INSY 7300-6 Reference : Chapter 3 of Montgomery(8e) Maghsoodloo

Single Factor Experiments

The objective is to determine if a single factor (or input, or the abscissa) has a significant effect on a response variable Y (the output, or the ordinate). If the input is found to have a significant impact on the output Y, then the 2nd objective is to identify which level of the factor (or input) will optimize the response Y. For an example see Table 3.1 p. 67 of Montgomery's 8th edition, where the factor is the radio frequency power (RFP) setting (in watts) and the output is Y = Etch rate of a tool measured in Å/min (Angstrom/minute).

In this example (the same as Example 3.1 of Montgomery p. 75), the experimenter knows that the range of the input is from 160-220W and uses 4 equispaced levels of this factor, namely 160, 180, 200, and 220W. The experimental layout in Table 3.1 and the randomization order are given on page 75 of Montgomery(8e), where n = 5 replications per treatment and the *N* = 20 total observations are taken completely at random order. Such an experiment is called a Completely Randomized Design (CRD). See also pp. 75-79 of Montgomery's 8th edition.

A procedure called Analysis of Variance (ANOVA) is used to test the equality of 3 or more process means simultaneously W/O affecting the pre-assigned LOS, α, of the test. To develop the ANOVA procedure, we start with the following identity

$$y_{ij} \equiv \mu - \mu + \mu_i - \mu_i + y_{ij} \equiv \mu + (\mu_i - \mu) + (y_{ij} - \mu_i) \equiv \mu + \tau_i + \in_{ij} \qquad (1)$$

This last identity (1) is called the linear additive model (LAM) for the one-factor CRD (completely randomized design). A factor is fixed if its levels are purposely designed (or set) by the experimenter and not selected at random from a population of levels. In such a case, the LAM (1) is called the fixed-effects model. If the levels of the factor are selected at random from a population of levels, then each $\tau_i = \mu_i - \mu$ (i = 1, 2, … *a*) is a random variable and the LAM (1) is called the random-effects model. For the fixed- effects model, $\mu = \sum_{i=1}^{a} \mu_i / a$, while for the random-effects model $\mu = E(\mu_i)$.

ANOVA ASSUMPTIONS

Fixed-Effects: The parameters μ and τ_i are fixed whether the design is balanced (i.e., all n_i = n) or not, and as a result we always have $\sum_{i=1}^{a} \tau_i = 0$. However, the rvs (random variables) $\in_{ij}$ are assumed NID(0, $\sigma^2_{\in}$), where NID stands for normally and independently distributed. Henceforth, we will also use σ^2 in lieu of $\sigma^2_{\in}$. In the case of fixed-effects, conclusions from ANOVA pertain only to the treatments that have been studied and no other levels.

Random-effects: Only μ is a parameter, τ_i 's ~ NID (0, σ^2_{τ}) and independent of $\in_{ij}$ which are N(0, σ^2). Conclusions from ANOVA pertain to the entire population of treatments from which only a random sample of size "*a*" has been studied.

THE FISHER'S F TEST FOR TESTING $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_a = \mu$, ($a \geq 2$) versus the alternative H_1 : At least two μ_i's are significantly different.

First, we replace each parameter in Eq. (1) by its point least-squares estimator as shown below.

$$y_{ij} \equiv \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.}) \equiv \bar{y}_{..} + \hat{\tau}_i + e_{ij}, \qquad (2)$$

where $\hat{\tau}_i \equiv (\bar{y}_{i.} - \bar{y}_{..})$ is called the effect of the i^{th} treatment and $e_{ij} \equiv (y_{ij} - \bar{y}_{i.})$ is called the residual of the $(ij)^{th}$ cell.

Exercise 1. Show that for fixed effects the e_{ij} 's ~ N[0, $(n-1)\sigma^2/n$], and for an unbalanced design $\sum_{i=1}^{a} n_i \hat{\tau}_i \equiv 0$. Further, the $V(y_{ij}) = \sigma^2 = \sigma^2_{\in}$, while the $V(e_{ij}) = (n-1)\sigma^2/n$.

