



# BIOESTATÍSTICA

Marcos Roberto Queiroga

# Caros alunos

Esse ebook é um pdf interativo. Para conseguir acessar todos os seus recursos, é recomendada a utilização do programa *Adobe Reader 11*.

Caso não tenha o programa instalado em seu computador, segue o link para download:

<http://get.adobe.com/br/reader/>

Para conseguir acessar os outros materiais como vídeos e sites, é necessário também a conexão com a internet.

O menu interativo leva-os aos diversos capítulos desse ebook, enquanto a barra superior ou inferior pode lhe redirecionar ao índice ou às páginas anteriores e posteriores.

Nesse pdf, o professor da disciplina, através de textos próprios ou de outros autores, tece comentários, disponibiliza links, vídeos e outros materiais que complementarão o seu estudo.

Para acessar esse material e utilizar o arquivo de maneira completa, explore seus elementos, clicando em botões como flechas, linhas, caixas de texto, círculos, palavras em destaque e descubra, através dessa interação, que o conhecimento está disponível nas mais diversas ferramentas.

**Boa leitura!**



# ÍNDICE

---



# APRESENTAÇÃO

Olá, alunos

Este e-book é uma ferramenta de suporte à realização de atividades da disciplina de Bioestatística. O objetivo será o de fornecer informações que possibilitem compreender e utilizar os recursos estatísticos básicos apresentados no desenvolvimento de projetos de pesquisa e leituras científicas. Além disso, procuramos oferecer subsídios para que o aluno(a) possa dar continuidade ao processo de aprendizagem em outros recursos em bioestatística. Embora quando se fala em estatística a maioria das pessoas se apavora, devemos ter em mente que as informações que consumimos no dia a dia estão recheadas de dados estatísticos. Para esclarecer essa afirmação, observem os dados da pesquisa Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico-VIGITEL do Ministério da Saúde (Brasil, 2013). O resultado estatístico de um determinado assunto tem por finalidade permitir a compreensão de um dado fenômeno.

O que vamos abordar nesta disciplina é um resumo dos recursos estatísticos, iniciando com fundamentos gerais sobre o método estatístico na Unidade 1. Essa unidade visa familiarizar o aluno(a) com alguns termos estatísticos bem como fornecer bases para os conteúdos das próximas unidades.

Na Unidade 2, apresentamos os recursos estatísticos descritivos que não permitem estabelecer nenhum grau de associação com outras variáveis ou grupos. A estatística descritiva é muito útil para apresentação dos dados e nos dá uma ideia das características da amostra que está sendo investigada (média e desvio padrão da idade, estatura, %G, IMC, informações sobre o nível socioeconômico, estado civil etc).

A Unidade 3 conduz o aluno(a) a entrar no mundo indutivo dos métodos estatísticos. Nessa unidade pretende-se apresentar comparações para se determinar igualdade ou diferenças entre grupos/valores. Considerando a grande quantidade de conteúdo e possibilidades de análises estatísticas existentes, bem como o tempo para desenvolver os conteúdos, vamos nos limitar a realizar comparações até dois grupos (dependente e independente).

Será fornecido para cada método um banco de dados que nos permitirá seguir relacionando teoria e prática durante o processo ensino-aprendizagem da disciplina. O que é importante para o adequado aprendizado é o conhecimento mínimo do uso do Microsoft Office Excel, que é um editor de planilhas produzido pela Microsoft para computadores que utilizam o sistema operacional Microsoft Windows. Existem outros programas no mercado, porém o acesso seria mais difícil para a maioria dos alunos(as) deste curso.

O tema do módulo é demasiadamente abrangente e dinâmico e, por isso, aguardamos contatos de vocês com o professor e os tutores no fórum tira-dúvidas.



# FUNDAMENTOS GERAIS

## 1.1. DEFINIÇÃO DE ESTATÍSTICA

A estatística compreende um conjunto de técnicas metódicas por meio das quais se pode, de modo apropriado, coletar (uniformizar a coleta), organizar, apresentar (descrever), analisar e interpretar os dados de uma ou mais variáveis (Barros; Reis, 2003).

Esse conjunto de técnicas pode ser subdividido em dois tipos principais, **Estatística Descritiva** e **Estatística Inferencial (Indutiva)**.

