

Família de Distribuições Log-*Skew*-Multivariadas: Definição, Entropia e outras propriedades

Aluna: Marina Muniz de Queiroz
Orientadora: Rosangela Helena Loschi
Co-orientador: Roger William Câmara Silva

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
Belo Horizonte, 15 de Novembro de 2013

Agradecimentos

Agradeço aos meus pais, Ilze e Cristiano, pelo apoio durante meus estudos e preocupação com minha educação e cultura ao longo de todos estes anos. Por terem me incentivado e acreditado em minhas escolhas. À minha irmã, Thaís, por ter escutado tantas leituras e ajudado na correção deste trabalho.

Aos meus amigos, pelos dias de estudo, pelas tardes passadas na UFMG, pela ajuda nos problemas da dissertação e principalmente pelos momentos de descontração e relaxamento também necessários para este trabalho. Em especial, agradeço às amigas Luciana, Gabi, Lorena e Juliane, por participarem de tantas discussões, ouvirem reclamações e principalmente pelos cafés compartilhados.

À minha orientadora, Rosangela H. Loschi, pelo comprometimento com este trabalho, pelas diversas leituras e correções que tanto contribuíram para o texto final, pela paciência e entusiasmo durante os dias de trabalho. Ao meu orientador, Roger William Câmara Silva, pelo incentivo para que eu continuasse os estudos em probabilidade, pela confiança e pelo comprometimento com esta dissertação.

Aos membros da banca examinadora, professores Reinaldo Boris Arellano-Valle (PUC Chile), Bernardo Nunes Borges de Lima (UFMG) e Fredy Walther Castellares Cáceres (UFMG) pela participação, contribuições e sugestões.

Aos professores dos departamentos de Estatística e Matemática que me acompanharam durante os anos de graduação e mestrado e ajudaram na construção deste conhecimento. À minha professora do ensino médio Neila, que foi a principal responsável pela escolha deste caminho.

À CAPES e ao CNPq pela bolsa de mestrado que me possibilitou dedicação integral neste programa.

Resumo

O estudo de famílias assimétricas é muito importante, já que a suposição de normalidade, muitas vezes considerada na descrição do comportamento de dados estatísticos é, em algumas situações, inadequada. Além disto, em alguns casos, estamos interessados em variáveis que assumem somente valores positivos e, assim, as distribuições *skew*-normais podem não ser apropriadas. Esta dissertação apresenta uma extensão da família de distribuições log-*skew* normal multivariada introduzida por Marchenko e Genton (2010), chamada de família de distribuições log-*skew*-normal fundamental canônica, onde a função usada para assimetrizar a distribuição normal passa a ser uma função de distribuição acumulada m -variada. Ela é obtida através da transformação da família de distribuições *skew*-normal fundamental canônica multivariada, definida por Arellano-Valle e Genton (2005), e algumas de suas propriedades, tais como momentos, distribuições marginais e condicionais e representações estocásticas, são discutidas neste texto.

Outra contribuição desta dissertação é o estudo do comportamento da entropia nas famílias assimétricas. Calculamos a entropia e informação mútua na família de distribuições log-*skew*-normal fundamental canônica multivariada, relacionando estes resultados com os já existentes para as distribuições normal e *skew*-normal multivariadas. Com o objetivo de compararmos esta nova família e a definida por Marchenko e Genton (2010), também encontramos uma expressão para a entropia relativa entre elas.

A nova distribuição é usada no ajuste de bancos de dados que assumem somente valores positivos, através da abordagem bayesiana. Dois modelos são propostos e comparados quanto à qualidade de ajuste quando variamos a dimensão m da função responsável por assimetrizar a distribuição normal na família log-*skew*-normal fundamental canônica.

Abstract

The study of asymmetric families is very important, since the normality assumption often considered in describing the statistical behavior of data is, in some cases, inadequate. Moreover, sometimes, we are interested in variables that do not take values in the negative reals and thus the *skew*-normal distributions may not be appropriate. This work introduces a generalization of the family of multivariate log-*skew* normal distributions defined by Marchenko and Genton (2010), called canonical fundamental log-*skew* normal distribution, where the function used to asymmetrise the normal distribution becomes a m -variate cumulative distribution function. It is obtained via a transformation of the family of multivariate canonical fundamental *skew* normal distribution, defined by Arellano-Valle and Genton (2005). Some of its properties, such as moments, marginal and conditional distributions and stochastic representations, are discussed in this text.

Another contribution of this work is the study of entropy in asymmetric families. We calculate the entropy and mutual information for the canonical fundamental log-*skew* normal distribution, relating these results with existing ones for multivariate normal and *skew*-normal distributions. Aiming to compare this new family and the one introduced by Marchenko and Genton (2010), we also find an expression for the relative entropy between them.

The new distribution is used to model databases that take only positive values, using bayesian inference. Different models are proposed and compared by the goodness of fit when we change the dimension m of the function responsible for asymmetrise normal distribution in the family of canonical fundamental log-*skew* normal distributions.

Sumário

1	Distribuições Univariadas Originadas de Transformações da Distribuição Normal	8
1.1	Distribuição <i>Skew</i> -Normal Univariada	9
1.2	Distribuição Log- <i>Skew</i> -Normal Univariada	12
2	Distribuição Log-<i>Skew</i>-Elíptica Multivariada	15
2.1	Distribuições Esféricas e Elípticas	16
2.2	Distribuição <i>Skew</i> -Elíptica Multivariada	19
2.3	Distribuição Log- <i>Skew</i> -Elíptica Multivariada	20
3	Distribuição Log-<i>Skew</i>-Normal Fundamental Multivariada	25
3.1	Distribuição <i>Skew</i> -Normal Fundamental Multivariada	26
3.2	Distribuição Log- <i>Skew</i> -Normal Fundamental Canônica Multivariada	35
4	Entropia e Informação Mútua em Famílias de Distribuições Assimétricas	41
4.1	Entropia nas Famílias de Distribuições Univariadas	42
4.2	Entropia nas Famílias de Distribuições Multivariadas	45
4.3	Entropia na Família de Distribuições Log- <i>Skew</i> -Normal Fundamental Canônica Multivariada	50
5	Inferência Bayesiana em Modelos Assimétricos	55
5.1	Modelo Estatístico	55
5.2	Estudo de Casos: Precipitação Mensal nos EUA	60

Introdução

A família de distribuições normal é amplamente utilizada para descrever o comportamento de fenômenos diversos, mas, muitas vezes, a suposição de normalidade não é adequada. Existem dados que apresentam características que não são apropriadamente descritas pela distribuição normal, tais como caudas pesadas, bimodalidade e assimetria. Por esta razão, a construção de outras famílias de distribuições com tais características vem sendo abordada na literatura. Azzalini (1985) constrói uma família de distribuições assimétrica, denominada por família de distribuições *skew-normal* univariada, que, além de incluir a distribuição normal como caso particular, é capaz de captar graus de assimetria relativamente fortes. Azzalini e Dalla Valle (1996) generalizam esta ideia e constroem a família de distribuições *skew-normal* multivariada. Branco de Dey (2001) generalizam a classe de distribuições elípticas, introduzindo a família de distribuições *skew-elíptica* multivariada. Em todos estes casos, a assimetria é introduzida pela multiplicação da função densidade de probabilidade de uma variável aleatória n -dimensional por uma função de distribuição acumulada univariada. Arellano-Valle e Genton (2005) definem uma família ainda mais ampla de distribuições *skew-normal*, chamada fundamental canônica.

O suporte destas distribuições assimétricas com núcleo normal é o conjunto dos números reais, no caso univariado, ou o \mathbb{R}^n , no caso multivariado. Para situações em que as variáveis não podem assumir valores negativos, o uso de tais distribuições pode não ser adequado. Com o objetivo de tentar descrever mais apropriadamente o comportamento de fenômenos desta natureza, surge a necessidade de obter distribuições de probabilidade que colocam massa nula em valores negativos. Marchenko e Genton (2010) introduzem uma família de distribuições *log-skew* elíptica multivariada, gerada a partir da transformação da família de distribuições *skew-elíptica* multivariada, e estudam algumas de suas propriedades. A família generaliza em alguns sentidos a família de distribuições *log-skew* normal definida em Azzalini *et al.* (2003) e a família *log-normal* que é construída a partir de transformações similares da família de distribuições normal.

A motivação para definirmos uma nova família de distribuições do tipo *log* partindo de transformações mais gerais que as de famílias *skew-normal* vem dos resultados obtidos em Santos *et al.* (2013). Eles estão interessados na interpretação dos parâmetros do modelo logístico misto quando distribuições mais flexíveis que a normal são assumidas para modelar o comportamento dos efeitos aleatórios. No modelo logístico misto, a razão de chances (OR) usada para interpretar os parâmetros é uma quantidade aleatória que é função dos efeitos aleatórios (ver Larsen *et al.* (2000) para detalhes). Portanto, sua interpretação dependerá da distribuição adotada para tais efeitos. Se assumimos que os efeitos aleatórios são independentes e identicamente distribuídos com distribuição *skew-normal* com locação μ , escala σ^2 e parâmetro de assimetria α (a qual será definida no Capítulo 1) e se assumimos que apenas o intercepto do modelo logístico é aleatório, Santos *et al.* (2013) provam que a distribuição da razão de chances para comparar dois indivíduos com mesmas características localizados em *clusters* diferentes tem a seguinte função densidade de probabilidade (fdp):

$$f_{OR}(r) = \frac{4}{r} \phi(\ln r | 2\sigma^2) \Phi_2 \left(\frac{\delta \ln r}{2\sigma} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \middle| I_2 - \frac{\delta^2}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right), \quad (0.1)$$

onde $\delta = \alpha(1 + \alpha)^{-1/2}$ e $\phi_n(\cdot | \Sigma)$ e $\Phi_n(\cdot | \Sigma)$ denotam a fdp e a função de distribuição acumulada

(fda) de uma distribuição normal n -variada com média zero e matriz de covariâncias Σ . A família definida na Expressão (0.1) não pertence à classe de distribuições *log-skew-normal* definidas em Marchenko e Genton (2010), uma vez que a função que introduz assimetria na fdp não é a fda de uma distribuição normal univariada. Sendo assim, a família obtida na Expressão (0.1) precisa ser formalmente definida e suas propriedades necessitam ser exploradas. Da mesma forma que em Marchenko e Genton (2010), definimos uma nova família de distribuições a partir da transformação da família proposta por Arellano-Valle e Genton (2005), chamada de *log-skew-normal fundamental canônica*. Esta família generaliza a família proposta por Marchenko e Genton (2010) no caso de normalidade e inclui a família encontrada em Santos *et al.* (2013) como caso particular.

Uma das características das famílias de distribuições é a entropia, que pode ser entendida como uma medida matemática da informação necessária, em média, para descrever o comportamento de uma quantidade aleatória. Arellano-Valle *et al.* (2011) estudaram o conceito de entropia nas famílias de distribuições elíptica e *skew-elíptica* multivariadas, apresentando resultados particulares para as famílias de distribuições *skew-normal* multivariada. Neste trabalho, encontraremos expressões e estudaremos as propriedades da entropia nas famílias de distribuições *skew-normal fundamental canônica* e *log-skew-normal fundamental canônica* multivariadas. Com o objetivo de compararmos as famílias de distribuições *log-skew-normal fundamental canônica* multivariada e *log-skew-normal*, também encontraremos uma expressão para a entropia relativa entre estas duas famílias.

Alguns aspectos relacionados à questão da inferência na família de distribuições *log-skew-normal fundamental canônica* univariada, do ponto de vista bayesiano, são apresentados nesta dissertação. Usaremos esta distribuição para modelar o comportamento do índice de precipitação (PCP) nos Estados Unidos. Ajustaremos modelos considerando alguns valores distintos para o componente m da função que assimetriza a distribuição normal nesta família. O objetivo é avaliar se há melhoria na qualidade de ajuste ao aumentarmos o valor de m . Desta forma, como subproduto, também discutimos procedimentos bayesianos para fazer inferência na família *log-skew-normal* introduzida por Marchenko e Genton (2010). Os modelos estatísticos descritos foram implementados no software gratuito *Winbugs* e os resultados obtidos foram analisados utilizando o software gratuito R.

Esta dissertação está organizada da seguinte forma. No Capítulo 1 apresentamos as famílias de distribuições *skew-normal* e *log-skew-normal* univariadas, originadas de transformações da distribuição normal. No Capítulo 2, faremos uma revisão sobre as distribuições elípticas e esféricas e apresentaremos as famílias de distribuições *skew-elíptica* e *log-skew-elíptica* multivariadas e algumas de suas propriedades. O Capítulo 3 é destinado a introduzir a família de distribuições *skew-normal fundamental canônica* e definir a nova família de distribuições *log-skew-normal fundamental canônica* e ao estudo de suas propriedades. O Capítulo 4 apresenta resultados sobre a entropia e informação mútua nas famílias assimétricas. No Capítulo 5 abordaremos a questão da inferência na família de distribuições *log-skew-normal fundamental canônica* univariada, do ponto de vista bayesiano. Também ajustaremos modelos assimétricos para um banco de dados reais sobre o índice de precipitação nos EUA.

Capítulo 1

Distribuições Univariadas Originadas de Transformações da Distribuição Normal

A suposição de normalidade, muitas vezes considerada na descrição do comportamento de dados estatísticos é, em algumas situações, inadequada. Não é pouco comum encontrarmos dados que indicam a presença de caudas pesadas, assimetria, bimodalidade e outras características que não são apropriadamente descritas ou capturadas pela distribuição normal.

Para tratar adequadamente dados estatísticos com tais características, alguns caminhos vêm sendo considerados amplamente na literatura. Entre eles estão métodos não paramétricos (ver Escobar e West (1998), por exemplo), a consideração de mistura discreta de distribuições (West, 1987) e a busca por famílias de distribuições de probabilidade flexíveis o bastante para capturar tais características. Esta última abordagem é bastante atrativa, pois, em geral, obtemos modelos mais parcimoniosos. Existem vários procedimentos encontrados na literatura para construir distribuições paramétricas mais flexíveis. De uma forma geral, Azzalini (1985) prova que uma função densidade de probabilidade (fdp) assimétrica pode ser construída a partir de uma fdp f_o simétrica ao redor de zero e de uma função de distribuição acumulada (fda) H , com fdp h também simétrica ao redor de zero, da seguinte forma:

$$f(x) = 2f_o(x)H(\alpha x), x \in \mathbb{R}, \quad (1.1)$$

onde $\alpha \in \mathbb{R}$ é o parâmetro responsável por controlar o grau de assimetria da nova fdp. Ao fazermos isso, estamos construindo uma família mais geral de probabilidade que inclui a distribuição f_o como caso particular, quando $\alpha = 0$. Partindo desta ideia, Azzalini (1985) introduz a família normal assimétrica, denominada família *skew-normal* (SN) ao longo deste texto, ao assumir, na Expressão (1.1), que f_o e H são, respectivamente, a fdp e a fda da distribuição normal padrão.

Se f_o é também unimodal com segundo momento finito dado por $b_2 = \int x^2 f_o(x) dx$, por exemplo, uma família assimétrica bimodal pode ser construída a partir das funções f_o e H , com fdp dada por

$$f(x) = 2 \left(\frac{1 + \alpha x^2}{1 + \alpha b_2} \right) f_o(x) H(\alpha_0 x), x \in \mathbb{R}, \quad (1.2)$$

onde $\alpha_0 \in \mathbb{R}$ e $\alpha \geq 0$. Se $\alpha_0 = \alpha = 0$, então voltamos à distribuição $f_o(x)$ (Elal-Olivero *et al.*, 2009).

Neste capítulo faremos uma revisão detalhada de algumas distribuições de probabilidade univariadas originadas de transformações da distribuição normal univariada, as quais são relevantes para o estudo que faremos nos capítulos seguintes. Na Seção 1.1 definiremos, formalmente, a família de distribuições *skew-normal* univariada introduzida por Azzalini (1985) e mostraremos algumas de suas propriedades. Na Seção 1.2, apresentaremos a distribuição *log-skew-normal* obtida a partir de uma transformação da variável *skew-normal*, estudada por Azzalini *et al.* (2003).

Ao longo deste trabalho, denotaremos por $\phi_n(\cdot | \boldsymbol{\mu}, \boldsymbol{\Sigma})$ e $\Phi_n(\cdot | \boldsymbol{\mu}, \boldsymbol{\Sigma})$, respectivamente, a fdp e a fda de uma variável aleatória com distribuição normal n -variada com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$. Se o vetor de médias é zero, $\boldsymbol{\mu}$ será omitido da notação. Se, além disso, a matriz de covariâncias for a identidade, $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ serão omitidos. No caso univariado, denotaremos da mesma forma, apenas omitindo n da notação.

1.1 Distribuição *Skew-Normal* Univariada

Em 1985, Azzalini constrói uma família de distribuições de probabilidade com núcleo normal que, além de incluir a família normal como caso particular, também é capaz de descrever o comportamento de variáveis aleatórias cujas distribuições possuem grau de assimetria razoavelmente forte se comparadas à distribuição normal. Esta família é denominada *skew-normal* univariada e é definida a seguir.

Definição 1.1. *Se uma variável aleatória X tem função densidade dada por*

$$f(x | \alpha) = 2\phi(x)\Phi(\alpha x), \quad x \in \mathbb{R}, \quad (1.3)$$

*com parâmetro de assimetria α , dizemos que X tem distribuição *skew-normal* padrão, a qual denotamos por $X \sim SN(\alpha)$.*

Para tornarmos a classe de distribuições dada na Expressão (1.3) ainda mais flexível, é possível adicionarmos parâmetros de locação e escala a tal família através da transformação $Z = \sigma X + \mu$, onde $\mu \in \mathbb{R}$ e $\sigma > 0$. Neste caso, a fdp da variável transformada Z será dada por

$$f(z | \mu, \sigma, \alpha) = \frac{2}{\sigma} \phi\left(\frac{z - \mu}{\sigma}\right) \Phi\left(\alpha \left(\frac{z - \mu}{\sigma}\right)\right), \quad z \in \mathbb{R}, \quad (1.4)$$

onde μ é parâmetro de locação, σ é parâmetro de escala e $\alpha \in \mathbb{R}$ é parâmetro de assimetria. A distribuição definida na Expressão (1.4) é denotada ao longo deste texto por $Z \sim SN(\mu, \sigma^2, \alpha)$.

A Figura 1.1 a seguir mostra o comportamento da fdp de uma variável Z com distribuição *skew-normal* para diferentes parâmetros de assimetria quando assumimos a forma padrão dada na Expressão (1.3) em que $\mu = 0$ e $\sigma^2 = 1$. Percebemos que quanto maior o valor absoluto de α , maior o grau de assimetria da distribuição. Além disto, valores positivos de α induzem uma assimetria positiva na curva e valores negativos, uma assimetria negativa. Destaca-se que quando $\alpha \rightarrow -\infty$, obtemos uma distribuição assimétrica que coloca quase certamente a totalidade de sua massa não-nula em valores negativos.

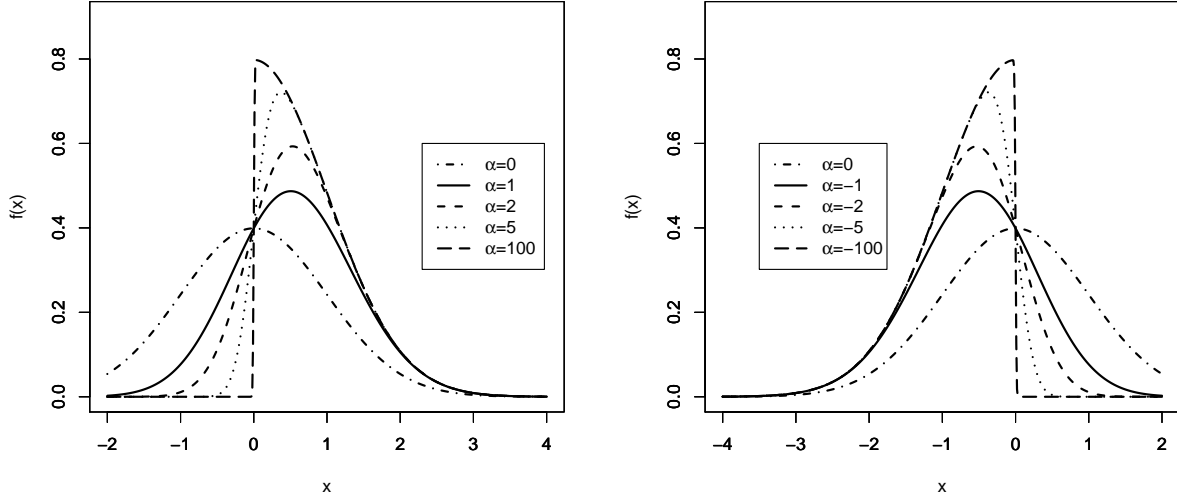


Figura 1.1: Densidades da distribuição *skew*-normal padrão para diferentes parâmetros de assimetria.

A distribuição *skew*-normal definida na Expressão (1.4) possui propriedades importantes, algumas das quais, estudadas por Azzalini (1985), são enumeradas a seguir.

1. Se $Z \sim SN(\mu, \sigma^2, 0)$, então $Z \sim N(\mu, \sigma^2)$. Isto decorre diretamente da Expressão (1.4), por notarmos que $\Phi(0) = 1/2$.
2. Quando $\alpha \rightarrow \infty$, temos que $f(z | \mu, \sigma, \alpha) \rightarrow \frac{2}{\sigma} \phi\left(\frac{z-\mu}{\sigma}\right) \mathbf{I}_{\{z>0\}}$, que é a fdp de uma variável aleatória com distribuição *half*-normal. Isto ocorre pois $\Phi(\alpha z) \rightarrow 1$, se $z > 0$, e $\Phi(\alpha z) \rightarrow 0$, se $z < 0$.
3. A fdp dada na Expressão (1.4) é fortemente unimodal, isto é, sua fdp é log-côncava.

As propriedades 4, 5 e 6, que apresentaremos mais a frente neste texto, são consequência da identidade que apresentaremos agora. Denotemos por $\Phi_{SN}(\cdot | \alpha)$ a fda da distribuição *skew*-normal padrão com parâmetro de assimetria α . Para todo $z \in \mathbb{R}$, tal fda pode ser escrita como

$$\Phi_{SN}(z | \alpha) = \int_{-\infty}^z 2\phi(t)\Phi(\alpha t)dt = \int_{-\infty}^z \int_{-\infty}^{\alpha t} 2\phi(t)\phi(u)dudt. \quad (1.5)$$

Note que a região de integração da Expressão (1.5) é denotada por A na Figura 1.2, para z e α positivos.

Para obtermos propriedades especiais de $\Phi_{SN}(z | \alpha)$, vamos relacioná-la com a função $T(z, \alpha)$ que é definida em Owen (1956) como a integral de uma distribuição normal padrão bivariada com correlação nula na região dada por $t \geq z$, $u = 0$ e $u = \alpha t$, para z e α positivos. Desta forma, Owen (1956) define

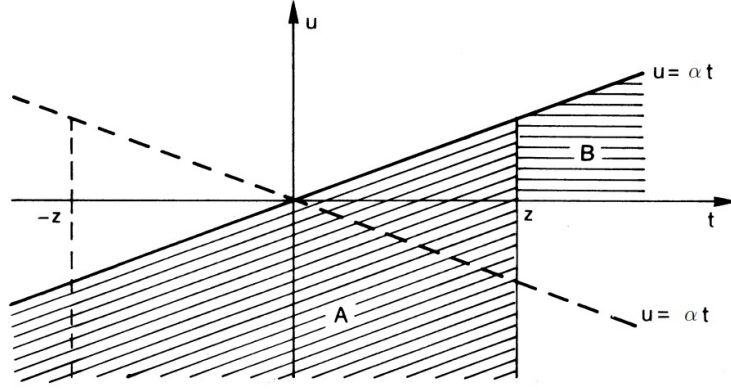


Figura 1.2: Região de integração para $\Phi_{SN}(z | \alpha)$.

$$2T_B(z, \alpha) = \int_z^\infty \int_0^{\alpha t} 2\phi(u)\phi(t)du dt. \quad (1.6)$$

Consequentemente, segue das Expressões (1.5) e (1.6) que

$$\begin{aligned} 2T_B(z, \alpha) &= \int_z^\infty 2\phi(t)(\Phi(\alpha t) - \Phi(0))dt \\ &= \int_z^\infty 2\phi(t)\Phi(\alpha t)dt - \int_z^\infty \phi(t)dt \\ &= 1 - \Phi_{SN}(z) - 1 + \Phi(z) \\ &= \Phi(z) - \Phi_{SN}(z | \alpha). \end{aligned} \quad (1.7)$$

Algumas propriedades de $T(z, \alpha)$ são conhecidas, tais como

$$-T(z, \alpha) = T(z, -\alpha), \quad T(-z, \alpha) = T(z, \alpha), \quad 2T(z, 1) = \Phi(z)\Phi(-z). \quad (1.8)$$

Sendo assim, como consequência, obtemos das Expressões (1.7) e (1.8) que, na família de distribuições *skew-normal*, as seguintes propriedades são verificadas:

4. $1 - \Phi_{SN}(-z | \alpha) = \Phi_{SN}(z | -\alpha)$.
5. $\Phi_{SN}(z | 1) = (\Phi(z))^2$.
6. Se $Z \sim SN(0, 1, \alpha)$, então $Z^2 \sim \chi^2$.

Para calcularmos os momentos da variável $Z \sim SN(\mu, \sigma^2, \alpha)$, primeiramente, calculamos os momentos da variável $X \sim SN(\alpha)$ e, em seguida, usamos a transformação $Z = \sigma X + \mu$ para obtermos o resultado geral. Para isso, será necessário o seguinte lema que pode ser encontrado em Zacks (1981).

Lema 1.1. *Seja $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$. Então, para qualquer vetor $\mathbf{u} \in \mathbb{R}^m$ e qualquer matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$, temos*

$$E_{\mathbf{Z}}(\Phi_m(\mathbf{u} + \mathbf{AZ}|\Sigma)) = \Phi_m(\mathbf{u}|\Sigma + \mathbf{AA}'). \quad (1.9)$$

Os momentos de X , e conseqüentemente de Z , podem ser mais facilmente obtidos a partir da função geradora de momentos (fgm) e suas propriedades. Sendo assim, inicialmente, vamos calcular a função geradora de momentos de $X \sim SN(\alpha)$. Para todo $t \in \mathbb{R}$, segue que

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \int_{-\infty}^{\infty} 2e^{tx} \phi(x) \Phi(\alpha x) dx = \int_{-\infty}^{\infty} 2e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Phi(\alpha x) dx \\ &= 2e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} \Phi(\alpha x) dx. \end{aligned}$$

Considerando a seguinte transformação $y = x - t$ e aplicando o Lema (1.1), obtemos que a fgm de X é

$$M_X(t) = 2e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \Phi(\alpha(y+t)) dy = 2e^{\frac{t^2}{2}} E(\Phi(\alpha(Y+t))) = 2e^{\frac{t^2}{2}} \Phi\left(\frac{\alpha t}{\sqrt{1+\alpha^2}}\right). \quad (1.10)$$

Usando propriedades da fgm, segue da Expressão (1.10) que $E(X) = \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1+\alpha^2}}$ e $\text{Var}(X) = 1 - \frac{2\alpha^2}{\pi(1+\alpha^2)}$.

Agora, considere $Z \sim SN(\mu, \sigma^2, \alpha)$. Segue, de propriedades da esperança e variância, que

$$E(Z) = \mu + \sigma \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1+\alpha^2}} \text{ e } \text{Var}(Z) = \sigma^2 \left(1 - \frac{2\alpha^2}{\pi(1+\alpha^2)}\right). \quad (1.11)$$

1.2 Distribuição Log-Skew-Normal Univariada

Em algumas situações, estamos interessados em variáveis que assumem somente valores positivos. Nestes casos, a distribuição *skew-normal* pode não ser apropriada, mesmo quando seu comportamento se aproxima da distribuição *half-normal*. Fazendo um paralelo com a teoria desenvolvida a partir da distribuição normal, Azzalini *et al.* (2003) constrói, através da transformação $Y = e^Z$, onde $Z \sim SN(\mu, \sigma^2, \alpha)$, uma família de distribuições que coloca massa apenas em valores positivos, a qual é chamada distribuição *log-skew-normal* e é, formalmente, definida como segue.

