Universidade Estadual da Paraíba
Centro de Ciências e Tecnologia
Departamento de Estatística

Leomir Ferreira Sousa

ALGUMAS RELAÇÕES ENTRE DISTRIBUIÇÕES UNIVARIADAS

Campina Grande,
28 de Novembro 2014.

Leomir Ferreira Sousa

ALGUMAS RELAÇÕES ENTRE DISTRIBUIÇÕES UNIVARIADAS

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador:
Sílvio Fernando Alves Xavier Júnior

Campina Grande,
28 de Novembro 2014.

S725a Sousa, Leomir Ferreira.
Algumas relações entre distribuições univariadas [manuscrito] / Leomir Ferreira Sousa. - 2014.
39 p. : il.

Digitado.
Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2014.
"Orientação: Prof. Me. S lvio Fernando Alves Xavier J unior, Departamento de Estatística".

1. Distribui ções univariadas. 2. Probabilidade. 3. Fun ção caracter ística. I. Título.

21. ed. CDD 519.5

Leomir Ferreira Sousa

ALGUMAS RELAÇÕES ENTRE DISTRIBUIÇÕES UNIVARIADAS

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 28 / 11 / 2014

Banca Examinadora:

Prof. Me. Sílvio Fernando Alves Xavier Júnior
Universidade Estadual da Paraíba - DE/CCT
Orientador

Prof. Me. Juarez Fernandez de Oliveira
Universidade Estadual da Paraíba - DE/CCT
Examinador

Tiago Almeida de Oliveira

Prof. Dr. Tiago Almeida de Oliveira
Universidade Estadual da Paraíba - DE/CCT
Examinador

Dedicatória

Dedico este trabalho primeiramente a Deus, a minha mãe Maria do Carmo, heroína que me deu apoio e incentivo nas horas difíceis tanto de desânimo quanto de cansaço. Ao meu pai onde sua presença significou segurança e certeza de que não estou sozinho nessa caminhada. A meu avô Tirbutino, que mesmo lá de cima, sempre esteve presente comigo.

Agradecimentos

Quero agradecer em primeiro lugar a Deus pela força e coragem durante toda esta longa caminhada. A minha família, em especial a minha vó Lourdes, minhas tias, tios e primos, a minha namorada pela força que me deu nos momentos difíceis que passei.

Ao meu orientador pela dedicação na elaboração deste trabalho. Aos chefes do Departamento, coordenadores, professores que me proporcionaram o conhecimento não apenas racional, mas a manifestação do trabalho e afetividade da educação no processo de formação do profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, meus sinceros agradecimentos.

Resumo

As distribuições univariadas comuns são geralmente discutidas separadamente em livros-textos de probabilidade em nível intermediário, proporcionando dessa forma, dificuldades no entendimento das relações existentes entre essas distribuições. Neste contexto, pretende-se com o presente trabalho, estudar as técnicas de verificação das relações de distribuições de probabilidade: a técnica de função de distribuição acumulada, função geradora de momentos, função característica. Além disso, utilizou-se *software R* para solucionar problemas práticos e como instrumento didático na ilustração da teoria estatística.

Palavras-chave: Distribuições Univariadas, Probabilidade, Função Característica.

Abstract

The common univariate distributions are generally discussed separately in probability of textbooks in intermediate level, thereby providing, difficulties in understanding the relationship between these distributions. In this context, it is intended with this project, studying the verification techniques of relations of probability distributions: the cumulative distribution function technique, moment generating function, characteristic function. Also, use the it R software to solve practical problems and as an educational tool in the illustration of statistical theory.

Sumário

Lista de Figuras

Lista de Tabelas

1 Introdução

A produção científica em estatística está completamente relacionada com o conhecimento do pesquisador em várias disciplinas, como por exemplo, o cálculo diferencial e integral, a álgebra linear, a teoria das probabilidades, a inferência, etc. Neste contexto, um dos elementos de estudos de grande importância, o qual está presente em quase tudo que se faz para se produzir algo novo em estatística é o estudo das relações entre as distribuições univariadas mais comuns. Do ponto de vista do *cálculo diferencial e integral*, associado a *teoria das probabilidades*, estabelecer tais relações de uma ou mais de uma distribuição contínua ou discreta, com seus respectivos domínios e contradomínios, em alguns casos, não é uma tarefa trivial e requer do pesquisador certa habilidade em *Estatística Matemática*.

Livros de estatística e de introdução à probabilidade normalmente inserem distribuições univariadas comuns de maneira individual, e raramente relatam todas as relações entre as distribuições. Elucidaremos uma atualização de um escopo apresentado por (LEEMIS, 1986) que mostra as principais propriedades e relações entre as distribuições univariadas mais comumente utilizadas. Mais detalhes sobre estas distribuições, tratamentos mais concisos podem ser vistos em (BALAKRISHNAN; NEVZOROV, 2004), (EVANS *et al.*, 2000).

Em geral, a literatura especializada no tema, apresenta poucas ou quase nenhuma técnica de obtenção das relações e propriedades das distribuições univariadas. A obtenção de tais relações foi feita utilizando-se das abordagens sugeridas em (LEEMIS, 1986), (BALAKRISHNAN; NEVZOROV, 2004) e, em seguida, fazer uma análise para indicar qual delas é mais conveniente em cada caso específico.

O domínio de um software estatístico é de suma importância para viabilizar as soluções de problemas práticos visando dar um suporte computacional ao tema. Nesta direção, associam-se soluções de problemas teóricos numa visão prática, utilizando-se o *software* R. Assim sendo, é preciso buscar informações em textos especializados, tais como: (PETER-

NELLI; MELLO, 2007), (FOX, 2002), (EVERITT; HOTHORN, 2006), (MAINDONALD; BRAUN, 2007).

De início, iniciamos com a distribuição exponencial onde vimos sua f.d.p de duas formas, ou seja, com reparametrização, logo depois vimos a sua f.d.a, e seus respectivos momentos e suas propriedades. Após verificamos a distribuição gama, onde descrevemos a função de densidade, suas propriedades e suas relações entre exponencial, weibull e rayleigh.

Portanto, a proposta deste trabalho foi utilizar as relações univariadas pertencentes a família Gama: Exponencial, Weibull e Rayleigh. Não obstante, a função geradora de momentos, a função característica e suas principais propriedades foram elucidadas.

2 Fundamentação Teórica

Para o desenvolvimento deste trabalho, a metodologia utilizada foi baseada na ideia proposta por (LEEMIS; MCQUESTON, 2008).

Os mesmos elucidam muitas propriedades que são aplicadas as distribuições individuais listadas na Figura 1.

