Estimação e Intervalos de Confiança

10.1 X: "tempo de reação de um condutor de camião da amostra de 140, escolhido aleatoriamente".

A amostra é suficientemente grande para que, segundo o teorema do limite central, $\overline{X}$ siga aproximadamente uma distribuição normal.

$\mu_{\overline{X}} = \mu_X = 0,85 \qquad \sigma_{\overline{X}}^2 = \frac{1}{140} \times \sigma_X^2 = \frac{1}{140} \times 0,20^2 = 2,86 \times 10^{-4} \qquad \sigma_{\overline{X}} = 0,017$

$\overline{X} \sim N(0,85; 1,43 \times 10^{-3})$ A amostra é de grande dimensão $(s \simeq \sigma_{\overline{X}})$

$\frac{\alpha}{2} = (1 - 0,95)/2 = 0,025$ O intervalo de confiança é:

$$\left[\overline{X} - Z_{(\alpha/2)} \times \frac{\sigma}{\sqrt{N}}, \overline{X} + Z_{(\alpha/2)} \times \frac{\sigma}{\sqrt{N}}\right] = \left[0,85 - Z_{(0,025)} \times \frac{0,038}{\sqrt{140}}, 0,85 + Z_{(0,025)} \times \frac{0,038}{\sqrt{140}}\right] =$$

$$= \left[0,85 - 1,96 \times \frac{0,20}{\sqrt{140}}, 0,85 + 1,96 \times \frac{0,20}{\sqrt{140}}\right] =$$

$$= [0,817; 0,883]$$

10.2 X: "tensão de rotura de 10 cabos do lote" $X \sim N(\mu, \sigma^2)$

O intervalo aberto à direita é dado por:

$$\left[\bar{X} - t_{N-1(\alpha/2)} \times \frac{s}{\sqrt{N}}, +\infty\right[\quad (1) \qquad \alpha = 0{,}05 \text{ (intervalo aberto).}$$

$$(1) \quad \left[4537 - t_{9(0,05)} \times \frac{112}{\sqrt{10}}, +\infty\right[= \left[4537 - 1{,}833 \times \frac{112}{\sqrt{10}}, +\infty\right[= \left[4472, +\infty\right[$$

10.3 O intervalo é dado por $\left[\frac{(N-1)\times s^2}{\chi^2_{N-1(\alpha/2)}}, \frac{(N-1)\times s^2}{\chi^2_{N-1}(1-\frac{\alpha}{2})}\right] = \quad \frac{\alpha}{2} = \frac{1-0{,}99}{2} = 0{,}005$

$$(1) \quad \left[\frac{24\times 0{,}3^2}{\chi^2_{24}(0{,}005)}, \frac{24\times 0{,}3^2}{\chi^2_{24}(0{,}995)}\right] = \left[\frac{24\times 0{,}3^2}{45{,}56}, \frac{24\times 0{,}3^2}{9{,}89}\right] = [0{,}047 ; 0{,}218] \leftarrow 99\% \text{ confiança}$$

$$(2) \quad \left[\frac{24\times 0{,}3^2}{\chi^2_{24}(0{,}025)}, \frac{24\times 0{,}3^2}{\chi^2_{24}(0{,}975)}\right] = \left[\frac{24\times 0{,}3^2}{39{,}36}, \frac{24\times 0{,}3^2}{12{,}4}\right] = [0{,}055 ; 0{,}174] \leftarrow 95\% \text{ confiança}$$

$$(3) \quad \left[\frac{24\times 0{,}3^2}{\chi^2_{24}(0{,}05)}, \frac{24\times 0{,}3^2}{\chi^2_{24}(0{,}95)}\right] = \left[\frac{24\times 0{,}3^2}{36{,}42}, \frac{24\times 0{,}3^2}{13{,}85}\right] = [0{,}059 ; 0{,}156] \leftarrow 90\% \text{ confiança}$$

10.4 Y: "número de operários que de entre os 300 seleccionados aleatoriamente uma semana após a emissão do mesmo, os conhece suficientemente bem."

