# CAP 9 - ESTIMAÇÃO POR INTERVALO

**Conceitos e resultados a serem apresentados neste capítulo:**

**9.1- Intervalo de confiança (IC) para média populacional**
- **9.1.1- IC para média populacional com DP conhecido**
- **9.1.2- IC para média populacional com DP desconhecido**
  **Distribuição t de *Student.* Número de graus de liberdade**

**9.2- Intervalo de confiança para proporção populacional**

**Limites de confiança**

No Capítulo 8 vimos que, se $\hat{\theta}$ é um estimador pontual do parâmetro $\theta$, então $\hat{\theta}$ é uma variável aleatória e, por isso, há uma certa dose de incerteza inerente a esse processo de estimação. De fato, **o valor de $\hat{\theta}$ varia de amostra para amostra**.

Mais informativo do que gerar somente uma estimativa pontual é, portanto, **com base nos dados amostrais**, construir um **intervalo ao qual o verdadeiro valor de $\theta$ tenha uma alta chance de pertencer**.

A partir da distribuição amostral de $\hat{\theta}$, serão determinadas duas estatísticas $L_{inf}$ e $L_{sup}$, tais que para uma pequena probabilidade $\alpha$ pré-fixada, se verifique :

$$P(L_{inf} \leq \theta \leq L_{sup}) = 1 - \alpha \text{ , onde } 0 < \alpha < 1.$$

Se $l_{inf}$ e $l_{sup}$ são, respectivamente, os valores de $L_{inf}$ e $L_{sup}$, para **a amostra** considerada, então o intervalo $(l_{inf}, l_{sup})$ é chamado de **intervalo de confiança de 100 $(1-\alpha)$%** **para $\theta$**. $l_{inf}$ e $l_{sup}$ são, respectivamente, o limite inferior e o limite superior do IC de nível $100(1-\alpha)$% obtido com a amostra observada.

**É claro que quanto menor for a amplitude do intervalo obtido maior será a força dessa afirmação**. IC mais informativo

# Estimação por Intervalo

Já vimos que se $\hat{\theta}$ é um estimador pontual do parâmetro $\theta$, é uma variável aleatória e por isso há uma certa dose de incerteza inerente a esse processo de estimação.

O objetivo aqui é obter, com base nos dados, **um intervalo de valores** ao qual **o valor correto do parâmetro deve ter grande chance de pertencer**.

$$\mathbf{P}\left[\,|\hat{\theta}-\theta| < \mathbf{d}\,\right] = 1-\alpha$$

- Fixado uma tolerância **d**, do parâmetro distar de sua estimativa;
- Fixado uma probabilidade **1-** $\alpha$;
- Achar um **LS** e um **LI** (limites superiores e inferiores)

Digamos que, em uma dada situação, se pretende usar a média amostral como estimativa da média populacional de uma certa variável.

Qual deveria ser o tamanho n da amostra a ser utilizada para que se possa garantir uma boa precisão na estimativa?

$$\left|\overline{X} - \mu\right| < d$$

Suponha que $\mu$ e $\sigma$ são respectivamente a média e o desvio padrão populacionais.

Admita também que, nesse processo de estimação, o erro absoluto máximo considerado tolerável com uma probabilidade pré-fixada $1 - \alpha$, é igual a d, ou seja:

$$P\left[\left|\overline{X} - \mu\right| < d\right] = 1 - \alpha.$$

**I.C. Para μ :** $\quad |\overline{X} - \mu| < d \quad \Longrightarrow \quad P\left[|\overline{X} - \mu| < d\right] = 1 - \alpha.$

Como $\mathrm{Var}(\overline{X}) = \dfrac{\sigma^2}{n}$, admitindo que n é suficientemente grande para que o Teorema Central do Limite seja aplicável, temos

$$P\left[\frac{|\overline{X} - \mu|}{\sigma/\sqrt{n}} < \frac{d}{\sigma/\sqrt{n}}\right] = 1 - \alpha$$

Então, se $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$, esta v.a. tem distribuição aproximadamente Normal(0;1), e a igualdade acima implica que

$$z_{1-\frac{\alpha}{2}} = \frac{d}{\sigma/\sqrt{n}}$$

onde $z_{1-\frac{\alpha}{2}}$ é o quantil $1 - \alpha/2$ da Normal(0,1)(*)

## 9.1 Intervalo de Confiança para a média populacional

O processo de construção de um IC para a média populacional μ segue uma linha de raciocínio semelhante à que foi usada no Dimensionamento da Amostra.

