Geometry, Statistics and Decision Making in Gene Therapy

Victor Pătrângenaru

Abstract

One emphasizes the role of geometry in normal parametric statistics and introduces the pi-beta distributions. Using a standard geometric interpretation of a random effect model, one explains the origin of tests for means useful in a statistical analysis of the effect of adenovirus treatments of tumors in vitro and in vivo.

Introduction

This paper is addressed to a large audience, some of which at times has to teach an introductory course in analytic geometry. It the past years some departments of mathematics, have eliminated such courses from their undergraduate curriculum.

The present paper comes with a very concrete example as of why a multidimensional geometry course is necessary both in a mathematics or statistics department.

One of the most important results in probability is the central limit theorem (CLT). The CLT states that variates which are sums of many independent and identically distributed effects tend to be normally distributed as the number of effects becomes large [10].

A normally distributed random variable has the probability density function

$$dp = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu^2)\right) dx. \tag{0.1}$$

When possibly related measurements such as heights and weights of individuals of same group are made, the quadratic term under the exponential in (0.1) is substituted with a quadratic form in more variables. It does not take an eye of an expert to see that the hidden correlation parameters (coefficients of the quadratic forms) of these measurements can be expressed in terms of some Euclidean objects such as various projections, volumes, etc. It is less obvious , still straightforward [2], [10] to find the distribution function of these measurements, under certain linear hypothesis. This will be taken care of in the first two sections. The *Wishart distributions* are derived in section 1. In section 2, we will make standard geometric assumptions on the covariance

Balkan Journal of Geometry and Its Applications, Vol.2, No.1, 1997, pp. 83-100

matrices, to write computable tests in terms of ANOVA (analysis of variance) tables. These tests are used to decide to what extent the mean value of a multivariate normal distribution of an orthogonal model, lies in a vector subspace L versus being in a larger subspace M. In very few cases such a test may be based on the popular F- *distribution;* in general they are unknown, and under additional splitting constraints on L and M, one is lead to *pi-beta* tests, that is products tests based on powers of independently *beta distributed* r.v.'s, which are rarely tabulated, leading to interesting numerical problems.

Which brings us to the second goal of this paper, a question regarding decision making in recent *gene therapy* models experimented at Dr. Milton W. Taylor's biology lab [11]. We will put it in two ways:

One is do decide if there is a significant difference at the end of the application of various treatments in *vitro* of randomly the **MDA-MB-435** breast cancer cells. This leads to an F test.

The other concerns the difference *in time* between treatments with *recombinant adenovirus* on **K562** *melanoma cells in vivo*, in animal models. We will explain in details the design we use, which is called by some a 3 way ANOVA. This leads to a pi-beta test.

It is interesting to note that other modern approved immunotherapeutical protocols [9], using TIL (tumor infiltrating leukocytes) who have an established 40% response in patients with *melanoma* or *kidney cancer*, were not successful in breast carcinoma [6], which makes the successful preclinical trials in [11] even more important.

The relationship between geometry and statistics, goes beyond this introductory paper [5]. Parametric statistics is viewed by some researchers as the study of *statistical manifolds,* which are Riemannian manifolds of distributions, where the metric is the Fisher information on the space of parameters [1] .

1 The geometry of multivariate normal distribution

Multivariate analysis deals with ordered lists of data, representing a number of aspects of the same phenomenon. These numbers or measurements are indexed by the ordered set I, of size p, so that one such list can be regarded as a matrix-valued random variable $(r.v.)$ $X = (X_n)$, where X_n are independent multivariate normal distributed r.v.'s, with same covariance matrix $\Sigma = \left(\sigma^{ij}\right)_{(i,j)\in I\times I}$, that is has the probability density function $(p.d.f.)$:

$$(1.1) \qquad (2\pi)^{-\frac{p}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} < x_n - \mu_n, \quad \Sigma^{-1}(x_n - \mu_n) >\right),$$

where $<,>$ is the usual scalar product on $\mathbf{R}^p$.

Assume X_n are identically distributed. We would like to make inferences about the "hidden parameters" (μ, Σ) in $R^p \times P(I)$ from the given numerical data $(x_n)_{n=\overline{1,N}}$.

In this respect we will follow the *maximum likelihood estimator* (MLE) [2] approach
.

The joint distribution function of $X = (X_n)_{n=\overline{1,N}}$, is a multiple of the so called the *likelihood function*

$$(1.2) \qquad L(\mu, \Sigma; x) = \det(\Sigma)^{-\frac{N}{2}} \exp\left(-\frac{1}{2} \sum_{n=\overline{1,N}} < x_n - \mu,\ \Sigma^{-1}(x_n - \mu) >\right).$$

The MLE of $(\hat{\mu}, \hat{\Sigma}) = (\hat{\mu}, \hat{\Sigma})(x)$ is the point of maximum of $L(\cdot, \cdot; x)$ and may be easily determined with a geometric argument. Indeed, for Σ fixed, define on $\mathbf{R}^p$ the scalar product $< \cdot, \cdot >_\Sigma := < \cdot, \Sigma^{-1}(\cdot) >$, which obviously extends to a scalar product on $(\mathbf{R}^p)^N$, labeled $<< \cdot, \cdot >>_\Sigma$.

Let $\Delta = \Delta((\mathbf{R}^p)^N)$ be the diagonal of elements of the form $(x, x, ..., x),\ x \in \mathbf{R}^p$, and let P be the orthoprojection of $((\mathbf{R}^p)^N, << \cdot, \cdot >>_\Sigma)$ on Δ, which is actually independent of Σ.

Let $\overline{x}$ be the baricenter with equal weights of the $\mathbf{R}^p$ - components of $x = (x_n)_{n=\overline{1,N}}$. Then

$$(1.3) \qquad P((x_n)_{n=\overline{1,N}}) = (\overline{x})_{n=\overline{1,N}}.$$

$Q = Id - P$ is the orthoprojection on the orthocomplement $\Delta^\perp$ (again w.r.t. any of the scalar products $<< \cdot, \cdot >>_\Sigma$). It is obvious that

$$(1.4) \qquad \|x - (\mu, \mu, ..., \mu)\|_\Sigma^2 \geq \|Q(x)\|_\Sigma^2$$

and the equality holds iff $\mu = \overline{x}$. The value of the likelihood function at such points is

$$(1.5) \qquad L(\overline{x}, \Sigma; x) = \det(\Sigma)^{-\frac{N}{2}} \exp(-\frac{1}{2} \|Q(x)\|_\Sigma^2).$$

