1 Distribução t de Student-Prof. Maurício-2021.2

Definição. Dizemos que a v.a. X tem uma distribuição t-Student se sua fdp é da forma

$$f_X(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} \; I_{\mathbb{R}}\,(x). \tag{1}$$

O parâmetro k é chamado "grau de liberdade ".

Notação: $X \sim\ t(k)$

Como $\Gamma(1/2) = \sqrt{\pi}$ a f.d.p. de X pode aparecer na forma:

$$f_X(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\ \Gamma(\frac{1}{2})\ \sqrt{k}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} \; I_{\mathbb{R}}\,(x). \tag{2}$$

Ou em termos da função beta:

$$f_X(x) = \frac{1}{\mathbb{B}\left(\frac{k}{2},\frac{1}{2}\right)\ \sqrt{k}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} \; I_{\mathbb{R}}\,(x). \tag{3}$$

Na Figura 1, apresentamos o gráfico para a função densidade de probabilidade de (1) para certos valores de k.

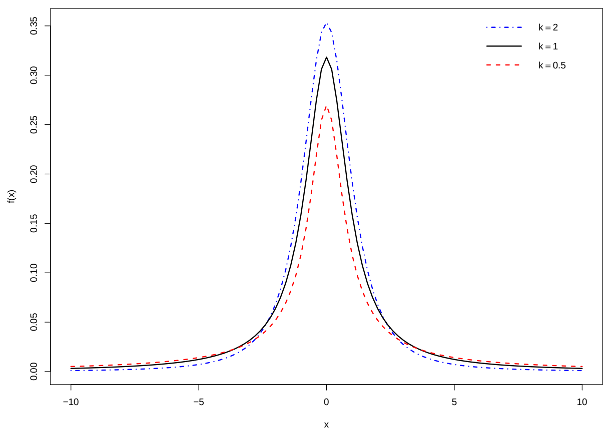

Figura 1: Gráfico da função densidade t-Student

Observação 1. Lê-se a notação acima do seguinte modo: X *segue distribuição t- Student com k graus de liberdade.*

Observação 2. Quando $k = 1$, temos

$$f(x) = \frac{\Gamma(1)}{\Gamma\left(\frac{1}{2}\right)\sqrt{\pi}} \cdot \frac{1}{(1+x^2)^1} = \frac{1}{\pi(1+x^2)} \quad I_{\mathbf{R}}(x)$$

e portando, X tem distribuição Cauchy padrão, e consequentemente, $E(X)$ não existe.

1.1 Fato 1

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Prova:

$$I = \int_{-\infty}^{\infty} \frac{dx}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} = 2\int_0^{\infty} \left(1+\frac{x^2}{k}\right)^{\frac{-(k+1)}{2}} dx$$

fazendo

$$u = \frac{x^2}{k} \Rightarrow x^2 = ku \Rightarrow x = \sqrt{ku} \Rightarrow dx = \frac{\sqrt{k}}{2\sqrt{u}}\, du$$

assim

$$\begin{aligned} I &= \int_0^{\infty} (1+u)^{\frac{-(k+1)}{2}} \frac{\sqrt{k}}{2\sqrt{u}}\, du \\ &= \sqrt{k} \int_0^{\infty} u^{-\frac{1}{2}} (1+u)^{\frac{-(k+1)}{2}}\, du \end{aligned}$$

Sabemos que

$$\int_0^{\infty} \frac{x^{a-1}}{(1+x)^{a+b}} dx = I\!B(a,b)$$

então

$$\int_0^{\infty} \frac{u^{\frac{1}{2}-1}}{(1+u)^{\frac{k+1}{2}}} du = I\!B\left(\frac{1}{2}, \frac{k}{2}\right)$$

Analisando agora a parte constante de $f(x)$ temos

$$\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} = \frac{1}{\sqrt{k}}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)} = \frac{1}{\sqrt{k}\cdot I\!B\left(\frac{1}{2},\frac{k}{2}\right)}$$

Assim

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \frac{1}{\sqrt{k}\cdot I\!B\left(\frac{1}{2},\frac{k}{2}\right)}\underbrace{\int_{-\infty}^{\infty}\left(1+\frac{x^2}{k}\right)^{\frac{-(k+1)}{2}}}_{I}\\ &= \frac{1}{\sqrt{k}\cdot I\!B\left(\frac{1}{2},\frac{k}{2}\right)}\cdot\sqrt{k}\cdot I\!B\left(\frac{1}{2},\frac{k}{2}\right) = 1\end{aligned}$$

1.2 Aparecimento

A distribuição T aparece na teoria estatística como a razão entre duas variáveis aleatórias independentes

$$Z \sim N(0,1) \quad e \quad V \sim \chi^2(k),$$

$$X = \frac{Z}{\sqrt{\frac{V}{k}}}.$$

O parâmetro k é chamado de graus de liberdade.

