Distribuição de $\bar{X}$

Exemplo

Uma variável aleatória X tem distribuição normal, com média 100 e desvio padrão 10.

(a) Qual a $P(90 < X < 110)$?

(b) Se $\bar{X}$ for a média de uma amostra de 16 elementos retirados dessa população, calcule $P(90 < \bar{X} < 110)$.

(c) Represente, num único gráfico, as distribuições de X e $\bar{X}$.

(d) Que tamanho deveria ter a amostra para que $P(90 < \bar{X} < 110) = 0,95$?

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 274.

Distribuição de $\bar{X}$

(a) Devemos padronizar os valores, para comparar com a distribuição normal padrão, ou seja:

$$P(90 < X < 110) = P\left(\frac{90 - 100}{10} < \frac{X - 100}{10} < \frac{110 - 100}{10}\right)$$

$$= P(-1 < Z < 1) = P(Z < 1) - P(Z < -1) = \Phi(1) - \Phi(-1)$$

Consultando a tabela disponível na página da disciplina, vemos que $\Phi(1) = 0,8413$. Para encontrar $\Phi(-1)$, note que a distribuição normal padrão é simétrica em torno do zero e portanto $\Phi(-z) = 1 - \Phi(z)$. Então $\Phi(-1) = 0,1586$ e portanto $\Phi(1) - \Phi(-1) = 0,6827$.

Distribuição de $\bar{X}$

(b) Se temos uma amostra e consideramos a média, note que

$$\mathrm{Var}\left(\bar{X}\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \overset{(1)}{=} \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) \overset{(2)}{=} \frac{\mathrm{Var}\left(X\right)}{n}$$

onde a igualdade (1) vale por independência, e a igualdade (2) vale por serem identicamente distribuidas. Consequentemente, o desvio padrão novo será $\sigma/\sqrt{n}$, ou $10/4$. Temos então que

$$P(90 < \bar{X} < 110) = P\left(\frac{90-100}{10/4} < \frac{\bar{X}-100}{10/4} < \frac{110-100}{10/4}\right)$$

Distribuição de $\bar{X}$

(b) Continuando,

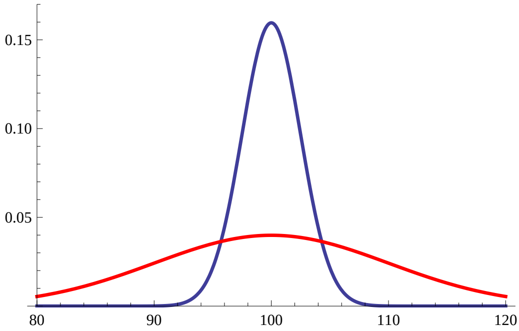
$$= P(-4 < Z < 4) = P(Z < 4) - P(Z < -4) = \Phi(4) - \Phi(-4)$$

Se consultarmos a tabela agora, veremos que a probabilidade $P(Z < 4)$ é tão grande que nem está listada. Ela então pode ser considerada como aproximadamente igual a 1.

De fato, com a ajuda de algum método de integração numérica, podemos verificar que $\Phi(4) - \Phi(-4)$ é igual a $0,9999367$.

Distribuição de $\bar{X}$

(c) No gráfico, a função de densidade de X está em vermelho, e a de $\bar{X}$ em azul:



Distribuição de $\bar{X}$

(d) Queremos resolver a seguinte equação:

$$P\left(\frac{90-100}{10/\sqrt{n}} < \frac{\bar{X}-100}{10/\sqrt{n}} < \frac{110-100}{10/\sqrt{n}}\right) = 0,95$$

Consultando a tabela, vemos que $P(-q < Z < q) = 0,95$ se $q = 1,96$. Então a equação que queremos resolver pode ser reescrita como:

$$\frac{110-100}{10/\sqrt{n}} = 1,96 \Leftrightarrow \sqrt{n}\frac{110-100}{10} = 1,96 \Leftrightarrow n = 1,96^2$$

Portanto, $n = 4$ é suficiente para obtermos a confiança desejada.