Quando o objetivo é determinar o tamanho n da amostra, fixam-se a margem de erro d e probabilidade α e calcula-se n tal que

$$\mathbf{P}\left[\,|\hat{\theta}-\theta| < \mathbf{d}\,\right] = 1-\alpha$$

equivale a :

$$\mathrm{P}(-\mathbf{d} \le \bar{\mathbf{X}} - \mu \le \mathbf{d}) = 1 - \alpha \qquad (*)$$

Na construção do IC, a amostra aleatória já foi coletada e, conseqüentemente, **n** já foi escolhido. Então fixa-se apenas α e calcula-se **d** , de forma a que a expressão (*) acima seja atendida.

## 9.1.1 Intervalo de Confiança para a média populacional, com o desvio padrão conhecido.

Seja $(X_1, X_2, \ldots, X_n)$ uma amostra aleatória de uma v.a. X, com $E(X) = \mu$ desconhecida e $Var(X) = \sigma^2$, conhecida .

Sabemos que, para n suficientemente grande, pelo TCL, a média amostral $\bar{X}$ segue uma distribuição que se aproxima da Normal( $\mu$ , $\sigma^2/n$); ou que é exatamente Normal( $\mu$ , $\sigma^2/n$), no caso em que a distribuição de X é Normal.

Então

$$P(-d \leq \bar{X} - \mu \leq d) = 1 - \alpha \Rightarrow P\left(-\frac{d}{\sigma/\sqrt{n}} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{d}{\sigma/\sqrt{n}}\right) = 1 - \alpha$$

$$\Rightarrow P\left(-\frac{d}{\sigma/\sqrt{n}} \leq Z \leq \frac{d}{\sigma/\sqrt{n}}\right) = 1 - \alpha$$

Isso implica que $z_{1-\frac{\alpha}{2}} = \frac{d}{\sigma/\sqrt{n}}$ . Logo, $d = z_{1-\frac{\alpha}{2}} . \sigma/\sqrt{n}$ .

Substituindo d na expressão (*) acima temos

$$P(-z_{1-\frac{\alpha}{2}} . \sigma/\sqrt{n} \leq \bar{X} - \mu \leq z_{1-\frac{\alpha}{2}} . \sigma/\sqrt{n}) = 1 - \alpha$$ ou, equivalentemente,

$$P(\bar{X} - z_{1-\frac{\alpha}{2}} . \sigma/\sqrt{n} \leq \mu \leq \bar{X} + z_{1-\frac{\alpha}{2}} . \sigma/\sqrt{n}) = 1 - \alpha$$

Podemos escrever, então $P(L_{inf} \leq \mu \leq L_{sup}) = 1 - \alpha$ , onde

$$L_{inf} = \bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n} \qquad e \qquad L_{sup} = \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n}$$

**Se considerarmos os valores particulares da amostra coletada teremos**

$$l_{inf} = \bar{x} - z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n} \qquad e \qquad l_{sup} = \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n}$$

que, no caso de σ conhecido, são os limites de confiança de 100(1-α)% para a média populacional μ , obtidos para a particular amostra.

Portanto,

O intervalo de confiança (IC) de 100(1-α)% para μ, obtido com base em uma particular amostra de tamanho n, quando o desvio padrão populacional é conhecido, é expresso por

$$\left(\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n} \; , \; \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n}\right)$$

onde $z_{1-\frac{\alpha}{2}}$ é um quantil da distribuição Normal padrão.

## Exemplo 9.1 : Resistência de uma fibra têxtil

Deseja-se estimar a resistência média de um certo tipo de fibra usada na fabricação de um tecido. Uma particular amostra aleatória de 40 espécimes da fibra tem uma média aritmética de 12,4 bar. Se o desvio padrão $\sigma$ da população é conhecido e igual a 2,1 bar, determine um intervalo de confiança de 95% para o verdadeiro valor da média populacional $\mu$ .