- A propriedade da combinação Linear(L) indica que as combinações lineares de variáveis aleatórias independentes possuindo uma distribuição particular são oriundas da mesma família de distribuição.

  Exemplo: Se $X_i \sim N\left(\mu_i, \sigma_i^2\right), i = 1, 2, ..., n; a_1, a_2, ..., a_n$ são constantes reais, e $X_1, X_2, ..., X_n$ são independentes, então

  $$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

- A propriedade da Convolução(C) indica que a soma de va's possuindo distribuição particular origina-se da mesma família de distribuição

  Exemplo: Se $X_i \sim \chi^2(n_i), i = 1, 2, ..., n; X_1, ..., X_n$ são independentes, então

  $$\sum_{i=1}^{n} X_i \sim \chi^2\left(\sum_{i=1}^{n} n_i\right)$$

- A propriedade do Produto(p) indica que va's independentes possuindo esta distribuição particular, origina-se da mesma família de distribuição

  Exemplo: Se $X_i \sim lognormal(\mu_i, \sigma_i^2), i = 1, 2, ..., n$ e $X_1, X_2, ..., X_n$ são independentes, então

  $$\prod_{i=1}^{n} X_i \sim lognormal\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

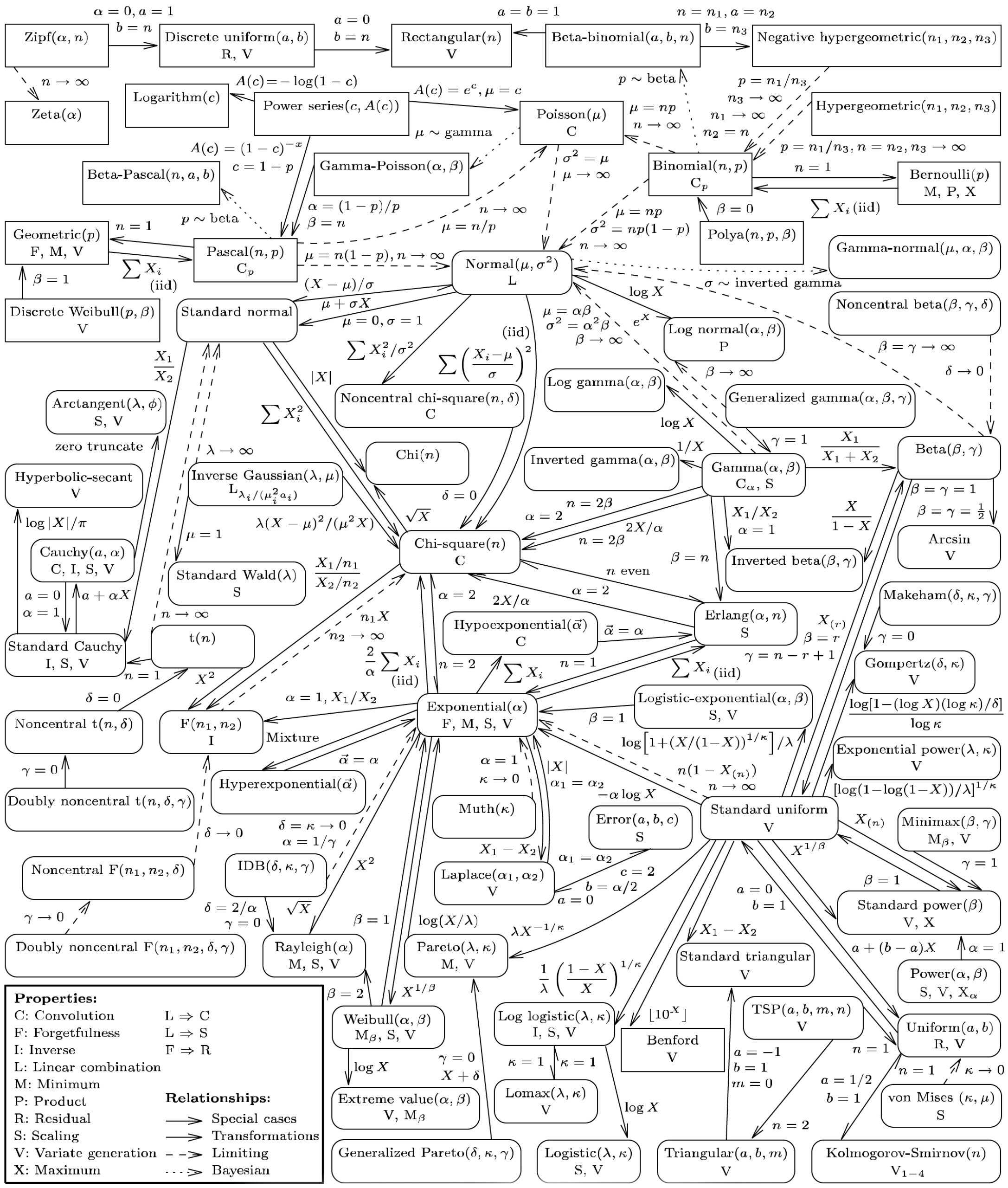

Figura 1: Relações entre as distribuições univariadas de acordo com (LEEMIS; MCQUESTON, 2008)

As propriedades da *inversão*, *mínimo*, *máximo* e *resíduos* também estão descritas no artigo.

A Figura 1 contém 76 distribuições de probabilidades univariadas. Há 19 modelos discretos e 57 modelos contínuos. As distribuições discretas são apresentadas em caixas retangulares; distribuições contínuas são apresentadas em caixas arredondadas. As distribuições discretas estão na parte superior da Figura 1, com a exceção da distribuição Benford. Uma distribuição é descrita por duas linhas de texto em cada caixa. A primeira linha dá o nome da distribuição e seus parâmetros. A segunda linha contém as propriedades que a distribuição assume.

Existem três tipos de linhas(procedimentos) usadas para conectar as distribuições entre si. A linha **sólida** é utilizada em *casos especiais* e *transformações.* Transformações tipicamente possuem um X nos rótulos para diferenciá-las dos casos especiais. A linha **tracejada** é usada para relações assintóticas, as quais estão tipicamente no limite de um ou mais parâmetros aproximarem-se do limite do espaço paramétrico. A linha **pontilhada** é usada para *relações Bayesianas* (Beta – binomial, Beta–Pascal, Gama–normal, e Gama–Poisson). Há certos casos especiais, em que as distribuições se sobrepõem para apenas uma única configuração de seus parâmetros. Exemplos incluem (a) a distribuição exponencial com média dois e a distribuição qui–quadrado com dois graus de liberdade, (b) a distribuição qui–quadrado com um número par de graus de liberdade e a distribuição Erlang com dois parâmetro de escala, e (c) a de distribuição Kolmogorov-Smirnov (caso todos os parâmetros sejam conhecidos) para uma amostra de tamanho $n = 1$ e de distribuição $U \sim (1/2, 1)$. Cada um destes casos é indicado por uma seta de duas pontas.

