INTERVALOS DE CONFIANÇA

INTERVALOS DE CONFIANÇA

- Estabelecer um intervalo de confiança para o parâmetro θ .

$$P(\hat{\theta}_I < \theta < \hat{\theta}_S) = 1 - \alpha$$

- Determinar os dois limites que definem o intervalo,

$$\hat{\theta}_I < \theta < \hat{\theta}_S$$

limites que dependem da distribuição amostral de θ e são, respectivamente, os limites inferior e superior do intervalo.

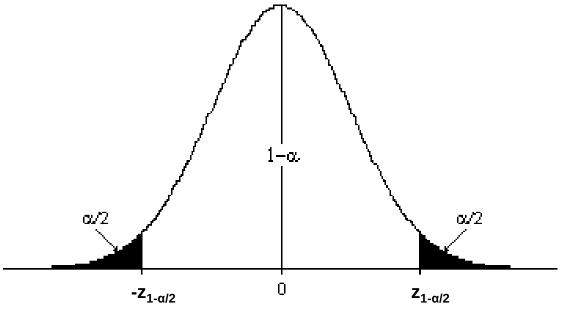
INTERVALO DE CONFIANÇA

- A média de uma amostra possui uma distribuição, σ^2 conhecido.

$$\mu_{\bar{x}} = \mu \qquad \sigma_{\bar{x}}^2 = \sigma^2 / n \qquad Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

$$P\left(-z_{1-\alpha/2} < Z < z_{1-\alpha/2}\right) = 1 - \alpha$$

INTERVALO DE CONFIANÇA



INTERVALO DE CONFIANÇA

$$P\left(-z_{1-\alpha/2} < \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} < z_{1-\alpha/2}\right) = 1-\alpha$$

$$P\left(\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

INTERVALO DE CONFIANÇA: MÉDIA

- σ^2 conhecido

$$\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\overline{x} \pm z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \approx \overline{x} \pm z_{1-\alpha/2}\frac{s}{\sqrt{n}}$$

EXEMPLO 1

- Suponha que era conhecido que a média e o desvio padrão das alturas dos rapazes com 20 anos era $\mu = 170\ cm, \sigma = 10\ cm$
- Considere que foram recolhidas 5 amostras de 25 rapazes, tendo sido observadas as seguintes médias

Amostra	**1**	**2**	**3**	**4**	**5**
Média (*cm*)	172	168	171	165	172

SOLUÇÃO 1

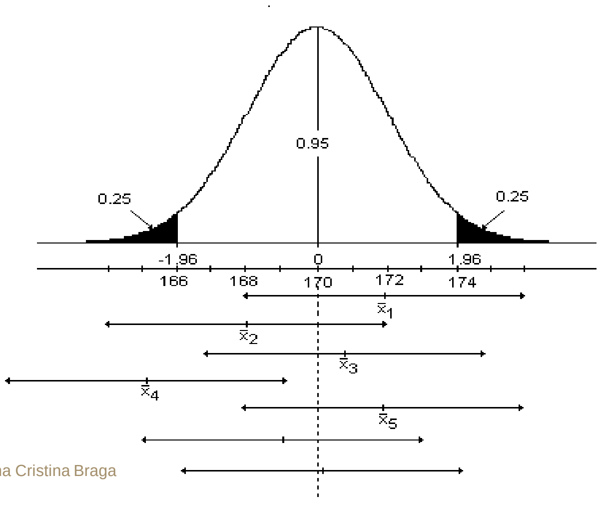
- Intervalo de Confiança de 95%

$$\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\bar{x} - 1.96\frac{10}{\sqrt{25}} < \mu < \bar{x} + 1.96\frac{10}{\sqrt{25}}$$

$$\bar{x} \pm 4cm$$

SOLUÇÃO 1



EXEMPLO 2

- O peso ao nascer é uma das variáveis mais importantes na avaliação do bem-estar de um recém nascido.
- Suponha que o valor do desvio padrão para os bebés de sexo masculino é 562 gramas. Num determinado centro de saúde, uma amostra de 19 recém nascidos apresentou uma média 3222 gramas.
- Construa um intervalo de confiança de 95% para média do peso dos bebés.

