Estatística Matemática

Amostra aleatória

Prof. Paulo Cerqueira Jr - cerqueirajr@ufpa.br
Faculdade de Estatística - FAEST
Programa de Pós-Graduação em Matemática e Estatística - PPGME
Instituto de Ciências Exatas e Naturais - ICEN

https://github.com/paulocerqueirajr

Introdução

Introdução

- Os avanços científicos são, na maioria das vezes, atribuídos aos experimentos realizados.
- Um pesquisador realiza o experimento e obtém dados.
- Baseado nos dados, algumas conclusões podem ser retiradas.
- Estas conclusões vão, geralmente, além dos que foi observado nos dados.
- Dessa forma, o pesquisador generaliza, partindo de um experimentos para os demais que são similares.
- Esta generalização é denominada de `Inferência`.

Uma função da Estatística:

Fornecer um conjunto de metodologias para realizar a inferência e medir o grau de incerteza dessa inferência, através da `teoria das probabilidades`.

Introdução

Exemplo 1

- Suponha um recepiente com 10 milhões de sementes de flores.
- Cada semente pode produzir flores brancas ou vermelhas.
- Pergunta-se: Qual a porcentagem de flores brancas que serão geradas?
- Para saber o resultado real, teríamos que plantar todas as sementes.
- Seria uma tarefa muito trabalhosa!
- Solução: Plantar algumas sementes, e baseando-se nos resultados podemos obter alguma informação para a porcentagem de flores brancas.

População e amostra

População e amostra

Definição 1 (População) É um conjunto que contém todos os elementos do problema a ser discutido, com pelo menos uma característica em comum. Desejamos obter informação sobre esta população.

Exemplo 2

- Preços da carne em um mês na região metropolitana de Belém.
- Preços do pão em certo dia em Belém.
- Produção de leite por animal em uma fazenda.
- Queremos estudar a proporção de votos para um determinado candidato ao governo do Estado do Pará.
- Queremos estudar o grau de satisfação dos usuários de uma determinada operadora de telefonia celular.

População e amostra

Definição 2 (Amostra aleatória - a.a) Sejam $X_1, X_2, \ldots, X_n$ uma sequência de variáveis aleatórias com distribuição conjunta $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ que fatora como na seguinte igualdade:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \times f_{X_2}(x_2) \times \cdots \times f_{X_n}(x_n) = \prod_{i=1}^{n} f_{X_i}(x_i),$$

em que $f(\cdot)$ é a função de probabilidade (f.p) ou função de densidade de probabilidade (f.d) para cada X_i . Então, $X_1, X_2, \ldots, X_n$ é uma amostra aleatória de tamanho n retirada de uma população com p.d/f.d.

População e amostra

Exemplo 3 Imagine 10 milhões de sementes em um recepiente e a produção de flores brancas e vermelhas.

- `População:` Sementes dentro do recipiente.
- `Unidade experimental:` Uma semente.
- `Característica:` Flor branca ou vermelha.

- Não temos um valor numérico associado a cada elemento, mas podemos definir o seguinte tipo de resposta:

$$\text{Flor branca} = 1 \text{ e Flor vermelha} = 0.$$

`Variável aleatória:` $X_i = 1$ ou $X_i = 0$, para $i = 1, 2, \ldots, n$.

Poupulação e amostra

- A variável aleatória X_i é uma representação do valor numérico que a $i-$ésima unidade amostral irá assumir.
- Depois que a amostra $X_1, X_2, \ldots, X_n$ observada os valores serão conhecidos e denotados por $x_1, x_2, \ldots, x_n$.
- Logo:

Suponha que X pode assumir apenas os valores 0 ou 1 com probabilidades $1-\theta$ e θ, respectivamente. Então, sejam $X_1, X_2, \ldots, X_n$ uma a.a de $X \sim Ber(\theta)$, sua distribuição conjunta $P(X_1 = x_1; X_2 = x_2; \ldots; X_n = x_n)$ é dada por

$$= \theta^{x_1}(1-\theta)^{1-x_1} \times \theta^{x_2}(1-\theta)^{1-x_2} \times \cdots \times \theta^{x_n}(1-\theta)^{1-x_n}$$

$$= \theta^{x_1+x_2+\cdots+x_n}(1-\theta)^{(1-x_1)+(1-x_2)+\cdots+(1-x_n)}$$

$$= \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{\sum_{i=1}^{n}(1-x_i)}$$

$$= \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{(n-\sum_{i=1}^{n} x_i)}$$

Estatísticas e Parâmetros

Estatísticas e Parâmetros - Introdução

Um dos problemas principais da Estatística envolve o seguinte:

- Estudar uma população com f.p/f.d $f(\cdot \mid \theta)$, onde a forma da f.p/f.d é conhecida com parêmetro desconhecido θ.
- Se θ fosse conhecido f.p/f.d estaria completamente especificada.

Procedimento de inferência envolverá:

- A obtenção de uma amostra aleatória $X_1, X_2, \ldots, X_n$ desta f.p/f.d.
- O uso de uma função $T(x_1, x_2, \ldots, x_n)$ como estimativa para o parâmetro θ (desconhecido).

Estatísticas

- O problema aqui consiste em determinar qual será a melhor função para estimar θ.
- Iremos avaliar certas funções (funções amostrais) de uma amostra aleatória.

Definição 3 (Estatísticas) É uma função da amostra, $T(x) = f(X_1, X_2, \ldots, X_n)$, representando uma característica da amostra.

Importante:

A formulação de uma estatística `não pode envolver quantidades desconhecidas`.

Estatísticas

> Os exemplos mais comuns:

- Média amostral: $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$.
- Variância amostral: $\hat{\sigma^2} = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.
- Mediana amostral: $\tilde{X} = \text{med}(X_1, X_2, \ldots, X_n)$.
- Mínimo amostral: $X_{(1)} = \min(X_1, X_2, \ldots, X_n)$.
- Máximo amostral: $X_{(n)} = \max(X_1, X_2, \ldots, X_n)$.
- Ponto médio amostral: $\dfrac{1}{2}(X_{(1)} + X_{(n)})$.

Parâmetros

Definição 4 (Parâmetro) Uma parâmetro é uma medida (desconhecida) usada para descrever uma característica da população.

- As relação das estatísticas com seus respectivos parâmetros:

Medida	**Estatística**	**Parâmetro**
Média	$\overline{X}$	μ
Variância	$\hat{\sigma}^2$	σ^2
N de elementos	n	N
Proporção	$\hat{\theta}$	θ

Estatísticas e Parâmetros

Exemplo 4 Considere uma variável aleatória observável com f.d:

- $f(x) = N\left[x \mid \mu, \sigma^2\right]$, com μ e σ desconhecidos.
- Logo,

$$X - \mu \text{ e } X/\sigma \text{ são Estatísticas??}$$

- `Não são`, pois contém elementos desconhecidos.
- $X, X + 3$ e $X^2 + \log X^2$ `são estatísticas`.

Estatísticas e Parâmetros

Exemplo 5 Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória com f.p/f.d $f(\cdot;\theta)$ então:

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{e} \quad \frac{1}{2}\{\min(X_1,\ldots,X_n) + \max(X_1,\ldots,X_n)\}$$

`são exemplos` de estatísticas.

Exemplo 6 Se $f(x;\theta) = N\,[x \mid \theta, 1]$, com θ desconhecido.

$$\overline{X}_n - \theta \quad \text{é uma Estatística?}$$

- `Não` é uma estatística, pois depende de θ.

Momentos amostrais

Momentos amostrais

Seja $X_1, \ldots, X_n$ uma a.a com f.p/f.d $f(\cdot)$. O r-ésimo momento amostral em relação à 0 é definido por

$$M_r^{'} = \frac{1}{n} \sum_{i=1}^{n} (X_i - 0)^r = \frac{1}{n} \sum_{i=1}^{n} (X_i)^r.$$

- Em particular, quando $r = 1$, temos a média amostral $\overline{X}$ ou $\overline{X}_n$, em que

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} (X_i) .$$

O r-ésimo momento em relação à $\overline{X}_n$ é dado por

$$M_r = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X}_n \right)^r$$

Momentos amostrais

> Momentos amostrais são exemplos de estatísticas!

Teorema 1 (Momentos Amostrais) Seja $X_1, \ldots, X_n$ uma a.a com f.p/f.d $f(\cdot)$. O valor esperado do r-ésimo momento amostral (em relação à 0) é igual ao r-ésimo momento populacional, isto é,

$$E(M_r^{'}) = \mu_r^{'}, \text{ se } \mu_r^{'} \text{ existir.}$$

- Temos que $\mu_r^{'} = E(X^r)$ é o r-ésimo momento populacional de uma população com f.p/f.d $f(x) = f_X(x)$.
- Além disso:

$$Var(M_r^{'}) = \frac{1}{n}\left[E(X^{2r}) - E^2(X^r)\right]$$

$$= \frac{1}{n}\left[\mu_{2r}^{'} - (\mu_r^{'})^2\right].$$

Momentos amostrais

`Demonstração:(cont.)`

> A média:

$$E(M_r^{'}) = E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i)^r\right] = \frac{1}{n}E\left[\sum_{i=1}^{n}(X_i)^r\right] =$$

$$= \frac{1}{n}\sum_{i=1}^{n}E\left[(X_i)^r\right] = \frac{1}{n}\sum_{i=1}^{n}\mu_r^{'} = \mu_r^{'}.$$

> A variância:

$$Var(M_r^{'}) = Var\left[\frac{1}{n}\sum_{i=1}^{n}(X_i)^r\right] = \frac{1}{n^2}Var\left[\sum_{i=1}^{n}(X_i)^r\right]$$

Momentos amostrais

Demonstração:(cont.)

Supondo independência, temos

$$Var(M_r^{'}) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left[(X_i)^r\right] = \frac{1}{n^2}\sum_{i=1}^{n}\left[E(X_i)^{2r} - E^2\left(X_i^r\right)\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\left[\mu_{2r}^{'} - (\mu_r^{'})^2\right] = \frac{1}{n}\left[\mu_{2r}^{'} - (\mu_r^{'})^2\right].$$

Momentos amostrais

Quando $r = 1$, temos o seguinte corolário.

