Expectância de Quadrados Médios em Experimentos Aleatorizados com Estrutura Balanceada e Completa

Este exemplar corresponde à redação final da tese, devidamente corrigida e defendida por André Luís Santos de Pinho e aprovada pela Comissão Julgadora.

Campinas, 9 de janeiro de 1996.

Prof. Dr. Euclydes Custódio de Lima Filho
Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para a obtenção do título de Mestre em Estatística.

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Pinho,André Luís Santos de

P655e Expectância de quadros médios em experimentos aleatorizados com estrutura balanceada e completa / André Luís Santos de Pinho-- Campinas, [S.P. :s.n.], 1986.

Orientador : Euclydes Custódio de Lima Filho.

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Ciência da Computação.

1. Amostragem (Estatística). 2. Média (Matemática). 3. Estatística. I. Lima Filho,Euclydes Custódio de. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Ciência da Computação. III. Título.

Tese de Mestrado defendida e aprovada em de de 199

pela Banca Examinadora composta pelos Profs. Drs.

Prof (a). Dr (a).

Prof (a). Dr (a).

Prof (a). Dr (a).

Agradecimentos

Seria uma grande injustiça, chegando onde eu cheguei, não mencionar a ajuda de pessoas que de uma forma ou de outra contribuíram nesta caminhada. Em particular gostaria de citar os meus pais, responsáveis pela minha formação, meus tios Lauro e Nilza, que em momentos difíceis deram uma força e tanto, aos amigos de mestrado, principalmente os da minha turma de 93(verdadeiros cúmplices), aos professores Euclydes, José de Carvalho (pela sugestão do tema e empréstimo de material) e Ademir Petenate e um muito especial a Carla Almeida Vivacqua, que desde 89 tem trilhado, junto comigo, momentos de grande importância e decisão.

Índice

Capítulo II - Amostragem e Função Indicadora

Capítulo III - Experimentos Aleatorizados

Capítulo IV - O Poder dos Testes de Aleatorização

Apêndice A

Introdução

Este trabalho tem como objetivo apresentar uma forma para obtenção das expectâncias dos quadrados médios em experimentos com estrutura balanceada e completa, e tem como base as referências [7], [19], [21].

No Capítulo I, estão as definições básicas, que serão usadas no decorrer de toda a dissertação, também será vista uma abordagem geométrica dos espaços gerados pelas médias admissíveis em uma estrutura de uma população de respostas, e a relação entre estas médias e os componentes que elas dão origem. Os cap sigmas[1] ($\sum$), que são a base deste trabalho, são definidos também neste Capítulo.

No Capítulo II, está a parte teórica de experimentos aleatorizados. No entanto, são consideradas apenas as estruturas populacionais. É nele, que são introduzidas as variáveis aleatórias indicadoras para estruturas de amostra aleatória simples sem reposição, cruzadas e hierárquicas, com suas respectivas funções de probabilidade, expectâncias de primeira e segunda ordem, e as covariâncias. E, por fim, é dada a forma de escrever estas expectâncias em função dos componentes de variação (σ^2) e dos cap sigmas, para que estes possam ser usados na ANOVA populacional.

No Capítulo III, introduz-se a aleatorização, isto é, a atribuição aleatória de tratamentos às unidades experimentais, que devido a este processo, ficam embutidas nos tratamentos. Os cap sigmas amostrais necessários para a construção da ANOVA amostral também são definidos neste Capítulo. Uma representação gráfica para estruturas completas é apresentada em uma seção, com o objetivo de auxiliar na visualização de estruturas presentes em uma população ou em um experimento.

---

[1] A palavra cap sigma vem do inglês *capital sigma*, ou seja, sigma maiúsculo.

O Capítulo IV apresenta algumas simulações, com o intuito de (i) averiguar o comportamento da função poder do experimento aleatorizado; (ii) estudar a qualidade de aproximação da distribuição F para a distribuição de referência gerada pelo processo de aleatorização. Em ambos os casos, são utilizados experimentos aleatorizados em blocos de tamanho 2.

No apêndice A, estão conceitos de álgebra linear usados principalmente no Capítulo 1.

No apêndice B, encontram-se os programas relativos às simulações do Capítulo IV.

Chama-se a atenção para o fato deste trabalho não ter explorado as propriedades dos estimadores via aleatorização, pois isto se encontra nas referências [11], [13], que também são teses de mestrado e estão em português. Portanto, não há a necessidade de repetir toda esta parte.

Capítulo I

Estruturas Populacionais

Neste capítulo serão dadas as definições necessárias para o desenvolvimento da dissertação. A visão geométrica também será abordada para o melhor entendimento e servir de motivação para o trabalho. Alguns conceitos de álgebra linear (Apêndice A) serão precisos para uma boa fluência na leitura do texto. Os cap sigmas, que são uma das bases desse trabalho, serão definidos aqui.

1.1 - Conceitos Básicos

1.1.1 - Indivíduos e População de Indivíduos

Indivíduos são elementos que possuem determinada característica em comum, por exemplo: parafusos produzidos em uma fábrica, pessoas portadoras da bactéria do cólera, etc. O conjunto de todos os indivíduos forma a *população de indivíduos*. Pode-se dizer que a menor unidade, objeto de estudo, em uma população é o indivíduo, ou seja, medidas são feitas a partir deles para se fazer inferências e tirar conclusões.

1.1.2 - Fatores e Níveis

Um *fator* é a entidade que particiona a população em subconjuntos disjuntos, exaustivos e não-vazios, chamados de níveis do fator. Com isso, pode-se dizer que os

fatores classificam os indivíduos segundo os seus níveis. Por exemplo: o fator sexo particiona a populção de pessoas em dois subconjuntos disjuntos cuja interseção é vazia, que são os níveis masculino e feminino.

1.1.3 - Combinação de Níveis de um Conjunto de Fatores ou Tratamento

Seja Z um determinado conjunto de fatores. *Tratamento* é uma combinação dos níveis dos fatores em Z, e será denotado por letras minúsculas. Por exemplo, z denotará a combinação de níveis do conjunto de fatores em Z. Se Z consistir apenas de um fator, então z será um nível desse fator.

1.1.4 - Resposta e População Conceitual de Respostas

A cada indivíduo estarão associadas medidas, denominadas respostas, que são características numéricas obtidas a partir da aplicação de tratamentos sobre o indivíduo. As respostas serão denotadas pela letra maiúscula Y e dependerão inteiramente de um número finito de fatores, tais como: temperatura, pressão, etc. Esses fatores serão indicados por índices subscritos em Y e a variação de cada indíce estará de acordo com o número de níveis do respectivo fator.

Formalmente, a população conceitual de respostas é definida da seguinte forma: seja z uma combinação de níveis que representa um tratamento, e seja i o índice que denota o i-ésimo indivíduo na população de inivíduos. Considere Y_{iz} a resposta do i-ésimo indivíduo ao tratamento z. Então, o conjunto dos Y_{iz} (considerando todos os i's e z's) forma a população conceitual de respostas.

1.1.5 - Estrutura Hierárquica e Cruzada

Um fator B é dito ser embutido dentro de um fator A, se a identificação plena de um nível de B requer, também, a especificação de um nível de A. Por causa dessa ordem na especificação é que se dá o nome de *estrutura hierárquica.*

A estrutura hierárquica tem a propriedade de transitividade, ou seja, se A embute B e B embute C, então A embute C. E essa relação valerá para qualquer número de fatores. A estrutura entre dois fatores será cruzada quando todos os níveis de um dos fatores tiverem como par todos os níveis do outro fator, na referência[17] está uma definição rigorosa de tais estruturas.

A seguir, é dada uma notação para ambas as estruturas. O símbolo : indicará a presença da estrutura hierárquica, e os fatores que estiverem à sua esquerda embutirão os que estiverem à sua direita. Quando a relação dos fatores for de cruzamento, um virá ao lado do outro. Parêntesis serão usados para facilitar a visualização das estrututras. Para ilustrar, dois exemplos são dados:

1° ◊ Considere a estrutura existente entre três fatores (B, T e P), onde B embute P e T está cruzado com B. Na notação previamente definida tem-se que:

(B : P) (T);

2 ° ◊ Considere a estrutura onde quatro fatores (P, Q, R e S) têm a seguinte relação: Q está embutido em S e R está embutido no cruzamento SP, simbolicamente:

(S : Q) (P) e (SP : R).

1.1.6 - Estrutura de uma Populacão de Respostas

A *estrutura de uma população de respostas* é definida pelo número de seus fatores e pela relação existente entre eles, que pode ser ou de embutimento, ou cruzada ou uma mistura deles.

1.1.7 - Estrutura Completa

Seja Z um conjunto de fatores em determinada estrutura de resposta. Suponha que todos os z em um experimento ocorram[1]. Se isto for verdadeiro para cada conjunto de fatores, isto é, cada subconjunto de X, o conjunto de todos os fatores, a estrutura é chamada de *completa* . Um exemplo simples de uma estrutura completa em um experimento de dois fatores (linhas e colunas), é uma matriz de duas entradas, onde cada casela contém pelo menos uma resposta e o número de colunas é maior ou igual ao número de linhas. Esta estrutura é um exemplo de blocos completos, onde a coluna faz o papel dos blocos. A estrutura de blocos incompletos, ou seja, aquela em que um ou mais tratamentos não ocorrem em um ou mais blocos, é um exemplo de estruturas incompletas.

1.1.8 - Estrutura Balanceada

Seja Z um conjunto de fatores de determinada estrutura de uma população de

[1] Quando uma combinação de níveis dos fatores tem pelo menos uma resposta no experimento, diz-se que esta combinação ocorre. Isso fará mais sentido no capítulo seguinte quando se tiver lidando com uma amostra da população conceitual de respostas.

resposta. Se todos os z contiverem o mesmo número de respostas, então Z é um conjunto *balanceado* .

A estrutura balanceada e completa são características independentes. Uma estrutura pode ter ambas, uma ou nenhuma delas. Apenas as estruturas balanceadas serão consideradas neste trabalho. Novamente, ambos os conceitos farão mais sentido quando usados em uma amostra da população conceitual de respostas, a parte de amostragem será introduzida no capítulo seguinte.

1.1.9 - Conjunto de Fatores e Médias Admissíveis

Seja T um conjunto de fatores em determinada estrutura de uma população de resposta. Se T contém cada fator que embute pelo menos um de seus fatores, então T é um conjunto *admissível*. Por definição, o conjunto vazio é, também, um conjunto admissível. É fácil ver que cada combinação de níveis de um conjunto de fatores é uma combinação de algum conjunto admissível. Por exemplo: se W é um conjunto de fatores (e não necessariamente todos os fatores de uma população de respostas) e V_w é o conjunto de todos os fatores que embutem um ou mais fatores em W, além do próprio W, então $V_w \cup W$ é um conjunto admissível.

Uma *média admissível* é uma média artimética de todas as respostas populacionais que estão contidas em w (uma combinação de W), e é denotada por Y_w. Neste trabalho se considerará apenas as médias admissíveis.

1.1.10 - O Conjunto do Parêntese mais à Direita (CPMD)[2]

[2]O termo técnico *parêntese mais à direita* vem do inglês right most bracket. Este termo juntamente com o termo admissível foram propostos por Zyskind (1958) - Error Structures in Experimental Designs. Tese de Ph.D não publicada. Biblioteca, Iowa State University of Science and

Sejam T e W_t conjuntos, tais que o primeiro é admissível e o segundo contém todos os fatores em T que não possuem nenhum outro fator de T embutido em si. Então, W_t é chamado de *CPMD* de T. Seja $V_t = T - W_t$, então T será escrito como T = V(W), para ressaltar o fato de que W_t é o CPMD de T. Note que V_t e W_t são conjuntos disjuntos e cada fator em V_t embute pelo menos um fator em W_t. Observe que se W_t = { } ↔ T = { }. Por outro lado, se V_t = { } → T é o seu próprio CPMD, isto é, T = (T). É conveniente indicar os índices pertencentes ao CPMD entre parêntesis.

1.2 - O Modelo ou Identidade Populacional

1.2.1 - Média Parcial Admissível

Antes de concentrar a atenção nos modelos ou identidades populacionais, precisa-se da definição de médias parciais, de fundamental importância no desenvolvimeto dessa teoria.

Média parcial é a média aritmética simples sobre a dimensão de um particular conjunto de índices. A média parcial não é uma observação da população de respostas, mas, sim, uma medida dela. Elas serão denotadas pelo símbolo usual de uma resposta, porém com os índices onde os cálculos foram feitos omitidos. Restringir-se-á a atenção somente às médias parciais admissíveis. Por exemplo:

1° ◊ Considere a seguinte estrutura (B : P) (T), também conhecida como estrutura Aleatorizada em Blocos, B (blocos), T (tratamentos) e P (parcelas). Existem seis tipos de médias parciais, denotadas por:

Technology, Ames, Iowa.

$$Y,\ Y_{(i)},\ Y_{(k)},\ Y_{(ik)},\ Y_{i(j)},\ Y_{i(jk)},\ \textit{onde}\quad i = 1,...,b\ ;\quad j = 1,...,p\ \textit{e}\quad k = 1,...,t.$$

2° ◊ Considere a seguinte estrutura (S : Q) (P) e (SP : R), nela oito tipos de médias parciais estão presentes:

$$Y,\ Y_{(i)},\ Y_{(j)},\ Y_{(ij)},\ Y_{i(k)},\ Y_{i(jk)},\ Y_{ij(l)},\ Y_{ij(kl)},\quad \textit{onde}\quad i = 1,..,s;\quad j = 1,...,p;\quad k = 1,...,q;\ \textit{e}\quad l = 1,...,r.$$

1.2.1.1 - Subespaços Vetoriais Relativos aos Vetores de Médias Parciais Admissíveis

Seria aconselhável que o leitor que não esteja familiarizado com os conceitos de, espaço e subespaço vetoriais, projeções, ortogonalidade e outras definições de álgebra linear, lesse antes o apêndice A, que dá uma breve introdução a esses conceitos.

Cada tipo de média parcial admissível tem um significado maior que uma simples conta de média aritmética, isto é, cada uma delas está imersa em um subespaço vetorial. Por exemplo: na estrutura (B : P) (T), cada um dos seis tipos de médias parciais Y, $Y_{(i)}$, $Y_{(k)}$, $Y_{(ik)}$, $Y_{i(j)}$, $Y_{i(jk)}$, representam um vetor em um determinado subespaço. Os subespaços relativos aos tipos de médias parciais não são disjuntos entre si, pois, há uma interseção entre eles, além do subespaço trivial nulo. Existem duas maneiras de se mostrar tal interseção. A primeira delas é através da relação entre médias parciais, que pertencem à mesma estrutura de resposta.

Exemplo 1) Sejam S_1 e S_2 conjuntos de índices de duas médias parciais, de tal forma

que $S_1 \subset S_2$, isto é, S_2 tem um número maior de índices que S_1. Se se pegar a média parcial com índices S_2 e somar sobre os índices em S_2 que excedem S_1 (índices excedentes), e depois dividir o resultado da soma pelo número de elementos somados, obter-se-á exatamente a média parcial com índices S_1. Isto significa, que o subespaço relativo a média parcial com índices S_1 está contido no subespaço relativo à outra média com índices S_2.

Exemplo 2) Sejam S_1 e S_2 conjuntos de índices de duas médias parciais, onde $S_1 \not\subset S_2$. Mesmo nesse caso, as duas médias parciais terão uma interseção, além do subespaço nulo. Somando-se os índices em S_1 e dividindo pela dimensão de S_1, obter-se-á a média geral. O mesmo acontecerá se o procedimento acima for aplicado na média com índices S_2, de modo que, ambos os espaços contêm a média geral.

A segunda maneira usa conceitos de álgebra linear.

Exemplo 1) As condições são as mesmas do exemplo 1 supra citado. Em outras palavras, somando-se todos os vetores que pertençam ao subespaço da média parcial com índices S_2, que excedam os índices em S_1, e multiplicando o vetor resultante da soma pelo inverso do número de vetores que o compõem, ter-se-á o vetor que representa a média parcial com índices S_1. Como a soma de vetores que pertençam ao mesmo subespaço e a multiplicação de um vetor por um escalar não fazem com que o vetor resultante saia do subespaço em questão, conclui-se que o subespaço da média parcial com índices S_1 está contido no subespaço da média parcial com índices S_2.

O uso de matrizes e suas notações vêm simplificar bastante o manuseio das expressões, pois, apesar do espaço ser finito, lida-se frequentemente com grandes dimensões.

Através do produto de matrizes específicas pelo vetor de respostas consegue-se

chegar às médias parciais. É possível criar uma matriz aumentada, cujas colunas são formadas pelas colunas das matrizes que geram as médias parciais. Tal matriz será denominada matriz geradora de médias (mgm). No caso da estrutura aleatorizada em blocos (B : P) (T), onde cada um dos índices (i, j e k) tem dois níveis, a matriz está especificada abaixo. Pode-se notar, por aqui também, que os espaços vetoriais pertinentes às médias parciais não são disjuntos entre si, pois basta somar algumas colunas geradoras de uma média parcial para se obter colunas geradoras de outras médias parciais. Por exemplo: somando-se as duas primeiras colunas da matriz geradora da média $Y_{(ik)}$ e depois repetindo o mesmo procedimento para as duas últimas, o resultado será uma matriz com duas colunas, que, a menos de uma constante, é idêntica à matriz geradora da média Y_i.

$$mgm_{(i:j)(k)} = \begin{bmatrix}
\frac{1}{8} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{8} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{8} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{8} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\frac{1}{8} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\frac{1}{8} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

$$Y \quad Y_{(i)} \quad Y_{(k)} \quad Y_{(ik)} \quad Y_{i(j)} \quad Y_{i(jk)}$$

1.2.1.2 - Construção de Subespaços Disjuntos[3]

Já foi visto, anteriormente, que os subespaços relativos às médias parciais possuem interseção entre si. Com isso, usam-se projeções ortogonais de vetores, diferenças entre vetores projetados e as suas projeções ortogonais, para gerar subespaços ortogonais, e portanto, disjuntos. Raciocínio análogo será empregado no vetor de respostas populacionais, isto é, ele será fatorado em componentes, ortogonais entre si, cuja soma dará o próprio vetor de observações. Os componentes ortogonais serão funções lineares das médias parciais admissíveis. É também importante observar que ao se projetar ortogonalmente um componente em um subespaço, esse subespaço tem que ser disjunto ao subespaço do componente. Daí, não se projetar ortogonalmente o componente diretamente no subespaço de uma média parcial , mas sim no subespaço dessa média parcial que não contenha nenhum outro subespaço, aonde já se fez projeções ortogonais de outros componentes. É importante salientar que essa decomposição é única. O exemplo ilustra essa idéia da decomposição:

Exemplo 1) Considere a estrutura (B : P) (T). As médias parciais admissíveis são:

$$Y,\ Y_{(i)},\ Y_{(k)},\ Y_{(ik)},\ Y_{i(j)},\ Y_{i(jk)},\ \textit{onde}\quad i = 1,\ldots,b\ ;\quad j = 1,\ldots,p\ \textit{e}\quad k = 1,\ldots,t.$$

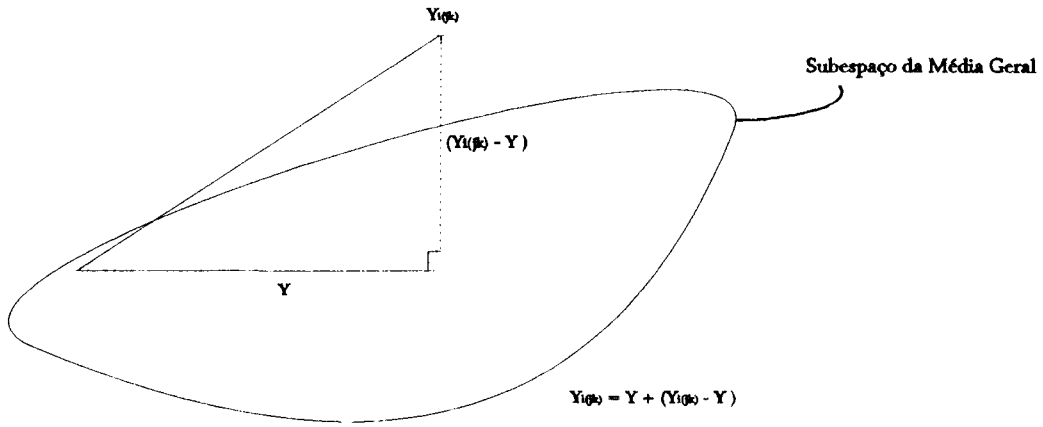
O vetor de respostas será fatorado ou decomposto em uma soma de componentes ortogonais entre si, como mostra o esquema abaixo. Pode-se notar que o número de componentes, ao final da decomposição, é exatamente igual ao número de médias parciais. O mesmo raciocínio, pode perfeitamente, ser estendido para qualquer estrutura.

As duas seções anteriores foram necessárias para que fosse dada uma motivação,

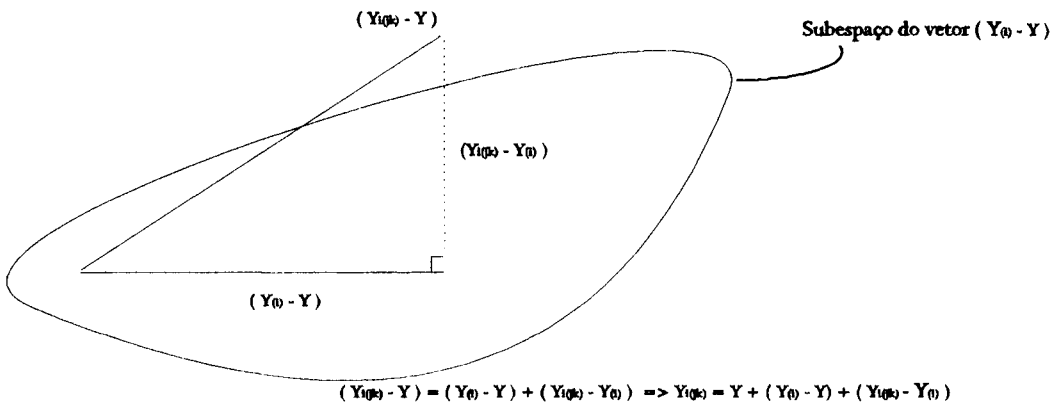
[3] Entende-se subespaços disjuntos, àqueles subespaços cujas interseções entre si são o subespaço trivial nulo.

visão e interpretação geométrica do modelo ou identidade populacional, e também para que esse modelo não fosse entendido apenas como uma expansão matemática e algébrica.