Definição 1.2. Dizemos que Y tem distribuição *log-skew-normal* com parâmetro de locação $\mu \in \mathbb{R}$, parâmetro de escala $\sigma > 0$ e parâmetro de assimetria $\alpha \in \mathbb{R}$ se, e somente se, $Z = \ln Y \sim SN(\mu, \sigma^2, \alpha)$. Denotamos por $Y \sim LSN(\mu, \sigma^2, \alpha)$.

Considere a transformação $Y = e^Z$. Então, considerando a fdp da Expressão (1.4) e o método jacobiano, mostramos que Y tem fdp dada por

$$f(y | \mu, \sigma, \alpha) = \frac{2}{\sigma y} \phi\left(\frac{\ln y - \mu}{\sigma}\right) \Phi\left(\alpha \left(\frac{\ln y - \mu}{\sigma}\right)\right), y > 0. \quad (1.12)$$

Assim como a família de distribuições *skew-normal* contém a família de distribuições normal como caso particular, a família de distribuições *log-skew-normal* apresentada na Expressão (1.12) contém a família de distribuições *log-normal* como caso particular quando assumimos $\alpha = 0$. Dizemos que uma variável aleatória X tem distribuição *log-normal* com parâmetro de locação μ e parâmetro de escala σ^2 , denotada por $X \sim LN(\mu, \sigma^2)$, se sua fdp é dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \mu)^2}{\sigma^2}}, x > 0. \quad (1.13)$$

A Figura 1.3 apresenta a fdp da distribuição *log-skew-normal* definida na Expressão (1.12), com parâmetros $\mu = 0$ e $\sigma^2 = 1$ e alguns valores distintos para α . Percebemos que a fdp pode assumir diversas formas quando variamos o parâmetro de assimetria, colocando mais massa próximo do zero se $\alpha \rightarrow 0$ ou massa praticamente nula próximo da origem se $\alpha \rightarrow \infty$.

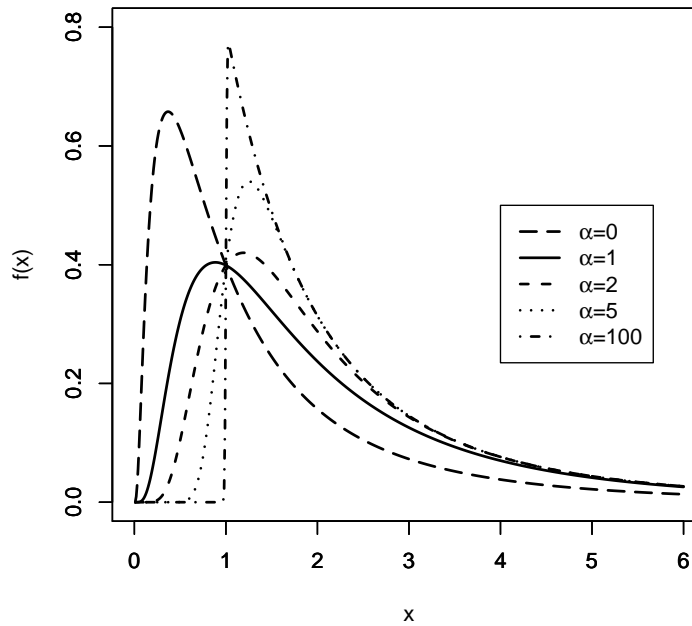


Figura 1.3: Densidades das distribuições *log-skew-normal* com $\mu=0$ e $\sigma^2=1$ e diferentes valores para o parâmetro de assimetria α .

Usando o Lema 1.1 e alguns resultados do cálculo de probabilidade, decorre que a esperança de uma variável aleatória com a distribuição *log-skew-normal* dada na Expressão (1.12) é

$$\begin{aligned}
E(Y) &= \int_0^\infty \frac{2}{\sigma} \phi\left(\frac{\ln y - \mu}{\sigma}\right) \Phi\left(\alpha\left(\frac{\ln y - \mu}{\sigma}\right)\right) dy \\
&= \int_{-\infty}^\infty \frac{2}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) \Phi\left(\alpha\left(\frac{x - \mu}{\sigma}\right)\right) e^x dx \\
&= \frac{2}{\sigma} e^{(\mu + \frac{\omega^2}{2})} \int_{-\infty}^\infty \phi\left(\frac{x - (\mu + \sigma^2)}{\omega}\right) \Phi\left(\alpha\left(\frac{x - \mu}{\sigma}\right)\right) e^x dx \\
&= 2e^{(\mu + \frac{\sigma^2}{2})} \Phi\left(\frac{\alpha\sigma}{\sqrt{1 + \alpha^2}}\right). \tag{1.14}
\end{aligned}$$

Usando argumentos similares, também se prova que o segundo momento de $Y \sim LSN(\mu, \sigma^2, \alpha)$ é $E(Y^2) = 2e^{(2\mu + 2\sigma^2)} \Phi\left(\frac{2\alpha\sigma}{\sqrt{1 + \alpha^2}}\right)$ e, conseqüentemente, temos que a variância de Y é

$$\text{Var}(Y) = 2e^{(2\mu + \sigma^2)} \left[e^{\sigma^2} \Phi\left(\frac{2\alpha\sigma}{\sqrt{1 + \alpha^2}}\right) - 2 \left[\Phi\left(\frac{\alpha\sigma}{\sqrt{1 + \alpha^2}}\right) \right]^2 \right].$$

No próximo capítulo apresentaremos a família de distribuições log-*skew*-elíptica multivariada introduzida por Marchenko e Genton (2010). Alguns dos conceitos e propriedades discutidos neste capítulo serão utilizados no estudo de tal família.

Capítulo 2

Distribuição Log-*Skew*-Elíptica Multivariada

Nos últimos anos, cresce o interesse em novas famílias de distribuições com parâmetros adicionais que nos permitem regular assimetria e peso das caudas. Generalizações de famílias já existentes na literatura, como aquela definida por Azzalini (1985) apresentada na Expressão (1.3), são feitas para que tenhamos famílias ainda mais flexíveis, capazes de descrever comportamentos de objetos aleatórios em um número maior de situações do que as famílias previamente definidas. Azzalini e Dalla Valle (1996) constroem uma família de distribuições *skew*-normal multivariada, de forma análoga ao que foi feito no caso univariado. Branco e Dey (2001) e Azzalini e Capitanio (2003) introduzem, independentemente, a família de distribuição *skew*-t multivariada, a qual foi construída multiplicando-se a fdp de uma distribuição t-Student n -variada pela fda de uma distribuição t-Student univariada. Como ocorre na distribuição t-Student, na família *skew*-t, o número de graus de liberdade também é um parâmetro que nos permite controlar o peso das caudas. Branco e Dey (2001) vão mais além, construindo a família de distribuições *skew*-elípticas. Outros exemplos de distribuições assimétricas podem ser encontrados em Genton (2004), por exemplo.

O suporte destas distribuições assimétricas pode ser o conjunto de números reais ou \mathbb{R}^n , nos casos multivariados. Para dados que não podem ser negativos, o uso de tais distribuições pode não ser adequado. Para tentar descrever mais apropriadamente o comportamento de dados com este perfil, distribuições de probabilidade que não colocam massa nos números reais negativos são construídas. Marchenko e Genton (2010) introduzem uma família de distribuições log-*skew*-elíptica multivariada e estudam algumas de suas propriedades. Azzalini *et al.* (2003) já havia trabalhado com casos particulares desta família, definindo as famílias de distribuições log-*skew*-normal e a log-*skew*-t univariadas, mas sem explorar formalmente suas características e propriedades.

Neste capítulo, revisaremos algumas famílias de distribuições multivariadas e algumas de suas propriedades. Na Seção 2.1, estudaremos as principais características das distribuições esféricas e elípticas, baseadas em Fang *et al.* (1990). Na Seção 2.2, apresentamos formalmente a família de distribuições *skew*-elíptica multivariada, introduzida por Branco e Dey (2001), restringindo-nos apenas àquela sub-família de distribuições que tem fdp. Na Seção 2.3, apresentaremos a família de distribuições log-*skew*-elíptica multivariada, definida por Marchenko e Genton (2010).

2.1 Distribuições Esféricas e Elípticas

Entre as famílias de distribuições mais flexíveis que a família de distribuições normal estão, por exemplo, as famílias de distribuições esféricas e elípticas. Ambas incluem a família de distribuições normal como caso particular. As distribuições esféricas são aquelas invariantes com respeito a transformações ortogonais. Elas formam uma classe de distribuições com simetria esférica, na qual se encontra a distribuição normal padrão multivariada. Já as distribuições elípticas são obtidas a partir de transformações lineares das distribuições esféricas. Uma maneira de definir as distribuições esféricas pode ser encontrada em Fang *et al.* (1990) e é feita da seguinte forma:

Definição 2.1. *Um vetor aleatório \mathbf{X} de dimensão $n \times 1$ tem distribuição esférica n -variada se, para toda matriz \mathbf{A} ortogonal, de dimensão $n \times n$, tem-se que $\mathbf{X} \stackrel{d}{=} \mathbf{A}\mathbf{X}$, onde $\stackrel{d}{=}$ denota igualdade em distribuição.*

Nem toda distribuição esférica possui fdp. Segundo Fang *et al.* (1990), quando a fdp existe, esta é da forma

$$p_{\mathbf{X}}(\mathbf{x}) = g(\mathbf{x}'\mathbf{x}), \mathbf{x} \in \mathbb{R}^n,$$

onde g é uma função não negativa definida de \mathbb{R} em \mathbb{R}^+ . Além disso, para que $p_{\mathbf{X}}(x)$ seja, de fato, uma fdp n -dimensional, a expressão $r^{n-1}g(r^2)$ deve ser integrável em $[0, \infty)$.

Da Definição 2.1, derivam-se condições, baseadas na função característica, para que um vetor aleatório tenha distribuição esférica. Na demonstração do Teorema 2.1 a seguir utilizaremos o conceito de invariante maximal apresentada na página 15 de Fang *et al.* (1990).

Teorema 2.1. *Um vetor aleatório \mathbf{X} de dimensão $n \times 1$ tem distribuição esférica se, e somente se, a sua função característica $\varphi_{\mathbf{X}}(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^n$, satisfaz a uma das seguintes condições equivalentes:*

- i) $\varphi_{\mathbf{X}}(\mathbf{A}'\mathbf{t}) = \varphi_{\mathbf{X}}(\mathbf{t})$ para toda matriz \mathbf{A} , $n \times n$, ortogonal,
- ii) Existe uma função escalar ψ tal que $\varphi_{\mathbf{X}}(\mathbf{t}) = \psi(\mathbf{t}'\mathbf{t})$.

Demonstração. i) Se \mathbf{X} é esférica, então $\varphi_{\mathbf{X}}(\mathbf{t}) = E(e^{i\mathbf{t}'\mathbf{X}}) = E(e^{i\mathbf{t}'\mathbf{A}\mathbf{X}}) = E(e^{i(\mathbf{A}'\mathbf{t})'\mathbf{X}}) = \varphi_{\mathbf{X}}(\mathbf{A}'\mathbf{t})$. Por outro lado, se $\varphi_{\mathbf{X}}(\mathbf{A}'\mathbf{t}) = \varphi_{\mathbf{X}}(\mathbf{t})$, então $\varphi_{\mathbf{A}\mathbf{X}}(\mathbf{t}) = \varphi_{\mathbf{X}}(\mathbf{t})$ e, pela unicidade da função característica, segue que $\mathbf{A}\mathbf{X} \stackrel{d}{=} \mathbf{X}$.

ii) Admita que exista uma função ψ tal que $\psi(\mathbf{t}'\mathbf{t}) = \varphi_{\mathbf{X}}(\mathbf{t})$. Como \mathbf{A} é uma matriz ortogonal, segue que

$$\varphi_{\mathbf{X}}(\mathbf{A}'\mathbf{t}) = \psi((\mathbf{A}'\mathbf{t})'(\mathbf{A}'\mathbf{t})) = \psi(\mathbf{t}'\mathbf{A}\mathbf{A}'\mathbf{t}) = \psi(\mathbf{t}'\mathbf{t}) = \varphi_{\mathbf{X}}(\mathbf{t}).$$

Logo, por (i), temos que \mathbf{X} tem distribuição esférica. Por outro lado, do resultado dado em (i), temos que $\varphi_{\mathbf{X}}(\mathbf{t})$ é uma função invariante no grupo $O(n)$ (ver Fang *et al.* (1990), Seção 1.3), cujo invariante maximal é $\mathbf{t}'\mathbf{t}$. Por isso, $\varphi_{\mathbf{X}}(\mathbf{t})$ deve ser uma função de $\mathbf{t}'\mathbf{t}$, o que implica (ii). \square

Exemplo 1.1 Seja $\mathbf{X} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. Então \mathbf{X} tem função característica dada por

$$\varphi_{\mathbf{X}}(\mathbf{t}) = e^{\frac{-1}{2}\sigma^2 \mathbf{t}'\mathbf{t}} = \psi(\mathbf{t}'\mathbf{t}), \mathbf{t} \in \mathbb{R}^n.$$

Logo, pelo item (ii) do Teorema 2.1 concluímos que o vetor aleatório \mathbf{X} tem distribuição esférica, com $\psi(\mathbf{t}'\mathbf{t}) = e^{\frac{-1}{2}\sigma^2 \mathbf{t}'\mathbf{t}}$.

Suponha, agora, um vetor aleatório \mathbf{X} de dimensão $m \times 1$ com distribuição esférica e cuja função característica é $\varphi_{\mathbf{X}}(\mathbf{s})$, $\mathbf{s} \in \mathbb{R}^m$. Neste caso, definimos a distribuição elíptica através de uma transformação linear deste vetor como se segue.

Definição 2.2. *Sejam \mathbf{Z} um vetor aleatório de dimensão $n \times 1$ definido pela transformação linear $\mathbf{Z} = \boldsymbol{\mu} + \mathbf{A}\mathbf{X}$, onde \mathbf{A} é uma matriz $n \times m$, $\boldsymbol{\mu} \in \mathbb{R}^n$, e \mathbf{X} é um vetor com distribuição esférica. Então o vetor \mathbf{Z} tem distribuição elíptica n -variada com vetor de locação $\boldsymbol{\mu}$ e matriz de dispersão $\boldsymbol{\Sigma} = \mathbf{A}'\mathbf{A}$, positiva definida, denotada por $\mathbf{Z}_n \sim El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \varphi)$.*

Quando a fdp de \mathbf{Z} existe, segundo Fang *et al.* (1990), ela é dada por

$$p_{\mathbf{Z}}(\mathbf{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} g((\mathbf{z} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{z} - \boldsymbol{\mu})), \mathbf{z} \in \mathbb{R}^n,$$

onde g é uma função não negativa, definida de \mathbb{R}^n em \mathbb{R}^+ . Neste caso, podemos substituir φ pela função geradora de densidade g na nossa notação.

A função característica de \mathbf{Z} pode ser calculada facilmente, como pode ser visto na seguinte proposição:

Proposição 2.1. *Se o vetor aleatório $\mathbf{Z} \sim El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \varphi)$, então a função característica de \mathbf{Z} é da forma*

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = e^{it'\boldsymbol{\mu}} \psi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}), \quad (2.1)$$

para alguma função ψ tal que $\psi(\mathbf{u}) \in \mathbb{R}$, $\mathbf{u} > 0$.

Demonstração. Da Definição 2.2 e usando propriedades da função característica, temos que

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = E(e^{it'(\boldsymbol{\mu} + \mathbf{A}\mathbf{X})}) = e^{it'\boldsymbol{\mu}} E(e^{it'\mathbf{A}\mathbf{X}}) = e^{it'\boldsymbol{\mu}} \varphi_{\mathbf{X}}(\mathbf{A}'\mathbf{t}).$$

Como \mathbf{X} tem distribuição esférica, existe uma função ψ tal que $\varphi_{\mathbf{X}}(\mathbf{t}) = \psi(\mathbf{t}'\mathbf{t})$. Desta forma, podemos escrever $\varphi_{\mathbf{Z}}(\mathbf{t}) = e^{it'\boldsymbol{\mu}} \psi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})$, onde $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'$. □

Exemplo 1.2 Se $\mathbf{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, então sua função característica é

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = e^{it'\boldsymbol{\mu}} e^{\frac{-1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}, \mathbf{t} \in \mathbb{R}^n.$$

Portanto, segue da Proposição 2.1 que \mathbf{Z} é um vetor aleatório com distribuição elíptica, onde $\psi(\mathbf{t}'\mathbf{t}) = e^{\frac{-1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$.

A Figura 2.1 mostra as curvas de nível de duas distribuições normais bivariadas centradas em zero. O gráfico à esquerda é de uma distribuição elíptica cuja matriz de correlação é

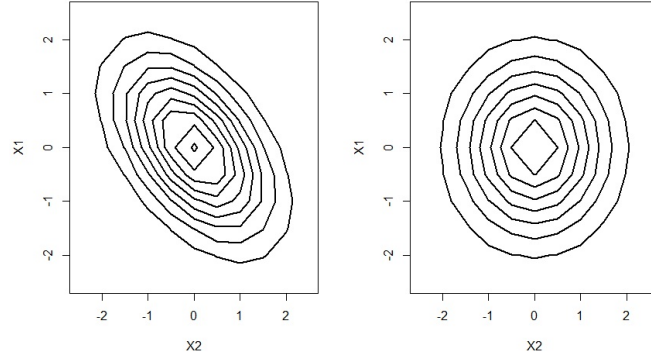


Figura 2.1: Curvas de nível de distribuições normais bivariadas centradas em zero, com $\rho = 0.5$ (esquerda) e $\rho = 0$ (direita).

$\Sigma = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}$ e, o gráfico à direita é de uma distribuição esférica, com $\Sigma = \mathbf{I}_2$.

Muitas das propriedades associadas à família de distribuições normal também são estendidas para a família de distribuições elípticas. Vamos enumerar algumas destas propriedades que julgamos serem importantes para o desenvolvimento deste trabalho e que podem ser encontradas em Fang *et al.* (1990).

Assuma que $\mathbf{X} \sim El_n(\boldsymbol{\mu}, \Sigma, \varphi)$, onde $\boldsymbol{\mu} \in \mathbb{R}^n$, Σ é uma matriz de dispersão, $n \times n$, não-nula e com $\text{posto}(\Sigma) = k \leq n$, e φ é a função característica que define a família de distribuições elíptica de interesse. Para esta família, as seguintes propriedades são verificadas:

1. $E(\mathbf{X}) = \boldsymbol{\mu}$ e $\text{Cov}(\mathbf{X}) = \kappa \Sigma$, onde $\kappa = 2\varphi'(0)$ e $\varphi'(0) = \left. \frac{d\varphi(u)}{du} \right|_{u=0}$, desde que existam.
2. Sejam $\boldsymbol{\eta} \in \mathbb{R}^m$ um vetor e \mathbf{A} uma matriz $m \times n$. Então, a transformação linear $\mathbf{Y} = \boldsymbol{\eta} + \mathbf{A}\mathbf{X} \sim El_m(\boldsymbol{\eta} + \mathbf{A}\boldsymbol{\mu}, \mathbf{A}\Sigma\mathbf{A}', \varphi)$.
3. Todas as distribuições marginais de \mathbf{X} pertencem à família de distribuições elíptica. Considere a partição de \mathbf{X} , $\boldsymbol{\mu}$ e Σ como sendo

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \text{ e } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

onde \mathbf{X}_1 e $\boldsymbol{\mu}_1$ são vetores de dimensão $m \times 1$ e Σ_{11} é matriz $m \times m$ com $1 \leq m < n$. Então, $\mathbf{X}_1 \sim El_m(\boldsymbol{\mu}_1, \Sigma_{11}, \varphi)$.

4. Assumindo a partição definida na Propriedade 3, a distribuição condicional de \mathbf{X}_1 dado $\mathbf{X}_2 = \mathbf{x}_2$, é a seguinte distribuição elíptica m -variada:

$$\mathbf{X}_1 \mid \mathbf{X}_2 = \mathbf{x}_2 \sim El_m(\boldsymbol{\mu}_1(\mathbf{x}_2), \Sigma_{11.2}, g_q(x_2)),$$

onde $\boldsymbol{\mu}_1(\mathbf{x}_2) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$, $\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{12}$ e $g_{q(\mathbf{x}_2)}$ é uma função de \mathbf{x}_2 através da forma quadrática $q(\mathbf{x}_2) = (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$, no caso em que a fdp de \mathbf{X} existe.

2.2 Distribuição *Skew*-Elíptica Multivariada

Com o objetivo de generalizar a classe das distribuições elípticas de forma a acomodar distribuições assimétricas com caudas mais pesadas, Branco e Dey (2001) introduzem as distribuições *skew*-elípticas multivariadas. A assimetria é introduzida pela multiplicação da fdp de uma variável aleatória n -dimensional com distribuição elíptica multivariada pela fda de uma variável com a mesma distribuição elíptica univariada.

Como na Seção 2.1, denotaremos a família de distribuições elíptica n -variada por $El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)})$, onde $g^{(n)}(u)$ é função geradora de densidade que identifica cada família. Nos casos em que a fdp existe, temos que esta será da forma $f_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)}) = |\boldsymbol{\Sigma}|^{-1/2} g^{(n)}((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}))$, $\mathbf{x} \in \mathbb{R}^n$ (Fang *et al.* (1990)). Seguindo Branco e Dey (2001), a família de distribuições *skew*-elípticas que possui fdp é definida como se segue (ver também Genton (2004)).

Definição 2.3. *Um vetor aleatório \mathbf{X} , de dimensão $n \times 1$, tem distribuição *skew*-elíptica multivariada, com vetor de forma $\boldsymbol{\alpha} \in \mathbb{R}^n$, vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^n$ e matriz de dispersão $\boldsymbol{\Sigma}$ de ordem $n \times n$, se sua fdp é dada por*

$$f_{SE_n}(\mathbf{x}) = 2f_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)})F(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\mathbf{x} - \boldsymbol{\mu}); g_{Q(\mathbf{x})}), \quad \mathbf{x} \in \mathbb{R}^n, \quad (2.2)$$

onde $\boldsymbol{\omega} = \text{diag}(\boldsymbol{\Sigma})^{1/2}$, $f_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)})$ denota a fdp de um vetor aleatório n -variado com distribuição $El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)})$ e $F(u; g_{Q(\mathbf{x})})$ é a fda de uma variável aleatória com distribuição $El_1(0, 1, g_{Q(\mathbf{x})})$, onde função geradora é definida por $g_{Q(\mathbf{x})}(u) = g^{(n+1)}(u + Q(\mathbf{x}))/g^{(n)}(Q(\mathbf{x}))$ e $Q(\mathbf{x}) = Q(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})$. Esta distribuição é denotada por $SEl_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, g^{(n+1)})$.

Dois casos particulares importantes desta família são as famílias de distribuição *skew*-normal multivariada também e *skew*-t multivariada, encontradas, respectivamente, em Azzalini e Dalla-Valle (1996) e Azzalini e Capitanio (2003).

A construção de um vetor aleatório \mathbf{Z} com distribuição *skew*-normal multivariada partindo da Equação (2.2) é feita ao tomarmos $g(u) = c \exp(-u/2)$, onde c é uma constante positiva. Neste caso, $g_{Q(\mathbf{x})}(u) = c \exp(-(u + Q(\mathbf{x}))/2)/c \exp(-Q(\mathbf{x})/2) = \exp(-u/2)$ e chegamos à seguinte fdp:

$$f_{SN_n}(\mathbf{z}) = 2\phi_n(\mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\mathbf{z} - \boldsymbol{\mu})), \quad \mathbf{z} \in \mathbb{R}^n, \quad (2.3)$$

onde $\phi_n(\cdot)$ e $\Phi(\cdot)$ são definidas como no Capítulo 1. Denotamos a família de distribuições *skew*-normal multivariada por $\mathbf{Z} \sim SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$.

A Figura (2.2) apresenta as curvas de nível de distribuições *skew*-normal bivariadas padrão, considerando graus diferentes de assimetria, e também nos mostra como a escolha do vetor $\boldsymbol{\alpha}$ influencia na assimetria da curva. Quando as entradas de $\boldsymbol{\alpha}$ são negativas, a curva é deformada colocando mais massa no quadrante negativo e, quando suas entradas são positivas, a massa se concentra no quadrante positivo. Vemos também que quanto maior as entradas do vetor em módulo, maior é a deformação da curva.

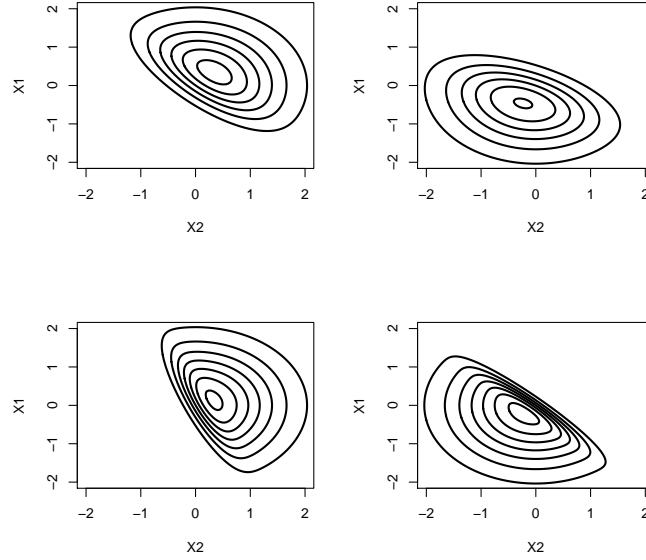


Figura 2.2: Curvas de nível de distribuições skew-normais bivariadas padrões, com $\boldsymbol{\alpha} = (2, 2)'$ (esquerda superior), $\boldsymbol{\alpha} = (-1, -2)'$ (direita superior), $\boldsymbol{\alpha} = (5, 2)'$ (esquerda inferior) e $\boldsymbol{\alpha} = (-5, -5)'$ (direita inferior).

De forma análoga se dá a construção de um vetor aleatório \mathbf{W} com distribuição *skew-t* multivariada. Para isto, tomamos $g(u) = c \left(1 + \frac{u}{\nu}\right)^{-(\nu+n)/2}$, onde c é uma constante positiva e $\nu > 0$. Desta forma, a fdp de \mathbf{W} é dada por

$$f_{LST_n}(\mathbf{w}) = 2t_n(\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) T \left[\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} (\mathbf{w} - \boldsymbol{\mu}) \left(\frac{\nu + n}{\nu + Q(\mathbf{w})} \right)^{1/2}; \nu + d \right], \mathbf{w} \in \mathbb{R}^n, \quad (2.4)$$

onde $t_n(\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \Gamma\{(\nu + n)/2\} [1 + (\mathbf{w} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{w} - \boldsymbol{\mu}) / \nu]^{-(\nu+n)/2} \{|\boldsymbol{\Sigma}|^{1/2} (\nu\pi)^{n/2} \Gamma(\nu/2)\}^{-1}$ é a fdp de uma distribuição t-Student n -variada com ν graus de liberdade, com vetor de forma $\boldsymbol{\alpha} \in \mathbb{R}^n$, vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^n$, matriz de escala $\boldsymbol{\Sigma}$ de ordem $n \times n$, e $T(\cdot; \nu + n)$ denota a função de distribuição acumulada da distribuição t-Student padrão univariada com $n + \nu$ graus de liberdade. Denotamos a família de distribuições *skew-t* multivariada, dada na Expressão (2.4), por $ST_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \nu)$. Esta notação será usada ao longo de todo este texto.

2.3 Distribuição Log-Skew-Elíptica Multivariada

Em algumas situações práticas, estamos interessados na modelagem do comportamento de dados que assumem apenas valores positivos. Já vimos na Seção 1.2 que as famílias de distribuições log-normal e log-skew-normal univariadas definidas nas Expressões (1.12) e (1.13) satisfazem tal requisito. Com o objetivo de criar uma classe de distribuições multivariada que não coloca massa em valores negativos, geral e flexível, e que é capaz de descrever diferentes graus de assimetria, Marchenko e Genton (2010), partindo da família *skew-elíptica* definida por Branco

e Dey (2001), introduzem a família de distribuições log-*skew*-elíptica multivariada, a qual vamos rever nesta seção.

