Probabilidade e Estatística, 2011/2

CCT - UDESC

Prof. Fernando Deeke Sasse

Intervalos de Confiança

Problemas Resolvidos em Maple

1. Os dados relativos a cargas de falha sobre amostras de um tipo de aço fornecem os seguintes resultados (megapascal):
12.4 , 16.5 , 12.7 , 11.9 , 11.4 , 11.4 , 14.1 , 17.6 , 16.7 , 15.8, 19.5 , 8.8 , 13.6 , 11.9 , 11.4.
.Determine o intervalo de confiança de 95% sobre a média amostral.

Solução
A média amostral é

```
> restart
> with(Statistics) :
> L := [ 12.4 , 16.5 , 12.7 , 11.9 , 11.4 , 11.4 , 14.1 , 17.6 , 16.7 , 15.8, 19.5 , 8.8 , 13.6 ,
    11.9 , 11.4 ]
```

$$L := [12.4, 16.5, 12.7, 11.9, 11.4, 11.4, 14.1, 17.6, 16.7, 15.8, 19.5, 8.8, 13.6, 11.9, 11.4] \quad \textbf{(1.1)}$$

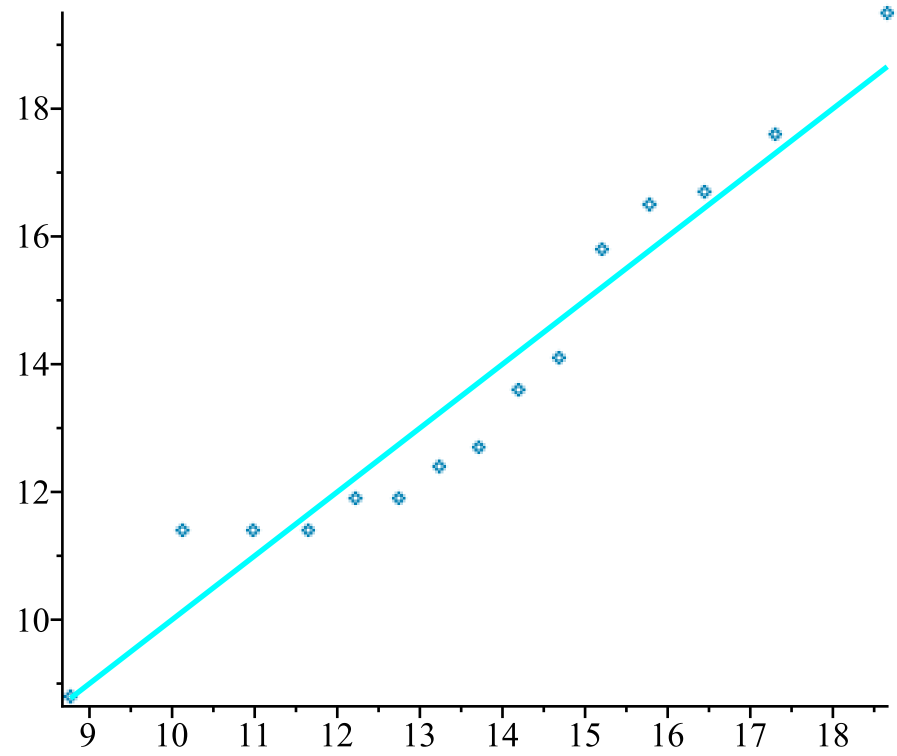
```
> XM := Mean(L)
```

$$XM := 13.71333333 \quad \textbf{(1.2)}$$

```
> ;
```

Um plot normal mostra que a distribuição é aproximadamente normal:

```
> NormalPlot(L, thickness = 2)
```



O desvio padrão da população é desconhecido, de modo que devemos estimá-lo:

> $n := nops(L)$

$$n := 15 \tag{1.3}$$

> $s := \sqrt{\frac{sum((L[i]-XM)^2, i=1..n)}{n-1}}$

$$s := 2.919360660 \tag{1.4}$$

> ;

ou

> $StandardDeviation(L)$

$$2.919360661 \tag{1.5}$$

> ;

Como a amostra é pequena, podemos utilizar a distribuição t para construir o intervalo de confiança. Ou seja, supomos que a variável

