CAPÍTULO 7

ESTIMATIVA DE PARÂMETROS

Introdução

Uma variável aleatória é caracterizada ou descrita pela sua distribuição de probabilidade.

Em aplicações industriais, as distribuições de probabilidade são usadas para modelar tempos de processo ou características de qualidade tais como dimensionais críticos ou percentuais de não conformes.

A distribuição de probabilidade, por sua vez, é descrita pelos seus parâmetros populacionais. Por exemplo, a média μ e o desvio-padrão σ são os parâmetros populacionais da distribuição Normal, enquanto λ é o parâmetro da distribuição de Poisson.

Introdução

Assim, existe interesse em conhecer os parâmetros da distribuição e é preciso desenvolver procedimentos para estimar esses valores.

As estimativas dos parâmetros da distribuição serão feitas a partir dos resultados (dados) de uma variável aleatória de uma amostra coletada. As estimativas podem ser pontuais, ou podem ser feitas através de um intervalo de confiança.

Introdução

Assim, existe interesse em conhecer os parâmetros populacionais da distribuição de probabilidade.

Como geralmente os parâmetros populacionais da distribuição de probabilidade não são conhecidos, é preciso desenvolver procedimentos para estimar esses parâmetros.

As estimativas dos parâmetros populacionais da distribuição são realizadas a partir dos resultados (dados) de uma variável aleatória de uma amostra representativa extraída dessa população.

Introdução

Esse procedimento é chamado de estatística inferencial, pois estima-se um parâmetro populacional desconhecido da distribuição de probabilidade através de uma amostra representativa extraída dessa população.

A estatística inferencial compreende a estimação de parâmetros populacionais e testes de hipótese a respeito da população.

Na verdade, a estatística inferencial forma a base das atividades de controle da qualidade e também pode auxiliar na tomada de decisão e em muitas outras situações.

7.1. Estimativas Pontuais

A estimação de parâmetros populacionais pode ser por ponto (pontual) ou por intervalo de confiança.

A estimativa pontual é um valor obtido a partir dos resultados (dados) de uma variável aleatória de uma amostra representativa extraída da população.

Seja a variável aleatória *X*, com distribuição de probabilidade *f(X)*, e seja que o valor dos parâmetros populacionais da média μ e da variância σ^2 são desconhecidos.

7.1. Estimativas Pontuais

Se uma amostra representativa da variável aleatória X é extraída da população, a média $\bar{X}$ e a variância S^2 dessa amostra podem ser usadas como estimadores pontuais dos parâmetros populacionais μ e σ^2.

Por exemplo, pode haver interesse em estimar a média e a variância da produtividade.

Se uma amostra de 15 unidades indica

$$\bar{X} = 5{,}026\text{cm e } S^2 = 0{,}0012\text{cm}^2$$

então esses valores são tomados como estimativas pontuais dos parâmetros populacionais μ e σ^2.

Há várias propriedades que fazem um estimador ser um bom estimador. Dentre elas citamos:

1. Um estimador deve ser não tendencioso, isto é, ele não deve subestimar ou superestimar sistematicamente o valor do parâmetro que está sendo estimado.

2. Ele deve apresentar variância mínima, isto é, sua variabilidade deve ser menor que a variabilidade de qualquer outro estimador que possa ser concebido.

Distribuição Amostral

⇨ **Parâmetros e Estatísticas:**

•**Parâmetros:** São medidas estatísticas obtidas através do censo para descrever uma característica da população.

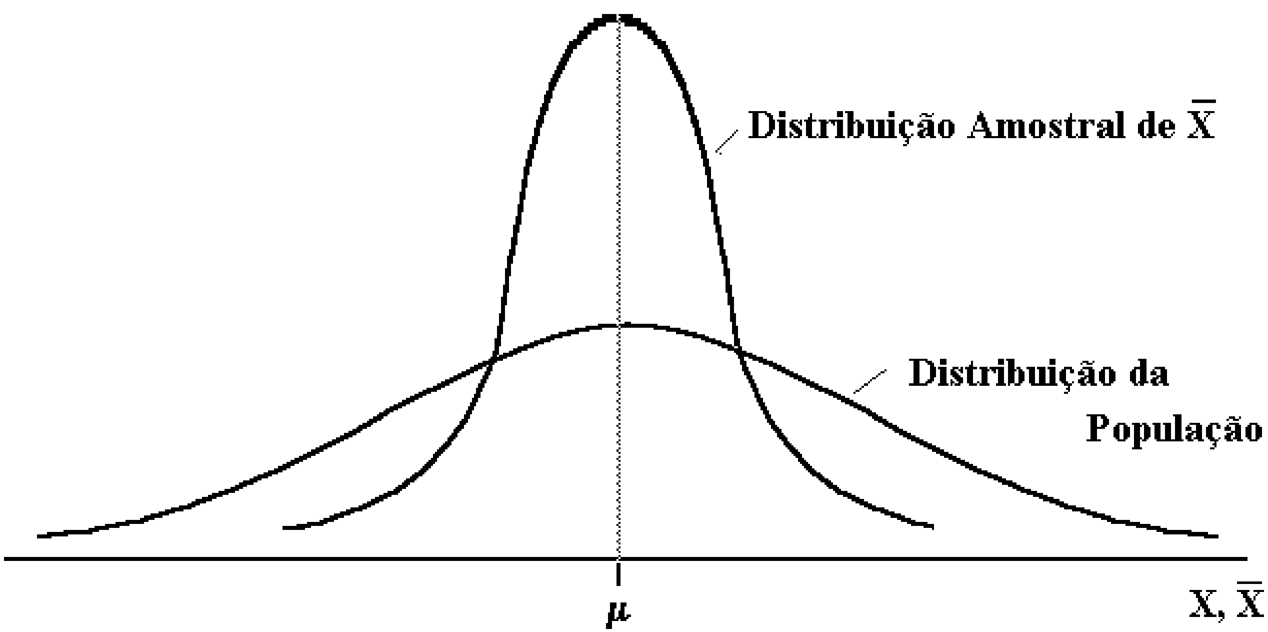
•**Estatística:** São medidas características obtidas através de uma amostra.

Medida	Parâmetro	Estatística
Média	μ	$\bar{x}$
Desvio padrão	σ	s
Variância	σ^2	s^2
Proporção	P	p

Distribuição Amostral da Média

Uma distribuição amostral das médias indica a probabilidade de ocorrência de uma média amostral.