Solução:

Temos os valores n=40, $\bar{x}$ = 12,4 bar , $\sigma$ = 2,1 bar.

Como n é relativamente grande podemos aplicar o Teorema Central do Limite e construir um intervalo de confiança conforme discutido acima.

Para 1-$\alpha$ = 0,95 temos $z_{1-\frac{\alpha}{2}}$ = 1,96 . Portanto,

$$d = z_{1-\frac{\alpha}{2}} . \sigma / \sqrt{n} = \frac{1{,}96 \times 2{,}1}{\sqrt{40}} = 0{,}65$$

Os limites de confiança são:

$$l_{inf} = \bar{x} - d = 12{,}4 - 0{,}65 = 11{,}75$$

$$l_{sup} = \bar{x} + d = 12{,}4 + 0{,}65 = 13{,}05$$

Portanto, com base na amostra selecionada, o intervalo de confiança de 95% para a verdadeira resistência média da fibra, em bar, é (11,75 ; 13,05)

Neste ponto, é inevitável perguntarmos **qual é a interpretação correta dos 95% , percentual com base no qual foi construído esse intervalo de confiança**.

Estaria correto afirmar que para a resistência da fibra, em bar, usada na fabricação do tecido se verifica $P(11{,}75 \leq \mu \leq 13{,}05) = 0{,}95$?

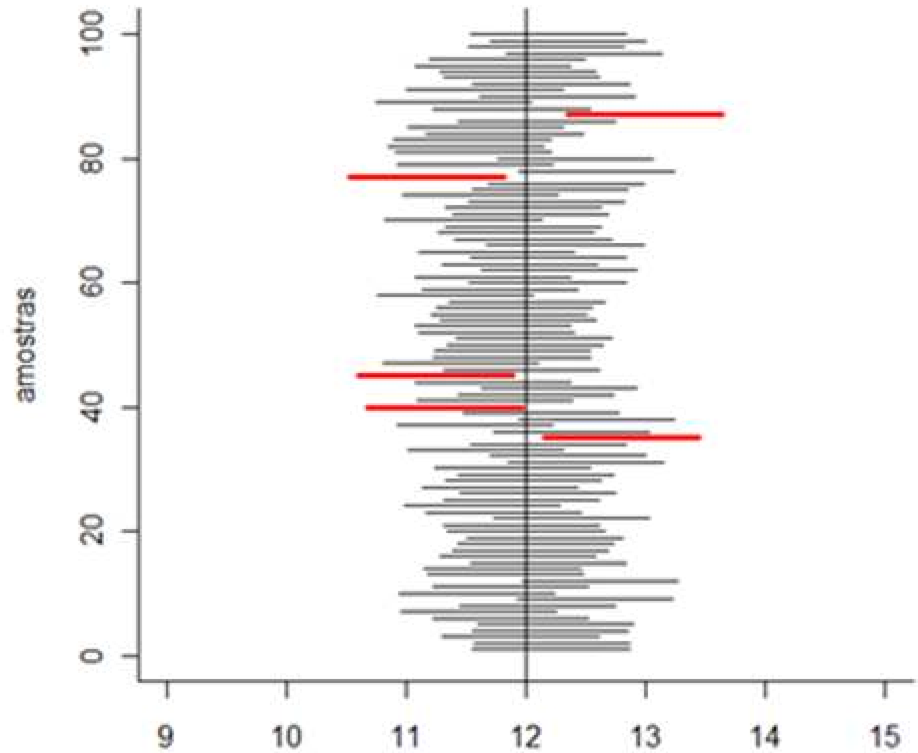
A resposta é NÃO , porque a média populacional é constante e não uma variável aleatória, embora o seu valor seja desconhecido. Assim, ou $\mu$ pertence ou não pertence a um intervalo cujos extremos são valores conhecidos. Portanto a probabilidade acima só poderia ser igual a 1 ou igual a 0.

