UMA DEMONSTRAÇÃO GEOMÉTRICA PARA UMA IDENTIDADE DE FISHER PARA O MODELO DE DOIS FATORES

Fernanda Gomes da SILVEIRA[1]
Luzia Aparecida da COSTA[1]
Leandro da Silva PEREIRA[1]
Lucas Monteiro CHAVES[1]
Devanil Jaques de SOUZA[1]

1 Introdução

"A análise de variância não é um teorema matemático, mas um método simples de se arranjar a aritmética de forma a isolar e ressaltar os fatos fundamentais contidos no conjunto de dados".

Esta afirmação de Fisher em uma carta a Snedecor (Searle, 1987) coloca o problema de como obter esta decomposição simples da aritmética. Como exemplo considere um experimento fatorial com modelo $y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$ em que $i = 1, 2, ..., I$ e $j = 1, 2, ..., J$ indexam os níveis de dois fatores F e G, e $k = 1, 2, ..., K$ indexa as repetições. Fisher observou a identidade elementar:

$$\left(y_{ijk} - \overline{y}_{\bullet\bullet\bullet}\right) = \left[\left(\overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}\right) + \left(\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}\right) + \left(\overline{y}_{ij\bullet} - \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet j\bullet} + \overline{y}_{\bullet\bullet\bullet}\right) + \left(y_{ijk} - \overline{y}_{ij\bullet}\right)\right]. \quad (1)$$

Tal identidade tem uma justificativa estatística razoável uma vez que os termos estão relacionados com a diferença de médias das várias hipóteses possíveis como médias dos tratamentos, médias dos efeitos de cada fator individualmente, média geral. Mas tal fato por si só não justifica a identidade, pois certamente Fisher procurou uma identidade que tivesse a propriedade de uma decomposição em quadrados, como de fato ocorre, pois:

$$\sum_{i,j,k}\left(y_{ijk} - \overline{y}_{\bullet\bullet\bullet}\right)^2 = \sum_{i,j,k}\left[\left(\overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}\right)^2 + \left(\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}\right)^2 + \left(\overline{y}_{ij\bullet} - \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet j\bullet} + \overline{y}_{\bullet\bullet\bullet}\right)^2 + \left(y_{ijk} - \overline{y}_{ij\bullet}\right)^2\right]. \quad (2)$$

[1] Universidade Federal de Lavras – UFLA, Departamento de Ciências Exatas, Caixa Postal 3037, CEP: 37200-000. Lavras, MG, Brasil. E-mail: *fernandaphn@yahoo.com.br* / *luziamatematica@hotmail.com* / *lesptec@bol.com.br* / *Lucas@dex.ufla.br* / *DevanilJaques@dex.ufla.br*

Como ele visualizou tal fato? A demonstração algébrica desta identidade entre quadrados é no mínimo assustadora, bem como a demonstração matricial (Apêndice II). Portanto pode-se supor que a álgebra não foi a fonte de inspiração para a obtenção desta identidade e, generalizando, talvez não seja a fonte de inspiração para a maioria das identidades envolvendo soma de quadrados que aparecem de forma recorrente em todas as áreas da estatística. A fonte para tais identidades parece ser, como dito na biografia de Fisher, escrita por sua filha Joan Fisher Box (Saville e Wood, 1991), o fato de Fisher ter sido um excelente geômetra, e era pela via da geometria que obtinha seus resultados originais. A álgebra aparece mais como uma necessidade prática, pois como tinha dificuldade de explicar seus intrincados raciocínios geométricos, teve de recorrer à álgebra para ser entendido. Pode-se dizer então que a geometria tem direitos a serem reivindicados na teoria e no ensino da Estatística em geral. Um exemplo muito bem sucedido é a abordagem geométrica para o estudo dos delineamentos experimentais adotada no livro de Bailey (2008). No entanto a abordagem algébrica, incluindo aí a álgebra de matrizes, se tornou hegemônica com suas vantagens e certamente também com suas desvantagens. O fato é que uma abordagem geométrica da estatística é no mínimo um procedimento válido e que pode ser vantajoso para uma ampla gama de futuros estatísticos, principalmente aqueles com formação mais matemática. Neste trabalho uma demonstração geométrica para a identidade de Fisher é apresentada, baseada nos conceitos geométricos desenvolvidos em Bailey (2008). No Apêndice I é apresentada a demonstração algébrica. Uma demonstração bastante pormenorizada, utilizando matrizes, é apresentada no Apêndice II. Tais demonstrações não foram localizadas, pelos autores, na literatura. Além disso, no Apêndice II ainda são obtidas, utilizando-se o teorema de Fisher-Cochran, as distribuições de cada um dos termos da decomposição.

2 Geometria dos delineamentos fatoriais

Considere o espaço vetorial $\mathbb{R}^n$ em que n é igual ao produto IJK. Vamos ordenar as respostas y_{ijk} de tal forma que cada resposta corresponda a uma coordenada e para isto será utilizada a ordem lexicográfica em relação a i, j, k. Desta forma as respostas podem ser consideradas como um vetor

$$\vec{y} = (y_{111}, y_{112}, \ldots, y_{11K}, y_{121}, y_{122}, \ldots, y_{12K}, \ldots, y_{1J1}, \ldots, \\ y_{1JK}, y_{211}, y_{212}, \ldots, y_{21K}, \ldots, y_{IJ1}, \ldots, y_{IJK}).$$

Utilizando a notação usual de modelos lineares têm-se as somas e médias em relação aos vários índices, como por exemplo:

$$y_{...} = \sum_i \sum_j \sum_k y_{ijk} \qquad \bar{y}_{...} = \frac{1}{IJK} \sum_i \sum_j \sum_k y_{ijk}$$

$$y_{i..} = \sum_j \sum_k y_{ijk} \qquad \bar{y}_{i..} = \frac{1}{JK} \sum_j \sum_k y_{ijk}.$$

Da mesma forma pode-se então construir vetores n dimensionais em relação a estas várias somas e médias:

$$\overrightarrow{y}_{...} = (y_{...}, y_{...}, ..., y_{...}, y_{...}) \qquad \overline{\overrightarrow{y}}_{...} = (\overline{y}_{...}, ..., \overline{y}_{...}, ..., \overline{y}_{...}, \overline{y}_{...})$$

$$\overrightarrow{y}_{.jk} = (y_{.11}, y_{.12}, y_{.13}, ..., y_{.1K}, y_{.21}, y_{.22}, y_{.23}, ...y_{.2K}, ..., y_{.J1}, y_{.J2}, ..., y_{.JK}).$$

Considere o subespaço vetorial de $\mathbb{R}^n$ constituído pelos vetores que, para tratamentos iguais, possuem coordenadas iguais. Tais vetores, representando respostas, ocorreriam caso não houvesse a aleatoriedade experimental pois, repetições de um mesmo tratamento deveriam ter respostas iguais. Tal subespaço será denominado espaço de tratamentos e denotado por:

$$V_T = \{ (\underbrace{a,a,...a}_{K}, \underbrace{b,b,...b}_{K}, \underbrace{c,c,...c}_{K}, ..., \underbrace{d,d,...d}_{K}), a,b,c,...,d \in \mathbb{R} \};$$

em que temos IJ sequências com respostas iguais. Tem-se que a dimensão de $V_T = IJ$.

Pode-se agora considerar os subespaços vetoriais V_F formado pelos vetores que possuem coordenadas iguais para entradas com o mesmo índice i, isto é,

$$V_F = \{ (\underbrace{a, a, \ldots, a}_{JK}, \underbrace{b, b, \ldots, b}_{JK}, \underbrace{c, c, \ldots, c}_{JK}, \ldots, \underbrace{d, d, \ldots d}_{JK}) \};$$

em que cada uma das sequências de letras iguais é de tamanho JK e se tem I sequências.