- **As <u>médias amostrais</u> tendem a agrupar-se em torno da <u>média populacional</u>.**



Distribuição amostral de $\overline{X}$ - População Normal

Distribuição Amostral da Média

A média das <u>médias amostrais</u> é igual a verdadeira média populacional.

$$E[\overline{X}] = \mu$$

- **E o desvio padrão da <u>distribuição amostral das médias</u> será dado por:**

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

⇨ <u>Teorema do limite central:</u>

- **Se a população sob amostragem tem distribuição normal, a distribuição das médias amostrais também será normal.**
- **Mesmo que a população não seja considerada uma distribuição normal, a distribuição das médias amostrais será aproximadamente normal para grandes amostras.**

Distribuição Amostral da Média

$$\bar{x} \cong N(\mu, \sigma/\sqrt{n})$$

Logo; $$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Exemplo:

Certas válvulas produzidas por uma empresa têm vida média de 900h e desvio-padrão de 80h. A empresa envia 1000 lotes de 100 válvulas cada um. Em quantos lotes pode-se esperar que a vida média exceda a 910h?

Distribuição Amostral da Média

⇨**Distribuição Amostral da média quando é desconhecido:**

- **Quando <u>desconhecemos o desvio-padrão</u> <u>populacional</u> utilizamos como estimativa o valor de s. Desta forma, o desvio-padrão das médias (ou erro padrão) será dado por:**

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}, \quad \textbf{onde} \quad s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

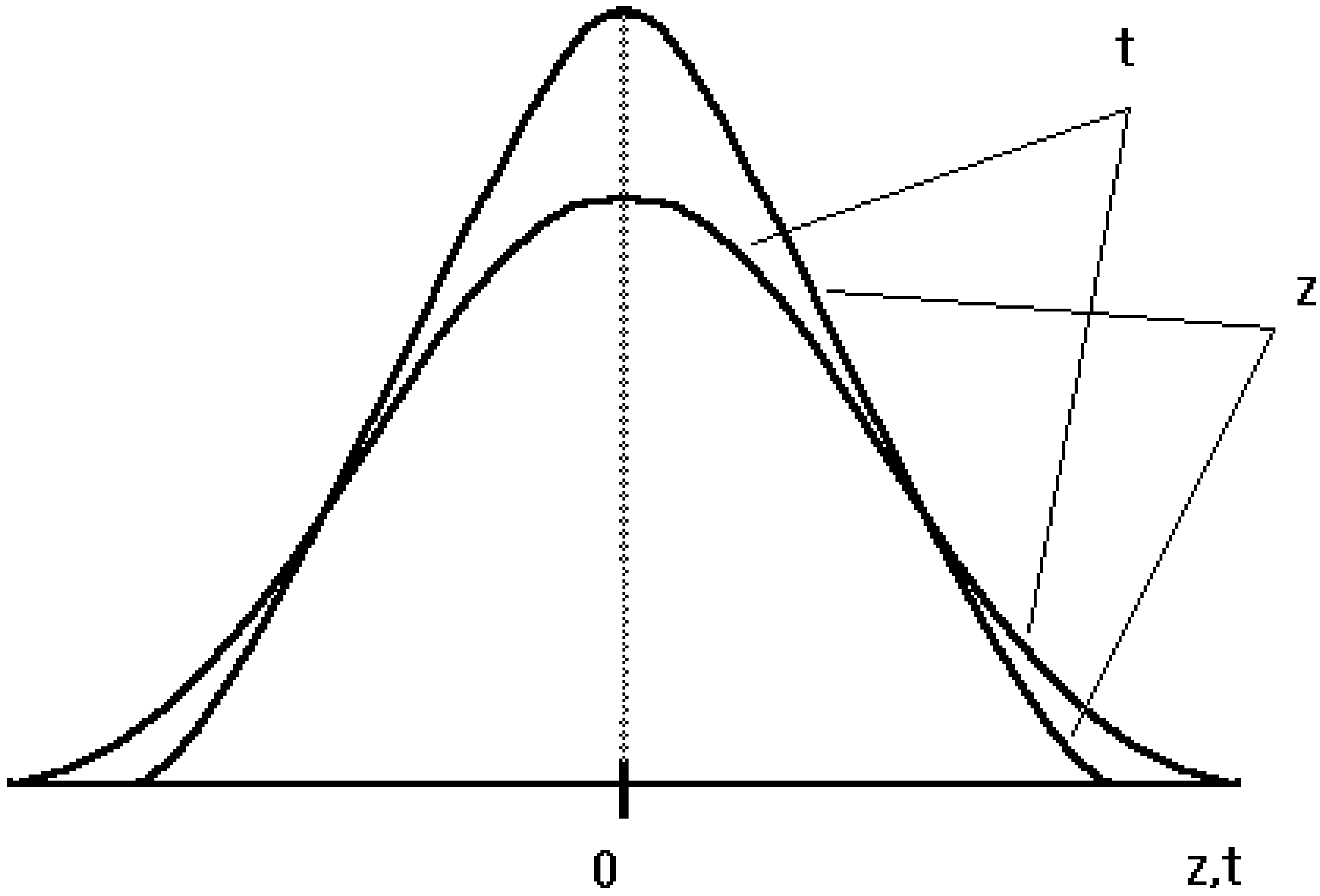
- **Para grandes amostras, podemos admitir que a variação dos valores observados na amostra seje semelhante a variação da população. Porém, para pequenas amostras isso pode não ser verdadeiro. Neste caso, a distribuição adequada é a distribuição t-student.**

Assim, utilizamos a estatística: $$\frac{(\bar{x}-\mu)}{\frac{s}{\sqrt{n}}} \approx t-\text{student} \quad ;$$

com (n-1) graus de liberdade.