Corolário 1 Seja $X_1, \dots, X_n$ uma a.a com f.p/f.d $f(\cdot)$ e se $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n}(X_i)$ é a média amostral, então,

$$E(\overline{X}_n) = \mu, \text{ e } Var(\overline{X}_n) = \frac{\sigma^2}{n}.$$

em que μ e σ^2 são a média e a variância de $f(\cdot)$.

Momentos amostrais

- O Teorema 1 fornece a média e a variância, em termos de momentos populacionais, do r-ésimo momento amostral em relação a 0.
- Um resultado similar, porém mais complicado, pode ser derivado para a média e variância do r-ésimo momento amostral em relação a média amostral.
- Considere $r = 2$, tal que $M_2 = \dfrac{1}{n} \sum_i (X_i - \overline{X})^2$.
- M_2, as vezes, é chamado de `variância amostral`.
- Entretanto, definiremos a variância amostral de outra forma.

Momentos amostrais

Definição 5 Seja $X_1, \ldots, X_n$ uma a.a com f.p/f.d $f(\cdot)$,

$$S_n^2 = S^2 = \frac{1}{n-1}\sum_i (X_i - \overline{X})^2, \text{ para } n > 1,$$

é definida como **variância amostral**.

> A razão para considerarmos S^2 ao invés de M_2 como variância é devido
>
> $$E(S^2) = \sigma^2 \text{ e } E(M_2) \neq \sigma^2$$

- Revisando: $\mu_r^{'} = E(X^r)$ é o r-ésimo momento de X (em relação a 0).

Momentos amostrais

Definição 6 (Momento central) O r-ésimo momento central de uma variável aleatória X com relação ao ponto $\mathbf{a}$ é definido por

$$\mu_r = E\left[(X - \mathbf{a})^r\right]$$

- Se $\mathbf{a} = E(X) = \mu$, temos $\mu_r = E\left[(X - \mathbf{a})^r\right]$, então

$$\mu_1 = E\left[(X-\mu)^1\right] = 0 \quad \text{e} \quad \mu_2 = E\left[(X-\mu)^2\right] = Var(X) = \sigma^2.$$

Teorema 2 Seja $X_1, \ldots, X_n$ uma a.a com f.p/f.d $f(\cdot)$ e seja

$$S^2 = \frac{1}{n-1}\sum_i (X_i - \overline{X})^2,$$

Então,

$$E(S^2) = \sigma^2 \quad \text{e} \quad Var(S^2) = \frac{1}{n}\left[\mu_4 - \frac{n-3}{n-1}\sigma^2\right].$$

Momentos amostrais

Prova: Para $E(S^2) = \sigma^2$, temos que $\sigma^2 = E\left[(X-\mu)^2\right]$ e $\mu_r = E\left[(X-\mu)^r\right]$. Note que,

$$\begin{aligned}
\sum_i (X_i - \mu)^2 &= \sum_i \left(X_i + \overline{X} - \overline{X} - \mu\right)^2 \\
&= \sum_i \left[\left(X_i - \overline{X}\right)^2 - 2\left(X_i - \overline{X}\right)\left(\overline{X} - \mu\right) + \left(\overline{X} - \mu\right)^2\right] \\
&= \sum_i \left(X_i - \overline{X}\right)^2 - 2\left(\overline{X} - \mu\right)\sum_i \left(X_i - \overline{X}\right) + n\left(\overline{X} - \mu\right)^2 \\
&= \sum_i \left(X_i - \overline{X}\right)^2 - 2\left(\overline{X} - \mu\right)\underbrace{\left(n\overline{X} - n\overline{X}\right)}_{0} + n\left(\overline{X} - \mu\right)^2 \\
&= \sum_i \left(X_i - \overline{X}\right)^2 + n\left(\overline{X} - \mu\right)^2
\end{aligned}$$

Momentos amostrais

Prova (cont.):

Assim,

$$
\begin{aligned}
E(S^2) &= E\left[\frac{1}{n-1}\sum_i (X_i - \overline{X})^2\right] \\
&= E\left[\frac{1}{n-1}\sum_i (X_i - \mu)^2 - n\left(\overline{X} - \mu\right)^2\right] \\
&= \frac{1}{n-1} E\left[\sum_i (X_i - \mu)^2 - n\left(\overline{X} - \mu\right)^2\right] \\
&= \frac{1}{n-1} E\left[\sum_i (X_i - \mu)^2\right] - \frac{n}{n-1} E\left[\left(\overline{X} - \mu\right)^2\right] \\
&= \frac{n}{n-1}\sigma^2 - \frac{n}{n-1} Var(\overline{X}) \\
&= \frac{n}{n-1}\sigma^2 - \frac{n}{n-1}\frac{\sigma^2}{n} \\
&= \sigma^2
\end{aligned}
$$

Momentos amostrais

- De forma similar,

$$
\begin{aligned}
E(M_2) &= E\left[\frac{1}{n}\sum_i (X_i - \overline{X})^2\right] \\
&= \frac{1}{n}\sigma^2 - \frac{n}{n}Var(\overline{X}) \\
&= \sigma^2\left(\frac{n-1}{n}\right)
\end{aligned}
$$

- Momentos amostrais são exemplos de **exemplos de estatísticas** que podem ser usados para estimar quantidades populacionais.
- Por exemplo:
  - $M_r^{'}$ para estimar $\mu_r^{'}$;
  - $\overline{X}$ para estimar μ;
  - S^2 para estimar σ^2.

Função de verossimilhança

Função de verossimilhança

Definição 7 (Função de verossimilhança) A f.p/p.d.f conjunta é denominada função de verossimilhança de θ, correspondente a amostra observada $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e será denotada por

$$L(\theta \mid \mathbf{x}) = \prod_{i=1}^{n} f(x_i \mid \theta) = f(x_1 \mid \theta) \times f(x_2 \mid \theta) \times \cdots \times f(x_n \mid \theta).$$

Dada a amostra $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, podemos encontrar o `ponto mais verossímil` para θ.

Função de verossimilhança

Exemplo 7 (Caso discreto) Sejam $X_1, X_2, \ldots, X_n$ uma a.a de $X \sim Pois(\theta)$. Temos que a função de verossimilhança é dada por

$$
\begin{aligned}
L(\theta \mid \mathbf{x}) &= \prod_{i=1}^{n} \frac{\exp\{-\lambda\}\lambda^{x_i}}{x_i!} \\
&= \frac{\exp\{-\lambda\}\lambda^{x_1}}{x_1!} \times \frac{\exp\{-\lambda\}\lambda^{x_2}}{x_2!} \times \cdots \times \frac{\exp\{-\lambda\}\lambda^{x_n}}{x_n!} \\
&= \frac{\exp\{-n\lambda\}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}.
\end{aligned}
$$

Função de verossimilhança

Exemplo 8 (Caso contínuo) Sejam $X_1, X_2, \ldots, X_n$ uma a.a de $X \sim N(\mu, \sigma^2)$. Temos que a função de verossimilhança é dada por

$$
\begin{aligned}
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{1}{2\sigma^2} (x_i - \mu)^2 \right\} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{1}{2\sigma^2} (x_1 - \mu)^2 \right\} \times \cdots \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{1}{2\sigma^2} (x_n - \mu)^2 \right\} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left\{ \frac{1}{2\sigma^2} \textstyle\sum_{i=1}^{n} (x_i - \mu)^2 \right\} \\
&= \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{ \frac{1}{2\sigma^2} \textstyle\sum_{i=1}^{n} (x_i - \mu)^2 \right\} \\
&= \left(2\pi\sigma^2 \right)^{-n/2} \exp\left\{ \frac{1}{2\sigma^2} \textstyle\sum_{i=1}^{n} (x_i - \mu)^2 \right\}
\end{aligned}
$$

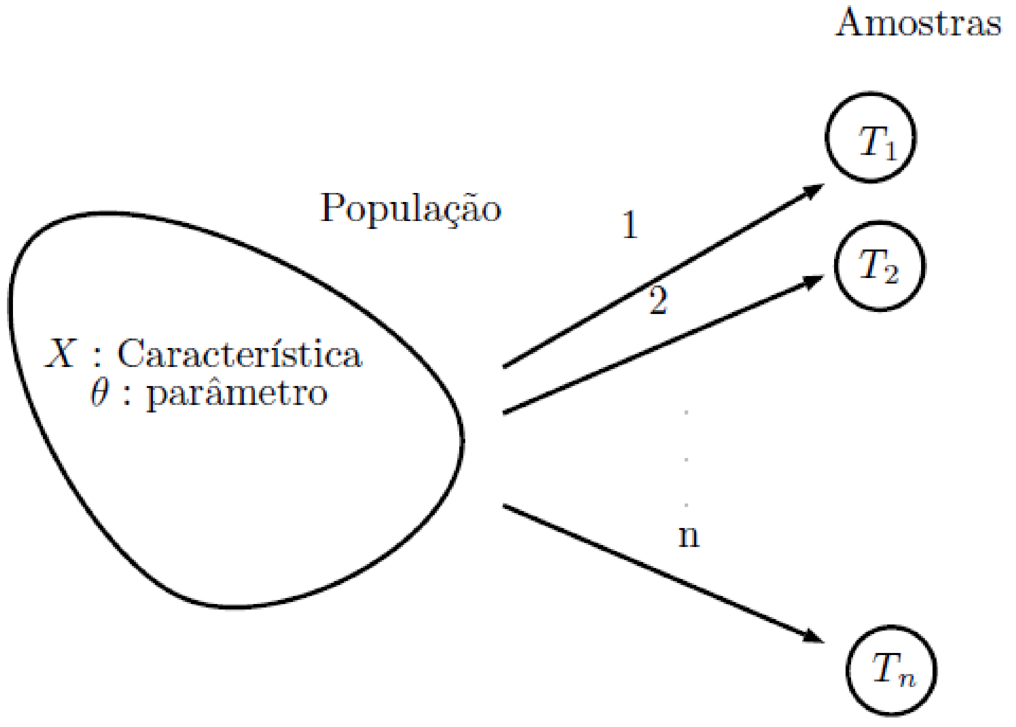
Distribuição amostral

Introdução

Para este caso temos:

- X : Variável de interesse;
- θ : parâmetro de interesse;
- $T = f(X_1, X_2, \ldots, X_n)$ é uma função da amostra que vai fornecer informação sobre θ.