Precisão e Tamanho Amostral

Exemplo

Qual deve ser o tamanho de uma amostra se cuja a população da qual ela será sorteada possui um desvio-padrão igual a 10, de sorte que a diferença entre a média amostral para e a média da população, em valor absoluto, seja menor que 1, com coeficiente de confiança igual a:

(a) 95%

(b) 99%

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 308.

Precisão e Tamanho Amostral

(a) Note que se $X \sim N(\mu, \sigma^2)$, então $X - \mu \sim N(0, \sigma^2)$. Sabemos que $\sigma = 10$, e que o desvio-padrão do estimador da média, $\bar{X}$, será $10/\sqrt{n}$. Queremos que $P(|\bar{X} - \mu| < 1) = 0,95$. Mas este evento é equivalente a

$$P\left(-\sqrt{n}/10 < Z < \sqrt{n}/10\right)$$

Como $P(-1,96 < Z < 1,96) = 0.95$, então $\sqrt{n}/10 = 1,96$ ou $n \approx 385$.

(b) De modo análogo, temos que $P(-2,57 < Z < 2,57) = 0,99$, então $\sqrt{n}/10 = 2,57$ ou $n \approx 665$.

Intervalo de Confiança para proporções

Exemplo

Suponha que $p = 30\%$ dos estudantes de uma escola sejam mulheres. Colhemos uma amostra aleatória simples de $n = 10$ estudantes e calculamos $\hat{p} =$ proporção de mulheres na amostra. Qual a probabilidade de que $\hat{p}$ difira de p em menos de $0,01$? E se $n = 50$?

Adaptado de: Morettin & Bussab, Estatística Básica 5ª edição, pág 276.

Intervalo de Confiança para proporções

Temos que a probabilidade que desejamos encontrar é dada por

$$P\left(|\hat{p} - p| < 0,01\right) = P\left(-0,01 < \hat{p} - p < 0,01\right)$$

Onde p é o valor verdadeiro da proporção de mulheres, e $\hat{p}$ a proporção observada na amostra. Sabemos que se n é grande, $\hat{p} - p$ pode ser aproximada por uma normal $N\left(0, p(1-p)/n\right)$. Como $p = 0,3$, temos que

$$\text{Var}\left(\hat{p} - p\right) = \frac{0,3 \times 0,7}{10} = 0,021.$$

Intervalo de Confiança para proporções

Portanto, a probabilidade pedida é igual a

$$P\left(\frac{-0,01}{\sqrt{0,021}} < Z < \frac{0,01}{\sqrt{0,021}}\right) = P(-0,07 < Z < 0,07) = 0,056$$

. Mas $n = 10$ é suficientemente grande? Podemos comparar essa probabilidade com o resultado exato.

Não sabemos a distribuição de $\hat{p}$, mas o evento $\hat{p} = \alpha$ é igual ao evento $\sum X_i = n\alpha$, onde X_i são v.a. independentes e identicamente distribuidas Bernoulli(0.3). A soma é portanto distribuição binomial$(10; 0,3)$.

Intervalo de Confiança para proporções

O evento $\{|\hat{p} - p| < 0.01\}$ é igual ao evento $\{|\sum X_i - 10 \times 0,3| < 0,1\}$.

Como $\sum X_i$ assume somente valores inteiros, temos que

$$\left\{\left|\sum_{i=1}^{10} X_i - 10 \times 0,3\right| < 0,1\right\} = \left\{\sum_{i=1}^{10} X_i = 3\right\}.$$

Portanto,

$$P\left(\left\{\sum_{i=1}^{10} X_i = 3\right\}\right) = \binom{10}{3} 0,3^3 0,7^7 = 0,267.$$

Temos uma probabilidade que é 5 vezes maior que a aproximação.