- **Esta distribuição é muito parecida com a distribuição normal, sendo simétrica em torno da média zero, porém tem maior dispersão comparado com a normal.**

- **A forma da distribuição t-student depende do tamanho da amostra. Quanto menor o tamanho da amostra, menor serão os graus de liberdade e mais dispersa ("achatada") será a curva.**



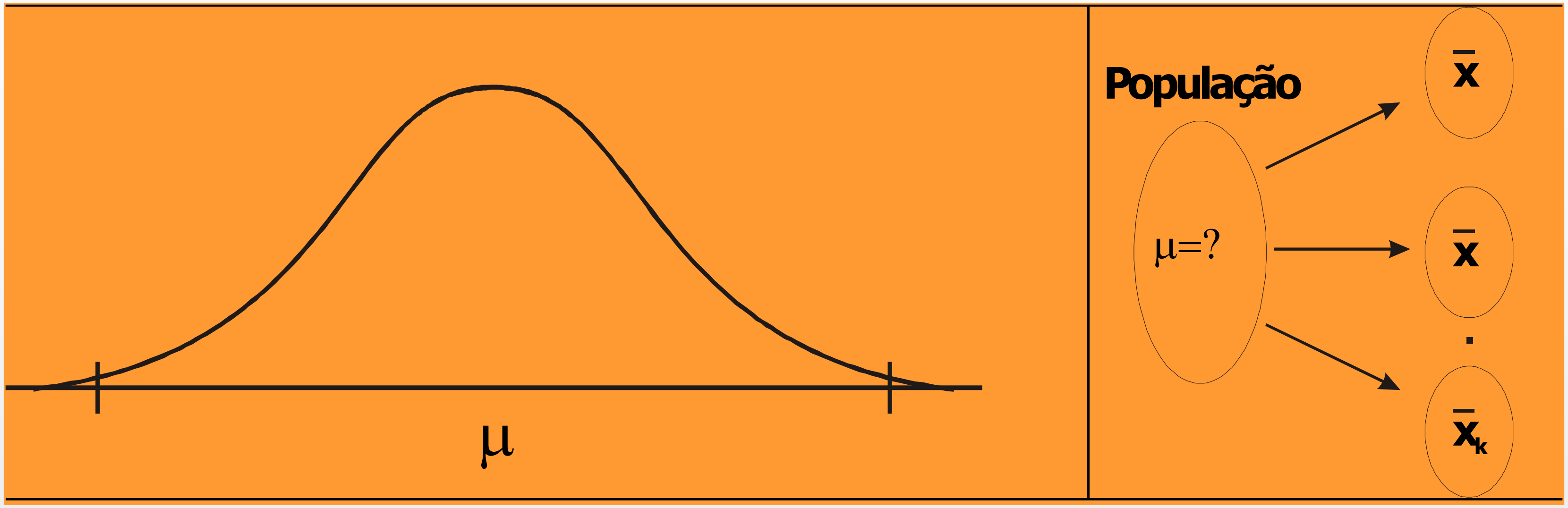
Distribuição t e Distribuição Normal Reduzida

7.2. Estimativas através de Intervalos de Confiança

A estimação de parâmetros populacionais por intervalo de confiança consiste em gerar um intervalo, centrado na estimativa pontual, no qual se admite que esteja o parâmetro populacional.

A estimativa pontual é calculada a partir de uma amostra extraída da população. No entanto, pode-se extrair várias amostras de uma população.

Por exemplo, para estimar a média populacional μ pode-se retirar várias amostras diferentes que podem gerar várias estimativas pontuais diferentes.

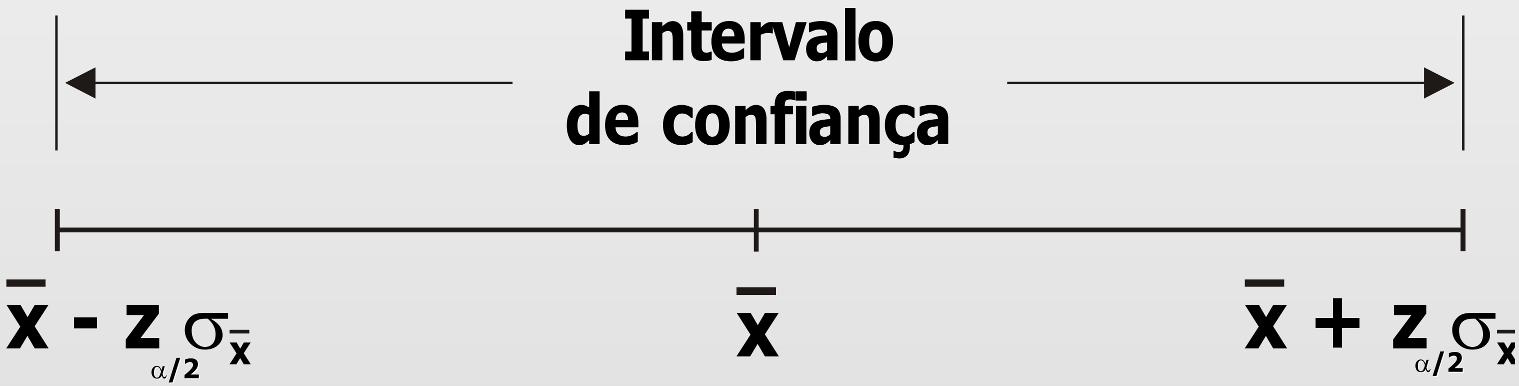


Se a amostra for representativa da população, ela tende a gerar valor próximo do parâmetro populacional, mas não igual.

Como a estimativa é baseada em uma única amostra, o quão próximo o valor encontrado nessa amostra está do verdadeiro parâmetro populacional?? Não há como saber se a amostra coletada foi extraída da cauda superior ou inferior da distribuição.

Logo, para se ter confiança de estimar o verdadeiro parâmetro populacional, gera-se um intervalo de possíveis valores para o parâmetro populacional, a partir do valor encontrado da amostra.

Quanto maior a amplitude do intervalo, maior a confiança (probabilidade) de estimar corretamente o verdadeiro parâmetro populacional.



Conforme a amplitude do intervalo, existe uma probabilidade (1-α) de que o parâmetro populacional esteja contido no intervalo.

Essa probabilidade (1-α) é chamada nível de confiança, sendo α a probabilidade do erro, ou seja, a probabilidade do intervalo não conter o verdadeiro parâmetro populacional.

Um intervalo de confiança de *100(1-α)%* é estabelecido a partir de dois limites, tais que a probabilidade do verdadeiro valor do parâmetro estar incluído dentro do intervalo é *100(1-α)%*.

Por exemplo, para construir um intervalo de confiança de 95% para a média, nós precisamos achar os limites L e U tais que:

$$P\{L \leq \mu \leq U\} = 95\%$$

A interpretação do intervalo de confiança é a seguinte:

se um grande número desses intervalos fosse construído, a partir de diversas amostras aleatórias, então *95%* desses intervalos iria incluir o verdadeiro valor da média populacional μ, ou seja, 5% dos intervalos de confiança estimados baseados nas médias amostrais não conteria o verdadeiro parâmetro populacional μ.

O intervalo de confiança apresentado acima é um intervalo bilateral. Também pode haver interesse em construir intervalos unilaterais.