$Y \sim H(4000p, 4000(1-p), 300) \quad \mu_Y = 300p \quad \sigma_Y^2 = 300p(1-p) \times \frac{4000-300}{4000-1} =$
$= 277{,}6\, p(1-p)$ Como $4000 \geq 10 \times 300$, H pode ser aproximada por $B(300, p)$

Como $300 \geq 20$, esta pode ser aproximada por uma distribuição normal.

$Y \sim N(300p, 300p(1-p)) \qquad \hat{p} = \frac{Y}{300}$, sendo $\hat{p}$ o estimador de p

$\hat{p} \sim N\left(p, \frac{277{,}6}{300^2} p(1-p)\right) \qquad \hat{p} = \frac{75}{300} = 0{,}25$

$$\left[0{,}25 - Z_{(0,005)} \times \sqrt{\frac{0{,}25\times 0{,}75}{300} \times \frac{4000-300}{4000-1}}, 0{,}25 + Z_{(0,005)} \times \sqrt{\frac{0{,}25 \times 0{,}75}{300} \times \frac{4000-300}{4000-1}}\right] =$$

$$= [0{,}25 - 1{,}96 \times 0{,}024 ; 0{,}25 + 1{,}96 \times 0{,}024] = [0{,}203 ; 0{,}297]$$

10.5 Amostras Normais e de pequena dimensão. Admitindo que têm variâncias diferentes,

$$GL = \frac{\left(\frac{2{,}74^2}{17} + \frac{3{,}91^2}{13}\right)^2}{\frac{\left(\frac{2{,}74^2}{17}\right)^2}{16} + \frac{\left(\frac{3{,}91^2}{13}\right)^2}{12}} = \frac{2{,}617}{0{,}127} = 20{,}6 \simeq 20$$

95%: $\left[(48,2-41,5)-t_{20(0,025)}\times\sqrt{\frac{2,74^2}{17}+\frac{3,91^2}{13}}, (48,2-41,5)+t_{20(0,025)}\times\sqrt{\frac{2,74^2}{17}+\frac{3,91^2}{13}}\right]=$

$=\left[6,7-2,086\times1,272, 6,7+2,086\times1,272\right]=\left[4,05;9,35\right]$

90% $\left[(48,2-41,5)-t_{20(0,05)}\times\sqrt{\frac{2,74^2}{17}+\frac{3,91^2}{13}}, (48,2-41,5+t_{20(0,05)}\times\sqrt{\frac{2,74^2}{17}+\frac{3,91^2}{13}}\right]=$

$=\left[6,7-1,725\times1,272; 6,7+1,725\times1,272\right]=\left[4,51;8,89\right]$

Assumindo que as populações têm variâncias iguais:

$GL=17+13-2=28 \qquad s=\sqrt{\frac{16\times2,74^2+12\times3,91^2}{28}}=3,3$

95%: $\left[(48,2-41,5)-t_{28(0,025)}\times3,3\times\sqrt{\frac{1}{17}+\frac{1}{13}}, (48,2-41,5)+t_{28(0,025)}\times3,3\times\sqrt{\frac{1}{17}+\frac{1}{13}}\right]=$

$=\left[6,7-2,048\times3,3\times0,37, 6,7+2,048\times3,3\times0,37\right]=\left[4,20;9,20\right]$

90%: $\left[(48,2-41,5)-t_{28(0,05)}\times3,3\times\sqrt{\frac{1}{17}+\frac{1}{13}}, (48,2-41,5)+t_{28(0,05)}\times3,3\times\sqrt{\frac{1}{17}+\frac{1}{13}}\right]=$

$=\left[6,7-1,701\times3,3\times0,37, 6,7+1,701\times3,3\times0,37\right]=\left[4,62;8,78\right]$

10.6 Q: "número de quadros superiores, de entre os entrevistados, que leram a última edição do semanário".

G: "número de gestores de empresas, de entre os entrevistados, que leram a última edição do semanário."