O vetor $Y_{i(jk)}$ será projetado ortogonalmente sobre o subespaço da média geral

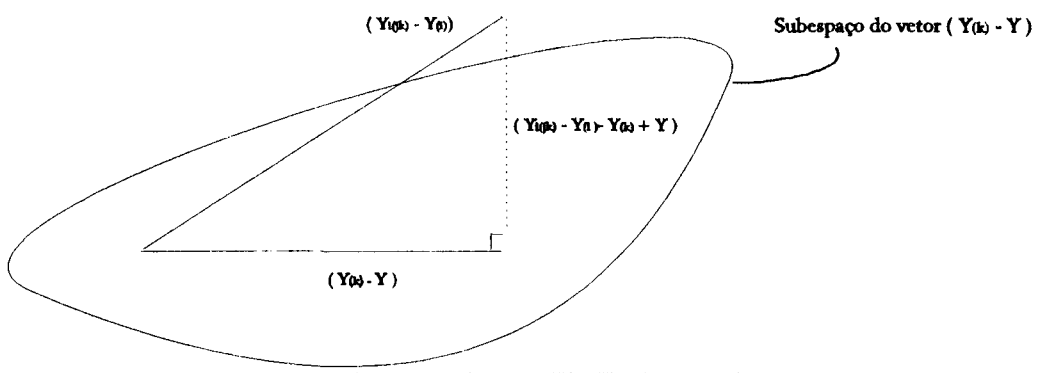


Agora, o vetor $(Y_{i(jk)} - Y)$ será projetado ortogonalmente no subespaço $(Y_{(i)} - Y)$



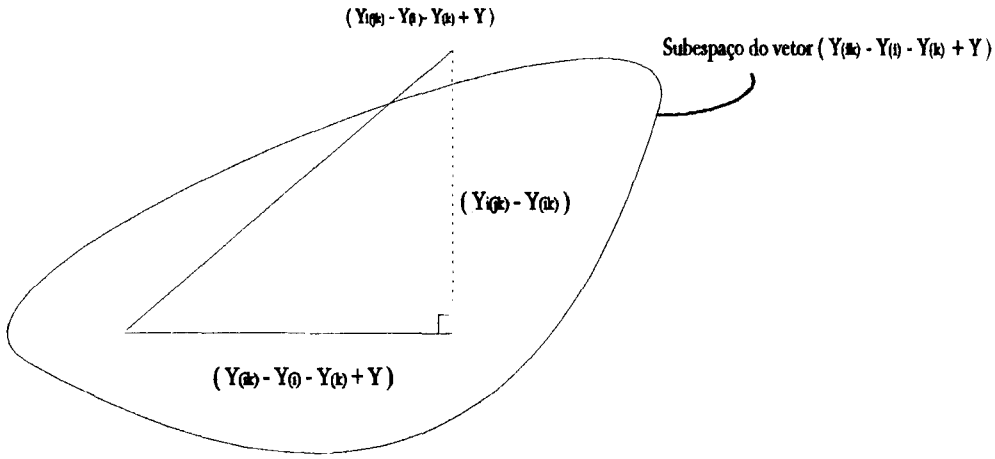
$(Y_{i(jk)} - Y) = (Y_{(i)} - Y) + (Y_{i(jk)} - Y_{(i)}) \Rightarrow Y_{i(jk)} = Y + (Y_{(i)} - Y) + (Y_{i(jk)} - Y_{(i)})$

Agora, o vetor $(Y_{i(jk)} - Y_{(i)})$ será projetado no subespaço $(Y_{(k)} - Y)$



$(Y_{i(jk)} - Y_{(i)}) = (Y_{(k)} - Y) + (Y_{i(jk)} - Y_{(i)} - Y_{(k)} + Y) \Rightarrow Y_{i(jk)} = Y + (Y_{(i)} - Y) + (Y_{(k)} - Y) + (Y_{i(jk)} - Y_{(i)} - Y_{(k)} + Y)$

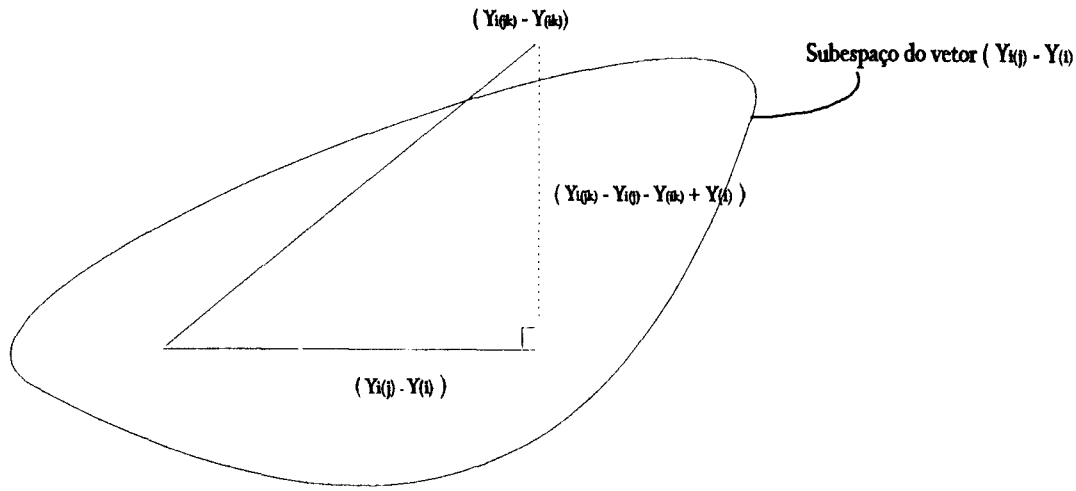
Agora, o vetor ($Y_{i(jk)} - Y_{(i)} - Y_{(k)} + Y$) será projetado no subespaço ($Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y$)



$$(Y_{i(jk)} - Y_{(i)} - Y_{(k)} + Y) = (Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y) + (Y_{i(jk)} - Y_{(ik)}) \Rightarrow$$

$$Y_{i(jk)} = Y + (Y_{(i)} - Y) + (Y_{(k)} - Y) + (Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y) + (Y_{i(jk)} - Y_{(ik)})$$

Agora, o vetor ($Y_{i(jk)} - Y_{(ik)}$) será projetado no subespaço ($Y_{i(j)} - Y_{(i)}$)



$$(Y_{i(jk)} - Y_{(ik)}) = (Y_{i(j)} - Y_{(i)}) + (Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{(i)}) \Rightarrow$$

$$Y_{i(jk)} = Y + (Y_{(i)} - Y) + (Y_{(k)} - Y) + (Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y) + (Y_{i(j)} - Y_{(i)}) + (Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{(i)})$$

1.2.2 - Algoritmo de Obtenção dos Componentes Através das Médias Parciais

O processo que utiliza projeções ortogonais para a obtenção de componentes ortogonais vai se complicando e ficando tedioso, à medida que o número de fatores cresce. Zyskind (*1962 - On structure, relation, Σ, and expectation of mean square, Sankhya Ser. A, 24:115-148*) desenvolveu um algoritmo automatizando o processo de criação dos componentes usando as médias parciais admissíveis e o CPMD.

Combinações lineares de médias parciais podem ser obtidas de cada média parcial. Estas combinações, até aqui chamadas de componentes , são formadas pela seleção de todas as médias parciais produzidas pela média em questão (termo principal), quando um, alguns ou todos os índices do CPMD são omitidos, considerando todas as possibilidades.

Algoritmo: Sempre que um número ímpar de índices pertencentes ao CPMD for omitido, a média terá o coeficiente (-1), se um número par for omitido, a média terá o coeficiente mais um (+1). O zero é considerado um número par. Por exemplo:

1° ◊ Considere a estrutura aleatorizada em blocos, ou seja, (B : P) (T) ou (i : j) (k);

Média Parcial (Termo Principal)	Conduz	Componente
Y	→	(Y)
$Y_{(i)}$	→	($Y_{(i)}$ - Y)
$Y_{(k)}$	→	($Y_{(k)}$ - Y)

$Y_{(ik)}$	→	$(Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y)$
$Y_{i(j)}$	→	$(Y_{i(j)} - Y_{(i)})$
$Y_{i(jk)}$	→	$(Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{(i)})$

2° ◇ Considere a estrutura (S : Q) (P) e (SP : R) ou (i : j) (k) e (ik : l)

Média Parcial (Termo Principal)	Conduz	Componente
Y	→	(Y)
$Y_{(i)}$	→	$(Y_{(i)} - Y)$
$Y_{(k)}$	→	$(Y_{(k)} - Y)$
$Y_{(ik)}$	→	$(Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y)$
$Y_{i(j)}$	→	$(Y_{i(j)} - Y_{(i)})$
$Y_{i(jk)}$	→	$(Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{(i)})$
$Y_{ik(l)}$	→	$(Y_{ik(l)} - Y_{(ik)})$
$Y_{ik(jl)}$	→	$(Y_{ik(jl)} - Y_{i(jk)} - Y_{ik(l)} + Y_{(ik)})$

As propriedades, mostradas a seguir, são consequências imediatas da definição do algoritmo:

(1) Se o CPMD do termo principal contiver p índices, então o número de médias presentes no componente é 2^p.

Prova: $\sum_{n=0}^{p} \binom{n}{p} = (1 + 1)^p = 2^p$

(2) Excluindo o componente (Y), cuja a soma do coeficiente é um, a soma dos

coeficientes das médias em qualquer outro componente é zero.

Prova: $\sum_{n=0}^{p} (-1)^n \binom{n}{p} = (1-1)^p = 0$, *para* $p \neq 0$.

Os componentes populacionais satisfazem algumas propriedades bastante úteis, quando a população considerada for balanceada. Logo abaixo estão enumeradas tais propriedades. Detalhes das demostrações podem ser encontradas em Zyskind(1958).

1 ⇨ A soma das respostas de um componente qualquer sobre os índices do CPMD do termo principal é zero. Por exemplo:

$$\sum_{i} (Y_{(i)} - Y) = 0.$$

2 ⇨ A soma de quadrados de uma resposta sobre todos os índices é igual a soma de quadrados de seus componentes sobre os mesmos índices. Por exemplo:

$$\sum_{ij} Y_{i(j)}^2 = \sum_{ij} Y^2 + \sum_{ij} (Y_{(i)} - Y)^2 + \sum_{ij} (Y_{i(j)} - Y_{(i)})^2.$$

3 ⇨ O número de valores linearmente independentes (ou graus de liberdade) de um componente é igual ao produto da variação populacional dos índices que não pertençam ao CPMD vezes a variação reduzida[4] dos índices pertencentes ao CPMD do termo principal. Assim, a soma destes números sobre os componentes é igual a N. Por exemplo: considere a estrutura (I), as médias parciais são Y e $Y_{(i)}$, com i = 1,...,I. Y tem um grau

[4] Variação reduzida significa a variação populacional menos uma unidade.

de liberdade e ($Y_{(i)}$ - Y) tem (I - 1) graus de liberdade, de modo que a soma dos graus de liberdade dá o número total de elementos, I.

4 ⇨ Para qualquer componente o quadrado da soma sobre os índices do CPMD é igual a soma de quadrados sobre o mesmo conjunto de índices, com os coeficientes das médias parciais sendo mantidos inalterados. Por exemplo:

$$\sum_{jk} (Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{(i)})^2 = \sum_{jk} (Y^2_{i(jk)} - Y^2_{i(j)} - Y^2_{(ik)} + Y^2_{(i)}).$$

Corolário: a propriedade 4 também é válida quando a soma é feita sobre todos os índices de uma resposta.

Para a definição dos cap sigmas (Σ) precisa-se antes definir os componentes de variação (c.v.). Um componente de variação de um componente nada mais é do que a soma de quadrados das médias parciais desses componentes divididos pelos seus graus de liberdade. O c.v. será simbolicamente representado por σ^2, com o mesmo subscrito do respectivo componente.

Os cap sigmas, definidos abaixo, são funções lineares dos componentes de variação. Para cada componente da identidade populacional haverá um cap sigma correspondente.

Definição: O Σ de um determinado componente é a combinação linear dos σ^2 que tiverem os índices do termo principal (desse determinado componente) como um subconjunto, e que os índices excedentes sejam exclusivamente do CPMD dos σ^2. O coeficiente de cada c.v. que tem k índices excedentes é:

$$(- 1)^k \frac{1}{\textit{produto da variação populacional dos índices excedentes}}$$

Obs.: O componente de variação relativo à media geral será denotado por $\sigma^2_e = Y^2$. Para fixar as idéias considere os exemplos abaixo:

1° ⇨ (B : P) (T);

Identidade populacional:

$Y_{i(jk)} = Y + (Y_{(i)} - Y) + (Y_{(k)} - Y) + (Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y) + (Y_{i(j)} - Y_{(i)}) + (Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{i}) = \mu + B_i + T_k + (BT)_{ik} + P_{i(j)} + (TP)_{i(jk)}$.

$$\Sigma_{\varnothing} = \sigma^2_{\varnothing} - \frac{1}{B}\sigma^2_B - \frac{1}{T}\sigma^2_T + \frac{1}{BT}\sigma^2_{BT}$$

$$\Sigma_B = \sigma^2_B - \frac{1}{P}\sigma^2_{B(P)} - \frac{1}{T}\sigma^2_{BT} + \frac{1}{PT}\sigma^2_{B(PT)}$$

$$\Sigma_T = \sigma^2_T - \frac{1}{B}\sigma^2_{BT}$$

$$\Sigma_{BT} = \sigma^2_{BT} - \frac{1}{P}\sigma^2_{B(PT)}$$

$$\Sigma_{B(P)} = \sigma^2_{B(P)} - \frac{1}{T}\sigma^2_{B(PT)}$$

$$\Sigma_{B(PT)} = \sigma^2_{B(PT)}.$$

2 ° ⇨ (S : Q) (P) e (SP : R);

Identidade populacional:

$Y_{ik(jl)} = Y + (Y_{(i)} - Y) + (Y_{(k)} - Y) + (Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y) + (Y_{i(j)} - Y_{(i)}) + (Y_{i(jk)} - Y_{(ik)} - Y_{i(j)} + Y_{(i)}) + (Y_{ik(l)} - Y_{(ik)}) + (Y_{ik(jl)} - Y_{i(jk)} - Y_{ik(l)} + Y_{(ik)}) =$

$= \mu + S_i + P_k + (SP)_{ik} + S(Q)_{i(j)} + S(PQ)_{i(jk)} + SP(R)_{ik(l)} + SP(QR)_{ik(jl)}$.

$$\Sigma_{\varnothing} = \sigma^2_{\varnothing} - \frac{1}{S}\sigma^2_S - \frac{1}{P}\sigma^2_P + \frac{1}{SP}\sigma^2_{SP}$$

$$\Sigma_S = \sigma^2_S - \frac{1}{P}\sigma^2_{SP} - \frac{1}{Q}\sigma^2_{S(Q)} + \frac{1}{PQ}\sigma^2_{S(PQ)}$$

$$\Sigma_P = \sigma^2_P - \frac{1}{S}\sigma^2_{SP}$$

$$\Sigma_{SP} = \sigma^2_{SP} - \frac{1}{Q}\sigma^2_{S(PQ)} - \frac{1}{R}\sigma^2_{SP(R)} + \frac{1}{QR}\sigma^2_{SP(QR)}$$

$$\Sigma_{S(Q)} = \sigma^2_{S(Q)} - \frac{1}{P}\,\sigma^2_{S(QP)}$$

$$\Sigma_{S(PQ)} = \sigma^2_{S(PQ)} - \frac{1}{R}\,\sigma^2_{SP(QR)}$$

$$\Sigma_{SP(R)} = \sigma^2_{SP(R)} - \frac{1}{Q}\,\sigma^2_{SP(QR)}$$

$$\Sigma_{SP(QR)} = \sigma^2_{SP(QR)}.$$

Fisher (*1918 - The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. Edin. 52:339 - 433.*) introduziu a análise de variância baseado na quebra do total da soma de quadrados em partes distintas. Cada uma delas, supostamente, com um significado físico, no sentido de descrever o acesso a uma ou complexas fontes de variações. Em cada linha da ANOVA o quadrado médio é:

$$\textit{Quadrado Médio} = (\textit{produtos das dimensões populacionais dos índices que não estão envolvidos no componente da linha em questão}) * * \frac{\sum_{\textit{índices do componente}} (\textit{componente})^2}{G.L.}$$

$$= (\textit{nº de respostas que entram como termo principal em um componente}) * \sigma^2_{\textit{componente}}.$$

Capítulo II

Amostragem e Função Indicadora

Neste capítulo será introduzido o processo que representa uma amostra da população conceitual de respostas. Neste contexto, cada elemento da população é um valor fixo e desconhecido, e portanto, não é uma variável aleatória. Através da identidade populacional, não é difícil ver que a amostra dos elementos populacionais induzem, também, a amostras dos componentes populacionais. Por exemplo: considere a estrutura i(j). Dessa forma, $Y_{i(j)}$ pode ser expresso como $Y_{i(j)} = Y + (Y_{(i)} - Y) + (Y_{i(j)} - Y_{(i)})$. Com isso, cada particular $(Y_{(i)} - Y)$ e $(Y_{i(j)} - Y_{(i)})$ selecionados dependem exclusivamente dos $Y_{i(j)}$'s escolhidos, sendo assim, diz -se que as amostras daqueles são induzidas pela amostras destes.

As variáveis aleatórias utilizadas serão funções indicadoras, que determinam o tipo de amostragem realizada sobre a população. Desse modo, cada elemento da amostra será uma função das variáveis aleatórias indicadoras e dos elementos da população conceitual de respostas. Esse processo vai permitir o cálculo das expectâncias necessárias para montar a tabela de análise de varância amostral em função dos cap sigmas.

2.1 - Amostragem Aleatória Simples Sem Reposição (AASSR) de uma População de Tamanho N

Considere uma população de respostas com N elementos que estão representados por Y_i, $i = 1,...,N$. Se uma AASSR for efetuada sobre esta população, então cada conjunto de n elementos dos Y_i's terá a mesma probabilidade igual à $\frac{1}{\binom{N}{n}}$ de constituir a

amostra.

Seja y_{i^*}, $i^* = 1,...,n$; a i-ésima escolha amostral. Dessa forma, $P(y_{i^*} = Y_i) = \frac{1}{N}$, onde $i^* = 1,...,n$ e $i = 1,...,N$ e $P(y_{i^*} = Y_i , y_{i^{*'}} = Y_{i'}) = \frac{1}{N(N-1)}$. Pode-se ligar a observação amostral com os elementos da população, através da variável aleatória indicadora definida da seguinte forma:

$$\begin{cases} \delta_i^{i^*} = 1, \textit{ se a } i^*\textit{-ésima escolha da amostra for o } i\textit{-ésimo elemento populacional}; \\ \delta_i^{i^*} = 0, \textit{ caso contrário } (c.c.). \end{cases} \tag{2.1}$$

Então, Nn variáveis aleatórias $\delta_i^{i^*}$'s têm a mesma função de probabilidade, induzidas pelo tipo de amostragem. A função de probabilidade é dada abaixo, juntamente com algumas propriedades importantes dessas variáveis aleatórias.

$$\begin{aligned} &P(\delta_i^{i^*} = 1) = \frac{1}{N} = 1 - P(\delta_i^{i^*} = 0) \Rightarrow E(\delta_i^{i^*}) = E[(\delta_i^{i^*})^2] = \frac{1}{N}, \ \forall i, i^*; \\ &P(\delta_i^{i^*} \delta_{i'}^{i^*} = 1) = 0 \Rightarrow E(\delta_i^{i^*} \delta_{i'}^{i^*}) = 0, \ \forall i \neq i' \ e \ i^*; \\ &P(\delta_i^{i^*} \delta_i^{i^{*'}} = 1) = 0 \Rightarrow E(\delta_i^{i^*} \delta_i^{i^{*'}}) = 0, \ \forall i \ e \ i^* \neq i^{*'}; \\ &P(\delta_i^{i^*} \delta_{i'}^{i^{*'}} = 1) = \frac{1}{N(N-1)} \Rightarrow E(\delta_i^{i^*} \delta_{i'}^{i^{*'}}) = \frac{1}{N(N-1)}, \ \forall i \neq i' \ e \ i^* \neq i^{*'}. \end{aligned} \tag{2.2}$$

Obs.: os símbolos P e E representam respectivamente a probabilidade e a expectância.

Assim, cada observação amostral é dada como uma função linear das variáveis aleatórias indicadoras, bem como dos elementos da população, através da relação:

$$y_{i^*} = \sum_{i=1}^{N} \delta_i^{i^*} Y_i = \sum_{i=1}^{N} \delta_i^{i^*} [Y + (Y_i - Y)] = \mu + \sum_{i=1}^{N} \delta_i^{i^*} A_i = \mu + e_{i^*}, \tag{2.3}$$

onde $\mu = Y = \sum_{i=1}^{N} \frac{Y_i}{N}$, $A_i = (Y_i - Y)$ e $e_{i^*} = \sum_{i=1}^{N} \delta_i^{i^*} A_i$.

Note que o modelo acima é o modelo mais trivial e simples na teoria de modelos lineares, todavia, ele foi construído com base num processo físico de amostragem e não suposto. Desse modo,

$$E(y_i^*) = \mu + \frac{1}{N}\sum_{i=1}^{N} A_i, \tag{2.4}$$

como $\sum_{i=1}^{N} A_i = \sum_{i=1}^{N} (Y_i - Y) = 0$, tem-se que:

$$E(y_i^*) = \mu \Leftrightarrow E(e_i^*) = 0, \tag{2.5}$$

por construção. E ainda,

$$E\left(y_i^{*2}\right) = E\left[\mu + \sum_{i=1}^{N} \delta_i^{i*}(Y_i - Y)\right]^2 = E\left\{\mu^2 + 2\mu\sum_{i=1}^{N} \delta_i^{i*}(Y_i - Y) + \left[\sum_{i=1}^{N} \delta_i^{i*}(Y_i - Y)\right]^2\right\} \tag{2.6}$$

Calculando as expectâncias dos termos com $\delta_{i^*}^{i}$'s usando (2.2) obtém-se:

$$\begin{aligned}
E\left[2\mu\sum_{i=1}^{N} \delta_i^{i*}(Y_i - Y)\right] &= 2\mu\sum_{i=1}^{N} (Y_i - Y)\,E(\delta_i^{i*}) = 2\mu\frac{1}{N}\sum_{i=1}^{N} (Y_i - Y) = 0 \\
E\left[\sum_{i=1}^{N} \delta_i^{i*}(Y_i - Y)\right]^2 &= E\left[\sum_{i=1}^{N} \delta_i^{i*}(Y_i - Y)^2 + 2\sum_{1\le i<i'\le N} \delta_i^{i*}\delta_{i'}^{i'*}(Y_i - Y)(Y_{i'} - Y)\right] = \\
&= \sum_{i=1}^{N} (Y_i - Y)^2 E(\delta_i^{i*}) + 2\sum_{1\le i<i'\le N} (Y_i - Y)(Y_{i'} - Y)\,E(\delta_i^{i*}\delta_{i'}^{i'*}) = \\
&= \frac{1}{N}\sum_{i=1}^{N} (Y_i - Y)^2.
\end{aligned} \tag{2.7}$$

Substituindo (2.7) em (2.6) obtém-se:

$$E\left(y_{i*}^2\right) = \mu^2 + \frac{1}{N}\sum_{i=1}^{N}(Y_i - Y)^2 = \mu^2 + \frac{(N-1)}{N}\frac{\sum_{i=1}^{N}(Y_i - Y)^2}{N-1} = \tag{2.8}$$

$$= \mu^2 + \left(1 - \frac{1}{N}\right)\frac{\sum_{i=1}^{N}(Y_i - Y)^2}{N-1},$$

fazendo $\frac{\sum_{i=1}^{N}(Y_i - Y)^2}{N-1} = \sigma_A^2$, tem-se:

$$E\left(y_{i*}^2\right) = \left(\mu^2 - \frac{1}{N}\sigma_A^2\right) + \sigma_A^2 = \Sigma_o + \Sigma_A. \tag{2.9}$$

Segue-se, então, que:

$$V(y_{i*}) = E\left(y_{i*}^2\right) - [E(y_{i*})]^2 = \left(1 - \frac{1}{N}\right)\sigma_A^2. \tag{2.10}$$

Ao contrário do que se pensa, a covariância entre as observações obtidas de uma amostra aleatória não é zero, como demonstrado abaixo:

$$cov(y_{i*}, y_{i*'}) = cov(e_{i*}, e_{i*'}) \Rightarrow E(y_{i*}, y_{i*'}) - E(y_{i*})E(y_{i*'}) = E(e_{i*}, e_{i*'}) - E(e_{i*})E(e_{i*'}) =$$

$$= E\left[\sum_{i \neq i'} \delta_i^{i*}\delta_{i'}^{i*'}A_iA_{i'}\right] - 0 = \sum_{i \neq i'} A_iA_{i'}E(\delta_i^{i*}\delta_{i'}^{i*'}) = \frac{1}{N(N-1)}\sum_{i \neq i'} A_iA_{i'}. \tag{2.11}$$