O limite inferior para um intervalo unilateral de *100(1-α)%* é definido calculando-se o valor de L tal que:

$$P\{L \leq \mu\} = 1 - \alpha$$

O limite superior para um intervalo unilateral de *100(1-α)%* é definido calculando-se o valor tal que:

$$P\{\mu \leq U\} = 1 - \alpha$$

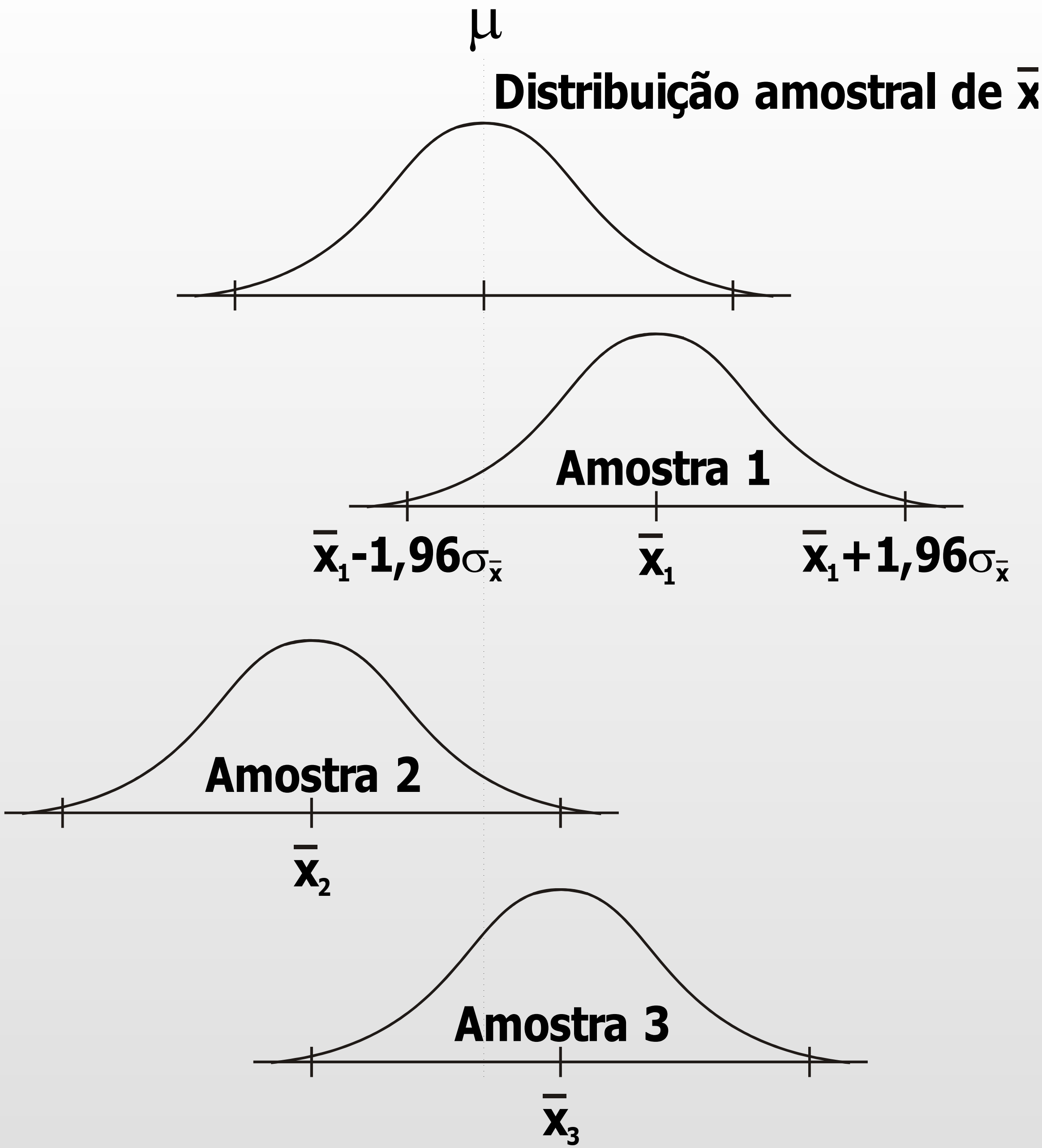
7.3 Intervalo de confiança para a média, variância conhecida

Seja X uma variável aleatória qualquer que siga a distribuição Normal X → N(μ, σ) e seja $x_1, \ldots, x_n$ uma amostra aleatória desse processo.

A partir do teorema do limite central, sabe-se que a distribuição da média segue a distribuição Normal $\bar{X} \rightarrow N(\mu, \sigma/\sqrt{n})$.

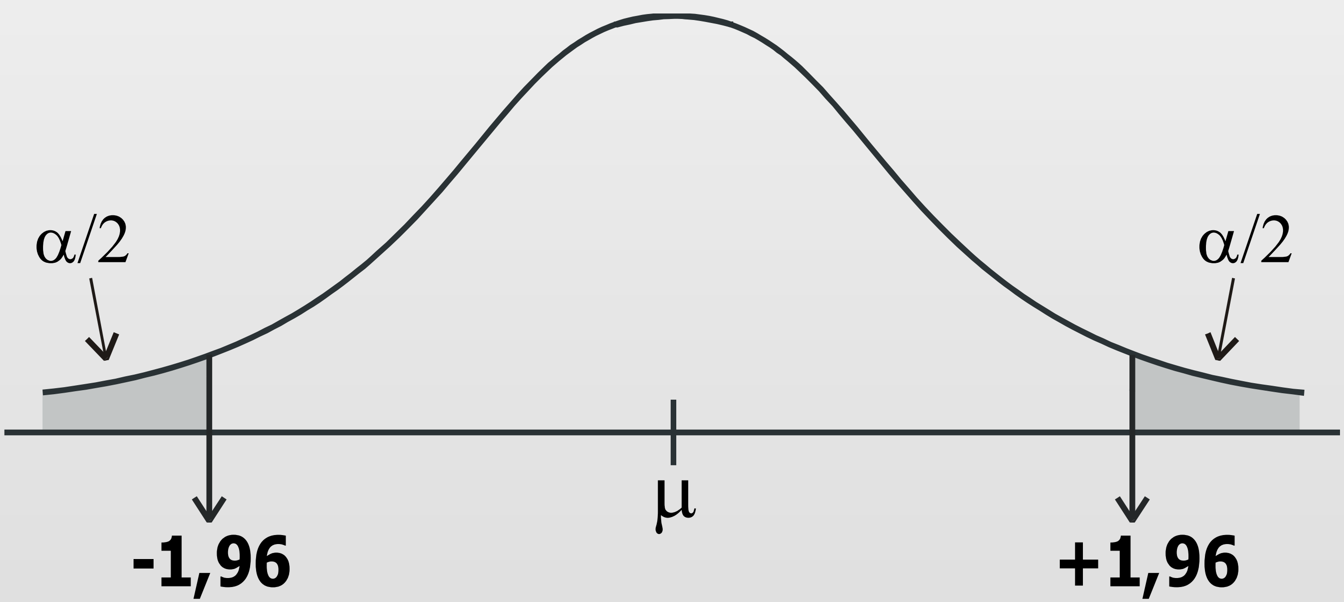
Mais ainda, para n suficientemente grande este resultado é válido mesmo que a distribuição de origem não seja Normal

Seja que uma variável aleatória X tenha média desconhecida e variância conhecida. E seja que amostras dessa população apresentem média igual a $\bar{X}$.


