Distribuições contínuas: Lognormal, Gama, Weibull e Beta

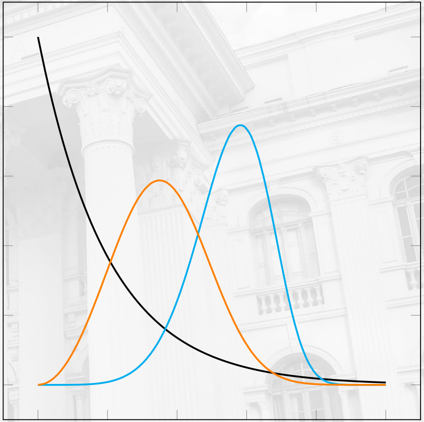
Prof. Walmes M. Zeviani

Departamento de Estatística
Universidade Federal do Paraná

Conteúdo

Neste vídeo

- Modelos probabilísticos contínuos.
  - Lognormal.
  - Gama.
  - Weibull.
  - Beta.
- Fundamentação e propriedades.
- Exemplos de aplicação.



Distribuição Lognormal

Características de uma v.a. com distribuição Lognormal

- Seja Y_{N} uma variável com distribuição Normal. Então, $Y = \exp\{Y_{\mathrm{N}}\}$ tem distribuição Lognormal.
- Diferente da Normal, a Lognormal tem suporte no conjunto dos **reais positivos** e apresenta assimetria.
- A distribuição Lognormal, assim como outras neste material, tem aplicações na modelagem de variáveis na área de **confiabilidade** e análise de **sobrevivência**.

$$f(y) = \frac{1}{y\omega\sqrt{2\pi}} \exp\left\{-\frac{(\ln(y)-\theta)^2}{2\omega^2}\right\}$$

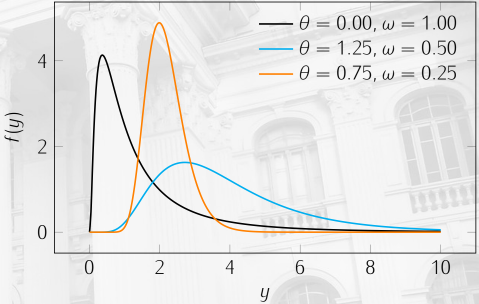


Figura 1. Distribuição Lognormal.

Exemplos de v.a. com distribuição Lognormal

1. Qualquer variável que é o exponencial de uma v.a. Normal.
2. Tempo para a falha de um equipamento eletrônico.
3. Tempo de vida de um paciente após um tratamento médico.
4. Intervalo de tempo entre acionamentos de um gerador de energia.
5. Intervalo de tempo entre quedas de um serviço web.
6. Duração de um processo judicial.

Figura 2. Interruptores em um quadro de força. Foto de Pixabay no Pexels.

Distribuição Lognormal

Seja $Y_{\text{N}} \sim \text{N}(\theta, \omega^2)$. Então, $Y = \exp\{Y_{\text{N}}\}$ tem distribuição Lognormal com função densidade de probabilidade dada por

$$f(y) = \frac{1}{y\omega\sqrt{2\pi}} \exp\left\{-\frac{(\ln(y) - \theta)^2}{2\omega^2}\right\}, \quad y > 0.$$

Denotamos por $Y \sim \text{LNor}(\theta, \omega)$. $\theta \in \mathbb{R}$ e $\omega > 0$ são a média e a variância de uma v.a. Lognormal e são chamados de parâmetros de **locação** e **dispersão**.

A média e variância de Y são

$$\mu = \text{E}(Y) = \exp\left\{\theta + \frac{\omega^2}{2}\right\} \quad \text{e} \quad \sigma^2 = \text{V}(Y) = \exp\left\{2\theta + \omega^2\right\} \cdot \left(\exp\left\{\omega^2\right\} - 1\right).$$

A distribuição não possui forma fechada para a função de distribuição acumulada.

Gráfico de densidade da Lognormal

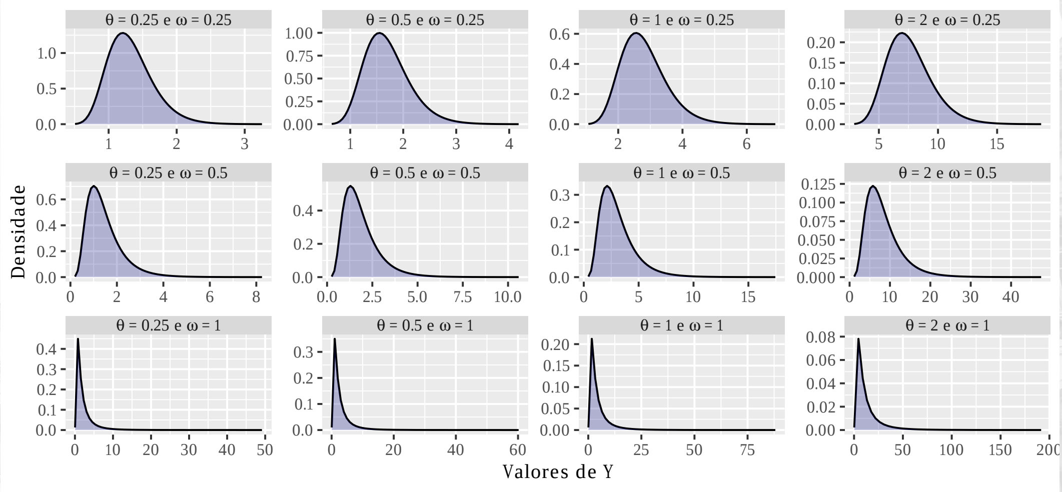


Figura 3. Gráficos de densidade para a distribuição Lognormal.

