

Métodos Estocásticos da Engenharia II

Capítulo 6 - Inferência: outros problemas de estimação intervalar e testes não-paramétricos

Prof. Magno Silvério Campos

2024/2



Bibliografia

Essas notas de aulas foram baseadas nas seguintes obras:

- ❶ BORNIA, A. C.; BARBETTA, P. A.; REIS, M. M. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2009.
- ❷ HINES, W.W.; et al. *Probabilidade e Estatística na Engenharia*. 4. ed. Rio de Janeiro: LTC, 2006.
- ❸ MONTGOMERY, D.C.; RUNGER, G.C. *Estatística Aplicada e Probabilidade para Engenheiros*. 4. ed. Rio de Janeiro: LTC, 2009.
- ❹ MONTGOMERY, D.C.; RUNGER, G.C. *Estatística Aplicada e Probabilidade para Engenheiros*. 6. ed. Rio de Janeiro: LTC, 2016.

Aconselha-se pesquisá-las para se obter um **maior aprofundamento** e um **melhor aproveitamento** nos estudos.

Conteúdo Programático

1 Outros problemas de estimação intervalar

- Intervalos de Predição
- Intervalos de Tolerância

2 Testes não-paramétricos

- Testes de Aderência:
 - Teste χ^2 para adequação de ajuste
 - Teste de Kolmogorov-Smirnov
 - Teste de Lilliefors
- Testes de Associação:
 - Testes χ^2 de Independência
- Testes para duas populações:
 - Teste dos sinais
 - Teste dos Postos Sinalizados de Wilcoxon
 - Teste de Wilcoxon-Mann-Whitney



Intervalos de Predição

Introdução

Até o presente momento, foram apresentados estimadores intervalares para parâmetros populacionais, tais como a média, μ .

Há muitas situações em que há interesse em prever **uma única observação futura** para a variável aleatória de interesse, em vez de prever ou estimar a média dessa variável aleatória. Um **intervalo de predição** pode ser construído para qualquer observação única em algum tempo futuro.



Considere uma amostra aleatória dada de tamanho n , X_1, X_2, \dots, X_n , de uma população normal com média μ e variância σ^2 .

Denote por \bar{X} a média amostral. Suponha que desejemos prever a observação futura X_{n+1} .

Como \bar{X} é o preditor pontual para essa observação, o erro de predição é dado por

$$X_{n+1} - \bar{X}.$$

O valor esperado e a variância do erro de predição são

$$E(X_{n+1} - \bar{X}) = E(X_{n+1}) - E(\bar{X}) = \mu - \mu = 0$$

$$Var(X_{n+1} - \bar{X}) = Var(X_{n+1}) + Var(\bar{X}) = \sigma^2 + \frac{\sigma^2}{n} = \sigma^2(1 + \frac{1}{n}).$$

Como X_{n+1} e \bar{X} são variáveis aleatórias independentes, distribuídas normalmente, o erro de predição é também distribuído normalmente e

$$Z = \frac{(X_{n+1} - \bar{X}) - 0}{\sqrt{\sigma^2(1 + \frac{1}{n})}} = \frac{(X_{n+1} - \bar{X})}{\sqrt{\sigma^2(1 + \frac{1}{n})}} \sim N(0, 1).$$

Se σ^2 é desconhecida, pode ser estimada pela variância amostral, S^2 , e, então,

$$T = \frac{(X_{n+1} - \bar{X})}{\sqrt{S^2(1 + \frac{1}{n})}} \sim t_{(n-1)}.$$



Seguindo o procedimento usual para a construção de intervalos de confiança, o intervalo de predição bilateral de $100(1 - \alpha)\%$ de confiança é

$$-t_{\frac{\alpha}{2}, n-1} \leq \frac{(X_{n+1} - \bar{X})}{\sqrt{S^2(1 + \frac{1}{n})}} \leq t_{\frac{\alpha}{2}, n-1}.$$

Rearranjando a desigualdade acima, obtemos a forma final para o intervalo de predição bilateral de $100(1 - \alpha)\%$:

$$\bar{X} - t_{\frac{\alpha}{2}, n-1} \sqrt{S^2(1 + \frac{1}{n})} \leq X_{n+1} \leq \bar{X} + t_{\frac{\alpha}{2}, n-1} \sqrt{S^2(1 + \frac{1}{n})}$$



Intervalo de Predição

Um intervalo de predição de $100(1 - \alpha)\%$ para uma observação futura a partir de uma distribuição normal é dado por

$$\bar{X} - t_{\frac{\alpha}{2}, n-1} \sqrt{S^2 \left(1 + \frac{1}{n}\right)} \leq X_{n+1} \leq \bar{X} + t_{\frac{\alpha}{2}, n-1} \sqrt{S^2 \left(1 + \frac{1}{n}\right)}$$



Exemplo 1 - [Hines e outros(2006)]

As forças máximas experimentadas por um avião de transporte de uma linha aérea para 10 voos em determinada rota são (em unidades de gravidade, g)

1,15; 1,23; 1,56; 1,69; 1,71; 1,83; 1,83; 1,85; 1,90; 1,91.

