UNIDADE II

Estatística Inferencial

Dr. José Roberto Andrade do Nascimento Junior

Aula 01

Estatística Inferencial: Teste de Hipótese e Significância Estatística

Introdução

Pense no seguinte objetivo de uma pesquisa: comparar o nível de estresse de homens e mulheres que trabalham em uma empresa. Para se chegar à conclusão de que o nível de estresse é maior entre os homens ou entre as mulheres, é necessário comparar dois grupos extraídos da população e formular hipóteses a respeito da distribuição dos dados em comparação ao comportamento da população. Para isso, é preciso testar uma hipótese por meio de um teste estatístico, o qual, por sua vez, permite verificar se as hipóteses são verdadeiras ou falsas. A partir dos resultados do teste estatístico, as hipóteses são confirmadas ou rejeitadas e, consequentemente, sabe-se se existe alguma diferença no nível de estresse entre os homens e as mulheres. Este procedimento é fundamental para se chegar a alguma conclusão e para pautar as possíveis tomadas de decisão.

Ao final desta aula, você será capaz de:

- compreender o conceito e a utilização dos testes de hipóteses;
- reconhecer a significância estatística;
- identificar os erros do tipo I e II nos testes de hipóteses.

Hipóteses Nula e Alternativa

A estatística inferencial permite ao pesquisador estabelecer inferências sobre uma população a partir dos dados da amostra. Para isso, é necessário testar quanta evidência estes dados demonstram contra uma suposição específica em relação a uma determinada variável. O processo para confirmar se as hipóteses formuladas a respeito dos dados são verdadeiras ou falsas é chamado de teste de hipóteses ou teste de significância.

SAIBA MAIS

Quando se fala em teste de hipóteses, é fundamental compreender o conceito de hipótese de pesquisa. A hipóteses se refere à previsão de como duas variáveis podem se relacionar entre si, ou como uma variável pode se comportar de forma diferente entre dois grupos.

Fonte: Dancey, Reidy e Rowe (2017).

De acordo com Barros et al. (2012), o teste de hipóteses ocorre por meio dos procedimentos apontados no Quadro 1:

Estabelecimento das hipóteses nula (H0) e alternativa (H1) da pesquisa.
Coleta dos dados do fenômeno investigado na amostra selecionada.
Cálculo do teste estatístico para a H0.
Comparação do valor do teste estatístico com os valores de uma distribuição de probabilidade já conhecida.
Compreensão do valor da significância (p).

Quadro 1 - Procedimentos para a realização dos testes de hipóteses
Fonte: Barros et al. (2012).

A hipótese nula (H_0) corresponde a uma afirmação conservadora a respeito do fenômeno que está sendo investigado, isto é, se uma variável está sendo comparada entre dois grupos, a H_0 diz que não existe diferença entre os grupos em relação a esta variável (DANCEY; REIDY; ROWE, 2017). Ainda, na análise da relação entre duas variáveis independentes a H0 é de que não existe relação entre as variáveis. Vamos tomar como exemplo a comparação da média da idade entre homens e mulheres de uma empresa, na qual a H_0 será de que não existe diferença na idade entre os homens e as mulheres da empresa.

Já a hipótese alternativa (H_1) é uma hipótese estabelecida como alternativa para a H_0. Assim, a H_1 sempre será de que existe diferença entre os grupos comparados ou de que existe relação entre duas variáveis analisadas. Isso acontece quando existe evidência suficiente de que a H_0 não é verdadeira, a qual é rejeitada e a H_1 se torna a resposta para o fenômeno que está sendo pesquisado (BARROS et al., 2012).

Para exemplificar as duas hipóteses, vamos considerar uma situação hipotética que pretende comprar as notas de uma prova de homens e mulheres. As hipóteses do estudos ficariam da seguinte forma:

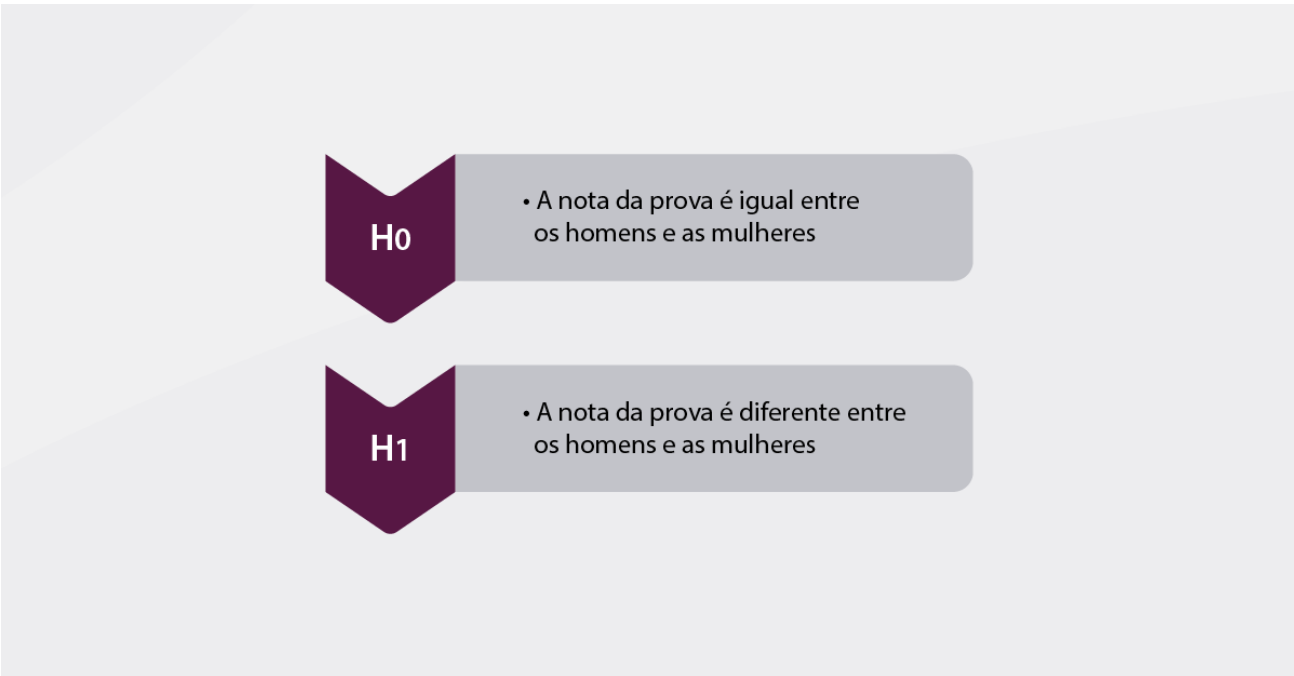

Figura 1 - H_0 e H_1 da situação hipotética
Fonte: Elaborada pelo autor.

Pode-se dizer que os testes de hipóteses podem ser considerados como uma competição entre as hipóteses: a hipótese da pesquisa (H_1 - existe diferença nas notas entre os homens e as mulheres) e a H_0 (não existe diferença na nota entre os grupos). Se o pesquisador concluir que a hipótese nula deve ser rejeitada, pode-se dizer que a probabilidade de obter os resultados encontrados se a hipótese nula fosse verdadeira é tão pequena que é mais viável aceitar a hipótese alternativa. Logo, na situação hipotética da Figura 1, se a hipótese nula for rejeitada, isso significa que os dados evidenciaram diferença na nota da prova entre os homens e as mulheres

O teste estatístico utilizado é que determinará a probabilidade de observar o padrão dos dados se a hipótese nula for verdadeira. Nesse tipo de situação nunca se sabe se existem diferenças entre os grupos nem a distância entre as médias dos grupos. Com isso, as diferenças podem ocorrer em qualquer direção, isto é, tanto os homens quanto as mulheres podem ter maior nota na prova. Nessa perspectiva, a utilização de um teste bicaudal é importante, uma vez que analisa qualquer posição da curva de probabilidade (BARROS et al., 2012).

Nível de Significância

Após saber que o teste de hipóteses é o processo estatístico para descobrir se a hipótese nula é verdadeira ou falsa, um ponto fundamental é compreender a probabilidade que deve ser calculada no teste de hipótese para que se chegue à conclusão de que a hipótese nula deve ser rejeitada ou aceita. Para isso, os dados coletados das amostras são analisados por meio de equações matemáticas que fornecem o valor do teste estatístico e a probabilidade de a hipótese nula ser aceita ou rejeitada.

Após a determinação do valor do teste utilizado, este é comparado com uma distribuição teoricamente conhecida na população, a qual permite identificar o valor entre $0 \ e \ 1$, que indica o valor exato da probabilidade de os resultados encontrados na amostra serem idênticas à distribuição da população (BARROS et al., 2012).

De acordo com Dancey, Reidy e Rowe (2017), não existe um valor definitivo para o cálculo dos testes de hipóteses, entretanto, existe um consenso na literatura científica que uma probabilidade de 5 é suficientemente pequena para servir de ponto de corte. Isto significa que, se a hipótese nula for verdadeira, a probabilidade de um determinado efeito é menor do que 5, ou seja, se a pesquisa for conduzida 20 vezes, somente uma vez uma diferença ou relacionamento tão grande quanto a que foi observada surgirá no caso de a hipótese nula for verdadeira. Essa probabilidade de corte associada aos testes estatísticos é denominada nível de significância e representada pelo alfa (α).

SAIBA MAIS

Alfa (α) representa o critério de significância fixado nos testes estatísticos e se refere à probabilidade utilizada como ponto de corte em um teste de hipótese. Se o resultado do teste for abaixo desse ponto de corte, é possível assumir que o resultado é improvável para que a hipótese alternativa seja aceita, logo, a hipótese nula será verdadeira.

Fonte: Dancey, Reidy e Rowe (2017).

Atualmente, os pesquisadores frequentemente têm utilizado os termos significativo e não significativo ao realizar testes de hipóteses. Isso significa que, quando a hipótese nula é falsa e que a probabilidade de obter um efeito seja menor do que 5, o resultado é considerado como significativo, enquanto que, nos casos em que a probabilidade for maior do que 5, o resultado é considerado não significativo.

Quando se pretende relatar o valor exato da probabilidade de um teste estatístico, é utilizado o valor de p, que é um valor entre $0 \ e \ 1$, e deve sempre estar associado aos testes realizados. Considerando um nível de significância de 5, se o valor de ρ for menor que $5 \ (p \ < \ 0,05)$, o resultado indica que a hipótese alternativa é verdadeira e a hipótese nula é falsa.

SAIBA MAIS

O valor de p de um determinado teste estatístico corresponde à probabilidade exata de se encontrar a diferença observada ou uma diferença ainda maior quando a hipótese nula é verdadeira.

Fonte: Dancey, Reidy e Rowe (2017).

Barros et al. (2012) apontam quatro considerações importantes a respeito do valor de p:

1. Na área da saúde, um nível de significância de $5 \ (p \ < \ 0,05)$ é suficiente para se rejeitar a hipótese nula e estabelecer que os resultados foram estatisticamente significativos a esse nível de significância.
2. Quanto menor o valor de p, maior a evidência para se rejeitar a hipótese nula.

3. Quando o valor de p é maior ou igual a $0,05$, não há evidência suficiente para rejeitar a hipótese nula.

4. Investigações que exigem evidências mais fortes e precisas dos resultados adotam um valor de p mais rigoroso, como 1 ou até $0,1$. Pesquisas que testam a cura de doenças ou o desenvolvimento de novos medicamentos adotam um valor de ρ mais rigoroso, visto que nesses casos a probabilidade de o resultado encontrado ter sido ao acaso deve ser muito pequena.

ATENÇÃO

Embora se saiba que, quanto menor o valor de p, maior a evidência para se rejeitar a hipótese nula, deve-se tomar cuidado com a interpretação do valor de p, principalmente com a equiparação do nível de significância com o tamanho real do efeito experimental. Alguns estudantes podem associar o nível de significância com a força da relação entre duas variáveis, ou seja, dizer que, quanto menor o valor de p, maior a força a relação entre as variáveis. No entanto, esta afirmação é equivocada, visto que o nível de significância proporciona somente uma indicação da probabilidade de entrar o relacionamento entre as variáveis, por acaso, se a hipótese nula for verdadeira.

Fonte: Dancey, Reidy e Rowe (2017).

Para exemplificar a utilização do valor de p em um estudo, imagine que um pesquisador comparou a frequência cardíaca de dois grupos (homens e mulheres) após uma sessão de exercício físico de alta intensidade, verificando que a média da frequência cardíaca dos homens foi maior que a média da frequência cardíaca das mulheres. No entanto, após realizar o teste de hipótese adequado, o pesquisador encontrou um valor de $p = 0,25$. O que esse valor de p significa?

Este valor de p significa que a chance de essa diferença entre as médias ser devida ao acaso (e não uma diferença real entre os grupos) é de 25, isto é, se o pesquisador afirmar que as diferenças entre as médias ocorreram por causa dos tratamentos, ele tem 25 de chances de estar errado. Em outras palavras, podemos dizer que, se o pesquisador realizar a mesma comparação 100 vezes, ele irá encontrar resultados semelhantes em 75 experimentos.

Nesses casos, um erro que muitos pesquisadores cometem é realizar a seguinte inferência: “Embora não tenha sido encontrada diferença significativa entre os grupos $(p > 0,05)$, a média da frequência cardíaca das mulheres foi 10 superior à média da frequência cardíaca dos homens”. Este é um erro grave, pois o pesquisador tem 25 de chances de que essa diferença de 10 entre as médias não seja uma diferença real entre os sexos, e sim de outro fator que não está sendo estudado, sendo que o risco aceitável é de apenas 5. Enfim, o pesquisador estará atribuindo aos sexos uma diferença que tem grandes chances de ser obra do acaso, de ser um falso-positivo.

SAIBA MAIS

Para saber mais sobre o valor de p e as compreensões importantes relacionadas a ele, leia o artigo de Ferreira e Patino O que realmente significa o valor-p?, disponível em: <https://bit.ly/2QjlGXC>. Acesso em: 23 maio 2019.