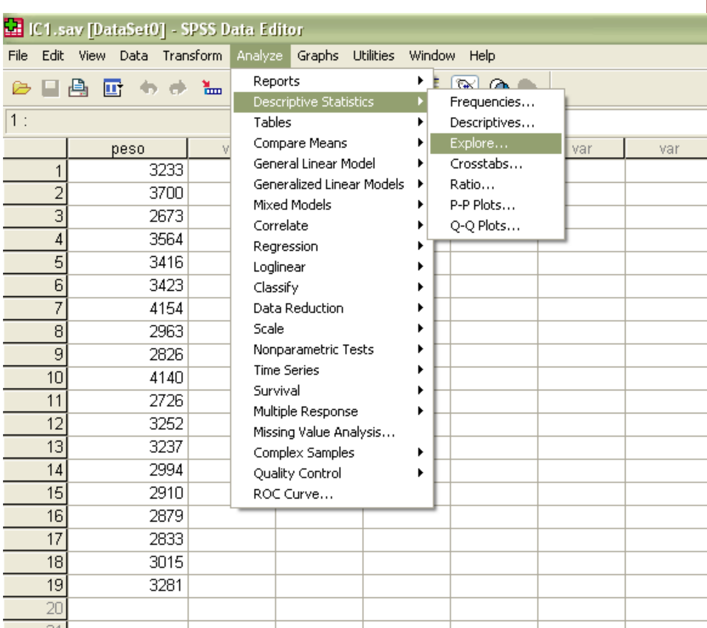
SOLUÇÃO 2

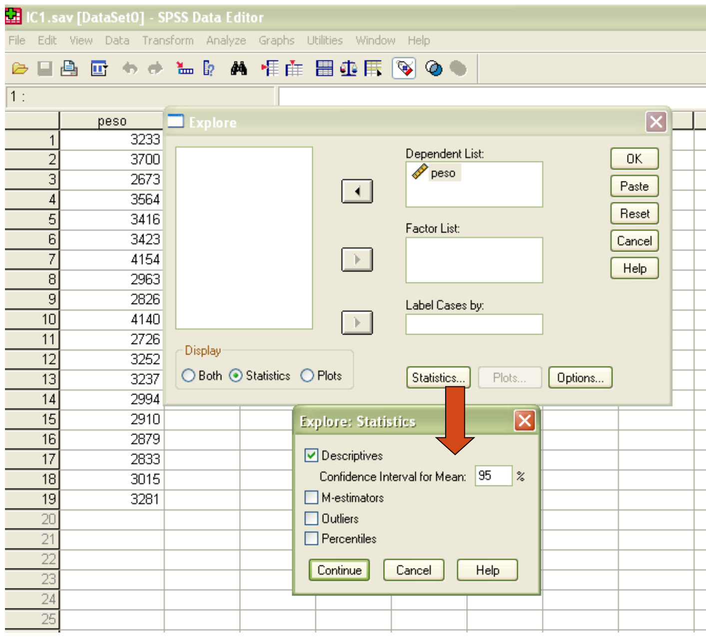
$$\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

