Auburn University NCAT Document **July 18, 2016**

Comparing Several Simple Regression functions for both Slopes and Intercepts Simultaneously **By S. Maghsoodloo**

For the sake of simplicity, suppose we have 3 mixtures and wish to ascertain if Asphalt Content (AC = X) impacts Air-Voids (y = AVs) in a similar manner for all 3 mixes (AV = y = the response). The step-by-step procedure is outlined below. The reader must be cognizant of the fact that most of what is being documented herein by this author has been known in Statistical Literature for well more than 15 years.

1. For Mixture 1, obtain the simple regression of y_1 on x_1. The terminology **Simple**-Regression implies that the regression function has exactly one independent variable (or regressor on the RHS of the regression function). Assuming that we have $n_1 = 18$ samples (or pairs of responses), then the Residual Sum of Squares of this regression function has $n_1 - 2 = 16$ degrees of freedom (*df*), subtraction of 2 because 2 parameters (slope & intercept) are being estimated. Document this as SS_{RES1}.
2. Repeat the above regression procedure separately for mixes 2 and 3. That is, obtain the regressions of y_2 on x_2 and y_3 on x_3 separately. Suppose mix 2 has $n_2 = 15$ samples (pairs of responses) so that its SS_{RES2} has 13 *df* , and mix 3 has $n_3 = 17$ samples so that its simple-regression function has SS_{RES3} with 15 *df*. Denote the total overall sample size by $n = n_1 + n_2 + n_3$, which for our example will equal to n = 50. Note that none of the above 3 separate simple-regression functions contain the impact of differences in mixes.

3. Then, add the three SS_{RES}'s from the 3 separate simple-regression functions and denote this as SS_{TRES} (i.e., T for the Total RES SS's from all the individual regression functions). For our example SS_{TRES} has 16 + 13 + 15 = 44 *df* .

4. Now pool all the data together and obtain the pooled regression of y on x with the n = 50 pairs of (x, y). Then obtain the pooled simple regression function of y on x and denote its Residual sum of squares as SS_{PRES} (P for pooled). For our example, this SS_{PRES} will have 50 −2 = 48 degrees of freedom, and in general will be larger than $SS_{TRES} = SS_{RES1} + SS_{RES2} + SS_{RES3}$ because SS_{PRES} also contains the effects of differences in SS's of the 3 separate mixes. If fact, if $SS_{PRES} \leq SS_{TRES}$, immediately accept the composite hypotheses H_0: Slope 1 = Slope 2 = Slope 3, and H_0: Intercept 1 = Intercept 2 = Intercept 3, and analyses must be ended (*P-value* will exceed 50%). Otherwise, proceed as follows:

5. Now compute the extra residual SS's due to pooling as $SS_{EXTRA} = SS_{PRES} - SS_{TRES}$. Then form the Sir Ronald A. Fisher F-statistic as

$$F_0 = \frac{MS_{EXTRA}}{MS_{TRES}},$$

where $MS_{EXTRA} = SS_{EXTRA}/(\text{Extra } df) = SS_{EXTRA}/(48 - 44) = SS_{EXTRA}/4$, and $MS_{TRES} = SS_{TRES}/(n-2\times3) = SS_{TRES}/44$. Thus, the numerator *df* of F_0 for our 3 mixes has $\nu_1 = 4 = 3\times2 - 2 = 4$ and its denominator *df* $\nu_2 = 44$. If F_0 exceeds the 5 percentage point of Sir Ronald A. Fisher's F_{ν_1,ν_2}, then reject H_0 and ascertain that the 3 simple regression functions differ convincingly at the 5% level. For our example, the 5-percentage point (or the 95th percentile) of Fisher's F is $F_{0.05,4,44}$ = 2.58366743 so that if F_0 > 2.5837, then reject H_0: Slope 1 = Slope 2 = Slope 3, and also H_0: Intercept 1 = Intercept 2 = Intercept 3. The *P-value* (or Pr level) of the above statistical test is given by $\hat{\alpha} =$ Pr($F_{4,44} \geq F_0$). The smaller this last *P-value* is, the more strongly the composite null hypotheses must be rejected. For real-life situations, this author recommends, perhaps, one should consider setting the prior LOS (Level of Significance) of the test at α equal to 0.10 = 10%, instead of the nominal value of 5%. This will help identify practical differences, in some experiments, if they do exist.

6. Unfortunately, the test of significance outlined above in steps 1 thru 5 does not specify where the differences are once H_0: Slope 1 = Slope 2 = Slope 3, H_0: Intercept 1 = Intercept 2 = Intercept 3 is rejected. The differences could be due any two slopes, any two intercepts, or combinations thereof. In such situations, it will be helpful to obtain the overall Bonferroni confidence intervals for difference in pairs of coefficients. Those 95% Bonferroni confidence intervals which do not enclose zero, their corresponding two coefficients should be declared significantly different at the 5% level (assuming that the prior LOS α = 0.05).

7. In general, for a maximum of 8 mixes the SS_{TRES} will have ν_2 = n −2×8 *df*, the pooled Residual SS's, SS_{PRES} , will have n −2 *df* , while SS_{EXTRA} will have ν_1 = (n−2) –(n −2×8) = 2×8 −2 = 2(8–1) *df*, etc.

Comparing Several Simple Regression Slopes Against a Standard slope

Step 1. Use Minitab to obtain the simple Regression function of y1 on x1; then document the value of Lane 1 slope and its *se* (standard error); repeat this procedure for lanes 2, 3, 4, 5 ,6 ,7 ,8, 9, and 11. The objective is to compare slopes from lanes 2, 3, 4, 5 ,6 ,7 ,8, 9, and 11 against lane 1 slope.

Step 2. Obtain the absolute differences, d_j, between slopes j = 2, 3, 4, 5 ,6 ,7 ,8, 9, 11 and slope 1. Then compute the $se(d_j)$ as follows:

$$se(d_j) = \sqrt{se(d_1)^2 + se(d_j)^2}$$

Step 3. Now obtain the widths $wj = t_{0.025,\nu} \times se(d_j)$, where $\nu = \nu_1 + \nu_j$; if dj exceeds Wj, then declare jth slope significantly different from the standard. The reader should be cognizant that from a statistical standpoint it might be best to use the 5 percentage points of Dunnett in step 2 in lieu of percentage points of Student's t.

For example, the $t_{0.025,14}$ =2.145 while that of D(0.05, 14) =3.07 (2-sided); unfortunately Dunnett's percentage points will make this procedure too conservative and will not often allow for practical differences to be detected. Dunnett's percentage points can be obtained from Appendix VIII of D. c. Montgomery's 8th edition on page 703.