Erros ao Testar Hipóteses

Quando realizamos testes de hipóteses para estabelecer conclusões a respeito da H_0, é importante lembrar que estas conclusões são obtidas a partir de uma amostra de informações, as quais podem conduzir aos erros nos testes de hipóteses. Os erros de hipóteses podem ser classificados como erro do tipo I e erro do tipo II (BARROS et al., 2012).

O erro do tipo I se refere à probabilidade de o teste estatístico demonstrar uma relação estatisticamente significativa quando na verdade ela não existe (ou que você rejeita a hipótese nula quando ela não deveria ser rejeitada). A probabilidade de este tipo de erro acontecer é chamada de alfa, que é o nível de significância do teste em que será rejeitada ($p < \alpha$). Por exemplo: ao testar o resultado de dois medicamentos semelhantes contra um nível de significância de $\alpha = 0,05$, admite-se que a probabilidade da ocorrência de um erro do tipo I é de 5, ou seja, existe uma uma chance de 5 de dizer que um medicamento é melhor do que o outro apesar de os dois serem semelhantes. O erro do tipo I aumenta na medida em que o pesquisador realiza mais testes de significância dentro do conjunto de dados. Casos em que muitas variáveis são analisadas simultaneamente favorecem o aumento do erro do tipo I.

O erro do tipo II corresponde à probabilidade de o teste estatístico indicar que não existe diferença estatisticamente significativa quando, na verdade, essa diferença existe. Em outras palavras, é quando se aceita H_0 quando esta é falsa, ou seja, conclui-se que não existe efeito quando, na verdade, existe. A probabilidade de ocorrência do erro do tipo II é denominada beta (β), que convencionalmente é estabelecida em 20.

Este erro também ocorre em decorrência do poder do teste estatístico, o qual é obtido por meio do seguinte cálculo: $1 - \beta$. Em um teste perfeito não existe a possibilidade de o erro do tipo II acontecer, apresentando poder um poder de 100, entretanto, sempre existe a possibilidade de este erro ocorrer. Por isso, um teste adequado deve apontar pelo menos 80 de poder. Existem diversos fatores que influenciam o poder de um teste estatístico, sendo os principais citados no Quadro 2.

Tamanho da amostra	O poder do teste aumenta na medida em que o tamanho da amostra também aumenta. Amostras maiores sempre proporcionam maior chance de encontrar os efeitos de um teste, enquanto que em amostras menores estes efeitos são mais difíceis de serem encontrados.
Variabilidade dos dados	O poder do teste aumenta na mesma proporção que a variabilidade dos dados diminui. Por exemplo, quanto menor for o desvio-padrão dos dados, maior o poder do teste.
Nível de significância	Quanto maior o nível de significância, maior o poder do teste. Por exemplo: a possibilidade de observar um efeito real aumenta quando se adota um nível de significância menor que , em detrimento a um nível de .

Quadro 2 - Fatores que influenciam o poder de um teste
Fonte: Elaborado pelo autor.

O Quadro 3 apresenta os dois tipos de erro em relação à conclusão sobre a H_0:

H_0	Conclusão sobre H_0	
	Rejeitar	**Aceitar**
Verdadeira	Erro do tipo I	Sem erro
Falsa	Sem erro	Erro do tipo II

Quadro 3 - Resumos dos erros do tipo I e II em relação à conclusão sobre H_0
Fonte: Adaptado de Barros et al. (2012).

SAIBA MAIS

Para saber mais sobre os erros do tipo I e tipo II dos testes estatísticos, assista ao vídeo disponível em: <https://bit.ly/2JC96m2>. Acesso em: 23 maio 2019.

Tipos de Variáveis e Análise Estatística

Dependendo da variável que for investigada, o teste de hipótese selecionado é diferente. Por exemplo: as variáveis quantitativas requerem testes estatísticos diferentes das variáveis qualitativas. Em relação aos dados quantitativos, serão abordadas nas próximas aulas as seguintes situações:

- 1 grupo comparado a um valor de referência;
- 2 grupos pareados ou dependentes que podem ser comparados;
- 3 ou mais grupos independentes que podem ser comparados;
- 2 variáveis que podem apresentar associação entre si.

QUESTÃO OBJETIVA

A estatística inferencial permite ao pesquisador estabelecer inferências sobre uma população a partir dos dados da amostra. Para isso, é necessário testar quanta evidência estes dados demonstram contra uma suposição específica em relação a uma determinada variável. O processo para confirmar se as hipóteses formuladas a respeito dos dados são verdadeiras ou falsas é chamado de teste de hipóteses. Qual é o raciocínio dos testes de hipóteses?

Determinar a probabilidade de obter um efeito devido ao erro amostral quando a hipótese nula é verdadeira.

- Determinar a probabilidade de ocorrer os erros do tipo I e II.
- Determinar a probabilidade de obter um efeito devido ao erro amostral quando a hipótese nula é falsa.
- Determinar a probabilidade de obter um efeito devido ao erro amostral quando a hipótese alternativa é verdadeira.
- Todas as alternativas estão corretas.

QUESTÃO OBJETIVA

Sabe-se que um nível de significância de 5% é suficiente para se rejeitar a hipótese nula e estabelecer que os resultados foram estatisticamente significativos a esse nível de significância. Mas, por qual motivo a maioria dos pesquisadores adota esse nível de significância?

- É um ponto de corte tradicional utilizado na área da saúde e que representa um bom balanceamento entre os erros do tipo I e II.

- Os resultados significativos são obtidos mais facilmente.
- O nível de significância não influencia a análise dos dados.
- Porque é um valor que representa todo o conjunto de dados.
- Nenhuma das alternativas anteriores.

Fechamento

Nesta aula vimos os principais conceitos e pressupostos relacionados aos testes de hipóteses e que serão fundamentais para os testes mais específicos que serão apresentados nas próximas aulas desta Unidade. Especificamente, vimos os seguintes aspectos: compreendemos o raciocínio lógico dos testes de hipóteses e da significância estatística; diferenciamos as hipóteses nula e alternativa; compreendemos a interpretação do valor de p para verificar a probabilidade de que o efeito em uma investigação ocorra devido ao erro amostral se a hipótese nula é verdadeira; reconhecemos os erros do tipo I e tipo II e como eles interferem no teste de significância. Além disso, vimos que os testes de significância possuem limitações que devem ser levadas em consideração no momento de estabelecer conclusões a respeito de um fenômeno investigado.

Nesta aula, você teve a oportunidade de:

- compreender o conceito e a lógica dos dos testes de hipóteses;
- reconhecer a importância e a interpretação da significância estatística;
- identificar os erros do tipo I e II nos testes de hipóteses.

Aula 02

Estatística Inferencial: Comparação de Médias

Introdução

Nesta aula abordaremos os primeiros testes de hipóteses que permitirão analisar as diferenças entre as médias em distintas condições, os quais são chamados de testes paramétricos. Estas condições são as seguintes: um grupo único, dois grupos independentes e dois grupos dependentes. Por exemplo, veremos a condição na qual é possível comparar a pressão arterial de adultos sedentários com o valor da população de referência (grupo único), a condição na qual pode-se comparar a pressão arterial entre homens e mulheres sedentários (dois grupos independentes) e a condição em que se pode comparar a pressão arterial de adultos sedentários antes e depois de um programa de atividade física (dois grupos dependentes). É importante lembrar que todos estes testes requerem o principal pressuposto quando se pretende comparar as médias de variáveis quantitativas, que é a distribuição normal. Quando este pressuposto é violado, deve-se adotar testes paramétricos, os quais serão abordados em outro momento.

Ao final desta aula, você será capaz de:

- reconhecer o teste estatístico adequado para se comparar as médias em diferentes condições;

- compreender a lógica dos testes *t* para um grupo, t independente e *t* independente;
- interpretar os resultados dos testes.

Comparando Grupos Independentes e Dependentes

Quando pensamos em utilizar um teste de hipóteses para chegar à conclusão em relação a algum problema de pesquisa que tenha alguma variável quantitativa, o primeiro pressuposto a ser analisado é a distribuição dos dados e saber se será utilizada a estatística paramétrica ou a estatística não paramétrica. Nesta aula, serão abordados os testes paramétricos para as diferentes condições em que podemos comparar grupos. Em todas as condições utilizamos o teste t, que foi criado por William Gosset, em 1908, quando realizava experimentos na cervejaria Guinness (DANCEY; REIDY, 2019). Os testes não paramétricos serão abordados mais à frente nesta Unidade.

É possível comparar três condições para se comparar médias: comparação de um grupo, de dois grupos independentes e de dois grupos dependentes. A primeira condição é quando queremos comparar as diferenças da média de uma variável de um grupo com um valor de referência na população.

Os testes para comparar dois grupos que não estão relacionados são chamados de testes independentes. Por exemplo, se queremos comparar a pressão arterial sistólica e diastólica entre um grupo de homens e um grupo de mulheres, utilizamos os testes independentes.

Já se o objetivo é comparar variáveis de um mesmo grupo em dois momentos distintos, os testes utilizados são os testes dependentes. Por exemplo, se queremos comparar a pressão arterial sistólica e diastólica de um grupo de mulheres antes e depois de um programa de intervenção de atividade física, utilizamos os testes dependentes.

Para fazer as comparações entre dois grupos relacionados ou independentes, faz-se a comparação da média ou da distribuição dos valores de um grupo em relação à média (testes paramétricos) ou da distribuição observada em outro grupo (testes não paramétricos).

Comparação de um Grupo

O teste para se comparar a média da variável de um grupo com o valor de uma população de referência é o teste t para um grupo. Nesses casos, como as variâncias da população de referência são desconhecidas, a distribuição normal não é possível de ser utilizada. Assim, deve-se assumir que o comportamento da variável na população de referência (no caso, a pressão arterial) segue uma distribuição t (curva t), sendo utilizado como valor de referência a média da população (BARROS et al., 2012). Retomando o exemplo demonstrado no início desta aula, suponha que uma amostra de 50 adultos sedentários foi avaliada quanto à pressão arterial sistólica e diastólica, observando que a média da pressão sistólica foi $125\ mmHg$, com desvio-padrão de $5,25$, enquanto que a média da pressão diastólica foi $85\ mmHg$, com desvio padrão de $3,10$. Esta situação hipotética é ilustrada na Figura 1.

Figura 1 - Exemplo de utilização do teste t para um grupo.
Fonte: Elaborada pelo autor.

O valor de referência contra o qual devemos comparar a média da pressão arterial do grupo de 50 adultos sedentários é $1200\ mmHg$ (pressão arterial sistólica) e

$80\ mmHg$ (pressão arterial diastólica), que são os valores de referência para a pressão arterial sistólica e diastólica normal na população. Nesse caso, teríamos as seguintes hipóteses estatísticas:

- H_0: a média da pressão arterial sistólica e diastólica na amostra de 50 adultos sedentários é igual à da população de referência.
- H_1: a média da pressão arterial sistólica e diastólica na amostra de 50 adultos sedentários é diferente do valor de referência da população.

Para testar essa H_0, deve-se efetuar a seguinte equação matemática:

$$t = \frac{\bar{x} - \mu 0}{\frac{8}{\sqrt{n}}}$$

Sabendo da equação e dos valores da amostra e da população de referência, agora é só efetuar a equação. Vamos exemplificar o cálculo da comparação da pressão sistólica:

$$t = (125 - 120) / (5,25 / \sqrt{50})$$

$$t = 5 / (5,25 / 7,07)$$

$$t = 5 / 0,74$$

$$t = 6,76$$

Após descobrir o valor de t, deve-se compará-lo com os valores críticos (Figura 2), considerando o nível de significância $(\rho < 0,05)$ e os graus de liberdade (gl) do teste estatístico $(n - 1)$, que nesse caso será $50 - 1 = 49$. Como o valor de $t$nesta comparação foi $6,76$ e supera o valor tabelado que é $2,01$, pode-se concluir que há evidência suficiente para rejeitar H_0 e que a média da pressão sistólica da amostra de 50 adultos sedentários apresentou diferença (é superior) com o valor de referência.

SAIBA MAIS

Após ver a aplicação da equação para a comparação da média da pressão arterial sistólica com o valor de referência, faça você o cálculo da comparação para a pressão arterial diastólica. Calcule o valor do teste t e compare com o valor crítico. Lembre-se dos valores a serem utilizados:

Amostra: $n = 50; média = 85\ mmHg;$
$desvio - padrão = 3,10\ mmHg.$

Média da população: $80\ mmHg$

Distribuição "t" de Student									
Nível de significância									
gl	↑0,800 0,200	↑0,850 0,150	↑0,900 0,100	↑0,950 0,050	**↑0,975 (a) 0,025**	↑0,980 0,020	↑0,985 0,015	↑0,990 0,010	**↑0,995 (h) 0,005**
1	1,37638	1,96261	3,07768	6,31375	**12,70615**	15,89447	21,20505	31.82096	**63,65590**
2	1,06066	1,38621	1,88562	2,91999	**4,30266**	4,84873	5,64280	6,96455	**9,92499**
3	0,97847	1,24978	1,63775	2,35336	**3,18245**	3,48191	3,89606	4,54071	**5,84085**
4	0,94096	1,18957	1,53321	2,13185	**2,77645**	2,99583	3,29763	3,74694	**4,60408**