A dimensão de V_I é I . Da mesma forma, o subespaço V_G é formado pelos vetores que possuem coordenadas iguais para entradas com o mesmo índice j, isto é,

$$V_G = \{ (\underbrace{\underbrace{a, a, \ldots, a}_{K}, \underbrace{b, b, \ldots, b}_{K}, \underbrace{c, c \ldots, c}_{K}, \ldots, \underbrace{d, d, \ldots d}_{K}}_{J},$$

$$\underbrace{\underbrace{a, a, \ldots, a}_{K}, \underbrace{b, b, \ldots, b}_{K}, \underbrace{c, c \ldots, c}_{K}, \ldots, \underbrace{d, d, \ldots d}_{K}}_{J}, \ldots,$$

$$\underbrace{\underbrace{a, a, \ldots, a}_{K}, \underbrace{b, b, \ldots, b}_{K}, \underbrace{c, c \ldots, c}_{K}, \ldots, \underbrace{d, d, \ldots d}_{K}}_{J}) \}.$$

A dimensão de V_G é J . Note que $V_F \cap V_G = V_0 = \{(a,a,a,...,a) \ , \ a \in \mathbb{R}\}$. Sejam W_F e W_G subespaços de V_F e V_G definido pelos vetores ortogonais ao subespaço V_0. Tem-se que:

$$W_F = \{(a,a,...,a,b,b,...,b,....,c,c....,c,d,d,...d), JKa + JKb + ... + JKc + JKd = 0\}.$$

$$W_G = \{(a,a,...,a,b,b,...,b,....,c,c....,c,d,d,...d,a,a,...,a,b,b,...,b,...$$
$$...,c,c....,c,d,d,...d,...,a,a,...,a,b,b,...,b,....,c,c....,c,d,d,...$$
$$...d), Ka+Kb+...+Kc+Kd+Ka+Kb+...+Kc+Kd+...+$$
$$+Ka+Kb+...+Kc+Kd = IKa+IKb+...+IKc+IKd = 0\}.$$

Em termos vetoriais as equações (1) e (2) são expressas por:

$$\vec{y}-\vec{\bar{y}}_{...} = \left(\vec{\bar{y}}_{i..}-\vec{\bar{y}}_{...}\right)+\left(\vec{\bar{y}}_{.j.}-\vec{\bar{y}}_{...}\right)+\left[\left(\vec{\bar{y}}_{ij.}-\vec{\bar{y}}_{i..}\right)-\left(\vec{\bar{y}}_{.j.}-\vec{\bar{y}}_{...}\right)\right]+\left(\vec{y}-\vec{\bar{y}}_{ij.}\right). \quad (3)$$

$$\left\|\vec{y}-\vec{\bar{y}}_{...}\right\|^2 = \left\|\left(\vec{\bar{y}}_{i..}-\vec{\bar{y}}_{...}\right)\right\|^2+\left\|\left(\vec{\bar{y}}_{.j.}-\vec{\bar{y}}_{...}\right)\right\|^2+\left\|\left[\left(\vec{\bar{y}}_{ij.}-\vec{\bar{y}}_{i..}\right)-\left(\vec{\bar{y}}_{.j.}-\vec{\bar{y}}_{...}\right)\right]\right\|^2+\left\|\left(\vec{y}-\vec{\bar{y}}_{ij.}\right)\right\|^2. \quad (4)$$

A idéia para demonstrar (4) é mostrar que cada um dos 4 fatores da soma (3) estão em subespaços mutuamente ortogonais. Será usada a notação $V \perp W$ para designar que V e W são subespaços mutuamente ortogonais, e $V^{\perp}$ para denotar o complementar ortogonal de V, que é o subespaço formado pelos vetores que são ortogonais a todos os vetores do subespaço V.

Por inspeção direta é verificado que:

$$V_T^{\perp} = \{(\underbrace{x_{11},x_{12},...,x_{1K}}_{K},\underbrace{x_{21},x_{22},...,x_{2K}}_{K},...,\underbrace{x_{IJ1},x_{IJ2},...,x_{IJK}}_{K}), \sum_{j=1}^{K} x_{ij} = 0\}.$$

Todos os quatro vetores da decomposição (3) têm média zero e portanto são perpendiculares a V_0.

Tem-se, então:

$$\vec{\bar{y}}_{i..} - \vec{\bar{y}}_{...} \in W_F \qquad \vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...} \in W_G \qquad \vec{\bar{y}}_{ijk} - \vec{\bar{y}}_{ij.} \in V_T^{\perp}.$$

Proposição: $W_F \perp W_G$

Fazendo o produto interno entre 2 vetores com $\vec{v} \in W_F$ e $\vec{u} \in W_G$:

$$\vec{v} = (\underbrace{a,a,...,a}_{JK},\underbrace{b,b,...,b}_{JK},....,\underbrace{c,c,...,c}_{JK},\underbrace{d,d,...d}_{JK}),\ a+b+...+c+d=0$$

$$\vec{u} = (\underbrace{\alpha,\alpha,...,\alpha}_{K},\underbrace{\beta,\beta,...,\beta}_{K},...,\underbrace{\delta,\delta...,\delta}_{K},\underbrace{\lambda,\lambda,...\lambda}_{K},$$
$$\underbrace{\alpha,\alpha,...,\alpha}_{K},\underbrace{\beta,\beta,...,\beta}_{K},...,\underbrace{\delta,\delta...,\delta}_{K},\underbrace{\lambda,\lambda,...\lambda}_{K},...,$$
$$\underbrace{\alpha,\alpha,...,\alpha}_{K},\underbrace{\beta,\beta,...,\beta}_{K},...,\underbrace{\delta,\delta...,\delta}_{K},\underbrace{\lambda,\lambda,...\lambda}_{K}), \qquad \alpha+\beta+...+\delta+\lambda=0$$

$$\vec{v} \bullet \vec{u} = Ka\alpha + Ka\beta + ... + Ka\delta + Ka\lambda + Kb\alpha + Kb\beta + ... + Kb\delta +$$
$$+ Kb\lambda + ... + Kc\alpha + Kc\beta + ... + Kc\delta + Kc\lambda + Kd\alpha + Kd\beta + ... + Kd\delta + Kd\lambda =$$
$$Ka(\alpha + \beta + ... + \delta + \lambda) + Kb(\alpha + \beta + ... + \delta + \lambda) + ... +$$
$$+ Kc(\alpha + \beta + ... + \delta + \lambda) + Kd(\alpha + \beta + ... + \delta + \lambda) = 0. \qquad \square$$

Resta, portanto, analisar em qual subespaço se encontra o vetor $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{i..}\right) - \left(\vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...}\right)$.

Observe que $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{i..}\right) \in V_F^{\perp} \cap V_T$ e $\left(\vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...}\right) \in W_G$. Mas como, $W_G \subset V_F^{\perp} \cap V_T$ tem-se $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{i..}\right) \in V_F^{\perp} \cap V_T$.

Como se pode expressar $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{i..}\right) - \left(\vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...}\right)$ como $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{.j.}\right) - \left(\vec{\bar{y}}_{i..} - \vec{\bar{y}}_{...}\right)$ temos que $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{.j.}\right) \in V_G^{\perp} \cap V_T$ e $\left(\vec{\bar{y}}_{i..} - \vec{\bar{y}}_{...}\right) \in W_F$. Mas, $W_F \subset V_G^{\perp} \cap V_T$, logo $\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{.j.}\right) \in V_G^{\perp} \cap V_T$. Observe que $V_F^{\perp} \cap V_T \cap V_G^{\perp} \cap V_T = \left(V_F + V_G\right)^{\perp} \cap V_T$. Fica então demonstrado que a soma (3) está associada à soma direta de subespaços mutuamente ortogonais

$$W_F \oplus W_G \oplus \left(V_F \oplus V_G\right)^{\perp} \cap V_T \oplus V_T^{\perp}.$$

As Figuras 1 e 2 abaixo ilustram a decomposição do vetor de dados em termos do vetor de erros e em componentes no subespaço dos tratamentos.

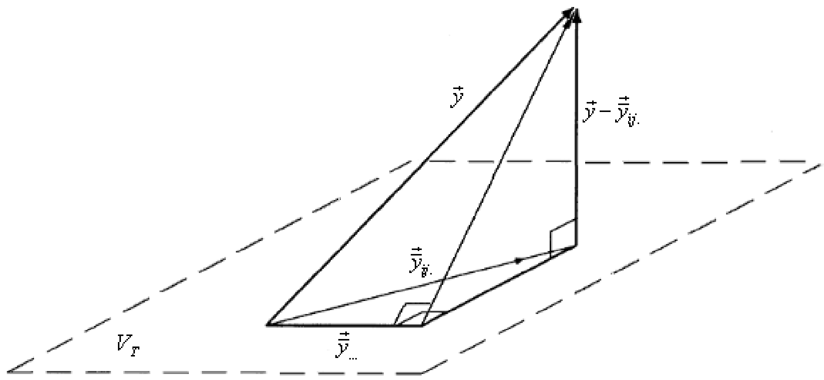

Figura 1 - Decomposição do vetor de dados em termos do vetor de erros e em componentes no subespaço dos tratamentos.

