

# Probabilidade e Estatística

Paulo de Souza

2022-05-13



# Sumário

<b>Informações Gerais</b>	<b>5</b>
Sobre o Livro . . . . .	5
Uso . . . . .	5
 <b>I Básico</b>	 <b>7</b>
<b>1 Introdução</b>	<b>9</b>
1.1 Conceitos Principais . . . . .	9
1.2 Testes de Hipótese . . . . .	9
1.3 Testes de Comparação Amostral . . . . .	9
 <b>2 Estatística</b>	 <b>11</b>
2.1 Conceitos Básicos de Estatística . . . . .	11
 <b>3 Probabilidade</b>	 <b>15</b>
3.1 Análise Combinatória . . . . .	15
3.2 Distribuições de Probabilidade . . . . .	15
 <b>II Testes Amostrais</b>	 <b>21</b>
 <b>4 Dois Grupos Independentes e Paramétricos</b>	 <b>23</b>
4.1 Intervalo e limite de confiança . . . . .	23
4.2 t de Student . . . . .	23
4.3 Comparação entre 2 proporções . . . . .	23
 <b>5 Dois Grupos Independentes e Não-Paramétricos</b>	 <b>25</b>
5.1 Qui-Quadrado . . . . .	25
5.2 U de Mann Whitney . . . . .	25
5.3 Prova de Fischer . . . . .	29

<b>6</b>	<b>Dois grupos Pareados e Paramétricos</b>	<b>31</b>
6.1	Teste de t-Student pareado . . . . .	31
<b>7</b>	<b>Dois grupos Pareados e Não - Paramétricos</b>	<b>33</b>
7.1	Prova de MacNemar . . . . .	33
7.2	Prova de Wilcoxon . . . . .	33
<b>8</b>	<b>Três ou mais grupos Independentes e Paramétricos</b>	<b>35</b>
8.1	Teste de Tuckey . . . . .	35
8.2	ANOVA de 1 ou 2 vias . . . . .	35
<b>III</b>	<b>Testes de Normalidade</b>	<b>37</b>
<b>9</b>	<b>Testes de Normalidade</b>	<b>39</b>
9.1	Shapiro-Wilk . . . . .	39
9.2	Kolmogorov - Smirnov . . . . .	39
9.3	Anderson - Darling . . . . .	39
9.4	Cramer Von-Mises . . . . .	39
<b>IV</b>	<b>ACA</b>	<b>41</b>
<b>10</b>	<b>Análise de Concordância de Atributos</b>	<b>43</b>
<b>11</b>	<b>Estática Kappa</b>	<b>45</b>
11.1	Teste Kappa de Cohen . . . . .	45
11.2	Teste Kappa de Fleiss . . . . .	45

# Informações Gerais

## Sobre o Livro

Este livro, é apenas um resumo baseado em anotações do autor, com o que diz respeito ao estudo de temas referentes a **probabilidade** e **estatística**.

## Uso

O livro pode ser usado pelos entusiastas nos assuntos supracitados.



**Parte I**

**Básico**





# Capítulo 1

## Introdução

### 1.1 Conceitos Principais

Grupos independentes

Grupos pareados

Tipo paramétrico

Tipo não paramétrico

### 1.2 Testes de Hipótese

#### 1.2.1 Hipótese nula e alternativa

#### 1.2.2 O significado de p-valor

### 1.3 Testes de Comparação Amostral

São diversos os modelos de dados que são analisados, e cada um destes tem suas características probabilísticas; quando queremos comparar grupos amostrais de nossos dados, são necessários testes para entender melhor como essa amostra se comporta.

Na Tabela abaixo são apresentados alguns dos principais testes de **Comparação entre Amostras**, cada um dos termos da tabela, assim como os métodos, serão explicados ao longo deste livro/resumo.