Assume $\{x_n - \overline{x}\}_{n=\overline{1,N}}$ spans $\mathbf{R}^p$, which happens iff the rows of $Q(x)$ are l.i., or if the Gramm determinant $\det(Q(x)\, Q(x)^t)$ is positive. Then if Ω is a positive definite square root of Σ^{-1}, then

$$(1.6) \qquad L(\overline{x}, \Sigma; x) =$$

$$= \det(Q(x)\, Q(x)^t)^{-\frac{N}{2}} \det(\Omega Q(x)\, Q(x)^t \Omega)^{\frac{N}{2}} \exp(-\frac{1}{2} Tr(\Omega Q(x)\, Q(x)^t \Omega)).$$

If we set $\Lambda = \Omega Q(x)\, Q(x)^t \Omega$, then $L(\overline{x}, \Sigma; x)$ is proportional with

$$G(\Lambda) = \det(\Lambda)^{\frac{N}{2}} \exp(-\frac{1}{2} Tr \Lambda)$$

and since Λ is symmetric, there is an orthogonal matrix T, such that $T \Lambda T^t = \Delta$ is diagonal. Then $G(\Lambda) = G(\Delta)$ and w.l.o.g., one may suppose that $\Lambda = \operatorname{diag} \Lambda_i$. Therefore

$$g(\Lambda) = \ln G(\Lambda) = \sum_{i \in I} \left(\frac{N}{2} \ln \Lambda_i - \frac{1}{2} \Lambda_i\right).$$

The max of $u(t) = \dfrac{1}{2}(N \ln t - t)$ occurs for $t = N$, and therefore the max of $g(\Lambda)$ occurs when $\Lambda_i = N$, that is $\Lambda = N\, Id$, or $\Sigma^{-1} Q(x)\, Q(x)^t = N\, Id$.

Let $\Phi : \mathbf{R}^p \otimes \mathrm{R}^N \to (\mathbf{R}^p)^N$ be the isomorphism $\Phi(x \otimes y) = (y^1 x, ..., y^N x)$, and for a vector subspace V of $\mathbf{R}^p$ consider $S = \Phi(\mathbf{R}^p \otimes V)$. Since $\Sigma S = S$, for any Σ in $P(I)$, $S^\perp$ is independent of Σ.

Assume the mean $\underline{\mu} = (\mu_n)$ of the distribution in (1.1) is in S. Along the same lines, if P_S is the orthoprojector on S, and $Q_S = 1 - P_S$, we get at no additional effort the following classical result:
Theorem 1.1 *If $N \geq \dim M + p$. Then the MLE of $L(\underline{\mu}, \Sigma; x)$ exists in general and* $\hat{\underline{\mu}} = P_S$, $\hat{\Sigma} = \frac{1}{N} Q_S(x) Q_S(x)^t$.

Let $D_S = \{x \in \mathbf{R}^p \mid Q_S(x) Q_S(x)^t \text{is positive definite}\}$ and assume F_S associates the MLE with data, that is

$$F_S : D_S \to S \times P(I)$$

$$F_S(x) = (P_S(x), \frac{1}{N} Q_S(x) Q_S(x)^t).$$

In order to answer specific inference questions, one has to find the distribution function of $F_S(x)$. Behind the answer to this standard question, there is again a nice geometric idea. If Ω is positive definite on $W \oplus U$ and W, U are Ω -invariant then for any $W \oplus U$-valued multivariate normally distributed r.v. Y of covariance matrix Ω, the U and W parts are normally distributed and independent as r.v.'s. It turns out that $P_S(x)$ is normally distributed and furthermore $\frac{1}{N} Q_S(x) Q_S(x)^t$ and $P_S(x)$ are independent. Moreover if we select a basis in $\mathbf{R}^N$, the first vectors of which are in M, then $S = (\mathbf{R}^p)^d \times \{0\}$ and $S^{\perp} = \{0\} \times (\mathbf{R}^p)^{N-d}$. As such, w.l.o.g. one may assume that $z = Q_S(x) = (x_1, x_2, ..., x_{N-d})$ is a sample of $N - d$ independent $N(0, \Sigma)$-distributed r.v's. As such we are interested only in a distribution of $\theta = t(z) = z\, z^t$, given that the joint distribution of z is

$$d\, N(z) = \det(\Sigma)^{-\frac{m}{2}} \exp\left(-\frac{1}{2} \mathrm{Tr}(\Sigma^{-1} z\, z^t)\right) d\lambda(z),\ m = N - d.$$

By the change of variable formula, θ has the density

$$\det(\Sigma)^{-\frac{m}{2}} \exp\left(-\frac{1}{2} \mathrm{Tr}(\Sigma^{-1} \Theta)\right) dt(\lambda)(\Theta). \tag{1.7}$$

In order to determine $dt(\lambda)$, we make a digression on invariant measures we learned from the senior author of [4].We are looking for *GL(p) -invariant measure* on $P(I)$, given that $GL(p)$ acts transitively on $P(I)$ by

$$(A,\, \Sigma) \to A \Sigma A^t. \tag{1.8}$$

Let G be a group, acting on Y, on the left. If μ is a measure on Y, and g is a transformation of Y, $g^{-1}\mu$ is the measure defined by $g^{-1}\mu\,(M) =: \mu(gM)$.
Definition a. A measure μ on Y *is relatively invariant, with multiplicator* χ, if for any g in G, $g^{-1}\mu = \chi(g)\mu$. Since χ has to be positive, and it follows that $\chi(1_G) = 1$.
b. A measure μ on Y is *invariant,* if μ is relatively invariant with multiplicator 1.
Example 1.1 The Lebesgue measure on $\mathbf{R}^p$ is relatively invariant w.r.t. the natural action of the affine group $GL(p) \times \mathbf{R}^p$ on $\mathbf{R}^p$ with multiplicator

$$\chi((g, u)) = \det(g).$$

Remark 1.1. Assume μ is a relatively invariant measure on Y with multiplicator χ, and $n : Y \to \mathbf{R}$ is a measurable function, such that $n(gx) = \chi(g) n(x)$. Then $n^{-1}\mu$ is an invariant measure.