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

tem distribuição t com n - 1 graus de liberdade. A distribuição t é dada por

$$f(x) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\,\Gamma(k/2)} \cdot \frac{1}{[(x^2/k)+1]^{(k+1)/2}} \quad -\infty < x < \infty$$

ou seja,

> $f := (x,k) \rightarrow \frac{\Gamma\left(\frac{(k+1)}{2}\right)}{\sqrt{\pi \cdot k}\,\Gamma\left(\frac{k}{2}\right)} \cdot \frac{1}{\left(\left(\frac{x^2}{k}\right)+1\right)^{\frac{(k+1)}{2}}}$

$$f := (x,k) \rightarrow \frac{\Gamma\left(\frac{1}{2}\,k+\frac{1}{2}\right)}{\sqrt{\pi\, k}\;\Gamma\left(\frac{1}{2}\,k\right)\left(\frac{x^2}{k}+1\right)^{\frac{1}{2}\,k+\frac{1}{2}}} \qquad \textbf{(1.6)}$$

> ;

Aqui

> $k := n - 1$

$$k := 14 \qquad \textbf{(1.7)}$$

Os limites de confiança de 95% sobre μ são

$$\bar{x} - t_{\alpha/2,n-1}s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1}s/\sqrt{n}$$

ou seja,

> $\alpha := 0.05$

$$\alpha := 0.05 \qquad \textbf{(1.8)}$$

Devemos agora determinar o valor de t que corresponde à probabilidade acumulada α/2=0.025:

```
> CD := 0 : t0 := -3 :
> while CD < 0.025 do
    CD := evalf(int(f(t, k), t = -∞..t0));
    t0 := t0 + 0.001;
  od:
> t0
```

$$-2.143 \qquad \textbf{(1.9)}$$

De fato,

```
> X := RandomVariable(StudentT(k)) :
> CDF(X, t0)
```

$$0.0250832674255764311 \qquad \textbf{(1.10)}$$

> $x0 := -evalf\left(\frac{t0 \cdot s}{\sqrt{n}}\right)$

$$x0 := 1.615341285 \qquad \textbf{(1.11)}$$

Os limites de confiança de 95% para a média são então

> XM

$$13.71333333 \qquad \textbf{(1.12)}$$

> $L1 := XM + x0;\ L2 := XM - x0$

$$L1 := 15.32867462$$
$$L2 := 12.09799204 \qquad \textbf{(1.13)}$$

> ;

2 Um teste de impacto foi realizado em 20 amostras de tubo PVC. A média amostral de força de ruptura é $xm = 1.25$ e o desvio padrão amostral é $s=0.25$. Encontre um intervalo de confiança limitado inferiormente de 99% na média da força de ruptura. Suponha que a distribuição é normal.

Solução

Lembramos a definição

A $100(1-\alpha)\%$ **upper-confidence bound** for μ is

$$\mu \leq u = \bar{x} + z_\alpha \sigma/\sqrt{n}$$

and a $100(1-\alpha)\%$ **lower-confidence bound** for μ is

$$\bar{x} - z_\alpha \sigma/\sqrt{n} = l \leq \mu$$

Como a amostra é pequena devemos utilizar a distribuição t para construir o intervalo de confiança.

```
> restart
> with(Statistics) :
> CD := 0 : t0 := -3 : k := 19 :
> X := RandomVariable(StudentT(k)) :
> while CD < 0.010 do
    CD := evalf(CDF(X, t0));
    t0 := t0 + 0.010;
  od:
> t0
```

$$-2.520 \qquad (2.1)$$

```
> CDF(X, t0)
```

$$0.01042087338023369245 \qquad (2.2)$$

$$> L := -\frac{t0 \cdot 0.25}{\text{sqrt}(20.)}$$

$$L := 0.1408722826 \qquad (2.3)$$

```
> 1.25 - L
```

$$1.109127717 \qquad (2.4)$$

```
>
```

O intervalo de confiança de 99% na média é, portanto, $[1.109, +\infty]$.

3 Um teste de impacto foi realizado em 100 amostras de tubo PVC. A média amostral de força de ruptura é $xm = 1.35$ e o desvio padrão amostral é $s=0.2$. Encontre um intervalo de confiança de 95% na média da força de ruptura. Suponha que a distribuição é normal.