Ideia da distribuição amostral



T é uma variável aleatória.

`Pergunta:` Qual a distribuição de T quando $X_1, X_2, \ldots, X_n$ assume valores observados?

Distribuição amostral da média

Distribuição amostral da média

Exemplo 9 Suponha que selecionamos todas as amostras de tamanho 2, com reposição, da população $\{1, 3, 5, 5, 7\}$

X	**1**	**3**	**5**	**7**
$P(X = x)$	1/5	1/5	2/5	1/5

- Encontrar a distribuição conjunta da v.a. $(X_1, X2)$, sendo X_1 sendo o número selecionado na primeira extração e X_2 o número da segunda.
- Encontre a distribuição de $\overline{X} = \dfrac{X_1 + X_2}{2}$.

Distribuição amostral da média

Combinação	Prob.	X_1	X_2	$\overline{X}$
$(1,1)$	1/25	1	1	1
$(1,3)$	1/25	1	3	2
$(1,5)$	2/25	1	5	3
$(1,7)$	1/25	1	7	4
$(3,1)$	1/25	3	1	2
$(3,3)$	1/25	3	3	3
$(3,5)$	2/25	3	5	4
$(3,7)$	1/25	3	7	5
$(5,1)$	2/25	5	1	3
$(5,3)$	4/25	5	3	4
$(5,5)$	2/25	5	5	5
$(5,7)$	1/25	5	7	6
$(7,1)$	1/25	5	7	4
$(7,3)$	1/25	5	7	5
$(7,5)$	1/25	5	7	6
$(7,7)$	1/25	5	7	7

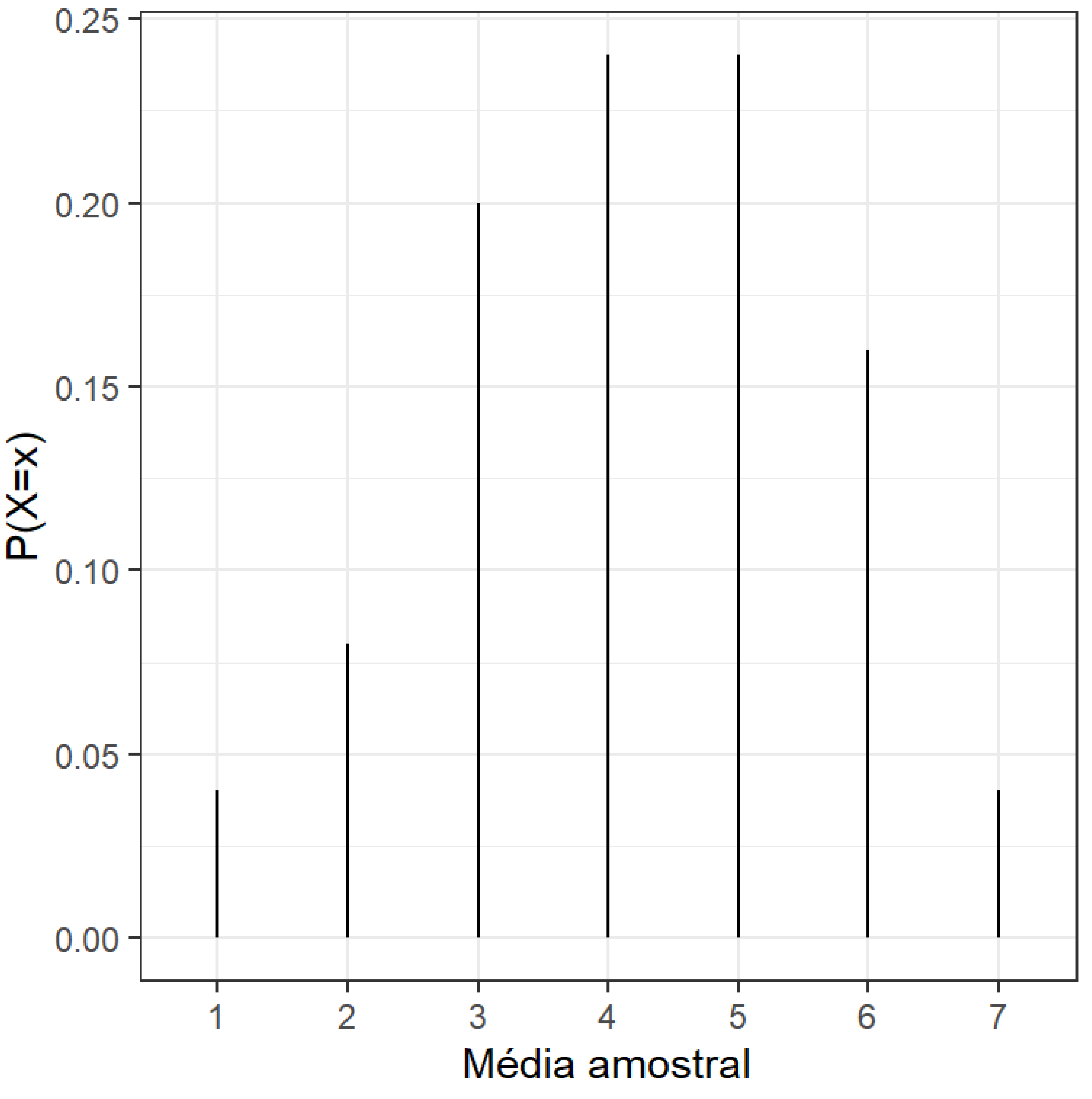
`Distribuição conjunta:`

X_1/X_2	**1**	**3**	**5**	**7**	**Total**
1	1/25	1/25	2/25	1/25	1/5
3	1/25	1/25	2/25	1/25	1/5
5	2/25	2/25	4/25	2/25	2/5
7	1/25	1/25	2/25	1/25	1/5
Total	1/5	1/5	2/5	1/5	Total

Distribuição amostral da média

Distribuição amostral da média $\overline{X}$:

$\overline{X}$	1	2	3	4	5	6	7
$P(\overline{X} = \overline{x})$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{5}{25}$	$\frac{6}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{1}{25}$



Distribuição amostral da média

- O primeiro momento amostral é a média definida como:

$$\overline{X} = \overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} (X_i) \, .$$

onde $X_1, X_2, \ldots, X_n$ é uma amostra aleatória com f.p/f.d $f(\cdot)$.

- $\overline{X}$ é função das v.a $X_1, X_2, \ldots, X_n$ e, portanto a distribuição pode ser encontrada teoricamente.
- Pode ser útil pensar na média amostral $\overline{X}$ como uma estimativa da média μ da f.p/f.d $f(\cdot)$ a partir de qual amostra foi selecionada.

> Um dos objetivos da amostragem é estimar μ a partir de $\overline{X}$.

Distribuição amostral da média

Teorema 3 Seja $X_1, X_2, \ldots, X_n$ uma a.a com f.p/f.d $f(\cdot)$, média μ e variância σ^2. Considere:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n}(X_i)\,.$$

Então, $E(\overline{X}) = \mu_{\overline{X}} = \mu$ e $Var(\overline{X}) = \sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$.

> $E(\overline{X}) = \mu$: diz que em média $\overline{X}$ é igual ao parâmetro μ sendo estimado ou que a distribuição de $\overline{X}$ está centrada em μ.

> $Var(\overline{X}) = \dfrac{\sigma^2}{n}$: diz que a dispersão dos valoers de $\overline{X}$ em torno de μ é pequena para amostras grandes em comparação com tamanhos menores.

Teoremas de convergência

Teoremas de convergência

- Para amostras grandes, os valores de $\overline{X}$ (que são usados para estimar μ) tendem a estar mais concentrados de μ do que em amostras pequenas.
- Esta noção será definida pela **Lei dos Grandes Números**
- Seja $E(X) = \mu$ para a f.p/f.d $f(\cdot)$. Desejamos estimar μ.
- De maneira não rigorosa, $E(X)$ é a média de um número infinito de valores da variável aleatória X.
- Em qualquer problema real podemos observar apenas um número finito de valores da variável aleatória X.
- Questão: Usando apenas um número finito de valores de X (uma amostra aleatória de tamanho n) pode ser feita qualquer inferência confiável sobre $E(X)$? **A resposta é sim**.
- Usaremos isso através da Lei Fraca dos Grandes Números.

Teoremas de convergência

Teorema 4 (Lei fraca dos Grandes Números) Seja $X_1, X_2, \ldots, X_n$ uma a.a de tamanho n de uma população com variável X, com média $E(X) = \mu$ e $Var(X) = \sigma^2 < \infty$. Sejam, $\epsilon > 0$ e $0 < \delta < 1$. Se, $n > \dfrac{\sigma^2 \epsilon^2}{\delta}$, então,

$$P(\mid \overline{X}_n - \mu \mid < \epsilon) \geq 1 - \delta,$$

ou seja, $\overline{X}_n$ converge em probabilidade para μ.