Várias dessas relações sugerem outras distribuições que ainda não foram desenvolvidas. Primeiramente, o valor extremo e a distribuição log–gamma indicam que o logaritmo de quaisquer resultados da distribuição de sobrevivência em uma distribuição com o suporte ao longo de todo eixo real. Em segundo lugar, a distribuição gamma invertida indica que o recíproco de quaisquer resultados da distribuição de sobrevivência em outra distribuição de sobrevivência. Em terceiro lugar, os papéis de mudança $F(x)$ e $F^{-1}(u)$ para uma variável aleatória com suporte sobre (0, 1) resulta em uma distribuição complementar (por exemplo, (JONES, 2002)).

Ficar claro para o aluno que existem muitas outras distribuições, consideradas elementares, que mantém relacionamento entre si, que são muito importante na **Estatística**.

2.1 Teorema Central do Limite

Vamos apresentar apenas uma pequena ideia sobre o Teorema Central do Limite, visto que algumas propriedades decorrem deste teorema, considerado um dos mais importantes da Estatística. (MAGALHÃES, 2006) em essência, trata-se da convergência em distribuição para o modelo Normal de uma soma de variáveis aleatórias independentes, após uma conveniente padronização.

2.1.1 Teorema Central do Limite para variáveis i.i.d.

Sejam $X_n : n \geq 1$ variávies aleatórias independentes, identicamente distribuídas e com esperança μ e variância σ^2, com $0 < \sigma^2 <$. Então, para $S_n = X_1 + X_2 + ... + X_n.$, temos,

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \rightarrow N(0,1).$$

3 Demonstração

Nesta seção demonstraremos como se dá a relação entre algumas distribuições univariadas.

3.1 Definição

A distribuição exponencial desempenha importante papel na distribuição de uma grande classe de fenômenos, particulamente nos assuntos da teoria da confiabilidade (MEYER, 1983). Uma variável aleatória contínua X, que tome todos os valores não-negativos, terá uma distribuição exponencial com parâmetro α, se sua f.d.p for dada por

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, \text{ caso contrário} \end{cases}$$

Por outro lado, a f.d.p pode ser escrita ainda da seguinte forma $\alpha > 0$ e $\alpha \in \mathbb{R}$

$$f(x) = \begin{cases} \frac{1}{\alpha} e^{\frac{-x}{\alpha}}, & x > 0 \\ 0, \text{ caso contrário.} \end{cases}$$

3.2 Função de distribuição Acumulada

A função de distribuição acumulada (FDA) de uma variável aleatória contínua X é definida por

$$F(X) = \int_{-\infty}^{x} f(x)dx. \tag{3.1}$$

Desta forma, a FDA da distribuição exponencial é definida por

$$F(X) = P(X < x) = \begin{cases} \int_0^x \alpha e^{-\alpha t} dt = 1 - e^{-\alpha x} \\ 0, \text{ caso contrário} \end{cases}$$

Portanto, $P(X > x) = e^{-\alpha x}$

3.2.1 Valor Esperado

O valor esperado ou esperança de uma variável aleatória X é por definição a média de sua distribuição de probabilidades. Isto é, se X for uma variável aleatória *contínua*, então

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx, \qquad (3.2)$$

onde $f(X)$ é a f.d.p da variável aleatória X.
Como $X \sim exp(\alpha)$, temos que

$$E(X) = \int_0^{\infty} x \alpha e^{-\alpha x} dx. \qquad (3.3)$$

O resultado da expressão acima é obtido pelo método de integração por partes.

$u = x \Rightarrow du = dx$;

logo, temos que $dv = \alpha e^{-\alpha x} dx$;

segue que $v = \int_0^{\infty} \frac{-\alpha}{\alpha} e^{-\alpha x} = -e^{-\alpha x}$.

Em que,

$$E(X) = -xe^{-\alpha x} + \int_0^{\infty} e^{-\alpha x} dx$$
$$E(X) = \left[-xe^{-\alpha x} - \frac{e^{-\alpha x}}{\alpha} \right]_0^{\infty} = \frac{1}{\alpha}.$$

Por conseguinte, a esperança é igual ao inverso do parâmetro α. Pelo simples "rebatismo" do parâmetro $\alpha = \frac{1}{\beta}$ poderíamos escrever a f.d.p de X como sendo $f(x) = \frac{1}{\beta} e^{\frac{-x}{\beta}}$. Nesta forma o parâmetro β fica igual ao valor da esperança de X.

3.2.2 Variância

A variância de uma variável aleatória é por definição a variância de sua distribuição de probabilidades, isto é, a variância de X é dada por

$$Var(X) = E\{[X - E(X)]^2\}. \tag{3.4}$$

Se X for uma variável contínua, temos que

$$Var(X) = \int_{-\infty}^{\infty} [X - E(X)]^2 f(x)dx,$$

em que $f(X)$ é a f.d.p da variável aleatória X. Alternativamente, podemos calcular a variância da seguinte forma

$$Var(X) = E(X^2) - [E(X)]^2. \tag{3.5}$$

Calculando $E(X^2)$, ou seja, $E(X^2) = \int x^2 f(x)dx$. A variância de X pode ser obtida por uma integração semelhante.

$$E(X^2) = \int_0^{\infty} x^2 \alpha e^{-\alpha x}$$

Fazendo $u = x^2 \Rightarrow du = 2xdx$, então $dv = \alpha e^{-\alpha x}dx \Rightarrow v = \int \alpha e^{-\alpha e}dx$;

logo, $v = \frac{-\alpha e^{-\alpha x}}{\alpha} \Rightarrow v = -e^{-\alpha x}$.Então, $E(X^2) = -x^2 e^{-\alpha x} + \int e^{-\alpha x} 2xdx$.

Fazendo, $u = x \Rightarrow du = dx$ e $dv = e^{-\alpha x}dx$, logo, $v = -e^{-\alpha x}$.

Portanto,

$$\begin{aligned}
E(X^2) &= -x^2 e^{-\alpha x} + 2\int_0^{\infty} e^{-\alpha x}dx \\
E(X^2) &= \left(\frac{-xe^{-\alpha x}}{\alpha} - 2\int \frac{e^{-\alpha x}}{\alpha}dx\right)_0^{\infty} = \frac{2}{\alpha^2} \\
Var(X) &= \frac{2}{\alpha^2} - \frac{1}{\alpha^2} = \frac{1}{\alpha^2}
\end{aligned}$$

3.2.3 Falta de Memória

Uma característica interessante da distribuição exponencial é a falta de memória. Segundo Magalhães (2006), dizemos que uma variável aleatória não tem memória se

$P(X > s + t)|X > t) = P(X > s)$, para todo $s, t \geq 0$.

