Capítulo 1

Intervalos de Confiança

Esta é uma técnica de inferência estatística, que a partir de um intervalo de confiança, construido com os elemntos amostrais, pode-se inferir sobre um parâmetro da população.

Seja θ um parâmetro populacional e $\bar{\theta}$ o estimado de θ. Conhecendo-se a distribuição de probabilidade de $\bar{\theta}$, é possível construir um intervalo $\bar{\theta}_1 \leq \theta \leq \bar{\theta}_2$ que contém θ, e se exigir que a probabilidade do intervalo seja de $(1-\alpha)$ que é o nível de confiança. Geralmente, $(1-\alpha) \times 100 = 90\%, 95\%, 99\%$, e α é chamado de nível de significância, isto é, representa o erro que se está cometendo quando se afirma que a probabilidade do intervalo $[\bar{\theta}_1; \bar{\theta}_2]$ conter o verdadeiro parâmetro populacional θ é $(1-\alpha)$.

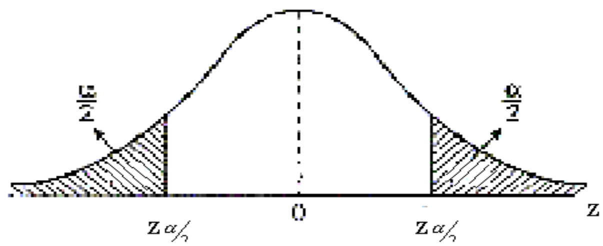
1.1 Intervalo de Confiança para a Média Populacional (μ), quando a Variância Populacional (σ^2) é conhecida.

Seja uma população $X \stackrel{\mathrm{d}}{=} N(\mu, \sigma^2)$. Sabemos que

$$\bar{X} \stackrel{\mathrm{d}}{=} N(\mu, \frac{\sigma^2}{n}) \quad \text{e} \quad Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \stackrel{\mathrm{d}}{=} N(0,1).$$

Com base na figura temos que

$$P(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}) = 1-\alpha \quad \text{ou} \quad P(-Z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}) = 1-\alpha.$$



Isolando o parâmetro populacional μ teremos:

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

Portanto, $IC_{\mu,\sigma^2} = [\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$ é o intervalo de confiança para a média populacional μ, com σ^2 conhecida, com nível de significância $\alpha\%$.

Exemplo: Feito um ensaio de corrosão com 64 peças de um lote de produção verificou-se que o tempo médio que a peça suportou nesse teste foi de $\bar{X} = 200$ horas. Calcular um intervalo de 95% de confiança para a verdadeira média populacional μ, sabendo-se que $\sigma = 16$ horas.

Solução: *Temos que como* $\alpha = 0,05 \Rightarrow\ Z_{\alpha/2} \approx 1,96.$

$$P(200 - 1,96\frac{16}{\sqrt{64}} \leq \mu \leq 200 + 1,96\frac{16}{\sqrt{64}}) = 0,95 \ \Rightarrow\ P(196,08 \leq \mu \leq 203,92) = 0,95.$$

Logo, $IC_{\mu,\sigma^2} = [196,08h; 203,92h]$ *contém a duração média, em horas, das peças com 95% de confiança.*

Exemplo: Sabe-se que os comprimentos das barras de ferro produzidas por uma siderúrgica têm uma distribuição normal com variância $\sigma^2 = 1,69\ m^2$. Numa amostra de 5 barras encontrou-se os seguintes valores: 20,1;21,0;21,4;22,1;23,3 m. Determinar o intervalo de comfiança para a média, com $\alpha = 0,06$.

Solução: *Temos que como* $\alpha = 0,06 \Rightarrow\ Z_{\alpha/2} \approx 1,881$ *e* $\sigma = 1,3m$.