Ocorre que para a amostra dada os extremos do intervalo são valores particulares das estatísticas $L_{inf}$ e $L_{sup}$ . Ou seja , as variáveis aleatórias envolvidas nesse processo são os extremos do Intervalo de Confiança.

Admita que foram extraídas 100 amostras (e não apenas uma), com n = 40 espécimes da fibra cada. Então, para cada uma dessas amostras poderia ser aplicado exatamente o mesmo procedimento aqui descrito e obteríamos 100 intervalos de confiança para $\mu$ .

Como o nível de confiança que caracterizou o processo de construção desses intervalos é de 95% , espera-se que cerca de 95 desses intervalos contenham o verdadeiro valor do parâmetro $\mu$ , enquanto que os outros 5 não o conterão.

Para ilustrar, foi feita uma simulação com 100 amostras, cada uma de tamanho 40, extraídas de uma população com $\mu = 12$ e $\sigma = 2{,}1$ bar. Obtivemos um IC para cada amostra e os IC's foram marcados no gráfico a seguir.

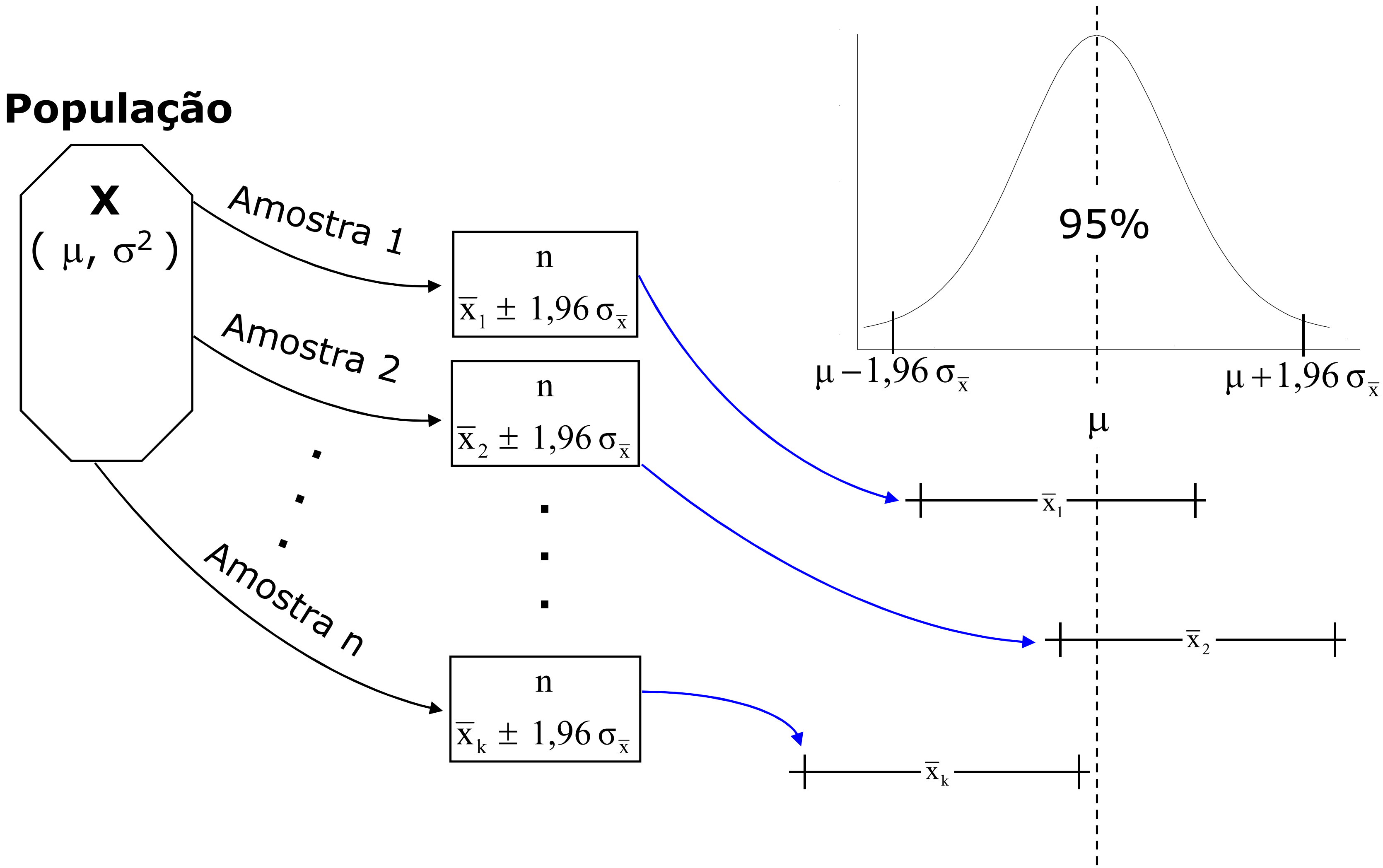



Simulando a construção de intervalos de confiança para a média populacional

o valor do parâmetro $\mu$ é conhecido, o que em geral não acontece nas situações reais.

Além disso, é claro que em uma situação real só se utiliza uma amostra.

# Esquema do Conceito de Intervalo de Confiança




**Aproximadamente 95% dos intervalos comtém μ**

### 9.1.2 Intervalo de Confiança para a média populacional, com o desvio padrão desconhecido. A distribuição t de *Student*

No exemplo 9.1 o desvio padrão populacional σ foi suposto conhecido. Mas, e se σ não for conhecido?

Na prática, quando se pretende construir um intervalo de confiança para a média populacional μ o que ocorre na maior parte dos casos é que o desvio padrão populacional σ também é desconhecido. Em uma tal situação, a solução usual é substituir σ por sua estimativa pontual derivada da amostra, s.

O procedimento é o seguinte:

No caso em que σ era conhecido recorremos à v.a. $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ cuja distribuição é a Normal padrão.