Vamos obter a densidade de X através do teorema da mudança de variáveis.

Prova:

A conjunta de (Z, V) é dada por:

$$g(z,v) = \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}\ \frac{1}{\Gamma\left(\frac{k}{2}\right)\ 2^{k/2}}\ \ v^{k/2-1}\ \ e^{-v/2}\ \ I_{\mathbb{R}}\ (z)\ I_{\mathbb{R}^+}(v),$$

que pode ser posta na forma:

$$g(z,v) = \frac{1}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)\ 2^{(k+1)/2}}\ \ v^{k/2-1}\ \ e^{-(v+z^2)/2}\ \ I_{\mathbb{R}}\ (z)\ I_{\mathbb{R}^+}(v),$$

Considere a variável auxiliar $A = \sqrt{V}$. Vamos obter a conjunta de:

$$X = \frac{Z}{\sqrt{\frac{V}{k}}} = \frac{\sqrt{k}\,Z}{\sqrt{V}} \quad e \quad A = \sqrt{V}.$$

Assim as funções inversas são dadas por:

$$w_2(x,a) = v = a^2, \;\; e \;\; w_1(x,a) = \frac{ax}{\sqrt{k}}.$$

O jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial w_1}{\partial x} & \frac{\partial w_1}{\partial a} \\ \frac{\partial w_2}{\partial x} & \frac{\partial w_2}{\partial a} \end{vmatrix} = \begin{vmatrix} \frac{a}{\sqrt{k}} & \frac{x}{\sqrt{k}} \\ 0 & 2a \end{vmatrix} = \frac{2a^2}{\sqrt{k}}.$$

A densidade de $X = h_1(Z,V)$ e $A = h_2(Z,V)$ com suporte B é dada por:

$$g(x,a) = f(w_1(x,a), w_2(x,a)) \;\; |J| \;\; I_B(x,a).$$

Logo,

$$g(x,a) = \frac{1}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)\; 2^{(k+1)/2}} \;\; (a^2)^{k/2-1} \;\; e^{-(a^2+\frac{a^2x^2}{k})/2} \;\; \frac{2a^2}{\sqrt{k}} \;\; I_{\mathbb{R}}\,(\frac{ax}{\sqrt{k}})\; I_{\mathbb{R}^+}(a^2),$$

que pode ser posta na forma:

$$g(x,a) = \frac{1}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)\; 2^{(k-1)/2}\sqrt{k}} \;\; a^k \;\; e^{-\frac{1}{2}\;\left(1+\frac{x^2}{k}\right)a^2} \;\; I_{\mathbb{R}}\;(x)\; I_{\mathbb{R}^+}(a),$$

A marginal de X é dada por:

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} g(x,a)\;\; da \\ &= c \int_{0}^{\infty} \;\; a^k \;\; e^{-\frac{1}{2}\;\left(1+\frac{x^2}{k}\right)a^2} \; da \\ &= cI, \end{aligned}$$

onde

$$c = \frac{1}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)\; 2^{(k-1)/2}\;\sqrt{k}}.$$

Por outro lado

$$
\begin{aligned}
I &= \int_0^\infty a^k \; e^{-\frac{1}{2}\left(1+\frac{x^2}{k}\right)a^2} \, da \\
&= IGG(k+1, \frac{1}{2}\left(1+\frac{x^2}{k}\right)a^2, 2) \\
&= \frac{\Gamma\left(\frac{k+1}{2}\right)}{\frac{2\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}}{2^{\frac{k+1}{2}}}} \\
&= \frac{\Gamma\left(\frac{k+1}{2}\right)}{\frac{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}}{2^{\frac{k-1}{2}}}}
\end{aligned}
$$

A densidade de X, $f(x) = cI$, é dada por:

$$
f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)\sqrt{k}} \; \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} \; I_{\mathbb{R}}\,(x).
$$

1.3 Função de Distribuição

A função de distribuição (fd) de $X \sim t_{(k)}$ é dada por

$$
F(x) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{1}{\left(1+\frac{t^2}{k}\right)^{\frac{k+1}{2}}} \, dt
$$

Na Figura 2, apresentamos o gráfico da função de distribuição t-Student para certos valores de k.