A média e o desvio-padrão amostrais são calculados como $\bar{x} = 1,666$ e $s = 0,273$, respectivamente. Pode ser importante predizer a próxima força máxima a ser experimentada pelo avião. Então, qual será o intervalo de predição para x_{11} , com 95% de confiança?



Exemplo 2 - [Montgomery e Runger(2016)]

Um artigo no periódico *Materials Engineering* (1989, Vol. II, N^o. 4, pp. 275-281) descreve os resultados de testes trativos de adesão em 22 corpos-de-prova de liga U-700. A média amostral para a carga no ponto de falha do corpo-de-prova é $\bar{x} = 13,71$ megapascal, e o desvio-padrão da amostra é $s = 3,55$ megapascal.

- 1 Sabendo-se que a população segue uma distribuição normal, encontre um IC de 95% para μ .
- 2 Planejamos testar um vigésimo terceiro corpo de prova. Construa um intervalo de 95% de confiança para a carga na falha para esse corpo-de-prova.





Comparações

Já vimos na PARTE I deste capítulo que o intervalo de confiança para μ , com coeficiente de confiança $100(1 - \alpha)\%$ é dado por:

$$\bar{X} - t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}.$$

Por outro lado, um intervalo de predição de $100(1 - \alpha)\%$ para uma observação futura a partir de uma distribuição normal é dado por

$$\bar{X} - t_{\frac{\alpha}{2}, n-1} \sqrt{S^2 \left(1 + \frac{1}{n}\right)} \leq X_{n+1} \leq \bar{X} + t_{\frac{\alpha}{2}, n-1} \sqrt{S^2 \left(1 + \frac{1}{n}\right)}$$



Observação 1

O intervalo de predição para X_{n+1} será sempre **maior** do que o intervalo de confiança para μ , pelo fato de haver mais variabilidade associada com o erro de previsão do que com o erro de estimação.

Observação 2

Quando n torna-se grande ($n \rightarrow \infty$), o comprimento do IC diminui a zero, tornando-se essencialmente o valor único μ ; porém, o comprimento do intervalo de predição se aproxima de $2t_{\frac{\alpha}{2}, n-1}S$.

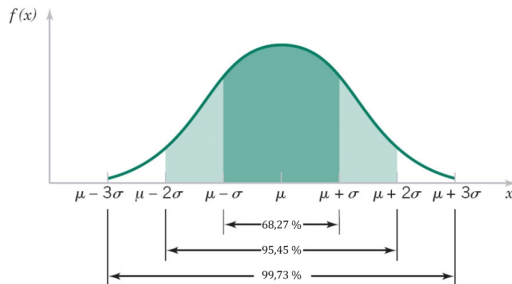
Assim, quando n aumenta, a incerteza em estimar μ vai para zero, embora sempre haverá incerteza sobre o valor futuro X_{n+1} .



Intervalos Tolerância para uma distribuição normal

Introdução

Os intervalos de confiança são os intervalos nos quais esperamos que contenham o verdadeiro parâmetro populacional, tal como a média μ . Em contraposição, os **intervalos de tolerância** são intervalos nos quais esperamos que contenham uma **porcentagem** dos valores populacionais.



Fonte: Montgomery & Runger (2009), p.74

Intervalo de Tolerância

Um intervalo de tolerância para capturar no mínimo $q\%$ dos valores de uma distribuição normal, com nível de confiança de $100(1 - \alpha)\%$, é

$$\bar{x} \pm ks$$

sendo k um fator do intervalo de tolerância que se encontra em tabelas estatísticas.



Exemplo 1 - [Hines e outros(2006)]

Reconsidere as forças máximas para o exemplo do avião de transporte. Deseja-se um intervalo de tolerância bilateral que cubra 99% de todas as forças máximas, com 95% de confiança.



Exemplo 2 - [Montgomery e Runger(2016)]

Reconsidere os testes tratativos de adesão, já apresentados no slide 10. A carga na falha para $n = 22$ corpos-de-prova foi observada e encontramos que $\bar{x} = 13,71$ e $s = 3,55$. Deseja-se um intervalo de tolerância bilateral que cubra 90% de todas as cargas na falha na população, com 95% de confiança.



Observação 1

Da tabela para o fator k do intervalo de tolerância, notamos que quando $n \rightarrow \infty$, o valor de k vai para o valor $Z_{\frac{1-q}{2}}$. Por exemplo, se quisermos que 90% da população caiam no intervalo bilateral de tolerância, k se aproxima de $Z_{0,05} = 1,645$ quando $n \rightarrow \infty$.

Observação 2

Quando $n \rightarrow \infty$, um intervalo de predição de $100(1 - \alpha)\%$ para um valor futuro se aproxima de um intervalo de tolerância que contém $100(1 - \alpha)\%$ da distribuição.



Testes não-paramétricos

Introdução: Os testes descritos nos capítulos 4 e 5 são ditos *paramétricos*, porque supõem que os dados seguem determinada distribuição de probabilidades.

Imagine que as suposições necessárias para a aplicação dos testes paramétricos não sejam satisfeitas. De acordo com Bornia *et al.* (2009), suponha que ocorram alguns dos casos abaixo:

- ❶ Os dados sob análise têm um nível de mensuração **qualitativo**;
- ❷ Os dados sob análise têm nível de mensuração quantitativo, mas há indícios de que a distribuição populacional não é a normal;
- ❸ Há interesse em realizar inferência sobre outras características da população, além dos parâmetros de sua distribuição, como a própria forma da distribuição.

