TP 547- Princípios de Simulação de Sistemas de Comunicação

Prof. Samuel Baraldi Mafra

Inatel
Instituto Nacional de Telecomunicações

Suponha que $X_1,X_2,X_3,...,X_n$ são variáveis aleatórias independentes de uma distribuição comum tendo média θ e variância σ^2;

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

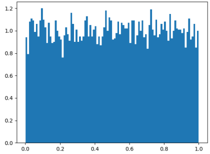
$$E[\bar{X}] = \theta$$

$$Var[\bar{X}] = \frac{\sigma^2}{n}$$

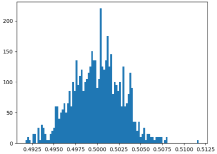
Para n grande através do teorema do limite central, podemos dizer que a média segue uma distribuição Normal $N(\theta,\sigma^2/n)$. Portanto, podemos escrever a distribuição da média como a Normal padrão fazendo Z igual:

$$Z = \sqrt{n}\frac{\bar{X} - \theta}{\sigma} \sim N(0,1)$$

PDF distribuição uniforme

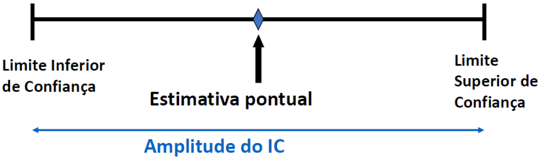


PDF médias da distribuição uniforme



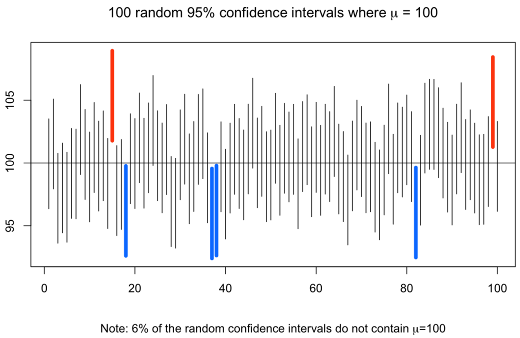
Código: geradoruniformeic.ipynb

- Uma estimativa pontual é um número;
- Um intervalo de confiança provê informação adicional sobre variabilidade.



Intervalo de Confiança

- Um intervalo de confiança (IC) é um intervalo estimado de um parâmetro de interesse de uma população. Em vez de estimar o parâmetro por um único valor, é dado um intervalo de estimativas prováveis.
- O quanto estas estimativas são prováveis será determinado pelo coeficiente de confiança $(1-\alpha)$, para $\alpha \in (0, 1)$;

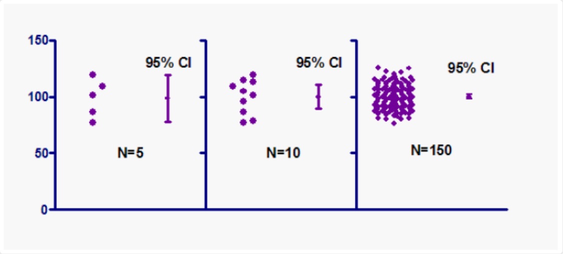


Intervalo de Confiança

- Intervalo de confiança como um intervalo que contém os valores "plausíveis" que o parâmetro pode assumir. Assim, a amplitude do intervalo está associada a incerteza que temos a respeito do parâmetro.
- Um intervalo de confiança é simplesmente uma maneira de medir o quão bem sua amostra representa a população que você está estudando.

Intervalo de Confiança

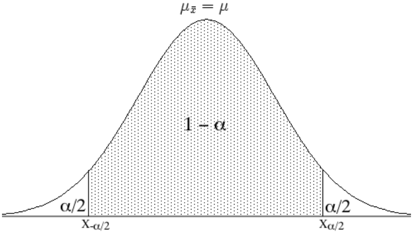
- Intervalos de confiança são usados para indicar a confiabilidade de uma estimativa. Por exemplo, um IC pode ser usado para descrever o quanto os resultados de uma pesquisa são confiáveis.
- Uma pesquisa que resulte num IC pequeno é mais confiável do que uma que resulte num IC maior.



$$\Pr[-x_{\alpha/2} < X < x_{\alpha/2}] = 1 - \alpha.$$
$$\bar{X} - x_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \theta < \bar{X} + x_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\begin{cases} x_{\alpha/2} = 1.96 \text{ para } \alpha = 0.05 \\ x_{\alpha/2} = 2.576 \text{ para } \alpha = 0.01 \end{cases}$$

O valor de $x_{\alpha/2}$ pode ser determinado através de tabelas ou pelo uso de integração numérica. Em python usar norm.cdf da biblioteca scipy.stats.



Exemplo:Variância conhecida

O projetista de uma indústria tomou uma amostra de 36 funcionários para verificar o tempo médio gasto para montar um determinado equipamento. Lembrando que foi verificado que $\bar{x} = 19.9$ e $\sigma = 5.73$, construir um intervalo de confiança de nível 95% para θ.

Uma pessoa chega a um prédio e vai até o elevador. Depois de acionado, o elevador demora entre 0 e 40 segundos para chegar. Considerando que o intervalo entre 0 e 40 segundos é uniformemente distribuído. Qual é o tempo médio que o elevador leva para chegar? Calcular o intervalo de confiança de 95% para alguns valores de N.

