)^3 \arctan \sqrt{2}}{\pi\sqrt{\pi}} \\
&\quad - \frac{10p(1-p)^4}{\pi} \left[ \frac{3 \arctan \sqrt{2}}{2\sqrt{\pi}} - \frac{\sqrt{\pi}}{4} \right] - \sum_{n=5}^{\infty} \frac{p(1-p)^n n(n+1)}{2\pi} J_{n-1}(1).
\end{aligned}$$

A Tabela 3.1, apresenta resultados numéricos para o primeiro momento da distribuição MON, para alguns valores de  $p$ , calculados de duas formas. A primeira delas, denotada por  $E(X)$ , é feita utilizando-se a expressão

$$E(X) = \int_{-\infty}^{\infty} x \frac{p\phi(x)}{\{1 - (1-p)[1 - \Phi(x)]\}^2} dx. \quad (3.21)$$

Já a esperança de  $X$  aproximada,  $\mathbf{E}(\mathbf{X})$  foi calculada pelas séries (3.18) e (3.21) somadas até os menores valores de  $n$  onde se observou uma boa aproximação quando comparado com o cálculo direto de  $E(X)$  pela expressão (3.21). A aproximação foi feita com três casas decimais. Observe-se que para alguns valores de  $p$  a aproximação foi satisfatória utilizando apenas alguns termos da série que representa o primeiro momento, quando comparados com os valores calculados diretamente pela expressão (3.21).

Tabela 3.1: Primeiro momento da distribuição MON padrão para alguns valores de  $p$ .

$p$	$n$	$E(X)$	$\mathbf{E}(\mathbf{X})$ aprox.
0.1	60	-1.244	-1.241
0.4	15	-0.513	-0.513
0.6	5	-0.287	-0.282
0.7	10	-0.201	-0.201
0.9	4	-0.059	-0.059
3	15	0.613	0.610
5	25	0.882	0.882
8	52	1.132	1.130
12	72	1.334	1.330
15	102	1.442	1.440

### Segundo momento

De forma similar como foi calculado para o primeiro momento da distribuição MON padrão, nesta subseção calcula-se os primeiros termos da expansão em série do segundo momento para uma amostra de  $n + 1$  observações da distribuição MON padrão.

Substituindo  $s = 2$  na expressão dos momentos (2.12), apresentada no Capítulo 2, o segundo momento da distribuição MON padrão pode ser escrito como

$$E(X^2) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [1 - \Phi(x)]^n dx, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [\Phi(x)]^n dx, & \text{se } q^{-1} = p > 1. \end{cases} \quad (3.22)$$

Desenvolvendo a expansão do segundo momento quando  $q^{-1} = p > 1$  tem-se que

$$E(X^2) = \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [\Phi(x)]^n dx \quad (3.23)$$

$$\begin{aligned} &= q \int_{-\infty}^{\infty} x^2 \phi(x) dx + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [\Phi(x)]^n dx \\ &= q + \sum_{n=1}^{\infty} q(1-q)^n \mu'_2(n+1, n+1), \end{aligned} \quad (3.24)$$

em que  $\mu'_2(n+1, n+1) = (n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [\Phi(x)]^n dx$  representa o segundo momento do máximo de  $n+1$  variáveis com distribuição normal padrão. A integral  $\int_{-\infty}^{\infty} x^2 \phi(x) dx = 1$ , pois denota o segundo momento da distribuição normal padrão, a variância.

Por Bose e Gupta (1959), tem-se que o segundo momento do máximo de  $n+1$  variáveis com distribuição normal padrão,  $\mu'_2(n+1, n+1)$ , pode ser expresso por

$$\mu'_2(n+1, n+1) = 1 + \frac{n(n-1)(n+1)}{4\pi} \int_{-\infty}^{\infty} \frac{[\Phi(x)]^{n-2} e^{-\frac{3}{2}x^2}}{\sqrt{2\pi}} dx. \quad (3.25)$$



Substituindo, respectivamente,  $n = 1$  e  $n = 2$  em (3.25), tem-se

$$\begin{aligned}\mu'_2(2, 2) &= 1, \\ \mu'_2(3, 3) &= 1 + \frac{3}{2\pi\sqrt{3}} \int_{-\infty}^{\infty} \frac{\sqrt{3}e^{-\frac{3}{2}x^2}}{\sqrt{2\pi}} dx = \frac{2\pi + \sqrt{3}}{2\pi}.\end{aligned}$$

Segundo Jones (1948), o segundo momento do máximo de quatro variáveis aleatórias normalmente distribuídas, com média 0 e variância 1, é dado por

$$\mu'_2(4, 4) = \frac{\pi + \sqrt{3}}{\pi}.$$

Substituindo os valores encontrados para  $\mu'_2(2, 2)$ ,  $\mu'_2(3, 3)$  e  $\mu'_2(4, 4)$  na expressão (3.24), o segundo momento da distribuição MON padrão para  $p > 1$  pode ser representado como

$$\begin{aligned}E(X^2) &= q + \sum_{n=1}^{\infty} q(1-q)^n \mu'_2(n+1, n+1) \\ &= q + q(1-q)\mu'_2(2, 2) + q(1-q)^2\mu'_2(3, 3) + q(1-q)^3\mu'_2(4, 4) \\ &\quad + \sum_{n=4}^{\infty} q(1-q)^n \mu'_2(n+1, n+1) \\ &= q + q(1-q) + q(1-q)^2 \left( \frac{2\pi + \sqrt{3}}{2\pi} \right) + q(1-q)^3 \left( \frac{\pi + \sqrt{3}}{\pi} \right) \\ &\quad + \sum_{n=4}^{\infty} q(1-q)^n \mu'_2(n+1, n+1).\end{aligned}$$

Analogamente, para  $0 < p \leq 1$  tem-se que

$$\begin{aligned}E(X^2) &= \sum_{n=0}^{\infty} p(1-p)^n (n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [1 - \Phi(x)]^n dx \\ &= p \int_{-\infty}^{\infty} x^2 \phi(x) dx + \sum_{n=1}^{\infty} p(1-p)^n (n+1) \int_{-\infty}^{\infty} x^2 \phi(x) [1 - \Phi(x)]^n dx \\ &= p + \sum_{n=1}^{\infty} p(1-p)^n \mu'_2(n+1, 1),\end{aligned}$$

em que  $\mu'_2(n+1, 1) = \int_{-\infty}^{\infty} x^2 \phi(x) [1 - \Phi(x)]^n dx$  denota o segundo momento do mínimo de  $n+1$  variáveis aleatórias com distribuição normal padrão. Novamente por Bose e Gupta (1959) o segundo momento do mínimo entre  $n+1$  variáveis aleatórias com distribuição normal padrão pode ser escrito como

$$\mu'_2(n+1, 1) = 1 + \frac{n(n-1)(n+1)}{4\pi} \int_{-\infty}^{\infty} \frac{[1 - \Phi(x)]^{n-2} e^{-\frac{3}{2}x^2}}{\sqrt{2\pi}} dx. \quad (3.26)$$

Substituindo  $n = 1, n = 2, n = 3$  em (3.26), respectivamente, tem-se que

$$\begin{aligned}\mu'_2(2, 1) &= 1, \\ \mu'_2(3, 1) &= \mu'_2(3, 3) = \frac{2\pi + \sqrt{3}}{2\pi} \\ \mu'_2(4, 1) &= 1 + \frac{6}{\pi} \int_{-\infty}^{\infty} \frac{e^{-\frac{3}{2}x^2}}{\sqrt{2\pi}} dx - \frac{6}{\pi} \int_{-\infty}^{\infty} \frac{\Phi(x)e^{-\frac{3}{2}x^2}}{\sqrt{2\pi}} dx \\ &= 1 + \frac{2\sqrt{3}}{\pi} - [\mu'_2(4, 4) - 1] \\ &= \frac{\sqrt{3} + \pi}{\pi}.\end{aligned}$$

Portanto, o segundo momento para a distribuição MON padrão, para  $0 < p \leq 1$ , é dado pela expressão

$$\begin{aligned}E(X^2) &= p + \sum_{n=1}^{\infty} p(1-p)^n \mu'_2(n+1, 1) \\ &= p + p(1-p)\mu'_2(2, 1) + p(1-p)^2\mu'_2(3, 1) + p(1-p)^3\mu'_2(4, 1) \\ &\quad + \sum_{n=4}^{\infty} p(1-p)^n \mu'_2(n+1, 1) \\ &= p + p(1-p) + p(1-p)^2 \left( \frac{2\pi + \sqrt{3}}{2\pi} \right) + p(1-p)^3 \left( \frac{\pi + \sqrt{3}}{\pi} \right) \\ &\quad + \sum_{n=4}^{\infty} p(1-p)^n \mu'_2(n+1, 1).\end{aligned}$$

### Terceiro e quarto momentos

Seguindo o mesmo procedimento das subseções anteriores, obteve-se os primeiros termos das expansões para o terceiro e quarto momentos da distribuição MON padrão. As demonstrações são apresentadas no Apêndice A.

O terceiro momento para a distribuição MON padrão, para  $0 < p \leq 1$ , pode ser representado como

$$\begin{aligned}E(X^3) &= -p(1-p) \frac{5\sqrt{\pi}}{2\pi} - p(1-p)^2 \frac{15\sqrt{\pi}}{4} - p(1-p)^3 \left[ \frac{15 \arctan \sqrt{2}}{\pi\sqrt{\pi}} + \frac{\sqrt{2\pi}}{2\pi^2} \right] \\ &\quad + \sum_{n=4}^{\infty} p(1-p)^n \mu'_3(n+1, 1),\end{aligned}$$

e para  $p > 1$ , usando a reparametrização  $q = p^{-1}$ , pode ser escrito como

$$E(X^3) = \frac{5\sqrt{\pi}q(1-q)}{2} + \frac{15q(1-q)^2}{4\sqrt{\pi}} + q(1-q)^3 \left[ \frac{15 \arctan \sqrt{2}}{\pi\sqrt{\pi}} + \frac{\sqrt{2\pi}}{2\pi^2} \right] + q(1-q)^4 \left[ \frac{75 \arctan \sqrt{2} + 5}{\sqrt{(2\pi)^3}} - \frac{25}{4\sqrt{\pi}} \right] + \sum_{n=5}^{\infty} q(1-q)^n \mu'_3(n+1, n+1).$$

O quarto momento para a distribuição MON padrão para  $0 < p \leq 1$  pode ser expresso por

$$E(X^4) = 3p + 3p(1-p) + p(1-p)^2 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{2\pi + \sqrt{3}}{2\pi} \right) \right] + p(1-p)^3 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{\pi + \sqrt{3}}{\pi} \right) \right] + p(1-p)^4 \left[ -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 1) + \frac{\sqrt{5}}{4\pi^2} \right] + \sum_{n=5}^{\infty} p(1-p)^n \mu'_4(n+1, 1).$$

Para  $p > 1$  utilizando a reparametrização  $q = p^{-1}$ , o quarto momento pode ser representado como

$$E(X^4) = 3q + 3q(1-q) + q(1-q)^2 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{2\pi + \sqrt{3}}{2\pi} \right) \right] + q(1-q)^3 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{\pi + \sqrt{3}}{\pi} \right) \right] + q(1-q)^4 \left( -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 5) + \frac{\sqrt{5}}{4\pi^2} \right) + \sum_{n=5}^{\infty} q(1-q)^n \mu'_4(n+1, n+1).$$

### 3.5 Variância

A variância de uma variável aleatória é uma medida de dispersão que indica o quão distante os seus valores estão afastados do valor esperado. Utilizando a expressão da variância como função dos dois primeiros momentos centrais, se  $X$  é uma variável aleatória com distribuição MON padrão, então sua variância para  $0 < p \leq 1$  é expressa por

$$\begin{aligned} \text{Var}(X) &= E(X^2) - [E(X)]^2 \\ &= p + \sum_{n=1}^{\infty} p(1-p)^n \mu'_2(n+1, 1) - \left[ \sum_{n=1}^{\infty} p(1-p)^n \mu'_1(n+1, 1) \right]^2, \end{aligned}$$

em que  $\mu'_2(n+1, 1)$  e  $\mu'_1(n+1, 1)$  denotam, respectivamente, o segundo e o primeiro momento do mínimo de  $n+1$  variáveis com distribuição normal padrão. Analogamente, se  $p > 1$ , considerando a reparametrização  $q = p^{-1}$ , tem-se que

$$\text{Var}(X) = q + \sum_{n=1}^{\infty} q(1-q)^n \mu'_2(n+1, n+1) - \left[ \sum_{n=1}^{\infty} q(1-q)^n \mu'_1(n+1, n+1) \right]^2,$$

com  $\mu'_2(n+1, n+1)$  e  $\mu'_1(n+1, n+1)$  representando, respectivamente, o segundo e o primeiro

momento do máximo de  $n + 1$  variáveis com distribuição normal padrão.

A Figura 3.6 ilustra a variância da distribuição MON padrão para alguns valores de  $p$  calculada numericamente pela expressão

$$\text{Var}(X) = \int_{-\infty}^{\infty} x^2 \frac{p\phi(x)}{\{1 - (1-p)[1 - \Phi(x)]\}^2} dx - \left( \int_{-\infty}^{\infty} x \frac{p\phi(x)}{\{1 - (1-p)[1 - \Phi(x)]\}^2} dx \right)^2.$$

Observa-se que a variância é crescente para  $0 < p < 1$  e decrescente para  $p > 1$ .

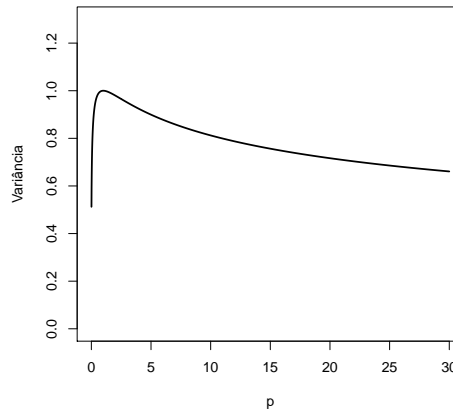


Figura 3.6: Variância de uma variável aleatória com distribuição MON padrão para diferentes valores de  $p$ .