Solução

Como a amostra é suficientemente grande podemos utilizar a distribuição normal para construir o intervalo de confiança.

```
> restart
```

```
> with(Statistics) :
> σ := 0.2 : μ := 1.35 : n := 100 :
> X1 := RandomVariable(Normal(mu, 0.2/sqrt(n))) :
> Quantile(X1, 0.025)
```

$$1.310800720 \quad (3.1)$$

```
> Quantile(X1, 1 - 0.025)
```

$$1.389199280 \quad (3.2)$$

Cálculo usando padronização:

```
> CD := 0 : t0 := -3 :
>  X := RandomVariable(Normal(0, 1)) :
> while CD < 0.025 do
     CD := evalf(CDF(X, t0));
      t0 := t0 + 0.00015;
    od:
> t0
```

$$-1.95975 \quad (3.3)$$

```
> 2· CDF(X, t0)
```

$$0.05002501792 \quad (3.4)$$

$$> X1 := -\frac{t0 \cdot \sigma}{\mathrm{sqrt}(n)}$$

$$X1 := 0.03919500000 \quad (3.5)$$

```
> L1 := μ - X1; L2 := μ + X1
```

$$L1 := 1.310805000$$
$$L2 := 1.389195000 \quad (3.6)$$

Portanto, o intervalo de confiança de 95% é [1.31, 1.39].

4 Um fabricante produz anéis de pistão com diâmetro normalmente distribuído com $\sigma = 0.01$ mm. Uma amostra aleatória de 40 anéis tem diâmetro médio $xm = 74.036$ mm.

(a) Construa um intervalo de 99% de confiança na média do diâmetro do pistão.

(b) Construa um intervalo de 99% de confiança inferiormente limitado na média do diâmetro do pistão

Solução

Como o número de amostras é suficientemente grande, podemos utilizar a distribuição normal para construir o intervalo de confiança.

```
> restart
> with(Statistics) :
>
> σ := 0.01 : μ := 74.036 :
> CD := 0 : t0 := -3 :
>  X := RandomVariable(Normal(0, 1)) :
> while CD < 0.005 do
     CD := evalf(CDF(X, t0));
      t0 := t0 + 0.0005;
    od:
> t0
```

$$-2.5750 \quad (4.1)$$

> $2 \cdot CDF(X, t0)$

$$0.01002400866 \quad \textbf{(4.2)}$$

> $X1 := -\frac{t0 \cdot \sigma}{\mathrm{sqrt}(40.)}$

$$X1 := 0.004071432488 \quad \textbf{(4.3)}$$

> $L1 := \mu - X1; L2 := \mu + X1$

$$L1 := 74.03192857$$
$$L2 := 74.04007143 \quad \textbf{(4.4)}$$

Portanto, o intervalo de confiança é [74.0319, 74.0401].

> ;

(b)

> $CD := 0 : T0 := -3 :$

> $X := RandomVariable(Normal(0, 1)) :$

```
> while CD < 0.01 do
    CD := evalf(CDF(X, T0));
     T0 := T0 + 0.001;
  od:
```

> $T0$

$$-2.325 \quad \textbf{(4.5)}$$

> $CDF(X, T0)$

$$0.0100359801002740511 \quad \textbf{(4.6)}$$

> $Y1 := -\frac{T0 \cdot \sigma}{\mathrm{sqrt}(40.)}$

$$Y1 := 0.003676147780 \quad \textbf{(4.7)}$$

> $LI1 := \mu - Y1$

$$LI1 := 74.03232385 \quad \textbf{(4.8)}$$

Portanto, o intervalo de confiança é [74.0323, ∞].

5 Uma máquina produz barras de metal utilizadas em um sistema automobilístico. Uma amostra aleatória de 15 elementos é selecionada e o diâmetro é medido. Os dados resultantes (em mm) são os seguintes: 8.24, 8.25, 8.20, 8.23, 8.24, 8.21, 8.26, 8.26, 8.20, 8.25, 8.23, 8,23, 8.19, 8.28, 8.24.

(a) Verifique a suposição de normalidade para os diâmetros da barras.

(b) Determine um intervalo de confiança de 95% sobre o diâmetro médio das barras.

Solução

(a)

> $restart$

> $with(Statistics) :$

> $L := [8.24, 8.25, 8.20, 8.23, \quad 8.24, 8.21, 8.26, 8.26, \quad 8.20, 8.25, 8.23, 8.23, \quad 8.19, 8.28, 8.24]$

$$L := [8.24, 8.25, 8.20, 8.23, 8.24, 8.21, 8.26, 8.26, 8.20, 8.25, 8.23, 8.23, 8.19, 8.28, 8.24] \quad \textbf{(5.1)}$$

