Outline of Bonferroni Procedure

1 Multiple Confidence Intervals

Recall the significance level α is the probability of wrongly rejecting the null hypothesis if the null hypothesis is true. In the context of confidence intervals, α is the proportion of random samples that will result in a confidence interval that does not contain the true value of the population parameter.

Consider a simple linear regression equation: $E(y|x) = \beta_0 + \beta_1 x$. Suppose we want to construct $(1-\alpha)100\%$ confidence intervals for both coefficients β_0 and β_1, i.e.,

$$\hat{\beta}_0 \pm t_{1-\alpha/2;n-2} se(\hat{\beta}_0) \tag{1}$$

and

$$\hat{\beta}_1 \pm t_{1-\alpha/2;n-2} se(\hat{\beta}_1). \tag{2}$$

respectively. When conducting multiple inferences on our data, we are concerned whether we still have $(1-\alpha)100\%$ confidence that **both** (1) and (2) will contain the population parameter β_0 and β_1. Analysis of data frequently consists of a series of estimates where we want assurances about the correctness of the entire series of estimates.

2 Family Confidence

We want to maintain the same level of confidence for our entire series, or **family**, of estimates. This is the **family confidence coefficient**: the proportion of families of estimates that are **entirely** correct when repeated samples are selected and specified CIs for the entire family are calculated for each sample.

Consider two events A_1 and A_2, where

$$\begin{aligned} A_1 &= \{(1) \text{ does not cover } \beta_0\} \\ A_2 &= \{(2) \text{ does not cover } \beta_1\} \end{aligned}$$

and we know that $P(A_1) = \alpha$ and $P(A_2) = \alpha$ based on the definition of α. Therefore, the probability of both intervals being correct, i.e., the family confidence, is denoted as $P\left(A_1^c \cap A_2^c\right)$.

$$\begin{aligned} P\left(A_1^c \cap A_2^c\right) &= 1 - P\left(A_1 \cup A_2\right) \\ &= 1 - P\left(A_1\right) - P\left(A_2\right) + P\left(A_1 \cap A_2\right). \end{aligned} \tag{3}$$

Since $P\left(A_1 \cap A_2\right) \geq 0$, (3) becomes

$$\begin{aligned} P\left(A_1^c \cap A_2^c\right) &\geq 1 - P\left(A_1\right) - P\left(A_2\right) \\ &= 1 - 2\alpha. \end{aligned} \tag{4}$$

We can see from (4) the family confidence coefficient is now **at least** $(1-2\alpha)100\%$ instead of being $(1-\alpha)100\%$.

3 Bonferroni Procedure

From (4), we can see that if we want the family confidence coefficient to be at least $(1-\alpha)100\%$ for our two confidence intervals, each confidence interval can be constructed at $(1-\alpha/2)100\%$ confidence. This means that (1) and (2) become

$$\hat{\beta}_0 \pm t_{1-\frac{\alpha}{2}/2;n-2} se(\hat{\beta}_0) \tag{5}$$

and

$$\hat{\beta}_1 \pm t_{1-\frac{\alpha}{2}/2;n-2} se(\hat{\beta}_1). \tag{6}$$

In fact, (4), (5), and (6) can be generalized to any number, g, of confidence intervals we want in our family of estimates since

$$P\left(\bigcap_{i=1}^{g} A_i^c\right) \geq 1 - g\alpha.$$

Therefore, the Bonferroni procedure to ensure that we have at least $(1-\alpha)100\%$ confidence in our family of estimates is

$$\hat{\beta}_j \pm t_{1-\alpha/(2g);n-p} se(\hat{\beta}_j) \tag{7}$$

for any $j = 0, 1, \cdots, k$.

Other procedures exist that are less conservative but are more complicated to implement. For the Bonferroni procedure, all we do is divide the significance level α by the number of confidence intervals we wish to construct, g.

Note: In page 102 of the textbook, the author has a slightly different form for the Bonferroni procedure as I have written in (7):

$$\hat{\beta}_j \pm t_{1-\alpha/(2p);n-p} se(\hat{\beta}_j). \tag{8}$$

The textbook author divides α by p, the number of parameters in the model, whereas I divide α by g, the number of confidence intervals you need to create.