μ
Distribuição amostral de $\bar{x}$
Amostra 1
$\bar{x}_1 - 1{,}96\sigma_{\bar{x}}$
$\bar{x}_1$
$\bar{x}_1 + 1{,}96\sigma_{\bar{x}}$
Amostra 2
$\bar{x}_2$
Amostra 3
$\bar{x}_3$

Para obter-se um intervalo de confiança de 95% (1-α = 0,95, α = 0,05), ou seja, 95% dos intervalos, construídos a partir das amostras coletadas, contenham o verdadeiro parâmetro populacional, é preciso gerar um intervalo cuja amplitude contenha 95% das possíveis amostras coletadas.

Ou seja, um intervalo proporcional a $\pm\ Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ **, onde** $Z_{\alpha/2}$ **neste caso é** $Z_{0,025}$**=1,96.**



7.3 Intervalo de confiança para a média, variância conhecida

Então o intervalo bilateral de confiança de *100(1-α)%* para μ :

$$\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Observa-se que, para *n* suficiente grande, as médias seguem a distribuição Normal independentemente da distribuição original de *X*.

Consequentemente a equação anterior é o intervalo de confiança para a média de observações que apresentam uma distribuição de origem qualquer.

Exemplo 7.1: A variabilidade do tempo de atendimento em um caixa bancário é conhecida σ = 0,10 min. Uma amostragem com 20 pessoas indicou tempo médio de atendimento $\overline{X} = 1{,}5\,\text{min}$ Construa um intervalo de confiança de 95% para o tempo médio de atendimento.

$$\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$1{,}5 - 1{,}96\frac{0{,}10}{\sqrt{20}} \le \mu \le 1{,}5 + 1{,}96\frac{0{,}10}{\sqrt{20}}$$

$$1{,}5 - 1{,}96\frac{0{,}10}{\sqrt{20}} \le \mu \le 1{,}5 + 1{,}96\frac{0{,}10}{\sqrt{20}}$$

$$1{,}46 \le \mu \le 1{,}54$$

Exercícios 7.1 e 7.3

Um intervalo unilateral de *100(1-α)%* com limite superior é estabelecido a partir de:

$$\mu \leq \overline{X} + Z_{\alpha} \frac{\sigma}{\sqrt{n}}$$

Um intervalo unilateral de *100(1-α)%* com limite inferior é:

$$\overline{X} - Z_{\alpha} \frac{\sigma}{\sqrt{n}} \leq \mu$$

7.4 Erro de Estimação

O intervalo de confiança bilateral tem a forma

$$\overline{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Aumentando a amplitude do intervalo, aumenta-se o nível de confiança do intervalo, no entanto, aumenta-se o erro máximo de estimação que é o valor absoluto da diferença entre o parâmetro amostral ($\overline{X}$) e o parâmetro populacional μ, representado como $\varepsilon = \left|\overline{X} - \mu\right|$ **.**

Como o intervalo de confiança tem centro na média amostral, o erro máximo provável é igual a metade da amplitude do intervalo.

Como $\bar{X} \pm Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ **, pode-se escrever** $\bar{X} \pm erro$

Logo $e = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

Para determinar o tamanho da amostra mínimo para estimar um parâmetro populacional, basta isolar o valor de *n* na equação acima.

$$n = \left(Z_{\alpha/2}\frac{\sigma}{e} \right)^2$$

Logo, o tamanho da amostra dependerá:

- do grau de confiança desejado α **;**

- da dispersão na população σ**;**

- de certo valor especificado para o erro tolerável.

Exemplo 7.2: **Qual o tamanho da amostra necessário para estimar a média populacional de uma característica dimensional de um processo cujo desvio-padrão σ = 3 cm, com 95% de confiança e precisão de 0,5 cm ?**

$\alpha = 0{,}05 \Longrightarrow z_{\alpha/2} = 1{,}96;\ \sigma = 3{,}0$ cm; $e = 0{,}5$cm;

$$n = \left(\frac{1{,}96 \times 3{,}0}{0{,}5}\right)^2 = 138{,}3$$

Logo são necessários n = 139 peças

Exercício 7.12

7.6. Intervalo de confiança para a média, variância desconhecida

Seja X uma variável aleatória Normal com média e variância desconhecidas. Se uma amostra de n valores indica média $\overline{X}$ e variância S^2, o intervalo de confiança de *100(1-α)%* é calculado usando-se a distribuição de Student.

$$\overline{X} - t_{\alpha/2,n-1}\,\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{\alpha/2,n-1}\,\frac{S}{\sqrt{n}}$$

Os intervalos unilaterais de confiança de *100(1-α)%* com limites superior e inferior são respectivamente:

$$\mu \le \overline{X} + t_{\alpha,n-1}\,\frac{S}{\sqrt{n}}$$

$$\overline{X} - t_{\alpha,n-1}\,\frac{S}{\sqrt{n}} \le \mu$$

Exemplo 7. 3: A quantidade mensal de produtos entregues por uma empresa segue uma distribuição Normal com média e variância desconhecidas. Analise os dados a seguir, que representam uma amostra de 20 meses e construa um intervalo de 95% para a média.