Gráfico de distribuição da Lognormal

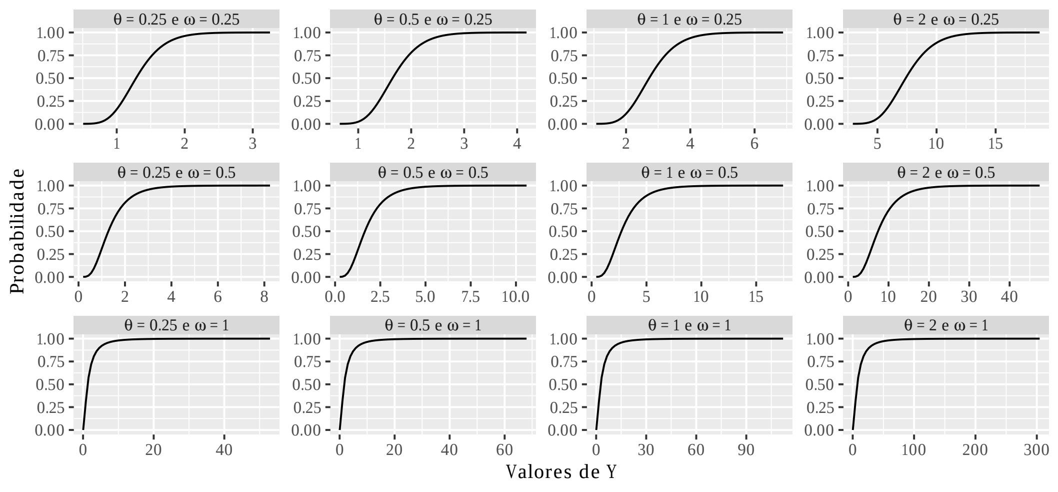


Figura 4. Gráficos de distribuição para a Lognormal.

Exemplo: duração de atendimentos

A duração do atendimento de cada cliente no caixa de um supermercado tem distribuição Lognormal com parâmetro de locação $\theta = 1.5$ e parâmetro de dispersão $\omega = 0.5$.

1. Qual a média e variância da duração dos atendimentos?
2. Qual a probabilidade de um atendimento durar menos de 5 minutos?

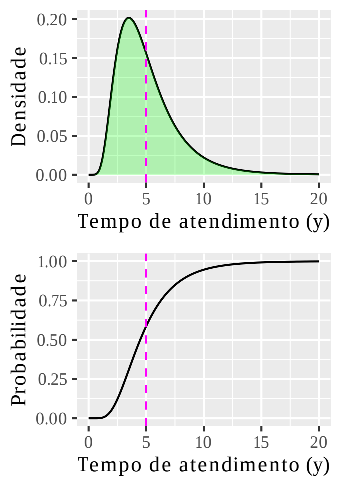


Figura 5. Gráficos da distribuição para o exercício.

Solução

1. Aplicam-se as expressões para média e variância,

$$\mu = \exp\left\{1.5 + \frac{0.5^2}{2}\right\} = 5.078.$$
$$\sigma^2 = \exp\left\{2 \cdot 1.5 + 0.5^2\right\} \cdot \left(\exp\left\{0.5^2\right\} - 1\right) = 7.325.$$

2. A probabilidade do referido evento é calculada por

$$\mathrm{P}(Y \leq 5) = \int_0^5 f(y)\,\mathrm{d}y = 0.587.$$

Usa-se algum software de integração numérica, planilha eletrônica ou linguagem de programação, para obtê-la.

Distribuição Gama

Características de uma v.a. com distribuição Gama

- Seja $Y_{\text{E}i} \sim \text{Exp}(\lambda)$ $(i \in \{1, \ldots, k\})$ uma variável com distribuição Exponencial. Então, $Y = Y_{\text{E}1} + Y_{\text{E}2} + \cdots + Y_{\text{E}k}$ tem distribuição Gama.
- A Gama tem suporte no conjunto dos **reais positivos**, assumindo formas assimétricas.
- Ela tem aplicações na área de **confiabilidade** e análise de **sobrevivência**, assim como a Lognormal.

$$f(y) = \frac{\lambda^r}{\Gamma(r)} \cdot y^{r-1} \cdot \exp\{-\lambda y\}$$

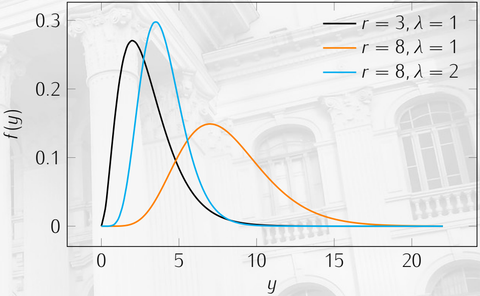


Figura 6. Distribuição Gama.

