Variáveis aleatórias: modelos contínuos especiais (parte 2)

Notas de Aula da Professora Verónica González-López, digitadas por Beatriz Cuyabano, Pós-Graduação IMECC/UNICAMP, com modificações do Prof. Caio Azevedo

Modelos probabilísticos contínuos especiais

- Veremos agora modelos contínuos especiais, no sentido de que eles foram motivados por problemas reais na Estatística, Probabilidade e, principalmente, em outras áreas.
- Apresentaremos os modelos mais básicos (conhecidos) e indicaremos onde encontrar outros modelos.
- É virtualmente impossível dar indicações de todos os modelos existentes também, porque regularmente, novos modelos são disponibilizados na literatura.

Modelos probabilísticos discretos especiais

- Apresentaremos a concepção dos modelos (como surgiram) e suas principais características (em relação ao que apresentamos anteriormente).
- Adiante (no curso) apresentaremos outras características e propriedades.

Modelo Uniforme

- Dizemos que a v.a. X tem distribuição uniforme no intervalo $[a, b]$, $a < b$ se sua f.d.p. (função densidade de probabilidade ou função densidade ou simplesmente densidade) f_X é dada por:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases}$$

ou

$$f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x).$$

- Notação: $X \sim U[a, b]$ ou $X \sim U(a, b)$
- Exercício: provar que $f_X(.)$ é, de fato, uma densidade.

Modelo Uniforme

- Cálculo da f.d.a.:

$$F_X(x) = \begin{cases} 0, & x < a \\ \displaystyle\int_a^x \frac{1}{b-a}dt = \frac{1}{b-a}t\Big|_a^x = \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- Em termos de funções indicadoras:

$$F_X(x) = \left(\frac{x-a}{b-a}\right) \mathbb{1}_{[a,b]}(x) + I_{(b,\infty)}(x),$$

$$S_X(x) = \mathbb{1}_{(-\infty,a)}(x) + \left(\frac{b-x}{b-a}\right) \mathbb{1}_{[a,b]}(x).$$

Teorema

- Teorema de transformação de variáveis aleatórias contínuas: Se X uma vac com densidade ($f_X(.)$) e defina $Y = g(X)$, em que $g(.)$ é uma transformação 1 a 1. Então:

$$f_Y(y) = f_X\left(g^{-1}(y)\right)|J|\,\mathbb{1}_B(y),$$

  em que B é suporte de Y, o qual é induzido por A e g e $J = \frac{d}{dy}g^{-1}(y)$ (conhecido como Jacobiano).

- Demonstração: exercício.

Teorema (Cont.)

- Obs: se $g(.)$ não for um a um mas, puder ser "particionada" em r transformações (em função do suporte de X, digamos A) 1 a 1, o teorema ainda é válido. Com efeito, seja $A_1, ..., A_r$ essa partição ($A = \dot{\bigcup}_{i=1}^{r} A_i$), assim:

$$f_Y(y) = \sum_{i=1}^{r} f_X\left(g_i^{-1}(y)\right)|J_i|\,\mathbb{1}_{B_i}(y),$$

em que B_i é suporte de Y, o qual é induzido por A_i e g_i e $J_i = \dfrac{d}{dy} g_i^{-1}(y)$, $i = 1, 2, ..., r$.

- Demonstração: exercício.

Exemplo

- Seja X tal que $f_X(x) = 2x\mathbb{1}_{(0,1)}(x)$ e defina $Y = X^2$. Encontre $f_Y(y)\mathbb{1}_B(y)$. Assim $x = \sqrt{y}$ e $B = (0,1)$. Logo $J = \dfrac{d}{dy}\sqrt{y} = \dfrac{1}{2}y^{-1/2}$ e, portanto:

$$f_Y(y) = f_X(\sqrt{y})\frac{1}{2}y^{-1/2}\mathbb{1}_{(0,1)}(y) = \mathbb{1}_{(0,1)}(y),$$

ou seja $Y \sim U(0,1)$.