Teorema de convergência

Seja $X_1, X_2, \ldots, X_n$ uma a.a. de $X \sim Ber(0.5)$. Observe que,

$$X_i = \begin{cases} 0, & \text{fracasso} \\ 1, & \text{sucesso} \end{cases}, i = 1, 2, \ldots, n.$$

A proporção amostral é determinada por

$$\hat{p}_n = \overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Para $n = 1 \quad \Rightarrow \quad \hat{p}_1 = \dfrac{X_1}{1}.$

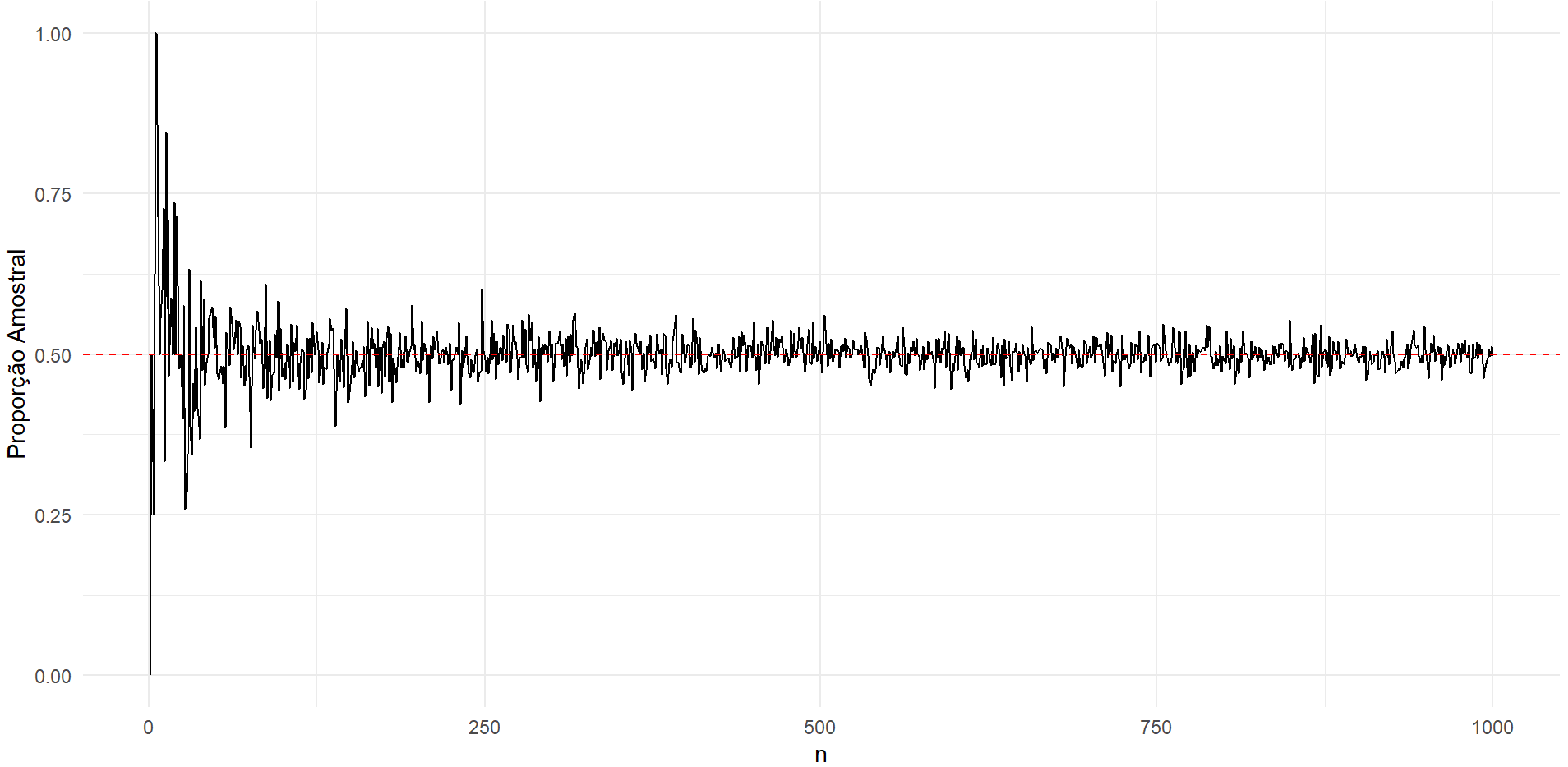
Para $n = 2 \quad \Rightarrow \quad \hat{p}_2 = \dfrac{X_1 + X_2}{2}.$

$\vdots$

Para $n = n \quad \Rightarrow \quad \hat{p}_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}.$

Teorema de convergência

- Ao gerar uma amostra de tamanho $n = 1, 2, 3, \ldots, 1000$ e calcular a proporção amostral em cada uma delas tem-se o seguinte resultado:



Teorema central do limite

> O Teorema Central do Limite é um dos mais importantes resultados em toda área de Probabilidade e Estatística. Ele nos diz aproximadamente como a **média amostral** é distribuída.

Teorema 5 (Teorema Central do Limite - TCL) Seja $X_1, X_2, \ldots, X_n$ uma sequência de v.a.'s independentes com $E(X_i) = \mu_i$ e $Var(X_i) = \sigma^2) < \infty, i = 1, 2, \ldots, n$. Tome $S_n = X_1 + X_2 + \cdots + X_n$, então, sob determinadas condições gerais,

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \to N(0, 1).$$

A distribuição de Z_n se aproxima da $N(0, 1)$ quando $n \to \infty$.

O Teorema 5 nos diz que a ditribuição limite de Z_n (S_n padronizado) será a distribuição $N(0, 1)$.

Distribuição amostral da média

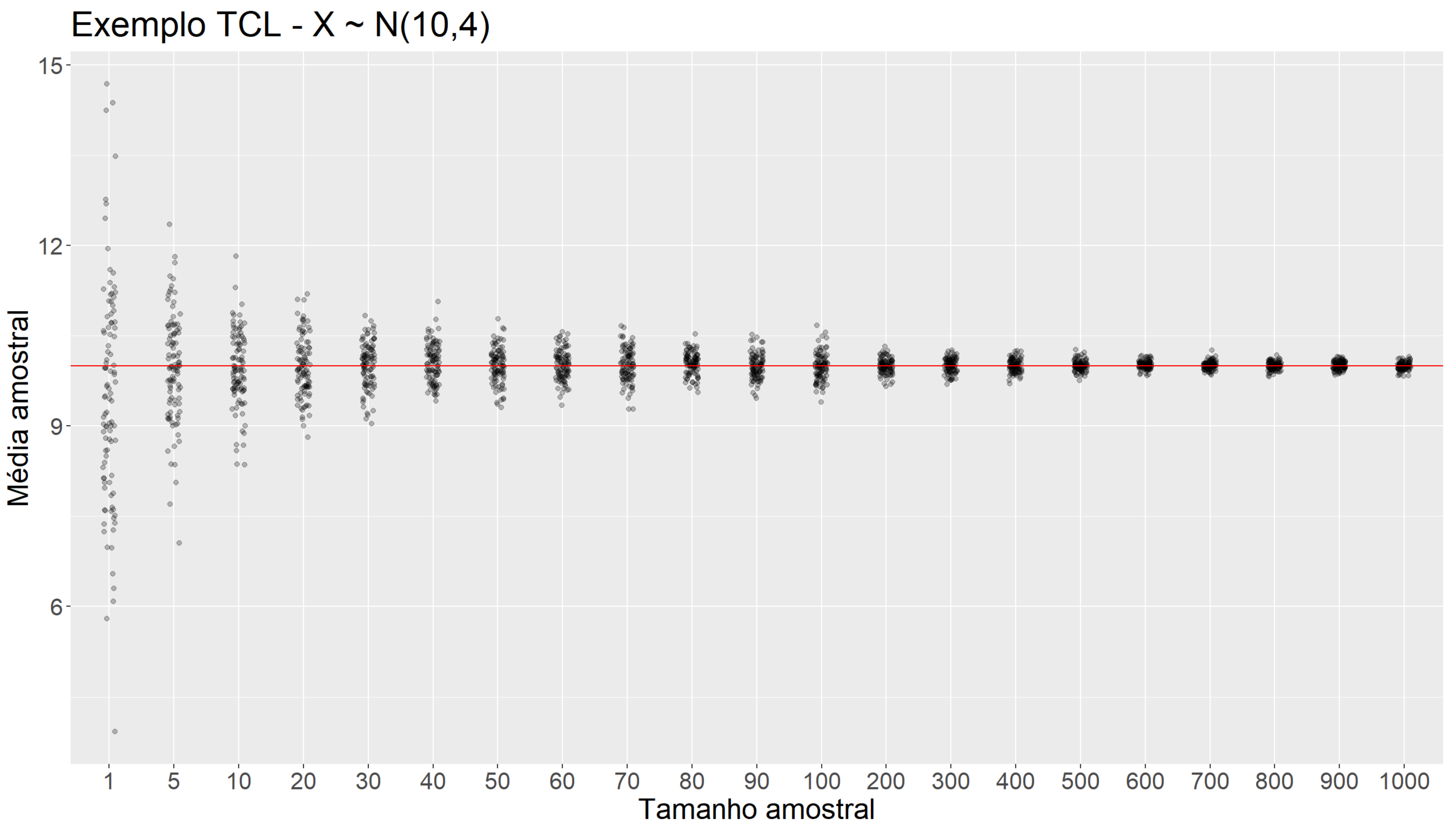
Corolário 2 Seja $(X_1, X_2, \ldots, X_n)$ uma a.a. de X com $E(X) = \mu$ e $Var(X) = \sigma^2 < \infty$. Então, para $n \to \infty$,

$$Z_n = \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \to N(0,1).$$

- Em outras palavras $\overline{X}_n$ é assintoticamente distribuído como uma Normal com média μ e variância σ^2/n.
- Um aspecto importante sobre o Teorema 5 é o fato de que nada é dito sobre a forma da **f.p ou f.d original**. Qualquer que seja a distribuição, dado que possui variância **finita**, a média amostral terá **aproximadamente** distribuição Normal para amostras grandes.