Sabendo que $\sum_{i=1}^{N} A_i = \sum_{i=1}^{N}(Y_i - Y) = 0$, então

$$\left(\sum_{i=1}^{N} A_i\right)^2 = 0 = \sum_{i=1}^{N} A_i^2 + \sum_{i \neq i'} A_i A_{i'} \Rightarrow \sum_{i \neq i'} A_i A_{i'} = -\sum_{i=1}^{N} A_i^2. \tag{2.12}$$

Substituindo (2.12) em (2.11) obtém-se:

$$cov(y_{i*}, y_{i*'}) = -\frac{1}{N(N-1)} \sum_{i=1}^{N} A_i^2 = -\frac{1}{N}\sigma_A^2. \tag{2.13}$$

$$MatCov(n \times n) = \begin{bmatrix} 1 & \cdots & -\frac{1}{N}\sigma_A^2 \\ -\frac{1}{N}\sigma_A^2 & 1 \cdots & -\frac{1}{N}\sigma_A^2 \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ -\frac{1}{N}\sigma_A^2 & \cdots 1 & -\frac{1}{N}\sigma_A^2 \\ -\frac{1}{N}\sigma_A^2 & \cdots & 1 \end{bmatrix}, \tag{2.14}$$

de modo que a matriz de covariância induzida pelo procedimento amostral não é diagonal. A $E(y_{i*}y_{i*'})$ pode ser obtida através de (2.13), da seguinte forma:

$$E(y_{i*}, y_{i*'}) = E(y_{i*})E(y_{i*'}) + cov(y_{i*}, y_{i*'}) = \mu^2 - \frac{1}{N}\sigma_A^2. \tag{2.15}$$

Se se pegar, agora,

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_{i*} = \mu + \frac{1}{n}\sum_{i*=1}^{n}\sum_{i=1}^{N} \delta_i^{i*} A_i, \tag{2.16}$$

não é difícil mostrar que E(y) = μ. Calculando a E(y^2), obtém-se:

$$E(y^2) = E\left(\frac{1}{n}\sum_{i_*=1}^{n} y_{i_*}\right)^2 = \frac{1}{n^2}E\left(\sum_{i_*=1}^{n} y_{i_*}\right)^2 = \frac{1}{n^2}E\left(\sum_{i_*=1}^{n} y_{i_*}^2 + 2\sum_{1\le i_*<i_*'\le n} y_{i_*}\,y_{i_*'}\right) =$$
$$= \frac{1}{n^2}\left[\sum_{i_*=1}^{n} E\left(y_{i_*}^2\right) + 2\sum_{1\le i_*<i_*'\le n} E(y_{i_*}\,y_{i_*'})\right]. \qquad (2.17)$$

Usando (2.9) e (2.15), tem-se que:

$$E(y^2) = \frac{1}{n^2}\left\{\sum_{i_*=1}^{n}\left[\mu^2 + \left(1-\frac{1}{N}\right)\sigma_A^2\right] + 2\sum_{1\le i_*<i_*'\le n}\left(\mu^2 - \frac{\sigma_A^2}{N}\right)\right\} =$$
$$= \frac{1}{n^2}\left\{n\left[\mu^2 + \left(1-\frac{1}{N}\right)\sigma_A^2\right] + \frac{2n(n-1)}{2}\left(\mu^2 - \frac{\sigma_A^2}{N}\right)\right\} =$$
$$= \frac{1}{n}\left[\left(\mu^2 - \frac{\sigma_A^2}{N}\right) + \sigma_A^2 + (n-1)\left(\mu^2 - \frac{\sigma_A^2}{N}\right)\right] = \qquad (2.18)$$
$$= \frac{1}{n}\left[\left(\mu^2 - \frac{\sigma_A^2}{N}\right)(n-1+1) + \sigma_A^2\right] = \frac{1}{n}\left[\left(\mu^2 - \frac{\sigma_A^2}{N}\right)n + \sigma_A^2\right] =$$
$$= \mu^2 - \frac{\sigma_A^2}{N} + \frac{\sigma_A^2}{n} = \mu^2 + \left(1-\frac{n}{N}\right)\frac{\sigma_A^2}{n}.$$

Dessa forma, a variância da média amostral é dada por:

$$V(y) = E(y^2) - [E(y)]^2 = \left(1-\frac{n}{N}\right)\frac{\sigma_A^2}{n}. \qquad (2.19)$$

Todavia, a atenção será concentrada na expectância do quadrado da média amostral de tamanho n,

$$E(y^2) = \mu^2 + \left(1 - \frac{n}{N}\right)\frac{\sigma_A^2}{n} = \left(\mu^2 - \frac{\sigma_A^2}{N}\right) + \frac{\sigma_A^2}{n} = \Sigma_e + \frac{\Sigma_A}{n}. \qquad (2.20)$$

Com isso, foi ilustrado através do exemplo de AASSR, o uso da variável aleatória indicadora ou variável de delineamento de amostragem, que será bastante usada na formulação dos modelos lineares deduzidos.

2.2 - Amostragem Aleatória Cruzada

Considere, agora, elementos com valores P_{ij} , dispostos em uma tabela de dupla entrada com A linhas e B colunas, onde os P_{ij}'s estão sujeitos à condição de que

$$\sum_{i=1}^{A} P_{ij} = \sum_{j=1}^{B} P_{ij} = 0. \qquad (2.21)$$

Uma amostra aleatória desses elementos é escolhida da seguinte forma: selecionam-se **a** linhas dentre as A existentes e, independentemente desta seleção **b** colunas são selecionadas dentre as B. A interseção entre as linhas e colunas escolhidas darão origem aos P_{ij}'s na amostra, tal esquema recebe a denominação de *amostragem cruzada*. Esta amostra pode ser representada pela figura abaixo:

Entidade B

Entidade A

	0	0			0				
					0				
	0	0							
	0	0			0				

Novamente, o esquema pode ser representado por variáveis aleatórias de amostragem. Sejam

$$\begin{cases} \delta_i^{i*} = 1, \text{ se a } i*\text{-ésima escolha seleciona todos os } P_{ij}\text{/s com subscrito } i; \\ \delta_i^{i*} = 0, \text{ c.c.} \end{cases}$$
$$\begin{cases} \beta_j^{j*} = 1, \text{ se a } j*\text{-ésima escolha seleciona todos os } P_{ij}\text{/s com subscrito } j; \\ \beta_j^{j*} = 0, \text{ c.c.} \end{cases} \tag{2.22}$$

Desse modo, a observação amostral y_{i*j*} é expressa como uma função dos P_{ij}'s da seguinte forma:

$$y_{i*j*} = \sum_{ij} \delta_i^{i*} \beta_j^{j*} P_{ij}. \tag{2.23}$$

Algumas propriedades dessas variáveis aleatórias estão apresentadas abaixo:

$$P(\delta_i^{i*} = 1) = \frac{1}{A} \Leftrightarrow E(\delta_i^{i*}) = \frac{1}{A}, \ \forall\ i \ e\ i*;$$

$$P(\beta_j^{j*} = 1) = \frac{1}{B} \Leftrightarrow E(\delta_j^{j*}) = \frac{1}{B}, \ \forall\ j \ e\ j*;$$

$$P(\delta_i^{i*}\ \delta_{i'}^{i*} = 1) = P(\beta_j^{j*}\ \beta_{j'}^{j*} = 1) = 0, \ \forall\ i \neq i', i*, j \neq j', e\ j*;$$

$$P(\delta_i^{i*}\ \delta_i^{i*'} = 1) = P(\beta_j^{j*}\ \beta_j^{j*'} = 1) = 0, \ \forall\ i, i* \neq i*', j, j* \neq j*'; \tag{2.24}$$

$$P(\delta_i^{i*}\ \delta_{i'}^{i*'} = 1) = P(\delta_{i'}^{i*'} = 1/\delta_i^{i*} = 1)\,P(\delta_i^{i*} = 1) = \frac{1}{A(A-1)}, \ \forall\ i \neq i' \ e\ i* \neq i*';$$

$$P(\beta_j^{j*}\ \beta_{j'}^{j*'} = 1) = \frac{1}{B(B-1)}, \ \forall\ j \neq j' \ e\ j* \neq j*';$$

$$E(\delta_i^{i*}\,\delta_{i'}^{i*'}\,\beta_j^{j*}\,\beta_{j'}^{j*'}) = E(\delta_i^{i*}\,\delta_{i'}^{i*'})\,E(\beta_j^{j*}\,\beta_{j'}^{j*'}) = \frac{1}{AB(A-1)(B-1)},\ \forall\ i \neq i',\ i* \neq i*',\ j \neq j' \text{ e } j* \neq j*'.$$

Usando (2.21) obtêm-se as relações:

$$\left(\sum_{ij} P_{ij}\right)^2 = 0 = \sum_{ij} P_{ij}^2 + \sum_{\substack{i \\ j \neq j'}} P_{ij}P_{ij'} + \sum_{\substack{j \\ i \neq i'}} P_{ij}P_{i'j} + \sum_{\substack{i \neq i' \\ j \neq j'}} P_{ij}P_{i'j'}$$

$$\sum_{\substack{i \\ j \neq j'}} P_{ij}P_{ij'} = -\sum_{ij} P_{ij}^2$$

$$\sum_{\substack{j \\ i \neq i'}} P_{ij}P_{i'j} = -\sum_{ij} P_{ij}^2$$

$$\Rightarrow \sum_{\substack{i \neq i' \\ j \neq j'}} P_{ij}P_{i'j'} = \sum_{ij} P_{ij}^2 \quad (2.25)$$

A $\text{cov}(y_{i^*j^*}, y_{i^{*'}j^{*'}})$, $\text{cov}(y_{i^*j^*}, y_{i^*j^{*'}})$ e $\text{cov}(y_{i^*j^*}, y_{i^{*'}j^*})$, da mesma forma que em (2.13), não são nulas, como demonstrado abaixo, usando (2.24) e (2.25):

$$cov(y_{i*j*}, y_{i*'j*'}) = E(y_{i*j*}\,y_{i*'j*'}) - E(y_{i*j*})\,E(y_{i*'j*'}) =$$

$$= E\left[\left(\sum_{ij} \delta_i^{i*}\,\beta_j^{j*}\,P_{ij}\right)\left(\sum_{i'j'} \delta_{i'}^{i*'}\,\beta_{j'}^{j*'}\,P_{i'j'}\right)\right] =$$

$$= \sum_{ij}\sum_{i'j'} P_{ij}\,P_{i'j'}\,E(\delta_i^{i*}\,\beta_j^{j*}\,\delta_{i'}^{i*'}\,\beta_{j'}^{j*'}) = \frac{1}{AB(A-1)(B-1)}\sum_{ij}\sum_{i'j'} P_{ij}\,P_{i'j'} \quad (2.26)$$

$$= \frac{\sum_{ij} P_{ij}^2}{AB(A-1)(B-1)} = \frac{\sigma_{AB}^2}{AB},\ \textit{onde}\ \sigma_{AB}^2 = \frac{\sum_{ij} P_{ij}^2}{(A-1)(B-1)}.$$

mesmo acontece com a $\text{cov}(y_{i^*j^*}, y_{i^*j^{*'}})$ e $\text{cov}(y_{i^*j^*}, y_{i^{*'}j^*})$, usando (2.24) e (2.25):

$$var(y_{i*j*}) = \sum_{ij}\sum_{i'j'} P_{ij}\,P_{i'j'}\,E(\delta_i^{i*}\,\delta_{i'}^{i*'}\,\beta_j^{j*}\,\beta_{j'}^{j*'}) = \frac{\sum_{ij} P_{ij}^2}{AB} = \sigma_{AB}^2\left(1 - \frac{1}{A}\right)\left(1 - \frac{1}{B}\right).$$

$$cov(y_{i_*j_*}, y_{i_*j_*'}) = \sum_{\substack{i \\ j \neq j'}} P_{ij}P_{ij'}E(\delta_i^{i_*}\beta_j^{j_*}\beta_{j'}^{j_*'}) = \frac{\sum_{i,j\neq j'} P_{ij}P_{ij'}}{AB(B-1)} = -\frac{\sum_{ij} P_{ij}^2}{AB(B-1)} = -\frac{\sigma_{AB}^2}{B}\left(1 - \frac{1}{A}\right); \quad (2.27)$$

$$cov(y_{i_*j_*}, y_{i_*'j_*}) = \sum_{\substack{j \\ i \neq i'}} P_{ij}P_{i'j}E(\delta_i^{i_*}\delta_{i'}^{i_*'}\beta_j^{j_*}) = \frac{\sum_{j,i\neq i'} P_{ij}P_{i'j}}{AB(A-1)} = -\frac{\sum_{ij} P_{ij}^2}{AB(A-1)} = -\frac{\sigma_{AB}^2}{A}\left(1 - \frac{1}{B}\right);$$

Trabalhando, agora, com a média amostral y e usando (2.26) e (2.27), tem-se que:

$$E(y^2) = E\left(\frac{1}{ab}\sum_{i_*j_*} y_{i_*j_*}\right)^2 = \frac{1}{(ab)^2}E\left(\sum_{i_*j_*} y_{i_*j_*}\right)^2 =$$

$$= \frac{1}{(ab)^2}E\left(\sum_{i_*j_*} y_{i_*j_*}^2 + \sum_{\substack{i_* \\ j_*\neq j_*'}} y_{i_*j_*}y_{i_*j_*'} + \sum_{\substack{j_* \\ i_*\neq i_*'}} y_{i_*j_*}y_{i_*'j_*} + \sum_{\substack{i_*\neq i_*' \\ j_*\neq j_*'}} y_{i_*j_*}y_{i_*'j_*'}\right) =$$

$$= \frac{1}{(ab)^2}\left[\sum_{i_*j_*} E(y_{i_*j_*}^2) + \sum_{\substack{i_* \\ j_*\neq j_*'}} E(y_{i_*j_*}y_{i_*j_*'}) + \sum_{\substack{j \\ i_*\neq i_*'}} E(y_{i_*j_*}y_{i_*'j_*}) + \sum_{\substack{i_*\neq i_*' \\ j_*\neq j_*'}} E(y_{i_*j_*}y_{i_*'j_*'})\right] =$$

$$= \frac{1}{(ab)^2}\left[ab\frac{\sigma_{AB}^2}{AB}(A-1)(B-1) - ab(b-1)\frac{\sigma_{AB}^2}{AB}(A-1) + \right. \quad (2.28)$$

$$\left. - a(a-1)b\frac{\sigma_{AB}^2}{AB}(B-1) + a(a-1)b(b-1)\frac{\sigma_{AB}^2}{AB}\right] =$$

$$= \frac{\sigma_{AB}^2}{ABab}[(A-1)(B-1) - (A-1)(b-1) - (a-1)(B-1) + (a-1)(b-1)] =$$

$$= \frac{\sigma_{AB}^2}{ABab}(A-a)(B-b) = \frac{\sigma_{AB}^2}{ab}\left(1 - \frac{a}{A}\right)\left(1 - \frac{b}{B}\right) \diamond$$

$$E(y^2) = \frac{\sigma_{AB}^2}{ab}\left(1 - \frac{a}{A}\right)\left(1 - \frac{b}{B}\right).$$

Então, o fator de correção acima envolve cada classificação e o σ^2_{AB} é dividido pelo tamanho da amostra. Uma observação de extrema importância, é que a condição (2.21) é uma consequência da estrutura balanceada da população de respostas. Logo não é uma imposição assumida por um modelo matemático. Uma propriedade bastante interessante para (2.28) é o fato de que quando uma seleção envolver todos os elementos, de pelo menos uma das classificações, a E(y^2) será zero. Implicando que y é identicamente igual a zero. O que significa, na terminologia estatística, que o(s) fator(es) de classificação(ões) é(são) fixo(s). Aplicar-se-á, agora, os resultados, vistos até aqui, em uma situação clássica de planejamento de experimentos. Seja $Y_{(ij)}$, i=1, . . . , A; j=1, . . . , B, uma resposta da casela (i,j) em uma classificação balanceada com dois fatores cruzados, A e B. Usando a decomposição de $Y_{(ij)}$ em componentes ortogonais obtém-se:

$$\begin{aligned} Y_{ij} &= Y + (Y_{(i)} - Y) + (Y_{(j)} - Y) + (Y_{(ij)} - Y_{(i)} - Y_{(j)} + Y) = \\ &= \mu + A_{(i)} + B_{(j)} + (AB)_{(ij)}, \end{aligned} \tag{2.29}$$

de forma que a (i*j*)-ésima observação amostral está relacionada com os $Y_{(ij)}$'s por:

$$y_{(i^*j^*)} = \mu + \sum_i \delta_i^{i^*} A_{(i)} + \sum_j \beta_j^{j^*} B_{(j)} + \sum_{ij} \delta_i^{i^*} \beta_j^{j^*} (AB)_{(ij)}. \tag{2.30}$$

Aproveitando os resultados (2.18) e (2.28), previamente calculados, tem-se que a expressão para a E(y^2) é da forma:

$$\begin{aligned} E(y^2) &= \mu^2 + \left(1 - \frac{a}{A}\right)\frac{\sigma^2_{(A)}}{a} + \left(1 - \frac{b}{B}\right)\frac{\sigma^2_{(B)}}{b} + \left(1 - \frac{a}{A}\right)\left(1 - \frac{b}{B}\right)\frac{\sigma^2_{(AB)}}{ab} = \\ &= \Sigma_e + \frac{1}{a}\Sigma_{(A)} + \frac{1}{b}\Sigma_{(B)} + \frac{1}{ab}\Sigma_{(AB)}, \end{aligned} \tag{2.31}$$

onde $\sigma_A^2 = \dfrac{\sum_{i=1}^{A} A_i^2}{(A-1)} = \dfrac{\sum_{i=1}^{A} (Y_{(i)} - Y)^2}{(A-1)}$, $\sigma_B^2 = \dfrac{\sum_{i=1}^{B} B_i^2}{(B-1)}$ e $\sigma_{AB}^2 = \dfrac{\sum_{ij} P_{ij}^2}{(A-1)(B-1)}$.

O fato do Σ ser uma consequência dos σ^2 será visto adiante com mais detalhes, onde o caso exposto acima é uma situação particular.

2.3 - Amostragem Aleatória com Estrutura Hierárquica

Finalmente, um último exemplo, agora com uma estrutura hierárquica. Considere a estrutura A(B) ou i(j). Cada $Y_{i(j)}$ denotará o j-ésimo elemento do i-ésimo grupo. A identidade populacional para este caso é:

$$Y_{i(j)} = Y + (Y_{(i)} - Y) + (Y_{i(j)} - Y_{(i)}) = \mu + A_{(i)} + B_{i(j)}. \tag{2.32}$$

Como a população é balanceada, as condições abaixo são satisfeitas:

$$\sum_{i=1}^{A} A_{(i)} = \sum_{j=1}^{B} B_{i(j)} = 0, \ \forall i. \tag{2.33}$$

A amostragem dos $Y_{i(j)}$'s segue o seguinte procedimento: selecionar aleatoriamente **a** dentre os A existentes, e para cada um dos **a** selecionados escolher **b** dentre os B existentes. A observação amostral será denotada por $y_{i^*(j^*)}$, isto é, a j*-ésima escolha dentro do i*-ésimo grupo. A ligação entre a observação amostral $y_{i^*(j^*)}$ e a observação populacional será feita através de abAB variáveis aleatórias do tipo $\delta_i^{i^*}$, $\beta_j^{j^*}$, onde i=1, . . ., A ; i*= 1, . . ., a ; j= 1, . . .,B e j*= 1, . . .,b.

$$\begin{cases} \beta_{i^* j}^{i^* j^*} = 1, \textit{ se a } j^*\textit{-ésima escolha do } i^*\textit{-ésimo grupo for a } j\textit{-ésima unidade populacional deste grupo}; \\ \beta_{i^* j}^{i^* j^*} = 0, \textit{ c.c.} \end{cases} \tag{2.34}$$

e os δ_{i*}^{i}'s são definidos exatamente como em (2.1), de modo que:

$$y_{i*(j*)} = \sum_{ij} \delta_i^{i*} \beta_{i*j}^{i*j*} Y_{i(j)}, \tag{2.35}$$

substituindo (2.32) em (2.35) obtém-se:

$$y_{i*(j*)} = \sum_{ij} \delta_i^{i*} \beta_{i*j}^{i*j*} (\mu + A_{(i)} + B_{i(j)}) = \mu + \sum_{i} \delta_i^{i*} A_{(i)} + \sum_{ij} \delta_i^{i*} \beta_{i*j}^{i*j*} B_{i(j)}. \tag{2.36}$$

A função de probabilidade dos (β_{i*j}^{i*j*}) e dos $(\delta_i^{i*} \beta_{i*j}^{i*j*})$ estão definidas abaixo, com:

$$P(\beta_{i*j}^{i*j*} = 1) = \frac{1}{AB}, \ \forall\ j,\ i*\ e\ j*;$$

$$P(\delta_i^{i*} \beta_{i*j}^{i*j*} = 1) = \frac{1}{B} \diamond E(\delta_i^{i*} \beta_{i*j}^{i*j*}) = \frac{1}{B}\ ,\ \forall\ i,\ j,\ i*\ e\ j*;$$

$$P(\delta_i^{i*} \delta_{i'}^{i*} \beta_{i*j}^{i*j*} = 1) = 0 \diamond E(\delta_i^{i*} \delta_{i'}^{i*} \beta_{i*j}^{i*j*}) = 0\ ,\ \forall\ i \neq i',\ j,\ i*\ e\ j*;$$

$$P(\delta_i^{i*} \beta_{i*j}^{i*j*} \beta_{i*j'}^{i*j*} = 1) = 0 \diamond E(\delta_i^{i*} \beta_{i*j}^{i*j*} \beta_{i*j'}^{i*j*}) = 0\ ,\ \forall\ i,\ j \neq j',\ i*\ e\ j*;$$

$$P(\delta_i^{i*} \delta_{i'}^{i*} \beta_{i*j}^{i*j*} \beta_{i*j'}^{i*j*} = 1) = 0 \diamond E(\delta_i^{i*} \delta_{i'}^{i*} \beta_{i*j}^{i*j*} \beta_{i*j'}^{i*j*}) = 0\ ,\ \forall\ i \neq i',\ j \neq j',\ i*\ e\ j*; \tag{2.37}$$

$$P(\delta_i^{i*} \beta_{i*j}^{i*j*} \beta_{i*j'}^{i*j*'} = 1) = \frac{1}{B(B-1)} \diamond E(\delta_i^{i*} \beta_{i*j}^{i*j*} \beta_{i*j'}^{i*j*'}) = \frac{1}{B(B-1)}\ ,\ \forall\ i,\ j \neq j',\ i*\ e\ j* \neq j*';$$

$$P(\delta_i^{i*} \beta_{i*j}^{i*j*} \beta_{i*j}^{i*j*'} = 1) = 0 \diamond E(\delta_i^{i*} \beta_{i*j}^{i*j*} \beta_{i*j}^{i*j*'}) = 0\ ,\ \forall\ i,\ j,\ i*\ e\ j* \neq j*';$$

$$P(\delta_i^{i*} \delta_{i'}^{i*'} \beta_{i*j}^{i*j*} \beta_{i*'j}^{i*'j*} = 1) = P(\delta_i^{i*} \beta_{i*j}^{i*j*}) P(\delta_{i'}^{i*'} \beta_{i*'j}^{i*'j*})\ ,\ \forall\ i \neq i',\ j,\ i* \neq i*'\ e\ j*;$$

$$P(\delta_i^{i*} \delta_{i'}^{i*'} \beta_{i*j}^{i*j*} \beta_{i*'j}^{i*'j*'} = 1) = P(\delta_i^{i*} \beta_{i*j}^{i*j*}) P(\delta_{i'}^{i*'} \beta_{i*'j}^{i*'j*'})\ ,\ \forall\ i \neq i',\ j,\ i* \neq i*'\ e\ j* \neq j*'.$$