Para este fim, defina $\ln(\mathbf{X}) = (\ln(X_1), \dots, \ln(X_n))'$ e $\exp(\mathbf{X}) = (\exp(X_1), \dots, \exp(X_n))'$, respectivamente, como sendo o logaritmo e a exponencial do vetor aleatório $\mathbf{X} = (X_1, \dots, X_n)'$.

Definição 2.4. Dizemos que o vetor aleatório \mathbf{X} tem distribuição log-*skew* elíptica com vetor de forma $\boldsymbol{\alpha} \in \mathbb{R}^n$, vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^n$, matriz de escala $\boldsymbol{\Sigma}$ de dimensão $n \times n$ e função geradora de densidade $g^{(n+1)}$, denotada por $LSE\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, g^{(n+1)})$, se, e somente se, $\mathbf{Y} = \ln \mathbf{X} \sim SE\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, g^{(n+1)})$ com fdp dada na Expressão (2.2).

Através da Expressão (2.2) e do método jacobiano, obtemos que a fdp de \mathbf{X} é da forma

$$f_{LSE_n}(\mathbf{x}) = 2 \left(\prod_{i=1}^n x_i \right)^{-1} f_n(\ln \mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)}) F(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} (\ln \mathbf{x} - \boldsymbol{\mu}); g_Q(\mathbf{x})), \mathbf{x} > 0. \quad (2.5)$$

Quando tomamos $\boldsymbol{\alpha} = \mathbf{0}$, a família de distribuições dada na Expressão (2.5) se reduz à família de distribuições log-elíptica.

Da Definição 2.4, obtemos a fdp da distribuição log-*skew*-normal multivariada se considerarmos $g^{(n)}$ como a função geradora da distribuição normal, como na Expressão (2.3). Esta família de distribuições é denotada por $LSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ e sua fdp é dada por

$$f_{LSN_n}(\mathbf{x}) = 2 \left(\prod_{i=1}^n x_i \right)^{-1} \phi_n(\ln \mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} (\ln \mathbf{x} - \boldsymbol{\mu})), \mathbf{x} > 0, \quad (2.6)$$

onde os parâmetros são os mesmos especificados na Expressão (2.3).

Quando $\boldsymbol{\alpha} = \mathbf{0}$, a Expressão (2.6) se reduz a

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\prod_{i=1}^n x_i \right)^{-1} \phi_n(\ln \mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}), \mathbf{x} > 0, \quad (2.7)$$

que é a fdp de um vetor da família de distribuições log-normal multivariada com vetor de locação $\boldsymbol{\mu}$ de dimensão $n \times 1$ e matriz de escala positiva definida $\boldsymbol{\Sigma}$ de dimensão $n \times n$, denotado por $\mathbf{X} \sim LN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

A Figura 2.3 mostra os gráficos com as curvas de níveis da fdp de algumas distribuições log-*skew*-normal padrão bivariadas considerando diferentes valores para o parâmetro de assimetria $\boldsymbol{\alpha}$. Notamos que quanto maior o valor das entradas de $\boldsymbol{\alpha}$, maior a dispersão. Quando a assimetria é positiva, a tendência é termos uma distribuição colocando massa quase nula em valores próximos de zero. Quando a assimetria é negativa, há concentração de massa próximo da origem.

Da mesma forma, se considerarmos $g^{(n+1)}$ como a função geradora da família de distribuições t-Student, como na Expressão (2.4), obtemos a distribuição log-*skew*-t, denotada por $LST_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \nu)$, cuja fdp é dada por

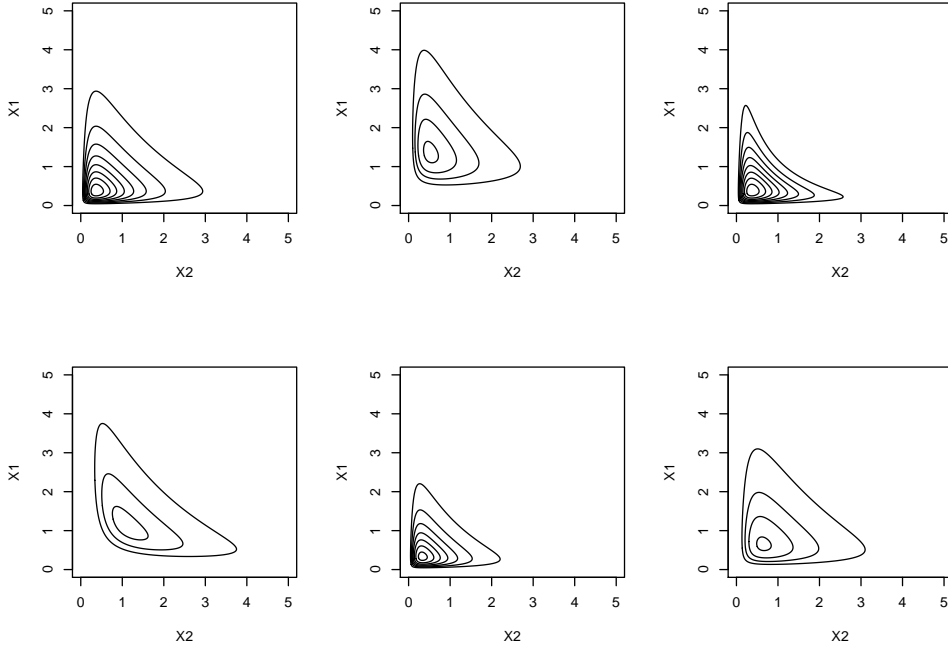


Figura 2.3: Curvas de nível da distribuição log-*skew*-normais padrão bivariadas com $\boldsymbol{\alpha} = (0, 0)'$ (esquerda superior), $\boldsymbol{\alpha} = (0.5, 2)'$ (centro superior), $\boldsymbol{\alpha} = (-2, -2)'$ (direita superior), $\boldsymbol{\alpha} = (2, 2)'$ (esquerda inferior), $\boldsymbol{\alpha} = (-0.5, -0.5)'$ (centro inferior) e $\boldsymbol{\alpha} = (0.5, 0.5)'$ (direita inferior).

$$f_{LST_n}(x) = 2 \left(\prod_{i=1}^n x_i \right)^{-1} t_n(\ln \mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) T \left(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} (\ln \mathbf{x} - \boldsymbol{\mu}) \left[\frac{\nu + n}{\nu + Q(\ln \mathbf{x})} \right]^{1/2}; \nu + n \right), \mathbf{x} > 0,$$

onde os parâmetros são os mesmos especificados na Expressão (2.4).

Exploraremos algumas propriedades que nos ajudarão a compreender o comportamento das distribuições log-*skew*-elípticas. Tais propriedades foram apresentadas em Marchenko e Genton (2010), onde suas demonstrações podem ser encontradas.

Proposição 2.2. *Se $\mathbf{X} \sim LSE\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, g^{(n+1)})$ e se os momentos mistos existem para $\mathbf{t} = (t_1, \dots, t_d)'$, $t_i \in \mathbb{N}$, então*

$$E\left(\prod_{i=1}^n X_i^{t_i}\right) = M_{\ln \mathbf{X}}(\mathbf{t}),$$

onde $M_{\ln \mathbf{X}}(\cdot)$ é a função geradora de momentos da variável aleatória $\ln \mathbf{X}$, que tem distribuição $SE\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, g^{(n+1)})$.

Uma consequência importante da Proposição 2.2 é que os momentos de primeira e segunda ordem de uma variável aleatória com distribuição log-*skew*-normal são facilmente obtidos. Seja $\mathbf{X} \sim LSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$. Então, de resultados obtidos por Genton (2004) para a distribuição *skew*-normal e da Proposição 2.2, segue que

$$E\left(\prod_{i=1}^n X_i^{t_i}\right) = 2 \exp(\boldsymbol{\mu}'\mathbf{t} + \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2) \Phi(\boldsymbol{\alpha}'\bar{\boldsymbol{\Sigma}}\boldsymbol{\omega}\mathbf{t}/(1 + \boldsymbol{\alpha}'\bar{\boldsymbol{\Sigma}}\boldsymbol{\alpha})^{1/2}),$$

onde $\bar{\boldsymbol{\Sigma}} = \boldsymbol{\omega}^{-1}\boldsymbol{\Sigma}\boldsymbol{\omega}^{-1}$ e $\boldsymbol{\omega} = \text{diag}(\boldsymbol{\Sigma})^{1/2}$. Como consequência, temos que os dois primeiros momentos de uma variável aleatória $X \sim LSN(\mu, \omega^2, \alpha)$ são

$$E(X) = 2 \exp(\mu + \omega^2/2) \Phi(\alpha\omega/(1 + \alpha^2)^{1/2}),$$

$$E(X^2) = 2 \exp(2\mu + 2\omega^2) \Phi(2\alpha\omega/(1 + \alpha^2)^{1/2}).$$

Algumas características da família de distribuições log-*skew*-elípticas são herdadas da família de distribuições elíptica como, por exemplo, o fato de que se um vetor tem distribuição log-*skew*-elíptica, então as distribuições marginais de todos os seus componentes também têm distribuições log-*skew*-elípticas. Este resultado é formalmente apresentado na seguinte proposição.

Proposição 2.3. *Considere as seguintes partições dos vetores \mathbf{X} , $\boldsymbol{\mu}$ e $\boldsymbol{\alpha}$ e da matriz de escala $\boldsymbol{\Sigma}$*

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix} \text{ e } \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

onde \mathbf{X}_1 , $\boldsymbol{\alpha}_1$ e $\boldsymbol{\mu}_1$ são vetores de dimensão $m \times 1$, \mathbf{X}_2 , $\boldsymbol{\alpha}_2$ e $\boldsymbol{\mu}_2$ são vetores de dimensão $(n - m) \times 1$ e $\boldsymbol{\Sigma}_{11}$ e $\boldsymbol{\Sigma}_{22}$ são, respectivamente, matrizes de ordens $m \times m$ e $(n - m) \times (n - m)$, com $1 \leq m < n$. Defina as matrizes quadradas $\boldsymbol{\omega}_1 = \text{diag}(\omega_{11}^{1/2}, \dots, \omega_{mm}^{1/2})$ e $\boldsymbol{\omega}_2 = \text{diag}(\omega_{m+1, m+1}^{1/2}, \dots, \omega_{nn}^{1/2})$ de ordens $m \times m$ e $(n - m) \times (n - m)$, respectivamente. Então, $\mathbf{X}_1 \sim LSE_m(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \boldsymbol{\alpha}_1^*, g^{n+1})$ e $\mathbf{X}_2 \sim LSE_{n-m}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22}, \boldsymbol{\alpha}_2^*, g^{n+1})$, onde

$$\boldsymbol{\alpha}_1^* = \frac{\boldsymbol{\alpha}_1 + \boldsymbol{\omega}_1 \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}^{-1} \boldsymbol{\omega}_2^{-1} \boldsymbol{\alpha}_2}{(1 + \boldsymbol{\alpha}_2' \bar{\boldsymbol{\Sigma}}_{22.1} \boldsymbol{\alpha}_2)^{1/2}},$$

$$\boldsymbol{\alpha}_2^* = \frac{\boldsymbol{\alpha}_2 + \boldsymbol{\omega}_2 \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}^{-1} \boldsymbol{\omega}_1^{-1} \boldsymbol{\alpha}_1}{(1 + \boldsymbol{\alpha}_1' \bar{\boldsymbol{\Sigma}}_{11.2} \boldsymbol{\alpha}_1)^{1/2}},$$

$$\bar{\boldsymbol{\Sigma}}_{22.1} = \boldsymbol{\omega}_2^{-1} (\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}) \boldsymbol{\omega}_2^{-1},$$

$$\bar{\boldsymbol{\Sigma}}_{11.2} = \boldsymbol{\omega}_1^{-1} (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}) \boldsymbol{\omega}_1^{-1}.$$

Apesar de ser fechada sob marginalização, a família de distribuições log-*skew*-elíptica não é fechada sob condicionamento. Para determinarmos a distribuição de $\mathbf{X}_2 \mid \mathbf{X}_1$, precisaremos da definição da família de distribuições *skew*-elíptica estendida (Arellano-Valle e Genton (2009), Arellano-Valle e Azzalini (2006), Arellano-Vale *et al.*(2006)), dada a seguir:

Definição 2.5. Um vetor aleatório \mathbf{X} tem distribuição *skew*-elíptica estendida multivariada, denotada por $\mathbf{X} \sim SEE_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \tau; g^{(n+1)})$, se sua densidade é da forma

$$f_{SEE_n}(\mathbf{x}) = \frac{1}{F(\tau/\sqrt{1 + \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}}, g^{(n)})} f_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)}) F(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\mathbf{x} - \boldsymbol{\mu}) + \tau; g_Q(\mathbf{x})), \mathbf{x} \in \mathbb{R}^n, \quad (2.8)$$

onde $\tau \in \mathbb{R}$ e os demais parâmetros são definidos como na Expressão (2.2).

A partir da Expressão (2.8), Marchenko e Genton (2010) definem a família de distribuições log-*skew*-elíptica estendida, denotada por $\mathbf{X} \sim LSEE_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \tau; g^{(n+1)})$ e cuja fdp é

$$f_{LSEE_n}(\mathbf{x}) = \left(\prod_{i=1}^n x_i \right)^{-1} \frac{1}{F(\tau/\sqrt{1 + \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}}, g^{(n)})} f_n(\ln \mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n)}) F(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\ln \mathbf{x} - \boldsymbol{\mu}) + \tau; g_Q(\mathbf{x})), \mathbf{x} \in \mathbb{R}^{n+}.$$

Proposição 2.4. Seja $\mathbf{X} \sim LSEE_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \tau; g^{(n+1)})$. Considere as partições definidas na Proposição 2.3. Então, a distribuição de \mathbf{X}_2 dado $\mathbf{X}_1 = \mathbf{x}_1$ é

$$\mathbf{X}_2 \mid \mathbf{X}_1 = \mathbf{x}_1 \sim LSEE_n(\boldsymbol{\mu}_{2,1}, \boldsymbol{\Sigma}_{22,1}, \boldsymbol{\alpha}_2, \tau; g_{Q(\mathbf{x}_1)}^{(n+1)}),$$

onde $\boldsymbol{\mu}_{2,1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)$, $\boldsymbol{\Sigma}_{22,1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$, $\tau = (\boldsymbol{\alpha}_1 + \boldsymbol{\omega}_1\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\omega}_2^{-1}\boldsymbol{\alpha}_2)'\boldsymbol{\omega}_1^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)$ e $g_{Q(\mathbf{x}_1)}^{(n+1)}$ é definido como na Expressão (2.2).

No próximo capítulo, apresentaremos a família de distribuições *skew*-normal fundamental canônica multivariada, definida por Arellano e Genton (2005), e definiremos uma nova família de distribuições denotada por log-*skew*-normal fundamental canônica multivariada, obtida pela transformação adequada de um vetor com distribuição *skew*-normal fundamental canônica multivariada. Também exploraremos as principais propriedades desta nova família.

Capítulo 3

Distribuição Log-*Skew*-Normal Fundamental Multivariada

Existem famílias de distribuições assimétricas muito mais gerais do que aquelas definidas por Azzalini (1985) ou por Marchenko e Genton (2010) e que foram revistas nos Capítulos 1 e 2. Generalizações dos processos de assimetrizar funções de densidade de probabilidade foram propostas por diversos autores, entre os quais citamos Genton e Loperfido (2005) que introduziram uma classe mais geral de distribuições *skew*-elípticas com fdp da forma

$$f(\mathbf{z} \mid Q) = 2f_{0_n}(\mathbf{z})Q(\mathbf{z}), \mathbf{z} \in \mathbb{R}^n, \quad (3.1)$$

onde f_{0_n} é a fdp de uma variável com distribuição elíptica n -dimensional (ver Capítulo 2) e Q é a função responsável por introduzir assimetria e tal que $Q(\mathbf{z}) \geq 0$ e $Q(-\mathbf{z}) = 1 - Q(\mathbf{z})$ para todo $\mathbf{z} \in \mathbb{R}^n$. Note que a Expressão (3.1) generaliza a família introduzida por Branco e Dey (2001), uma vez que Q não é necessariamente a fda univariada associada à família a qual pertence f_{0_n} . Outras discussões sobre famílias de distribuições mais gerais podem ser encontradas em Gonzáles-Farías *et al.* (2004), onde também se define a família de distribuições normal assimétrica canônica, cuja fdp é dada por

$$f(\mathbf{z} \mid \Phi_m, \mathbf{D}) = \frac{\phi_n(\mathbf{z})\Phi_m(\mathbf{D}\mathbf{z})}{\Phi_m(\mathbf{0} \mid \mathbf{I}_m + \mathbf{D}\mathbf{D}')} , \mathbf{z} \in \mathbb{R}^n, \quad (3.2)$$

onde \mathbf{D} é uma matriz de dimensão $m \times n$. A família de distribuições definida na Expressão (3.2) contém as famílias definidas na Expressão (1.3), quando tomamos $m = 1$ e $n = 1$, e na Expressão (2.3), quando tomamos $m = 1$.

Uma maneira ainda mais geral de obtermos a fdp de uma distribuição assimétrica é dada por

$$f(\mathbf{z} \mid Q_m) = K_m^{-1} f_{0_n}(\mathbf{z})Q_m(\mathbf{z}), \mathbf{z} \in \mathbb{R}^n, \quad (3.3)$$

onde $K_m = P(\mathbf{X} > 0)$, $Q_m(\mathbf{z}) = P(\mathbf{X} > 0 \mid \mathbf{Z} = \mathbf{z})$ e \mathbf{X} e \mathbf{Z} são vetores com dimensões $m \times 1$ e $n \times 1$, respectivamente, com distribuição conjunta tal que a distribuição marginal de \mathbf{Z} é f_{0_n} . Este procedimento de construção de distribuições assimétricas é conhecido como mecanismo

de seleção e foi amplamente discutido em Arellano-Valle *et al.*(2006). Este mecanismo é utilizado por Arellano-Valle e Genton (2005) para construir a família de distribuições *skew-normal* fundamental canônica.

Neste capítulo, faremos uma revisão da distribuição *skew-normal* fundamental canônica, denotada por CFUSN, definida em Arellano-Valle e Genton (2005). Esta família pode ser vista como uma generalização da família *skew-normal* multivariada definida na Expressão (2.3) do Capítulo 2. Partindo desta família e seguindo Marchenko e Genton (2010), definiremos uma nova família de distribuições chamada família log-*skew-normal* fundamental canônica, que será denotada por LCFUSN. A família LCFUSN é obtida através da transformação apropriada da família de distribuições CFUSN. Além disso, estudaremos algumas propriedades e características desta família, tais como sua função geradora de momentos, suas distribuições marginais e condicionais e critérios para independência.

Na Seção 3.1 apresentaremos a definição da família CFUSN considerando o caso padrão em que o vetor de médias $\boldsymbol{\mu}$ de dimensão $n \times 1$ é igual a zero e matriz de dispersão $\boldsymbol{\Sigma}$ de dimensão $n \times n$ é igual à identidade e suas propriedades mais relevantes. A seguir, definiremos uma transformação que gera uma distribuição CFUSN com vetor de médias $\boldsymbol{\mu}$ e matriz positiva definida de escala $\boldsymbol{\Sigma}$. Na Seção 3.2 apresentaremos uma das principais contribuições deste trabalho. Considerando a família de distribuições CFUSN e a transformação de variáveis usada, por exemplo, em Marchenko e Genton (2010), construiremos uma família de distribuições que coloca massa apenas em valores positivos e que generaliza a família de distribuições log-*skew-normal* multivariada definida no Capítulo 2, Expressão (2.6). Nesta seção também estabeleceremos algumas de suas propriedades.

3.1 Distribuição *Skew-Normal* Fundamental Multivariada

Arellano-Valle e Genton (2005) introduzem uma classe mais ampla de distribuições *skew-normal*, chamada de fundamental, a qual é denotada por FUSN. Esta nova família de distribuições é construída utilizando o mecanismo de seleção discutido em Arellano-Valle *et al.*(2006) dado na Expressão (3.3). O uso do mecanismo de seleção permite a construção de uma nova distribuição de probabilidade por truncamento e, para a família FUSN, é definido a seguir.

Definição 3.1. *Seja $\mathbf{Z}^* = [\mathbf{Z} | \mathbf{X} > 0]$, onde \mathbf{Z} é um vetor aleatório $n \times 1$, tal que $\mathbf{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e \mathbf{X} é um vetor aleatório $m \times 1$. Dizemos que \mathbf{Z}^* tem uma distribuição *skew-normal* fundamental n -variada (FUSN), denotada por $\mathbf{Z}^* \sim FUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, Q_m)$, se sua fdp é da forma*

$$f_{\mathbf{Z}^*}(z) = K_m^{-1} \phi_n(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) Q_m(\mathbf{z}), \mathbf{z} \in \mathbb{R}^n, \quad (3.4)$$

onde $Q_m(\mathbf{z}) = P(\mathbf{X} > 0 | \mathbf{Z} = \mathbf{z})$ e $K_m^{-1} = E(Q_m(\mathbf{Z})) = P(\mathbf{X} > 0)$.

É possível verificar que, através do condicionamento $[\mathbf{Z} | \mathbf{X} > 0]$, realmente obtemos a fdp dada na Expressão (3.4). De maneira geral, usando o mecanismo de seleção, segue que a fda de \mathbf{Z}^* é

$$\begin{aligned}
F_{\mathbf{Z}^*}(\mathbf{z}) &= P(\mathbf{Z}^* \leq \mathbf{z}) \\
&= \frac{P(\mathbf{Z} \leq \mathbf{z}, \mathbf{X} > 0)}{P(\mathbf{X} > 0)} \\
&= \frac{\int_{-\infty}^{\mathbf{z}} \int_0^{\infty} f_{\mathbf{Z}, \mathbf{X}}(\mathbf{t}, \mathbf{w}) d\mathbf{w} dt}{P(\mathbf{X} > 0)} \\
&= \frac{\int_{-\infty}^{\mathbf{z}} \int_0^{\infty} f_{\mathbf{X}|\mathbf{Z}=\mathbf{t}}(\mathbf{w}) f_{\mathbf{Z}}(\mathbf{t}) d\mathbf{w} dt}{P(\mathbf{X} > 0)} \\
&= \frac{\int_{-\infty}^{\mathbf{z}} f_{\mathbf{Z}}(\mathbf{t}) \left[\int_0^{\infty} f_{\mathbf{X}|\mathbf{Z}=\mathbf{t}}(\mathbf{w}) d\mathbf{w} \right] dt}{P(\mathbf{X} > 0)} \\
&= \frac{\int_{-\infty}^{\mathbf{z}} f_{\mathbf{Z}}(\mathbf{t}) P(\mathbf{X} > 0 | \mathbf{Z} = \mathbf{t}) dt}{P(\mathbf{X} > 0)}. \tag{3.5}
\end{aligned}$$

Derivando a função de distribuição acumulada $F_{\mathbf{Z}^*}(\mathbf{z})$ na Expressão (3.5), temos, como consequência, que a fdp de \mathbf{Z}^* é

$$f_{\mathbf{Z}^*}(\mathbf{z}) = f_{\mathbf{Z}}(\mathbf{z}) \frac{P(\mathbf{X} > 0 | \mathbf{Z} = \mathbf{z})}{P(\mathbf{X} > 0)}.$$

Então, se $\mathbf{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, segue que $f_{\mathbf{Z}^*}(\mathbf{z}) = \phi_n(\mathbf{z}) P(\mathbf{X} > 0 | \mathbf{Z} = \mathbf{z}) [P(\mathbf{X} > 0)]^{-1}$.

Um caso particular da distribuição definida na Expressão (3.4) é obtido quando tomamos o vetor $\begin{pmatrix} \mathbf{Z} \\ \mathbf{X} \end{pmatrix} \sim N_{n+m} \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{I}_n & \boldsymbol{\Delta} \\ \boldsymbol{\Delta} & \mathbf{I}_m \end{pmatrix} \right)$, onde $\boldsymbol{\Delta}$ é uma matriz $n \times m$ tal que $\|\boldsymbol{\Delta}\mathbf{a}\| < 1$, para todo vetor unitário $\mathbf{a} \in \mathbb{R}^m$. Conseqüentemente, $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ e $\mathbf{X} | \mathbf{Z} = \mathbf{z} \sim N_m(\boldsymbol{\Delta}'\mathbf{z}, \mathbf{I}_m - \boldsymbol{\Delta}'\boldsymbol{\Delta})$. A partir destas distribuições, obtemos a versão canônica da família de distribuições FUSN, introduzida por Arellano e Genton (2005), que é definida como se segue.

Definição 3.2. Dizemos que \mathbf{Z}^* tem distribuição skew-normal fundamental canônica n -variada (CFUSN) com matriz de assimetria $\boldsymbol{\Delta}$ de ordem $n \times m$, denotada por $\mathbf{Z}^* \sim CFUSN_{n,m}(\boldsymbol{\Delta})$, se sua fdp é da forma

$$f_{\mathbf{Z}^*}(\mathbf{z}) = 2^m \phi_n(\mathbf{z}) \Phi_m(\boldsymbol{\Delta}'\mathbf{z} | \mathbf{I}_m - \boldsymbol{\Delta}'\boldsymbol{\Delta}), \mathbf{z} \in \mathbb{R}^n, \tag{3.6}$$

onde $\boldsymbol{\Delta}$ é tal que $\|\boldsymbol{\Delta}\mathbf{a}\| < 1$, para todo vetor unitário $\mathbf{a} \in \mathbb{R}^m$.

A família definida na Expressão (3.6) unifica algumas classes de distribuições skew-normal conhecidas na literatura. Por exemplo, a família skew-normal definida na Expressão (1.4) por Azzalini (1985) é obtida da Expressão (3.6), se considerarmos $\mu = 0$, $\omega = 1$ e $n=m=1$. Neste caso, apenas reparametrizamos a família, assumindo $\boldsymbol{\Delta} = \delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$. Como a relação entre $\alpha \in \mathbb{R}$ e $\delta \in (-1, 1)$ é injetiva, temos que a família de distribuições skew-normal univariada, com fdp dada por $f(z) = 2\phi(z)\Phi(\alpha z)$, $z \in \mathbb{R}$, $\alpha \in \mathbb{R}$, e a família de distribuições skew-normal fundamental canônica univariada, com fdp dada por $f(z) = 2\phi(z)\Phi(\delta z | 1 - \delta^2)$, $z \in \mathbb{R}$, $|\delta| < 1$, são equivalentes.

Além disto, note que a matriz Δ pode ser construída como sendo $\Delta = \Lambda(\mathbf{I}_m + \Lambda'\Lambda)^{-1/2}$ para alguma matriz real Λ de dimensão $n \times m$ e com entradas finitas. Neste caso, temos a identidade $\mathbf{I}_m - \Delta'\Delta = (\mathbf{I}_m + \Lambda'\Lambda)^{-1}$ e podemos reescrever a fdp associada à família de distribuições CFUSN da seguinte forma:

$$f_{\mathbf{z}^*}(\mathbf{z}) = 2^m \phi_n(\mathbf{z}) \Phi_m((\mathbf{I}_m + \Lambda'\Lambda)^{-1} \Lambda' \mathbf{z} | (\mathbf{I}_m + \Lambda'\Lambda)^{-1}), \mathbf{z} \in \mathbb{R}^n. \quad (3.7)$$

A Figura 3.1 mostra a fdp de uma variável aleatória com distribuição na família CFUSN quando $n = 1$ e quando consideramos quatro valores distintos de m . Percebemos que é possível obter diferentes graus de assimetria nesta distribuição modificando a dimensão m .

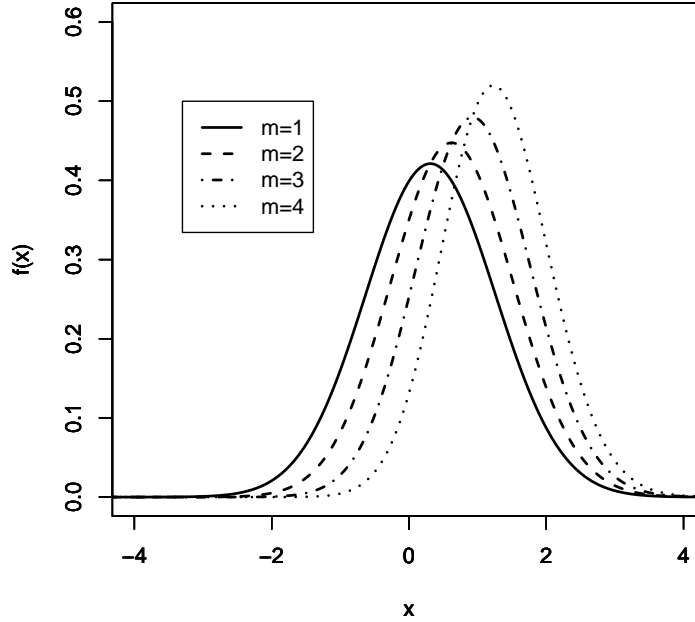


Figura 3.1: Função densidade de probabilidade da CFUSN com $n=1$, para $m=1$ e $\Delta = 0.4$, $m=2$ e $\Delta = (0.4, 0.4)$, $m=3$ e $\Delta = (0.4, 0.4, 0.4)$ e $m=4$ e $\Delta = (0.4, 0.4, 0.4, 0.4)$.