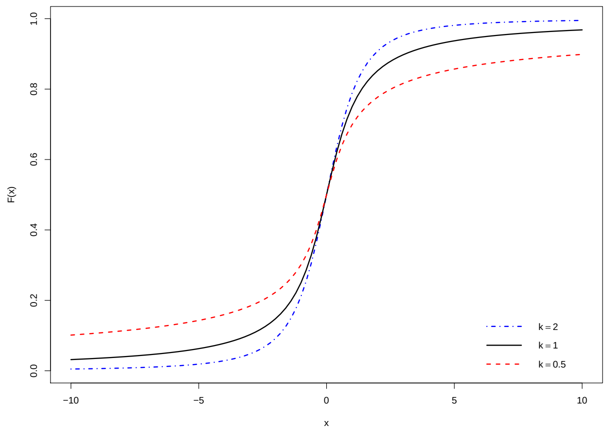

Figura 2: Gráfico da função de distribuição t-Student

Esta função vem tabelada em alguns livros de Estatística, para alguns valores de k e valores de $p = P(X \leq t_0)$. Vamos utilizar o R para se calcular as seguintes probabilidades

a) $P[t_{(8)} \leq 3,36]$

```
> pt(3.36,8)
[1] 0.9950341
```

b) $P[t_{(10)} \geq 2,90]$

```
> pt(2.9,10,lower.tail=FALSE)
[1] 0.0079168
```

c) $P[0,159 \leq t_{(27)} \leq 0,263]$

```
> pt(0.263,27)-pt(0.159,27)
[1] 0.04015254
```

um outro resultado importante que podemos obter no R é o valor de t_0 em que $P(t_{(10)} \leq t_0) = 0,95$, da forma

```
> t0=qt(0.95,10);t0
[1] 1.812461
```

Seja também $P(t_{(k)} \leq t_{(k)}) = 0,975$, analisando essa probabilidade observamos que a medida que k cresce $t_{(k)}$ decresce de 12,71 para 1,96 que é o valor da normal padrão de modo que

$$P(Z \leq 1,96) = 0,975$$

essa probabilidade é muito importante na estatística aplicada, pois costuma-se aproximar-se a distribuição t de Student para a distribuição normal padrão sempre que $k \geq 30$, isto é,

$$P(t_{(k)} \leq t_0) \approx P(Z \leq t_0) = \Phi(t_0)$$

Esta propriedade pode ser verificada diretamente, ou seja, seja $X \sim t_{(k)}$ então

$$\lim_{k\to\infty} f(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

Prova

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}}$$

vamos analisar os seguintes limites

$$\lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}} = \lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{-\frac{k}{2}} \cdot \lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{-\frac{1}{2}}$$

mas

$$\lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{-\frac{1}{2}} = \left[\lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)\right]^{-\frac{1}{2}} = 1^{-\frac{1}{2}} = 1$$

por outro lado, sabemos que

$$\lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{k} = e^{x^2}$$

assim

$$\lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{-\frac{k}{2}} = e^{-\frac{x^2}{2}}$$

este limite é um caso especial do seguinte limite

$$\lim_{n\to\infty}\left[1+\frac{b}{n}+\frac{\psi(n)}{n}\right]^{cn}=e^{bc}$$

em que, b e c não dependem de n e $\lim\limits_{n\to\infty}\psi(n)=0$. Na aplicação acima temos: $b=x^2$, $c=-\frac{1}{2}$, $\psi(n)=0$ e assim $e^{bc}=e^{-\frac{x^2}{2}}$. Para o outro limite vamos usar a aproximação

$$\Gamma(k+1)\approx(2\pi)^{\frac{1}{2}}\ k^{\frac{2k+1}{2}}\ e^{-\frac{k}{2}}\quad\text{(fórmula de Stirling).}$$

Como

$$\begin{aligned}\Gamma\left(\frac{k+1}{2}\right) &= \Gamma\left(\frac{k-1}{2}+1\right)\\ &\approx (2\pi)^{\frac{1}{2}}\left(\frac{k-1}{2}\right)^{\frac{k}{2}}\ e^{-\frac{k-1}{2}}\\ &= (2\pi)^{\frac{1}{2}}\ 2^{-\frac{k}{2}}\,((k-1))^{\frac{k}{2}}\quad e^{-\frac{k}{2}}.\end{aligned}$$