### 1.1.1. ESTATÍSTICA DESCRITIVA:

compreende procedimentos da coleta de dados, tabulação e descrição das informações (média de idade, peso corporal e altura, quantos homens e mulheres, tempo de treinamento etc). Visa

sumariar e descrever os atributos mais proeminentes aos dados. Esse assunto será tema da nossa unidade 2.

### 1.1.2. ESTATÍSTICA INFERENCIAL (INDUTIVA):

Compreende os métodos de análise dos dados, a relação, associação e diferença entre as variáveis (relação entre massa corporal e gordura; diferença entre força muscular de treinados e não treinados etc) (Barros; Reis, 2003). Em outras palavras, é um conjunto de métodos estatísticos que visam caracterizar (ou inferir sobre) uma população a partir de uma parte dela (a amostra). Esse assunto será tema da nossa unidade 3.

O aspecto mais básico e vital de um tratamento estatístico de dados passa



pela necessidade de se observar que tipo de dados que o objeto de estudo permite que sejam coletados.

Vale a pena lembrar que um bom planejamento para a coleta e análise de dados poupa muitas horas de trabalho para se saber como se poderá tratar os dados ou, ainda, pode-se poupar a necessidade de retomada de dados (coletar novamente). Dessa forma, é importante destacar os parâmetros que devem ser observados no planejamento de um estudo.

### **1.2. PARÂMETROS BÁSICOS NO PLANEJAMENTO DE UM ESTUDO**

Dentre os diversos testes e métodos estatísticos que se apresentam no contexto de um trabalho de pesquisa, é comum verificar um desespero inicial do interessado para identificar os métodos

que são ou não aplicáveis a sua situação. Para se realizar a escolha adequada, é importante considerar e conhecer alguns parâmetros básicos no planejamento de um estudo, tais como:

*Dê duplo clique para abrir o arquivo*

**1.2.1. Número de amostras (grupos):** refere-se ao número de grupos que fazem parte do estudo (um ou mais). Entende-se por grupo, sexo, idade, equipes etc.

**1.2.2. Relações entre amostras:** se as medidas forem repetidas, após um intervalo de tempo, no mesmo grupo (antes e depois) as amostras são “pareadas” ou “dependentes” (você comparado a você mesmo); se as medidas forem realizadas em grupos distintos, não no mesmo, as amostras são “independentes” ou “não pareadas” (você comparado com outro; uma equipe versus outra).

**1.2.3. Escalas de medida:** cada tipo de variável, seja ela qualitativa



ou quantitativa, pode ser mensurada em diferentes escalas que definirão se o modelo estatístico se aplica aos métodos estatísticos paramétricos ou não paramétricos (Nominal, Ordinal, Intervalar e de Razão) (Barros; Reis, 2003).

**Nominal:** não existe ordenação dentre as categorias. Exemplos: sexo, cor dos olhos, etnia, estado civil, fumante/não fumante, doente/sadio.

**Ordinal:** existe uma ordenação entre as categorias. Exemplos: escolaridade (1º, 2º, 3º graus), estágio da doença (inicial, intermediário, terminal), mês de observação (janeiro, fevereiro, ..., dezembro), nível social (A, B, C, ..., alto, médio, baixo), faixa etária.

**Intervalar:** o valor zero é arbitrário e não representa a ausência da característica mensurada (temperatura, conhecimento). Nessa escala, a diferença entre duas

medidas permite conhecer quando um caso é menor ou maior do que o outro.

**Razão:** é semelhante à escala intervalar, porém, o valor zero representa o valor mínimo ou a ausência de característica mensurada (massa corporal, % de gordura, altura, distâncias, tempo).

**1.2.4. Tipo de variável:** está associada à escala de medida.

### **VARIÁVEIS QUANTITATIVAS**

**(numéricas):** atribuem gradações (escala intervalar e de razão). Apresentam valores numéricos que fazem sentido. Podem ser variáveis contínuas ou discretas.

**Variáveis contínuas:** são variáveis que teoricamente podem assumir qualquer valor. Usualmente devem ser medidas por meio de algum instrumento. Exemplos: peso corporal (balança), altura (estadiômetro), tempo (relógio), pressão



arterial, idade, distância, força muscular, tempo etc.

