MIDTERM

Statistics 207

Winter Quarter, 2016

You must show all work in order to get credit. Notations such as $\bar{Y}_{i\cdot}, \bar{Y}_{\cdots}$, SSA, SSE, MSA are the same as those used in the handouts. If anything is unclear, please ask for clarifications.

1. An article in the journal Industrial Quality Control describes an experiment to investigate the effect the type of glass (factor A) and the type of phosphor (factor B) on the brightness of a television tube. The response variable is the current necessary (in microamps) to obtain a specified brightness level. The experiment had 18 units with three at each combination of type of glass and type of phosphor. A table of sample means and some summary statistics are given below.

	Phosphor			
Glass	$j = 1$	$j = 2$	$j = 3$	Row mean
$i = 1$	285.00	301.67	288.33	291.67
$i = 2$	235.00	245.00	225.00	235.00
Column mean	260.00	273.33	256.67	$\bar{Y}_{\cdots} = 263.33$

$SSTO = 16150.0, SSE = 633.3,$

$\sum(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdots})^2 = 1605.56, \sum(\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdots})^2 = 155.55.$

(a) Plot the means in order to investigate if interaction effects are present. Does this plot reveal anything about the main effects of the factors? What are your conclusions?

(b) Set up the ANOVA table (including the F-statistics).

(c) Carry out tests (at $\alpha = 0.05$) to decide if the interaction effects and main effects of the factors are present. State your conclusions. Are they consistent with your assessments in part (a)?

(d) The investigator also wanted to fit an additive model to this data. Set up the ANOVA table (including the F-statistics) assuming an additive model.

2. A large company employs a large number of personnel officers who interview job applicants. At the end of the interview, the personnel officer assigns a rating between 0 and 100 to indicate the applicant's potential value on the job. In a study, five personnel officers were randomly selected, and each was assigned four applicants at random. Note that in this study there were a total of 20 applicants. A summary of the data is given below:

$\bar{Y}_{1\cdot} = 75$, $\bar{Y}_{2\cdot} = 70.5$, $\bar{Y}_{3\cdot} = 54.75$, $\bar{Y}_{4\cdot} = 77.25$, $\bar{Y}_{5\cdot} = 77.75$,

$\sum(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 = 364.3075$, $\sum\sum(Y_{ij} - \bar{Y}_{i\cdot})^2 = 1099.3$.

[Here Y_{ij} is the rating by the i^{th} personnel officer for the j^{th} applicant, $j = 1, ..., 4, i = 1, ..., 5$]

(a) Write down an ANOVA model that is appropriate for the analysis of this data. Carry out a test to decide if there is variability among personnel officers at a $\alpha = 0.05$ level of significance.

(b) Obtain 95% confidence interval for the mean rating ($\mu_{\cdot}$) of all personnel officers in this company.

(c) Obtain a 95% confidence interval for the proportion of variability in ratings due to personnel officers.

(d) Obtain the best linear unbiased predictor for α_3, the random effect associated with personnel officer 3 in the sample.

3. Consider a two factor study where Y_{ij} is the observation when factor A is at level i and factor B is at level j, $j = 1, \ldots, b, i = 1, \ldots, a$. An additive model with normal errors is considered appropriate here. For parts (a-b), assume that the factors are fixed.

(a) Show $E(MSA) = \sigma^2 + b\sum\alpha_i^2/(a-1)$, where MSA is the mean square error of factor A.

(b) Let $L = \sum c_i\alpha_i$ be a contrast in the main effects of factor A. Obtain the estimate $\hat{L}$ of L and find its distribution after deriving its mean and variance in terms of the model parameters. Obtain an estimate of $Var(\hat{L})$

For parts (c-d) assume that factor A is fixed but factor B is random.

(c) Let $L = \sum c_i\alpha_i$ be a contrast in the main effects of factor A. Obtain the estimate $\hat{L}$ of L and find its distribution after deriving its mean and variance in terms of the model parameters. Obtain an estimate of $Var(\hat{L})$.

(d) Show that $E(MSB) = \sigma^2 + a\sigma_\beta^2$, where σ_β^2 is the variance the main effects of factor B. Use this result to obtain an estimate of σ_β^2.

4. The impurity present in a chemical product is affected by two factors - temperature (factor A) and pressure (factor B). An experiment was carried out with three different levels of temperature and four different levels of pressure, and the impurities were measured. Note that there is only one observation for each combination of temperature and pressure. The data set is given below.

	Pressure					
Temperature	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	Row mean
$100^o F$ $(i = 1)$	5	4	6	3	5	4.6
$125^o F$ $(i = 2)$	3	1	4	2	3	2.6
$150^o F$ $(i = 3)$	1	1	3	1	2	1.6
Column mean	3	2	4.33	2	3.33	$\bar{Y}_{..} = 2.93$

$\sum\sum(Y_{ij} - \bar{Y}_{..})^2 = 36.93$, $\sum(\bar{Y}_{i.} - \bar{Y}_{..})^2 = 4.67$, $\sum(\bar{Y}_{.j} - \bar{Y}_{..})^2 = 3.87$,
$\sum\sum Y_{ij}(\bar{Y}_{i.} - \bar{Y}_{..})(\bar{Y}_{.j} - \bar{Y}_{..})^2 = 1.33$

For parts (a-c), assume an additive model (i.e., a model without interaction).

(a) Set up the analysis of variance table (including the F-statistics).

(b) Is the impurity affected by temperature? Is the impurity affected by pressure? Answer these questions by carrying out appropriate tests at a 0.01 level of significance. Find the p-values of your tests.

(c) Use Tukey's method to compare the mean impurities at the three temperatures. What are your findings?

(d) Carry out Tukey's test at a 0.05 level of significance to determine if an additive model is reasonable for this data.