Exemplos de v.a. com distribuição Gama

1. Soma de v.a. com distribuição Exponencial.
2. Tempo de carregamento de um navio.
3. Volume de chuva em dias com precipitação.
4. Tempo de permanência de um usuário em um site.
5. Distribuição de idade de animais em ambiente natural.
6. Tempo de vida de um paciente após transplante.
7. Distância dos passes de bola em um jogo de futebol.

Figura 7. Navio sendo carregado com containers. Foto de Kai Pilger no Pexels.

Distribuição Gama

A variável aleatória Y tem distribuição Gama de parâmetros $r > 0$ e $\lambda > 0$ se sua função densidade de probabilidade é dada por

$$f(y) = \frac{\lambda^r}{\Gamma(r)} \cdot y^{(r-1)} \cdot \exp\{-\lambda y\}, \quad y > 0,$$

em que

$$\Gamma(r) = \int_0^\infty u^{r-1} \exp\{-u\}\, \mathrm{d}u \quad \text{com} \quad \Gamma(r) = (r-1)! \quad \text{quando } r \in \mathbb{N},$$

é a **função gama** (por isso distribuição Gama).

Denotamos por $Y \sim \text{Gama}(r, \lambda)$. Os parâmetros λ e r são chamados de **taxa** e **forma**, respectivamente.

A média e a variância de Y são

$$\mu = \text{E}(Y) = \frac{r}{\lambda} \quad \text{e} \quad \sigma^2 = \text{V}(Y) = \frac{r}{\lambda^2}.$$

Gráfico de densidade da Gama

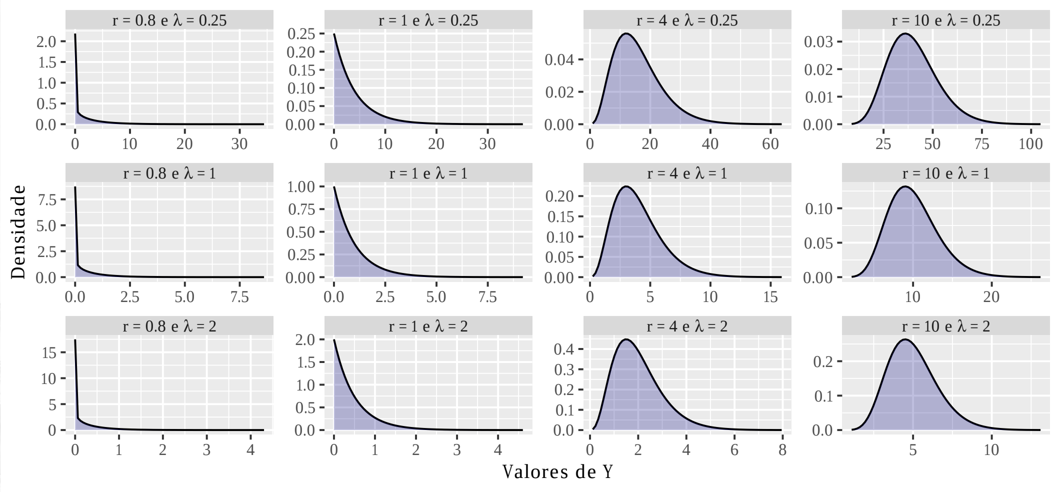


Figura 8. Gráficos de densidade para a Gama.

Gráfico de distribuição da Gama

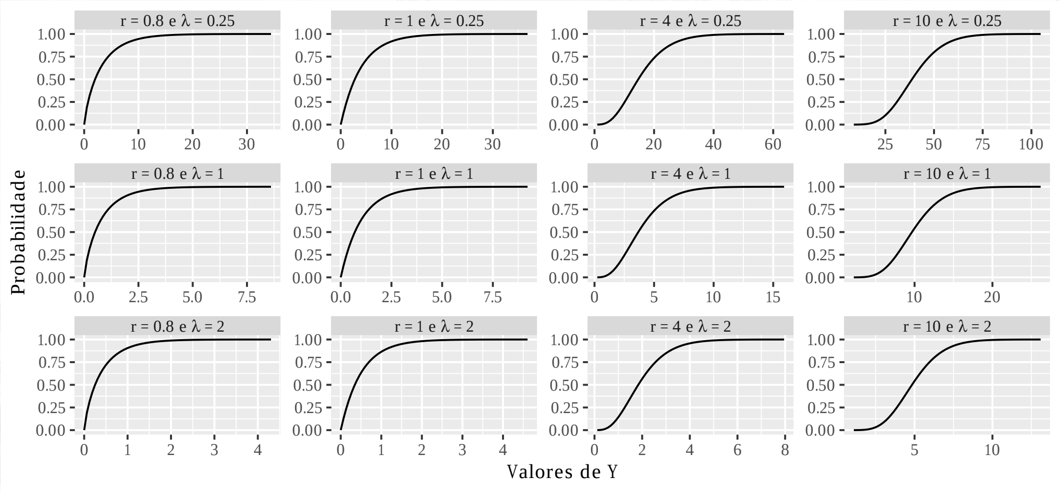


Figura 9. Gráficos de distribuição para a Gama.

