Preliminary Examination - Multivariate Models

August 16, 2021, Monday, 1:00-3:30pm

Answer all questions and show all your work.

1. A state highway department studied the wear characteristics of five different paints on eight locations in the state. The standard, currently used paint (paint 1) and four experimental paints (paints 2, 3, 4, 5) were included in the study. The eight locations were randomly selected, thus reflecting variations in traffic densities throughout the state. At each location, a random ordering of the paints to the chosen road surface was employed. After a suitable period of exposure to weather and traffic, a combined measure of wear, considering both durability and visibility, was obtained. The data on wear follow.

		Location (i)							
		1	2	3	4	5	6	7	8
	1	11	20	8	30	14	25	43	13
	2	13	28	10	35	16	27	46	14
Paint (j)	3	10	15	8	27	13	26	41	12
	4	18	30	16	41	22	33	55	20
	5	15	18	12	28	16	25	42	13

Source	df	SS	MS	F
Blocks (Location)		4826.375		
Paint types		531.350		
Error				
Total		5479.975		

a) State an appropriate statistical model including model assumptions.

b) Obtain the complete ANOVA table except the column of F.

c) Test whether or not the mean wear differs for the five paints; Use $\alpha = 0.05$.
State the alternatives, decision rule, and conclusion.

d) Paints 1, 3, and 5 are white, whereas paints 2 and 4 are yellow. Estimate the difference in the mean wear for the two groups of paints with a 95% confidence interval. Interpret your findings.
($\bar{Y}_{.1} = 20.5, \bar{Y}_{.2} = 23.625, \bar{Y}_{.3} = 19.0, \bar{Y}_{.4} = 29.375, \bar{Y}_{.5} = 21.125$)

2. Consider a split block experiment with Factor A at 3 levels, Factor B at 4 levels, and six blocks. Assume that the three levels of Factor A are randomly assigned to three homogeneous parts (whole plots) within each block, and that the four levels of Factor B are assigned at random to four split-plots within each whole plot. Let y_{ijk} be the measurement of a response variable y_{ijk} for i-th level of A, k-th level of B in j-th block.
Suppose that SS(A)= 170, SSE(Block)=1550, SS(A*Block)=600, SS(B)=2000, SS(A*B)=30, and SSE=800.

a) Write down the statistical model for y_{ijk} based on the split-plot design described above, and state the assumption and distributions (if any) of the components in your model.

b) Write down the ANOVA table showing the Source, Sum of Squares, Degrees of Freedom, and Mean Square columns.

c) Give the value of the test statistic for testing whether there is a significant interaction effect between the factors A and B. What distribution would you use to get the p-value for this test?

d) Give the value of the test statistic for testing whether there is significant effect of Factor A. What distribution would you use to get the p-value for this test?

3. Consider the model for three stage nested design, $Y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(ij)} + \epsilon_{ijkl}$, i=1,…,a; j=1,…,b; k=1,…,c; l=1,…, n where τ_i is the effect of A, $\beta_{j(i)}$ is the effect of B within A, $\gamma_{k(ij)}$ is the effect of C within A and B. Assume that A is fixed, B & C are random factors. Construct the ANOVA table including the form of SS (e.g. SST = $\sum_i \sum_j \sum_k \sum_l \left(Y_{ijkl} - \bar{Y}_{....}\right)^2$), df, MS, F-ratio, and Expected Mean Squares.

4. Consider the one-way random effect model, i.e., $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$ where Y_{ij} is the value of the response variable in the j[th] trial for the i[th] treatment, μ is fixed but unknown, τ_i's are independent $N(0, \sigma_\tau^2)$ and ϵ_{ij}'s are independent $N(0, \sigma^2)$ for $i = 1, \ldots, r, j = 1, \ldots, n$. And assume τ_i's and ϵ_{ij}'s are mutually independent. Let $\bar{Y}_{i.} = \sum_{j=1}^{n} Y_{ij}/n$ and $\bar{Y}_{..} = \sum_{i=1}^{r} \sum_{j=1}^{n} Y_{ij}/rn$. We can write the log-likelihood function as

$$l(\mu, \sigma_\tau^2, \sigma^2 \mid y) = -\frac{rn}{2}\log(2\pi) - \frac{1}{2}r(n-1)\log\sigma^2 - \frac{r}{2}\log\lambda - \frac{SSE}{2\sigma^2} - \frac{SSTR}{2\lambda} - \frac{rn(\bar{Y}_{..} - \mu)^2}{2\lambda},$$

where $\lambda = \sigma^2 + n\,\sigma_\tau^2$.

a) Find the maximum likelihood estimator of $(\mu, \sigma_\tau^2, \sigma^2)$,

b) Find the restricted maximum likelihood (REML) estimators of σ_τ^2 and σ^2. First, show the restricted likelihood function of $(\sigma_\tau^2, \sigma^2)$ is

$$l_R\left(\sigma_\tau^2, \sigma^2 \middle| SSTR, SSE\right) = -\frac{1}{2}(rn-1)\log(2\pi) - \frac{1}{2}\log(rn) - \frac{1}{2}r(n-1)\log\sigma^2 - \frac{1}{2}(r-1)\log\lambda - \frac{SSE}{2\sigma^2} - \frac{SSTR}{2\lambda}.$$

Then, find the REML estimator of $(\sigma_\tau^2, \sigma^2)$.

5. Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$ be independent $MVN_p(\mu_{p\times 1}, \Sigma_{p\times p})$ random vectors. Define

$$\boldsymbol{V}_1 = \frac{1}{3}\boldsymbol{X}_1 - \frac{1}{3}\boldsymbol{X}_2 + \frac{1}{3}\boldsymbol{X}_3 \quad \text{and} \quad \boldsymbol{V}_2 = \frac{1}{2}\boldsymbol{X}_1 + \frac{1}{2}\boldsymbol{X}_2$$

a) Find the joint distribution of the random vectors $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$. Are $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ statistically independent?

b) Find the marginal distributions for each of $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$.