5	0,91954	1,15577	1,47588	2,01505	**2,57058**	2,75651	3,00288	3,36493	**4,03212**
6	0,90570	1,13416	1,43976	1,94318	**2,44691**	2,61224	2,82893	3,14267	**3,70743**
7	0,89603	1,11916	1,41492	1,89458	**2,36462**	2,51675	2,71457	2,99795	**3,49948**
8	0,88889	1,10815	1,39682	1,85955	**2,30601**	2,44899	2,63381	2,89647	**3,35538**
9	0,88340	1,09972	1,38303	1,83311	**2,26216**	2,39844	2,57381	2,82143	**3,24984**
10	0,87906	1,09306	1,37218	1,81246	**2,22814**	2,35931	2,52749	2,76377	**3,16926**
11	0,87553	1,08767	1,36343	1,79588	**2,20099**	2,32814	2,49067	2,71808	**3,10582**
12	0,87261	1,08321	1,35622	1,78229	**2,17881**	2,30272	2,46070	2,68099	**3,05454**
13	0,87015	1,07947	1,35017	1,77093	**2,16037**	2,28160	2,43585	2,65030	**3,01228**
14	0,86805	1,07628	1,34503	1,76131	**2,14479**	2,26378	2,41490	2,62449	**2,97685**
15	0,86624	1,07353	1,34061	1,75305	**2,13145**	2,24854	2,39701	2,60248	**2,94673**
16	0,86467	1,07114	1,33676	1,74588	**2,11990**	2,26536	2,38155	2,58349	**2,92079**
17	0,86328	1,06903	1,33338	1,73961	**2,10982**	2,22384	2,36805	2,56694	**2,89823**
18	0,86205	1,06717	1,33039	1,73406	**2,10092**	2,21370	2,35618	2,55238	**2,87844**
19	0,86095	1,06551	1,32773	1,72913	**2,09302**	2,20470	2,34565	2,53948	**2,86094**
20	0,85996	1,06402	1,32534	1,72472	**2,08596**	2,19666	2,33625	2,52798	**2,84534**
21	0,85907	1,06267	1,32319	1,72074	**2,07961**	2,18943	2,32779	2,51765	**2,83137**
22	0,85827	1,06145	1,32124	1,71714	**2,07388**	2,18289	2,32016	2,50832	**2,81876**
23	0,85753	1,06034	1,31946	1,71387	**2,06865**	2,17696	2,31323	2,49987	**2,80734**
24	0,85686	1,05932	1,31784	1,71088	**2,06390**	2,17155	2,30692	2,49216	**2,79695**
25	0,85624	1,05838	1,31635	1,70814	**2,05954**	2,16659	2,30113	2,48510	**2,78744**
26	0,85567	1,05752	1,31497	1,70562	**2,05553**	2,16203	2,29581	2,47863	**2,77872**
27	0,85514	1,05673	1,31370	1,70329	**2,05183**	2,15782	2,29092	2,47266	**2,77068**
28	0,85465	1,05599	1,31253	1,70113	**2,04841**	2,15394	2,28638	2,46714	**2,76326**
29	0,85419	1,05530	1,31143	1,69913	**2,04523**	2,15033	2,28218	2,46202	**2,75639**
30	0,85377	1,05466	1,31042	1,69726	**2,04227**	2,14697	2,27827	2,45726	**2,74998**
35	0,85201	1,05202	1,30621	1,68957	**2,03011**	2,13316	2,26219	2,43772	**2,72381**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

40	0,85070	1,05005	1,30308	1,68385	**2,02107**	2,12291	2,25027	242326	**2,70446**
45	0,84968	1,04852	1,30065	1,67943	**2,01410**	2,11500	2,24109	2,41212	**2,68959**
50	0,84887	1,04729	1,29871	1,67591	**2,00856**	2,10872	2,23378	2,40327	**2,67779**
60	0,84765	1,04547	1,29582	1,67065	**2,00030**	2,09936	2,22292	2,39012	**2,66027**
70	0,84679	1,04417	1,29376	1,66692	**1,99444**	2,09273	2,21523	2,38080	**2,64790**
80	0,84614	1,04319	1,29222	1,66413	**1,99007**	2,08778	2,20949	2,37387	**2,63870**
90	0,84563	1,04244	1,29103	1,66196	**1,98667**	2,08394	2,20504	2,36850	**2,63157**
100	0,84523	1,04184	1,29008	1,66023	**1,98397**	2,08088	2,20150	2,36421	**2,62589**
110	0,84490	1,04134	1,28930	1,65882	**1,98177**	2,07839	2,19860	2,36072	**2,62127**
120	0,84463	1,04093	1,28865	1,65765	**7,97993**	2,07631	2,19620	2,35783	**2,61742**
140	0,84420	1,04029	1,28763	1,65581	**7,97706**	2,07306	2,19244	2,35328	**2,61140**
160	0,84387	1,03980	1,28686	1,65443	**1,97490**	2,07063	2,18962	2,34988	**2,60690**
180	0,84362	1,03943	1,28627	1,65336	**1,97323**	2,06874	2,18743	2,34724	**2,60341**
200	0,84342	1,03913	1,28580	1,65251	**1,97189**	2,06723	2,18569	2,34513	**2,60063**
∞	0,84178	1,03697	1,28240	1,64638	**1,96234**	2,05643	2,17319	2,33008	**2,58075**

Tabela construída usando o programa Microsoft Excel. (a) Nível crítico para p < 0,05 numa prova bicaudal (b) Nível crítico para p < 0,01 numa prova bicaudal

Figura 2 - Distribuição dos valores críticos do teste *i*
Fonte: Adaptado de Barros et al. (2012).

SAIBA MAIS

Apesar de ser possível realizar o cálculo do test *t* manualmente, diversos testes estatísticos efetuam a equação automaticamente, já apontando se existe diferença significativa ou não entre o valor de referência. Para ver como efetuar a comparação do teste t para um grupo no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=u2FrR3rSWZs>. Acesso em: 23 maio 2019.

SAIBA MAIS

Para ver como efetuar a comparação do teste t para um grupo no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=u2FrR3rSWZs&t=131s>. Acesso em: 23 maio 2019.

Comparação de Dois Grupos Independentes

O teste paramétrico utilizado para comparar as médias de variáveis quantitativas entre dois grupos independentes (não relacionados) é o teste *t* para amostras independentes. Em outras palavras, no teste *t* independente os participantes fazem parte de apenas um dos dois grupos. O seu respectivo não paramétrico é o teste “U” de Mann-Whitney, que será visto posteriormente nesta Unidade.

Este teste requer alguns pressupostos, como a utilização de amostras acima de 30 indivíduos, variáveis quantitativas, distribuição normal dos dados e variâncias semelhantes. Quando a amostra tem menos de 30 indivíduos, mas os dados apresentam distribuição normal, é possível aplicar o teste t para comparar as médias de dois grupos (DANCEY; REIDY, 2019).

Quando se pretende comparar dois grupos independentes, é necessário utilizar duas equações, uma para calcular o desvio-padrão agrupado dos dois grupos, e outra para calcular o valor de *t*, conforme ilustrado a seguir.

$$S = \sqrt{\frac{(n_1 - 1)\,S_1^2 + (n_2 - 1)\,S_2^2}{n_1 + n_2 - 2}}$$

S = desvio padrão agrupado da amostra

n_1 e n_2 = tamanho das amostras 1 e 2

S_1 e S_2 = desvios padrão das amostras 1 e 2

$$t = \frac{(Média\,1 - Média\,2)}{S\,.\,\sqrt{(1/n_1) + (1/n_2)}}$$

t = teste t, a estatística a ser calculada

n_1 e n_2 = tamanho das amostras 1 e 2

S = desvio padrão agrupado

Vamos adotar a seguinte situação, a comparação da pressão arterial sistólica de dois grupos: 30 homens e 20 mulheres sedentários. Esta situação hipotética é ilustrada na Figura 4.

Figura 3 - Exemplo de utilização do teste t para comparação entre dois grupos independentes.
Fonte: Elaborada pelo autor.

Nesse caso, as hipóteses da pesquisa ficariam da seguinte forma:

- H_0: a média da pressão arterial sistólica é estatisticamente igual entre os homens e as mulheres.
- H_1: a média da pressão arterial sistólica não é estatisticamente igual entre os homens e as mulheres.

Vamos considerar que os homens apresentaram média de $127\ mmHg$, com desvio-padrão de $4,00\ mmHg$, e as mulheres apresentaram média de $135\ mmHg$, com desvio-padrão de $5,00\ mmHg$. Supondo que os dados apresentaram distribuição normal e o tamanho da amostra de 50 indivíduos, podemos utilizar o teste t independente. Vamos efetuar o cálculos das equações:

$$s = \sqrt{}(30 - 1).42 + (20 - 1).52 \, / \, 30 + 20 - 2$$

$$s = \sqrt{}29.16 + 19.25 \, / \, 48$$

$$s = \sqrt{}464 + 475 \, / \, 48$$

$$s = \sqrt{939} / 48$$

$$s = \sqrt{19,56}$$

$$s = 4,42$$

$$t = (127 \;\; 135) \; / \; 4,42.\sqrt{(1/30)} \; + \; (1/20)$$

$$t = -8 \; / \; 4,42.\sqrt{0,03} \; + \; 0,05$$

$$t = -8 \; / \; 4,42.\sqrt{0,08}$$

$$t = -8 \; / \; 4,42. \; 0,28$$

$$t = -8 \; / \; 1,24$$

$$t = -6,45$$

Após descobrir o valor de t, deve-se compará-lo com os valores críticos, considerando o nível de significância $(p < 0,05)$ e os graus de liberdade (gl) do teste estatístico $(n1 + n2 - 2)$, que nesse caso será $50 - 2 = 48$. Como o valor de t nesta comparação foi $6,45$ e supera o valor tabelado que é $2,01$, pode-se concluir que há evidência suficiente para rejeitar H_0 e que a média da pressão sistólica dos homens é estatisticamente diferente da média da pressão sistólica das mulheres e que as mulheres apresentaram média superior.

SAIBA MAIS

Após ver a aplicação da equação para a comparação da média da pressão arterial sistólica entre homens e mulheres, faça você o cálculo da comparação para a pressão arterial diastólica. Calcule o valor do teste t e compare com o valor crítico. Lembre-se dos valores a serem utilizados:

Amostra: $n = 50$ (30 homens e 20 mulheres);
$média\ dos\ homens = 82\ mmHg\ desvio - padrão = 1,50\ mmHg$;
$média\ das\ mulheres = 86\ mmHg$ e
$desvio - padrão\ de\ 2,00\ mmHg$.

SAIBA MAIS

Para ver como efetuar a comparação do teste t para dois grupos independentes no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=Lt_8EbOHDic>. Acesso em: 23 maio 2019.

Uma das análises mais utilizadas na área da saúde e biológica é a comparação de medidas repetidas, na qual o pesquisador pretende comparar dois conjuntos dependentes (relacionados) de dados, isto é, o pesquisador pretende comparar uma variável quantitativa de uma mesma amostra em dois momentos distintos. Assim como nos testes anteriores, o principal pressuposto a ser analisado é a distribuição normal. Se os dados apresentam distribuição normal, recorremos ao testet dependente. Quando os dados não apresentam distribuição normal, o teste adotado é o teste de Wilcoxon, que será visto posteriormente nesta unidade.

O teste t dependente é mais sensível do que o independente, uma vez que cada participantes faz parte das duas condições (momentos), sendo comparado com ele mesmo. Outra condição importante para a realização do testet dependente é que a amostra a ser comparada tenha o mesmo tamanho nos dois momentos. Caso algum indivíduo não participe da avaliação em algum dos momentos, este deve ser excluído do conjunto de dados. Considere uma situação na qual se pretende comparar a pressão arterial sistólica 30 adultos sedentários antes e depois de um programa de 6 semanas de atividade física. Esta situação hipotética é ilustrada na Figura 5.

Figura 4 - Exemplo de utilização do teste t dependente para comparação entre dois momentos.
Fonte: Elaborada pelo autor.

Nesse caso, as hipóteses da pesquisa ficariam da seguinte forma:

- H_0: a média da pressão arterial sistólica é estatisticamente igual antes e depois do programa de atividade física.
- H_1: a média da pressão arterial sistólica é estatisticamente diferente antes e depois do programa de atividade física.

SAIBA MAIS

Para ver o cálculo manual das equações do teste t dependente, leia o capítulo 10 do livro da seguinte obra: BARROS et al. **Análise de dados em saúde**. 3. ed. Londrina/PR: Midiograf, 2012.

SAIBA MAIS

Para ver como efetuar a comparação do teste t dependente no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=5SccsRtT47c>. Acesso em: 23 maio 2019.

QUESTÃO OBJETIVA

A utilização dos testes t para comparação de amostras independentes e dependentes requer alguns pressupostos básicos. Qual é o pressuposto mais importante para a utilização dos testes t?

- Os valores devem ser retirados de uma população com distribuição normal.
- A amostra deve ser superior a 100 indivíduos.
- As condições devem ter a mesma média.
- O tamanho da amostra nas condições deve ser sempre igual.
- Todas as alternativas estão corretas.

QUESTÃO OBJETIVA

Um pesquisador pretende testar o nível de estresse de 50 funcionários de uma empresa antes e depois de um programa de seis semanas de sessões de relaxamento. Os valores foram obtidos a partir de uma população com distribuição normal. Qual é o teste estatístico mais apropriado para ser utilizado nesta situação?

- Teste t dependente.
- Test t independente.
- Teste de Wilcoxon.
- Teste t para um grupo único.
- Nenhuma das alternativas anteriores.

Fechamento

Ao longo desta aula, você teve a oportunidade de conhecer os testes paramétricos para a comparação das médias de variáveis quantitativas em diferentes condições de pesquisa. O teste t é utilizado para comparar uma variável quantitativa em um grupo único, dois grupos e dois momentos. Para se comparar a média de um grupo com uma referência da população, deve-se empregar o teste t para grupo único. Quando se quer comparar as médias entre dois grupos independentes, deve-se utilizar o teste t independente, ao passo que, quando se quer comparar as médias de uma amostra em dois momentos, deve-se utilizar o teste t dependente. Além disso, foi possível compreender a lógica matemática do teste t e sua forma de interpretação em cada situação de pesquisa.

Nesta aula, você teve a oportunidade de:

- reconhecer o teste estatístico adequado para se comparar as médias em diferentes condições;
- compreender a lógica dos testes *t* para um grupo, *t* independente e *t* independente;
- interpretar os resultados dos testes para cada condição apresentada.

Aula 03

Estatística Inferencial: Análise de Variância

Introdução

Nesta aula avançaremos nos testes de hipóteses que permitem comparar grupos independentes e dependentes. Nós já vimos os testes t que podem ser utilizados para a comparação de um grupo único, dois grupos independentes e dois grupos dependentes. Mas, e se o objetivo for comparar a nota de uma prova entre os estudantes solteiros, casados e divorciados (três grupos independentes)? Ou se desejar analisar uma variável quantitativa de uma amostra em três momentos distintos? Quais testes eu posso utilizar? Nesses casos, nós vamos utilizar a Análise de Variância (ANOVA) para 1 Fator ou a ANOVA de Medidas Repetidas, que são testes paramétricos considerados como uma extensão do teste t. As ANOVAs também possuem seus equivalentes não paramétricos, que serão vistos na próxima aula juntamente com os outros testes não paramétricos.

Ao final desta aula, você será capaz de:

- identificar os testes utilizados para comparar os valores de uma variável quantitativa de mais de dois grupos;
- identificar os testes utilizados para comparar os valores de uma variável quantitativa de uma amostra em mais de dois momentos distintos;

- interpretar a lógica dos testes e a forma de interpretação dos resultados.

Considerações Gerais Sobre a ANOVA

Assim como nos testes t, a ANOVA analisa as diferenças entre as médias dos grupos, determinando a média geral e verificando a sua diferença de cada média individual (DANCEY; REIDY; ROWE, 2017). Por esse motivo, a ANOVA é considerada uma extensão do teste t e apresentaria resultados semelhantes ao do teste t em uma comparação com duas condições (grupos ou momentos). A principal diferença da ANOVA para os testes t é que ela verifica as diferenças entre três ou mais grupos (amostras independentes) e entre três ou mais momentos (amostra dependente). Ao comparar três ou mais grupos ou momentos de uma só vez, ANOVA é mais robusta que o teste t e reduz as chances de erro tipo.

SAIBA MAIS

A Análise de Variância (ANOVA) verifica as diferentes fontes de variação que podem ocorrer em um conjunto de dados de uma variável quantitativa, ou seja, ela verifica em qual dos grupos (três ou mais) ou momentos (três ou mais) a variância é maior.

Fonte: Dancey, Reidy e Rowe (2017).

Existem dois tipos básicos de ANOVA: a ANOVA independente, conhecida como Anova 1 fator, e a ANOVA relacionada, chamada de Anova de Medidas Repetidas. O teste não paramétrico para a ANOVA 1 fator é o teste de Kruskal-Wallis, enquanto o teste não paramétrico para a Anova de Medidas Repetidas é o Teste de Friedman, os quais serão vistos juntamente com os demais testes não paramétricos.

Na ANOVA 1 fator, os participantes são avaliados somente em uma condição, ou seja, entre participantes ou entre um delineamento independente. Essa condição é representada por um delineamento de um fator com pelo menos três níveis (DANCEY; REIDY; ROWE, 2017). É um delineamento de um fator, pois apenas uma variável entra na análise e tem três níveis porque cada participante aparece somente em um nível. A Tabela 1 ilustra um delineamento de um fator, que é a nota da prova, com três níveis, que são os grupos (solteiros, casados e divorciados).

Nota dos Solteiros	Nota dos Casados	Nota dos Divorciados
I1	I7	I13
I2	I8	I14
I3	I9	I15
I4	I10	I16
I5	I11	I17
I6	I12	I18

Tabela 1 - Delineamento de 1 fator (nota da prova) com três níveis (estudantes solteiros, casados e divorciados) independentes
Nota: I = indivíduo.
Fonte: Elaborada pelo autor.

Na ANOVA de medidas repetidas, os participantes são avaliados em todas as condições em um delineamento intraparticipantes. Essa condição é representada por um delineamento de um fator com pelo menos três níveis ou condições de medidas repetidas (DANCEY; REIDY; ROWE, 2017). É um delineamento de um fator, pois apenas uma variável entra na análise e tem três níveis (momentos) porque cada participante apresenta um valor em um nível. A Tabela 2 ilustra um delineamento de um fator, que é a nota da prova de um grupo de alunos, com três níveis, que são os momentos (1º bimestre, 2º bimestre e 3º bimestre).

Nota do 1º bimestre	Nota do 2º bimestre	Nota do 3º bimestre
I1	I1	I1
I2	I2	I2
I3	I3	I3
I4	I4	I4

I5	I5	I5
I6	I6	I6

Tabela 2 - Delineamento de 1 fator (nota da prova) com três níveis (1° bimestre, 2° bimestre e 3° bimestre) dependentes
Nota: I = indivíduo.
Fonte: Elaborada pelo autor.

Os delineamentos da ANOVA independente e relacionada estão ilustrados na Figura 1, na qual primeiramente se pretende comparar a nota da prova entre três grupos independentes e depois se pretende comparar a nota da prova de uma amostra em três momentos distintos.

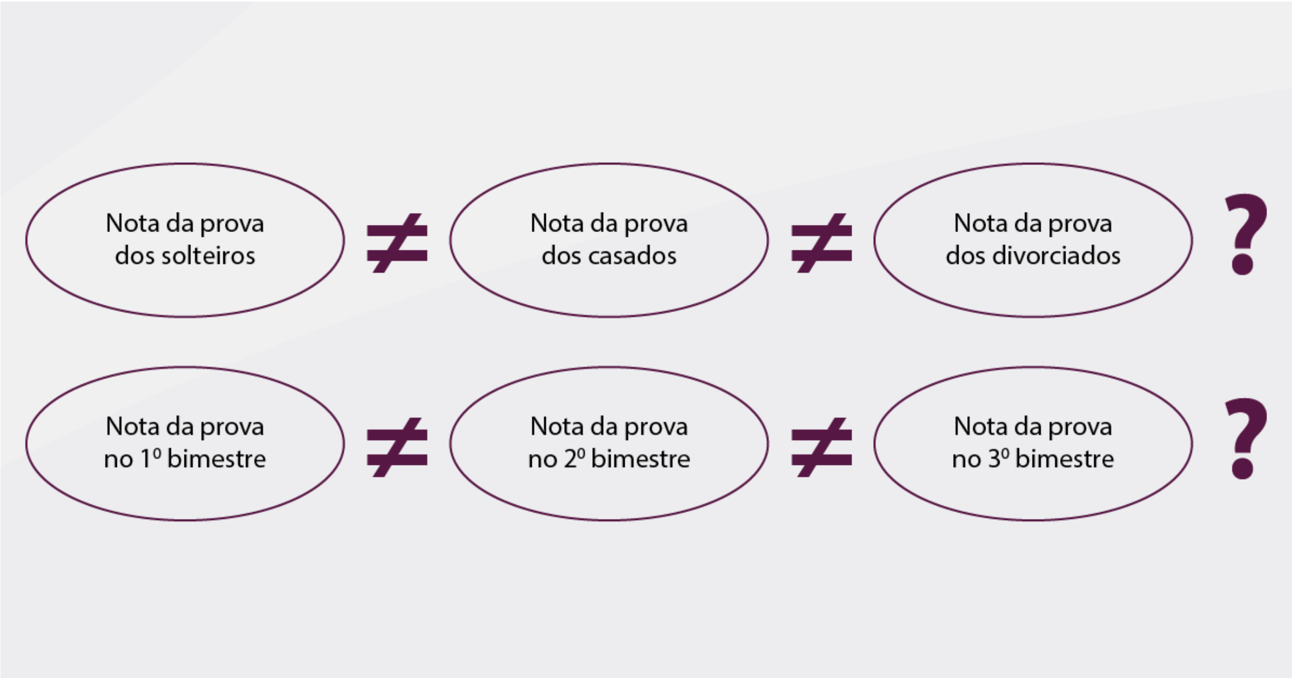

Figura 1 - Delineamentos da ANOVA independente e relacionada
Fonte: Elaborada pelo autor.

Nesses casos, as hipóteses da pesquisa tanto na ANOVA independente quanto na ANOVA relacionada ficariam da seguinte forma:

- H_0: a média da nota da prova é estatisticamente igual entre três grupos ou momentos

- H_1 : a média da nota da prova não é estatisticamente igual entre três grupos ou momentos.

Por ser um teste de hipótese mais complexo e robusto do que os testes t, a ANOVA (independente e relacionada) requer uma lógica mais complexa para o cálculo da diferença entre os grupos ou momentos. A ANOVA analisa a igualdade das variâncias dos dados comparando dois componentes, a variância intergrupos e intragrupos, verificando se a variância entre os grupos é semelhante à variância dentro dos grupos (BARROS et al., 2012).

Variância Entre os Grupos (intergrupos)

A variância entre os grupos se refere à diferença entre cada indivíduo em comparação à média de todas as observações (grupos). Assim, quando as médias dos grupos são diferentes, pode-se dizer que a variação entre os grupos é alta, enquanto que, se não existir diferença entre as médias dos grupos, não existe variação. Esta variação pode ser ilustrada a partir do exemplo da média da nota da prova dos solteiros, casados e divorciados na Tabela 3, que variaram entre 5,3 e 8,3. Pode-se dizer que a variação entre os grupos é uma variação entre as colunas da Tabela 3.

Nota dos Solteiros	Nota dos Casados	Nota dos Divorciados
I1 = 7,0	I7 = 8,5	I13 = 5,0
I2 = 7,5	I8 = 8,5	I14 = 5,5
I3 = 7,0	I9 = 8,5	I15 = 6,0
I4 = 6,5	I10 = 8,0	I16 = 4,5
I5 = 7,0	I11 = 8,0	I17 = 5,0
I6 = 6,0	I12 = 8,5	I18 = 6,0
Média = 6,8	Média = 8,3	Média = 5,3

Tabela 3 - Nota da prova dos alunos solteiros, casados e divorciados
Nota: I = Indivíduo.

Fonte: Elaborada pelo autor.

Variância Intragrupos (Dentro dos Grupos)

A variância intragrupos se refere às diferenças e variações dentro dos grupos, ou seja, às diferenças entre cada participante em comparação à média do grupo que faz parte. Essa variação pode ser considerada como a variação dentro de cada coluna da tabela. Ao analisar os dados da Tabela 3, perceba que as notas do grupo 2 (casados) têm pequena variação interna, já que todos os indivíduos tiveram nota entre 8,0 e 8,5. Em contrapartida, o grupo 3 (divorciados) apresenta uma variação maior, entre 3,5 e 6,5.

Calculando as Variâncias e o Teste F

Para descobrir se existe diferença entre as médias de dois ou mais grupos ou momentos, a ANOVA executa os seguintes cálculos:

- Média de cada um dos três ou mais grupos ou momentos.
- Média geral da amostra (soma das médias de cada grupo ou momento e divisão pela quantidade de grupos ou momentos).
- Em cada grupo é calculada a variação total de cada indivíduo em comparação à média do grupo (variação intragrupos).
- A variação da média de cada grupo em relação à média geral (variação entre os grupos).

Percebe-se que todos os cálculos na ANOVA envolvem a média das observações. Assim, a variância total de um grupo revela a distância entre a média geral e o valor mais distante, conforme ilustra a Figura 2. A variância total é particionada devido à diferença entre a variância entre os grupos e a variância intragrupos (DANCEY; REIDY, 2019).

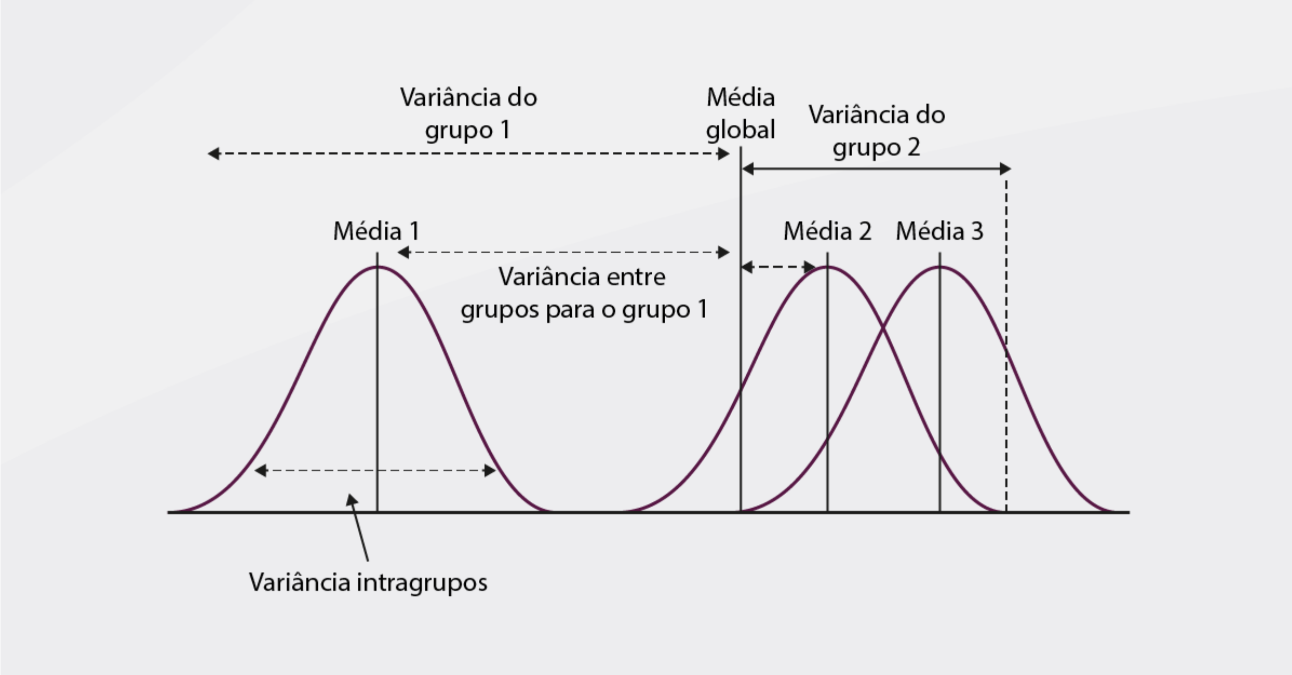

Figura 2 - Simulação de três grupos mostrando a decomposição da variância e as distâncias da média geral das observações
Fonte: Dancey e Reidy (2019, p. 305).

Para verificar as diferenças entre as variâncias intergrupos e intragrupos, utiliza-se o teste F. O valor obtido no teste F é comparado com os valores críticos da distribuição de F (BARROS et al., 2012). Assim, o resultado da ANOVA pode ser representado pela razão entre a variância intergrupos e a variância intragrupos, conforme a equação abaixo:

$$F \ = \ Variância\ intergrupos \ / \ Variância\ intragrupos$$

Se o valor de F calculado for maior do que o F tabelado, deve-se rejeitar a H_0, ao passo que se o valor de F calculado for menor do que o F tabelado, deve-se aceitar a H_0.

