ANOVA - ANÁLISE DE VARIÂNCIA

1 Factor, Modelo de Efeito Fixo

OBJECTIVO Comparar I grupos (tratamentos, populações) representados por J indivíduos (observações) cada um.

PRESSUPOSTOS Os I grupos são normais; têm variâncias iguais e as amostras são aleatórias e independentes.

HIPÓTESES H_0: $\mu_i = \mu \quad i = 1, \ldots, I$
H_1: $\mu_i \neq \mu$ para pelo menos um i

NOTAÇÃO $\bar{x}_{i\bullet} = \frac{x_{i\bullet}}{J}$ e $\bar{x} = \frac{x_{\bullet\bullet}}{IJ}$ com: $x_{i\bullet}$ - soma das J observações do grupo i e $x_{\bullet\bullet}$ - soma geral de todas as observações.

MODELO Para cada observação x_{ij}, o seu desvio em relação à média geral $\bar{x}$ será:

$$x_{ij} - \bar{x} = (\bar{x}_{i\bullet} - \bar{x}) + (x_{ij} - \bar{x}_{i\bullet})$$

$$\sum_{ij}(x_{ij} - \bar{x})^2 = J\sum_{i}(\bar{x}_{i\bullet} - \bar{x})^2 + \sum_{ij}(x_{ij} - \bar{x}_{i\bullet})^2$$

$$SQD_t = SQD_\alpha + SQD_r$$

Desvio Total=Desvio entre grupos+Desvio dentro dos grupos

$$SQD_r = \sum_{ij} x_{ij}^2 - \frac{1}{J}\sum_{i} x_{i\bullet}^2 \qquad SQD_\alpha = \frac{1}{J}\sum_{i} x_{i\bullet}^2 - \frac{x_{\bullet\bullet}^2}{IJ} \qquad SQD_t = \sum_{ij} x_{ij}^2 - \frac{x_{\bullet\bullet}^2}{IJ}$$

Tabela ANOVA 1 Factor de efeito fixo

Origem Variação	SQD's	GL's	QM's	Estatistica Teste
Entre Populações (α)	SQD_α	$I-1$	$QM_\alpha = \frac{SQD_\alpha}{GL_\alpha}$	$F_{obs} = \frac{QM_\alpha}{QM_r}$
Dentro Populações (r)	SQD_r	$I(J-1)$	$QM_r = \frac{SQD_r}{GL_r}$	
Total	SQD_t	$IJ-1$		

F é significativo ao nível de significância α se $F_{obs} > F_{crit}$ com $F_{crit} = F^{\alpha}_{GL_n,GL_d} = F^{\alpha}_{GL_\alpha,GL_r}$

N de observações desiguais

Em cada tratamento ou grupo temos ni observações ($i = 1, \ldots, I$), e seja $N = \sum_i ni$

$$SQD_t = \sum_{ij}(x_{ij} - \bar{x})^2 = \sum_{i=1}^{I}\sum_{j=1}^{ni} x_{ij}^2 - \frac{x_{\bullet\bullet}^2}{N} \qquad SQD_\alpha = \sum_{i=1}^{I}\frac{x_{i\bullet}^2}{ni} - \frac{x_{\bullet\bullet}^2}{N}$$

$$SQD_r = SQD_t - SQD_\alpha$$

Tabela ANOVA 1 Factor de efeito fixo, observações desiguais

Origem Variação	SQD's	GL's	QM's	Estatistica Teste
Entre Populações (α)	SQD_α	$I-1$	$QM_\alpha = \frac{SQD_\alpha}{GL_\alpha}$	$F_{obs} = \frac{QM_\alpha}{QM_r}$
Dentro Populações (r)	SQD_r	$N-I$	$QM_r = \frac{SQD_r}{GL_r}$	
Total	SQD_t	$N-1$		

F é significativo ao nível de significância α se $F_{obs} > F_{crit}$ com $F_{crit} = F^{\alpha}_{GL_n,GL_d} = F^{\alpha}_{GL_\alpha,GL_r}$

ANOVA - ANÁLISE DE VARIÂNCIA

2 Factores, Modelo de Efeito Fixo, sem interacção

OBJECTIVO Pretende-se testar a influência de 2 origens de variação "cruzadas"sobre um parâmetro. Temos uma única observação para cada modalidade das duas origens de variação (i e j).