### 3.6 Assimetria

As medidas de assimetria possibilitam analisar uma distribuição de acordo com as relações entre suas medidas de tendência central, tais como moda, média e mediana. Seguindo o artigo de Rubio e Stell (2012), para avaliar o comportamento do parâmetro  $p$  como mecanismo de introduzir assimetria na distribuição MON, foram consideradas três medidas de assimetria: o terceiro momento padronizado (EM), a medida de assimetria de Pearson (PM) e a medida de assimetria Arnold-Groeneveld (AG) introduzida por Arnold e Groeneveld (1995).

Conceitualmente, a assimetria descreve qual lado de uma distribuição tem uma cauda mais longa. Se a cauda longa está à direita, a assimetria é para a direita ou positiva, e se a cauda é mais longa do lado esquerdo, então a assimetria é à esquerda ou negativa.

A assimetria é usualmente medida pelo terceiro momento padronizado definido por

$$\text{EM} = \frac{\mu_3}{\mu_2^{3/2}},$$

em que  $\mu_2$  e  $\mu_3$  são, respectivamente, o segundo e o terceiro momentos centrais. Essa medida de assimetria pode ser enganadora, uma vez que pode ser dominada pelo comportamento no ex-

tremo das caudas. Além disso, não é sempre que está definida e pode ser de difícil interpretação pois não é limitada.

Uma outra medida de assimetria é a medida de assimetria de Pearson (PM) (Pearson, 1895), definida por

$$PM = \frac{Me - Mo}{\mu_2^{1/2}},$$

com Me e Mo representando, respectivamente, a média e a moda da distribuição. Essa medida permite calcular o grau de assimetria de uma distribuição. Segundo Ferreira (2005), se  $|PM| < 0.15$  a assimetria da distribuição é considerada fraca. A distribuição será considerada moderadamente assimétrica se  $0.15 \leq |PM| < 1$ , e para  $|PM| \geq 1$  a assimetria é considerada forte.

Outra medida de verificação de assimetria em termos da moda foi introduzida por Arnold e Groeneveld (1995). Os autores assumem nessa medida que a moda existe e é única, e a definem como

$$\begin{aligned} AG &= [1 - F(Mo)] - F(Mo) \\ &= 1 - 2F(Mo), \end{aligned}$$

com F sendo a função de distribuição acumulada e Mo denotando a moda. A medida AG calcula a diferença entre as massas alocadas em ambos os lados da moda. Ela satisfaz  $-1 < AG < 1$ , em que  $AG = 0$  significa que a distribuição é simétrica. Valores negativos de AG são associados com assimetria à esquerda e positivos refletem assimetria à direita. Além de ser limitada, uma das vantagens desta medida é que ela não requer a existência de qualquer momento.

A Figura 3.7 ilustra numericamente as medidas de assimetria EM, PM e AG para a distribuição MON padrão. O valor máximo, em módulo, assumido pela medida de assimetria de Pearson foi de 0.2606, o que indica que a distribuição MON, de acordo com essa medida, pode ser considerada moderadamente assimétrica. Já para a medida de assimetria AG, o valor máximo em módulo encontrado foi 0.1445, o que sugere que a distribuição MON não possui assimetria significativa. Pelos gráficos tem-se que o terceiro momento padronizado, a medida de assimetria de Pearson e medida Arnold-Groeneveld, atribuem à distribuição MON, respectivamente, uma assimetria forte, moderada e fraca. Como a medida AG é limitada e de fácil interpretação, utilizando essa medida para medir o grau de assimetria, conclui-se que a distribuição MON não é recomendada para modelar dados com assimetria alta ou mesmo moderada. Portanto, a distribuição MON não compete com distribuições assimétricas como afirmaram García et al. (2010).

### 3.7 Curtose

A curtose é uma medida de dispersão que caracteriza o pico ou achatamento da curva da função de distribuição de probabilidade. Essa medida pode ser definida como a razão entre o quarto

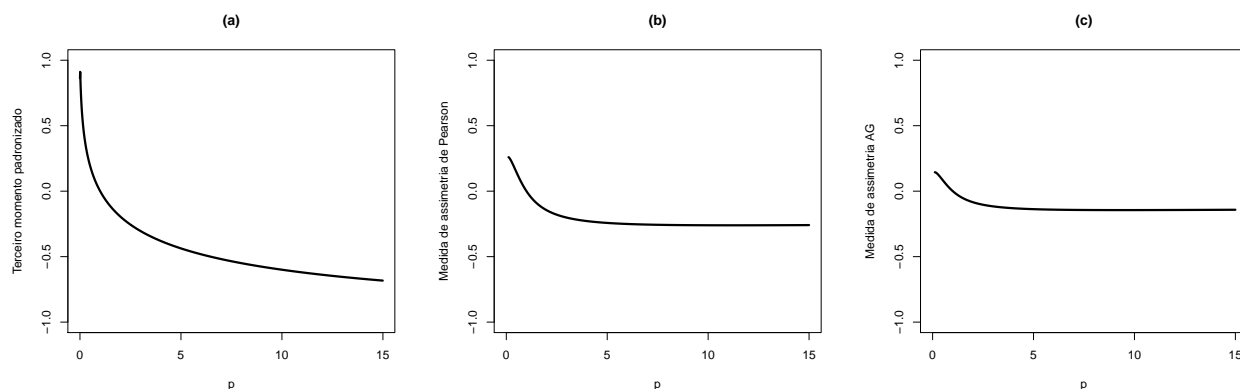


Figura 3.7: Medidas de assimetria da distribuição MON: (a) terceiro momento padronizado; (b) medida de assimetria de Pearson; (c) medida de assimetria AG .

momento central e o quadrado do segundo momento central (variância), ou seja,

$$\begin{aligned} \text{Curtose}(X) &= \frac{E(X - E(X))^4}{\{E[X - E(X)]^2\}^2} \\ &= \frac{E(X^4) - 4E(X^3)E(X) + 6E(X^2)E^2(X) - 3E^4(X)}{[\text{Var}(X)]^2}. \end{aligned}$$

Sob essa definição, a curtose da distribuição normal é igual a 3. Como a curtose da distribuição MON padrão é uma expressão longa e complicada, ela foi calculada numericamente diretamente pelas integrais para alguns valores de  $p$ , representada na Figura 3.8. Observa-se que os valores para a curtose são maiores do que 3 e, portanto, a distribuição MON padrão é mais alta (afunilada) e concentrada que a distribuição normal. Diz-se então que a distribuição MON padrão é leptocúrtica e desta forma possui caudas pesadas. Isso implica que a MON pode competir com a  $t$ -Student.

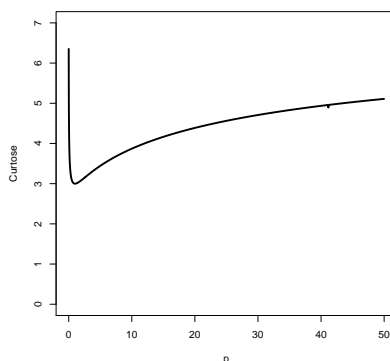


Figura 3.8: Curtose da distribuição MON padrão para determinados valores de  $p$ .

### 3.8 Entropia

A entropia de uma variável aleatória  $X$  com função densidade de probabilidade  $f(x)$  é uma medida que quantifica a incerteza ou aleatoriedade de um sistema. Essa medida é muito utilizada nas áreas de engenharia e ciências em geral. Nesta seção encontra-se duas medidas de entropia para a distribuição MON padrão, as entropias de Shannon e de Rényi, as quais até o presente trabalho não estavam disponíveis na literatura.

A entropia de Shannon, é definida por

$$H_x = - \int_{T(x)} f(x) \ln f(x) dx,$$

em que  $T(x)$  é o suporte de  $X$  e  $f(x)$  é função de densidade de probabilidade. Considerando a distribuição MON padrão, a entropia de Shannon pode ser escrita como

$$H_x = - \int_{-\infty}^{\infty} \frac{p\phi(x)}{\{1 - (1-p)[1 - \Phi(x)]\}^2} \ln \left( \frac{p\phi(x)}{\{1 - (1-p)[1 - \Phi(x)]\}^2} \right) dx. \quad (3.27)$$

Utilizando a expansão do binômio de Newton para expoentes negativos, o denominador da expressão (3.27) pode ser reescrito como

$$\begin{aligned} \{1 - (1-p)[1 - \Phi(x)]\}^{-2} &= \sum_{m=0}^{\infty} \binom{-2}{m} (-1)^m (1-p)^m [1 - \Phi(x)]^m \\ &= \sum_{m=0}^{\infty} (m+1)(1-p)^m [1 - \Phi(x)]^m. \end{aligned}$$

Para  $0 < p < 1$ , tem-se que  $|(1-p)[1 - \Phi(x)]| < 1$  e, portanto, a expansão do logaritmo pode ser expressa como

$$-2 \ln \{1 - (1-p)[1 - \Phi(x)]\} = 2 \sum_{n=1}^{\infty} \frac{(1-p)^n [1 - \Phi(x)]^n}{n}.$$

Portanto, a entropia de Shannon para a distribuição MON padrão pode ser representada por

$$\begin{aligned} H_x &= - \int_{-\infty}^{\infty} p\phi(x) \sum_{m=0}^{\infty} (m+1)(1-p)^m [1 - \Phi(x)]^m \left\{ \ln [p\phi(x)] - 2 \ln \{1 - (1-p)[1 - \Phi(x)]\} \right\} dx \\ &= - \int_{-\infty}^{\infty} p\phi(x) \sum_{m=0}^{\infty} (m+1)(1-p)^m [1 - \Phi(x)]^m \left\{ \ln [p\phi(x)] + 2 \sum_{n=1}^{\infty} \frac{(1-p)^n [1 - \Phi(x)]^n}{n} \right\} dx. \end{aligned}$$

Para  $p \geq 1$  considerando a reparametrização  $p = q^{-1}$  tem-se que a entropia de Shannon para a distribuição MON é expressa por

$$H_x = - \int_{-\infty}^{\infty} \frac{q\phi(x)}{\{1 - (1-q)\Phi(x)\}^2} \ln \left( \frac{q\phi(x)}{\{1 - (1-q)\Phi(x)\}^2} \right) dx. \quad (3.28)$$

Pela expansão do binômio de Newton para expoentes negativos, o denominador da expressão (3.28) pode ser reescrito como

$$\{1 - (1 - q)\Phi(x)\}^{-2} = \sum_{m=0}^{\infty} (m+1)(1-q)^m \Phi^m(x).$$

Para  $p > 1$  tem-se que  $0 < q < 1$ , e desta forma  $|(1 - q)\Phi(x)| < 1$ . Sendo assim, a expansão do logaritmo pode ser feita obtendo-se

$$-2 \ln \{1 - (1 - q)\Phi(x)\} = 2 \sum_{n=1}^{\infty} \frac{(1 - q)^n \Phi^n(x)}{n}.$$

Desta forma, quando  $p > 1$ , a entropia de Shannon para a distribuição MON padrão pode ser representada por

$$H_x = - \int_{-\infty}^{\infty} q\phi(x) \sum_{m=0}^{\infty} (m+1)(1-q)^m \Phi^m(x) \left\{ \ln [q\phi(x)] + 2 \sum_{n=1}^{\infty} \frac{(1-q)^n \Phi^n(x)}{n} \right\} dx.$$

Uma outra medida de entropia chamada de entropia de Rényi é definida por

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \left[ \int f^\gamma(x) dx \right], \quad \text{com } \gamma > 0, \gamma \neq 1.$$

Quando  $\gamma \rightarrow 1$ , a entropia de Rényi converge para a entropia de Shannon.

Considerando a função densidade de probabilidade apresentada na expressão (3.2) da distribuição MON padrão, tem-se que a entropia de Rényi para essa distribuição pode ser escrita por

$$\begin{aligned} I_R &= \frac{1}{1-\gamma} \log \left[ \int_{-\infty}^{\infty} \left( \frac{p\phi(x)}{\{1 - (1-p)[1 - \Phi(x)]\}^2} \right)^\gamma dx \right] \\ &= \frac{1}{1-\gamma} \log \left[ \int_{-\infty}^{\infty} p^\gamma \phi^\gamma(x) \{1 - (1-p)[1 - \Phi(x)]\}^{-2\gamma} dx \right]. \end{aligned}$$

Utilizando a expansão do binômio de Newton para expoentes negativos tem-se que a entropia de Rényi para a distribuição MON padrão pode ser expressa como

$$\begin{aligned} I_R &= \frac{1}{1-\gamma} \log \left[ \int_{-\infty}^{\infty} p^\gamma \phi^\gamma(x) \sum_{m=0}^{\infty} \binom{-2\gamma}{m} (-1)^m (1-p)^m [1 - \Phi(x)]^m dx \right] \\ &= \frac{1}{1-\gamma} \log \left[ \sum_{m=0}^{\infty} \binom{-2\gamma}{m} p^\gamma (-1)^m (1-p^{-1})^m \int_{-\infty}^{\infty} [1 - \Phi(x)]^m \phi^\gamma(x) dx \right]. \end{aligned}$$



Fazendo a substituição  $x = \sqrt{\frac{2}{\gamma}}y$  a entropia de Rényi pode ser reescrita como

$$\begin{aligned} I_R &= \frac{1}{1-\gamma} \log \left[ \sum_{m=0}^{\infty} \binom{-2\gamma}{m} p^\gamma (-1)^m (1-p^{-1})^m \int_{-\infty}^{\infty} \frac{[1 - \Phi(\sqrt{\frac{2}{\gamma}}y)]^m e^{-y^2} dx}{\sqrt{\gamma\pi}} \right] \\ &= \frac{1}{\sqrt{\gamma\pi}(1-\gamma)} \log \left[ \sum_{m=0}^{\infty} \binom{-2\gamma}{m} p^\gamma (-1)^m (1-p^{-1})^m J_m \left( \sqrt{\frac{2}{\gamma}} \right) \right], \end{aligned} \quad (3.29)$$

em que  $J_m(\sqrt{\frac{2}{\gamma}})$  foi dado anteriormente em (3.18).

### 3.9 Identificabilidade

A suposição de identificabilidade é fundamental quando se considera a teoria da estatística clássica. Uma família de distribuição de probabilidade paramétrica é dita ser não identificável se existem dois parâmetros diferentes que determinam a mesma distribuição de probabilidade. Famílias não identificáveis aparecem em muitos modelos estatísticos tais como misturas de distribuições, distribuições generalizadas, análise de variância, modelos de sobrevivência entre outros. Nesta seção mostra-se que o modelo MON é identificável, o que até o presente trabalho não estava disponível na literatura.