SAIBA MAIS

Para saber como calcular o valor do teste F manualmente e verificar se existe diferença entre grupos independentes, assista ao vídeo do link a seguir:

<https://www.youtube.com/watch?v=9DMKUL3QYpI>. Acesso em: 29 abr. 2019.

O ideal em um estudo que se pretende comparar três ou mais grupos ou momentos é que a variância dentro dos grupos seja mínima para que o valor do teste F seja alto. Ainda, quando a variância entre os grupos é maior do que a variância dentro dos grupos, o valor do teste F será alto e a probabilidade de rejeitar a hipótese nula também é maior (DANCEY; REIDY, 2019).

A Figura 2 demonstra que existe uma sobreposição entre os grupos 2 e 3, indicando que eles não apresentam valores muito diferentes, enquanto que o grupo 1 não apresenta nenhuma sobreposição com os grupos 2 e 3. Além disso, a distância das médias dos grupos 2 e 3 é pequena quando comparada à distância entre a média do grupo 1 e a média geral.

Dessa forma, quanto maior for a variância entre os grupos em relação à variância média intragrupos, maior será o valor de F e a diferença entre os grupos. A Figura 3 ilustra situações de grande sobreposição (pequena diferença) e pequena sobreposição (grande diferença) entre os grupos.

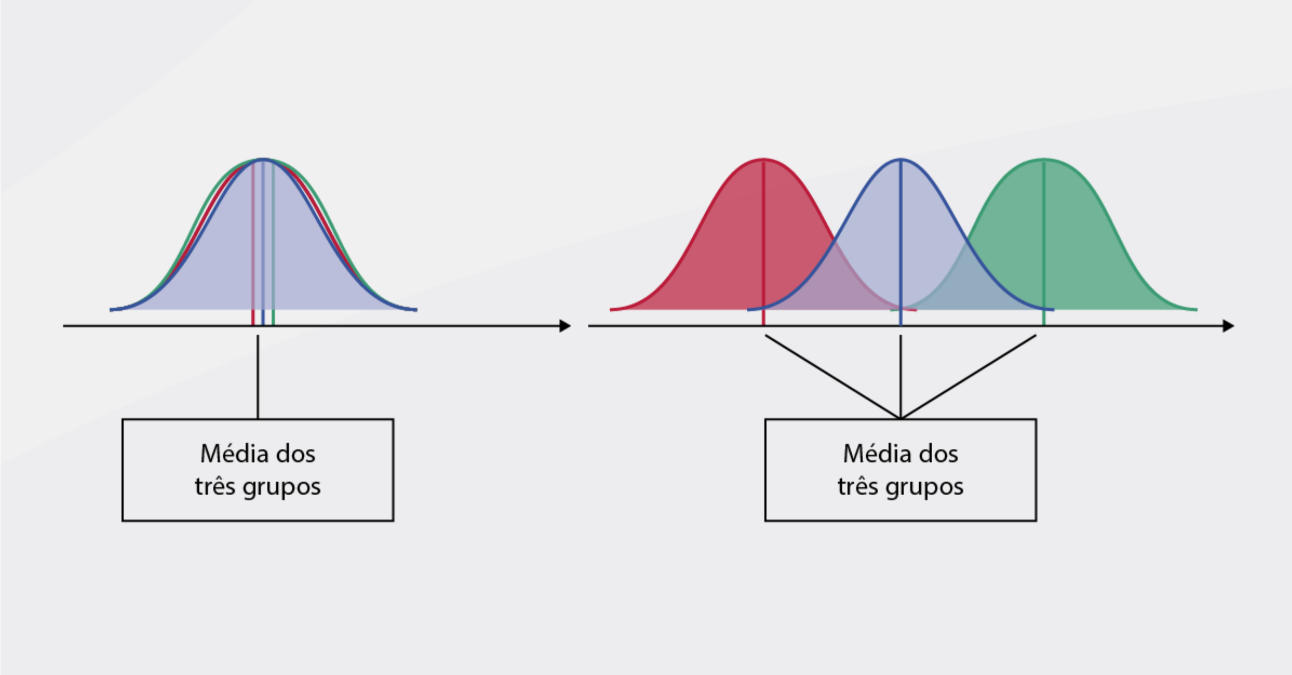

Figura 3 - Representação de grande sobreposição e pequena sobreposição entre três grupos
Fonte: Elaborada pelo autor.

Embora a ANOVA e o cálculo do teste F sejam possíveis de serem realizados manualmente, os softwares estatísticos efetuam a ANOVA independente (1 Fator) e a ANOVA dependente (Medidas repetidas) rapidamente e sem a necessidade de cálculos manuais.

SAIBA MAIS

Para ver como efetuar a ANOVA de um fator para três grupos independentes no software SPSS, assista ao vídeo do link a seguir: <https://www.youtube.com/watch?v=EsPAX1Au8yc>. Acesso em: 29 abr. 2019.

SAIBA MAIS

Para ver como efetuar a ANOVA de medidas repetidas para três momentos no software SPSS, assista ao vídeo do link a seguir:

<https://www.youtube.com/watch?v=x0AZr9-U7Ss>. Acesso em: 29 abr. 2019.

Comparações Múltiplas

A ANOVA indica se existe alguma diferença entre as médias dos três grupos ou momentos, entretanto, não indica entre quais grupos foi encontrada a diferença significativa. Para isso, é necessário escolher um dos procedimentos de comparações múltiplas, chamados de Post-Hoc.

A escolha destes testes de comparações múltiplas deve levar em consideração o tamanho da amostra, número de grupo ou momentos, e o balanceamento entre a ocorrência do erro tipo I e tipo II. Os testes de Post-Hoc são similares ao teste t, mas são mais robustos por reduzirem a ocorrência do erro tipo I. Os principais Post-Hoc são o teste de Scheffé, o teste de Tukey e o teste de Bonferroni, os quais devem ser devem ser utilizados em algumas situações particulares (BARROS et al., 2012), como:

- Teste de Tukey: amostras com mesmo tamanho e variâncias semelhantes.

- Teste de Scheffé: é mais conservador e é indicado quando se tem muitos grupos.
- Teste de Bonferroni: mais indicado quando se tem menor quantidade de grupos.

SAIBA MAIS

Para verificar mais características dos testes de comparações múltiplas, leia o texto do link a seguir: <http://www.portalaction.com.br/anova/teste-de-comparacoes-multiplas>. Acesso em: 29 abr. 2019.

QUESTÃO OBJETIVA

Um pesquisador pretende comparar três diferentes tipos de tratamento para uma determinada doença. Para isso, ele aloca aleatoriamente 100 participantes em cada uma das três condições. Como ele é um pesquisador cuidadoso, ele verificou por meio do histograma e do teste de Kolmogorov-Smirnov que os dados apresentaram uma distribuição normal. Para analisar as diferenças entre as três condições, qual é o teste mais adequado?

- ANOVA de 1 fator.
- Teste t.
- Correlação linear simples.
- ANOVA de medidas repetidas.
- Curva de Gauss.

QUESTÃO OBJETIVA

A ANOVA é o teste estatístico mais apropriado quando se quer comparar as médias de três ou mais grupos ou momentos. Para se chegar ao resultado da ANOVA, é necessário calcular o valor do teste F. Qual é a equação para se obter o valor do teste F?

- Variância entre os grupos / variância intragrupos.
- Variância dentro dos grupos / variância intragrupos.
- Variância entre os grupos x variância intragrupos.
- Variância entre os grupos + variância intragrupos.
- Variância entre os grupos - variância intragrupos.

Fechamento

Nesta aula aprendemos sobre o teste que é considerado a extensão do teste t, a ANOVA, que é o teste paramétrico que deve ser utilizado sempre que desejamos comparar três ou mais grupos ou momentos. Quando se quer comparar as médias entre três grupos independentes, deve-se utilizar a ANOVA 1 fator, ao passo que, quando se quer comparar as médias de uma amostra em três ou mais momentos, deve-se utilizar a ANOVA de medidas repetidas. Ambos os testes requerem o pressuposto da distribuição normal dos dados. No entanto, a ANOVA não indica entre quais grupos foi encontrada diferença nas médias, sendo necessário aplicar um teste de comparações múltiplas, os Post-Hoc. Os testes não paramétricos da ANOVA 1 fator e da ANOVA de Medidas Repetidas são o teste de Kruskal-Wallis e o Teste de Friedman.

Nesta aula, você teve a oportunidade de:

- reconhecer a ANOVA de 1 Fator como o teste adequado para comparar os valores de uma variável quantitativa em mais de dois grupos;

- reconhecer a ANOVA de Medidas Repetidas como o teste adequado para comparar os valores de uma variável quantitativa de uma amostra em mais de dois momentos distintos;
- interpretar a lógica dos testes na forma de interpretação dos resultados.

Aula 04

Correlação e Regressão

Introdução

Todos os dias nós associamos acontecimentos e situações cotidianas. Frases como “se chover vai diminuir o calor” ou “se meu salário aumentar eu vou viajar no final do ano” são exemplos claros de relação entre duas variáveis e que evidenciam que se uma variável aumentar a outra vai diminuir, ou que se uma variável aumentar a outra vai aumentar na mesma proporção. Essas situações são exemplos de relação entre duas variáveis quantitativas, indicando que o resultado de uma variável pode ser diretamente ou indiretamente proporcional ao resultado da outra.

Nesta aula, veremos as duas medidas estatísticas que permitem este tipo de interpretação. Ao passo que a correlação apenas indica o grau da relação entre duas variáveis, a regressão tenta prever o resultado de uma variável a partir do conhecimento do resultado da outra.

Ao final desta aula, você será capaz de:

- conhecer os conceitos de correlação e regressão;
- compreender as formas de interpretação da correlação e da regressão;
- compreender a aplicação da correlação e da regressão em trabalhos científicos.

Considerações Gerais a Respeito da Correlação e Regressão

Um dos principais interesses quando se estuda duas ou mais variáveis é saber se elas têm algum relacionamento entre si, isto é, se os valores altos ou baixos de uma das variáveis implicam em valores altos ou baixos da outra variável. Por exemplo, pode-se analisar se existe associação entre a taxa de investimento financeiro na saúde e a taxa de mortalidade em um país, ou entre a renda mensal média da população e o gasto mensal familiar com a saúde. Para descobrirmos se tal relação existe, é necessário selecionar um número de famílias e registrar a renda mensal de cada uma delas e o quanto cada família gasta por mês com a saúde.

O estudo deste tipo de associação entre variáveis é denominado correlação e regressão. Quando o estudo possui apenas duas variáveis tem-se a correlação e a regressão simples, ao passo que, quando envolve mais de duas variáveis, tem-se a correlação e a regressão múltiplas. Além disso, a correlação proporciona apenas o grau da relação entre duas variáveis, enquanto a regressão permite a previsão do resultado de uma variável a partir do conhecimento do resultado da outra.

SAIBA MAIS

Assista ao vídeo do link a seguir para ver mais exemplos da utilização da correlação e da regressão: <https://www.youtube.com/watch?v=89M97CWnkik>. Acesso em: 29 abr. 2019.

Correlação

A correlação é uma medida estatística que mede a relação entre duas variáveis e indica a força e a direção do relacionamento linear entre duas variáveis aleatórias, entretanto, esta medida não possibilita dizer se uma variável é causa ou consequência da outra. Quando duas variáveis estão correlacionadas, significa que, quando os valores em uma variável mudam, os valores da outra variável mudam de forma previsível (DANCEY; REIDY, 2019).

A medida que proporciona informações a respeito da força e da direção da relação entre duas variáveis é denominada coeficiente de correlação de Pearson ou correlação linear simples, o qual é representado pela letra “r”. A palavra simples que compõe o nome correlação linear simples indica que estão envolvidas no cálculo somente duas variáveis. O coeficiente de Pearson não possui unidade de medida e indica apenas o grau de proximidade entre as duas observações (BARROS et al., 2012).

Duas variáveis podem estar associadas de forma positiva (direta) ou de forma negativa (inversa), dessa forma, pode-se avaliar o quanto os valores destas variáveis estão correlacionados. A seguir estão elencadas informações importantes a respeito do coeficiente de correlação:

- O coeficiente de correlação varia entre $-1~e~+1$.
- Quanto mais próximo de zero o coeficiente de correlação, mais fraca a relação entre as variáveis.
- Quando o coeficiente de correlação é zero, assume-se que não há relação linear entre as variáveis.
- Uma correlação positiva indica que as duas variáveis se movem no mesmo sentido, enquanto uma correlação negativa indica que as duas variáveis se movem em sentidos opostos.
- Quanto mais próximo de $+1$ o coeficiente de correlação, mais forte a relação positiva entre as variáveis, enquanto que, quanto mais próximo de -1 a correlação, mais forte a relação positiva entre as variáveis.
- Duas variáveis que estão perfeitamente correlacionadas de forma positiva $(r = 1)$ movem-se essencialmente em perfeita proporção na mesma direção.
- Duas variáveis que estão perfeitamente correlacionados de forma negativa $(r = -1)$ movem-se em perfeita proporção em direções opostas.

Além do coeficiente de correlação (r), uma das principais formas de representar a associação entre duas variáveis é o gráfico de dispersão, que se refere a um diagrama no qual os pontos x e y no espaço cartesiano são usados para representar simultaneamente os valores de duas variáveis quantitativas medidas em cada elemento do conjunto de dados. Este gráfico é utilizado para demonstrar a intensidade, o sentido e tipo de correlação entre duas variáveis.

Conforme as informações da Figura 1, percebe-se que, quanto mais próxima de 1 (positivo ou negativo), mais linear é a relação entre as variáveis. Embora seja uma medida interessante, não é possível afirmar que o valor do coeficiente de correlação será o mesmo em uma amostra diferente.