Se, por ser σ desconhecido, simplesmente o substituímos por S na expressão acima teremos uma variável aleatória que é função de duas outras, $\bar{X}$ e S cuja distribuição não é a Normal padrão, mas que pode se aproximar dela para valores grandes do tamanho da amostra.

Contudo, para valores pequenos de n $(n < 30)$ o erro cometido ao usarmos a distribuição Normal padrão é muito grande e será necessário alterar a metodologia

Se $(X_1, X_2, \ldots, X_n)$ é uma amostra aleatória de uma variável X distribuída conforme uma Normal( $\mu$ , $\sigma^2$) , a distribuição de probabilidade da variável aleatória

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

é um caso particular do modelo chamado t de Student que descreveremos brevemente a seguir.

## Noções sobre a distribuição t de *Student*, Algumas observações gerais sobre a distribuição t de Student :

1. A distribuição t de *Student* é, juntamente com a Qui-quadrado e a F de *Fisher-Snedecor*, uma das distribuições de probabilidade **derivadas da Normal**.
2. A t de *Student* depende do parâmetro $\nu$ , um número inteiro e positivo chamado **número de graus de liberdade**.
3. Assim como acontece com a Normal padrão, a função de densidade da t de *Student* é também uma **curva simétrica e centrada em zero**, porém ela é **mais dispersa** em torno de zero que a Normal Padrão.
4. O cálculo de probabilidades associadas à t de *Student* também pode ser feito com o auxílio de um software apropriado, ou através de consulta a uma tabela de probabilidade como a Tabela II do Apêndice II.

5. À medida que o número de graus de liberdade $\nu$ tende a infinito, a curva da t de Student se aproxima cada vez mais da curva da Normal Padrão. (Basta observar uma tabela da t de *Student* para constatar que, dado um valor fixo de α, quando $\nu$ cresce, o quantil 1 – α da t com $\nu$ graus de liberdade se aproxima cada vez mais do quantil 1 – α da Normal padrão.)

Na figura a seguir as 3 linhas representam a distribuição t de Student com nº de graus de liberdade igual a 1, 2 e 5. A título de comparação foi desenhada, com uma linha mais grossa, uma Normal padrão.

