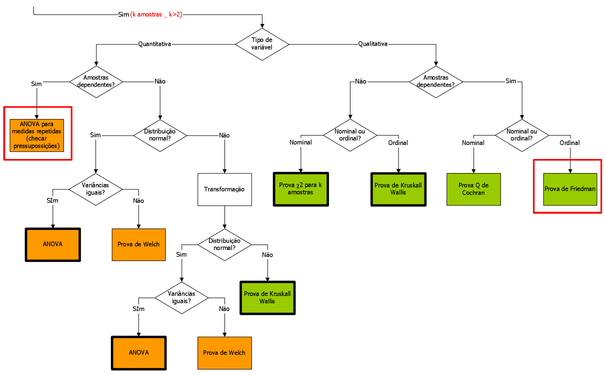
Comparando k amostras (k>2) dependentes

- Abordagem paramétrica e não-paramétrica

Testes de Tendência Central (média, mediana, proporção)	Classificação	Variável 1	Variável 2	Número Grupos	Dependência	Premissas
Teste Z	Paramétrico	Quantitativa	-	1	-	Variância pop. * conhecida
Teste t	Paramétrico	Quantitativa	-	1	-	Distribuição normal
Wilcoxon (teste dos sinais, Wilcoxon p/ 1 amostra)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	-	1	-	
Teste t p/ 2 amostras	Paramétrico	Quantitativa	Nominal	2	Independentes	Distribuição normal; variâncias iguais
Teste t p/ 2 amostras com variâncias diferentes	Paramétrico	Quantitativa	Nominal	2	Independentes	Distribuição normal; variâncias diferentes
Teste t pareado	Paramétrico	Quantitativa	Nominal	2	Dependentes	Distribuição normal
ANOVA	Paramétrico	Quantitativa	Nominal	3 ou mais	Independentes	Distribuição normal; variâncias iguais
ANOVA de Welch	Paramétrico	Quantitativa	Nominal	3 ou mais	Independentes	Distribuição normal
ANOVA p/ medidas repetidas	Paramétrico	Quantitativa	Nominal	3 ou mais	Dependentes	Distribuição normal; Esfericidade
Mann-Whitney (Wilcoxon não-pareado)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	2	Independentes	
Wilcoxon (Wilcoxon pareado, teste dos sinais)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	2	Dependentes	
Kruskal-Wallis	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	3 ou mais	Independentes	
Friedman	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	3 ou mais	Dependentes	
Teste p/ 1 proporção (Prova Binomial)	Paramétrico	Nominal	-	1	-	
Teste p/ 2 proporções	Paramétrico	Nominal	Nominal	2	Independentes	
Testes de associação						
Qui-quadrado	Não-Paramétrico	Nominal	Nominal	2 ou mais	Independentes	Células possuem valor esperado > 5
Teste exato de Fisher	Não-Paramétrico	Nominal	Nominal	2	Independentes	

Comparando k amostras (k>2) dependentes



Anova para medidas repetidas

- **Para que serve?**
  Serve para comparar as medias que se obtém de diferentes medidas (mais de duas) de um único grupo.
- **Porque não podemos usar uma ANOVA?**
  Porque se viola a hipótese de independência.

- **Qual dos seguintes é um exemplo de amostras dependentes:**
  - **A)** Administrar um medicamento ativo a um grupo de animais e um placebo inativo a um segundo grupo e comparar as pressões arteriais entre os grupos.
  - **B)** Amostrar a pressão arterial dos mesmos animais antes e depois de receber uma dose do medicamento.

Tipos de estudos

- **Tempo**: Quando se quer avaliar a efetividade de um tratamento.
  - Pode se medir antes, durante e depois da intervenção.
- **Condição**: Quando se quer medir a mesma coisa em diferentes estímulos ou contextos.
  - Ritmo cardíaco em condição estática, exercício leve, exercício moderado.

Pressuposições da ANOVA de medidas repetidas

- **Normalidade**
  - Kolmogorov Smoirnoff ou Shapiro Wilks
- **Esfericidade**
  - As variâncias das diferenças entre todos os pares de medidas devem ser similares.
  - Se comprova com a prova de esfericidade de Mauchly.

Que acontece se não se cumprem os pressupostos?

- **Normalidade**
  - A ANOVA é flexível frente a esta violação, pois pode manter o erro tipo 1 ajustado.
  - <u>Alternativas</u>
    - Eliminar valores extremos
    - Transformar os dados
    - Usar a prova de Friedman (Não parametrica)

Huynh y Feldt (1970) e Rouanet y Lepine (1970)

O que acontece se não se cumprem os pressupostos?

- **Esfericidade**
  - <u>Alternativas</u>
    - Se os dados violam a condição de esfericidade, há algumas correções que se podem aplicar para conseguir um valor de F adequado ajustando os graus de liberdade:
      - Correção de Greenhouse-Geisser.
      - Correção de Huynh-Feldt.

Como se faz?

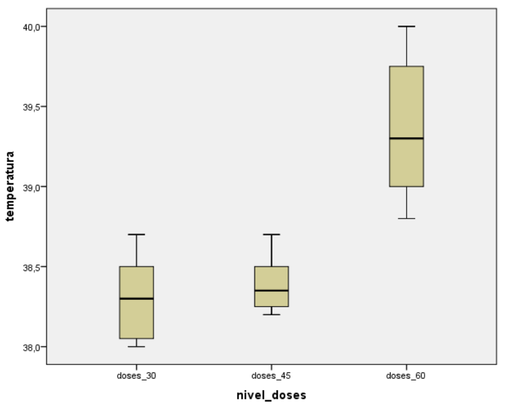
Exemplo

- Um investigador quer avaliar se a doses de certa droga causam efeitos na temperatura dos cães. Para isso selecionou uma amostra de cães e administrou uma doses da droga em 3 ocasiões distintas, com 3 diferentes quantidades (30mg/kg, 45mg/kg e 60mg/kg). Ele mediu a temperatura em cada ocasião e agora quer saber se há diferença entre o valor da temperatura nas 3 ocasiões.

Banco de dados

cão	Doses_30mg	Doses_45mg	Doses_60mg
1	38,2	38,3	38,8
2	38,4	38,4	38,9
3	38,5	38,5	39,2
3	38	38,2	39,1
4	38,1	38,3	39,4
5	38,5	38,5	40
6	38,7	38,7	39,7
7	38	38,2	39,4
8	38,2	38,2	39,8

- Que tipo de variáveis e quantas condições temos no estudo:
  - **A)** variáveis categóricas ordinais e 3 condições
  - **B)** variáveis categóricas nominais e 2 condições
  - **C)** variáveis quantitativas discretas e 3 condições
  - **D)** variáveis quantitativas continuas e 3 condições



Método

- Premissas:
  - Normalidade
    - H0 = Os dados seguem uma distribuição normal.
    - H1 = Os dados não seguem uma distribuição normal.

Distribuição Normal

- As amostras para as 3 condições seguem uma distribuição normal (Shapiro-Wilk - $p > 0,05$)

Método

- Premissas:
  - Normalidade
    - H0 = Os dados seguem uma distribuição normal.
    - H1 = Os dados não seguem uma distribuição normal.
  - Esfericidade
    - H0 = Há homogeneidade na diferença das variâncias entre todos os pares de medidas.
    - H1 = Não há homogeneidade na diferença das variâncias entre todos os pares de medidas.

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
doses	,128	12,335	2	,002	,534	,552	,500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: doses

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

De acordo com o teste, o nível critico associado ao estatístico W (sig. = 0,002) é menor que 0.05, então rejeitamos a hipótese nula de esfericidade.

No caso de rejeição da esfericidade, podemos usar o estatístico F univariado aplicando um índice corretor chamado de "épsilon".

Este índice expressa o grau em que a matriz de variâncias -covariâncias se afasta da esfericidade. (Em condições de esfericidade perfeita os valores de épsilon são iguais a 1)

Para poder utilizar o estatístico F em condições de não esfericidade é necessário corrigir os graus de liberdade de F multiplicando pelo valor de épsilon.

Método

- Premissas:
  - Normalidade
    - H0 = Os dados seguem uma distribuição normal.
    - H1 = Os dados não seguem uma distribuição normal.
  - Esfericidade
    - H0 = Há homogeneidade na diferença das variâncias entre todos os pares de medidas.
    - H1 = Não há homogeneidade na diferença das variâncias entre todos os pares de medidas.

Método

- Formulação da Hipótese
  - H0: As medias das temperaturas dos cães é a mesma nas 3 diferentes doses.
  - H1: As medias das temperaturas dos cães não é a mesma nas 3 diferentes doses.

Como funciona?

cão	Doses_30mg	Doses_45mg	Doses_60mg
1	38,2	38,3	38,8
2	38,4	38,4	38,9
3	38,5	38,5	39,2
3	38	38,2	39,1
4	38,1	38,3	39,4
5	38,5	38,5	40
6	38,7	38,7	39,7
7	38	38,2	39,4
8	38,2	38,2	39,8

$$SQ_I = \sum S^2{}_n (n_n - 1)$$

$$SQ_I = 0{,}1(2) + 0{,}08(2) + 0{,}16(2) + 0{,}49(2) + 0{,}75(2) + 0{,}33(2) + 0{,}57(2) + 0{,}85(2) + = 6{,}7$$

$$GL_I = (n-1)\,\#\,pessoas$$

$$GL_I = (3-1)8 = 16$$

$$SQ_M = \sum n_K (\overline{x_K} - x_{total})^2$$

$$SQ_M = 8(38{,}325 - 38{,}70)^2 + 8(38{,}387 - 38{,}70)^2 + 8(39{,}4 - 38{,}70)^2 = 5{,}826$$

$$GL_M = (n-1)$$

$$GL_M = (3-1) = 2$$

$$SQR = SQ_I - SQ_M$$

$$SQR = 6{,}7 - 5{,}826 = 0{,}874$$

$$GLR - GL_I - GL_M$$

$$GLR = 16 - 2 = 14$$

Como funciona?

$$F = \frac{MQ_M}{MQ_R}$$

$$MQ_M = \frac{SQ_M}{GL_M} \qquad MQ_R = \frac{SQ_R}{GL_R}$$

$$MQ_M = \frac{5{,}825}{2} = 2{,}91 \qquad MQ_R = \frac{0{,}874}{14} = 0{,}062$$

$$F = \frac{2{,}91}{0{,}062} = 46{,}651$$

Greenhouse-Geisser
0,534

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
doses	Sphericity Assumed	5,826	2	2,913	46,651	,000
	Greenhouse-Geisser	5,826	1,068	5,453	46,651	,000
	Huynh-Feldt	5,826	1,104	5,278	46,651	,000
	Lower-bound	5,826	1,000	5,826	46,651	,000
Error(doses)	Sphericity Assumed	,874	14	,062		
	Greenhouse-Geisser	,874	7,479	,117		
	Huynh-Feldt	,874	7,726	,113		
	Lower-bound	,874	7,000	,125		

As quatro versões do estatístico F chegam na mesma conclusão. Dado que o nível critico (Sig.) é menor que 0,05, podemos rejeitar a hipótese de igualdade das médias e concluir que a temperatura dos cães não é a mesma nos 3 diferentes doses.

Post hoc

Pairwise Comparisons

Measure: MEASURE_1

(I) doses	(J) doses	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-,063	,032	,285	-,164	,039
	3	-1,075*	,153	,001	-1,554	-,596
2	1	,063	,032	,285	-,039	,164
	3	-1,013*	,149	,001	-1,480	-,545
3	1	1,075*	,153	,001	,596	1,554
	2	1,013*	,149	,001	,545	1,480

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Resultado final

- A media de temperatura dos cães se vê afetada pela doses do medicamento administrado, $F(1.068, 5.453) = 46.651$, $p<0,05$.
- Resultado final se reporta desde uma aproximação uni-variada com um ajuste de Greenhouse-Geisser (0.534).

Resultado final

- Nas comparações por pares, não existe diferença estatisticamente significativa na media de temperatura dos cães com uma dose de 30mg e 45mg ($p<0,05$); Entretanto, existe diferença entre a temperatura dos cães com uma doses de 30mg e 60mg ($p<0,05$) e entre 45mg e 60mg ($p<0,05$).

E se nossos dados não cumprem as premissas ou não são adequados para a ANOVA de medida repetidas?

Friedman

- Para que serve?

Serve para testar a hipótese de que existe diferença entre mais de dois tratamentos, com base em uma amostra de grupos dependentes.

Friedman

- Para que serve?

Serve para testar a hipótese de que existe diferença entre mais de dois tratamentos, com base em uma amostra de grupos dependentes.

O teste pressupõe que a variável em análise seja medida em escala ordinal ou numérica.

- **Quais dos seguintes estudos seriam adequadas para o analise de Friedman**

Estudos que:

- A) avaliaram a glicemia (mg/dl) de um grupo de cães antes e depois de um tratamento.
- B) avaliaram o comportamento (tranquilo e exaltado) de três diferentes grupos de cães (abaixo do peso, peso normal e acima do peso).
- C) avaliaram a glicemia (mg/dl) de um grupo de cães antes, durante e depois de um tratamento.

Como se faz?

Exemplo

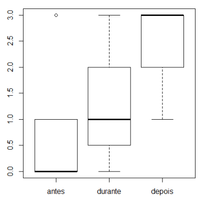
- Para verificar a aprendizagem de seus alunos, um professor de radiologia fez uma perfuração na mandíbula de um crânio canino usando uma broca nº10. Depois radiografou a mandíbula usando técnica digital. Pediu, então, aos seus alunos (8 alunos) que examinassem a radiografia e dissessem se:
  a) não havia lesão
  b) provavelmente não havia lesão
  c) provavelmente havia lesão
  d) havia lesão.

Exemplo

- Os alunos não sabiam, mas era sempre a mesma radiografia e, evidentemente, deveria ser registrada uma lesão. Os alunos examinaram a radiografia antes, durante e depois do curso e fizeram suas observações.
- Para análise, o professor conferiu valores 0, 1, 2 ou 3, de acordo com os critérios: 0: lesão ausente; 1:lesão provavelmente ausente; 2: lesão provavelmente presente; 3: lesão presente.

Aluno	Período do curso		
	Antes	Durante	Depois
Indiv 1	0	1	3
Indiv 2	0	1	2
Indiv 3	0	2	3
Indiv 4	1	0	3
Indiv 5	1	2	3
Indiv 6	0	3	2
Indiv 7	0	1	3
Indiv 8	3	0	1

3 tempos



0: lesão ausente;
1:lesão provavelmente ausente;
2: lesão provavelmente presente;
3: lesão presente.

Método

1. Hipótese e o valor de significância alfa.
   - H0: As notas dadas pelos alunos em períodos distintos têm idêntica distribuição.
   - H1: As notas dadas pelos alunos em períodos distintos não têm idêntica distribuição.
   - Alfa = 0,05

Método

- Para a prova de Friedman, os dados se dispõem em uma tabela de dupla entrada com N linhas e k colunas.
- N = linhas = indivíduos/conjuntos
- K = colunas = condições/tratamentos

Aluno	Período do curso		
	Antes	Durante	Depois
Indiv 1	0	1	3
Indiv 2	0	1	2
Indiv 3	0	2	3
Indiv 4	1	0	3
Indiv 5	1	2	3
Indiv 6	0	3	2
Indiv 7	0	1	3
Indiv 8	3	0	1

3 tempos

0: lesão ausente;
1:lesão provavelmente ausente;
2: lesão provavelmente presente;
3: lesão presente.

Método

2. Atribuição dos postos:
   - Temos que trabalhar por bloco (por linha)
   - O valor mínimo e máximo dos postos depende do numero de tempos que vamos a trabalhar.

Aluno	Período do curso		
	Antes	Durante	Depois
Indiv x	0	2	1
Indiv y	3	1	0

Aluno	Postos		
	Antes	Durante	Depois
Indiv x	1	3	2
Indiv y	3	2	1

   - Temos 3 tempos. Por tanto, os postos em cada linha, vão de 1 (o dado de menor valor) a 3 (o dado de maior valor)

Aluno	Período do curso		
	Antes	Durante	Depois
Indiv 1	0	1	3
Indiv 2	0	1	2
Indiv 3	0	2	3
Indiv 4	1	0	3
Indiv 5	1	2	3
Indiv 6	0	3	2
Indiv 7	0	1	3
Indiv 8	3	0	1

Aluno	Postos		
	Antes	Durante	Depois
Indiv 1	1	2	3
Indiv 2	1	2	3
Indiv 3	1	2	3
Indiv 4	2	1	3
Indiv 5	1	2	3
Indiv 6	1	3	2
Indiv 7	1	2	3
Indiv 8	3	1	2

0: lesão ausente;
1:lesão provavelmente ausente;
2: lesão provavelmente presente;
3: lesão presente.

Aluno	Período do curso		
	Antes	Durante	Depois
Indiv 1	0	1	3
Indiv 2	0	1	2
Indiv 3	0	2	3
Indiv 4	1	0	3
Indiv 5	1	2	3
Indiv 6	0	3	2
Indiv 7	0	1	3
Indiv 8	3	0	1

Aluno	Postos		
	Antes	Durante	Depois
Indiv 1	1	2	3
Indiv 2	1	2	3
Indiv 3	1	2	3
Indiv 4	2	1	3
Indiv 5	1	2	3
Indiv 6	1	3	2
Indiv 7	1	2	3
Indiv 8	3	1	2

0: lesão ausente;
1:lesão provavelmente ausente;
2: lesão provavelmente presente;
3: lesão presente.

- 4. calcular X_r^2 :
  - Conforme a formula

$$X_r^2 = \frac{12}{N\,k\,(k+1)}\left(\sum R_1^2 + \sum R_2^2 + \sum R_3^2\right) - 3N(k+1)$$

$$X_r^2 = \frac{12}{8 \times 3\,(3+1)}\left(11^2 + 15^2 + 22^2\right) - 3 \times 8 \times (3+1)$$

$$X_r^2 = \frac{12}{8 \times 3 \times 4} \times 830 - 3 \times 8 \times 4 = 7{,}75$$

Aluno	Postos		
	Antes	Durante	Depois
Indiv 1	1	2	3
Indiv 2	1	2	3
Indiv 3	1	2	3
Indiv 4	2	1	3
Indiv 5	1	2	3
Indiv 6	1	3	2
Indiv 7	1	2	3
Indiv 8	3	1	2

- Como o valor calculado de x^2 de Friedman é maior do que 5,99 (valor critico da tabela de valores de x^2, com dois graus de liberdade e ao nível de 5% de significância). O professor afirma que houve mudança no grau de conhecimento de seus alunos.

bloco	tratamento	observação
indv 1	antes	0
indv 2	antes	0
indv 3	antes	0
indv 4	antes	1
indv 5	antes	1
indv 6	antes	0
indv 7	antes	0
indv 8	antes	3
indv 1	durante	1
indv 2	durante	1
indv 3	durante	2
indv 4	durante	0
indv 5	durante	2
indv 6	durante	3
indv 7	durante	1
indv 8	durante	0
indv 1	depois	3
indv 2	depois	2
indv 3	depois	3
indv 4	depois	3
indv 5	depois	3
indv 6	depois	2
indv 7	depois	3
indv 8	depois	1

bloco	antes	durante	depois
indv 1	0	1	3
indv 2	0	1	2
indv 3	0	2	3
indv 4	1	0	3
indv 5	1	2	3
indv 6	0	3	2
indv 7	0	1	3
indv 8	3	0	1

Saída do Minitab

Teste de Friedman: observação versus tratamento blocado por bloco

```
S = 7,75  GL = 2  P = 0,021

                  Mediana  Soma dos
tratamento  N       Est     Postos
antes       8     0,000      11,0
depois      8     2,000      22,0
durante     8     1,000      15,0

Mediana global = 1,000
```

Saída do SPSS

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
antes	8	,6250	1,06066	,00	3,00	,0000	,0000	1,0000
durante	8	1,2500	1,03510	,00	3,00	,2500	1,0000	2,0000
depois	8	2,5000	,75593	1,00	3,00	2,0000	3,0000	3,0000

Friedman Test

Ranks

	Mean Rank
antes	1,38
durante	1,88
depois	2,75

Test Statistics[a]

N	8
Chi-Square	7,750
df	2
Asymp. Sig.	,021

a. Friedman Test

Saída do Rcmdr

```
Medians:
  antes  depois durante
      0       3       1

        Friedman rank sum test

data:  .Responses
Friedman chi-squared = 7.75, df = 2, p-value = 0.02075
```

Pergunta?

- O teste de Friedman seria um método alternativo para:
  - A) Wilcoxon
  - B) ANOVA
  - C) ANOVA para medidas repetidas
  - D) Kruskall wallis

Post hoc

Não existe uma técnica estatística não paramétrica exata para comparações de grupos dois a dois, mas nesses casos, é utilizado o teste de **Wilcoxon pareado** ajustado com o método de **Bonferroni** ou o teste de **Dunn**.

Post hoc

Wilcoxon Bonferroni

- Antes – Durante (p = 0.3142)
- Antes – Depois (p = 0.03636)
- Durante – Depois (p = 0.03709)

Teste de Dunn

- Antes – Durante (p = 0.4199)
- Antes – Depois (p = 0.0030)
- Durante – Depois (p = 0.0659)

Conclusão

- Os alunos revelaram ter mais conhecimento do assunto quando terminaram o curso, em relação ao que sabiam no inicio do curso.