We now rewrite the identity (2) by transposing $\bar{y}_{..}$ to the LHS (left hand side) of the equation:

$$y_{ij} - \bar{y}_{..} \equiv (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.}) \equiv \hat{\tau}_i + e_{ij} \qquad (3)$$

Next we square both sides of Eq. (3), then sum both sides over i and j, and use the fact that the resulting sum of cross-product term vanishes:

$$\sum_{i=1}^{\boldsymbol{a}}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2 \equiv \sum_{i=1}^{\boldsymbol{a}}\sum_{j=1}^{n_i}(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum_{i=1}^{\boldsymbol{a}}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2$$

$$SS_{Total} \equiv \sum_{i=1}^{\boldsymbol{a}}\sum_{j=1}^{n_i}\hat{\tau}_i^2 + \sum_{i=1}^{\boldsymbol{a}}\sum_{j=1}^{n_i}e_{ij}^2 \equiv SS_{Treatments} + SS_{Residuals}, \quad (4)$$

where $SS_{RES} = SS_{Residuals}$. Identity (4) is fundamental to ANOVA because it breaks down the total sum of squares, [$SS_{Total} = SS_T = \sum_{i=1}^{\boldsymbol{a}}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2$], on the LHS into 2 additive components: (1) the 1st term on the RHS of Eq. (4) gives the SS (sum of squares) among (or due to) treatments, and (2) the 2nd term on the RHS of Eq. (4) gives the SS within treatments (or experimental error SS). Further, both SS's on the RHS have independent $(\sigma^2)\chi^2$ distributions under H_0: $\tau_i = \mu_i - \mu = 0$. In fact, SS_{Error}/σ^2 has also a χ^2 distribution under H_1. As a result the null distribution of $\frac{(SS_{Treatment}/\sigma^2)/(\boldsymbol{a}-1)}{(SS_{Error}/\sigma^2)/(\boldsymbol{N}-\boldsymbol{a})} = MS_{Treatments}/MS_{Error}$ is the Fisher's F with $\nu_1 = a - 1$ *df* for the numerator and $\nu_2 = N - a$ *df* for the denominator. Note that Montgomery uses the notation SS_E for SS_{Error}. Letting $N = \sum_{i=1}^{\boldsymbol{a}} n_i$, then the *df* (degrees of freedom) for the identity (4) are as follows:

df: $N - 1 = (a - 1)$ + Error *df* $\rightarrow$ *df* of Error = $N - a$;

If the design is balanced, i.e., $n_i = n$ for all i, then Error or Residual *df* = $\sum_{i=1}^{\boldsymbol{a}}(n_i - 1) = a(n-1)$.

Exercise 2. Show that under the LAM (1) and balanced design $E(SS_{Error}) = a(n-1)\sigma^2 \rightarrow$ An unbiased estimator of σ^2 is given by $MS_{Error} = SS_{Error}/[a(n-1)]$. Further, $E(SS_\tau) = E(SS_{Treatments}) = n\sum\tau_i^2 + (a-1)\sigma^2 \rightarrow E(MS_{Treatments}) = \sigma^2 + n\sum\tau_i^2/(a-1)$. Thus, if H_0 :" zero treatment effects" is true, i.e., all τ_i's are 0, then $MS_{Treatments} = MS_\tau$ is also an unbiased estimator of σ^2.

From Exercise 2 above we deduce that the ratio MS(Treatments)/MS(Error) has the Fisher F distribution with $\nu_1 = a - 1$ and $\nu_2 = N - a$ *df*. Since $E(MS_{Treatments}) = E(MS_{Error}) + \sum_{i=1}^{a} n_i \tau_i^2 / (a-1)$, this constitutes a right-tailed test for the statistic F_0 = MS(Treatments)/ MS(Error). For example, "H_0 : all τ_i's = 0" must be rejected at the 5% LOS if F_0 exceeds the 5 percentage point of $F_{a-1,N-a}$, denoted by $F_{0.05,a-1,N-a}$.

Exercise 3. Show that the computational formulas for SS_T = SS(Total) and Treatment SS are given by $SS_{Total} = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - CF = SS_T$ and $SS(Treatments) = SS_\tau = \sum_{i=1}^{a} n_i(\overline{y}_{i.} - \overline{y}_{..})^2 = \sum_{i=1}^{a}(y_{i.}^2 / n_i) - CF,$ where $CF = y_{..}^2 / N$.