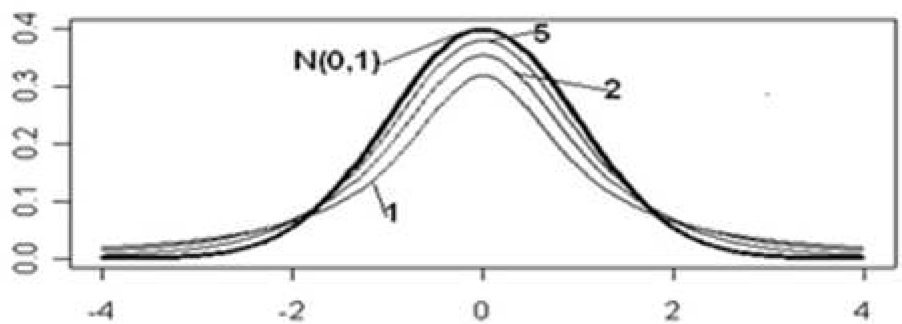



Figura 9.2 – A densidade da t de *Student* para diversos valores de $\nu$

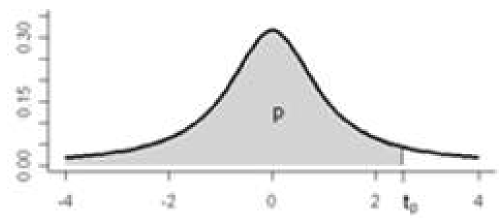
**Uso da Tabela da t de *Student*:**

Admita que a v.a. T segue uma distribuição t de *Student* com $\nu$ graus de liberdade. A figura abaixo mostra a função densidade de T.

Notação:

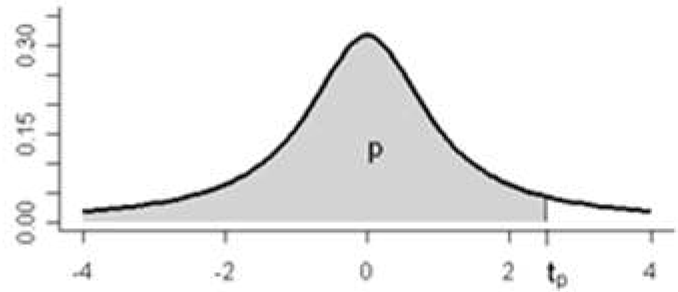
- p representa a área sob a curva e à esquerda de $t_p$.
- $t_p$ representa um valor da variável T, também chamado quantil de T, à esquerda do qual há uma área p.

Em símbolos : $p = P[T \leq t_p]$




A rigor, a notação mais informativa a ser usada neste caso seria $t_{p;\ \nu}$, onde apareceriam simultaneamente a probabilidade p e o número de graus de liberdade $\nu$. Ela será usada sempre que, para evitar ambigüidades, for necessário especificar claramente o valor de $\nu$. Entretanto, para não sobrecarregar a notação, optamos por escrever somente $t_p$.

$$t_{p;\ \nu} = t_p$$




1. Determinação de $t_p$ a partir do nº de graus de liberdade $\nu$ e da probabilidade p

Exemplos:

- Se $\nu = 6$ e $p = 0{,}90 \rightarrow t_{0,90} = 1{,}440$ (ou, mais precisamente, $t_{0,90;\,6} = 1{,}440$)
- Se $\nu = 3$ e $p = 0{,}995 \rightarrow t_{0,995} = 5{,}841$ (ou, mais precisamente, $t_{0,995;3} = 5{,}841$)

2. Determinação de $p = P[T < t_p]$ a partir de $\nu$ e $t_p$.

Exemplos:

- Se $\nu = 10$ e $t_p = 2{,}31 \rightarrow$ Pela tabela, vemos que $p = P[T < 2{,}31]$ está entre 0,975 e 0,98
- Se $\nu = 4$ e $t_p = 0{,}88 \rightarrow$ Pela tabela, $p = P[T < 0{,}88]$ está entre 0,70 e 0,80

<u>Observação</u>: Para uma maior precisão pode ser feita uma interpolação ou então pode ser usado um software apropriado.

