A Note on UMPI F Tests

Ronald Christensen
Professor of Statistics
Department of Mathematics and Statistics
University of New Mexico

May 22, 2015

Abstract

We examine the transformations necessary for establishing that the linear model F test is a uniformly most powerful invariant (UMPI) test. We also note that the Studentized range test for equality of groups means in a balanced one-way ANOVA is not invariant under all of these transformations so the UMPI result says nothing about the relative powers of the ANOVA F test and the Studentized range test.

KEY WORDS: F tests, Linear models, UMPI tests, Uniformly most powerful invariant tests.

1. Introduction

It has been well-known for a long time that the linear model F test is a uniformly most powerful invariant (UMPI) test. Lehmann (1959) discussed the result in the first edition of his classic test and in all subsequent editions, e.g. Lehmann and Romano (2005). But the exact nature of this result is a bit convoluted and may be worth looking at with some simpler and more modern terminology.

Consider a (full) linear model

$$Y = X\beta + e, \qquad e \sim N(0, \sigma^2 I)$$

where Y is an n vector of observable random variables and consider a reduced (null) model

$$Y = X_0\gamma + e, \qquad C(X_0) \subset C(X),$$

where $C(X)$ denotes the column (range) space of X. Let M be the perpendicular projection operator (ppo) onto $C(X)$ and M_0 be the ppo onto $C(X_0)$. The usual F test statistic, which is equivalent to the generalized likelihood test statistic, is

$$F(Y) \equiv F \equiv \frac{Y'(M - M_0)Y/[r(X) - r(X_0)]}{Y'(I - M)Y/[n - r(X)]},$$

where $r(X)$ denotes the rank of X.

Consider a group of transformations $\mathcal{G}$ that map $\mathbf{R}^n$ into $\mathbf{R}^n$. A test statistic $T(Y)$ is invariant under $\mathcal{G}$ if

$$T(Y) = T[G(Y)],$$

for any $G \in \mathcal{G}$ and any Y. It is not too surprising that the F statistic is invariant under location and scale transformations. Specifically, if we define $G(Y) = a(Y + X_0\delta)$ for any positive real number a and any vector δ, it is easy to see using properties of ppos that $F(Y) = F[G(Y)]$. Unfortunately, this is not the complete set of transformations required to get the UMPI result. Note also that the location invariance is defined with respect to the reduced model, that is, it involves X_0. Given that the alternative hypothesis is the existence of a location $X\beta \neq X_0\gamma$ for any γ, one would not want a test statistic that is invariant to changes in the alternative, particularly changes that could turn the alternative into the null.

Before discussing the third group of transformations required for a UMPI F test, let's look at the best known alternative to the linear model F test. Consider a balanced one-way ANOVA model,

$$y_{ij} = \mu_i + \varepsilon_{ij}, \qquad \varepsilon_{ij} \text{ iid } N(0, \sigma^2),$$

$i = 1, \ldots, a$, $j = 1, \ldots, N$. The F statistic for $H_0 : \mu_1 = \cdots = \mu_a$ is

$$F(Y) \equiv F \equiv \frac{N \sum_{i=1}^{a} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 / [a-1]}{\sum_{i=1}^{a} \sum_{j=1}^{N} (\bar{y}_{ij} - \bar{y}_{i\cdot})^2 / [a(N-1)]} \equiv \frac{MSGrps}{MSE}.$$

The best known competitor to an F test for H_0 is the test that rejects for large values of the studentized range,

$$Q(Y) \equiv Q \equiv \frac{\max_i \bar{y}_{i\cdot} - \min_i \bar{y}_{i\cdot}}{\sqrt{MSE/N}}.$$

We already know that F is location and scale invariant and it is easy to see that Q is too. In this case, location invariance means that the test statistic remains the same if we add a constant to every observation. Moreover, it is reasonably well known that neither of these tests is uniformly superior to the other, which means that Q must not be invariant under the full range of transformations that are required to make F a UMPI test.