Relação da Gama com outras distribuições

- A distribuição Gama tem como **caso particular** a distribuição exponencial (λ) ao fixarmos $r = 1$.
- Dessa relação, a Gama pode ser obtida como o **tempo acumulado** para k eventos de Poisson, uma vez que o intervalo entre eventos é Exponencial.
- A Gama tem mais variedades de formas por ter 2 parâmetros, permitindo modelar adequadamente um maior número de variáveis aleatórias que a Exponencial.
- A **soma** de v.a. Gama é Gama, ou seja, se $Y_1, Y_2, \cdots, Y_k$ são variáveis aleatórias independentes, com distribuição Gama de parâmetros r e λ, então

$$Y_{\text{soma}} = Y_1 + Y_2 + \cdots + Y_k \sim \text{Gama}(kr, \lambda).$$

- A distribuição Erlang é um **caso particular** da Gama quando r é um número natural, $r \in \{1, 2, \ldots\}$.

Exemplo: duração de atendimentos

A duração do atendimento de cada cliente no caixa de um supermercado tem distribuição Gama com parâmetro de forma $r = 12$ e parâmetro de taxa $\lambda = 2$.

1. Qual a média e variância da duração dos atendimentos?
2. Qual a probabilidade de um atendimento durar menos de 5 minutos?
3. Se um cliente entra na fila tendo 4 clientes antes dele, considerando independência entre os atendimentos, qual a probabilidade de ser atendido antes de 20 minutos?

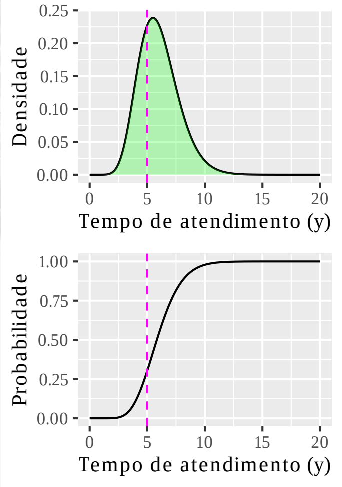


Figura 10. Gráficos da distribuição para o exercício.

Solução

1. Aplicam-se as expressões para média e variância,

$$\mu = \frac{12}{2} = 6 \quad \text{e} \quad \sigma^2 = \frac{12}{2^2} = 3.$$

2. A probabilidade do referido evento é dada por

$$\mathrm{P}(Y \leq 5) = \int_0^5 f(y)\,\mathrm{d}y = 0.303.$$

3. A soma de k v.a. $Y \sim \text{Gama}(r, \lambda)$ independentes é uma v.a. Gama, ou seja, $Y_{\text{comb}} = Y_1 + \cdots + Y_k \sim \text{Gama}(kr, \lambda)$. Dessa forma, sendo $r = 12$ e $\lambda = 2$.

$$\mathrm{P}(Y_{\text{comb}} \leq 20) = \int_0^{20} \frac{2^{4\cdot 12}}{\Gamma(4 \cdot 12)} \cdot y^{(4\cdot 12-1)} \cdot \exp\{-2y\}\,\mathrm{d}y = 0.120.$$

Parametrizações adicionais da Gama

- Parametrização de **forma** (r) e **escala** (θ)

$$f_{(}y) = \frac{\theta^{-r} \cdot y^{r-1} \cdot \exp\{-y/\theta\}}{\Gamma(\alpha)}, \quad \text{em que } \lambda = 1/\theta.$$

- Parametrização da **média** (μ) e **forma** (r)

$$f(y) = \left(\frac{r}{\mu}\right)^r \cdot \frac{y^{(r-1)} \cdot \exp\{-ry/\mu\}}{\Gamma(r)}, \quad \text{em que } \mu = r/\lambda.$$

- Parametrização de **moda** (γ) e **desvio-padrão** (σ)

$$f(y) = \frac{\lambda^r}{\Gamma(r)} \cdot y^{(r-1)} \cdot \exp\{-\lambda y\}, \text{ em que } \lambda = (\gamma + \sqrt{\gamma^2 + 4\sigma^2})/(2\sigma^2) \text{ e } r = 1 + \gamma \cdot \lambda.$$

- E tem ainda a parametrização na forma da **família exponencial** de distribuições.

Distribuição de Weibull

Características de uma v.a. com distribuição de Weibull

- A Weibull tem suporte no conjunto dos **reais positivos**, assumindo formas assimétricas, com assimetria à **direita** e à **esquerda**.
- Ela tem aplicações na área de **confiabilidade** e análise de **sobrevivência**, assim como a Lognormal e a Gama.
- Tem como **caso particular** a distribuição Exponencial.

$$f(y) = \frac{\beta}{\delta}\left(\frac{y}{\delta}\right)^{\beta-1} \exp\left\{-\left(\frac{y}{\delta}\right)^{\beta}\right\}$$

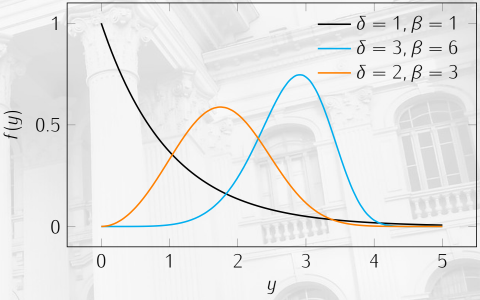


Figura 11. Distribuição Weibull.