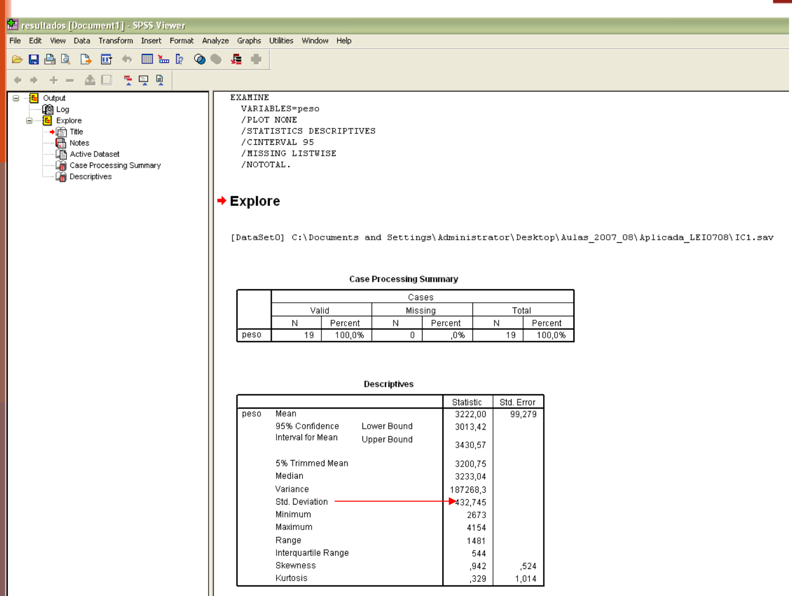
$$3222 - 1.96\frac{562}{\sqrt{19}} < \mu < 3222 + 1.96\frac{562}{\sqrt{19}}$$