As Figuras 3.2 e 3.3 mostram os gráficos com as curvas de nível de um vetor aleatório com distribuição CFUSN quando consideramos $n = 2$ e $m = 2$ e $m = 3$, respectivamente, e seis valores distintos do parâmetro de assimetria Δ . A primeira figura mostra as curvas de nível da distribuição CFUSN bivariada com $m=2$, para $\Delta = \begin{pmatrix} 0.3 & 0.3 \\ 0.3 & 0.3 \end{pmatrix}$ (esquerda superior), $\Delta = \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}$ (centro superior), $\Delta = \begin{pmatrix} 0.4 & 0.8 \\ 0.3 & 0.3 \end{pmatrix}$ (direita superior), $\Delta = \begin{pmatrix} -0.3 & -0.3 \\ -0.3 & -0.3 \end{pmatrix}$ (esquerda inferior), $\Delta = \begin{pmatrix} -0.1 & -0.1 \\ -0.1 & -0.1 \end{pmatrix}$ (centro inferior) e $\Delta = \begin{pmatrix} -0.4 & -0.8 \\ -0.3 & -0.3 \end{pmatrix}$ (direita

inferior). Já a segunda mostra as curvas de nível da distribuição CFUSN bivariada com $m=3$, para $\Delta = \begin{pmatrix} 0.3 & 0.3 & 0.3 \\ 0.3 & 0.3 & 0.3 \end{pmatrix}$ (esquerda superior), $\Delta = \begin{pmatrix} 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 \end{pmatrix}$ (centro superior), $\Delta = \begin{pmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{pmatrix}$ (direita superior), $\Delta = \begin{pmatrix} -0.3 & -0.3 & -0.3 \\ -0.3 & -0.3 & -0.3 \end{pmatrix}$ (esquerda inferior), $\Delta = \begin{pmatrix} -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 \end{pmatrix}$ (centro inferior) e $\Delta = \begin{pmatrix} -0.1 & -0.3 & -0.2 \\ -0.1 & -0.3 & -0.2 \end{pmatrix}$ (direita inferior). Percebemos que quando o vetor Δ é composto de entradas positivas, a curva se desloca para a direita e coloca mais massa onde X_2 é positivo. Já para entradas negativas, a curva se desloca para a esquerda. Além disto, obtemos assimetria mais acentuada para $m = 3$.

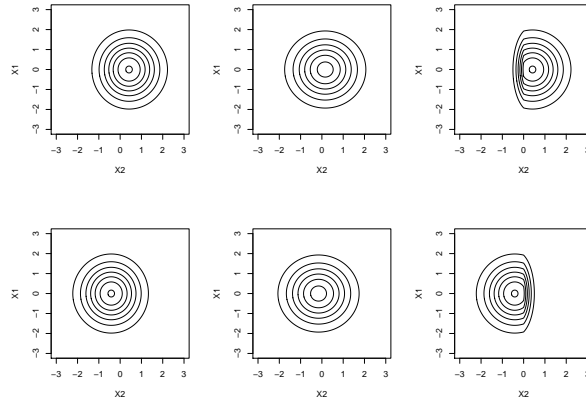


Figura 3.2: Curvas de nível da distribuição CFUSN bivariada com $m=2$ e diferentes valores para Δ .

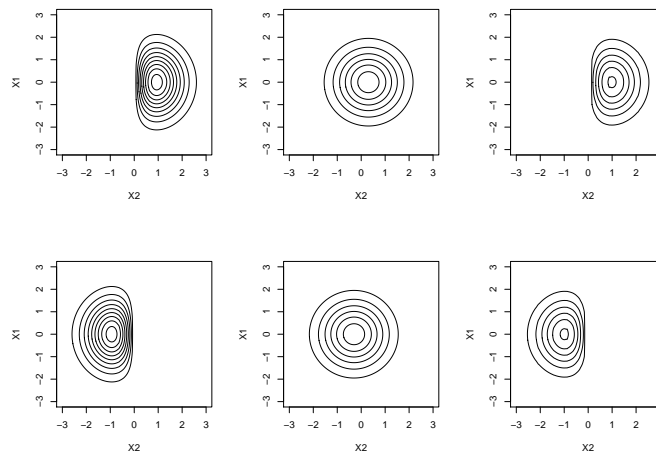


Figura 3.3: Curvas de nível da distribuição CFUSN bivariada com $m=3$ e diferentes valores para Δ .

Apresentaremos agora algumas propriedades desta família de distribuições que foram introduzidas por Arellano-Valle e Genton (2005). Todas estas são extensões das propriedades apresentadas no Capítulo 2 para a fdp definida na Expressão (2.3). A proposição seguinte fornece a função de distribuição acumulada de uma distribuição na família CFUSN.

Proposição 3.1. *Se $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$, então sua função de distribuição acumulada é dada por*

$$F_{\mathbf{Z}^*}(\mathbf{z}) = 2^m \Phi_{n+m}((\mathbf{z}', \mathbf{0}')' | \Omega), \mathbf{z} \in \mathbb{R}^n, \quad (3.8)$$

onde $\Omega = \begin{pmatrix} \mathbf{I}_n & -\Delta \\ -\Delta' & \mathbf{I}_m \end{pmatrix}$.

Demonstração. Considere \mathbf{U} e \mathbf{V} , tais que $\begin{pmatrix} \mathbf{U} \\ \mathbf{V} \end{pmatrix} \sim N_{n+m} \left(\begin{pmatrix} -\mathbf{z} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{I}_n & -\Delta \\ -\Delta' & \mathbf{I}_m \end{pmatrix} \right)$. Então, como consequência, temos que $\mathbf{V} | \mathbf{U} = \mathbf{u} \sim N_m(-\Delta(\mathbf{u} + \mathbf{z}), \mathbf{I}_m - \Delta' \Delta)$, $\mathbf{U} \sim N_n(-\mathbf{z}, \mathbf{I}_n)$ e que a fdp conjunta de (\mathbf{U}, \mathbf{V}) é $f_{(\mathbf{U}, \mathbf{V})}(\mathbf{u}, \mathbf{v}) = \phi_n(\mathbf{u} + \mathbf{z}) \phi_m(\mathbf{v} + \Delta'(\mathbf{u} + \mathbf{z}) | \mathbf{I}_m - \Delta' \Delta)$.

Agora, usando o resultado da Expressão (3.6), temos que $F_{\mathbf{Z}^*}(\mathbf{z}) = \int_{\mathbf{t} < \mathbf{z}} 2^m \phi_n(\mathbf{t}) \Phi_m(\Delta' \mathbf{t} | \mathbf{I}_m - \Delta' \Delta) d\mathbf{t}$. Considerando a transformação $\mathbf{t} = \mathbf{u} + \mathbf{z}$ e a representação estocástica que acabamos de mencionar, $F_{\mathbf{Z}^*}(z)$ torna-se

$$\begin{aligned} F_{\mathbf{Z}^*}(\mathbf{z}) &= \int_{\mathbf{u} < \mathbf{0}} 2^m \phi_n(\mathbf{u} + \mathbf{z}) \Phi_m(\Delta'(\mathbf{u} + \mathbf{z}) | \mathbf{I}_m - \Delta' \Delta) d\mathbf{u} \\ &= 2^m \int_{\mathbf{u} < \mathbf{0}} \int_{\mathbf{v} < \mathbf{0}} \phi_n(\mathbf{u} + \mathbf{z}) \phi_m(\mathbf{v} + \Delta'(\mathbf{u} + \mathbf{z}) | \mathbf{I}_m - \Delta' \Delta) d\mathbf{v} d\mathbf{u} \\ &= 2^m P(\mathbf{U} \leq \mathbf{0}, \mathbf{V} \leq \mathbf{0}). \end{aligned}$$

Portanto, segue que

$$F_{\mathbf{Z}^*}(\mathbf{z}) = 2^m \Phi_{n+m} \left(\begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{I}_n & -\Delta \\ -\Delta' & \mathbf{I}_m \end{pmatrix} \right).$$

□

Se $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$, então, segundo Arellano-Valle e Genton (2005), \mathbf{Z}^* pode ser representado estocasticamente da seguinte forma:

$$\mathbf{Z}^* \stackrel{d}{=} \Delta |\mathbf{X}| + (\mathbf{I}_n - \Delta \Delta')^{1/2} \mathbf{Y}, \quad (3.9)$$

onde $\mathbf{X} \sim N_m(\mathbf{0}, \mathbf{I}_m)$, $\mathbf{Y} \sim N_n(\mathbf{0}, \mathbf{I}_n)$, \mathbf{X} e \mathbf{Y} são vetores aleatórios independentes e $|\mathbf{X}| = (|X_1|, \dots, |X_m|)'$. A representação dada na Expressão (3.9) é chamada de **representação marginal**. Além disso, uma outra representação estocástica para \mathbf{Z}^* pode ser obtida se definirmos $\mathbf{Z} \stackrel{d}{=} \Delta \mathbf{X} + (\mathbf{I}_n - \Delta \Delta')^{1/2} \mathbf{Y}$ e considerarmos $\mathbf{Z}^* \stackrel{d}{=} [\mathbf{Z} | \mathbf{X} > \mathbf{0}]$, onde $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$. Esta representação é denominada **representação condicional**.

Da representação estocástica dada na Expressão (3.9) podemos encontrar a esperança e a variância de \mathbf{Z}^* . Usando o fato de que $E(|\mathbf{X}|) = \sqrt{\frac{2}{\pi}} \mathbf{1}_m$ e $\text{Var}(|\mathbf{X}|) = \mathbf{I}_m - \frac{2}{\pi} \Delta \Delta'$, onde $\mathbf{1}_m$ é um vetor de uns de ordem $m \times 1$, segue que

$$E(\mathbf{Z}^*) = \sqrt{\frac{2}{\pi}} \Delta \mathbf{1}_m \text{ e } \text{Var}(\mathbf{Z}^*) = \mathbf{I}_n - \frac{2}{\pi} \Delta \Delta'. \quad (3.10)$$

A função geradora de momentos de um vetor aleatório \mathbf{Z}^* com distribuição CFUSN é apresentada a seguir. Propriedades adicionais da família de distribuições CFUSN, tais como condições para a independência e distribuições de transformações lineares nesta família podem ser obtidas através desta função. Além disto, a fgm nos fornece um procedimento alternativo para obtermos a média e a variância de \mathbf{Z}^* .

Proposição 3.2. *Se $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$, então sua função geradora de momentos é dada por*

$$M_{\mathbf{Z}^*}(\mathbf{t}) = 2^m e^{(1/2)\mathbf{t}'\mathbf{t}} \Phi_m(\Delta'\mathbf{t}), \mathbf{t} \in \mathbb{R}^n. \quad (3.11)$$

Demonstração. Usando a definição de fgm e uma vez que $e^{\mathbf{t}'\mathbf{z}} \phi_n(\mathbf{z}) = e^{(1/2)\mathbf{t}'\mathbf{t}} \phi_n(\mathbf{z} - \mathbf{t})$, temos que

$$\begin{aligned} M_{\mathbf{Z}^*}(t) = E[e^{\mathbf{t}'\mathbf{Z}^*}] &= 2^m \int_{\mathbb{R}^n} e^{\mathbf{t}'\mathbf{z}} \phi_n(\mathbf{z}) \Phi_m(\Delta'\mathbf{z} | \mathbf{I}_m - \Delta'\Delta) d\mathbf{z} \\ &= 2^m e^{\mathbf{t}'\mathbf{t}/2} \int_{\mathbb{R}^n} \phi_n(\mathbf{z} - \mathbf{t}) \Phi_m(\Delta'\mathbf{z} | \mathbf{I}_m - \Delta'\Delta) d\mathbf{z} \\ &= 2^m e^{\mathbf{t}'\mathbf{t}/2} \int_{\mathbb{R}^n} \phi_n(\mathbf{x}) \Phi_m(\Delta'(\mathbf{x} + \mathbf{t}) | \mathbf{I}_m - \Delta'\Delta) d\mathbf{x} \\ &= 2^m e^{(1/2)\mathbf{t}'\mathbf{t}} E[\Phi_m(\Delta\mathbf{t} + \Delta'\mathbf{X} | \mathbf{I}_m - \Delta'\Delta)], \end{aligned}$$

onde $\mathbf{X} \sim N_n(\mathbf{0}, \mathbf{I}_n)$. Então, segue do Lema 1.1, Capítulo 1, que $E[\Phi_m(\Delta\mathbf{t} + \Delta'\mathbf{Y} | \mathbf{I}_m - \Delta'\Delta)] = \Phi_m(\Delta'\mathbf{t})$, concluindo a prova. \square

Em muitas situações, o uso de transformações lineares de vetores aleatórios se faz necessário. Caracterizaremos agora a distribuição da transformação linear $\mathbf{AZ}^* + \mathbf{b}$ por meio da função geradora de momentos de \mathbf{Z}^* e, quando \mathbf{A} for uma matriz não singular, também através de sua fdp.

Proposição 3.3. *Se $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$, então, para quaisquer matriz \mathbf{A} de dimensão $k \times n$ e \mathbf{b} vetor de dimensão $k \times 1$, teremos*

$$M_{\mathbf{AZ}^* + \mathbf{b}}(\mathbf{t}) = 2^m e^{\mathbf{t}'\mathbf{b} + (1/2)\mathbf{t}'\mathbf{A}\mathbf{A}'\mathbf{t}} \Phi_m(\Delta'\mathbf{A}'\mathbf{t}), \mathbf{t} \in \mathbb{R}^k. \quad (3.12)$$

Além disso, se $k = n$ e \mathbf{A} for uma matriz não singular, então a fdp de $\mathbf{AZ}^ + \mathbf{b}$ existe e é dada por*

$$f_{\mathbf{AZ}^* + \mathbf{b}}(\mathbf{z}) = 2^m |\det(\mathbf{A})|^{-1} \phi_n(\mathbf{A}^{-1}(\mathbf{z} - \mathbf{b})) \Phi_m(\Delta'\mathbf{A}^{-1}(\mathbf{z} - \mathbf{b}) | \mathbf{I}_m - \Delta'\Delta), \mathbf{z} \in \mathbb{R}^n. \quad (3.13)$$

Demonstração. Usando o resultado da Proposição 3.2 e propriedades da fgm, temos que para todo $\mathbf{t} \in \mathbb{R}^k$,

$$\begin{aligned} M_{\mathbf{AZ}^*+\mathbf{b}}(\mathbf{t}) &= E[e^{\mathbf{t}'(\mathbf{AZ}^*+\mathbf{b})}] = E[e^{\mathbf{t}'\mathbf{AZ}^*+\mathbf{t}'\mathbf{b}}] = e^{\mathbf{t}'\mathbf{b}} E[e^{\mathbf{t}'\mathbf{AZ}^*}] = e^{\mathbf{t}'\mathbf{b}} M_{\mathbf{Z}^*}(\mathbf{A}'\mathbf{t}) \\ &= 2^m e^{\mathbf{t}'\mathbf{b}+(1/2)\mathbf{t}'\mathbf{AA}'\mathbf{t}} \Phi_m(\Delta'\mathbf{A}'\mathbf{t}). \end{aligned}$$

Agora, usando o fato de que $f_{\mathbf{AZ}^*+\mathbf{b}}(\mathbf{z}) = |\det(\mathbf{A})|^{-1} f_{\mathbf{Z}^*}(\mathbf{A}^{-1}(\mathbf{z} - \mathbf{b}))$, obtemos a fdp dada na Expressão (3.13). □

Um resultado de grande interesse no estudo da estrutura ds dados que estamos modelando são as distribuições marginais. Assim como todas as famílias apresentadas nos Capítulos 1 e 2, a família de distribuições CFUSN também é fechada sob marginalização. O resultado é consequência direta da Proposição 3.3 e é dado na seguinte proposição.

Proposição 3.4. *Seja $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$. Considere as partições $\mathbf{Z}^* = \begin{pmatrix} \mathbf{Z}_1^* \\ \mathbf{Z}_2^* \end{pmatrix}$ e $\Delta = \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix}$, onde \mathbf{Z}_i^* e Δ_i tem dimensões $n_i \times 1$ e $n_i \times m$, $i = 1, 2$, respectivamente, e $n_1 + n_2 = n$. Então o componente \mathbf{Z}_i^* , $i = 1, 2$, da partição é tal que $\mathbf{Z}_i^* \sim CFUSN_{n_i,m}(\Delta_i)$, $i = 1, 2$, com fdp dada por*

$$f_{\mathbf{Z}_i^*}(\mathbf{z}_i) = 2^{m_i} \phi_{n_i}(\mathbf{z}_i) \Phi_m(\Delta'_i \mathbf{z}_i | \mathbf{I}_m - \Delta'_i \Delta_i), \mathbf{z}_i \in \mathbb{R}^{n_i}.$$

Demonstração. O resultado segue diretamente da Proposição 3.3, se tomarmos $\mathbf{A} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}$ para $i = 1$ e $\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}$ para $i = 2$. □

Suponha agora que $m > 1$ e particione as matrizes $\Delta_i, i = 1, 2$ como $\Delta_i = (\Delta_{i1}, \Delta_{i2})$, onde Δ_{ij} tem dimensão $n_i \times m_j, j = 1, 2$ e $m_1 + m_2 = m$. É possível obtermos independência entre os componentes \mathbf{Z}_1^* e \mathbf{Z}_2^* do vetor aleatório \mathbf{Z}^* a partir da matriz Δ , como pode ser visto na Proposição 3.5.

Proposição 3.5. *Seja $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$ e considere a partição assumida na Proposição 3.4. Para cada uma das condições seguintes sobre a matriz Δ , os vetores \mathbf{Z}_1^* e \mathbf{Z}_2^* são independentes:*

- i) $\Delta_{12} = \Delta_{21} = \mathbf{0}$ e, neste caso, $\mathbf{Z}_i^* \sim CFUSN_{n_i,m_i}(\Delta_{ii}), i = 1, 2$;
- ii) $\Delta_{ii} = \mathbf{0}, i = 1, 2$ e, neste caso, $\mathbf{Z}_1^* \sim CFUSN_{n_1,m_2}(\Delta_{12})$ e $\mathbf{Z}_2^* \sim CFUSN_{n_2,m_1}(\Delta_{21})$.

Demonstração. Como $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$, segue da Expressão (3.6) que a fdp conjunta de $(\mathbf{Z}_1^*, \mathbf{Z}_2^*)$ é dada por

$$f_{\mathbf{Z}_1^*, \mathbf{Z}_2^*}(\mathbf{z}_1, \mathbf{z}_2) = 2^{m_1+m_2} \phi_{n_1}(\mathbf{z}_1) \phi_{n_2}(\mathbf{z}_2) \Phi_{m_1+m_2}(\Delta'_1 \mathbf{z}_1 + \Delta'_2 \mathbf{z}_2 | \mathbf{I}_{m_1+m_2} - \Delta'_1 \Delta_1 - \Delta'_2 \Delta_2). \quad (3.14)$$

Além disso, se considerarmos $\Delta_i = (\Delta_{i1}, \Delta_{i2}), i = 1, 2$, temos que

$$\Delta'_1 \mathbf{z}_1 + \Delta'_2 \mathbf{z}_2 = \begin{pmatrix} \Delta'_{11} \mathbf{z}_1 + \Delta'_{21} \mathbf{z}_2 \\ \Delta'_{12} \mathbf{z}_1 + \Delta'_{22} \mathbf{z}_2 \end{pmatrix},$$

e, também que

$$\mathbf{I}_{m_1+m_2} - \Delta'_1 \Delta_1 - \Delta'_2 \Delta_2 = \begin{pmatrix} \mathbf{I}_{m_1} - (\Delta'_{11} \Delta_{11} + \Delta'_{21} \Delta_{21}) & -(\Delta'_{11} \Delta_{12} + \Delta'_{21} \Delta_{22}) \\ -(\Delta'_{12} \Delta_{11} + \Delta'_{22} \Delta_{21}) & \mathbf{I}_{m_2} - (\Delta_{12} \Delta_{12} + \Delta'_{22} \Delta_{22}) \end{pmatrix}.$$

Para que a Expressão (3.14) fatore no produto das distribuições marginais de Z_1^* e Z_2^* , é necessário que $\Delta_{12} = \mathbf{0}$ e $\Delta_{21} = \mathbf{0}$ ou que $\Delta_{11} = \mathbf{0}$ e $\Delta_{22} = \mathbf{0}$. No primeiro caso, temos $\Phi_m(\Delta' \mathbf{z} | \mathbf{I}_m - \Delta' \Delta) = \Phi_{m_1}(\Delta'_{11} \mathbf{z}_1 | \mathbf{I}_{m_1} - \Delta'_{11} \Delta_{11}) \Phi_{m_2}(\Delta'_{22} \mathbf{z}_2 | \mathbf{I}_{m_2} - \Delta'_{22} \Delta_{22})$, e no segundo caso, temos $\Phi_m(\Delta' \mathbf{z} | \mathbf{I}_m - \Delta' \Delta) = \Phi_{m_2}(\Delta'_{12} \mathbf{z}_1 | \mathbf{I}_{m_2} - \Delta'_{12} \Delta_{12}) \Phi_{m_1}(\Delta'_{21} \mathbf{z}_2 | \mathbf{I}_{m_1} - \Delta'_{21} \Delta_{21})$, concluindo a prova. \square

Apesar de ser fechada sob marginalização, a família de distribuições CFUSN não é fechada sob condicionamento. Porém, as distribuições condicionais derivadas a partir da família CFUSN pertencem a uma família mais geral de distribuições denominada família de distribuições *skew-normal*, a qual é denotada por FUSN, e cuja fdp é dada na Expressão (3.2).

Proposição 3.6. *Se $\mathbf{Z}^* = (\mathbf{Z}_1^*, \mathbf{Z}_2^*) \sim CFUSN_{n,m}(\Delta)$, então a fdp de \mathbf{Z}_1^* dado $\mathbf{Z}_2^* = \mathbf{z}_2$ é*

$$f_{\mathbf{Z}_1^* | \mathbf{Z}_2^* = \mathbf{z}_2}(\mathbf{z}_1) = \frac{\phi_{n_1}(\mathbf{z}_1) \Phi_m(\Delta'_1 \mathbf{z}_1 | -\Delta'_2 \mathbf{z}_2, \mathbf{I}_m - \Delta' \Delta)}{\Phi_m(\Delta'_2 \mathbf{z}_2 | \mathbf{I}_m - \Delta'_2 \Delta_2)}, \quad (3.15)$$

onde as matrizes Δ_1 e Δ_2 são definidas na Proposição 3.4.

Demonstração. Temos que $\Phi_m(\Delta' \mathbf{z} | \mathbf{I}_m - \Delta' \Delta) = \Phi_m(\Delta'_1 \mathbf{z}_1 + \Delta'_2 \mathbf{z}_2 | \mathbf{I}_m - \Delta' \Delta) = \Phi_m(\Delta'_1 \mathbf{z}_1 | -\Delta'_2 \mathbf{z}_2, \mathbf{I}_m - \Delta' \Delta)$. Então, considerando o resultado da Proposição 3.4,

$$\begin{aligned} f_{\mathbf{Z}_1^* | \mathbf{Z}_2^* = \mathbf{z}_2}(\mathbf{z}_1) &= \frac{2^m \phi_{n_1}(\mathbf{z}_1) \phi_{n_2}(\mathbf{z}_2) \Phi_m(\Delta'_1 \mathbf{z}_1 | -\Delta'_2 \mathbf{z}_2, \mathbf{I}_m - \Delta' \Delta)}{2^m \phi_{n_2}(\mathbf{z}_2) \Phi_m(\Delta'_2 \mathbf{z}_2 | \mathbf{I}_m - \Delta'_2 \Delta_2)} \\ &= \frac{\phi_{n_1}(\mathbf{z}_1) \Phi_m(\Delta'_1 \mathbf{z}_1 | -\Delta'_2 \mathbf{z}_2, \mathbf{I}_m - \Delta' \Delta)}{\Phi_m(\Delta'_2 \mathbf{z}_2 | \mathbf{I}_m - \Delta'_2 \Delta_2)}. \end{aligned}$$

\square

Dado que $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$, é possível também adicionarmos locação e escala à distribuição através da transformação linear

$$\mathbf{W}^* = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2} \mathbf{Z}^*, \quad (3.16)$$

onde $\boldsymbol{\mu}$ denota o vetor de locação de ordem $n \times 1$ e $\boldsymbol{\Sigma}$ denota a matriz de escala, positiva definida de dimensão $n \times n$. Neste caso, dizemos que \mathbf{W}^* tem distribuição *skew-normal* fundamental canônica com locação e escala, denotada por $\mathbf{W}^* \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \Delta)$. Se $\boldsymbol{\Sigma} > \mathbf{0}$, a fdp de \mathbf{W}^* existe e é obtida através da Expressão (3.13). Sua função geradora de momentos pode ser obtida da Expressão (3.12). Estas são dadas, respectivamente, por

$$f_{\mathbf{W}^*}(\mathbf{w}) = 2^m |\boldsymbol{\Sigma}|^{-1/2} \phi_n(\boldsymbol{\Sigma}^{-1/2}(\mathbf{w} - \boldsymbol{\mu})) \Phi_m(\Delta' \boldsymbol{\Sigma}^{-1/2}(\mathbf{w} - \boldsymbol{\mu}) | \mathbf{I}_m - \Delta' \Delta), \quad \mathbf{w} \in \mathbb{R}^n, \quad (3.17)$$

e

$$M_{\mathbf{W}^*}(\mathbf{t}) = 2^m e^{\mathbf{t}'\boldsymbol{\mu} + (1/2)\mathbf{t}'\boldsymbol{\Sigma}'\mathbf{t}} \Phi_m(\boldsymbol{\Delta}'\boldsymbol{\Sigma}^{1/2}\mathbf{t}), \mathbf{t} \in \mathbb{R}^n. \quad (3.18)$$

Segue da Expressão (3.10) que a esperança e a variância de um vetor aleatório $\mathbf{W}^* \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{\Delta})$ são, respectivamente, dadas por

$$E(\mathbf{W}^*) = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}} \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Delta} \mathbf{1}_m \text{ e } \text{Var}(\mathbf{W}^*) = \boldsymbol{\Sigma} - \frac{2}{\pi} \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Delta} \boldsymbol{\Delta}' \boldsymbol{\Sigma}^{1/2}. \quad (3.19)$$

A escolha da transformação linear na Expressão (3.16) envolve a raiz quadrada da matriz $\boldsymbol{\Sigma}$. Isto pode apresentar alguns problemas, já que esta raiz não é única. A unicidade ocorre, por exemplo, quando $\boldsymbol{\Sigma}$ é diagonal. Além disto, a derivação de distribuições marginais e condicionais no caso geral é bastante complicada. Apresentaremos aqui as propriedades de $\mathbf{W}^* \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{\Delta})$ para o caso particular em que $\boldsymbol{\Sigma}$ é uma matriz positiva definida diagonal.

Para este fim, considere as partições

$$\mathbf{W}^* = \begin{pmatrix} \mathbf{W}_1^* \\ \mathbf{W}_2^* \end{pmatrix}, \boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_1 \\ \boldsymbol{\Delta}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix},$$

onde \mathbf{W}_i^* , $\boldsymbol{\mu}_i$ e $\boldsymbol{\Delta}_i$ têm dimensões $n_i \times 1$, $n_i \times 1$ e $n_i \times m$, $i = 1, 2$, respectivamente, e $n_1 + n_2 = n$. Considere também a partição

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

onde $\boldsymbol{\Sigma}_{ij}$ tem dimensão $n_i \times n_j$.