HIPÓTESES $H_0^{(1)}$: $\alpha_i = 0$ $\quad H_0^{(2)}$: $\alpha_i = 0$

$H_1^{(1)}$: nem todos os α_i são 0 $\quad H_1^{(2)}$: nem todos os α_i são 0

MODELO $x_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$

com α_i - efeito da modalidade i; β_j - efeito da modalidade j e ε_{ij} - erro aleatório

$$\begin{cases} \mu = \bar{x} \\ \alpha_i = \bar{x}_{i\bullet} - \bar{x} \\ \beta_j = \bar{x}_{\bullet j} - \bar{x} \\ \varepsilon_{ij} = x_{ij} - \bar{x}_{i\bullet} - \bar{x}_{\bullet j} + \bar{x} \end{cases}$$

$$x_{ij} - \bar{x} = (\bar{x}_{i\bullet} - \bar{x}) + (\bar{x}_{\bullet j} - \bar{x}) + (x_{ij} - \bar{x}_{i\bullet} - \bar{x}_{\bullet j} x\bar{x})$$

$$\sum_{ij}(x_{ij} - \bar{x})^2 = \sum_{ij}(\bar{x}_{i\bullet} - \bar{x})^2 + \sum_{ij}(\bar{x}_{\bullet j} - \bar{x})^2 + \sum_{ij}(x_{ij} - \bar{x}_{i\bullet} - \bar{x}_{\bullet j} x\bar{x})^2$$

$$SQD_t = SQD_\alpha + SQD_\beta + SQD_r$$

$$SQD_t = \sum_{ij} x_{ij}^2 - \frac{x_{\bullet\bullet}^2}{IJ} \qquad SQD_\alpha = \frac{1}{J}\sum_i x_{i\bullet}^2 - \frac{x_{\bullet\bullet}^2}{IJ} \qquad SQD_\beta = \frac{1}{I}\sum_j x_{\bullet j}^2 - \frac{x_{\bullet\bullet}^2}{IJ}$$

$$SQD_r = SQD_t - SQD_\alpha - SQD_\beta$$

Tabela ANOVA 2 factores de efeito fixo, sem interacção

Origem Variação	SQD's	GL's	QM's	Estatistica Teste
α	SQD_α	$I-1$	$QM_\alpha = \frac{SQD_\alpha}{GL_\alpha}$	$F_{obs}^{(\alpha)} = \frac{QM_\alpha}{QM_r}$
β	SQD_β	$J-1$	$QM_\beta = \frac{SQD_\beta}{GL_\beta}$	$F_{obs}^{(\beta)} = \frac{QM_\beta}{QM_r}$
variação residual	SQD_r	$(I-1)(J-1)$	$QM_r = \frac{SQD_r}{GL_r}$	
Total	SQD_t	$IJ-1$		

O efeito de α é declarado significativo ao nível de significância α se $F_{obs}^{(\alpha)} > F_{crit}$ com $F_{crit} = F_{(GL_n, GL_d, \alpha)} = F_{(GL_\alpha, GL_r, \alpha)}$

O efeito de β é declarado significativo ao nível de significância α se $F_{obs}^{(\beta)} > F_{crit}$ com $F_{crit} = F_{(GL_n, GL_d, \alpha)} = F_{GL_{(\beta}, GL_r, \alpha)}$

ANOVA - ANÁLISE DE VARIÂNCIA

2 Factores, Modelo de Efeito Fixo, com interacção

OBJECTIVO: No caso em que dispomos de n valores x_{ij} na tabela ($n > 1$) podemos testar a interacção entre α e β. A existência desta interacção significa que a classificação das modalidades de α não é independente de β.

HIPÓTESES:

$H_0^{(1)}$: $\alpha_i = 0$	$H_0^{(2)}$: $\alpha_i = 0$	$H_0^{(3)}$: $(\alpha\beta)_{ij} = 0$
$H_1^{(1)}$: nem todos os α_i são 0	$H_1^{(2)}$: nem todos os β_i são 0	$H_1^{(3)}$: nem todos os $(\alpha\beta)_{ij}$ são 0

MODELO:

$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

com α_i - efeito da modalidade i; β_j - efeito da modalidade j; k - índice de repetição em cada caso $k = 1, \ldots, n$;$(\alpha\beta)_{ij}$ - termo de interacção e ε_{ij} - erro aleatório

$$\begin{cases} \mu = \bar{x} \\ \alpha_i = \bar{x}_{i\bullet} - \bar{x} \\ \beta_j = \bar{x}_{\bullet j} - \bar{x} \\ (\alpha\beta)_{ij} = \bar{x}_{ij\bullet} - \bar{x}_{i\bullet\bullet} - \bar{x}_{\bullet j\bullet} + \bar{x} \\ \varepsilon_{ij} = x_{ijk} - \bar{x}_{ij\bullet} \end{cases}$$

$$\sum_{ijk}(x_{ijk} - \bar{x})^2 = \sum_{ijk}(\bar{x}_{i\bullet\bullet} - \bar{x})^2 + \sum_{ijk}(\bar{x}_{\bullet j\bullet} - \bar{x})^2 + \sum_{ijk}(\bar{x}_{ij\bullet} - \bar{x}_{i\bullet\bullet} - \bar{x}_{\bullet j\bullet} + \bar{x})^2 + \sum_{ijk}(x_{ijk} - \bar{x}_{ij\bullet})$$