Tabela 1.1: Testes Para Comparação de Amostras

Quantidade	Tipo	Método de Teste
2 grupos independentes	<i>paramétricos</i>	Int. e lim. de confiança (1 ou 2 grupos) t de Student (1 ou 2 grupos)
	<i>não paramétricos</i>	Comparação entre 2 proporções Qui-quadrado $\chi^2$ U de Mann Whitney
2 grupos pareados	<i>paramétrico</i>	Prova de Fischer t de Student pareado
	<i>não paramétricos</i>	Prova de MacNemar

Quantidade	Tipo	Método de Teste
$\geq 3$ grupos independentes	<i>paramétrico</i>	Prova de Wilcoxon
	<i>não paramétricos</i>	ANOVA de 1 ou 2 vias Qui-quadrado $\chi^2$ Kruskall Wallis
$\geq 3$ grupos pareados	<i>paramétrico</i>	ANOVA p/ medidas repetidas
	<i>não paramétrico</i>	Teste de Friedman

Na linha 1 da tabela 1.1 as abreviações **Int** e **lim** significam **intervalo** e **limite**, respectivamente.

# Capítulo 2

## Estatística

Em probabilidade e estatística, existem diversos conceitos e axiomas que são fundamentais para o entendimento e a resolução dos problemas. Neste capítulo serão desenvolvidos os pontos que serão mais aplicados ao decorrer do livro, demais conceitos que sejam considerados extras, serão apenas indicados e referências para estes são deixadas a disposição.

### 2.1 Conceitos Básicos de Estatística

Entre os conceitos mais básicos da estatística, estão a **média**, **moda** e **mediana**, de forma direta a explicação de cada uma é dada na sequência

**Média** - Valor médio

**Mediana** - O valor central

**Moda** - O valor que mais se repete

#### 2.1.1 Média

A **média** como citado anteriormente, é o valor médio de uma sequência de dados, matematicamente isso significa a soma de todos os termos, dividido pela quantidade dos termos, como apresentado na equação (2.1)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

Para fixar melhor este conceito, vejamos o exemplo abaixo.

---

**Exemplo 1.** Dado o seguinte registro da velocidade de 13 carros:

$$vel = [99, 86, 87, 88, 111, 86, 103, 87, 94, 78, 77, 85, 86]$$

calcular a média desses dados.

**Resolução:** Para calcular a média, basta somarmos todos os termos e dividirmos pela quantidade de termos, isto é

$$\bar{x} = \frac{1}{13} (99 + 86 + 87 + 111 + 86 + 103 + 87 + 94 + 78 + 77 + 85 + 86) = 89.77$$

Portanto, a média das velocidades coletadas é  $\bar{x} = 89.77$

Outro conceito que usualmente aparece, é o de **média ponderada**, neste caso é associado um determinado “peso” a cada um dos termos da amostra.

### 2.1.2 Mediana

### 2.1.3 Moda

### 2.1.4 Variância

A **Variância** é um parâmetro que compara o quão distantes estão os valores de determinado grupo de dados com relação a média deste mesmo grupo. A mesma pode ser do tipo **Amostral** ou **Populacional** e a diferença fica mais explícita na equação que as definem.

#### Variância Amostral

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} [\sum x_i^2 - n\bar{x}^2] \quad (2.2)$$

#### Variância Populacional

$$\sigma^2 = \sum_{i=1}^N \frac{(x_i - \bar{\mu})^2}{N} = \frac{\sum x_i^2}{N} - \bar{\mu}^2 \quad (2.3)$$

#### 2.1.4.1 Demonstração das relações de Variância

Seja a **variância amostral** dada pela relação inicial:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

vamos mostrar que a mesma pode ser escrita como sendo

$$s^2 = \frac{1}{n-1} [\sum x_i^2 - n\bar{x}^2]$$

*Demonstração.*

$$\begin{aligned}
 s^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \\
 &= \frac{\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{n-1} \\
 &= \frac{1}{n-1} \left( \sum x_i^2 - 2 \sum x_i \bar{x} + \sum \bar{x}^2 \right) \\
 &= \frac{1}{n-1} \left[ \sum x_i^2 - 2 \sum x_i \left( \frac{1}{n} \sum x_i \right) + \sum \left( \frac{1}{n} \sum x_i \right)^2 \right] \\
 &= \frac{1}{n-1} \left[ \sum x_i^2 - \frac{2}{n} \left( \sum x_i \right)^2 + n \left( \frac{1}{n} \right)^2 \left( \sum x_i \right)^2 \right] \\
 &= \frac{1}{n-1} \left[ \sum x_i^2 - \frac{2n}{n^2} \left( \sum x_i \right)^2 + n \left( \frac{1}{n} \right)^2 \left( \sum x_i \right)^2 \right] \\
 &= \frac{1}{n-1} \left[ \sum x_i^2 + \frac{1}{n^2} (n-2n) \left( \sum x_i \right)^2 \right] \\
 &= \frac{1}{n-1} \left[ \sum x_i^2 + \left( \frac{1}{n} \sum x_i \right)^2 (-n) \right] \\
 &= \frac{1}{n-1} \left[ \sum x_i^2 - n\bar{x}^2 \right]
 \end{aligned}$$