Uma alternativa para essas situações é a utilização dos testes **não-paramétricos**, ou *testes livres de distribuição*.

Testes de Aderência

De acordo com Bornia *et al.* (2009), o objetivo de um teste de aderência é verificar se os dados de uma amostra comportam-se de acordo com uma distribuição teórica, que pode ser uma distribuição de probabilidades clássica (como normal, exponencial, gama, etc.), ou proporções definidas especificamente para o problema.

Serão apresentados os seguintes testes de aderência:

- Teste χ^2 para adequação de ajuste;
- Teste de Kolmogorov-Smirnov;
- Teste de Lilliefors.



Teste χ^2 para adequação de ajuste

Segundo Montgomery *et al.* (2009), um tipo de hipótese é frequentemente encontrado: não conhecemos a distribuição sobre consideração da população e desejamos testar a hipótese de que uma distribuição particular será satisfatória como um modelo para a população.

Considere uma amostra de tamanho n , proveniente de uma população cuja distribuição de probabilidades é desconhecida, arranjadas em k intervalos de classe.

Seja O_i a **frequência observada** no i -ésimo intervalo de classe. A partir da distribuição de probabilidades hipotetizada, calcula-se a **frequência esperada** E_i no i -ésimo intervalo de classe. Então, seja testar as seguintes hipóteses:

$$H_0 : O_i = E_i, \quad \forall \quad i = 1, 2, \dots, k;$$

$$H_1 : O_i \neq E_i, \quad \text{para algum } i = 1, 2, \dots, k.$$

Estatística de Teste:

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \sim \chi_{k-p-1}^2 \text{ se } H_0 \text{ é verdadeira,}$$

onde p representa o número de parâmetros da distribuição utilizada na hipótese, estimados pelas estatísticas amostrais.

Critério de Rejeição de H_0 :

Se $\chi_0^2 > \chi_{\alpha, k-p-1}^2$.

Montgomery *et al.* (2009) sugerem que para melhores resultados, $E_i \geq 3$. Caso isso não ocorra, frequências esperadas de classes adjacentes podem ser agrupadas. Assim, as frequências observadas correspondentes seriam combinadas também, reduzindo o valor de k .

Exemplo 1 - [Bornia e outros(2009)]

Determinado veículo utilitário está sofrendo pesadas críticas de seus proprietários, com relação à grande frequência de defeitos em certas posições dos pneus. Por isso, o fabricante do veículo resolveu coletar informações sobre 152 ocorrências de defeitos, classificando-as por posição do pneu. Os resultados estão na tabela abaixo:

Posição do pneu	Dianteiro Esquerdo	Dianteiro Direito	Traseiro Esquerdo	Traseiro direito	Total
Frequência	35	32	57	28	152

Usando nível de significância de 5%, há razão para acreditar que a probabilidade de defeito é diferente para alguma das posições?





Exemplo 2 - [Montgomery e Runger(2016)]

O número de defeitos nas placas de circuito impresso é suposto seguir a distribuição de Poisson. Uma amostra aleatória de $n = 60$ placas impressas foi coletada e o número de defeitos observado:

Número de Defeitos	Frequência Observada
0	32
1	15
2	9
3	4

Podemos afirmar, a um nível de 5% de significância, que o número de defeitos nessas placas seguem uma distribuição de Poisson?





Observação

Segundo Hines *et al.* (2006), uma prática comum na construção dos intervalos de classe é escolher os limites das células de modo que as frequências esperadas ($E_i = np_i$) sejam iguais para todas as classes.

Para se usar esse método, desejamos escolher as fronteiras a_0, a_1, \dots, a_k para as k classes, de modo que todas as probabilidades

$$p_i = P(a_{i-1} \leq X \leq a_i) = \int_{a_{i-1}}^{a_i} f(x)dx, \quad \forall \quad i = 1, 2, \dots, k.$$

sejam iguais.



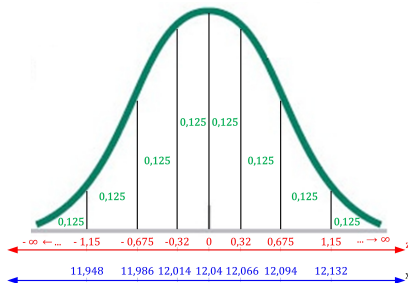
Exemplo 3 - [Hines e outros(2006)]

Um engenheiro de produção está testando uma fonte de energia usada em uma estação de trabalho de processamento. Ele deseja determinar se a voltagem de saída é descrita adequadamente por uma distribuição normal.

De uma amostra aleatória de $n = 100$ unidades ele obtém estimativas amostrais da média e do desvio-padrão, $\bar{x} = 12,04$ V e $s = 0,08$ V.

Que conclusões podem ser tiradas, usando $\alpha = 5\%$? Suponha que ele decida utilizar $k = 8$ classes de frequências para realizar o teste.