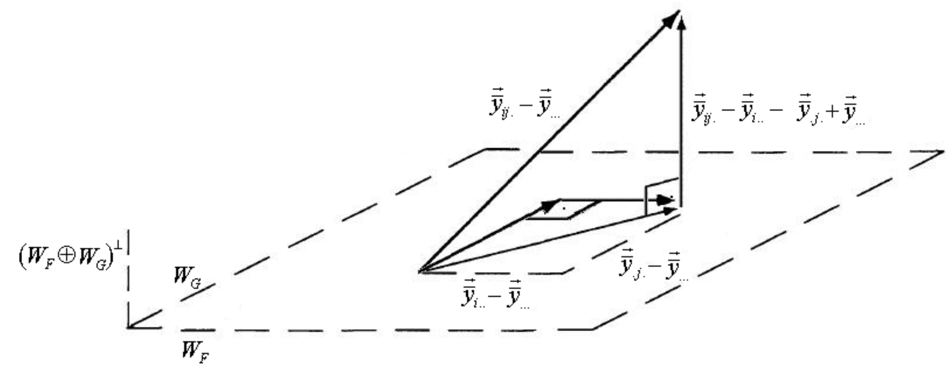

Figura 2 - Decomposição do vetor de dados corrigido pela sua média em termos do vetor de erros e em componentes no subespaço dos tratamentos.

Uma demonstração geométrica da identidade de Fisher diferente da apresentada neste artigo pode ser encontrada no Apêndice C de Saville e Wood (1991).

Para exemplificar as construções expostas segue um exemplo:

Delineamento fatorial com dois fatores com dois níveis cada e com duas repetições.

$$\vec{y} = (y_{111}, y_{112}, y_{121}, y_{122}, y_{211}, y_{212}, y_{221}, y_{222})$$
$$\vec{y}_{.jk} = (y_{.11}, y_{.12}, y_{.21}, y_{.22}, y_{.11}, y_{.12}, y_{.21}, y_{.22})$$
$$\vec{y}_{i.k} = (y_{1.1}, y_{1.2}, y_{1.1}, y_{1.2}, y_{2.1}, y_{2.2}, y_{2.1}, y_{2.2})$$
$$\vec{y}_{ij.} = (y_{11.}, y_{11.}, y_{12.}, y_{12.}, y_{21.}, y_{21.}, y_{22.}, y_{22.})$$
$$\vec{y}_{i..} = (y_{1..}, y_{1..}, y_{1..}, y_{1..}, y_{2..}, y_{2..}, y_{2..}, y_{2..})$$
$$\vec{y}_{.j.} = (y_{.1.}, y_{.1.}, y_{.2.}, y_{.2.}, y_{.1.}, y_{.1.}, y_{.2.}, y_{.2.}).$$

Consideremos valores numéricos para o vetor

$$\vec{y} = \left(y_{111}, y_{112}, y_{121}, y_{122}, y_{211}, y_{212}, y_{221}, y_{222}\right) = \left(1,2,1,4,3,1,6,9\right).$$

Assim temos: $y_{.11} = (1+3)$, $y_{.12} = (2+1)$, $y_{.21} = (1+6)$, $y_{.22} = (4+9)$ e $\vec{y}_{.jk} = (4,3,7,13,4,3,7,13)$. Analogamente, $y_{1.1} = (1+1)$, $y_{1.2} = (2+4)$, $y_{2.1} = (3+6)$, $y_{2.2} = (1+9)$, $\vec{y}_{i.k} = (2,6,2,6,9,10,9,10)$, $y_{11.} = (1+2)$, $y_{12.} = (1+4)$, $y_{21.} = (3+1)$, $y_{22.} = (6+9)$ e $\vec{y}_{ij.} = (3,3,5,5,4,4,15,15)$.

Vetores de médias:

$$\vec{\bar{y}}_{.jk} = \left(2, \frac{3}{2}, \frac{7}{2}, \frac{13}{2}, 2, \frac{3}{2}, \frac{7}{2}, \frac{13}{2}\right) \qquad \vec{\bar{y}}_{i.k} = \left(1,3,1,3, \frac{9}{2}, 5, \frac{9}{2}, 5\right)$$

$$\vec{\bar{y}}_{...} = \left(\frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}\right) \qquad \vec{\bar{y}}_{.j.} = \left(\frac{7}{4}, \frac{7}{4}, 5, 5, \frac{7}{4}, \frac{7}{4}, 5, 5\right).$$

$$\vec{\bar{y}}_{ij.} = \left(\frac{3}{2}, \frac{3}{2}, \frac{5}{2}, \frac{5}{2}, 2, 2, \frac{15}{2}, \frac{15}{2}\right) \qquad \vec{\bar{y}}_{i..} = \left(2,2,2,2, \frac{19}{4}, \frac{19}{4}, \frac{19}{4}, \frac{19}{4}\right)$$

Queremos mostrar que é válida a relação:

$$\vec{y} - \vec{\bar{y}}_{...} = \left(\vec{\bar{y}}_{i..} - \vec{\bar{y}}_{...}\right) + \left(\vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...}\right) + \left[\left(\vec{\bar{y}}_{ij.} - \vec{\bar{y}}_{i..}\right) - \left(\vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...}\right)\right] + \left(\vec{y}_{ijk} - \vec{\bar{y}}_{ij.}\right).$$

Faremos isso por partes, comparando o primeiro lado da igualdade com o segundo lado.

$$\left(\vec{y} - \vec{\bar{y}}_{...}\right) = (1,2,1,4,3,1,6,9) - \left(\frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}\right)$$
$$= \left(-\frac{19}{8}, -\frac{11}{8}, -\frac{19}{8}, \frac{5}{8}, -\frac{3}{8}, -\frac{19}{8}, \frac{21}{8}, \frac{45}{8}\right).$$

Desenvolvendo o segundo lado da igualdade tem-se:

$$\left(\vec{\bar{y}}_{i..} - \vec{\bar{y}}_{...}\right) = \left(2,2,2,2, \frac{19}{4}, \frac{19}{4}, \frac{19}{4}, \frac{19}{4}\right) - \left(\frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}\right)$$
$$= \left(-\frac{11}{8}, -\frac{11}{8}, -\frac{11}{8}, -\frac{11}{8}, \frac{11}{8}, \frac{11}{8}, \frac{11}{8}, \frac{11}{8}\right).$$

$$\left(\vec{\bar{y}}_{.j.} - \vec{\bar{y}}_{...}\right) = \left(\frac{7}{4}, \frac{7}{4}, 5, 5, \frac{7}{4}, \frac{7}{4}, 5, 5\right) - \left(\frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}, \frac{27}{8}\right) =$$
$$= \left(-\frac{13}{8}, -\frac{13}{8}, \frac{13}{8}, \frac{13}{8}, -\frac{13}{8}, -\frac{13}{8}, \frac{13}{8}, \frac{13}{8}\right).$$

Logo:

$$\left(\vec{\bar{\bar{y}}}_{i..}-\vec{\bar{\bar{y}}}_{...}\right)+\left(\vec{\bar{\bar{y}}}_{.j.}-\vec{\bar{\bar{y}}}_{...}\right)=\left(-\frac{11}{8},-\frac{11}{8},-\frac{11}{8},-\frac{11}{8},\frac{11}{8},\frac{11}{8},\frac{11}{8},\frac{11}{8},\right)$$
$$+\left(-\frac{13}{8},-\frac{13}{8},\frac{13}{8},\frac{13}{8},-\frac{13}{8},-\frac{13}{8},\frac{13}{8},\frac{13}{8},\right)$$
$$=\left(-3,-3,\frac{1}{4},\frac{1}{4},-\frac{1}{4},-\frac{1}{4},3,3\right).$$

$$\left(\vec{\bar{\bar{y}}}_{ij.}-\vec{\bar{\bar{y}}}_{i..}\right)=\left(\frac{3}{2},\frac{3}{2},\frac{5}{2},\frac{5}{2},2,2,\frac{15}{2},\frac{15}{2}\right)-\left(2,2,2,2,\frac{19}{4},\frac{19}{4},\frac{19}{4},\frac{19}{4}\right)$$
$$=\left(-\frac{1}{2},-\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{11}{4},-\frac{11}{4},\frac{11}{4},\frac{11}{4}\right).$$

$$\left(\vec{\bar{\bar{y}}}_{ij.}-\vec{\bar{\bar{y}}}_{i..}\right)-\left(\vec{\bar{\bar{y}}}_{.j.}-\vec{\bar{\bar{y}}}_{...}\right)=\left(-\frac{1}{2},-\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{11}{4},-\frac{11}{4},\frac{11}{4},\frac{11}{4}\right)$$
$$-\left(-\frac{13}{8},-\frac{13}{8},\frac{13}{8},\frac{13}{8},-\frac{13}{8},-\frac{13}{8},\frac{13}{8},\frac{13}{8},\right)$$
$$=\left(\frac{9}{8},\frac{9}{8},-\frac{9}{8},-\frac{9}{8},-\frac{9}{8},-\frac{9}{8},\frac{9}{8},\frac{9}{8}\right).$$

$$\left(\vec{y}-\vec{\bar{\bar{y}}}_{ij.}\right)=(1,2,1,4,3,1,6,9)-\left(\frac{3}{2},\frac{3}{2},\frac{5}{2},\frac{5}{2},2,2,\frac{15}{2},\frac{15}{2}\right)$$
$$=\left(-\frac{1}{2},\frac{1}{2},-\frac{3}{2},\frac{3}{2},1,-1,-\frac{3}{2},\frac{3}{2}\right).$$