**Variáveis discretas:** é uma grandeza que não pode ser subdividida, usualmente representada por números inteiros. Geralmente é o resultado de contagens. Exemplos: número de filhos, número de cigarros fumados por dia, número de pontos (gols, sets), número de livros lidos em um ano, número de alunos em uma sala de aula, frequência cardíaca etc.

**VARIÁVEIS QUALITATIVAS (categóricas):** são as características que não possuem valores quantitativos, mas, ao contrário, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos (escala nominal e ordinal).

**1.2.5. Distribuição:** também conhecida como distribuição gaussiana, está associada à distribuição dos dados

(distribuição normal/simetria ou não-normal/assimetria). Existem testes estatísticos para se determinar a distribuição (normalidade) dos dados. Isso é realizado antes de qualquer análise com dados numéricos. Vamos discutir esse assunto na unidade 3.

### 1.3. ESTATÍSTICA PARAMÉTRICA E NÃO PARAMÉTRICA

Os critérios para utilizar a estatística paramétrica ou não paramétrica fundamentam-se no tipo da variável e na distribuição (normalidade) dos dados. Estatísticas paramétricas normalmente são provenientes de dados quantitativos (numéricos), porém devem atender ao pressuposto de normalidade. Estatísticas não paramétricas normalmente são provenientes de dados qualitativos





(categóricos) e apresentam variáveis mensuradas na escala ordinal e/ou nominal.

#### 1.4. POPULAÇÃO E AMOSTRA

**População:** qualquer conjunto de indivíduos que apresentam pelo menos uma característica em comum (escolares de um município; árbitros de futebol FIFA etc).

**Amostra:** um grupo de sujeitos selecionados de uma população (amostra de escolares; de árbitros). Para que as informações de uma pesquisa obtidas de uma amostra possam ser generalizadas para a população, a seleção da amostra é muito importante. Para tanto, existem formas de selecionar uma amostra que determinarão as previsões ou achados observados em uma escala maior.

Dessa forma, uma amostra pode ser selecionada de forma **não casual**, no qual o pesquisador escolhe de acordo com que lhe convém (uma equipe de futebol, uma turma escolar) e, **casual**, onde é realizada uma seleção (sorteio) para formar o grupo de sujeitos que farão parte da pesquisa (pesquisa eleitoral, censo populacional, sorteio de escolares por sala etc).

#### 1.5. VARIÁVEIS DEPENDENTES E VARIÁVEIS INDEPENDENTES

Variáveis independentes são aquelas que são manipuladas enquanto que variáveis dependentes são apenas medidas ou registradas.

**Variável independente (VI):** É uma variável que está totalmente livre para variar e que não é afetada por outras variáveis (variável preditora).



Exemplos:

- Condição climática (VI)
- Consumo de lipídios (VI)
- Atividade física (VI)
- Escolaridade (VI)
- Preço das hortaliças (VD)
- Morte por doenças cardíacas (VD)
- Obesidade (VD)
- Nível educacional (VD)

**Variável dependente (VD):** é uma variável que não tem liberdade para variar, ou seja, depende dos efeitos de uma ou mais variáveis (variável critério).

Variável dependente (é o consequente) é alterada ou influenciada pela variável independente (é o antecedente). Exemplo:

**Dependente:** predisposição a problemas cardíacos;

**Independente:** sexo; consumo de lipídios; prática de exercícios etc.

## 1.6. VARIÁVEIS

As variáveis são aspectos, propriedades, características individuais ou fatores observáveis ou mensuráveis de um fenômeno. Em outras palavras, é a característica ou propriedade da população que está sendo medida (o que coletamos; dados). Exemplo: massa corporal (kg);

altura (m/cm); consumo de oxigênio (ml. kg.min/L.min); frequência cardíaca (bpm); resultados de testes motores (repetições/tempo/distância/massa), tempo de treinamento (anos, meses), sensação de esforço, classe econômica etc.