17,4	18,2	18,3	18,8	19,0	19,2	19,3	19,6	19,6	19,9
20,2	20,2	20,5	20,7	20,9	21,0	21,3	21,5	21,9	22,6

$$\overline{X} = 20{,}01$$

$$S = 1{,}34$$

$$t_{0{,}025;19} = 2{,}093$$

$$20{,}01 - 2{,}093\frac{1{,}34}{\sqrt{20}} \le \mu \le 20{,}01 + 2{,}093\frac{1{,}34}{\sqrt{20}}$$

ou

$$19{,}38 \le \mu \le 20{,}64$$

Exemplo 7.4: A empresa pode estar preocupada exclusivamente com a quantidade mensal de produtos entregues muito baixa. Construa um intervalo de confiança unilateral com 95% de confiança no limite inferior.

$$\overline{X} = 20{,}01$$

$$S = 1{,}34$$

$$t_{0,05;19} = 1{,}729$$

$$20{,}01 - 1{,}729\frac{1{,}34}{\sqrt{20}} \leq \mu \qquad ou \qquad 19{,}49 \leq \mu$$

Exercício 7.2, 7.4, 7.5

Exemplo 7.5: **Qual o tamanho da amostra necessário para estimar a média populacional de uma característica dimensional de um processo com 95% de confiança e precisão de 0,5cm ?**

Sem conhecimento da variabilidade populacional, estima-se o desvio-padrão populacional através de uma amostra piloto.

A partir de uma amostra de 20 peças, calculou-se o desvio-padrão S.

7	11	12	11	13	8	15	8	11	16
10	12	9	6	11	10	11	10	12	9

Como a variabilidade não é previamente conhecida, mas calculada a partir da amostra, usa-se a distribuição t.

α = 0,05 ==> $t_{0.025,19}$ = 2,093 e = 0,5 cm, S=2,45

$$n = \left(\frac{t_{\alpha/2,n-1} S}{e} \right)^2 = \left(\frac{2{,}093x2{,}46}{0{,}5} \right)^2 = 106$$

Logo é necessário coletar mais 86 (106-20) peças.

Exercício 7.13

7.7. Intervalo de confiança para a diferença entre duas médias, variância conhecida

Sejam X_1 **e** X_2 **duas variáveis aleatórias com médias** μ_1 **e** μ_2 **desconhecidas e variâncias** $\sigma_1^2 \, e \, \sigma_2^2$ **conhecidas.**

Um intervalo de confiança ***100(1-α)%*** **para a diferença entre as médias pode ser construído a partir dos resultados de amostras aleatórias de cada uma dessas populações.**

Pode ser demonstrado que a variância das diferenças entre as médias vem dada por:

$$\sigma^2 = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)$$

Assim o intervalo de confiança bilateral de *100(1-α)%* será:

$$(\bar{X}_1 - \bar{X}_2) - Z_{\alpha/2}\sigma \leq (\mu_1 - \mu_2) \leq (\bar{X}_1 - \bar{X}_2) + Z_{\alpha/2}\sigma$$

E os correspondentes intervalos unilaterais serão:

$$(\mu_1 - \mu_2) \leq \left(\bar{X}_1 - \bar{X}_2\right) + Z_{\alpha}\,\sigma$$
$$(\mu_1 - \mu_2) \geq \left(\bar{X}_1 - \bar{X}_2\right) - Z_{\alpha}\,\sigma$$

7.8. Intervalo de confiança para a diferença entre duas médias, variância desconhecida

Sejam X_1 e X_2 duas variáveis aleatórias Normais com médias μ_1 e μ_2 e variâncias $\sigma_1{}^2 e\, \sigma_2{}^2$ desconhecidas.

Se for possível assumir que as variâncias sejam iguais, ou seja, $\sigma_1{}^2 = \sigma_2{}^2$ uma estimativa da variância pode ser obtida como:

$$S_p{}^2 = \frac{(n_1 - 1)S_1{}^2 + (n_2 - 1)S_2{}^2}{n_1 + n_2 - 2}$$

7.8. Intervalo de confiança para a diferença entre duas médias, variância desconhecida

Uma vez encontrada a estimativa da variância dos valores individuais, pode ser demonstrado que a estimativa da variância da diferença entre as médias será:

$$S^2 = \left(\frac{S_p{}^2}{n_1} + \frac{S_p{}^2}{n_2} \right) = S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

com graus de liberdade $v = n_1 + n_2 - 2$

De modo que o intervalo de confiança bilateral ***100(1-α)%*** **será:**

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2,\nu} S \leq (\mu_1 - \mu_2) \leq (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2,\nu} S$$

Os correspondentes intervalos de confiança unilaterais serão:

$$\left(\mu_1 - \mu_2\right) \leq \left(\bar{X}_1 - \bar{X}_2\right) + t_{\alpha,\nu}\,\mathrm{S}$$

$$\left(\mu_1 - \mu_2\right) \geq \left(\bar{X}_1 - \bar{X}_2\right) - t_{\alpha,\nu}\,\mathrm{S}$$

Exemplo 7.6: **Um eixo deve ser montado no interior de um rolamento. Uma amostra de doze unidades indicou para o diâmetro interno do rolamento** $\overline{X}_1 = 2{,}538cm \text{ e } S_1 = 0{,}008$ **; e para o diâmetro do eixo** $\overline{X}_2 = 2{,}520cm \text{ e } S_2 = 0{,}006$ **. Calcule o intervalo de confiança de 99% para a folga de montagem.**

Solução: **Supondo variâncias iguais têm-se:**

$$S_p^{\,2} = \frac{(11)0{,}008^2 + (11)0{,}006^2}{12+12-2} = 0{,}000050$$

$$S = \sqrt{S_p^{\,2}\left(\frac{1}{12}+\frac{1}{12}\right)} = 0{,}00289$$

$$\nu = 12+12-2 = 22$$

$$t_{0{,}005;22} = 2{,}82$$

Exercício 7.8

$$(2{,}538 - 2{,}52) - 2{,}82(0{,}00289) \le \text{folga} \le (2{,}538 - 2{,}52) + 2{,}82(0{,}00289)$$

$$0{,}00986 \le \text{folga} \le 0{,}026$$

7.9. Intervalo de confiança para a diferença entre observações

No caso em que se deseja comparar dois sistemas é possível, e as vezes necessário, trabalhar com a diferença entre as observações.

Por exemplo, para comparar dois métodos de tratamento contra corrosão, pode-se escolher diversos blocos de terreno, colocar dois tubos (de marcas diferentes) em cada bloco e observar as diferenças.

7.9. Intervalo de confiança para a diferença entre observações

Seja X_1 **os resultados do sistema 1;**

X_2 **os resultados do sistema 2;**

$d = X_1 - X_2$ **as diferenças medidas bloco a bloco;**

A partir dos resultados de n blocos, calcula-se e usa-se a distribuição t para construir o intervalo de confiança para a média da diferença μd :

$$\bar{d} - t_{\alpha/2}\frac{S_d}{\sqrt{n}} \le \mu_d \le \bar{d} + t_{\alpha/2}\frac{S_d}{\sqrt{n}}$$

Se o valor zero estiver contido neste intervalo, então não pode ser descartada a hipótese que o desempenho dos dois sistemas seja o mesmo.