A seguir, estão alguns resultados que serão usados em futuras demonstrações:

$$(y_{i*j*})^2 = \left(\mu + \sum_{i} \delta_i^{i*} A_{(i)} + \sum_{ij} \delta_i^{i*} \beta_{i*j}^{i*j*} B_{i(j)} \right)^2 = \mu^2 + 2\mu \sum_{i} \delta_i^{i*} A_{(i)} + 2\mu \sum_{ij} \delta_i^{i*} \beta_{i*j}^{i*j*} B_{i(j)} + \tag{2.38}$$

$$+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)^{2}+2\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)+\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)^{2}.$$

$$(y_{i*j*})(y_{i*'j*})=\left(\mu+\sum_{i}\delta_{i}^{i*}A_{(i)}+\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\mu+\sum_{i}\delta_{i}^{i*'}A_{(i)}+\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*}B_{i(j)}\right)=$$

$$=\mu^{2}+\mu\sum_{i}\delta_{i}^{i*}A_{(i)}+\mu\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*}B_{i(j)}+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{i}\delta_{i}^{i*'}A_{(i)}\right)+$$

$$+\left(\sum_{i}\delta_{i}^{i*'}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*}B_{i(j)}\right)+ \tag{2.39}$$

$$+\mu\sum_{i}\delta_{i}^{i*'}A_{(i)}+\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*}B_{i(j)}\right)+\mu\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}.$$

$$(y_{i*j*})(y_{i*j*'})=\left(\mu+\sum_{i}\delta_{i}^{i*}A_{(i)}+\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\mu+\sum_{i}\delta_{i}^{i*}A_{(i)}+\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*'}B_{i(j)}\right)=$$

$$=\mu^{2}+2\mu\sum_{i}\delta_{i}^{i*}A_{(i)}+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)^{2}+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*'}B_{i(j)}\right)+$$

$$+\mu\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*'}B_{i(j)}+\mu\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)+ \tag{2.40}$$

$$+\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*'}B_{i(j)}\right).$$

$$(y_{i*j*})(y_{i*'j*'})=\left(\mu+\sum_{i}\delta_{i}^{i*}A_{(i)}+\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\mu+\sum_{i}\delta_{i}^{i*'}A_{(i)}+\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*'}B_{i(j)}\right)= \tag{2.41}$$

$$=\mu^{2}+\mu\sum_{i}\delta_{i}^{i*}A_{(i)}+\mu\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*'}B_{i(j)}+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{i}\delta_{i}^{i*'}A_{(i)}\right)+\mu\sum_{i}\delta_{i}^{i*'}A_{(i)}+$$

$$\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)+\mu\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}+$$

$$+\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)\left(\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*'}B_{i(j)}\right)+\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\sum_{ij}\delta_{i}^{i*'}\beta_{i*'j}^{i*'j*'}B_{i(j)}\right).$$

$$\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)^{2}=\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}^{2}+\sum_{\substack{i\\ j\neq j'}}\delta_{i}^{i*}\beta_{i*j}^{i*j*}\beta_{i*j'}^{i*j*}B_{i(j)}B_{i(j')}+\sum_{\substack{j\\ i\neq i'}}\delta_{i}^{i*}\delta_{i'}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}B_{i'(j)}+$$

$$+\sum_{\substack{i\neq i'\\ j\neq j'}}\delta_{i}^{i*}\delta_{i'}^{i*}\beta_{i*j}^{i*j*}\beta_{i*j'}^{i*j*}B_{i(j)}B_{i'(j')}. \tag{2.42}$$

$$\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*'}B_{i(j)}\right)=\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}\beta_{i*j}^{i*j*'}B_{i(j)}^{2}+\sum_{\substack{i\\ j\neq j'}}\delta_{i}^{i*}\beta_{i*j}^{i*j*}\beta_{i*j}^{i*j*'}B_{i(j)}B_{i(j')}+$$

$$\sum_{\substack{j\\ i\neq i'}}\delta_{i}^{i*}\delta_{i'}^{i*}\beta_{i*j}^{i*j*}\beta_{i*j}^{i*j*'}B_{i(j)}B_{i'(j)}+\sum_{\substack{i\neq i'\\ j\neq j'}}\delta_{i}^{i*}\delta_{i'}^{i*}\beta_{i*j}^{i*j*}\beta_{i*j'}^{i*j*'}B_{i(j)}B_{i'(j')}. \tag{2.43}$$

É fácil mostrar que $E(y_{i^*j^*}) = \mu$. Resta, então, calcular os momentos cruzados dessas variáveis aleatórias. Usando (2.2), (2.36), (2.37) e (2.38) tem-se que:

$$E\left[(y_{i*(j*)})^{2}\right]=\mu+E\left(\sum_{i}A_{(i)}\delta_{i}^{i*}\right)^{2}+E\left(\sum_{ij}\delta_{i}^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)^{2}. \tag{2.44}$$

Sabe-se a partir de (2.9), que $E\left(\sum_{i}\delta_{i}^{i*}A_{(i)}\right)^{2}=\left(1-\frac{1}{A}\right)\sigma_{(A)}^{2}$, de modo que, para o cálculo de (2.44) falta apenas a última parcela da sua soma. Usando (2.33),(2.37) e (2.42) obtém-se:

$$E\left(\sum_{ij}\delta_i^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)^2 = E\left(\sum_{ij}\delta_i^{i*}\beta_{i*j}^{i*j*}B_{i(j)}^2\right) = \sum_{ij}B_{i(j)}^2 E\left(\delta_i^{i*}\beta_{i*j}^{i*j*}\right) =$$
$$= \frac{1}{B}\sum_{ij}B_{i(j)}^2 = \left(1 - \frac{1}{B}\right)\sigma_{A(B)}^2, \tag{2.45}$$

onde $\sigma_{A(B)}^2 = \dfrac{\sum_{ij}B_{i(j)}^2}{B-1}$. O resultado final de (2.44), usando (2.45), é da forma:

$$E\left[(y_{i*j*})^2\right] = \mu^2 + \left(1 - \frac{1}{A}\right)\sigma_{(A)}^2 + \left(1 - \frac{1}{B}\right)\sigma_{A(B)}^2. \tag{2.46}$$

A $E(y_{i^*j^*}\, y_{i^*j^{*\prime}})$ também é calculada usando (2.33), (2.37)e (2.40), e ainda (2.43):

$$E(y_{i*j*}y_{i*j*'}) = \mu^2 + E\left(\sum_i \delta_i^{i*}A_{(i)}\right) + E\left[\left(\sum_{ij}\delta_i^{i*}\beta_{i*j}^{i*j*}B_{i(j)}\right)\left(\sum_{ij}\delta_i^{i*}\beta_{i*j}^{i*j*'}B_{i(j)}\right)\right] =$$
$$= \mu^2 + \left(1 - \frac{1}{A}\right)\sigma_{(A)}^2 + \frac{1}{B(B-1)}\sum_{\substack{i \\ j*j'}} B_{i(j)}B_{i(j')} = \tag{2.47}$$
$$= \mu^2 + \left(1 - \frac{1}{A}\right)\sigma_{(A)}^2 - \frac{1}{B(B-1)}\sum_{ij}B_{i(j)}^2 = \mu^2 + \left(1 - \frac{1}{A}\right)\sigma_{(A)}^2 - \frac{\sigma_{A(B)}^2}{B}.$$

Usando (2.12), (2.37), (2.39),(2.41) e o fato de que $E(y_{i^*j^*}) = \mu$, tem-se que:

$$E(y_{i*j*}\, y_{i*'j*}) = E(y_{i*j*}\, y_{i*'j*'}) = \mu^2 - \frac{\sigma_{(A)}^2}{A}. \tag{2.48}$$

Assim, a expectância do quadrado da média amostral y é:

$$E\left[(y)^2\right] = \frac{1}{(ab)^2}E\left(\sum_{i*j*}y_{i*j*}\right)^2 = \tag{2.49}$$

$$= E\left(\sum_{i_* j_*} y_{i_* j_*}^2 + \sum_{\substack{i_* \\ j_* \neq j_*'}} y_{i_* j_*} y_{i_* j_*'} + \sum_{\substack{j_* \\ i_* \neq i_*'}} y_{i_* j_*} y_{i_*' j_*} + \sum_{\substack{i_* \neq i_*' \\ j_* \neq j_*'}} y_{i_* j_*} y_{i_*' j_*'} \right).$$

Substituindo (2.46), (2.47) e (2.48) em (2.49) obtém-se:

$$E[(y)^2] = \frac{1}{(ab)^2}\left\{ ab\left[\mu^2 + \left(1 - \frac{1}{A}\right)\sigma_{(A)}^2 + \left(1 - \frac{1}{B}\right)\sigma_{A(B)}^2\right] + \left[\mu^2 + \left(1 - \frac{1}{A}\right)\sigma_{(A)}^2 - \frac{\sigma_{A(B)}^2}{B}\right] x \right.$$

$$\left. x\ ab(b-1) + a(a-1)b\left[\mu^2 - \frac{\sigma_{(A)}^2}{A}\right] + a(a-1)b(b-1)\left[\mu^2 - \frac{\sigma_{(A)}^2}{A}\right]\right\} = \quad (2.50)$$

$$= \mu^2 + \left(1 - \frac{a}{A}\right)\sigma_{(A)}^2 + \left(1 - \frac{b}{B}\right)\frac{\sigma_{A(B)}^2}{ab} = \Sigma_e + \frac{1}{a}\Sigma_{(A)} + \frac{1}{ab}\Sigma_{A(B)}.$$

É importante salientar que os fatores de correção $\left(1 - \frac{a}{A}\right)$ e $\left(1 - \frac{b}{B}\right)$ correspondem aos índices subscritos pertencentes exclusivamente ao CPMD.

2.4 - As Expectâncias dos Quadrados das Médias Amostrais em Função dos σ^2's e dos Σ's

A seguir, uma série de teoremas, corolários e resultados serão dados para estabelecer a relação entre a expectância do quadrado de uma média amostral e os σ^2's e Σ's.

Teorema 2.1 (Zyskind - 1958): A expectância do quadrado de qualquer média parcial de uma amostra balanceada é igual a uma função linear de todos os componentes de variação da população. O coeficiente de cada componente é o produto de fatores de correção finitos, um fator para cada índice do CPMD do componente, denotado por f_i^a.

O significado de f_i^a é o seguinte: **a** indicará o número de valores amostrais do índice **i** que entram na formação da média parcial em questão, e **i** é o tamanho do índice populacional. Além disso, o componente de variação será dividido por 1 quando o índice desse componente também aparecer na média parcial que se está fazendo a expectância. Caso contrário, será dividido pelo produto dos tamanhos amostrais dos índices do componente que excedem os índices da média parcial que se está fazendo a expectância.

As expectâncias de y^2 e $y_{i..}^2$, previamente calculadas, exemplificam o teorema 2.1. Um outro exemplo, agora concreto, é dado a seguir: Há a necessidade da investigação da variabilidade de um tratamento de um determinado material. Suponha que este material seja estratificado em S fontes, com cada uma tendo um cruzamento de R linhas e C colunas, onde as colunas, dentro das fontes, sejam subdivididas em M faixas. Note que as faixas não estão embutidas nas linhas. Simbolicamente, a representação da estrutura populacional é (S:(R)(C:M)). A amostra consiste da seleção aleatória de **s** fontes das S, e em cada uma das selecionadas, selecionam-se aleatoriamente **r** linhas das R e **c** colunas das C e por último, de cada coluna escolhida **m** faixas são selecionadas das M existentes. Pelo teorema 2.1 a expressão para a $E(y^2)$ em termos dos σ^{2}'s é:

$$E(y^2) = \mu^2 + \left(1 - \frac{s}{S}\right)\frac{\sigma^2_{(S)}}{s} + \left(1 - \frac{r}{R}\right)\frac{\sigma^2_{S(R)}}{sr} + \left(1 - \frac{c}{C}\right)\frac{\sigma^2_{S(C)}}{sc} +$$
$$+ \left(1 - \frac{r}{R}\right)\left(1 - \frac{c}{C}\right)\frac{\sigma^2_{S(RC)}}{src} + \left(1 - \frac{m}{M}\right)\frac{\sigma^2_{SC(M)}}{scm} + \qquad (2.51)$$
$$+ \left(1 - \frac{r}{R}\right)\left(1 - \frac{m}{M}\right)\frac{\sigma^2_{SC(RM)}}{scrm}.$$

Por outro lado, a expressão para o quadrado da média amostral da m-ésima faixa da c-ésima coluna da s-ésima fonte é:

$$E\left(y^2_{s\bullet\bullet m\bullet}\right) = \mu^2 + \left(1 - \frac{1}{S}\right)\frac{\sigma^2_{(S)}}{1} + \left(1 - \frac{r}{R}\right)\frac{\sigma^2_{S(R)}}{r} + \left(1 - \frac{1}{C}\right)\frac{\sigma^2_{S(C)}}{1} + \left(1 - \frac{r}{R}\right)\left(1 - \frac{1}{C}\right)\frac{\sigma^2_{S(RC)}}{r} + \left(1 - \frac{1}{M}\right)\frac{\sigma^2_{SC(M)}}{1} + \left(1 - \frac{r}{R}\right)\left(1 - \frac{1}{M}\right)\frac{\sigma^2_{SC(RM)}}{r}. \quad (2.52)$$

Agora, a E(y^2) será dada em função dos Σ's.

1º Passo: considere um particular σ^2 na expressão da E(y^2). Seja X o conjunto dos índices que não pertençam ao CPMD de σ^2 e Y o conjunto complementar, isto é, os subscrito pertencentes ao CPMD. Seja Z um subconjunto arbitrário tal que $Z \subseteq Y$. Seja $q=\#^1 Z$, e para q=0 o produto das dimensões dos subscritos em Z é 1. Seja $N_{X(Y-Z)}$ o número de diferentes componentes que entram na formação da média amostral parcial, expecificado pelo conjunto de índices X +(Y - Z). A forma expandida de escrever σ^2 é:

$$\sigma^2_{X(Y)} \sum_{Z \subseteq Y} \frac{(-1)^q}{\prod_{i \in Z} \textit{dimensão populacional}} \times \frac{1}{N_{X+(Y-Z)}}. \quad (2.53)$$

Logo abaixo, está uma tabela com um exemplo bem detalhado da aplicação dos resultados acima. A primeira coluna da tabela representa o conjunto X, a segunda coluna representa o conjunto Y, a terceira coluna representa todas as possibilidades do conjunto Z para um X e Y fixados, e a quarta e última coluna traz a expressão do componente de variação para estes conjuntos.

Exemplo para a estrutura (S:(R) (C:M)):

[1] #, significa a cardinalidade do conjunto.

$\notin$ *CPMD*	$\in$ *CPMD*	$Z \subseteq Y$	*Média amostral parcial*
X	Y	Z	
$\varnothing$	$\varnothing$	$\varnothing$	$y^2 = \mu^2 = \sigma^2_{\varnothing}$
$\varnothing$	{S}	$\varnothing$ {S}	$\left.\begin{matrix} y_{(s\bullet)} \\ \\ y \end{matrix}\right\} \sigma^2_{(S)}\left(\frac{1}{s} - \frac{1}{S}\right) + \left(1 - \frac{s}{S}\right)\frac{\sigma^2_{(S)}}{s}$
{S}	{R}	$\varnothing$ {R}	$\left.\begin{matrix} y_{s\bullet(r\bullet)} \\ \\ y_{(s\bullet)} \end{matrix}\right\} \sigma^2_{S(R)}\left(\frac{1}{sr} - \frac{1}{R}\frac{1}{s}\right) + \left(1 - \frac{r}{R}\right)\frac{\sigma^2_{S(R)}}{sr}$
{S}	{C}	$\varnothing$ {C}	$\left.\begin{matrix} y_{s\bullet(c\bullet)} \\ \\ y_{(s\bullet)} \end{matrix}\right\} \sigma^2_{S(C)}\left(\frac{1}{sc} - \frac{1}{C}\frac{1}{s}\right) + \left(1 - \frac{c}{C}\right)\frac{\sigma^2_{S(C)}}{sc}$
{S}	{R,C}	$\varnothing$ {R} {C} {R,C}	$\left.\begin{matrix} y_{s\bullet(rc\bullet)} \\ y_{s\bullet(c\bullet)} \\ y_{s\bullet(r\bullet)} \\ y_{(s\bullet)} \end{matrix}\right\} \sigma^2_{S(RC)}\left(\frac{1}{src} - \frac{1}{R}\frac{1}{sc} - \frac{1}{C}\frac{1}{sr} + \frac{1}{RC}\frac{1}{s}\right) + \left(1 - \frac{r}{R}\right)\left(1 - \frac{c}{C}\right)\frac{\sigma^2_{S(R)}}{src}$
{S,C}	{M}	$\varnothing$ {M}	$\left.\begin{matrix} y_{sc\bullet(m\bullet)} \\ \\ y_{sc\bullet(c\bullet)} \end{matrix}\right\} \sigma^2_{SC(M)}\left(\frac{1}{scm} - \frac{1}{M}\frac{1}{sc}\right) + \left(1 - \frac{m}{M}\right)\frac{\sigma^2_{SC(M)}}{scm}$
{S,C}	{R,M}	$\varnothing$ {R} {M} {R,M}	$\left.\begin{matrix} y_{sc\bullet(rm\bullet)} \\ y_{sc\bullet(m\bullet)} \\ y_{sc\bullet(cr\bullet)} \\ y_{sc\bullet(c\bullet)} \end{matrix}\right\} \sigma^2_{SC(RM)}\left(\frac{1}{srcm} - \frac{1}{R}\frac{1}{scm} - \frac{1}{M}\frac{1}{scr} + \frac{1}{RM}\frac{1}{sc}\right) + \left(1 - \frac{r}{R}\right)\left(1 - \frac{m}{M}\right)\frac{\sigma^2_{SC(R}}{src}$

2º Passo: seja R um conjunto fixo de índices subscritos. Selecionam-se todos os termos na E(y^2) cujos σ^2 tenham o mesmo X e R=Y-Z, isto é, variam-se Y e Z sujeitos à restrição Y-Z=R. Neste caso, X pode ser também o conjunto de índices do σ^2 que seja o seu próprio CPMD, por exemplo: considere o $\sigma^2_{(S)}$ na expectância (2.51), lá o conjunto de índices do componente de variação é o seu próprio CPMD.

$$\frac{1}{N_{X\,+\,(Y\,-\,Z)}} \times \sum_{\substack{Z \subset Y \\ Y-Z=R}} \sigma^2_{X(Y)} \frac{(-1)^q}{\prod_{i \in Z} (\textit{dimensão populacional})} = \frac{1}{N_{X\,+\,(Y\,-\,Z)}} \Sigma_{X(R)}. \qquad (2.54)$$

Assim,

$$E(y^2) = \sum_{X,R} \left(\frac{\Sigma_{X(R)}}{N_{X+R}} \right) = \textit{soma sobre todos os pares admissíveis de } X \textit{ e } R \textit{ de } \frac{\Sigma_{X(R)}}{N_{X+R}}. \qquad (2.55)$$

Pode-se escrever $N_{X+R} = \prod_{i \in X+R} a_i$, onde a_i é o tamanho amostral do índice i, e $a_i = 1$ quando i pertence ao conjunto vazio, exemplo:

$$\begin{aligned} E(y^2) = {} & \Sigma_{\phi} + \frac{1}{s}\Sigma_{(S)} + \frac{1}{sr}\Sigma_{S(R)} + \frac{1}{sc}\Sigma_{S(C)} + \frac{1}{sr}\Sigma_{S(R)} + \\ & + \frac{1}{src}\Sigma_{S(RC)} + \frac{1}{sr}\Sigma_{S(R)} + \frac{1}{scm}\Sigma_{SC(M)}. \end{aligned} \qquad (2.56)$$

A próxima etapa é escrever a expectância do quadrado de uma média parcial qualquer em função dos Σ's.

Seja S_1 o conjunto de índices em uma média parcial admissível. Então, relativo a esta média, o número de diferentes componentes, cujo tipo é especificado pelo conjunto de índices S_2=X+R é N= $\prod_{i \in (S_2 - S_1)} a_i$.

Teorema 2.2: Sob as condições do teorema 2.1 e com a notação desenvolvida acima, a expansão da expectância do quadrado de uma média parcial qualquer em função dos Σ's é:

$$E\left(y_{S_1}^2\right) = \sum_{X,R}\left(\frac{\Sigma_{X(R)}}{\prod_{i\in S_2 - S_1} a_i}\right).$$

Exemplos destes teoremas podem ser encontrados em (2.9). Para a expectância (2.52) obtém-se:

$$E\left(y_{s\circ\circ m\circ}^2\right) = \Sigma_{\varnothing} + \Sigma_S + \frac{\Sigma_{S(R)}}{r} + \Sigma_{S(C)} + \frac{\Sigma_{S(RC)}}{r} + \Sigma_{SC(M)} + \frac{\Sigma_{SC(RM)}}{r}. \qquad (2.57)$$

Por causa da amostragem pura, isto é, apenas combinações de amostragens cruzadas e/ou embutidas, a estrutura amostral é a mesma da estrutura populacional, a identidade amostral é a mesma da populacional. As definições dos σ^2's, Σ's, graus de liberdade amostrais são completamente análogas às da população.

Por conveniência de notação, se uma letra maiúscula denotar um conjunto de índices qualquer na população, a respectiva letra minúscula indicará o equivalente conjunto na amostra, exemplo: se S denota o conjunto de todos os índices na população, então, **s** indica o conjunto de todos os índices na amostra.

Teorema 2.3[2]: Se x é um subconjunto qualquer de **s** tal que Σ_x é admissível na estrutura amostral ⇔ $E(\Sigma_x)=\Sigma_X$.

Como as outras definições, a análise de variância amostral também tem a sua análoga à populacional. A decomposição da soma de quadrados para uma amostra

[2]Este teorema é devido a Throckmorton, T.N. 1961. Structure of classification data. Tese de Ph.D. não publicada. Iowa State University and Technology, Ames, Iowa.

balanceada está garantida pela existência da identidade e pela estrutura amostral.

Sejam S o conjunto de todos os subscritos, S_1 os subscritos do termo principal em uma linha na ANOVA amostral e S_2 os subscritos de um particular σ^2. S_i pode ser escrito como $S_i=X_i+Y_i$, i=1, 2; onde os X's e Y's são definidos como em (2.53) e (2.54). Sejam $W=Y_2 - (Y_2 \cap S_1)$, e a_i e A_i o tamanho amostral e populacional do índice i respectivamente, ambos serão iguais a 1 quando pertencerem ao conjunto vazio.

Teorema 2.3: O valor esperado do quadrado da média em uma linha da ANOVA amostral, onde os subscritos do termo principal formam o conjunto S_1, tem a forma:

$$\sum_{S_2 \subset S} R(S_1, S_2)\,\sigma^2_{S_2} = \sum_{S_2 \subset S} P(S_1, S_2)\,\Sigma^2_{S_2},$$

onde os R's e P's são constantes com valores:

i) $R(S_1, S_2)=P(S_1, S_2)=0 \leftrightarrow S_1 \not\subset S_2$;

ii) Se $S_1 \subset S_2 \diamond P(S_1, S_2)=P(S_2)$ e $R(S_1, S_2)=P(S_2)Q(W)$, onde $P(S_2)= \prod_{i \in (S - S_2)} a_i$ = número de vezes que o conjunto especificado pelos índices S_2 entram na amostra e $Q(W)= \prod_{i \in W} f_i^{a_i} =$

$$\prod_{i \in W}\left(1 - \frac{\textit{tamanho amostral do índice } i}{\textit{tamanho populacional do índice } i}\right) \qquad (2.58)$$

Corolários:

1. A média de quadrados na ANOVA populacional pode ser obtida, tanto para σ^2 quanto para Σ, a partir da ANOVA amostral substituindo todas as quantidades amostrais pelas correspondentes quantidades na população. Exceto o coeficiente do σ^2 onde $S_1=S_2$ é diferentes de zero.