## Construção do IC para $\mu$ de uma distribuição Normal, com $\sigma$ desconhecido

Seja $(X_1, X_2, \ldots, X_n)$ uma amostra aleatória de uma v.a. X, com

$$E(X) = \mu \quad e \quad Var(X) = \sigma^2.$$

Vimos que, quando $\sigma$ é conhecido, o intervalo de confiança $(\bar{x} - d, \bar{x} + d)$ foi construído calculando o erro da estimativa como $d = z_{1-\frac{\alpha}{2}} . \sigma/\sqrt{n}$ ,

Quando $\sigma$ é desconhecido o intervalo de confiança para $\mu$ segue uma configuração semelhante, sendo que agora o erro da estimativa, d, é expresso por $d = t_{1-\frac{\alpha}{2}} . s/\sqrt{n}$ ,

onde $t_{1-\frac{\alpha}{2}}$ é um quantil da curva t de *Student* com n – 1 graus de liberdade.

Portanto,

O intervalo de 100(1-α)% de confiança para a média populacional $\mu$ de uma população Normalmente distribuída, obtido a partir de uma amostra aleatória de tamanho n, quando $\sigma$ é desconhecido, é dado por

$$(\bar{x} - t_{1-\frac{\alpha}{2}} . s/\sqrt{n} \; , \bar{x} + t_{1-\frac{\alpha}{2}} . s/\sqrt{n} \;),$$

onde:

- s é a estimativa do desvio padrão populacional $\sigma$; e
- $t_{1-\frac{\alpha}{2}}$ é o quantil $1 - \frac{\alpha}{2}$ da distribuição t de Student com n-1 graus de liberdade.

**Exemplo 9.2**: **Voltando ao carbono na liga de ferro**

Considere novamente os dados do Exemplo 8.5 referentes ao conteúdo de carbono em uma amostra de 10 espécimes da liga de ferro. Suponha que o objetivo é determinar, com base nessa amostra, um intervalo de confiança de 90% para o verdadeiro conteúdo médio de carbono na liga de ferro produzida pela Empresa.

A partir dos dados obtemos :

n = 10 , $\sum_{i=1}^{10} x_i$ = 43,12 ; $\sum_{i=1}^{10} x_i^2$ = 201,75.

Daí, tem-se :

$\bar{x}$ = 4,31 g/(100g da liga) ,

s = 1,33 g/(100g da liga)

Como σ não é conhecido e n é menor que 30, usaremos o intervalo de confiança baseado na distribuição t de *Student*.

Para um nível de confiança (1- α) = 0,90 e (n-1) = 9 graus de liberdade, temos

$t_{1-\frac{\alpha}{2}} = t_{0,95}$ = 1,833 ,

logo $d = 1{,}833 \cdot \frac{1{,}33}{\sqrt{10}}$ = 0,77

Assim sendo , $l_{inf}$ = 4,31 – 0,77 = 3,54 e $l_{sup}$ = 4,31 + 0,77 = 5,08

Portanto, o intervalo de confiança de 90% para o conteúdo médio de carbono, obtido com a amostra dada, é (3,53 g/100g ; 5,08 g/100g).

**Observação:**

Para o resultado acima ter validade devemos ter boas razões para supor que o conteúdo de carbono segue uma distribuição Normal, ou muito próxima da Normal. Caso contrário o uso da distribuição t de Student pode conduzir a resultados imprecisos. Sendo assim, diante de uma tal situação, é recomendável que se verifique a plausibilidade da premissa de Normalidade da variável em questão. Uma forma de verificar a Normalidade através da amostra é o uso do gráfico de probabilidade Normal (ver Capítulo 12).

Caso seja evidente a falta de Normalidade da distribuição de X, uma solução seria aplicar uma transformação que conduza a uma Normalidade aproximada (Algumas possibilidades seriam: logX, $\sqrt{X}$, 1/X, etc). Se isso não der resultado, a solução é recorrer a métodos não paramétricos. (Ver Capítulos 11 e 12.)