A Definição 1 apresenta o conceito de parâmetro identificável.

**Definição 1.** *O parâmetro  $\theta_1$  pertencente ao espaço paramétrico  $\Theta$  é globalmente identificável se não existe um outro valor  $\theta_2$  em  $\Theta$  tal que  $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2}$ .*

*O parâmetro  $\theta_1$  será chamado de localmente identificável se existe uma vizinhança  $\mathbb{V}$  de  $\theta_1$  tal que*

$$\forall \theta_2 \neq \theta_1, \quad \theta_2 \in \mathbb{V} \Rightarrow \mathbb{P}_{\theta_1} \neq \mathbb{P}_{\theta_2}.$$

Ou seja, um parâmetro  $\theta_1$  no espaço paramétrico é dito identificável quando não existe outro parâmetro  $\theta_2$  tal que quando aplicados na função de distribuição tenham o mesmo valor. A Definição 2 estende o conceito de identificabilidade de um parâmetro para identificabilidade de um modelo.

**Definição 2.** *Dados  $\theta_1, \theta_2 \in \Theta$ , um modelo é não identificável se*

$$\theta_1 \neq \theta_2 \Rightarrow f(x, \theta_1) = f(x, \theta_2), \quad \forall x \in \Omega,$$

*com  $\Omega$  denotando o espaço amostral.*

Em outras palavras, um modelo é dito não identificável se existem dois parâmetros diferentes que determinam o mesmo valor da função de densidade de probabilidade. Quando se tem um problema de identificabilidade não é possível encontrar qual dos parâmetros realmente gerou a amostra.

Existe na literatura um critério geral para identificabilidade de um parâmetro proposto por Bowden (1973). Para isso, é necessário a definição da informação de Kullback-Leibler dada por

$$H(\theta, \theta_0) = E_{\theta_0} \left\{ \ln \left[ \frac{f(x, \theta_0)}{f(x, \theta)} \right] \right\} = \int_{-\infty}^{\infty} \ln \left[ \frac{f(x, \theta_0)}{f(x, \theta)} \right] f(x, \theta_0) dx. \quad (3.30)$$

A informação de Kullback-Libler mede a informação perdida quando o modelo  $f(x, \theta)$  é aproximado por  $f(x, \theta_0)$ . O Teorema 5 a seguir mostra duas propriedades muito úteis da informação de Kullback-Leibler.

**Teorema 5.** *A informação de Kullback-Leibler  $H$  satisfaz*

$$\begin{aligned} H(\theta_1, \theta_2) &\geq 0, \\ H(\theta_1, \theta_2) &= 0 \Leftrightarrow \mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2}. \end{aligned}$$

O critério geral de identificabilidade proposto por Bowden (1973) é apresentado no Teorema 6 o qual utiliza a informação de Kullback-Leibler para provar a identificabilidade de um parâmetro.

**Teorema 6.** *Seja  $\mathbb{P}_{\theta}$  uma distribuição absolutamente contínua para todo  $\theta \in \Theta$ . O parâmetro  $\theta_0$  é globalmente identificável se, e somente se, a equação*

$$H(\theta, \theta_0) = 0$$

*tem como solução única  $\theta_0$  em todo  $\Theta$ . O parâmetro será dito localmente identificável se, somente se,  $\theta_0$  é única solução em alguma vizinhança de  $\theta_0$ .*

As demonstrações dos teoremas 5 e 6 podem ser encontradas em Bowden (1973). As definições e os teoremas apresentados serão utilizados para mostrar que a distribuição MON padrão é identificável no ponto  $p = 1$ . Utilizando a expressão (3.2) para a função densidade de probabilidade quando  $0 < p \leq 1$ , tem-se que a informação de Kullback-Leibler para a distribuição MON padrão no ponto  $p = 1$  é expressa por

$$\begin{aligned} H(p, 1) &= \int_{-\infty}^{\infty} \ln \left( \frac{\{1 - (1 - p)[1 - \Phi(x)]\}^2}{p} \right) \phi(x) dx \\ &= -\ln p + 2 \int_{-\infty}^{\infty} \ln \{1 - (1 - p)[1 - \Phi(x)]\} \phi(x) dx. \end{aligned}$$

Como  $|(1 - p)[1 - \Phi(x)]| < 1$  a expansão do logaritmo pode ser feita obtendo-se

$$\begin{aligned} H(p, 1) &= -\ln p + 2 \int_{-\infty}^{\infty} \left( -\sum_{n=1}^{\infty} \frac{(1 - p)^n [1 - \Phi(x)]^n}{n} \right) \phi(x) dx \\ &= -\ln p - 2 \sum_{n=1}^{\infty} \frac{(1 - p)^n}{n} \int_{-\infty}^{\infty} [1 - \Phi(x)]^n \phi(x) dx. \end{aligned}$$

Tem-se que  $(n + 1)[1 - \Phi(x)]^n \phi(x)$  representa a densidade do mínimo de  $n + 1$  variáveis aleatórias com distribuição normal padrão. Portanto,  $\int_{-\infty}^{\infty} (n + 1)[1 - \Phi(x)]^n \phi(x) dx = 1$ . Desta forma, a

informação de Kullback-Leibler pode ser reescrita como

$$\begin{aligned}
H(p, 1) &= -\ln p - 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n(n+1)} \int_{-\infty}^{\infty} (n+1)[1-\Phi(x)]^n \phi(x) dx \\
&= -\ln p - 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n(n+1)} \\
&= -\ln p - 2 \sum_{n=1}^{\infty} \left[ \frac{(1-p)^n}{n} - \frac{(1-p)^n}{n+1} \right] \\
&= -\ln p - 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n} + 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n+1} \\
&= -\ln p - 2(-\ln p) + 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n+1} \\
&= \ln p + 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n+1} \\
&= -\sum_{n=1}^{\infty} \frac{(1-p)^n}{n} + 2 \sum_{n=1}^{\infty} \frac{(1-p)^n}{n+1} \\
&= \sum_{n=1}^{\infty} (1-p)^n \left[ \frac{2}{n+1} - \frac{1}{n} \right] \\
&= \sum_{n=1}^{\infty} \frac{(1-p)^n (n-1)}{n(n+1)}. \tag{3.31}
\end{aligned}$$

Como  $n \geq 1$ , a série (3.31) possui apenas termos positivos. Sendo assim, a informação de Kullback-Leiber é maior ou igual que zero para  $0 < p \leq 1$ , com a igualdade se  $p = 1$ . Portanto, a distribuição MON é identificável no ponto  $p = 1$ .

Analogamente, para  $p \geq 1$ , a informação de Kullback-Leibler para a distribuição MON padrão é representada por

$$\begin{aligned}
H(p, 1) &= \int_{-\infty}^{\infty} \ln \left( \frac{[1 - (1-p^{-1})\Phi(x)]^2}{p^{-1}} \right) \phi(x) dx \\
&= -\ln p^{-1} + 2 \int_{-\infty}^{\infty} \ln [1 - (1-p^{-1})\Phi(x)] \phi(x) dx.
\end{aligned}$$

Utilizando a expansão do logaritmo e o fato de que  $(n+1)\Phi(x)^n \phi(x)$  representa a densidade do máximo de  $n+1$  variáveis aleatórias com distribuição normal padrão, tem-se que  $H(p, 1)$

pode ser reescrita como

$$\begin{aligned}
H(p, 1) &= -\ln p^{-1} + 2 \int_{-\infty}^{\infty} \left( -\sum_{n=1}^{\infty} \frac{(1-p^{-1})^n \Phi(x)^n}{n} \right) \phi(x) dx \\
&= -\ln p^{-1} - 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n(n+1)} \int_{-\infty}^{\infty} (n+1) \Phi(x)^n \phi(x) dx \\
&= -\ln p^{-1} - 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n(n+1)} \\
&= -\ln p^{-1} - 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n} + 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n+1} \\
&= -\ln p^{-1} - 2(-\ln p^{-1}) + 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n+1},
\end{aligned}$$

continuando, tem-se que

$$\begin{aligned}
H(p, 1) &= \ln p^{-1} + 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n+1} \\
&= -\sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n} + 2 \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n}{n+1} \\
&= \sum_{n=1}^{\infty} (1-p^{-1})^n \left[ \frac{2}{n+1} - \frac{1}{n} \right] \\
&= \sum_{n=1}^{\infty} \frac{(1-p^{-1})^n (n-1)}{n(n+1)}. \tag{3.32}
\end{aligned}$$

Como  $n \geq 1$ , a série da pela expressão (3.32) possui apenas termos positivos e portanto a informação de Kullback-Leibler para  $p \geq 1$  é não negativa. Novamente,  $H(1, 1) = 0$  e portanto a distribuição MON padrão é identificável no ponto  $p = 1$ .

Para mostrar que a distribuição MON padrão é identificável em todo ponto do espaço paramétrico é necessário provar que  $H(\alpha, \beta) \geq 0$ , com a igualdade se, e somente se,  $\alpha = \beta$ . Sejam  $0 < \alpha \leq 1$  e  $0 < \beta \leq 1$ . A informação de Kullback-Leibler  $H(\alpha, \beta)$  é dada por

$$H(\alpha, \beta) = \int_{-\infty}^{\infty} \ln \left[ \frac{\alpha}{\beta} \left\{ \frac{1 - (1-\beta)[1 - \Phi(x)]}{1 - (1-\alpha)[1 - \Phi(x)]} \right\}^2 \right] \frac{\alpha \phi(x)}{\{1 - (1-\alpha)[1 - \Phi(x)]\}^2} dx.$$

Utilizando as expansões do binômio de Newton para expoentes negativos e do logaritmo, a

informação de Kullback-Leibler,  $H(\alpha, \beta)$ , pode ser reescrita como

$$\begin{aligned}
H(\alpha, \beta) &= \ln\left(\frac{\alpha}{\beta}\right) + 2\alpha \int_{-\infty}^{\infty} \ln\{1 - (1 - \beta)[1 - \Phi(x)]\} \phi(x) \sum_{m=0}^{\infty} (m+1)(1-\alpha)^m [1 - \Phi(x)]^m dx \\
&\quad - 2\alpha \int_{-\infty}^{\infty} \ln\{1 - (1 - \alpha)[1 - \Phi(x)]\} \phi(x) \sum_{m=0}^{\infty} (m+1)(1-\alpha)^m [1 - \Phi(x)]^m dx \\
&= \ln\left(\frac{\alpha}{\beta}\right) - 2\alpha \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{(1-\beta)^n [1 - \Phi(x)]^n}{n} \phi(x) \sum_{m=0}^{\infty} (m+1)(1-\alpha)^m [1 - \Phi(x)]^m dx \\
&\quad + 2\alpha \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{(1-\alpha)^n [1 - \Phi(x)]^n}{n} \phi(x) \sum_{m=0}^{\infty} (m+1)(1-\alpha)^m [1 - \Phi(x)]^m dx \\
&= \ln\left(\frac{\alpha}{\beta}\right) - 2\alpha \sum_{n=1}^{\infty} \frac{(1-\beta)^n}{n} \sum_{m=0}^{\infty} \frac{(m+1)(1-\alpha)^m}{(m+n+1)} \int_{-\infty}^{\infty} (m+n+1)[1 - \Phi(x)]^{m+n} \phi(x) dx \\
&\quad + 2\alpha \sum_{n=1}^{\infty} \frac{(1-\alpha)^n}{n} \sum_{m=0}^{\infty} \frac{(m+1)(1-\alpha)^m}{(m+n+1)} \int_{-\infty}^{\infty} (m+n+1)[1 - \Phi(x)]^{m+n} \phi(x) dx.
\end{aligned}$$

Observa-se que  $\int_{-\infty}^{\infty} (m+n+1)[1 - \Phi(x)]^{m+n} \phi(x) dx = 1$  pois  $(m+n+1)[1 - \Phi(x)]^{m+n} \phi(x)$  representa a densidade do mínimo de  $(m+n+1)$  variáveis aleatórias com distribuição normal padrão. Portanto,  $H(\alpha, \beta)$  pode ser reescrita como

$$\begin{aligned}
H(\alpha, \beta) &= \ln\left(\frac{\alpha}{\beta}\right) + 2\alpha \sum_{m=0}^{\infty} \frac{(m+1)(1-\alpha)^m}{(m+n+1)} \sum_{n=1}^{\infty} \frac{(1-\alpha)^n - (1-\beta)^n}{n} \\
&= 2\alpha \sum_{m=0}^{\infty} \frac{(m+1)(1-\alpha)^m}{(m+n+1)} \sum_{n=1}^{\infty} \frac{(\alpha - \beta) [(1-\beta)^{n-1} + (1-\beta)^{n-2}(1-\alpha) + \dots + (1-\alpha)^{n-1}]}{n} \\
&\quad + \ln\left(\frac{\alpha}{\beta}\right).
\end{aligned}$$

Como  $[(1-\beta)^{n-1} + (1-\beta)^{n-2}(1-\alpha) + \dots + (1-\alpha)^{n-1}]$  contém apenas termos positivos, considerando  $\alpha > \beta$ , tem-se que  $H(\alpha, \beta) \geq 0$ , com a igualdade se, e somente se,  $\alpha = \beta$ . Portanto, a distribuição MON padrão é identificável em todo o espaço paramétrico. Para valores de  $\alpha, \beta \geq 1$ ,  $0 < \alpha \leq 1$  e  $\beta > 1$  ou  $0 < \beta \leq 1$  e  $\alpha > 1$  as demonstrações são análogas.

No próximo capítulo um breve estudo sobre inferência é apresentado. Expressões para os parâmetros estimados da distribuição MON pelo método da máxima verossimilhança são calculadas e uma revisão sobre testes paramétricos e não paramétricos e alguns critérios de seleção de modelos também são feitos. A matriz de informação esperada é encontrada e intervalos de confiança para os parâmetros são construídos.