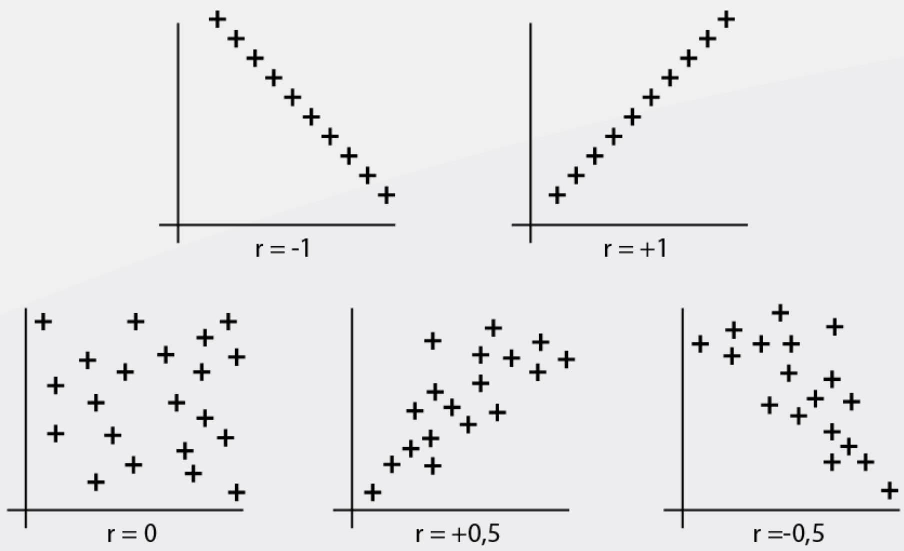

Figura 1 - Diagramas de dispersão com diferentes valores de "r"
Fonte: Barros et al. (2012, p.186).

O cálculo do coeficiente de correlação é possível de ser efetuado quando se tem duas variáveis quantitativas de uma amostra com n indivíduos. A equação para o cálculo do coeficiente de correlação linear simples é obtida a partir da seguinte equação.

$$r = \frac{\Sigma\,(x - Média\ x)\ (y - Média\ Y)}{\sqrt{\Sigma(x - Média\ x)^2\ \Sigma(y - Média\ y)^2}}$$

SAIBA MAIS

Assista ao vídeo do link abaixo para saber mais sobre os conceitos apresentados sobre a correlação e para ver exemplo do cálculo do coeficiente de correlação linear de forma manual: <https://www.youtube.com/watch?v=jiK6E5jrjPM>. Acesso em: 29 abr. 2019.

Após calcular o coeficiente de correlação (r), é importante analisar a magnitude (intensidade) da associação entre as variáveis. Embora a interpretação deste coeficiente varie na literatura, os resultados podem ser analisados no intervalo 1 até $+1$ tomando por base a tabela a seguir:

Valor de r	Classificação
$0,01$ *até* $0,39$	Fraca
$0,40$ *até* $0,69$	Moderada
$0,70$ a $1,00$	Forte

Tabela 1 - Classificação do Coeficiente de Correlação de Pearson
Fonte: Adaptado de Dancey e Reidy (2019).

A Figura 2 ilustra detalhadamente os vários graus dos coeficientes de correlação de acordo com o sentido (positivo ou negativo) da relação entre as variáveis:

Perfeito	+1	-1
	+0,9	-0,9
	+0,8	-0,8
	+0,7	-0,7
	+0,6	-0,6
Moderado	+0,5	-0,5
	+0,4	-0,4
	+0,3	-0,3
	+0,2	-0,2
	+0,1	-0,1
Zero	0	

Figura 2 - Intensidade dos coeficientes de correlação positivos e negativos
Fonte: Dancey e Reidy (2019, 187).

É importante ressaltar que o valor da correlação (r) não indica uma relação de causa e efeito entre as variáveis. Outra medida para interpretar o valor do coeficiente de correlação é o coeficiente de determinação $(r2)$, que é a potência ao quadrado do valor de r multiplicado por 100, e que será muito útil nas análises de regressão (DANCEY; REIDY, 2019). O coeficiente de determinação indica o percentual de variação de uma variável que pode ser atribuída à relação com a outra variável.

ATENÇÃO

As correlações negativas ao quadrado têm como resultado um valor positivo. Por exemplo, duas variáveis apresentaram uma correlação de $-0,6$ ao quadrado $(-0,6\ x\ -0,6) = 0,36$. Logo, estas variáveis compartilham 16% de variância, resultado semelhante se a correlação fosse positiva $(+0,6)$.

Para ilustrar a explicação do coeficiente de determinação, vamos tomar como exemplo as variáveis massa corporal e estatura de crianças, as quais estão associadas de forma positiva, isto é, quanto maior a estatura, maior a massa corporal das crianças. Sempre que duas variáveis se correlacionam, pode-se dizer que elas compartilham variância. Por exemplo, os círculos do infográfico a seguir representam as variáveis independentes massa corporal e estatura.

INFOGRÁFICO INTERATIVO

Para consultar o Infográfico Interativo, acesse a **versão digital** deste material

Pode-se dizer que 49 da variância da massa corporal pode ser explicada pela variância da estatura, enquanto que 49 da variância da estatura também pode ser explicada pela variância da massa corporal. Se as variáveis compartilham 49 de variância, temos 51 de variância não compartilhada, que também é conhecida como variância exclusiva, ou seja, $25,5$ é exclusiva da massa corporal e $25,5$ é exclusiva da estatura. Isso quer dizer que 51 da variância de cada uma das variáveis se deve a outros fatores. Sempre que a variância compartilhada for maior que a variância exclusiva, o valor do coeficiente de correlação será alto, ao passo que, quando a variância exclusiva for maior que a variância compartilhada, a correlação será baixa. A Tabela 2 apresenta os valores do coeficiente de determinação a partir dos coeficientes de correlação.

Correlação	Coeficiente de determinação (r^2)	Variância compartilhada
$0,0$	$0,0^2$	$0,00$
$0,1$	$0,1^2$	0,01~\left(1% \right)
$0,2$	$0,2^2$	0,04~\left(4% \right)
$0,3$	$0,3^2$	0,09~\left(9% \right)
$0,4$	0,4^2	0,16~\left(16% \right)
$0,5$	$0,5^2$	0,25~\left(25% \right)
$0,6$	$0,6^2$	0,36~\left(36% \right)
$0,7$	$0,7^2$	0,49~\left(49% \right)
$0,8$	$0,8^2$	0,64~\left(64% \right)
$0,9$	$0,9^2$	0,81~\left(81% \right)
$1,0$	$1,0^2$	1,0~\left(100% \right)

Tabela 2 - Coeficientes de correlação e respectivos coeficientes de determinação
Fonte: Adaptada de Barros et al. (2012).

SAIBA MAIS

Para ver como efetuar a correlação linear de Pearson para associar duas variáveis no software SPSS, assista ao vídeo do link a seguir: <https://www.youtube.com/watch?v=9a1ova9v03Y>. Acesso em: 29 abr. 2019.

Regressão

Enquanto a análise de correlação permite apenas identificar a intensidade e o sentido da associação entre duas variáveis, a análise de regressão permite verificar o efeito de uma variável sobre a outra. Em outras palavras, a partir da análise de regressão é possível saber quanto o valor de uma variável pode mudar a partir da alteração no valor da outra variável. O tipo mais comum de regressão é a regressão simples, que analisa o efeito de uma variável sobre a outra (DANCEY; REIDY, 2019).

Vamos imaginar que temos dados sobre o preço de um medicamento e a quantidade de venda do produto. Por meio da correlação podemos encontrar que o preço do medicamento e a quantidade de vendas têm uma correlação negativa, entretanto, não é possível afirmar o quanto a venda do produto reduzirá a partir do aumento do preço.

Ao usar a análise de regressão, podemos verificar o efeito que a variável x (aumento do preço) causa na variável y (quantidade de vendas). Na análise de regressão existe uma equação matemática que permite prever quanto o valor da variável y mudará a partir da mudança no resultado da variável x. Com isso, seria possível prever a seguinte situação fictícia: se o preço do medicamento aumenta 30, a quantidade vendas cairá 25. Dessa forma, podemos sugerir que o valor de uma variável influenciou o valor da outra variável.

Outros exemplos da utilização da regressão envolvem o efeito do estresse na produtividade no trabalho, o efeito do volume de exercício sobre a redução do percentual de gordura ou a influência da motivação do aluno para os estudos sobre o desempenho nas provas.

SAIBA MAIS

Leia o artigo científico a seguir para ver a aplicação da análise de regressão em uma pesquisa na área da saúde:

RAMOS, A. C. V. et al. Estratégia Saúde da Família, saúde suplementar e desigualdade no acesso à mamografia no Brasil. **Revista Panamericana de Salud Pública**, v. 42, p. 166, 2018.

Para entender a fórmula da análise de regressão, é importante entender o conceito da linha de regressão, que é uma linha traçada no gráfico de dispersão (já visto quando falamos de correlação) para ver o quanto os pontos de dados estão agrupados ao redor da linha. Quanto mais próximos da linha estão os pontos, maior a previsão de x em y. A partir do momento que temos uma linha reta representando os dados, é possível afirmar que para cada unidade de mudança em $x,\ y$ muda por uma quantidade específica.

O resultado da análise de regressão é estimado pela seguinte fórmula: $Y\ =\ a\ +\ bx.$ Esta fórmula indica a maneira pela qual $y\ )mudacomoresultadodamudançaem\backslash(x.$ X é a variável que está prevendo a outra variável, sendo chamada de preditora ou independente, enquanto a variável y é a variável que está sendo explicada pela variável x, sendo chamada de variável critério ou dependente. Já a representa o valor de y quando x é igual a zero, a qual é chamada de linha de interceptação. A inclinação da linha (chamada de b) proporciona uma medida de quanto y muda a partir da mudança em x (DANCEY; REIDY, 2019).

Um exemplo seria usar a equação de regressão para tentar prever a nota da prova de um aluno (y) a partir da quantidade de horas de estudo (x). A partir da análise de dados chegou-se aos resultados da Figura 4. A equação obtida na análise foi a seguinte:

$$y\ =\ a\ +\ bx$$

$$y\ =\ 0,74\ (a\ -\ >\ \ quando\ x\ é\ zero,\ y\ é\ igual\ a\ 0,74)\ +\ 0,92\,(b)\,.\,(hora\ de\ estudo)$$

-> Vamos supor que o aluno estudou 8 horas para a prova:

$$y\ =\ 0,74\ +\ 0,92.8$$

$$y\ =\ 0,74\ +\ 7,36$$

$$y\ =\ 8,1$$

Logo, podemos prever que um aluno que estudou 8 horas tem uma nota prevista de $8,1$ na prova. Ao olhar o gráfico, você verá que a equação foi capaz de confirmar essa afirmação. Além disso, percebe-se que foi obtido um coeficiente de determinação de 0,98~\left(98% \right), indicando que as variáveis compartilham 98 de variância.

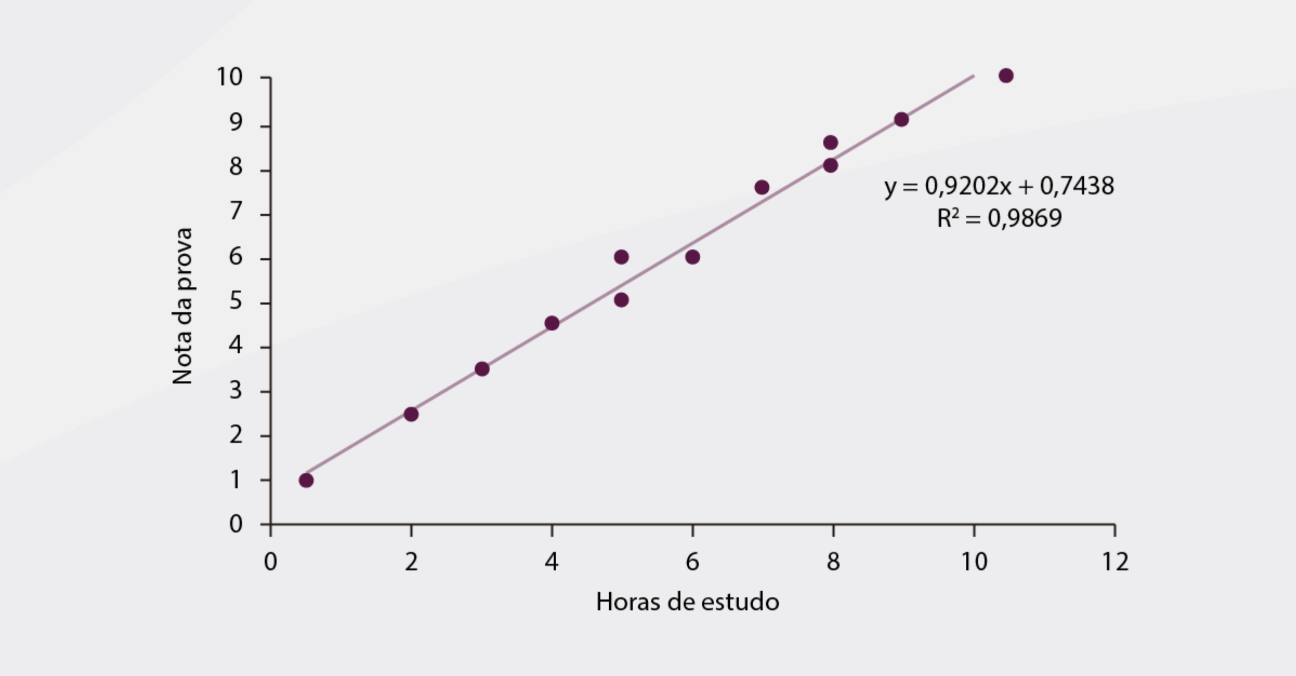

Figura 3 - Gráfico de dispersão com a linha de regressão
Fonte: Elaborada pelo autor.

Quando o objetivo é analisar o efeito de mais de uma variável preditora sobre uma variável desfecho, é necessário empregar a regressão linear múltipla. A regressão múltipla possui características semelhantes às da regressão linear simples. A equação da regressão múltipla é uma extensão da regressão simples.

SAIBA MAIS

Leia o artigo científico *Regressão múltipla stepwise e hierárquica em Psicologia Organizacional: aplicações, problemas e soluções* para saber mais sobre a aplicação da análise de regressão linear múltipla.

Disponível em: <https://bit.ly/2QkWf84>. Acesso em: 29 abr. 2019.