Exemplo

- Seja X tal que $f_X(x) = \frac{3}{2}x^2 \mathbb{1}_{(-1,1)}(x)$ e defina $Y = X^2$. Encontre a densidade de Y.
- Note que $B = (0,1)$, em que B é o suporte de Y. Defina $A_1 = (-1,0)$ e $A_2(0,1)$. Logo $g(x) = x^2$, mas $g_1^{-1}(y) = -\sqrt{y}$ e $g_2^{-1}(y) = \sqrt{y}$ e, assim $J_1 = -\dfrac{d}{dy}\sqrt{y} = -\dfrac{1}{2}y^{-1/2}$ e $J_2 = \dfrac{d}{dy}\sqrt{y} = \dfrac{1}{2}y^{-1/2}$, portanto:

$$\begin{aligned} f_Y(y) &= f_X(\sqrt{y})\frac{1}{2}y^{-1/2}\mathbb{1}_{(0,1)}(y) + f_X(\sqrt{y})\frac{1}{2}y^{-1/2}\mathbb{1}_{(0,1)}(y) \\ &= \frac{3}{2}y^{1/2}\mathbb{1}_{(0,1)}(y). \end{aligned}$$

Teorema

- Teorema da transformada integral: Se X uma vac tal que sua fda ($F_X(.)$) seja estritamente crescente. Então $Y = F_X(X) \sim U(0,1)$.
- Demonstração: Note, primeiramente, que $B = (0,1)$ (suporte de Y). Além disso, temos que

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) = P(F_X(X) < y) = P(X \leq F_X^{-1}(y)) \\
&= F_X(F_X^{-1}(y)) = y \rightarrow f_Y(y) = F_Y'(y) = 1.
\end{aligned}
$$

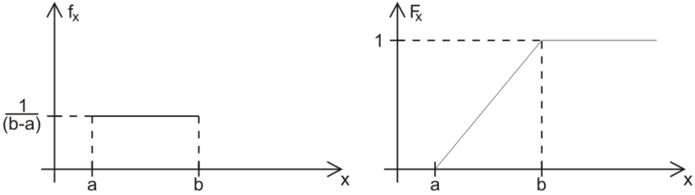
- Assim $f_Y(y) = \mathbb{1}_{(0,1)}(y)$. Analogamente, $X = F_X^{-1}(Y) \sim F_X$.

Simulação

- Logo, para simular de uma vac X com fda $F_X(.)$ basta simular $u \sim U(0,1)$ e fazer $x = F_X^{-1}(u)$.
- Com efeito, não entraremos em detalhes sobre as simulações dos modelos que veremos.
- Eventualmente, há outros algoritmos mais eficientes para gerar de certas variáveis contínuos.

Modelo Uniforme

- Gráficos da f.d.p. e f.d.a.



- No R:
  - Densidade: dunif(x,min=a,max=b).
  - Fda: punif(x,min=a,max=b)(fda).
  - Simulação: runif(n,min=a,max=b) (mais detalhes, veja aqui).

Modelo Uniforme

- Cálculo da $E(X)$:

$$
\begin{aligned}
E(X) &= \int_a^b x\frac{1}{(b-a)}dx = \frac{1}{b-a}\left(\frac{b^2}{2}-\frac{a^2}{2}\right) = \frac{(b+a)(b-a)}{2(b-a)} \\
&= \frac{a+b}{2}.
\end{aligned}
$$

- Cálculo da $Var(X)$:

$$
\begin{aligned}
E(X^2) &= \int_a^b x^2\frac{1}{(b-a)}dx = \frac{1}{3(b-a)}\left[b^3-a^3\right] \\
&= \frac{(b-a)(a^2+ab+b^2)}{3(b-a)} = \frac{(a^2+ab+b^2)}{3}.
\end{aligned}
$$

Modelo Uniforme

- Logo

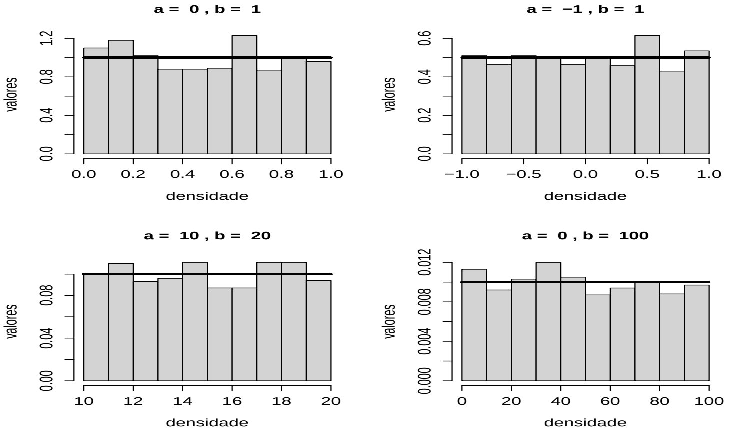
$$
\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 = \frac{(a^2+ab+b^2)}{3} - \frac{(b+a)^2}{4} \\
&= \frac{(a^2+ab+b^2)}{3} - \frac{b^2+2ab+a^2}{4} = \frac{(b-a)^2}{12}.
\end{aligned}
$$