Exemplos de v.a. com distribuição Weibull

1. Duração de uma cirurgia.
2. Quantidade de resíduos industriais na água.
3. Altura do nível da água de um rio após as chuvas.
4. Tempo para eclosão de larvas de insetos.
5. Tempo para aparecimento de sintomas de doença em frutos.
6. Produtividade de leite de um rebanho.
7. Intervalo de tempo entre corridas de táxi.

Figura 12. Médicos em uma cirurgia. Foto de Pixabay no Pexels.

Distribuição de Weibull

A variável aleatória Y tem distribuição de Weibull de parâmetros $\theta > 0$ e $\beta > 0$ se sua função densidade de probabilidade é dada por

$$f(y) = \frac{\beta}{\delta}\left(\frac{y}{\delta}\right)^{\beta-1} \exp\left\{-\left(\frac{y}{\delta}\right)^{\beta}\right\}, \quad y > 0.$$

Denotamos por $Y \sim \text{Wei}(\beta, \delta)$. β é chamado de parâmetro de **forma** e δ é o parâmetro de **escala**.

A média e a variância de Y são dadas por

$$\mu = \text{E}(Y) = \delta\Gamma\left(1 + \frac{1}{\beta}\right) \quad \text{e} \quad \sigma^2 = \text{V}(Y) = \delta^2 \left\{\Gamma\left(1+\frac{2}{\beta}\right) - \left\{\Gamma\left(1+\frac{1}{\beta}\right)\right\}^2\right\}.$$

Diferente das duas anteriores, a Weibull tem expressão fechada para a função de distribuição **acumulada** que é

$$F(y) = 1 - \exp\{-(y/\delta)^{\beta}\}.$$

Gráfico da densidade da Weibull

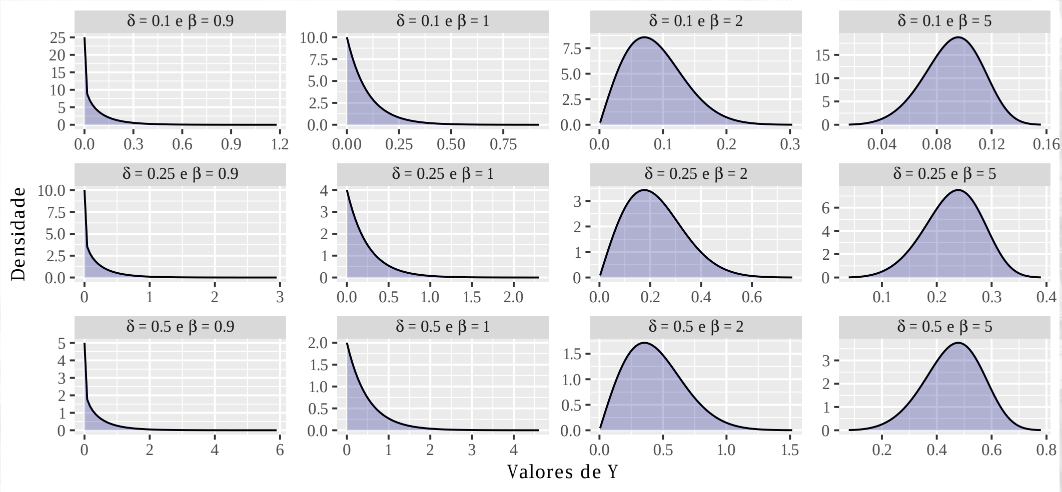


Figura 13. Gráfico da densidade para a Weibull.

Gráfico de distribuição da Weibull

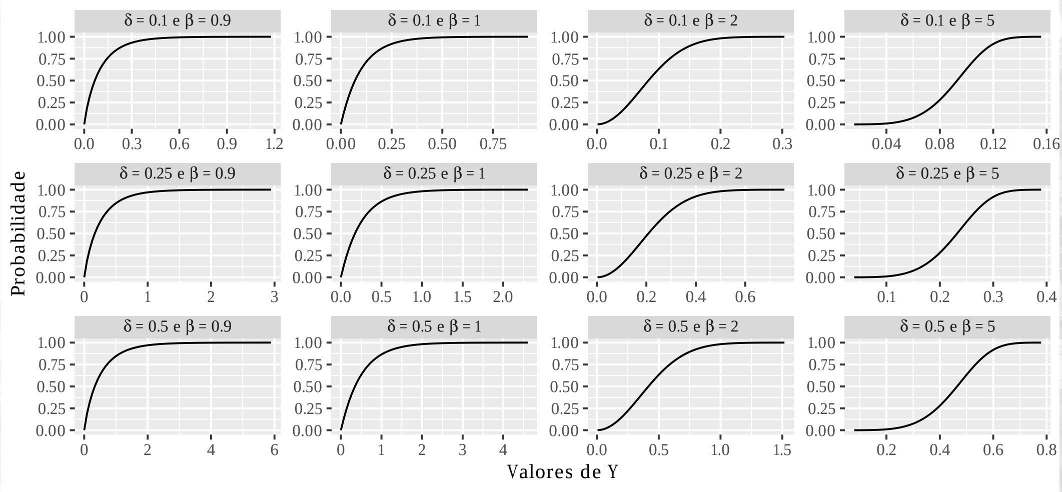


Figura 14. Gráficos de distribuição para a Weibull.