RESIDUALS

For all statistical models, a residual "e" is the difference between the actual observed value and the corresponding value estimated by the model, i.e., $e_{ij} = y_{ij} - \hat{y}_{ij}$. Under the LAM (1), the fitted value $\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\in}_{ij} = \overline{y}_{..} + (\overline{y}_{i.} - \overline{y}_{..}) + 0 = \overline{y}_{i.}$, and hence $e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \overline{y}_{i.}$. Residuals play an important role in model adequacy checking (study pp. 80-84 of Montgomery). However, the equation (3.18) of Montgomery's 8th edition in the middle of page 82 is not accurate because the estimated $V(e_{ij}) \neq MS_{Error}$ but is equal to $v(e_{ij}) = (n-1)MS_{Error}/n$ and hence (3.18) must be revised to

$$r_{ij} = d_{ij} = \frac{e_{ij}}{\sqrt{(n-1)MS_{Error}/n}} \qquad \text{(Revised 3.18, p. 82)}$$

As a result the largest Studentized residual is $r_{32} = d_{32} = 25.6/(0.8 \times 333.7)^{0.5} = 1.56681$, not 1.40 as reported on p. 82 of Montgomery. Further, there seems to exist some discrepancies in Figure 3.6 atop p. 83 of Montgomery's 8th edition.

For testing equality of variances within "*a*" treatments, Study pp. 80-81 of Montgomery on Bartlett's test and verify that indeed q = 0 when all S_i^2 are equal to the same value S^2.

If residual analysis reveals that normality assumption of e_{ij}'s is violated, then data should be transformed according to pp. 87-88 of Montgomery. However, the conclusions from the ANOVA Table pertain only to the transformed data $y^* = y^{\lambda}$.

If the factor is found to have a significant impact on the response Y, then regression can be used to find the best model that describes the relationship between y and the factor within the range of the factor space. The regression models on pages 89-90 of Montgomery(8e) are valid only for $160 \leq x \leq 220$ W, where y represents the etch rate of a tool.

POST-ANOVA For The FIXED-EFFECTS MODEL

If the F-test from the ANOVA table rejects $H_0 : \tau_i = 0$ for all i at the 5% level, then the objective will be to identify which pairs of means are significantly different. If the experimenter decides prior to data gathering which treatments should be compared, then s/he should use orthogonal contrasts; otherwise, after experimentation there are several post-ANOVA multiple comparison procedures that can be applied. Montgomery presents the Scheffe''s Method (pp. 96-97 for contrasts), Tukey's procedure (pp. 98-99), and Dunnett's test for comparison with a control. All these 3 procedures control the overall type I error rate α. Emphasis will be only on Tukey's Studentized Range Test and Dunnett's test because most statistical packages (Minitab) provide these procedures. The explanation in the text (pp. 91-101) is clear, and a short time will be spent in class to review Tukey's multiple comparison's procedure for the data of Montgomery's plasma etching experiment of Example 3.1 given in my notes under ANOVA-Logic.

ORTHOGONAL CONTRASTS (ORCs)

The number of orthogonal contrasts possible is always equal to the *df* of SS(Treatments), or treatment SS. For the Problem 3.10 on page 132 of Montgomery's 8th edition, we can define a maximum of 4 ORCs prior to data gathering. Say, we wish to compare μ_1 vs μ_2, μ_3 vs μ_5, (μ_1

+ $\mu_2)/2$ vs μ_4, and $(\mu_1 + \mu_2 + \mu_4)/3$ vs $(\mu_3 + \mu_5)/2$. Then the 4 corresponding contrasts will be given as follows:

$C_1 = y_{1.} - y_{2.} = -28$, or $C_1 = \bar{y}_{1.} - \bar{y}_{2.}$

$C_2 = y_{3.} - y_{5.} = 34$, or $C_2 = \bar{y}_{3.} - \bar{y}_{5.}$

$C_3 = y_{1.} + y_{2.} - 2y_{4.} = -90$, or $C_3 = \bar{y}_{1.} + \bar{y}_{2.} - \bar{y}_{4.}$

$C_4 = 2y_{1.} + 2y_{2.} - 3y_{3.} + 2y_{4.} - 3y_{5.} = 42$, or $C_4 = 2\bar{y}_{1.} + 2\bar{y}_{2.} - 3\bar{y}_{3.} + 2\bar{y}_{4.} - 3\bar{y}_{5.}$

where $y_{1.} = 49$, $y_{2.} = 77$, $y_{3.} = 88$, $y_{4.} = 108$ and $y_{5.} = 54$. Each contrast carries exactly 1 *df* and their SS's is defined as $SS(C_k) = (C_k)^2 / \sum_{i=1}^{a} n_i c_{ik}^2$, k = 1, 2,..., a –1, $\sum_{i=1}^{a} n_i c_{ik} = 0$, and to guarantee orthogonality it is necessary that $\sum_{i=1}^{a} n_i c_{ik} c_{im} = 0$, for all $k \neq m$, and as a result we will have $\sum SS(C_k) = SS(\text{Treatments})$. Note that Montgomery defines a contrast, C, on page 92 in terms of $\bar{y}_{i.}$, while my definition is in terms of $y_{i.}$; therefore, my above formulas and Montgomery's atop page 89 are the same. We now compute the SS's of the above 4 contrasts.