Exemplo 7.7: **Uma empresa quer verificar se o conhecimento de seus alunos a respeito de um determinado assunto melhorou após 30 horas de treinamento. Para isso foi realizado com os quinze alunos do treinamento um teste antes e após o treinamento. Os dados a seguir representam as notas obtidas pelos alunos. Conclua a respeito da eficiência do treinamento com 95% de confiança.**

Alunos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Antes	6,5	6,7	7,0	7,0	6,5	7,3	7,8	6,9	6,7	7,2	7,5	7,5	7,2	7,0	6,8
Depois	7,5	7,7	7,9	8,0	7,4	8,3	8,8	8,9	7,7	8,2	8,5	8,5	8,2	8,0	8,8
Difer.	1,0	1,0	0,9	1,0	0,9	1,0	1,0	2,0	1,0	1,0	1,0	1,0	1,0	1,0	2,0

$$\overline{d} = 1{,}12$$

$$S_d = 0{,}36$$

$$t_{0,025;14} = 2{,}145$$

$$1{,}12 - 2{,}145\frac{0{,}36}{\sqrt{15}} \leq \mu \leq 1{,}12 + 2{,}145\frac{0{,}36}{\sqrt{15}}$$

ou

$$0{,}92 \leq \mu \leq 1{,}32$$

Como o valor zero não está incluído no intervalo, rejeita-se a hipótese de que as notas antes e depois sejam as mesmas, logo concluí-se que o treinamento foi eficiente.

Exercício 7.9

7.10. Intervalo de confiança para a variância

Suponha que X é uma variável aleatória Normal com média e variância desconhecidas. Seja que a variância amostral S^2 é calculada para uma amostra de n observações. Então, um intervalo bilateral de confiança *100(1-α)%* é obtido usando-se a distribuição do chi-quadrado:

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}$$

No caso do interesse residir em intervalos unilaterais de

***100(1-α)%* teremos:**

Limite inferior: $$\sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{\alpha,n-1}}$$

Limite superior: $$\sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha,n-1}}$$

Exemplo 7.8. **Ache o intervalo de 95% para a variância no exemplo da quantidade mensal de produtos entregues:**

$$S^2 = 1{,}34^2 = 1{,}80$$

$$X^2{}_{0{,}025;19} = 32{,}85 \,;\; X^2{}_{0{,}975;19} = 8{,}91$$

$$\frac{19(1{,}80)}{32{,}85} \leq \sigma^2 \leq \frac{19(1{,}80)}{8{,}91}$$

$$1{,}04 \leq \sigma^2 \leq 3{,}84$$

ou

$$1{,}02 \leq \sigma \leq 1{,}96$$

Exercício 7.6 e 7.7

7.11. Intervalo de confiança para o quociente entre duas variâncias

Para comparar duas variâncias, $\sigma_1^2 \, e \, \sigma_2^2$, oriundas de populações com distribuição Normal, é vantajoso trabalhar com o quociente σ_1^2/σ_2^2 uma vez que este se distribui conforme a distribuição *F*.

O intervalo de confiança para este quociente virá dado por:

$$\frac{S_1^2}{S_2^2} F_{1-\alpha/2;n_1-1;n_2-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} F_{\alpha/2;n_1-1;n_2-1}$$

Onde $F_{\alpha,u,v}$ são os pontos percentuais da distribuição *F* com *u* e *v* graus de liberdade, tais que $P\{F \geq F_{\alpha,u,v}\} = \alpha$.

Se o valor *1* (um) estiver contido neste intervalo, então não pode ser descartada a hipótese de que a variância das duas populações seja a mesma.

Os respectivos intervalos unilaterais serão dados por:

Limite inferior: $$\frac{\sigma_1^2}{\sigma_2^2} \geq \frac{S_1^2}{S_2^2} F_{1-\alpha;n_1-1;n_2-1}$$

Limite superior: $$\frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} F_{\alpha;n_1-1;n_2-1}$$

As tabelas da distribuição F costumam fornecer apenas os valores de F_α , mas $F_{1-\alpha}$ pode ser obtido a partir da seguinte relação:

$$1-\alpha,u,v \qquad F_{\alpha,v,u}$$

Exemplo 7.9. **Os valores a seguir representam os tempos de produção de duas máquinas. Analise os dados e conclua a respeito da variabilidade das máquinas A e B:**

A	91,0	90,3	90,2	92,1	91,8	91,3	89,3,	91,0	91,2	89,6
B	91,8	91,2	89,4	89,2	90,7	92,6	91,3	91,2		

$$F_{0,025;9,7} = 4,82$$

Máquina A: $s_1^2 = 0,8307$

$$F_{0,975;9,7} = \frac{1}{F_{0,025;7,9}} = \frac{1}{4,20} = 0,238$$

Máquina B: $s_2^2 = 1,316$

$$\frac{0,8307}{1,316}(0,238) \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{0,8307}{1,316}(4,82)$$

$$0,1502 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3,0425$$

O intervalo inclui o valor 1, assim não pode ser descartada a hipótese de que a variabilidade das duas máquinas seja a mesma.

Além de servir para a comparação direta de duas variâncias, a distribuição F é a chave para a comparação de vários grupos, o que é feito usando o procedimento conhecido como Análise de Variância. Esse assunto será abordado em um capítulo posterior.

Exercício 7.10

7.12 Intervalo de confiança para o parâmetro da Binomial

Intervalos de confiança para proporções, por exemplo, fração de não conformes em um processo, podem ser estabelecidos utilizando-se a distribuição Binomial.

Se *n* é grande ($n \geq 30$) e $p \geq 0{,}1$, então a aproximação Normal para a Binomial pode ser usada, resultando no seguinte intervalo de confiança de $100(1-\alpha)\%$:

$$p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

Se *n* é pequeno, o problema deve ser resolvido usando tabelas da distribuição Binomial. Se *p* é pequeno, é possível usar a distribuição de Poisson.