□

Para a **Variância Populacional** segue do resultado anterior

*Demonstração.*

$$\begin{aligned}
 \sigma^2 &= \frac{\sum_{i=1}^N (x_i - \bar{\mu})^2}{N} \\
 &= \frac{1}{N} \left[ \sum x_i^2 - N\bar{\mu}^2 \right] \\
 &= \frac{\sum x_i^2}{N} - \bar{\mu}^2
 \end{aligned}$$

□

Segue agora alguns exemplos da aplicação da variância.

Os mesmos foram retirados da referência...

### 2.1.5 Desvio Padrão



# Capítulo 3

## Probabilidade

Neste capítulo serão apresentados os seguintes tópicos:

- Axiomas da Probabilidade
- Análises Combinatórias
- Distribuições de Probabilidade

### 3.1 Análise Combinatória

### 3.2 Distribuições de Probabilidade

São diversos os tipos de distribuições para análise de dados, podendo ser separado em dois grupos, o de distribuições **discretas** e o de distribuições **contínuas**; as mesmas ainda apresentam características importantes, são algumas delas:

- Função de Densidade de Probabilidade (**PDF**)
- Função de Densidade Acumulada (**CDF**)
- Função Percentil (**PPF**)
- Esperança e Variância da Distribuição (**E(x)** e **V(x)**)

Na sequência são apresentadas várias dessas distribuições e suas características, além disso, é disposto implementações em *Octave* para se obter resultados de estudo. Na próxima seção, é feita uma bateria de exemplos que mostram como aquelas são utilizadas.

#### 3.2.1 Normal

##### Densidade de Probabilidade

A fórmula geral para a **Função Densidade de Probabilidade** de uma **Distribuição Normal** é

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (3.1)$$

Nos casos em que  $\mu = 0$  e  $\sigma = 1$ , temos a chamada **função normal padrão**, costumeiramente representado por  $N(1,0)$ . A equação anterior se reduz a:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (3.2)$$

O seguinte gráfico é referente a **PDF** da normal padrão.

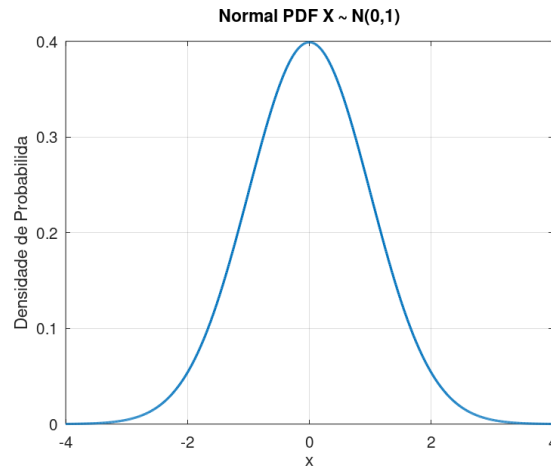


Figura 3.1: Função Densidade de Probabilidade da Normal Padrão

### Densidade Acumulada

A fórmula para o cálculo da **Função Densidade Acumulada** para uma distribuição normal padrão é dado por:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (3.3)$$

O seguinte gráfico representa os valores de **CDF** para uma distribuição normal padrão:

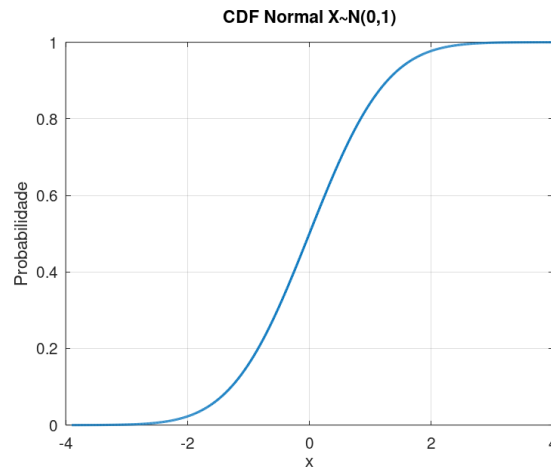


Figura 3.2: Função Densidade Acumulada da Normal Padrão



### Função Percentil

Não existe uma forma fechada de se calcular a **função percentil** para a distribuição normal; no entanto sua interpretação é que dado um valor de probabilidade  $p$  obtêm-se o valor de  $x$ , isto é, ela é a inversa da **CDF**. No gráfico a seguir é apresentada a **PPF** da distribuição normal padrão.

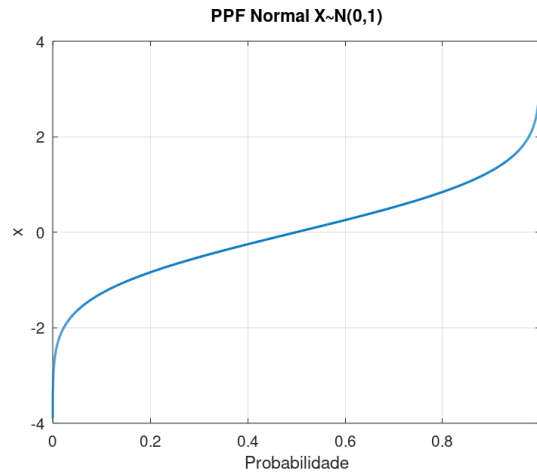


Figura 3.3: Função Percentil de Probabilidade da Normal Padrão

### 3.2.2 Uniforme

#### Densidade de Probabilidade

A **Distribuição Uniforme** tem sua **Densidade de Probabilidade** dada por:

$$f(x) = \frac{1}{B - A} \quad A \leq x \leq B \quad (3.4)$$

Em que  $A$  é o parâmetro locação (ou desvio) e  $B - A$  é o parâmetro de escala. O gráfico a seguir mostra o caso em que  $A = 1$  e  $B = 3$ .

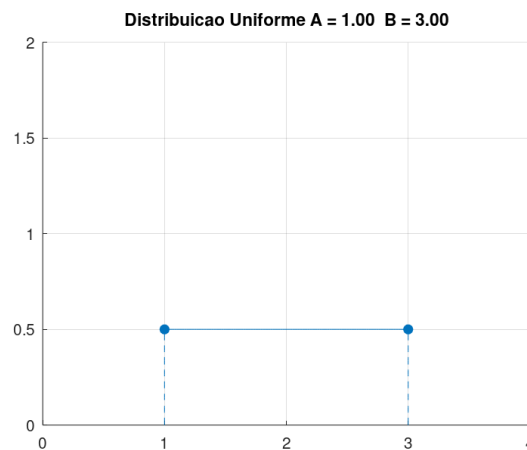


Figura 3.4: Função Densidade de Probabilidade da Uniforme

Na ocasião em que  $A = 0$  e  $B = 1$ , temos a chamada **distribuição uniforme padrão**, e a equação anterior se reduz a:

$$f(x) = 1 \quad 0 \leq x \leq 1 \quad (3.5)$$

O gráfico a seguir mostra a **PDF** da uniforme padrão.

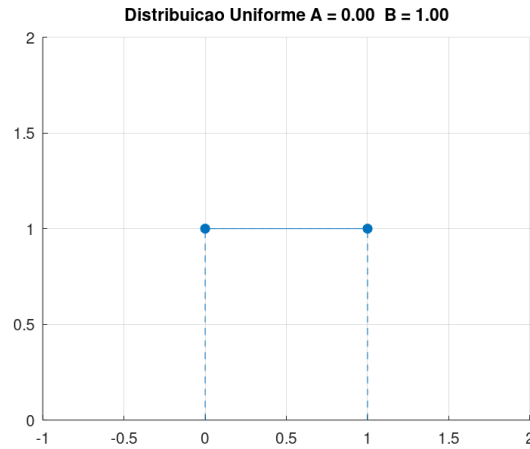


Figura 3.5: Função Percentil de Probabilidade da Normal Padrão

### Densidade Acumulada

A **Densidade Acumulada** para uma distribuição normal padrão, é simplesmente:

$$F(x) = x \quad 0 \leq x \leq 1 \quad (3.6)$$

O gráfico a seguir apresenta a curva da **CDF** para a normal padrão.