2. Se algum i $\in$ W tem o seu tamanho amostral igual ao seu tamanho populacional, a contribuição do respectivo σ^2 desaparece, então a classificação correspondente a este índice é fixa.

3. Quando o tamanho populacional de um índice for infinito, o seu respectivo fator de correção será 1.

A seguir, um exemplo da ANOVA amostral para a estrutura(S:(R)(C:M)):

Fonte	G.L.	E.Q.M.
$Y_{\varnothing}$	1	$mrcs\Sigma_{\varnothing} + mrc\Sigma_{(S)} + mc\Sigma_{S(R)} + mr\Sigma_{S(C)} + m\Sigma_{S(RC)} + r\Sigma_{SC(M)} + \Sigma_{SC(RM)} -$ $- mrcs\sigma^2_{\varnothing} + mrc\left(1 - \frac{s}{S}\right)\sigma^2_{(S)} + mc\left(1 - \frac{r}{R}\right)\sigma^2_{S(R)} + mr\left(1 - \frac{c}{C}\right)\sigma^2_{S(C)} +$ $+ m\left(1 - \frac{r}{R}\right)\left(1 - \frac{c}{C}\right)\sigma^2_{S(RC)} + r\left(1 - \frac{m}{M}\right)\sigma^2_{SC(M)} + \left(1 - \frac{r}{R}\right)\left(1 - \frac{m}{M}\right)\sigma^2_{SC(RM)}$
$Y_{(S)}$	s-1	$mrc\Sigma_{(S)} + mc\Sigma_{S(R)} + mr\Sigma_{S(C)} + m\Sigma_{S(RC)} + r\Sigma_{SC(M)} + \Sigma_{SC(RM)} -$ $- mrc\sigma^2_{(S)} + mc\left(1 - \frac{r}{R}\right)\sigma^2_{S(R)} + mr\left(1 - \frac{c}{C}\right)\sigma^2_{S(C)} + r\left(1 - \frac{m}{M}\right)\sigma^2_{SC(M)} +$ $+ m\left(1 - \frac{r}{R}\right)\left(1 - \frac{c}{C}\right)\sigma^2_{S(RC)} + \left(1 - \frac{r}{R}\right)\left(1 - \frac{m}{M}\right)\sigma^2_{SC(RM)}$
$Y_{S(C)}$	s(c-1)	$mr\Sigma_{S(C)} + m\Sigma_{S(RC)} + r\Sigma_{SC(M)} + \Sigma_{SC(RM)} -$ $- mr\sigma^2_{S(C)} + m\left(1 - \frac{r}{R}\right)\sigma^2_{S(RC)} + r\left(1 - \frac{m}{M}\right)\sigma^2_{SC(M)} +$ $+ \left(1 - \frac{r}{R}\right)\left(1 - \frac{c}{C}\right)\sigma^2_{SC(RM)}$

$Y_{S(R)}$	s(r-1)	$mc\Sigma_{S(R)} + m\Sigma_{S(RC)} + \Sigma_{SC(RM)} -$ $- mc\sigma^2_{S(R)} + m\left(1 - \frac{c}{C}\right)\sigma^2_{S(RC)} + \left(1 - \frac{m}{M}\right)\sigma^2_{SC(RM)}$
$Y_{S(RC)}$	s(r-1)(c-1)	$m\Sigma_{S(RC)} + \Sigma_{SC(RM)} -$ $- m\sigma^2_{S(RC)} + \left(1 - \frac{m}{M}\right)\sigma^2_{SC(RM)}$
$Y_{SC(M)}$	sc(m-1)	$r\Sigma_{SC(M)} + \Sigma_{SC(RM)} -$ $- r\sigma^2_{SC(M)} + \left(1 - \frac{r}{R}\right)\sigma^2_{SC(RM)}$
$Y_{SC(RM)}$	sc(r-1)(m-1)	$\Sigma_{SC(RM)} - \sigma^2_{SC(RM)}$

Capítulo III

Experimentos Aleatorizados

Neste capítulo estudar-se-ão as consequências da aleatorização na estrutura experimental, bem como novas definições, como por exemplo os cap sigmas amostrais. Também será apresentada uma representação gráfica de estruturas completas, úteis para sua melhor visualização. E por fim, exemplos serão dados para ilustração dos conceitos e fixação das idéias.

3.1 - Consequências da Aleatorização

Até aqui, estudaram-se as consequências da amostragem pura. Todavia, em planejamento de experimentos, o ato de aleatorizar já é um outro tipo de amostragem, que modifica a estrutura entre as entidades que participam do processo de aleatorização. Com isso, a relação entre os fatores no experimento difere da relação dos mesmos na população. Então, se uma resposta Y_{ik} é determinada pelo tratamento **i** e a unidade **k**, a qual o tratamento **i** é aplicado, tem-se que na amostra a unidade **k** só pode ser usada uma única vez. Uma amostra obtida da aplicação aleatória de tratamentos às unidades experimentais (u.e.), com **r** aplicações de cada tratamento (replicações) é efetuada. Esse esquema recebe o nome de amostra fracionada simples com representação estrutural indicada na figura abaixo, onde T e P são respectivamente os tratamentos e às unidades.

Entidade P

Entidade T							
	0			0			
					0		0
		0	0				

A formalização do que foi visto acima, no caso de duas dimensões, é feita agora. Considere AP itens com valores P_{iu}, onde i=1, . . . , A; u=1, . . . , P e

$$\sum_{i=1}^{A} P_{iu} = \sum_{u=1}^{P} P_{iu} = 0. \tag{3.1}$$

Particionam-se os P_{iu}'s pelo índice **i** e selecionam-se aleatoriamente **a** conjuntos. Dentro de cada conjunto escolhido, realiza-se outra seleção aleatória de tamanho **r**, com a restrição de que nenhum P_{iu} possa ter o mesmo índice **u**. Usando as variáveis aleatórias, descritas abaixo, consegue-se reproduzir o esquema de amostragem, citado acima.

Sejam δ_i^{i*} e ρ_u^{i*f} com i*=1, . . ., a ; i=1, . . ., A; f=1, . . ., r e u=1, . . ., P variáveis aleatórias tais que:

$$\begin{cases} \delta_i^{i*} = 1, \textit{ se a i*-ésima escolha seleciona todos os } P_{iu}\textit{'s com subscrito i;} \\ \delta_i^{i*} = 0 \end{cases}$$

$$\begin{cases} \rho_u^{i*f} = 1, \textit{ se a f-ésima seleção dentro do i*-ésimo grupo corresponde ao } P_{iu} \textit{ tendo subscrito u;} \\ \rho_u^{i*f} = 0, \textit{ c.c.} \end{cases} \tag{3.2}$$

As propriedades das variáveis aleatórias, que serão úteis, estão descritas abaixo:

$$E\left(\delta_i^{i*}\right) = E\left(\delta_i^{i*}\right)^2 = \frac{1}{A},\ \forall\ i\ e\ i*;$$

$$E\left(\delta_i^{i*}\,\delta_{i'}^{i*'}\right) = \frac{1}{A(A-1)},\forall\ i\neq i'\ e\ i*\neq i*';$$

$$E\left(\rho_u^{i*f}\right) = \frac{1}{P}\quad e\quad E\left(\rho_u^{i*f}\,\rho_{u'}^{i*f'}\right) = \frac{1}{P(P-1)},\ \forall\ i*,\ f\neq f'\ e\ u\neq u'; \tag{3.3}$$

$$E\left(\rho_u^{i*f}\,\rho_{u'}^{i*'f'}\right) = \frac{1}{P(P-1)},\ \forall\ i*\neq i*'\ u\neq u'\ e\ \forall\ \textit{valores de}\ f\ e\ f'.$$

A conexão da observação amostral com a populacional é da forma:

$$y_{i*f} = \sum_{i,u} \delta_i^{i*}\,\rho_u^{i*f}\,P_{iu}. \tag{3.4}$$

Para o cálculo da $E(y^2)$ precisa-se dos seguintes resultados:

$$E\left(y_{i*f}^2\right) = E\left(\sum_{i,u}\delta_i^{i*}\,\rho_u^{i*f}P_{iu}\right)^2 = E\left(\sum_{i,u}\delta_i^{i*}\,\rho_u^{i*f}P_{iu}^2\right) + E\left(\sum_{\substack{i\\ u\neq u'}}\delta_i^{i*}\,\rho_u^{i*f}\,\rho_{u'}^{i*f}P_{iu}P_{iu'}\right) +$$

$$+ E\left(\sum_{\substack{u\\ i\neq i'}}\delta_i^{i*}\,\delta_{i'}^{i*}\,\rho_u^{i*f}P_{iu}P_{i'u}\right) + E\left(\sum_{\substack{i\neq i'\\ u\neq u'}}\delta_i^{i*}\,\delta_{i'}^{i*}\,\rho_u^{i*f}\,\rho_{u'}^{i*f}P_{iu}P_{i'u'}\right) = \tag{3.5}$$

$$= \sum_{i,u}P_{iu}^2\,E\left(\delta_i^{i*}\,\rho_u^{i*f}\right) = \frac{1}{AP}\sum_{i,u}P_{iu}^2 = \frac{\sigma_{AP}^2}{AP}(A-1)(P-1),$$

onde $\sigma_{AP}^2 = \dfrac{\sum_{i,u}P_{iu}^2}{(A-1)(P-1)}$.

$$E(y_{i\bullet f}\,y_{i\bullet f'}) = E\left[\left(\sum_{i,u}\delta_i^{i\bullet}\,\rho_u^{i\bullet f}P_{iu}\right)\left(\sum_{i,u}\delta_i^{i\bullet}\,\rho_u^{i\bullet f'}P_{iu}\right)\right] = \sum_{\substack{i\\ u\neq u'}} P_{iu}P_{iu'}E\left(\delta_i^{i\bullet}\,\rho_u^{i\bullet f}\,\rho_{u'}^{i\bullet f'}\right) =$$

$$= \frac{1}{AP(P-1)}\sum_{\substack{i\\ u\neq u'}} P_{iu}P_{iu'} = -\frac{1}{AP(P-1)}\sum_{iu}P_{iu}^2 = -\frac{\sigma_{AP}^2}{AP}(A-1). \tag{3.6}$$

$$E(y_{i\bullet f}\,y_{i'\bullet f}) = E\left[\left(\sum_{i,u}\delta_i^{i\bullet}\,\rho_u^{i\bullet f}P_{iu}\right)\left(\sum_{i,u}\delta_i^{i'\bullet}\,\rho_u^{i'\bullet f}P_{iu}\right)\right] = \sum_{\substack{i\neq i'\\ u\neq u'}} P_{iu}P_{i'u'}E\left(\delta_i^{i\bullet}\,\delta_{i'}^{i'\bullet}\,\rho_u^{i\bullet f}\,\rho_{u'}^{i'\bullet f}\right) =$$

$$= \frac{1}{A(A-1)P(P-1)}\sum_{\substack{i\neq i'\\ u\neq u'}} P_{iu}P_{i'u'} = \frac{1}{A(A-1)P(P-1)}\sum_{iu}P_{iu}^2 = \frac{\sigma_{AP}^2}{AP}. \tag{3.7}$$

$$E(y_{i\bullet f}\,y_{i'\bullet f'}) = E\left[\left(\sum_{i,u}\delta_i^{i\bullet}\,\rho_u^{i\bullet f}P_{iu}\right)\left(\sum_{i,u}\delta_i^{i'\bullet}\,\rho_u^{i'\bullet f'}P_{iu}\right)\right] = \sum_{\substack{i\neq i'\\ u\neq u'}} P_{iu}P_{i'u'}E\left(\delta_i^{i\bullet}\,\delta_{i'}^{i'\bullet}\,\rho_u^{i\bullet f}\,\rho_{u'}^{i'\bullet f'}\right) =$$

$$= \frac{1}{A(A-1)P(P-1)}\sum_{\substack{i\neq i'\\ u\neq u'}} P_{iu}P_{i'u'} = \frac{1}{A(A-1)P(P-1)}\sum_{iu}P_{iu}^2 = \frac{\sigma_{AP}^2}{AP}. \tag{3.8}$$

Usando (3.5), (3.6), (3.7) e (3.8), a $E(y^2)$ fica da forma:

$$E\left[\frac{1}{ar}\sum_{i\bullet f}y_{i\bullet f}\right]^2 = \frac{1}{(ar)^2}E\left(\sum_{i\bullet f}y_{i\bullet f}\right)^2 =$$

$$= \frac{1}{(ar)^2}E\left(\sum_{i\bullet f}y_{i\bullet f}^2 + \sum_{\substack{i\bullet\\ f\neq f'}}y_{i\bullet f}\,y_{i\bullet f'} + \sum_{\substack{f\\ i\bullet\neq i'\bullet}}y_{i\bullet f}\,y_{i'\bullet f} + \sum_{\substack{i\bullet\neq i'\bullet\\ f\neq f'}}y_{i\bullet f}\,y_{i'\bullet f'}\right) = \tag{3.9}$$

$$= \frac{1}{(ar)^2}\left[ar\frac{\sigma^2_{AP}}{AP}(A-1)(P-1) - ar(r-1)\frac{\sigma^2_{AP}}{AP}(A-1) + \right.$$

$$\left. + ar(a-1)\frac{\sigma^2_{AP}}{AP} + a(a-1)r(r-1)\frac{\sigma^2_{AP}}{AP}\right] =$$

$$= \frac{\sigma^2_{AP}}{AP(ar)}(AP - Ar + ar - P) = \frac{\sigma^2_{AP}}{AP(ar)}(P(A-1) - r(A-a)) =$$

$$= \frac{\sigma^2_{AP}}{ar}\left[\left(1 - \frac{1}{A}\right) - \frac{r}{P}\left(1 - \frac{a}{A}\right)\right].$$

Este resultado admite generalizações para dimensões maiores. Por exemplo: se os tratamentos A e B tiverem uma estrutura fatorial a generalização de (3.9) é:

$$E(y^2) = \left[\left(1 - \frac{1}{A}\right)\left(1 - \frac{1}{B}\right) - \frac{r}{P}\left(1 - \frac{a}{A}\right)\left(1 - \frac{b}{B}\right)\right]\frac{\sigma^2_{B(PT)}}{rab}.$$

Novamente, com no caso da amostragem pura, o fator de correção na amostragem fracionada envolve somente subscritos do CPMD do componente em questão. Por exemplo, no caso da estrutura aleatorizada em blocos, o termo relativo à interação do *plot* com o tratamento na expressão da $E(y^2)$ é

$$\left[\left(1 - \frac{1}{T}\right) - \frac{r}{P}\left(1 - \frac{t}{T}\right)\right]\frac{\sigma^2_{B(PT)}}{rbt}, \text{ onde } \sigma^2_{B(PT)} = \frac{\sum_{i,j,k}(PT)^2_{i(jk)}}{B(P-1)(T-1)}.$$

A seguir, será dado um exemplo para ilustrar os conceitos vistos até aqui. Já é de conhecimento, que a estrutura populacional aleatorizada em blocos é (B : P)(T) ou (i : j)(k) e a identidade populacional é $Y_{i(jk)} = Y + (Y_{(i)} - Y) + (Y_{(k)} - Y) + (Y_{(ik)} - Y_{(i)} - Y_{(k)} + Y) + (Y_{i(j)} - Y_{(i)}) + (Y_{i(jk)} - Y_{i(j)} - Y_{(ik)} + Y_{(i)}) = \mu + B_{(i)} + T_{(k)} + (BT)_{(ik)} + B(P)_{i(j)} + (TP)_{i(jk)}$.

O procedimento amostral consiste da escolha aleatória de **b** blocos dos B, e em cada um dos escolhidos, **rt** u.e. são selecionadas das P existentes e, por fim, **t** tratamentos são escolhidos de forma aleatória dos T tratamentos na população. Dentro de cada um dos blocos sorteados aplica-se aleatoriamente **r** vezes cada um dos **t** tratamentos às u.e. O valor observado da f-ésima réplica sorteada do k*-ésimo tratamento selecionado dentro do i*-ésimo bloco escolhido é denotado por $y_{i^* k^* f}$. A estrutura amostral é i*k*(f) e a identidade é

$$y_{i^* k^*(f)} = y + (y_{(i^*)} - y) + (y_{(i^* k^*)} - y) + (y_{(i^* k^*)} - y_{(i^*)} - y_{(k^*)} + y) + (y_{i^* k^*(f)} - y_{(i^* k^*)}).$$

As variáveis indicadoras que vão relacionar as observações amostrais com as populacionais são:

$$\begin{cases} \delta_i^{i*} = 1, \textit{ se o i*-ésimo bloco selecionado for o i-ésimo bloco populacional}; \\ \delta_i^{i*} = 0, \textit{ c.c.} \end{cases}$$

$$\begin{cases} u_{i*j}^{i*j*} = 1, \textit{ se o j*-ésimo plot selecionado do i*-ésimo bloco escolhido for o j-ésimo plot populacional do mesmo bloco}; \\ u_{i*j}^{i*j*} = 0, \textit{ c.c.} \end{cases}$$

$$\begin{cases} \beta_k^{k*} = 1, \textit{ se o k*-ésimo tratamento selecionado for o k-ésimo tratamento populacional}; \\ \beta_k^{k*} = 0, \textit{ c.c} \end{cases} \qquad (3.10)$$

$$\begin{cases} \lambda_{i*j*}^{i*k*f} = 1, \textit{ se a f-ésima réplica do k*-ésimo tratamento selecionado do i*-ésimo bloco escolhido for o j*-ésimo plot sorteado do mesmo bloco}; \\ \lambda_{i*j*}^{i*k*f} = 0, \textit{ c.c.} \end{cases}$$

e

$$\begin{cases} \rho_{i_* j}^{i_* k_* f} = 1, \textit{ se a } f\textit{-ésima réplica do } k_*\textit{-ésimo tratamento selecionado do } i_*\textit{-ésimo bloco escolhido} \\ \qquad\qquad \textit{for o } j\textit{-ésimo plot populacional do mesmo bloco}; \\ \rho_{i_* j}^{i_* k_* f} = 0, \textit{ c.c.}, \end{cases}$$

onde $\rho_{i_* j}^{i_* k_* f} = \sum_{j_*=1}^{rt} \mu_{i_* j}^{i_* j_*} \lambda_{i_* j_*}^{i_* k_* f}$. As propriedades das variáveis aleatórias supra citadas não serão enunciadas, como nos casos anteriores, pois são obtidas de modo análogo e podem ser feitas pelo leitor sem muita dificuldade.

O modelo estatístico é:

$$\begin{aligned} y_{i_* k_*(f)} &= \sum_{i,j,j_*,k} \delta_i^{i_*} \beta_k^{k_*} \mu_{i_* j}^{i_* j_*} \lambda_{i_* j_*}^{i_* k_* f} Y_{i(jk)} = \sum_{i,j,k} \delta_i^{i_*} \beta_k^{k_*} \rho_{i_* j}^{i_* k_* f} Y_{i(jk)} = \\ &= \mu + \sum_i \delta_i^{i_*} B_{(i)} + \sum_k \beta_k^{k_*} T_{(k)} + \sum_{ik} \delta_i^{i_*} \beta_k^{k_*} (BT)_{(ik)} + \\ &+ \sum_{ij} \delta_i^{i_*} \rho_{i_* j}^{i_* k_* f} B(P)_{i(j)} + \sum_{ijk} \delta_i^{i_*} \beta_k^{k_*} \rho_{i_* j}^{i_* k_* f} B(TP)_{i(jk)}. \end{aligned} \tag{3.11}$$

Usando os resultados do capítulo II sobre as variáveis aleatórias indicadoras e (3.9) encontra-se facilmente que:

$$\begin{aligned} E(y^2) = \mu^2 &+ \left(1 - \frac{b}{B}\right)\frac{\sigma^2_{(B)}}{b} + \left(1 - \frac{t}{T}\right)\frac{\sigma^2_{(T)}}{t} + \left(1 - \frac{b}{B}\right)\left(1 - \frac{t}{T}\right)\frac{\sigma^2_{(BT)}}{b} + \\ &+ \left(1 - \frac{rt}{P}\right)\frac{\sigma^2_{B(P)}}{rtb} + \left[1 - \frac{1}{T} - \frac{r}{P}\left(1 - \frac{t}{T}\right)\right]\frac{\sigma^2_{B(PT)}}{RBT}. \end{aligned} \tag{3.12}$$

Reagrupando os termos em (3.12) obtém-se:

$$
\begin{aligned}
E(y^2) = & \left(\mu^2 - \frac{1}{B}\sigma^2_{(B)} - \frac{1}{T}\sigma^2_{(T)} + \frac{1}{BT}\sigma^2_{(BT)} \right) + \\
& + \frac{1}{b}\left(\sigma^2_{(B)} - \frac{1}{T}\sigma^2_{(BT)} - \frac{1}{P}\sigma^2_{B(P)} + \frac{1}{PT}\sigma^2_{B(PT)} \right) + \\
& + \frac{1}{t}\left(\sigma^2_{(T)} - \frac{1}{B}\sigma^2_{(BT)} \right) + \frac{1}{bt}\left(\sigma^2_{(BT)} - \frac{1}{P}\sigma^2_{B(PT)} \right) + \\
& + \frac{1}{rtb}\left(\sigma^2_{B(P)} - \frac{1}{T}\sigma^2_{B(PT)} \right) + \frac{1}{rtb}\sigma^2_{B(PT)},
\end{aligned}
\tag{3.13}
$$

ou equivalentemente, em função dos Σ's:

$$
E(y^2) = \Sigma_{\varnothing} + \frac{1}{b}\Sigma_{(B)} + \frac{1}{t}\Sigma_{(T)} + \frac{1}{bt}\Sigma_{(BT)} + \frac{1}{(rtb)}\left(\Sigma_{B(P)} + \Sigma_{B(PT)}\right). \tag{3.14}
$$

Vale a pena observar, que a expansão da $E(y^2)$ em função dos Σ's mantém a mesma regra que a enunciada no final do capítulo 1. Da mesma forma acontece com o coeficiente de cada Σ, ou seja, ele é o inverso do número de diferentes valores do componente, em questão, que entram na formação da média amostral y.

Para se obter o valor esperado de uma média parcial a partir de (3.14), supondo esta expansão válida, é bastante simples. Considerando a estrutura acima, se y_{i*} denota a média observada do i*-ésimo bloco selecionado, então a $E(y_{i*}^2)$ é obtida de (3.14) substituindo o valor b=1, isto é:

$$
E(y_{i*}^2) = \Sigma_{\varnothing} + \Sigma_{(B)} + \frac{\Sigma_{(T)}}{t} + \frac{\Sigma_{B(T)}}{t} + \frac{\Sigma_{B(P)}}{rt} + \frac{\Sigma_{B(PT)}}{rt}. \tag{3.15}
$$

Para ressaltar a simplicidade da forma das expectâncias calculadas, far-se-á analogia com o modelo em que os componentes são aleatórios e não correlacionados. Por exemplo: considere o modelo $Y_{ij} = \mu + a_i + b_j + (ab)_{ij}$, i = 1, . . ., a e j = 1, . . .,b; onde μ é uma constante, e os a_i's são aleatórios com tamanho da amostra igual a a, os b_j's com

tamanho de amostra igual a **b** e os $(ab)_{ij}$ com tamanho de amostra **ab**, com todas as amostras independentes e de populações infinitas com média zero e variância finita.

