BSCB 694: Theory of Multivariate Statistics

Homework 2

Due: Thursday, March 13

1. Does multivariate get you anything?

   A non-central $F(p, n, \mu)$ distribution is defined to be the distribution of

   $$\frac{\sum_{j=1}^{p} \nu_j^2}{\sum_{i=1}^{n} \tau_i^2}$$

   where the τ_i are independently $N(0, 1)$ and the ν_j are independent $N(\mu, 1)$.

   Now we observe that for $\gamma_1 > \gamma_2$, an $F(p, n, \gamma_1)$ random variable is stochastically larger than an $F(p, n, \gamma_2)$ random variable. That is, it has lower probability of taking small values.

   (a) For a two-sample T^2 test, find the distribution of the likelihood ratio test for $\mu_1 = \mu_2$ when $\mu_1 \neq \mu_2$.

   (b) Generalize this result to consider tests of the form $R(\mu_1 - \mu_2) = 0$.

   (c) Are there any conditions under which conducting the above test with $R = e_1^T$ (ie, 1 in the first entry and zeros elsewhere) is more powerful than than conducting a two-sample t-test?

2. (MKB 5.3.3)

   (a) Consider the hypothesis $H_0 : \mu = k\mu_0$ where Σ and μ_0 are known. Derive the maximum likelihood estimate for k. Show that minus twice the likelihood ratio is given by

   $$n\bar{x}^T \Sigma^{-1} \big(\Sigma - (\mu_0^T \Sigma^{-1} \mu_0)^{-1} \mu_0 \mu_0^T\big) \Sigma^{-1} \bar{x}$$

   Deduce that this is distributed exactly as a χ^2_{p-1} random variable.

   (b) Now consider the same problem with Σ unknown. Derive a T^2 test for the hypothesis.

3. Union-Intersection tests for MANOVA.

   A union-intersection test for a MANOVA involves constructing an ANOVA test for the univariate data Xa for each vector a, giving a test statistic F_a that has an F-distribution under the null hypothesis. The union-intersection test statistic is $\max_a F_a$. Show that this statistic is the largest eigenvalue of $W^{-1}B$.

4. Singular Design Matrices

   When X is singular, Mardia, Kent and Bibby suggest using a linearly-independent set of the columns of X in a multivariate regression. However, for designed experiments, alternative ways of interpreting the regression are sometimes desired.

   (a) If X is singular, demonstrate that a General Linear Hypothesis may be conducted by replacing $(X^T X)^{-1}$ with $(X^T X)^{-}$ in the formulae for H and E. Verify that these matrices remain independent and calculate the degrees of freedom to use in a likelihood ratio test. You will need to demonstrate that if $(A^T A)^{-}$ is a generalized inverse for $A^T A$, then $(A^T A)^{-} A$ is a generalized inverse for A.

   (b) Verify Timm's theorem 2.5.5 (6) and his calculation of the degrees of freedom for $\Gamma = 0$ and for parallelism in MANCOVA.

5. Estimable Functions

   (a) Characterize the estimable functions for a two-way MANOVA with factors explicitly given by

   $$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_{ij} + \epsilon_{ijk}$$

   (b) What functions are estimable in the additive two-way MANOVA (ie, $\tau_{ij} = 0$)?

6. Consider a MANOVA using n samples of p-variate observations and q groups with $n - q < p$. Show that the maximum likelihood estimate of Σ is singular and hence a test for the means being equal cannot be conducted with a Λ-test. Characterize the set of tests that can be conducted about the means.

   Bonus: Is there a generalized union-intersection test that will allow all the above tests to be conducted simultaneously? Provide a test statistic, but not a null distribution.