Devemos encontra $\bar{X}$, *então:*

$$\bar{X} = \frac{20,1 + 21,0 + 21,4 + 22,1 + 23,3}{5} = 21,58.$$

$P(21,58-1,88\frac{1,3}{\sqrt{5}} \leq \mu \leq 21,58+1,88\frac{1,3}{\sqrt{5}}) = 0,94 \Rightarrow P(20,49 \leq \mu \leq 22,67) = 0,96.$

Logo, $IC_{\mu,\sigma^2} = [20,49m; 22,67m]$ *contém o comprimento médio das barras de ferro com 94% de confiança.*

1.2 Intervalo de Confiança para Médias Populacionais (μ), quando a Variância Populacional (σ^2) é desconhecida.

Neste caso devemos estimar a variância populacional pela variância amostral S^2, onde

$$S^2 = \frac{1}{n-1}[\sum x_i^2 - \frac{(\sum x_i)^2}{n}] = \frac{1}{n-1}[\sum x_i^2 - n\bar{X}^2]$$

Já é conhecido que $t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$, com $(n-1)$ graus de liberdade. Portanto,

$$P(-t_{\alpha/2} \leq t \leq t_{\alpha/2}) = 1-\alpha \quad \text{ou} \quad P(-t_{\alpha/2} \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq t_{\alpha/2}) = 1-\alpha.$$

Isolando μ teremos:

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

Portanto, $IC_{\mu,S^2} = [\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}}; \bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}]$ é o intervalo de confiança para a média populacional μ, com σ^2 desconhecida, com nível de significância $\alpha\%$.

Exemplo: A seguinte amostra,6;6;7;8;9;9;9;10;11;12, foi retirada de uma população normal. Construir um intervalo para μ, ao nível de significância de 10%.

Solução: *Temos* $\alpha = 0,10, \quad \phi = n-1 = 10-1 = 9 \quad \Rightarrow \quad t_{\alpha/2} = 1,833.$

Então, devemos calcular $\bar{X}$ *e* S*, considerando* $\sum x_i = 87 \quad \sum x_i^2 = 793$*:*

$$\bar{X} = 8,7, \qquad S^2 = \frac{1}{9}[793 - 10(8,7)^2] = 4 \Rightarrow S = 2.$$

Logo,

$$P(8,7 - 1,833\frac{2}{\sqrt{10}} \leq \mu \leq 8,7 + 1,833\frac{2}{\sqrt{10}}) = 0,90.$$

Portanto, $IC_{\mu,S} = [7,54; 9,86]$ *com 90% de confiança.*

1.3 Intervalo de Confiança para a Proporção (p)

Como já estudado em distribuição amostral, $\hat{p}$ que é o estimador de p, têm distribuição amostral dada por:

$$\hat{p} \stackrel{\mathrm{d}}{=} N(p, \frac{pq}{n}) \Rightarrow \quad \text{população infinita.}$$

$$\hat{p} \stackrel{\mathrm{d}}{=} N(p, \frac{pq}{n}\left(\frac{N-n}{N-1}\right)) \Rightarrow \quad \text{população finita.}$$

A variável padronizada de $\hat{p}$, para população infinita, é dada por:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}q}{n}}}.$$

Então, temos o seguinte intervalo de confiança:

$$P\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{pq}{n}} \leq p \leq \hat{p} + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right) = 1 - \alpha.$$

Substituir p por $\hat{p}$ e $\hat{q}$ por $1 - \hat{p}$, teremos:

1- População infinita ($n \geq 30$):

$$IC_p = \left[\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right].$$

2- População finita:

$$IC_p = \left[\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}\left(\frac{N-n}{N-1}\right)}; \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}\left(\frac{N-n}{N-1}\right)}\right].$$

ao nível $(1 - \alpha)$ de confiança.

Exemplo: Retirada uma amostra de 1000 peças da produção de uma máquina, verificou-se que 35 eram defeituosas. Construir um intervalo de confiança, ao nível de 95%, para a proporção real das peças defeituosas fornecidas por tal máquina.