QUESTÃO OBJETIVA

Um pesquisador investigou se o tempo gasto vendo televisão tem alguma relação com o percentual de gordura de adolescentes. Após a análise dos dados, o pesquisador verificou uma correlação linear de 0,75 entre as variáveis. A partir desse resultado, qual foi a conclusão do pesquisador?

- O tempo gasto vendo televisão apresentou uma associação forte e positiva com o percentual de gordura. Além disso, as variáveis compartilham 56% de variância.

- O tempo gasto vendo televisão apresentou um efeito forte e positivo com o percentual de gordura.

- O tempo gasto vendo televisão apresentou uma associação forte e negativa com o percentual de gordura. Além disso, as variáveis compartilham pequeno percentual de variância.

- O tempo gasto vendo televisão não tem associação com o percentual de gordura dos adolescentes.

- O tempo gasto vendo televisão apresentou uma associação forte e negativa com o percentual de gordura. Além disso, as variáveis compartilham 56% de variância.

QUESTÃO OBJETIVA

A análise de correlação e regressão linear simples proporcionam informações a respeito da associação entre duas variáveis, entretanto, ambas possuem particularidades. Qual é a principal diferença entre as duas análises?

A análise de correlação proporciona informações sobre a intensidade e o sentido da associação entre duas variáveis, enquanto a regressão linear simples proporciona informações a respeito do efeito de uma variável sobre a outra.

- A análise de correlação proporciona informações sobre a intensidade, o sentido e a causalidade da associação entre duas variáveis, enquanto a regressão linear simples proporciona informações a respeito do efeito de uma variável sobre a outra.

- A análise de correlação proporciona informações sobre a intensidade e o sentido da associação entre duas variáveis, enquanto a regressão linear simples proporciona informações a respeito do efeito de muitas variáveis sobre apenas outra variável.

- As análises não possuem diferenças e proporcionam as mesmas informações.

- A análise de regressão linear simples proporciona informações sobre a intensidade e o sentido da associação entre duas variáveis, enquanto a correlação proporciona informações a respeito do efeito de uma variável sobre a outra.

Fechamento

Nesta aula vimos que a análise de correlação é útil quando temos como objetivo analisar a associação entre duas variáveis, ao passo que a análise de regressão vai além da correlação e permite verificar o efeito de uma variável sobre a outra. A correlação entre duas variáveis quantitativas é amplamente utilizada quando se deseja saber o grau e o sentido de interdependência no aspecto de variação conjunta que uma variável tem sobre outra. Já a regressão proporciona informações a respeito do quanto o valor de uma variável pode influenciar a mudança no valor da outra variável.

Nesta aula, você teve a oportunidade de:

- conhecer os conceitos de correlação e regressão linear.
- compreender as formas de interpretação da correlação e da regressão linear simples.
- compreender a aplicação da correlação e da regressão linear simples em trabalhos científicos.

Aula 05

Estatística Inferencial: Testes não Paramétricos

Introdução

Imagine que um pesquisador pretende comparar as médias do nível de ansiedade de homens e mulheres com diagnóstico de câncer, entretanto, ao analisar o principal pressuposto para a utilização dos testes de comparação de médias (testes t e ANOVA), ele percebe que o conjunto de dados possui muitos valores extremos e que a distribuição dos dados é assimétrica, fugindo da distribuição normal. E agora? O que se deve fazer já que não é recomendado utilizar um teste para comparação de médias? Nesse caso e em todas as situações em que este pressuposto básico não é atendido, deve-se adotar a estatística não paramétrica, a qual possui testes alternativos para todos os testes paramétricos vistos nas aulas anteriores desta Unidade. Nessa aula, veremos os testes “U” de Mann-Whitney, Wilcoxon, Kruskal-Wallis, Friedman e a Correlação de Spearman.

Ao final desta aula, você será capaz de:

- compreender as características e aplicabilidade da estatística não paramétrica;
- identificar os testes não paramétricos alternativos para os testes paramétricos;
- compreender a interpretação dos testes não paramétricos.

Estatística não Paramétrica

Embora a estatística paramétrica seja mais poderosa e robusta do que a estatística não paramétrica, os testes paramétricos requerem alguns pressupostos específicos, sendo o principal deles a distribuição normal do conjunto de dados. Em alguns casos, este pressuposto é violado, fazendo com que sejam adotados testes não paramétricos.

A Estatística não paramétrica representa um conjunto de testes de hipóteses apropriados para pesquisas nas quais não se conhece muito bem a distribuição da população e seus parâmetros, ou seja, quando os dados de conjunto de observações não atendem o pressuposto da distribuição normal (DANCEY; REIDY; ROWE, 2017). A aplicação dos testes não paramétricos possui vantagens e desvantagens para a análise de um conjunto de dados, as quais estão listadas no Quadro 1.

Vantagens	Desvantagens
Os testes não paramétricos podem ser aplicados a diferentes situações, uma vez que não requerem populações com distribuição normal.	A estatística não paramétrica pode perder informações durante os cálculos matemáticos, uma vez que os dados numéricos são frequentemente reduzidos a uma forma qualitativa (postos).
Diferentemente dos testes paramétricos, os testes não paramétricos podem ser aplicados em variáveis qualitativas.	Os testes não paramétricos são menos robustos do que os testes paramétricos e requerem amostras maiores para aumentar a probabilidade de rejeitar a hipótese nula.
A estatística não paramétrica utiliza cálculos matemáticos mais simples e de fácil compreensão.	

Quadro 1 - Vantagens e desvantagens da utilização da estatística não paramétrica
Fonte: Elaborada pelo autor.

Como nas situações em que os dados não apresentam distribuição normal o valor da média é distorcido pelos valores extremos do conjunto de dados, quando utilizamos os testes não paramétricos é recomendada a utilização da mediana como medida de tendência central para representar os dados.

SAIBA MAIS

Para ver mais aplicações e características da estatística não paramétrica, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=RRx-S4aDLPY>. Acesso em: 29 abr. 2019.

Nos tópicos a seguir abordaremos os testes não paramétricos para cada um dos testes paramétricos vistos nas aulas anteriores desta Unidade, começando pelos testes não paramétricos para o teste *t*.

Alternativas não Paramétricas ao Teste *t*

Os testes "U" de Mann-Whitney e Wilcoxon são os testes equivalentes não paramétricos dos testes t independente e dependente. O teste "U" de Mann-Whitney é apropriado para duas amostras independentes (dois grupos), enquanto o teste de Wilcoxon é utilizado quando se têm os mesmos participantes ou participantes emparelhados em duas condições ou dois momentos (DANCEY; REIDY; ROWE, 2017).

Diferentemente do teste *t* independente, o teste "U" não se baseia nos valores médios e seu cálculo é mais simples, sendo obtido por meio da comparação da soma dos postos (posições) que os valores do conjunto de dados ocupam em cada um dos dois grupos (BARROS et al., 2012).

SAIBA MAIS

Para saber como é calculado o ranking dos postos nos testes "U" de Mann-Whitney e Wilcoxon, leia os capítulos 10 e 11 do livro:

BARROS et al. **Análise de dados em saúde.** 3. ed. Londrina/PR: Midiograf, 2012.

Considere uma situação na qual um pesquisador pretende comparar densidade óssea entre crianças com desnutrição e crianças sem desnutrição. Ao verificar que os dados não apresentaram distribuição normal, o pesquisador teve que recorrer ao teste não paramétrico "U" de Mann-Whitney. As hipóteses estatísticas nula e alternativa são formuladas no mesmo raciocínio dos testes paramétricos, sendo a hipótese nula sempre conservadora:

- H_0: os dois grupos têm a mesma distribuição.
- H_1: os dois grupos não têm a mesma distribuição.

SAIBA MAIS

Para ver como efetuar o teste não paramétrico "U" de Mann-Whitney para comparar dois grupos no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=SJT-k6VaHiU>. Acesso em: 29 abr. 2019.

O teste de Wilcoxon é o equivalente não paramétrico do teste *t* dependente e também utiliza o ranking dos postos para se calcular a estatística do teste. No entanto, este teste determina a diferença entre as medidas de cada par de observações, e estas diferenças são classificadas de acordo com o sinal (positivo ou negativo) e ordenadas pela magnitude. Por último, os escores são transformados pelo valor do posto que ocupam no ordenamento dos desvios (BARROS et al., 2012).

Considere uma situação na qual um pesquisador pretende comparar o percentual de gordura de uma amostra de crianças antes e depois de um programa de oito semanas de alimentação saudável. Ao verificar que os dados não apresentaram distribuição normal, o pesquisador teve que recorrer ao teste não paramétrico de Wilcoxon. As hipóteses estatísticas nula e alternativa são formuladas no mesmo raciocínio dos testes paramétricos, ficando da seguinte forma:

- H_0: a mediana das diferenças entre as observações é igual a zero.
- H_1: a mediana das diferenças entre as observações não é igual a zero.

SAIBA MAIS

Para ver como efetuar o teste não paramétrico de Wilcoxon para comparar dois momentos no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=VOVEESAEGnA>. Acesso em: 29 abr. 2019.

Alternativas não Paramétricas à ANOVA

Os testes de Kruskal-Wallis e Friedman são os testes equivalentes não paramétricos da ANOVA 1 fator e da ANOVA de medidas repetidas. O teste de Kruskal-Wallis é apropriado para três amostras independentes (três grupos), enquanto o teste de Friedman é utilizado quando se tem os mesmos participantes ou participantes emparelhados em três condições ou três momentos.

O teste de Kruskal-Wallis é uma extensão do teste “U” de Mann-Whitney e não considera as distribuições dos valores, mas sim o posicionamento das observações em cada grupo, as quais são comparadas por meio da soma dos postos. As principais condições em que o teste de Kruskal-Wallis é utilizado são quando os dados não apresentam distribuição normal, com amostras reduzidas e com variáveis em escala ordinal (DANCEY; REIDY; ROWE, 2017).

SAIBA MAIS

Para saber como é calculada manualmente a equação matemática do teste de Kruskal-Wallis, leia o capítulo 12 do livro:

BARROS et al. **Análise de dados em saúde**. 3. ed. Londrina-PR: Midiograf, 2012.

Considere uma situação na qual um professor pretende comparar o nível de conhecimento em bioestatística de estudantes do primeiro, segundo e terceiro anos de uma faculdade. Além de ter uma amostra reduzida, os dados apresentaram uma distribuição assimétrica, impossibilitando a aplicação da ANOVA 1 fator. Dessa forma, é necessário aplicar o teste não paramétrico de Kruskal-Wallis. As hipóteses estatísticas nula e alternativa são formuladas no mesmo raciocínio dos testes paramétricos, ficando da seguinte forma:

- H_0: os três grupos têm a mesma distribuição dos valores.
- H_1: os três grupos não têm a mesma distribuição dos valores.

SAIBA MAIS

Para ver como efetuar o teste não paramétrico de Kruskal-Wallis para comparar três ou mais grupos no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=LROp-jeaa7g>. Acesso em: 29 abr. 2019.

O teste de Friedman é utilizado nos casos em que os dados não atendem os pressupostos para a utilização da ANOVA de medidas repetidas. Este teste é uma extensão do teste de Wilcoxon, entretanto, analisa as diferenças entre os mesmos participantes ou o emparelhamento dos participantes em três ou mais condições (momentos). Assim como o teste de Kruskal-Wallis, o teste de Friedman não considera as distribuições dos valores, mas sim o posicionamento das observações em cada grupo, as quais são comparadas por meio da soma dos postos (DANCEY; REIDY; ROWE, 2017).

Considere uma situação na qual um professor pretende comparar o nível de conhecimento em bioestatística de uma amostra de estudantes de uma faculdade em três condições (momentos): começo do ano, meio do ano e fim do ano. Considerando que os dados não apresentaram distribuição normal, não é recomendada a aplicação da ANOVA de medidas repetidas, sendo necessária a utilização do teste não paramétrico de Friedman. As hipóteses estatísticas nula e alternativa são formuladas no mesmo raciocínio dos testes paramétricos, ficando da seguinte forma:

- H_0 : as três condições têm a mesma distribuição dos valores.
- H_1 : as três condições não têm a mesma distribuição dos valores.

SAIBA MAIS

Para ver como efetuar o teste não paramétrico de Friedman para comparar três ou mais momentos no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=MHLoWbQ09qQ>. Acesso em: 29 abr. 2019.

Alternativa à Correlação de Pearson

Em algumas situações não é recomendado empregar a correlação de Pearson, uma vez que podem levar a interpretações errôneas a respeito da relação entre as variáveis do conjunto de observações investigado (BARROS et al., 2012). As principais situações em que a correlação de Pearson (r) não deve ser empregada estão apontadas na Figura 1.

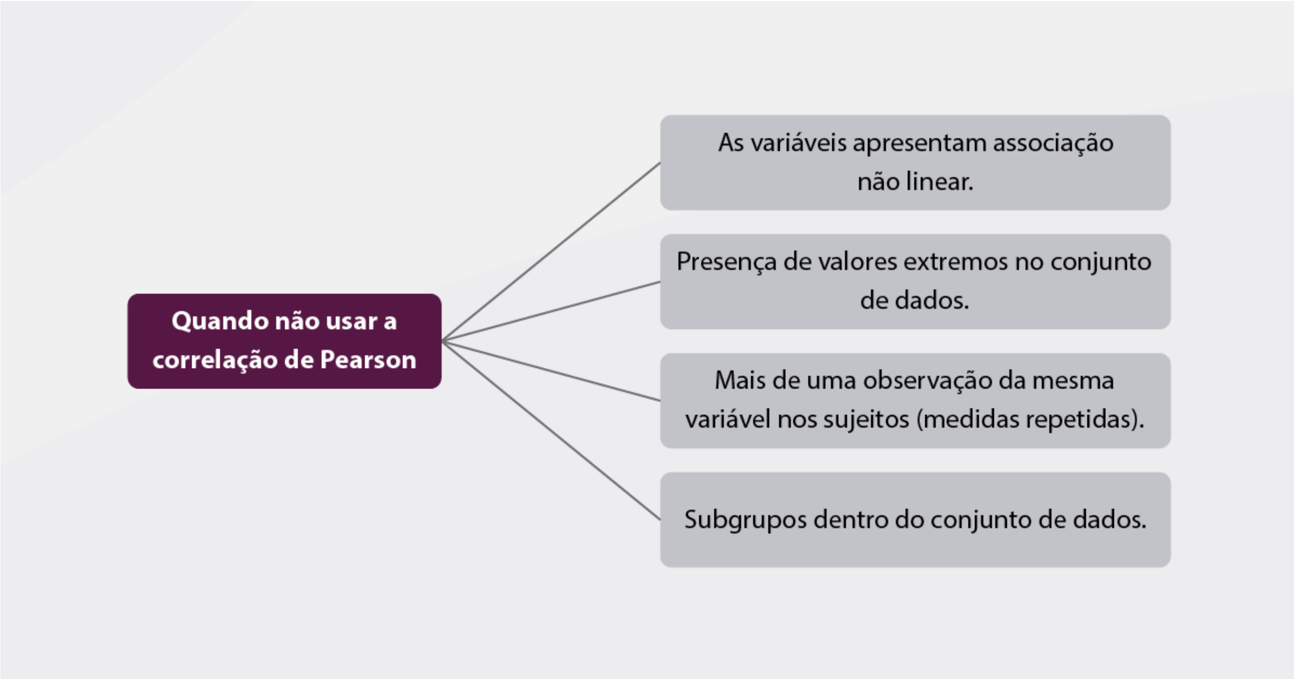

Figura 1 - Situações nas quais se deve evitar o uso da correlação de Pearson
Fonte: Elaborada pelo autor.

