CK0033 - INTRODUÇÃO A COMPUTAÇÃO
Universidade Federal do Ceará
Daniel Magalhães Nunes, 376163
Francilene da Silva Sales, 485249
2020.2

Trabalho1 - Latex - Escrita do arquivo Trabalho3-LAtex.pdf

Inferência Estatística para Uma Única Amostra

Definição: Intervalo de Confiança para a Média com Variância Conhecida

Se $\overline{x}$ for a média de uma amostra aleatória n, de uma população com variância conhecida σ^2, um intervalo com $100\,(1-\alpha)\,\%$ de confiança para μ é dado por:

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

Sendo $z_{\alpha/2}$ o ponto superior com $100\alpha/2\%$ da distribuição normal padrão.

EXEMPLO 8.6

Considere o problema do foguete do exemplo 8.2. Suponha que queiramos achar um intervalo com 95% de confiança para a taxa média de queima. Podemos usar a Eq. 8.35 para construir o intervalo de confiança. Um intervalo de 95% implica $1 - \alpha = 0,95; logo, \alpha = 0,05$, e da Tabela II no Apêndice, $z_{\alpha/2} = z_{0,05/2} = z_{0,025} = 1,96$.

$$\begin{aligned} l &= \overline{x} - z_{\alpha/2}\sigma/\sqrt{n} \\ &= 51,3 - 1,96(2)/\sqrt{25} \\ &= 51,3 - 0,78 \\ &= 50,52 \end{aligned}$$

e o limite superior é

$$\begin{aligned} u &= \overline{x} + z_{\alpha/2}\sigma/\sqrt{n} \\ &= 51,3 + 1,96(2)/\sqrt{25} \\ &= 51,3 + 0,78 \\ &= 52,08 \end{aligned}$$

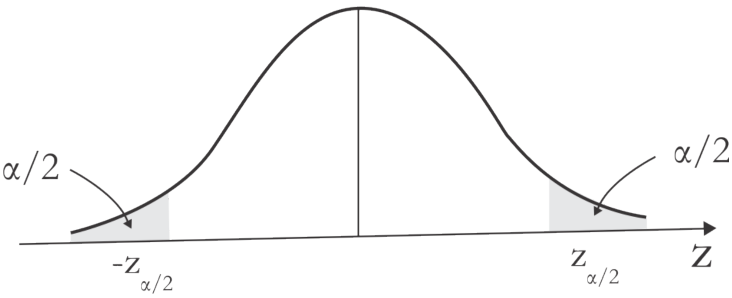

Figura 1: A distribuição Z.

Desse modo, o intervalo bilateral com 95% de confiança é

$$50,52 \leq \mu \leq 52,08$$

Sendo nosso intervalo de valores razoáveis, para a taxa média de queima, com 95% de confiança.

Escolha do tamanho da amostra

Definição

Se $\overline{x}$ for usada como uma estimativa de μ, podemos estar $100(1-\alpha)\%$ confiantes de que o erro $|\overline{x}-1|$ não excederá um valor especificado E quando o tamanho da amostra for

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$$

Note a relação geral entre o tamanho da amostra, o comprimento desejado do intervalo de confiança de $100(1-\alpha)\%$ e o desvio-padrão σ:

- À medida que o comprimento desejado do intervalo $2E$ diminui, o tamanho requerido n da amostra aumenta para um valor fixo de σ e confiança especificada;
- À medida que σ aumenta, o tamanho requerido n da amostra aumenta para comprimento desejado fixo σ $2E$ e confiança especificada;
- À medida que o nível de confiança aumenta, o tamanho requerido n da amostra aumenta para o comprimento desejado fixo $2E$ e desvio-padrão σ.

Intervalo Unilateral de Confiança

É também possível obter intervalos unilaterais de confiança para μ, estabelecendo $l = -\infty$ ou $\mu = \infty$ e trocando $z_{\alpha/2}$ por z_α.

O intervalo superior com $100(1-\alpha)\%$ de confiança para μ é

$$\mu \leq u = \overline{x} + z_\alpha$$

e o intervalo inferior com $100(1-\alpha)\%$ de confiança para μ é

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

8.2.7. Método Geral para reduzir um intervalo de Confiança

É fácil dar um método geral para encontrar um intervalo de confiança para parâmetro desconhecio θ. Faça $X1, X2, ...X,$ ser uma amostra aleatória com n observações. Suponha que possamos encontrar uma estatística $g(X1, X2, ...X, ; \theta)$ com as seguintes propriedades:

1. $g(X1, X2, ..., Xn : \theta)$ depende da amostra e de θ e
2. a distribuição de probabilidade de $g(X1, X2, ..., X : \theta)$ não depende de θ ou de qualquer outro parâmetro desconhecido.