Se pensarmos em X como sendo o tempo de vida de algum instrumento, então a probabilidade de que o equipamento funcione por pelo menos $s + t$ horas dado que já funcionou por pelo menos t horas é a mesma probabilidade inicial que ele funcione por s horas. Em outras palavras, se o instrumento está funcionando no tempo t, então a distribuição da quantidade de tempo que ele ainda irá funcionar é a mesma distribuição do tempo de vida inicial. Isto é, o instrumento "não lembra" que ele já foi usado por um tempo t.

Daí, $P(X > t, X > s + t) = P(X > s)\ P(X > t)$

o que implica que,

$P(X > s + t) = P(X > t)P(X > s)$

No caso da distribuição exponencial, vimos que

$P(X > s + t) = e^{-\alpha(t+s)} = e^{-\alpha s - \alpha t} = e^{-\alpha t} e^{-\alpha s} = P(X > t)P(X > s)$

e portanto , a distribuição exponencial não tem memória. Na verdade, a distribuição poisson tem tal propriedade equivalente.

3.3 Função Geradora de Momentos

A função geradora de momentos da variável aleatória X, (DANTAS, 2008),

$$M(t) = E(e^{tx}) \tag{3.6}$$

$\forall t, -\infty < t < \infty$; em que a esperança seja finita. No caso da função exponencial, a função geradora é dada por

$$\begin{aligned}
M(t) = E(e^{tx}) &= \int_0^\infty e^{tx}\alpha e^{-\alpha x}dx \\
M(t) = E(e^{tx}) &= \alpha\int_0^\infty e^{-(\alpha-t)x}dx \\
&= -\alpha\left(\frac{e^{(-\alpha-t)x}}{\alpha-t}\right)_0^\infty \\
M(t) = E(e^{tx}) &= \frac{\alpha}{\alpha-t}
\end{aligned}$$

3.4 Função Característica

É sempre válido ter maneiras alternativas de representar o mesmo objeto matemático. E a função característica é uma delas, algumas vantagens do uso da função característica são:

- Pode-se calcular os momentos de uma variável aleatória X diferenciando-se a função característica (o que geralmente é mais simples que usar diretamente as definições de momento que envolvem integrais);
- Pode-se calcular mais facilmente a distribuição de soma de variáveis aleatórias independentes;
- e finalmente, o uso de funções características ajuda na prova de uma família de Teoremas Centrais do Limite que ajudam a explicar a prevalência da distribuição normal ou Gaussiana na natureza.

3.5 Definição

A função característica Y de uma Variável aleatória X é dada por

$$\varphi_X(t) = Ee^{itx} = E(\cos tX) + E(sentX) \tag{3.7}$$

onde, $i = \sqrt{-1}$. Note que como $cos(tX)$ e $sen(tX)$ são variável aleatórias limitadas, a esperança na definição acima é finita e, consequentemente, a função caracteristica de qualquer variável é bem definida. Note que também de acordo com esta definição, a função de distribuição acumulada determina a função característica de uma variável aleatória (RêGO,).

No caso particular de uma variável aleatória discreta, temos

$$\varphi_X(t) = \sum_k e^{itx_k} p(x_k)$$

onde $p(x_k)$ é a função probabilidade de X.

Analogamente, se X for uma variável aleatória contínua, temos

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itX} f(x) dx$$

onde $f(X)$ é a função densidade de probabilidade de X.

Suponhamos que $X \sim Exp(\alpha)$. Então

$$\begin{aligned}
\varphi_X(t) &= \int_0^{\infty} e^{itx} \alpha e^{-\alpha x} dx \\
\varphi_X(t) &= \alpha \int_0^{\infty} e^{x(-\alpha+it)} dx \\
&= \left[\frac{\alpha}{-\alpha + it} e^{x(-\alpha+it)} \right]_0^{\infty} \\
\varphi_X(t) &= \frac{\alpha}{\alpha - it}.
\end{aligned}$$

Derivando a função característica da distribuição exponencial em relação a t, encontramos a $E(X)$, ou seja:

$$\begin{aligned}
\varphi_X(t)' &= \left(\frac{\alpha}{\alpha - it}\right)' \\
\varphi_X(t)' &= \frac{\alpha'(\alpha - it) - \alpha(\alpha - it)'}{(\alpha - it)^2} \\
&= \frac{\alpha i}{(\alpha - it)^2} \\
&\vdots \\
&= \frac{\varphi_X(0)'}{i} \\
&= \frac{\alpha i}{\alpha^2} \\
&\vdots \\
E(X) &= \frac{1}{\alpha},
\end{aligned}$$

com $t = 0$ e com $i = 1$.

A partir da primeira derivada, encontramos a $E(X^2) = \frac{\alpha i}{(\alpha - it)^2}$

$$\begin{aligned}
\varphi_X(t)'' &= \frac{\alpha i'(\alpha - it)^2 - \alpha i(\alpha - it)'^2}{(\alpha - it)^4} \\
&= \frac{2\alpha i^2}{(\alpha - it)^3} \\
&= \frac{\varphi_X(0)''}{i^2} \\
&= \frac{-2\alpha i^2}{\alpha^3} \\
E(X^2) &= \frac{2}{\alpha^2},
\end{aligned}$$

com $i^2 = -1$ e $t = 0$.

Agora vamos calcular $E(X^3)$

$$
\begin{aligned}
\varphi_X(t)'' &= \frac{2\alpha i^2}{(\alpha - it)^3} \\
\varphi_X(t)''' &= \frac{(2\alpha i^2)'(\alpha - it)^3 - 2\alpha i^2(\alpha - it)'^3}{(\alpha - it)^6} \\
&\vdots \\
\varphi_X(t)''' &= \frac{6\alpha i^3}{(\alpha - it)^4} \\
&\vdots \\
E(X^3) &= \frac{\varphi_X(0)'''}{i^3} \\
&\vdots \\
E(X^3) &= \frac{6i^3}{\alpha^3}.
\end{aligned}
$$