Finalmente:

$$\left(\vec{\bar{\bar{y}}}_{i..}-\vec{\bar{\bar{y}}}_{...}\right)+\left(\vec{\bar{\bar{y}}}_{.j.}-\vec{\bar{\bar{y}}}_{...}\right)+\left[\left(\vec{\bar{\bar{y}}}_{ij.}-\vec{\bar{\bar{y}}}_{i..}\right)-\left(\vec{\bar{\bar{y}}}_{.j.}-\vec{\bar{\bar{y}}}_{...}\right)\right]+\left(\vec{\bar{\bar{y}}}_{ijk}-\vec{\bar{\bar{y}}}_{ij.}\right)$$
$$=\left(-\frac{19}{8},-\frac{11}{8},-\frac{19}{8},\frac{5}{8},-\frac{3}{8},-\frac{19}{8},\frac{21}{8},\frac{45}{8}\right).$$

A ortogonalidade entre os vetores pode ser comprovada por inspeção.

SILVEIRA, F. G.; COSTA, L. A.; PEREIRA, L. S.; CHAVES, L. M.; SOUZA, D. J. A geometric proof for a Fisher's Identity. *Rev. Bras. Biom.*, São Paulo, v.30, n.2, p.199-222, 2012.

- *ABSTRACT: A totally geometrical proof for a Fisher's identity, for a two factors model, is constructed based in orthogonal projections in appropriated model subspaces. Another proof, based only in matrix algebra, is presented in appendix..*

- *KEYWORDS: Orthogonal projection; factor subspaces; projection matrices; Fisher-Cochran theorem.*

Referências

BAILEY, R. A. *Design of comparative experiments.* Cambridge: Cambridge University, 2008. 255p.

RAO, C. R. *Linear statistical inference and its applications*. John Wiley & Sons, 2002. 625p.

SEARLE, S. R. *Linear models for unbalanced data.* New York: J. Wiley & Sons, 1987. 536p.

SEARLE, S. R. *Matrix algebra for the biological sciences.* New York: J. Wiley & Sons, 1966. 442p.

SAVILLE, D. J., WOOD, G. R. *Statistical methods:* the geometric approach. New York: Springer-Verlag, 1991. 560p.

SAVILLE, D. J., WOOD, G. R. A method for teaching statistics using n-dimensional geometry. *The American Statistician*, v.40, n.3, p.205–214. 1986.

Recebido em 17.11.2011

Aprovado após revisão 29.08.2012

Apêndice I

<u>Demonstração algébrica da identidade de Fisher</u>:

Claramente, temos que:

$$y_{ijk} - \bar{y}_{\bullet\bullet\bullet} \equiv (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}) + (y_{ijk} - \bar{y}_{ij\bullet})$$

Elevando ambos os termos ao quadrado e aplicando somatório temos:

$$\sum_{ijk}(y_{ijk} - \bar{y}_{\bullet\bullet\bullet})^2 \equiv \sum_{ijk}[(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}) + (y_{ijk} - \bar{y}_{ij\bullet})]^2$$

Desenvolvendo o segundo termo desta equação vem:

$$\sum_{ijk}\left[(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}) + (y_{ijk} - \bar{y}_{ij\bullet})\right]^2 =$$

$$= \underbrace{\sum_{ijk}(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})^2}_{1} + \underbrace{\sum_{ijk}(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{2} + \underbrace{\sum_{ijk}(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})}_{3}$$

$$+ \underbrace{\sum_{ijk}(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})(y_{ijk} - \bar{y}_{ij\bullet})}_{4} + \underbrace{\sum_{ijk}(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{5} + \underbrace{\sum_{ijk}(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})^2}_{6}$$

$$+ \underbrace{\sum_{ijk}(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})}_{7} + \underbrace{\sum_{ijk}(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})(y_{ijk} - \bar{y}_{ij\bullet})}_{8}$$

$$\underbrace{\sum_{ijk}(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{9} + \underbrace{\sum_{ijk}(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{10}$$

$$+ \underbrace{\sum_{ijk}(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})^2}_{11} + \underbrace{\sum_{ijk}(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})(y_{ijk} - \bar{y}_{ij\bullet})}_{12} +$$

$$+ \underbrace{\sum_{ijk}(y_{ijk} - \bar{y}_{ij\bullet})(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{13} + \underbrace{\sum_{ijk}(y_{ijk} - \bar{y}_{ij\bullet})(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{14} +$$

$$+ \underbrace{\sum_{ijk}(y_{ijk} - \bar{y}_{ij\bullet})(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})}_{15} + \underbrace{\sum_{ijk}(y_{ijk} - \bar{y}_{ij\bullet})^2}_{16} \qquad (*)$$

Como, por exemplo, $\bar{y}_{ij\bullet} = \frac{1}{K}\sum_{k} y_{ijk}$ e $\bar{y}_{i\bullet\bullet} = \frac{1}{JK}\sum_{jk} y_{ijk}$ temos:

$$2: \sum_{ijk}\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) = K\sum_{ij}\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)$$

$$= K\sum_{i}\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\sum_{j}\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) = K\left[\sum_{i}\bar{y}_{i\bullet\bullet} - \sum_{i}\bar{y}_{\bullet\bullet\bullet}\right]\sum_{j}\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)$$

$$= K\left[\sum_{i}\frac{1}{JK}.\frac{I}{I}\sum_{jk} y_{ijk} - I.\bar{y}_{\bullet\bullet\bullet}\right]\sum_{j}\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) = K\left[I.\frac{1}{IJK}\sum_{ijk} y_{ijk} - I.\bar{y}_{\bullet\bullet\bullet}\right]\sum_{j}\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)$$

$$= K.0.\sum_{j}\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) = 0$$

$$3: \sum_{ijk}\left[\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right)\right]$$

$$= \sum_{ijk}\left[\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet}\right) + \left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\right]$$

$$= \sum_{ijk}\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet}\right) + \underbrace{\sum_{ijk}\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)}_{=0\,(conforme\ demonstrado\ em\ 2)}$$

$$= K\sum_{ij}\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet}\right) = K\sum_{i}\left[\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\sum_{j}\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet}\right)\right]$$

$$= K\sum_{i}\left\{\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left[\sum_{j}\bar{y}_{ij\bullet} - \sum_{j}\bar{y}_{i\bullet\bullet}\right]\right\}$$

$$= K\sum_{i}\left\{\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left[\sum_{j}\frac{1}{K}\sum_{k} y_{ijk} - J.\bar{y}_{i\bullet\bullet}\right]\right\}$$

$$= K\sum_{i}\left\{\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left[\frac{J}{JK}\sum_{jk} y_{ijk} - J.\bar{y}_{i\bullet\bullet}\right]\right\}$$

$$= K\sum_{i}\left\{\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left[J.\bar{y}_{i\bullet\bullet} - J.\bar{y}_{i\bullet\bullet}\right]\right\} = K\sum_{i}\left\{\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left[\,0\,\right]\right\} = 0$$

$$4: \sum_{ijk} (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})(y_{ijk} - \bar{y}_{ij\bullet}) = \sum_{ij} \left[(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) \sum_{k} (y_{ijk} - \bar{y}_{ij\bullet}) \right]$$

$$= \sum_{ij} \left\{ (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) \left[\sum_{k} y_{ijk} - \sum_{k} \bar{y}_{ij\bullet} \right] \right\} = \sum_{ij} \left\{ (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) \left[\frac{K}{K} \sum_{k} y_{ijk} - K.\bar{y}_{ij\bullet} \right] \right\}$$

$$= \sum_{ij} \left\{ (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) \left[K.\bar{y}_{ij\bullet} - K.\bar{y}_{ij\bullet} \right] \right\} = \sum_{ij} \left\{ (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})[0] \right\} = 0$$