Lemma 1.1. *The Lebesgue measure λ on $P(I)$, is relatively invariant w.r.t. the natural action* (1.6) *of $GL(p)$ on $P(p)$ with multiplicator $\chi(g) = (\det g)^{p+1}$.*
Proof (sketch). Assume $A = cId$, then $A \cdot \Theta = A\Theta A^t = c^2\Theta$. In this case

$$A^{-1}\lambda(S) = \int_{c^2 S} d\tau_1^1 ... d\tau_p^p = c^{2(p+1)p/2} \int_S d\tau_1^1 ... d\tau_p^p = (\det A)^{p+1}\lambda(S).$$

It is obvious then that $n(\Theta) = (\det\,\Theta)^{\frac{p+1}{2}}$ satisfies the conditions in Remark 1.1. and therefore
Corollary 1.1. *Let $(A, \Theta) \to A\Theta A^t$ be the transitive action of $GL(p)$ on $P(I)$ and λ be the Lebesgue measure on $P(I) = GL(p)/O(p)$. Then density function relatively to the Lebesgue measure given by* $(\det\,\Theta)^{-\frac{(p+1)}{2}}$ *defines $GL(p)$ invariant measure on $P(I)$.*

We recall that the Lebesgue measure λ is $GL(p)$ invariant, that is if A is in $GL(p)$ transformation will justify the formula:

$$A^{-1}\lambda = |\det\,A|^p\lambda.$$

Notice that $GL(p)$ acts on the left on $(\mathbf{R}^p)^m$. Since t is equivariant, $t(\lambda)$ is relatively invariant with multiplicator

$$\chi(A) = |\det\,A|^m$$

since

$$A^{-1}t(\lambda)\,(G) = t\lambda(A\,G) = \int_{\{z, zz^t \in AGA^t\}} d\lambda_z = \int_{\{z, A^{-1}z(A^{-1}z)^t \in G\}} d\lambda_z =$$

$$= |\det\,A|^m \int_{\{u, uu^t \in G\}} d\lambda_u = |\det\,A|^m t(\lambda)\,(G).$$

Let then $n : P(I) \to \mathbf{R}$, $n(\Theta) := (\det\Theta)^{\frac{m}{2}}$, satisfies to the condition in Remark 1.1, and thus $(\det\Theta)^{-\frac{m}{2}}t(\lambda)$ is an invariant measure on $P(I)$. Two invariant measures are equal up to a constant on each orbit of the group action. In our case since $P(I)$ is homogeneous for the action (1.6) is transitive, it follows that

$$dt(\lambda)_\Sigma = (\det\,\Theta)^{\frac{m}{2}}(\det\,\Theta)^{-\frac{p+1}{2}}d\lambda_{P(I)}.$$

Then, due to (1.8), we have shown that
Proposition 1.2. *If z is a sample of $N(0, \Sigma)$, then the p.d.f. of $z\,z^t := \Theta$ is proportional to*

$$\det(\Sigma)^{-\frac{m}{2}} \det(\Theta)^{\frac{m-p-1}{2}} \exp(-\frac{1}{2}Tr(\Sigma^{-1}\Theta)). \tag{1.9}$$

The probability distribution proportional to (1.9), is named *Wishart distribution* [2] *with* n *degrees of freedom, and form parameter* Σ, and is labeled as $W(\Sigma, m)$.
Corollary 1.2. *Assume $X = (X_n)_{n=\overline{1,N}}$ is a sample of $N(\underline{\mu}, \Sigma)$, $\mu \in S$. Then the sample variance $\frac{N}{N-d}\hat{\Sigma}$ is distributed as $W(\frac{1}{N-d}\Sigma,\, N-d)$.*

2 Orthogonal designs

Many problems in multivariate statistics are concerned with test for means. One would like to estimate from experimental data the subspace S, where the parameter $\underline{\mu}$ of the $N(\underline{\mu}, \Sigma)$ distributed r.v. $X = (X_i)_{i=\overline{1,N}}$. Such tests can be expressed in terms of some Wishart distributions. Still given the peculiar form of the subspace S, we would like to relax the conditions on Σ. We will assume that $X = (X_i)_{i=\overline{1,N}}$ is a sample from $\mathbf{R}^p$-valued r.v. that is $N(\mu, \Sigma)$ - distributed.

a. To start with let us first assume that Σ has the simplest possible form, that is $\Sigma = \sigma^2 Id$. In this case the likelihood function is

$$L(\mu, \sigma; x) = \sigma^{-Np} \exp\left(-\frac{\sigma^{-2}}{2} \sum \|x_n - \mu\|^2\right). \tag{2.1}$$

Let $L \subseteq M$ be two vector subspace of R^p. We would like to test to what extent the hypothesis $H : \mu$ is in L, versus μ is in M, holds true ([10], p.33).
Let pr_L denote the orthoprojection of $\mathbf{R}^p$ on L and $\overline{x}$ as before denote the baricenter with equal weights of $(x_n)_{n=\overline{1,N}}$.
Remark 2.1. Since there is only one variance component, one may substitute list the data in a single observation of size $N\,p$, and substitute L by the diagonal of L^N, as subspace of $(\mathbf{R}^p)^N$. As such w.l.o.g., one may assume that $N = 1$.

A straightforward computation shows that under the hypothesis that μ is in L, the MLE of $(\mu\,, \sigma^2)$ is

$$(\hat{\mu}_L, \hat{\sigma}_L^2) = \left(pr_L\overline{x}, \frac{1}{p}\|Q_L(x)\|^2\right) \tag{2.2}$$

and the corresponding maximum of the likelihood function is

$$\hat{L}_L = (\hat{\sigma}_L^2)^{-\frac{p}{2}} \exp(-\frac{p}{2}). \tag{2.3}$$

The components of x are uncorrelated, and being normally distributed, they are independent. If $d = dim\, L$, since $\sigma^{-1}Q_L(x)$ is a vector valued r.v. all of that splits into a sum of independent unit normally distributed r.v.'s $\frac{1}{p-d}\sigma^{-2}\|Q_L(x)\|^2$ follows a *chi-square distribution with* $p-d$ *degrees of freedom,* χ^2_{p-d}. The *likelihood ratio* associated whit the hypothesis H, is

$$\hat{L}_L/\hat{L}_M = (\hat{\sigma}_M^2/\hat{\sigma}_L^2)^{p/2}.$$

Before we proceed further we recall some facts on beta distributions. The *beta integrals with form parameters* $(a,\, b)$ are $B(a,b) = \int_0^1 r^{a-1}(1-r)^{b-1}dr$. The *beta distribution with parameters* $(a,\, b)$ has the density function on $(0,\, 1)$:

$$B(a,b)^{-1}r^{a-1}(1-r)^{b-1}.$$

Recall also that $\Gamma(a) = \int_0^\infty x^{a-1}\exp(-x)dx$ and the *gamma distribution with form parameter* **a** has the density function on $(0,\, \infty)$:

$$\Gamma(a)^{-1}x^{a-1}\exp(-x).$$

A straightforward calculation shows that if $x,\, y$ are independent r.v's with gamma distributions of form parameters $a,\, b$, then $x+y$, $r = x/(x+y)$ are also independent and r has a beta distribution with form parameters (a, b). In particular, since a χ^2_m distributed r.v, follows a gamma distribution with form parameter $\frac{m}{2}$, from the above argument, it turns out that

Proposition 2.1. *$\hat{\sigma}^2_M/\hat{\sigma}^2_L$ has a distribution function g on the interval $(0, \mathrm{codim} M/\mathrm{codim} L)$ given by*

$$g(u) = \mathrm{codim} L/\mathrm{codim} M f(u\, \mathrm{codim} L/\mathrm{codim} M), \tag{2.4}$$

where f is a beta distribution with form parameters $(\mathrm{codim} M, \mathrm{codim}_M L)$.