*Em síntese,*

*As variáveis são os dados que coletamos, que podem ser de natureza **quantitativa** (números) ou **qualitativa** (categorias). De acordo com o tipo de pesquisa (experimental), as variáveis podem demonstrar algum grau de **dependência (variável dependente)** ou **independência (variável independente)**. Para analisar os dados com os **métodos paramétricos**, as variáveis quantitativas devem seguir uma **distribuição normal** (simétrica). Distribuição assimétrica e variáveis categóricas (qualitativo) a análise deve ser realizada por meio de métodos **não paramétricos**.*



# ESTATÍSTICA DESCRITIVA (DESCREVENDO OS DADOS)

É um método estatístico para organização e síntese dos dados obtidos. Consiste basicamente na organização e descrição dos dados, na identificação de valores que traduzem o elemento típico e na quantificação da variabilidade presente nos dados. Todos esses cálculos podem ser realizados a partir de programas estatísticos, incluindo o Excel. O arquivo anexo ao e-book fornece os fundamentos da estatística descritiva bem como um passo-a-passo para a realização de cálculos para a estimativa de medidas de tendência central e de dispersão de dados.

*Dê duplo clique para abrir o arquivo*

As principais ou mais empregadas medidas de **tendência central e dispersão** são:

## 2.1. MEDIDAS DE TENDÊNCIA CENTRAL (POSIÇÃO):

Para extrair qualquer informação de um conjunto de dados quantitativos é importante utilizar medidas de posição para sintetizar as observações.

**2.1.1. Média aritmética:** é a soma das medidas dividida pelo número de casos. Exemplo: Na amostra 35; 45; 32; 40 anos a média é:  $35+45+32+40 = 152/4 = 38$  anos.

**2.1.2. Mediana:** é o valor que se encontra na posição central da série de dados, depois de ordenada a amostra (crescente/decrescente). Se a amostra apresentar dimensão ímpar, coincide com a observação central. Exemplo: 32; 35; **39**; 40; 45 anos a mediana é = **39** anos. Se a amostra apresentar dimensão par, a



mediana toma o valor da média das duas observações mais centrais. Exemplo: 35; 32; 40; 45 anos  $(32+40/2) = 36$  anos.

**2.1.3. Moda:** é aquele que ocorre com maior frequência em uma série de dados. Exemplo: Na amostra 32; 35; 35; 39; 40; 45 anos a moda é: = **35 anos** (esta idade apareceu duas vezes enquanto as outras apenas uma vez no conjunto de dados).

## 2.2. MEDIDAS DE VARIABILIDADE OU DISPERSÃO:

É determinação da variabilidade ou dispersão dos dados, relativamente à medida de localização do centro da amostra (média).

**2.2.1. Amplitude:** é a diferença entre o valor máximo e o mínimo. Exemplo: Na amostra 35; 45; 32; 39; 40; 39 anos a amplitude: 45 (maior) – 32 (menor) = **13 anos**.

**Máximo:** é o valor mais elevado da amostra: **45 anos**.

**Mínimo:** é o valor mais reduzido da amostra: **32 anos**.

**2.2.2. Variância:** é a média dos quadrados dos desvios das observações em relação à média da amostra. A soma dos desvios entre cada valor e a média sempre será zero (0). Por isso, para o cálculo da variância, a soma dos desvios é elevada ao quadrado.

Equação simplificada da variância: define-se a variância como a medida que se obtém somando os quadrados dos desvios das observações da amostra, relativamente à sua média, e dividindo pelo número de observações da amostra menos um.

$$\frac{\sum (x_i - \text{média})^2}{n - 1}$$



Exemplo: Na amostra 35; 45; 32; 40 anos ( $152/4 = 38$  anos/média) a variância é o produto de cada variável, ao quadrado, menos a média do grupo: ex.:  $(35-38)^2+(45-38)^2+(32-38)^2+(40-38)^2 = (-3)^2+(7)^2+(-6)^2+(2)^2 = 9+49+36+4 = 98 / 3 (4-1) = 32,66$  ou **32,7**.

2.2.3. **Desvio padrão:** O desvio padrão é a raiz quadrada da variância ( $\sqrt{32,7} = 5,7$ ) e quanto maior for, maior será a dispersão dos dados.