Exemplo: duração de atendimentos

A duração do atendimento de cada cliente no caixa de um supermercado tem distribuição de Weibull com parâmetro de forma $\delta = 7$ e parâmetro de escala $\beta = 4$.

1. Qual a média e variância da duração dos atendimentos?
2. Qual a probabilidade de um atendimento durar menos de 5 minutos?

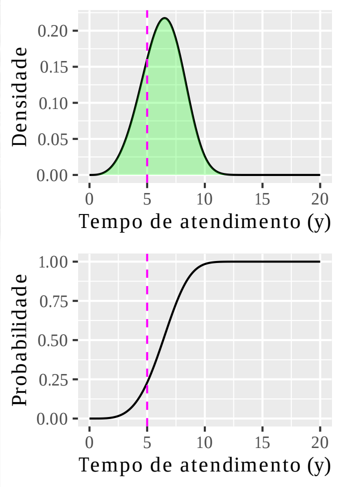


Figura 15. Gráficos da distribuição para o exercício.

Solução

1. Aplicam-se as expressões para média e variância

$$\mu = 7\Gamma\left(1 + \frac{1}{4}\right) = 6.345.$$

$$\sigma^2 = 7^2 \left\{\Gamma\left(1 + \frac{2}{4}\right) - \left\{\Gamma\left(1 + \frac{1}{4}\right)\right\}^2\right\} = 3.168.$$

2. Como a Weibull tem expressão para a função de distribuição acumulada, tem-se

$$\text{P}(Y \leq 5) = \int_0^5 f(y)\,\text{d}y = (1 - \exp\{-(5/7)^4\}) - (1 - \exp\{-(0/7)^4\}) = 0.229.$$

Relação da Weibull com outras distribuições e parametrizações

- A Weibull tem como caso particular a distribuição Exponencial quando $\beta = 1$.
- A distribuição de Raleigh também é caso particular quando $\beta = 2$.
- A parametrização de **média** (μ) é

$$f(y) = \frac{\beta}{\delta}\left(\frac{y}{\delta}\right)^{\beta-1} \cdot \exp\{-(y/\delta)^{\beta}\}, \text{ em que } \delta = \mu/(\Gamma(1/\beta)+1).$$

- Existem outras parametrizações e extensões para mais parâmetros.

Distribuição Beta

Características de uma v.a. com distribuição Beta

- A Beta é uma v.a. **contínua limitada** ao intervalo (0, 1). Dessa forma, serve para variáveis que têm essa característica.
- É uma distribuição com 2 parâmetros que pode apresentar muitas formas
  - Forma simétrica.
  - Assimétrica com moda no interior.
  - Assimétrica com moda na borda.
  - Forma de banheira.
- As aplicações incluem modelagem de variáveis que são **frações** de números reais como **teores** vindos de **concentrações** expressas em m/m ou v/v.

$$f(y) = \frac{y^{\alpha-1} \cdot (1-y)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}$$

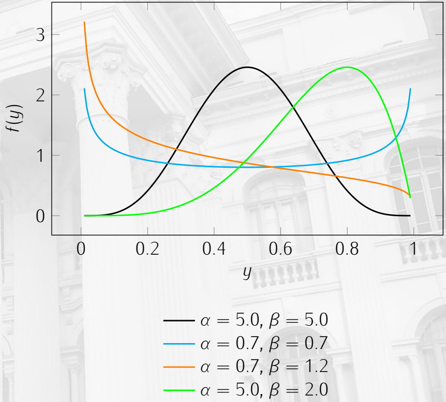


Figura 16. Distribuição Beta.

Exemplos de v.a. com distribuição Beta

1. Teor de gordura no leite.
2. Teor de proteína de uma ração animal.
3. Teor de matéria orgânica do solo.
4. Fração do tempo nadando em uma prova de triathlon.
5. Fração de tempo de um jogo de futebol com bola parada.
6. Concentração de álcool em uma cerveja.
7. Concentração de açúcar em um chocolate.
8. Índices de qualidade de vida.

Figura 17. Cerveja sendo servida. Foto de ELEVATE no Pexels.

Distribuição Beta

A variável aleatória Y tem distribuição Beta de parâmetros $\alpha > 0$ e $\beta > 0$ se sua função densidade de probabilidade é dada por

$$f(y) = \frac{y^{\alpha-1} \cdot (1-y)^{\beta-1}}{\mathrm{B}(\alpha, \beta)} \quad 0 < y < 1,$$

em que

$$\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)},$$

é a **função beta** (por isso o nome distribuição Beta). Denotamos por $Y \sim \mathrm{Beta}(\alpha, \beta)$.