Proof. We have to add only the remark that $\|Q_M(x)\|^2$ and $\|Q_L(x) - Q_M(x)\|^2$ are independent and orthogonal.

Still the beta distributions are seldom used in this case. Historically another distribution related to the beta distribution appeared first. The larger the ratio $\hat{\sigma}^2_L\ /\ \hat{\sigma}^2_M$, the less likely the hypothesis that μ is in L. This ratio is $1+ \parallel Q_L(x) - Q_M(x) \parallel^2\ /\ \parallel Q_M(x) \parallel^2$.

Assume $\dim M = m$. The hypothesis H is rejected if the ratio of orthoprojections $R(x) = \parallel Q_L(x) - Q_M(x) \parallel^2\ /\ \parallel Q_M(x) \parallel^2$ is large enough. Since $Q_L(x) - Q_M(x)$, $Q_M(x)$ are independent as random variables, and $\frac{1}{m-d}\sigma^{-2} \parallel Q_L(x) - Q_M(x) \parallel^2$ follows a $\chi^2(m-d)$, and since the denominator and numerator of $R(x)$ are independent r.v.'s, the ratio $F(x) = \dfrac{\mid I \mid -d}{m-d}R(x)$, as a quotient of two independent chi-square distributions with prescribed degrees of freedom, will follow an F distribution with bidegree $(\mid I \mid -m, m-d)$, $F_{(\mathrm{codim} M, \mathrm{codim}_M L)}$.

An $F_{a,b}$ distribution function on $(0, \infty)$ is given by

$$f(x) = B(a,b)^{-1} a^a b^b x^{a-1}/(ax+b)^{a+b}\ ,\ x > 0$$

for small bidegrees, their cumulative distributions are tabulated in textbooks.

Proposition 2.2. *Let $\alpha \in (0, 1)$. H is rejected at level the confidence level $1-\alpha$, if $F(x) > F_\alpha$, where F_α is such that $\int_{F_\alpha}^\infty f(x)dx = \alpha$ and f is the distribution function of $F_{\mathrm{codim}\, M, \mathrm{codim}_M L}$.*

b. The symbol $\perp$ stands for a direct sum of orthogonal subspaces. Assume that the sample space has an orthogonal decomposition $R^I = S_1 \perp S_2 \perp \ldots \perp S_c$, and w.r.t. this decomposition, the components of the vector valued random variable are uncorrelated. Also assume for each j, the S_j - component has the covariance matrix $\Sigma_j = \sigma_j^2\ Id$. Such a decomposition is usually called in statistics *orthogonal design.* Note that for an orthogonal design, the probability distribution function obviously factors

$$L(\mu, \Sigma; x) = \Pi_j L_j(\mu_j, \Sigma_j; x_j) \tag{2.5}$$

and each of the factors L_j is given by a (2.1) type of expression. It is then obvious that if L has a decomposition $L = \perp L_j$, with $L_j \subseteq S_j$, then by a similar argument as in Rem. 2.1, one may assume that $N = 1$. For each j ,

$$\left(\hat{\mu}_{L_j}\ ,\ \hat{\sigma}^2_{jL_j}\right) = \left(\mathrm{pr}_{L_j}\overline{x}\ ,\ \frac{1}{|I_j|}\|{}_jQ_{L_j}(x)\|^2\right). \tag{2.6}$$

In (2.6), ${}_jQ_{L_j}$ represents the orthoprojection on $S_j \cap L_j^\perp$. Note that ${}_jQ_{L_j}(x)$ is a vector valued "unitary" normally distributed random variable of rank $\mid I_j \mid -d_j$, where $d_j = \dim L_j$ and $\mid I_j \mid = \dim S_j$ and $\frac{1}{|I_j|-d_j}\sigma_j^{-2}\left\|{}_jQ_{L_j}(x)\right\|^2$ follows a $\mathcal{X}^{\in}_{|\mathcal{I}_|| - \lceil_|}$ distribution. The likelihood ratio associated with the hypothesis H in **a.**, where M is also assumed to have an orthogonal splitting $M = \perp M_j$,

$$p(x) = \hat{L}_L / \hat{L}_M = \Pi_j \left(\hat{\sigma}^2_{jM_j} / \hat{\sigma}^2_{jL_j}\right)^{|I_j|/2}. \tag{2.7}$$

According to Proposition 2.1, each of the factors $\hat{\sigma}^2_{jM_j}/\hat{\sigma}^2_{jL_j}$ involved in the likelihood ratio, has up to a constant a beta distribution of form parameters $(\mid I_j \mid - m_j,\ m_j - d_j)$. It is also obvious that these factors are independent. We set

$$F_j(x) = \frac{\mid I_j \mid - d_j}{m_j - d_j}\left(\left\|Q_{L_j}(x) - Q_{M_j}(x)\right\|^2 / \left\|Q_{M_j}(x)\right\|^2\right).$$

There is only one situation in which the question whether H is true has a straightforward answer, that is when L and M differ only on one component of the decomposition of S. Then we may use an F-test.

Proposition 2.3. *Let $\alpha \in (0,1)$. H_{j_0} be the hypothesis H in which $L_j = M_j$, if $j \neq j_0$. H_{j_0} is rejected at level the confidence level $1-\alpha$, if $F_{j_0}(x) > F_\alpha$, where F_α is such that $\int_F^\infty f(x)dx = \alpha$, where f is the distribution of a*

$$F_{|I_{j_0}|} - m_{j_0},\ m_{j_0} - d_{j_0}.$$

The situation encountered in Proposition 2.3 occurs seldom in practical situations. For the general case, we need a definition. Let s be a notation for the multi index $(s_j)_{j=\overline{1,r}}$.

Definition 2.1. A *pi-beta* distribution with parameters the multi indices *(a,b)* and powers the multi index n is a distribution of a product r independent powers of beta distributed random variables, the j- factor, being the n_j power of a r.v. following a beta distribution with form parameters (a_j, b_j). Such a distribution is denoted by $\pi\beta(a,b;n)$.

Let b be the p.d.f. of $\pi\beta(a,b;n)$ distributed r.v. and let $p_{a,b;n}(\alpha) = t$ be a value such that $\int_0^t b(s)ds = \alpha$. We proved the following

Theorem 2.1.. *Assume $L = \perp L_j \subseteq M = \perp M_j$ and $\alpha \in (0,1)$. Then H is rejected at level the confidence level $1-\alpha$, if $p(x) < p_{a,b;n}(\alpha)$.*

In concrete situations, since tables of $\pi\beta_(a,b;n)$ distributions are scarce one runs into numerical problems, solvable on a computer algebra system.