Similarmente,

$$\begin{aligned}\Gamma\left(\frac{k}{2}\right) &= \Gamma\left(\frac{k-2}{2}+1\right)\\ &\approx (2\pi)^{\frac{1}{2}}\left(\frac{k-2}{2}\right)^{\frac{k-1}{2}}\ e^{-\frac{k-2}{2}}\\ &= (2\pi)^{\frac{1}{2}}\ 2^{-\frac{k-1}{2}}\,((k-2))^{\frac{k-1}{2}}\quad e^{-\frac{k-2}{2}}.\end{aligned}$$

assim

$$
\begin{aligned}
C(k) &= \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k}} \\
&\approx \frac{(2\pi)^{\frac{1}{2}}\ 2^{-\frac{k}{2}}\,(k-1)^{\frac{k}{2}}\quad e^{-\frac{k-1}{2}}}{(2\pi)^{\frac{1}{2}}\ 2^{-\frac{k-1}{2}}\,(k-2)^{\frac{k-1}{2}}\quad e^{-\frac{k-2}{2}}\sqrt{k}} \\
&= \frac{2^{-\frac{k}{2}}\,(k-1)^{\frac{k}{2}}\quad e^{-\frac{k}{2}}\ \ e^{\frac{1}{2}}}{2^{-\frac{k}{2}}2^{\frac{1}{2}}\,(k-2)^{\frac{k}{2}}\,(k-2)^{-\frac{1}{2}}\quad e^{-\frac{k}{2}}\quad e^{1}\,\sqrt{k}} \\
&= \frac{(k-1)^{\frac{k}{2}}\quad e^{-\frac{1}{2}}}{\sqrt{2}\,(k-2)^{\frac{k}{2}}\,(k-2)^{-\frac{1}{2}}\ \ \sqrt{k}} \\
&= \frac{e^{-\frac{1}{2}}}{\sqrt{2}}\ \sqrt{\frac{k-2}{k}}\ \left(\frac{k-1}{k-2}\right)^{\frac{k}{2}} \\
&= \frac{e^{-\frac{1}{2}}}{\sqrt{2}}\ \sqrt{\frac{1-\frac{2}{k}}{1}}\ \left(\frac{k-1}{k-2}\right)^{\frac{k}{2}}
\end{aligned}
$$

portanto

$$
\begin{aligned}
\lim_{k\to\infty} C(k) &= \lim_{k\to\infty}\frac{e^{-\frac{1}{2}}}{\sqrt{2}}\ \sqrt{\frac{1-\frac{2}{k}}{1}}\ \left(\frac{k-1}{k-2}\right)^{\frac{k}{2}} \\
&= \frac{e^{-\frac{1}{2}}}{\sqrt{2}}\ \sqrt{\lim_{k\to\infty}\left(\frac{1-\frac{2}{k}}{1}\right)}\ \ \lim_{k\to\infty}\left(\frac{k-1}{k-2}\right)^{\frac{k}{2}} \\
&= \frac{e^{-\frac{1}{2}}}{\sqrt{2}}\ \sqrt{1}\ \ \lim_{k\to\infty}\left(\frac{k-2+1}{k-2}\right)^{\frac{k}{2}} \\
&= \frac{e^{-\frac{1}{2}}}{\sqrt{2}}\ \lim_{k\to\infty}\left(1+\frac{1}{k-2}\right)^{\frac{k}{2}}
\end{aligned}
$$

Fazendo a mudança $j = k-2$, temos $k = j+2$. Portanto,

$$
\begin{aligned}
\lim_{k\to\infty}\left(1+\frac{1}{k-2}\right)^{\frac{k}{2}} &= \lim_{j\to\infty}\left(1+\frac{1}{j}\right)^{\frac{j+2}{2}}\\
&= \lim_{j\to\infty}\left(1+\frac{1}{j}\right)^{\frac{j}{2}} \times \lim_{j\to\infty}\left(1+\frac{1}{j}\right)^{\frac{2}{2}}\\
&= e^{\frac{1}{2}} \times 1 = e^{\frac{1}{2}}.
\end{aligned}
$$

logo

$$
\lim_{k\to\infty} C(k) = \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{\frac{1}{2}} = \frac{1}{\sqrt{2\pi}}
$$

assim

$$
\lim_{k\to\infty} f_X(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}
$$

conhecida como a função densidade de probabilidade de distribuição normal padrão.