Seja $f_X(x) = \frac{1}{\alpha} e^{\frac{-x}{\alpha}}$. A função característica é dada por

$$
\begin{aligned}
\varphi_X(t) &= E(e^{itX}) = \int_0^\infty e^{itx} \frac{1}{\alpha} e^{\frac{-x}{\alpha}} dx \\
&= \frac{1}{\alpha} \int_0^\infty e^{x(it - \frac{1}{\alpha})} dx. \\
&\vdots \\
&= \left[\frac{e^{x(it - \frac{1}{\alpha})}}{(\alpha it - 1)} \right]_0^\infty \\
\varphi_X(t) &= E(e^{itX}) = \frac{1}{-\alpha it + 1}
\end{aligned}
$$

Com isso calculamos a $E(X)$,$E(X^2)$ e $E(X^3)$.

$$\begin{aligned}
\varphi_X(t) &= \frac{1}{-\alpha it+1} \\
\varphi_X(t)' &= \frac{1'(-\alpha it+1)-1(-\alpha it+1)'}{(-\alpha it+1)^2} \\
&= \frac{\alpha i}{(-\alpha it+1)^2} \\
&\vdots \\
E(X) &= \frac{\varphi_X(0)'}{i} \\
E(X) &= \frac{\alpha i}{1}
\end{aligned}$$

com i=1 e t=0.

Calculando $E(X^2)$

$$\begin{aligned}
\varphi_X(t)^{'} &= \frac{\alpha i}{(-\alpha it+1)^2} \\
\varphi_X(t)^{''} &= \frac{(-\alpha i)'(-\alpha it+1)^2-(-\alpha i)(-\alpha it+1)'^2}{(-\alpha it+1)^4} \\
&\vdots \\
&= \frac{-2\alpha^2 i^2}{(-\alpha it+1)^3} \\
&\vdots \\
E(X^2) &= \frac{\varphi_X(0)^{''}}{i^2} \\
E(X^2) &= -2(\alpha^2 i^2) \\
E(X^2) &= 2\alpha^2
\end{aligned}$$

com $i^2=-1$ e $t=0$. Concluindo, calculamos a $E(X^3)$,

$$
\begin{aligned}
\varphi_X(t)'' &= \frac{-2\alpha^2 i^2}{(-\alpha i t + 1)^3} \\
\varphi_X(t)''' &= \frac{(-2\alpha^2 i^2)'(-\alpha i t + 1)^3 - (-2\alpha^2 i^2)(-\alpha i t + 1)'^3}{(-\alpha i t + 1)^6} \\
&\vdots \\
&= \frac{6\alpha^3 i^3}{(-\alpha i t + 1)^4} \\
&\vdots \\
E(X^3) &= \frac{\varphi_X(0)'''}{i^3} \\
E(X^3) &= 6\alpha^3 i^3.
\end{aligned}
$$

com $i^3 = 1$ e $t = 0$.

Na figura 2, temos a representação gráfica da distribuição exponencial com $\alpha = 1$

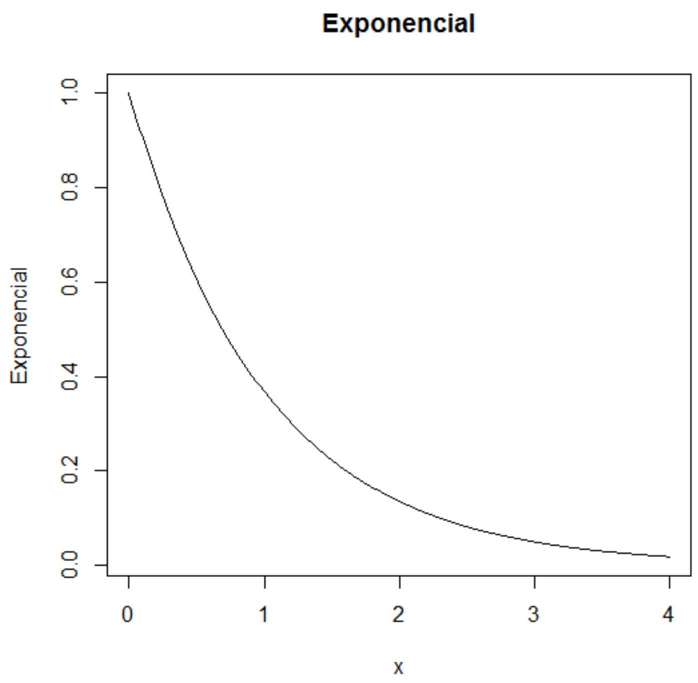

Figura 2: Gráfico da Função Exponencial com o parâmetro $\alpha = 1$.

3.6 Distribuição Gama

Diz-se que uma variável aleatória tem distribuição gama com parâmentros $(\alpha, \lambda), \lambda > 0, \alpha > 0$, se sua função densidade é dada por

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

onde $\Gamma(\alpha)$, chamada de função gama, é definida como

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy$$

A integração de $\Gamma(\alpha)$ por partes resulta em

$$\begin{aligned} \Gamma(\alpha) &= -e^{-y} y^{\alpha-1}|_0^\infty + \int_0^\infty e^{-y}(\alpha-1) y^{\alpha-2} dy \\ &= (\alpha-1)\int_0^\infty e^{-y} y^{\alpha-2} dy \\ &= (\alpha-1)\Gamma(\alpha-1) \end{aligned} \tag{3.8}$$

Para valores inteiros de α, digamos $\alpha = n$, obtemos, aplicando a Equação (3.8) repetidamente,

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = (n-1)(n-2)\ldots 3 \cdot 2\Gamma(1)$$

3.7 Propriedades da função Gama

Como $\Gamma(1) = \int_0^\infty e^{-x} dx = 1$, tem-se que, para valores inteiros de n,

$$\Gamma(n) = (n-1)!$$

A função gama tem a seguinte propriedade recursiva: $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Para Demonstrar esse resultado, iremos usar integração por partes

$$
\begin{aligned}
\Gamma(\alpha+1) &= \int_0^{\infty} e^{-x}x^{\alpha}dx \\
u &= x \\
du &= dx \\
dv = \int_0^{\infty} e^{-x} &\Rightarrow v = -e^{-x} \\
\Gamma(\alpha+1) &= [-x^{\alpha}e^{-x}]_0^{\infty} + \int_0^{\infty} e^{-x}\alpha x^{\alpha+1}dx \\
&= [-x^{\alpha}e^{-x}]_0^{\infty} + \alpha\int_0^{\infty} e^{-x}x^{\alpha+1}dx \\
&= 0 + \alpha\int_0^{\infty} e^{-x}x^{\alpha+1}dx \\
\Gamma(\alpha+1) &= \alpha\Gamma(\alpha)
\end{aligned}
$$

Vamos trabalhar agora, com $\alpha = n$ inteiro.

$$\Gamma(1) = \int_0^{\infty} e^{-x}x^{1-1}dx = \int_0^{\infty} e^{-x}dx = 1$$

$$\Gamma(2) = \int_0^{\infty} e^{-x}x^{2-1}dx = -xe^{-x} + \int_0^{\infty} e^{-x}dx = 0 + 1 = 1!$$

$$\Gamma(3) = \int_0^{\infty} e^{-x}x^{3-1}dx = -x^2e^{-x} + 2\int_0^{\infty} e^{-x}dx = 0 + 2 = 2!$$

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{1-1}}{\Gamma(1)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Logo encontramos a função exponencial se mudarmos λ por α.