Então, segundo Arellano-Valle e Azzalini (2006), temos que a distribuição marginal de \mathbf{W}_i^* é dada por

$$f_{\mathbf{W}_i^*}(\mathbf{w}) = 2^m \phi_{n_i}(\mathbf{w} - \boldsymbol{\mu}_i | \boldsymbol{\Sigma}_{ii}) \Phi_m(\boldsymbol{\Delta}_i' \boldsymbol{\Sigma}_{ii}^{-1/2} (\mathbf{w} - \boldsymbol{\mu}_i) | I_m - \boldsymbol{\Delta}_i' \boldsymbol{\Delta}_i), \mathbf{w} \in \mathbb{R}^{n_i}, \quad (3.20)$$

e a distribuição condicional de $\mathbf{W}_1^* | \mathbf{W}_2^* = \mathbf{w}_2$ é dada por

$$f_{\mathbf{W}_1^* | \mathbf{W}_2^* = \mathbf{w}_2}(\mathbf{w}_1) = \frac{\phi_{n_1}(\mathbf{w}_1 - \boldsymbol{\mu}_1 | \boldsymbol{\Sigma}_{11}) \Phi_m(\boldsymbol{\Delta}_2' \boldsymbol{\Sigma}_{22}^{-1/2} (\mathbf{w}_2 - \boldsymbol{\mu}_2) + \boldsymbol{\Delta}_1' \boldsymbol{\Sigma}_{11}^{-1/2} (\mathbf{w}_1 - \boldsymbol{\mu}_1) | I_m - \boldsymbol{\Delta}' \boldsymbol{\Delta})}{\Phi_m(\boldsymbol{\Delta}_2' \boldsymbol{\Sigma}_{22}^{-1/2} (\mathbf{w}_2 - \boldsymbol{\mu}_2) | I_m - \boldsymbol{\Delta}_2' \boldsymbol{\Delta}_2)}, \quad (3.21)$$

$\mathbf{w}_1 \in \mathbb{R}^{n_1}$ e $\mathbf{w}_2 \in \mathbb{R}^{n_2}$.

Na próxima seção, alguns resultados obtidos aqui serão utilizados na construção de uma nova família de distribuições: a família log-skew-normal fundamental canônica. Também estudaremos suas principais propriedades.

3.2 Distribuição Log-*Skew*-Normal Fundamental Canônica Multivariada

Nesta seção introduziremos uma nova família de distribuições log-*skew*-normal fundamental, a qual chamaremos de família de distribuições log-*skew*-normal fundamental canônica. Esta nova família é obtida através da transformação da família de distribuições *skew*-normal fundamental canônica apresentada na Expressão (3.6) e generaliza algumas famílias de distribuições log-*skew*-normal já existentes na literatura. A família de distribuições log-*skew*-normal fundamental canônica também é capaz de descrever o comportamento de variáveis aleatórias com diferentes graus de assimetria e que colocam massa apenas em valores positivos. Derivaremos algumas de suas propriedades probabilísticas partindo dos resultados apresentados nas proposições demonstradas na Seção 3.1, incluindo distribuições marginais, condicionais, momentos e representações estocásticas. Veremos que a distribuição log-*skew*-normal multivariada definida por Marchenko e Genton (2010) na Expressão (2.6) no Capítulo 2 é um caso particular dessa distribuição, quando tomamos $m = 1$.

Como definido anteriormente, sejam $\ln(\mathbf{Y}) = (\ln(Y_1), \dots, \ln(Y_n))'$ e $\exp(\mathbf{X}) = (\exp(Y_1), \dots, \exp(Y_n))'$, respectivamente, o logaritmo e a exponencial do vetor aleatório $\mathbf{Y} = (Y_1, \dots, Y_n)'$.

Definição 3.3. Dizemos que \mathbf{Y} tem distribuição log-*skew*-normal fundamental canônica padrão com vetor de assimetria Δ , que denotaremos por $LCFUSN_{n,m}(\Delta)$, se, e somente se, $\mathbf{Z}^* = \ln \mathbf{Y}$ é um vetor aleatório $n \times 1$ com distribuição $CFUSN_{n,m}(\Delta)$.

Considerando o método jacobiano e a transformação $\mathbf{Y} = \exp(\mathbf{Z}^*)$, temos que a fdp de \mathbf{Y} é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = 2^m \left(\prod_{i=1}^n y_i \right)^{-1} \phi_n(\ln \mathbf{y}) \Phi_m(\Delta' \ln \mathbf{y} | \mathbf{I}_m - \Delta' \Delta), \mathbf{y} \in \mathbb{R}^{n^+}, \quad (3.22)$$

onde Δ é tal que $\|\Delta \mathbf{a}\| < 1$, para todo vetor unitário $\mathbf{a} \in \mathbb{R}^m$.

A família log-*skew*-normal multivariada definida por Marchenko e Genton (2010) (ver Capítulo 2, Expressão (2.6)) é um caso particular da família definida na Expressão (3.22), quando tomamos $m = 1$ e fazemos $\alpha = (\mathbf{I}_m - \Delta' \Delta)^{-\frac{1}{2}} \Delta'$.

Ilustramos na Figura 3.4 o comportamento da fdp de uma variável com distribuição na família LCFUSN univariada para diferentes valores de m . O caso $m=1$ é incluído, permitindo a comparação com a família estudada no Capítulo 2. Percebemos da Figura 3.4 que a assimetria introduzida pela multiplicação de uma função de distribuição acumulada m -variada à distribuição log-*skew*-normal convencional possui formas mais flexíveis, com caudas mais pesadas à direita.

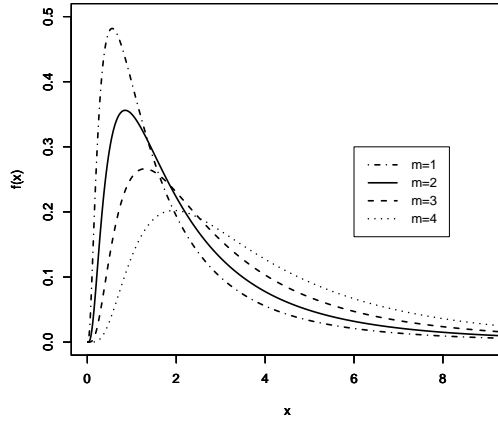


Figura 3.4: Função densidade de probabilidade da distribuição LCFUSN padrão com $n = 1$, para $m = 1$ e $\Delta = 0.4$, $m = 2$ e $\Delta = (0.4, 0.4)$, $m = 3$ e $\Delta = (0.4, 0.4, 0.4)$ e $m = 4$ e $\Delta = (0.4, 0.4, 0.4, 0.4)$.

As Figuras 3.5 e 3.6 mostram os gráficos com as curvas de nível de um vetor aleatório com distribuição LCFUSN quando consideramos $n = 2$ e, respectivamente, $m = 2$ e $m = 3$. Seis valores distintos do parâmetro de assimetria Δ são considerados. A Figura 3.5 mostra as curvas de nível da distribuição LCFUSN bivariada com $m=2$, para $\Delta = \begin{pmatrix} 0.3 & 0.3 \\ 0.3 & 0.3 \end{pmatrix}$ (esquerda superior), $\Delta = \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}$ (centro superior), $\Delta = \begin{pmatrix} 0.4 & 0.8 \\ 0.3 & 0.3 \end{pmatrix}$ (direita superior), $\Delta = \begin{pmatrix} -0.3 & -0.3 \\ -0.3 & -0.3 \end{pmatrix}$ (esquerda inferior), $\Delta = \begin{pmatrix} -0.1 & -0.1 \\ -0.1 & -0.1 \end{pmatrix}$ (centro inferior) e $\Delta = \begin{pmatrix} -0.4 & -0.8 \\ -0.3 & -0.3 \end{pmatrix}$ (direita inferior). Já a Figura 3.6 mostra as curvas de nível da distribuição LCFUSN bivariada com $m=3$, para $\Delta = \begin{pmatrix} 0.3 & 0.3 & 0.3 \\ 0.3 & 0.3 & 0.3 \end{pmatrix}$ (esquerda superior), $\Delta = \begin{pmatrix} 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 \end{pmatrix}$ (centro superior), $\Delta = \begin{pmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{pmatrix}$ (direita superior), $\Delta = \begin{pmatrix} -0.3 & -0.3 & -0.3 \\ -0.3 & -0.3 & -0.3 \end{pmatrix}$ (esquerda inferior), $\Delta = \begin{pmatrix} -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 \end{pmatrix}$ (centro inferior) e $\Delta = \begin{pmatrix} -0.1 & -0.3 & -0.2 \\ -0.1 & -0.3 & -0.2 \end{pmatrix}$ (direita inferior). Assim como vimos na Figura 2.3 do Capítulo 2, quando as entradas do parâmetro Δ são positivas, a curva se afasta da origem a medida que as entradas de Δ assumem valores maiores. Quando as entradas da matriz Δ são negativas, a curva tende a colocar mais massa próximo da origem, havendo maior concentração próximo de zero quanto menor forem os componentes de Δ .

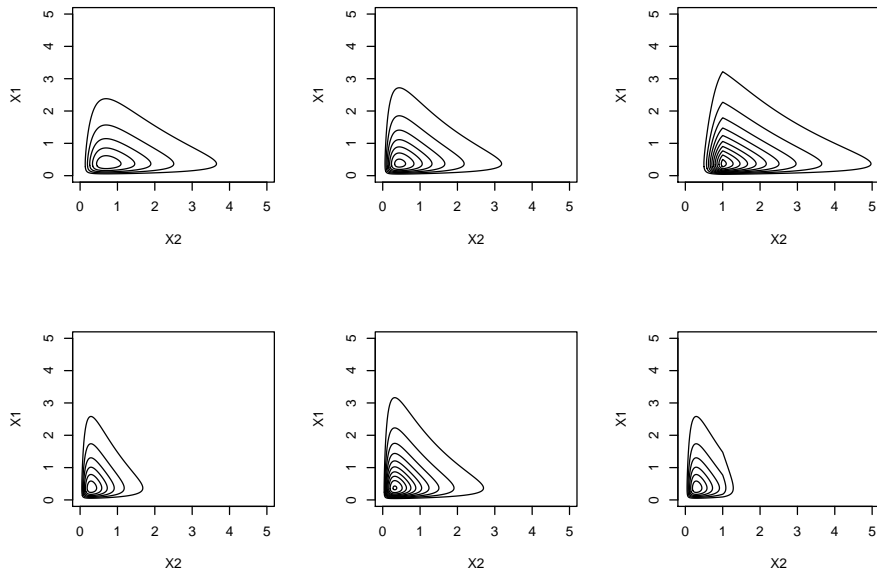


Figura 3.5: Curvas de nível da distribuição LCFUSN bivariada com $m=2$, para diferentes valores de Δ .

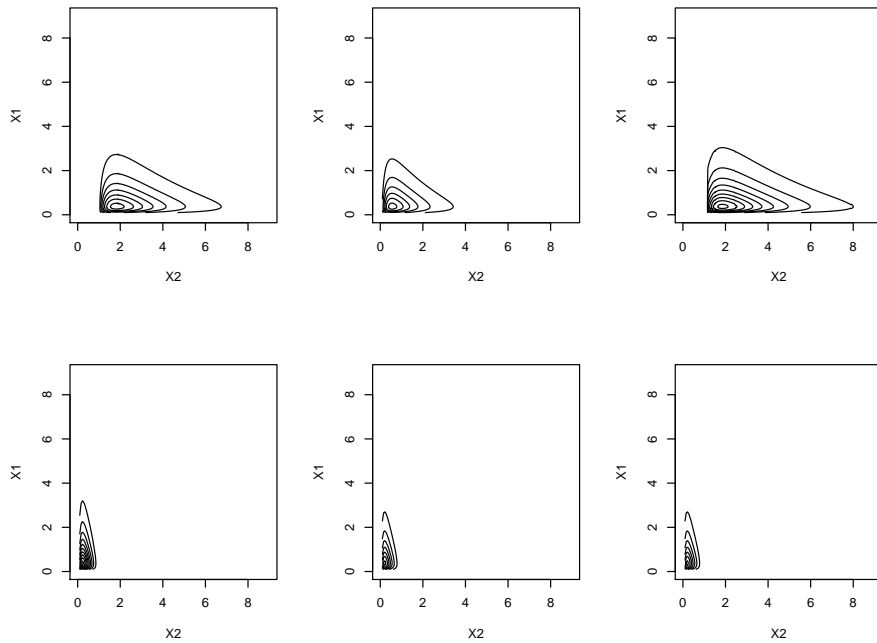


Figura 3.6: Curvas de nível da distribuição LCFUSN bivariada com $m=3$, para diferentes valores de Δ .

Uma representação estocástica para o vetor aleatório \mathbf{Y} pode ser obtida através da representação estocástica para o vetor \mathbf{Z}^* , dada na Expressão (3.9), e da transformação exponencial.

Assim, temos que \mathbf{Y} pode ser escrita como

$$\mathbf{Y} \stackrel{d}{=} e^{(\Delta|\mathbf{x}|+(\mathbf{I}_n-\Delta\Delta')^{1/2}\mathbf{w})}, \quad (3.23)$$

onde $\mathbf{X} \sim N_m(\mathbf{0}, \mathbf{I}_m)$, $\mathbf{W} \sim N_n(\mathbf{0}, \mathbf{I}_n)$, \mathbf{X} e \mathbf{W} são vetores aleatórios independentes e $|\mathbf{X}| = (|X_1|, \dots, |X_m|)'$.

A partir dos resultados vistos na Seção 3.1, é possível derivar a forma da função de distribuição acumulada para a família LCFUSN, como mostrado na Proposição 3.7.

Proposição 3.7. *Se $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$, então sua função de distribuição acumulada é dada por*

$$F_{\mathbf{Y}}(\mathbf{y}) = 2^m \Phi_{n+m}((\ln \mathbf{y}', \mathbf{0}')' | \Omega), \mathbf{y} \in \mathbb{R}^{n^+}, \quad (3.24)$$

$$\text{onde } \Omega = \begin{pmatrix} \mathbf{I}_n & -\Delta \\ -\Delta' & \mathbf{I}_m \end{pmatrix}.$$

Demonstração. Temos que $P(\mathbf{Y} \leq \mathbf{y}) = P(e^{\mathbf{Z}^*} \leq \mathbf{y}) = P(\mathbf{Z}^* < \ln \mathbf{y}) = F_{\mathbf{Z}^*}(\ln \mathbf{y})$. O resultado segue imediatamente da Proposição 3.1. □

Os momentos mistos do vetor \mathbf{Y} podem ser expressos usando a função geradora de momentos da família de distribuições *skew-normal* fundamental canônica, dada na Expressão (3.12). Este resultado é provado na seguinte proposição:

Proposição 3.8. *Se $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$ e $\mathbf{t} = (t_1, \dots, t_d)'$, $t_i \in \mathbb{N}$, então*

$$E\left(\prod_{i=1}^n Y_i^{t_i}\right) = 2^m e^{(1/2)\mathbf{t}'\mathbf{t}} \Phi_m(\Delta'\mathbf{t}). \quad (3.25)$$

Demonstração. Temos que $E(\prod_{i=1}^n Y_i^{t_i}) = E(\prod_{i=1}^n e^{t_i \ln Y_i}) = E(e^{\sum_{i=1}^n t_i \ln Y_i}) = E(e^{\mathbf{t} \ln \mathbf{Y}}) = M_{\ln \mathbf{Y}}(\mathbf{t})$. Como $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$, temos que $\ln \mathbf{Y} \sim CFUSN_{n,m}(\Delta)$ e o resultado da Proposição 3.2 conclui a prova. □

Utilizando o resultado na Expressão (3.25), podemos calcular os momentos de um vetor aleatório com distribuição na família LCFUSN. Para o caso particular em que a variável aleatória $\mathbf{Y} \sim LCFUSN_{1,m}(\Delta)$, temos que

$$\begin{aligned} E(Y) &= 2^m e^{1/2} \Phi_m(\Delta) \\ E(Y^2) &= 2^m e^2 \Phi_m(2\Delta) \\ E(Y^3) &= 2^m e^{9/2} \Phi_m(3\Delta) \\ E(Y^4) &= 2^m e^8 \Phi_m(4\Delta). \end{aligned}$$

Como ocorre com a família de distribuições CFUSN, é razoável pensarmos que as distribuições marginais dos componentes de $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$ também terão distribuição LFUSN, uma vez que cada marginal de $\mathbf{Z}^* \sim CFUSN_{n,m}(\Delta)$ também tem distribuição na

família CFUSN. A Proposição 3.9 formaliza este fato e sua prova será omitida, pois é consequência direta da Definição 3.3 e da Proposição 3.4 da Seção 3.1, onde se prova que \mathbf{Z}^* é fechada sob marginalização.

Proposição 3.9. *Seja $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$. Considere as partições $\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}$ e $\Delta = \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix}$, onde \mathbf{Y}_i e Δ_i têm dimensões $n_i \times 1$ e $n_i \times m$, respectivamente, e $n_1 + n_2 = n$. Então, para $i = 1, 2$, $\mathbf{Y}_i \sim LCFUSN_{n_i,m}(\Delta_i)$, e a fdp de \mathbf{Y}_i é dada por*

$$f_{\mathbf{Y}_i}(\mathbf{y}_i) = 2^m \left(\prod_{j=1}^{n_i} y_j \right)^{-1} \phi_{n_i}(\ln \mathbf{y}_i) \Phi_m(\Delta'_i \ln \mathbf{y}_i | \mathbf{I}_m - \Delta'_i \Delta_i), \mathbf{y}_i \in \mathbb{R}^{n_i^+}. \quad (3.26)$$

É possível obtermos independência entre os componentes \mathbf{Y}_1 e \mathbf{Y}_2 de uma partição qualquer do vetor aleatório \mathbf{Y} a partir da matriz Δ , como pode ser visto na Proposição 3.10. A prova deste resultado segue diretamente da Proposição 3.5, onde se mostram as condições sobre Δ para que \mathbf{Z}_1^* e \mathbf{Z}_2^* sejam independentes, e, por isso, será omitida.

Proposição 3.10. *Seja \mathbf{Y} um vetor aleatório $n \times 1$ tal que $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$ e considere a partição assumida na Proposição 3.9. Para cada uma das seguintes condições sobre a matriz Δ , os vetores \mathbf{Y}_1 e \mathbf{Y}_2 são independentes:*

i) $\Delta_{12} = \Delta_{21} = \mathbf{0}$ e, neste caso, $\mathbf{Y}_i \sim LCFUSN_{n_i,m_i}(\Delta_{ii})$, $i = 1, 2$;

ii) $\Delta_{ii} = \mathbf{0}$, $i = 1, 2$ e, neste caso, $\mathbf{Y}_1 \sim LCFUSN_{n_1,m_2}(\Delta_{12})$ e $\mathbf{Y}_2 \sim LCFUSN_{n_2,m_1}(\Delta_{21})$.

Assim como ocorre na família de distribuições CFUSN, a família de distribuições LCFUSN também não é fechada sob condicionamento. Isto pode ser visto na seguinte proposição:

Proposição 3.11. *Seja $\mathbf{Y} \sim LCFUSN_{n,m}(\Delta)$. Considere as partições $\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}$ e $\Delta = \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix}$, onde \mathbf{Y}_i e Δ_i tem dimensões $n_i \times 1$ e $n_i \times m$, respectivamente, e $n_1 + n_2 = n$. Então a densidade condicional de \mathbf{Y}_1 dado $\mathbf{Y}_2 = \mathbf{y}_2$, $\mathbf{y}_2 \in \mathbb{R}^{n_2^+}$, é dada por*

$$f_{\mathbf{Y}_1|\mathbf{Y}_2=\mathbf{y}_2}(\mathbf{y}_1) = \left(\prod_{i=j}^{n_1} y_j \right)^{-1} \frac{\phi_{n_1}(\ln \mathbf{y}_1) \Phi_m(\Delta'_1 \ln \mathbf{y}_1 | -\Delta'_2 \ln \mathbf{y}_2, \mathbf{I}_m - \Delta'_1 \Delta)}{\Phi_m(\Delta'_2 \ln \mathbf{y}_2 | \mathbf{I}_m - \Delta'_2 \Delta_2)}, \mathbf{y}_1 \in \mathbb{R}^{n_1^+}. \quad (3.27)$$

Demonstração. Considerando a partição do vetor \mathbf{Y} e da matriz Δ acima mencionados, dado $\mathbf{y}_2 \in \mathbb{R}^{n_2^+}$, temos que $\Phi_m(\Delta'_1 \ln \mathbf{y} | \mathbf{I}_m - \Delta'_1 \Delta) = \Phi_m(\Delta'_1 \ln \mathbf{y}_1 + \Delta'_2 \ln \mathbf{y}_2 | \mathbf{I}_m - \Delta'_1 \Delta) = \Phi_m(\Delta'_1 \ln \mathbf{y}_1 | -\Delta'_2 \ln \mathbf{y}_2, \mathbf{I}_m - \Delta'_1 \Delta)$. Então, usando os resultados da Proposição 3.9 segue que

$$\begin{aligned} f_{\mathbf{Y}_1|\mathbf{Y}_2=\mathbf{y}_2}(\mathbf{y}_1) &= \frac{2^m \left(\prod_{i=j}^n y_j \right)^{-1} \phi_{n_1}(\ln \mathbf{y}_1) \phi_{n_2}(\ln \mathbf{y}_2) \Phi_m(\Delta'_1 \ln \mathbf{y}_1 | -\Delta'_2 \ln \mathbf{y}_2, \mathbf{I}_m - \Delta'_1 \Delta)}{2^m \left(\prod_{j=n_1+1}^n y_j \right)^{-1} \phi_{n_2}(\ln \mathbf{y}_2) \Phi_m(\Delta'_2 \ln \mathbf{y}_2 | \mathbf{I}_m - \Delta'_2 \Delta_2)} \\ &= \frac{\left(\prod_{j=1}^{n_1} y_j \right)^{-1} \phi_{n_1}(\ln \mathbf{y}_1) \Phi_m(\Delta'_1 \ln \mathbf{y}_1 | -\Delta'_2 \ln \mathbf{y}_2, \mathbf{I}_m - \Delta'_1 \Delta)}{\Phi_m(\Delta'_2 \ln \mathbf{y}_2 | \mathbf{I}_m - \Delta'_2 \Delta_2)}, \end{aligned} \quad (3.28)$$

concluindo a prova. □

Famílias de distribuições sem locação e escala podem ser bastante limitadas, não descrevendo o comportamento de dados adequadamente em muitas situações. Para tornarmos a família que construímos mais flexível, vamos incorporar estes parâmetros a ela. Faremos isto considerando a transformação de uma variável com distribuição CFSUN com locação e escala. Seja \mathbf{W}^* um vetor aleatório $n \times 1$ tal que $\mathbf{W}^* \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{\Delta})$. Definimos a transformação $\mathbf{U} = e^{\mathbf{W}^*}$. Então $\mathbf{U} \sim LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{\Delta})$, onde $\boldsymbol{\mu}$ é um vetor de locação $n \times 1$ e $\boldsymbol{\Sigma}$ é uma matriz de escala positiva definida de dimensão $n \times n$. A fdp de \mathbf{U} é obtida através da Expressão (3.17) e, para todo $\mathbf{u} \in \mathbb{R}^{n^+}$, é dada por

$$f_{\mathbf{U}}(\mathbf{u}) = 2^m |\boldsymbol{\Sigma}|^{-1/2} \left(\prod_{j=1}^n u_j \right)^{-1} \phi_n(\boldsymbol{\Sigma}^{-1/2}(\ln \mathbf{u} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-1/2}(\ln \mathbf{u} - \boldsymbol{\mu}) | \mathbf{I}_m - \boldsymbol{\Delta}' \boldsymbol{\Delta}). \quad (3.29)$$

Pelas mesmas razões mencionadas na Seção 3.1, a determinação de distribuições marginais e condicionais associados a um vetor \mathbf{U} com distribuição log-*skew*-normal fundamental canônica multivariada com locação e escala não é simples. Da mesma forma como foi feito na Seção 3.1, podemos estudar o caso particular da fdp dada na Expressão (3.29) quando a matriz $\boldsymbol{\Sigma}$ é diagonal. Considerando as mesmas partições usadas nas Expressões (3.20) e (3.21) para \mathbf{U} e para os parâmetros e considerando a transformação $e^{\mathbf{U}_i}$, temos que a distribuição marginal do vetor aleatório \mathbf{U}_i é dada por

$$f_{\mathbf{U}_i}(\mathbf{u}) = 2^m \left(\prod_{j=1}^{n_i} u_j \right)^{-1} \phi_{n_i}(\ln \mathbf{u} - \boldsymbol{\mu}_i | \boldsymbol{\Sigma}_{ii}) \Phi_m(\boldsymbol{\Delta}'_i \boldsymbol{\Sigma}_{ii}^{-1/2}(\ln \mathbf{u} - \boldsymbol{\mu}_i) | \mathbf{I}_m - \boldsymbol{\Delta}'_i \boldsymbol{\Delta}_i), \mathbf{u} \in \mathbb{R}^{n_i^+}, \quad (3.30)$$

e a distribuição condicional de $\mathbf{U}_1 | \mathbf{U}_2 = \mathbf{u}_2$ é dada por

$$f_{\mathbf{U}_1 | \mathbf{U}_2 = \mathbf{u}_2}(\mathbf{u}_1) = \frac{\phi_{n_1}(\ln \mathbf{u}_1 - \boldsymbol{\mu}_1 | \boldsymbol{\Sigma}_{11}) \Phi_m(\boldsymbol{\Delta}'_2 \boldsymbol{\Sigma}_{22}^{-1/2}(\ln \mathbf{u}_2 - \boldsymbol{\mu}_2) + \boldsymbol{\Delta}'_1 \boldsymbol{\Sigma}_{11}^{-1/2}(\ln \mathbf{u}_1 - \boldsymbol{\mu}_1) | \mathbf{I}_m - \boldsymbol{\Delta}' \boldsymbol{\Delta})}{\left(\prod_{j=1}^{n_1} u_j \right) \Phi_m(\boldsymbol{\Delta}'_2 \boldsymbol{\Sigma}_{22}^{-1/2}(\ln \mathbf{u}_2 - \boldsymbol{\mu}_2) | \mathbf{I}_m - \boldsymbol{\Delta}'_2 \boldsymbol{\Delta}_2)}, \quad (3.31)$$

$\mathbf{u}_1 \in \mathbb{R}^{n_1^+}$ e $\mathbf{u}_2 \in \mathbb{R}^{n_2^+}$.

Perceba que a família definida na Expressão (3.22) é caso particular desta família quando $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \mathbf{I}_n$.

No próximo capítulo, apresentaremos os cálculos para a entropia e o índice de informação mútua para a família de distribuições log-*skew*-normal fundamental canônica multivariada, definida na Seção 3.2, e para a família de distribuições log-*skew*-normal multivariada, definida por Marchenko e Genton (2010).