$SS(C_1) = (-28)^2/[5(1+1)] = 78.4$, $SS(C_2) = 34^2/[5(1+1)] = 115.6$,

$SS(C_3) = (-90)^2/[5(1+1+4)] = 270$, and $SS(C_4) = 42^2/[5(4+4+9+4+9)] = 11.76$.

Note that $\sum_{k=1}^{4} SS(C_k) \equiv SS(\text{Treatments}) = 475.7600$, which is in complete agreement with the ANOVA of Table 3.4 at the bottom of page 71 of Montgomery's 5th edition. It seems that if contrasts are defined in terms of treatment means, then the orthogonality condition should be revised to $\sum_{i=1}^{a} c_{ik} c_{im} / n_i = 0$. Further, the contrast SS is given by $SS(C_k) = (C_k)^2 / \sum_{i=1}^{a} (c_{ik}^2 / n_i)$.

Further, as Montgomery points out in his Eq. (3.26, p.92), the variance of a contrast, such as C_3 above, is computed as $V(C_3) = V(y_{1.} + y_{2.} - 2y_{4.}) = V(y_{1.}) + V(y_{2.}) + 4V(y_{4.}) = 5\sigma_\epsilon^2 + 5\sigma_\epsilon^2 + 5\times 16\sigma_\epsilon^2 =$ $(\sum_{i=1}^{a} n_i c_{ik}^2)\sigma_\epsilon^2$, i.e., $(C_k)^2 / (\sigma_\epsilon^2 \sum_{i=1}^{a} n_i c_{ik}^2) = Z^2$ and hence, $(C_k)^2 / (MS_{Error} \sum_{i=1}^{a} n_i c_{ik}^2) = t_{\nu_2}^2 = F_{1,\nu_2}$.

THE RANDOM-EFFECTS MODEL

This model is valid only when the "*a*" levels of a factor are randomly selected from a

population of levels (e.g., selecting 6 operators at random from a population of 100 operators working in a plant). Operators would then form a random qualitative factor and the conclusions drawn from the ANOVA Table will pertain to the entire population of 100 operators. Note that there will not be any interest whatsoever in determining which 2 of the actually selected 6 treatments (or operators) are significantly different (i.e., Tukey's SRT and ORCs will not be applicable) but rather the objective is to determine if there is significant variation in the entire population of treatments (or operators). Therefore, our null hypothesis is $H_0: \sigma_\tau^2 = 0$ versus $H_1: \sigma_\tau^2 > 0$, which again is clearly a right-tailed test. The ANOVA Table is obtained exactly as in the case of fixed-effects, but the post-ANOVA procedure is quite different. In the case of random-effects model, if the F-test rejects the null hypothesis that $\sigma_\tau^2 = 0$, i.e., F_0 exceeds the threshold value of F_{α,ν_1,ν_2}, then it is essential to estimate the components of variance in the LAM : $y_{ij} = \mu + \tau_i + \in_{ij}$, where $\mu = E(y_{ij})$ for all i = 1, 2, ..., *a*, and j = 1, 2, ..., n.

Note that we are considering only the simpler balanced case (i.e., $n_i = n$ for all i) for the random-effects model. Further, τ_i's are assumed NID(0, σ_τ^2) so that the constraint $\sum_{i=1}^{a} \tau_i = 0$ no longer applies to the random-effects model, but $E(\tau_i) = 0$ for all i, which implies that $E(\mu_i) = \mu$. The experimental errors, $\in_{ij}$'s, are as before NID(0, σ^2) and independent of τ_i's. This leads to $V(y_{ij}) = \sigma_\tau^2 + \sigma^2$, which shows that there are 2 components of variance.

For an example of obtaining the ANOVA Table for a random-effects model, see the Example 3.11, pp. 119-122 of Montgomery's 8th edition. In order to estimate the 2 components of variance, $\sigma_\in^2$ and σ_τ^2, we must derive $E(MS_E)$ and $E(MS_{Treatments})$. As in the case of fixed-

effects, it can easily be verified that $E(MS_E) = \sigma^2$. Therefore, for the Example 3.11 of Montgomery, $\hat{\sigma}^2_{\in} = \hat{\sigma}^2$ = MS(Error) = MS_{PE} = 1.896 = MS_E, where PE stands for pure error.

