2015S

Fountain*, Crain

2015F

2016S

Fountain, Tableman*

2016F

Jong Sung Kim*, Brad Crain

2017S

Brad Crain, Jong Sung Kim*

2017F

Robert Fountain*, Daniel Taylor-Rodriguez

2018S

Robert Fountain*, Daniel Taylor-Rodriguez

2018S1

A company has developed three specialized training workshops for their employees. Each of the 12 employees is randomly assigned to one of the three workshops. The company would like to develop a model that could be used to predict each employee's performance score (a number from 0 to 100) based on their attendance at the workshops. The previous year's performance score is also available for use as a predictor. The following table shows and ordered pair for each employee, consisting of the current performance score and the previous year's score.

a) Write an appropriate model for the situation described above, allowing for different slopes and different intercepts for the three workshops. Hint: there are 12 observations, and you will need to use at least one indicator variable.
b) Write the matrix form of the appropriate model. Show the contents and dimensions of all matrices.
c) Suppose that you wish to test for equality of the three slopes. Write the matrix form of the reduced model. What will be the numerator and denominator degrees of freedom for the additional sum of squares F test?

2018S2

A company that produces textiles is trying to determine if the final quality is determined by production site (A), machine operator (B), and thread type (C). The company operates 3 production sites, and all 3 will participate in the experiment. At each of the 3 sites, 5 machine operators will be randomly selected. The company uses two different types of thread. Each of the operators will produce 3 samples of cloth using each type of thread, yielding a total of $3 \times 5 \times 2 \times 3 = 90$ observations. a) State which effects are fixed at which effects are random. b) State which effects are nested within others and which effects are crossed. c) Create an abbreviated ANOVA table that has two columns: one column that lists the effects in the model (including the appropriate main effects, interactions, and nested effects), and a second column that gives the degrees of freedom for each item in the first column.

2018S3

The multiple linear regression model $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \varepsilon_i$ was fit to a data set of 75 observations. The regression SS's (SSR) were partitioned sequentially into the following:
$SSR(X_1) = 108$
$SSR(X_2|X_1) = 163$
$SSR(X_3|X_1X_2) = 29$
$SSR(X_4|X_1X_2X_3) = 41$
$SSR(X_5|X_1X_2X_3X_4) = 26$
The model $y_i = \beta_0 + \beta_1 x_{i1} + \beta_3 X_{i3} + \beta_5 X_{i5} + \varepsilon_i$ was also fit to the same data and the following ANOVA was calculated:
Source SS Regression 214
Residual Error 489
Total 703
Answer the following from the above information:

(a) Calculate the F-statistic for testing the hypothesis (H_0) that X_3, X_4, and X_5 have no significant effect on the response Y.
(b) Calculate R^2 for the model $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 X_{i2}\varepsilon_i$
(c) Calculate the R^2_{adj} for the model in part (b).
(d) Calculate the F-statistic for testing $H_0 : \beta_2 = \beta_4 = 0$.

2018S4

2019S1
Assume the model $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$, $i = 1, .., n$ with the additional restrictions that $\beta_1 = 0$, $\beta_0 = 2\beta_2$. Find the least-squares estimators of β_0, β_1, and β_2.

2018F

Robert Fountain*, Daniel Taylor-Rodriguez

2018F1

The weights (y_i, kilograms) and corresponding heights (x_i, centimeters) of 10 randomlysampled adolescents (i= 1,..,10) are recorded, and the following summary statistics are computed:
$\sum_{i=1}^{10}(x_i - \bar{x})^2 = 472$,$\sum_{i=1}^{10}(y_i - \bar{y})^2 = 731$, $\sum_{i=1}^{10}(x_i - \bar{x})(y_i - \bar{y}) = 274$
You will perform a simple linear regression of weight on height, under the usual assumption of independent, identically distributed, normal errors.

a) Compute the least squares estimates for the intercept and slope parameters.
b) Compute the usual unbiased estimate of the error variance.
c) Compute unbiased estimates of the variances of the least squares estimates in part (a).
d) Perform a two-sided test for whether or not height and weight are related (assuming the simple linear regression model holds). State the null and alternative hypotheses, and use $\alpha = 0.05$.
e) Compute 95% simultaneous two-sided confidence intervals for the intercept andslope parameters, using the Bonferroni method.