# Capítulo 4

## Inferência

Neste capítulo aborda-se alguns tópicos relacionados à inferência da distribuição Marshall-Olkin normal (MON). Expressões para os estimadores dos parâmetros de uma amostra independente da distribuição MON são obtidas utilizando o método da máxima verossimilhança. A matriz de informação de Fisher do modelo em estudo também é apresentada, e as propriedades assintóticas dos estimadores de máxima verossimilhança são utilizadas para a construção de intervalos de confiança e testes de hipóteses. Os testes adotados foram da razão de verossimilhança, da razão de verossimilhança generalizada e de bondade de ajuste. Alguns critérios de seleção de modelos também são considerados.

### 4.1 Estimação

#### 4.1.1 Estimação pontual

Seja  $\mathbf{x} = (x_1, \dots, x_n)^T$  uma amostra aleatória de tamanho  $n$  da distribuição MON com vetor de parâmetros  $\lambda = (p, \mu, \sigma)^T$ . Considere  $t_i = \frac{x_i - \mu}{\sigma}$ .

A função de verossimilhança é dada por

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^n f(t_i; \lambda) \\ &= \prod_{i=1}^n \frac{p\phi(t_i)}{\sigma[\Phi(t_i)(1-p) + p]^2}. \end{aligned} \quad (4.1)$$

Observe que  $\Phi(t_i)(1-p) + p$  é sempre positivo e diferente de um para  $p \neq 1$ . Para o caso  $p = 1$  recaímos na distribuição normal. Desta forma, o logaritmo pode ser aplicado e a função de log-verossimilhança é escrita como

$$\begin{aligned} l(\lambda) &= \log[L(\lambda)] = \sum_{i=1}^n \log \left\{ \frac{p\phi(t_i)}{\sigma[\Phi(t_i)(1-p) + p]^2} \right\} \\ &= n \log p - n \log \sigma + \sum_{i=1}^n \log \phi(t_i) - 2 \sum_{i=1}^n \log [\Phi(t_i)(1-p) + p]. \end{aligned} \quad (4.2)$$

O estimador de máxima verossimilhança  $\hat{\lambda} = (\hat{p}, \hat{\mu}, \hat{\sigma})$  do vetor de parâmetros  $\lambda = (p, \mu, \sigma)$  é obtido através da maximização da função de verossimilhança (4.2), solucionando o sistema de equações não-lineares  $U = \left( \frac{\partial l}{\partial p}, \frac{\partial l}{\partial \mu}, \frac{\partial l}{\partial \sigma} \right)^T = \mathbf{0}$ . O vetor  $U$ , chamado de função ou vetor escore, para a distribuição MON é dado pelas expressões

$$\begin{aligned} \frac{\partial l(\lambda)}{\partial p} &= \frac{n}{p} - \sum_{i=1}^n \frac{2[1 - \Phi(t_i)]}{\Phi(t_i)(1-p) + p}, \\ \frac{\partial l(\lambda)}{\partial \mu} &= \sum_{i=1}^n \frac{t_i}{\sigma} + \frac{2(1-p)}{\sigma} \sum_{i=1}^n \frac{\phi(t_i)}{\Phi(t_i)(1-p) + p}, \\ \frac{\partial l(\lambda)}{\partial \sigma} &= -\frac{n}{\sigma} + \sum_{i=1}^n \frac{t_i^2}{\sigma} + \frac{2(1-p)}{\sigma} \sum_{i=1}^n \frac{\phi(t_i)t_i}{\Phi(t_i)(1-p) + p}. \end{aligned}$$

As equações  $\frac{\partial l(\lambda)}{\partial p} = 0$ ,  $\frac{\partial l(\lambda)}{\partial \mu} = 0$ ,  $\frac{\partial l(\lambda)}{\partial \sigma} = 0$  não podem ser resolvidas analiticamente e desta forma métodos iterativos podem ser utilizados para encontrar estes estimadores numericamente. As Figuras 4.1, 4.2 e 4.3 representam, respectivamente, as superfícies da função de verossimilhança da distribuição MON fixando os parâmetros  $p, \mu$  e  $\sigma$ . Observa-se que existe apenas um ponto de máximo e sendo assim tem-se evidências numéricas que os máximos da função de verossimilhança podem ser encontrados computacionalmente, como por exemplo utilizando o algoritmo de Newton-Raphson descrito a seguir.

### 4.1.2 Algoritmo de Newton-Raphson

Expandindo a função escore  $U(\lambda)$  em série de Taylor até a primeira ordem em torno de um ponto inicial  $\lambda^{(0)}$ , obtém-se

$$U(\lambda) \cong U(\lambda^{(0)}) + (\lambda - \lambda^{(0)})U'(\lambda^{(0)}) = U(\lambda^{(0)}) + (\lambda - \lambda^{(0)})H(\lambda^{(0)}), \quad (4.3)$$

em que  $H(\lambda^{(0)})$  é a matriz de derivadas segundas da função log-verossimilhança avaliada em  $\lambda = \lambda^{(0)}$ . Fazendo  $U(\lambda) = 0$ , da equação (4.3) resulta o processo iterativo

$$\lambda^{(k+1)} = \lambda^{(k)} - H(\lambda^{(k)})^{-1}U(\lambda^{(k)}), \quad k = 0, 1, \dots,$$

que é repetido até que o processo se estabilize, ou seja, até que a distância entre  $\lambda^{(k+1)}$  e  $\lambda^{(k)}$  seja menor do que o nível desejado. Neste caso, o ponto  $\hat{\lambda}$  que estabiliza o processo é então o estimador de máxima verossimilhança.





Figura 4.1: Superfície da função de verossimilhança fixando  $p = 0,5$  e  $p = 2$



Figura 4.2: Superfície da função de verossimilhança fixando  $\mu = -2$  e  $\mu = 2$ .

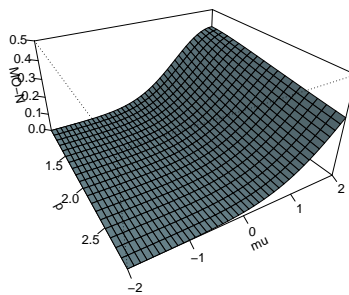


Figura 4.3: Superfície da função de verossimilhança fixando  $\sigma = 2$ .

### 4.1.3 Função escore e informação

Seja  $\mathbf{x} = (x_1, \dots, x_n)^T$  o valor observado de uma variável aleatória  $X = (X_1, \dots, X_n)^T$  caracterizada por uma função de densidade de probabilidade  $f(x; \theta)$  com  $\theta = (\theta_1, \dots, \theta_p)^T$  o vetor de parâmetros desconhecidos.

A primeira derivada da função log-verossimilhança  $U(\theta) = \frac{\partial l(\theta)}{\partial \theta}$  é chamada de função (vetor) escore com dimensão  $p \times 1$ . A matriz de informação esperada para  $\theta \in \mathbb{R}^p$   $J(\theta)$  obtida através dos dados  $\mathbf{x}$  é uma matriz quadrada  $p \times p$  definida por

$$K(\theta) = E[U(\theta)U(\theta)^T].$$

Quando as observações são independentes, a função escore e a informação esperada são as somas de contribuições individuais sobre  $\theta$ .

As condições de regularidade usadas na teoria assintótica estão relacionadas sobretudo com as propriedades diferenciais da função densidade de probabilidade e com a troca de diferenciação por integração. Essas condições são necessárias para provar as propriedades assintóticas dos estimadores de máxima verossimilhança, tais como consistência, unicidade, normalidade, eficiência e suficiência. As condições de regularidade a seguir foram retiradas de Casella e Benger (2011, cap. 10):

1. Observa-se  $X_1, \dots, X_n$  realizações independentes de uma variável aleatória  $X$ , com  $X_i \sim f(x|\theta)$ .
2. O parâmetro é identificável, ou seja, se  $\theta \neq \theta'$ , então  $f(x|\theta) \neq f(x|\theta')$ .
3. As densidades  $f(x|\theta)$  têm suporte comum e são diferenciáveis em  $\theta$ .
4. O espaço paramétrico  $\Theta$  contém um conjunto aberto no qual o verdadeiro valor do parâmetro  $\theta_0$  é ponto interior.
5. Para cada  $x$  a densidade  $f(x|\theta)$  é três vezes diferenciável com relação a  $\theta$ , cuja terceira derivada é contínua em  $\theta$ , e  $\int f(x|\theta)dx$  pode ser diferenciado três vezes.
6. Para qualquer  $\theta \in \Theta$ , existe um número positivo  $c$  e uma função  $M(x)$  de modo que

$$\left| \frac{\partial^3}{\partial \theta^3} \log f(x|\theta) \right| \leq M(x)$$

com  $\theta_0 - c < \theta < \theta_0 + c$  e  $E_{\theta_0}[M(x)] < \infty$ .

Sob estas condições, a esperança do vetor escore é zero e sua covariância é dada por

$$Cov(U) = E \left( -\frac{\partial U^T}{\partial \theta} \right) = E \left( -\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} \right) = K,$$

em que a matriz de primeiras derivadas do vetor escore com sinal negativo  $J = -\frac{\partial U^T}{\partial \theta} = -\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T}$  é denominada de matriz de informação observada. A matriz hessiana é dada por  $-J$  e sendo assim  $E(J) = K$ .

Casella e Berger (2011, cap.10) mostraram que se as condições de regularidade são satisfeitas então o estimador de máxima verossimilhança  $\hat{\theta}$  é consistente, assintoticamente eficiente e sua distribuição é aproximadamente normal  $p$ -variada:  $\hat{\theta} \xrightarrow{D} N_p(\theta, K(\theta)^{-1})$ .

#### 4.1.4 Estimação intervalar

Sejam  $\lambda = (p, \mu, \sigma)^T$  o vetor de parâmetros da distribuição MON e  $\hat{\lambda} = (\hat{p}, \hat{\mu}, \hat{\sigma})^T$  seu estimador de máxima verossimilhança. A matriz de informação (esperada) de Fisher é definida por

$$K(\lambda) = -E \left[ \frac{\partial^2 l(\lambda)}{\partial \lambda \partial \lambda^T} \right],$$

e sua inversa será representada por  $K(\lambda)^{-1}$ .

Utilizando o fato que a esperança do vetor escore é zero, das equações (4.3), encontra-se as relações

$$E[v_i(T_i)] = \frac{1}{2p}, \quad E[w_i(T_i)] = -\frac{E(T_i)}{2(1-p)} \quad \text{e} \quad E[w_i(T_i)T_i] = \frac{1 - E(T_i^2)}{2(1-p)}, \quad (4.4)$$

em que

$$v_i(t_i) = \frac{1 - \Phi(t_i)}{\Phi(t_i)(1-p) + p} \quad \text{e} \quad w_i(t_i) = \frac{\phi(t_i)}{\Phi(t_i)(1-p) + p}.$$

As segundas derivadas da função de log-verossimilhança são apresentadas no Apêndice B. A matriz de informação esperada de Fisher para os parâmetros  $(p, \mu, \sigma)$  da distribuição MON é dada por

$$K(\lambda) = \begin{bmatrix} k_{p,p} & k_{p,\mu} & k_{p,\sigma} \\ k_{\mu,p} & k_{\mu,\mu} & k_{\mu,\sigma} \\ k_{\sigma,p} & k_{\sigma,\mu} & k_{\sigma,\sigma} \end{bmatrix},$$

em que os elementos de cada entrada são

$$\begin{aligned} k_{p,p} &= \frac{n}{p^2} - 2nE[v^2(T)], \\ k_{p,\mu} &= -\frac{n}{\sigma(1-p)}E(T) + \frac{2n(1-p)}{\sigma}E[v(T)w(T)], \\ k_{p,\sigma} &= \frac{n}{\sigma(1-p)}[1 - E(T^2)] + \frac{2n(1-p)}{\sigma}E[Tv(T)w(T)], \\ k_{\mu,p} &= k_{p,\mu}, \\ k_{\mu,\mu} &= \frac{n}{\sigma^2}E(T^2) - \frac{2n(1-p)^2}{\sigma^2}E[w^2(T)], \\ k_{\mu,\sigma} &= \frac{n}{\sigma}E(T) - \frac{2n(1-p)}{\sigma^2} \{E[(T^2 - 1)w(T)] + E[Tw^2(T)]\}, \\ k_{\sigma,p} &= k_{p,\sigma}, \\ k_{\sigma,\mu} &= k_{\mu,\sigma}, \\ k_{\sigma,\sigma} &= \frac{n}{\sigma^2} + \frac{n}{\sigma^2}E(T^2) - \frac{2n(1-p)}{\sigma^2}E \{T^2w(T)[T + w(T)]\}, \end{aligned}$$

em que  $v(T) = v_i(T_i)$  para todo  $i = 1, 2, \dots, n$ . A esperança  $E[v^2(T)]$  é uma integral do tipo beta e portanto é finita. A proposição a seguir mostra que a matriz de informação esperada de Fisher está bem definida, ou seja, as outras esperanças que aparecem na matriz  $K(\lambda)$  são finitas.

**Proposição 2.** *Seja  $f$  a função densidade de probabilidade da distribuição MON com parâmetros  $p, \mu, \sigma$  e  $t = \frac{x - \mu}{\sigma}$ . As integrais da forma*

$$\int_{-\infty}^{+\infty} \phi^k(t) t^m f(t) dt = E[T^m \phi^k(T)]$$

são finitas para todo  $k, m \in \mathbb{N}$ .

*Demonstração.* Pelo Teorema 4 para  $p < 1$  tem-se as desigualdades

$$E[T^m \phi^k(T)] = \int_{-\infty}^{+\infty} \phi^k(t) t^m \frac{p\phi(x)}{\{\Phi(x) + p[1 - \Phi(x)]\}^2} dt \quad (4.5)$$

$$< \int_{-\infty}^{+\infty} \frac{\phi^{k+1}(t) t^m}{p} dt = \int_{-\infty}^{+\infty} \frac{e^{-\frac{t^2(k+1)}{2}} t^m}{p(\sqrt{2\pi})^{(k+1)}} dt. \quad (4.6)$$

Tem-se que

$$\int_{-\infty}^{+\infty} \frac{e^{-\frac{t^2(k+1)}{2}} t^m}{p(\sqrt{2\pi})^{(k+1)}} dt = \frac{M_m(t)}{p\sqrt{k+1}}$$

em que  $M_m(t)$  representa o  $m$ -ésimo momento de uma distribuição normal com média 0 e variância  $\frac{1}{k+1}$ . Desta forma,  $E[T^m \phi^k(T)]$  é finita. Para  $p \geq 1$  a demonstração é análoga.  $\square$

Para amostras grandes e sob as condições de regularidade apresentadas para a função de verossimilhança, se  $\hat{\lambda}$  é um estimador de máxima verossimilhança da distribuição MON então

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N_3(0, k(\lambda)^{-1}),$$

com  $\xrightarrow{D}$  denotando a convergência em distribuição. Em outras palavras, a distribuição de  $\hat{\lambda}$  é assintoticamente normal multivariada com média  $\lambda$  e matriz de covariância  $K(\lambda)^{-1} = n^{-1}k(\lambda)^{-1}$ .