1.4 Momentos, Assimetria e Curtose

O r-ésimo momento em relação à origem de $X \sim t_{(k)},\ k > 1$ é dado por

$$
\mathbb{E}(X^r) = \begin{cases} 0 & \text{, se } r \text{ for ímpar,} r < k;\\[2ex] \frac{k^{\frac{r}{2}}\Gamma(\frac{r+1}{2})\Gamma(\frac{k-r}{2})}{\Gamma(\frac{k}{2})\sqrt{\pi}} & \text{, se } r \text{ for par, } r < k. \end{cases}
$$

Prova

$$
\mathbb{E}(X^r) = \int_{-\infty}^{\infty} x^r \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}}\ dx
$$

se r for um número ímpar temos que o valor de $\mathbb{E}(X^r)$ será nulo, pois obteremos uma função ímpar. Agora para o caso em que r é par obtemos

$$\mathbb{E}(X^r) = 2\int_0^\infty x^r \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} dx$$

fazendo $c = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}}$ e fazendo $u = \frac{x^2}{k} \Rightarrow x^2 = ku$
e $\Rightarrow x = \sqrt{ku} \Rightarrow dx = \sqrt{k}\, u^{-1/2}\frac{1}{2}\, du$
temos

$$\begin{aligned}
\mathbb{E}(X^r) &= 2c\int_0^\infty \frac{(uk)^{\frac{r}{2}}}{(1+u)^{\frac{k+1}{2}}} \cdot \frac{k}{\sqrt{uk}}\,\frac{1}{2};\ du \\
&= c\, k^{\frac{r+1}{2}} \int_0^\infty \frac{u^{\frac{r+1}{2}-1}}{(1+u)^{\frac{r+1}{2}+\frac{k-r}{2}}} du \\
&= c2c\, k^{\frac{r+1}{2}} \mathbb{B}\left(\frac{r+1}{2}, \frac{k-r}{2}\right) \\
&= \frac{\Gamma\left(\frac{k+1}{2}\right) k^{\frac{r+1}{2}}}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{\Gamma(\frac{r+1}{2})\Gamma(\frac{k-r}{2})}{\Gamma(\frac{k+1}{2})} \\
&= \frac{k^{\frac{r}{2}}\Gamma(\frac{r+1}{2})\Gamma(\frac{k-r}{2})}{\Gamma(\frac{k}{2})\sqrt{\pi}},
\end{aligned}$$

com $k - r > 0$ ou $r < k$.
Foi usado o fato:
Para $r = 1$ temos:

$$\mathbb{E}(X) = \mu = 0$$

Para $r = 2$ temos:

$$\begin{aligned}
\mathbb{E}(X^2) &= \frac{k^{\frac{2}{2}}\Gamma(\frac{3}{2})\Gamma(\frac{k-2}{2})}{\Gamma(\frac{k}{2})\sqrt{\pi}} \\
&= \frac{k\frac{\sqrt{\pi}}{2}}{\left(\frac{k}{2}-1\right)\sqrt{\pi}} \\
&= \frac{\frac{k}{2}}{\frac{k-2}{2}} \\
&= \frac{k}{k-2},\ k > 2
\end{aligned}$$

Para $r = 3$ temos:

$$\mathbb{E}(X^3) = 0, \quad k > 3$$

Para $r = 4$ temos:

$$\begin{aligned}
\mathbb{E}(X^4) &= \frac{k^{\frac{4}{2}}\Gamma(\frac{5}{2})\Gamma(\frac{k-4}{2})}{\Gamma(\frac{k}{2})\sqrt{\pi}} \\
&= \frac{k^{\frac{4}{2}}\ \frac{3}{2}\ \ \frac{1}{2}\ \Gamma(\frac{1}{2})\ \Gamma(\frac{k}{2}-2)}{\Gamma(\frac{k}{2})\sqrt{\pi}} \\
&= \frac{k^2\ \ \frac{3\sqrt{\pi}}{4}}{\left(\frac{k}{2}-1\right)\ \left(\frac{k}{2}-2\right)\sqrt{\pi}} \\
&= \frac{3k^2}{4\left(\frac{k}{2}-1\right)\ \left(\frac{k}{2}-2\right)} \\
&= \frac{3k^2}{(k-2)(k-4)}.\ k > 4
\end{aligned}$$