Representação do TCL graficamente

- Ao gerar um 100 amostras de tamanho $n = 1, 2, 3, \ldots, 1000$ e calcular a média amostral em cada uma delas, supondo $X \sim N(10, 4)$, tem-se o seguinte resultado:



Algumas distribuições exatas

- Se $X \sim Ber(\theta)$, então:

$$P(\bar{X} = \bar{x}_n) = \binom{n}{n\bar{x}_n} \theta^{n\bar{x}_n} (1-\theta)^{n - n\bar{x}_n}, \bar{x}_n = 0, 1/n, 2/n, \ldots, 1.$$

- Se $X \sim Pois(\theta)$, então:

$$P(\bar{X} = \bar{x}_n) = \frac{e^{-n\theta}\theta^{n\bar{x}_n}}{n\bar{x}_n!}, \bar{x}_n = 0, 1/n, 2/n, \ldots$$

- Se $X \sim Normal(\mu, \sigma^2)$, então:

$$\bar{X}_n \sim N(\mu, \sigma^2).$$

- Se $X \sim Exp(\theta)$, então:

$$\bar{X}_n \sim Gama(n, n\theta).$$

Distribuição amostral da proporção

- Seja $X_1, X_2, \ldots, X_n$ uma a.a. de $X \sim Ber(\theta)$.
- Em que θ representa a proporção de elementos com uma determinada característica na população.
- Temos que

$$E(X_i) = \theta \quad \text{e} \quad Var(X_i) = \theta(1-\theta).$$

- Seja $S_n = X_1 + X_2 + \cdots + X_n$, então a proporção amostral é definida por

$$\hat{p} = \frac{S_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \bar{X}.$$

Distribuição amostral da proporção

`Distribuição exata:`

- Temos que $S_n = X_1 + X_2 + \cdots + X_n \sim Bin(n, \theta)$, então

$$P\left(\hat{p} = \frac{k}{n}\right) = \binom{n}{k}\theta^k(1-\theta)^{n-k}, k = 0, 1, 2, \ldots, n.$$

`Distribuição aproximada pelo TCL:`

- Temos que $\hat{p} = \dfrac{S_n}{n} = \dfrac{X_1 + X_2 + \cdots + X_n}{n} = \bar{X}_n.$, então

$$\hat{p} \sim N\left(\theta, \frac{\theta(1-\theta)}{n}\right).$$

Distribuição amostral de estatísticas de ordem

Estatísticas de ordem amostrais

- Seja $X_1, X_2, \ldots, X_n$ sendo uma sequência de variáveis aleatórias i.i.d. com função de distribuição $F(\cdot)$.
- Podemos reordenar (de forma crescente) essa sequência da seguinte forma

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}.$$

- No caso em que F seja contínua, temos que

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}.$$

- Uma vez que $P(X_i = X_j) = 0$ para todo $i \neq j$, para variáveis aleatórias contínuas.

Estatísticas de ordem amostrais

- Para uma sequência de v.a's $X_1, X_2, \ldots, X_n$ i.i.d., $X_{(k)}$ é denominada de $k-$ésima estatística de ordem.
- O mínimo é denotado por $X_{(1)}$:

$$X_{(1)} = \min(X_1, X_2, \ldots, X_n).$$

- De maneira similar, o máximo é denotado por $X_{(n)}$:

$$X_{(n)} = \max(X_1, X_2, \ldots, X_n).$$

Estatísticas de ordem amostrais

Seja $g(x)$ e $G(X)$ as funções de densidade (probabilidade) e distribuição de X, respectivamente.

- Para o $X_{(1)}$:
  - A função de distribuição é dada por

$$F_{X_{(1)}}(x) = 1 - (1 - G(x))^n.$$

- A função de densidade é dada por

$$f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) = ng(x)(1 - G(x))^{(n-1)}.$$

Estatísticas de ordem amostrais

Seja $g(x)$ e $G(X)$ as funções de densidade (probabilidade) e distribuição de X, respectivamente.

- Para o $X_{(n)}$:
  - A função de distribuição é dada por

$$F_{X_{(n)}}(x) = G(x)^n.$$

- A função de densidade é dada por

$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x) = ng(x)G(x)^{(n-1)}.$$

Amostrando da Normal - Média amostral

Média amostral

- Esta seção lida com propriedades das quantidades amostrais provenientes de uma população Normal.
- A distribuição Normal tem papel importante nos estudos estatísticos.
- Muitas populações seguem a distribuição Normal com um bom grau de aproximação.
- Modelos estatísticos utilizando a distribuição Normal são amplamente considerados na literatura científica.
- Amostrar de uma população Normal leva a muitas propriedades úteis e também a muitas distribuições amostrais conhecidas.

`Média amostral e a população Normal:`

- A média amostral é uma das mais simples funções de uma amostra aleatória.
- Para uma amostra aleatória da distribuição Normal, a distribuição exata da média amostral também é Normal (para qualquer tamanho amostral n, ou seja, não precisamos do TCL para dar suporte a esta afirmação).

Média amostral

Teorema 6 Seja $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n}(X_i)$ a média amostral de uma amostra aleatória $X_1, \ldots, X_n$ obtida da distribuição Normal com média μ e variância σ^2. Então, $\overline{X}_n \sim N\left(\mu, \sigma^2/n\right)$.

Prova: Usaremos a função geradora de momentos (f.g.m),

$$
\begin{aligned}
m_{\overline{X}_n}(t) &= E\left(e^{t\overline{X}_n}\right) = E\left[\exp\left\{\frac{t}{n}\sum_{i=1}^{n}X_i\right\}\right] \\
&= E\left[\prod_{i=1}^{n}\exp\left\{\frac{t}{n}X_i\right\}\right] \underbrace{=}_{X_i^{'}s\ ind.} \prod_{i=1}^{n}E\left[\exp\left\{\frac{t}{n}X_i\right\}\right] \\
&= \prod_{i=1}^{n}m_{X_i}(t/n) = \prod_{i=1}^{n}\exp\left\{\mu\frac{t}{n} + \frac{\sigma^2}{2}\frac{t^2}{n^2}\right\} \\
&= \exp\left\{\mu t + \frac{\sigma^2/n}{2}t^2\right\}. \Rightarrow \text{f.g.m da } N\left(\mu, \sigma^2/n\right)
\end{aligned}
$$

Média amostral

> Dado que temos a distribuição exata de $\overline{X}_n$ quando estimamos μ com $\overline{X}_n$, seremos capazes de calcular, por exemplo, a probabilidade exata de que nosso estimador $\overline{X}_n$ esteja dentro de uma distância fixada do parâmetro desconhecido μ.

Amostrando da Normal - Variância amostral

Variância amostral

- A distribuição Normal possui dois parâmetros desconhecidos μ e σ^2.
- Vimos anteriormente a distribuição amostral de $\overline{X}_n$ que estima μ.
- Procuremos agora pela distribuição de

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2,$$

que estima o parâmetro σ^2.

- A distribuição Qui-Quadrado desempenha um papel fundamental na determinação da distribuição de S^2.

Variância amostral

Definição 8 (Distribuição Qui-Quadrado) Seja X uma v.a. com f.d

$$f(x) = \frac{1}{\Gamma(k/2)}\left(\frac{1}{2}\right)^{k/2} x^{k/2-1}\exp\left\{-\frac{1}{2}x\right\} \times I_{(0,\infty)}(x),$$

então dizemos que X tem distribuição Qui-Quadrado com k graus de liberdade (k é um número inteiro positivo).

`Notação:` χ^2_k = Qui-Quadrado com k graus de liberdade.

- A densidade da Qui-Quadrado é um caso particular da densidade $Gama(r, \lambda)$, onde $r = k/2$ e $\lambda = 1/2$.
- Se $X \sim Ga(r, \lambda)$ temos $f(x) = \dfrac{\lambda^r}{\Gamma(r)} x^{r-1} e^{\lambda x}, I_{(0,\infty)}(x)$.
- Fazendo $r = k/2$ e $\lambda = 1/2$, tem-se na densidade definida anteriormente.

Variância amostral

Temos também que:

$$E(X) = \frac{r}{\lambda} = \frac{k/2}{1/2} = k \quad \text{e} \quad Var(X) = \frac{r}{\lambda^2} = \frac{k/2}{1/2^2} = 2k.$$

A f.g.m é dada por

$$m_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-r} = \left(1 - \frac{t}{1/2}\right)^{-k/2} = \left(\frac{1}{1-2t}\right)^{k/2}, \quad \text{para todo} \quad t < \lambda = 1/2$$

Teorema 7 Se $X_1, \ldots, X_n$ são v.a. Normais independentes com médias μ_i e variâncias σ_i^2 (Note que não prescisar ser i.i.d). Então,

$$U = \sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi_k^2.$$

Variância amostral

`Prova:` Seja $Z_i = \dfrac{X_i - \mu_i}{\sigma_i}$, em que $Z_i \sim N(0,1)$. Assim,

$$\begin{aligned} m_U(t) &= E\left[e^{tU}\right] &= E\left[e^{t\sum_{i=1}^{k} Z_i^2}\right] \\ &= E\left[\prod_{i=1}^{k} e^{tZ_i^2}\right] &\underbrace{=}_{Z_i' sind.} \prod_{i=1}^{k} E\left[e^{tZ_i^2}\right]. \end{aligned}$$

Variância amostral

Mas,

$$
\begin{aligned}
E\left[e^{tZ_i^2}\right] &= \int_{-\infty}^{\infty} e^{tZ^2}(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}Z^2} dZ = \int_{-\infty}^{\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}(Z^2-2tZ^2)} dZ \\
&= \int_{-\infty}^{\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}(1-2t)Z^2} dZ \\
&= \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \sqrt{1-2t}(2\pi)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\left(\frac{1}{(1-2t)}\right)} Z^2\right\} dZ \\
&= \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \left[2\pi\left(\frac{1}{(1-2t)}\right)\right]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\left(\frac{1}{(1-2t)}\right)} Z^2\right\} dZ
\end{aligned}
$$

Variância amostral

- A expressão de dentro da integral é a densidade da $N(0, V)$, onde $V = \frac{1}{(1-2t)}$.
- O resultado da integral é, portanto, igual a 1.
- Conclusão:

$$E\left[e^{tZ_i^2}\right] = m_{Z_i^2}(t) = \left(\frac{1}{1-2t}\right)^{\frac{1}{2}}, \text{ para } t < 1/2.$$

`Nota:` O resultado acima determina que $Z_i^2 \sim \chi_1^2$, ou seja, se $Z_i \sim N(0,1)$ temos $Z_i^2 \sim \chi_1^2$.

$$m_U(t) = \prod_{i=1}^{k} E\left[e^{tZ_i^2}\right] = \prod_{i=1}^{k} \left(\frac{1}{1-2t}\right)^{\frac{1}{2}} = \left(\frac{1}{1-2t}\right)^{\frac{k}{2}}$$

Portanto, $U \sim \chi_2^k$ finalizando a prova.