Capítulo 4

Entropia e Informação Mútua em Famílias de Distribuições Assimétricas

A teoria da informação, também conhecida como teoria da comunicação, foi introduzida por Shannon (1948) para lidar com problemas das áreas de sistemas da comunicação, transmissão de dados criptografados e compressão de dados, entre outros. Esta área do conhecimento busca quantificar a informação de um sistema e, para isto, Shannon definiu uma medida chamada entropia, que é atribuída à incerteza da informação e pode ser entendida como uma medida matemática da informação necessária, em média, para descrever o comportamento de uma quantidade aleatória. A entropia de um vetor aleatório $\mathbf{Z} \in \mathbb{R}^n$, cuja fdp é $f_{\mathbf{Z}}(\mathbf{z})$, é definida como

$$H_{\mathbf{Z}} = -E[\ln(f_{\mathbf{Z}}(\mathbf{Z}))] = - \int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{z}) \ln(f_{\mathbf{Z}}(\mathbf{z})) d\mathbf{z}. \quad (4.1)$$

A necessidade de comparar a informação trazida por duas quantidades aleatórias fez com que se buscassem medidas mais efetivas para compará-las. Uma destas medidas é o índice de informação mútua, que mede a quantidade de informação que um vetor aleatório contém acerca do outro. Formalmente, sejam $\mathbf{X} \in \mathbb{R}^n$ e $\mathbf{Y} \in \mathbb{R}^m$ vetores aleatórios com distribuições marginais dadas, respectivamente, por $f_{\mathbf{X}}(\mathbf{x})$ e $f_{\mathbf{Y}}(\mathbf{y})$, e cuja distribuição conjunta é $f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})$. A informação mútua entre \mathbf{X} e \mathbf{Y} é definida como

$$I_{\mathbf{X}\mathbf{Y}} = -E \left[\ln \left(\frac{f_{\mathbf{X},\mathbf{Y}}(\mathbf{X}, \mathbf{Y})}{f_{\mathbf{X}}(\mathbf{X})f_{\mathbf{Y}}(\mathbf{Y})} \right) \right] = - \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) \ln \left(\frac{f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})}{f_{\mathbf{X}}(\mathbf{x})f_{\mathbf{Y}}(\mathbf{y})} \right) dx dy. \quad (4.2)$$

Note que a informação mútua fornece o ganho esperado de informação por se incluir a correlação entre \mathbf{X} e \mathbf{Y} se comparado à independência entre eles. Mostra-se que $I_{\mathbf{X}\mathbf{Y}}$ pode ser expressa em função das entropias marginais de \mathbf{X} e \mathbf{Y} e da entropia conjunta de (\mathbf{X}, \mathbf{Y}) , da seguinte forma:

$$I_{\mathbf{X}\mathbf{Y}} = H_{\mathbf{X}} + H_{\mathbf{Y}} - H_{\mathbf{X}\mathbf{Y}}, \quad (4.3)$$

Da Expressão (4.2), vemos que se \mathbf{X} e \mathbf{Y} são independentes, então $I_{\mathbf{X}\mathbf{Y}}=0$. Por outro lado, $I_{\mathbf{X}\mathbf{Y}}$ cresce com o aumento da dependência entre \mathbf{X} e \mathbf{Y} . Além disto, $I_{\mathbf{X}\mathbf{Y}} \geq 0$, para quaisquer vetores \mathbf{X} e \mathbf{Y} .

Outro conceito importante, neste contexto, é o de entropia relativa entre duas distribuições de probabilidade $F_{\mathbf{X}}$ e $F_{\mathbf{Y}}$, a qual mede a ineficiência em se assumir que uma distribuição é $f_{\mathbf{X}}$ quando a verdadeira distribuição é $f_{\mathbf{Y}}$. A entropia relativa, denotada por $D(f_{\mathbf{X}}||f_{\mathbf{Y}})$, é uma generalização do conceito apresentado em (4.2), e sua expressão é dada por

$$D(f_{\mathbf{X}}||f_{\mathbf{Y}}) = E_{\mathbf{X}} \left[\ln \left(\frac{f_{\mathbf{X}}(\mathbf{X})}{f_{\mathbf{Y}}(\mathbf{X})} \right) \right] = \int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) \left[\ln \left(\frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{Y}}(\mathbf{x})} \right) \right] d\mathbf{x}. \quad (4.4)$$

Muitos autores já calcularam os valores das entropias de Shannon e os índices de informação mútua para as distribuições mais conhecidas na literatura. Kullback (1978) faz os cálculos para a distribuição normal, Ahmed e Gokhale (1989) encontram expressões para distribuições multivariadas, Javier e Gupta (2008, 2009) estudam a informação mútua para distribuições não-normais que pertencem à família de distribuições contínuas multivariadas de locação e escala. Arellano-Valle *et al.* (2011) estudam os conceitos de entropia e informação mútua na família de distribuições elíptica multivariada e também na família de distribuições *skew*-elíptica multivariada, apresentando resultados particulares para as famílias de distribuições *skew*-normal e *skew*-t multivariadas. Além disto, aplicam o conceito de informação mútua dentro das famílias assimétricas para modelar o comportamento de estações de monitoramento de ozônio em Santiago de Chile.

Neste capítulo, faremos uma revisão sobre a entropia das famílias de distribuições estudadas ao longo deste trabalho. Na Seção 4.1, apresentaremos os resultados para famílias de distribuições univariadas e tentaremos relacionar sua entropia com a entropia de uma variável aleatória com distribuição normal. Na Seção 4.2, faremos uma revisão sobre a entropia nas famílias de distribuições elípticas e *skew*-elípticas, estudadas por Arellano-Valle *et al.* (2011), e também calcularemos a entropia para a família de distribuições *log-skew*-normal multivariada, estudada por Marchenko e Genton (2010). A seguir, encontraremos expressões para a entropia da família de distribuições *skew*-normal fundamental canônica multivariada. Na Seção 4.3, daremos nossa maior contribuição nesta área, calculando a entropia e informação mútua na família de distribuições *log-skew*-normal fundamental canônica multivariada, introduzida no Capítulo 3 e definida na Expressão (3.17), e buscaremos relacioná-la com a entropia de algumas famílias de distribuições, tais como a CFUSN, SN e normal, por exemplo. Com o objetivo de compararmos as famílias de distribuições *log-skew*-normal fundamental canônica multivariada e *log-skew*-normal, também encontraremos uma expressão para a entropia relativa entre estas duas famílias e estudaremos seu comportamento.

Para simplificar, ao longo deste capítulo, usaremos a notação $\Delta^* = I_m - \Delta' \Delta$. Também consideraremos as notações usadas nos capítulos anteriores.

4.1 Entropia nas Famílias de Distribuições Univariadas

Nesta seção vamos rever os resultados do cálculo da entropia para variáveis univariadas já vistas ao longo deste trabalho e buscaremos entender a relação entre elas. Nesta seção H_D denotará a entropia de um objeto aleatório com distribuição D .

Nosso interesse está em obter a entropia de distribuições que se relacionam com a distribuição normal, como as famílias de distribuição *log*-normal, *skew*-normal e *log-skew*-normal, por

exemplo. Começaremos mostrando um resultado para a entropia da família de distribuições normal univariada e buscaremos relacioná-lo com os resultados seguintes. Se X é uma variável aleatória com distribuição normal padrão, então a entropia de X é dada por

$$H_{N(0,1)} = \frac{1}{2}(1 + \ln(2\pi)). \quad (4.5)$$

Podemos adicionar escala e locação à distribuição normal e calcular a entropia desta variável. Então, se $X \sim N(\mu, \sigma^2)$, temos que

$$H_{N(\mu, \sigma^2)} = \frac{1}{2} \ln \sigma^2 + \frac{1}{2}(1 + \ln(2\pi)). \quad (4.6)$$

Note que a entropia da distribuição normal com escala e locação não depende do parâmetro de locação μ .

Nosso interesse está particularmente voltado para distribuições que não colocam massa em valores negativos, como é o caso da distribuição log-normal univariada, definida na Expressão (1.13). Na seguinte proposição obtemos o valor da entropia para esta variável.

Proposição 4.1. *Se $X \sim LN(\mu, \sigma^2)$, então a entropia da variável aleatória X é dada por*

$$H_{LN(\mu, \sigma^2)} = H_{N(\mu, \sigma^2)} + \mu, \quad (4.7)$$

onde $H_{N(\mu, \sigma^2)}$ é a entropia da distribuição normal univariada com média μ e variância σ^2 .

Demonstração. Considere a definição de entropia dada na Expressão (4.1), a fdp da variável log-normal dada na Expressão (1.13) e a transformação $y = \ln x$. Segue que

$$\begin{aligned} H_{LN(\mu, \sigma^2)} &= - \int_0^\infty \frac{1}{x} \phi(\ln x | \mu, \sigma^2) \ln \left[\frac{1}{x} \phi(\ln x | \mu, \sigma^2) \right] dx \\ &= - \int_{-\infty}^\infty \phi(y | \mu, \sigma^2) \ln [e^{-y} \phi(y | \mu, \sigma^2)] dy \\ &= - \int_{-\infty}^\infty \phi(y | \mu, \sigma^2) \ln [\phi(y | \mu, \sigma^2)] dy + \int_{-\infty}^\infty y \phi(y | \mu, \sigma^2) dy \\ &= H_{N(\mu, \sigma^2)} + \mu. \end{aligned}$$

□

Como pode ser visto, a entropia da distribuição log-normal univariada é obtida a partir da entropia da distribuição normal univariada. Contrário ao que ocorria na família de distribuições normal, a entropia da distribuição log-normal univariada dependerá também do parâmetro de locação μ .

Considere agora a família de distribuições *skew-normal* univariada definida na Expressão 1.3 do Capítulo 1. A Proposição 4.2 nos mostra o resultado sobre a entropia nesta família de distribuições.

Proposição 4.2. Se $X \sim SN(\mu, \sigma^2, \alpha)$, então a entropia da variável aleatória X é dada por

$$H_{SN(\mu, \sigma^2, \alpha)} = H_{N(\mu, \sigma^2)} - E[\ln(2\Phi(\alpha X_0))], \quad (4.8)$$

onde $X_0 \sim SN(\alpha)$ e $H_{N(\mu, \sigma^2)}$ é a entropia da distribuição normal univariada com média μ e variância σ^2 .

Demonstração. Considere a definição de entropia dada na Expressão (4.1), a fdp dada na Expressão (1.12) e a transformação $x_0 = \frac{x-\mu}{\sigma}$. Segue que

$$\begin{aligned} H_{SN(\mu, \sigma^2, \alpha)} &= - \int_{-\infty}^{\infty} \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\alpha \left(\frac{x-\mu}{\sigma}\right)\right) \ln \left[\frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\alpha \left(\frac{x-\mu}{\sigma}\right)\right) \right] dx \\ &= - \int_{-\infty}^{\infty} 2\phi(x_0)\Phi(\alpha x_0) \ln \left[\frac{2}{\sigma} \phi(x_0)\Phi(\alpha x_0) \right] dx_0 \\ &= - \int_{-\infty}^{\infty} 2\phi(x_0)\Phi(\alpha x_0) \ln[\phi(x_0)] dx_0 - \int_{-\infty}^{\infty} 2\phi(x_0)\Phi(\alpha x_0) \ln \left[\frac{2}{\sigma} \Phi(\alpha x_0) \right] dx_0 \\ &= - \int_{-\infty}^{\infty} 2\phi(x_0)\Phi(\alpha x_0) \left(\frac{-x_0^2}{2} \right) dx_0 + \frac{1}{2} \ln 2\pi \\ &\quad - \int_{-\infty}^{\infty} 2\phi(x_0)\Phi(\alpha x_0) \ln[\Phi(\alpha x_0)] dx_0 + \frac{1}{2} \ln \sigma^2 \\ &= \frac{1}{2} \ln 2\pi + \frac{1}{2} \ln \sigma^2 + \frac{1}{2} E(X_0^2) - E[\ln(2\Phi(\alpha X_0))] \end{aligned}$$

Como $X_0 \sim SN(\alpha)$, então $E(X_0^2) = 1$, o que conclui a prova. □

Novamente vemos que a entropia da distribuição *skew*-normal univariada é função da entropia da distribuição normal univariada e, além disto, depende da função $\Phi(\cdot)$ responsável por assimetrizar a distribuição.

Outra distribuição que coloca massa apenas em valores não-negativos é a distribuição *log-skew*-normal univariada, definida na Expressão (1.12) quando $n = 1$. A entropia desta distribuição é dada a seguir na Proposição (4.3).

Proposição 4.3. Se $X \sim LSN(\mu, \sigma^2, \alpha)$, então a entropia da variável aleatória X é dada por

$$H_{LSN(\mu, \sigma^2, \alpha)} = H_{SN(\mu, \sigma^2, \alpha)} + E(Y), \quad (4.9)$$

onde $H_{SN(\mu, \sigma^2)}$ é a entropia da distribuição normal univariada com média μ e variância σ^2 , $Y \sim SN(\mu, \sigma^2, \alpha)$ e $E(Y)$ é dada na Expressão (1.11).

Demonstração. Considere a definição de entropia dada na Expressão (4.1) e a transformação

$y = \ln x$. Segue que

$$\begin{aligned}
H_{LSN(\mu, \sigma^2, \alpha)} &= - \int_0^\infty \frac{2}{x} \phi(\ln x | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(\ln x - \mu)) \ln \left[\frac{2}{x} \phi(\ln x | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(\ln x - \mu)) \right] dx \\
&= - \int_{-\infty}^\infty 2\phi(y | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(y - \mu)) \ln [2e^{-y} \phi(y | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(y - \mu))] dy \\
&= - \int_{-\infty}^\infty 2\phi(y | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(y - \mu)) \ln [2\phi(y | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(y - \mu))] dy \\
&\quad + \int_{-\infty}^\infty 2y\phi(y | \mu, \sigma^2) \Phi(\alpha \sigma^{-1}(y - \mu)) dy \\
&= H_{SN(\mu, \sigma^2, \alpha)} + E(Y).
\end{aligned}$$

□

Percebemos que esta entropia está relacionada com a entropia da distribuição *skew*-normal univariada e, conseqüentemente, está relacionada com a entropia da distribuição normal univariada. Ela segue o mesmo padrão encontrado na entropia da distribuição log-normal dada na Expressão (4.15), sendo a entropia da distribuição original sem a transformação logarítmica adicionada da esperança de uma variável aleatória também com distribuição na família original. Substituindo a expressão para a entropia da distribuição *skew*-normal univariada e o valor de $E(Y)$ dado no Capítulo 1 na Expressão (1.11), temos que a Expressão (4.9) pode ser escrita como

$$H_{LSN(\mu, \sigma^2, \alpha)} = H_{N(\mu, \sigma^2)} - E[\ln(2\Phi(\alpha X_0))] + \mu + \sigma \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1 + \alpha^2}}.$$

Note também que a entropia da distribuição log-*skew*-normal univariada se relaciona com a entropia da distribuição log-normal, como se segue.

$$H_{LSN(\mu, \sigma^2, \alpha)} = H_{LN(\mu, \sigma^2)} - E[\ln(2\Phi(\alpha X_0))] + \sigma \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1 + \alpha^2}}.$$

4.2 Entropia nas Famílias de Distribuições Multivariadas

Apresentaremos, nesta seção, resultados relacionados à entropia das famílias de distribuições multivariadas que discutimos ao longo deste trabalho. Assim como na Seção 4.1, procuraremos relacioná-las com a entropia da família de distribuições normal multivariada e avaliar a sua dependência em relação aos parâmetros.

Lema 4.1. *Seja \mathbf{Z} vetor aleatório com distribuição normal multivariada com escala e locação. Então a entropia de $\mathbf{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ é dada por*

$$H_{N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} = \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{n}{2} (1 + \ln(2\pi)). \quad (4.10)$$

Temos também que $H_{N_n(\mathbf{0}, \mathbf{I}_n)} = \frac{n}{2}(1 + \ln(2\pi))$ e $H_{N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} = \frac{1}{2} \ln |\boldsymbol{\Sigma}| + H_{N_n(\mathbf{0}, \mathbf{I}_n)}$, onde $H_{N_n(\mathbf{0}, \mathbf{I}_n)}$ é a entropia de um vetor aleatório com distribuição normal padrão multivariada. Note que a entropia da distribuição normal não depende do parâmetro de locação $\boldsymbol{\mu}$, assim como ocorre no caso univariado.

No Capítulo 1, vimos que Azzalini (1985) mostra que se $Z \sim SN(\alpha)$, então $Z^2 \sim \chi_1^2$ para qualquer valor do parâmetro de assimetria α . Este resultado também foi generalizado por Azzalini e Dalla Valle (1996) para a família de distribuições *skew*-normal multivariada, onde se prova que se $\mathbf{Z} \sim SN_n(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$, então a forma quadrática $\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z} \sim \chi_n^2$. Note que, em ambos os casos, a forma quadrática não depende da função de assimetria $\Phi(\cdot)$. Segundo Genton e Loperfido (2005), esta propriedade se estende para famílias de distribuições ainda mais gerais, como, por exemplo, para a família de distribuições *skew*-elíptica generalizada definida a seguir.

Definição 4.1. *Um vetor aleatório $\mathbf{Z} \in \mathbb{R}^n$ tem distribuição skew-elíptica generalizada multivariada com vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^n$, matriz de escala positiva definida $\boldsymbol{\Sigma}$ de ordem $n \times n$ e função de assimetria π , denotada por $\mathbf{Z} \sim GSE_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \pi, g)$, se sua fdp é da forma*

$$f_{\mathbf{Z}}(\mathbf{z}) = 2|\boldsymbol{\Sigma}|^{-\frac{1}{2}}g(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{z} - \boldsymbol{\mu}))\pi(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{z} - \boldsymbol{\mu})), \mathbf{z} \in \mathbb{R}^n, \quad (4.11)$$

onde g é a fdp de um vetor com distribuição esférica e π é tal que $0 \leq \pi(\mathbf{z}) \leq 1$ e $\pi(-\mathbf{z}) = 1 - \pi(\mathbf{z}), \forall \mathbf{z} \in \mathbb{R}^n$.

Genton e Loperfido (2005) mostram que para um vetor aleatório $\mathbf{Z} \sim GSE_n(\mathbf{0}, \boldsymbol{\Sigma}, \pi, \mathbf{g})$ e qualquer função par $\tau(\cdot)$ de \mathbf{Z} (isto é, qualquer função tal que $\tau(-\mathbf{z}) = \tau(\mathbf{z}), \forall \mathbf{z} \in \mathbb{R}^n$), a distribuição do objeto aleatório $\tau(\mathbf{Z})$ não dependerá da função de assimetria π . A Proposição 4.4 apresenta este resultado.

Proposição 4.4. *Se $\mathbf{Z} \sim GSE_n(\mathbf{0}, \boldsymbol{\Sigma}, \mathbf{g}, \pi)$, então a função de distribuição de $\tau(\mathbf{Z})$, onde τ é uma função par, não depende da função de assimetria π .*

Como consequência da Proposição 4.4, segue que, para qualquer vetor aleatório \mathbf{Z} da família de distribuições *skew*-elíptica generalizada, a distribuição da forma quadrática $\mathbf{Z}'\mathbf{Z}$ não depende da função de assimetria π . Por exemplo, se $\mathbf{X}_0 \sim El_n(\mathbf{0}, \mathbf{I}_n, g^{(n)})$ e $\mathbf{Z}_0 \sim SEl_n(\mathbf{0}, \mathbf{I}_n, \boldsymbol{\alpha}, g^{(n)})$, então $\mathbf{X}_0'\mathbf{X}_0 \stackrel{d}{=} \mathbf{Z}_0'\mathbf{Z}_0$. Este resultado é importante para a obtenção da entropia de um vetor aleatório com distribuição *skew*-elíptica multivariada, encontrada por Arellano-Valle *et al.* (2011), e que é fornecido na Proposição 4.5.

Proposição 4.5. *A entropia de um vetor aleatório $n \times 1$ com distribuição skew-elíptica multivariada $\mathbf{Z} \sim SEl_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, g^{(n+1)})$ é dada por*

$$H_{SEl_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} = H_{El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} - E[\ln[2F(\bar{\boldsymbol{\alpha}}'\mathbf{Z}_0; g^{(1)})]], \quad (4.12)$$

onde $H_{El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})}$ é a entropia de um vetor aleatório com distribuição $\mathbf{X} \sim El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n+1)})$, $\mathbf{Z}_0 \sim SEl_n(\mathbf{0}, \mathbf{I}_n, \bar{\boldsymbol{\alpha}}, g^{(n+1)})$ e $\bar{\boldsymbol{\alpha}} = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\alpha}$.

Demonstração. Considere a definição de entropia dada na Expressão (4.1) e transformação $\mathbf{Z}_0 = \boldsymbol{\Sigma}^{-1/2}(\mathbf{Z} - \boldsymbol{\mu})$. Temos que

$$\begin{aligned}
H_{SE\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} &= - \int_{-\infty}^{\infty} 2f_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n+1)})F(\boldsymbol{\alpha}'(\mathbf{x} - \boldsymbol{\mu}); g^{(1)}) \\
&\quad \ln[2f_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, g^{(n+1)})F(\boldsymbol{\alpha}'(\mathbf{x} - \boldsymbol{\mu}); g^{(1)})]d\mathbf{x} \\
&= - \int_{-\infty}^{\infty} 2f_n(\mathbf{z}_0; g^{(n+1)})F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{z}_0; g^{(1)}) \ln[f_n(\mathbf{z}_0; g^{(n+1)})]d\mathbf{z}_0 - \\
&\quad \int_{-\infty}^{\infty} 2f_n(\mathbf{z}_0; g^{(n+1)})F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{z}_0; g^{(1)}) \ln[2F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{z}_0; g^{(1)})]d\mathbf{z}_0.
\end{aligned}$$

Como $f_n(\mathbf{z}_0) = h^{(n+1)}(\mathbf{z}'_0\mathbf{z}_0)$ é uma função par, segue da Proposição 4.4 que a sua distribuição não depende da função de assimetria $F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{z}_0; g^{(1)})$. Então, temos que

$$\begin{aligned}
H_{SE\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} &= - \int_{-\infty}^{\infty} 2f_n(\mathbf{z}_0; g^{(n+1)}) \ln[f_n(\mathbf{z}_0; g^{(n+1)})]d\mathbf{z}_0 - \\
&\quad \int_{-\infty}^{\infty} 2f_n(\mathbf{z}_0; g^{(n+1)})F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{z}_0; g^{(1)}) \ln[2F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{z}_0; g^{(1)})]d\mathbf{z}_0 \\
&= H_{E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} - E[\ln[2F(\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{1/2}\mathbf{Z}_0; g^{(1)})]] \\
&= H_{E\ell_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} - E[\ln[2F(\bar{\boldsymbol{\alpha}}'\mathbf{Z}_0; g^{(1)})]],
\end{aligned}$$

onde $\bar{\boldsymbol{\alpha}} = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\alpha}$. □

Segue da Proposição 4.5 que para calcularmos a entropia de um vetor aleatório com distribuição *skew*-elíptica multivariada basta conhecermos as distribuições de $U = \bar{\boldsymbol{\alpha}}'\mathbf{Z}_0$ e de $S = \mathbf{Z}'_0\mathbf{Z}_0$. Particularmente, se consideramos a família de distribuições *skew*-normal multivariada definida na Expressão (2.2), decorre diretamente de (4.12) que a entropia de um vetor aleatório $\mathbf{X} \sim SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ é dada por

$$H_{SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} = H_{N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} - E[\ln(2\Phi(\bar{\boldsymbol{\alpha}}'Z_0))], \quad (4.13)$$

onde $H_{N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})}$ é a entropia de $\mathbf{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e $\mathbf{Z}_0 \sim SN_n(\mathbf{0}, \mathbf{I}_n, \bar{\boldsymbol{\alpha}})$.

Nosso interesse está também voltado para famílias de distribuições multivariadas que colocam massa nula em valores negativos, como as famílias de distribuições log-normal e log-*skew*-normal multivariadas. Na proposição seguinte obtemos a entropia de um vetor de dimensão $n \times 1$ com distribuição log-*skew*-normal multivariada, definida na Expressão (2.6).

Proposição 4.6. *A entropia de um vetor aleatório $n \times 1$ com distribuição log-skew-normal multivariada $\mathbf{X} \sim LSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ é dada por*

$$H_{LSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} = H_{SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} + \sum_{i=1}^n E(Y_i), \quad (4.14)$$

onde $H_{SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})}$ é a entropia de um vetor aleatório \mathbf{X} com distribuição na família $SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$ e Y_i é a i -ésima componente do vetor aleatório $\mathbf{Y} \sim SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$.

Demonstração. Considere a definição de entropia dada na Expressão (4.1), a fdp dada na Expressão (2.6) e a transformação $\mathbf{y} = \ln \mathbf{x}$. Segue que

$$\begin{aligned}
H_{LSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} &= - \int_0^\infty 2 \left(\prod_{i=1}^n x_i \right)^{-1} \phi_n(\ln \mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\ln \mathbf{x} - \boldsymbol{\mu})) \\
&\quad \ln \left[2 \left(\prod_{i=1}^n x_i \right)^{-1} \phi_n(\ln \mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\ln \mathbf{x} - \boldsymbol{\mu})) \right] d\mathbf{x} \\
&= - \int_{-\infty}^\infty 2 \phi_n(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})) \ln \left[2 e^{-\sum_{i=1}^n y_i} \phi_n(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})) \right] d\mathbf{y} \\
&= - \int_{-\infty}^\infty 2 \phi_n(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})) \ln \left[2 \phi_n(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})) \right] d\mathbf{y} + \\
&\quad \int_{-\infty}^\infty 2 \phi_n(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})) \left[\sum_{i=1}^n y_i \right] d\mathbf{y} \\
&= H_{SN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}) + \sum_{i=1}^n E(Y_i).
\end{aligned}$$

□

Segue diretamente das Expressões (4.13) e (4.14) a relação entre a entropia da família de distribuições log-skew-normal e normal multivariada. Como a família de distribuições log-normal multivariada, definida na Expressão (2.7), é caso particular desta família quando $\boldsymbol{\alpha} = \mathbf{0}$, obtemos que a entropia de $\mathbf{X} \sim LN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ é

$$H_{LN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})} = H_{N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} + \sum_{i=1}^n \mu_i. \quad (4.15)$$

Como pode ser visto, a entropia de um vetor na família de distribuições log-normal multivariada é obtido a partir da entropia da distribuição normal padrão multivariada e, contrário ao que ocorria na família de distribuições normal, dependerá dos parâmetros de locação μ_i 's da distribuição normal da qual foi obtido.

Outra família importante neste estudo é a família de distribuições skew-normal fundamental canônica n -variada, definida na Expressão (3.6). O cálculo da entropia para a família CFUSN é introduzido na proposição que se segue.

Proposição 4.7. *Se $\mathbf{X} \sim CFSUN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{\Delta})$, então sua entropia é dada por*

$$H_{CFSUN(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} = \frac{n}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{1}{2} \sum_{i=1}^n E(X_{i0}^2) - E_{X_0}[\ln(2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))], \quad (4.16)$$

onde X_{0i} é a i -ésima componente do vetor aleatório $\mathbf{X}_0 \sim CFUSN_{n,m}(\mathbf{0}, \mathbf{I}_n, \boldsymbol{\Delta})$.

Demonstração. Considere a definição de entropia dada na Expressão (4.1), a fdp dada na Expressão (3.17) e a padronização $\mathbf{x}_0 = \Sigma^{-1/2}(\mathbf{x} - \boldsymbol{\mu})$. Segue que

$$\begin{aligned}
H_{CFSUN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}) &= - \int_{-\infty}^{\infty} 2^m |\boldsymbol{\Sigma}|^{-1/2} \phi_n(\boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\
&\quad \ln \left[2^m |\boldsymbol{\Sigma}|^{-1/2} \phi_n(\boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \right] d\mathbf{x} \\
&= - \int_{-\infty}^{\infty} 2^m \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln \left[2^m |\boldsymbol{\Sigma}|^{-1/2} \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \right] d\mathbf{x}_0 \\
&= - \int_{-\infty}^{\infty} 2^m \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln [\phi_n(\mathbf{x}_0)] d\mathbf{x}_0 + \frac{1}{2} \ln |\boldsymbol{\Sigma}| - \\
&\quad \int_{-\infty}^{\infty} 2^m \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln [2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*)] d\mathbf{x}_0 \\
&= - \int_{-\infty}^{\infty} 2^m \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln \left[(2\pi)^{-n/2} e^{-\frac{x'_0 x_0}{2}} \right] d\mathbf{x}_0 - \\
&\quad E(\ln [2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*)]) + \frac{1}{2} \ln |\boldsymbol{\Sigma}| \\
&= \frac{n}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{1}{2} \sum_{i=1}^n E(X_{i0}^2) - E_{X_0}[\ln(2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))].
\end{aligned}$$

□

Na Expressão (4.16), podemos substituir os valores para $E(X_{i0}^2)$ obtidos da Expressão (3.10) e obtemos o seguinte resultado

$$\begin{aligned}
H_{CFSUN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}) &= \frac{n}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{n}{2} + \frac{1}{\pi} \left[- \sum_{i=1}^n \sum_{j=1}^m \Delta_{ij}^2 + \sum_{i=1}^n \left(\sum_{j=1}^m \Delta_{ij} \right)^2 \right] - \\
&\quad E_{X_0}[\ln(2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))],
\end{aligned} \tag{4.17}$$

onde Δ_{ij} são as componentes da matriz $\boldsymbol{\Delta}$.