Equação do desvio padrão: para obter o desvio padrão de um grupo de dados, calcula-se a raiz quadrada da variância.

$$\sqrt{\frac{\sum (x_i - \text{média})^2}{n - 1}}$$

É uma medida de dispersão e vem na mesma unidade das observações e significa que 34,13% das variáveis encontra-se acima (+34,13%) e abaixo (-34,13%) do valor médio (68,26% total). Utilizando o

exemplo acima (média de 38 anos e desvio padrão de 5,7 anos), pode-se afirmar que a amostra possui idade variando entre 32,3 anos ( $38 - 5,7$ ) a 43,7 anos ( $38 + 5,7$ ). Os dados são apresentados como  $38 \pm 5,7$  anos (o sinal  $\pm$  refere-se ao desvio padrão).

A **média** e o **desvio** padrão são os recursos estatísticos descritivos (para variáveis quantitativas com distribuição normal) mais empregados em estudos científicos. Média aritmética amostral ( $\bar{x}$ )  $\pm$  desvio padrão (dp).



# ESTATÍSTICA INFERENCIAL

Dê duplo clique para abrir o arquivo

Os procedimentos de análise envolvendo a estatística inferencial (**Indutiva**) são quase ilimitados e com frequência surgem métodos de análise de dados mais robustos. Para o uso de um método estatístico indutivo (procurar verificar se há diferenças entre variáveis/grupos), é preciso respeitar e identificar os parâmetros básicos para o planejamento de um estudo apresentados na unidade 1. Em outras palavras, deve-se ter em mente quantos grupos possui o estudo, se existe alguma relação entre os grupos (pareado/independente), se as variáveis são quantitativas ou qualitativas e, por fim, se a distribuição dos dados é normal (apresentam simetria). Esses procedimentos devem nos indicar se a análise estatística será **paramétrica** ou **não paramétrica**.

## 3.1. INSTALANDO E CONFIGURANDO SUPLEMENTOS ESTATÍSTICOS PARA O EXCEL

Considerando que estamos utilizando o programa Excel para realização dos nossos cálculos, é possível que alguns pacotes do Microsoft Office gravados em seus computadores não se encontrem habilitados/configurados para todas as análises necessárias.

Assim, sugerimos acessar o link para habilitar o suplemento de análise estatística no Excel, caso ainda não esteja configurado.

Esse suplemento não permite análises não paramétricas, o que nos leva a sugerir outro dispositivo. Nesse sentido, para as próximas análises (e também



para estatística paramétrica) será necessário à instalação de um software de estatística para realização das análises não paramétricas integradas ao Excel. O software é gratuito e poderá ser localizado e baixado no seguinte link:

*Ainda, sugere-se a visualização de um vídeo para auxiliar na instalação e configuração deste suplemento. Lembramos que esse último suplemento (ferramenta Action) seria suficiente para a realização de todos os testes (descritivos e indutivos) que vamos estudar na disciplina e outros que podem se aprofundar.*

### **3.2. TESTE DE NORMALIDADE DOS DADOS NUMÉRICOS**

Antes de decidirmos sobre qual o método estatístico utilizado, é importante testar a normalidade dos dados numéricos.

Interessante destacar que o resultado significativo, ou seja,  $P < 0,05$  indica que não há normalidade na distribuição dos dados. Nesse caso, a opção será utilizar um teste não paramétrico. A seguir, sugerimos conferir um vídeo demonstrativo sobre a realização de um teste de normalidade

### **3.3. SIGNIFICÂNCIA ESTATÍSTICA**

Para nosso objetivo, vamos tratar de casos (estudos) de natureza quantitativo e qualitativo com até dois grupos. Nesse sentido, vamos apresentar recursos estatísticos que analisam/comparam a média amostral (dados numéricos) ou os postos (dados categóricos) de até dois grupos, tanto grupos pareados (ele mesmo em dois momentos) ou como



independentes (dois grupos distintos). Se o aluno entender a comparação de dois grupos poderá se aventurar em comparações com três ou mais grupos.