Solução: $\hat{p} = \frac{35}{1000} = 0.035 \ (3,5\%) \qquad (1 - \hat{p}) = 0,965 \ (96,5\%).$

$\alpha = 0,05 \quad \Rightarrow \quad Z_{0,025} \approx 1,960.$

Portanto, $P\left(0,035 - 1,96\sqrt{\frac{0,035.0,965}{1000}} \leq p \leq 0,035 + 1,96\sqrt{\frac{0,035.0,965}{1000}}\right) = 0,95.$

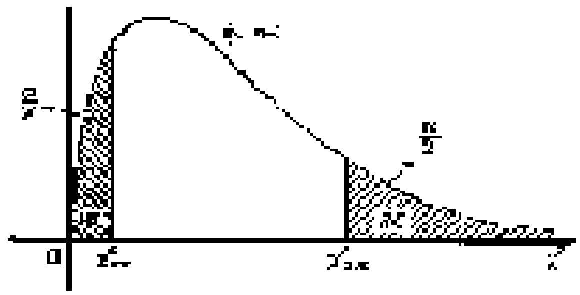
Daí,

$$IC_p = [2,36\%; 4,64\%]$$

ao nível de 95% de confiança.

1.4 Intervalo de Confiança para a variância (σ^2)

O estimador de σ^2 é S^2. Já foi visto que $\frac{(n-1)S^2}{\sigma^2}$ têm distribuição qui-quadrado (χ^2) com $(n-1)$ graus de liberdade. Ou seja, $\chi^2 \stackrel{d}{=} \frac{(n-1)S^2}{\sigma^2}$.



Lembrando que: $\chi^2_{inf} = \chi^2_{(1-\alpha/2)}$ e $\chi^2_{sup} = \chi^2_{\alpha/2}$.

Portanto, $P(\chi^2_{inf} \leq \chi^2 \leq \chi^2_{sup}) = 1 - \alpha$. Substituindo o valor de χ^2 e isolando σ^2 teremos:

$$P\left(\frac{(n-1)S^2}{\chi^2_{sup}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{inf}}\right) = 1 - \alpha.$$

Então, temos o

$$IC_{\sigma^2} = \left[\frac{(n-1)S^2}{\chi^2_{sup}}; \frac{(n-1)S^2}{\chi^2_{inf}}\right],$$

ao nível $(1-\alpha)$ de confiança, onde $(n-1)$ é o grau de liberdade.

Exemplo: Admita-se $n = 10$, $S^2 = 4$ e que se deseja construir um intervalo de confiança para a variância ao nível 90%.

Solução: *Conhecemos os seguintes dados:* $n = 10$, $S^2 = 4 \Rightarrow S = 2$, $\alpha = 10\%$, $\varphi = (n-1) = 9$.

$$\chi^2_{inf} = \begin{cases} 1 - \alpha/2 = 0.95 \\ \\ \varphi = 9 \end{cases} \Rightarrow \chi^2_{0{,}95} = 3{,}33.$$

$$\chi^2_{sup} = \begin{cases} \alpha/2 = 0.05 \\ \\ \varphi = 9 \end{cases} \Rightarrow \chi^2_{0{,}05} = 16{,}9.$$

Logo,

$$IC_{\sigma^2} = \left[\frac{9 \times 4}{16,9}; \frac{9 \times 4}{3,33}\right] = [2,13; 10,83] \quad \textit{ao nível de 90\% de confiança.}$$

Exemplo: De uma população normal foi retirada uma amostra de 15 elementos e calculou-se $\sum x_i = 8,7$ e $\sum x_i^2 = 27,3$. Determinar o intervalo de confiança de 80% para a variância populacional.

