Multiple hypothesis testing worksheet

Ryan T. Godwin

Multiple hypotheses

1. What is a multiple hypothesis?

   A multiple hypothesis involves more than one parameter, e.g. $H_0 : \beta_2 = 0$ and $\beta_3 = 0$.

2. Write a multiple hypothesis using the R and $\boldsymbol{q}$ matrices.

   For the above example, and for $k = 4$, the R and $\boldsymbol{q}$ matrices are:

$$R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad ; \quad \boldsymbol{q} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   and the null hypothesis is written as $R\boldsymbol{\beta} = \boldsymbol{q}$.

3. What is the expected value of $\boldsymbol{m}$ if the null hypothesis is true?

   $\boldsymbol{m}$ is a vector of differences between the estimates and the values under the null: $\boldsymbol{m} = R\boldsymbol{b} - \boldsymbol{q}$. If the null is true then $E[R\boldsymbol{b}] = R\boldsymbol{\beta} = \boldsymbol{q}$ and $E[m] = \boldsymbol{q} - \boldsymbol{q} = \boldsymbol{0}$.

4. What is the variance of $\boldsymbol{m}$?

$$V[\boldsymbol{m}] = V[R\boldsymbol{b} - \boldsymbol{q}] = V[R\boldsymbol{b}] = RV[\boldsymbol{b}]R' = \sigma^2 R(X'X)^{-1}R'$$

5. What is the distribution of $\boldsymbol{m}$ under the null?

$$\boldsymbol{m} \sim N(\boldsymbol{0}, \sigma^2 R(X'X)^{-1}R')$$

Wald test

6. What is the Wald test statistic?

$$W = \boldsymbol{m}'V[\boldsymbol{m}]^{-1}\boldsymbol{m}$$

7. What is the distribution of the Wald test statistic?

$$W \sim \chi^2_{(J)}$$

   where J is the number of restrictions.

8. What if σ^2 is unknown?

   In the more realistic case that σ^2 is unknown, it must be estimated. This changes the distribution of the test statistic.

9. What is the F-test?

   The F-test is what we get when we replace the unknown σ^2 with the estimator s^2. (We also divide by the number of restrictions J). That is,

$$F = \left(\frac{W}{J}\right)\left(\frac{\sigma^2}{s^2}\right)$$

   The F-test statistic follows the F-distribution with J and $(n-k)$ degrees of freedom.

RLS

10. What are nested models?

    Nested models are when one model can be obtained from another by imposing restrictions on the parameter values (restricting the β).

11. What is the RLS estimator?

    The RLS estimator are the values for $\boldsymbol{\beta}$ that minimize the sum of squared residuals subject to the constraint imposed by the null hypothesis. It is solved for by minimizing the Lagrangian:

$$\mathcal{L} = (\boldsymbol{y} - X\boldsymbol{b}_*)'(\boldsymbol{y} - X\boldsymbol{b}_*) + 2\boldsymbol{\lambda}'(R\boldsymbol{b}_* - \boldsymbol{q})$$

    Solving the minimization problem leads to:

$$\boldsymbol{b}_* = \boldsymbol{b} - \left(X'X\right)^{-1} R'\left[R\left(X'X\right)^{-1} R'\right]^{-1}(R\boldsymbol{b} - \boldsymbol{q})$$

    where we see that RLS = LS + an "adjustment".

12. What is the purpose of studying the RLS estimator?

    The purpose is not to calculate estimates under restrictions (if we wanted to do this we just impose the restrictions and estimate the restricted model). Rather, we use the RLS estimator to evaluate the costs and benefits of imposing restrictions on a model.

13. What does the equation for the RLS estimator tells us?

    It tells us that imposing restrictions which numerically change all of the estimates, except for the silly case where we set restrictions to be exactly equal to what we estimate (set $R\boldsymbol{b} = \boldsymbol{q}$).

14. When is the RLS estimator unbiased (when does imposing restrictions on a model lead to an unbiased estimator)?

$$\begin{aligned} E\left(\boldsymbol{b}_*\right) &= E(\boldsymbol{b}) - \left(X'X\right)^{-1} R'\left[R\left(X'X\right)^{-1} R'\right]^{-1}(\boldsymbol{R}E(\boldsymbol{b}) - \boldsymbol{q}) \\ &= \boldsymbol{\beta} - \left(X'X\right)^{-1} R'\left[R\left(X'X\right)^{-1} R'\right]^{-1}(R\boldsymbol{\beta} - \boldsymbol{q}) \end{aligned}$$

    So $E(\boldsymbol{b}_*) = \boldsymbol{\beta}$ only when the restrictions are true (when $R\boldsymbol{\beta} = \boldsymbol{q}$).

15. What is the variance of RLS compared to LS? What is the intuitive reason for this?

    The variance of the restricted estimator is always less than the unrestricted estimator. Having fewer parameters to estimate means that all available information can focus on the remaining estimators. It is as if the sample size has increased.