Deste ponto em diante, quando procuramos verificar se há diferenças entre grupos/variáveis, precisamos ter alguma familiaridade com **nível de significância**. O papel da análise estatística é testar se os resultados obtidos têm ou não significância estatística, ou seja, se diferem ou não diferem (são iguais). Para isto, é necessário a definição das hipóteses nula e alternativa. Por exemplo: se conjecturarmos que "*meninas de 8 anos demonstram flexibilidade de quadril igual (não diferem) aos meninos com mesma idade*" estamos definindo uma hipótese de pesquisa [**Hipótese nula (H<sub>0</sub>)** é a hipótese que assumimos como verdade para a construção do teste]. Por sua vez, se afirmarmos que "*as meninas*

*de 8 anos apresentam diferenças (maior/menor) na flexibilidade de quadril em relação aos meninos de mesma idade*", estamos definindo uma hipótese de pesquisa identificada como hipótese alternativa [**Hipótese alternativa (H<sub>1</sub>)** é o que consideramos caso a hipótese nula não tenha evidência estatística que a defenda]. De acordo com o resultado do teste estatístico, corremos o risco de cometer erro do tipo I ou II.

**Erro tipo I:** Rejeitar H<sub>0</sub> quando de fato H<sub>0</sub> é verdadeiro.

**Erro tipo II:** Não rejeitamos H<sub>0</sub> quando de fato H<sub>0</sub> é falsa.

Para as afirmações realizadas no texto, poderíamos exemplificar os erros da seguinte forma:

**Erro tipo I:** Não aceitar que as meninas diferem na flexibilidade quando na verdade diferem (mais ou menos).





**Erro tipo II:** Aceitar que as meninas diferem na flexibilidade quando na verdade não diferem (mais ou menos).

Especificamente, o nível-p representa a probabilidade de erro envolvida em aceitar o resultado observado como válido, isto é, como “representativo da população”. Por exemplo, um nível-p de 0,05 (1/20) indica que há 5% de probabilidade de que a relação entre as variáveis, encontrada na amostra, seja um “acaso feliz”. Em outras palavras, assumindo que não haja relação entre aquelas variáveis na população, e o experimento de interesse seja repetido várias vezes, poderia se esperar que em aproximadamente 20 realizações do experimento haveria apenas uma em que a relação entre as variáveis em questão seria igual ou mais forte do que a que foi observada naquela amostra anterior. Em muitas áreas de pesquisa, o nível-p de 0,05 é costumeiramente tratado como um “limite aceitável” de erro.

### 3.4. MÉTODO ESTATÍSTICO PARAMÉTRICO PARA DUAS AMOSTRAS

Já demonstramos que uma característica bem definida para o uso da **estatística paramétrica** é que seja proveniente de dados quantitativos (numéricos), porém, a distribuição dos dados deve atender ao pressuposto de normalidade.

Como estabelecemos para a disciplina comparação de até dois grupos, vamos exemplificar a análise para grupos pareados (dependente) e não pareados (independente). Na estatística paramétrica, o *teste t de student* é o padrão para comparação de dois grupos, ou seja, *teste t* para amostras *pareadas (mesmo grupo)* e para amostras *independentes (dois grupos)*.

3.4.1. Comparações entre grupos pareados (Teste T de Student)



Dê duplo clique para abrir o arquivo

.....[planilha grupo pareado](#)), queremos verificar se um grupo de meninas difere em relação à massa corporal (MC) que foi mensurada em dois momentos (pré e pós). Vamos aceitar diferença se o valor do teste estatístico (calculado) for menor do que 5% ( $P < 0,05$ ). O [arquivo anexo](#) fornece um passo-a-passo para a realização do teste t.

Dê duplo clique para abrir o arquivo

*Para maiores informações sobre Teste t pareado, sugerimos a consulta deste vídeo.*

#### 3.4.2. Comparações entre grupos independentes (Teste T de Student)

Dê duplo clique para abrir o arquivo

.....[planilha grupo independente](#)), queremos verificar se meninas e meninos diferem em relação à massa corporal (MC). Vamos aceitar

diferença se o valor do teste estatístico (calculado) for menor do que 5% ( $P < 0,05$ ). O [arquivo anexo](#) fornece um passo-a-passo para a realização do teste t.

Dê duplo clique para abrir o arquivo

*Para maiores informações sobre Teste t independente, sugerimos a consulta deste vídeo.*

### 3.5. MÉTODOS ESTATÍSTICOS NÃO PARAMÉTRICOS PARA DUAS AMOSTRAS

Estatísticas não paramétricas (livre distribuição) são provenientes de dados qualitativos de variáveis mensuradas na escala ordinal e/ou nominal. Outra possibilidade em empregar os métodos não paramétricos é quando as variáveis numéricas apresentam distribuição assimétrica.