Se $\boldsymbol{\Sigma}$ for matriz de covariâncias, a Expressão (4.17) ainda pode ser escrita em função da entropia da distribuição normal multivariada como

$$H_{CFSUN_{n,m}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}) = H_{N_n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) + \frac{1}{\pi} \left[- \sum_{i=1}^n \sum_{j=1}^m \Delta_{ij}^2 + \sum_{i=1}^n \left(\sum_{j=1}^m \Delta_{ij} \right)^2 \right] - E_{X_0}[\ln(2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))].$$

Em muitas situações práticas buscamos a construção de modelos mais parcimoniosos. Uma maneira de fazermos isto é considerando estruturas particulares para a matriz de covariâncias $\boldsymbol{\Sigma}$ ou para a matriz $\boldsymbol{\Delta}$, responsável pela assimetria no modelo. Um caso particular interessante é apresentado no seguinte corolário:

Corolário 4.1. Se $\mathbf{Z} \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$, onde $\boldsymbol{\Delta}$ é tal que $\boldsymbol{\Delta}'\boldsymbol{\Delta}$ é uma matriz diagonal de ordem $m \times m$, então a entropia de \mathbf{Z} é

$$H_{CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} = \frac{n}{2} \ln 2\pi + \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{n}{2} - E[\ln(2^m \Phi_m(\boldsymbol{\Delta}'\mathbf{X}_0 | \mathbf{I}_m - \boldsymbol{\Delta}'\boldsymbol{\Delta}))], \quad (4.18)$$

onde $\mathbf{X}_0 \sim CFUSN_{n,m}(\mathbf{0}, \mathbf{I}_n, \boldsymbol{\Delta})$.

Demonstração. Note que

$$-\sum_{i=1}^n \sum_{j=1}^m \Delta_{ij}^2 + \sum_{i=1}^n \left(\sum_{j=1}^m \Delta_{ij} \right)^2 = -\sum_{i=1}^n \sum_{j=1}^m \Delta_{ij}^2 + \sum_{i=1}^n \left(\sum_{j=1}^m \Delta_{ij}^2 + 2 \sum_{k=1}^m \Delta_{ik} \Delta_{jk} \right). \quad (4.19)$$

Se $\boldsymbol{\Delta}'\boldsymbol{\Delta}$ é diagonal, a parcela $2 \sum_{k=1}^m \Delta_{ik} \Delta_{jk}$ é nula para todo i e j . Logo, a Expressão (4.19) também vale zero e substituindo este resultado na Expressão (4.16), concluímos a prova. \square

4.3 Entropia na Família de Distribuições Log-*Skew*-Normal Fundamental Canônica Multivariada

Nesta seção apresentaremos alguns resultados relacionados ao estudo da entropia na família log-*skew* normal fundamental canônica multivariada, as quais, juntamente com os resultados da Proposição 4.7, são as principais contribuições deste capítulo. Vamos calcular a entropia de um vetor aleatório com distribuição log-*skew*-normal fundamental canônica multivariada, introduzido na Seção 3.2, e investigar sua relação com as entropias das famílias de distribuições normal, *skew*-normal, *skew*-normal fundamental canônica e log-normal multivariadas.

O resultado sobre a entropia de um vetor na família de distribuições log-*skew*-normal fundamental canônica multivariada é exibido na seguinte proposição.

Proposição 4.8. Seja $\mathbf{Z} \sim LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$. Então, a entropia do vetor aleatório \mathbf{Z} de dimensão $n \times 1$ é dada por

$$H_{LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} = H_{CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} + \sum_{i=1}^n E(X_i), \quad (4.20)$$

onde X_i é a i -ésima componente do vetor aleatório $\mathbf{X} \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$ e $H_{CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})}$ é a entropia de um vetor $\mathbf{X} \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$.

Demonstração. Considere a definição de entropia dada na Expressão (4.1), a fdp dada na

Expressão (3.29) e a transformação $\mathbf{x} = \ln \mathbf{z}$. Segue que

$$\begin{aligned}
H_{LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} &= - \int_0^\infty 2^m \left(\prod_{i=1}^n z_i \right)^{-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{z} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{z} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\
&\quad \ln \left[2^m \left(\prod_{i=1}^n z_i \right)^{-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{z} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{z} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \right] d\mathbf{z} \\
&= - \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\
&\quad \ln \left[2^m \left(e^{-\sum_{i=1}^n x_i} \right) |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \right] d\mathbf{x} \\
&= - \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\
&\quad \ln \left[2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \right] d\mathbf{x} \\
&\quad + \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \left(\sum_{i=1}^n x_i \right) d\mathbf{x} \\
&= H_{CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} + \sum_{i=1}^n E(X_i)
\end{aligned}$$

concluindo a prova. \square

Substituindo na Expressão (4.20) a entropia encontrada na Proposição (4.7) e a expressão para $\sum_{i=1}^n E(X_i)$ obtida através da Expressão (3.19), segue que

$$\begin{aligned}
H_{LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} &= \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{n}{2} \ln 2\pi + \frac{n}{2} + \frac{1}{\pi} \left[\sum_{i=1}^n \sum_{j=1}^m \Delta_{ij}^2 + \sum_{i=1}^n \left(\sum_{j=1}^m \Delta_{ij} \right)^2 \right] + \\
&\quad \sum_{i=1}^n \mu_i + \left(\frac{2}{\pi} \right)^{1/2} (\mathbf{1}_m \boldsymbol{\Delta} \boldsymbol{\Sigma}^{1/2})' \mathbf{1}_n - E_{X_0}[\ln(2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))] \quad (4.21)
\end{aligned}$$

onde $\mathbf{1}_n$ é um vetor de uns de ordem $n \times 1$.

Se a matriz $\boldsymbol{\Sigma}$ é de covariâncias, ainda é possível relacionar a Expressão (4.21) com a entropia da distribuição normal multivariada.

$$\begin{aligned}
H_{LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} &= H_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) + \frac{1}{\pi} \left[\sum_{i=1}^n \sum_{j=1}^m \Delta_{ij}^2 + \sum_{i=1}^n \left(\sum_{j=1}^m \Delta_{ij} \right)^2 \right] + \\
&\quad \sum_{i=1}^n \mu_i + \left(\frac{2}{\pi} \right)^{1/2} (\mathbf{1}_m \boldsymbol{\Delta} \boldsymbol{\Sigma}^{1/2})' \mathbf{1}_n - E_{X_0}[\ln(2^m \Phi_m(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))].
\end{aligned}$$

Da mesma forma como feito no caso da distribuição CFUSN, a entropia na família de distribuições LCFUSN é simplificada quando a matriz $\Delta'\Delta$ é diagonal. O corolário a seguir mostra este resultado e a prova é omitida por ser igual à prova do Corolário 4.1.

Corolário 4.2. *Se $\mathbf{Z} \sim LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \Delta)$, onde Δ é tal que $\Delta'\Delta$ é uma matriz diagonal de ordem $m \times m$, então a entropia de \mathbf{Z} é*

$$H_{LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \Delta)} = H_{N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})} + \sum_{i=1}^n E(X_i) - E[\ln(2^m \Phi_m(\Delta'\mathbf{X}_0 | \mathbf{I}_m - \Delta'\Delta))], \quad (4.22)$$

onde $\mathbf{X}_0 \sim CFUSN_{n,m}(\mathbf{0}, \mathbf{I}_n, \Delta)$ e X_i é a i -ésima componente do vetor aleatório $\mathbf{X} \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \Delta)$.

Além da entropia das famílias de distribuições multivariadas, estamos interessados em obter medidas que nos auxiliam a entender o que ocorre com o nosso conhecimento sobre uma partição de um vetor quando temos informação sobre a outra partição. Dado um vetor aleatório \mathbf{Z} e uma partição \mathbf{Z}_1 e \mathbf{Z}_2 deste vetor, a informação mútua $I_{\mathbf{Z}_1\mathbf{Z}_2}$ mede a quantidade de informação que o vetor aleatório \mathbf{Z}_1 contém acerca do vetor aleatório \mathbf{Z}_2 , ou seja, quantifica a redução da incerteza de \mathbf{Z}_2 por conhecimento de \mathbf{Z}_1 . A Proposição 4.9 nos mostra um resultado sobre o cálculo da informação de um vetor aleatório com distribuição log-skew-normal fundamental canônica multivariada.

Proposição 4.9. *Sejam $\mathbf{Z} \sim LCFUSN_{n,m}(\Delta)$, \mathbf{Z}_1 e \mathbf{Z}_2 partição do vetor \mathbf{Z} , tais que $\mathbf{Z}_1 \sim LCFUSN_{n_1,m}(\Delta_1)$ e $\mathbf{Z}_2 \sim LCFUSN_{n_2,m}(\Delta_2)$. Então, a informação do vetor aleatório $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)$ é dada por*

$$I_{\mathbf{Z}_1\mathbf{Z}_2} = E_X \left[\ln \left(\frac{\Phi_m(\Delta'\mathbf{X} | \mathbf{I}_m - \Delta'\Delta)}{2^m \Phi_m(\Delta'_1\mathbf{X}_1 | \mathbf{I}_m - \Delta'_1\Delta_1) \Phi_m(\Delta'_2\mathbf{X}_2 | \mathbf{I}_m - \Delta'_2\Delta_2)} \right) \right], \quad (4.23)$$

onde $\mathbf{X} \sim CFUSN_{n,m}(\Delta)$ e \mathbf{X}_1 e \mathbf{X}_2 formam uma partição de \mathbf{X} , tais que $\mathbf{X}_1 \sim CFUSN_{n_1,m}(\Delta_1)$ e $\mathbf{X}_2 \sim CFUSN_{n_2,m}(\Delta_2)$.

Demonstração. Temos, da Expressão (4.3) que a informação mútua do vetor aleatório \mathbf{Z} é dada por $I_{\mathbf{Z}_1\mathbf{Z}_2} = H_{\mathbf{Z}_1} + H_{\mathbf{Z}_2} - H_{\mathbf{Z}}$. Substituindo as entropias pelo resultado da Expressão (4.20), temos que

$$\begin{aligned} I_{\mathbf{Z}_1\mathbf{Z}_2} &= \frac{n_1}{2} \ln 2\pi + \frac{1}{2} \sum_{i=1}^{n_1} E(X_i^2) + \sum_{i=1}^{n_1} E(X_i) - E_{X_1}[\ln(2^m \Phi_m(\Delta'_1\mathbf{X}_1 | \mathbf{I}_m - \Delta'_1\Delta_1))] \\ &\quad + \frac{n_2}{2} \ln 2\pi + \frac{1}{2} \sum_{i=n_1+1}^{n_2} E(X_i^2) + \sum_{i=n_1+1}^{n_2} E(X_i) - E_{X_2}[\ln(2^m \Phi_m(\Delta'_2\mathbf{X}_2 | \mathbf{I}_m - \Delta'_2\Delta_2))] \\ &\quad - \frac{n}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^n E(X_i^2) - \sum_{i=1}^n E(X_i) + E_X[\ln(2^m \Phi_m(\Delta'\mathbf{X} | \mathbf{I}_m - \Delta'\Delta))] \\ &= E_X \left[\ln \left(\frac{\Phi_m(\Delta'\mathbf{X} | \mathbf{I}_m - \Delta'\Delta)}{2^m \Phi_m(\Delta'_1\mathbf{X}_1 | \mathbf{I}_m - \Delta'_1\Delta_1) \Phi_m(\Delta'_2\mathbf{X}_2 | \mathbf{I}_m - \Delta'_2\Delta_2)} \right) \right]. \end{aligned}$$

□

Uma medida para comparação entre duas distribuições é dada pela entropia relativa, que mede a ineficiência de se assumir que uma distribuição é $f_{\mathbf{X}}$ quando a verdadeira distribuição é $f_{\mathbf{Y}}$. Podemos calcular a entropia relativa entre a família de distribuições log-skew-normal fundamental canônica multivariada, definida na Expressão (3.17), e a família de distribuições log-skew-normal multivariada, definida em (2.5). A Proposição 4.10 traz este resultado.

Proposição 4.10. *Sejam $\mathbf{Z} \sim LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$ e $\mathbf{Y} \sim LSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$. Então, a entropia relativa $D(f_{\mathbf{Z}}||f_{\mathbf{Y}})$ entre \mathbf{Y} e \mathbf{Z} é dada por*

$$D(f_{\mathbf{Y}}||f_{\mathbf{Z}}) = E_{\mathbf{X}_0} \left[\ln \left(\frac{2^{m-1} \Phi_m(\boldsymbol{\Delta} \mathbf{X}_0 | \boldsymbol{\Delta}^*)}{\Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{X}_0)} \right) \right], \quad (4.24)$$

onde $\mathbf{X}_0 \sim CFUSN_{n,m}(\boldsymbol{\Delta})$.

Demonstração. Da Expressão (4.4), temos que a entropia relativa entre as distribuições de \mathbf{Z} e \mathbf{Y} é

$$\begin{aligned} D(f_{\mathbf{Z}}||f_{\mathbf{Y}}) &= \int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{w}) \ln \left[\frac{f_{\mathbf{Z}}(\mathbf{w})}{f_{\mathbf{Y}}(\mathbf{w})} \right] d\mathbf{w} \\ &= \int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{w}) \ln[f_{\mathbf{Z}}(\mathbf{w})] d\mathbf{w} - \int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{w}) \ln[f_{\mathbf{Y}}(\mathbf{w})] d\mathbf{w} \\ &= -H_{LCFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})} - \int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{w}) \ln[f_{\mathbf{Y}}(\mathbf{w})] d\mathbf{w}. \end{aligned}$$

O primeiro termo da Expressão (4.25) é dado na Proposição (4.20). O segundo termo é calculado como se segue. Considere a transformação $\mathbf{x} = \ln \mathbf{w}$. Com isto, segue que

$$\begin{aligned} \int_0^\infty f_{\mathbf{Z}}(\mathbf{w}) \ln[f_{\mathbf{Y}}(\mathbf{w})] d\mathbf{w} &= \int_0^\infty 2^m \left(\prod_{i=1}^n w_i \right)^{-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{w} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{w} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\ &\quad \ln \left[2^m \left(\prod_{i=1}^n w_i \right)^{-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\ln \mathbf{w} - \boldsymbol{\mu})) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\ln \mathbf{w} - \boldsymbol{\mu})) \right] d\mathbf{w} \\ &= \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\ &\quad \ln \left[2^m e^{-(\sum_{i=1}^n x_i)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{x} - \boldsymbol{\mu})) \right] d\mathbf{x} \\ &= \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\ &\quad \ln \left[\phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \right] d\mathbf{x} - \sum_{i=1}^n E_{X_i}(X_i) - \frac{1}{2} \ln |\boldsymbol{\Sigma}| \\ &\quad + \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})) \Phi_m(\boldsymbol{\Delta}' \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu}) | \boldsymbol{\Delta}^*) \\ &\quad \ln \left[2^m \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1}(\mathbf{x} - \boldsymbol{\mu})) \right] d\mathbf{x}. \end{aligned} \quad (4.25)$$

Segue de (4.25) e da padronização $\mathbf{X}_0 = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{X} - \boldsymbol{\mu})$ que

$$\begin{aligned}
\int_0^\infty f_{\mathbf{Z}}(\mathbf{w}) \ln[f_{\mathbf{Y}}(\mathbf{w})] d\mathbf{w} &= \int_{-\infty}^\infty 2^m |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln \left[(2\pi)^{-n/2} e^{\frac{x_0' \mathbf{x}_0}{2}} \right] d\mathbf{x} - \sum_{i=1}^n E_{X_i}(X_i) \\
&\quad - \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \int_{-\infty}^\infty 2^m \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln \left[2\Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{x}_0) \right] d\mathbf{x}_0. \\
&= -\frac{1}{2} \sum_{i=1}^n E(X_{0i}^2) - \frac{n}{2} \ln 2\pi - \sum_{i=1}^n E_{X_i}(X_i) - \frac{1}{2} \ln |\boldsymbol{\Sigma}| \\
&\quad + \int_{-\infty}^\infty 2^m \phi_n(\mathbf{x}_0) \Phi_m(\boldsymbol{\Delta}' \mathbf{x}_0 | \boldsymbol{\Delta}^*) \ln \left[2\Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{x}_0) \right] d\mathbf{x}_0. \tag{4.26}
\end{aligned}$$

onde $\mathbf{X}_0 \sim CFUSN_{n,m}(\boldsymbol{\Delta})$, X_{0i} é a i -ésima componente do vetor \mathbf{X}_0 e X_i é a i -ésima componente de um vetor $\mathbf{X} \sim CFUSN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$.

Substituindo o resultado dado na Expressão (4.26) na Expressão (4.25) e considerando o resultado sobre a entropia da família de distribuições log-*skew*-normal fundamental canônica multivariada dado na Expressão (4.20), temos:

$$\begin{aligned}
D(f_{\mathbf{Y}} || f_{\mathbf{Z}}) &= -H_{LCFUSN} - \int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{x}) \ln[f_{\mathbf{Y}}(\mathbf{x})] d\mathbf{x} \\
&= -\frac{1}{2} \ln |\boldsymbol{\Sigma}| - \frac{n}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^n E(X_{i0}^2) - \sum_{i=1}^n E(X_i) + E_{X_0}[\ln(2^m \Phi(\boldsymbol{\Delta}' \mathbf{X}_0 | \boldsymbol{\Delta}^*))] \\
&\quad + \frac{1}{2} \sum_{i=1}^n E(X_{0i}^2) + \frac{n}{2} \ln 2\pi + \sum_{i=1}^n E(X_i) - E_{\mathbf{X}_0} \left[\ln \Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{X}_0) \right] + \frac{1}{2} \ln |\boldsymbol{\Sigma}| \\
&= E_{\mathbf{X}_0} \left[\ln \left(\frac{2^{m-1} \Phi_m(\boldsymbol{\Delta} \mathbf{X}_0 | \boldsymbol{\Delta}^*)}{\Phi(\boldsymbol{\alpha}' \boldsymbol{\omega}^{-1} \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{X}_0)} \right) \right].
\end{aligned}$$

□

No próximo capítulo apresentaremos um breve estudo sobre inferência bayesiana nas famílias CFUSN e LCFUSN. Serão obtidas as distribuições condicionais completas para os parâmetros de interesse e variáveis latentes. Também ajustaremos dois modelos para um banco de dados reais.

Capítulo 5

Inferência Bayesiana em Modelos Assimétricos

Este capítulo se destina a discutir alguns aspectos relacionados à questão da inferência na família de distribuições log-*skew*-normal fundamental canônica, no caso em que $n = 1$, do ponto de vista bayesiano. Aplica-se a distribuição proposta para modelar o comportamento do índice de precipitação (PCP) nos Estados Unidos. Dados semelhantes já foram analisados em Marchenko e Genton (2010). Ajustaremos modelos considerando alguns valores distintos para o componente m da função que assimetriza a distribuição normal nesta família. O objetivo é avaliar se há melhoria na qualidade de ajuste ao aumentarmos o valor de m . O modelo, quando assumimos $m = 1$, retorna ao modelo log-*skew*-normal multivariado. Assim, também se fornece, neste trabalho, uma abordagem para estimar os parâmetros em tal modelo diferente daquela considerada por Marchenko e Genton (2010).

Do ponto de vista teórico, na Seção 5.1, apresentaremos o modelo e serão calculadas as distribuições condicionais completas para todos os parâmetros da família LCFUSN e da CFUSN. Na Seção 5.2, ilustramos a metodologia proposta aplicando-a a dados de precipitação nos EUA. Foram usados métodos MCMC para obter uma amostra da distribuição *a posteriori* dos parâmetros.

5.1 Modelo Estatístico

Considere uma amostra $\mathbf{x} = (x_1, x_2, \dots, x_n)$ que foi aleatoriamente selecionada de uma população que possui distribuição $LCFUSN_{1,m}(\mu, \sigma^2, \Delta)$, cuja fdp é dada na Expressão (3.29). Desta forma, tem-se que a função de verossimilhança de \mathbf{x} é dada por

$$\begin{aligned} f(\mathbf{x}|\mu, \sigma^2, \Delta) &= \prod_{i=1}^n [2^m \sigma^{-1}(x_i)^{-1} \phi(\sigma^{-1}(\ln x_i - \mu)) \Phi_m(\Delta' \sigma^{-1}(\ln x_i - \mu) | \mathbf{I}_m - \Delta' \Delta)] \\ &= 2^{nm} (2\pi\sigma^2)^{-n/2} \left(\prod_{i=1}^n x_i \right)^{-1} e^{-\sum_{i=1}^n \frac{(\ln x_i - \mu)^2}{2\sigma^2}} \prod_{i=1}^n \Phi_m(\Delta' \sigma^{-1}(\ln x_i - \mu) | \mathbf{I}_m - \Delta' \Delta). \end{aligned} \tag{5.1}$$

Para completar a especificação do modelo, é necessário especificar a distribuição *a priori* para os parâmetros. Neste trabalho, a dimensão m é fixada. O parâmetro μ pode assumir qualquer valor real e σ^2 qualquer valor real positivo. Por isso, *a priori*, se pode considerar que $\mu \sim N(m, v)$ e $\sigma^2 \sim GI(\alpha, \beta)$. O parâmetro Δ é um vetor de dimensão $1 \times m$ e deve ser tal que $\|\Delta \mathbf{a}\| < 1$, para todo vetor unitário $\mathbf{a} \in \mathbb{R}^m$, o que torna o processo de escolha da distribuição *a priori* para Δ bastante difícil. Para contornar este problema, podemos usar a parametrização alternativa dada na Expressão (3.7) do Capítulo 3, onde $\Delta = \Lambda(\mathbf{I}_m + \Lambda' \Lambda)^{-1/2}$ para alguma matriz real Λ de ordem $n \times m$ e com entradas finitas. Uma escolha possível para a distribuição *a priori* deste parâmetro é $\Lambda \sim N_m(\mathbf{A}, \mathbf{B})$. Considera-se também que todos os parâmetros são independentes entre si.

Para se fazer inferência considerando a nova parametrização, a função de verossimilhança deve ser reescrita em função de Λ e torna-se

$$f(\mathbf{x}|\mu, \sigma^2, \Lambda) = 2^{nm}(\sigma^2)^{-n/2}(2\pi)^{-n/2} \left(\prod_{i=1}^n x_i \right)^{-1} \exp\left(\frac{-\sum_{i=1}^n (\ln x_i - \mu)^2}{2\sigma^2}\right) \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \Lambda' \Lambda)^{-1} \Lambda' \sigma^{-1} (\ln x_i - \mu) | (\mathbf{I}_m + \Lambda' \Lambda)^{-1}). \quad (5.2)$$

Considerando as especificações *a priori* anteriormente mencionadas e a função de verossimilhança dada na Expressão (5.2), segue que as distribuições condicionais completas para μ , σ^2 e Λ são dadas, respectivamente, por

$$\begin{aligned} f(\mu|\mathbf{x}, \sigma^2, \Lambda) &\propto e^{-\sum_{i=1}^n \frac{(\ln x_i - \mu)^2}{2\sigma^2}} e^{-\frac{(\mu - m)^2}{2v^2}} \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \Lambda' \Lambda)^{-1} \Lambda' \sigma^{-1} (\ln x_i - \mu) | (\mathbf{I}_m + \Lambda' \Lambda)^{-1}) \\ &\propto \exp\left(\frac{-\mu^2(nv^2 + \sigma^2) + 2\mu(v^2 \sum_{i=1}^n \ln x_i + m\sigma^2)}{2v^2\sigma^2}\right) \\ &\quad \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \Lambda' \Lambda)^{-1} \Lambda' \sigma^{-1} (\ln x_i - \mu) | (\mathbf{I}_m + \Lambda' \Lambda)^{-1}) \\ &\propto \exp\left(-\frac{nv^2 + \sigma^2}{2v^2\sigma^2} \left(\mu - \frac{v^2 \sum_{i=1}^n \ln x_i + m\sigma^2}{nv^2 + \sigma^2}\right)^2\right) \\ &\quad \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \Lambda' \Lambda)^{-1} \Lambda' \sigma^{-1} (\ln x_i - \mu) | (\mathbf{I}_m + \Lambda' \Lambda)^{-1}) \\ &\propto \phi\left(\mu \left| \frac{v^2 \sum_{i=1}^n \ln x_i + m\sigma^2}{nv^2 + \sigma^2}, \frac{v^2\sigma^2}{nv^2 + \sigma^2}\right.\right) \Phi_{mn}((\mathbf{V}_1, \dots, \mathbf{V}_n) | \mathbf{I}_m \otimes (\mathbf{I}_m + \Lambda' \Lambda)^{-1}), \quad (5.3) \end{aligned}$$

onde $V_i = (\mathbf{I}_m + \Lambda' \Lambda)^{-1} \Lambda' \sigma^{-1} (\ln x_i - \mu)$ é um vetor de dimensão $m \times 1$ e $\mathbf{A} \otimes \mathbf{B}$ representa o produto de Kronecker entre as matrizes \mathbf{A} e \mathbf{B} ,

$$\begin{aligned}
f(\sigma^2|\mathbf{x}, \mu, \mathbf{\Lambda}) &\propto (\sigma^2)^{-(\alpha+1)} \exp\left(\frac{\beta}{\sigma^2}\right) \cdot (\sigma^2)^{-n/2} \exp\left(-\sum_{i=1}^n \frac{(\ln x_i - \mu)^2}{2\sigma^2}\right) \\
&\quad \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\sigma^{-1}(\ln x_i - \mu)|(\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}) \\
&\propto (\sigma^2)^{-(\frac{n}{2}+\alpha+1)} \exp\left(\frac{-\sum_{i=1}^n (\ln x_i - \mu)^2 + 2\beta}{2\sigma^2}\right) \\
&\quad \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\sigma^{-1}(\ln x_i - \mu)|(\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}) \\
&\propto (\sigma^2)^{-(\frac{n}{2}+\alpha+1)} \exp\left(\frac{-\sum_{i=1}^n (\ln x_i - \mu)^2 + 2\beta}{2\sigma^2}\right) \\
&\quad \Phi_{mn}((\mathbf{V}_1, \dots, \mathbf{V}_n)|\mathbf{I}_m \otimes (\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}), \tag{5.4}
\end{aligned}$$

$$\begin{aligned}
f(\mathbf{\Lambda}|\mathbf{x}, \mu, \sigma^2) &\propto \prod_{i=1}^n \Phi_m((\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\sigma^{-1}(\ln x_i - \mu)|(\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}) \\
&\quad \exp\left(-\frac{(\mathbf{\Lambda} - \mathbf{A})'\mathbf{B}^{-1}(\mathbf{\Lambda} - \mathbf{A})}{2}\right) \\
&\propto \phi(\mathbf{\Lambda}|\mathbf{A}, \mathbf{B})\Phi_{mn}((\mathbf{V}_1, \dots, \mathbf{V}_n)|\mathbf{I}_m \otimes (\mathbf{I}_m + \mathbf{\Lambda}'\mathbf{\Lambda})^{-1}). \tag{5.5}
\end{aligned}$$

Note que as distribuições condicionais completas de μ e $\mathbf{\Lambda}$ pertencem a famílias de distribuições normal assimetrizadas e a distribuição condicional completa de σ^2 está na família de distribuições gama invertida assimetrizada, onde a função de assimetização em todos os casos é a fda de uma distribuição normal mn -variada.

Apesar de ser possível o reconhecimento das famílias às quais pertencem as distribuições condicionais completas, suas dependências da fda de distribuições normais multivariadas tornam os métodos de amostragem da distribuição *a posteriori* muito caros do ponto de vista computacional. Para se contornar este problema, uma possibilidade é utilizar uma representação estocástica para a família LCFUSN.

Seja o vetor $\mathbf{z} = (z_1, \dots, z_n)$, onde cada componente z_i possui distribuição $CFUSN_{1,m}(\mu, \sigma^2, \mathbf{\Delta})$. Pode-se estimar os parâmetros da distribuição $LCFUSN_{1,m}(\mu, \sigma^2, \mathbf{\Delta})$ a partir do modelo construído para a distribuição $CFUSN_{1,m}(\mu, \sigma^2, \mathbf{\Delta})$. Para isto, basta aplicar a transformação logarítmica nos dados originais e ajustar um modelo CFUSN com os dados transformados.