Solução: *Conhecemos os seguintes dados:* $n = 15, \alpha = 20\%$.

$$\chi^2_{inf} = \begin{cases} 1 - \alpha/2 = 0,9 \\ \\ \varphi = 14 \end{cases} \Rightarrow \chi^2_{0,9} = 7,790.$$

$$\chi^2_{sup} = \begin{cases} \alpha/2 = 0,1 \\ \\ \varphi = 14 \end{cases} \Rightarrow \chi^2_{0,1} = 21,064.$$

Como, $S^2 = \frac{1}{4}\left[27,3 - \frac{(8,7)^2}{15}\right] = 1,59.$

Então,

$$IC_{\sigma^2} = \left[\frac{14 \times 1,59}{21,064}; \frac{14 \times 1,59}{7,790}\right] = [1,03; 2,86] \quad \textit{ao nível de 80\% de confiança.}$$

1.5 Intervalo de Confiança para o Desvio Padrão (σ^2)

Como o desvio padrão é a raiz quadrada da variância, pode-se usar a seguinte fórmula:

$$P\left(S\sqrt{\frac{(n-1)}{\chi^2_{sup}}} \leq \sigma \leq S\sqrt{\frac{(n-1)}{\chi^2_{inf}}}\right) = 1 - \alpha.$$

Então, temos o

$$IC_\sigma = \left[S\sqrt{\frac{(n-1)}{\chi^2_{sup}}}; S\sqrt{\frac{(n-1)}{\chi^2_{inf}}}\right],$$

ao nível $(1 - \alpha)$ de confiança, onde $(n - 1)$ é o grau de liberdade.

1.6 Exercícios

1- De uma produção diária de parafusos foi retirada uma amostra de 25 parafusos, encontrando-se uma média de 5,2mm de diâmetro.Sabendo-se que os diâmetros têm distribuição normal com desvio-padrão populacional 1,2mm, construir intervalos de confiança para a média aos nível de 95% e 99%.

2- Suponha que as alturas dos alunos da UFRuralRJ tenham distribuição normal com $\sigma = 15\ cm$. Foi retirada uma amostra de 100 alunos obtendo-se $\bar{X} = 175\ cm$. Construir o intervalo de confiança para a verdadeira altura média dos alunos, ao nível de 95% de confiança.

3- Sejam $n = 10$, $\bar{X} = 110$ e $S = 10$. Determinar os intervalos de confiança para μ aos níveis de 90% e 95%.

4- Supondo populações normais, construir o intervalo de confiança para a variância ao nível de 90% para as amostras:

a) 43,9; 43,1; 44; 43,9; 42,2; 45,5.

b) 2; 2; 2; 3; 3; 4; 5; 5; 5; 5; 6; 6; 7; 7; 8.

5- Supondo, que de uma população Normal com μ e σ^2 desconhecido, foi retirada uma amostra de tamanho 15 fornecendo os valores $\sum x_i = 8,7$ e $\sum x_i^2 = 27,3$. Determinar para σ^2 um intervalo de confiança de 95%.

6- Tomando-se uma amostra de 100 componentes elétricos verificou-se que 93 deles funcionaram mais de 500 horas. Determinar um intervalo de confiança de 95% para a proporção.

7- Em uma pesquisa comercial em uma cidade foram selecionados 400 domicílios mostrando que 25% deles são financiados pela CEF. Determinar o intervalo de confiança, com 98%, da proporção de casas financiadas.

8-(Jairo Simon) Para verificar se um dado é viciado, jogou-se o mesmo 120 vezes, obtendo-se 25 vezes o númeo cinco. Calcular um intervalo de confiança para a proporção $\alpha = 1\%$. Pode-se dizer que o dado é viciado?

Respostas:

1- $IC = [4,73; 5,67]$ e $IC = [4,58; 5,82]$.

2- $IC = [172,06cm; 177,94cm]$.

3- $IC = [104,2; 115,8]$.

4- $a)$ $IC = [0,32; 3,13]$ $b)$ $IC = [2,25; 8,13]$.

5- $IC = [0,85; 3,95]$.

6- $IC = [0,88; 0,98]$.

7- $IC = [19,96\%; 30\%]$.

8- $IC = [0,11; 0,31]$ Não se pode dizer que o dado é viciado ao nível de 99% de confiança.