Para cada intervalo, $p_i = \frac{1}{8} = 0,125$. Assim, podemos construir a seguinte tabela:

Intervalo de Classe	Frequência Observada (O_i)	Frequência Esperada ($E_i = np_i$)
$x < 11,948$	10	12,5
$11,948 \leq x < 11,986$	14	12,5
$11,986 \leq x < 12,014$	12	12,5
$12,014 \leq x < 12,040$	13	12,5
$12,040 \leq x < 12,066$	11	12,5
$12,066 \leq x < 12,094$	12	12,5
$12,094 \leq x < 12,132$	14	12,5
$12,132 < x$	14	12,5
Total	100	100



Teste de Kolmogorov-Smirnov

De acordo com Bornia *et al.* (2009), quando se deseja verificar a aderência de um conjunto de valores em relação a uma distribuição de probabilidades específica (discreta ou contínua), uma alternativa é utilizar o Teste de Kolmogorov-Smirnov, sendo mais poderoso do que o teste do χ^2 , nestas situações.

Esse teste pode ser aplicado para avaliar a aderência a qualquer distribuição, desde que seus parâmetros sejam especificados.

Hipóteses:

Seja $F(x)$ a FDA, com parâmetros **especificados**, para a qual se quer verificar a aderência dos dados. As hipóteses são:

H_0 : os dados provém de $F(x)$ (há aderência);

H_1 : os dados **não** provém de $F(x)$ (não há aderência); .

Procedimento de Teste:

- 1 Seja a FDA empírica $S(x_i)$, definida como

$$S(x_i) = \frac{\text{número de valores } \leq x_i}{n}, \quad \forall i = 1, 2, \dots, n,$$

onde n é o tamanho da amostra e x_i é um valor qualquer da amostra.

- 2 Obtenha os valores teóricos de $F(x_i)$ ($\forall i = 1, 2, \dots, n$), calculados pela FDA, hipotetizada em H_0 .
- 3 Calcule as discrepâncias entre $S(X)$ e $F(x)$ através de:

$$|F(x_i) - S(x_i)| \quad \text{e} \quad |F(x_i) - S(x_{i-1})| \quad \forall i = 1, 2, \dots, n.$$

- 4 Estatística de Teste:

$$D = \max\{|F(x_i) - S(x_i)|, |F(x_i) - S(x_{i-1})|\} \quad \forall i = 1, 2, \dots, n.$$

- 5 Critério de Rejeição de H_0 : Se $D \geq d_c$, onde d_c é um valor

Exemplo 4 - [Bornia e outros(2009)]

A metodologia usada para calcular os índices de confiabilidade de um sistema de transmissão de energia elétrica exige que os tempos para a falha dos componentes sigam distribuições exponenciais.

Observações anteriores indicaram a validade de tal suposição, **MAS** um engenheiro recém-contratado decidiu verificar se o tempo para a falha (em horas) de um componente, especificamente crítico, pode ser admitido com distribuição exponencial com média de 500 horas.

Para testar essa hipótese, utilizando nível de significância de 1%, coletou-se uma amostra de 20 observações do tempo de falha desse componente:

7,55; 25,20; 41,00; 133,59; 146,77; 157,55; 158,07; 206,08; 385,09; 426,89;
 555,86; 639,43; 816,11; 847,57; 924,63; 945,66; 968,66; 1.130,39;
 1.143,93; 1.365,69

Que conclusões podem ser tiradas?

Solução: As hipóteses são:

H_0 : os dados provêm de $F(x)$ exponencial com média $\mu = 500$;

H_1 : os dados **não** provêm de $F(x)$ exponencial com média $\mu = 500$;

Os cálculos para o teste são apresentados na tabela a seguir:

x_i	$S(x_i)$	$F(x_i)$	$ F(x_i) - S(x_i) $	$ F(x_i) - S(x_{i-1}) $
	0	0		
7,55	0,050	0,015	0,035	0,015
25,20	0,100	0,049	0,051	0,001
41,00	0,150	0,079	0,071	0,021
133,59	0,200	0,234	0,034	0,084
146,77	0,250	0,254	0,004	0,054
157,55	0,300	0,27	0,03	0,02
158,07	0,350	0,271	0,079	0,029
206,08	0,400	0,338	0,062	0,012
385,09	0,450	0,537	0,087	0,137
426,89	0,500	0,574	0,074	0,124
555,86	0,550	0,671	0,121	0,171
639,43	0,600	0,722	0,122	0,172
816,11	0,650	0,805	0,155	0,205
847,57	0,700	0,816	0,116	0,166
924,63	0,750	0,843	0,093	0,143
945,66	0,800	0,849	0,049	0,099
968,66	0,850	0,856	0,006	0,056
1130,39	0,900	0,896	0,004	0,046
1143,93	0,950	0,899	0,051	0,001
1365,69	1,000	0,935	0,065	0,015



Teste de Lilliefors

De acordo com Bornia *et al.* (2009), esse teste é usado para verificar a aderência dos dados a uma distribuição normal qualquer, isto é, sem a especificação de seus parâmetros.

É bastante parecido com o teste de aderência de Kolmogorov-Smirnov, pois também avaliamos:

- As distribuições acumuladas $S(x)$ e $F(x)$;
- Obtemos a distância máxima D entre elas;
- Comparamos D com um valor tabelado, em função do nível de significância e do tamanho da amostra;

As diferenças residem na forma de obtenção de $F(x)$ (pois a média e o desvio padrão são calculados com base na amostra) e na tabela utilizada para a decisão do teste.