$$3222 \pm 253g$$

$$2969 < \mu < 3475$$








```
EXAMINE
  VARIABLES=peso
  /PLOT NONE
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Explore

[DataSet0] C:\Documents and Settings\Administrator\Desktop\Aulas_2007_08\Aplicada_LEI0708\IC1.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
peso	19	100,0%	0	,0%	19	100,0%

Descriptives

			Statistic	Std. Error
peso	Mean		3222,00	99,279
	95% Confidence Interval for Mean	Lower Bound	3013,42	
		Upper Bound	3430,57	
	5% Trimmed Mean		3200,75	
	Median		3233,04	
	Variance		187268,3	
	Std. Deviation		432,745	
	Minimum		2673	
	Maximum		4154	
	Range		1481	
	Interquartile Range		544	
	Skewness		,942	,524
	Kurtosis		,329	1,014

INTERVALO DE CONFIANÇA: MÉDIA

- σ^2 desconhecido, n<30

$$T = \frac{\overline{x} - \mu}{s/\sqrt{n}} \qquad P\left(-t_{\alpha/2,n-1} < T < t_{\alpha/2,n-1}\right) = 1 - \alpha$$

$$P\left(-t_{\alpha/2,n-1} < \frac{\overline{x} - \mu}{s/n} < t_{\alpha/2,n-1}\right) = 1 - \alpha$$

$$P\left(\overline{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right) = 1 - \alpha$$

INTERVALO DE CONFIANÇA: MÉDIA

- σ^2 desconhecido, n<30

$$\overline{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

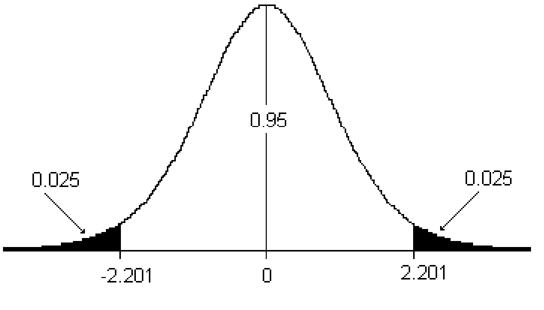
$$\overline{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

EXEMPLO 3

- Numa universidade, uma amostra de 12 estudantes foi seleccionada.
- O comprimento médio da mão encontrado foi de 19.92 cm com um desvio padrão de 0.17cm.
- Construa um intervalo de confiança de 95% para o verdadeiro valor do comprimento médio.

SOLUÇÃO 3

- *t*-Student, com 11 graus de liberdade



SOLUÇÃO 3

$$19.92 - 2.201\frac{0.17}{\sqrt{12}} < \mu < 19.92 + 2.201\frac{0.17}{\sqrt{12}}$$

$$19.92 \pm 0.108$$

$$19.812 < \mu < 20.028$$

INTERVALO DE CONFIANÇA DIFERENÇA DE MÉDIAS

- $\overline{x}_1$, $\overline{x}_2$ médias de amostras aleatórias independentes, de dimensão n_1, n_2
- Populações normais com médias μ_1 e μ_2 e variância comum desconhecida σ^2

$$T = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

INTERVALO DE CONFIANÇA DIFERENÇA DE MÉDIAS

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2, n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2$$
$$< (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2, n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2, n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

EXEMPLO 5