### Função Percentil

A fórmula da **Função Percentil** para uma distribuição uniforme padrão é bem definida, e é expressa por:

$$G(p) = p \quad 0 \leq p \leq 1 \quad (3.7)$$

O gráfico da **PPF** da uniforme padrão é apresentado a seguir:

**3.2.3 T-de-Student****3.2.4 F de Fisher - Snedecor****3.2.5 Qui - Quadrado****3.2.6 Exponencial****3.2.7 Weidbull****3.2.8 Geométrica****3.2.9 Hipergeométrica****3.2.10 Gama****3.2.11 Beta****3.2.12 Bernoulli****3.2.13 Binomial**

A **Distribuição Binomial** é um tipo de distribuição discreta, e uma decorrência dos ensaios de Bernoulli, quando o número de eventos *sucesso* é maior do que 1.

**Densidade de Probabilidade**

O cálculo referente a função **Densidade de Probabilidade** é dado pela função:

$$f(x; p, n) = \binom{n}{x} (p)^x (1 - p)^{n-x} \quad (3.8)$$

Em que

- $x$  é o número de vezes que o meu sucesso deve ocorrer, na ocasião  $x$  é um número inteiro positivo, isto é,  $x = 0, 1, 2, \dots$ ;
- $p$  é a probabilidade do sucesso ocorrer uma única vez;
- $n$  quantidade de eventos avaliados.

Sendo ainda o termo  $\binom{n}{x}$  a **Combinação**  $C(n, x)$ , calculada por:

$$C(n, x) = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

**Densidade Acumulada**

**Função Percentil**

### **3.2.14 Binomial - Negativa**

### **3.2.15 Poisson**

**Densidade de Probabilidade**

A **Distribuição de Poisson**, é um tipo de distribuição discreta que tem como função de probabilidade a seguinte equação

$$f(x, \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (3.9)$$

Em que

- $x$  é o número de ocorrências no estudo em questão, sendo este ainda um número inteiro não negativo, isto é,  $x = 0, 1, 2, \dots$ ;
- $\lambda$  é o número esperado (médio) de ocorrências no intervalo de estudo.

**Densidade Acumulada**

**Função Percentil**

### **3.2.16 Pareto**

# Parte II

## Testes Amostrais



## Capítulo 4

# Dois Grupos Independentes e Paramétricos

4.1 Intervalo e limite de confiança

4.2  $t$  de Student

4.3 Comparação entre 2 proporções





## Capítulo 5

# Dois Grupos Independentes e Não-Paramétricos

### 5.1 Qui-Quadrado

### 5.2 U de Mann Whitney

O teste de **U de Mann Whitney**, também conhecido como **Soma do Posto de Wilcoxon** é utilizado na comparação de dois grupos amostrais que tenham preferencialmente o mesmo tamanho.

O método funciona com os seguintes passos:

1. Coloca-se em ordem crescente todos os dados;
2. Calcula-se o **posto** referente a cada um dos valores;
3. Atribui-se este posto a cada um dos valores na amostra original;
4. Soma-se o posto de cada uma das duas amostras;
5. Calcula-se o valor  $U_1$  e  $U_2$ , e toma-se  $U = \min(U_1, U_2)$ . Define-se as seguintes equações (5.1) e (5.2) para o cálculo de  $U_1$  e  $U_2$ :

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad (5.1)$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \quad (5.2)$$

Caso a quantidade de valores coletados seja menor que 20, isto é, a soma de  $n_1$  e  $n_2$  sejam menores que 20, deve ser feito o comparativo do valor de  $U_{calculado}$  com o valor de  $U_{tabelado}$ , consultar a tabela **Valores Críticos U de Mann-Whitney**<sup>1</sup>.

Se a população for maior que 20, é necessário usar a **tabela z-normal**; nesta ocasião é efetuado mais um passo, que é o cálculo de  $z$ .