We now prove that for the random-effects model, under the LAM (1), $E(MS_{Treatments}) = E(MS_\tau) = n\sigma^2_\tau + \sigma^2$. **Proof.** $E(MS_{Treatments}) = E\sum_{i=1}^{a}\sum_{j=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2/(a-1)$

$$= nE\sum_{i=1}^{\boldsymbol{a}}(\bar{y}_{i.} - \bar{y}_{..})^2/(\boldsymbol{a}-1) = \frac{n}{a-1}E\sum\left[(\mu+\tau_i+\bar{\varepsilon}_{i.}) - (\mu+\bar{\tau}+\bar{\varepsilon}_{..})\right]^2$$

$$= \frac{n}{\boldsymbol{a}-1}E\sum\left[(\tau_i-\bar{\tau}) + (\bar{\varepsilon}_{i.}-\bar{\varepsilon}_{..})\right]^2 = \frac{n}{\boldsymbol{a}-1}\left[E\sum(\tau_i-\bar{\tau})^2 + E\sum(\bar{\varepsilon}_{i.}-\bar{\varepsilon}_{..})^2\right]$$

$$= \frac{n}{\boldsymbol{a}-1}\left[(a-1)\sigma^2_\tau + (a-1)\sigma^2_{\bar{\varepsilon}}\right] = n\sigma^2_\tau + \sigma^2 = E(MS_\tau).$$

As before an unbiased estimator of σ^2 is clearly MS_E, and the above proof shows that $E(MS_{Treatments}) - \sigma^2 = n\sigma^2_\tau$. Inserting $E(MS_E)$ for σ^2 in this last equation results in $E(MS_{Treatments}) - E(MS_E) = n\sigma^2_\tau$, or $E[(MS_{Treatments} - MS_E)/n] = \sigma^2_\tau$. This last equality clearly shows that an unbiased estimator of σ^2_τ is indeed $(MS_{Treatments} - MS_E)/n$. You should now be in a good position to understand the point estimates obtained near the bottom of p. 119 of Montgomery(8e) for the data of Example 3.11, where $MS(Treatments) \cong n\hat{\sigma}^2_\tau + \hat{\sigma}^2 \rightarrow \hat{\sigma}^2_\tau = \frac{MS_{Treatments} - MS_{Error}}{n} = \frac{MS_{Treatments} - MS_{RES}}{n} = \frac{29.72917 - 1.89583}{4} \cong 6.95833 \rightarrow \hat{\sigma}^2_y = \hat{\sigma}^2_\tau + \hat{\sigma}^2_{\in} = 6.95833 + 1.89583 = 8.85417$. The % of variation attributed to all the looms in the textile company is estimated to be approximately 78.59%.

Exercise 4. Work problems 3.30 & 3.31 on page 135, coding the data by 23 and 480, respectively.

GENERAL REGRESSION SIGNIFICANCE TEST APPROACH TO ANOVA (GRST)

If DOX (design of experiment or experiments) involves 2 or more factors and is unbalanced (n_i's are different at different factor level combinations), then the procedures presented so far for computing SS's will no longer be valid. In fact, when there are 2 or more factors, some SS's will always become negative, if the standard ANOVA procedure for SS's computations is used! Thus, in the case of 2 or more factors impacting the response, resort must be made to the GRST approach which is very complicated and cumbersome at best. Fortunately, most statistical packages (such as SAS, Minitab and SPSS) have a built-in routine to obtain the ANOVA table for an unbalanced factorial experiments involving more than one factor (such as Minitab's GLM = General Linear Models). Before presenting the GRST approach, for the readers understanding, we should show that $SS_{Error} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - R(\mu, \tau)$, where $R(\mu, \tau)$ is called the reduction of the USS = $\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2$ due to fitting the LAM "$\mu + \tau_i + \in_{ij}$" to the y_{ij}'s.

Exercise 5. Show that the $R(\mu, \tau) = \hat{\mu}\ y_{..} + \sum_{i=1}^{a} \hat{\tau}_i\ y_{i.}$.