A variância de $X \sim t_{(k)}$ é dada por

$$Var(X) = \frac{k}{k-2},\ k > 2$$

Prova

$$\begin{aligned}
Var(X) &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \\
&= \frac{k}{k-2},\ k > 2
\end{aligned}$$

A assimetria de $X \sim t_{(k)}$ é dada por

$$\alpha_3 = 0, \quad k > 3$$

Prova

Sabemos que

$$\alpha_3 = \frac{\mu_3}{\sigma^3}$$

em que $\mu_3 = \mathbb{E}[(X-\mu)^3] = \mathbb{E}(X^3) = 0$.Logo

$$\alpha_3 \;=\; 0$$

O Coeficiente de curtose de $X \sim t_{(k)}$ é dada por

$$\alpha_4 = 3\,\frac{k-2}{k-4} = 3 + \frac{6}{k-4},\ k > 4$$

Prova

Sabemos que

$$\alpha_4 = \frac{\mu_4}{\sigma^4}$$

em que $\mu_4 = \mathbb{E}[(X-\mu)^4] = \mathbb{E}(X^4) = 3\frac{k^2}{(k-2)(k-4)},\ k > 4$. Calculando μ_4. Assim

$$\begin{aligned}
\alpha_4 &= 3\,\frac{\frac{k^2}{(k-2)(k-4)}}{\frac{k^2}{(k-2)^2}} \\
&= 3\,\frac{k-2}{k-4} \\
&= 3\,\frac{k-4+2}{k-4} \\
&= 3 + \frac{6}{k-4},\ k > 4.
\end{aligned}$$

A distribuição t-student com $k > 4$ graus de liberdade é sempre leptocúrtica.

1.5 Moda "M_o"

A moda de $X \sim t_{(k)}$ é dada por

$$M_o = 0$$

Prova

$$\begin{aligned}
f(x) &= \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} \\
g(x) &= \ln\Gamma\left(\frac{k+1}{2}\right) - \ln\Gamma\left(\frac{k}{2}\right) - \frac{1}{2}\ln(k\pi) - \frac{k+1}{2}\ln\left(1+\frac{x^2}{k}\right) \\
g'(x) &= -\frac{\frac{k+1}{2}}{1+\frac{x^2}{k}} \cdot \frac{2x}{k} \\
g'(x) &= -\frac{k+1}{k}\,\frac{x}{1+\frac{x^2}{k}} \\
g''(x) &= -\frac{k+1}{k}\,\frac{1-\frac{x^2}{k}}{1+\frac{x^2}{k}}
\end{aligned}$$

Assim, $g'\left(M_o\right) = 0$

$$-\frac{k+1}{k} \cdot \frac{Mo}{1+\frac{Mo^2}{k}} = 0$$

Portanto,

$$M_o = 0,$$

pois $g”(0) = -\frac{k+1}{k} < 0$.

1.6 Relações entre Distribuições

1.6.1 Transformação 1

Se $X \sim t_{(k)}$ então
$$Y = \frac{1}{1+\frac{X^2}{k}} \sim Beta(\frac{k}{2}, \frac{1}{2}).$$

Prova

$x \in I\!R \Rightarrow x^2 \geq 0 \Rightarrow \frac{x^2}{k} \geq 0.$

$1+\frac{x^2}{k} \geq 1 \Rightarrow \dfrac{1}{1+\frac{x^2}{k}} \leq 1 \Rightarrow 0 \leq y \leq 1.$ partindo de

$$\begin{aligned}
G(y) &= P(Y \leq y) = P\left(\frac{1}{1+\frac{X^2}{k}} \leq y\right) \\
&= P\left(1+\frac{X^2}{k} \leq \frac{1}{y}\right) = P\left(X^2 \geq \frac{k(1-y)}{y}\right) \\
&= P\left(|X| \geq \sqrt{\frac{k(1-y)}{y}}\right) = 2P\left(X \geq \sqrt{\frac{k(1-y)}{y}}\right) \\
&= 2\left[1 - F_X\left(\sqrt{k}\sqrt{\frac{1}{y}-1}\right)\right]
\end{aligned}$$

derivando a expressão acima obtemos

$$\begin{aligned}
g(y) &= 2\cdot(-1)\cdot\frac{\frac{-\sqrt{k}}{y^2}}{2\sqrt{\frac{1}{y}-1}} f\left(\sqrt{k}\sqrt{\frac{1}{y}-1}\right) \\
&= \frac{y^{-2}\sqrt{k}\sqrt{y}}{\sqrt{1-y}}\frac{y^{\frac{k+1}{2}}\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}} \\
&= \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}} y^{\frac{k}{2}-1}(1-y)^{\frac{-1}{2}} \underset{(0,1)}{I}(y)
\end{aligned}$$

logo $Y \sim Beta\left(\frac{k}{2}, \frac{1}{2}\right)$

1.6.2 Transformação 2

Se $X \sim t_{(k)}$ então $Y = X^2 \sim F(1,k)$.