6. O calculo de  $z$  é dado por:

$$z = \frac{U - \mu_R}{\sigma_R} \quad (5.3)$$

---

<sup>1</sup>Tabela de Mann Whitney

em que

$$\mu_R = \frac{n_1 \cdot n_2}{2} \quad \sigma_R = \sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 + 1)}{12}}$$

Vamos resolver um exemplo, para que fique mais clara a aplicação do método.

---

**Exemplo 2.** Na investigação da eficiência de um novo remédio para asma, um grupo de 10 pacientes aleatórios são submetidos ao teste, sendo metade utilizando o novo remédio e a outra parte um placebo. Após uma semana os mesmos são questionados sobre a quantidade de crises que tiveram durante o período, os dados são apresentados na sequência.

<i>Placebo</i>	<i>Novo Remédio</i>
7	3
5	6
6	4
4	2
12	1

Tome um nível de 5% de significancia para o teste e as seguintes hipóteses nula e alternativa

$H_0$ : A duas populações são iguais

$H_1$ : A duas populações não são iguais.

**Resolução** Vamos tomar como **Pl** a coluna do **Placebo** e **NR** a coluna do **Novo Remédio**, então  $n_{Pl} = 5$  e  $n_{NR} = 5$ ; seguindo o passo a passo do método, iremos primeiro colocar todos os dados em ordem crescente, então fazemos:

**Passo 1** Colocando todos os dados em ordem crescente

# ordem	1	2	3	4	4	5	6	6	7	12
---------	---	---	---	---	---	---	---	---	---	----

**Passo 2** Deve ser calculado o posto de cada valor; o posto de uma amostra é dado de acordo com a posição na qual os dados de mesmo valor estão localizados na sequência crescente e a quantidade dos mesmos. Por exemplo, na ocasião o primeiro valor repetido é o número 4, o mesmo está localizado na posição 4 e 5 (sendo então duas repetições) da lista ordenada, então o posto do valor 4 será

$$posto_4 = \frac{4 + 5}{2} = 4.5$$

o mesmo procedimento é feito para o valor 6, que se encontra na posição 7 e 8, logo:

$$posto_6 = \frac{7 + 8}{2} = 7.5$$

os demais valores irão assumir os postos de suas posições, sendo assim:

# ordem	1	2	3	4	4	5	6	6	7	12
# postos	1	2	3	4.5	4.5	6	7.5	7.5	9	10

**Passo 3** Agora deve-se atribuir o valor dos postos encontrados, em cada uma das amostras originais

Placebo	Posto Pl	Novo Remédio	Posto NR
7	9	3	3
5	6	6	7.5
6	7.5	4	4.5
4	4.5	2	2
12	10	1	1

**Passo 4** Agora somaremos o posto de cada uma das amostras

$$R_{Pl} = 9 + 6 + 7.5 + 4.5 + 10 = 37 R_{NR} = 3 + 7.5 + 4.5 + 2 + 1 = 18$$

**Passo 5** Iremos calcular o valor de  $U$ , o que segue:

Primeiro  $U_{Pl}$

$$U_{Pl} = n_{Pl} \cdot n_{NR} + \frac{n_{Pl}(n_{Pl} + 1)}{2} - R_{Pl} \quad \therefore$$

$$U_{Pl} = 5 \cdot 5 + \frac{5(5 + 1)}{2} - 37 \Rightarrow U_{Pl} = 3$$

e agora  $U_{NR}$

$$U_{NR} = n_{Pl} \cdot n_{NR} + \frac{n_{NR}(n_{NR} + 1)}{2} - R_{NR} \quad \therefore$$

$$U_{NR} = 5 \cdot 5 + \frac{5(5 + 1)}{2} - 18 \Rightarrow U_{NR} = 22$$

Com ambos os valores calculados, tomaremos o menor, sendo assim  $U = 3$ , como a amostra só tem 10 valores, podemos então olhar a tabela de valor crítico  $U$  de Mann Whitney, uma parte da mesma é apresentada na figura a seguir

n <sub>2</sub>	α	n <sub>1</sub>																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42

Figura 5.1: Parte da Tabela de Valores Críticos de  $U$

Como nosso  $n_1 = 5$ ,  $n_2 = 5$  e  $\alpha = 5\%$ , temos  $U_{\text{tabelado}} = 2$ ; sendo o  $U$  calculado maior que o tabelado,  $2 < 3$ , então a hipótese nula é aceita.