No caso considerado nessa seção, o parâmetro $\theta = \mu$. A variável aleatória $g(X1, X2, ..., X_n : \mu) = (\overline{X} - \mu)/(\sigma/\sqrt{n})$ satisfaz ambas as condições anteriores; ela depende da amostra e de μ e tem uma distribuição normal padrão desde que θ seja conhecido. Agora, tem-se de encontrar as constantes C_L e C_U de modo a

$$P[C_L \leq g(X_1, X_2, ..., X_n; \theta) \leq C_U] = 1 - \alpha$$

Devido à propriedade $2, C_L$ e C_U não depende de θ. Em nosso exemplo, $C_L = z_{\alpha/2}$ e $C_U = z_{\alpha/2}$. Finalmente, você tem de manipular as desigualdades no enuciado de probabilidade, de modo a

$$P[L(X_1, X_2, ..., X_n) \leq \theta \leq U(X_1, X_2, ..., X_n)] = 1 - \alpha$$

Isso fornece $L(X_1, X_2, ..., X_n)$ e $U(X_1, X_2, ..., X_n)$ como os limites inferiores e superiores de confiança, definindo o intervalo de confiança de $100(1-\alpha)\%$ para θ. Em nosso exemplo, encontramos $L(X_1, X_2, ..., X_n) = \overline{X} - z_\alpha \sigma/\sqrt{n}$ e $U(X_1, X_2, ..., X_n) = \overline{X} + z_\alpha \sigma/\sqrt{n}$.

8.3 INFERÊNCIA SOBRE A MÉDIA DE UMA POPULAÇÃO COM VARIÂNCIA DESCONHECIDA

Quando estamos testando hipóteses ou construíndo intervalos de confiaça para a média μ de uma população quando σ^2 for deseconhecida, podemos usar os procedimentos de testes da Seção 8.2. Entretanto, quando a amostra for pequena e σ^2 for desconhecida, teremos de fazer uma suposição sobre a forma de distribuição em estudo de modo a obter um procedimento de teste. Uma suposição razoável, em muitos casos, é que a distribuição sob consideração seja normal.

8.3.1 Testes de Hipóteses para a Média

Suponha que a população de interesse tenha distribuição normal, com a média μ e variância σ^2 desconhecidas. Desejamos testar a hipótese de que μ seja igual a uma constante μ_o. Note que essa situação é similar àquela da Seção 8.2, exeto que agora ambas, μ e σ^2, são desconhecidas. Considere que uma da amostra aleatória de tamanho n, como $X1, X2, ...Xn$. seja disponível e sejam $\overline{X}$ e S^2 a média e a variância amostral, respectivamente. Desejamos testar a hipótese alternativa bilateral.

Trabalho1 - Latex - Escrita do arquivo Trabalho4-LAtex.pdf

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \pm \mu_0$$

Se a variância σ^2 for conhecida, a estatística de teste será a Eq. 8.10:

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

Quando σ^2 for desconhecida, um procedimento lógico será trocar σ na Eq. 9.10 pelo desvio-padrão, S, da amostra. A estatística de teste é agora

$$T_0 = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \tag{8.39}$$

Uma questão lógica é qual o efeito de trocar σ por S na distribuição da estatística T_0? Se n for grande, a resposta a essa questão é "muito pouco"e podemos usar o procedimento de teste baseado na distribuição normal da seção 8.2. Entretanto, n é geralmente pequeno na maioria dos problemas de engenharia e nessa situação uma distribuição diferente tem de ser empregada.

Definição

Faça X_1, X_2, ... , X_n ser uma amostra aleatória para uma distribuição normal, com média μ e variância σ^2 desconhecida. A grandeza

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

tem uma distribuição t, com $n-1$ graus de liberdade.

A função densidade da probabilidade t é

$$f(x) = \frac{\Gamma\left[(k+1)/2\right]}{\sqrt{\pi k}\Gamma\left(k/2\right)}.\frac{1}{\left[(x^2/k)+1\right]^{(k+1)/2}} - \infty < x < \infty \tag{8.40}$$

sendo k o número de graus de liberdade. A média e a variância da distribuição t são iguais a zero e $k/(k-2)$(para $k > 2$), respectivamente.