$$5: \sum_{ijk} (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) = \underbrace{\sum_{ijk} (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{=0(\text{conforme demonstrado em } 2)} = 0$$

$$7: \sum_{ijk} \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right) = \sum_{ijk} \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{\bullet j\bullet}\right)$$

$$-\sum_{ijk} \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) = K \sum_{j} \left[\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) \sum_{i} \left(\bar{y}_{ij\bullet} - \bar{y}_{\bullet j\bullet}\right) \right]$$

$$-\underbrace{\sum_{ijk} \left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)}_{=0(\text{conforme demonstrado em } 2)} = K \sum_{j} \left\{ \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) \left[\sum_{i} \bar{y}_{ij\bullet} - \sum_{i} \bar{y}_{\bullet j\bullet} \right] \right\}$$

$$= K \sum_{j} \left\{ \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) \left[\sum_{i} \frac{1}{K} \sum_{k} y_{ijk} - I\bar{y}_{\bullet j\bullet} \right] \right\}$$

$$= K \sum_{j} \left\{ \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) \left[\frac{1}{K} \frac{I}{I} \sum_{ik} y_{ijk} - I\bar{y}_{\bullet j\bullet} \right] \right\}$$

$$= K \sum_{j} \left\{ \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) \left[I\bar{y}_{\bullet j\bullet} - I\bar{y}_{\bullet j\bullet} \right] \right\} = K \sum_{j} \left\{ \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)[0] \right\} = 0$$

$$8: \sum_{ijk} \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(y_{ijk} - \bar{y}_{ij\bullet}\right) = \sum_{ij} \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) \underbrace{\sum_{k} \left(y_{ijk} - \bar{y}_{ij\bullet}\right)}_{=0(\text{conforme em } 4)}$$

$$9: \sum_{ijk} \left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)$$

$$= \underbrace{\sum_{ijk} \left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right)}_{=0(\text{conforme demonstrado em } 3)} = 0$$

$$10:\sum_{ijk}\left(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)$$

$$=\underbrace{\sum_{ijk}\left(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)\left(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet}\right)}_{=0(conforme\,demonstrado\,em\,7)}=0$$

$$12:\sum_{ijk}\left(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet}\right)\left(y_{ijk}-\bar{y}_{ij\bullet}\right)=\sum_{ijk}\left(\bar{y}_{ij\bullet}-\bar{y}_{\bullet j\bullet}\right)\left(y_{ijk}-\bar{y}_{ij\bullet}\right)$$

$$-\underbrace{\sum_{ijk}\left(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)\left(y_{ijk}-\bar{y}_{ij\bullet}\right)}_{=0(conforme\,demonstrado\,em\,4)}=\sum_{ij}\left(\bar{y}_{ij\bullet}-\bar{y}_{\bullet j\bullet}\right)\sum_{k}\left(y_{ijk}-\bar{y}_{ij\bullet}\right)$$

$$=\sum_{ij}\left\{\left(\bar{y}_{ij\bullet}-\bar{y}_{\bullet j\bullet}\right)\left[\frac{K}{K}\sum_{k}y_{ijk}-\sum_{k}\bar{y}_{ij\bullet}\right]\right\}=\sum_{ij}\left\{\left(\bar{y}_{ij\bullet}-\bar{y}_{\bullet j\bullet}\right)\left[K\bar{y}_{ij\bullet}-K\bar{y}_{ij\bullet}\right]\right\}$$

$$=\sum_{ij}\left\{\left(\bar{y}_{ij\bullet}-\bar{y}_{\bullet j\bullet}\right)[0]\right\}=0$$

$$13:\underbrace{\sum_{ijk}\left(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)\left(y_{ijk}-\bar{y}_{ij\bullet}\right)}_{=0(conforme\,demonstrado\,em\,4)}=0$$

$$14:\underbrace{\sum_{ijk}\left(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)\left(y_{ijk}-\bar{y}_{ij\bullet}\right)}_{=0(conforme\,demonstrado\,em\,8)}=0$$

$$15:\underbrace{\sum_{ijk}\left(y_{ijk}-\bar{y}_{ij\bullet}\right)\left(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet}\right)}_{=0(conforme\,demonstrado\,em\,12)}=0$$

Desta forma, os únicos termos que não se anulam em (*) são os termos 1, 6, 11 e 16. Portanto,

$$\boxed{\begin{aligned}\sum_{ijk}\left(y_{ijk}-\bar{y}_{\cdots}\right)^2 &\equiv \sum_{ijk}\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)^2+\sum_{ijk}\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)^2\\ &+\sum_{ijk}\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)^2+\sum_{ijk}\left(y_{ijk}-\bar{y}_{ij\cdot}\right)^2.\end{aligned}}$$

Apêndice II

<u>Demonstração matricial da identidade de Fisher</u>:

Seja $N = IJK$, em que *I*, *J*, e *K* são inteiros positivos. Considere matrizes $N \times N$ com colunas identificadas por C_{stu} $\left(s = 1..I,\ t = 1..J,\ u = 1..K \right)$ e linhas por L_{pqr} $\left(p = 1..I,\ q = 1..J,\ r = 1..K \right)$, com os índices ordenados lexicograficamente. Uma matriz *A* terá, então, seus elementos identificados, como $A = \left[a_{(pqr,stu)} \right]$. Ou seja,

$$\begin{array}{c|ccccccccc}
 & C_{111} & C_{112} & \cdots & C_{11K} & C_{121} & \cdots & C_{12K} & \cdots & C_{IJK} \\
\hline
L_{111} & a_{(111,111)} & a_{(111,112)} & & & & & & & a_{(111,IJK)} \\
L_{112} & a_{(112,111)} & & & & & & & & \\
\vdots & & & & & & & & & \\
L_{11K} & & & & & & & & & \\
L_{121} & & & & & & & & & \\
\vdots & & & & & & & & & \\
L_{12K} & & & & & & & & & \\
 & & & & & & & & & \\
L_{IJK} & a_{(IJK,111)} & & & & & & & & a_{(IJK,IJK)}
\end{array}$$

Construa as matrizes:

(i) V_0 uma matriz $N \times N$ com todas as entradas iguais a $1/N$. Isto é, se 1_N é uma matriz $N \times N$ com todas as entradas iguais a 1, então $V_0 = \dfrac{1}{N} 1_N$. Observe que V_0 é simétrica de posto 1 e, como cada elemento da matriz $\left(V_0 \right)^2$ é $\sum_{i=1}^{N} \dfrac{1}{N} \dfrac{1}{N} = \dfrac{1}{N}$, V_0 é também idempotente. Além disso, como $\left(I_N - V_0 \right)^T = I_N^T - V_0^T = I_N - V_0$ e $\left(I_N - V_0 \right)^2 = I_N^2 - I_N V_0 - V_0 I_N + V_0^2 = I_N - V_0$, a matriz $I_N - V_0$ é também idempotente e simétrica.

(ii) V_T uma matriz com entradas iguais a $\frac{IJ}{N} = \frac{1}{K}$ nas posições em que os dois primeiros índices das colunas são iguais aos dois primeiros índices das linhas, e zero nas demais posições. Isto é:

$$V_T = \left[b_{(pqr,stu)} \right] \text{ em que } b_{(pqr,stu)} = \begin{cases} 1/K & \text{caso } p = s \text{ e } q = t \\ 0 & \text{caso } p \neq s \text{ ou } q \neq t \end{cases}$$

A matriz V_T é bloco-diagonal de posto IJ, simétrica $\left(V_T^T = V_T\right)$ com *IJ* blocos, todos de dimensão $K \times K$. A matriz V_T é também idempotente $\left(V_T^2 = V_T\right)$, pois na matriz $\left(V_T\right)^2$ só são diferentes de zero as entradas correspondentes aos blocos da diagonal de V_T, com cada entrada igual a $\sum_{i=1}^{K} \frac{1}{K}\frac{1}{K} = \frac{1}{K}$.

Observe que, se $V_R = I_N - V_T$, então:

$$V_R^T = \left(I_N - V_T\right)^T = I_N^T - V_T^T = I_N - V_T = V_R$$

$$V_R^2 = \left(I_N - V_T\right)^2 = I_N^2 + V_T^2 - I_N V_T - V_T I_N = I_N + V_T - V_T - V_T$$
$$= I_N - V_T = V_R$$

$$V_R V_T = \left(I_N - V_T\right) V_T = I_N V_T - V_T V_T = V_T - V_T = [0]$$

$$V_T V_R = V_T \left(I_N - V_T\right) = V_T V_T - I_N V_T = V_T - V_T = [0]$$

O posto de V_R é, portanto, $N - IJ = IJ\left(K - 1\right)$.