A média e a variância de Y são dadas por

$$\mu = \mathrm{E}(Y) = \frac{\alpha}{\alpha + \beta} \quad \text{e} \quad \sigma^2 = \mathrm{V}(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Gráfico da densidade para a Beta

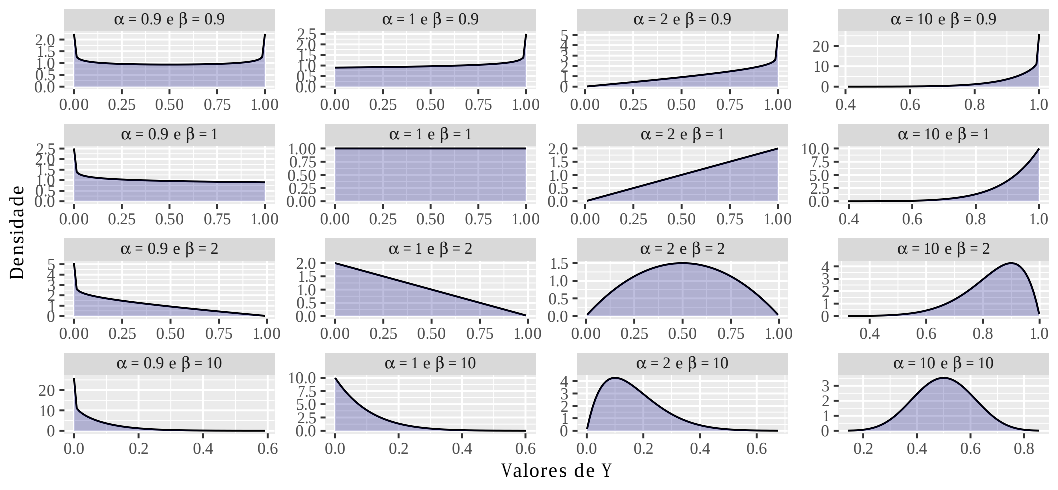


Figura 18. Gráfico da densidade para a Beta.

Gráfico de distribuição da Beta

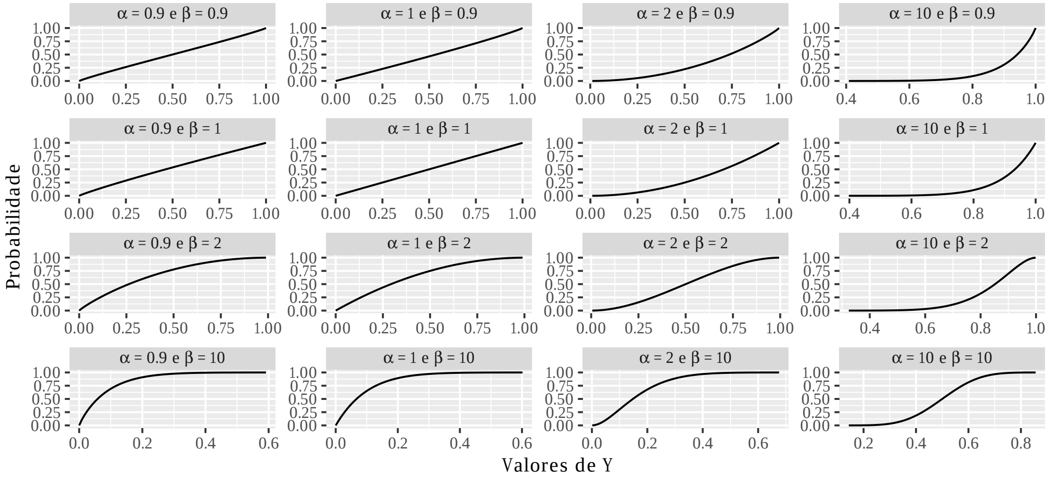


Figura 19. Gráficos de distribuição para a Beta.

Exemplo: teor de gordura no leite

O teor de gordura no leite de um rebanho bovino é uma variável com distribuição Beta com parâmetros $\alpha = 2$ e $\beta = 50$.

1. Qual o teor médio de gordura no leite?
2. Qual o percentual de amostras que terá teor de gordura menor que 10%?

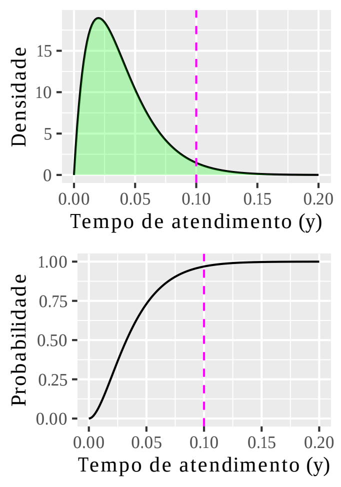


Figura 20. Gráficos da distribuição para o exercício.