Sendo assim, intervalos de confiança assintóticos para  $p, \mu$  e  $\sigma$  podem ser construídos como

$$\begin{aligned} IC(p, 1 - \gamma) &= \left( \hat{p} - z_{\gamma/2} \sqrt{k(\hat{\lambda})^{(p,p)}}, \quad \hat{p} + z_{\gamma/2} \sqrt{k(\hat{\lambda})^{(p,p)}} \right) \\ IC(\mu, 1 - \gamma) &= \left( \hat{\mu} - z_{\gamma/2} \sqrt{k(\hat{\lambda})^{(\mu,\mu)}}, \quad \hat{\mu} + z_{\gamma/2} \sqrt{k(\hat{\lambda})^{(\mu,\mu)}} \right) \\ IC(\sigma, 1 - \gamma) &= \left( \hat{\sigma} - z_{\gamma/2} \sqrt{k(\hat{\lambda})^{(\sigma,\sigma)}}, \quad \hat{\sigma} + z_{\gamma/2} \sqrt{k(\hat{\lambda})^{(\sigma,\sigma)}} \right), \end{aligned}$$

em que  $k(\hat{\lambda})^{(p,p)}$ ,  $k(\hat{\lambda})^{(\mu,\mu)}$  e  $k(\hat{\lambda})^{(\sigma,\sigma)}$  são as variâncias assintóticas de  $\hat{p}$ ,  $\hat{\mu}$  e  $\hat{\sigma}$  dadas pelos primeiro, segundo e terceiro elementos da diagonal da matriz  $n^{-1}K(\lambda)^{-1}$ , respectivamente. Os valores  $z_{\gamma/2}$  representam o  $(1 - \gamma)$ -ésimo quantil da distribuição. Os intervalos de confiança possuem comprimento  $100(1 - \gamma)\%$ .

## 4.2 Testes de hipóteses

Nesta seção é feita uma breve introdução dos testes paramétricos e não paramétricos. Na parte paramétrica os testes da razão de verossimilhança e da razão de verossimilhança generalizada são apresentados. São considerados na abordagem não paramétrica o teste de bondade do ajuste feito através das estatísticas de  $W^*$  e  $A^*$ , o teste de Kolmogorov-Smirnov e alguns critérios de seleção de modelos.

Para obter as propriedades assintóticas dos estimadores de máxima verossimilhança da distribuição MON, as condições de regularidade apresentadas devem ser cumpridas. Porém, não foi possível verificar todas essas condições analiticamente e sendo assim métodos computacionais foram utilizados nos quais evidências numéricas mostraram que a distribuição em estudo satisfazem todas as condições de regularidade. Seja  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  uma amostra independente de tamanho  $n$  de uma variável aleatória com distribuição MON, com vetor de parâmetros  $\lambda = (p, \mu, \sigma)$ . Tem-se que

1. A distribuição MON é identificável, ou seja, se  $\lambda \neq \lambda'$  então  $F_\lambda \neq F_{\lambda'}$ ;
2. A distribuição MON tem suporte independente do parâmetro  $\lambda$  e sua função densidade de probabilidade é diferenciável em  $\lambda$ ;
3. O espaço paramétrico contém um conjunto aberto no qual o verdadeiro valor do parâmetro  $\lambda_0$  é um ponto interior;
4. A matriz de informação de Fisher é finita e definida positiva em uma vizinhança aberta de  $\lambda_0$ ;
5. As terceiras derivadas da função log-verossimilhança são limitadas por uma função integrável de  $X \sim MON$  com esperança finita.

### 4.2.1 Teste da razão verossimilhança

O teste da razão de verossimilhança (TRV) pode ser feito para testar dois modelos quando um dos modelos é submodelo do outro. O TRV é feito através da estatística de razão de verossimilhança definida por

$$LR = 2[l(\hat{\lambda}) - l(\tilde{\lambda})], \quad (4.7)$$

em que  $\hat{\lambda}$  e  $\tilde{\lambda}$  são, respectivamente, os estimadores de máxima verossimilhança (EMVs) para o modelo representado na hipótese alternativa  $H_A$  e para o modelo representado na hipótese nula  $H_0$ . Tem-se que sob  $H_0$ , a estatística  $LR$  converge em distribuição para uma distribuição *chi-quadrado* com  $q$  graus de liberdade, com  $q$  sendo a diferença entre o número de parâmetros

dos modelos. Matematicamente,  $LR \xrightarrow{d} \chi_p^2$ . Rejeita-se a hipótese nula se  $LR > p_{1-\alpha}$ , em que  $p_{1-\alpha}$  representa o quantil  $1 - \alpha$  da distribuição *chi-quadrado* com  $q$  graus de liberdade e  $\alpha$  é o nível de significância desejado.

### 4.2.2 Teste da razão de verossimilhança generalizada

Para testar modelos *non-nested* (não aninhados) o teste da razão de verossimilhança descrito anteriormente não é mais apropriado. Vuong (1989) apresentou a seguinte definição para modelos *non-nested*:

**Definição 3.** *Dois modelos  $F_\theta$  e  $G_\gamma$ , são estritamente non-nested se, e somente se,  $F_\theta \cap G_\gamma = \emptyset$ .*

Considere a escolha entre dois modelos  $F_\theta$  e  $G_\gamma$  com funções de densidades de probabilidade correspondentes  $f(y_i|x_i, \theta)$  e  $g(y_i|x_i, \gamma)$ , respectivamente. A estatística da razão de verossimilhança generalizada para testar a hipótese nula que os modelos são equivalentes é definida como

$$T_{F,G} = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\gamma})} \right\} \times \left\{ \frac{1}{n} \sum_{i=1}^n \left( \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\gamma})} \right)^2 - \left( \frac{1}{n} \sum_{i=1}^n \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\gamma})} \right)^2 \right\}^{-\frac{1}{2}}. \quad (4.8)$$

A estatística  $T_{F,G}$  converge em distribuição sob a hipótese nula de equivalência dos modelos, para a distribuição normal padrão. Seja  $c$  um valor crítico da distribuição normal padrão para o nível de significância desejado. Se o valor da estatística  $T_{F,G}$  for maior que  $c$  então rejeita-se a hipótese nula que os modelos são equivalentes em favor do modelo  $F_\theta$  ser melhor do que  $G_\gamma$ . Se  $T_{F,G}$  for menor do que  $-c$ , então a hipótese nula é rejeitada em favor do modelo  $G_\gamma$  ser melhor do que o modelo  $F_\theta$ . Por fim, se  $|T_{F,G}| \leq c$ , então os modelos são equivalentes com base nos dados. Para mais detalhes ver Vuong (1989).

### 4.2.3 Teste bondade do ajuste

Para verificar qual distribuição entre as distribuições testadas se ajusta melhor a um determinado conjunto de dados, o teste de bondade de ajuste pode ser realizado. As estatísticas de Crámer-von Mises ( $W^*$ ), Anderson-Darling ( $A^*$ ) e o teste de Kolmogorov-Smirnov serão utilizados. Em geral, o modelo que apresentar menores valores para ( $W^*$ ), ( $A^*$ ) e para a estatística do teste de Kolmogorov-Smirnov, será considerado o melhor para o ajuste do conjunto de dados.

Para testar  $H_0: X_1, X_2, \dots, X_n$  é uma variável aleatória de uma distribuição contínua com função de distribuição acumulada  $F(x; \lambda)$ , em que a forma  $F$  é conhecida mas vetor de parâmetros  $\lambda$  é desconhecido, as estatísticas  $W^*$  e  $A^*$  são calculadas seguindo os procedimentos descritos a seguir:

1. Estime  $\hat{\lambda}$  um estimador eficiente de  $\lambda$  e calcule  $v_i = F_i(x_i; \hat{\lambda})$ , em que os  $x_i$ 's são colocados em ordem crescente;
2. Calcule  $y_i = \Phi^{-1}(v_i)$ , em que  $\Phi(\cdot)$  é a função de distribuição acumulada da distribuição normal padrão e  $\Phi^{-1}(\cdot)$  sua inversa;

3. Calcule  $u_i = \Phi\left(\frac{y_i - \bar{y}}{s_y}\right)$ , em que  $\bar{y} = n^{-1} \sum_{i=1}^n y_i$  e  $s_y^2 = (n-1)^{-1} \sum_{i=1}^n (y_i - \bar{y})^2$ ;

4. Calcule

$$W^2 = \sum_{i=1}^n \left\{ u_i - \frac{(2i-1)}{2n} \right\}^2 + \frac{1}{12n}$$

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n [(2i-1) \log(u_i) + (2n+1-2i) \log(1-u_i)];$$

5. Modifique  $W^2$  por  $W^* = W^2(1 + 0.5/n)$  e  $A^* = A^2(1 + 0.75/n + 2.25/n^2)$ , em que  $n$  é o tamanho da amostra.

A hipótese nula é rejeitada ao nível de significância  $\alpha$  se as estatísticas modificadas  $W^*$  e  $A^*$  excedem os limites superiores de significância apresentados em uma tabela dada por Chen e Balakrishnan (1995). Em geral, quanto menor as estatísticas  $W^*$  e  $A^*$ , melhor é o ajuste dos dados. Para mais detalhes ver os últimos autores.

O teste de Kolmogorov-Smirnov (KS) verifica se uma amostra pode ser considerada como proveniente de uma população com uma determinada função de distribuição acumulada. As hipóteses do teste são

$$H_0 : \text{a população tem uma determinada distribuição } D$$

$$H_1 : \text{a população não tem a distribuição } D.$$

O procedimento para calcular a estatística  $D_n$  do teste  $KS$ , a qual é definida como o máximo valor absoluto da diferença entre as funções de distribuição acumulada empírica e estimada, e o correspondente p-valor são descritos a seguir.

1. Ordenar os  $n$  valores dos dados  $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ ;
2. Calcular a estatística

$$D_n = \max_{j=1, \dots, n} \left\{ \frac{j}{n} - \hat{F}(y_{(j)}), \hat{F}(y_{(j)}) - \frac{j-1}{n} \right\},$$

em que  $\hat{F}(\cdot)$  é a função de distribuição acumulada  $D$  sob a hipótese nula;

3. Gerar números aleatórios de  $n$  distribuição  $U(0, 1)$  e ordenar,  $u_{(1)}^{(i)} \leq u_{(2)}^{(i)} \leq \dots \leq u_{(n)}^{(i)}$ ;
4. Calcular

$$d^{(i)} = \max_{j=1, \dots, n} \left\{ \frac{j}{n} - u_{(j)}^{(i)}, u_{(j)}^{(i)} - \frac{j-1}{n} \right\};$$

5. Seja  $I_i = 1$  se  $d^{(i)} \geq D_n$  e caso contrário zero. Repetir os passos 3 e 4  $N$  vezes,  $I_1, \dots, I_N$ . O p-valor é estimado por  $\sum_{i=1}^N \frac{I_i}{N}$ .

A hipótese nula é rejeitada se o p-valor for menor que o nível de significância utilizado para o teste. Para informações adicionais ver Lin et al. (2007).

## 4.3 Critérios de seleção de modelos

### 4.3.1 Critério de informação de Akaike - AIC

O critério de informação de Akaike (AIC) (Akaike, 1973), é uma ferramenta para seleção de modelos e foi desenvolvido a partir da distância de Kullback-Leibler, definida na expressão (3.30). O AIC é definido como

$$AIC = -2l(\hat{\lambda}) + 2k,$$

em que  $l(\hat{\lambda})$  é a função log-verossimilhança maximizada e  $k$  é o número de parâmetros do modelo. O termo acrescentado na função de log-verossimilhança tem a finalidade de corrigir o viés resultante da comparação de modelos com diferentes números de parâmetros. Dentre os modelos considerados o que possuir menor valor de AIC será considerado o modelo de melhor ajuste.

### 4.3.2 Critério de informação de Akaike corrigido - AICc

Segundo Sugiura (1978), o critério de informação de Akaike pode ser ruim se existem muitos parâmetros quando comparados com o tamanho da amostra. Desta forma ele propôs uma correção do viés do AIC através da expressão

$$AICc = AIC + \frac{2k(k+1)}{n-k-1},$$

em que  $n$  é tamanho da amostra e  $k$  o número de parâmetros do modelo. Segundo Burnham e Anderson (2004) sua utilização é indicada quando a  $\frac{n}{k} < 40$ . Se a razão for suficientemente grande os critérios AIC e AICc apresentam resultados semelhantes. O modelo com menor AICc é selecionado.

### 4.3.3 Critério Bayesiano de Schwarz - BIC

O critério de informação de Schwarz (BIC) (Schwarz, 1978), pode ser expresso por

$$BIC = -2l(\hat{\lambda}) + 2k \ln(n),$$

em que  $l(\hat{\lambda})$  é função log-verossimilhança generalizada,  $n$  o número de observações e  $k$  o número de parâmetros. Uma característica desse critério é penalizar os modelos com mais parâmetros. Um modelo com menor BIC é considerado o melhor modelo.

O próximo capítulo apresenta um estudo de simulação e um ajuste da distribuição MON para um banco de dados reais. Uma comparação desta distribuição é feita com as distribuições normal, Cauchy, logística e  $t$ -Student.



# Capítulo 5

## Simulação e Aplicação

Neste capítulo, apresenta-se um estudo de simulação em que amostras da distribuição MON são geradas. Para ilustrar a utilidade do modelo MON apresentado no Capítulo 3, um conjunto de dados reais será analisado para comparar o modelo em estudo com distribuições usuais que modelam dados com caudas pesadas e curtose maior do que a da distribuição normal.