Várias distribuições t são mostradas na Figura 2.

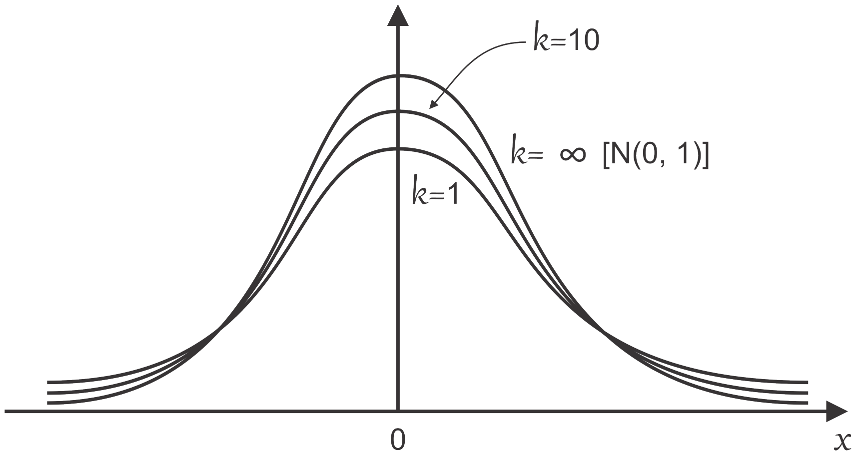

Figura 2: Funções densidade de probabilidade de várias distribuições t.

Agora, pode-se ver, de forma direta, que a distribuição da estatística de teste na Eq.8.39 é t, com $n-1$ graus de liberdade, se a hipótese nula $H_0 : \mu = \mu_0$ for verdadeira. Para testar $H_0 : \mu = \mu_0$, o valor da estatística de teste t_0 na Eq.8.39 é calculado e H_0 é rejeitada se

$$t_0 > t_{\alpha/2,n-1} \tag{8.41a}$$

ou

$$t_0 < -t_{\alpha/2,n-1} \tag{8.41b}$$

em que $t_0 > t_{\alpha/2,n-1}$ e $t_0 < -t_{\alpha/2,n-1}$ são pontos $100\alpha/2\%$ superior e inferior da distribuição t, com $n-1$ graus de liberdade, definidos previamente.

Para a hipótese alternativa unilateral

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0 \tag{8.42}$$

Calculamos a estatística de teste t_0, a partir da Eq. 8.9, e rejeitamos

$$t_0 > t_{\alpha,n-1} \tag{8.43}$$

Para a outra hipótese alternativa unilateral

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0 \tag{8.44}$$

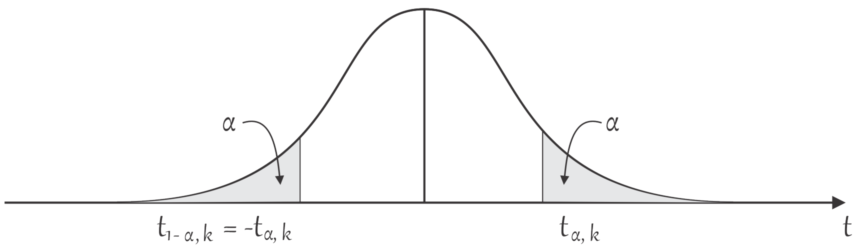

Figura 3: pontos percentuais da distribuições t.

rejeitamos H_0 se

$$t_0 < -t_{\alpha,n-1} \tag{8.45}$$

EXEMPLO 8.8

Um artigo no periódico Materials Engineering (1989, Vol.II, No. 4, pp. 275-281) descreve os resultados de testes de tensão quanto à adesão em 22 corpos de prova de liga U-700. A carga no ponto de falha do corpo de prova é dada a seguir (em MPa):

19,8	18,5	17,6	16,7	15,8
15,4	14,1	13,6	11,9	11,4
11,4	8,8	7,5	15,4	15,4
19,5	14,9	12,7	11,9	11,4
10,1	7,9			

A média amostral é $\overline{x} = 13,71$ e o desvio-padrão é $s = 3,55$. Os dados sugerem que a carga média na falha excede 10MPa? Considere que a carga na falha tenha uma distribuição normal e use $\alpha = 0,05$.

A solução, usando o procedimento de 8 etapas para o teste de hipóteses, é dada a seguir:

1. O parâmetro de interesse é a carga média na falha, μ.

2. H_0:$\mu = 10$

3. H_1:$\mu > 10$. Queremos rejeitar H_0 se a carga média na falha exceder 10MPa.

4. $\alpha = 0,05$.