We can decompose Y into three orthogonal pieces,

$$Y = M_0 Y + (M - M_0) Y + (I - M) Y = X_0 \hat{\gamma} + (X\hat{\beta} - X_0 \hat{\gamma}) + (Y - X\hat{\beta}). \quad (1)$$

The first term of the decomposition contains the fitted values for the reduced model. The second term is the difference between the fitted values of the full model and those of the reduced model. The last term is the residual vector from the full model. Intuitively we can think of the transformations that define the invariance as relating to the three parts of this decomposition. The residuals are used to estimate σ, the scale parameter of the linear model, so we can think of scale invariance as relating to $(I - M)Y$. The translation invariance of adding vectors $X_0\delta$ modifies $M_0 Y$. To get the UMPI result we need another group of transformations that relate to $(M - M_0)Y$. Specifically, we need to incorporate *rotations* of the vector $(M-M_0)Y$ that keep the vector *within* $C(M-M_0) = C(X_0)^{\perp}_{C(X)}$, the orthogonal complement of $C(X_0)$ with respect to $C(X)$. If we allow rotations of $(M-M_0)Y$ within $C(M-M_0)$, the end result can be any vector in $C(M-M_0)$ that has the same length as $(M-M_0)Y$. The length of a vector v is $\|v\| \equiv \sqrt{v'v}$. The end result of a rotation within $C(M-M_0)$ can be, for any n vector v with $\|(M-M_0)v\| \neq 0$,

$$\frac{\|(M-M_0)Y\|}{\|(M-M_0)v\|}(M-M_0)v.$$

Finally, the complete set of transformations to obtain the UMPI result is for any positive number a, any appropriate size vector δ, and any n vector v with $\|(M-M_0)v\| \neq 0$,

$$G(Y) = a\left[M_0Y + X_0\delta + (I-M)Y + \frac{\|(M-M_0)Y\|}{\|(M-M_0)v\|}(M-M_0)v\right].$$

Again, it is not difficult to see that $F(Y) = F[G(Y)]$.

However, in the balanced ANOVA problem, there exist such transformations G with $Q(Y) \neq Q[G(Y)]$, so Q is not invariant under these transformations and when we say that F is UMPI, it says nothing about the relative powers of F and Q. We know that Q is invariant to location and scale changes, so it must be the rotation that Q is not invariant to. Let J_m be an m dimensional vector of 1s. In a one-way ANOVA, write $Y = [y_{11}, y_{12}, \ldots, y_{aN}]'$ and

$$\begin{aligned}(M-M_0)Y &= X\hat{\beta} - X_0\hat{\gamma} \\ &= \begin{bmatrix} J_N & & & 0 \\ & J_N & & \\ & & \ddots & \\ 0 & & & J_n \end{bmatrix}\begin{bmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix} - J_{aN}\,\bar{y}_{\cdot\cdot} = \begin{bmatrix} (\bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot})J_N \\ (\bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot})J_N \\ \vdots \\ (\bar{y}_{a\cdot} - \bar{y}_{\cdot\cdot})J_N \end{bmatrix}. \end{aligned} \tag{2}$$

Since Y is an arbitrary vector, $(M-M_0)v$ must display a similar structure. Also

$$\|(M-M_0)Y\|^2 = N\sum_{i=1}^{a}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 \equiv SSGrps. \tag{3}$$

Thinking about the decomposition in (1), if $Q(Y)$ were invariant we should get the same test statistic if we replace M_0Y with $M_0Y + X_0\delta$ (which we do) and if we replace $(M-M_0)Y$ with $[\|(M-M_0)Y\|/\|(M-M_0)v\|](M-M_0)v$ (which we do not). The numerator of Q is a function of $(M-M_0)Y$, namely, it takes the difference between the largest and smallest components of $(M-M_0)Y$. For $Q(Y)$ to be invariant, the max minus the min of $(M-M_0)Y$ would have to be the same as the max minus the min of $[\|(M-M_0)Y\|/\|(M-M_0)v\|](M-M_0)v$ for any Y and v. Alternatively, the max minus the min of $[1/\|(M-M_0)Y\|](M-M_0)Y$ would have to be the same as the max minus the min of $[1/\|(M-M_0)v\|](M-M_0)v$ for any Y and v. In other words, given (2) and (3),

$$\frac{\max_i \bar{y}_{i\cdot} - \min_i \bar{y}_{i\cdot}}{\sqrt{SSGrps}}$$

would have to be a constant for any data vector Y, which it is not.

Reference

Lehmann, E.L. (1959). *Testing Statistical Hypotheses.* John Wiley and Sons, New York.

Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, Third Edition. Springer, New York.