### 5.1 Simulação

Para se obter uma variável aleatória com distribuição MON é suficiente gerar um número aleatório  $u$  da distribuição Uniforme(0,1) e em seguida efetuar o quantil de ordem  $u$  da distribuição MON:

$$x = \Phi^{-1} \left[ \frac{pu}{1 - (1-p)u} \right] \sigma + \mu.$$

A Figura 5.1 mostra alguns histogramas para dados simulados de uma distribuição MON com a linha contínua representando a função de densidade estimada da frequência dos dados, dada pelo comando *density* do R em sua versão 2.15.2..

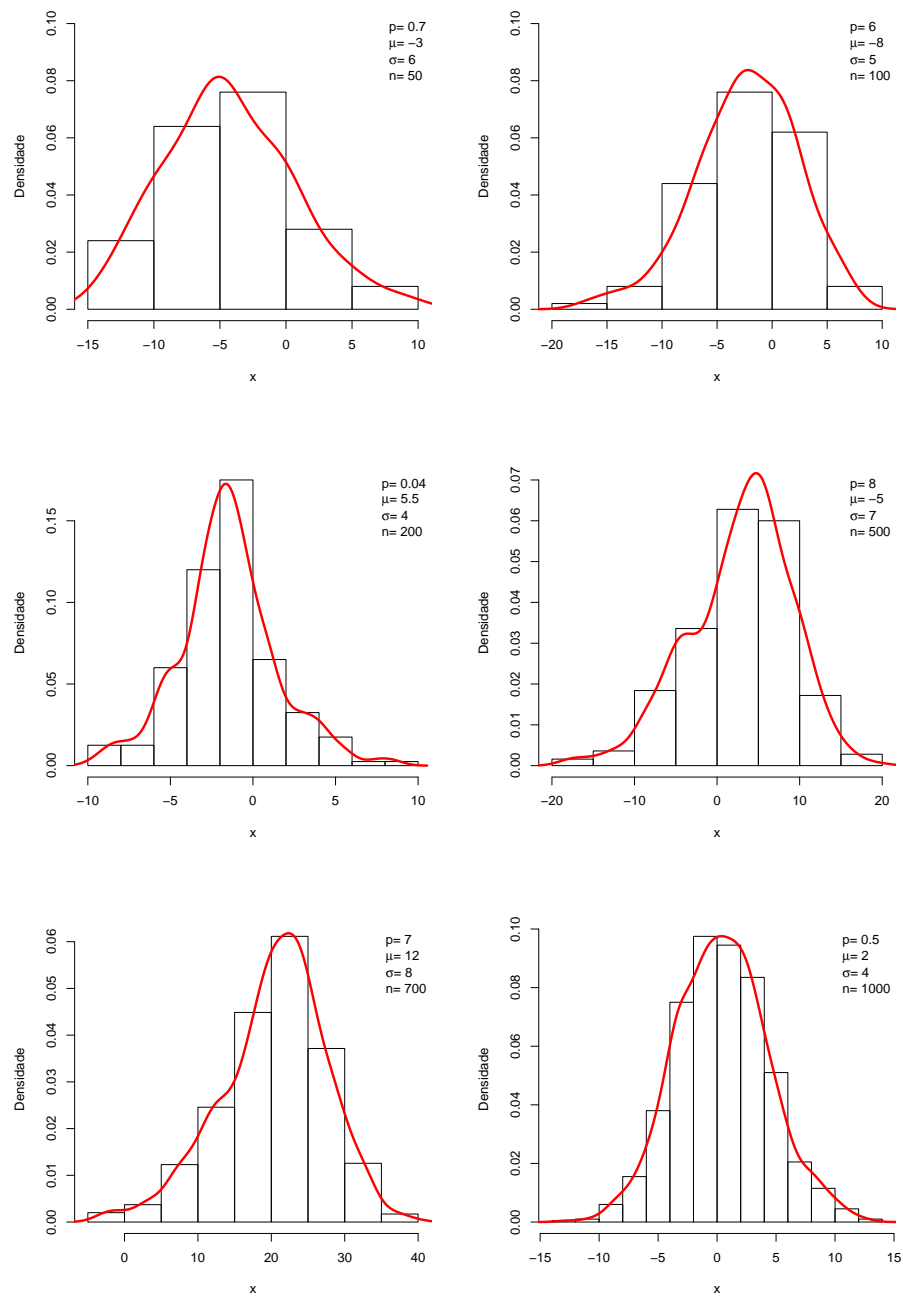


Figura 5.1: Gráficos da função densidade exata da distribuição MON com histogramas para dados simulados.

## 5.2 Aplicação

Nesta seção apresenta-se uma análise de um conjunto de dados reais para mostrar a flexibilidade da distribuição MON. Foram comparadas com a distribuição MON as distribuições normal,

logística, Cauchy e t-Student. O banco de dados utilizado foi originalmente considerado por Badar e Priest (1982) e representa 63 valores de resistência para amostras de fibras de carbono. Os dados são apresentados a seguir.

1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474
2.518	2.522	2.525	2.532	2.575	2.614	2.616	2.618	2.624	2.659	2.675	2.738
2.740	2.856	2.917	2.928	2.937	2.937	2.977	2.996	3.030	3.125	3.139	3.145
3.220	3.223	3.235	3.243	3.264	3.272	3.294	3.332	3.346	3.377	3.408	3.435
3.493	3.501	3.537	3.554	3.562	3.628	3.852	3.871	3.886	3.971	4.024	4.027
4.225	4.395	5.020.									

A Tabela 5.1, lista as estimativas de máxima verossimilhança (EMVs) com os correspondentes erros padrões entre parênteses, o logaritmo da verossimilhança e as estatísticas AIC, BIC e AICc. Os resultados encontrados indicam que o modelo MON teve os menores valores

Tabela 5.1: EMVs dos parâmetros dos modelos MON, normal, logística, t-Student e Cauchy para dados de fibras de carbono, correspondentes erros-padrão (entre parênteses), log-verossimilhança e valores dos AIC, BIC, AICc.

Distribuição	Estimativas (Erro)	Log-vero	AIC	BIC	AICc
MON	$\hat{p} = 0.1786$ (0.2009)				
	$\hat{\mu} = 3.6647$ (0.4503)	-57.2830	120.5660	126.9954	120.9728
	$\hat{\sigma} = 0.6519$ (0.0867)				
Normal	$\hat{\mu} = 3.0596$ (0.0776)	-58.8664	121.7329	126.0192	121.9329
	$\hat{\sigma} = 0.6159$ (0.0549)				
Logística	$\hat{\mu} = 3.0244$ (0.0780)	-59.3295	122.6591	126.9453	122.8591
	$\hat{\sigma} = 0.3525$ (0.0365)				
t-Student	$\hat{\mu} = 3.0475$ (0.0773)				
	$\hat{\tau} = 0.6150$ (0.0556)	- 58.7922	123.5845	130.0139	123.9913
	$\hat{\nu} = 28.7203$ (1.2437)				
Cauchy	$\hat{\mu} = 2.9902$ (0.0916)	-71.4473	146.8945	151.1808	147.0945
	$\hat{\lambda} = 0.4033$ (0.0611)				

para AIC e AICc dentre os modelos considerados, e então poderia ser escolhido como o melhor modelo. Em relação ao BIC, o da distribuição MON foi menor que o das distribuições t-Student e Cauchy, sendo considerado o melhor modelo quando comparado com as mesmas. Porém, quando comparado com as distribuições normal e logística, o BIC da MON foi maior do que o encontrado para estas distribuições. A diferença foi pequena (na primeira e na segunda casas decimais). Este resultado era de se esperar pois o BIC penaliza as distribuições com mais parâmetros.

A Figura 5.2 ilustra o histograma dos dados com as densidades dos modelos comparados substituindo os valores encontrados para as estimativas de máxima verossimilhança. Nota-se que a distribuição MON proporciona um ajuste satisfatório aos dados.

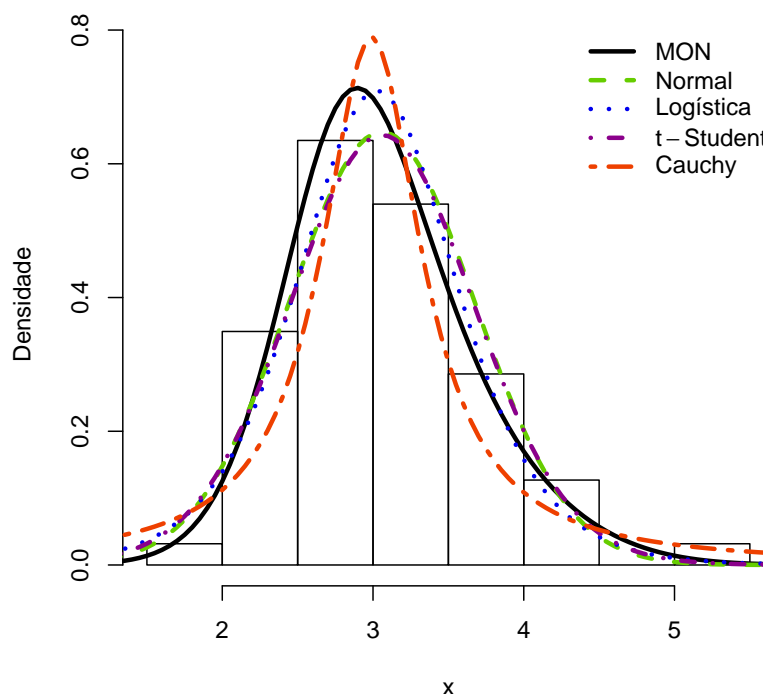


Figura 5.2: Histograma do conjunto de dados de fibras de carbono com as funções de densidades de probabilidades dos modelos MON, normal, logística, t-Student e Cauchy.

Para verificar se o modelo MON se ajusta melhor ao banco de dados do que o modelo normal pode-se testar as hipóteses  $H_0: p = 1$  versus  $H_1: H_0$  não é verdadeira, equivalentemente,  $H_0$ : normal  $\times H_1$ : MON. A estatística  $LR$  encontrada para testar estas hipóteses foi 3.1668 (p-valor  $< 0.10$ ). Desta forma, rejeita-se a hipótese nula em favor da distribuição MON ao nível de significância de 10%, ou seja, o modelo MON é significativamente melhor que o modelo normal baseado na estatística da razão de verossimilhança.

Testou-se, via estatística dada na expressão (4.8) para modelos *non-nested* (“não aninhados”), o modelo MON com os modelos logística, Cauchy e t-Student. A estatística encontrada para testar  $H_0$ : logística versus  $H_A$ : MON foi de 0.2588. Desta forma, a hipótese nula não é rejeitada ao nível de significância de 5%, ou seja, os modelos MON e logística são equivalentes para o conjunto de dados. Para testar as hipóteses  $H_0$ : Cauchy versus  $H_A$ : MON, a estatística encontrada foi de 0.8951. Ao nível de 5% de significância a hipótese nula não é rejeitada e portanto, os modelos MON e Cauchy são equivalentes para o ajuste dos dados. Por fim, a estatística encontrada para testar as hipóteses  $H_0$ : t-Student versus  $H_A$ : MON foi de 0.3827. Aceita-se a hipótese nula o que significa que os modelos MON e t-Student são significativamente equivalentes para o ajuste do conjunto de dados.

Os valores para as estatísticas ( $W^*$ ) e ( $A^*$ ) encontrados para todos os modelos estão listados na Tabela 5.2. Pela tabela de valores críticos para  $W^*$  e  $A^*$ , apresentada em Chen e Balakrishnan (1995), todas as distribuições testadas, exceto a distribuição Cauchy, se adequam aos

dados. Entretanto, as estatísticas encontradas para o modelo MON foram as menores e desta forma tem-se que este modelo se encaixa melhor ao conjunto de dados do que os outros modelos testados.

Tabela 5.2: Teste bondade do ajuste

Distribuição	Estatística $W^*$	Estatística $A^*$
MON	0.0051	0.1734
Normal	0.0069	0.3311
Logística	0.0094	0.3966
Cauchy	0.0657	2.4666
t-Student	0.0071	0.3281

A Tabela 5.3 mostra os resultados do teste de Kolmogorov-Smirnov (KS) para os cinco modelos ajustados. Dentre os modelos comparados, o melhor é o modelo MON com o maior p-valor igual a 0.7130. Isto é, o teste KS sugere que os dados para fibra de carbono seguem uma distribuição MON.

Tabela 5.3: Teste de Kolmogorov-Smirnov para as distribuições testadas

Distribuição	$D_n$	p-valor
MON	0.0868	0.7130
Normal	0.0987	0.5210
Logística	0.0944	0.6000
Cauchy	0.1239	0.2570
t-Student	0.0903	0.6810



# Capítulo 6

## Conclusões e discussões

Neste trabalho, estudou-se a distribuição Marshall-Olkin normal, a qual é obtida através da transformação da família definida em Marshall e Olkin (1997). Expansões em séries convergentes foram calculadas para as funções de distribuição acumulada e de densidade de probabilidade, expressões para alguns momentos foram deduzidas com os primeiros termos encontrados explicitamente. As entropias de Shannon e de Rényi foram calculadas e também estudou-se a identificabilidade do modelo. Na parte inferencial, expressões para os estimadores dos parâmetros de uma amostra independente da distribuição Marshall-Olkin normal foram obtidas através do método de máxima verossimilhança. A matriz de informação esperada de Fisher foi apresentada e intervalos de confiança foram construídos.

Investigou-se o comportamento do parâmetro adicional como uma medida de assimetria e foi verificado que o resultado de Rubio e Stell (2012) é verdadeiro, ou seja, o parâmetro adicional da MON não modela assimetria forte ou moderada. Para constatar esse resultado, um estudo com o mesmo conjunto de dados utilizado por García et al. (2010) foi realizado (os resultados se encontram no Apêndice C). Desta forma, verificou-se que a distribuição MON não compete com distribuições assimétricas como afirmaram García et al. (2010).

A distribuição MON possui curva de densidade de probabilidade deslocada (para a direita ou esquerda) e é mais afunilada (alta) quando comparada com a distribuição normal, deduziu-se que esta distribuição poderia competir com distribuições usuais que modelam dados com caudas pesadas e curtose maior do que a da distribuição normal. Para verificar esse pensamento intuitivo, utilizou-se um banco de dados reais ajustando-se além da distribuição MON as distribuições normal, logística, Cauchy e t-Student. As distribuições ajustadas foram comparadas utilizando-se as estatísticas AIC, BIC, CAIC, os testes da razão de verossimilhança, da razão de verossimilhança generalizada e de bondade de ajuste. De acordo com os critérios citados, a MON se ajustou melhor aos dados do que as outras distribuições testadas.