Prova

Seja $y = x^2 \geq 0$.

A função de distribuição de Y é da forma

$$\begin{aligned} G(y) &= P(Y \leq y) = P(X^2 \leq y) \\ &= P(|X| \leq \sqrt{y}) = F(\sqrt{y}) - F(-\sqrt{y}) \end{aligned}$$

derivando a equação acima

$$\begin{aligned} g(y) &= \frac{1}{2\sqrt{y}}[f(\sqrt{y}) + f(-\sqrt{y})] = \frac{f(\sqrt{y})}{\sqrt{y}} \\ &= \frac{1}{\sqrt{y}} \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \frac{1}{\left(1+\frac{y}{k}\right)^{\frac{k+1}{2}}} \end{aligned}$$

assim a densidade de Y é da forma

$$g(y) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi}} k^{-\frac{1}{2}} y^{-\frac{1}{2}} \frac{1}{\left(1+\frac{y}{k}\right)^{\frac{k+1}{2}}} I_{(0,\infty)}(y)$$

logo $Y \sim F(1,k)$.

Outra Prova:

Sabemos que : Se $X \sim t_{(k)}$ então

$$X = \frac{Z}{\sqrt{\frac{V}{k}}},$$

com

$$Z \sim N(0,1) \;\; e \;\; V \sim \chi^2(k), \;\; \text{Z e V independentes.}$$

.

Assim,

$$X^2 = \frac{Z^2}{\frac{V}{k}} = \frac{\frac{Z^2}{1}}{\frac{V}{k}},$$

Com $Z^2 \sim \chi^2(1)$ e $V \sim \chi^2(k)$, Z^2 e V independentes.
Assim,

$$X^2 \sim F(1, k).$$

1.7 Aplicações

A distribuição t de Student têm várias aplicações em inferência estatística, tanto em teste de hipóteses e intervalos de confiança para as médias quanto em estimação de variâncias. Vamos tratar de algumas delas.

1.7.1 Uma População com distribuição Normal

Considere uma amostra aleatória de tamanho n de $X \sim \mathcal{N}(\mu, \sigma^2)$ e seja $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ e $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$, em que $S^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$ é um estimador não viesado para a variância e independente de $\bar{x}$. Assim

$$T = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\bar{X}-\mu}{S\sqrt{n}}$$

tem distribuição $t_{(n-1)}$, já que em população Normal $\bar{X}$ e S^2 são independentes.

1.7.2 Duas Populações Normais

Vamos considerar inicialmente duas populações normais independentes com mesma variância.

Sejam $X \sim N(\mu_1, \sigma^2)$ *e* $Y \sim N(\mu_2, \sigma^2)$.

Sejam $X_1, X_2, \cdots, X_n$ uma amostra aleatória de X e $Y_1, Y_2, \cdots, Y_n$ uma amostra aleatória de Y, independentes. Assim se fizermos

$$(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)$$

temos uma distribuição normal com média zero e variância $\sigma^2\left(\frac{1}{m} + \frac{1}{n}\right)$.

Para estimarmos a variância comum σ^2 usaremos:

$$S_p^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}$$

e como S_1^2 é independente de $\bar{X}$ e S_2^2 é independente de $\bar{Y}$. Além disso

$\bar{X}$ e $\bar{Y}$ são independentes bem como S_1^2 e S_2^2 também o são.

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

tem distribuição $t_{(m+n-2)}$. vamos estudar agora o caso de duas populações normais dependentes.

1.7.3 Dados Pareados

Agora (X, Y) uma variável aleatória normal bidimensional com parâmetros μ_1, μ_2, σ_1^2 ,σ_2^2 e ρ.

Considere a amostra aleatória $(X_i, Y_i), \forall i = 1, 2, \cdots, n$.