3 Two and three way layouts with applications to genetherapy

In this last section, we give two examples of orthogonal designs, as an application of the theory considered in the previous section.

a. The *two way layout (with one observation per cell)* is a classic example of orthogonal model [7]. $I = R \times C$, where R stands for rows and C for columns; identify $\mathbf{R}^I$

with $R \times C$ real matrices, with the convention $(r, c) = rc$. The natural projection $R \times C \to R,\ rc \to r$, induces the inclusion of $\mathbf{R}^R \to \mathbf{R}^I$, given by $(x_r) \to (x_{rc})$, where for each $c, X_{rc} = x_r$. The image of $\mathbf{R}^R$ under this monomorphism is L_R, the set of all $R \times C$ matrices with equal components within each row. Similarly, L_C is set of all $R \times C$ matrices with equal components within each column, and $L_0 =$ the set of $R \times C$ matrices with equal entries. Also let $L_{R+C} = L_R + L_C$. Thus the main subspaces for this design are the vertices of the lattice

(3.1)

The scalar product in $L = \mathbf{R}^I$ is in matrix notation $x \cdot y = \mathrm{Tr}(xy^t)$ which induces the following orthogonal decomposition w.r.t.:

$$\mathbf{R}^I = L_0 \perp (L_R \cap L_0^\perp) \perp (L_C \cap L_0^\perp) \perp L_{R+C}^\perp. \tag{3.2}$$

Note the following useful decompositions :

$$\begin{aligned} L_R &= L_0 \perp (L_R \cap L_0^\perp) \\ L_C &= L_0 \perp (L_C \cap L_0^\perp) \\ L_{R+C} &= L_0 \perp (L_R \cap L_0^\perp) \perp (L_C \cap L_0^\perp). \end{aligned}$$

If we set $\mathbf{I}$ to be a matrix with all entries $= 1$, the projection $R_0 := P_0$, is given by

$$R_0(x) = \frac{\mathrm{Tr}(x\mathbf{I}^t)}{\mid R \mid \mid C \mid}\mathbf{I} = \bar{x}_{..}\mathbf{I} = (\bar{x}_{..}). \tag{3.3}$$

One dot as index means average of the data, w.r.t. that index.

The orhtoprojections onto the other summands in (3.2) are

$R_R := P_R - P_0$,
$R_C := P_C - P_0$,
$R_1 := I_{\mathbf{R}^I} - P_{R+C} = I_{\mathbf{R}^I} - (P_R + P_C - P_0)$,
$R_0 x := (\bar{x}_{..})$, *is the contrast vector for the level*
$R_R x = (\bar{x}_{r.} - \bar{x}_{..})$, *is the contrast vector for the rows*
$R_C x = (\bar{x}_{.c} - \bar{x}_{..})$, *is the contrast vector for the columns*
$R_1 x = (x_{rc} - \bar{x}_{r.} - \bar{x}_{.c} + \bar{x}_{..})$, *is the contrast vector for the interaction.*

Their lengths are called *effects*, and are key elements for the test statistics, are usually found in the so called ANOVA tables under various abbreviations, coming from their analytical formulae.

degrees of freedom	SSD
$\mid I \mid - \mid R \mid - \mid C \mid +1$	$\Vert R_1 x \Vert^2 = SSD_1$
$\mid R \mid -1$	$\Vert R_R x \Vert^2 = SSD_R$
$\mid C \mid -1$	$\Vert R_C x \Vert^2 = SSD_C$
1	$\Vert R_0 x \Vert^2 = SSD_0$

The covariance matrix has the decomposition

$$\Sigma(\sigma_1^2,\ \sigma_R^2,\ \sigma_C^2,\ \sigma_0^2) = \Sigma = \sigma_1^2 R_1 + \sigma_R^2 R_R + \sigma_C^2 R_C + \sigma_0^2 R_0,$$

where $\sigma_1^2, \sigma_R^2, \sigma_C^2, \sigma_0^2$ are the covariance components, and various orthogonals submodels , can be obtained by setting them equal in accordance with the partial order of the lattice (3.1). They are

1. $(\sigma_1^2 = \sigma_R^2 = \sigma_C^2 = \sigma_0^2)$
2. $(\sigma_1^2 = \sigma_R^2 = \sigma_C^2,\ \sigma_0^2 > 0)$
3. $(\sigma_1^2 > 0,\ \sigma_R^2 = \sigma_C^2 = \sigma_0^2)$
4. $(\sigma_1^2,\ \sigma_R^2 > 0,\ \sigma_C^2 = \sigma_0^2)$
5. $(\sigma_1^2,\ \sigma_C^2 > 0,\ \sigma_0^2 = \sigma_R^2)$
6. $(\sigma_1^2 = \sigma_R^2,\ \sigma_C^2 = \sigma_0^2 > 0)$
7. $(\sigma_1^2 = \sigma_C^2,\ \sigma_R^2 = \sigma_0^2)$
8. $(\sigma_1^2 = \sigma_R^2,\ \sigma_C^2,\ \sigma_0^2 > 0)$
9. $(\sigma_1^2 = \sigma_C^2,\ \sigma_R^2,\ \sigma_0^2 > 0)$
10. $(\sigma_1^2,\ \sigma_R^2,\ \sigma_C^2,\ \sigma_0^2 > 0)$

Note the covariances are:

$$\Sigma_{rc,\,r'c'} = \begin{cases} \mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) & r \neq r',\ c \neq c' \\ \mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) + \mid c \mid^{-1} \left(\sigma_R^2 - \sigma_1^2\right) & r = r',\ c \neq c' \\ \mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) + \mid R \mid^{-1} \left(\sigma_C^2 - \sigma_1^2\right) & r \neq r',\ c = c' \\ \mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) + \mid c \mid^{-1} \left(\sigma_R^2 - \sigma_1^2\right) & \\ \qquad + \mid R \mid^{-1} \left(\sigma_C^2 - \sigma_1^2\right) + \sigma_1^2 & r = r',\ c = c'\ . \end{cases}$$

The covariance structures together with the five subspaces define all the orthogonal models in this variance component design. There are 50 of them. There is a total of 13 models with unique solution for the MLE.
As far as the likelihood ratio test is concerned there are even less cases.

Out of these models those that are not boxed, are uninteresting (general linear model), those boxed are ill posed from mathematical or *random effect* standpoint, those double boxed will be studied.
The names of these diagrams show which of the variance component models are equal, and tell the subspace of means. For example $0R, C1, C$ is a distribution, with $\sigma_1^2 = \sigma_C^2$, $\sigma_R^2 = \sigma_0^2$, $\mu \in L_C$.