Exemplo 5 - [Bornia e outros(2009)]

Um fabricante de autopeças está próximo de fechar um grande contrato com uma montadora. O ponto-chave é a garantia da qualidade de seus produtos, especialmente do diâmetro (em *mm*) dos eixos produzidos que ele supõe seguir uma distribuição normal

A montadora selecionou uma amostra aleatória de 15 eixos, para testar as especificações a 5% de significância. Os valores observados foram:

93,45; 94,46; 94,93; 96,17; 96,74; 97,07; 97,68; 97,93; 99,10; 99,30;
100,73; 103,29; 103,60; 103,83; 105,20

Que conclusões podem ser tiradas?



Solução: As hipóteses são:

H_0 : os dados provêm de $F(x)$ normal

H_1 : os dados **não** provêm de $F(x)$ normal

Os cálculos para o teste são apresentados na tabela a seguir:

x_i	$S(x_i)$	z_i	$F(x_i) = \Phi(z_i)$	$ F(x_i) - S(x_i) $	$ F(x_i) - S(x_{i-1}) $
	0		0		
93,45	0,067	-1,47	0,070	0,003	0,070
94,46	0,133	-1,20	0,115	0,018	0,048
94,93	0,200	-1,07	0,142	0,058	0,009
96,17	0,267	-0,74	0,230	0,037	0,030
96,74	0,333	-0,58	0,280	0,053	0,013
97,07	0,400	-0,49	0,310	0,090	0,023
97,68	0,467	-0,33	0,371	0,096	0,029
97,93	0,533	-0,26	0,397	0,136	0,070
99,10	0,600	0,05	0,522	0,078	0,011
99,30	0,667	0,11	0,543	0,124	0,057
100,73	0,733	0,49	0,690	0,043	0,023
103,29	0,800	1,19	0,882	0,082	0,149
103,60	0,867	1,27	0,898	0,031	0,098
103,83	0,933	1,33	0,909	0,024	0,042
105,20	1,000	1,70	0,956	0,044	0,023



Testes de Associação

De acordo com Bornia *et al.* (2009), dizemos que duas variáveis qualitativas não têm associação entre si, quando são independentes. Podemos testar essa independência através de testes de hipóteses a partir de uma estrutura denominada tabela de contingência.

Exemplo 6 - [Montgomery e Runger(2016)]

Uma empresa opera três máquinas com possibilidade de duas mudanças de configuração todo dia. De acordo com registros de produção, são coletados dados do número de interrupções:

	Máquinas		
Configuração	A	B	C
1	41	31	16
2	46	17	21

Deseja-se testar a hipótese, a um nível de 5% de significância, de que as interrupções são independentes das configurações adotadas.

Testes χ^2 de Independência

Considerando duas variáveis categóricas X e Y , podemos construir as seguintes hipóteses:

H_0 : X e Y são independentes;

H_1 : Não há independência entre X e Y (*hassocio*);.

Alternativamente, podemos escrever:

$$H_0 : f(x, y) = p(x)q(y)$$

$$H_1 : f(x, y) \neq p(x)q(y).$$

Isto é, duas variáveis são independentes \Leftrightarrow a probabilidade conjunta for igual ao produto das probabilidades marginais.

Considerando:

- O_{ij} a frequência observada na célula (i, j) da tabela de contingência;
- E_{ij} a frequência esperada na célula (i, j) , segundo H_0 , vem:

$$H_0 : O_{ij} = E_{ij}$$

$$H_1 : O_{ij} \neq E_{ij},$$

onde $E_{ij} = np_{ij} = n\hat{p}_i\hat{p}_j$, se H_0 é verdadeira, e p_{ij} é a proporção observada na célula (i, j) da tabela de contingência.

Estatística de Teste

$$\chi_0^2 = \sum_i \sum_j \left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}} \right] \sim \chi_{(n^\circ \text{ de linhas} - 1)(n^\circ \text{ de colunas} - 1)}^2$$

Agora, se a tabela de contingência for 2×2 ,

$$\chi_0^2 = \sum_i \sum_j \left[\frac{(|O_{ij} - E_{ij}| - 0,5)^2}{E_{ij}} \right] \sim \chi_1^2$$

Critério de Rejeição Se $\chi_0^2 \geq \chi_{\alpha, (n^\circ \text{ de linhas} - 1)(n^\circ \text{ de colunas} - 1)}^2$

Voltando ao exemplo 6, temos a seguinte solução:

Hipóteses:

H_0 : as interrupções na produção independem da configuração adotada;

H_1 : há associação entre as interrupções e as configurações.

Estatística de teste

Frequências Observadas O_{ij}

Configuração	Máquinas		
	A	B	C
1	41	31	16
2	46	17	21

Frequências Esperadas E_{ij}

Configuração	Máquinas			Total
	A	B	C	
1	$172 \cdot \frac{88}{172} \cdot \frac{87}{172}$ = 44,51	24,56	18,93	88
2	42,49	23,44	18,07	84
Total	87	48	37	172

Logo,

$$\chi_0^2 = \sum_i \sum_j \left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}} \right] =$$

Exemplo 7 - Adaptado de [Bornia e outros(2009)]

Uma empresa produz um certo tipo de produtos, trabalhando em três turnos. O setor de qualidade deseja verificar se o desempenho dos turnos é semelhante, no que tange às proporções de peças aprovadas e rejeitadas.