Variância amostral

Corolário 3 Se $X_1, \ldots, X_n$ são v.a. Normais independentes com médias μ e variâncias σ^2

$$U = \sum_{i=1}^{k} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_k^2.$$

Em palavras, o Teorema 7 diz que a soma de quadrados de variáveis aleatórias N(0,1) independentes possui distribuição Qui-Quadrado com grau de liberdade igual ao número de termos na soma.

Variância amostral

Teorema 8 Se $Z_1, Z_2 \ldots, Z_n$ é uma amostra aleatória da distribuição $N(0,1)$, então:

i. $\overline{Z} \sim N(0, 1/n)$;

ii. $\overline{Z}$ e $\sum_{i=1}^{n}(Z_i - \overline{Z})^2$ são independentes;

iii. $\sum_{i=1}^{n}(Z_i - \overline{Z})^2 \sim \chi^2_{n-1}$.

`Prova:` (i) é um caso especial do Teorema 7. A prova da parte (ii) será incompleta (somente o caso $n = 2$). Se $n = 2$ temos $\overline{Z} = (Z_1 + Z_2)/2$ e

$$
\begin{aligned}
\sum_{i=1}^{2}(Z_i - \overline{Z})^2 &= (Z_1 - (Z_1 + Z_2)/2)^2 + (Z_2 - (Z_1 + Z_2)/2)^2 \\
&= \frac{(Z_1 - Z_2)^2}{4} + \frac{(Z_2 - Z_1)^2}{4} = \frac{(Z_2 - Z_1)^2}{2}
\end{aligned}
$$

Variância amostral

- Então, $\overline{Z}$ é função de $Z_1 + Z_2$ e $\sum_{i=1}^{2}(Z_i - \overline{Z})^2$ é função de $Z_2 - Z_1$.
- Para mostrar que $\overline{Z}$ e $\sum_{i=1}^{2}(Z_i - \overline{Z})^2$ são independentes basta mostrar que $Z_1 + Z_2$ e $Z_2 - Z_1$ são independentes.

Veja que

$$\begin{aligned} m_{Z_1+Z_2}(t_1) &= E\left[e^{t_1(Z_1+Z_2)}\right] = E\left[e^{t_1 Z_1} e^{t_1 Z_2}\right] \\ &\underbrace{=}_{Z_i' ind.} E\left[e^{t_1 Z_1}\right] E\left[e^{t_1 Z_2}\right] = \exp\left\{\frac{1}{2}t_1^2\right\} \exp\left\{\frac{1}{2}t_1^2\right\} \\ &= \exp\{t_1^2\}. \end{aligned}$$

Similarmente, $m_{Z_2-Z_1}(t_2) = \exp\{t_2^2\}$. Se $Z_1 \sim N(0,1)$, então $-Z_1 \sim N(0,1)$.

Variância amostral

Além disso, temos que a $m_{Z_1+Z_2,Z_2-Z_1}(t_1,t_2)$

$$
\begin{aligned}
&= E\left[e^{t_1(Z_1+Z_2)+t_2(Z_2-Z_1)}\right] = E\left[e^{(t_1-t_2)Z_1}e^{(t_1+t_2)Z_2}\right] \\
&\underbrace{=}_{ind.} E\left[e^{(t_1-t_2)Z_1}\right]E\left[e^{(t_1+t_2)Z_2}\right] = \exp\left\{\frac{1}{2}(t_1-t_2)^2\right\}\exp\left\{\frac{1}{2}(t_1+t_2)^2\right\} \\
&= \exp\left\{\frac{1}{2}(t_1-t_2)^2\right\}\exp\left\{\frac{1}{2}(t_1+t_2)^2\right\} \\
&= \exp\left\{\frac{1}{2}\left[t_1^2-2t_1t_2+t_2^2+t_1^2+2t_1t_2+t_2^2\right]\right\} = \exp\left\{t_1^2\right\}\exp\left\{t_2^2\right\} \\
&= m_{Z_1+Z_2}(t_1)m_{Z_2-Z_1}(t_2).
\end{aligned}
$$

Visto que a f.g.m. conjunta pode ser fatorada no produto das f.g.m.'s Z_1+Z_2 e Z_2-Z_1 são independentes.

Variância amostral

iii. Para provar essa parte, iremos supor que $\overline{Z}$ e $\sum_{i=1}^{n}(Z_i - \overline{Z})^2$ para um n arbitrário.

Note que:

$$
\begin{aligned}
\sum_{i=1}^{n} Z_i^2 &= \sum_{i=1}^{n}(Z_i - \overline{Z} + \overline{Z})^2 \\
&= \sum_{i=1}^{n}(Z_i - \overline{Z})^2 + 2\overline{X}\underbrace{\sum_{i=1}^{n}(Z_i - \overline{Z})}_{n\overline{Z}-n\overline{Z}=0} + n\overline{X}^2 \\
&= \sum_{i=1}^{n}(Z_i - \overline{Z})^2 + n\overline{Z}^2
\end{aligned}
$$

Variância amostral

Usando o resultado da parte (ii) temos que que $\overline{Z}$ e $\sum\limits_{i=1}^{n}(Z_i - \overline{Z})^2$ são independentes.

Então:

$$m_{\sum_{i=1}^n Z_i^2}(t) \quad = \quad m_{\sum_{i=1}^n (Z_i - \overline{Z})^2 + n\overline{Z}^2}(t) \quad \underbrace{=}_{ind.} \quad m_{\sum_{i=1}^n (Z_i - \overline{Z})^2}(t) m_{n\overline{Z}^2}(t)$$

e

$$m_{\sum_{i=1}^n (Z_i - \overline{Z})^2}(t) = \frac{m_{\sum_{i=1}^n Z_i^2}(t)}{m_{n\overline{Z}^2}(t)} = \frac{\overbrace{\left(\frac{1}{1-2t}\right)^{\frac{n}{2}}}^{\chi_n^2}}{\underbrace{\left(\frac{1}{1-2t}\right)^{\frac{1}{2}}}_{\chi_1^2}} = \left(\frac{1}{1-2t}\right)^{\frac{n-1}{2}}.$$

O resultado da f.g.m. da distribuição χ_{n-1}^2.

Variância amostral

- O Teorema 8 foi definido para uma amostra aleatória da distribuição $N(0, 1)$, entretanto se desejamos fazer inferência sobre μ e σ^2 devemos considerar uma amostra da $N(\mu, \sigma^2)$.
- Se $X_1, X_2, \ldots, X_n$ uma a.a. da $N(\mu, \sigma^2)$. Neste caso definimos $Z_i = (X_i - \mu)/\sigma$. (ver `Teorema 8`)
- Parte (i) do Teorema 8 se torna:

1. $\overline{Z} = \dfrac{1}{n} \sum_{i=1}^{n} (X_i - \mu)/\sigma = \dfrac{\overline{X} - \mu}{\sigma} \sim N(0, 1/n)$

Variância amostral

- Parte (ii) do Teorema 8 se torna:

2. $\overline{X} = \dfrac{\overline{X} - \mu}{\sigma}$ e

$\sum_{i=1}^{n} (Z_i - \overline{Z})^2 = \sum_{i=1}^{n} \left[\dfrac{X_i - \mu}{\sigma} - \dfrac{\overline{X} - \mu}{\sigma} \right]^2 = \sum_{i=1}^{n} \dfrac{(X_i - \overline{X})^2}{\sigma^2}$

são independentes, implicando que $\overline{X}$ e $\sum_{i=1}^{n} (X_i - \overline{X})^2$ também são.

- Parte (iii) do Teorema 8 se torna

$$\sum_{i=1}^{n} (Z_i - \overline{Z})^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

conforme mostrado em ii.

Variância amostral

Corolário 4 Considere $X_1, X_2, \ldots, X_n$ uma a.a. da $N(\mu, \sigma^2)$. Seja $S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$ a variâcia amostral. Então,

$$U = \frac{(n-1)}{\sigma^2} S^2 \sim \chi^2_{n-1}.$$

`Prova:` Considere a parte (iii) da extensão do `Teorema 8`.

Observação: Como S^2 é uma função linear de U no corolário acima, a densidade de S^2 pode ser obtida a partir da densidade de U. $U \sim \chi^2_{n-1}$ e S^2 é dada por uma função monótona de U.

$$S^2 = \frac{\sigma^2}{n-1} U, \; U = \frac{(n-1)}{\sigma^2} S^2 \quad \text{e} \quad \frac{dU}{dS^2} = \frac{n-1}{\sigma^2}$$

Variância amostral

A função de densidade de S^2 é dada por

$$
\begin{aligned}
f_{S^2}(s^2) &= f_U\left[\frac{(n-1)}{\sigma^2}s^2\right] \times \left|\frac{n-1}{\sigma^2}\right|, \quad \text{para} \quad s^2 > 0 \\
&= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{1}{2}\right)^{\frac{n-1}{2}}\left[\frac{(n-1)}{\sigma^2}s^2\right]^{\frac{n-1}{2}-1}\exp\left\{-\frac{1}{2}\frac{(n-1)}{\sigma^2}s^2\right\}\frac{(n-1)}{\sigma^2}I_{(0,\infty)} \\
&= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{1}{2}\right)^{\frac{n-1}{2}}\left[\frac{(n-1)}{\sigma^2}\right]^{\frac{n-1}{2}}\left[\frac{(n-1)}{\sigma^2}\right]^{-1}(s^2)^{\frac{n-1}{2}-1}\times \\
&\times \quad \frac{(n-1)}{\sigma^2}\exp\left\{-\frac{(n-1)}{2\sigma^2}s^2\right\} \times I_{(0,\infty)}(s^2) \\
&= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\left[\frac{(n-1)}{2\sigma^2}\right]^{\frac{n-1}{2}}(s^2)^{\frac{n-3}{2}}\exp\left\{-\frac{(n-1)}{2\sigma^2}s^2\right\} \times I_{(0,\infty)}(s^2)
\end{aligned}
$$

Variância amostral

- Todos os resultados desta subseção são desenvolvidos para o caso de populações Normais.
- Pode ser provado que para nenhuma outra distribuição:

1. A média e a variância amostral são independentemente distribuídas.
2. A média amostral possui exatamente a distribuição Normal

Distribuição F

Distribuição F

- Apresenta considerável interesse prático.
- É a distribuição da razão de duas variáveis aleatórias Qui-Quadrado independentes e divididas pelos seus respectivos graus de liberdade.
- Suponha que $U \sim \chi^2_m$ e $V \sim \chi^2_n$ são variáveis aleatórias independentes. A densidade conjunta será dada por:

$$f_{U,V}(u,v) = f_U(u) f_V(v)$$

$$\begin{aligned}
= \quad & \frac{1}{\Gamma\left(\frac{m}{2}\right)}\left(\frac{1}{2}\right)^{\frac{m}{2}} u^{\frac{m}{2}-1} \exp\left(-\frac{1}{2}u\right) I_{(0,\infty)}(u) \\
& \times \frac{1}{\Gamma\left(\frac{n}{2}\right)}\left(\frac{1}{2}\right)^{\frac{n}{2}} v^{\frac{n}{2}-1} \exp\left(-\frac{1}{2}v\right) I_{(0,\infty)}(v) \\
= \quad & \frac{1}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)2^{\frac{m+n}{2}}} u^{\frac{m}{2}-1} v^{\frac{n}{2}-1} \exp\left(-\frac{1}{2}(u+v)\right) I_{(0,\infty)}(u) I_{(0,\infty)}(v)
\end{aligned}$$

Distribuição F

- Desejamos encontrar a distribuição da quantidade

$$X = \frac{\frac{U}{m}}{\frac{V}{n}}$$

conhecida como `razão de variâncias`.

- Para encontrar a distribuição de X, considere:

$$X = \frac{\frac{U}{m}}{\frac{V}{n}} = \frac{n}{m}\frac{U}{V} \quad \text{e} \quad Y = V$$

$$U = \frac{m}{n} Y X$$

- Jacobiano da Transformação:

$$\begin{vmatrix} \frac{\partial U}{\partial X} & \frac{\partial U}{\partial Y} \\ \frac{\partial V}{\partial X} & \frac{\partial V}{\partial Y} \end{vmatrix} = \begin{vmatrix} \frac{mY}{n} & \frac{mX}{n} \\ 0 & 1 \end{vmatrix} = \frac{mY}{n}$$

Distribuição F

- A distribuição conjunta de X e Y é dada por:

$$
\begin{aligned}
f_{X,Y}(x,y) &= f_{U,V}\left(\frac{m}{n}yx, y\right) \cdot \left|\frac{my}{n}\right| \cdot I_{(0,\infty)}(x) \cdot I_{(0,\infty)}(y) \\
&= \frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})2^{\frac{m+n}{2}}} \left(\frac{m}{n}yx\right)^{\frac{m}{2}-1} y^{\frac{n}{2}-1} \\
&\times \exp\left\{-\frac{1}{2}\left[\frac{m}{n}yx + y\right]\right\} \frac{my}{n} I_{(0,\infty)}(x) I_{(0,\infty)}(y)
\end{aligned}
$$

Distribuição F

- A distribuição marginal de X é

$$
\begin{aligned}
&= \int_0^{\infty} f_{X,Y}(x,y)\,dy = \frac{1}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)2^{\frac{m+n}{2}}}\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m-2}{2}} \\
&\times \int_0^{\infty} y^{\frac{m+n-2}{2}} \exp\left\{-\frac{1}{2}y\left(\frac{m}{n}x+1\right)\right\}dy \\
&= \frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} 2^{-\frac{m+n}{2}} \left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m-2}{2}} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\left[\frac{1}{2}\left(\frac{m}{n}x+1\right)\right]^{\frac{m+n}{2}}} \\
&\times \int_0^{\infty} \frac{\left[\frac{1}{2}\left(\frac{m}{n}x+1\right)\right]^{\frac{m+n}{2}}}{\Gamma\left(\frac{m+n}{2}\right)} y^{\frac{m+n}{2}-1} \exp\left\{-y\frac{1}{2}\left(\frac{m}{n}x+1\right)\right\}dy
\end{aligned}
$$

Distribuição F

- Dentro da integral temos a densidade da

$$Ga\left[\frac{m+n}{2}, \frac{1}{2}\left(\frac{m}{n}x+1\right)\right]$$

- Portanto, esta integral resulta em 1. Logo,

$$f_X(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{x^{\frac{m-2}{2}}}{\left(1+\frac{m}{n}x\right)^{\frac{m+n}{2}}} I_{(0,\infty)}(x)$$

Distribuição F

Definição 9 (Distribuição F) Se X é uma variável aleatória com a f.d $f_x(x)$ do slide anterior, então X é definida como uma variável aleatória apresentando distribuição F com m e n graus de liberdade.

- `Notação:` $X \sim F_{m,n}$

Importante:

- A ordem na qual os graus de liberdade são indicados é importante visto que a f.d da distribuição não é simétrica em relação a m e n.
- Se a variável aleatória F é obtida a partir da razão de duas variáveis aleatórias Qui-Quadrado independentes, divididas pelos seus respectivos graus de liberdade, então o grau de liberdade da variável no numerador deve ser indicado primeiro na notação da F.

Distribuição F

Teorema 9 Seja U uma variável aleatória χ^2_m e V uma variável aleatória χ^2_n. Assuma também que U e V são independentes. Então:

$$X = \frac{U/m}{V/n} \quad \text{tem distribuição} \quad F_{m,n}.$$

Corolário 5 Se $X_1, \ldots, X_{m+1}$ é uma amostra aleatória de tamanho $m+1$ de uma população $N(\mu_X, \sigma^2)$. Considere $Y_1, \ldots, Y_{n+1}$ uma amostra aleatória de tamanho $n+1$ de uma população $N(\mu_Y, \sigma^2)$.

- Assuma que as duas amostras sejam independentes, então:

$$\frac{1}{\sigma^2}\sum_{j=1}^{m+1}(X_j - \bar{X})^2 \sim \chi^2_m \quad \text{e} \quad \frac{1}{\sigma^2}\sum_{i=1}^{n+1}(Y_i - \bar{Y})^2 \sim \chi^2_n$$

e a estatística

$$\frac{\sum_{j=1}^{m+1}(X_j - \bar{X})^2/m}{\sum_{i=1}^{n+1}(Y_i - \bar{Y})^2/n} \sim F_{m,n}.$$

Distribuição F

Observação 1:

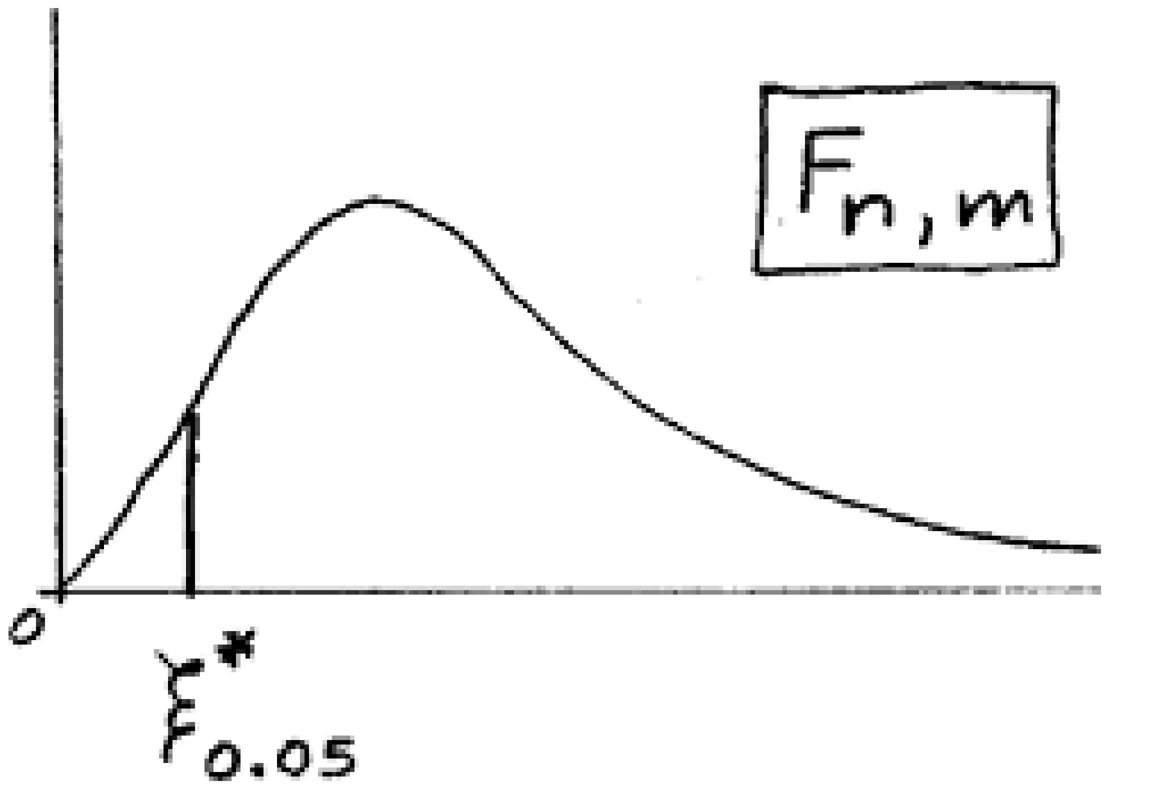
- Se $X \sim F_{m,n}$, então $E(X) = \frac{n}{n-2} \quad \text{para} \quad n > 2$

e

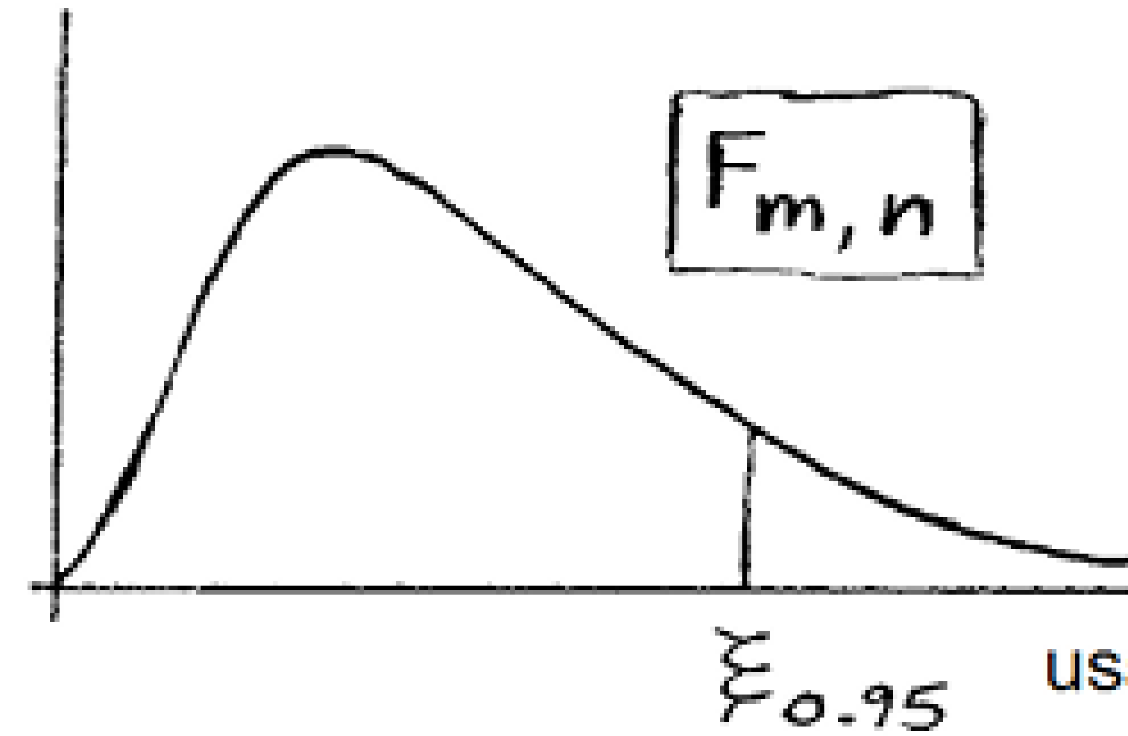
$$\text{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad \text{para} \quad n > 4.$$

Observação 2:

Distribuição F



Este valor não é encontrado na Tabela da $F_{n,m}$



usamos a tabela $F_{m,n}$ para determinar $\xi^*_{0.05}$

$$\xi^*_{0.05} = 1/\xi_{0.95}$$

Distribuição t-Student

Distribuição t-Student

- Outra distribuição de considerável importância prática.
- Seja $Z \sim N(0,1)$ e $U \sim \chi_k^2$ v.a. independentes.
- Deseja-se encontrar a distribuição de

$$X = \frac{Z}{\sqrt{U/k}}.$$

- A densidade conjunta de Z e U é dada por

$$\begin{aligned} f_{Z,U}(z,u) &= f_Z(z) f_U(u) \\ &= (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}z^2\right\} \frac{1}{\Gamma(k/2)} \\ &\times \left(\frac{1}{2}\right)^{k/2} u^{\frac{k}{2}-1} \exp\left\{-\frac{1}{2}u\right\} I_{(-\infty,\infty)}(v) I_{(0,\infty)}(u) \end{aligned}$$

Distribuição t-Student

- Considere a seguinte transformação:

$$X = \frac{Z}{\sqrt{U/k}}, \quad Y = U \quad \text{e} \quad Z = X\sqrt{Y/k}.$$

Então,

$$\begin{vmatrix} \frac{\partial Z}{\partial X} = \sqrt{\frac{Y}{k}} & \frac{\partial U}{\partial X} = 0 \\ \frac{\partial Z}{\partial X} = \frac{X}{2k}\left(\frac{Y}{k}\right)^{-1/2} & \frac{\partial U}{\partial Y} = 1 \end{vmatrix} = \sqrt{\frac{Y}{k}}$$

- Logo,

$$\begin{aligned} f_{XY}(x,y) &= f_{Z,U}(x\sqrt{y/k}, y) \times \sqrt{\frac{y}{k}} \\ &= (2\pi)^{-1/2}\exp\left\{-\frac{1}{2}\frac{x^2 y}{k}\right\}\frac{1}{\Gamma(k/2)}\left(\frac{1}{2}\right)^{k/2} y^{\frac{k}{2}-1}\exp\left\{-\frac{1}{2}y\right\}\left(\frac{y}{k}\right)^{1/2} I_{(-\infty,} \end{aligned}$$

Distribuição t-Student

$$
\begin{aligned}
f_X(x) &= \int_0^{\infty} f_{XY}(x,y)dy \\
&= (2k\pi)^{-1/2}\frac{1}{\Gamma(k/2)}\left(\frac{1}{2}\right)^{k/2}\int_0^{\infty} y^{\frac{k}{2}+\frac{1}{2}-1}\exp\left\{-\frac{1}{2}y\left[\frac{x^2}{k}+1\right]\right\}dy
\end{aligned}
$$

- Dentro da integral temos o núcleo da: $Ga\left[\frac{k+1}{2}; \frac{1}{2}\left(\frac{x^2}{k}+1\right)\right]$

Distribuição t-Student

$$
\begin{aligned}
f_X(x) &= (2k\pi)^{-1/2}\frac{1}{\Gamma(k/2)}\left(\frac{1}{2}\right)^{k/2}\frac{\Gamma\left(\dfrac{k+1}{2}\right)}{\left[\dfrac{1}{2}\left(\frac{x^2}{k}+1\right)\right]^{\dfrac{k+1}{2}}}\times \\
&\times \underbrace{\int_0^\infty \frac{\left[\dfrac{1}{2}\left(\frac{x^2}{k}+1\right)\right]^{\dfrac{k+1}{2}}}{\Gamma\left(\dfrac{k+1}{2}\right)} y^{\frac{k+1}{2}-1}\exp\left\{-\frac{1}{2}y\left[\frac{x^2}{k}+1\right]\right\}dy}_{1}
\end{aligned}
$$

Distribuição t-Student

- Assim,

$$
\begin{aligned}
f_X(x) &= (2k\pi)^{-1/2}\frac{1}{\Gamma(k/2)}\left(\frac{1}{2}\right)^{k/2}\frac{\Gamma\left(\dfrac{k+1}{2}\right)}{\left[\dfrac{1}{2}\left(\frac{x^2}{k}+1\right)\right]^{\dfrac{k+1}{2}}} \\
&= \frac{\Gamma\left(\dfrac{k+1}{2}\right)}{\Gamma\left(\dfrac{k}{2}\right)}\frac{1}{\sqrt{k\pi}}\frac{1}{\left(\frac{x^2}{k}+1\right)^{\dfrac{k+1}{2}}}I_{(-\infty,\infty)}(x).
\end{aligned}
$$

Distribuição t-Student

Teorema 10 (Distribuição t-Student) Se X é uma v.a. com f.d. igual a anterior, então dizemos que X segue a distribuição t-Student com k-graus de liberdade.

- `Notação:` t_k.

Teorema 11 Se $Z \sim N(0,1)$ e $U \sim \chi^2_k$ são v.a's independentes, então

$$X = \frac{Z}{\sqrt{U/k}} \sim t_k.$$

Distribuição t-Student

O seguinte corolário mostra como o resultado do **Teorema 10** é aplicável quando consideramos a amostragem a partir da distribuição Normal.

Corolário 6 Considere $X_1, X_2, \ldots, X_n$ uma a.a. da $N(\mu, \sigma^2)$, então $Z = (\overline{X} - \mu)/(\sigma/\sqrt{n}) \sim N(0,1)$ e $\dfrac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X})^2 \sim \chi^2_{n-1}$.

- Além disso Z e U são independentes (`Teorema 8`). Sendo assim, a seguinte v.a

$$\frac{Z}{\sqrt{U/n-1}} = \frac{(\overline{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{\dfrac{1}{\sigma^2}\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}}} = \frac{\sqrt{n(n-1)}(\overline{X}-\mu)}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}} = \frac{(\overline{X}-\mu)}{\sqrt{S^2/n}} \sim t_{n-1}.$$

Distribuição t-Student

Observações:

- Quando o grau de liberdade é igual a 1, a `distribuição t-Student` se torna a `distribuição Cauchy`.
- Conforme o grau de liberdade cresce, a distribuição t-Student se aproxima da distribuição $N(0, 1)$.
- O quadrado de uma variável aleatória com distribuição t_k terá distribuição $F_{1,k}$. Isto é, se $Y \sim t_k$ então $Y^2 \sim F_{1,k}$.
- Se X é uma variável aleatória com distribuição t_k, então

$$E(X) = 0, \quad \text{se } k > 1 \text{ e } Var(X) = \frac{k}{k-2}, \quad \text{se } k > 2.$$