Obs: $\mu^2 = \sigma^2_{\mu} = \sigma^2_{\varnothing}$, $Var(a_i) = \sigma^2_a$, $Var(b_j) = \sigma^2_b$ e $Var[(ab)_{ij})] = \sigma^2_{ab}$.

$$E\left(Y^2_{ij}\right) = \sigma^2_{\varnothing} + \sigma^2_a + \sigma^2_b + \sigma^2_{ab};$$

$$E\left(Y^2_i\right) = E\left(\frac{1}{b}\sum_{j=1}^{b} Y_{ij}\right)^2 = \sigma^2_{\varnothing} + \sigma^2_a + \frac{\sigma^2_b}{b} + \frac{\sigma^2_{ab}}{b};$$

$$E\left(Y^2_j\right) = E\left(\frac{1}{a}\sum_{i=1}^{a} Y_{ij}\right)^2 = \sigma^2_{\varnothing} + \frac{\sigma^2_a}{a} + \sigma^2_b + \frac{\sigma^2_{ab}}{a};$$

$$E\left(Y^2\right) = E\left(\frac{1}{ab}\sum_{ij} Y_{ij}\right)^2 = \sigma^2_{\varnothing} + \frac{\sigma^2_a}{a} + \frac{\sigma^2_b}{b} + \frac{\sigma^2_{ab}}{ab}.$$

Note, que cada σ^2_i, i=∅, a, b e ab; é dividido, também, pelo nº de diferentes i's que entram na média amostral em questão.

O próximo passo é a formalização desses resultados. Sejam $\Sigma_{X(R)}$ e $N_{X(R)}$ definidos como em (2.54). Então a $E(y^2)$ admite a expansão padrão dos Σ's se

$$E(y^2) = \sum_{X,R} \frac{\Sigma_{X(R)}}{N_{X(R)}}. \qquad (3.16)$$

Seja S_{1*}, o conjunto de subscritos de uma particular média parcial e $N_{X(R),S_{1*}}$, o número de diferentes valores do componente especificado por X(R) que entram na formação da média de $y_{S_{1*}}$. Se $E(y^2)$satisfaz (3.16), a seguinte expectância é válida:

$$E\left(y^2_{S_{1*}}\right) = \sum_{X,R} \frac{\Sigma_{X(R)}}{N_{X(R),S_{1*}}}. \qquad (3.17)$$

Uma condição suficiente para (3.16) ser verdadeira é que toda a amostragem balanceada empregada seja do tipo pura e/ou fracionada simples. Existe, também, outra condição suficiente: usando a expansão da $E(y^2)$ em função dos componentes de variação σ^2, que será exemplificada a seguir com a estrutura aleatorizada em blocos. O termo envolvendo $\sigma^2_{B(PT)}$ na expansão da $E(y^2)$ em função dos σ^2 considerando a estrutura aleatorizada em blocos é:

$\left[\left(1 - \frac{1}{T}\right) \quad \frac{r}{P}\left(1 - \frac{t}{T}\right)\right]\frac{\sigma^2_{B(PT)}}{rbt}$, que pode também ser escrita como:

$$\frac{1}{rbt}\frac{\sigma^2_{B(PT)}}{1} - \frac{1}{rbt}\frac{\sigma^2_{B(PT)}}{T} - \frac{1}{bt}\frac{\sigma^2_{B(PT)}}{P} + \frac{1}{b}\frac{\sigma^2_{B(PT)}}{PT}.$$

Se se denotar por Z, o conjunto de letras no denominador, cujo numerador é $\sigma^2_{B(PT)}$, e por $N_{B(PT-Z)}$, o número de valores do componente do tipo B(PT-Z) que entram na formação da média amostral experimental que se está calculando a expectância, então os termos da expectância, acima, são da forma:

$$\frac{(-1)^q}{N_{B(PT-Z)}}\frac{\sigma^2_{B(PT)}}{Z}, \tag{3.18}$$

onde q é o número de letras em Z.

3.2 - Os Cap Sigmas Amostrais

O propósito desta seção é introduzir notações e conceitos suficientes para a montagem da ANOVA amostral em função das expectâncias dos quadrados. Os novos Σ's terão subscritos com letras minúsculas, que corresponderão exatamente aos subscritos das

médias amostrais parciais.

No caso visto anteriormente, estrutura aleatorizada em blocos, ambos os cap sigmas, $\Sigma_{B(P)}$ e $\Sigma_{B(PT)}$, possuem o mesmo coeficiente, $\frac{1}{btp}$. Isto sugere, então, que a função $\Sigma_{B(P)} + \Sigma_{B(PT)}$ represente uma única entidade na amostra. Se se fizer:

$$\Sigma_{\varnothing}=\Sigma_{\varnothing}\ ,\ \Sigma_{(b)}=\Sigma_{(B)}\ ,\ \Sigma_{(t)=}\Sigma_{(T)}\ ,\ \Sigma_{(b\,t)}=\Sigma_{(BT)}\ ,\ \Sigma_{b(t\,p)}=\Sigma_{B(P)} + \Sigma_{B(PT)},$$

cada conjunto de índices num Σ amostral corresponderá a uma média amostral parcial. A expressão (3.14) é reescrita para a forma:

$$E(y^2) = \Sigma_{\varnothing} + \frac{1}{b}\Sigma_{(b)} + \frac{1}{t}\Sigma_{(t)} + \frac{1}{bt}\Sigma_{(bt)} + \frac{1}{(rtb)}\Sigma_{b(pt)}, \qquad (3.19)$$

onde **p** representa o número de u.e. em cada bloco aonde são aplicados os tratamentos. O exemplo acima serve de ilustração para a definição abaixo dos Σ's amostrais.

Definição 3.2.1: Se $y_{h(k)}$ é uma média amostral qualquer, então, o $\Sigma_{h(k)}$ amostral é definido como a soma de todos os Σ's populacionais admissíveis, cujos subscritos contenham como um subconjunto os índices populacionais do conjunto k, e que os mesmos Σ's populacionais não tenham índices excedentes com relação ao conjunto h+k.

Obs.: Ambos os teoremas, descritos abaixo, podem ser encontrados em **Zyskind-(1958)**, com suas respectivas demonstrações.

Teorema 3.1: Para qualquer experimento balanceado e completo a expectância do quadrado da média amostral é:

$$E(y^2) = \sum_{h,k} \frac{1}{n_{h+k}} \Sigma_{h(k)},$$

onde $\frac{1}{n_{h+k}}$ é o produto das dimensões amostrais dos índices em (h+k).

Teorema 3.2: Considere um experimento balanceado e completo, e denote por **s** o conjunto de todos os índices amostrais. Se x_k é uma média amostral parcial admissível, então a expectância do quadrado em uma linha da ANOVA amostral é:

$$\sum_{h \subset s} n_{s-h} \Sigma_h,$$

onde k⊂h. Este teorema é o mais importante enunciado neste capítulo, pois permite montar a ANOVA amostral sem o cálculo de todas as expectâncias de quadrados.

Antes de se prosseguir com a aplicação dos resultados dos teoremas em alguns experimentos padrões será apresentado, na seção seguinte, uma representação gráfica para as estruturas completas.

3.3 - Representação Gráfica de uma Estrutura de Resposta Completa

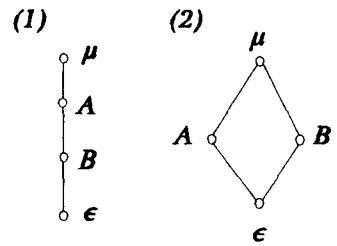
Definição 3.3.1: A representação gráfica de cada fator de classificação, incluindo μ e ϵ, recebe a denominação de diagrama de estrutura[1]. Os o representam os fatores de classificação. Se houver uma relação de hierarquia entre dois fatores, por exemplo A embute B, então haverá uma linha descendente de A para B. Se a relação for de cruzamento, então um fator virá ao lado do outro. Se por exemplo um fator D estiver

[1]Este método foi desenvolvido por Throckmorton,T.N. - Structures of classification data . Tese de Ph.D. não publicada. Biblioteca de Iowa State University of Science and Technology, Ames, Iowa, 1961.

embutido em um cruzamento, haverá uma linha descendente para cada fator que embute o fator D. A seguir, exemplos serão dados para ilustrar a definição acima.

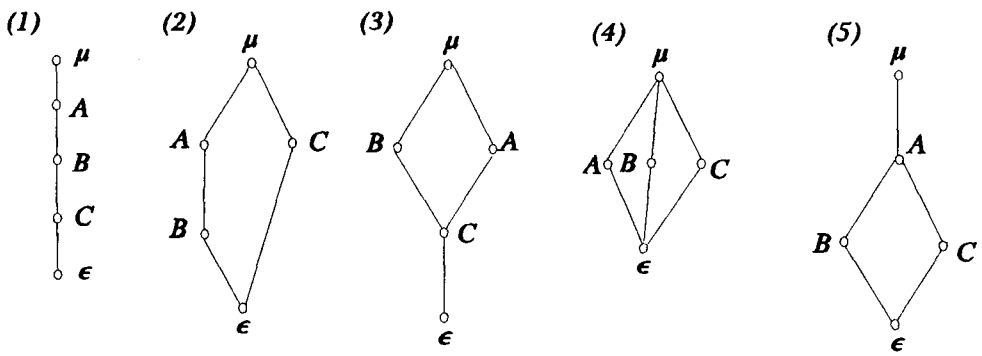
Exemplo: Como a designação de letras aos fatores de classificação é arbitrária, não se fará distinção entre A embutir B ou B embutir A na representação gráfica.

1) Dois fatores de classificação, A e B.



Na primeira estrutura A embute B; já na segunda, A e B estão cruzados.

2) A relação entre três fatores é uma situação mais rica, abaixo estão as possíveis representações dos diagramas.



Na situação (1), A embute B que embute C; na (2), A e C estão cruzados, e B está

embutido em A; na (3) A e B estão cruzados, e C está embutido neste cruzamento; na (4) A, B e C estão cruzados e na última, (5) A embute o cruzamento de B com C.

Vale a pena lembrar que não existe um método de cálculo para predizer o número de relações para n fatores.

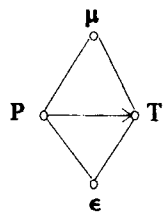
Obs.:

1ª) A média de todas as respostas μ poder ser considerada como um fator de apenas um nível que ocorre com cada resposta. Assim, cada nível de cada fator tem que ocorrer junto com este nível de μ. Deste modo, todos os fatores estão embutidos em μ.

2ª) Se se pegar todas as respostas, provenientes de várias medições, com os mesmos tratamentos e se tirar a média , obter-se-á o valor "verdadeiro" deste conjunto de respostas. A diferença de cada resposta, do conjunto acima, para este valor verdadeiro é, então, decorrente do efeito do fator ϵ. Isto significa, que ϵ é um erro técnico ou um erro de medida.

3ª) Quando uma seta partir de uma entidade para outra indicará que a origem da seta foi aleatoriamente embutida na entidade de destino no experimento em questão.

4ª) O diagrama de uma estrutura experimental representará tanto a estrutura populacional quanto a observacional, e exibirá a atribuição aleatória que ocorre no experimento, como demonstrado no exemplo abaixo:

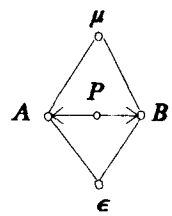


No presente caso, P e T estão cruzados na estrutura populacional, enquanto que na estrutura observacional P está embutido em T.

3.4 - Aplicações

3.4.1 - Esquema com dois Fatores Completamente Aleatorizados

Será estudada a situação em que os tratamentos são caracterizados por dois fatores, A e B, e as u.e. por P. O experimento consiste da escolha aleatória de **a** dentre os A níveis de A, **b** dentre os B níveis de B e **pab** dentre os P níveis de P; e então aleatoriamente atribuir às **pab** u.e. selecionadas aos tratamentos sorteados, de modo que, existam **p** réplicas para cada tratamento. O diagrama experimental é:



Médias	*Populacionais*	$Y, Y_A, Y_B, Y_{AB}, Y_P, Y_{AP}, Y_{BP}, Y_{ABP}$
Admissíveis	*Amostrais*	$y, y_a, y_b, y_{ab}, y_{ab(p)}$

Como o experimento é balanceado e completo, tem-se que:

$$E(y^2) = \mu^2 + \left(1 - \frac{a}{A}\right)\frac{\sigma_A^2}{a} + \left(1 - \frac{b}{B}\right)\frac{\sigma_B^2}{b} + \left(1 - \frac{a}{A} - \frac{b}{B} + \frac{ab}{AB}\right)\frac{\sigma_{AB}^2}{ab} +$$

$$+ \left(1 - \frac{pab}{P}\right)\frac{\sigma_P^2}{pab} + \left(1 - \frac{1}{A} - \frac{pb}{P} + \frac{pab}{AP}\right)\frac{\sigma_{AP}^2}{pab} +$$

$$+ \left(1 - \frac{1}{B} - \frac{pa}{P} + \frac{pab}{PB}\right)\frac{\sigma_{BP}^2}{pab} +$$

$$+\left(1-\frac{1}{A}-\frac{1}{B}+\frac{1}{AB}-\frac{p}{P}+\frac{pa}{AP}+\frac{pb}{BP}-\frac{pab}{ABP}\right)\frac{\sigma^2_{ABP}}{pab}=$$

$$=\Sigma_{\varnothing}+\frac{1}{a}\Sigma_A+\frac{1}{b}\Sigma_B+\frac{1}{ab}\Sigma_{AB}+\frac{1}{pab}(\Sigma_P+\Sigma_{AP}+\Sigma_{BP}+\Sigma_{ABP}),$$

onde

$$\Sigma_{\varnothing}=\mu^2-\frac{1}{A}\sigma^2_A-\frac{1}{B}\sigma^2_B+\frac{1}{AB}\sigma^2_{AB}-\frac{1}{P}\sigma^2_P+\frac{1}{AP}\sigma^2_{AP}+\frac{1}{BP}\sigma^2_{BP}-\frac{1}{ABP}\sigma^2_{ABP}$$

$$\Sigma_A=\sigma^2_A-\frac{1}{B}\sigma^2_{AB}-\frac{1}{P}\sigma^2_{AP}+\frac{1}{BP}\sigma^2_{ABP}$$

$$\Sigma_B=\sigma^2_B-\frac{1}{A}\sigma^2_{AB}-\frac{1}{P}\sigma^2_{BP}+\frac{1}{AP}\sigma^2_{ABP}$$

$$\Sigma_{AB}=\sigma^2_{AB}-\frac{1}{P}\sigma^2_{ABP}$$

$$\Sigma_P=\sigma^2_P-\frac{1}{A}\sigma^2_{AP}-\frac{1}{B}\sigma^2_{BP}+\frac{1}{AB}\sigma^2_{ABP}$$

$$\Sigma_{AP}=\sigma^2_{AP}-\frac{1}{B}\sigma^2_{ABP}$$

$$\Sigma_{BP}=\sigma^2_{BP}-\frac{1}{A}\sigma^2_{ABP}$$

$$\Sigma_{ABP}=\sigma^2_{ABP}.$$

Pode-se verificar que todos os cap sigmas populacionais envolvendo P têm o mesmo coeficiente na $E(y^2)$, e pela definição 3.2.1,

$$\Sigma_{ab(p)}=\Sigma_P+\Sigma_{AP}+\Sigma_{BP}+\Sigma_{ABP}.$$

Os demais Σ's amostrais têm os mesmos subscritos, com letras minúsculas, dos Σ's populacionais. Assim, de acordo com teorema 3.1, a $E(y^2)$ é

$$E(y^2) = \Sigma_{\varnothing} + \frac{1}{a}\Sigma_a + \frac{1}{b}\Sigma_b + \frac{1}{ab}\Sigma_{ab} + \frac{1}{pab}\Sigma_{ab(p)}.$$

Substituindo por 1 os subscritos apropriados na $E(y^2)$, obtêm-se as expectâncias dos quadrados das médias parciais, como exibido abaixo.

$$E\left(y_a^2\right) = \Sigma_{\varnothing} + \Sigma_a + \frac{1}{b}\Sigma_b + \frac{1}{b}\Sigma_{ab} + \frac{1}{pb}\Sigma_{ab(p)}$$

$$E\left(y_b^2\right) = \Sigma_{\varnothing} + \frac{1}{a}\Sigma_a + \Sigma_b + \frac{1}{a}\Sigma_{ab} + \frac{1}{pa}\Sigma_{ab(p)}$$

$$E\left(y_{ab}^2\right) = \Sigma_{\varnothing} + \Sigma_a + \Sigma_b + \Sigma_{ab} + \frac{1}{p}\Sigma_{ab(p)}$$

$$E\left(y_{ab(p)}^2\right) = \Sigma_{\varnothing} + \Sigma_a + \Sigma_b + \Sigma_{ab} + \Sigma_{ab(p)}$$

Na ANOVA amostral, cada linha é representada por uma fonte de variação. O cálculo de suas expectâncias de quadrados serão exemplificados através de y_a. A soma de quadrados é dada por

$$\sum_{abp}(y_a - y)^2 = pb\sum_{abp}(y_a^2 - y^2) \diamond$$

$$\diamond\ pab\left(E\left(y_a^2\right) - E\left(y^2\right)\right) = pb(a-1)\Sigma_a + p(a-1)\Sigma_{ab} + (a-1)\Sigma_{ab(p)}.$$

Dividindo a expectância da soma dos quadrados pelos seus respectivos graus de liberdade, obtém-se a expectância do quadrado médio, ou seja,

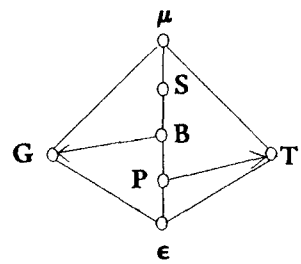
$$pb\Sigma_a + p\Sigma_{ab} + \Sigma_{ab(p)}.$$

Da mesma forma é feito para as demais fontes de variação, obtendo, então a ANOVA amostral abaixo:

Fonte	*G.L.*	*E.Q.M.*
y	1	$\Sigma_{ab(p)}+p\Sigma_{ab}+pa\Sigma_{b}+pb\Sigma_{pab}+pab\Sigma_{\varnothing}$
y_a	a-1	$\Sigma_{ab(p)}+p\Sigma_{ab}+pb\Sigma_{a}$
y_b	b-1	$\Sigma_{ab(p)}+p\Sigma_{ab}+pa\Sigma_{b}$
y_{ab}	(a-1)(b-1)	$\Sigma_{ab(p)}+p\Sigma_{ab}$
$y_{ab(p)}$	ab(p-1)	$\Sigma_{ab(p)}$

3.4.2 - Esquema com Split-plot Generalizado

Neste experimento as u.e. estão hierarquicamente classificadas em S fontes, B blocos dentro de fontes e P unidades dentro de blocos. Os tratamentos, T e G, têm uma estrutura fatorial. O procedimento experimental consiste em alocar **g** níveis escolhidos aleatoriamente, dos G existentes, aos **bg** blocos sorteados aleatoriamente, dos B. Isso dentro de cada uma das **s** fontes selecionadas de forma aleatória, com a restrição de que cada nível de G escolhido tenha exatamente **b** blocos; e por fim, atribuir aleatoriamente os **t** níveis de T aos **pt** plots dentro de cada bloco, de modo que, cada nível de **t** receba **p** plots. O diagrama experimental é



Médias	*Populacionais*	$Y, Y_S, Y_{S(B)}, Y_{SB(P)}, Y_G, Y_{SG}, Y_{S(GB)}, Y_{SB(GP)},$ $Y_T, Y_{ST}, Y_{S(BT)}, Y_{GT}, Y_{SGT}, Y_{S(BGT)}, Y_{SB(GTP)}$
Admissíveis	*Amostrais*	$y, y_s, y_g, y_{sg}, y_{sg(b)}, y_t,$ $y_{st}, y_{gt}, y_{sg(bt)}, y_{sbgt(p)}$

Da mesma forma que na seção 3.4.1 este experimento é balanceado e completo, implicando que a $E(y^2)$ é:

$$E(y^2) = \mu^2 + \left(1 - \frac{s}{S}\right)\frac{\sigma_S^2}{s} + \left(1 - \frac{g}{G}\right)\frac{\sigma_G^2}{g} + \left(1 - \frac{t}{T}\right)\frac{\sigma_T^2}{t} + \left(1 - \frac{s}{S}\right)\left(1 - \frac{g}{G}\right)\frac{\sigma_{SG}^2}{sg} +$$

$$+\left(1 - \frac{s}{S}\right)\left(1 - \frac{t}{T}\right)\frac{\sigma_{ST}^2}{st} + \left(1 - \frac{g}{G}\right)\left(1 - \frac{t}{T}\right)\frac{\sigma_{GT}^2}{gt} + \left(1 - \frac{s}{S}\right)\left(1 - \frac{g}{G}\right)\left(1 - \frac{t}{T}\right)\frac{\sigma_{SG}^2}{sgt}$$

$$+\left(1 - \frac{bg}{B}\right)\frac{\sigma_{S(B)}^2}{sbg} + \left(1 - \frac{pt}{P}\right)\frac{\sigma_{SB(P)}^2}{sbgpt} + \left[1 - \frac{1}{G} - \frac{b}{B}\left(1 - \frac{g}{G}\right)\right]\frac{\sigma_{S(BG)}^2}{sbg} +$$

$$+\left(1 - \frac{bg}{B}\right)\left(1 - \frac{t}{T}\right)\frac{\sigma_{S(BT)}^2}{sbgt} + \left(1 - \frac{1}{G}\right)\left(1 - \frac{pt}{P}\right)\frac{\sigma_{SB(GP)}^2}{sbgpt} + \left[1 - \frac{1}{T} - \frac{p}{P}\left(1 - \frac{t}{T}\right)\right]\frac{\sigma_{SB(PT)}^2}{sbgpt} +$$

$$+\left(1 - \frac{t}{T}\right)\left[1 - \frac{1}{G} - \frac{b}{B}\left(1 - \frac{g}{G}\right)\right]\frac{\sigma_{S(BGT)}^2}{sbgt} + \left(1 - \frac{1}{G}\right)\left[1 - \frac{1}{T} - \frac{p}{P}\left(1 - \frac{t}{T}\right)\right]\frac{\sigma_{SB(GPT)}^2}{sbgpt} =$$

$$= \Sigma_{\varnothing} + \frac{\Sigma_S}{s} + \frac{\Sigma_G}{g} + \frac{\Sigma_T}{t} + \frac{\Sigma_{SG}}{sg} + \frac{\Sigma_{ST}}{st} + \frac{\Sigma_{GT}}{gt} + \frac{\Sigma_{SGT}}{sgt} + \frac{1}{sbg}\left(\Sigma_{S(B)} + \Sigma_{S(BG)}\right) +$$

$$+ \frac{1}{sbgt}\left(\Sigma_{S(BT)} + \Sigma_{S(BGT)}\right) + \frac{1}{sbgt}\left(\Sigma_{SB(P)} + \Sigma_{SB(PG)} + \Sigma_{SB(PT)} + \Sigma_{SB(PGT)}\right),$$

onde

$$\Sigma_{e} = \mu^2 - \frac{\sigma_S^2}{S} - \frac{\sigma_G^2}{G} - \frac{\sigma_T^2}{T} + \frac{\sigma_{SG}^2}{SG} + \frac{\sigma_{ST}^2}{ST} + \frac{\sigma_{GT}^2}{GT} - \frac{\sigma_{STG}^2}{STG}$$

$$\Sigma_{S} = \sigma_S^2 - \frac{\sigma_{SG}^2}{G} - \frac{\sigma_{ST}^2}{T} + \frac{\sigma_{STG}^2}{GT} - \frac{\sigma_{S(B)}^2}{B} - \frac{\sigma_{S(BG)}^2}{BG} + \frac{\sigma_{S(BT)}^2}{BT} - \frac{\sigma_{S(BGT)}^2}{BGT}$$

$$\Sigma_{G} = \sigma_G^2 - \frac{\sigma_{SG}^2}{S} - \frac{\sigma_{GT}^2}{T} + \frac{\sigma_{SGT}^2}{ST}$$

$$\Sigma_{T} = \sigma_T^2 - \frac{\sigma_{ST}^2}{S} - \frac{\sigma_{GT}^2}{G} + \frac{\sigma_{SGT}^2}{SG}$$

$$\Sigma_{SG} = \sigma_{SG}^2 - \frac{\sigma_{SGT}^2}{T} - \frac{\sigma_{S(GB)}^2}{B} + \frac{\sigma_{S(BGT)}^2}{BT}$$

$$\Sigma_{ST} = \sigma_{ST}^2 - \frac{\sigma_{SGT}^2}{G} - \frac{\sigma_{S(BT)}^2}{B} + \frac{\sigma_{S(BGT)}^2}{BG}$$

$$\Sigma_{GT} = \sigma_{GT}^2 - \frac{\sigma_{SGT}^2}{S}$$

$$\Sigma_{SGT} = \sigma_{SGT}^2 - \frac{\sigma_{S(BGT)}^2}{B}$$

$$\Sigma_{S(B)} = \sigma_{S(B)}^2 - \frac{\sigma_{SB(P)}^2}{P} - \frac{\sigma_{S(BG)}^2}{G} - \frac{\sigma_{S(BT)}^2}{T} + \frac{\sigma_{SB(GP)}^2}{GP} + \frac{\sigma_{SB(PT)}^2}{PT} + \frac{\sigma_{S(BGT)}^2}{GT} - \frac{\sigma_{SB(GTP)}^2}{GTP}$$

$$\Sigma_{SB(P)} = \sigma_{SB(P)}^2 - \frac{\sigma_{SB(GP)}^2}{G} - \frac{\sigma_{SB(PT)}^2}{T} + \frac{\sigma_{SB(GTP)}^2}{GT}$$

$$\Sigma_{S(BG)} = \sigma_{S(BG)}^2 - \frac{\sigma_{SB(GP)}^2}{P} - \frac{\sigma_{S(BGT)}^2}{T} + \frac{\sigma_{SB(GPT)}^2}{PT}$$

$$\Sigma_{S(BT)} = \sigma_{S(BT)}^2 - \frac{\sigma_{SB(PT)}^2}{P} - \frac{\sigma_{S(BGT)}^2}{G} + \frac{\sigma_{SB(GPT)}^2}{GP}$$

$$\Sigma_{SB(GP)} = \sigma^2_{SB(GP)} - \frac{\sigma^2_{SB(GPT)}}{T}$$

$$\Sigma_{SB(PT)} = \sigma^2_{SB(PT)} - \frac{\sigma^2_{SB(GPT)}}{G}$$

$$\Sigma_{S(BGT)} = \sigma^2_{S(BGT)} - \frac{\sigma^2_{SB(GPT)}}{P}$$

$$\Sigma_{SB(GPT)} = \sigma^2_{SB(GPT)}.$$

Assim, de acordo com o teorema 3.1 tem-se:

$$E(y^2) = \Sigma_{\varnothing} + \frac{\Sigma_s}{s} + \frac{\Sigma_g}{g} + \frac{\Sigma_t}{t} + \frac{\Sigma_{sg}}{sg} + \frac{\Sigma_{st}}{st} + \frac{\Sigma_{gt}}{gt} + \frac{\Sigma_{sgt}}{sgt} + \frac{\Sigma_{sg(b)}}{sbg} + \frac{\Sigma_{sg(bt)}}{sbgt} + \frac{\Sigma_{sbgt(p)}}{sbgtp}.$$

