Qualifying Exam II

Sample 2

1. Let $X_1, ..., X_n$ be observations of a random sample from pdf

$$f_\sigma(x) = \frac{1}{\sigma} e^{-|x|/\sigma}, \quad x \in \mathcal{R},$$

where $\sigma > 0$ is an unknown parameter.

(a) Find the pdf of $X_1 - X_2$ and the pdf of $|X_1 - X_2|$.

(b) Find the UMVUE of σ. Does the variance of the UMVUE achieve the Cramer-Rao lower bound?

(c) Find the UMP level α test of the hypotheses $H_0 : \sigma \leq \sigma_0$ versus $H_1 : \sigma > \sigma_0$, where $\sigma_0 > 0$ is a known constant. Write down the power function of this test in terms of the cumulative distribution function of a known distribution whose parameters do not depend on σ.

(d) Let

$$\theta = P(|X_1| \geq 1) = e^{-1/\sigma}$$

Find $\hat{\theta}_{mls}$, the MLE of θ, and its asymptotic distribution.

(e) Let $\hat{\theta}_{prop}$ be the proportion of $|X_1|, ..., |X_n|$ exceeding 1. Find the asymptotic relative efficiency of $\hat{\theta}_{mls}$ with respect to $\hat{\theta}_{prop}$.

(f) Find $\hat{\theta}_{umvue}$, the UMVUE of θ, and show that

$$\sqrt{n}(\hat{\theta}_{mls} - \hat{\theta}_{umvue}) = o_p(1)$$

2. Let $X_1, ..., X_n$ be observations of a random sample from the Bernoulli distribution with probability θ, i.e., $P(X_1 = 1) = \theta$ and $P(X_1 = 0) = 1 - \theta$, where $\theta \in (0, 1)$ is unknown.

   (a) Let k and m be two integers with $1 \leq k < m \leq n$, $U = \sum_{i=1}^{m} X_i$, and $W = \sum_{i=k+1}^{n} X_i$. Calculate the covariance between U and W and show that they are not independent.

   (b) Let $Y = \sum_{i=k+1}^{m} X_i$. Identify the conditional distribution of U given $Y = y$, $1 \leq y \leq m - k$.

   (c) Show that, conditional on Y, U and W are independent, i.e.,

$$P(U = u, W = w|Y = y) = P(U - u|Y = y)P(W = w|Y = y) \quad \text{for all } u,\ w,\ y$$

   (d) Invert the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, where θ_0 is a known constant, to find a confidence interval with asymptotic coverage probability $1 - \alpha$ (don't solve for the endpoints explicitly; just write down the equations the endpoints must satisfy).

   (e) Suppose that θ has a prior that is a known Beta distribution. Consider testing the one-sided hypotheses $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$. We reject H_0 if the posterior probability of H_0 is less than the posterior probability of H_1. Write the rejection region of this test in terms of θ_0 and the median of a certain distribution.

   (f) Using the same Beta prior for θ, find the shortest $1 - \alpha$ credible interval for θ (don't solve for the endpoints explicitly; just write down the equations the endpoints must satisfy). For the Uniform[0,1] prior on θ, compare the form of the credible interval to the form of the confidence interval derived in part (d) and comment on any similarities and differences.

3. A scale has two pans. The measurements given by the scale is the difference between the weights in pan # 1 and pan # 2 plus a random error. Thus, if a weight μ_1 is put in pan # 1, a weight μ_2 is put in pan # 2, then the measurement is $Y = \mu_1 - \mu_2 + \epsilon$. Suppose that $E[\epsilon] = 0$ and $\text{Var}(\epsilon) = \sigma^2$, and that in repeated uses of the scale, observations Y_i are independent.

   Suppose that two objects, #1 and #2, have weights β_1 and β_2. Measurements are taken as follows:

   (i) Object #1 is put on pan #1, nothing on pan # 2.

   (ii) Object #2 is put on pan # 2, nothing on pan # 1.

   (iii) Object #1 is put on pan # 1, object #2 on pan #2.

   (vi) Objects #1 and #2 both put on pan # 1.

   Answer the following questions based on the above measurements.

   (a) Let $\mathbf{Y} = (Y_1, Y_2, Y_3, Y_4)'$ be the vector of observations. Formulate this as a linear model.