## 9.2 Intervalo de Confiança para a proporção populacional

Sabemos que a proporção p dos elementos de uma população que possuem determinada característica pode ser vista como a esperança de uma variável binária, ou seja, uma variável que assume o valor 1 para os elementos em que a característica está presente e o valor 0 para aqueles em que ela está ausente. Já vimos também que neste caso, uma vez extraída uma amostra aleatória com n elementos dessa população, o estimador natural de p é a proporção amostral

$\hat{p} = \dfrac{\sum_{i=1}^{n} X_i}{n}$, onde os $X_i$'s são v.a.'s iid com distribuição Bernoulli(p).

Como obter um intervalo de confiança para p a partir dos dados?

Já que a proporção é um caso particular da média, e, além disso, cada $X_i$ tem esperança p e desvio padrão $\sqrt{p(1-p)}$, podemos concluir também que, (para n suficientemente grande, de modo a ser válido o Teorema Central do Limite)

$$1-\alpha = P(\hat{p} - d \leq p \leq \hat{p} + d) \ , \quad \text{onde}$$

$$d = z_{1-\frac{\alpha}{2}} \times DP(\hat{p})$$

Como $DP(\hat{p}) = \dfrac{DP(X_i)}{\sqrt{n}} = \sqrt{\dfrac{p(1-p)}{n}}$ , obtemos :

$$1-\alpha = P\left[\hat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right]$$

A questão agora é que o próprio parâmetro p está aparecendo no cálculo dos extremos desse intervalo. Queremos então substituí-lo por algo que só dependa dos dados.

Dois caminhos são possíveis para resolvermos essa dificuldade:

- Lembrando mais uma vez que o produto p(1-p) é menor ou igual a ¼, para todo p,:

O intervalo de confiança conservativo para p ao nível $(1 - \alpha)$ é

$$\left(\hat{p} - z_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{4n}}, \hat{p} + z_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{4n}}\right).$$

- Substituindo simplesmente p por sua estimativa $\hat{p}$ na expressão acima:

O intervalo de confiança não conservativo para p ao nível $(1 - \alpha)$ é

$$\left[\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ ,\ \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Evidentemente, a opção conservativa sempre nos levará a intervalos de maior amplitude do que a opção não conservativa.

Note que ambas as opções acima são soluções aproximadas para a construção do IC.

**Exemplo 9.3: Pentes de memória defeituosos**

Um fabricante de pentes de memória RAM para computadores, que produz em grandes quantidades, deseja estimar a fração p de unidades defeituosas elaboradas por sua indústria.

Para isso ele selecionou uma amostra aleatória de 200 unidades e verificou que, entre elas, 5 eram defeituosas. Construa um intervalo de confiança de 95% para p .

Solução.

Seja $X_i$ a variável aleatória que representa a i-ésima unidade selecionada, onde

$X_i \sim$Bernoulli (p) , i=1, 2, ...., 200.

Sabemos que $\sum_{i=1}^{200} x_i = 5$

Então a estimativa pontual da fração de unidades defeituosas é $\hat{p} = \frac{5}{200} = 0{,}025$

Como n é grande, para obtermos o intervalo de confiança de 95% usamos a aproximação Normal.

Para $(1 - \alpha) = 0{,}95$ , temos $1 - \frac{\alpha}{2} = 0{,}975$ . Portanto, da Tabela da Normal padrão, encontramos $z_{1-\frac{\alpha}{2}} = 1{,}96$ .

Usando o intervalo não conservativo temos $\hat{p} - d \leq p \leq \hat{p} + d$ , onde a margem de erro d é dada por

$$d = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1{,}96 \times \sqrt{\frac{(0{,}025)(0{,}975)}{200}} = 0{,}022$$

Portanto, os limites de confiança de 95% para p, com base na amostra dada são:

$l_{inf}$ = 0,025 -0,022 = 0,003 e $l_{sup}$ = 0,025 + 0,022 = 0,047

Ou seja, com 95% de confiança espera-se que o intervalo (0,3% ; 4,7%) inclua no seu interior a verdadeira proporção de pentes defeituosos.

Na prática, neste tipo de problema, o mais relevante é o limite superior, que estabelece um limite máximo para a proporção de peças defeituosas produzidas pela fábrica (no caso, aproximadamente 5%).