- Pretende-se testar duas formulações alimentares no crescimento de frangos de aviário. Os frangos, distribuídos por dois pavilhões A e B, foram alimentados durante cinco semanas com a respetiva ração. No fim do período de crescimento, foram selecionadas duas amostras.

Grupo	n	Média (g)	Desvio Padrão (g)
Pav. A	16	1623,7500	192,7131
Pav. B	10	1588,0000	167,1194

SOLUÇÃO 5

- *t*-Student

$$s_p^2 = \frac{(16-1)(192.7131)^2 + (10-1)(167.1194)^2}{16+10-2}$$

$$s_p^2 = 33684.7970$$

$$t_{0.025,24} = 2.06$$

$$(1623.75 - 1588.00) \pm (2.06)(183.5342)\sqrt{\frac{1}{16} + \frac{1}{10}}$$

$$35.75 \pm 152.4091$$

$$-116.6591 < \mu_1 - \mu_2 < 188.1591$$

INTERVALO DE CONFIANÇA DIFERENÇA DE MÉDIAS

- $\bar{x}_1$, $\bar{x}_2$ médias de amostras aleatórias independentes
- Populações normais com médias μ_1 e μ_2 e variâncias desconhecidas e diferentes

INTERVALO DE CONFIANÇA DIFERENÇA DE MÉDIAS

$$n_1 = n_2 = n$$

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2,2(n-1)} \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}}$$

$$n_1 + n_2 - 2 = 2(n-1)$$

$$n_1 \neq n_2$$

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2,\nu} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

INTERVALO DE CONFIANÇA DIFERENÇA DE MÉDIAS

- Amostras emparelhadas

$$\mu_d = \left(\mu_1 - \mu_2\right)$$

$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) < \mu_d < \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$\bar{d} \pm t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

EXEMPLO 6

- Uma amostra de dez trabalhadores de uma fábrica onde existe a manipulação de dioxinas foi seleccionada aleatoriamente.
- Nestes trabalhadores foi determinada a concentração (em ppm, partes por milhão) de dioxinas no plasma e no tecido gordo.
- Construa um intervalo de confiança para a diferença entre as concentrações de dioxina no plasma e no tecido gordo.

EXEMPLO 6

Trabalhador	Plasma	Tecido Gordo
1	2.5	4.9
2	3.5	6.9
3	1.8	4.2
4	4.7	4.4
5	7.2	7.7
6	4.1	2.5
7	3.0	5.5
8	3.3	2.9
9	3.1	5.9
10	2.5	2.3

SOLUÇÃO 5

- *t*-Student

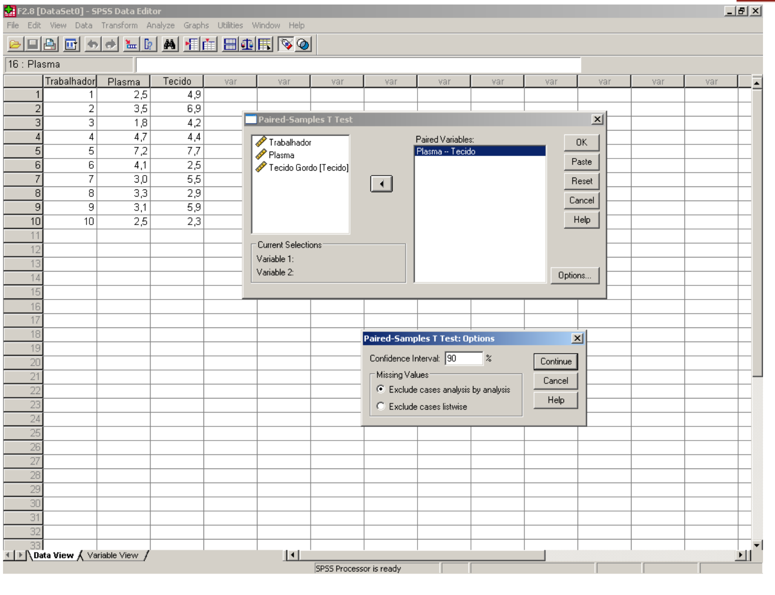
$$\bar{d} = -1.1500 \qquad t_{0.025,9} = 2.262$$
$$s_d = 1.7335$$

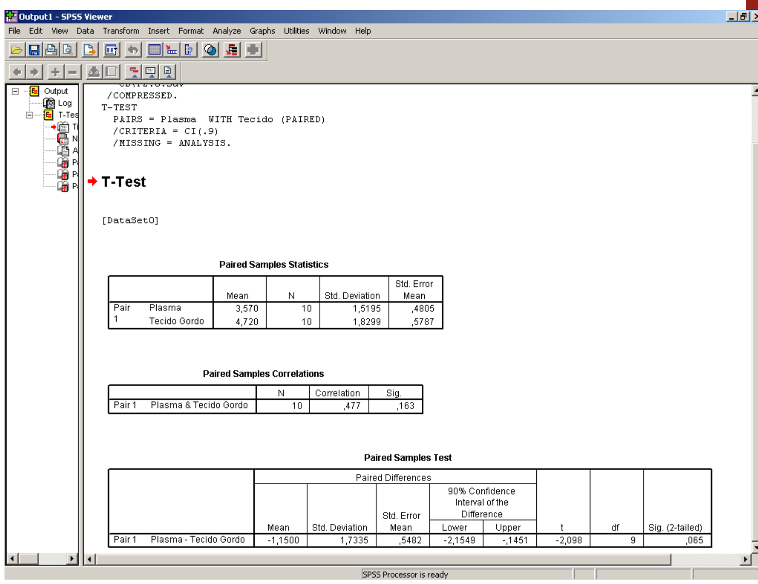
$$-1.1500 - 2.262\frac{1.7335}{\sqrt{10}} < \mu_d < -1.1500 + 2.262\frac{1.7335}{\sqrt{10}}$$

$$-1.1500 \pm 2.262\frac{1.7335}{\sqrt{10}}$$

$$-1.1500 \pm 1.2400$$

$$-2.3900 < \mu_d < 0.0900$$





Output1 - SPSS Viewer