Faça $D_i = X_i - Y_i,\ i = 1, 2, \ldots, n$. Logo

$\mu_D = E(D_i) = E(X_i) - E(Y_i) = \mu_1 - \mu_2$,

$$\sigma_D^2 = V(D_i) = V(X_i) + V(Y_i) - 2cov(X_i, Y_i) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2.$$

Assim podemos pensar que $D_1, D_2, \ldots, D_n$ é uma amostra aleatória de tamanho n de $D \sim N(\mu_D, \sigma_D^2)$. Assim temos o caso de uma única população normal novamente. A diferença retira a dependência entre as populações.

Sejam $\bar{D} = \bar{X} - \bar{Y}$, a média amostral e S_D^2 a variância amostral de nossa amostra de diferenças.

A quantidade pivotal apropriada é

$$T = \frac{\bar{D} - \mu_D}{S_D} = \frac{\bar{D} - (\mu_1 - \mu_2)}{S_D},$$

tem distribuição t de Student com $(n - 1)$ graus de liberdade.

1.8 Intervalos de confiança

Em determinados intervalos de confiança e teste de hipóteses usando a distribuição t de Student definimos como

$$F(t_{\alpha,k}) = \int_{-\infty}^{t_{\alpha,k}} f(t; k)dt = 1 - \alpha$$

isto é, é a probabilidade que a variável com distribuição $t_{(k)}$ exceda $t_{\alpha,k}$. Observe que, como a distribuição é simétrica em torno so zero, temos que $t_{\alpha,k}{=}{-}t_{1-\alpha,k}$. No caso de uma amostra com distribuição normal descrita nas subseções anteriores e se adotarmos um nível de confiança de $1-\alpha$, o intervalo de confiança para μ é dado por

$$\bar{x} - \frac{s}{\sqrt{n}} t_{\alpha/2,k-1} \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}} t_{\alpha/2,k-1}$$

1.9 Teste de hipóteses

Supondo que as amostras são obtidas de uma distribuição normal, podemos utilizar as estatísticas de teste expostas anteriormente. Vamos exemplicar um possível teste, no caso em temos uma única amostra e desejamos, usando o teste bilateral, testar a hipótese $H_0 : \mu = \mu_0$ e a hipótese alternativa $H_1 : \mu \neq \mu_0$. Utilizamos, supondo H_0 verdade, a estatística $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ e rejeitaremos H_0 com um nível de significância α se $|t| > t_{\alpha/2,k-1}$. Se o teste unilateral, a hipótese alternativa é da forma $H_1 : \mu < \mu_0$ ou $H_1 : \mu > \mu_0$

1.9.1 Distribuição amostral

Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória de $\mathcal{N}(\mu, \sigma^2)$. Ilustrar no `R` o resultado que justifica o teste-t para a média de uma amostra,

$$\frac{\bar{x} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

em que S é o desvio padrão e n o tamanho da amostra.

i) Escolha os parâmetros de uma distribuição normal;

ii) Escolha o tamanho da amostra n e á quantidade de simulações N;

iii) Agora simule N amostras de tamanho n;

iv) Para cada valor da amostra calcule $V = \frac{\bar{x}-\mu}{S/\sqrt{n}}$

v) reproduza um histograma com os valores de V e compare com a curva da distribuição t_{n-1}.

Um exemplo para a simulação acima é

```
n=18
N=1000

y=matrix(rnorm(n*N,100,6),nc=N)
T=apply(y,2, function(x) {(mean(x)-100)*sqrt(n)/sd(x)});T
mean(T)
var(T)
hist(T, prob=TRUE, main="Histograma de T",ylab="f(t)",xlab="t")
curve(dt(x,n-1),add=T)
```

1.10 Exercícios

1. Se $X_k \sim t(k)$ calcule usando a tabela do Bussab & Morettin as seguintes probabilidades:

   a. $P(-3,365 \leq X_5 \leq 3,365)$.

   b. $P(|X_8| \leq 1,397)$.

   c. $P(-1,076 \leq X_{14} \leq 2,145)$.

   d. O valor de a tal que $P(X_9 > a) = 0,02$.

   e. O valor de b tal que $P(X_{16} \leq b) = 0,05$.

   f. O valor de c tal que $P(|X_{11}| \leq c) = 0,10)$.

   g. O valor de d tal que $P(|X_{21}| > d) = 0,05)$.

   h. O valor de b tal que $P(-b \leq X_{14} \leq b) = 0,90$.

2. Resolva a questão 1 usando o R: