Métodos Estocásticos da Engenharia II

Capítulo 3 - Distribuições Amostrais e Estimação Pontual de Parâmetros

Prof. Magno Silvério Campos

2024/2

Bibliografia

Essas notas de aulas foram baseadas nas seguintes obras:

1. CANCHO, V.G. *Notas de Aulas sobre Noções de Estatística e Probabilidade*. São Paulo: USP, 2010.
2. HINES, W.W.; *et al. Probabilidade e Estatística na Engenharia*. 4. ed. Rio de Janeiro: LTC, 2006.
3. MONTGOMERY, D.C.; RUNGER, G.C. *Estatística Aplicada e Probabilidade para Engenheiros*. 6. ed. Rio de Janeiro: LTC, 2016.

Aconselha-se pesquisá-las para se obter um maior aprofundamento e um melhor aproveitamento nos estudos.

Conteúdo Programático

1. Seção 1 - Introdução
2. Seção 2 - Conceitos gerais de estimação pontual
3. Seção 3 - Método da Máxima Verossimilhança
4. Seção 4 - Distribuições amostrais
    - Distribuição da média amostral;
    - Teorema do Limite Central;
    - Distribuição da diferença de duas médias amostrais;
    - Distribuição amostral de uma proporção amostral.
5. Seção 5* - Outras distribuições utilizadas na inferência estatística
    - Distribuição Qui-quadrado;
    - Distribuição t-Student;
    - Distribuição F-Snedecor;

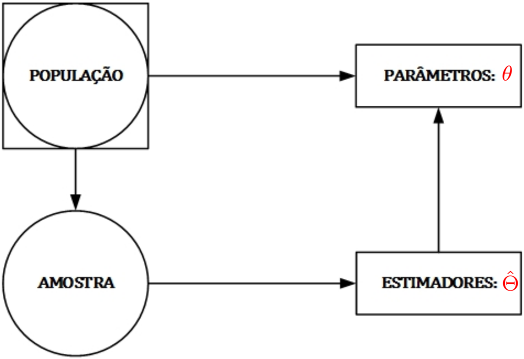
Introdução

Amostra

Se uma população for numerosa, talvez não seja possível estudar cada um dos seus elementos. Portanto, é muito difícil ter uma informação completa sobre ela. Nesses casos, recorre-se à informação proporcionada por uma parte finita da população, denominada **amostra.**

Uma amostra aleatória tem a propriedade de refletir as características da população da qual foi sorteada.

Esquema geral



Amostra aleatória - [Montgomery e Runger(2016)]

As variáveis aleatórias $X_1, X_2, \ldots, X_n$ constituem uma amostra aleatória de tamanho n de uma população $X \sim f(x, \theta)$, se: (a) cada X_i é uma variável aleatória independente e (b) cada X_i tem a mesma distribuição de probabilidade $f(x, \theta)$.

Amostra aleatória- [Cancho(2010)]

A definição de amostra aleatória é satisfeita nos seguintes casos:

1. Quando a amostra provem de uma população infinita e quando a amostra é sorteada ao acaso com reposição de uma população finita.
2. Quando se sorteia a amostra sem reposição de uma população finita, evidentemente não se satisfaz a definição de amostra aleatória, pois as variáveis aleatórias $X_1, \ldots, X_n$ não são independentes. Porém, se o tamanho da amostra é muito pequeno em comparação com o tamanho da população, a definição é satisfeita aproximadamente.

Estimadores pontuais - ([Montgomery e Runger(2016)]

Em geral, se X for uma variável aleatória com distribuição de probabilidades $f(x)$, caracterizada por um parâmetro desconhecido θ, e se $X_1, X_2, \ldots, X_n$ for uma amostra aleatória de tamanho n, então a estatística $\hat{\Theta} = h(X_1, X_2, \ldots, X_n)$ é chamada de um estimador pontual de θ.

Note que:

$\hat{\Theta}$ é uma variável aleatória, porque ela é função de variáveis aleatórias.

Exemplo - ([Montgomery e Runger(2016)]

Suponha que a variável aleatória X seja normalmente distribuída, com uma média desconhecida μ. A média da amostra é um estimador da média desconhecida μ da população. Isto é,

$$\hat{\mu} = \bar{X}$$

Estimativas pontuais

Depois da amostra ter sido escolhida, $\hat{\Theta}$ assume um valor numérico $\hat{\theta}$, denominado estimativa pontual de θ.

Exemplo

No exemplo anterior, depois da amostra ter sido escolhida, o valor numérico $\bar{x}$ é a estimativa pontual de μ. Assim, se $x_1 = 30, x_2 = 20, x_3 = 15, x_4 = 40$ e $X_5 = 25$, então a estimativa de μ é

$$\bar{x} = \frac{30+20+15+40+25}{5} = 26$$

Analogamente, se a variância da população σ^2 for também desconhecida, um estimador pontual para σ^2 será a variância da amostra S^2 e o valor numérico deste, calculado a partir dos dados da amostra, é chamado de estimativa pontual de σ^2.

Problemas de estimação frequentes em engenharia

Geralmente, necessitamos estimar:

- a média μ de uma população;
- a variância σ^2 de uma população;
- a proporção p de itens em uma população que pertencem a uma classe de interesse;
- a diferença nas médias de duas populações, $\mu_1 - \mu_2$;
- a diferença nas proporções de duas populações, $p_1 - p_2$;
- entre outros.

Estimativas pontuais razoáveis:

- Para μ, a estimativa é $\hat{\mu} = \bar{x}$, a média da amostra;
- Para σ^2, a estimativa é $\hat{\sigma^2} = s^2$, a variância da amostra;
- Para p, a estimativa é $\hat{p} = x/n$, a proporção da amostra;
- Para $\mu_1 - \mu_2$, a estimativa é $\hat{\mu_1} - \hat{\mu_1} = \bar{x_1} - \bar{x_2}$, a diferença entre as médias de duas AAI;
- Para $p_1 - p_2$, a estimativa é $\hat{p_1} - \hat{p_2}$, a diferença entre as proporções de duas AAI.

Conceitos gerais de estimação pontual

[1] - Tendência de um estimador

O estimador $\hat{\Theta}$ é um estimador não-tendencioso para o parâmetro θ, se

$$E(\hat{\Theta}) = \theta$$

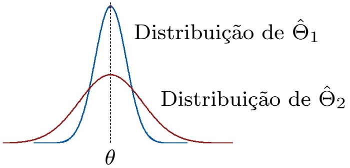
Se o estimador for tendencioso, então a tendência é dada por:

$$E(\hat{\Theta}) - \theta.$$

Exemplo: mostrar que a média amostral $\bar{X}$ é um estimador não tendencioso da média populacional, μ.

[2] - Variância de um estimador

Embora possam existir muitos estimadores $\hat{\Theta}$ não tendenciosos para o parâmetro θ, deve-se escolher aquele que tiver variância mínima, isto é, o Estimador Não Tendencioso de Variância Mínima (ENTVM) é o mais provável de produzir uma estimativa mais próxima do valor verdadeiro de θ.



Exemplo: considere dois estimadores para a média populacional, μ: a média amostral, $\bar{X}$, e uma única observação, X_i, proveniente da amostra. Neste caso, qual dos dois é melhor?

[3] - Erro-padrão de um estimador

O erro-padrão de um estimador $\hat{\Theta}$ é o seu desvio-padrão, isto é,

$$\sigma_{\hat{\Theta}} = \sqrt{V(\hat{\Theta})}.$$

Essa é uma medida de precisão da estimação muito útil. Caso sejam utilizadas estimativas amostrais, a notação muda para $\hat{\sigma}_{\hat{\Theta}} = SE(\hat{\Theta})$.

Exemplo: considere uma amostra aleatória proveniente de uma população normal com média μ e variância σ^2. Determine o erro-padrão da média amostral, $\bar{X}$, para os casos de variância σ^2 conhecida ou desconhecida.

Método da Máxima Verossimilhança

Existem alguns métodos de estimação pontual conhecidos na literatura estatística. Um dos melhores métodos para obter estimadores pontuais de parâmetros populacionais é o Método da Máxima Verossimilhança, que veremos a seguir.

Seja X uma variável aleatória que segue uma distribuição de probabilidades $f(x, \boldsymbol{\Theta})$, onde $\boldsymbol{\Theta}$ é um vetor de parâmetros θ desconhecidos.

Sejam $x_1, x_2, \dots, x_n$ os valores observados em uma amostra aleatória de tamanho n. A **função de verossimilhança** da amostra é definida como:

$$L(\boldsymbol{\Theta}) = f(x_1, \boldsymbol{\Theta}) \cdot f(x_2, \boldsymbol{\Theta}) \cdot \ldots \cdot f(x_n, \boldsymbol{\Theta}).$$

A função acima informa sobre a possibilidade (verossimilhança) de a variável X assumir os valores $x_1, x_2, \dots, x_n$; Para o caso de VAD, a verossimilhança é um valor de probabilidade.

A função de verossimilhança é função apenas do parâmetro desconhecido $\boldsymbol{\Theta}$. O **Estimador de Máxima Verossimilhança (EMV)** de $\boldsymbol{\Theta}$ é o valor de $\boldsymbol{\Theta}$ que maximiza $L(\boldsymbol{\Theta})$. Tal valor é obtido derivando $L(\boldsymbol{\Theta})$ em relação a $\boldsymbol{\Theta}$ e igualando o resultado a 0.

Fato: $L(\boldsymbol{\Theta})$ e $\ln L(\boldsymbol{\Theta})$ apresentam seus máximos no mesmo valor de $\boldsymbol{\Theta}$. Assim,

$$\frac{\partial L(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}} = 0 \quad \Rightarrow \quad \frac{\partial \ln L(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}} = 0$$

O Estimador de Máxima Verossilmilhança ($\hat{\boldsymbol{\Theta}}$) apresenta, em geral, propriedades assintóticas favoráveis:

- $\hat{\boldsymbol{\Theta}}$ é não tendencioso para valores grandes de n;
- $\hat{\boldsymbol{\Theta}}$ apresenta uma variância tão pequena quanto possível de ser obtida com qualquer outro estimador;
- $\hat{\boldsymbol{\Theta}}$ tem distribuição aproximadamente normal.

Exemplo 1 - [Montgomery e Runger(2016)]

Tempos até a falha de certo equipamento eletrônico seguem uma distribuição exponencial com parâmetro λ e *fdp* dada por:

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

Obtenha o estimador de máxima verossimilhança de λ.

Exercício - [Montgomery e Runger(2016)]

Tempos até a falha de certo equipamento eletrônico seguem uma distribuição exponencial. Alguns destes tempos foram anotados de forma contínua, obtendo-se os seguintes valores:

48; 80; 122; 188; 189; 220; 253; 311; 325; 358; 490; 495; 513; 723; 773; 879; 1510; 1674; 1809; 2005; 2028; 2038; 2870; 3103; 3205.

Determine a estimativa de λ de acordo com o método da máxima verossimilhança.

Exemplo 2 - [Montgomery e Runger(2016)]

Considere uma variável aleatória normalmente distribuída, com parâmetros μ e σ^2 desconhecidos, com *fdp* dada por:

$$f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty. \tag{2}$$

Obtenha os estimadores de máxima verossimilhança para μ e σ^2.

Observação 1:

Para utilizar a estimação pelo Método da Máxima Verossimilhança, é necessário o conhecimento da distribuição de probabilidades da população.

Observação 2:

Complicações podem surgir na obtenção do máximo da função de verossimilhança, tornando necessário o emprego de métodos computacionais avançados.

Distribuição da média amostral

Se de uma população com média μ_X e variância σ_X^2 se extraem amostras aleatórias de tamanho n e para cada amostra determina-se a média

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

então a média e variância da variável $\bar{X}$ são dadas por:

a) Se a amostragem é com reposição de uma população finita (ou amostragem com ou sem reposição em uma população infinita).

$$\mu_{\bar{X}} = \mu_X \quad e \quad \sigma^2_{\bar{X}} = \frac{\sigma^2_X}{n}$$

b) Se a amostragem é sem reposição de uma população finita com N elementos.

$$\mu_{\bar{X}} = \mu_X \;\; e \;\; \sigma^2_{\bar{X}} = \frac{\sigma^2_X}{n}\left[\frac{N-n}{N-1}\right]$$

Observação

Se a fração de amostragem $f = \frac{n}{N}$ é pequena $(f < 0,1)$ e o tamanho da população (N) é grande, a variância da média amostral em (b) é aproximada com a expressão do caso (a), isto é,

$$\sigma^2_{\bar{X}} = \frac{\sigma^2_X}{n}$$

Exemplo

Considere os seguintes níveis X de estoque para 4 produtos em um determinado dia:

Produto	A	B	C	D
X	0	2	0	1

Obter a média e a variância da média amostral para amostragem com ou sem reposição, quando $n = 2$.

Forma da distribuição da média amostral quando a população é normal

Seja X uma variável aleatória que tem uma distribuição normal com média μ_X e variância σ_X^2. Se desta distribuição selecionam-se amostras aleatórias de tamanho n, a média amostral,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

é uma combinação linear de variáveis X_i, todas elas com distribuição $N(\mu_X, \sigma_X^2)$ e independentes entre si (o fato da distribuição de X ser normal presume, em rigor que a população é infinita e que, portanto, não há diferença entre escolher uma amostra com e sem reposição [Cancho(2010)].

Vale relembrar que:

Combinação linear de variáveis aleatórias normais

Sejam $X_1, \ldots, X_n$, n variáveis aleatórias independentes onde $X_i \sim N(\mu_i; \sigma_i^2)$ para $i = 1, \ldots, n$ e sejam $a_1, \ldots, a_n$ constantes reias. Seja a variável aleatória Y uma combinação linear das variáveis aleatórias normais, $X_1, \ldots, X_n$. Isto é,

$$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n.$$

Então a variável aleatória Y tem distribuição normal com média

$$\mu_Y = a_1\mu_1 + a_2\mu_2 \cdots + a_n\mu_n = \sum_{i=1}^{n} a_i\mu_i$$

e variância

$$\sigma_Y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 \cdots + a_n^2\sigma_n^2 = \sum_{i=1}^{n} a_i^2\sigma_i^2.$$

Ou seja:

Uma combinação linear de variáveis normais independentes também é normal!

Portanto,

a média amostral segue uma distribuição normal com média μ_X e variância, $\frac{\sigma_X^2}{n}$. Isto é,

$$\bar{X} \sim N(\mu_X, \sigma_X^2/n).$$

Embora este resultado seja de extrema importância, ele é relativamente limitado, já que, somente permite especificar a distribuição da média amostral no caso de uma população normal.

Forma da distribuição da média amostral quando a população é normal

Exemplo - [Montgomery e Runger(2016)]

Uma empresa fabrica resistores que têm uma resistência média de 100 ohms e um desvio-padrão de 10 ohms. A distribuição de resistência é normal. Encontre a probabilidade de uma amostra aleatória de $n = 25$ resistores ter uma resistência média menor que 95 ohms.

Forma da distribuição da média amostral quando a população não é normal

Observação 1

Muitas vezes não temos informação a respeito da distribuição das variáveis que constituem a amostra, fato que nos impede de utilizar o resultado apresentado.

Observação 2

Felizmente, satisfeitas certas condições, pode ser mostrado que para uma amostra suficientemente grande, a distribuição de probabilidade da média amostral pode ser aproximada por uma distribuição normal, com média e variância iguais àquelas calculadas anteriormente. Este fato é um dos teoremas mais importantes da estatística e probabilidade e é denominado o **Teorema do Limite Central** [Montgomery e Runger(2016)].

Forma da distribuição da média amostral quando a população não é normal

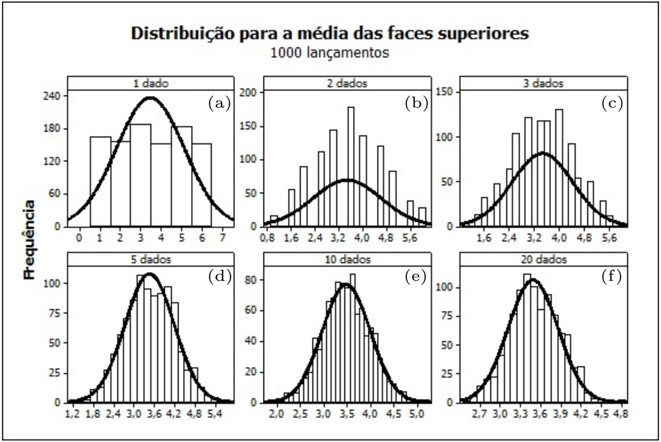
Teorema do Limite Central

Se $X_1, X_2, \ldots, X_n$ for uma amostra aleatória de tamanho n, retirada de uma população (finita ou infinita), com média μ e variância σ^2, e se $\bar{X}$ for a média da amostra, então $\bar{X}$ tem distribuição aproximadamente normal com média μ_X e variância σ_X^2/n, para n suficientemente grande $(n \to \infty)$. Isto é,

$$Z = \frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}} \xrightarrow{n \to \infty} N(0, 1).$$

Teorema do Limite Central

A aproximação normal para $\bar{X}$ depende do tamanho n da amostra. A figura a seguir mostra a distribuição obtida para 1000 arremessos de dados:



- as figuras de (a) a (f) mostram, respectivamente, a distribuição das médias obtidas para o arremesso de 1, 2, 3, 5, 10 e 20 dados.

Observações

1. Em muitos casos de interesse prático, se $n \geq 30$, a aproximação normal será satisfatória, independentemente da forma da população;
2. Se $n < 30$, o teorema do limite central funcionará, se a distribuição da população não for muito diferente da normal.

Quando está presente uma considerável assimetria, de acordo com [Montgomery e Runger(2016)], a aplicação do teorema do limite central para a descrição da distribuição amostral deve ser interpretada com cuidado!

Uma regra empírica que tem sido usada com sucesso em pesquisas por amostragem, onde é comum tal comportamento da variável (β_3), é que

$$n > 25.(\beta_3)^2.$$

Exemplo 1 - [Cancho(2010)]

Suponha que na produção em série de um artigo, a massa é uma variável aleatória com uma média de 950 g e uma variância de 1600 g^2. Selecionam-se aleatoriamente e com reposição 36 artigos. Calcular a probabilidade da média amostral ser maior que 965 g.

Exemplo 2 - [Montgomery e Runger(2016)]

Suponha que uma variável aleatória X tenha uma distribuição contínua uniforme

$$f(x) = \begin{cases} 1/2, & \text{se } 4 \leq x \leq 6 \\ 0, & \text{caso contrário.} \end{cases}$$

Encontre a distribuição da média de uma amostra aleatória de tamanho $n = 40$.

Distribuição da diferença de duas médias amostrais

Seja $X_1, X_2, \ldots, X_{n_1}$ uma amostra aleatória de tamanho n_1 de uma população com característica X que tem distribuição normal com média μ_1 e variância σ_1^2.

Seja também, $Y_1, Y_2, \ldots, Y_{n_2}$ outra amostra aleatória de tamanho n_2, de uma população com a característica Y que tem distribuição normal com média μ_2 e variância σ_2^2.

Se X e Y são independentes, então a diferença amostral $\bar{X} - \bar{Y}$ tem distribuição normal com média $\mu_1 - \mu_2$ e variância $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$. Isto é,

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1). \tag{3}$$

Observação

Se as populações onde foram retiradas as amostras não tiverem distribuição normal, pelo teorema do limite central, segue válido o resultado se os tamanhos amostrais n_1 e n_2 forem suficientemente grandes, isto é $n_1 \geq 30$ e $n_2 \geq 30$.

Exemplo - [Montgomery e Runger(2016)]

A vida de um componente usado em um motor de uma turbina de um avião é uma variável aleatória, com média de 5000 horas e desvio-padrão de 40 horas. A distribuição da vida é aproximada pela distribuição normal.

O fabricante do motor introduz uma melhoria no processo de fabricação para esse componente, que aumenta a vida média para 5050 horas e diminui o desvio-padrão para 30 horas. Considere novamente a normalidade para essa nova vida.

Suponha que uma amostra aleatória de $n = 16$ componentes seja selecionada do processo antigo e uma amostra aleatória de $m = 25$ componentes seja selecionada do processo novo.

Qual é a probabilidade de que a diferença nas duas médias amostrais $\bar{X}_2 - \bar{X}_1$ seja no mínimo de 25 horas? Considere que o processo antigo e o melhorado possam ser considerados como populações independentes.

Distribuição amostral de uma proporção amostral

Considere uma população dicotômica, constituída apenas por elementos de dois tipos, isto é, cada elemento pode ser classificado como *sucesso* ou *fracasso*. Suponha que a probabilidade de sucesso seja p e de fracasso seja $q = 1 - p$.

Dessa população retira-se uma amostra aleatória de n observações, $X_1, \ldots, X_n$, onde

$$X_i = \begin{cases} 1, & \text{se } \textit{sucesso} \\ 0, & \text{caso contrário.} \end{cases}$$

Seja a variável aleatória Y que conta o n^o de sucessos na amostra. Então:

1. $Y = \sum_{i=1}^{n} X_i$ tem distribuição Binomial com parâmetros n e p.
2. A proporção amostral de sucessos é: $\hat{p} = \frac{Y}{n} = \sum_{i=1}^{n} X_i/n = \bar{X}$. De (1), a distribuição de probabilidade de $\hat{p}$ é:

$$P(\hat{p} = \frac{y}{n}) = P(\frac{Y}{n} = \frac{y}{n}) = P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}.$$

E para n suficientemente grande (teorema do limite central), a proporção amostral tem distribuição aproximadamente normal com média p e variância $\frac{pq}{n}$. Isto é,

$$\hat{p} \sim N(p, \frac{pq}{n}).$$

Exemplo - [Cancho(2010)]

Uma empresa tem um número grande de funcionários. A probabilidade de que um empregado selecionado ao acaso, participe de um programa de treinamento é 0,40.

(a) Se 10 funcionários são escolhidos ao acaso, qual é a probabilidade que proporção de participantes seja
  (a1) exatamente 60%?
  (a2) pelo menos 80%?

(b) suponha que 100 funcionários são escolhidos ao acaso. Qual é a probabilidade de que a proporção amostral de participantes do programa seja maior que 50%?

Observação

Os resultados acima são válidos também nos seguintes casos:

1. Para uma população infinita, qualquer que seja o tipo de amostragem.
2. Para população finita, com amostragem com reposição.

Se a amostragem é sem reposição, em uma população finita de N elementos, a distribuição exata de probabilidade $\hat{p}$ é uma distribuição Hipergeométrica. Isto é,

$$P(\hat{p} = \frac{y}{n}) = \frac{\binom{M}{y}\binom{N-M}{n-y}}{\binom{N}{n}} \tag{4}$$

A variância de $\hat{p}$ é ajustada através do fator de correção de população finita, isto é,

$$Var(\hat{p}) = \frac{pq}{n}\left(\frac{N-n}{N-1}\right).$$

Se, n é suficientemente grande, pelo teorema central do limite, a variável aleatória,

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}\left(\frac{N-n}{N-1}\right)}},$$

tem distribuição aproximadamente normal padrão.

Exemplo - [Cancho(2010)]

Informações anteriores mostram que 10% do lote de peças para uma máquina são defeituosas. Suponha que um lote de 5000 peças foi adquirido. Seleciona-se uma amostra de 400 peças, ao acaso e sem reposição. Qual a probabilidade da amostra ter:

(a) entre 9% e 10% de peças defeituosas?

(b) menos de 8% de peças defeituosas?

Distribuição Qui-quadrado (χ^2)

Sejam $Z_1, Z_2, \ldots, Z_k$ variáveis aleatórias distribuídas normalmente e independentes com média $\mu = 0$ e variância $\sigma^2 = 1$. A variável aleatória

$$W = Z_1^2 + Z_2^2 + \cdots + Z_k^2 \tag{5}$$

tem distribuição Qui-quadrado com k graus de liberdade e sua função de densidade é dada por:

$$f(w) = \frac{1}{\Gamma(k/2)2^{k/2}} w^{\frac{k}{2}-1} e^{-\frac{w}{2}}, \quad w > 0 \tag{6}$$

onde $\Gamma(a)$ é uma função matemática definida

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \;\; para \;\; a > 0,$$

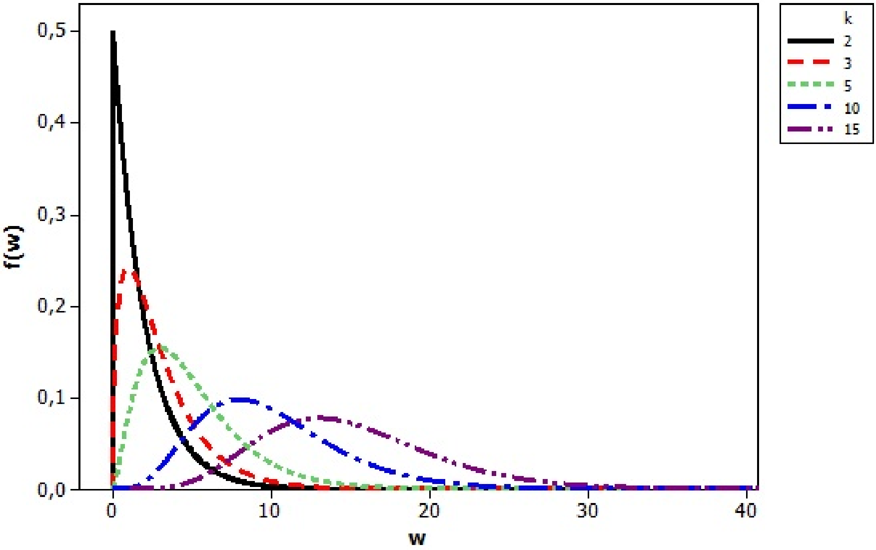
chamada de função gama.

Essa função satisfaz às seguintes propriedades:

$$\begin{aligned} \Gamma(a) &= (a-1)\Gamma(a-1) \\ \Gamma(a) &= (a-1)!, \quad \text{para } a \text{ inteiro} \\ \Gamma(1/2) &= \sqrt{\pi} \\ \Gamma(1) &= 0! = 1 \end{aligned}$$

O gráfico da distribuição Qui-quadrado para $k = 2, 3, 5, 10$ e 15 graus de liberdade é mostrado na figura a seguir:

Distribuição $\chi^2_{(k)}$



A notação $W \sim \chi^2_{(k)}$ é usada para indicar que a variável W tem distribuição Qui-quadrado com k graus de liberdade.

Propriedades

Se $W \sim \chi^2_{(k)}$

(a) $E(W) = k$ e $Var(W) = 2k$.

(b) A distribuição é assimétrica direita.

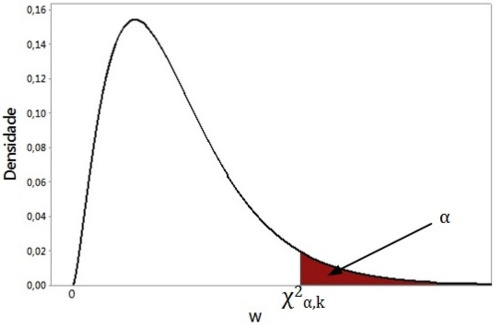
(c) A medida que aumentam-se os graus de liberdade, torna-se simétrica.

Uso da tabela Qui-quadrado

Na tabela $\chi^2_{(k)}$, tem-se os pontos críticos da distribuição $W \sim \chi^2_{(k)}$, denotado por $\chi^2_{\alpha,k}$ tal que a probabilidade

$$P(W > \chi^2_{\alpha,k}) = \int_{\chi^2_{\alpha,k}}^{\infty} f(w)dw$$

Pontos críticos $\chi^2_{\alpha,k}$ das distribuições $\chi^2_{(k)}$



A probabilidade

$$P(W > \chi^2_{\alpha,k}) = \int_{\chi^2_{\alpha,k}}^{\infty} f(w)dw = \alpha$$

é representada pela área sombreada da figura.

Para ilustrar o uso da tabela $\chi^2_{(k)}$, observe que as áreas α estão na primeira linha e na primeira coluna estão os graus de liberdade k. Portanto, o valor de χ^2 com 10 graus de liberdade e com área (probabilidade) 0,05 à direita é $\chi^2_{0{,}05{,}10} = 18,31$. Isto é,

$$P(W > \chi^2_{0{,}05{,}10}) = P(W > 18,31) = 0,05.$$

Exemplo 1 - [Cancho(2010)]

Se X é uma variável aleatória $\chi^2_{(17)}$, então obtenha:

(a) $P(X \geq 8,67)$;

(b) $P(X \leq 8,67)$;

(c) $P(6,41 < X < 27,59)$;

(d) o valor de a tal que $P(X < a) = 0,025$.

Exemplo 2 - [Cancho(2010)]

Se X é uma variável aleatória $\chi^2_{(4)}$, então obtenha $P(X \geq 7,932)$;

Propriedade reprodutiva

Se $W_1, W_2, \ldots, W_n$ são variáveis aleatórias independentes distribuídas cada uma com distribuição Qui-quadrado com $k_1, k_2, \ldots, k_n$ graus de liberdade respectivamente, então, a variável

$$W = W_1 + W_2 + \cdots + W_n$$

tem distribuição Qui-quadrado com $k = \sum_{i=1}^{n} k_i$ graus de liberdade

Exemplo

Se W_1, W_2 e W_3 são variáveis aleatórias independentes com distribuição Qui-quadrado respectivamente com 4, 7 e 9 graus de liberdade respectivamente, então $W = W_1 + W_2 + W_3 \sim \chi^2_{(20)}$.

Distribuição t-Student

Sejam Z e W duas variáveis independentes com distribuição normal padrão e Qui-quadrado com k graus de liberdade, respectivamente. A variável aleatória,

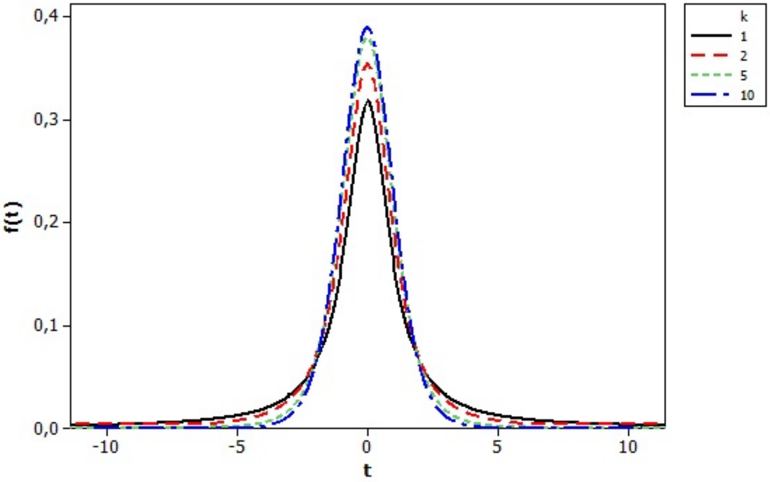
$$T = \frac{Z}{\sqrt{\frac{W}{k}}}$$

tem distribuição t-Student com k graus de liberdade. A função de densidade de probabilidade é dado por:

$$f(t) = \frac{\Gamma(\frac{k+1}{2})}{(k\pi)^{1/2}\Gamma(\frac{k}{2})} \left(1 + \frac{t^2}{k}\right)^{-(k+1)/2}$$

A notação $T \sim t(k)$ é usada para indicar que a variável T tem distribuição t-Student com k graus de liberdade.

Na figura abaixo é apresentado o gráfico da função de densidade de probabilidade, para $k = 1, 2, 5$ e 10 graus de liberdade.



Propriedades Se $T \sim t_{(k)}$.

(a)

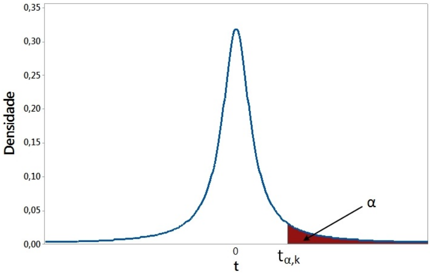
$$\begin{aligned} E(T) &= 0 \\ Var(T) &= \frac{k}{k-2}, \quad k > 2 \end{aligned}$$

(b) A distribuição é simétrica em torno de sua média.

(c) Se $k \to \infty$, $T \sim N(0,1)$.

Uso da tabela t-Student

A tabela t-Student proporciona os pontos críticos da distribuição t-Student. Seja $t_{\alpha,k}$ o valor da variável aleatória T com k graus de liberdade para o qual tem-se uma área (probabilidade) α. Portanto, $t_{\alpha,k}$ é um ponto crítico na cauda superior da distribuição t-Student com k graus de liberdade. Este ponto crítico aparece na figura abaixo.



A probabilidade

$$P(T > t_{\alpha,k}) = \int_{t_{\alpha,k}}^{\infty} f(t)dt = \alpha$$

é representada pela área sombreada da figura.

Na tabela t-Student, os valores de α encontram-se na primeira linha da tabela, enquanto os graus de liberdade aparecem na primeira coluna da parte esquerda. Para ilustrar o uso da tabela, observe que o valor de t-Student com 10 graus de liberdade que tem área de 0,05 à direita é $t_{0,05,10}$. Isto é,

$$P(T > t_{0,05,10}) = P(T > 1,812) = 0,05$$

Como, a distribuição t-Student é simétrica em relação a zero (média), tem-se que $t_{1-\alpha,k} = -t_{\alpha,k}$. Isto é, o valor da variável T que corresponde a uma área igual $(1-\alpha)$ à direita (e, portanto, uma área de α à esquerda) é igual ao negativo do valor de T, que tem área α na cauda direita da distribuição. Em conseqüência, $t_{0,95,10} = -t_{0,05,10} = -1,812$.

Exemplo - [Cancho(2010)]

Seja T uma variável aleatória com distribuição t-Student com 12 graus de liberdade. Determine:

(a) $P(T > -1,356)$

(b) $P(0,695 < T < 2,179)$

(c) $P(-2,179 < T < 2)$

(d) $P(-1,782 < T < 1,782)$

Distribuição F-Snedecor

Seja W_1 uma variável aleatória com distribuição Qui-quadrado com k_1 graus de liberdade e W_2 outra variável aleatória com distribuição Qui-quadrado com k_2 graus de liberdade. Se W_1 e W_2 são independentes, a variável aleatória

$$F = \frac{\frac{W_1}{k_1}}{\frac{W_2}{k_2}}$$

segue uma distribuição F-Snedecor com graus de liberdade: k_1 (numerador) e k_2 (denominador). A função de densidade de probabilidade é dada por:

$$h(f) = \frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\frac{\Gamma(k_1/2)}{\Gamma(k_2/2)}} \frac{\left(\frac{k_1}{k_2}\right)^{\frac{k_1}{2}} f^{\frac{k_1}{2}-1}}{\left(1+\frac{k_1}{k_2}f\right)^{\frac{k_1+k_2}{2}}}, \quad f > 0$$

A notação $F \sim F(k_1, k_2)$ indica que que a variável aleatória F tem distribuição F-Snedecor, com graus de liberdade k_1 e k_2.

Propriedades

Se $F \sim F(k_1, k_2)$ então

1. A distribuição é assimétrica direita.
2. A média e variância são respectivamente

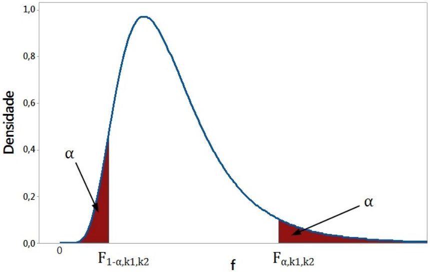
$$\mu = \frac{k_2}{k_2 - 2}, \quad k_2 > 2 \;\; e \;\; \sigma^2 = \frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2-2)^2(k_2-4)}, \quad k_2 > 4$$

Uso da tabela F-Snedecor

Os pontos críticos da distribuição F-Snedecor são apresentados na tabela F. Seja f_{α,k_1,k_2} o ponto crítico da distribuição F com k_1 graus de liberdade no numerador e k_2 graus de liberdade no denominador ,tal que a probabilidade de que variável aleatória F seja maior que este valor é

$$P(F > f_{\alpha,k_1,k_2}) = \int_{f_{\alpha,k_1,k_2}}^{\infty} h(f)df = \alpha$$

Isto é ilustrado na figura abaixo.



Por exemplo se $k_1 = 5$ e $k_2 = 10$, então da tabela F tem-se:

$$P(F > f_{0,05;5;10}) = P(F(5,10) > 3,33) = 0,05.$$

Isso é o ponto crítico do 5% superior de $F(5,10)$ é $f_{0,05;5;10} = 3,33$.

A tabela F contém somente pontos críticos na cauda superior (valores de f_{α,k_1,k_2}, para $\alpha \leq 0,25$) da distribuição F. Os pontos críticos na cauda inferior $f_{1-\alpha,k_1,k_2}$ podem ser obtidos da seguinte forma:

$$f_{1-\alpha,k_1,k_2} = \frac{1}{f_{\alpha,k_2,k_1}}.$$

Por exemplo, para determinar o ponto crítico na cauda inferior $f_{0,95;5;10}$ observe que:

$$f_{0,95;5;10} = \frac{1}{f_{0,05;10;5}} = \frac{1}{4,74} = 0,211.$$

Seja Y uma variável aleatória que segue uma distribuição F-Snedecor.

(a) Se $Y \sim F(8, 12)$ obtenha:

1. $P(Y > 2,85)$;

2. $P(2,85 < Y < 4,50)$;

3. y_1 se $P(y_1 < Y < 2,85) = 0,94$

(b) Se $Y \sim F(45, 24)$, achar y_1 tal que, $P(Y \leq y_1) = 0,95$

- Cancho, V., 2010. Notas de aulas sobre noções de estatística e probabilidade - São Paulo: USP.
- Montgomery, D., Runger, G., 2016. Estatística Aplicada e Probabilidade para Engenheiros. Rio de Janeiro: LTC.