$$SQD_t = \sum_{ijk} x_{ijk}^2 - \frac{x_{\bullet\bullet\bullet}^2}{IJK} \qquad SQD_r = \sum_{ijk} x_{ijk}^2 - \frac{1}{K}\sum_{ij} x_{ij\bullet}^2$$

$$SQD_\alpha = \tfrac{1}{JK}\sum_i x_{i\bullet\bullet}^2 - \frac{x_{\bullet\bullet\bullet}^2}{IJK} \qquad SQD_\beta = \tfrac{1}{IK}\sum_j x_{\bullet j\bullet}^2 - \frac{x_{\bullet\bullet\bullet}^2}{IJK}$$

$$SQD_{(\alpha\beta)} = SQD_t - SQD_\alpha - SQD_\beta - SQD_r$$

Tabela ANOVA 2 factores de efeito fixo, com interacção

Origem Variação	SQD's	GL's	QM's	Estatistica Teste
α	SQD_α	$I-1$	$QM_\alpha = \frac{SQD_\alpha}{GL_\alpha}$	$F_{obs}^{(\alpha)} = \frac{QM_\alpha}{QM_r}$
β	SQD_β	$J-1$	$QM_\beta = \frac{SQD_\beta}{GL_\beta}$	$F_{obs}^{(\beta)} = \frac{QM_\beta}{QM_r}$
interacção $\alpha\beta$	$SQD_{(\alpha\beta)}$	$(I-1)(J-1)$	$QM_{\alpha\beta} = \frac{SQD_{\alpha\beta}}{GL_{\alpha\beta}}$	$F_{obs}^{(\alpha\beta)} = \frac{QM_{\alpha\beta}}{QM_r}$
variação residual	SQD_r	$IJ(K-1)$	$QM_r = \frac{SQD_r}{GL_r}$	
Total	SQD_t	$IJK-1$		

O efeito de $\alpha\beta$ é declarado significativo ao nível de significância α se $F_{obs}^{(\alpha\beta)} > F_{crit}$ com $F_{crit} = F_{(GL_n, GL_d, \alpha)} = F_{(GL_{\alpha\beta}, GL_r, \alpha)}$

O efeito de α é declarado significativo ao nível de significância α se $F_{obs}^{(\alpha)} > F_{crit}$ com $F_{crit} = F_{(GL_n, GL_d, \alpha)} = F_{(GL_\alpha, GL_r, \alpha)}$

O efeito de β é declarado significativo ao nível de significância α se $F_{obs}^{(\beta)} > F_{crit}$ com $F_{crit} = F_{(GL_n, GL_d, \alpha)} = F_{(GL_\beta, GL_r, \alpha)}$

Caso I $H_0^{(3)}$ não pode ser rejeitada. Neste caso, as interacções não são muito grandes. Podemos então testar $H_0^{(1)}$e $H_0^{(2)}$ usando os F_{obs} conforme a tabela.

Caso II $H_0^{(3)}$ pode ser rejeitada. As interacções são significativamente grandes. As diferenças nos factores são importantes se forem grandes comparados com a interacção.

ANOVA - ANÁLISE DE VARIÂNCIA

Classificação de médias: comparação de pares de grupos

OBJECTIVO Quando H_0 em ANOVA é rejeitada e queremos testar:

HIPÓTESES $H_0: \ \mu_{i1} = \mu_{j2}$ para qualquer $i \neq j$

CÁLCULO Calcula-se a um nível de significância α o desvio mínimo que deve existir entre as médias dos grupos para as declararmos estatisticamente diferentes.

Caso de Efectivos Desiguais

Método do t

$$ds_{i1,i2} = t_{(\frac{\alpha}{I(I-1)};GL_r)} \times \sqrt{QM_r \times (\frac{1}{n_{i1}} + \frac{1}{n_{i2}})}$$

$ds_{i1,i2}$-diferença mínima entre 2 grupos $i1$ e $i2$, ou diferença significativa

t-valor do t de Student para um nível $\alpha/I(I-1)$ e para um n de graus de liberdade referente à diferença no interior dos grupos(QM_r)

QM_r-quadrado médio no interior dos grupos; corresponde à variância no interior dos grupos.

n_{i1} e n_{i2}-n de efectivos dos 2 grupos $i1$ e $i2$ que queremos comparar.

Caso de Efectivos Iguais

Método do t corrigido (temos o mesmo valor de ds para o conjunto de todas as comparações entre grupos)

$$ds = t_{(\frac{\alpha}{I(I-1)};GL_r)} \times \sqrt{QM_r \times \frac{2}{n}}$$

Teste de TUKEY

Este teste é utilizado especialmente para comparações de médias com efectivos iguais.

$$ds = T_{(\alpha;I;GL_r)} \times \sqrt{\frac{QM_r}{n}}$$

T-valor do teste de Tukey dado na tabela respectiva sendo α o risco do teste (em geral 0,05 ou 0,01).

I-n de grupos.