An example of random effect problem in two indices is the following. Assume the r.v.'s X_{rc} are given by

$$X_{rc} = Y + Y_r + Y_c + Y_{rc},$$

where Y, Y_r, Y_c, Y_{rc} are independent normally distributed, Y_r from $N(\alpha_r, \omega_r)$, Y_c from $N(\alpha_c, \omega_c)$, Y from $N(\alpha, \omega_0)$, and Y_{rc} from $N(\alpha_{rc}, \omega_1)$, ω_1 and the covariance matrix for X's is

$$\Sigma(X_{rc}, X_{r'c'}) = \begin{cases} \omega_0 & r \neq r',\ c \neq c' \\ \omega_0 + \omega_r & r = r',\ c \neq c' \\ \omega_0 + \omega_c & r \neq r',\ c = c' \\ \omega + \omega + \omega + \omega & r = r',\ c = c'. \end{cases}$$

We identify the two way layout with the random effect model,

$$\begin{aligned} &\mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) = \omega_0 \\ &\mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) + \mid c \mid^{-1} \left(\sigma_R^2 - \sigma_1^2\right) = \omega_0 + \omega_r \\ &\mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) + \mid R \mid^{-1} \left(\sigma_C^2 - \sigma_1^2\right) = \omega_0 + \omega_c \\ &\mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) + \mid c \mid^{-1} \left(\sigma_R^2 - \sigma_1^2\right) + \mid R \mid^{-1} \left(\sigma_C^2 - \sigma_1^2\right) + \sigma_1^2 = \\ &\qquad = \omega_0 + \omega_r + \omega_c + \omega_1. \end{aligned}$$

That is

$$\begin{aligned} &\mid I \mid^{-1} \left(\sigma_1^2 - \sigma_R^2 - \sigma_C^2 + \sigma_0^2\right) = \omega_0 \\ &\mid c \mid^{-1} \left(\sigma_R^2 - \sigma_1^2\right) = \omega_r \\ &\mid R \mid^{-1} \left(\sigma_C^2 - \sigma_1^2\right) = \omega_c \\ &\sigma_1^2 = \omega_1. \end{aligned}$$

Reading off this correspondence, we see that

$$\begin{array}{lcc} R1 & is & \omega_r = 0 \\ C1 & is & \omega_c = 0 \\ 0RC1 & is & \omega_r = \omega_c = \omega_0 \\ 0RC & is & \end{array}$$

$$\begin{aligned} &\mid I \mid^{-1} \left(\sigma_1^2 - \sigma_{RC0}^2\right) = \omega_0 \\ &- \mid c \mid^{-1} \left(\sigma_1^2 - \sigma_{RC0}^2\right) = \omega_r \\ &- \mid R \mid^{-1} \left(\sigma_1^2 - \sigma_{RC0}^2\right) = \omega_c, \end{aligned}$$

that is $\mid I \mid^{-1} \omega_0 = - \mid c \mid^{-1} \omega_r = - \mid R \mid^{-1} \omega_c$, which is quite strange;

$$\begin{array}{lcc} oC, R1 & is & \omega_r = \omega_0 = 0 \\ oR, C1 & is & \omega_c = \omega_0 = 0. \end{array}$$

As an example one may consider the following problem, in connection with the gene therapy data [11]: assume $\mid c \mid$ gene therapy treatments from the same category (modified adenovirus vectors Ad5-lfu, Ad5-ifn, etc) are injected in $\mid R \mid$ cultures of MDA-MB-435 breast carcinoma . The data, measures the area of the tumor per culture treated after a fixed period of time. We would like to test the following simple hypothesis H: *all the cultures grow in the same way independently on adenovirus type*, versus *K: cultures react the same way to each of the adenoviruses, but different adenoviruses yield different reactions.* Their reactions to these adenoviruses are the $\mid c \mid \mid R \mid$ observations which are assumed to be modeled in the random effect form:

$$X_{rc} = \alpha + Y_r + \alpha_c + Y_{rc},$$

or

$$X_{rc} = Y + Y_r + Y_c + Y_{rc}, \quad \text{with} \quad \omega_c = \omega_0 = 0.$$

We have seen before that this is the mathematical model $oR, C1$.

Let $\hat{\alpha}$ (respectively $\tilde{\alpha}$) be value of α associated with the MLE's of the regular model $oR, C1, C$ (respectively $oR, C1, 0$) .

$\hat{}$ The evaluation spaces are

- for $\hat{\xi} : L_0 \perp (L_C \cap L_0^{\perp})$ and $\hat{\xi} = P_C(x) = (\bar{x}_{.c})$,
- for $\hat{\sigma}^2_{Ro} : L_0 \perp (L_R \cap L_0^{\perp})$ and $\hat{\sigma}^2_{Ro} = \frac{1}{|R|}\|R_R x\|^2 = \frac{1}{|R|}\sum_{rc}(\bar{x}_{r.} - \bar{x}_{..})^2$,
- for $\hat{\sigma}^2_{1C}$ and $\hat{\sigma}^2_{1C} = \frac{1}{|I|-|R|} : L^{\perp}_{R+C}$ and

$$\|R_1 x\|^2 = \frac{1}{\mid R \mid (\mid C \mid -1)}\sum_{rc}(x_{rc} - \bar{x}_{r.} - \bar{x}_{.c} + \bar{x}_{..})^2.$$

$\tilde{}$ The evaluation spaces are

- for $\tilde{\xi} : L_0$ and $\tilde{\xi} = P_0(x) = (\bar{x}_{..})$,
- for $\tilde{\sigma}^2_{Ro} : L_R \cap L_0^{\perp}$ and $\tilde{\sigma}^2_{Ro} = \frac{1}{|R|}\|R_R x\|^2 = \frac{1}{|R|}\sum_{rc}(\bar{x}_{r.} - \bar{x}_{..})^2$,
- for $L_R^{\perp} : \tilde{\sigma}^2_{1C}$ and

$$\tilde{\sigma}^2_{1C} = \frac{1}{\mid I \mid - \mid R \mid}(\|R_1 x\|^2 + \|R_C x\|^2) = \frac{1}{\mid I \mid - \mid R \mid}\sum_{rc}(x_{rc} - \bar{x}_{r.} - \bar{x}_{.c} + \bar{x}_{..})^2.$$

(3.4) $$\mathbf{R}^I = L_0 \perp (L_R \cap L_0^\perp) \perp (L_C \cap L_0^\perp) \perp L_{R+C}^\perp.$$

In the notation of section 2, $\mathbf{R}^I = S_1 \perp S_2,\ S_1 = L_0 \perp (L_R \cap L_0^\perp) = L_R,$

$$S_2 = (L_C \cap L_0^\perp) \perp L_{R+C}^\perp,\ L_1 = L_0,\ L_2 = 0,\ M_1 = L_0 \perp (L_R \cap L_0^\perp),\ M_2 = 0.$$

Since the evaluation spaces differ only on one component, we may use in this case an F-test for means.