5. A estatística de teste é
$$t_0 = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

6. Rejeite H_0 se $t_0 > t_{0,05,21} = 3,55$, $\mu = 10$ e $n = 22$, temos
$$t_0 = \frac{13,71 - 10}{3,55/\sqrt{22}} = 4,90$$

7. Conclusões: Uma vez que $t_0 = 4,90 > 1,721$, rejeitamos H_0 e concluímos, com um nível de 0,05 de significância, que a carga média da falha excede 10 MPa.

8.3.5 Intervalo de Confiança para a Média

É fácil encontrar um intervalo de confiança de $100(1-\alpha)\%$ para a média de uma distribuição normal com variância desconhecida, procedendo como fizemos na Seção 8.2.6. Em geral, a distribuição de $T = \left(\overline{X} - \mu\right) / \left(S/\sqrt{n}\right)$ é t, com $n-1$ graus de liberdade. Fazendo $t_{\alpha/2,n-1}$ ser o ponto superior $100\alpha/2\%$ da dsitribuição t, com n-1 graus de liberdade, podemos escrever:
$$P\left(-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}\right)$$
ou
$$P\left(-t_{\alpha/2,n-1} \leq \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \leq t_{\alpha/2,n-1}\right)$$
Rearranjando essa última equação, resulta em

$$P\left(\overline{X} - t_{\alpha/2,n-1}S/\sqrt{n} \leq \mu \leq \overline{X} + t_{\alpha/2,n-1}S/\sqrt{n}\right) = 1 - \alpha \tag{8.49}$$

Isso conduz á seguinte definição de intervalo bilateral de confiança com $100\,(1-\alpha)\,\%$ para μ.

Definição: Intervalo de Confiança para Média de uma Distribuição Normal com Variância Desconhecida

Se $\overline{x}$ e s forem a média e o desvio-padrão de uma amostra aleatória proveniente de uma população normal, com variância desconhecida σ^2, então um intervalo de confiança de $100\,(1-\alpha)\,\%$ para a média μ é dado por

$$\overline{x} - t_{\alpha/2,n-1}S/\sqrt{n} \leq \mu \leq \overline{x} + t_{\alpha/2,n-1}S/\sqrt{n} \tag{8.50}$$

sendo $t_{\alpha/2,n-1}$ o ponto superior $100\alpha/2\%$ da distribuição t, com $n-1$ graus de liberdade.

Intervalos unilaterais de confiança para a média de uma distribuição normal são também de interesse e são fáceis de usar. Use simplesmente somente o limite inferior ou superior apropriado de (8.50) e troque $t_{\alpha/2,n-1}$ por $t_{\alpha,n-1}$.

EXEMPLO 8.10

Reconsidere o problema da tensão quanto à adesão no Exemplo 8.8. Sabemos que $n = 22$, $\overline{x} = 13,71$, $s = 3,55$. Encontraremos um intervalo de confiança de 95% para μ. Da Eq. 8.50, encontramos $\left(t_{\alpha/2,n-1} = t_{0,025;21} = 2,080\right)$:

$$\begin{gathered}
\overline{x} - t_{\alpha/2,n-1}s\sqrt{n} \leq \mu \leq \overline{x} + t_{\alpha/2,n-1}s\sqrt{n} \\
13,71 - 2,080\,(3,55)\,/\sqrt{22} \leq \mu \leq 13,71 + 2,080\,(3,55)\,/\sqrt{22} \\
13,71 - 1,57 \leq \mu \leq 13,71 + 1,57 \\
12,14 \leq \mu \leq 15,28
\end{gathered}$$

No exemplo 8.8, testamos uma hipótese alternativa unilateral para μ. Alguns engenheiros podem estar interessados em um intervalo inferior de confiança de 95% para a carga média na falha é encontrado usando o limite inferior de confiança de (8.50), com $t_{\alpha/2,n-1}$ trocado por $t_{\alpha,n-1}$, isso conduz a:

$$\begin{aligned}
\overline{X} - t_{0,05,n-1}S\sqrt{n} &\leq \mu \\
13,71 - 1,721\,(3,25)\,/\sqrt{22} &\leq \mu \\
12,52 &\leq \mu
\end{aligned}$$

Logo, podemos dizer com 95% de confiança que a carga média na falha excede 12,52 MPa.