Planejamento de Experimentos
Experimento com um fator aleatório

Enrico A. Colosimo/UFMG

Depto. Estatística - ICEx - UFMG

Modelo Estatístico para Efeitos Fixos

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

em que Y_{ij} é o valor da resposta da j-ésima observação ($j = 1, \ldots, n$) no i-ésimo nível do fator ($i = 1, \ldots, a$); μ: efeito geral da média: τ_i: efeito do i-ésimo tratamento.

- $Var(\epsilon_{ij}) = \sigma^2$
- $Y_{ij} \sim N(\mu + \tau_i, \sigma^2)$

Exemplo típico: Resistência de uma nova fibra sintética para camisas de homens. (Montgomery, 1997)

- Uma companhia têxtil utiliza vários teares para a produção de tecido. A engenharia de processos gostaria de alcançar uma produção mais homogênea de forma a obter uma fibra de resistência uniforme. O engenheiro de produto suspeita que além da variação usual na resistência do tecido, existe uma variação importante entre os teares.
- Para verificar tal hipótese ele selecionou de forma aleatória 4 teares e fez 4 medidas nas fibras provenientes de cada tear.
- Qual é a diferença do exemplo anterior (com 5 porcentagens de algodão)?
- A resposta é a resistência da tecido medida em libra/polegada^2.
- Houve aleatorização?

Exemplo: Resistência da fibra sintética

Tabela: Resistência do tecido.(em libra/polegada^2)

	Réplicas			
Teares	1	2	3	4
1	98	97	99	96
2	91	90	93	92
3	96	95	97	95
4	95	96	95	98
Média	97,5	91,50	95,75	97,00
DP	1,29	1,29	0,96	1,83

Média e desvio-padrão para o conjunto de dados são: 93,44 e 2,73.

- Em muitas situações, o fator de interesse pode ter um grande número de níveis possíveis, e o interesse do analista está em fazer inferências sobre a população inteira de níveis do fator.
- Se a níveis do fator de interesse são selecionados aleatoriamente, então dizemos que o fator é um fator aleatório.
- Neste caso, as conclusões alcançadas serão válidas para toda a população (de níveis do fator de interesse).
- Note que este tipo de situação é diferente daquele em que os níveis do fator são quantidades fixas e conhecidas.

Modelo Estatístico para Um Fator Aleatório

- Formule a pergunta em termos estatísticos.
- $$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

  em que Y_{ij} é o valor da resposta da j-ésima observação ($j = 1, \ldots, n$) no i-ésimo nível do fator ($i = 1, \ldots, a$); μ: efeito geral da média: τ_i: efeito do i-ésimo tratamento.
- $Var(\epsilon_{ij}) = \sigma^2$
- $\tau_i \sim N(0, \sigma_\tau^2)$
- $Y_{ij} \sim N(\mu, \sigma^2 + \sigma_\tau^2)$

- O modelo (linear) estatístico é

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij},$$

  com $i = 1, ..., a$ e $j = 1, ..., n$.

- Aqui, os efeitos dos tratamentos τ_i e os erros ϵ_{ij} são variáveis aleatórias independentes tal que $\tau_i \sim N(0; \sigma_\tau^2)$ e $\epsilon_{ij} \sim N(0; \sigma^2)$.
- As variâncias σ_τ^2 e σ^2 são chamadas de componentes de variância.
- Note que, ao assumirmos $\tau_i \sim N(0; \sigma_\tau^2)$, a suposição usual de $\sum_{i=1}^{a} \tau_i = 0$ dos modelos de efeitos fixos não se aplica ao modelo de efeitos aleatórios.

- Para o modelo de efeitos aleatórios, não faz sentido testar a hipótese de que os efeitos individuais dos tratamentos sejam zero.
- Note que:
  - Se $\sigma^2_\tau = 0$, então todos os tratamentos serão idênticos.
  - Se $\sigma^2_\tau > 0$, então existe variabilidade entre os tratamentos.
- Logo, para o modelo de efeitos aleatórios, as hipóteses de interesse são

$$\begin{cases} H_0 : \sigma^2_\tau = 0 \\ H_1 : \sigma^2_\tau > 0 \end{cases}$$

- A decomposição da ANOVA neste caso ainda é válida, isto é,

$$SQ_{Total} = SQ_T + SQ_E.$$

- Nos modelos de efeitos aleatórios com um fator, os valores esperados de QM_T e QM_E são, respectivamente, dados por

$$E(QM_T) = E\left(\frac{SQ_T}{a-1}\right) = \sigma^2 + n\sigma_\tau^2$$

e

$$E(QM_E) = E\left(\frac{SQ_E}{a(n-1)}\right) = \sigma^2,$$

em que SQ_T e SQ_E são obtidos de maneira idêntica a dos modelos de efeitos fixos.

- Note que, se $H_0 : \sigma_\tau^2 = 0$ é verdade, tanto QM_T quanto QM_E serão bons estimadores para σ^2.
- Pode ser mostrado que QM_T e QM_E são independentes, e que

$$F_0 = \frac{QM_T}{QM_E} \sim F_{a-1;a(n-1)}$$

sob a suposição de que $H_0 : \sigma_\tau^2 = 0$ é verdade.

- Além de testar $H_0 : \sigma^2_\tau = 0$ versus $H_1 : \sigma^2_\tau > 0$, em geral, também há o interesse em estimar-se as componentes de variância, isto é, σ^2_τ e σ^2.
- Obviamente, um estimador natural para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{a(n-1)} = QM_E.$$

- Já σ^2_τ é estimado a partir da as equações

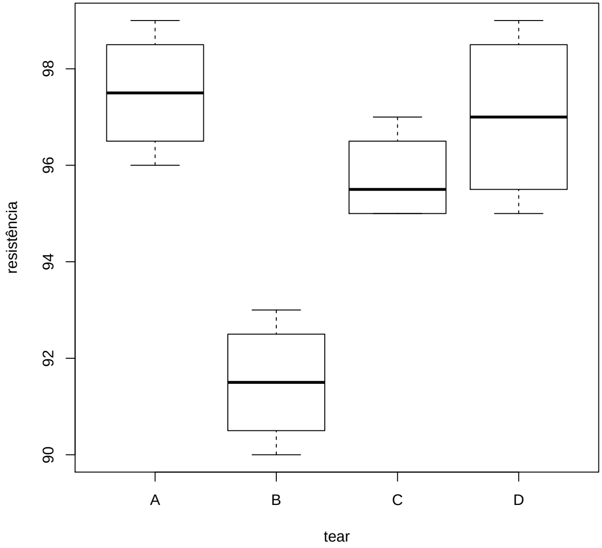
$$\begin{cases} E(QM_E) = \sigma^2 \\ E(QM_T) = \sigma^2 + n\sigma^2_\tau \end{cases}$$

- O método de estimação de σ^2_τ consiste em substituir $E(QM_T)$ por QM_T e $E(QM_E)$ por QM_E nas equações acima, resultando em

$$\hat{\sigma}^2_\tau = \frac{QM_T - QM_E}{n},$$

em que QM_T e QM_E são obtidos diretamente da tabela da ANOVA.

Boxplot - Exemplo



Fontes de variação

- De forma análoga, temos a seguinte tabela de análise de variância.

Tabela: Tabela da Análise de Variância - ANOVA.

Fontes de variação	SQ	GL	QM	*F*	*p-valor*
Entre Teares	89,19	3	$29,73$	15,68	< 0.001
Dentro dos grupos	22,75	12	$1,90$	-	-
Total	111,94	15	-	-	-