Sendo assim, a distribuição Marshall-Olkin normal pode modelar fenômenos com caudas pesadas, curtose maior que a da distribuição normal e um leve deslocamento, ou seja, modelar dados ligeiramente afastados da normalidade.





# Referências Bibliográficas

Alice, T. e Jose, K.K. *Bivariate semi-Pareto minification processes*. *Metrica*, 59:305-313, 2004.

Alice, T. e Jose, K.K. *Marshall-Olkin Pareto processes*. *Far East Journal of Theoretical Statistics*, 2:117-132, 2003.

Arnold, B.C. e Groeneveld, R.A. *Measuring skewness with respect to the mode*. *The American Statistician*, 49:34-38, 2006.

Bose, R.C. e Gupta, S.S. *Moments of order statistics from a normal population*. *Biometrika*, 46:433-440, 1959.

Bowden R. *The theory of parametric identification*. *Econometrica*, 41:1069-1074, 1973.

Caroni C. *Testing for the Marshall-Olkin extended form of the Weibull distribution*. *Stat Comput*, 51:325-336, 2010.

Castellares, F. e Lemonte, A. *On the Marshall-Olkin extended distributions*. Submetido para publicação, 2013.

Chen, G. e Balakrishnan, N. *A general purpose approximate goodness-of-fit test*. *Jornal of Quality Technology*, 27:154-161, 1995.

Economou, P. e Caroni, C. *Parametric proportional odds frailty models*. *Commun Stat Simulat*, 36:579-592, 2007.

Ferreira, D.F. *Estatística Básica*. UFLA, 2005.

García, V.J. , Gómez-Déniz, E. e Vázquez-Polo, F.J. *A new skew generalization of the normal distribution: properties and applications*. *Computational Statistics and Data Analysis*, 54:2021-2034, 2010.

Ghitany, M.E. *Marshall-Olkin extended Pareto distribution and its application*. *International Journal of Applied Mathematics*, 18:17-32, 2005.

Ghitany, M.E. , Al-Awadhi, F.A. e Alkhalfan, L.A. *Marshall-Olkin extended Lomax distribution and its application to censored data*. *Communication in Statistics - Theory and Methods*,

36:1855-1866, 2007.

Ghitany, M.E. , Al-Hussaini, E.K. e AlJarallah, R.A. *Marshall-Olkin extended Weibull distribution and its application to censored data*. Journal of Applied Statistics, 32:1025- 1034, 2005.

Ghitany, M.E. e Kotz, S. *Reliability properties of extended linear failure-rate distributions*. Probab Eng Inform Sc, 21:441-450, 2007.

Gómez-Déniz E. *Another generalization of the geometric distribution*. Test, 19:399-415, 2010.

Gupta, R.D. e Kundu, D. *Generalized exponential distributions*. Aust N Z J Stat, 41:173- 188, 1999.

Gupta, R.D. e Peng, C. *Estimating reliability in proportional odds ratio models*. Comput Stat Data An, 53:1495-1510, 2009.

Gupta, R.C., Lvin, S. e Peng, C. *Estimating turning points of the failure rate of the extended Weibull distribution*. Comput Stat Data An, 54:924-934, 2010.

Jose, K.K. e Uma, P. *On Marshall-Olkin Mittag-Leffler distributions and processes*. Far East Journal of Theoretical Statistics, 28:189-199, 2010.

Jose, K.K., Naik, S.R. e Ristic, M.M. *Marshall-Olkin  $q$ -Weibull distribution and max-min process*. Statistical Papers, 51:837-851, 2010.

Jose K.K., Ristic M.M. e Ancy Joseph. *Marshall-Olkin bivariate Weibull distributions and processes*. Statistical Papers (online version), 001 10.1007/S00362-009-0287-8, 2009.

Kirman S. e Gupta R.C. *On the proportional odds model in survival analysis*. Ann.Inst.Statist.Math, 53:203-216, 2001.

Kotz, S., Balakrishnan, N. e Johnson, N. *Continuous univariate distributions*, volume 1-2. Wiley series in probability and statistics, Chichester, 1994.

Kullback, S. e Leibler, R.A. *On information and sufficiency*. The Annals of Mathematical Statistics,, 22:79-86, 1951.

Lam, K.F. e Leung, T.L. *Marginal likelihood estimation for proportional odds models with right censored data*. Lifetime Data Anal, 7:39-54, 2001.

Lin, T.I. , Lee, J.C. e Hsieh, W.J. *Robust mixture modeling using the skew  $t$  distribution*. Stat Comput, 17: 81-92, 2007.

Jones, H.L. *Exact lower moments of order statistics in small samples from a normal distribution*. The Annals of Mathematical Statistics, 19:270-273, 1948.

Maiti, S.S. e Dey, M. *Tilted normal distribution and its survival properties*. Journal of Data Science, 10: 225-240, 2012.

Mardia, K. *Statistics of Directional Data*. Academic Press, 1972.

Mardia, K. e Jupp, P. *Directional statistics*. Wiley series in probability and statistics, Chichester, 2000.

Marshall, A.W. e Olkin, I. *A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families*. Biometrika, 84:641-652.

Ord, J. *Families of Frequency Distributions*. Griffin's Statistical Monographs and Courses, 1972.

Patil, G. , Kotz, S. e Ord, J. *A modern course on statistical distributions in scientific work : Vol. 1: Models and Structures Vol. 2: Model Building and Model Selection Vol. 3: Characterizations and Applications*, volume 79. Proceedings of the NATO Serie C. Advanced Study Institute held at the University of Calgary, Canada, 1975.

Patil, G., Taillie, C. e Baldessari, B. *A modern course on statistical distributions in scientific work : Vol. 4: Models, Structures, and Characterizations Vol. 5: Inferential Problems and Properties Vol. 6: Applications in Physical, Social, and Life Sciences.*, volume 17. Proceedings of the NATO Serie C. Advanced Study Institute held at the University of Calgary, Canada, 1981.

Rao G.S., Ghitany M.E. e Kantam R.R.L. *Reliability test plans for Marshall-Olkin extended exponential distribution*. Appl Math Sci, 55:2745-2755, 2009.

Ristic, M.M., Jose, K.K. e Ancy, J. *A Marshall-Olkin gamma distribution and minification process*. STARS: Stress and Anxiety Research Society, 11:107-117, 2007.

Roberts, H.V. *Data analysis for managers with minitab*. The American Statistician, 41:270-278, 1987.

Rubio, F.J. e Steel, M.F.J. *On the Marshall-Olkin transformation as a skewing mechanism*. Computacional Statistics and Data Analysis, 56:2251-2257, 2012.

Severini, T.A. *Elements of Distribution Theory*. Cambridge University Press, 2005.

Zhang, T. e Xie, M. *Failure data analysis with extended Weibull distribution*. Commun Stat Simulat, 36:579592, 2007.



# Apêndice A

## A.1 Terceiro momento

Neste apêndice os primeiros termos da expansão que representa os momentos de terceira e quarta ordem são calculados. Os resultados encontrados foram baseados no artigo de Bose e Gupta (1959).

O terceiro momento para  $X \sim MON(p, 0, 1)$  é dado por

$$E(X^3) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[1-\Phi(x)]^n dx, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)\Phi(x)^n dx, & \text{se } q^{-1} = p > 1. \end{cases} \quad (\text{A.1})$$

Para  $q^{-1} = p > 1$  tem-se

$$\begin{aligned} E(X^3) &= \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[\Phi(x)]^n dx \\ &= q \int_{-\infty}^{\infty} x^3 \phi(x) dx + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[\Phi(x)]^n dx \\ &= q \int_{-\infty}^{\infty} x^3 \phi(x) dx + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[\Phi(x)]^n dx \\ &= \sum_{n=1}^{\infty} q(1-q)^n \mu'_3(n+1, n+1), \end{aligned}$$

pois  $\int_{-\infty}^{\infty} x^3 \phi(x) dx$  é terceiro momento de uma distribuição normal padrão e é igual a zero, e

$$\mu'_3(n+1, n+1) = (n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[\Phi(x)]^n dx$$

é o terceiro momento do máximo de  $n+1$  variáveis com distribuição normal padrão. Bose e Gupta (1959) mostraram que

$$\mu'_3(n+1, n+1) = \frac{5}{2} \mu'_1(n+1, n+1) + \binom{n+1}{4} \int_{-\infty}^{\infty} [\Phi(x)]^{n-3} \frac{e^{-2x^2}}{\pi^2} dx. \quad (\text{A.2})$$

Substituindo alguns valores de  $n$  na expressão (A.2) tem-se

$$\begin{aligned}\mu'_3(2, 2) &= \frac{5}{2}\mu'_1(2, 2) = \frac{5\sqrt{\pi}}{2}, \\ \mu'_3(3, 3) &= \frac{5}{2}\mu'_1(3, 3) = \frac{15}{4\sqrt{\pi}}, \\ \mu'_3(4, 4) &= \frac{5}{2}\mu'_1(4, 4) + \frac{1}{\pi^2} \int_{-\infty}^{\infty} e^{-2x^2} dx \\ &= \frac{15I_2(1)}{\pi} + \frac{\sqrt{2\pi}}{2\pi^2} \\ &= \frac{15 \arctan \sqrt{2}}{\pi\sqrt{\pi}} + \frac{\sqrt{2\pi}}{2\pi^2}.\end{aligned}$$

O valor de  $\mu'_3(5, 5)$  foi apresentado por Bose e Gupta (1959):

$$\mu'_3(5, 5) = \frac{75 \arctan \sqrt{2} + 5}{\sqrt{(2\pi)^3}} - \frac{25}{4\sqrt{\pi}}.$$

Assim, a expansão do terceiro momento é dada por

$$\begin{aligned}E(X^3) &= \sum_{n=1}^{\infty} q(1-q)^n \mu'_3(n+1, n+1) \\ &= q(1-q)\mu'_3(2, 2) + q(1-q)^2\mu'_3(3, 3) + q(1-q)^3\mu'_3(4, 4) + q(1-q)^4\mu'_3(5, 5) + \\ &\quad \sum_{n=5}^{\infty} q(1-q)^n \mu'_3(n+1, n+1) \\ &= \frac{5\sqrt{\pi}q(1-q)}{2} + \frac{15q(1-q)^2}{4\sqrt{\pi}} + q(1-q)^3 \left[ \frac{15 \arctan \sqrt{2}}{\pi\sqrt{\pi}} + \frac{\sqrt{2\pi}}{2\pi^2} \right] + \\ &\quad + q(1-q)^4 \left[ \frac{75 \arctan \sqrt{2} + 5}{\sqrt{(2\pi)^3}} - \frac{25}{4\sqrt{\pi}} \right] + \sum_{n=5}^{\infty} q(1-q)^n \mu'_3(n+1, n+1).\end{aligned}$$

Quando  $0 < p \leq 1$  o terceiro momento da distribuição MON padrão pode ser representado como

$$\begin{aligned}E(X^3) &= \sum_{n=0}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[1-\Phi(x)]^n dx \quad (\text{A.3}) \\ &= p \int_{-\infty}^{\infty} x^3 \phi(x) dx + \sum_{n=1}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x^3 \phi(x)[1-\Phi(x)]^n dx \\ &= \sum_{n=1}^{\infty} p(1-p)^n \mu'_3(n+1, 1),\end{aligned}$$

em que  $\mu'_3(n+1, 1) = \int_{-\infty}^{\infty} (n+1)x^3\phi(x)[1-\Phi(x)]^n dx$  é o terceiro momento do mínimo de  $n+1$  variáveis normais padrão. Por Bose e Gupta (1959) tem-se

$$\mu'_3(n+1, 1) = \frac{5}{2}\mu'_1(n+1, 1) - \frac{\binom{n+1}{4}}{\pi^2} \int_{-\infty}^{\infty} [1-\Phi(x)]^{n-3} e^{-2x^2} dx.$$

Substituindo  $n = (1, 2, 3)$ , respectivamente, em (A.3), tem-se

$$\begin{aligned} \mu'_3(2, 1) &= \frac{5}{2}\mu'_1(2, 1) = -\frac{5\sqrt{\pi}}{2\pi}, \\ \mu'_3(3, 1) &= \frac{5}{2}\mu'_1(3, 1) = -\frac{15\sqrt{\pi}}{4}, \\ \mu'_3(4, 1) &= \frac{5}{2}\mu'_1(4, 1) - \frac{1}{\pi^2} \frac{\sqrt{2\pi}}{2} \\ &= -\frac{15 \arctan \sqrt{2}}{\pi\sqrt{\pi}} - \frac{\sqrt{2\pi}}{2\pi^2}. \end{aligned}$$

Desta forma, o terceiro momento da distribuição MON padrão para valores de  $p$  entre zero e um é dado pela expansão

$$\begin{aligned} E(X^3) &= \sum_{n=1}^{\infty} p(1-p)^n \mu'_3(n+1, 1) \\ &= p(1-p)\mu'_3(2, 1) + p(1-p)^2\mu'_3(3, 1) + p(1-p)^3\mu'_3(4, 1) + \sum_{n=4}^{\infty} p(1-p)^n \mu'_3(n+1, 1) \\ &= -p(1-p)\frac{5\sqrt{\pi}}{2\pi} - p(1-p)^2\frac{15\sqrt{\pi}}{4} - p(1-p)^3 \left[ \frac{15 \arctan \sqrt{2}}{\pi\sqrt{\pi}} + \frac{\sqrt{2\pi}}{2\pi^2} \right] \\ &\quad + \sum_{n=4}^{\infty} p(1-p)^n \mu'_3(n+1, 1). \end{aligned}$$

## A.2 Quarto momento

O quarto momento de uma variável aleatória com distribuição MON padrão pode ser representado por

$$E(X^4) = \begin{cases} \sum_{n=0}^{\infty} p(1-p)^n(n+1) \int_{-\infty}^{\infty} x^4\phi(x)[1-\Phi(x)]^n dx, & \text{se } 0 < p \leq 1, \\ \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^4\phi(x)\Phi(x)^n dx, & \text{se } q^{-1} = p > 1. \end{cases} \quad (\text{A.4})$$

Para  $q^{-1} = p > 1$  tem-se

$$\begin{aligned}
 E(X^4) &= \sum_{n=0}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [\Phi(x)]^n dx \\
 &= q \int_{-\infty}^{\infty} x^4 \phi(x) dx + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [\Phi(x)]^n dx \\
 &= 3q + \sum_{n=1}^{\infty} q(1-q)^n(n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [\Phi(x)]^n dx \\
 &= 3q + \sum_{n=1}^{\infty} q(1-q)^n \mu'_4(n+1, n+1),
 \end{aligned}$$

em que

$$\mu'_4(n+1, n+1) = (n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [\Phi(x)]^n dx$$

é o quarto momento do máximo de  $n+1$  variáveis com distribuição normal padrão, e  $\int_{-\infty}^{\infty} x^4 \phi(x) dx$  é o quarto momento (que representa a curtose) da distribuição normal padrão e é igual a 3.