Let us derive this test. The likelihood ratio is

(3.5) $$\tilde{L}/\hat{L} = \left(\|R_1 x\|^2/\|R_1 x\|^2 + \|R_C x\|^2\right)^{|I|-|R|/2}.$$

The hypothesis $H : x \in L_0$ is rejected if

$$\|R_C x\|^2 \|R_1 x\|^{-2} > \left(\frac{\mid C \mid -1}{\mid I \mid - \mid R \mid - \mid C \mid +1}\right)^{-1} F_{|C|-1,\ |I|-|R|-|C|+1}(1-\alpha).$$

b. Related to the gene therapy data [11], we consider now question of qualitative responses of the adenoviruses in *vivo* when equal amounts of identical breast cancer tumor cells are subcutaneously injected in breasts of $\mid M \mid \mid T \mid$ female mice and then they are locally treated with $\mid T \mid$ recombinant adenoviruses. One treatment is administered to an equal number of $\mid M \mid$ female mice. The superficial area of the tumor is measured daily. One would like to compare the reactions to different treatments. Like in the previous example the numerical computations are deferred to a future paper. The index set is therefore a subset of $I = D \times T \times M$, so that dtm is the index of the measurement on day d of the m-th mouse tested with the treatment t.

$$\mid I \mid = \mid D \mid \mid T \mid \mid M \mid,\ D = \text{days},\ T = \text{treatments},\ M = \text{mice}.$$

The area measured is modeled by a random variable X_{dtm}. The index ordered set is (dtm) and we under the experiment the mouse is assumed to be random, as such we use a random effect model

(3.6) $$X_{dtm} = \alpha + \alpha_d + \alpha_t + \alpha_{dt} + Y_{tm} + Y_{dtm},$$

where

$$Y_{tm} \sim N(\alpha_{tm}, \omega_{TM}), \quad \omega_{TM} \geq 0$$

and

$$Y_{dtm} \sim N(\alpha_{dtm}, \omega_{DTM}), \quad \omega_{DTM} > 0,$$

are i.d.r.v. 's (usually $\alpha_{tm} = \alpha_{dtm} = 0$). Since we have Y_{tm} and Y_{dtm} random, this is part of a three way layout

$$X_{dtm} = Y + Y_d + Y_m + Y_t + Y_{dt} + Y_{tm} + Y_{dm} + Y_{dtm}.$$

The correlation of the rv's in (3.6) is

(3.7) $$\Sigma_{dtm,\ d't'm'} = \begin{cases} 0 & tm \neq t'm' \\ \omega_{TM} & tm = t'm',\ d \neq d' \\ \omega_{TM} + \omega_{DTM} & dtm = d't'm'. \end{cases}$$

Like in the previous example, we would like to test the hypothesis
H_0: *the mice react the same way to all treatments, in time, that is* $E[X_{dtm}] = \alpha$,
versus
H_1: *the mice react the same way, but in time there is a treatment effect, that is* $E[X_{dtm}] = \alpha + \alpha_{dt}$.
The random effect model:

$$X_{dtm} = \alpha + \alpha_d + \alpha_t + \alpha_{dt} + Y_{tm} + Y_{dtm}$$

is obviously a submodel of

$$X_{dtm} = Y + Y_d + Y_m + Y_t + Y_{dt} + Y_{tm} + Y_{dm} + Y_{dtm}.$$

The idea is again to identify an orthogonal variance component submodel of a three way layout with the random effect model. We use the lattice formulation in [3], but our presentation is self contained. Consider the orthogonal design

This is a sub lattice of a lattice associated with a three way layout. In the subspace language, testing problem can be translated as follows:
H_0: The mean μ is in L_0;
versus
H_1: The mean μ in is L_{DT}.

The decomposition of $\mathbf{R}^I$ in terms of *contrast* (orthogonal) subspaces is

$$\begin{aligned} \mathbf{R}^I &= C_0 \perp C_D \perp C_T \perp C_{DT} \perp C_{TM} \perp C_{DTM} = \\ &= L_0 \perp (L_D \cap L_0^{\perp}) \perp (L_T \cap L_0^{\perp}) \perp ((L_D + L_T)^{\perp} \cap L_{DT}) \perp \\ &\perp (L_{DT}^{\perp} \cap L_{TM}) \perp (L_{DTM} \cap (L_{DT} + L_{TM})^{\perp}). \end{aligned} \tag{3.9}$$

Let $P_0, P_D, P_T, P_{DT}, P_{TM}, P_{DTM} = Id$ be the orthoprojections on the various subspaces. They are given by

$$
\begin{aligned}
&P_0(x) = (\bar{x}_{...}) = \mid I \mid^{-1} (s_0), \quad \text{where } s_0 = \textstyle\sum x_{dtm} \\
&P_D(x) = (\bar{x}_{d..}) = (\mid T \mid \mid M \mid)^{-1} (s_{d..}), \quad \text{where } s_{d..} = \textstyle\sum_{tm} x_{dtm} \\
&P_T(x) = (\bar{x}_{.t.}) = (\mid D \mid \mid M \mid)^{-1} (s_{.t.}), \quad \text{where } s_{.t.} = \textstyle\sum_{dm} x_{dtm} \\
&P_{DT}(x) = (\bar{x}_{dt.}) = (\mid M \mid)^{-1} (s_{dt.}), \quad \text{where } s_{dt.} = \textstyle\sum_{m} x_{dtm} \\
&P_{TM}(x) = (\bar{x}_{.tm}) = (\mid T \mid)^{-1} (s_{.tm}), \quad \text{where } s_{.tm} = \textstyle\sum_{d} x_{dtm} \\
&P_{DTM}(x) = x = (x_{dtm}).
\end{aligned} \tag{3.10}
$$

The contrast vectors

(3.11)

$$
\begin{aligned}
&R_0(x) = (\bar{x}_{...}) \\
&R_D(x) = P_D(x) - P_0(x) = (\bar{x}_{d..} - \bar{x}_{...}) \\
&R_T(x) = P_T(x) - P_0(x) = (\bar{x}_{.t.} - \bar{x}_{...}) \\
&R_{DT}(x) = P_{DT}(x) - P_D(x) - P_T(x) + P_0(x) = (\bar{x}_{dt.} - \bar{x}_{d..} - \bar{x}_{.t.} + \bar{x}_{...}) \\
&L_{TM} = (L_{DT} \cap L_{TM}) \perp (L_{DT}^{\perp} \cap L_{TM}) = L_T \perp (L_{DT}^{\perp} \cap L_{TM}) \\
&R_{TM}(x) = P_{TM}(x) - P_T(x) = (\bar{x}_{.tm} - \bar{x}_{.t.}) \\
&R_{DTM}(x) = (Id(x) + P_T(x) - P_{TM}(x) - P_{DT}(x)) = (x_{dtm} - \bar{x}_{.tm} - \bar{x}_{dt.} + \bar{x}_{.t.}).
\end{aligned}
$$