Em resumo: A matriz identidade se decompõe como $I_N = V_R + V_T$, em que V_T e V_R são, ambas, simétricas e idempotentes e $V_R V_T = V_T V_R = [0]$. Em termos de espaços vetoriais, as matrizes V_T e V_R são projetores em subespaços ortogonais de dimensões, respectivamente, IJ e $N - IJ = IJ\left(K - 1\right)$.

Continuando a construção de matrizes:

(iii) V_1 uma matriz com entradas iguais a $\frac{I}{N} = \frac{1}{JK}$ nas posições em que o primeiro índice das colunas é igual ao primeiro índice das linhas, e zero nas demais posições. Isto é:

$$V_1 = \left[c_{(pqr,stu)} \right] \text{ em que } c_{(pqr,stu)} = \begin{cases} 1/JK & \text{caso } p = s \\ 0 & \text{caso } p \neq s \end{cases}$$

A matriz V_1 é bloco-diagonal, simétrica de posto *I*, com *I* blocos, todos de dimensão $JK \times JK$. A matriz V_1 é também idempotente, pois na matriz $(V_1)^2$ só são diferentes de zero as entradas correspondentes aos blocos da diagonal de V_1, com cada entrada igual a $\sum_{i=1}^{JK} \frac{1}{JK}\frac{1}{JK} = \frac{1}{JK}$. Além disso, a soma de qualquer linha de V_1 é dada por:

$$\sum_{s=1}^{I}\sum_{t=1}^{J}\sum_{u=1}^{K} c_{(pqr,stu)} = \sum_{t=1}^{J}\sum_{u=1}^{K} c_{(pqr,ptu)} = \sum_{t=1}^{J}\sum_{u=1}^{K} \frac{1}{JK} = 1$$

A soma de qualquer coluna de V_1 é dada por

$$\sum_{p=1}^{I}\sum_{q=1}^{J}\sum_{r=1}^{K} c_{(pqr,stu)} = \sum_{t=1}^{J}\sum_{u=1}^{K} c_{(sqr,stu)} = \sum_{q=1}^{J}\sum_{r=1}^{K} \frac{1}{JK} = 1$$

Ou seja, qualquer linha ou qualquer coluna de V_1 soma 1.

(iv) V_2 uma matriz com entradas iguais a $\frac{J}{N} = \frac{1}{IK}$ nas posições em que o segundo índice das colunas é igual ao segundo índice das linhas, e zero nas demais posições. Isto é:

$$V_2 = \left[d_{(pqr,stu)} \right] \text{ em que } d_{(pqr,stu)} = \begin{cases} 1/IK & \text{caso } q = t \\ 0 & \text{caso contrário} \end{cases}$$

V_2 é uma matriz em blocos, simétrica, de posto *J*, com *J* blocos, todos de dimensão $IK \times IK$. A matriz V_2 é também idempotente, pois na matriz $(V_1)^2$ só são diferentes de zero as entradas correspondentes aos blocos não nulos de V_2, com cada entrada igual a $\sum_{i=1}^{IK} \frac{1}{IK}\frac{1}{IK} = \frac{1}{IK}$.

Qualquer linha ou coluna de V_2, à semelhança das de V_1, soma 1.

Essas quatro matrizes estão exemplificadas em seguida, para o caso $I = 2,\ J = 3$ e $K = 2$.

$$V_0 = \frac{1}{12}\begin{bmatrix} 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1&1 \end{bmatrix} \qquad V_T = \frac{1}{2}\begin{bmatrix} 1&1&0&0&0&0&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&1&0&0&0&0&0&0 \\ 0&0&0&0&1&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&1&0&0&0&0 \\ 0&0&0&0&0&0&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&0&0&0&0&1&1 \end{bmatrix}$$

$$V_1 = \frac{1}{6}\begin{bmatrix} 1&1&1&1&1&1&0&0&0&0&0&0 \\ 1&1&1&1&1&1&0&0&0&0&0&0 \\ 1&1&1&1&1&1&0&0&0&0&0&0 \\ 1&1&1&1&1&1&0&0&0&0&0&0 \\ 1&1&1&1&1&1&0&0&0&0&0&0 \\ 1&1&1&1&1&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&1&1&1&1&1 \\ 0&0&0&0&0&0&1&1&1&1&1&1 \\ 0&0&0&0&0&0&1&1&1&1&1&1 \\ 0&0&0&0&0&0&1&1&1&1&1&1 \\ 0&0&0&0&0&0&1&1&1&1&1&1 \\ 0&0&0&0&0&0&1&1&1&1&1&1 \end{bmatrix} \qquad V_2 = \frac{1}{4}\begin{bmatrix} 1&1&0&0&0&0&1&1&0&0&0&0 \\ 1&1&0&0&0&0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0&0&1&1&0&0 \\ 0&0&1&1&0&0&0&0&1&1&0&0 \\ 0&0&0&0&1&1&0&0&0&0&1&1 \\ 0&0&0&0&1&1&0&0&0&0&1&1 \\ 1&1&0&0&0&0&1&1&0&0&0&0 \\ 1&1&0&0&0&0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0&0&1&1&0&0 \\ 0&0&1&1&0&0&0&0&1&1&0&0 \\ 0&0&0&0&1&1&0&0&0&0&1&1 \\ 0&0&0&0&1&1&0&0&0&0&1&1 \end{bmatrix}$$

Observe que linhas e colunas dessas matrizes somam 1 e, sendo assim, $V_0V_* = \frac{1}{N}1_N V_* = V_*V_0 = V_0$, para $V_* = V_1,\ V_2$ ou V_T. Além disso,

$$V_1V_2 = \left[e_{(pqr,stu)} \right] = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{(pqr,ijk)} d_{(ijk,stu)}$$

$$= \begin{cases} K\dfrac{1}{JK}\dfrac{1}{IK} = \dfrac{1}{IJK} & \text{caso } i = p \text{ e } \mathrm{j} = s \\ 0 & \text{caso } i = p \text{ e } j \neq s \\ 0 & \text{caso } i \neq p \text{ e } j = s \\ 0 & \text{caso } i \neq p \text{ e } j \neq s \end{cases} = \left[\frac{1}{N} \right] = V_0 = V_2V_1$$

Por procedimentos semelhantes, mas tediosos, prova-se que $V_1V_T = V_1 = V_TV_1$ e $V_2V_T = V_2 = V_TV_2$.

(v) $W_1 = V_1 - V_0 = \left[e_{(pqr,stu)}\right]$ em que

$$e_{(pqr,stu)} = \begin{cases} \dfrac{1}{JK} - \dfrac{1}{N} = \dfrac{I-1}{N} & \text{caso } p = s \\ \dfrac{-1}{N} & \text{caso } p \neq s \end{cases}$$

$$W_1^T = \left(V_1 - V_0\right)^T = V_1^T - V_0^T = V_1 - V_0 = W_1 \rightarrow \text{A matriz } W_1 \text{ é simétrica.}$$

$$W_1^2 = \left(V_1 - V_0\right)^2 = V_1^2 - V_1V_0 - V_0V_1 + V_0^2 = V_1 - V_1V_0 - V_0V_1 + V_0$$

$$= V_1 - V_0 - V_0 + V_0 = V_1 - V_0 = W_1 \quad \rightarrow \quad \text{A matriz } W_1 \text{ é idempotente de}$$

posto $I-1$.