Por Bose e Gupta (1959), o quarto momento do máximo de  $n+1$  variáveis com distribuição normal padrão pode ser escrito como

$$\mu'_4(n+1, n+1) = -\frac{4}{3} + \frac{13}{3} \mu'_2(n+1, n+1) + \frac{5 \binom{n+1}{5}}{(2\pi)^{\frac{5}{2}}} \int_{-\infty}^{\infty} [\Phi(x)]^{n-4} e^{-\frac{5}{2}x^2} dx. \quad (\text{A.5})$$

Substituindo  $n = 1, 2, 3$  na expressão (A.5), respectivamente, tem-se que

$$\begin{aligned}
 \mu'_4(2, 2) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(2, 2) = 3, \\
 \mu'_4(3, 3) &= -\frac{9}{3} + \frac{13}{3} \mu'_2(3, 3) = -\frac{4}{3} + \frac{13}{3} \left( \frac{2\pi + \sqrt{3}}{2\pi} \right), \\
 \mu'_4(4, 4) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(4, 4) = -\frac{4}{3} + \frac{13}{3} \left( \frac{\pi + \sqrt{3}}{\pi} \right), \\
 \mu'_4(5, 5) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 5) + \frac{5}{(2\pi)^{\frac{5}{2}}} \int_{-\infty}^{\infty} e^{-\frac{5}{2}x^2} dx \\
 &= -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 5) + \frac{\sqrt{5}}{4\pi^2}.
 \end{aligned}$$



Portanto, o quarto momento da distribuição MON padrão pode ser representado como

$$\begin{aligned}
E(X^4) &= 3q + \sum_{n=1}^{\infty} q(1-q)^n \mu'_4(n+1, n+1) \\
&= 3q + q(1-q)\mu'_4(2, 2) + q(1-q)^2\mu'_4(3, 3) + q(1-q)^3\mu'_4(4, 4) + q(1-q)^4\mu'_4(5, 5) \\
&\quad + \sum_{n=5}^{\infty} q(1-q)^n \mu'_4(n+1, n+1) \\
&= 3q + 3q(1-q) + q(1-q)^2 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{2\pi + \sqrt{3}}{2\pi} \right) \right] + q(1-q)^3 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{\pi + \sqrt{3}}{\pi} \right) \right] \\
&\quad + q(1-q)^4 \left( -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 5) + \frac{\sqrt{5}}{4\pi^2} \right) + \sum_{n=5}^{\infty} q(1-q)^n \mu'_4(n+1, n+1).
\end{aligned}$$

Quando  $0 < p \leq 1$  o quarto momento da distribuição MON padrão pode ser escrito como

$$\begin{aligned}
E(X^4) &= \sum_{n=0}^{\infty} p(1-p)^n (n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [1 - \Phi(x)]^n dx \\
&= p \int_{-\infty}^{\infty} x^4 \phi(x) dx + \sum_{n=1}^{\infty} p(1-p)^n (n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [1 - \Phi(x)]^n dx \\
&= 3p + \sum_{n=1}^{\infty} p(1-p)^n \mu'_4(n+1, 1),
\end{aligned}$$

com  $\mu'_4(n+1, 1) = (n+1) \int_{-\infty}^{\infty} x^4 \phi(x) [1 - \Phi(x)]^n dx$  sendo o quarto momento do mínimo de uma amostra de  $n+1$  observações da distribuição normal padrão. Por Bose e Gupta (1959), esse momento é dado por

$$\mu'_4(n+1, 1) = -\frac{4}{3} + \frac{13}{3} \mu'_2(n+1, 1) + \frac{5 \binom{n+1}{5}}{(2\pi)^{\frac{5}{2}}} \int_{-\infty}^{\infty} [1 - \Phi(x)]^{n-4} e^{-\frac{5}{2}x^2} dx. \quad (\text{A.6})$$

Escrevendo  $n = 1, 2, 3$  na expressão (A.6) os primeiros termos dessa expansão são

$$\begin{aligned}
\mu'_4(2, 1) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(2, 1) = 3, \\
\mu'_4(3, 1) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(3, 1) = -\frac{4}{3} + \frac{13}{3} \left( \frac{2\pi + \sqrt{3}}{2\pi} \right), \\
\mu'_4(4, 1) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(4, 1) = -\frac{4}{3} + \frac{13}{3} \left( \frac{\pi + \sqrt{3}}{\pi} \right), \\
\mu'_4(5, 1) &= -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 1) + \frac{5}{(2\pi)^{\frac{5}{2}}} \int_{-\infty}^{\infty} e^{-\frac{5}{2}x^2} dx \\
&= -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 1) + \frac{\sqrt{5}}{4\pi^2}.
\end{aligned}$$

Portanto, o quarto momento para a distribuição MON padrão para valores de  $p$  entre zero e um pode ser escrito como

$$\begin{aligned}
E(X^4) &= 3p + \sum_{n=1}^{\infty} p(1-p)^n \mu'_4(n+1, 1) \\
&= 3p + p(1-p)\mu'_4(2, 1) + p(1-p)^2\mu'_4(3, 1) + p(1-p)^3\mu'_4(4, 1) \\
&\quad + p(1-p)^4\mu'_4(5, 1) + \sum_{n=5}^{\infty} p(1-p)^n \mu'_4(n+1, 1) \\
&= 3p + 3p(1-p) + p(1-p)^2 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{2\pi + \sqrt{3}}{2\pi} \right) \right] + p(1-p)^3 \left[ -\frac{4}{3} + \frac{13}{3} \left( \frac{\pi + \sqrt{3}}{\pi} \right) \right] \\
&\quad + p(1-p)^4 \left[ -\frac{4}{3} + \frac{13}{3} \mu'_2(5, 1) + \frac{\sqrt{5}}{4\pi^2} \right] + \sum_{n=5}^{\infty} p(1-p)^n \mu'_4(n+1, 1).
\end{aligned}$$

# Apêndice B

Neste apêndice são apresentadas as derivadas de segunda ordem da função de log-verossimilhança da distribuição MON. Lembrando que as representações usadas a seguir foram obtidas através da propriedade que a esperança do vetor escore é zero.

Fazendo-se as representações

$$v_i(t_i) = \frac{1 - \Phi(t_i)}{\Phi(t_i)(1-p) + p} \quad \text{e} \quad w_i(t_i) = \frac{\phi(t_i)}{\Phi(t_i)(1-p) + p},$$

tem-se que as derivadas parciais de segunda ordem em relação aos parâmetros da distribuição MON são dadas por

$$\begin{aligned} \frac{\partial^2 l}{\partial p^2} &= -\frac{n}{p^2} + 2 \sum_{i=1}^n v_i^2(t_i), \\ \frac{\partial^2 l}{\partial \mu^2} &= -\frac{n}{\sigma^2} + \frac{2(1-p)}{\sigma^2} \sum_{i=1}^n \{t_i w_i(t_i) + (1-p)w_i^2(t_i)\}, \\ \frac{\partial^2 l}{\partial \sigma^2} &= \frac{n}{\sigma^2} - \frac{3}{\sigma^2} \sum_{i=1}^n t_i^2 + \frac{2(1-p)}{\sigma^2} \sum_{i=1}^n \{(t_i^2 - 2)w_i(t_i)t_i + t_i^2 w_i^2(t_i)\}, \\ \frac{\partial^2 l}{\partial p \partial \mu} &= -\frac{2}{\sigma} \sum_{i=1}^n w_i(t_i) - \frac{2(1-p)}{\sigma} \sum_{i=1}^n v_i(t_i)w_i(t_i), \\ \frac{\partial^2 l}{\partial p \partial \sigma} &= -\frac{2}{\sigma} \sum_{i=1}^n t_i w_i(t_i) - \frac{2(1-p)}{\sigma} \sum_{i=1}^n t_i v_i(t_i)w_i(t_i), \\ \frac{\partial^2 l}{\partial \mu \partial \sigma} &= -\frac{2}{\sigma} \sum_{i=1}^n t_i + \frac{2(1-p)}{\sigma^2} \sum_{i=1}^n \{(t_i^2 - 1)w_i(t_i) + (1-p)t_i w_i^2(t_i)\}. \end{aligned}$$



# Apêndice C

Neste apêndice, uma análise do conjunto de dados utilizado por García et al. (2010) é realizada. Para verificar a bondade do ajuste, outros testes, além dos utilizados por eles, são apresentados.

Na Tabela C.1, tem-se as estimativas de máxima verossimilhança (com os correspondentes erros-padrão entre parênteses) e os valores das estatísticas AIC, BIC e CAIC para os modelos MON, normal, normal potência (NP) e skew-normal (SN).

Tabela C.1: EMVs dos parâmetros dos modelos MON, normal, NP, SN e SNB para dados de IQ Scores OITS, correspondentes erros-padrão (entre parênteses) e valores dos AIC, BIC e CAIC.

Distribuição	Estimativas (Erro)	Log-vero	AIC	BIC	CAIC
MON	$\hat{\rho} = 0.2507$ (0.2436) $\hat{\mu} = 112.8751$ (4.7352) $\hat{\sigma} = 8.3928$ (0.9503)	<b>-182.3138</b>	<b>370.6276</b>	<b>376.4813</b>	<b>391.6477</b>
Normal	$\hat{\mu} = 106.6538$ (1.1415) $\hat{\sigma} = 8.2296$ (0.7998)	-183.3872	370.7744	374.6769	384.7878
NP	$\hat{\lambda} = 42.3960$ (16.1669) $\hat{\mu} = 68.4544$ (5.3993) $\hat{\sigma} = 17.4787$ (2.0053)	-182.4298	370.8596	376.7133	379.7133
SN	$\hat{\lambda} = 2.5463$ (1.2147) $\hat{\mu} = 97.4548$ (1.9864) $\hat{\sigma} = 12.3430$ (1.9647)	<b>-182.1399</b>	<b>370.2798</b>	<b>376.1335</b>	<b>379.1335</b>

Na Tabela C.2 encontra-se os testes de hipóteses realizados para testar a distribuição MON com as distribuições normal e skew-normal.

Tabela C.2: Teste de Hipóteses

Teste	$H_0$	$H_1$	Estatística	Nível	p-valor
Razão de verossimilhança	normal	MON	2.1468	10%	0.1429
Razão de verossimilhança generalizada	SN	MON	-0.3505	5%	-

Desta forma, aceita-se a hipótese nula em favor da distribuição normal ao nível de significância de 5%, ou seja, o modelo normal se ajusta melhor ao conjunto de dados analisados do que o modelo MON, baseando-se na estatística da razão de verossimilhança. Já no teste da razão de verossimilhança generalizada, aceita-se a hipótese nula o que significa que os modelos MON e skew-normal são significativamente equivalentes para o ajuste do conjunto de dados.

A Tabela C.3 apresenta os resultados para o teste de Kolmogorov-Smirnov para os modelos testados. Dentre os quatro modelos analisados, o melhor ajuste é do modelo SN com p-valor igual a 0.8250. Em outras palavras, tem-se que o teste KS sugere que os dados seguem uma distribuição skew-normal.

Tabela C.3: Teste de Kolmogorov-Smirnov para as distribuições testadas.

Distribuição	$D_n$	p-valor
MON	0.0962	<b>0.6570</b>
Normal	0.1264	0.3450
NP	0.0932	0.734
SN	0.0859	<b>0.8250</b>

A Figura C.1 ilustra o histograma dos dados IQ scores OTIS. As linhas representam as densidades das distribuições MON, normal, normal potência e skew-normal, usando os valores encontrados para as estimativas de máxima verosimilhança encontradas para os parâmetros.

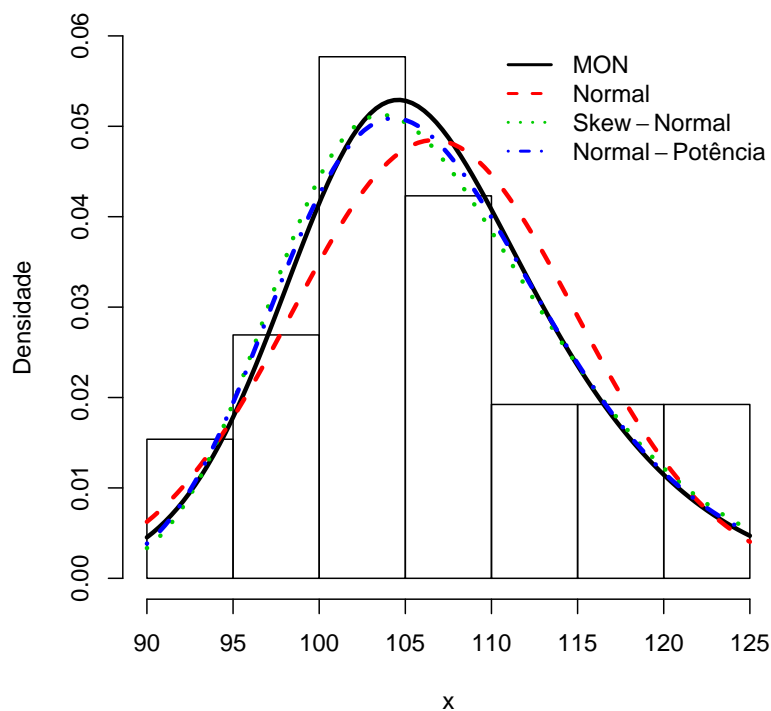


Figura C.1: Histograma dos dados de fibras de carbono com as funções de densidades de probabilidades substituindo-se as EMVs dos modelos MON, normal, skew-normal e normal-potência.