Hint: $SS_{Error} = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_{i.})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \hat{\tau}_i - \overline{y}_{..})^2$, where we have made use of the fact that $\hat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..}$. Thus, $SS_{Error} = \sum_{i=1}^{a}\sum_{j=1}^{n}[(y_{ij} - \overline{y}_{..}) - \hat{\tau}_i)]^2$ = … = $SS_{Total} - \sum_{i=1}^{a} \hat{\tau}_i\ y_{i.}$ + $n\,\overline{y}_{..} \times \sum_{i=1}^{a}\hat{\tau}_i$, but the $\sum_{i=1}^{a}\hat{\tau}_i \equiv 0$ so that the desired result now follows. If the n_i's are not equal, then the constraint changes to $\sum_{i=1}^{a} n_i\hat{\tau}_i = 0$. The conclusions that you must draw from this exercise is that the $SS(Total) = SS_{Error} + \sum_{i=1}^{a} \hat{\tau}_i\ y_{i.}$, which clearly shows that $SS(Treatments) = SS_\tau = \sum_{i=1}^{a} \hat{\tau}_i\ y_{i.}$. Further, $SS_{Error} = USS - CF - \sum_{i=1}^{a} \hat{\tau}_i\ y_{i.} = USS - \hat{\mu}\, y_{..} - \sum_{i=1}^{a} \hat{\tau}_i\ y_{i.}$, which shows that $R(\mu, \tau)$

$= \hat{\mu}\, y_{..} + \sum_{i=1}^{a} \hat{\tau}_i\, y_{i.}$.

THE GRST APPROACH

Our first objective is to estimate the parameters μ and the vector $\tau = [\tau_1 \quad \tau_2 \ldots \quad \tau_a]'$ such that the least squares function (LSF)

$$\text{LSF} = L(\mu, \tau_i) = \sum_{i=1}^{a}\sum_{j=1}^{n} \in_{ij}^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \mu - \tau_i)^2$$

is minimized wrt the (a +1) unknown parameters μ and τ. To accomplish this task, we must require that the partial derivatives of the LSF wrt the (a+1) parameters is zero.

$$\frac{\partial(\text{LSF})}{\partial\mu} = 2\sum_{j=1}^{n}\sum_{i=1}^{a} (y_{ij} - \mu - \tau_i)(-1) \xrightarrow{\text{set to}} = 0$$

$$\frac{\partial(\text{LSF})}{\partial\tau_k} = \sum_{j=1}^{n} \frac{\partial}{\partial\tau_k} \sum_{i=1}^{a} (y_{ij} - \mu - \tau_i)^2$$

$$= \sum_{j=1}^{n} 2(y_{kj} - \mu - \tau_k)(-1) \xrightarrow{\text{set to}} = 0, \text{ for } k = 1, 2, \ldots, a.$$

Each partial derivative above when set equal to 0 leads to exactly one LS (least squares) normal equation so that we have a system of (a+1) equations with (a+1) unknowns $\hat{\mu}$ and $\hat{\tau}_i$ (i = 1, 2, …, a), which are listed on pages 125-126 of Montgomery's 8th edition. As Montgomery notes, the (a+1) equations are not independent because the sum of the last "a" equations give rise to the 1st. However, for a balanced design we have the constraint that $\sum \hat{\tau}_i = 0$, and this leads to a set of unique solutions $\hat{\mu} = \overline{y}_{..}$ and $\hat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..}$, i = 1, 2, …, a. From the Exercise 5 on the previous page, it now follows that the contribution of the treatment i to SS(Treatments) is simply the product of $\hat{\tau}_i$ by the corresponding subtotal $y_{i.}$ on the RHS of the (i +1)th LS normal equation.

Before providing an example to illustrate the GRST approach, we should emphasize that there is generally a simpler method of obtaining the LSNEs for most ANOVA models. For the balanced LAM (1), to obtain the LSNE for μ, simply add both sides of Eq. (1) over all i and j

and use the assumptions that $\sum_{i=1}^{a} \tau_i = 0$, $\sum\sum \in_{ij} = 0$ and place a hat "^" on the remaining parameters. This leads to

$$\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij} = \sum\sum(\mu + \tau_i + \in_{ij}) \qquad \rightarrow \quad N\,\hat{\mu} = y_{..}$$

which is the 1st LSNE on page 125 of Montgomery. To obtain the LSNE for $\hat{\tau}_i$, keep the index i fixed and sum both sides of the LAM(1) from j = 1 to j = n:

$$\sum_{j=1}^{n} y_{ij} = \sum_{j=1}^{n}(\mu + \tau_i + \in_{ij}) \qquad \rightarrow \qquad n\hat{\mu} + n\hat{\tau}_i = y_{i.},$$

which for i = 1, 2, …, *a*, gives the last "*a*" LSNEs on pages 125-126 of Montgomery. To obtain the F-test, 1st fit the LAM: $\mu + \tau_i + \in_{ij}$ to the y_{ij}'s; then the reduction of the USS = $\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2$ resulting from fitting this model (as was shown in Exercise 5 of my notes) is given by $R(\mu, \tau) = \hat{\mu}\, y_{..} + \sum_{i=1}^{a} \hat{\tau}_i\, y_{i.}$.