Modelo Uniforme

- Exemplo: A temperatura T de destilação do petróleo é crucial na determinação da qualidade final do produto. T é considerada uma v.a. com distribuição $U[150^\circ, 300^\circ]$. Temos que

$$f_T(t) = \frac{1}{150} I_{[150^\circ, 300^\circ]}(t), E(T) = 75, V(X) = 1875.$$

fdp (linha) e valores simulados (histograma): uniforme



Modelo Exponencial

- Dizemos que uma v.a. X possui distribuição exponencial com parâmetro λ ($\lambda > 0$) se sua f.d.p. f_X é dada por:

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

ou

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x).$$

- Notação: $X \sim exp(\lambda)$. Exercício: provar que, de fato, $f_X(.)$ é uma densidade.

Modelo Exponencial

- Cálculo da f.d.a. ($y = \lambda t$):

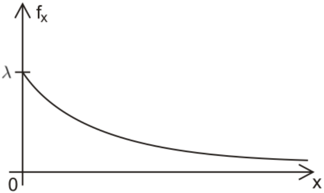
$$F_X(x) = \begin{cases} \displaystyle\int_0^x \lambda e^{-\lambda t} dt = \int_0^{\lambda x} e^{-y} dy = -e^{-y}\Big|_0^{\lambda x} = 1 - e^{-\lambda x}, & x > 0 \\ 0, \text{caso contrário} \end{cases}$$

- Logo:

$$\begin{aligned} F_X(x) &= \left(1 - e^{-\lambda x}\right) \mathbb{1}_{(0,\infty)}(x), \\ S_X(x) &= \mathbb{1}_{(-\infty,0]}(x) + e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x). \end{aligned}$$

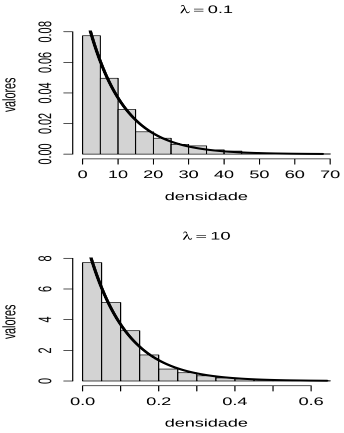
Modelo Exponencial

- Gráfico da f.d.p.



- No R:
  - Densidade: dexp(x,rate=lambda).
  - Fda: pexp(x,rate=lambda) (fda).
  - Simulação: rexp(n,rate=lambda) (mais detalhes veja aqui).

fdp (linha) e valores simulados (histograma): exponencial



Modelo Exponencial

- Introduzindo a função gama:

$$\Gamma(u) = \int_0^\infty t^{u-1} e^{-t} dt, \quad u > 0$$

- Propriedades:
  - $\Gamma(u+1) = u\Gamma(u)$.
  - $\Gamma(n) = (n-1)! \quad , \forall \; n \in \mathcal{N}$.
  - $\Gamma(1/2) = \sqrt{\pi}$.
- Essa função tem um papel importante na modelagem estatística e auxilia nos cálculos de várias quantidades relacionadas a várias distribuições.

Modelo Exponencial

- Cálculo da $E(X)$ (seja $u = x\lambda, x = \frac{u}{\lambda}, du = \lambda dx$):

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}\int_0^\infty u^{2-1} e^{-u} du = \frac{1}{\lambda}\Gamma(2) = \frac{1}{\lambda}.$$

- Cálculo da $Var(X)$ (seja $u = x\lambda, x = \frac{u}{\lambda}, du = \lambda dx$):

$$E(X^2) = \int_0^\infty x^2\lambda e^{-\lambda x} dx = \frac{1}{\lambda^2}\int_0^\infty u^{3-1} e^{-u} du = \frac{1}{\lambda^2}\Gamma(3) = \frac{2}{\lambda^2}.$$

- Logo

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Modelo Exponencial

- Exemplo: O tempo de vida (em horas) de um transistor é uma v.a. T com distribuição $exp(\lambda)$ em que $\lambda = \dfrac{1}{500}$. Assim
  - $E(T) = 500$ horas
  - $P(T \geq 500) = \displaystyle\int_{500}^{\infty} f_T(t)dt = e^{-1} = 0,3678$ (no R, 1-pexp(500,rate=1/500)).