```
 /COMPRESSED.
T-TEST
  PAIRS = Plasma  WITH Tecido (PAIRED)
  /CRITERIA = CI(.9)
  /MISSING = ANALYSIS.
```

T-Test

[DataSet0]

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Plasma	3,570	10	1,5195	,4805
	Tecido Gordo	4,720	10	1,8299	,5787

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Plasma & Tecido Gordo	10	,477	,163

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	90% Confidence Interval of the Difference: Lower	Upper			
Pair 1	Plasma - Tecido Gordo	-1,1500	1,7335	,5482	-2,1549	-,1451	-2,098	9	,065

SPSS Processor is ready

INTERVALO DE CONFIANÇA PROPORÇÃO

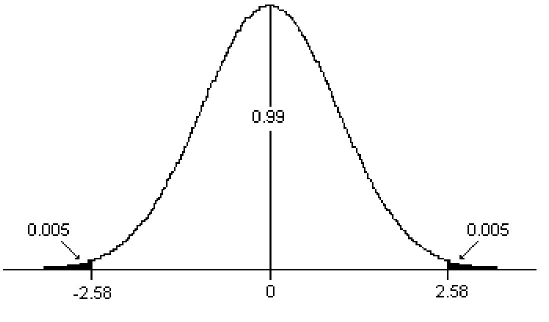
$$z = \frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \qquad P\left(-z_{1-\alpha/2} < \frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} < z_{1-\alpha/2}\right) = 1-\alpha$$

$$P\left(p - z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} < \pi < p + z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right)$$

EXEMPLO 7

- Para determinar a incidência de uma determinada doença genética no Norte de Portugal, foi recolhida uma amostra de gotas de sangue de 500 bebés, nascidos no ano de 1994.
- As análises permitiram detetar 37 bebés portadores da doença.
- Estime um intervalo de confiança de 99% para a proporção de portadores da doença.

SOLUÇÃO 7



SOLUÇÃO 7

$$p = \frac{37}{500} = 0.074$$

$$0.074 \pm 2.58\sqrt{\frac{0.074(1-0.074)}{500}}$$

$$0.074 \pm 0.030$$

$$0.044 < \pi < 0.104$$

Estimativa para a diferença de proporções $\pi_1 - \pi_2$

$$\boxed{Z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sigma_{p_1 - p_2}} \sim N(0,1)}$$

$$\sigma_{p_1 - p_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$$p_1 = \frac{x_1}{n_1} \text{ e } p_2 = \frac{x_2}{n_2}$$

$$\boxed{(p_1 - p_2) - z_{(1-\alpha/2)}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} < \pi_1 - \pi_2 < (p_1 - p_2) + z_{(1-\alpha/2)}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Exemplo:

- Quando um sinal de limite de velocidade de 50km/h foi colocado numa estrada, numa amostra de 100 veículos, 49 violaram o limite de velocidade. Quando o limite foi aumentado para 60 km/h, duma amostra de 100 veículos, 19 ultrapassaram o novo limite. Encontre um intervalo de confiança de 99% para $\pi_1 - \pi_2$ e interprete o seu resultado.

$$p_1 = \frac{49}{100} = 0{,}49 \text{ e } p_2 = \frac{19}{100} = 0{,}19$$

$$\sigma_{p_1 - p_2} = \sqrt{\frac{0{,}49(1-0{,}49)}{100} + \frac{0{,}19(1-0{,}19)}{100}} = 0{,}0635$$

$$z_{0.995} \overset{tab.5}{=} 2{,}575$$

$$0{,}30 - 0{,}164 < \pi_1 - \pi_2 < 0{,}30 + 0{,}164$$

$$0{,}136 < \pi_1 - \pi_2 < 0{,}464$$

Estimativa para o desvio padrão, σ

1. **Grandes amostras** $n \geq 100$

Se a variável em estudo apresenta uma distribuição Normal então o intervalo de confiança será:

$$\boxed{s_n - z_{1-\alpha/2} \cdot \frac{0{,}71.s_n}{\sqrt{n}} < \sigma < s_n + z_{1-\alpha/2} \cdot \frac{0{,}71.s_n}{\sqrt{n}}}$$

2. **Pequenas amostras** $n < 100$

Se a variável em estudo apresenta uma distribuição Normal então o intervalo de confiança será:

$$\boxed{\frac{(n-1)s_{n-1}^2}{\chi^2_{n-1,\alpha/2}} < \sigma^2 < \frac{(n-1)s_{n-1}^2}{\chi^2_{n-1,1-\alpha/2}}}$$

Para o desvio padrão poder-se-á escrever:

$$\boxed{\sqrt{\frac{(n-1)s_{n-1}^2}{\chi^2_{n-1,\alpha/2}}} < \sigma < \sqrt{\frac{(n-1)s_{n-1}^2}{\chi^2_{n-1,1-\alpha/2}}}}$$

Estimativa para o quociente das variâncias σ_1^2 / σ_2^2

Considerando duas populações, de onde são retiradas as amostras aleatórias, independentes e com distribuições aproximadamente normais, o intervalo de confiança (1-α)100% para o quociente das variâncias será dado por:

$$\boxed{\frac{s_1^2}{s_2^2}\frac{1}{F_{\alpha/2,n_1-1,n_2-1}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}F_{\alpha/2,n_2-1,n_1-1}}$$

onde $F_{\alpha/2,n_1-1,n_2-1}$ é o valor que localiza uma área de $\alpha/2$ na cauda superior da distribuição F com n_1-1 no numerador e n_2-1 graus de liberdade no denominador e $F_{\alpha/2,n_2-1,n_1-1}$ é o valor que localiza um área de $\alpha/2$ na cauda superior da distribuição F com n_2-1 no numerador e n_1-1 graus de liberdade no denominador.