E, por conseguinte, a ANOVA amostral:

Fonte	*G.L.*	*E.Q.M.*
y	1	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}+sbp\Sigma_{gt}+bpg\Sigma_{st}+pt\Sigma_{sg(b)}+bpt\Sigma_{sg}+$ $bpgt\Sigma_{s}+sbgp\Sigma_{t}+sbtp\Sigma_{g}+sbgpt\Sigma_{\varnothing}$
y_s	(s-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}+bpg\Sigma_{st}+pt\Sigma_{sg(b)}+bpt\Sigma_{sg}+bpgt\Sigma_{s}$
y_g	(g-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}+sbp\Sigma_{gt}+pt\Sigma_{sg(b)}+bpt\Sigma_{sg}+sbpt\Sigma_{g}$
y_{sg}	(s-1)(g-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+pt\Sigma_{sg(b)}+bpt\Sigma sg$
$y_{sg(b)}$	sg(b-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+pt\Sigma_{sg(b)}$
y_t	(t-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}+sbp\Sigma_{gt}+bpg\Sigma_{st}+sbpg\Sigma_{t}$
y_{st}	(s-1)(t-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}+bpg\Sigma_{st}$
y_{gt}	(g-1)(t-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}+sbp\Sigma_{gt}$
y_{sgt}	(s-1)(g-1)(t-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}+bp\Sigma_{sgt}$
$y_{sg(bt)}$	sg(b-1)(t-1)	$\Sigma_{sgbt(p)}+p\Sigma_{sg(bt)}$
$y_{sgbt(p)}$	sgbt(p-1)	$\Sigma_{sgbt(p)}$

Capítulo IV

O Poder dos Testes de Aleatorização

O frequente uso da distribuição F, no que concerne aos testes da tabela de ANOVA, requer originalmente a suposição de normalidade dos erros, provenientes do modelo matemático adequado à situação experimental. Considerações de experimentos aleatorizados, todavia, mostram que os erros experimentais seguem uma distribuição distinta daquela, ou seja, não normal.

Neste capítulo, ilustrar-se-á por meio de experimentos aleatorizados em blocos a diferença entre a distribuição de referência, gerada pelo processo de alocar aleatoriamente as u.e. aos tratamentos, e a distribuição F, baseada na suposição de normalidade dos erros, que nem sempre é uma hipótese sustentável. É importante lembrar que apesar das duas distribuições citadas acima não serem idênticas, à medida que o tamanho do experimento aumenta a qualidade da aproximação da distribuição F para a de referência melhora incrivelmente. O outro enfoque, é o comportamento do poder do teste de aleatorização, para tanto, gráficos serão feitos considerando situações distintas.

4.1 - Experimentos Aleatorizados em Blocos sob Aditividade de Blocos com Tratamentos

Um exemplo simples a ser considerado é o caso de experimentos aleatorizados em blocos sob aditividade de blocos com tratamentos. O procedimento amostral é representado, aqui, pelo uso da variável aleatória δ_{ij}^{k}, onde os δ_{ij}^{k}'s tomam valor 1 se o tratamento **k** está no plot **j** do bloco **i** na realização do experimento, e toma o valor 0 c.c.

Então, com t plots em cada bloco e com t tratamentos tem-se as seguintes propriedades:

$$P(\delta_{ij}^{k} = 1) = 1 - P(\delta_{ij}^{k} = 0) = \frac{1}{t}, \ \forall\ i, j, k;$$

$$P(\delta_{ij}^{k}\ \delta_{ij'}^{k'} = 1) = \frac{1}{t(t-1)}, \ \forall\ k \neq k', j \neq j' \text{ e } i; \tag{4.1}$$

$$P(\delta_{ij}^{k}\ \delta_{i'j'}^{k'} = 1) = \frac{1}{t^2}, \ \forall\ i \neq i' \text{ e } \textit{quaisquer valores de } j, j', k \text{ e } k'.$$

A conexão do valor observado com o elemento da população conceitual de respostas é da forma:

$$y_{ik} = \sum_{j=1}^{t} \delta_{ij}^{k} Y_{i(jk)}. \tag{4.2}$$

Mas, sob as condições de aditividade, $Y_{i(jk)}$ pode ser decomposto em

$$Y_{i(jk)} = Y + (Y_{(t)} - Y) + (Y_{i(jk)} - Y_{i(j)}) + (Y_{i(j)} - Y_{(i)}) = \mu + b_i + t_k + \epsilon_{ij}, \tag{4.3}$$

com $\sum_i b_i = \sum_k t_k = 0$. Substituindo (4.3) em (4.2) obtém-se:

$$y_{ik} = \mu + b_i + t_k + \sum_{j=1}^{t} \delta_{ij}^{k} \epsilon_{ij} = \mu + b_i + t_k + e_{ik}. \tag{4.4}$$

Qual a razão para que o modelo (4.4) seja chamado de aditivo? A resposta é muito simples. Por exemplo, se um incremento de t_2 implica em um aumento de 5 unidades na resposta e a influência do bloco aumenta 4 unidades na resposta, o acréscimo na resposta devido a ambos seria de 5+4=9 unidades. Apesar da simplicidade do modelo aditivo prover uma boa aproximação em muitas situações, há circunstâncias em que isto

não acontece. Daí a necessidade de usar recursos como testes para não aditividade e a análise gráfica dos resíduos para constatar a boa ou má aproximação do modelo aditivo. Para mais detalhes ver [10, seção 8.2].

Inspeções nos e_{ik}'s mostram que o e_{ik} toma diferentes valores ϵ_{ij}, com igual probabilidade $\frac{1}{t}$. Isto acarreta uma distribuição para os e_{ik} diferente da normal.

Sabe-se que para o caso dos e_{ik}'s serem não correlacionados e normalmente distribuídos, a F com sua respectiva região de rejeição de tamanho α gozam de uma série de propriedades ótimas. Este fato, juntamente com a implementação desta estatística em todos os pacotes estatísticos, incentivou o seu grande uso, deixando o teste exato de aleatorização numa espécie de dormência e esquecimento.

4.2 - Teste de Aleatorização e Função Poder

Segue-se o seguinte procedimento para a construção de um teste baseado numa estatística quando a distribuição dos e_{ik}'s é induzida pelo processo aleatório: calcula-se, no experimento que foi efetivamente observado, a razão entre o quadrado médio dos tratamentos e o quadrado médio dos resíduos, tal razão é denotada por F. Assumindo que os tratamentos não têm nenhum efeito diferenciador entre si, isto é, que eles são equivalentes, calculam-se todas as estatísticas F possíveis provenientes desta hipótese. Ordenam-se todos os F's em ordem crescente e contam-se quantas estatísticas são maiores ou iguais à F observada na amostra. Esta quantidade é o nível de significância do teste ou valor-p, que é reportado ao pesquisador ou ao responsável pelo experimento para tomar as devidas decisões.

O grande problema para o procedimento acima é o grande número de resultados possíveis, tornando-o proibitivo à medida que o tamanho do experimento aumenta.

As propriedades da função poder deste teste foram, empiricamente, estudadas

para muitas situações na Universidade Estadual de Iowa, E.U.A. Neste capítulo, simulações também são feitas para averiguação dos mesmos resultados.

Segundo as referências [14] e [18] quando a hipótese nula, de igualdade de efeito de tratamentos, é verdadeira, a estatística F, independentemente da distribuição dos erros ser normal ou a de referência, têm a mesma média e as variâncias muito próximas, o que já fornece alguma evidência , do uso na prática da F como um artifício de cálculo aproximado para a verdadeira distribuição de referência.

As simulações são baseadas nos experimentos aleatorizados em blocos de tamanho 2. Neste caso, dois tratamentos são testados e n pares com estes tratamentos constituem as u.e. É importante lembrar mais uma vez que as u.e. dentro de cada bloco são alocadas aleatoriamente aos tratamentos. De acordo com o modelo (4.4), i=1,...n e k=1,2; o estimador da diferença $t_1 - t_2$ do i-ésimo par é:

$$y_{i1} - y_{i2} = (t_{i1} - t_{i2}) + \left(\sum_j \delta_{ij}^1 \epsilon_{ij} - \sum_j \delta_{ij}^2 \epsilon_{ij} \right). \tag{4.5}$$

Antes da realização do experimento, cada uma das possibilidades abaixo tem probabilidade 0.5 de ocorrência.

$$\left(\sum_j \delta_{ij}^1 \epsilon_{ij} - \sum_j \delta_{ij}^2 \epsilon_{ij} \right) = \pm (\epsilon_{i1} - \epsilon_{i2}). \tag{4.6}$$

Com base nisto, constrói-se o teste segundo o procedimento descrito anteriormente.

O poder da função quando $(t_1 - t_2)$ é igual a uma particular constante δ_0 é calculado da seguinte maneira: cada atribuição de tratamentos às u.e. produzem um conjunto de n elementos do tipo $\delta_0 \pm (\epsilon_{i1} - \epsilon_{i1})$, i=1,...,n; onde em cada conjunto os sinais que precedem as quantidades $(\epsilon_{i1} - \epsilon_{i1})$ vêm numa particular sequência. Obviamente, a diferença que foi efetivamente observada no experimento faz parte deste conjunto. Na realização do teste descrito na seção (4.1), verifica-se a proporção de

estatísticas que caíram na região crítica, estabelecida por um α previamente fixado pelo pesquisador, tal quantidade representa o poder do teste quando $t_1 - t_2 = \delta_0$.

Uma característica inerente para os **n** valores da forma $\delta_0 \pm (\epsilon_{i1} - \epsilon_{i1})$, é que quando δ_0 é suficientemente grande e positivo todas as diferenças são também positivas, independente do sinal de $(\epsilon_{i1} - \epsilon_{i1})$. De posse desta informação, é de se esperar que a função poder seja igual a 1 para todo $\delta_0 = t_1 - t_2$ suficientemente grande.

4.3 - Simulação

Nesta seção, são ilustradas duas situações. A primeira, faz uso de uma coordenada da $F_\alpha(1,n)$, como delimitadora da região crítica à direita da distribuição, para construção da função poder considerando diferentes δ_0, diferentes distribuições e um $\alpha=0.05$. A segunda, vem mostrar a aproximação da distribuição F(1,n) para a distribuição de referência, para isso diferentes tamanhos de experimentos são considerados.

Vale a pena lembrar que $F_{1,\nu} = t_\nu^2$, onde a primeira é uma variável aleatória com distribuição F com 1 e ν graus de liberdade, e a segunda é uma variável aleatória t-de-Student com ν g.l. Percebe-se melhor a relação olhando as suas funções densidades:

$$f(x|\nu_1,\nu_2) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)\frac{x^{\frac{(\nu_1-2)}{2}}}{\left(1+\left(\frac{\nu_1}{\nu_2}\right)x\right)^{\frac{(\nu_1+\nu_2)}{2}}}; \; 0 \le x < \infty \; e \; \nu_i \ge 1, \; i = 1,2;$$

$$f(x|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \frac{1}{\left(1+\left(\frac{x^2}{\nu}\right)\right)^{\frac{(\nu+1)}{2}}}; \quad -\infty \leq x < \infty \ e \ \nu \geq 1.$$

Partindo da condição de que $t^2(\nu) = F(1,\nu)$ e o fato de que o critério para se testar a igualdade dos tratamentos segundo a t- de- Student é bem mais simples, opta-se pelo seu uso ao invés da F.

O programa em SAS para as simulações encontra-se no apêndice B.

As distribuições normal, gama e uniforme foram consideradas para a construção da função poder, todas com a mesma média e variância. Apenas os experimentos de tamanho 10 têm a sua representação gráfica impressa logo abaixo. Pode-se constatar, além da semelhança dos três, o que foi dito antes, que à medida que o delta aumenta o poder se aproxima de um. Um outro fato também importante, é que para experimentos pequenos o tamanho do teste proveniente da distribuição é em geral maior que os 5% fixado na simulação, na verdade, é aproximadamente 9%. Este fato está de acordo com o que foi encontrado na Universidade Estadual de Iowa. O contrário acontece para experimentos grandes, ou seja, há uma boa consonância entre as duas regiões críticas.

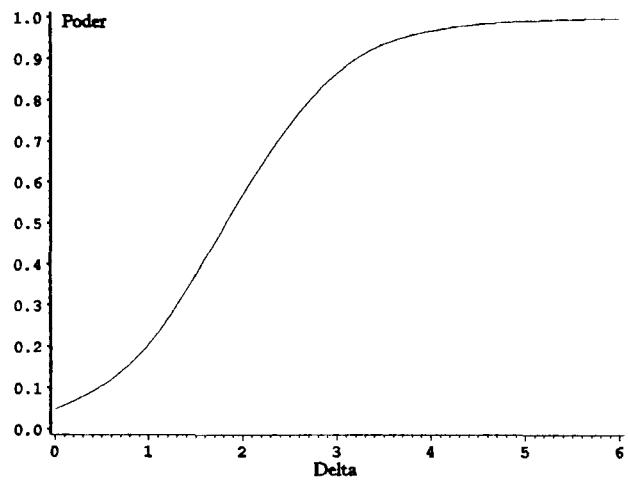


Gráfico 1- Distribuição normal

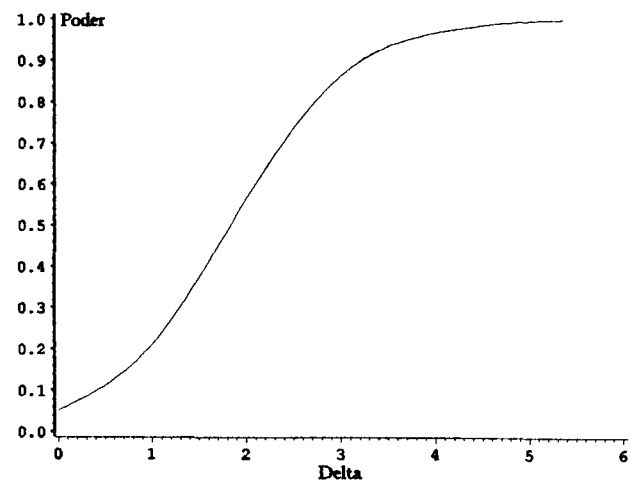


Gráfico 2 - Distribuição Gama

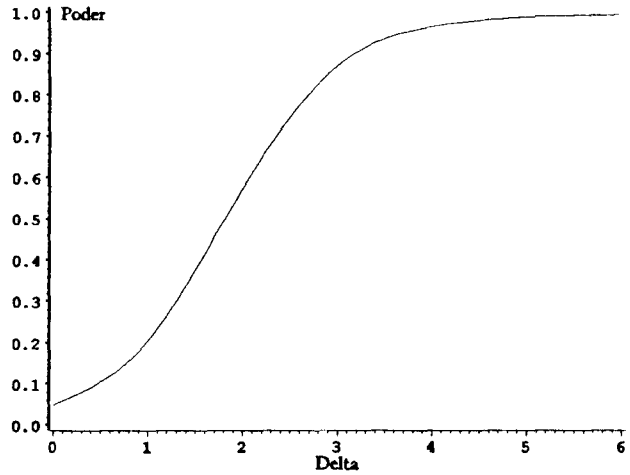
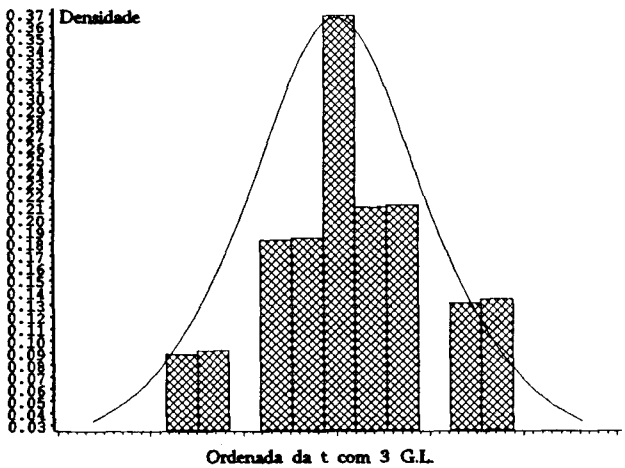
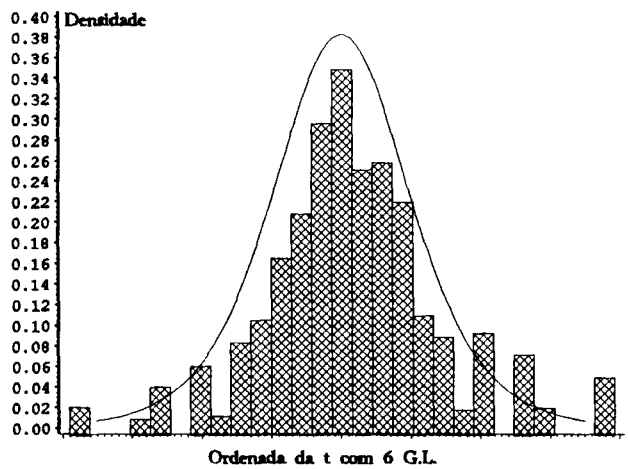
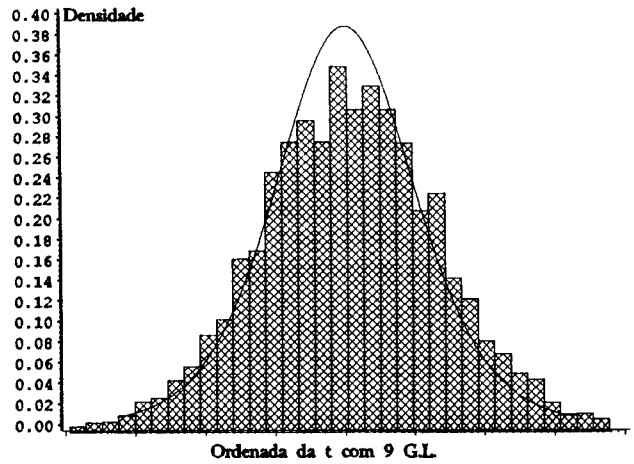


Gráfico 3- Distribuição Uniforme

Para a segunda parte nenhuma novidade, isto é, como era de se esperar a aproximação da distribuição t-de-Student para a distribuição de referência melhora conforme o tamanho do experimento aumenta.







Chama-se a atenção, aqui, para um fato que causa muita confusão e de extrema relevância, já que leva a resultados e interpretações errôneas, que é o nível de significância calculado sob H_0 com base na distribuição de referência gerada por um experimento aleatorizado. Primeiro, define-se o conjunto de referência:

Definição 4.1: ***Conjunto de referência*** = { ***todos os arranjos possíveis, sob a hipótese nula, e equiprováveis*** }.

Com base na definição acima, usar um elemento proveniente de um outro conjunto que satisfaça apenas uma das características do conjunto de referência, por exemplo, ser um dos arranjos, mas não ter a equiprobabilidade, é uma condição suficiente para que este elemento não possa ser usado, como se fosse um componente do conjunto acima, objetivando alguma inferência. Isto serve de ilustração para situações onde não se aleatoriza o experimento e se usa o argumento de que o resultado obtido é um dos possíveis resultados daquele experimento, e, portanto, não se deveria ter o trabalho adicional de aleatorizar. Bastaria, então, ver como o resultado observado está diante de todas as possibilidades, chegando, deste modo, a um p-valor. O procedimento acima está intrinsicamente errado, pois, como já foi dito, é necessário também garantir a

equiprobabilidade de todos os arranjos para o uso adequado da distribuição de referência nas condições da definição 4.1. O resultado desta má aplicação, é que o p-valor, e qualquer inferência, não é fidedigno à distribuição da qual é obtido, e consequentemente não pode ser usado para os dados observados, o verdadeiro valor-p pode ser tanto maior quanto menor que o valor calculado. A saída para este problema, quando não se conhece a probabilidade de alocar os tratamentos, é supor a distribuição normal para os erros, etc, e verificar se as hipóteses são razoáveis. Caso a probabilidade seja conhecida, mas as atribuições não sejam equiprováveis, basta levar esta informação em consideração no cálculo do p-valor, ou em qualquer outra inferência. Entretanto, este não é o procedimento usual, pois tem como premissa privilegiar uma certa sequência de atribuições de tratamentos. Para mais detalhes desta abordagem ver a referência[6].

Resumindo, sempre que possível deve-se aleatorizar o experimento.