Modelo Normal

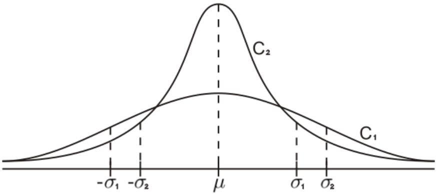
- Dizemos que uma v.a. X possui distribuição normal com parâmetros μ e σ^2, $\mu \in \mathcal{R}$ e $\sigma^2 > 0$, se sua f.d.p., f_X é dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \mathbb{1}_{(-\infty,\infty)}(x).$$

- Notação: $X \sim N(\mu, \sigma^2)$.
- Devido à impossibilidade de calcular a f.d.a. analiticamente, recorre-se à tabela da **Normal Padrão** ($N(0, 1)$).

Modelo Normal

- Gráfico da f.d.p.



  - $E(X) = \mu$: representa o ponto de simetria de f_X.
  - $Var(X) = \sigma^2$: representa a dispersão de f_X.

Modelo Normal

- **Afirmação:** $X \sim N(\mu, \sigma^2)$ se, e somente se $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$.

  Prova: Sabemos que $f_X(x) = \dfrac{d}{dx} F_X(x)$ (a volta fica como exercício).

  Por outro lado, temos que:

$$
\begin{aligned}
F_Z(z) &= P(Z \leq z) = P\left(\frac{X-\mu}{\sigma} \leq z\right) = P(X \leq z\sigma + \mu) \\
&= F_X(z\sigma + \mu) \rightarrow f_Z(z) = \frac{d}{dz} F_X(z\sigma + \mu) = f_x(z\sigma + \mu)\sigma \\
&= \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z^2}{2}\right)}.
\end{aligned}
$$

Modelo Normal

- Φ denota a f.d.a. da Normal padrão:

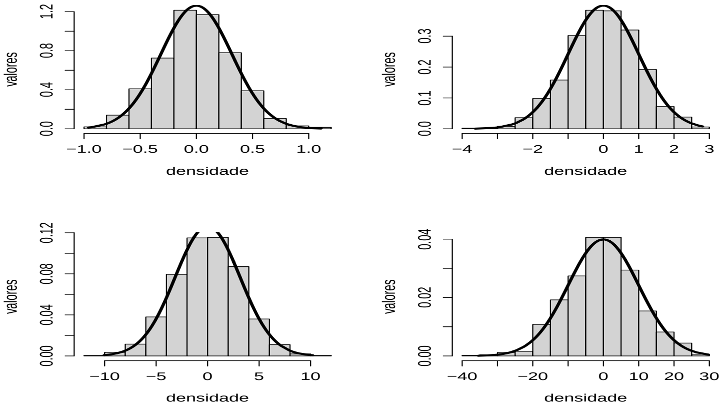
$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x^2}{2}\right)} dx.$$

(note que $\nexists\ G,\ \frac{d}{dx} G = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x^2}{2}\right)}$.)

- Temos, que se $X \sim N(\mu, \sigma^2)$, então:

$$F_X(a) = P(X \leq a) = P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right).$$

fdp (linha) e valores simulados (histograma): normal



no R

- Densidade: dnorm(x, mean = mu, sd = sigma).
- Fda: pnorm(x, mean = mu, sd = sigma).
- Simulação: rnorm(n, mean = mu, sd = sigma) (mais detalhes veja aqui).

Densidade, esperança e variância

- Exercício: Provar que $f_X(x) > 0, \forall x$.
- Resultados úteis: Além disso, seja $f(.)$ uma função real, então:
  - $\int_{-a}^{a} f(x)dx = 0$, se $f(.)$ for uma função ímpar $(f(x) = -f(-x))$.
  - $\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx$ se $f(.)$ for uma função par $(f(x) = f(-x))$.

Densidade, esperança e variância

- Temos que $(z = \dfrac{x-\mu}{\sigma}, x = z\sigma + \mu, dx = \sigma dz,\ y = z^2/2, z = \sqrt{2y}, dy = zdz$, e $f_X(.)$ é uma função par):

$$
\begin{aligned}
\int_{\mathcal{R}} f(x)dx &= \int_{\mathcal{R}} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \\
&= 2\int_0^\infty \frac{1}{\sqrt{2\pi}} exp\left[-\frac{z^2}{2}\right] dz \\
&= \frac{2}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{2}} y^{-1/2} e^{-y} dy = \frac{1}{\sqrt{\pi}} \underbrace{\Gamma(1/2)}_{\sqrt{\pi}} = 1.
\end{aligned}
$$

Densidade, esperança e variância

- $\mathcal{E}(X)$ e $\mathcal{V}(X)$. Note que se $X = \sigma Z + \mu$, em que $Z \sim N(0,1)$, então

$$\mathcal{E}(Z) = \int_{\mathcal{R}} zf(z)dz = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z^2}{2}\right)} dz = 0,$$

pois $z \times f(z)$ é uma função ímpar.