Assim, o modelo dado na Expressão (5.1) pode ser hierarquicamente representado como segue. Se $W_i \sim LCFUSN_{1,m}(\mu, \sigma^2, \mathbf{\Delta})$, então $Z_i = \ln W_i \sim CFUSN_{1,m}(\mu, \sigma^2, \mathbf{\Delta})$. De Arellano-Valle e Genton (2005), segue que

$$Z_i \stackrel{d}{=} \sigma\mathbf{\Delta}|\mathbf{X}_i| + \sigma(1 - \mathbf{\Delta}\mathbf{\Delta}')^{1/2}Y_i + \mu, \tag{5.6}$$

onde $\mathbf{X}_i \sim N_m(\mathbf{0}, \mathbf{I}_m)$, $Y_i \sim N(0, 1)$, \mathbf{X}_i e Y_i são vetores aleatórios independentes e $|\mathbf{X}_i| =$

$(|X_{i1}|, \dots, |X_{im}|)'$. Conseqüentemente, temos que o modelo assumido na Expressão (5.1) é equivalente a

$$\begin{aligned} Z_i &= \ln W_i \\ Z_i | \mathbf{X}_i = \mathbf{x}_i &\sim N(\mu + \sigma \mathbf{\Delta} | \mathbf{x}_i|, \sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')) \\ \mathbf{X}_i &\sim N_m(\mathbf{0}, \mathbf{I}_m), \end{aligned} \quad (5.7)$$

onde \mathbf{X}_i é uma variável latente (não-observada). Esta estratégia é conhecida como técnica de aumento de dados (van Dyk e Meng, 2001) e facilita bastante a parte computacional. Usando esta representação, a função de verossimilhança é, agora, dada por

$$\begin{aligned} f_{\mathbf{Z} | \mathbf{X}=\mathbf{x}}(\mathbf{z}) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma \sqrt{(1 - \mathbf{\Delta} \mathbf{\Delta}')}} \exp\left(\frac{-(z_i - \mu - \sigma \mathbf{\Delta} | \mathbf{x}_i|)^2}{2\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')} \right) \\ &= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')} \right)^{\frac{n}{2}} \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \mathbf{\Delta} | \mathbf{x}_i|)^2}{2\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')} \right). \end{aligned} \quad (5.8)$$

Pelos mesmos motivos considerados anteriormente, usaremos a parametrização $\mathbf{\Delta} = \mathbf{\Lambda}(\mathbf{I}_m + \mathbf{\Lambda}' \mathbf{\Lambda})^{-1/2}$. Desta forma, a Expressão (5.8) pode ser reescrita como

$$f_{\mathbf{Z} | \mathbf{X}=\mathbf{x}}(\mathbf{z}) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1 + \mathbf{\Lambda} \mathbf{\Lambda}'}{\sigma^2}\right)^{\frac{n}{2}} \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \mathbf{\Lambda}(\mathbf{I}_m + \mathbf{\Lambda}' \mathbf{\Lambda})^{-1/2} | \mathbf{x}_i|)^2}{2\sigma^2(1 + \mathbf{\Lambda} \mathbf{\Lambda}')^{-1}}\right).$$

Considerando as mesmas distribuições *a priori* para os parâmetros definidas anteriormente, as distribuições condicionais completas para σ^2 , $\mathbf{\Lambda}$ e μ são, respectivamente, dadas por:

$$\begin{aligned} f(\sigma^2 | \mu, \mathbf{\Delta}, \mathbf{Z}, \mathbf{X}) &\propto \left(\frac{1}{\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')} \right)^{\frac{1}{2}} \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \mathbf{\Delta} | \mathbf{x}_i|)^2}{2\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')} \right) \cdot \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left(\frac{-\beta}{\sigma^2}\right) \\ &\propto \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2} + \alpha + 1} \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \mathbf{\Delta} | \mathbf{x}_i|)^2}{2\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')} - \frac{\beta}{\sigma^2}\right) \\ &\propto \left(\frac{1}{\sigma^2}\right)^{\frac{3}{2} + \alpha} \exp\left(\frac{\sum_{i=1}^n 2(z_i - \mu) \mathbf{\Delta} | \mathbf{x}_i|}{2\sigma(1 - \mathbf{\Delta} \mathbf{\Delta}')} - \frac{2\beta(\mathbf{I}_n - \mathbf{\Delta} \mathbf{\Delta}') + \sum_{i=1}^n (z_i - \mu)^2}{2\sigma^2(1 - \mathbf{\Delta} \mathbf{\Delta}')}\right), \end{aligned} \quad (5.9)$$

$$\begin{aligned} f(\mathbf{\Lambda} | \mu, \sigma, \mathbf{Z}, \mathbf{X}) &\propto \left(\frac{1 + \mathbf{\Lambda} \mathbf{\Lambda}'}{\sigma}\right)^{\frac{n}{2}} \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \mathbf{\Lambda}(\mathbf{I}_m + \mathbf{\Lambda}' \mathbf{\Lambda})^{-1/2} | \mathbf{x}_i|)^2}{2\sigma^2(1 - \mathbf{\Lambda} \mathbf{\Lambda}')^{-1}}\right) \\ &\quad \exp\left(\frac{-(\mathbf{\Lambda} - A)' B^{-1} (\mathbf{\Lambda} - A)}{2}\right) \\ &\propto (1 + \mathbf{\Lambda} \mathbf{\Lambda}')^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^n (z_i - \mu - \sigma \mathbf{\Lambda}(\mathbf{I}_m + \mathbf{\Lambda}' \mathbf{\Lambda})^{-1/2} | \mathbf{x}_i|)^2}{2\sigma^2(1 - \mathbf{\Lambda} \mathbf{\Lambda}')^{-1}} - \frac{(\mathbf{\Lambda} - A)' B^{-1} (\mathbf{\Lambda} - A)}{2}\right), \end{aligned} \quad (5.10)$$

$$\begin{aligned}
f(\mu|\sigma, \Delta, \mathbf{Z}, \mathbf{X}) &\propto \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \Delta |\mathbf{x}_i|)^2}{2\sigma^2(1 - \Delta \Delta')}\right) \exp\left(\frac{-(\mu - m)^2}{2v^2}\right) \\
&\propto \exp\left(\frac{-\sum_{i=1}^n \mu^2 + 2\mu \sum_{i=1}^n (z_i - \sigma \Delta |\mathbf{x}_i|)}{2\sigma^2(1 - \Delta \Delta')} + \frac{-\mu^2 + 2m\mu}{2v^2}\right) \\
&\propto \exp\left(-\mu^2 \left(\frac{nv^2 + \sigma^2(1 - \Delta \Delta')}{2\sigma^2 v^2(1 - \Delta \Delta')}\right) + 2\mu \left(\frac{v^2 \sum_{i=1}^n (z_i - \sigma \Delta |\mathbf{x}_i|) + m\sigma^2(1 - \Delta \Delta')}{2\sigma^2 v^2(1 - \Delta \Delta')}\right)\right) \\
&\propto \exp\left(-\frac{nv^2 + \sigma^2(1 - \Delta \Delta')}{2\sigma^2 v^2(1 - \Delta \Delta')} \left(\mu - \frac{v^2 \sum_{i=1}^n (z_i - \sigma \Delta |\mathbf{x}_i|) + m\sigma^2(1 - \Delta \Delta')}{nv^2 + \sigma^2(1 - \Delta \Delta')}\right)^2\right).
\end{aligned} \tag{5.11}$$

Note da Expressão (5.11) que a distribuição condicional completa para μ é

$$\mu|\sigma^2, \Delta, \mathbf{Z}, \mathbf{X} \sim N\left(\frac{v^2 \sum_{i=1}^n (z_i - \sigma \Delta |\mathbf{x}_i|) + m\sigma^2(1 - \Delta \Delta')}{nv^2 + \sigma^2(1 - \Delta \Delta')}, \frac{\sigma^2 v^2(1 - \Delta \Delta')}{nv^2 + \sigma^2(1 - \Delta \Delta')}\right).$$

Além disto, para a variável latente \mathbf{X} , temos que o núcleo da distribuição condicional completa é

$$\begin{aligned}
f(\mathbf{X}|\mu, \sigma^2, \Delta, \mathbf{Z}) &\propto \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \Delta |\mathbf{x}_i|)^2}{2\sigma^2(1 - \Delta \Delta')}\right) \cdot \exp\left(-\frac{\mathbf{x}_i' \mathbf{x}_i}{2}\right) \\
&\propto \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \Delta |\mathbf{x}_i|)^2}{2\sigma^2(1 - \Delta \Delta')} - \frac{\mathbf{x}_i' \mathbf{x}_i}{2}\right).
\end{aligned}$$

As Expressões (5.9), (5.11) e (5.12) podem ser reescritas em termos do parâmetro Λ , como se segue:

$$f(\sigma^2|\mu, \Lambda, \mathbf{Z}, \mathbf{X}) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{3}{2}+\alpha} \exp\left(\frac{\sum_{i=1}^n (z_i - \mu) \Lambda (\mathbf{I}_m + \Lambda' \Lambda)^{-1/2} |\mathbf{x}_i|}{\sigma(1 + \Lambda \Lambda')^{-1}} - \frac{2\beta(1 + \Lambda \Lambda')^{-1} + \sum_{i=1}^n (z_i - \mu)^2}{2\sigma^2(1 + \Lambda \Lambda')^{-1}}\right),$$

$$f(\mu|\sigma^2, \Lambda, \mathbf{Z}, \mathbf{X}) \propto \exp\left(-V \left(\mu - \frac{v^2 \sum_{i=1}^n (z_i - \sigma \Lambda (\mathbf{I}_m + \Lambda' \Lambda)^{-1/2} |\mathbf{x}_i|) + m\sigma^2(1 + \Lambda \Lambda')^{-1}}{nv^2 + \sigma^2(1 + \Lambda \Lambda')^{-1}}\right)^2\right),$$

onde $V = \frac{nv^2 + \sigma^2(1 + \Lambda \Lambda')^{-1}}{2\sigma^2 v^2(1 + \Lambda \Lambda')^{-1}}$, e

$$f(\mathbf{X}|\mu, \sigma^2, \Lambda, \mathbf{Z}) \propto \exp\left(\frac{-\sum_{i=1}^n (z_i - \mu - \sigma \Lambda (\mathbf{I}_m + \Lambda' \Lambda)^{-1/2} |\mathbf{x}_i|)^2}{2\sigma^2(1 + \Lambda \Lambda')^{-1}} - \frac{\mathbf{x}_i' \mathbf{x}_i}{2}\right).$$

Perceba que as distribuições condicionais completas para μ , σ^2 , Λ e \mathbf{X} não dependem mais da fda das distribuições normais multivariadas, tornando o processo de amostragem destas distribuições um pouco mais simples. O amostrador de Gibbs pode ser usado para amostrar da distribuição condicionais completa de μ e o algoritmo Metropolis-Hastings pode ser usado para amostrar das demais distribuições. Usando a representação hierárquica dada em (5.7), também podemos usar o *Software Winbugs* para obtermos amostras das distribuições *a posteriori*. Isto será feito na próxima seção.

5.2 Estudo de Casos: Precipitação Mensal nos EUA

Nesta seção, apresentam-se os dados relacionados à precipitação mensal nos Estados Unidos no período de 1895 a 2007. Dados semelhantes foram analisados em Marchenko e Genton (2010). O banco de dados está disponível no National Climatic Data Center (NCDC) e é composto por 1344 observações sobre o índice de precipitação americana (PCP), coletadas por mês ao longo dos anos de 1895 e 2007. O principal objetivo aqui é ajustar modelos dentro da família de distribuições log-*skew* normal fundamental canônica univariada, variando a dimensão m da função de assimetria, e compará-los. Dois modelos são considerados: o Modelo 1, em que $m = 1$, e o Modelo 2, em que $m = 2$. Na Tabela 5.1 pode-se ver algumas estatísticas descritivas do índice de precipitação americano, como o mínimo (Min), mediana (Md), média (Me), a variância (Var) e máximo (Max).

Tabela 5.1: Estatísticas descritivas do PCP

PCP	Min	Md	Me	Var	Max
	0.54	2.44	2.42	0.30	4.21

Para facilitar o ajuste, trabalha-se com o logaritmo dos dados, e por isto ajusta-se um modelo dentro da família de distribuições *skew* normal fundamental canônica univariada. Assim, os dados não transformados terão distribuição na família log-*skew* normal fundamental canônica univariada, mantendo os mesmo parâmetros de locação μ e de escala σ . A Figura 5.1 mostra o histograma dos dados transformados (esquerda) e dos dados originais (direita). Percebe-se que existe assimetria a esquerda em ambos os casos, indicando que as curvas assimétricas estudadas podem ser um modelo promissor para análise destes dados.

As distribuições *a priori* escolhidas foram sempre as pouco informativas. Considera-se que, *a priori*, $\mu \sim N(0, 100)$ e $\sigma^2 \sim Gama(0.1, 0.1)$ para os dois modelos. Quando $m = 1$, assumiu-se $\Delta \sim U(-1, 1)$. Já para o caso $m = 2$, em decorrência da suposição de que Δ é tal que $\|\Delta \mathbf{a}\| < 1$, para todo vetor unitário $\mathbf{a} \in \mathbb{R}^m$, escolhe-se reduzir o suporte de $\Delta = (\Delta_1, \Delta_2)$. Considera-se que cada entrada do vetor pode assumir apenas valores entre -0.5 e 0.5, satisfazendo a condição imposta, para qualquer vetor unitário \mathbf{a} . Então, assume-se que $\Delta_1 \sim U(-0.5, 0.5)$ e $\Delta_2 \sim U(-0.5, 0.5)$.

Para o MCMC gera-se uma amostra de tamanho 10.000, descarta-se como *burn-in* as 1000 primeiras amostras, considera-se *lag* de 50, resultando em uma amostra *a posteriori* de tamanho 1000. A análise de convergência foi realizada através do teste de convergência de Geweke

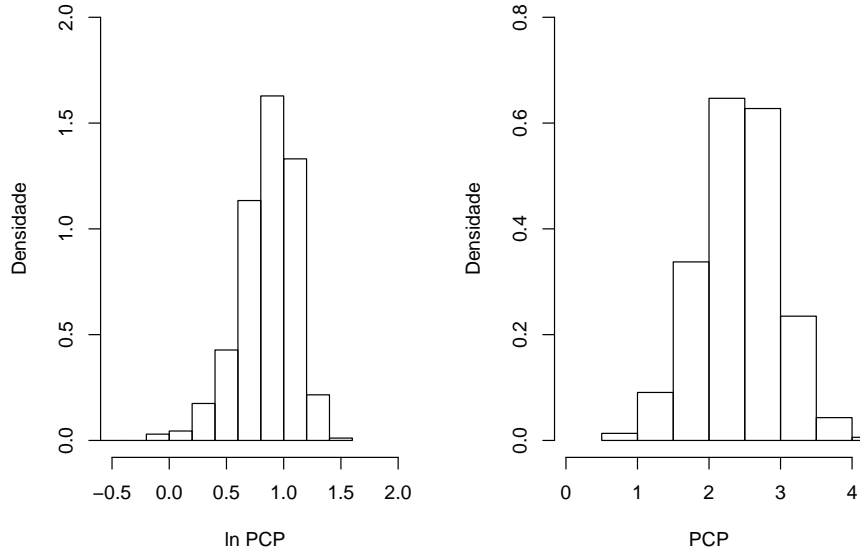


Figura 5.1: Histograma do \ln [PCP] (esquerda) e do PCP (direita) mensal nos Estados Unidos.

(Geweke (1992)), existente dentro do pacote CODA do R, do gráfico do traço e do gráfico da média ergódica das amostras das distribuições *a posteriori*.

A Tabela 5.2 apresenta média (Me), mediana (Md) e desvio padrão (DP) das distribuições marginais *a posteriori* de cada um dos parâmetros e, também, os respectivos intervalos de mais alta densidade (HPD) *a posteriori* com 95% de probabilidade, para ambos os modelos ajustados.

Tabela 5.2: Resumo das distribuições marginais *a posteriori* dos parâmetros nos Modelos 1 e 2

	Me	Md	DP	HPD
Modelo $m = 1$				
μ	1.139	1.140	0.010	[1.118, 1.156]
σ^2	0.375	0.375	0.010	[0.355, 0.395]
Δ	-0.947	-0.947	0.008	[-0.961, -0.928]
Modelo $m = 2$				
μ	1.082	1.083	0.015	[1.060, 1.105]
σ^2	0.291	0.291	0.007	[0.279, 0.304]
Δ_1	-0.481	-0.488	0.029	[-0.500, -0.455]
Δ_2	-0.480	-0.488	0.039	[-0.500, -0.443]

Da Tabela 5.2, observa-se que as estimativas para μ estão próximas nos dois modelos, assumindo valores positivos, como esperado da análise dos histogramas da Figura 5.1. Para o parâmetro de escala σ , as estimativas também se encontram próximas, sendo o Modelo 2 o que apresenta menor valor. Para o parâmetro de assimetria, como esperado, o valor estimado foi negativo e o HPD não inclui o valor zero, evidenciando a presença de assimetria no índice de

precipitação americano pelos dois modelos. No Modelo 1, este parâmetro é facilmente interpretável, pois quanto mais próximo de 1 ou -1 o valor de Δ , mais acentuada é a assimetria. Nota-se que o valor estimado *a posteriori* para este parâmetro é muito próximo de -1, indicando forte assimetria negativa no PCP como indicam os dados observados exibidos na Figura 5.1. Para o Modelo 2, percebe-se que os valores estimados para as entradas do vetor Δ sempre se aproximam do mínimo permitido e são sempre negativas.

Com o objetivo de se comparar os dois modelos, calculam-se a esperança e variância *a posteriori* do PCP para as duas distribuições. A Tabela 5.3 mostra que o índice de precipitação esperado estimado por ambos modelos é bem próxima e as variâncias são pequenas, sendo, em média, maiores sob o Modelo 2. Se comparadas com a média e variância dos dados observados, nota-se que os modelos estão subestimando estas duas medidas.

Tabela 5.3: Resumo *a posteriori* da esperança e variância nos modelos teóricos

	Me	Md	DP
Modelo 1			
E(Y)	1.352	1.352	0.013
Var(Y)	0.015	0.015	0.001
Modelo 2			
E(Y)	1.273	1.273	0.014
Var(Y)	0.035	0.033	0.001

A Figura 5.2 mostra as curvas da fdp estimada usando o método *plug-in* com as médias para os dois modelos sobrepostas ao histograma.

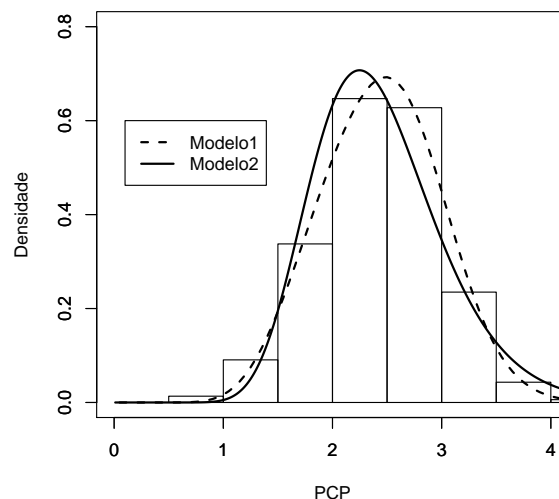


Figura 5.2: Histograma e funções de densidade ajustadas do PCP.

A Tabela 5.4 apresenta os resultados para o teste de Kolmogorov-Smirnov (KS). Este teste é usado para verificar a bondade do ajuste e consiste em calcular a estatística D_n , definida como

o valor absoluto da máxima diferença entre a distribuição empírica e a curva estimada. Neste caso, a hipótese nula é a de que os dados seguem a distribuição log-*skew*-normal fundamental canônica. O p-valor é calculado como em Lin *et al.* (2007).

Tabela 5.4: Teste de Kolmogorov-Smirnov para os modelos ajustados

Modelo	1	2
D_n	0.0227	0.0567
p-valor	0.66	0.01

Observa-se que as duas curvas são muito próximas e se adequam visualmente aos dados, indicando um bom ajuste. Percebe-se que a distribuição fornecida pelo modelo onde $m = 1$ está mais próxima da distribuição empírica, dando evidências de que o Modelo 1 é o mais indicado para descrever o comportamento dos dados. Isto é corroborado pelo teste de Kolmogorov-Smirnov (K-S), cujos resultados estão na Tabela 5.4. Porém, através da Figura 5.2, verificamos que a cauda do Modelo 2 é mais pesada e ele apresenta maior assimetria, o que pode ser interessante para o ajuste de dados com estas características.

Conclusões e Discussões

Neste trabalho, motivados por resultados obtidos em Santos *et al.* (2013), foi definida uma nova família de distribuições log-skew-normal fundamental canônica multivariada, obtida através da transformação da família definida em Arellano-Valle e Genton (2005). Esta é uma distribuição assimétrica com núcleo normal que coloca massa não nula apenas em valores positivos e generaliza a família definida em Marchenko e Genton (2010) no caso de normalidade. Também foram estudadas propriedades e características desta família.

Estudou-se a entropia para cada uma das distribuições consideradas ao longo deste trabalho e encontrou-se relações entre elas. Relacionaram-se todas as entropias estudadas com a entropia da distribuição normal. A principal contribuição nesta direção foi o estudo da entropia nas famílias CFUSN, LSN multivariada e LCFUSN aqui definida. Para obter um método de comparação entre as distribuições, obteve-se a entropia relativa entre as famílias de distribuições log-skew-normal fundamental canônica multivariada e log-skew-normal.

Em relação à inferência, utilizou-se uma abordagem bayesiana e encontrou-se as distribuições condicionais completas para os parâmetros de interesse. Para verificar a aplicabilidade da família de distribuições proposta, utilizou-se um banco de dados reais e ajustou-se dois modelos univariados variando a dimensão m da função responsável por assimetrizar a distribuição normal. No Modelo 1, quando foi usado $m = 1$, retornou-se à família definida por Marchenko e Genton (2010), e observou-se que esta distribuição se ajustou bem ao banco de dados. Para $m = 2$ não se verificou um bom ajuste pelo teste de Kolmogorov-Smirnov. Porém, ao construir a curva teórica, conclui-se que a segunda possui cauda mais pesada e parece ser mais assimétrica do que o primeiro modelo, podendo ser mais eficiente para o ajuste de dados com estas características.

Para trabalhos futuros, sugere-se a implementação de modelos utilizando as distribuições condicionais completas apresentadas no Capítulo 5. Desta forma, não haverá problemas na estimação do parâmetro Δ , como os que foram encontrados aqui, em decorrência da suposição de que Δ é uma matriz $n \times m$ tal que $\|\Delta \mathbf{a}\| < 1$, para todo vetor unitário $\mathbf{a} \in \mathbb{R}^m$. Pode-se utilizar a parametrização $\Delta = \Lambda(\mathbf{I}_m + \Lambda' \Lambda)^{-1/2}$ e obter uma amostra *a posteriori* para o parâmetro Λ , que é uma matriz com entradas finitas reais e não apresenta nenhum tipo de restrição em seu domínio. Pretendemos ainda estudar o comportamento teórico e comparar as distribuições log-skew-normal fundamental canônica, quando mudamos a dimensão m , em outros aspectos, tais como assimetria, entropia e peso das caudas. Acredita-se que em alguns casos o aumento de m pode trazer melhorias nos ajustes.

Além disto, esta dissertação deu origem a dois artigos que ainda estão em construção: *Entropy and Mutual Information in Asymmetric Distributions and its Applications* e *The Log Canonical Fundamental Skew Normal Distribution and its Applications*.

Referências Bibliográficas

- [1] Ahmed, N. A. e Gokhale, D. V. (1989). *Entropy expressions and their estimators for multivariate distributions*. IEEE Trans. Inform. Theory **35**, 688–692
- [2] Arellano-Valle, R. B., Contreras-Reyes, J. E. e Genton, M. G. (2011). *Shannon Entropy and Mutual Information for Multivariate Skew-Elliptical Distributions*. Scandinavian Journal of Statistics **40**, 42-62.
- [3] Arellano-Valle, R. B. e Genton, M. G. (2005). *On fundamental skew distributions*. Journal of Multivariate Analysis, **96** (1), 93-116.
- [4] Arellano-Valle, R. B. and Genton, M. G. (2010). *An invariance property of quadratic forms in random vectors with a selection distribution, with application to sample variogram and covariogram estimators*. Ann. Inst. Statist. Math., 363-381.
- [5] Arellano-Valle, R. B., M. D. Branco, and M. G. Genton (2006). *A unified view of skewed distributions arising from selections*. The Canadian Journal of Statistics **34**, 581-601.
- [6] Arellano-Valle, R.B. and Azzalini, A. (2006). *On unication of skew normal families*. Scandinavian Journal of Statistics **33**, 561-574.
- [7] Azzalini, A. (1985). *A class of distributions which includes the normal ones*. Scandinavian Journal of Statistics, **12**, 171-178.
- [8] Azzalini, A. e Dalla Valle, A. (1996). *The multivariate skew-normal distribution*. Biometrika, **83**, 715-726.
- [9] A. Azzalini, T. Dal Cappello and S. Kotz (2003). *Log-skew-normal and log-skew-t distributions as models for family income data*. J. Income Distribution, **11**, no. 3-4, 12-20.
- [10] Branco, M.D. and Dey, D.K. (2001). *A class of multivariate skew-elliptical distributions*. Journal of Multivariate Analysis, **79**, 99-113.
- [11] Cover, T. M. e Thomas, J. A. (2006). *Elements of information theory*. Second edn. Wiley, New Jersey.
- [12] Elal-Olivero, D., Gomez, H. W. and Quintana, F. A. (2009). *Bayesian modeling using a class of bimodal skew-elliptical distributions*. Journal of Statistical Planning and Inference, **139** (4), 1484-1492

- [13] Escobar, M. D. and West, Mike (1998). *Computing Nonparametric Hierarchical Models*. Practical nonparametric and semiparametric Bayesian statistics, **1**, Springer - Verlag (Berlin, New York).
- [14] Fang, K.T., Kotz, S and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.
- [15] Genton MG. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Chapman e Hall/CRC: Boca Raton.
- [16] M.G. Genton, N. Loperfido (2005). *Generalized skew-elliptical distributions and their quadratic forms*. Ann. Inst. Statist. Math.
- [17] Geweke, J. (1992). *Bayesian Statistics*, Clarendon Press, Oxford, UK, chapter Evaluating the accuracy of sampling-based approaches to calculating posterior moments.
- [18] G. González-Farías, J.A. Domínguez-Molina, A.K. Gupta (2004) *Additive properties of skew normal random vectors*. J. Statist. Plann. Inference **126**, 521–534
- [19] Javier, W. R. e Gupta, A. K. (2008). *Mutual information for the mixture of two multivariate normal distributions*. Far East J. Theor. Stat. **26**, 47–58.
- [20] Javier, W. R. e Gupta, A. K. (2009). *Mutual information for certain multivariate distributions*. Far East J. Theor. Stat. **29**, 39–51
- [21] Kullback, S. (1978). *Information theory and statistics*. Dover Edition, Gloucester.
- [22] Larsen, K., Petersen, J. H., Budtz-Jorgensen, E. e Endahl, L. (2000). *Interpreting parameters in the logistic regression model with random effects*, Biometrics **56**, 909-914.
- [23] Lin, TI and Lee, JC and Hsieh WJ (2007). *Robust mixture modeling using the skew t distribution*, *Statistics and Computing* **17**, 81-92.
- [24] Marchenko, Y. V. e Genton, M. G. (2010). *Multivariate log-skew-elliptical distributions with applications to precipitation data*. Environmetrics **21**, 318-340.
- [25] Queiroz, M. M., Silva, R. W. C. e Loschi, R. H. (2013) *Entropy and Mutual Information in Asymmetric Distributions and its Applications*. Não publicado.
- [26] Queiroz, M. M., Silva, R. W. C. e Loschi, R. H. (2013) *The Log Canonical Fundamental Skew Normal Distribution and its Applications*. Não publicado.
- [27] Shannon, C. E. (1948). *A mathematical theory of communication*. Bell Syst. Tech. J. **27**, 379–423, 623–656.
- [28] Santos, C. C., Loschi, R. H. e Arellano-Valle, R. B. *Parameter Interpretation in Skewed Logistic Regression with Random Intercept*. Bayesian Anal., V8, Number 2, 381-410.
- [29] van Dyk and X. L. Meng. (2001) *The art of data augmentation*. J. Comput. Graph. Statist. **10**, 59–68.

- [30] West, M. (1987). *On Scale Mixtures of Normal Distributions*. *Biometrika* **74**, 646.
- [31] Zacks, Shelemyahu (1981). *Parametric statistical inference: Basic theory and modern approaches*. Pergamon Press.