Ensaios Fatoriais

Universidade Estadual de Santa Cruz

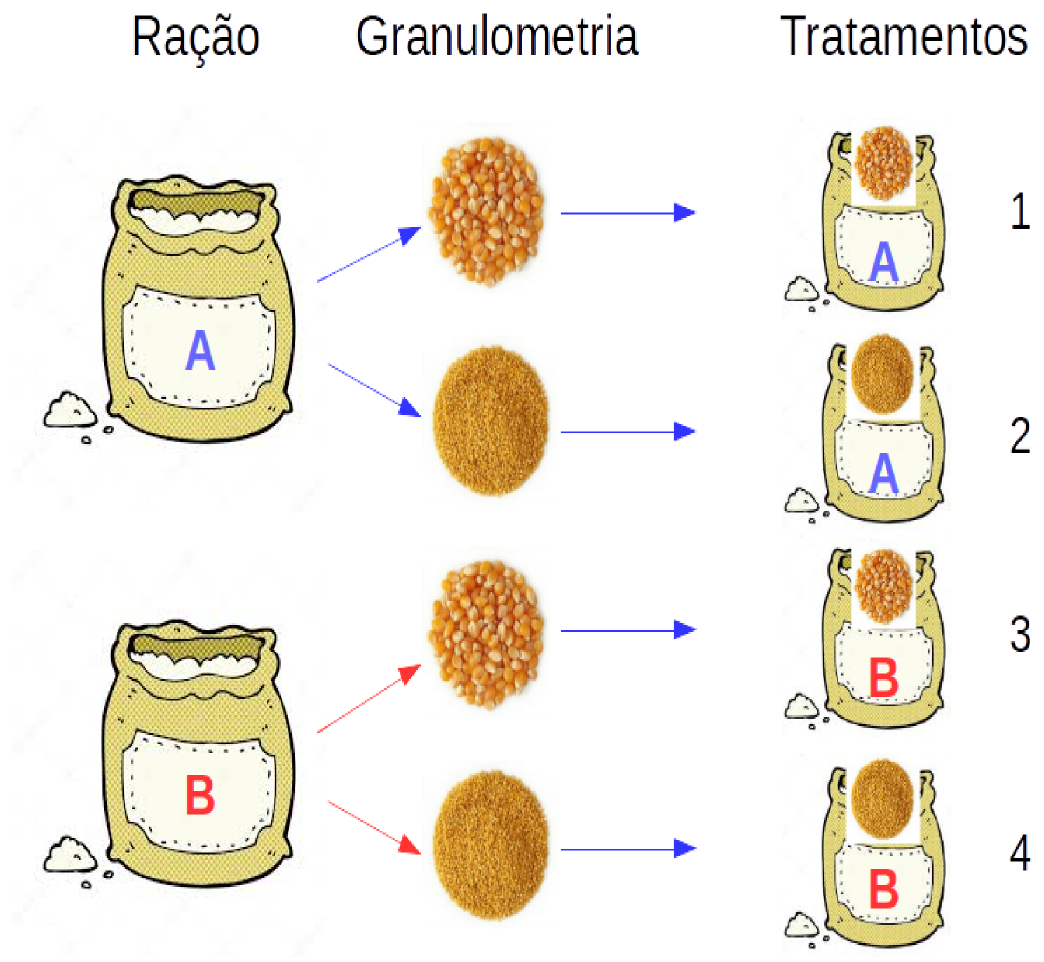
Ivan Bezerra Allaman

INTRODUÇÃO

- As estruturas fatorias são utilizadas quando se tem interesse em avaliar a relação entre dois ou mais fatores no estudo.
- Em uma estrutura fatorial completa, todos os níveis de um fator (digamos fator A) devem aparecer em todos os níveis do outro fator (digamos B).
- Quanto maior o número de fatores e de níveis dos fatores, maior será o número de combinações e mais oneroso se tornará o experimento. Logo, sempre que se pensar em uma estrurua fatorial, deve se pensar primeiro na exequibilidade do experimento.
- Tais estruturas podem ser utilizadas em qualquer um dos delineamentos já vistos (DIC, DBC ou DQL).