Intervalo de Confiança para proporções

Tome $n = 50$, agora, Podemos modificar rapidamente as contas da aproximação normal, A variância agora é $0,0042$, e portanto a probabilidade aproximada é

$$P\left(\frac{-0,01}{\sqrt{0,0042}} < Z < \frac{0,01}{\sqrt{0,0042}}\right) = P(-0,154 < Z < 0,154) = 0,12239$$

A probabilidade exata agora é dada pelo evento $\{|\sum X_i - 50 \times 0,3| < 0,5\}$, ou simplesmente $\{\sum_{i=1}^{50} X_i = 15\}$.

Intervalo de Confiança para proporções

Observe agora que

$$P\left(\sum_{i=1}^{50} X_i = 15\right) = \binom{50}{15} 0,3^{15} 0,7^{50-15} = 0,12237$$

A diferença agora é muito menor e, é possível demonstrar, que à medida que $n \to \infty$ ela tende à 0. É preciso contudo ter em mente que a aproximação só é válida para grandes tamanhos de amostra, independentes e identicamente distribuidas.

Intervalo de Confiança para proporções

Exemplo

Uma amostra aleatória de 625 donas de casa revela que 70% delas preferem a marca A de detergente. Construir um intervalo de confiança para $p =$ proporção das donas de casa que preferem A, com coeficiente de confiança $\gamma = 90\%$.

Fonte: Morettin & Bussab, Estatística Básica 5^a edição, pág 308.

Intervalo de Confiança para proporções

Temos que em nossa amostra aleatória, $\hat{p} = 0,7$. Como $\hat{p} \sim N(p, p(1-p)/n)$, então um intervalo de confiança é dado por

$$\left(\hat{p} - z_\gamma \sqrt{\hat{p}(1-\hat{p})/n}\ ;\ \hat{p} + z_\gamma \sqrt{\hat{p}(1-\hat{p})/n}\right)$$

Temos que para $\gamma = 0,90$, $z_\gamma = 1,64$ e portanto um intervalo de confiança para a proporção de donas de casa que preferem o detergente A é dado por

$$\left(0,7 - 1,64\sqrt{0,7 \times 0,3/625}\ ;\ 0,7 + 1,64\sqrt{0,7 \times 0,3/625}\right)$$

Intervalo de Confiança para proporções

Exercício

Suponha que estejamos interessados em estimar a porcentagem de consumidores de um certo produto. Se umaa amostra de tamanho 300 forneceu 100 indivíduos que consomem o dado produto, determine:

(a) Um intervalo de confiança de p, com c.c. de 95%; interprete o resultado.

(b) O tamanho da amostra para que o erro da estimativa não exceda 0,02 unidades com probabilidade de 95%; interprete o resultado.

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 309.

Intervalo de Confiança para proporções

(a) Um intervalo de confiança com 95% de confiabilidade é dado por:

$$IC(p; 0,95) = 0,333 \pm 1,96\sqrt{\frac{0,333 \times 0,667}{300}} = 0,333 \pm 0,053$$

Ou simplesmente $(0,280; 0,387)$.

Interpretação: Se pudéssemos construir um determinado número de intervalos aleatórios para p, todos baseados em amostras de tamanho n, 95% deles conteriam o parâmetro p.

Intervalo de Confiança para proporções

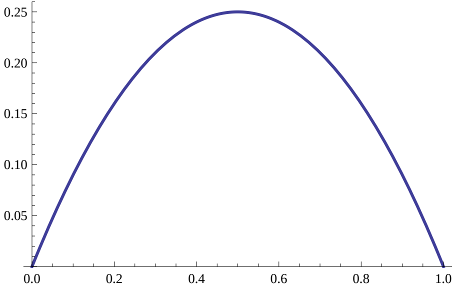
(b) Utilizando a estimativa da amostra observada ($\hat{p} = 0,333$), temos que n é dado por

$$n = \left(\frac{1,96}{0,02}\right)^2 \times 0,333 \times 0,667 \cong 2134.$$

Contudo, frequentemente devemos determinar o tamanho da amostra antes de realizar qualquer experimento, isto é, sem nenhuma informação prévia de p. Se esse for o caso, devemos considerar o caso em que a variância da amostra é a pior possível.