A Figura 2 ilustra as situações apontadas na Figura 1 no diagrama de dispersão para facilitar a compreensão das situações em que não se deve utilizar a correlação de Pearson.

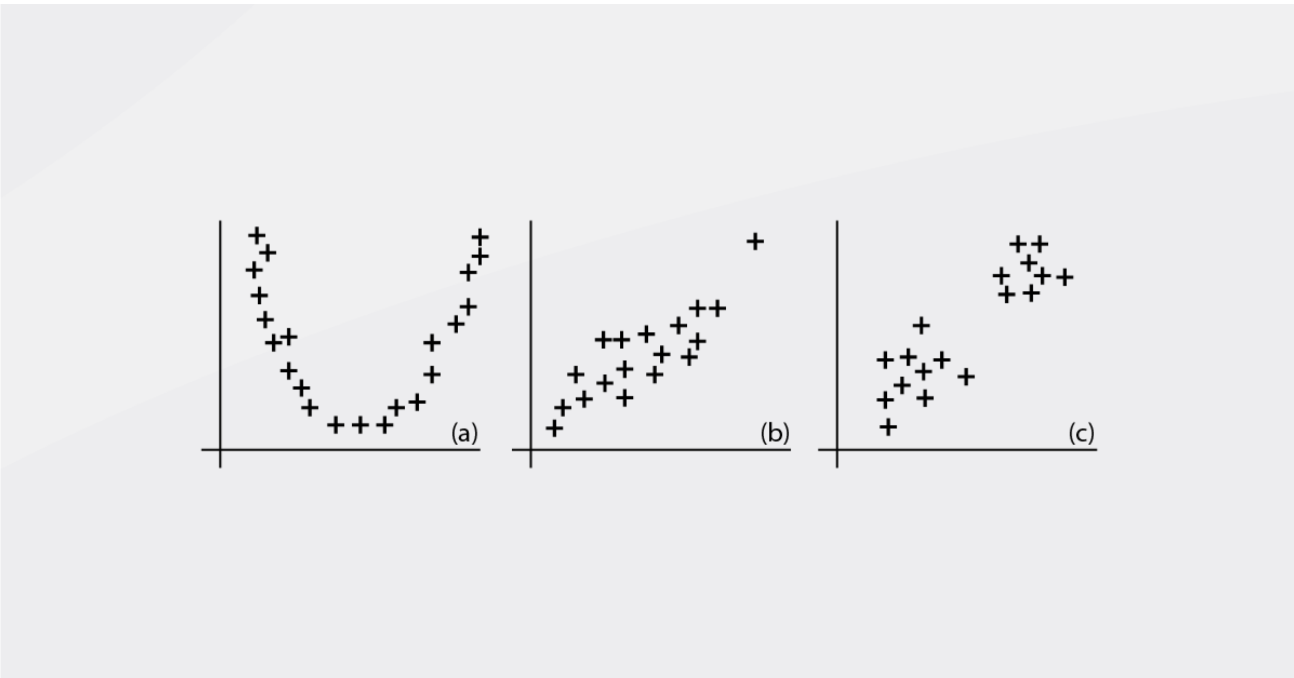

Figura 2 - Diagramas que ilustram relação não linear (a), presença de valores extremos (b) e presença de subgrupos (c).

Fonte: Barros et al. (2012, p.187).

Nesses casos deve-se empregar a correlação de Spearman (p), que é o equivalente não paramétrico da correlação de Pearson (r). Embora ambos os testes sejam similares e interpretados de forma semelhante, a correlação de Spearman é uma medida de correlação não linear e deve ser empregada nas situações listadas na Figura 3.

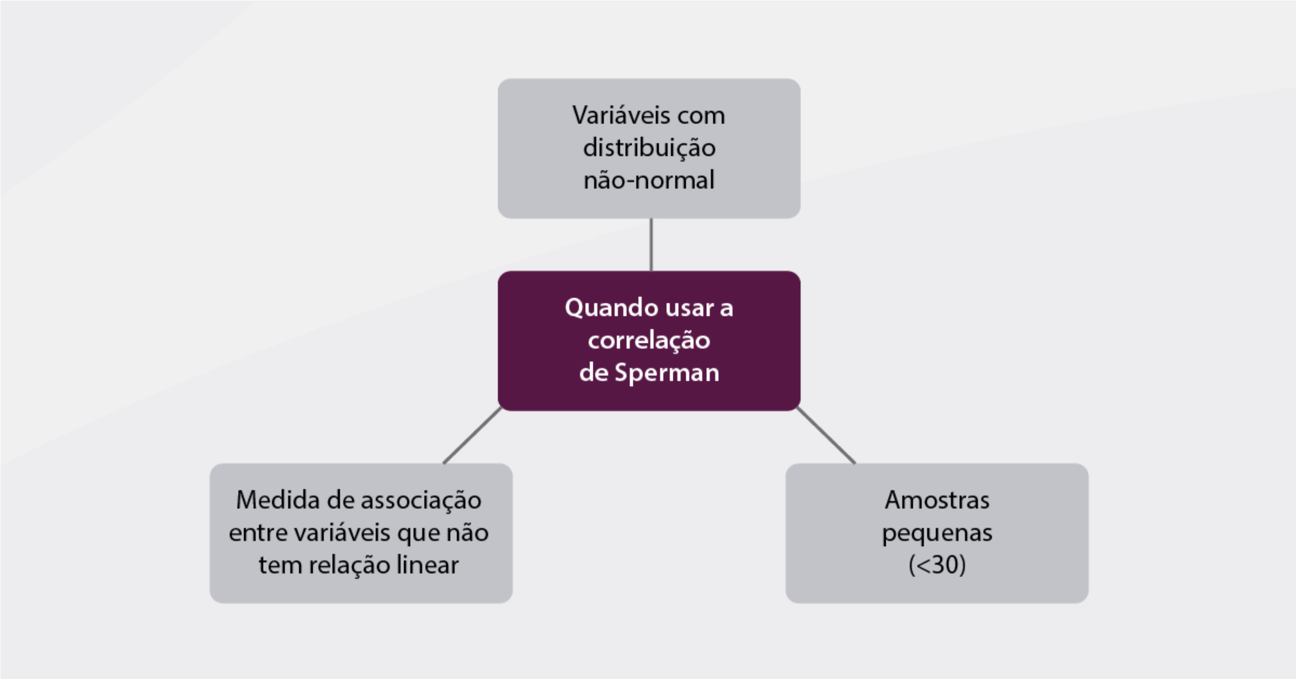

Figura 3 - Situações nas quais se deve utilizar a correlação de Spearman
Fonte: Elaborada pelo autor.

O cálculo da correlação de Spearman é obtido por meio transformação dos escores brutos de uma variável, os quais devem ser substituídos pelo valor da sua posição (posto) no conjunto de dados. A mesma operação deve ser realizada com a outra variável. Em seguida, as mesmas operações matemáticas da correlação de Pearson devem ser empregadas para a correlação de Spearman (BARROS et al., 2012). Dessa forma, pode-se dizer que a correlação de Spearman é uma correlação linear entre os postos das duas variáveis.

As hipóteses estatísticas nula e alternativa da correlação de Spearman são formuladas no mesmo raciocínio da correlação de Pearson, ficando da seguinte forma:

- H_0: o coeficiente de correlação é igual a zero.
- H_1: o coeficiente de correlação é não igual a zero.

SAIBA MAIS

Para ver como efetuar a Correlação de Pearson (paramétrica) e Spearman (não paramétrica) para associar duas variáveis quantitativas no software SPSS, assista ao vídeo disponível em: <https://www.youtube.com/watch?v=9a1ova9v03Y&t=379s>. Acesso em: 29 abr. 2019.

QUESTÃO OBJETIVA

Considere uma situação na qual um pesquisador pretende analisar a diferença do estado de humor dos funcionários de uma empresa do começo do expediente de trabalho (8h00min) para o final do expediente (17h0min). Para isso, ele aplicou um questionário no grupo de funcionários em dois momentos distintos, o qual forneceu um escore de humor para cada funcionário em cada momento (início e fim do expediente). Lembre-se de que, ao analisar a distribuição dos dados, o pesquisador percebeu que os dados fugiam da curva da distribuição normal. Qual o teste estatístico apropriado para esta situação?

- Teste de Wilcoxon
- Teste “U” de Mann-Whitney.
- Teste *t* independente.
- Teste *t* dependente.
- Anova de Medidas Repetidas.

QUESTÃO OBJETIVA

Para verificar a diferença nos valores de duas condições de delineamento entre participantes (amostras independentes), com dados organizados em postos, qual é o teste mais apropriado?

- "U" de Mann-Whitney.
- Kruskal-Wallis.
- Wilcoxon.
- Correlação de Spearman.
- Teste t independente.

Fechamento

A estatística não paramétrica é uma alternativa interessante nos casos em que os dados apresentem uma distribuição assimétrica e que fuja da distribuição normal, o que impossibilita a utilização dos testes paramétricos. Embora sejam úteis nesses casos e também possuam cálculos matemáticos mais simples e não exijam grandes amostras, os testes não paramétricos são menos poderosos e robustos do que os testes paramétricos, além de tornarem possível a perda de informação durante os cálculos, uma vez que os dados numéricos são frequentemente reduzidos a uma forma qualitativa (postos). Os testes não paramétricos vistos nesta aula foram o teste "U" de Mann-Whitney (equivalente ao teste *t* independente), Wilcoxon (equivalente ao teste *t* dependente), Kruskal-Wallis (equivalente à Anova de 1 fator), Friedman (equivalente à Anova de medidas repetidas) e Correlação de Spearman (equivalente à Correlação de Pearson).

Nesta aula, você teve a oportunidade de:

- compreender as características, vantagens e desvantagens da estatística não paramétrica;
- identificar os testes não paramétricos alternativos para os testes paramétricos;
- compreender a interpretação dos testes não paramétricos.

ATIVIDADE COMPLEMENTAR

Para complementar o aprendizado e aprofundar o conhecimento em relação aos testes paramétricos e não paramétricos discutidos ao longo desta Unidade, você pode ler o livro Descobrindo a estatística usando o SPSS, que consta na lista de referências ao final da Unidade. Durante a leitura, atente-se aos procedimentos necessários para efetuar cada um dos testes paramétricos e não paramétricos no software SPSS. Alguns dos testes abordados ao longo do livro são apresentados no vídeo do link a seguir, que também demonstra como efetuar testes estatísticos no SPSS.

Fonte: <https://www.youtube.com/watch?v=y4S_xEZ-8PM>. Acesso em: 29 abr. 2019.

Ao assistir ao vídeo e durante a leitura, atente-se aos pressupostos para a utilização dos testes paramétricos e não paramétricos e ao passo a passo para efetuar cada teste.

Teoria e Prática

Você sabia que pessoas calmas têm menor risco de ter demência durante a velhice? Esta é uma informação importante, considerando que vivemos em uma sociedade em que as pessoas vivem uma rotina de trabalho estressante e que afeta diretamente a saúde física e mental. Mas uma informação como essa não é obtida sem a condução de uma pesquisa e a utilização da estatística descritiva e inferencial.

Para se chegar a esta conclusão, um grupo de pesquisadores conduziu um experimento com 506 idosos que não sofriam de demência ao serem examinados inicialmente. O grupo recebeu questionários para apurar detalhes sobre sua personalidade e estilo de vida. Após acompanhar estes idosos por seis anos, o estudo concluiu que pessoas mais calmas e relaxadas têm 50% menor risco de desenvolver demência em comparação às pessoas com tendência a se estressar.

ESTUDO DE CASO

De acordo com a Organização Mundial de Saúde (OMS), o estresse é considerado o mal do século XXI, atingindo 90% da população mundial. Trabalho, problemas pessoais, relacionamentos amorosos, finanças, mudanças no estilo de vida, falta de tempo livre para o lazer e para a família são os principais fatores que desencadeiam o estresse. Para criar políticas públicas que auxiliem no combate ao estresse, uma equipe de pesquisadores procurou compreender a relação do estresse com alguns fatores do

cotidiano diário das pessoas, como a carga-horária de trabalho, o número de filhos, a renda mensal e as horas destinadas às atividades de lazer. Para se chegar a uma conclusão de que estas variáveis apresentam relação com o estresse das pessoas, a condução de uma pesquisa utilizando a estatística inferencial é fundamental, pois um teste de hipótese poderá proporcionar informações a respeito dos principais fatores que favorecem ou evitam o desencadeamento do estresse.

O pesquisador tem um problema a ser estudado (relação do estresse com fatores do cotidiano) e, após obter os dados de uma amostra por meio de um questionário, deve ser empregado um teste de hipóteses, no caso uma regressão linear, para verificar os possíveis fatores preditores do estresse. Com estas conclusões, é possível desenvolver políticas públicas e criar estratégias para reduzir o estresse na população.

Mapa Conceitual