$$W_1V_0 = \left(V_1 - V_0\right)V_0 = V_1V_0 - V_0^2 = V_0 - V_0 = [0]$$

$$V_0W_1 = V_0\left(V_1 - V_0\right) = V_0^2 - V_1V_0 = V_0 - V_0 = [0]$$

(vi) A matriz $W_2 = V_2 - V_0 = \left[f_{(pqr,stu)}\right]$ em que

$$f_{(pqr,stu)} = \begin{cases} \dfrac{1}{IK} - \dfrac{1}{N} = \dfrac{J-1}{N} & \text{caso } q = t \\ \dfrac{-1}{N} & \text{caso } q \neq t \end{cases}$$

também é, por razões semelhantes, simétrica e idempotente de posto $J-1$. Além disso, como $V_0V_2 = V_2V_0 = V_0$,

$$W_2V_0 = \left(V_2 - V_0\right)V_0 = V_2V_0 - V_0^2 = V_0 - V_0 = [0] \text{ e}$$

$$V_0W_2 = V_0\left(V_2 - V_0\right) = V_0^2 - V_2V_0 = V_0 - V_0 = [0]$$

Observe que:

$$W_1W_2 = \left(V_1 - V_0\right)\left(V_2 - V_0\right) = V_1V_2 - V_1V_0 - V_0V_1 + V_0V_0$$

$$= V_1V_2 - V_0 - V_0 + V_0 = V_1V_2 - V_0 = [0] = W_2W_1$$

(vii) A matriz $V_{12} = V_T - W_1 - W_2 - V_0 = V_T - V_1 - V_2 + V_0 = \left[h_{(pqr,stu)} \right]$ em que

$$\left[h_{(pqr,stu)} \right] = \begin{cases} \frac{1}{K} - \frac{1}{JK} - \frac{1}{IK} + \frac{1}{N} = \frac{1}{N}(IJ - I - J + 1) & \text{caso } p = s \text{ e } q = t \\ -\frac{1}{JK} + \frac{1}{N} = \frac{1}{N}(-I + 1) & \text{caso } p = s \text{ e } q \neq t \\ -\frac{1}{IK} + \frac{1}{N} = \frac{1}{N}(-J + 1) & \text{caso } p \neq s \text{ e } q = t \\ \frac{1}{N} & \text{caso } p \neq s \text{ e } q \neq t \end{cases}$$

Observe que:

$V_{12}^T = \left(V_T - W_1 - W_2 - V_0 \right)^T = V_T - W_1 - W_2 - V_0 = V_{12} \rightarrow$ A matriz V_{12} é simétrica.

$$\begin{aligned} V_{12}V_{12} &= \left(V_T - V_1 - V_2 + V_0 \right)\left(V_T - V_1 - V_2 + V_0 \right) \\ &= V_T^2 + V_1^2 + V_2^2 + V_0^2 - 2V_TV_1 - 2V_TV_2 + 2V_TV_0 + 2V_1V_2 - 2V_1V_0 - 2V_2V_0 \\ &= V_T + V_1 + V_2 + V_0 - 2V_1 - 2V_2 + 2V_0 + 2V_0 - 2V_0 - 2V_0 \\ &= V_T - V_1 - V_2 + V_0 = V_{12} \end{aligned}$$ $\rightarrow$ A matriz V_{12} é idempotente.

Como $V_{12}V_T = \left(V_T - V_1 - V_2 + V_0 \right)V_T = V_T - V_1 - V_2 + V_0 = V_{12}$, segue que:

$$V_{12}V_R = V_{12}\left(I_N - V_T \right) = V_{12} - V_{12}V_T = V_{12} - V_{12} = [0]$$

A soma de qualquer coluna de V_{12}:

$$\sum_{p=1}^{I}\sum_{q=1}^{J}\sum_{r=1}^{K} h_{(pqr,stu)} = \begin{cases} K\left(\frac{1}{K} - \frac{1}{JK} - \frac{1}{IK} + \frac{1}{N} \right) = K\left(\frac{JK - I - J + 1}{N} \right) & \text{caso } p = s \text{ e } q = t \\ K(J-1)\left(-\frac{1}{JK} + \frac{1}{N} \right) = \frac{K}{N}(J-1)(1-I) & \text{caso } p = s \text{ e } q \neq t \\ K(I-1)\left(-\frac{1}{IK} + \frac{1}{N} \right) = \frac{K}{N}(I-1)(1-J) & \text{caso } p \neq s \text{ e } q = t \\ K(J-1)(K-1)\frac{1}{N} = K\left(\frac{JK - I - J + 1}{N} \right) & \text{caso } p \neq s \text{ e } q \neq t \end{cases}$$

$$= K\left(\frac{JK - I - J + 1}{N} \right) + \frac{K}{N}(J-1)(1-I) + \frac{K}{N}(I-1)(1-J) + K\left(\frac{JK - I - J + 1}{N} \right)$$

$= 0 =$ soma de qualquer linha de V_{12}. Em razão disso: $V_{12}V_0 = [0] = V_0V_{12}$.

O posto da matriz V_{12} é $IJ-(I-1)-(j-1)-1=IJ-I-I+1$

No caso do produto $V_{12}W_1=\left[\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}h_{(pqr,ijk)}e_{(ijk,stu)}\right]$, há 4 situações a considerar:

(1) O caso $p=s$ e $q=t$

$$V_{12}W_1=\left[\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}h_{(pqr,ijk)}e_{(ijk,pqu)}\right]$$

$$=\left\{\left[\begin{array}{ll} K\frac{1}{N}(IJ-I-J+1)\left(\frac{I-1}{N}\right)=\frac{K}{N^2}(I-1)^2(J-1) & \text{caso } i=p \text{ e } j=q \\ K(J-1)\frac{1}{N}(-I+1)\frac{1}{N}(I-1)=\frac{-K}{N^2}(I-1)^2(J-1) & \text{caso } i=p \text{ e } j\neq q \\ K(I-1)\frac{1}{N}(-J+1)\left(\frac{-1}{N}\right)=\frac{K}{N^2}(I-1)(J-1) & \text{caso } i\neq p \text{ e } j=q \\ K(I-1)(J-1)\left(\frac{-1}{N}\right)\left(\frac{1}{N}\right)=\frac{-K}{N^2}(I-1)(J-1) & \text{caso } i\neq p \text{ e j} \neq q \end{array}\right]\right.$$

$$=[0]$$

(2) O caso $p=s$ e $q\neq t$

$$V_{12}W_1=\left[\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}h_{(pqr,ijk)}e_{(ijk,ptu)}\right]$$

$$=\left\{\left[\begin{array}{ll} K\frac{1}{N}(IJ-I-J+1)\left(\frac{I-1}{N}\right)=\frac{K}{N^2}(I-1)^2(J-1) & \text{caso } i=p \text{ e } j=q \\ K(J-1)\frac{1}{N}(-I+1)\frac{1}{N}(I-1)=\frac{-K}{N^2}(I-1)^2(J-1) & \text{caso } i=p \text{ e } j=t \\ K(I-1)\frac{1}{N}\frac{1}{N}(J-1)=\frac{K}{N^2}(I-1)(J-1) & \text{caso } i\neq p \text{ e } j=q \\ K(I-1)\frac{1}{N}\left(\frac{-1}{N}\right)=\frac{-K}{N^2}(I-1) & \text{caso } i\neq p \text{ e } j=t \\ K(I-1)(J-2)\frac{1}{N}\left(\frac{-1}{N}\right)=\frac{-K}{N^2}(I-1)(J-2) & \text{caso } i\neq p, j\neq q \text{ e } j\neq t \end{array}\right]\right.$$

$$=[0]$$

Por procedimento semelhante se mostra que:

(3) O caso $p \neq s$ e $q = t$

$$V_{12}W_1 = \left[\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} h_{(pqr,ijk)} e_{(ijk,squ)} \right] = [0]$$

(4) O caso $p \neq s$ e $q \neq t$

E mostra-se, também, que $W_1 V_{12} = [0]$.

Em resumo, a matriz identidade de dimensão N pode ser decomposta como $I_N = V_0 + W_1 + W_2 + V_{12} + V_R \iff I_N - V_0 = W_1 + W_2 + V_{12} + V_R$, em que cada parcela é simétrica e idempotente e, duas a duas, são todas mutuamente ortogonais.

Supondo um experimento fatorial com dois fatores, I níveis do primeiro fator, J níveis do segundo fator e K repetições de cada tratamento, isto é, cada combinação de um nível do primeiro fator com um nível do segundo fator, o vetor de dados pode ser indexado como Y_{ijk}, com $i = 1...I,\quad j = 1...J$ e $k = 1...K$ e escrito como uma matriz coluna $Y = \left[Y_{111}, Y_{112}, ..., Y_{11K}, Y_{121}, Y_{12K}, ..., Y_{IJK}\right]^T$. Observe que, utilizando a notação habitual para dados experimentais,

$$Y^T Y = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k-1}^{K} Y_{ijk}^2 = Y^T I_N Y = Y^T \left(V_0 + W_1 + W_2 + V_{12} + V_R\right) Y$$

$$= Y^T V_0 Y + Y^T W_1 Y + Y^T W_2 Y + Y^T V_{12} Y + Y^T V_R Y$$

$$Y^T I_N Y - Y^T V_0 Y = Y^T \left(I_N - V_0\right) Y = Y^T W_1 Y + Y^T W_2 Y + Y^T V_{12} Y + Y^T V_R Y$$

$$V_0 Y = \frac{1}{N} 1_N Y = \left[\overline{Y}_{\dots}, \overline{Y}_{\dots}, ..., \overline{Y}_{\dots}\right]^T$$

$$\left(I_N - V_0\right) Y = \left[Y_{111} - \overline{Y}_{\dots}, Y_{11K} - \overline{Y}_{\dots}, ..., Y_{IJK} - \overline{Y}_{\dots}\right]^T$$

$$Y^T \left(I_N - V_0\right) Y = Y^T \left(I_N - V_0\right)^T \left(I_N - V_0\right) Y$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(Y_{ijk} - \overline{Y}_{\dots}\right)^2$$

$$V_1 Y = \left[\overline{Y}_{1..}, ..., \overline{Y}_{2..}, ..., \overline{Y}_{I..}\right]^T$$

$$\left(V_1 - V_0\right) Y = \left[\overline{Y}_{1..} - \overline{Y}_{\dots}, ..., \overline{Y}_{2..} - \overline{Y}_{\dots}, ..., \overline{Y}_{I..} - \overline{Y}_{\dots}\right]^T$$

$$Y^T W_1 Y = Y^T \left(V_1 - V_0\right) Y = Y^T \left(V_1 - V_0\right)^T \left(V_1 - V_0\right) Y$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(Y_{i\bullet\bullet} - \overline{Y}_{\bullet\bullet\bullet}\right)^2$$