Solução

1. Aplica-se a expressão para a média,

$$\mu = \frac{2}{2+50} = 0.038.$$

2. A probabilidade de referido evento é dada por

$$\text{P}(Y \leq 0.1) = \int_0^{0.1} f(y)\,\text{d}y = 0.969.$$

Notas adicionais

Reparametrizações

- Como visto, é comum as distribuições terem várias **parametrizações**.
- A distribuição ainda é a mesma mas a interpretação dos parâmetros, expressões para média e variância, mudam conforme cada parametrização.
- Além de mudar o significado dos parâmetros e expressões, muda as propriedades dos **estimadores** para os parâmetros.
- Veja o caso da Exponencial

$$f_1(y) = \lambda \exp\{-\lambda y\} \quad \text{e} \quad f_2(y) = \frac{1}{\mu} \exp\left\{-\frac{y}{\mu}\right\}, \text{ pois } \mu = 1/\lambda.$$

- Esteja atento a isso, portanto, ao usar *softwares* que tem as distribuições implementadas.
- Faça testes até determinar a parametrização que está sendo utilizada para não cometer erros.

Função de risco

A função de risco de uma v.a. é definida por

$$h(y) = \frac{f(y)}{1 - F(y)}$$

em que f é a função densidade de probabilidade e F a função de distribuição.

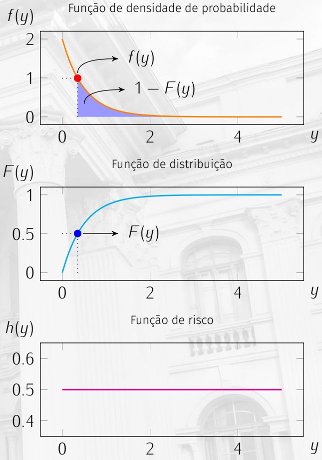


Figura 21. Função de risco para a distribuição Exponencial.

Funções de risco da Lognormal, Gama e Weibull

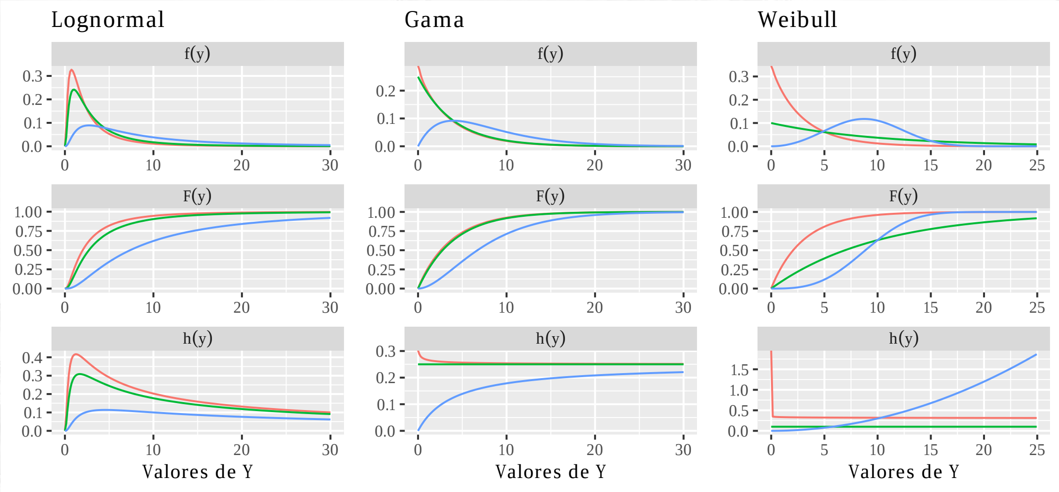


Figura 22. Gráfico das funções de risco das distribuições Lognormal, Gama e Weibull.

Distribuições contínuas adicionais

- Behrens-Fisher.
- Beta generalizada.
- Birnbaum-Saunders.
- Box-Cox.
- Gama Inversa.
- Gumbel.
- Kumaraswamy.
- Laplace.
- Log-logistic.
- Logistic.
- Normal Inversa.
- Simplex.
- Trapezoidal.
- Triangular.
- Tweedie.
- E muitas outras.

Relacionamento entre as distribuições vistas

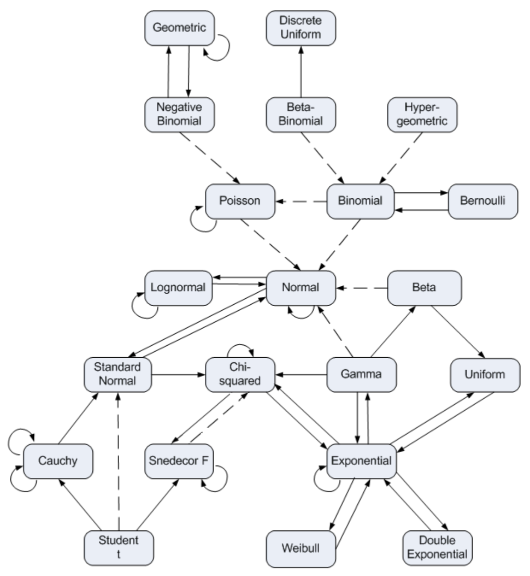


Figura 23. Relação entre variáveis aleatórias. Extraído de www.johndcook.com/.

Considerações finais

Considerações finais

Revisão

- Modelos probabilísticos contínuos.
  - Lognormal.
  - Gama.
  - Weibull.
  - Beta.
- Fundamentação e propriedades.
- Exemplos de aplicação.