- a=0, b=40.
- Tempo médio:$\mu = \frac{a+b}{2} = 20$ segundos;
- Desvio padrão: $\sigma = \sqrt{(b-a)^2/12} = 20/\sqrt{12} = 11.55$

Codigo: uniabgenexic.py

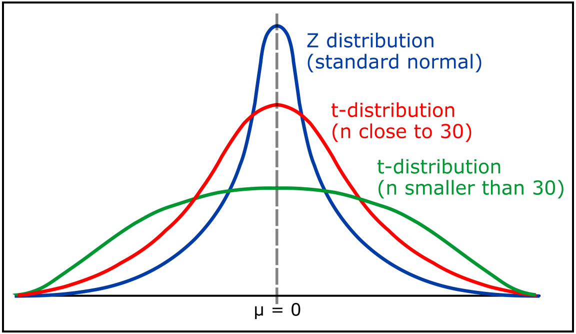
Variância desconhecida ou amostra pequena

Consideremos uma amostra aleatória simples $X_1, X_2, X_3, \ldots, X_n$, obtida de uma população com distribuição normal, com média θ e variância desconhecidas. Como neste caso a variância é desconhecida, utiliza-se a variância amostral s^2 no lugar de σ^2. Utiliza-se a distribução t-student com $n-1$ graus de liberdade.

$$\Pr[-t_{(n-1,\alpha/2)} < X < t_{(n-1,\alpha/2)}] = 1 - \alpha.$$

$$\bar{X} - t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}} < \theta < \bar{X} + t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$$

Utilizar a função ppf $t.ppf(1-\alpha/2, n-1)$ da biblioteca scipy.stats.


Z distribution
(standard normal)
t-distribution
(n close to 30)
t-distribution
(n smaller than 30)
μ = 0

Exemplo: Variância desconhecida ou amostra pequena

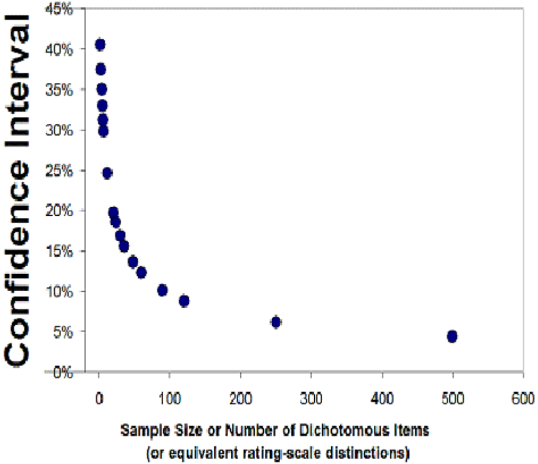
Foram realizados testes glicêmicos em 25 pacientes após um jejum de 8 horas. Os resultados são apresentados na tabela abaixo. Construir um intervalo de confiança de nível 95% para θ.

Teste glicêmico (mg/dL)				
80	118	100	90	83
117	95	84	102	80
112	78	102	121	82
77	88	73	104	88
132	91	103	140	101

Código: glicemia.ipynb

Intervalo de Confiança

- Qual o número mínimo de amostras que devo utilizar para obter um resultado consistente?



Número mínimo de amostras:

$$n = x_{\alpha/2}^{2}\frac{\sigma^2}{e^2}, \tag{1}$$

onde e representa a margem de erro aceitável em valores absolutos.

Exemplo:Variância conhecida

Uma pessoa chega a um prédio e vai até o elevador. Depois de acionado, o elevador demora entre 0 e 40 segundos para chegar. Considerando que o intervalo entre 0 e 40 segundos é uniformemente distribuído. Qual é o tempo médio que o elevador leva para chegar? Qual o número mínimo de amostras para uma margem de erro e=0.5, para um intervalo de confiança de 95%?

Intervalo de confiança para proporção

Consideremos $\hat{p}$ a proporção amostral. Pelo Teorema Central do Limite temos que, para um tamanho de amostra grande, podemos considerar a proporção amostral $\hat{p}$ como tendo aproximadamente distribuição normal com média p e variância p(1-p)/n. Desse modo segue que

$$IC(p,1-\alpha) = \left(\hat{p} - x_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + x_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right). \quad (2)$$

Exemplo

Numa amostra aleatória de tamanho n=700 foram encontrados 68 elementos defeituosos. Achar um intervalo de confiança de nível 95% para a proporção p de defeituosos.

Intervalo de confiança para variância

Consideremos uma amostra aleatória $X_1,...,X_n$ de tamanho n de uma população com distribuição normal com média μ e variância σ^2. Um estimador para σ^2 é a variância amostral s^2

$$IC(\sigma^2, 1-\alpha) = \left(\frac{(n-1)s^2}{Q_{1-\alpha/2}}, \frac{(n-1)s^2}{Q_{\alpha/2}} \right). \quad (3)$$

$Q_{1-\alpha/2} = chi2.ppf(1-\alpha/2, n-1)$: Função chi quadrado inversa ppf (chi2.ppf) em python.

Exemplo

O peso de componentes mecânicos produzidos por uma determinada empresa é uma variável aleatória que se supõe ter distribuição normal. Pretende-se estudar a variabilidade do peso dos referidos componentes. Para isso, uma amostra de tamanho 11 foi obtida,cujos valores em grama são:
98 97 102 100 98 101 102 105 95 102 100
Construa um intervalo de confiança para a variância do peso, com um grau de confiança igual a 95%.
Código: pecasmec.py