**OBS:** O exercício foi retirado e adaptado do site Mann-Whitney

Para automatizar o problema foi criada uma função em *Octave* na qual é apresentada na sequência

```

function testeU_MannWhitney(A,B)

display('Dados fornecidos')
display(A)
display(B)

nA = length(A);    %quantidade de observacoes em A
nB = length(B);    %quantidade de observacoes em B

n = nA+nB;         %quantidade de observacoes totais

C = [A,B];         %vetor auxiliar
C_cres = sort(C);  %vetor auxiliar em ordem crescente

%Pesos em A
for k=1:nA
    mA = find(C_cres == A(k));
    pesoA(k) = sum(mA)/length(mA);
end

RA = sum(pesoA);

%Pesos em B
for k=1:nB
    mB = find(C_cres == B(k));
    pesoB(k) = sum(mB)/length(mB);
end

RB = sum(pesoB);

for k=1:nA
    if k == 1
        fprintf('Valor A          rankA\n')
    end
    fprintf('%7.2f      %10.2f\n',A(k),pesoA(k))
    if k==nA
        fprintf('nA = %4.2f      RA = %5.2f\n\n',nA,RA)
    end
end

for k=1:nB
    if k == 1
        fprintf('Valor B          rankB\n')
    end
    fprintf('%7.2f      %10.2f\n',B(k),pesoB(k))
    if k==nB
        fprintf('nB = %4.2f      RB = %5.2f\n\n',nB,RB)
    end
end

%Estatistica para o teste de Mann Whitney
UA = nA*nB + 0.5*(nA*(nA+1))-RA;
UB = nA*nB + 0.5*(nB*(nB+1))-RB;

```

```

fprintf('UA = %.2f    UB = %.2f\n',UA,UB)
U = min(UA,UB);

%Para n>20 usa-se a tabela da distribuicao normal
if n>20
    display('Use a Tabela normal')
    mu_r = nA*nB/2;
    sig_r = sqrt((nA*nB)*(nA+nB+1)/12);
    z = (U-mu_r)/sig_r

%Para n<=20 usa-se a tabela de Valores Criticos de Mann-Whitney
else
    display('Use a Tabela de Mann-Whitney')
    fprintf('Sendo o valor calculado de U = %.2f\n',U)
end

```

Para o nosso exemplo então podemos definir  $P1 = [7 \ 5 \ 6 \ 4 \ 12]$ ,  $NR = [3 \ 6 \ 4 \ 2 \ 1]$  e usar o comando `testeU_MannWhitney(P1,NR)`, o resultado obtido é apresentado na sequência

```

## Dados fornecidos
## A =
##      7      5      6      4     12
##
## B =
##      3      6      4      2      1
##
## Valor A          rankA
##      7.00          9.00
##      5.00          6.00
##      6.00          7.50
##      4.00          4.50
##     12.00         10.00
## nA = 5.00      RA = 37.00
##
## Valor B          rankB
##      3.00          3.00
##      6.00          7.50
##      4.00          4.50
##      2.00          2.00
##      1.00          1.00
## nB = 5.00      RB = 18.00
##
## UA = 3.00    UB = 22.00
## Use a Tabela de Mann-Whitney
## Sendo o valor calculado de U = 3.00

```

## 5.3 Prova de Fischer



## Capítulo 6

# Dois grupos Pareados e Paramétricos

### 6.1 Teste de t-Student pareado





## Capítulo 7

# Dois grupos Pareados e Não - Paramétricos

### 7.1 Prova de MacNemar

### 7.2 Prova de Wilcoxon



## Capítulo 8

# Três ou mais grupos Independentes e Paramétricos

### 8.1 Teste de Tuckey

### 8.2 ANOVA de 1 ou 2 vias



## Parte III

# Testes de Normalidade



## Capítulo 9

# Testes de Normalidade

9.1 Shapiro-Wilk

9.2 Kolmogorov - Smirnov

9.3 Anderson - Darling

9.4 Cramer Von-Mises





## Parte IV

### ACA



## Capítulo 10

# Análise de Concordância de Atributos



## Capítulo 11

# Estática Kappa

### 11.1 Teste Kappa de Cohen

### 11.2 Teste Kappa de Fleiss