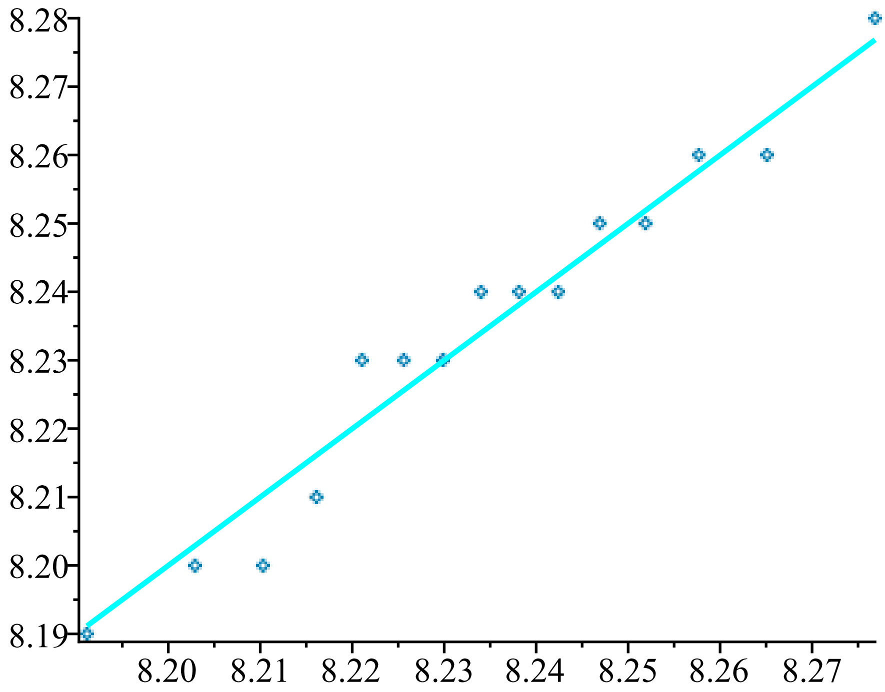
> $n := nops(L)$

$$n := 15 \quad \textbf{(5.2)}$$

>

Um plot normal mostra que a distribuição é aproximadamente normal :

> $NormalPlot(L, thickness = 2)$



(b)

A média é dada por

> $XM := Mean(L)$

$$XM := 8.234000000 \tag{5.3}$$

O desvio padrão da população é desconhecido, de modo que devemos estimá-lo:

> $s := \sqrt{\dfrac{sum\left((L[i]-XM)^2, i=1..n\right)}{n-1}}$

$$s := 0.02529822128 \tag{5.4}$$

> ;

Como, além ter o desvio padrão desconhecido, a amostra é pequena, devemos usar a distribuição t para estimar o intervalo de confiança de média.

> $CD := 0 : t0 := -3 : k := n-1 :$

> $X := RandomVariable(StudentT(k)) :$

> **while** $CD < 0.025$ **do**
$CD := evalf(CDF(X, t0));$
$t0 := t0 + 0.0001;$
od:

> $t0$

$$-2.1446 \tag{5.5}$$

> $X1 := -evalf\left(\dfrac{t0 \cdot s}{\text{sqrt}(n)}\right)$

$$X1 := 0.01400846854 \tag{5.6}$$

> $L1 := XM + X1; L2 := XM - X1$

$$L1 := 8.248008469$$
$$L2 := 8.219991531 \tag{5.7}$$

Portanto, o intervalo de confiança de 95% sobre a média é [8.219991531, 8.248008469]

6 Dados relativos à espessura, em angstroms, de óxido em semicondutores são listados a seguir:

325, 331, 318, 320, 321, 335, 318, 309, 332, 333, 323, 326, 310, 334, 335, 328, 312, 326, 311, 336, 321, 327, 318, 317, 322, 332

.. Determine o intervalo de confiança de 95% sobre a média amostral. Suponha que

Solução

A média amostral é

```
> restart
> with(Statistics) :
> L := [325, 331, 318, 320, 321, 335, 318, 309, 332, 333, 323, 326, 310, 334, 335, 328, 312, 326, 311, 336, 321, 327, 318, 317, 322, 332]
```

$L := [325, 331, 318, 320, 321, 335, 318, 309, 332, 333, 323, 326, 310, 334, 335, 328, 312, 326, 311, 336, 321, 327, 318, 317, 322, 332]$ **(6.1)**