Intervalo de Confiança para proporções

(b) Se olhamos a variância como função de p, obtemos o seguinte gráfico:



Note que a variância é máxima quando $p = 1/2$.

Intervalo de Confiança para proporções

(b) Utilizando o valor máximo de $p(1-p)$, isto é, $1/4$, obtemos

$$n = \left(\frac{1.96}{0.02}\right)^2 \times \frac{1}{4} \cong 2401$$

Interpretação: Utilizando o tamanho amostral encontrado, teremos uma probabilidade de 95% de que a proporção amostral não difira do verdadeiro valor de p em menos que 2%.

Obter amostras pequenas para examinar p, e então determinar o tamanho amostral sem utilizar o "pior caso", é no que consiste a idéia de *amostras piloto*.

Intervalo de Confiança para a média

- Seja $X \sim N(\mu, \sigma^2)$ e X_1, ..., X_n uma amostra casual simples de X.
- Considere σ^2 conhecido e que queremos construir um IC com cc $\gamma \times 100\%$ para μ.
- Neste caso, $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$. Então:

Intervalo de Confiança para a média

$$
\begin{aligned}
\gamma &= P\left(-z_\gamma \leq Z \leq z_\gamma\right) \\
&= P\left(-z_\gamma \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_\gamma\right) \\
&= P\left(-z_\gamma \frac{\sigma}{\sqrt{n}} \leq \overline{X}-\mu \leq z_\gamma \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\overline{X} - z_\gamma \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_\gamma \frac{\sigma}{\sqrt{n}}\right)
\end{aligned}
$$

$$IC(\mu,\gamma) = \left[\overline{X} - z_\gamma \tfrac{\sigma}{\sqrt{n}}; \overline{X} + z_\gamma \tfrac{\sigma}{\sqrt{n}}\right]$$

Intervalo de Confiança para a média

- Se σ^2 for desconhecido temos que $T = \frac{\overline{X}-\mu}{\frac{\widehat{\sigma}}{\sqrt{n}}} \sim t_{(n-1)}$, em que $\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ e $\widehat{\sigma} = \sqrt{\widehat{\sigma}^2}$. Então:

$$IC(\mu,\gamma) = \left[\overline{X} - t_\gamma \frac{\sigma^2}{\sqrt{n}}; \overline{X} + t_\gamma \frac{\sigma^2}{\sqrt{n}}\right]$$

  em que $P(T \leq t_\gamma) = \frac{1+\gamma}{2}$, $T \sim t_{(n-1)}$.

- Se X não tiver distribuição normal, mas n for suficientemente grande, com $E(X) = \mu$ e $Var(X) = \sigma^2$, os resultados continuam válidos, mas os IC's são aproximados.

Intervalo de Confiança para a média

a) Uma indústria química objetiva saber qual o tempo médio de reação de um determinado reagente. Para isso coletou uma amostra de 50 reagentes e monitorou o tempo de reacão (em minutos) de cada um, resultando em $\overline{x} = 14,5$. Supondo $\sigma^2 = 100$, construa um $IC(\mu, 0,99)$.

b) Repita o item *a*) considerando σ^2 desconhecido, e $\widetilde{\sigma}^2 = 110,87$.

Intervalo de Confiança para a média

a) $z_\gamma = 2,57,\ IC(\mu, 0,99) =$
$[14,5 - 2,57 \times 10/\sqrt{50}; 14,5 + 2,57 \times 10/\sqrt{50}] = [10,87; 18,13].$

b) $t_\gamma = 2,704,\ IC(\mu, 0,99) = [14,5 - 2,704 \times \sqrt{110,87}/\sqrt{50}; 14,5 + 2,704 \times \sqrt{110,87}/\sqrt{50}] = [10,47; 18,53].$

Intervalo de Confiança para diferenças entre médias

Exemplo

Estão sendo estudados dois processos para conservar alimentos, cuja principal variável de interesse é o tempo de duração destes. No processo A, o tempo X de duração segue a distribuição $N(\mu_A, 100)$, e no processo B o tempo Y obedece à distribuição $N(\mu_B, 100)$. Sorteiam-se duas amostras independentes: a de A, com 16 latas, apresentou tempo médio de duração igual a 50, e a de B, com 25 latas, duração média igual a 60.