Densidade, esperança e variância

- Além disso ($y = z^2/2, y = \sqrt{2\pi}, dy = zdz$ e $z^2 f(z)$ é uma função par),

$$\begin{aligned}
\mathcal{E}(Z^2) &= \int_{\mathcal{R}} z^2 f(z) dz = \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z^2}{2}\right)} dz \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \frac{2}{\sqrt{2}} y^{1/2} e^{-y} dy \\
&= \frac{2}{\sqrt{\pi}} \int_0^{\infty} y^{3/2-1} e^{-y} dy = \frac{2}{\sqrt{\pi}} \frac{1}{2} \Gamma(1/2) = 1
\end{aligned}$$

- Assim, como $\mathcal{E}(Z) = 0$, temos que $\mathcal{V}(X) = \mathcal{E}(X^2) = 1$.
- Portanto, $\mathcal{E}(X) = \sigma\mathcal{E}(Z) + \mu = \mu$ e $\mathcal{V}(X) = \sigma^2 V(Z) = \sigma^2$.

Modelo Normal

- Exercitando com a tabela da Normal:
  - $\Phi(0,2) = 0,5793$ (no R, pnorm(0.2, mean=0, sd=1))
  - $\Phi(0,45) = 0,6736$ (no R, pnorm(0.45, mean=0, sd=1))
  - $\Phi(1,98) = 0,9761$ (no R, pnorm(1.98, mean=0, sd=1))
  - $\Phi(-0,45) = 1 - \Phi(0,45) = 0,3264$ (no R, pnorm(-0.45, mean=0, sd=1))

Modelo Normal

Exemplos

- Sejam $Z \sim N(0,1)$ e $Y \sim N(4, 2^2)$.
  - $F_Y(6) = P(Y \leq 6) = P\left(\frac{Y-4}{2} \leq \frac{6-4}{2}\right) = P(Z \leq 1) = \Phi(1) = 0,8413$ (no R, pnorm(6,mean=4,sd=2)).
  - $P(2 < Y \leq 6) = P\left(\frac{2-4}{2} < \frac{Y-4}{2} \leq \frac{6-4}{2}\right) = P(-1 < Z \leq 1) = \Phi(1) - \Phi(-1) = \Phi(1) - [1 - \Phi(1)] = 2\Phi(1) - 1 = 2(0,8413) - 1 = 0,6826$ (no R, pnorm(6,mean=4,sd=2) - (pnorm(2,mean=4,sd=2))).

Modelo Normal

- Exemplo: As alturas dos indivíduos de uma população têm distribuição Normal com média $\mu = 170cm$ e desvio padrão $\sigma = 5cm$, ou seja, $X \sim N(170, 5^2)$:
- $P(X \leq 182) = P(Z \leq 2,4)$.
- $P(X \geq 167) =$ 1- P(Z ¡ -0,6).
- $P(165 \leq X \leq 178) = P(Z \leq 1,6) - P(Z \leq -1)$.
- $P(X > x) = 0,8754 \rightarrow 1 - P(Z \leq (x - 170)/5) = 0,1246$.
- No R pnorm(182,170,5) $=$ 0.9918, 1-pnorm(167,170,5) $=$ 0.7257, pnorm(178,170,5)-pnorm(165,170,5) $=$ 0.7865 e qnorm(0.1246,170,5) $=$ 164,24.

Modelo Qui-Quadrado

- Dizemos que uma v.a X tem distribuição qui-quadrado com k ($k > 0$) graus de liberdade, se sua f.d.p, f_X, é dada por:

$$f_X(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} \mathbb{1}_{(0,\infty)}(x).$$

- Notação: $X \sim \chi^2_{(k)}$.
- Exercício: provar que, de fato, $f_X(.)$ é uma densidade.
- Assim como no caso da Normal, a distribuição de qui-quadrado não apresenta forma analítica para sua fda. Assim, recorremos ao uso de uma tabela apropriada.