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, \text{ para quaisquer outros valores} \end{cases}$$

Tabela 1: Casos especiais da distribuição gama generalizada

$f(x\|a,b,c,d)$	Nome da distribuição
$f(x\|a,b,1,1)$	Distribuição exponencial com dois parâmetros
$f(x\|0,1,1,1)$	Distribuição exponecial reduzida
$f(x\|0,1,c,1)$	Distribuição WEIBULL reduzida
$f(x\|a,b,c,1)$	Distribuição WEIBULL de três parâmetros
$f(x\|a,b,2,1)$	Distribuição Rayleigh

Como $\Gamma(1) = \int_0^\infty e^{-x}dx = 1$, tem-se que, para valores inteiros de n,

$$\Gamma(n) = (n-1)!$$

Fazendo $z = \sqrt{2t} \Rightarrow z^2 = 2t$,tem-se

$$t = \frac{z^2}{2} \Rightarrow t^{\frac{1}{2}} = \frac{z}{\sqrt{2}}$$

Logo,

$$\frac{dz}{dt} = (2t)^{\frac{-1}{2}}$$

Daí,

$$dz = (2t)^{\frac{-1}{2}}dt \Rightarrow dt = (2t)^{\frac{1}{2}}dz$$

Portanto, $dt = zdz$. Com isso,

$$\Gamma(\frac{1}{2}) = \int_0^\infty \frac{\sqrt{2}}{z}e^{\frac{-z^2}{2}}zdz$$
$$\Gamma(\frac{1}{2}) = 2\sqrt{\pi}\int_0^\infty \frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz \Rightarrow \frac{2\sqrt{\pi}}{2} = \sqrt{\pi}$$

Assim,

$$\int_0^\infty \frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz$$

é f.d.p de uma $N(0,1)$.

Assim na tabela 1 tem-se a distribuição gama generalizada com 3 parâmetros

Na figura 3 tem-se a representação da distribuição gama com os parâmetros $\alpha = 10$ e

$\lambda = 3$ para diferentes valores de X.

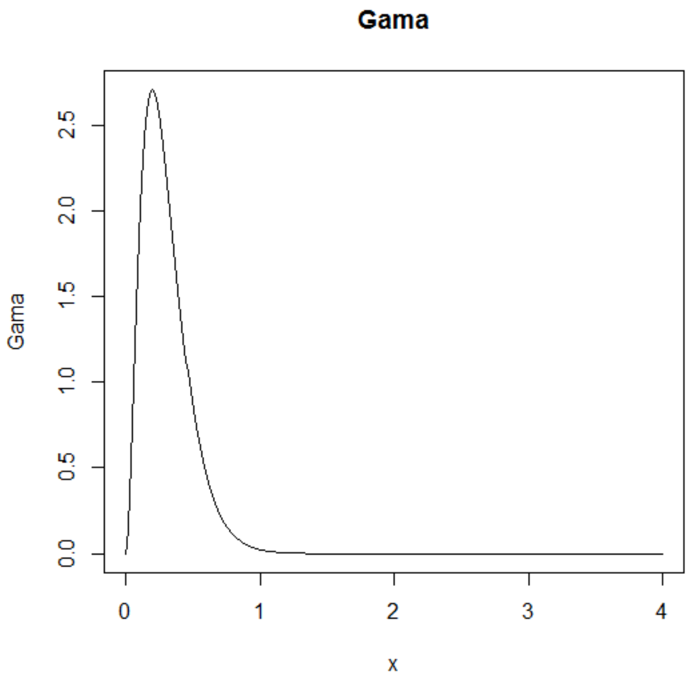

Figura 3: Gráfico da Função Gama com os parâmetros $\alpha = 10$ e $\lambda = 3$.

3.8 Função Característica

Para encontrarmos os momentos da função gama pode-se utilizar a função característica,

$$\varphi_X(t) = E(e^{itX}) = \int_0^\infty \frac{e^{itX}\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} dx$$

$$\varphi_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-x(\lambda - it)} x^{\alpha-1} dx$$

Fazendo $x = \frac{1}{(\lambda - it)} y$ e $dx = \frac{1}{(\lambda - it)} dy$ tem-se

$$\begin{aligned}
\varphi_X(t) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{y}{(\lambda - it)} \right)^{\alpha-1} e^{\frac{-y}{\lambda - it}(\lambda - it)} \frac{dy}{(\lambda - it)} \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \frac{y^{\alpha-1}}{(\lambda - it)^{\alpha-1}} e^{-y} \frac{dy}{(\lambda - it)} \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)(\lambda - it)^\alpha} \int_0^\infty e^{-y} y^{\alpha-1} dy
\end{aligned}$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)(\lambda - it)^\alpha}\Gamma\alpha$$
$$= \frac{\lambda^\alpha}{(\lambda - it)^\alpha}$$
$$= \left(\frac{\lambda}{\lambda - it}\right)^\alpha$$

Derivando a função característica da Distribuição Gama em relação a t, encontramos a $E(X)$, ou seja:

$$\varphi_X(t)' = \alpha \left(\frac{\lambda}{\lambda - it}\right)^{\alpha-1} \left(\frac{\lambda'(\lambda - it) - \lambda(\lambda - it)'}{(\lambda - it)^2}\right)$$
$$\varphi_X(t)' = \alpha \left(\frac{\lambda}{\lambda - it}\right)^{\alpha-1} \frac{\lambda i}{(\lambda - it)^2}$$

Com $i = 1$ e $t = 0$

$$E(X) = \frac{\varphi_X(0)'}{i} = \alpha\frac{\lambda}{\lambda}^{\alpha-1}\frac{\lambda i}{\lambda^2}$$
$$E(X) = \frac{\alpha 1^{\alpha-1} i}{\lambda}$$
$$E(X) = \frac{\alpha 1^{\alpha} 1^{-1} i}{\lambda}$$
$$E(X) = \frac{\alpha}{\lambda}$$