(a) Construa um IC para μ_A e μ_B, separadamente.

Exemplo

(b) Para verificar e os dois processos podem ter o mesmo desempenho, decidiu-se construir um IC para a diferença $\mu_A - \mu_B$. Caso o zero pertença ao intervalo, pode-se concluir que existe evidência de igualdade dos processos. Qual seria sua resposta?

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 318.

(a) Para o caso geral, um intervalo de confiança para μ, para variância conhecida, com coeficiente de confiabilidade γ, é dado por

$$\left(\bar{X} - z_\gamma\sqrt{\sigma^2/n}\ ;\ \bar{X} + z_\gamma\sqrt{\sigma^2/n}\right)$$

Note que $\sigma_A = \sigma_B$. Para um coeficiente de confiança $\gamma = 0,95$, por exemplo, temos $z_\gamma = 1,96$, e os intervalos de confiança serão, respectivamente:

$$\text{IC}(\mu_A) = \left(50 - 1,96\sqrt{100/16}\ ;\ 50 + 1,96\sqrt{100/16}\right)$$

$$\text{IC}(\mu_B) = \left(60 - 1,96\sqrt{100/25}\ ;\ 60 + 1,96\sqrt{100/25}\right)$$

Intervalo de Confiança

(a) (cont.) Fazendo as contas, obtemos que

$$\text{IC}(\mu_A) = (45,1; 54,9)$$

$$\text{IC}(\mu_B) = (56,08; 63,92)$$

Observe que os intervalos não se interceptam; temos evidência para dizer que as durações médias são diferentes, a 95% de confiança.

(b) Temos aqui duas amostras diferentes mas independentes. A diferença $\bar{X}_A - \bar{X}_B$ tem distribuição Normal, com média $\mu_A - \mu_B$ e variância $\sigma_A^2/n_A + \sigma_B^2/n_B$.

(b) (cont.) Então um intervalo de confiança para $\mu_A - \mu_B$ é dado por

$$\left(\bar{X}_A - \bar{X}_B - z_\gamma \sqrt{\sigma_A^2/n_A + \sigma_B^2/n_B}\ ;\right.$$

$$\left.\bar{X}_A - \bar{X}_B + z_\gamma \sqrt{\sigma_A^2/n_A + \sigma_B^2/n_B}\right)$$

Aplicando os valores conhecidos ou observados, e fixando a confiança em $\gamma = 0,95$ temos:

$$\mathrm{IC}(\mu_A - \mu_B) = \left(50 - 60 - 1,96\sqrt{100/16 + 100/25};\right.$$

$$\left.50 - 60 + 1,96\sqrt{100/16 + 100/25}\right)$$

Intervalo de Confiança

(b) (cont.) Executando as contas, obtemos finalmente que

$$\text{IC}(\mu_A - \mu_B) = (-16,27\ ;\ -3,72)$$

Em concordância com o item (a), vemos que 0 não está contido no intervalo e, portanto, rejeitamos a hipótese, a $\gamma = 0,95$ de confiança, das médias μ_A e μ_B serem iguais.