Para isso, amostras aleatórias desse produto foram coletadas em cada turno. Uma dessas amostras, é tabulada abaixo:

Situação das Peças	Turno			Total
	Matutino	Vespertino	Noturno	
Aprovadas	432	456	424	1312
Rejeitadas	230	238	219	687
Total	662	694	643	1999

A partir desses dados, é possível considerar semelhante o desempenho dos três turnos, usando um nível de significância de 5%?







Testes para duas populações

Introdução

De acordo com Bornia *et al.* (2009), muitas vezes, não é possível fazer uso de testes paramétricos para comparar duas populações, por causa da não observância de premissas básicas para a aplicação destes testes.

Por exemplo, alguns testes paramétricos apresentados na Parte I e II deste material requerem a suposição de normalidade para as populações comparadas.

Nesta seção, serão apresentados métodos alternativos para os casos onde as suposições não são ou não podem ser verificadas. Para isso, serão apresentados os seguintes testes:

- Teste dos Sinais;
- Teste dos Postos Sinalizados de Wilcoxon;
- Teste de Wilcoxon-Mann-Whitney.

Testes dos Sinais: Caso I: amostra pequena ($n < 15$)

Teste dos Sinais

Este teste é utilizado para comparar a mediana de duas distribuições populacionais, com base em [amostras pareadas](#).

Hipótese Nula

$$H_0 : Md_2 = Md_1$$



1 Estatística de Teste:

- Teste Unilateral Superior ($H_1 : Md_2 > Md_1$)

$$P = \sum_{x=n^+}^n \binom{n}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{n-x}$$

- Teste Unilateral Inferior ($H_1 : Md_2 < Md_1$)

$$P = \sum_{x=0}^{n^+} \binom{n}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{n-x}$$

- Teste Bilateral ($H_1 : Md_2 \neq Md_1$)

- Se $n^+ < \frac{n}{2}$,

$$P = 2 \cdot \sum_{x=0}^{n^+} \binom{n}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{n-x}$$

- Se $n^+ > \frac{n}{2}$,

$$P = 2 \cdot \sum_{x=n^+}^n \binom{n}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{n-x}$$

Região crítica

Hipótese Alternativa	Critério de rejeição de H_0
$H_1 : Md_2 \neq Md_1$	$\alpha \geq P$
$H_1 : Md_2 > Md_1$	$\alpha \geq P$
$H_1 : Md_2 < Md_1$	$\alpha \geq P$



Testes dos Sinais: Caso II: amostra grande ($n \geq 15$)

Hipótese Nula: $H_0 : Md_2 = Md_1$

Estatística de Teste

$$Z_0 = \frac{2n_+^* - n}{\sqrt{n}}, \text{ onde}$$

- $n_+^* = n^+ - 0,5$ (teste unilateral à direita);
- $n_+^* = n^+ + 0,5$ (teste unilateral à esquerda);
- $n_+^* = n^+$ (teste bilateral);

Hipótese Alternativa	Critério de rejeição de H_0
$H_1 : Md_2 \neq Md_1$	$ z_0 > z_{\frac{\alpha}{2}}$
$H_1 : Md_2 > Md_1$	$z_0 > z_{\alpha}$
$H_1 : Md_2 < Md_1$	$z_0 < -z_{\alpha}$

Exemplo 8 - Adaptado de [Bornia e outros(2009)]

Uma empresa está estudando a viabilidade de utilizar um novo tipo de equipamento (eletrônico), ao invés do equipamento utilizado atualmente (mecânico).

Após treinamento apropriado, 26 operários foram sorteados para realizar as medições das mesmas peças, com o equipamento eletrônico e o mecânico. Os tempos gastos (em segundos) foram registrados e estão na tabela a seguir:

Operário	Eletrônico	Mecânico	Sinal	Operário	Eletrônico	Mecânico	Sinal
1	27,0	27,0	0	14	22,0	29,0	-
2	25,0	30,1	-	15	16,0	16,0	0
3	22,0	28,0	-	16	22,0	20,6	+
4	34,0	34,0	0	17	29,5	25,0	+
5	23,0	24,5	-	18	36,0	33,0	+
6	22,0	28,0	-	19	35,0	41,0	-
7	25,0	28,0	-	20	35,0	35,0	0
8	32,3	30,0	+	21	27,5	28,0	-
9	34,0	36,0	-	22	29,0	31,8	-
10	23,5	29,0	-	23	27,0	28,0	-
11	34,2	39,0	-	24	24,3	29,3	-
12	31,8	30,0	+	25	30,1	30,8	-
13	28,4	20,0	+	26	29,3	25,6	-

Somente será viável a introdução dos novos equipamentos se o tempo mediano de medição for menor do que o obtido com os equipamentos mecânicos em uso.