Uma vez instalado o suplemento para análise no Excel (ferramenta Action), é preciso acessar uma planilha para analisarmos um grupo de dados. O quadro abaixo sintetiza testes paramétricos e não paramétricos que são correspondentes em suas categorias de análise.

#### Quadro. Testes paramétricos e não paramétricos correspondentes

TESTES PARAMÉTRICOS	TESTES NÃO-PARAMÉTRICOS
Coefficiente de Pearson para Correlação	Coefficiente de Spearman para Correlação
Teste-t 1 Amostra	Teste de Wilcoxon 1 Amostra
<u>Teste-t Pareado</u>	<u>Teste de Wilcoxon Pareado</u>
<u>Teste-t 2 Amostras</u>	<u>Teste de Wilcoxon-U de Mann-Whitney Amostras Independentes</u>
ANOVA	Teste de Kruskal-Wallis
ANOVA experimento fatorial em blocos	Teste de Friedman



### 3.5.1. Comparações entre grupos pareados (Teste de Wilcoxon Pareado)

O teste de Wilcoxon pareado é utilizado para verificar se as medidas de posição de duas amostras (dependentes) são iguais.

Antes de analisarmos, é preciso testar a normalidade dos dados. A [planilha em anexo](#), será utilizada para exemplificarmos o uso do teste de Wilcoxon pareado.

Nosso objetivo será o de verificar se um grupo de idosos difere ( $H_0$ ) ou não ( $H_1$ ) para o teste de agilidade após um período de prática de atividades físicas. Primeiramente precisamos testar a normalidade dos dados para este grupo

de dados (agilidade). O resultado do teste de normalidade indicará qual grupo de testes estatísticos para grupos pareados (paramétricos ou não paramétricos) serão utilizados. Como o exemplo (agilidade) não possui normalidade na distribuição dos dados será selecionado o teste de Wilcoxon para amostras pareadas. O [arquivo anexo](#) fornece um passo-a-passo para a realização do teste de normalidade e do teste Wilcoxon para amostras pareadas.

### 3.5.1. Comparações entre grupos independentes (Teste de Wilcoxon-U de Mann-Whitney - amostras independentes)

Para testar dois grupos independentes com variáveis qualitativas (categóricas) ou quando as variáveis numéricas apresentam distribuição assimétrica, utiliza-se o teste de Wilcoxon-U de Mann-Whitney. A [planilha em anexo](#), será

*Dê duplo clique para abrir o arquivo*

*Dê duplo clique para abrir o arquivo*

*Dê duplo clique para abrir o arquivo*



utilizada para exemplificarmos o uso do teste de Wilcoxon-U de Mann-Whitney para amostras independentes.

O objetivo será o de verificar se um grupo de homens idosos difere ( $H_0$ ) ou não ( $H_1$ ) de mulheres idosas para o desempenho no teste de dinamometria manual (força preensão manual em kg). Primeiramente precisamos testar a normalidade dos dados para este grupo de dados (dinamometria). Como o resultado (dinamometria) indica que não há normalidade na distribuição dos dados, será selecionado o teste de Wilcoxon para amostras independentes (U de Mann-Whitney). O [arquivo anexo](#) fornece um passo-a-passo para a realização do teste de Wilcoxon para amostras pareadas.

*Dê duplo clique para abrir o arquivo*



# REFERÊNCIAS

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Vigitel Brasil 2013: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico / Ministério da Saúde, Secretaria de Vigilância em Saúde. Brasília: Ministério da Saúde, 2014. 120p. Disponível em: <<https://biavati.files.wordpress.com/2014/05/vigitel-2013.pdf>>.

CALLEGARI-JACQUES, S.M. Bioestatística: Princípios e Aplicações. Porto Alegre: Artmed, 2006.

SIEGEL, S. Estatística não paramétrica para as ciências comportamentais. São Paulo: Makron, 1975.

BARROS, M.V.G. Análise de dados em atividade física e saúde. Demonstrando a utilização do SPSS. Londrina: Midiograf, 2003.

VIEIRA, S. Bioestatística. Rio de Janeiro: Campus, 1980.

Portal Action <<http://www.portalaction.com.br/>>