Apêndice A

Álgebra Linear

Neste apêndice estão alguns conceitos de Álgebra Linear usados no decorrer da dissertação, sobretudo no capítulo 1. As referências [3], [11] e [15] constituem ótimas fontes para consultas neste assunto, e é onde se encontram muitas das provas de resultados, teoremas e corolários expostos aqui. Com isso, nenhuma demonstração será efetuada.

1.Espaço Vetorial

Definição: Um *espaço vetorial* real é um conjunto **V**, não vazio, com duas operações:

$$\textit{soma}(+),\ V \times V \rightarrow V \ e$$

$$\textit{multiplicação por um escalar}(\bullet),\ \mathbb{R} \times V \rightarrow V,$$

tais que, para quaisquer **u**, **v** e **w** $\in$ **V** e a e b$\in$ $\mathbb{R}$, as propriedades abaixo sejam satisfeitas.

Propriedades:

i) (**u** + **v**) + **w** = **u** + (**v** + **w**);

ii) **u** + **v** = **v** + **u**;

iii) $\exists$ **0** $\in$ **v** tal que **u** + 0 = **u**;

iv) $\exists$ -**u** $\in$ **v** tal que **u** + (-**u**)=0;

v) a(**u** + **v**) = a**u** + a**v**;

vi) (a + b)**v** = a**v** + b**v**;

vii) (a b)**v** = a(b**v**);

viii) $1\mathbf{u} = \mathbf{u}$.

Obs.: Se na definição acima, ao invés de se ter números reais como escalares se tivesse números complexos, V seria um espaço vetorial complexo. Na presente dissertação é usado apenas o espaço vetorial real.

2.Subespaço Vetorial

Definição: Dado um espaço vetorial V, um subconjunto W, não vazio, será um *subespaço vetorial* de V se:

i) Para quaisquer $\mathbf{u}, \mathbf{v} \in W$ tem-se que $\mathbf{u} + \mathbf{v} \in W$;

ii) Para quaisquer $a \in \mathbb{R}$, $\mathbf{u} \in W$ tem-se $a\mathbf{u} \in W$.

Pode-se fazer três observações:

a) As condições da definição acima garantem que se operando em W, não se obterá um vetor fora de W. Isto é suficiente para afirmar que W é ele próprio um espaço vetorial, pois assim as operações ficam bem definidas e, além disso, não é preciso a verificação das propriedades de (i) a (viii) de espaço vetorial, porque elas são válidas em $V \supset W$.

b) Qualquer subespaço W de V precisa necessariamente conter o vetor nulo(por causa da condição (ii) quando a = 0).

c) Todo espaço vetorial admite pelo menos dois subespaços (que são chamados subespaços triviais), o conjunto formado somente pelo vetor nulo e o próprio espaço vetorial.

3.Combinação Linear

A finalidade da combinação linear é a obtenção de novos vetores a partir de vetores dados.

Definição: Sejam V um espaço vetorial real, $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$ e $a_1,\ldots,a_n$ números reais. Então, o vetor

$$\mathbf{v} = a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n$$

é um elemento de V ao que se chama *combinação linear* de $\mathbf{v}_1,\ldots,\mathbf{v}_n$.

Uma vez fixados os vetores $\mathbf{v}_1,\ldots,\mathbf{v}_n$ em V, o conjunto de W de todos os vetores de V que são combinações lineares destes, é um subespaço vetorial. W é chamado subespaço gerado por $\mathbf{v}_1,\ldots,\mathbf{v}_n$ e simbolicamente $W = [\mathbf{v}_1,\ldots,\mathbf{v}_n]$.

4.Dependência e Independência Linear

Em álgebra linear, é fundamental saber se um vetor é uma combinação linear de outros, por exemplo: se o espaço gerado por $\mathbf{v}_1$, $\mathbf{v}_2$ e $\mathbf{v}_3$ for o mesmo espaço gerado por $\mathbf{v}_1$ e $\mathbf{v}_2$, chega-se a conclusão de que $\mathbf{v}_3$ é uma combinação linear de $\mathbf{v}_1$ e $\mathbf{v}_2$, e portanto, é um vetor "supérfluo" para a descrição do espaço vetorial.

Para a identificação de tais vetores "supérfluos" é necessária a definição de dependência e independência linear.

Definição: Sejam V um espaço vetorial e $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$. Diz-se que o conjunto $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ é *linearmente independente* (LI), ou que os vetores $\mathbf{v}_1,\ldots,\mathbf{v}_n$ são LI, se a equação

$$a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n = 0 \Leftrightarrow a_1 = \ldots = a_n = 0.$$

No caso de existir algum $a_i \neq 0$ diz-se que $\{\mathbf{v}_1 + \ldots + \mathbf{v}_n\}$ é *linearmente dependente* (LD), ou que os vetores $\mathbf{v}_1,\ldots,\mathbf{v}_n$ são LD.

Vetores LD podem ser caracterizados de uma outra forma.

Teorema: $\{v_1, \ldots, v_n\}$ é LD se, e somente se um deste vetores for uma combinação linear dos outros.

5.Base de um Espaço Vetorial

Agora, o interesse está em encontrar, dentro de um espaço vetorial V, um conjunto finito de vetores, tais que qualquer outro vetor de V seja uma combinação linear deles. Em outras palavras, procura-se determinar um conjunto de vetores que gere V e tal que todos os elementos sejam realmente necessários para esta geração. Se tais vetores existirem, pode-se dizer que estes são os alicerces de V.

Definição: Um conjunto $\{v_1, \ldots, v_n\}$ de vetores de V será uma base de V se:

i) $\{v_1, \ldots, v_n\}$ é LI;

ii)$[v_1, \ldots, v_n]=V$.

6.Produto Interno

O interesse, agora, é formalizar conceitos que permitam falar em comprimento de um vetor e de ângulo entre dois vetores, que são conceitos importantes para a idéia de medição de área, volume e aproximação entre vetores. Daí a necessidade da definição de um instrumento para isso.

Definição: Seja V um espaço vetorial real. Um *produto interno* sobre V é uma função que a cada par de vetores, v_1 e v_2, associa um número real, denotado $< v_1 , v_2 >$, satisfazendo as propriedades:

i) $< v , v > \geq 0 \; \forall \, v \in V$ e $< v , v > = 0 \Leftrightarrow v = 0$;

ii)$< av_1 , v_2 > = a< v_1 , v_2 >, \forall \, a \in \mathbb{R}$;

iii) $< v_1 + v_2 , v_3 > = < v_1 , v_3 > + <v_2 , v_3 >$;

iv) $< v_1 , v_2 > =< v_2 , v_1 >$.

6.1.Ortogonalidade

Definição: Seja V um espaço vetorial com produto interno <,>. Diz-se que dois vetores v e w de V são ortogonais (em relação a este produto interno) se $< v , w > = 0$. No caso em que v e w são ortogonais, escreve-se $v \perp w$.

Propriedades:

i) $0 \perp v, \forall v \in V$;

ii) $v \perp w \Diamond w \perp v$;

iii) Se $v \perp w, \forall w \in V \Diamond v = 0$;

iv) Se $v_1 \perp w$ e $v_2 \perp w \Diamond v_1 + v_2 \perp w$;

v) Se $v \perp w$ e λ é um escalar $\Diamond$ $\lambda v \perp w$.

6.2.Norma

Definição: Seja V um espaço vetorial com produto interno <,>. Defini-se *norma* (comprimento) de um vetor v em relação a este produto interno por:

$\|v\| = \sqrt{< v , v >}$.

6.3.Ângulo Entre Dois Vetores

Desigualdade de Schwarz $\Diamond$ $|< v, w >| \leq ||v|| \, ||w||$.

A partir da desigualdade de Schwarz é possível definir ângulo entre dois vetores não nulos em um espaço vetorial V munido de um produto interno. Sejam v e $w \in V$ não nulos. Então, a desigualdade acima pode ser reescrita como:

$$\frac{|< v , w >|}{\|v\| \, \|w\|} \leq 1, \text{ ou seja, } \left|\frac{< v , w >}{\|v\| \, \|w\|}\right| \leq 1,$$

e, portanto, existe um ângulo θ entre 0 e π radiano tal que

$\cos\theta = \frac{\langle v, w \rangle}{\|v\| \|w\|}$. Este ângulo θ é chamado de ângulo entre **v** e **w**. Observe que esta noção é compatível com a de ortogonalidade pois, se $\langle \mathbf{v}, \mathbf{w} \rangle = 0$, então $\cos\theta = 0 \Leftrightarrow \theta = \pi/2$.

6.4.Complemento Ortogonal

Considere um espaço vetorial **V** munido de um produto interno <,> e um subconjunto não vazio **S** de **V** (**S** não é necessariamente um subespaço). Considere, então, o seguinte subconjunto de **V**:

$$\mathbf{S}^{\perp} = \{\mathbf{v} \in \mathbf{V} : \mathbf{v} \text{ é ortogonal a todos os vetores de } \mathbf{S}\}.$$

Os seguintes resultados são válidos em relação a **S**.

i) $\mathbf{S}^{\perp}$ é subespaço de V (mesmo que **S** não o seja);

ii) Se **S** é subespaço de V, então $\mathbf{V} = \mathbf{S} \oplus \mathbf{S}^{\perp}$ e $\mathbf{S}^{\perp}$ é chamado complemento ortogonal de **S**.

7.Transformação Linear

Definição: Sejam **V** e **W** dois espaços vetoriais. Uma *transformação linear* (aplicação linear) é uma função de **V** em **W**, F: $\mathbf{V} \to \mathbf{W}$, que satisfaz as seguintes condições:

i) Quaisquer que sejam **u** e **v** em **V**, $F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v})$;

ii) Quaisquer que sejam **v** em **V** e $k \in \mathbb{R}$, $F(k\mathbf{v}) = kF(\mathbf{v})$.

8.Inversa Generalizada

Definição: Seja A uma matriz m x n. Um matriz n x m A^- é dito ser uma *inversa generalizada* (ou G-inversa) se ela satisfaz a condição

$$AA^-A = A$$

Em geral A^- não é única, exceto para o caso em que A é uma matriz quadrada e não singular, e neste caso

Teorema: As seguintes propriedades valem para qualquer escolha da G-inversa.

a) $A(A'A)^{-}A'A = A$

b)$(A(A'A)^{-}A')' = A(A'A)^{-}A'$.

9.Projeção

9.1.Operador Projeção

Definição: Sejam **V** e **W** dois subespaços disjuntos que geram o espaço E^n, isto é, $E^n = V \oplus W$. Então qualquer vetor $x \in E^n$ pode ser unicamente decomposto como $x = y + z$, onde $y \in V$ e $z \in W$. A correspondência de **x** em **y** determina uma transformação linear **P** de E^n em **V**, que é expressa por

$$y = Px.$$

P é uma matriz quadrada de ordem n, e é o operador projeção, ou simplesmente projetor, em **V** na direção de **W**. P goza das seguintes propriedades:

i) P é linear e único em E^n, isto é, $P(a_1x_1 + a_2x_2) = a_1Px_1 + a_2Px_2$, $\forall x_1, x_2 \in E^n$;

ii) P é uma matriz idempotente, ou seja, $P^2 = P$.

9.2 Subespaços e Projetores

Um projetor pode ser definido sobre um subespaço na direção de um subespaço complementar.

Agora, serão consideradas algumas relações entre projetores e subespaços lineares.

Teorema 9.2.1: Sejam V_1 e V_2 espaços lineares em E^n, e seja W_1 e W_2 os correspondentes subespaços complementares de V_1 e V_2, respectivamente. Denote por P_1 e P_2 os projetores em V_1 na direção de W_1 e em V_2 na direção de W_2. Então, os seguintes resultados são equivalentes:

(a) $\mathbf{V}_1 \subset \mathbf{V}_2$ e $\mathbf{W}_1 \subset \mathbf{W}_2$;

(b) $P_1P_2 = P_2P_1$;

(c) $P_2 - P_1$ é o projetor em $\mathbf{V}_2 \cap \mathbf{W}_1$ na direção de $\mathbf{V}_1 \cap \mathbf{W}_2$.

Teorema 9.2.2: Nas condições do teorema 9.2.1, os seguintes resultados são equivalentes:

a) $\mathbf{W}_1 \supset \mathbf{V}_2$ e $\mathbf{W}_2 \supset \mathbf{V}_1$ $\diamond$ $\mathbf{V}_1$ e $\mathbf{V}_2$ são disjuntos;

b) $P_1 + P_2$ é o projetor em $\mathbf{V}_1 \oplus \mathbf{V}_2$ na direção de $\mathbf{W}_1 \cap \mathbf{W}_2$.

Teorema 9.2.3: Nas condições do teorema 9.2.2, P_1P_2 é um projetor em $\mathbf{V}_1 \cap \mathbf{V}_2$ na direção de $\mathbf{W}_1 + \mathbf{W}_2$, se, e somente se

$$P_1P_2 = P_2P_1 .$$

Definição: Seja $\mathbf{V}$ um subespaço em E^n, e seja $\mathbf{V}^\perp$ o complemento ortogonal de $\mathbf{V}$. O projetor de $\mathbf{V}$ na direção de $\mathbf{V}^\perp$ é dito ser o projetor ortogonal em $\mathbf{V}$, e além das propriedades (i) e (ii) para projetores quaisquer, o projetor ortogonal tem ainda:

iii) O projetor ortogonal é simétrico, isto é, $P' = P$.

Note que os teoremas 9.2.1, 9.2.2 e 9.2.3 podem ser modificados da seguinte forma: suponha que P_1 e P_2 sejam projetores ortogonais em $\mathbf{V}_1$ e $\mathbf{V}_2$ respectivamente.

Teorema 9.2.4: As seguintes afirmações são equivalentes:

a) $\mathbf{V}_1 \subset \mathbf{V}_2$;

b) $P_1P_2 = P_2P_1 = P_1$;

c) $P_1 - P_2$ é um projetor ortogonal em $\mathbf{V}_1^\perp \cap \mathbf{V}_2$.

Teorema 9.2.5: As seguintes afirmações são equivalentes:

a) $\mathbf{V}_1$ e $\mathbf{V}_2$ são ortogonais;

b) $P_1P_2 = P_2P_1$;

c) $P_1 + P_2$ é o projetor ortogonal em $V_1 \oplus V_2$.

Teorema 1.9.6: P_1P_2 é projetor em $V_1 \cap V_2$ se, e somente se, $P_1P_2 = P_2P_1$.

Forma explícita de um projetor. Considere uma matriz $X_{n\,x\,p}$. Então

$$P_x = X(X' X)^- X'.$$

Diz-se que P_x é uma projeção ortogonal de $\mathbb{R}^n$ sobre o espaço coluna de X. E $I-P_x$ é o projetor ortogonal no complemento ortogonal do espaço coluna de X.

Decomposição de projetores definidos em soma de subespaços lineares.

Considere dois subespaços V_1 e V_2 gerados por dois conjuntos de vetores n-dimensionais, $F = (f_1, \ldots, f_p)$ e $G = (g_1, \ldots, g_p)$, isto é, $V_1 = [F]$ e $V_2 = [G]$. Seja $Q_F = I - P_F$ e $Q_G = I - P_G$.

Teorema: $P_{F \cup G} = P_F + P_{G/F} = P_G + P_{F/G}$,
onde $P_{G/F} = Q_G G(G'Q_F G)^- G' Q_F$ e $P_{F/G} = Q_F F(F'Q_G F)^- F' Q_G$.

Corolário: No caso de um conjunto de vetores F ser particionado em $F=(F_1, \ldots, F_m)$ tem-se que;

$$P_F = P_{F1} + P_{F2/F1} + \ldots + P_{Fm/(F1 \cup \ldots \cup Fm-1)}.$$

Teorema: Suponha que $V_{F \cap G}$. Então, uma condição suficiente para $P_{F \cup G} = P_F + P_G - P_{F \cap G}$ ser verdadeira é que $P_FP_G = P_GP_F$.

Uma outra maneira é:

Teorema: $P_{F \cup G} = P_F + P_G - P_{F \cap G}$ se, e somente se, $P_FP_G = P_GP_F$.

Com isso, foi dada uma base para a associação de subespaços e projetores a cada um dos componentes, na decomposição do vetor de resposta no capítulo 1.

Apêndice B

Programas

Neste apêndice, consta o programa em *SAS* para a geração da distribuição de referência como também o programa para geração da função poder, ambos considerando o experimento aleatorizado em blocos de tamanho dois. Apenas o programa com dados gerados com distribuição normal e com 10 blocos é impresso, para as outras distribuições, basta mudar a função que gera os números aleatórios, e para mudar a quantidade de blocos basta aumentar ou diminuir o número de laços DO existentes no programa.

A versão do *software The SAS System* utilizada é a 6.08.

```
data dados;                    ************************************
   do i=1 to 10;               * Passo Data responsável pela geração *
     dif=rannor(0);            * de números aleatórios segundo a    *
     output;                   * distribuição Normal                *
   end;                        ************************************
  keep dif;

proc transpose out=saida;      ****************************************
   var dif;                    * Transforma as 10 observações da variável*
                               * dif em 10 variáveis com nome coli, com *
                               * i variando de 1 até 10                 *
                               ****************************************

data saida;                                        ********************
```

```
filename denst10'c:\users\carla\denst10.cgm'; * Serve para exportar um gráfico SAS para
                                              * o WordPerfect                          *
                                              ******************************************

proc gchart data=denst;          ************************************************
   vbar tcalc/                   * Faz o histograma das frequências segundo variável *
   space=0                       * tcalc                                            *
   freq=tcalc                    ************************************************
   type=freq
   noaxis;
   goptions device=cgmwpwa gsfmode=replace gsfname=vbar10; * Opcional, para *
                                                          *exportar para o WP*
symbol i=spline;

proc gplot;                   ******************************************************
   plot denst*tcalc;          * Faz o gráfico da densidade para os pontos calculados no*
   goptions gsfname=denst10; * passo data=t                                          *
                             ******************************************************

run;
```

O programa que calcula a função poder é basicamente o mesmo programa para a aproximação da distribuição T - de - Student para a distribuição de referência. Somente nos lugares que tiverem diferenças é que haverá comentários .

```
data boys;
   do i=1 to 10;
      dif=sqrt(3)+rannor(0); * A multiplicação por sqrt(3) é para garantir a mesma
média                          * e variância que foi usada em todas as distribuições
consideradas
     output;
   end;
   keep dif;

proc transpose out=saida;
   var dif;

data saida;
   set saida;
   tobs=mean(of col1-col10)/stderr(of col1-col10);

data t;
   set saida;
   do delta=0 to 7;
    k=0;
    k1=0;
    total=0;
    do i1=0,1;
```

```
do i2=0,1;
 do i3=0,1;
  do i4=0,1;
   do i5=0,1;
    do i6=0,1;
     do i7=0,1;
      do i8=0,1;
       do i9=0,1;
        do i10=0,1;
         tcalc=mean(col1*((-1)**i1),col2*((-1)**i2),col3*((-1)**i3),
               col4*((-1)**i4),col5*((-1)**i5),col6*((-1)**i6),
               col7*((-1)**i7),col8*((-1)**i8),col9*((-1)**i9),
               col10*((-1)**i10))/stderr(col1*((-1)**i1),
               col2*((-1)**i2),col3*((-1)**i3),
               col4*((-1)**i4),col5*((-1)**i5),col6*((-1)**i6),
               col7*((-1)**i7),col8*((-1)**i8),col9*((-1)**i9),
               col10*((-1)**i10))+delta;
         total+1;
         if tcalc>=tinv(0.95,9) then k1+1;
         poder=k1/total;                    * Cálculo do poder *
       if tcalc>=(tobs+delta) then k+1;
       pvalor=k/total;
      end;
     end;
    end;
   end;
  end;
 end;
```

```
        end;
       end;
      end;
     end;
    end;
    output;
   end;
   keep pvalor delta poder;

symbol i=spline;

filename pn10'c:\users\carla\pn10.cgm';

proc gplot;                          * Gráfico da função poder *
   plot poder*delta/vref=1;
   goptions device=cgmwpwa gsfmode=replace gsfname=pn10;

run;
```

Bibliografia

[1] - Amarante, A.R.(1992) - Um Curso em Modelos Lineares - Tese de Mestrado - Instituto de Matemática, Estatística e Ciência da Computação da Universidade Estadual de Campinas(IMECC - UNICAMP).

[2] - Basu, D.(1980) - Randomization Analysis of Experimental Data: The Fisher Randomization Test - Journal of American Statistical Assocition, 75, 591 - 593.

[3] - Boldrine, J.L., Costa, S.I.R., Figueredo, V.L. e Wetzler, H.G.(1984) - Álgebra Linear 3ª edição - HARBRA.

[4] - Box, G.E.P., Hunter, W.G. e Hunter, J.J. (1978) - Statistics for Experimenters - An Introduction to Design, Data analysis, and Model Building - John Wiley & Sons Inc.

[5] - Harville, D.A.(1975) - Experimental Randomization: Who Needs It? - The American Statistician, Vol. 29, N°1, 27 - 31.

[6] - Holland, P.W.(1986) - Statistics and Causal Inference - Journal of American Statistical Assocition, 81, 945 - 969

[7] - Kempthorne, O., Zyskind, G., Addelman, Throckmorton e White (1961) - Analysis of Variance Procedure - Aerospace Research Laboratory (ARL 149) - Iowa State University, Ames, Iowa.

[8] - Kempthorne, O.(1977) - Why Randomize? - Journal of Satatistical Planning and Inference 1, 1 - 25.

[9] - Kempthorne, O.(1978) - Logical, Epistemological and Statistical Aspects of Nature-Nurture Data Interpretation - Biometrics 34, 1- 23.

[10] - Kempthorne,O.(1983) - Design and Analysis of Experiments - Robert E. Krieger Publishing Company.

[11] - Mota, R.M.S.(1982) - Caracterização dos Melhores Estimadores Lineares não Tendenciosos no Modelo Linear Geral - Tese de Mestrado - Instituto de Matemática e Estatística da Universidade de São Paulo(IME - USP).

[12] - Neyman, J., Iwaszkiewicz, K. , Kolodziejczyk, St.(1935) - Statistical Problems in Agricultural Experimentation - Suplemento do Journal of the Royal Statistical Society, Vol. II, N°2, 107 - 54.

[13] - Petenate, A.J.(1979) - Modelos Lineares com Matriz de Covariância Arbitrária - Condições para que o Estimador Simples de Mínimos Quadrados seja BLUE - Tese de Mestrado - Instituto de Matemática, Estatística e Ciência da Computação da Universidade Estadual de Campinas(IMECC - UNICAMP).

[14] - Pitman, E.J.G.(1937) - Significance Tests Which May Be Applied To Samples From Any Populations III. The Analysis of Variance Test, Biometrika, 29, 322 - 335.

[15] - SAS Institute Inc. *SAS/GRAPH®* User's Guide, Version 5 edition. Cary,NC: SAS Institute., 1985. 596 pp.

[16] - Takeuchi, Yanai e Mukherjee(1984) - The Foundations of Multivariate Analysis - A Unified Approach by Means of Projections Onto Linear Subspaces - Second Reprint -

Wiley Eastern Limited.

[17] - Tjur,T.(1984) - Analysis of Variance Models in Orthogonal Designs - International Statistical Review 52, 1, 33 - 81.

[18] - Welch,B.L.(1937) - On The z-Test In Randomized Blocks And Latin Squares, Biometrika 29, 21 - 52.

[19] - White, R.F.(1970) - Randomization Analysis of the General Experiment - Aerospace Research Laboratory (ARL 70 - 0239) - Iowa State University, Ames, Iowa.

[20] - White, R.F.(1975) - Randomization and the Analysis of Variance - Biometrics 31, 552 - 572.

[21] - Zyskind, Kempthorne, White, Dayhoff, Doerfler(1964)- Research on Analysis of Variance and Related Topics - Aerospace Research Laboratory (ARL 64 - 199) - Iowa State University, Ames, Iowa.