Modelo Qui-Quadrado

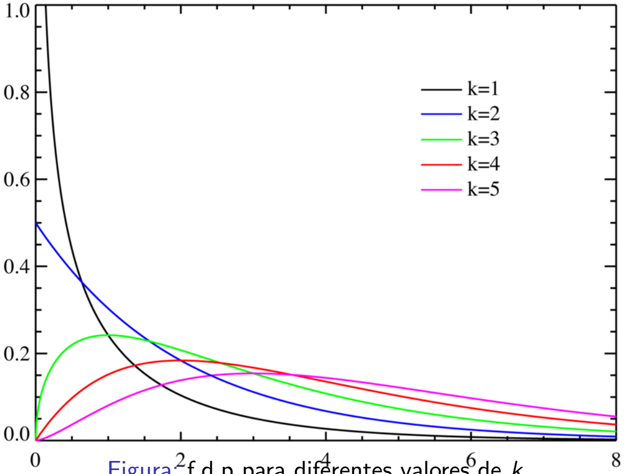
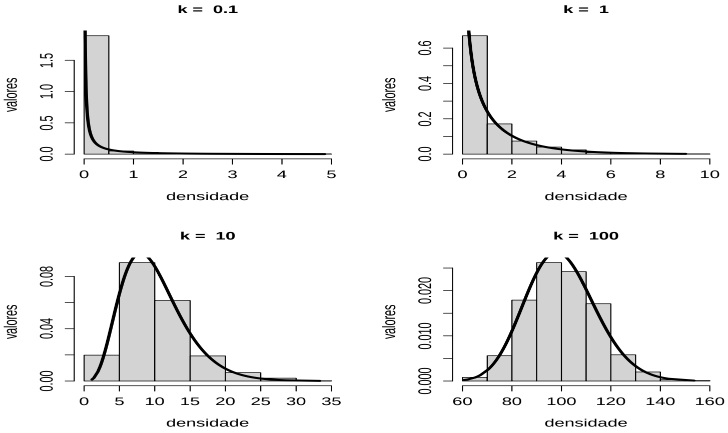


Figura: f.d.p para diferentes valores de k.

fdp (linha) e valores simulados (histograma): qui-quadrado



no R

- Densidade: dchisq(x, df = k).
- Fda: pchisq(x, df = k).
- Simulação: rchisq(n, df = k) (mais detalhes veja aqui).

Modelo Qui-Quadrado

- Esperança e variância. Temos que $(y = x/2, x = 2y, 2dy = dx)$:

$$
\begin{aligned}
E(X^r) &= \int_0^\infty \frac{1}{2^{k/2}\Gamma(k/2)} x^{r+(k/2)-1} e^{-x/2} dx \\
&= \frac{2}{2^{k/2}\Gamma(k/2)} \int_0^\infty (2y)^{r+(k/2)-1} e^{-y} dy \\
&= \frac{2^r}{\Gamma(k/2)} \int_0^\infty y^{r+(k/2)-1} e^{-y} dy = 2^r \frac{\Gamma(r+k/2)}{\Gamma(k/2)}.
\end{aligned}
$$

- Então, $E(X) = 2\frac{\Gamma(1+k/2)}{\Gamma(k/2)} = 2(k/2)\frac{\Gamma(k/2)}{\Gamma(k/2)} = k.$

Modelo Qui-Quadrado

- Além disso,

$$E(X^2) = 2^2 \frac{\Gamma(2 + k/2)}{\Gamma(k/2)} = 4(1 + k/2)(k/2)\frac{\Gamma(k/2)}{\Gamma(k/2)} = (2 + k)k.$$

- Portanto $V(X) = E(X^2) - E^2(X) = 2k + k^2 - k^2 = 2k$. A fda não possui forma analítica (é preciso utilizar uma tabela apropriada).
- **Propriedades:**
  - Se $X \sim N(0, 1)$, então $X^2 \sim \chi^2_{(1)}$. (exercício)
  - Se $X_i \overset{iid}{\sim} N(0, 1)$, então $\sum_{i=1}^{n} X_i^2 \sim \chi^2_{(n)}$. (provaremos adiante, no curso)

Modelo Qui-Quadrado

- Exercitando com a tabela da qui-quadrado:
  - Se $X \sim \chi^2_{(10)}$, $P(X > 2,558) = 0,99$ (1-pchisq(2.558,df=10)) e $P(X > 18,307) = 0,05$ (1-pchisq(18.307,df=10))
  - Se $X \sim \chi^2_{(30)}$, $P(X > 40,256) = 0,10$ (1-pchisq(40.256,df=30))
  - Se $X \sim \chi^2_{(5)}$, qual o valor de x_0 cuja $P(X \leq x_0) = 0,975$?

    R. $x_0 = 12,833$ (qchisq(0.975,df=5))
  - Se $X \sim \chi^2_{(3)}$, qual o valor de x_0 cuja $P(X < x_0) = 0,95$?

    R. $x_0 = 7,815$ (qchisq(0.95,df=3))

Modelo t-Student

- Dizemos que uma v.a X tem distribuição t-Student com $\nu, (\nu > 0)$ graus de liberdade, se sua f.d.p, f_X, é dada por:

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{(\nu+1)/2} \mathbb{1}_{(-\infty,\infty)}(x).$$

- Notação: $X \sim t_{(\nu)}$.
- Origem. Seja $Y \sim N(0,1)$ e $W \sim \chi^2_\nu$, $Y \perp W$ (independentes), então $X = \dfrac{Y}{\sqrt{W/k}}$.