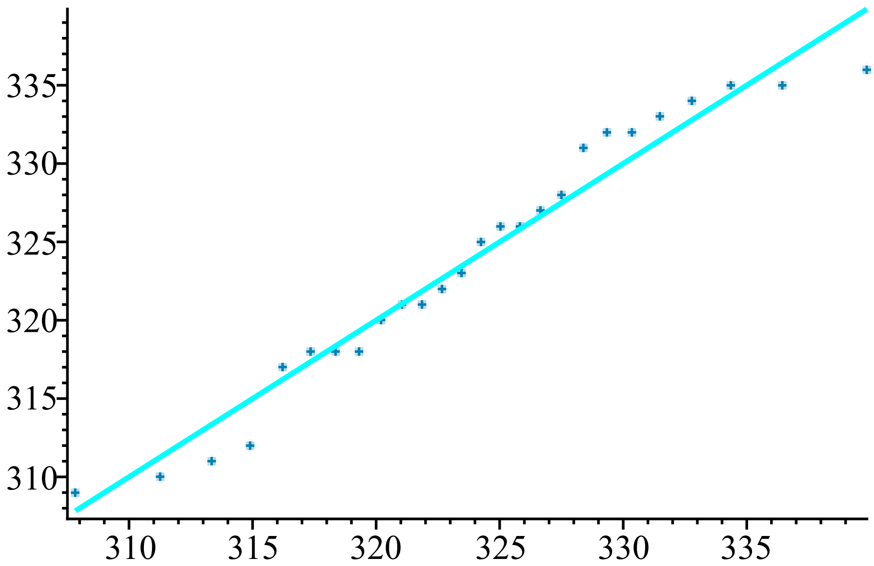
```
> XM := Mean(L)
```

$XM := 323.8461538$ **(6.2)**

```
> ;
```

Um plot normal mostra que a distribuição amostral é, de fato, aproximadamente normal:

```
> NormalPlot(L, thickness = 2)
```



O desvio padrão da população é desconhecido, de modo que devemos estimá-lo:

```
> n := nops(L)
```

$n := 26$ **(6.3)**

> $s := \sqrt{\dfrac{sum\left((L[i] - XM)^2, i = 1 ..n\right)}{n - 1}}$

$s := 8.264102165$ **(6.4)**

```
> ;
```

ou

```
> StandardDeviation(L)
```

8.264102166 **(6.5)**

```
> ;
```

Como a amostra é pequena, com média μ e variância σ^2 desconhecidos, podemos utilizar a distribuição t para construir o intervalo de confiança. A única suposição necessária é que a distribuição seja aproximadamente normal, uma hipótese corroborada acima pelo normal plot. Ou seja, supomos que a variável

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

tem distribuição t com n - 1 graus de liberdade. A distribuição t é dada por

$$f(x) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\Gamma(k/2)} \cdot \frac{1}{[(x^2/k)+1]^{(k+1)/2}} \quad -\infty < x < \infty$$

ou seja,

> $f := (x, k) \rightarrow \frac{\Gamma\left(\frac{(k+1)}{2}\right)}{\sqrt{\pi \cdot k}\,\Gamma\left(\frac{k}{2}\right)} \cdot \frac{1}{\left(\left(\frac{x^2}{k}\right)+1\right)^{\frac{(k+1)}{2}}}$

$$f := (x, k) \rightarrow \frac{\Gamma\left(\frac{1}{2}k + \frac{1}{2}\right)}{\sqrt{\pi k}\,\Gamma\left(\frac{1}{2}k\right)\left(\frac{x^2}{k}+1\right)^{\frac{1}{2}k+\frac{1}{2}}} \qquad \textbf{(6.6)}$$

> ;

Aqui

> $k := n - 1$

$$k := 25 \qquad \textbf{(6.7)}$$

Os limites de confiança de 95% sobre μ são

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

ou seja,

> $\alpha := 0.05$

$$\alpha := 0.05 \qquad \textbf{(6.8)}$$

Devemos agora determinar o valor de t que corresponde à probabilidade acumulada α/2=0.025:

```
> CD := 0 : t0 := -3 :
> while CD < 0.025 do
    CD := evalf(int(f(t, k), t = - ∞..t0));
    t0 := t0 + 0.0001;
  od:
> t0
```

$$-2.0594 \qquad \textbf{(6.9)}$$

De fato,

```
> X := RandomVariable(StudentT(k)) :
> 2· CDF(X, t0)
```

$$0.05001427102 \qquad \textbf{(6.10)}$$

> $x0 := -evalf\left(\frac{t0 \cdot s}{\sqrt{n}}\right)$

$$x0 := 3.337718547 \tag{6.11}$$

Os limites de confiança de 95% para a média são então

> XM

$$323.8461538 \tag{6.12}$$

> $L1 := XM + x0; L2 := XM - x0$

$$L1 := 327.1838723$$
$$L2 := 320.5084353 \tag{6.13}$$

> ;

>