2018F2

City planners are evaluating the effectiveness of a new "intelligent" traffic control system in reducing the amount of time motorists must spend on city streets. A total of 24 simulations are run: 4 simulations for each of the 6 combinations of control system (old or new) and traffic intensity (light, moderate, or heavy). All simulations use different random seeds, the combinations are run in a completely random order, and the median travel time (minutes) is recorded for each simulation. For each combination, the following table gives the average and sample standard deviation of the median travel times from the 4 simulations assigned that combination:

a) Write a (univariate) linear model equation of the usual full form for data from this experiment, with median travel time as the response. Explain each term and specify any conditions it satisfies. What crucial assumption are you making about the error variances?
b) Produce an ANOVA table with all appropriate sources of variation, including the (corrected) total. Include sums of squares, degrees of freedom, and appropriate mean squares.
c) Test whether your model in part (a) may be reduced to a model in which the effects of system and traffic intensity are purely additive. Remember to state the null and alternative hypotheses. Use $\alpha = 0.05$.
d) Form a two-sided 95% confidence interval for the difference in median travel time between the new system and the old system under moderate traffic conditions.

2018F3

Consider the linear mixed model

2018F4

Consider a randomized complete block design with 12 blocks and a single treatment factor having 3 levels. Let X_{ij} denote the response measured for an experimental unit in block j that receives treatment i for $i = 1, 2, 3$ and $j = 1, .., 12$. Suppose there is also a covariate whose value X_{ij} is measured for each experimental unit.
The following four models are fit to the data (using least squares), with the resulting residual (error) sums of squares as specified:
Model 1: $Y_{ij} = \mu + \gamma_j + \varepsilon_{ij}$ $SS(Res) = 660$
Model 2: $Y_{ij} = \mu + \alpha_i + \gamma_j + \varepsilon_{ij}$ $SS(Res) = 550$
Model 3: $Y_{ij} = \mu + \alpha_i + \gamma_j + \beta x_{ij} + \varepsilon_{ij}$ $SS(Res) = 300$
Model 4: $Y_{ij} = \mu + \gamma_j + \beta x_{ij} + \varepsilon_{ij}$ $SS(Res) = 420$
The treatment effects are $\alpha = (\alpha_1, \alpha_2, \alpha_3)'$ and the block effects are $\alpha = (\gamma_1, \gamma_2, .., \gamma_{12})'$. The corrected total sum of squares is 820.

a) Find the sequential sums of squares for γ_j, α_i, and β, in that order.
b) Form an ANOVA table for the randomized complete block design without the covariate X_{ij}, that is, based on Model 2. The table should include all appropriate sources of variation (including the corrected total), with degrees of freedom, sums of squares, and mean squares where appropriate. Then test whether or not there is any treatment effect based on this model. Use $\alpha = 0.05$.
c) Test whether there is any treatment effect, after accounting for both blocking and the covariate. Use $\alpha = 0.05$.
d) Suppose the (possibly incorrect) model $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$ is fit to the data. Compute the residual sum of squares for this model.

2019S

Robert Fountain*, Daniel Taylor-Rodriguez

2019S1

2018S4
Assume the model $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$, $i = 1, .., n$, with the restriction that $\beta_0 = 0$. Find the least-squares estimators of the regression coefficients.

2019S2

A manufacturer wishes to know if operator (A), material (B), and heat (C) affect the outcome of the product being produced. All of the machines being used operate at the same four heat settings. Five operators are randomly chosen. There are fifteen varieties of material that can be used, and three of these are assigned to each of the five operators. For each operator, that means that there are 12 combinations of heat and material that can be used. The operator produces two replications of each of these, for a total of $5 \times 3 \times 4 \times 2 = 120$ observations. Incorrectly treating all main effects as fixed and all combinations of factors as crossed, the statistician produces the ANOVA table shown below.
Produce the corrected ANOVA table, including the expected mean squares and the correct F statistics. If an exact F test is not available, construct an approximate F statistic (for which you need not compute the degrees of freedom).

2019S3

A company has developed two possible manufacturing processes. They will produce 5 items with each process, in a completely random order. Then they will measure the quality (Y) of each item on a 100-point scale. They suspect that the relative humidity during production might affect the outcome, so they also record it (X). They would like to develop a model that could be used to predict the outcome quality based on the process, with the relative humidity as a covariate. The following table shows and ordered pair for each item, consisting of the quality score and the humidity measurement.

a) Write an appropriate model for the situation described above, allowing for different slopes and different intercepts for the two processes. Hint: there are 10 observations, and you will need to use at least one indicator variable.
b) Write the matrix form of the appropriate model. Show the contents and dimensions of all matrices.
c) Suppose that you wish to test for equality of the two slopes. Write the matrix form of the reduced model. What will be the numerator and denominator degrees of freedom for the additional sum of squares F test?