Intervalo de Confiança

- Considere agora que as variâncias (σ^2_A, σ^2_B) são desconhecidas, porém iguais, ou seja $\sigma^2_A = \sigma^2_B = \sigma^2$. Neste caso, $\frac{\overline{X}_A - \overline{X}_B}{\sqrt{\widehat{\sigma}^2\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} \sim t_{(n_A+n_B-2)}$, em que $\widehat{\sigma}^2 = \frac{(n_A-1)\widehat{\sigma}^2_A + (n_B-1)\widehat{\sigma}^2_B}{n_A+n_B-2}$, $\widehat{\sigma}^2_A = \frac{1}{n_A-1}\sum_{i=1}^{n_A}\left(X_{iA} - \overline{X}_A\right)^2$, $\widehat{\sigma}^2_B = \frac{1}{n_B-1}\sum_{i=1}^{n_B}\left(X_{iB} - \overline{X}_B\right)^2$, $\overline{X}_A = \frac{1}{n_A}\sum_{i=1}^{n_A} X_{iA}$, $\overline{X}_A = \frac{1}{n_B}\sum_{i=1}^{n_B} X_{iB}$ e $X_{iA}, i = 1, ..., n_A$ e $X_{iB} = i = 1, ..., n_B$. Assim

$$
\begin{aligned}
IC(\mu_A - \mu_B, \gamma) \quad = \quad & \left[(\overline{X}_A - \overline{X}_B) - t_\gamma \sqrt{\widehat{\sigma}^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}\right. \\
; \quad & \left.(\overline{X}_A - \overline{X}_B) + t_\gamma \sqrt{\widehat{\sigma}^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}\right]
\end{aligned}
$$

Intervalo de Confiança

- Em que $P(T \leq t_\gamma) = \frac{1+\gamma}{2}$, $T \sim t_{(n_A+n_B-2)}$.
- Suponha que no item anterior, $\sigma_A^2 = \sigma_B^2 = \sigma^2$, $\widetilde{\sigma}_A^2 = 104,23$ e $\widetilde{\sigma}_B^2 = 98,50$.
- IC para cada média: $IC(\mu_A, 0,95) = [50 - 2,131 \times \frac{\sqrt{104,23}}{4}; 50 + 2,131 \times \frac{\sqrt{104,23}}{4}] = [44,56; 55,44]$ e $IC(\mu_B, 0,95) = [60 - 2,064 \times \frac{\sqrt{98,50}}{5}; 60 + 2,064 \times \frac{\sqrt{98,50}}{5}] = [55,90; 64,10]$.

Intervalo de Confiança

- $\widetilde{\sigma}^2 = \dfrac{15 \times 104,23 + 24 \times 98,50}{39} = 100,70,$

$$
\begin{aligned}
IC(\mu_A - \mu_B; 0,95) &= \Big[(50-60) - 2,021 \times \sqrt{100,70 \times (1/16 + 1/25)} \\
&; \quad (50-60) + 2,021 \times \sqrt{100,70 \times (1/16 + 1/25)}\Big] \\
&= [-16,49; -3,51]
\end{aligned}
$$

Intervalo de Confiança para a variância

Exemplo

Os dados abaixo referem-se às vendas diárias, em reais, durante uma semana, de carros de uma revendedora. Construa um $IC(\sigma^2; 90\%)$.

Vendas: 253, 187, 96, 450, 320, 105.

Fonte: Bussab & Morettin, Estatística Básica 6ª edição, pág 354.

Intervalo de Confiança para a variância

A construção do IC($\sigma^2; \gamma$) é feita a partir da seguinte expressão

$$P\left(q_{1\gamma} \leq \frac{(n-1)\widehat{\sigma}^2}{\sigma^2} \leq q_{2\gamma}\right) = \gamma, \tag{1}$$

que permite obter a seguinte desigualdade:

$$\frac{(n-1)S^2}{q_{2\gamma}} \leq \sigma^2 \leq \frac{(n-1)\sigma^2}{q_{1\gamma}}, \tag{2}$$

que será o IC procurado, em que $\frac{(n-1)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