Second, hypothesize that $H_0 : \tau_i = 0$ for all i, which reduces the LAM to $y_{ij} = \mu + \in_{ij}$. Now fit this reduced model $\mu + \in_{ij}$ to the data. Summing the reduced LAM over both i and j results in one LSNE: $y_{..} = N\hat{\mu}$. Therefore, the reduction of the USS = $\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2$ due to fitting the reduced model $y_{ij} = \mu + \in_{ij}$ is $R(\mu) = \hat{\mu}\, y_{..} = \bar{y}_{..}\, y_{..} = y_{..}^2 / N$ = the CF. Consequently, the net reduction of USS due to adding the τ_i's to the reduced model $y_{ij} = \mu + \in_{ij}$ must be $R(\mu, \tau) - R(\mu)$

$$= \sum_{i=1}^{a} \hat{\tau}_i\, y_{i.} = \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})\, y_{i.} = \sum_{i=1}^{a} (\bar{y}_{i.} y_{i.} - \bar{y}_{..}\, y_{i.}) = (\sum_{i=1}^{a} y_{i.}^2 / n) - \bar{y}_{..}\, y_{..} = (\sum_{i=1}^{a} y_{i.}^2 / n) - CF = SS_\tau.$$

As an example of the GRST approach, we work PROBLEM 3.27 on page 135 of Montgomery, and to make calculations less cumbersome, we change the value of y_{33} from 55.4 to y_{33} = 55.6; further, in order to ease computations we code the data by subtracting 50 from all N = 16 observations. This leads to the coded $y_{..} = 72.00$, $y_{1.} = 34.5$, $y_{2.} = 23.1$, $y_{3.} = 9.9$, $y_{4.} = 4.5$, USS = 444.64, and the CF = $72^2 /16 = 324$.

To apply the GRST procedure to this unbalanced one-factor experiment (n_1 = 5, $n_2 = n_4 = 4$ and $n_3 = 3$), we 1st sum the LAM: $\mu + \tau_i + \in_{ij} = y_{ij}$ over the 4 treatments and all the observations within treatments, leading to the LSNE for

$$\mu: \qquad 16\hat{\mu} + 5\hat{\tau}_1 + 4\hat{\tau}_2 + 3\hat{\tau}_3 + 4\hat{\tau}_4 = y_{..}$$

Next, in order to obtain the LSNE for $\hat{\tau}_1$, we put i = 1 in the LAM and sum over j = 1 to j = 5, and assume $\sum_{j=1}^{5} \hat{\in}_{1j} = 0$. (Note that $\mu + \tau_1 + \in_{1j} = y_{1j}$)

$$\tau_1: \; 5\hat{\mu} + 5\hat{\tau}_1 = y_{1.}$$

To obtain the LSNE for τ_2, we put i = 2 in the LAM and sum

from j = 1 to j = 4. $\rightarrow$ $\tau_2: \quad 4\hat{\mu} + 4\hat{\tau}_2 = y_{2.}$

In a similar manner we obtain the equations for τ_3 and τ_4 :

$$\tau_3: \quad 3\hat{\mu} + 3\hat{\tau}_3 = y_{3.}, \qquad \tau_4: \quad 4\hat{\mu} + 4\hat{\tau}_4 = y_{4.}$$

The above system of 5 equations with 5 unknowns are summarized below:

$$16\hat{\mu} + 5\hat{\tau}_1 + 4\hat{\tau}_2 + 3\hat{\tau}_3 + 4\hat{\tau}_4 = 72.0$$

$$5\hat{\mu} + 5\hat{\tau}_1 = 34.5$$

$$4\hat{\mu} + 4\hat{\tau}_2 = 23.1$$

$$3\hat{\mu} + 3\hat{\tau}_3 = 9.9$$

$$4\hat{\mu} + 4\hat{\tau}_4 = 4.5$$

But the above heterogeneous system has a determinant of 0 because the sum of the last 4 equations give rise to the 1st and hence there is no unique solution. In order to obtain a unique solution, we can impose either the constraint $\sum_{i=1}^{4} \hat{\tau}_i = 0$, or the constraint $\sum_{i=1}^{4} n_i \hat{\tau}_i = 0$. Both constraints lead to the same exact value of $R(\mu, \tau)$ but different parameter estimates! We arbitrarily select the constraint $\sum_{i=1}^{4} \hat{\tau}_i = 0$ and in the next exercise you will use the other constraint. Therefore, we now have a system of 5 independent equations with 5 unknowns as listed below:

$$\hat{\tau}_1 + \hat{\tau}_2 + \hat{\tau}_3 + \hat{\tau}_4 = 0$$

$$5\hat{\mu} + 5\hat{\tau}_1 = 34.5$$

$$4\hat{\mu} + 4\hat{\tau}_2 = 23.1$$

$$3\hat{\mu} + 3\hat{\tau}_3 = 9.90$$

$$4\hat{\mu} + 4\hat{\tau}_4 = 4.50$$

I wrote a simple Matlab program to solve the above system (or you may use MS Excel by inverting the matrix of the system and multiplying it by the column vector [34.5 23.1 9.90 4.50 0]'). The Matlab codes use Cramer's Rule to obtain the solution which are $\hat{\mu} = 4.275$, $\hat{\tau}_1 = 2.625$, $\hat{\tau}_2 = 1.50$, $\hat{\tau}_3 = -0.9750$, and $\hat{\tau}_4 = -3.150$. Thus, $R(\mu, \tau) = \hat{\mu} \times y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} = 409.185 \rightarrow$ SS(Error) = USS $-$ R(μ, τ) = 444.64 $-$ 409.185 = 35.455, which is in exact agreement with the SS(Error) of the Minitab program output that I used to verify the above results. Note that if we first square the identity (2), $y_{ij} \equiv \bar{y}_{..} + \hat{\tau}_i + e_{ij}$ given on page 11, and then sum from j = 1 to n_i and i = 1 to *a*, we obtain $USS \equiv \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i} (\bar{y}_{..} + \hat{\tau}_i + e_{ij})^2 =$

$\sum_{i=1}^{a}\sum_{j=1}^{n_i} (\bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n_i} \hat{\tau}_i^2 + \sum_{i=1}^{a}\sum_{j=1}^{n_i} e_{ij}^2$ because all 3 cross-product terms vanish. Therefore, the magnitude of USS is first due to the correction factor $CF = N(\bar{y}_{..})^2$, next is due to the $\sum_{i=1}^{a}\sum_{j=1}^{n_i} \hat{\tau}_i^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i} \hat{\tau}_i(\bar{y}_{i.} - \bar{y}_{..}) = \sum_{i=1}^{a}\sum_{j=1}^{n_i} \hat{\tau}_i \bar{y}_{i.} = \sum_{i=1}^{a} \hat{\tau}_i y_{i.}$ = $SS_{Treatments}$, and the rest is due to SS_{RES} = $\sum_{i=1}^{a}\sum_{j=1}^{n_i} e_{ij}^2$. Finally, $SS_{Treatments}$ can also be obtained by first hypothesizing that H_0: $\tau_i = 0$ for all i and using the model $y_{ij} \equiv \mu + \in_{ij}$ to obtain the LS estimate of μ. This yields y.. = N $\hat{\mu}$, or $\hat{\mu} = \bar{y}_{..}$ = 4.50, resulting in R(μ) = y.. $\bar{y}_{..}$ = 72×4.50 = 324.00. Thus, SS_{Tr} = R(μ, τ)– R(μ) = 409.185– 324 = 85.1850, as before.

Exercise 6. Repeat the above analysis using the 2nd constraint $\sum_{i=1}^{4} n_i \hat{\tau}_i = 0$.

Errata for Chapter 3 of Montgomery's 8th edition

1. Page 77, in the middle the page change the Eq. (3.18) to d_{ij} = $\frac{e_{ij}}{\sqrt{(n-1)MS_E / n}}$ and change d_1 to d_{32} and e_1 to e_{32}.

2. Page 83, in Figure 3.6 the abscissa value 500.15 should be 587.40, and 629.10 should be 625.4.

3. Page 86, in Table 3.7, change S_2 to 1.192, S_3 to 1.6471, and S_4 to 2.801.

4. Page 94, If contrasts are defined in terms of treatment means instead of subtotals ($y_{i.}$'s), then the orthogonality condition near the bottom of page 94 of Montgomery's 8th edition must be changed to $\sum_{i=1}^{a}(c_i d_i / n_i) = 0$.

5. Page 125, line 4 from the bottom of the page change the least squares normal equation $-2\sum_{j=1}^{n}(y_{ij} + \hat{\mu} - \hat{\tau}_i) = 0$ to $-2\sum_{j=1}^{n}(y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0$, i = 1, 2, …, *a*.