Sabe-se que os tempos de medição desses equipamentos não costumam seguir distribuições normais. Teste as seguintes hipóteses (use $\alpha = 0,05$):

$$H_0 : Md_e = Md_m$$

$$H_1 : Md_e < Md_m,$$





Teste dos postos sinalizados de Wilcoxon

Introdução

Segundo Bornia *et al.* (2009), o teste dos postos sinalizados de Wilcoxon, ou simplesmente, **Teste de Wilcoxon**, é uma alternativa mais poderosa do que o Teste dos Sinais, uma vez que este último não costuma detectar diferença entre dois tratamentos (populações) se as amostras não forem grandes, a menos que a diferença real entre os tratamentos seja muito grande.

Procedimento: Atribua postos às diferenças de cada par de observações, **independentemente do sinal**, alocando o posto 1 à menor diferença em **módulo**; o posto 2 à segunda menor diferença em módulo; \dots ; o posto n à maior diferença em módulo. Às observações empatadas atribua a média dos postos normais correspondentes. Por exemplo, considere três diferenças iguais (empatadas), que, se distintas, corresponderiam aos postos 7, 8 e 9. Então, atribua a cada uma delas a média desses postos, isto é, o posto 8.

É pressuposto que a variável estudada seja **contínua**.

Observada uma amostra pareada de $n > 20$ observações (já descontando os casos em que as diferenças entre os tratamentos for nula), defina s_+ como a *soma dos postos das diferenças positivas* da amostra observada.

Hipótese Nula

$$H_0 : Md_2 = Md_1$$



Estatística de Teste

$$Z_0 = \frac{\sqrt{24}}{4} \cdot \frac{4s_+ - n(n+1)}{\sqrt{n(n+1)(2n+1)}} \sim N(0,1) \text{ se } H_0 \text{ é verdadeiro}$$

Região Crítica

Hipótese Alternativa	Critério de rejeição de H_0
$H_1 : Md_2 \neq Md_1$	$ z_0 > z_{\frac{\alpha}{2}}$
$H_1 : Md_2 > Md_1$	$z_0 > z_{\alpha}$
$H_1 : Md_2 < Md_1$	$z_0 < -z_{\alpha}$



Exemplo 9 - Adaptado de [Bornia e outros(2009)]

Aplicando o Teste de Wilcoxon ao problema do exemplo 8.

Operário	Eletrônico	Mecânico	Sinal	Operário	Eletrônico	Mecânico	Sinal
1	27,0	27,0	0	14	22,0	29,0	-
2	25,0	30,1	-	15	16,0	16,0	0
3	22,0	28,0	-	16	22,0	20,6	+
4	34,0	34,0	0	17	29,5	25,0	+
5	23,0	24,5	-	18	36,0	33,0	+
6	22,0	28,0	-	19	35,0	41,0	-
7	25,0	28,0	-	20	35,0	35,0	0
8	32,3	30,0	+	21	27,5	28,0	-
9	34,0	36,0	-	22	29,0	31,8	-
10	23,5	29,0	-	23	27,0	28,0	-
11	34,2	39,0	-	24	24,3	29,3	-
12	31,8	30,0	+	25	30,1	30,8	-
13	28,4	20,0	+	26	29,3	25,6	-

Sabe-se que os tempos de medição desses equipamentos não costumam seguir distribuições normais. Teste as seguintes hipóteses (use $\alpha = 0,05$):

$$H_0 : Md_e = Md_m$$

$$H_1 : Md_e < Md_m,$$

Solução

Operário	Eletrônico	Mecânico	Diferença
1	27,0	27,0	0,0
2	25,0	30,1	-5,1
3	22,0	28,0	-6,0
4	34,0	34,0	0,0
5	23,0	24,5	-1,5
6	22,0	28,0	-6,0
7	25,0	28,0	-3,0
8	32,3	30,0	2,3
9	34,0	36,0	-2,0
10	23,5	29,0	-5,5
11	34,2	39,0	-4,8
12	31,8	30,0	1,8
13	28,4	20,0	8,4

Operário	Eletrônico	Mecânico	Diferença
14	22,0	29,0	-7,0
15	16,0	16,0	0,0
16	22,0	20,6	1,4
17	29,5	25,0	4,5
18	36,0	33,0	3,0
19	35,0	41,0	-6,0
20	35,0	35,0	0,0
21	27,5	28,0	-0,5
22	29,0	31,8	-2,8
23	27,0	28,0	-1,0
24	24,3	29,3	-5,0
25	30,1	30,8	-0,7
26	29,3	25,6	3,7

Operário	Diferença	Diferença	Posto (a)	Posto (b)	Sinal
21	-0,5	0,5	1	1	-
25	-0,7	0,7	2	2	-
23	-1,0	1,0	3	3	-
16	1,4	1,4	4	4	+
5	-1,5	1,5	5	5	-
12	1,8	1,8	6	6	+
9	-2,0	2,0	7	7	-
8	2,3	2,3	8	8	+
22	-2,8	2,8	9	9	-
7	-3,0	3,0	10	10,5	-
18	3,0	3,0	11	10,5	+

Operário	Diferença	Diferença	Posto (a)	Posto (b)	Sinal
17	4,5	4,5	12	12	+
11	-4,8	4,8	13	13	-
24	-5,0	5,0	14	14	-
2	-5,1	5,1	15	15	-
10	-5,5	5,5	16	16	-
3	-6,0	6,0	17	18	-
6	-6,0	6,0	18	18	-
19	-6,0	6,0	19	18	-
26	-6,3	6,3	20	20	-
14	-7,0	7,0	21	21	-
13	8,4	8,4	22	22	+



Teste de Wilcoxon-Mann-Whitney

Introdução

Esse teste é utilizado para comparar a posição central de duas populações, com base em amostras **independentes**, extraídas aleatoriamente dessas populações.