Intervalo de Confiança para a variância

Voltando ao exemplo, calculamos inicialmente a variância amostral, que é $s_0^2 = 18,460$; em seguida, os valores de $\mathcal{X}_1^2$ e $\mathcal{X}_2^2$ que satisfaçam (1):

$$P(1,145 \leq \mathcal{X}_5^2 \leq 11,070) = 0,90.$$

Substituindo em (2) obtemos

$$\mathrm{IC}(\sigma^2; \gamma) = [8,338; 80,611].$$

Intervalo de Confiança para a razão de variâncias

Exemplo

Uma companhia fabrica propulsores para o uso em motores de turbinas a jato. Uma das operações envolve dar um acabamento, esmerilhando uma determinada superfície de um componente feito com liga de titânio. Dois processos diferentes para esmerilhar podem ser usados, podendo produzir peças com iguais rugosidades médias na superfície.
Uma amostra aleatória de $n_1 = 11$ peças provenientes do primeiro processo resulta em um desvio-padrão de $\widetilde{\sigma}_1 = 5,1$ (desvio-padrão) micropolegadas. Uma amostra aleatória $n_2 = 16$ peças, proveniente do segundo processo, resulta em um desvio-padrão de $\widetilde{\sigma}_2 = 4,7$ (desvio-padrão) micropolegadas. Construa um intervalo de confiança de 90% para a razão de variâncias σ_1^2/σ_2^2.

Fonte: http://www.cin.ufpe.br/ rmcrs/ESAP/arquivos/cap10.pdf.

Intervalo de Confiança para a razão de variâncias

Já vimos que

$$U = \frac{(n-1)\widehat{\sigma}_1^2}{\sigma_1^2} \sim \mathcal{X}^2_{(n-1)},$$

$$V = \frac{(m-1)\widehat{\sigma}_2^2}{\sigma_2^2} \sim \mathcal{X}^2_{(m-1)},$$

e portanto

$$\frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2} = \frac{\dfrac{U}{(n-1)}}{\dfrac{V}{(m-1)}} \sim \mathcal{F}(n-1, m-1). \tag{3}$$

Para um dado $\gamma, 0 < \gamma < 1$, podemos encontrar dois valores f_1 e f_2, tais que

$$P(f_1 < \mathcal{F}(n-1, m-1) < f_2) = \gamma.$$

Intervalo de Confiança para a razão de variâncias

Dessa igualdade, segue-se que, com probabilidade γ,

$$f_1 < \frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2}\frac{\sigma_2^2}{\sigma_1^2} < f_2,$$

ou seja, um IC$(\sigma_2^2/\sigma_1^2;\gamma)$ será dado por

$$f_1\frac{\widehat{\sigma}_2^2}{\widehat{\sigma}_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < f_2\frac{\widehat{\sigma}_2^2}{\widehat{\sigma}_1^2}. \quad (4)$$

Intervalo de Confiança para a razão de variâncias

Voltando ao exemplo, temos que o IC dado por (4) ficaria, com $\gamma = 0,90$, da seguinte forma

$$\frac{1}{2,85}\left(\frac{22,09}{26,01}\right) < \frac{\sigma_2^2}{\sigma_1^2} < 2,54\left(\frac{22,09}{26,01}\right),$$

ou seja,

$$0,298 < \frac{\sigma_2^2}{\sigma_1^2} < 2,15.$$

Invertendo, obtemos o IC de confiança desejado, que é dado por:

$$0,463 < \frac{\sigma_1^2}{\sigma_2^2} < 2,159.$$

Intervalo de Confiança para a diferença entre proporções

- Semelhante ao caso do IC para uma proporção (e diferença de médidas), sendo $\widehat{p}_i$, $i = 1, 2$ a proporção amostral, p_i a propoção populacional e n_i o tamanho da amostra do grupo $i = 1, 2$, podemos construir os seguintes intervalos de confiança:

$$IC(p_1 - p_2, \gamma) = \left[\widehat{p}_1 - \widehat{p}_2 \pm z_\gamma \sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}\right]$$

$$IC(p_1 - p_2, \gamma) = \left[\widehat{p}_1 - \widehat{p}_2 \pm z_\gamma \sqrt{\frac{1}{4n_1} + \frac{1}{4n_2}}\right]$$

em que z_γ é obtido da mesma forma que para o IC para uma única propoção (dist. N(0,1)).