Exemplo 7.10. Um empresário deseja conhecer a satisfação de seus clientes em relação aos serviços prestados por sua empresa. Em uma amostra aleatória de n=100 clientes entrevistados, 4 pessoas demonstraram insatisfação com os serviços prestados. Construa um intervalo de 95% de confiança para a proporção de clientes insatisfeitos.

$$p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

$$0{,}04 - 1{,}96\sqrt{\frac{0{,}04(1-0{,}04)}{100}} \leq \pi \leq 0{,}04 + 1{,}96\sqrt{\frac{0{,}04(1-0{,}04)}{100}}$$

$$0{,}03 \leq \pi \leq 0{,}05$$

Exercício 7.11

Exemplo 7.11. **O fornecedor alega que entrega 10% de produtos defeituosos. Qual o tamanho de amostra suficiente para estimar a proporção de produtos defeituosos entregues por este fornecedor com precisão de 0,03 e 95% de confiança?**

Solução:

$\alpha = 0{,}05 \Longrightarrow z_{0,025} = 1{,}96$; $p = 0{,}10$; $e = 0{,}03$

Como deseja-se estimar uma variável do tipo percentual, utiliza-se a distribuição Binomial.

$$n = \frac{Z^2_{\alpha/2} \times p(1-p)}{e^2} \qquad n = \frac{1{,}96^2 \times 0{,}10 \times (1-0{,}10)}{0{,}03^2} = 384{,}16$$

Logo, é necessário uma amostra de 385 produtos.

Exercício 7.14

Exercícios

7.1 O tempo de atendimento em um restaurante apresenta variância $\sigma^2 = 0,0015$ **. Uma amostra aleatória de 12 mesas indicou tempo médio de atendimento de** $\overline{X} = 12,258mm$**. Construa um intervalo de 95% de confiança para o tempo médio de atendimento no restaurante.**

7.2 Recalcule o intervalo de confiança para o exercício 7.1, supondo que a variância não fosse conhecida e o valor $S^2 = 0,0015$ **tivesse sido medido diretamente na amostra.**

7.3 O peso de frangos apresenta variância conhecida igual a 900g. Uma amostra aleatória de 20 unidades indica $\overline{X} = 508g$ **. Construa um intervalo com 90% de confiança para o peso médio desses frangos.**

7.4 Em um processo químico, as características dimensionais do produto resultante segue o modelo normal. A partir da amostra apresentada a seguir, defina o limite inferior de um intervalo unilateral de 95% de confiança para a característica dimensional média.

35.2	36.7	37.5	38.2	38.7	39.5
36.3	37.3	37.8	38.3	39.3	40.1

7.5 Uma máquina é usada para encher pacotes de leite. O volume segue aproximadamente o modelo normal. Uma amostra de 16 pacotes indicou:

1021	1016	1012	1011	1014	1018	1022	1027
1008	1015	1013	1013	1017	1019	1007	1003

a) construa um intervalo unilateral de 99% com limite inferior para a média;

b) construa um intervalo de 95% para a média;

7.6 Considere os dados do exercício 7.4. Construa um intervalo de 90% para a variância da característica dimensional. Depois converta esse intervalo apresentando-o em termos de desvio padrão.

7.7 Considere os dados do exercício 7.5. Construa um intervalo de 95% para o desvio padrão do volume dos pacotes de leite.

7.8 Ainda em relação ao problema 7.5, imagine que há uma segunda máquina de enchimento para a qual uma amostra de 16 pacotes indicou:

1011	1015	1017	1015	1021	1021	1010	1007
1022	1018	1016	1015	1020	1022	1025	1030

Construa um intervalo de 95% para a diferença entre as duas médias das máquinas. Baseado nos resultados desses cálculos você concluiria que as duas máquinas fornecem o mesmo volume médio?

7.9 Em uma indústria química, os engenheiros desejam saber se o alongamento de um composto de borracha permanece inalterado ao passar por uma máquina extrusora. Como o alongamento do composto depende do lote de matéria prima usado na sua confecção, os dados foram coletados aos pares. Construa um intervalo de confiança para a diferença entre os pares de observações:

Lote	1	2	3	4	5	6	7	8	9	10
Antes	360	370	380	345	365	380	390	395	385	410
Depois	360	365	355	340	350	370	390	375	375	395

7.10 Em relação ao problema anterior, calcule o quociente entre as variâncias dos alongamentos medidos antes e depois do composto passar pela extrusora. Depois construa um intervalo de confiança para esse quociente.

7.11 Uma amostra aleatória de 250 dispositivos eletrônicos apresentou 27 unidades defeituosas. Estime a fração de não-conformes e construa um intervalo de 95% de confiança para o verdadeiro valor da fração de não-conformes.

7.12. Qual o tamanho da amostra necessário para estimar o tempo médio de atendimento de um serviço com desvio-padrão conhecido de σ=3 min com 95% de confiança e precisão de 2 min?

7.13. Qual o tamanho da amostra necessário para estimar o tempo médio de atendimento de um serviço com 95% de confiança e precisão de 2 min? Uma amostra de 20 tempos foi coletada para estimar o desvio-padrão S.

8	10	12	11	13	8	15	8	11	14
12	12	9	7	12	10	11	10	12	8

7.14. Em uma pesquisa eleitoral, 60 das 180 pessoas entrevistadas responderam que votariam no candidato da oposição. Essa amostra é suficiente para estimar a verdadeira proporção de eleitores desse candidato, com uma precisão de 0,04 e confiança 95%?

Referências

Esta apresentação baseia-se nas referências listadas abaixo, contendo partes das mesmas, com respectivas autorizações dos autores :

***Apostila de Estatística Industrial*, Programa de Pós-Graduação de Engenharia de Produção, UFRGS, Porto Alegre.**

***Apostila de Probabilidade e Estatística* professoras Márcia Echeveste e Liane Werner.**

***Apostila Noções de Probabilidade e Estatística para Concursos* professoras Márcia Echeveste e Ariane Porto Rosa.**

Bibliografia

1. Ribeiro, J. L. D. & Caten, C. t. 2000. *Apostila de Estatística Industrial*, Programa de Pós-Graduação de Engenharia de Produção, UFRGS, Porto Alegre.
2. Miller, I. & Freund, J.E. (1977), *Probability and Statistics for Engineers*. 2nd ed., Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.
3. Freund, J. E. & Simon, G. A. (2000) *Estatística Aplicada: Economia, Administração e Contabilidade*; tradução Alfredo de Farias. Bookman, Porto Alegre, Brasil.
4. Stevenson, W. J. (1981), *Estatística Aplicada à Administração*; tradução Alfredo de Farias. Harper & Raw do Brasil, São Paulo, SP, Brasil.
5. Spiegel, M. R. (1993) *Estatística*. Makron Books Brasil Editora, São Paulo, SP, Brasil.
6. Spiegel, M. R. (1978) Probabilidade e *Estatística*. *Coleção Schaum*. McGraw-Hill do Brasil, São Paulo, SP, Brasil.