Modelo t-Student

- Pesquise a respeito de como provar que f_X é uma legítima fdp.
- Assim como no caso da Normal, a distribuição de t de Student não apresenta forma analítica para sua fda. Assim, recorremos ao uso de uma tabela apropriada.

Modelo t-Student

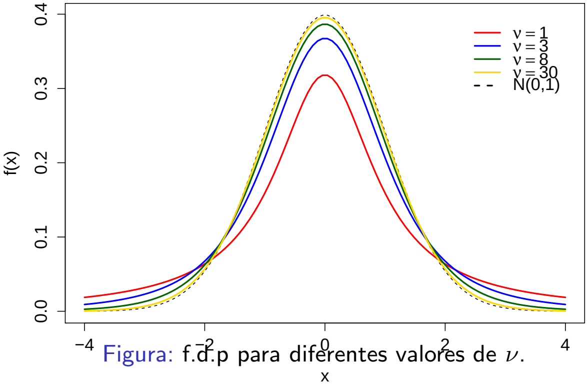
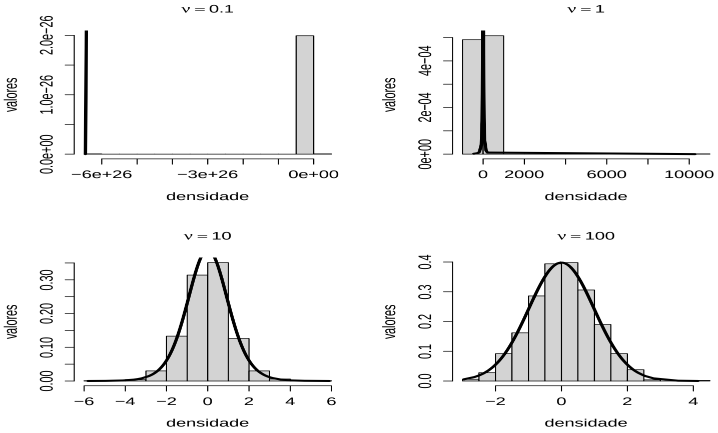


Figura: f.d.p para diferentes valores de ν.

fdp (linha) e valores simulados (histograma): t de Student



no R

- Densidade: dt(x, df $= \nu$).
- Fda: pt(x, df $= \nu$).
- Simulação: rt(n, df $= \nu$) (mais detalhes veja aqui).

Modelo t-Student

- Valor esperado:

$$E(X) = E\left(\frac{Y}{\sqrt{W/\nu}}\right) = \sqrt{\nu}E(Y)E(W^{-1/2}) = 0.$$

- Variância: $E(X^2) = \nu E(Y^2)E(W^{-1}) = \nu \times 1 \times \frac{2}{\nu - 2}$. Logo

$$Var(X) = E(X^2) - E^2(X) = \frac{\nu}{\nu - 2}, \nu > 2.$$

- OBS: para calcular $E(W^{-1/2})$ e $E(W^{-1})$ use a fórmula da $E(X^r)$, em que $X \sim \chi^2_{(\nu)}$.

Modelo t-Student

- Exercitando com a tabela da t de Student:

  - Se $X \sim t_{(6)}$, $P(X > 2,447) = 0,025$ (1-pt(2.447,df=6)) e $P(-1,943 < X < 1,943) = 0,90$ (pt(1.943,df=6)- pt(-1.943,df=6))

  - Se $X \sim t_{(11)}$, qual o valor de x_0 cuja $P(X < x_0) = 0,75$?

    R. $x_0 = 0,697$ (qt(0.75,df=11))

  - Se $X \sim t_{(20)}$, qual o valor de x_0 cuja $P(X > x_0) = 0,10$?