$Q \sim B(350, p_Q) \qquad G \sim B(325, p_G)$

Como $N \geq 20$ e $N\times p > 7$ para ambos os casos, ambas as distribuições binomiais podem ser aproximadas por distribuições normais.

$Q \sim N(350\times p_Q, 350\times p_Q\times(1-p_Q)) \qquad G \sim N(325\times p_G, 325\times p_G\times(1-p_G))$

$\hat{p}_Q=\frac{Q}{N_Q}=\frac{Q}{350} \qquad \hat{p}_G=\frac{G}{N_G}=\frac{G}{325} \qquad \hat{p}_Q \sim N\left(p_Q, \frac{p_Q\times(1-p_Q)}{350}\right) \quad \hat{p}_G \sim N\left(\hat{p}_G, \frac{p_G\times(1-p_G)}{325}\right)$

$p_Q=\frac{105}{350}=0,3 \qquad p_G=\frac{130}{325}=0,4$

$=\left[0,4-0,3-\overbrace{Z_{(0,005)}}^{2,575}\times\sqrt{\frac{0,4\times0,6}{350}+\frac{0,3\times0,7}{325}}, 0,4-0,3+\overbrace{Z_{(0,005)}}^{2,575}\times\sqrt{\frac{0,4\times0,6}{350}+\frac{0,3\times0,7}{325}}\right]=$

$=\left[0,006;0,194\right]$

10.7 X: "tensão da rotura de material usado num determinado tipo de tubo"

$\bar{X}$: "média amostral de X". N: "dimensão da amostra"

Admite-se que a dimensão da amostra é suficientemente grande para que, de acordo com o teorema do limite central e independentemente da distribuição da variável X, $\bar{X}$ seja uma distribuição normal.

Intervalo de confiança: $\left[\bar{X} - Z_{(0,005)} \times \frac{\sigma_x}{\sqrt{N}}, \bar{X} + Z_{(0,005)} \times \frac{\sigma_x}{\sqrt{N}}\right]$

Amplitude: $2 Z_{(0,005)} \times \frac{\sigma_x}{\sqrt{N}} = 2 \times 2,575 \times \frac{70}{\sqrt{N}}$

$2 \times 2,575 \times \frac{70}{\sqrt{N}} \leq 60 \Leftrightarrow \frac{70}{\sqrt{N}} \leq \frac{60}{2 \times 2,575} \Leftrightarrow N \geq \left(\frac{70 \times 2 \times 2,575}{60}\right)^2 \Leftrightarrow N \geq 36,1$

Ou seja, $N \geq 37$. R: No mínimo, 37 ensaios.

10.8 X: "número de alunos que, de entre os N inquiridos, com interesse no horário de manhã"

$X \sim H(3800p, 3800(1-p), N)$ Intervalo de confiança:

$$\left[\hat{p} - Z_{(0,025)} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{N} \times \frac{M-N}{M-1}}, \hat{p} + Z_{(0,025)} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{N} \times \frac{M-N}{M-1}}\right]$$

Como não se sabe $\hat{p}$, utiliza-se o valor de p que maximiza a amplitude:

$$\frac{d(\hat{p} \times (1-\hat{p}))}{d\hat{p}} = 1 - 2\hat{p} = 0 \Leftrightarrow \hat{p} = 0,5$$

Amplitude: $2 \overset{1,96}{Z_{(0,025)}} \times \sqrt{\frac{0,5 \times (1-0,5)}{N} \times \frac{3800-N}{3799}} \leq 0,1 \Leftrightarrow$

$\Leftrightarrow 2 \times 1,96 \times \sqrt{\frac{0,5 \times (1-0,5)}{N} \times \frac{3800-N}{3799}} \leq 0,1 \Leftrightarrow \sqrt{\frac{950 - 0,25N}{3799N}} \leq 0,0255 \Leftrightarrow$

$\Leftrightarrow 950 - 0,25N \leq 2,47N \Leftrightarrow 2,72N \geq 950 \Leftrightarrow N \geq 349,3$

Ou seja, $N \geq 350$. R: A dimensão mínima é 350.