In the three way layout the covariance matrix has 6 variance components

$$
\begin{aligned}
&\textstyle\sum = \sigma_0^2 R_0 + \sigma_D^2 R_D + \sigma_T^2 R_T + \sigma_{DT}^2 R_{DT} + \sigma_{TM}^2 R_{TM} + \sigma_{DTM}^2 R_{DTM} = \\
&= (\sigma_0^2 - \sigma_D^2 - \sigma_T^2 + \sigma_{DT}^2) P_0 + (\sigma_D^2 - \sigma_{DT}^2) P_D + \\
&+ (\sigma_T^2 - \sigma_{DT}^2 - \sigma_{TM}^2 + \sigma_{DTM}^2) P_T + (\sigma_{DT}^2 - \sigma_{DTM}^2) P_{DT} + \\
&+ (\sigma_{TM}^2 - \sigma_{DTM}^2) P_{TM} + \sigma_{DTM}^2 Id.
\end{aligned} \tag{3.12}
$$

If one identifies the matrix $\sum$ in (3.7) with a linear endomorphism of $\mathbf{R}^I$,

$$\sum = \mid D \mid \ \omega_{TM} P_{TM} + \omega_{DTM} Id. \tag{3.13}$$

The models are the same if

$$
\begin{aligned}
&\sigma_0^2 - \sigma_D^2 - \sigma_T^2 + \sigma_{DT}^2 = 0, \ \sigma_D^2 - \sigma_{DT}^2 = 0 \\
&\sigma_T^2 - \sigma_{DT}^2 - \sigma_{TM}^2 + \sigma_{DTM}^2 = 0, \ \sigma_{DT}^2 - \sigma_{DTM}^2 = 0 \\
&\sigma_{TM}^2 - \sigma_{DTM}^2 = \mid D \mid \omega_{TM}, \ \sigma_{DTM}^2 = \omega_{DTM}.
\end{aligned} \tag{3.14}
$$

or

$$\mid D \mid^{-1} (\sigma_{TM}^2 - \sigma_{DTM}^2) = \omega_{TM}, \ \sigma_{DTM}^2 = \omega_{DTM}.$$

One can treat the random effect model as a 2 variance components model,

$$
\begin{aligned}
&\sigma_0^2 = \sigma_T^2 = \sigma_{TM}^2 = \omega_{DTM} + \mid D \mid \omega_{TM} \\
&\sigma_D^2 = \sigma_{DT}^2 = \sigma_{DTM}^2 = \omega_{DTM}.
\end{aligned}
$$

In the notation of section 2, $\mathbf{R}^I = S_1 \perp S_2,\ S_1 = C_0 \perp C_T \perp C_{TM}$,

$$\begin{aligned} &S_2 = C_D \perp C_{DT} \perp C_{DTM} \\ &L_1 = C_0,\ L_2 = 0, \\ &M_1 = C_0 \perp C_T,\ M_2 = C_D \perp C_{DT}. \end{aligned}$$

As such although the question in the examples in 3.**a.** and 3.**b.** are the same, using the results in the previous section, we see that the test statistics has to be based on a product of powers two independent beta's.
Let h be the distribution function of a product of two independent distributions, a beta with form parameters $(\mid T \mid \mid M \mid - \mid T \mid, \mid T \mid -1)$ by the $\mid D \mid -1$ power of a beta with form parameters

$$(\mid D \mid \mid T \mid \mid M \mid - \mid T \mid \mid M \mid - \mid D \mid \mid T \mid + \mid T \mid, \mid D \mid\mid T \mid - \mid T \mid).$$

Theorem 3.1. *H_0 is rejected at level of confidence $1-\alpha$, if*

$$\left(\|R_{DTM}x\|^2 / \left(\|R_Dx\|^2 + \|R_{DT}x\|^2 + \|R_{DTM}x\|^2\right)\right)^{(|D|-1)} \cdot$$

$$\cdot \left(\|R_{TM}x\|^2 / \left(\|R_Tx\|^2 + \|R_{TM}x\|^2\right)\right) < h_\alpha,$$

where

$$\int_0^{h_\alpha} h(x)dx = \alpha.$$

Remark 3.1. Under this simple hypothesis testing problem, the main obstacles are numerical rather then conceptual. The key steps in solving such a problem are not only hidden in tables for the cumulative distributions of products of powers of beta distributed r.v.' s, but equally for concrete questions the numerical work stays behind the computation of various orthogonal components of the projections of the data x onto the corresponding summands L_j.

Acknowledgement. I am very grateful to Steen Andersson, for sharing his insight on multivariate normal statistics and orthogonal models and to Victor Goodman for useful conversations.

References

[1] S. Amari, *Differential geometry in statistical inference*, Bull. Inst. Internat. Stat., 52 (1987), no 2, 321-338.

[2] T.W. Anderson, *An introduction to multivariate statistical analysis*, 2-nd edition, Wiley Series in Mathematical Probability and Statistics, 1984.

[3] S.A. Andersson, *The lattice structure of orthogonal linear models and orthogonal variance component models*, Scand. J. Statist. 17 (1990), 287-319.

[4] S.A. Andersson, J. Madsen, *Symmetry and lattice conditional independence in a multivariate normal distribution*, to appear in Ann. of Statistics.

[5] O.E. Barndorff - Nielsen, P.E. Jupp, *Statistical inference and differential geometry - some recent developments*, Multivariate analysis: future direction (Barcelona, 1992), 385-396, North Holland Ser. in Statist. Prob., Amsterdam, 1993.

[6] B.J. Coventry, S.C.Weeks, S.E. Heckford, P.J. Sykes, J.M. Skinner, *Lack of IL-2 cytokine expression despite IL-2 messanger RNA transcription in TIL in primary human breast cancer*, J-Immunol. 1996, May 1, 156 (9): 3486-3492.

[7] I. Guttman, *Linear models: an introduction*, Wiley Ser. in Prob. and Math. Stat., 1982.

[8] D. Morrison, *Multivariate Statistical Methods*, McGraw Hill Ser. in Prob. and Stat., Third Ed., 1990.

[9] P.F. Robbins, M.Li.Y.F. El-Gamil, Y. Kawakami, D. Loftus, E. Appella, S.A. Rosenberg, *Breast cancers recognized by TIL*, J. Exp. Med., 1996 Mar. 1; 183 (3); 1185-1192.

[10] H. Scheffé, *The analysis of variance*, J. Wiley and Sons, Inc., 1959.

[11] M.W. Taylor, J.F. Zhang, C. Hu, Y. Geng, J. Selm, S.B. Klein, A. Orazi, *Treatment of a human breast cancer xenograft with an adenovirus vector containing an interferon gene results in rapid regression due to viral oncolysis and gene therapy*, Proc. Natl. Acad. Sci. USA, vol. 93, 4513-4518, April 1996.

Indiana University
Department of Mathematics
Bloomington, IN 47405-5701, USA
e-mail: vpatrang@indiana.edu