VISUALIZAÇÃO DO ESQUEMA FATORIAL

- Vamos supor que um pesquisador esteja interessado em estudar duas marcas de rações e dois diferentes tipos de granulometria da ração no ganho de peso de aves de corte, perfazendo-se um total de 4 tratamentos (2 x 2). Partindo do princípio que as condições biológicas e ambientais são homogêneas de tal forma que não interfira nos resultados da pesquisa, teremos o seguinte esquema:


Ração
Granulometria
Tratamentos
A
B
A
A
B
B
1
2
3
4

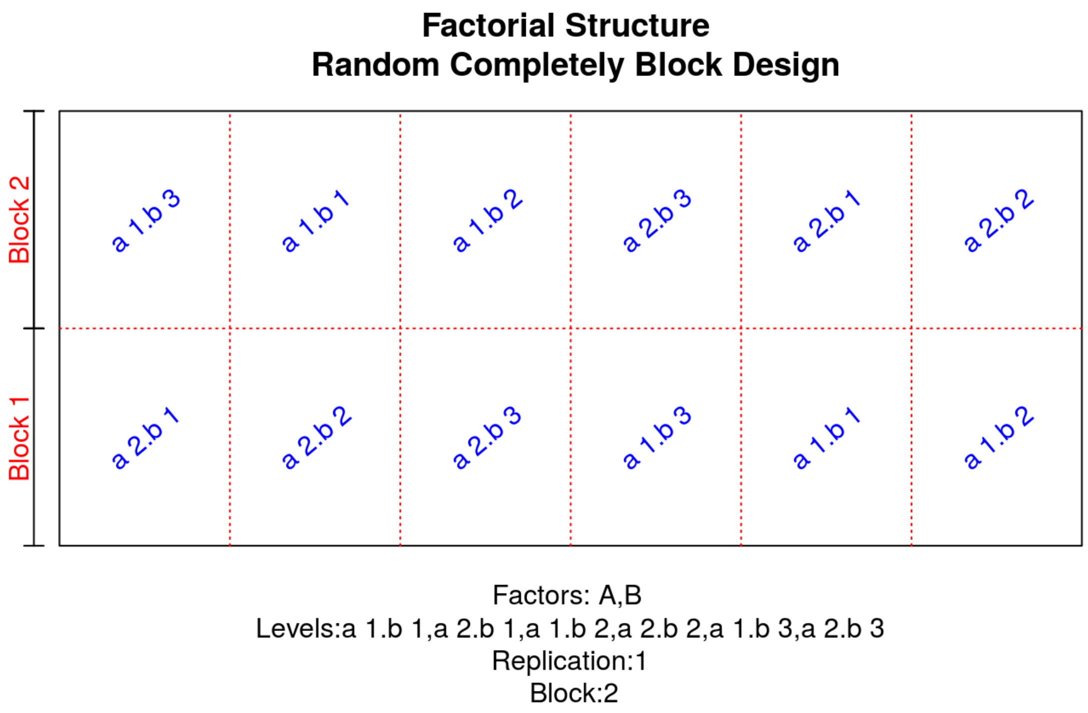
- A maneira com se distribui as combinações nas unidades experimentais irá depender do delineamento adotado.
- Utilizando o pacote `gexp` do `R` podemos facilmente planejar um experimento e ver como se distribui os tratamentos nas unidades experimentais.
- No caso de um esquema fatorial 2x2 em DIC com 3 repetições, tem-se o seguinte planejamento experimental:

Factorial Structure
Completely Random Design

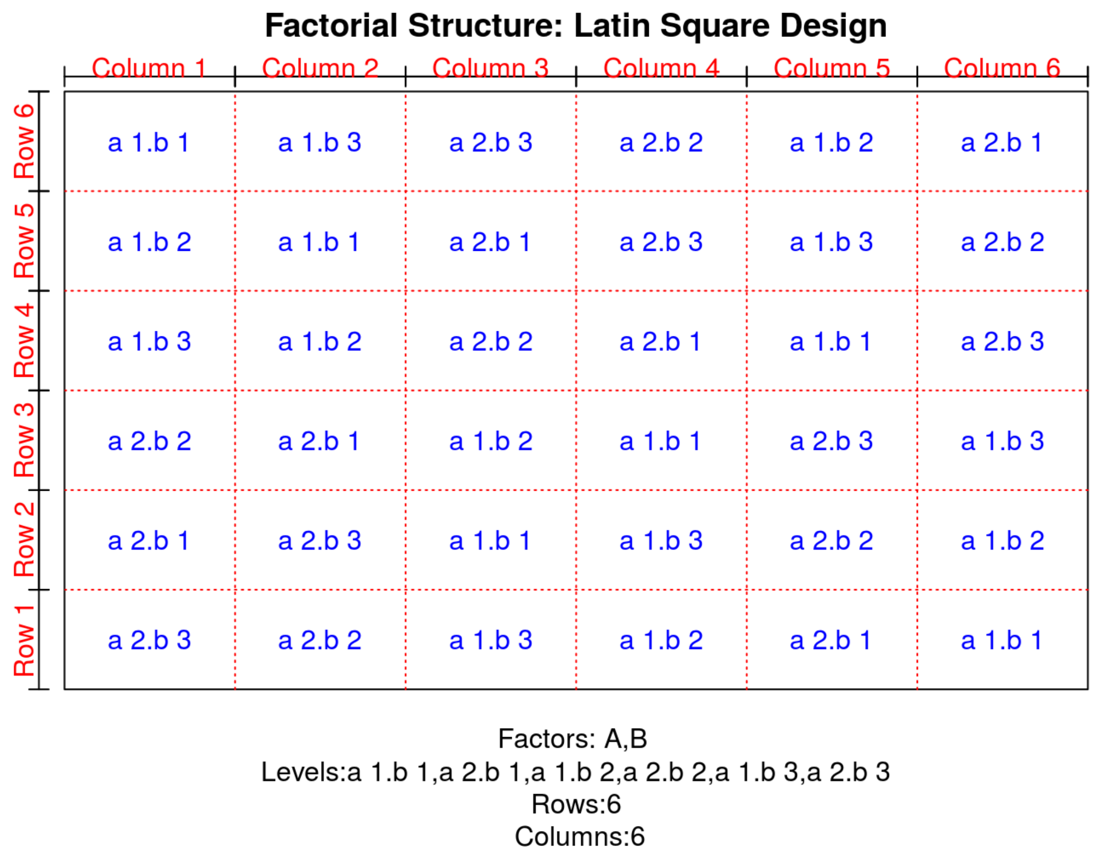
a 2.b 1	a 1.b 2	a 1.b 2
a 2.b 1	a 1.b 1	a 2.b 1
a 1.b 1	a 1.b 1	a 2.b 2
a 2.b 2	a 1.b 2	a 2.b 2

Factors: A,B
Levels:a 1.b 1,a 2.b 1,a 1.b 2,a 2.b 2
Replication:3

- No caso de um esquema fatorial 2x3 em DBC com 2 blocos e 1 repetições tem-se:



- No caso de um esquema fatorial 2x3 em DQL tem-se:



ANÁLISE DE VARIÂNCIA

Modelo estatístico

- O modelo estatístico irá depender do tipo de delineamento adotado e do número de fatores. Ser for um DIC com dois fatores, tem-se:

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk}$$

em que:

- y_{ijk} = é a obsevação que recebeu o nível i do fator τ e o nível j do fator γ na repetição k
- μ = é a média geral associada a todas as observações;
- τ_i = é o efeito do nível **i** do fator τ;
- γ_j = é o efeito do nível **j** do fator γ;
- $\tau\gamma$ = é o efeito da interação entre os fatores $\tau\gamma$;
- ε_{ijk} = é o erro da parcela k que recebeu o i-ésimo nível do fator τ e j-ésimo nível do fator γ.