    R. $x_0 = 1,325$ (qt(0.90,df=20))

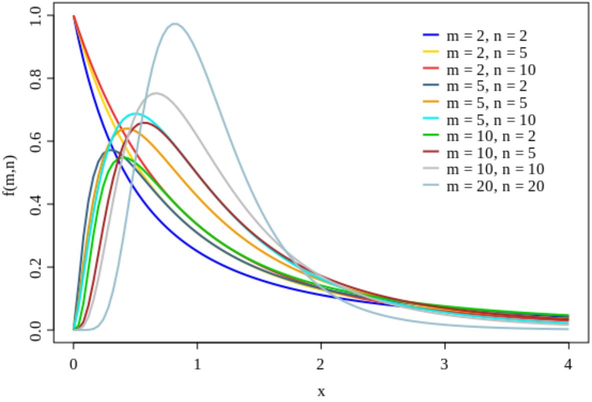
Modelo F de Snedecor

- Dizemos que uma v.a X tem distribuição F de Snedecor com m e n graus de liberdade se sua f.d.p, f_X, é dada por:

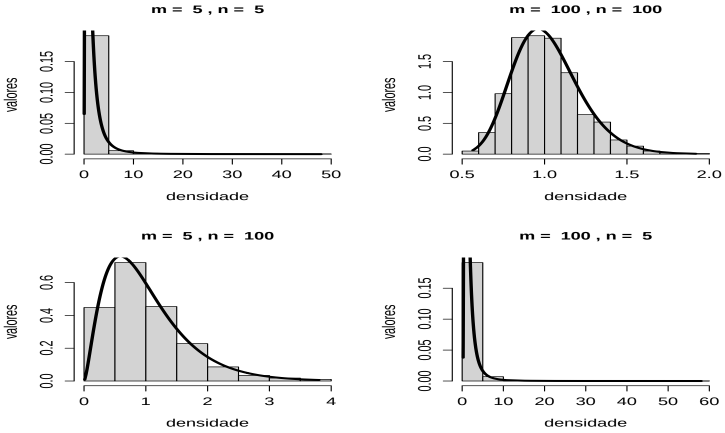
$$f_X(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\left[\left(\frac{m}{n}\right)x+1\right]^{\frac{m+n}{2}}} \mathbb{1}_{(0,\infty)}(x)(m, n > 0).$$

- Notação: $X \sim F(m, n)$.
- Teorema: Sejam $Y \sim \chi^2_m$ e $W \sim \chi^2_n$, $Y \perp W$, então $X = \dfrac{Y/m}{W/n}$.
- Assim como no caso da Normal, a distribuição F não apresenta forma analítica para sua fda. Assim, recorremos ao uso de uma tabela apropriada (aqui e aqui).

Modelo F de Snedecor



fdp (linha) e valores simulados (histograma): F de Snedcor



Modelo F de Snedecor

- Para o valor esperado e variância, utilizando $\mathcal{E}(X^r)$, $X \sim \chi^2_{(k)}$.
- Valor esperado:

$$E(X) = \frac{n}{m}E(Y)E(W^{-1}) = \frac{n}{m}m\frac{1}{n-2} = \frac{n}{n-2}, n > 2.$$

- Para a variância:

$$\begin{aligned} E(X^2) &= \frac{n^2}{m^2}E(Y^2)E(W^{-2}) = \frac{n^2}{m^2}(2+m)m\frac{1}{(n-2)(n-4)} \\ &= \frac{n^2}{m}\frac{(2+m)}{(n-2)(n-4)}, n > 4. \end{aligned}$$

Modelo F de Snedecor

- Assim:

$$V(X) = \mathcal{E}(X^2) - \mathcal{E}^S(X) = \frac{n^2}{m}\frac{(2+m)}{(n-2)(n-4)} - \frac{n^2}{(n-2)^2}$$
$$= \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, n > 4.$$

- **Observação:** Em geral, as tabelas (aqui e aqui) contém apenas os quantis da cauda superior ($F(\alpha, m, n)$, $P(X \leq F(\alpha, m, n)) = \alpha$, $X \sim F(\alpha, m, n)$, para $\alpha \geq 0,90$). Os quantis da cauda inferior $F(1-\alpha, m, n)$ podem ser obtidos a partir da seguinte relação:

$$F(1-\alpha, m, n) = \frac{1}{F(\alpha, n, m)}.$$

Modelo F de Snedecor

- Exercitando com a tabela da F de Snedecor:
  - Se $X \sim F(5,7)$, $P(X > 3,97) = 0,05$ (no R, 1-pf(3.97,df1=5, df2=7)) ou então $P(X \leq 3,97) = 0,95$ (no R, pf(3,97;5;7)).
  - Se $X \sim F(3,8)$, $P(X < 7,59) = 0,99$ (no R, pf(7,59;3;8)).
  - Se $X \sim F(5,7)$, qual o valor de f_0 cuja $P(X < f_0) = 0,05$?

    R. $0,05 = P(F(5,7) < f_0) = P[1/F(7,5) < f_0] = P[F(7,5) > 1/f_0]$. Consultando a tabela, obtemos que $1/f_0 = 4,88$ e, portanto, $f_0 = 0,205$ (no R, qf(0,05;5;7)).