De maneira semelhante

$$Y^T W_2 Y = Y^T \left(V_2 - V_0\right) Y = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(Y_{\bullet j\bullet} - \overline{Y}_{\bullet\bullet\bullet}\right)^2$$

$$Y^T V_{12} Y = Y^T \left(V_T - W_1 - W_2 - V_0\right) Y = Y^T \left(V_T - V_1 - V_2 + V_0\right) Y$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(Y_{ij\bullet} - Y_{i\bullet\bullet} - Y_{\bullet j\bullet} + \overline{Y}_{\bullet\bullet\bullet}\right)^2$$

$$Y^T V_R Y = Y^T \left(I_N - V_T\right) Y = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(Y_{ijk} - Y_{ij\bullet}\right)^2$$

A distribuição de formas quadráticas (RAO, 2002): Considere as variáveis aleatórias independentes $Y^T = \left[Y_1 \quad Y_2 \quad \cdots \quad Y_N\right]$, em que $Y_i \sim N\left(\mu_i, 1\right)$ e $\mu^T = \left[\mu_1 \quad \mu_2 \quad \cdots \quad \mu_N\right]$.

TEOREMA DE FISHER-COCHRAN

Se $Q_1, ..., Q_k$ são *k* formas quadráticas cujas matrizes $A_1, ..., A_k$ têm postos, respectivamente, $n_1, ..., n_k$ e sejam tais que:

$$Y^T Y = Q_1 + ... + Q_k;$$

então, uma condição necessária e suficiente para que as formas quadráticas tenham distribuição $Q_i \sim \chi^2\left(n_i, \lambda_i\right)$ e sejam independentes é que $N = \sum_{i=1}^{k} n_i$. Nesse caso, $\lambda_i = \mu^T A_i \mu$ e $\sum_{j=1}^{N} \mu_j^2 = \sum_{i=1}^{k} \lambda_i$.

#

Se Y_{ijk} é o resultado de um experimento fatorial com dois fatores, com *I* níveis do primeiro fator, *J* níveis do segundo fator e *K* repetições, o vetor de médias é dado por:

$$\mu^T = \Big[\underbrace{\mu_{11} \ \cdots \ \mu_{11}}_{k} \ \ \underbrace{\mu_{12} \ \cdots \ \mu_{12}}_{k} \ \cdots \ \cdots \ \cdots \ \underbrace{\mu_{IJ} \ \cdots \ \mu_{IJ}}_{k}\Big]$$

A decomposição $Y^TY = Y^TV_0Y + Y^TW_1Y + Y^TW_2Y + Y^TV_{12}Y + Y^TV_RY$ resulta em:

$$V_0\mu = \left[\bar{\mu}_{\bullet\bullet} \ \cdots \ \bar{\mu}_{\bullet\bullet}\right] \rightarrow \lambda_0 = \mu^TV_0\mu = N\bar{\mu}_{\bullet\bullet}^2 = \sum_{i+1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\bar{\mu}_{\bullet\bullet}^2$$

$$V_T\mu = \Big[\underbrace{\mu_{11} \ \cdots \ \mu_{11}}_{K} \ \ \underbrace{\mu_{12} \ \cdots \ \mu_{12}}_{K} \ \cdots \ \cdots \ \cdots \ \underbrace{\mu_{IJ} \ \cdots \ \mu_{IJ}}_{K}\Big]^T = \mu$$

$$\rightarrow \mu^TV_T\mu = K\left(\mu_{11}^2 + \mu_{12}^2 + \ldots + \mu_{IJ}^2\right) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\mu_{ij}^2$$

$$V_1\mu = \Big[\underbrace{\bar{\mu}_{1\bullet} \ \cdots \ \mu_{1\bullet}}_{JK} \ \ \underbrace{\mu_{2\bullet} \ \cdots \ \mu_{2\bullet}}_{JK} \ \cdots \ \cdots \ \cdots \ \underbrace{\mu_{J\bullet} \ \cdots \ \mu_{J\bullet}}_{JK}\Big]^T$$

$$\rightarrow \mu^TV_1\mu = JK\left(\bar{\mu}_{1\bullet}^2 + \bar{\mu}_{2\bullet}^2 + \ldots + \bar{\mu}_{I\bullet}^2\right) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\bar{\mu}_{i\bullet}^2$$

$$\rightarrow \mu^TV_2\mu = JK\left(\bar{\mu}_{\bullet 1}^2 + \bar{\mu}_{\bullet 2}^2 + \ldots + \bar{\mu}_{\bullet J}^2\right) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\bar{\mu}_{\bullet j}^2$$

$$W_1\mu = \left(V_1 - V_0\right)\mu$$

$$= \Big[\underbrace{\left(\bar{\mu}_{1\bullet} - \bar{\mu}_{\bullet\bullet}\right) \ \cdots \ \left(\bar{\mu}_{1\bullet} - \bar{\mu}_{\bullet\bullet}\right)}_{JK} \ \ldots \ \ldots \ \ldots \ \underbrace{\left(\bar{\mu}_{J\bullet} - \bar{\mu}_{\bullet\bullet}\right) \ \cdots \ \left(\bar{\mu}_{J\bullet} - \bar{\mu}_{\bullet\bullet}\right)}_{JK}\Big]^T$$

Então:

$$\lambda_1 = \mu^TW_1\mu = \mu^T\left(V_1 - V_0\right)\mu = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\bar{\mu}_{i\bullet} - \bar{\mu}_{\bullet\bullet}\right)^2$$

$$\lambda_2 = \mu^TW_2\mu = \mu^T\left(V_2 - V_0\right)\mu = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\bar{\mu}_{\bullet j} - \bar{\mu}_{\bullet\bullet}\right)^2$$

$$\lambda_{12} = \mu^TW_{12}\mu = \mu^T\left(V_T - V_1 - V_2)V_0\right)\mu$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\mu_{ij}^2 - \bar{\mu}_{i\bullet} - \bar{\mu}_{\bullet j} + \bar{\mu}_{\bullet\bullet}\right)^2$$

$$\lambda_R = \mu^T \left(I_N - V_T\right)\mu = \mu^T\mu - \mu^T V_T \mu = \mu^T\mu - \mu^T\mu = 0$$

Como:

$posto\left(V_0\right) = 1$,

$posto\left(W_1\right) = I - 1$,

$posto\left(W_2\right) = J - 1$,

$posto\left(V_{12}\right) = IJ - I - J - 1$ e

$posto\left(V_R\right) = IJ\left(K - 1\right)$,

segue que:

$$posto\left(V_0\right) + posto\left(W_1\right) + posto\left(W_2\right) + posto\left(V_{12}\right) + posto\left(V_R\right)$$
$$= IJK = N,$$

e, portanto, o teorema de Fisher-Cochran se aplica, ou seja, na decomposição:

$$Y^T Y = Y^T V_0 Y + Y^T W_1 Y + Y^T W_2 Y + Y^T V_{12} Y + Y^T V_R Y$$

as formas quadráticas à direita da igualdade têm distribuição, respectivamente,

$\chi^2\left(1, \lambda_0\right)$, $\chi^2\left(I - 1, \lambda_1\right)$,

$\chi^2\left(J - 1, \lambda_2\right)$, $\chi^2\left(IJ - I - J + 1, \lambda_{12}\right)$ e

$\chi^2\left(IJ\left(K - 1\right), \lambda_R\right)$.