- Se for um DBC com dois fatores, acrescenta-se no modelo acima apenas o efeito do bloco.

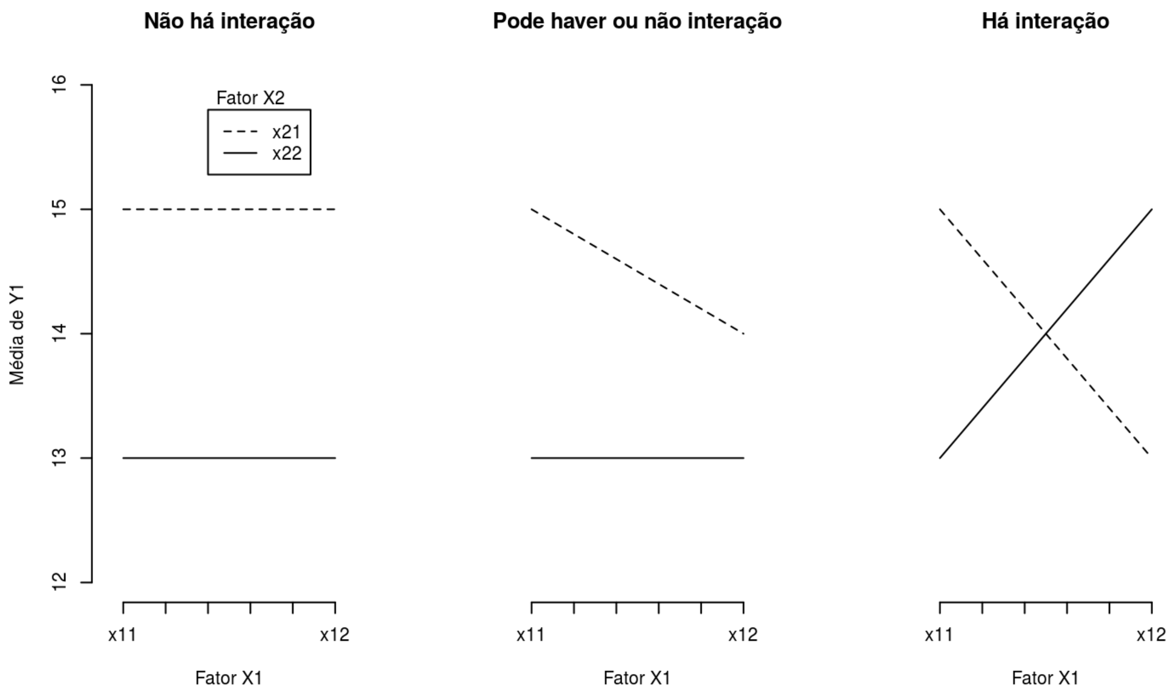
A tabela da ANOVA

- Considerando um DIC com dois fatores, tem-se:

Fontes de Variação	Graus de Liberdade	Soma de Quadrados	Quadrado Médio	F calculado
Fator A	a - 1	SQ_A	$QM_A = SQ_A/(a-1)$	$\frac{QM_A}{QM_{erro}}$
Fator B	b - 1	SQ_B	$QM_B = SQ_B/(b-1)$	$\frac{QM_B}{QM_{erro}}$
Interação (A x B)	(a-1)(b-1)	SQ_{AB}	$QM_{AB} = SQ_{AB}/(a-1)(b-1)$	$\frac{QM_{AB}}{QM_{erro}}$
Erro	ab(r-1)	SQ_{erro}	$QM_{erro} = SQ_{erro}/ab(r-1)$	
Total	abr - 1	SQ_{total}		

- Os pressupostos são os mesmos já abordados no assunto "introdução a análise de variância".

Visualizando os tipos de interações



Interpretação da ANOVA

- Quando se analisa uma estrutura fatorial, devemos primeiramente olhar para a **significância da interação**.
- Quando **não** há significância da interação entendemos que os fatores em estudo atuam de forma independente, ou seja, a influcência de um fator sobre a variável resposta não depende do outro fator em estudo. Nestes casos, deve-se olhar para a significância dos fatores isoladamente.
- Caso ela **seja** significativa, então os fatores em estudo são dependentes um do outro e, portanto, não podemos estudar os efeitos isoladamente. Neste caso, é necessário fazer o que chamamos de "desdobramento" da interação para verificarmos o comportamento dos níveis de um fator na presença de cada nível do outro fator.

Aplicação

1. Foi realizado um estudo cujo o delineamento foi o inteiramente ao acaso com 2 repetições em um esquema fatorial 2 (macho e fêmea) x 3 (90,180 e 270 dias de gestação) para investigar a atividade de "frutose-1-phosphatase-aldolase"(n-moles de substrato metabolizado/min/mg proteína), no terço superior da mucosa intestinal de bezerros, tomados através de secção cesariana em 12 vacas holandesas em 1ª gestação. Segue os dados simulados.

Sexo	Gesta	Rep	Moles
Macho	90	1	1.14
Femea	90	1	0.78
Macho	180	1	0.77
Femea	180	1	0.97
Macho	270	1	1.11
Femea	270	1	1.07
Macho	90	2	1.25
Femea	90	2	0.83
Macho	180	2	0.91
Femea	180	2	0.78
Macho	270	2	1.23
Femea	270	2	1.19

A soma de quadrados dos fatores é semelhante ao que já foi visto anteriormente, com pequenos detalhes nos cálculos. Seja **a** o número de níveis de sexo (2), **b** o número de níveis de gestação (3) e r o número de repetições (2), segue a soma de quadrados do exemplo em questão:

$$SQ_{sexo} = br\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2$$
$$= 3 \cdot 2\{(1,068 - 1,0025)^2 + (0,937 - 1,0025)^2\}$$
$$= 0,05201$$
$$SQ_{Gesta} = ar\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2$$
$$= 2 \cdot 2\{(1,00 - 1,0025)^2 + (0,8575 - 1,0025)^2 + (1,15 - 1,0025)^2\}$$
$$= 0,17115$$

$$SQ_{Interação} = r\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

$$= 2 \cdot \{(1,195 - 1,068 - 1,00 + 1,0025)^2 + (0,84 - 1,068 - 0,8575 + 1,0025)^2 +$$
$$+(1,17 - 1,068 - 1,15 + 1,0025)^2 + \cdots + (1,13 - 0,937 - 1,13 + 1,0025)^2\}$$
$$= 0,1029$$

Logo, tem-se a seguinte tabela da ANOVA:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Sexo	1	0.05	0.05	6.30	0.0459
Gesta	2	0.17	0.09	10.36	0.0113
Sexo:Gesta	2	0.10	0.05	6.23	0.0343
Residuals	6	0.05	0.01		

Logo, considerando um $\alpha = 0,05$ concluímos que há interação entre os fatores Sexo e Gestação, ou seja, a variável mensurada depende do sexo e da idade gestacional. Agora deve-se fazer o desdobramento dos fatores para estudar como cada fator responde em cada nível do outro fator.

Desdobrando (estudando) estágio de gestação dentro de cada nível de sexo. Neste caso, a soma de quadrados é desdobrada nos efeitos que estamos interessados. Neste caso, a soma de quadrados é calculada da seguinte forma:

$$SQ_{Masc/Gesta} = r \cdot \sum_{j=1}^{b} (\bar{y}_{1j} - \bar{y}_{1.})^2$$

$$= 2 \cdot \{(1,195 - 1,068)^2 + (0,840 - 1,068)^2 + (1,17 - 1,068)^2\}$$

$$= 0,157$$

$$SQ_{Fem/Gesta} = r \cdot \sum_{j=1}^{b} (\bar{y}_{2j} - \bar{y}_{2.})^2$$

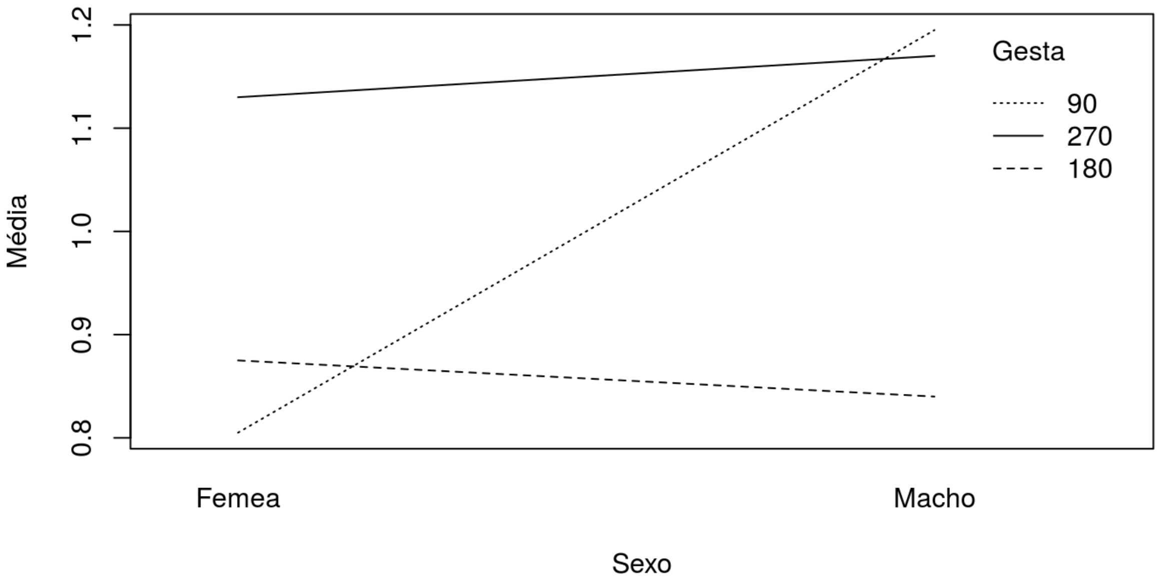
$$= 2 \cdot \{(0,805 - 0,9367)^2 + (0,875 - 0,9367)^2 + (1,13 - 0,9367)^2\}$$

$$= 0,117$$

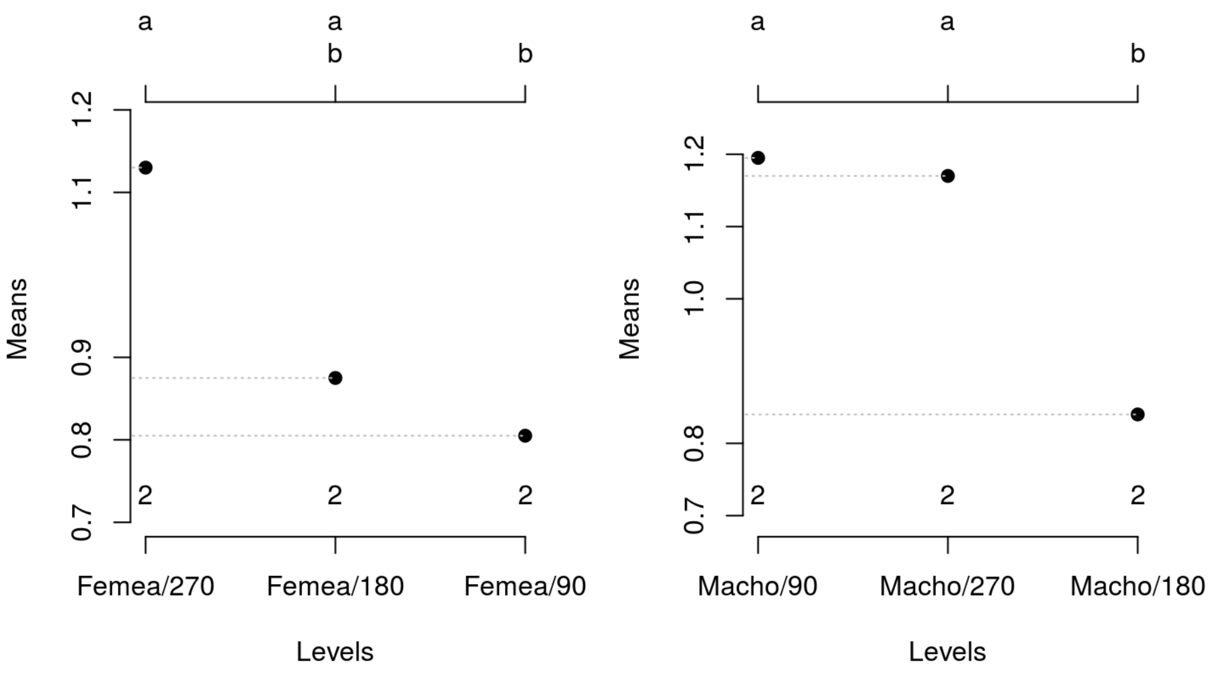
Segue então o quadro da ANOVA.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Sexo	1	0.05	0.05	6.30	0.0459
Sexo:Gesta	4	0.27	0.07	8.30	0.0127
Mac/gesta	2	0.16	0.08	9.51	0.0138
Fem/gesta	2	0.12	0.06	7.09	0.0263
Residuals	6	0.05	0.01		

Houve diferenças significativas de gestação nos dois sexos. Como são três níveis de gestação, será necessário um teste de médias para o ranqueamento. Primeiramente, segue um gráfico para ajudar a entendermos como se comporta os níveis de gestação dentro de cada sexo.



Abaixo o gráfico com o teste de Tukey.


a
a b
b
Means
1.2
1.1
0.9
0.8
0.7
2
2
2
Femea/270
Femea/180
Femea/90
Levels
a
a
b
Means
1.2
1.1
1.0
0.8
0.7
2
2
2
Macho/90
Macho/270
Macho/180
Levels

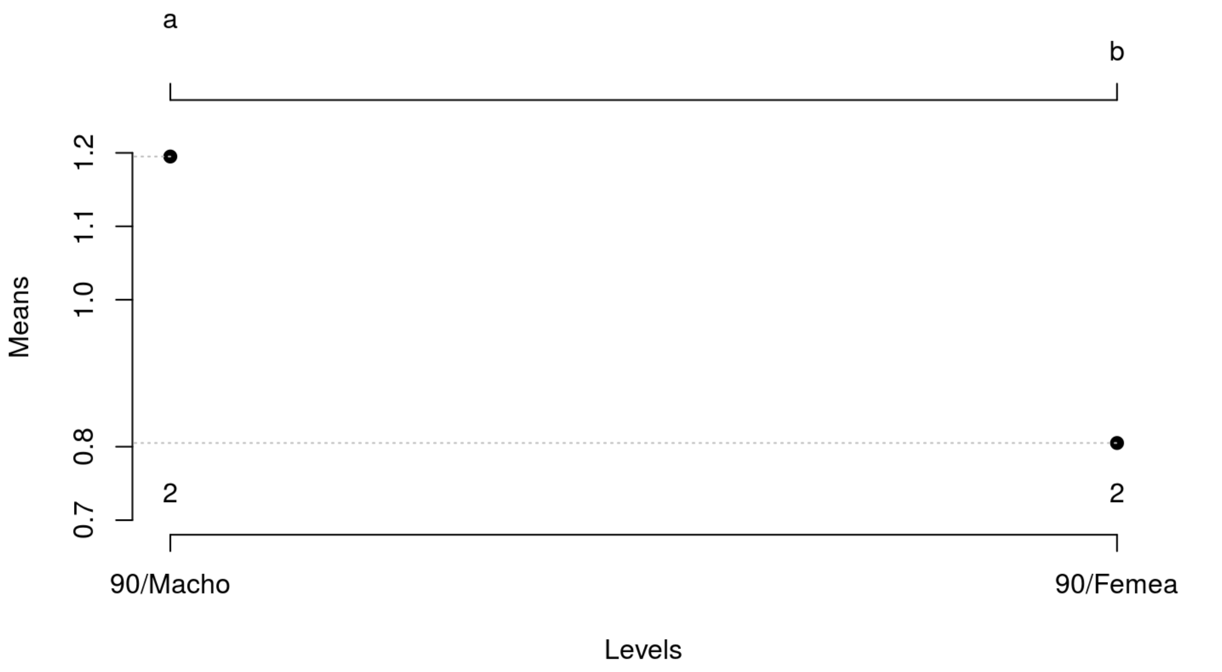
Agora vamos desdobrar sexo dentro de cada nível de estágio de gestação. Segue o cálculo da soma de quadrados.

$$SQ_{90/Sexo} = r \cdot \sum_{i=1}^{a} (\bar{y}_{i1} - \bar{y}_{.1})$$
$$= 2 \cdot \{(1,195 - 1,00)^2 + (0,805 - 1,00)^2\}$$
$$= 0,1521$$
$$SQ_{180/Sexo} = r \cdot \sum_{i=1}^{a} (\bar{y}_{i2} - \bar{y}_{.2})$$
$$= 2 \cdot \{(0,840 - 0,8575)^2 + (0,875 - 0,8575)^2\}$$
$$= 0,00123$$
$$SQ_{270/Sexo} = r \cdot \sum_{i=1}^{a} (\bar{y}_{i3} - \bar{y}_{.3})$$
$$= 2 \cdot \{(1,17 - 1,15)^2 + (1,13 - 1,15)^2\}$$
$$= 0,0016$$

Segue a tabela da ANOVA.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Gesta	2	0.17	0.09	10.36	0.0113
Gesta:Sexo	3	0.15	0.05	6.25	0.0281
90/sexo	1	0.15	0.15	18.42	0.0051
180/sexo	1	0.00	0.00	0.15	0.7134
270/sexo	1	0.00	0.00	0.19	0.6752
Residuals	6	0.05	0.01		

Houve diferenças significativas entre sexo apenas para a gestação 90. Neste caso, como há somente duas médias não há a necessidade de um teste de comparação múltiplas de médias. Segue o gráfico.


a
b
Means
1.2
1.1
1.0
0.8
0.7
2
2
90/Macho
90/Femea
Levels

2. An experiment was conducted to determine the effect of adding two vitamins (I and II) in feed on average daily gain of pigs. Two levels of vitamin I (0 and 4 mg) and two levels of vitamin II (0 and 5 mg) were used. The total sample size was 20 pigs, on which the four combinations of vitamin I and vitamin II were randomly assigned. The following daily gains were measured: http://nbcgib.uesc.br/lec/download/R/dados/kaps_pg318.txt

   As análises detalhadas estão no script nas notas de aula.

3. The objective of this experiment was to determine possible interactions of three types of protein source with increasing energy on milk yield in dairy cows. Three types of protein were used: rape seed + soybean, sunflower + soybean and sunflower + rape seed meal, and two energy levels: standard and increasing level. The base diet was the same for all cows. The following average daily milk yields were measured: http://nbcgib.uesc.br/lec/download/R/dados/kaps_pg322.txt

   As análises detalhadas estão no script nas notas de aula.

4. O consumo diário de ração em kg/dia, no período de crescimento/acabamento de suínos foi observado em um esquema envolvendo tipos de ração e formas de arraçoamento em um delineamento em blocos completos ao acaso. Considerando $\alpha = 0,05$ estudar o consumo médio diário em função dos dois fatores. Os dados estão na página 71 no link http://nbcgib.uesc.br/lec/download/material_didatico/pdf_files/est_experimental/padovani.pdf

   As análises detalhadas estão no script nas notas de aula.