A partir da primeira derivada encontramos a $E(X^2)$, portanto:

$$\varphi_X(t)'' = \alpha' \left(\frac{\lambda}{\lambda - it}\right)^{\alpha-1} \left(\frac{\lambda i}{(\lambda - it)^2}\right) + \alpha \left(\frac{\lambda}{\lambda - it}\right)^{\alpha-1'} \left(\frac{\lambda i}{(\lambda - it)^2}\right) + \alpha \left(\frac{\lambda}{\lambda - it}\right)^{\alpha-1} \left(\frac{\lambda i}{(\lambda - it)^2}\right)'$$

com $i^2 = -1$ e $t = 0$

$$E(X^2) = \frac{\varphi_X(t)''}{i^2}$$

$$Var(X) = E(X) - E^2(X) = \varphi_X(0)'' - \varphi_X(0)'^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}$$

Suponha que X_{i_i} seja uma sequência de variáveis aleatórias independentes, tais que $X_i \sim Gamma(\alpha_i, \lambda)$. Então

$$\sum_{i=1}^{k} X_i \sim Gama\left(\sum_{i=1}^{\lambda} \alpha_i, \lambda\right)$$

Para demonstrar esse fato vamos utilizar da função geradora de momentos, mas antes vamos encontrar a função geradora da distribuição gamma.

$$\begin{aligned} M_X(t) = E[e^{tX}] &= \int_0^{\infty} e^{tx} x^{\alpha-1} e^{-\lambda x} \frac{\lambda^\alpha}{\Gamma(\alpha)} \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} e^{-(\lambda-t)x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha} \\ &= \left(\frac{\lambda}{\lambda-t}\right)^\alpha \end{aligned}$$

Assim mostramos que $\sum_{i=1}^{k} X_i \sim Gama\left(\sum_{i=1}^{k} \alpha_i, \lambda\right)$. De fato

$$\begin{aligned} M_{\sum_{i=1}^{k}} X_i(t) &= E\left(e^{t(\sum_{i=1}^{k} x_i)}\right) \\ &= E\left(\prod_{i=1}^{k} e^{tx_i}\right) \\ &= \prod_{i=1}^{k} E\left(e^{tx_i}\right) \\ &= \prod_{i=1}^{k} \left(\frac{\lambda}{\lambda-t}\right)^{\alpha_i} \\ M_{\sum_{i=1}^{k}} X_i(t) &= \left(\frac{\lambda}{\lambda-t}\right)^{\sum_{i=1}^{k} \alpha_i} \end{aligned}$$

Portanto,

$$\sum_{i=1}^{k} X_i \sim Gama\left(\sum_{i=1}^{\lambda} \alpha_i, \lambda\right)$$

Com essa demonstração mostramos também que soma de exponencial é uma exponencial, basta somarmos os parâmetros e que soma de qui-quadrado é qui-quadrado, basta somarmos os graus de liberdade. Pois descrito anteriormente essas distribuições são casos particulares da distribuição gama.

A função exponencial na verdade é um caso particular da função Gama, pois se $X \sim Exp(\lambda)$, então

$$X \sim Gama(1, \lambda)$$

3.9 Weibull e outras distribuições

Distribuições estatísticas existem em grande número como pode ser visto a partir da abrangente documentação de (JOHNSON *et al.*, 1992, 1994, 1995). As relações entre as distribuições são como conexão como são as relações familiares entre as dinastias europeias.

3.10 Weibull e distribuições exponenciais

Começamos com talvez a relação mais simples entre a distribuição weibull e qualquer outro tipo de distribuição, ou seja, a distribuição exponencial. Olhando para as origens da distribuição Weibull na prática reconhecemos que Weibull bem como (ROSIN *et al.*, 1933), mas a adição de um terceiro parâmetro na distribuição exponencial de dois parâmetros chegamos na distribuição exponencial generalizada de função de densidade dada abaixo.

$$f(y|a,b) = \frac{1}{b} e^{[-\frac{(y-a)}{b}]}$$

com $y \geq a$, $a \in \mathbb{R}$ e $b \geq 0$.

Construindo o gráfico com dois parâmetros $a = 0.776$ e $b = 2.1$

Agora, vamos

$$Y = \left(\frac{X-a}{b} \right)^c$$

e temos a distribuição exponencial reduzida $(a = 0, b = 1)$ com densidade

$$f(y) = e^{-y}, y > 0$$

então X tem distribuição de Weibull com fdp.

$$f(x|a,b,c) = \frac{c}{b} \left(\frac{x-a}{b} \right)^{c-1} e^{[-(\frac{x-a}{b})^c]}; x \geq a, a \in \mathbb{R}, b, c \geq 0.$$

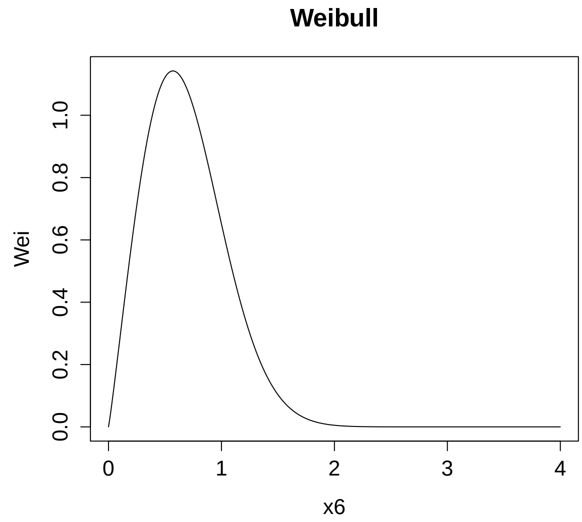

Figura 4: Representação gráfica da distribuição Weibull

É fácil ver que, com $c = 1$

$$f(x|a,b,1) = \frac{1}{b}\left(\frac{x-a}{b}\right)^{1-1} e^{\left[-\left(\frac{x-a}{b}\right)^{1}\right]}$$

$$f(x|a,b,1) = \frac{1}{b}e^{\left[-\left(\frac{x-1}{b}\right)^{1}\right]}$$

então, com $a = 1$, temos

$$f(x|1,b,1) = \frac{1}{b}e^{\left(\frac{-x}{b}\right)}, a > 0, a \in \mathbb{R}$$