   (b) Find vectors $\mathbf{c}_1$, $\mathbf{c}_2$ such that $\hat{\beta}_1 = \mathbf{c}_1'\mathbf{Y}$ and $\hat{\beta}_2 = \mathbf{c}_2'\mathbf{Y}$.

   (c) Find the covariance matrix of $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)'$.

   (d) Find the matrix $\mathbf{A}$ such that $s^2 = \mathbf{Y}'A\mathbf{Y}$, where s^2 is an unbiased estimator of σ^2.

   (e) For the observation $\mathbf{Y} = (7, 3, 1, 7)'$, find s^2, and estimate the covariance matrix of $\hat{\beta}$.

   (f) Show that four such weighings can be made in such a way that the least squares estimators of β_1 and β_2 have smaller variances than the experiment above.

4. (This problem has two parts concerning two different sets of experiments.)

(a) An experiment was conducted to evaluate the effect of two levels of fertilizer (denoted F_1 and F_2) and three levels of pesticide (denoted P_1, P_2, and P_3) on the yield of corn. The investigators are particularly interested in whether there is evidence of a fertilizer by pesticide interaction.

In the experiment, 36 plots were laid out in a 6 by 6 array in a large field. Because of concerns about possible trends in the field, although randomness was used, the experimenters were careful to arrange the treatment combinations such that each combination appeared once in each column and once in each row. The final treatment allocation looked like this:

F_2P_1	F_1P_3	F_2P_2	F_2P_3	F_1P_1	F_1P_2
F_1P_1	F_1P_2	F_2P_1	F_1P_3	F_2P_2	F_2P_3
F_1P_3	F_2P_2	F_2P_3	F_1P_1	F_1P_2	F_2P_1
F_1P_2	F_2P_1	F_1P_3	F_2P_2	F_2P_3	F_1P_1
F_2P_2	F_2P_3	F_1P_1	F_1P_2	F_2P_1	F_1P_3
F_2P_3	F_1P_1	F_1P_2	F_2P_1	F_1P_3	F_2P_2

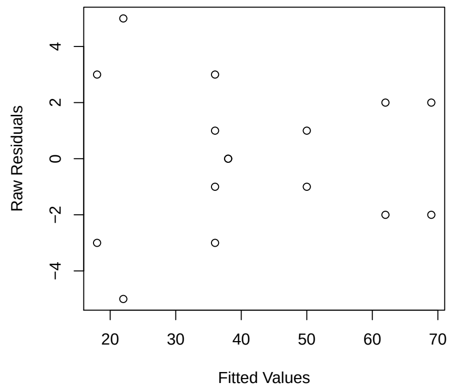
Write down the ANOVA table for this experiment, indicating Source and df only.

(b) An experiment was conducted to assess the effect of various factors on the quality of vanilla ice cream. Three factors were of interest: amount of sugar (high or low), blending time (short or long), and use of an egg flavoring substitute (present or absent). The experiment was randomized as follows: for each of the 8 treatment combinations, two tickets were made. All 16 tickets were put in a hat, and each morning a ticket would be drawn from the hat at random and a batch of ice cream would be made according to the combination on the ticket drawn from the hat. The experiment took 16 days to complete. On each day, after each batch was produced, an assessment of flavor was made on a 100 point scale — the larger the score, the better the flavor.

The data are given below:

Sugar:	Low	Low	Low	Low	High	High	High	High
Blending Time:	Short	Short	Long	Long	Short	Short	Long	Long
Egg Flavoring:	Yes	No	Yes	No	Yes	No	Yes	No
Results:	15, 21	60, 64	33, 39	35, 37	27, 17	71, 67	38, 38	49, 51

i. Write down the model for this experiment, defining all terms.
ii. Write down the ANOVA table for this experiment, indicating Source and df (only).
iii. Focus on the data where Sugar is Low, only. Using graphical methods, assess whether there is evidence of a two-way interaction between blending time and use of egg flavoring. (No formal analysis is required.)
iv. Now consider all the data (16 observations). Perform a formal test for the presence of a main effect for sugar amount. (You need only test the main effect — you are not required to test interactions first.) You may use the fact that, in the ANOVA table, SSError = 106.0.
v. Here is a residual plot for the experiment, based on fitting all main effects and all interactions (to all 16 observations).



Does this residual plot indicate any possible problems that need addressing? Choose: Yes or No, and explain briefly.