Procedimento

Sejam n_1 e n_2 os tamanhos das amostras 1 e 2. Os $n_1 + n_2$ elementos devem ser ordenados em **ordem crescente** em termos da variável observada.

Atribua o posto 1 à menor observação, posto 2 à segunda menor observação, e assim por diante, até atribuir o posto $n_1 + n_2$ à maior observação. Em caso de empate, adote o procedimento apresentado no Teste de Wilcoxon.



É pressuposto que os dados das duas amostras independentes provêm de populações com distribuições contínuas.

Observada uma amostra pareada de $n > 20$ observações, defina w_1 como sendo a *soma dos postos da amostra 1*.

Além disso, calcule

$$u = w_1 - \left[\frac{n_1(n_1 + 1)}{2} \right]$$

Hipótese Nula

$$H_0 : Md_2 = Md_1$$



Estatística de Teste

$$Z_0 = \frac{u - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim N(0, 1) \text{ se } H_0 \text{ é verdadeiro}$$

Região Crítica

Hipótese Alternativa	Critério de rejeição de H_0
$H_1 : Md_2 \neq Md_1$	$ z_0 > z_{\frac{\alpha}{2}}$
$H_1 : Md_2 > Md_1$	$z_0 > z_{\alpha}$
$H_1 : Md_2 < Md_1$	$z_0 < -z_{\alpha}$



Exemplo 10 - [Bornia e outros(2009)]

Um administrador de rede tem recebido insistentes reclamações de usuários de que os tempos de processamento dos dois servidores da rede são diferentes, no que tange ao acesso. Visando conferir tal afirmação dos clientes, ele registrou os tempos de acessos (em segundos) de 30 usuários em cada servidor. Os dados estão na tabela abaixo:

Servidor 1			Servidor 2		
5,83	3,78	6,79	2,27	6,24	2,44
0,99	1,40	2,70	7,41	4,73	4,17
6,07	5,88	3,05	3,21	9,34	5,01
6,53	3,52	2,44	7,76	4,33	16,68
0,04	3,42	3,74	2,24	4,63	2,97
4,96	0,99	2,66	1,93	3,97	13,45
6,86	1,72	3,09	6,07	4,61	5,35
2,55	4,05	2,03	3,80	5,02	1,80
2,63	1,70	4,65	2,93	6,40	2,97
1,97	6,48	4,26	9,04	4,51	10,75

Supondo-se que os tempos de acesso não seguem uma distribuição normal, teste as seguintes hipóteses (use $\alpha = 0,05$):

$$H_0 : Md_e = Md_m$$

$$H_1 : Md_e \neq Md_m,$$

Solução

Servidor	Tempo (s)	Posto (a)	Posto (b)	Servidor	Tempo (s)	Posto (a)	Posto (b)	Servidor	Tempo (s)	Posto (a)	Posto (b)	Servidor	Tempo (s)	Posto (a)	Posto (b)
1	0,04	1	1	1	2,63	16	16	1	4,05	31	31	1	6,07	46	46,5
1	0,99	2	2,5	1	2,66	17	17	2	4,17	32	32	2	6,07	47	46,5
1	0,99	3	2,5	1	2,70	18	18	1	4,26	33	33	2	6,24	48	48
1	1,40	4	4	2	2,93	19	19	2	4,33	34	34	2	6,40	49	49
1	1,70	5	5	2	2,97	20	20,5	2	4,51	35	35	1	6,48	50	50
1	1,72	6	6	2	2,97	21	20,5	2	4,61	36	36	1	6,53	51	51
2	1,80	7	7	1	3,05	22	22	2	4,63	37	37	1	6,79	52	52
2	1,93	8	8	1	3,09	23	23	1	4,65	38	38	1	6,86	53	53
1	1,97	9	9	2	3,21	24	24	2	4,73	39	39	2	7,41	54	54
1	2,03	10	10	1	3,42	25	25	1	4,96	40	40	2	7,76	55	55
2	2,24	11	11	1	3,52	26	26	2	5,01	41	41	2	9,04	56	56
2	2,27	12	12	1	3,74	27	27	2	5,02	42	42	2	9,34	57	57
1	2,44	13	13,5	1	3,78	28	28	2	5,35	43	43	2	10,75	58	58
2	2,44	14	13,5	2	3,80	29	29	1	5,83	44	44	2	13,45	59	59
1	2,55	15	15	2	3,97	30	30	1	5,88	45	45	2	16,68	60	60

Da tabela acima, temos que $w_1 = 754$.







Bornia, A., outros, 2009. Estatística para Cursos de Engenharia e Informática. São Paulo: Atlas.



Hines, W., outros, 2006. Probabilidade e Estatística na Engenharia. Rio de Janeiro: LTC.



Montgomery, D., Runger, G., 2016. Estatística Aplicada e Probabilidade para Engenheiros. Rio de Janeiro: LTC.