3.11 Weibull e distribuição Gama

Na última seção, vimos vários parentes de distribuições Weibull originários em algum tipo de transformação da variação Weibull. Aqui, vamos apresentar uma classe de distribuções, em especial a família gama, onde inclui a distribuição Weibull como um caso especial para valores de parâmetros distintos. A forma reduzida da distribuição gama com um único parâmetro tem a seguinte função de distribuição

$$f(x|d) = \frac{x^{d-1}e^{-x}}{\Gamma(d)}; x \geq, d > 0.$$

d é um parâmetro de forma $d = 1$ resultando na distribuição exponencial reduzida. A introdução de um parêmetro de escala b produz a distribuição gama de dois parâmetros, dada por

$$f(x|b,d) = \frac{(x-a)^{d-1}e^{-(x-a)/b}}{b^d\Gamma(d)}; x \geq; b,d > 0$$

A distribuição gama de três-parâmetros tem um parâmetro de localização adicional, temos

$$f(x|a,b,d) = \frac{(x-a)^{d-1}e^{-(x-a)/b}}{b^d\Gamma(d)}; x \geq 0, a \in \mathbb{R}, b,d > 0$$

Stacy e Mihram (1965) introduziram uma segundo parâmetro de forma $c(c > 0)$ para os dois parâmetros distribuição gama, chegando-se a distribuição gama de quatro parâmetros. Tal distribuição é também chamada de distribuição gama generalizada,

$$f(x|a,b,c,d) = \frac{c(x-a)^{cd-1}e^{\left(\frac{x-a}{b}\right)^c}}{b^{cd}\Gamma(d)},$$

com $x \geq a$, $a \in \mathbb{R}$ e $b,c,d > 0$.

3.12 Weibull e outras distribuições

No contexto da distribuição gama generalizada mencionamos seus casos particulares. Um deles é a distribuição qui-quadrado com graus de liberdade que possui a função de densidade, dada por

$$f\left(x|v\right) = \frac{2}{\sqrt{2}\Gamma\left(\frac{v}{2}\right)} \cdot \left(\frac{x}{\sqrt{2}}\right)^{v-1} \cdot \exp\left\{-\left(\frac{x}{\sqrt{2}}\right)^2\right\}, x \geq 0, v \in n$$

produz uma distribuição Weibull com $a = 0$,$b = \sqrt{2}$ e $c = 2$ para $v = 2$. Substituindo o fator de escala $\sqrt{2}$ por uma escala geral parâmetro $b > 0$ leva a seguinte densidade Weibull.

$$f\left(x|a,b,2\right) = \frac{2}{b}\left(\frac{x}{b}\right) \cdot \exp\left\{-\left(\frac{x}{b}\right)^{2}\right\}$$

que é facilmente reconhecido como a densidade de uma distribuição de R de Raylaigh.

A função geradora de momento de X é

$$M(t) = E[e^{tX}] = \int_0^\infty e^{tx}\frac{\beta}{\alpha}x^{\beta-1}e^{-(1/\alpha)}\frac{\beta}{\alpha}x^{\beta-1}e^{-(1/\alpha)x^{\beta}}dx \qquad t > 0.$$

A função característica de X é

$$\varphi(t) = E[e^{itX}] = \int_0^\infty e^{itx}\frac{\beta}{\alpha}x^{\beta-1}e^{-(1/\alpha)x^{\beta}}dx \qquad t > 0.$$

A esperança, segue

$$E[X] = \frac{\alpha}{\beta}\Gamma\left(\frac{1}{\beta}\right)$$

Logo a variância é

$$Var[X] = \alpha^2\{\frac{2}{\beta}\Gamma\left(\frac{2}{\beta}\right) - \left[\frac{1}{\beta}\Gamma\left(\frac{1}{\beta}\right)\right]^2\}$$

4 Conclusão

O estudo das relações existentes entre as Distribuições de Probabilidade possui grande importância para um aluno de graduação em Estatística, tendo em vista que grande parte da produção científica dentro deste contexto envolve exatamente estas relações. Assim, o desenvolvimento deste trabalho foi bastante proveitoso no sentido de dar ao aluno suporte acadêmico necessário para que no futuro, possa desenvolvidos diversos trabalhos na área em estudo.

O estudo teórico proposto neste trabalho de iniciação científica evidenciou as propriedades e relações entre as distribuições gama, exponencial, weibull e railaigh. O desenvolvimento do trabalho proporcionou uma retomada aos conceitos de cálculo e incentivando uma busca sempre maior no âmbito da construção do conhecimento.

Referências

BALAKRISHNAN, N.; NEVZOROV, V. B. *A primer on statistical distributions*. [S.l.]: John Wiley & Sons, 2004.

DANTAS, C. A. B. *Probabilidade: Um Curso Introdutório Vol. 10*. [S.l.]: Edusp, 2008.

EVANS, M.; HASTINGS, N.; PEACOCK, B. *Statistical Distributions*. New York: Wiley, 2000.

EVERITT, B. S.; HOTHORN, T. *Handbook of Statistical analysis using R*. New York: [s.n.], 2006.

FOX, J. *An R and S-Plus companion to applied regression*. [S.l.]: Sage, 2002.

JOHNSON, N. L.; KOTZ, S.; BALA. *Continuous Univariate Distributions*. New York: [s.n.], 1994.

JOHNSON, N. L.; KOTZ, S.; BALAKR. *Continuous Univariate Distributions*. New York: Wiley, 1995.

JOHNSON, N. L.; KOTZ, S.; KEMP, A. W. *Univariate Discrete Distributions*. New York: [s.n.], 1992.

JONES, M. C. The complementary beta distribution. *Journal of Statistical Planning and Inference*, n. 104, p. 329–337, 2002.

LEEMIS, L. M. Relationships among common univariate distributions. *The American Statistician*, Taylor & Francis, v. 40, n. 2, p. 143–146, 1986.

LEEMIS, L. M.; MCQUESTON, J. T. Univariate distribution relationships. *The American Statistician*, Taylor & Francis, v. 62, n. 1, p. 45–53, 2008.

MAGALHÃES, M. N. *Probabilidade e variáveis aleatórias*. So paulo. [S.l.]: 2e. EDUSP, 2006.

MAINDONALD, J.; BRAUN, J. *Data Analysis and Graphics using R- An Example-Based Approach*. New York: Cambridge University Press, 2007.

MEYER, P. L. *Probabilidade: Aplicações à Estatística*. Livraria técnicos científicos editora. [S.l.]: 2.ed., 1983.

PETERNELLI, L. A.; MELLO, M. d. Conhecendo o r: uma visão estatística. *Viçosa, MG: UVF*, 2007.

RêGO, L. C. Notas de aula do curso probabilidade 4. Departamento de Estatística -UFPE. Disponível em: <http://www.de.ufpe.br/leandro/AulasET5842010-1.pdf>.

ROSIN, P.; RAMMLER, E.; SPERLING, K. *Size of powdered coat and its meaning to grinding (in German)*. Berlin: Bericht C 52 des Reichskohlenrats, 1933.

STACY, E. W.; MIHRAM, G. A. Parameter estimation for a generalized gamma distribution. *Technometrics*, n. 7, p. 349–357, 1965.