Qualifying Exam II

Sample 1

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

$$f_\theta(x) = \theta x^{-(\theta+1)}, \quad x > 1,$$

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

   (a) Find the pdf of X_1/X_2.

   (b) Let k and j be two integers between 1 and n and $j < k$. Derive the joint pdf of $X_{(j)}$ and $X_{(k)}$, where $X_{(j)}$ is the jth order statistic.

   (c) Discuss under what condition on θ the law of large numbers can be applied to the sample mean $\bar{X} = n^{-1} \sum_{i=1}^n X_i$.

   (d) Discuss under what condition on θ the central limit theorem can be applied to the sample mean $\bar{X}$ and $\sqrt{n}(\bar{X}^{-1} - c)$ has a limit distribution. Identify c as a function of θ and the limit distribution.

   (e) Show that the gamma family of priors is conjugate. Derive the Bayes estimator of θ under the squared error loss and one of the prior from the gamma family.

   (f) For testing $H_0 : \theta = 3$ versus $H_1 : \theta \neq 3$, show that the likelihood ratio test is equivalent to the UMP test with size $\alpha \in (0, 1)$.

   (g) Obtain a $1-\alpha$ asymptotically correct confidence interval for θ by inverting the acceptance regions of likelihood ratio tests.

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

$$f_\theta(x) = \frac{c}{\sigma} \exp\left\{ -\left(\frac{x-\mu}{\sigma}\right)^4 \right\}, \quad x \in \mathcal{R}$$

where $\mu \in \mathcal{R}$ and $\sigma \in (0, \infty)$ are unknown parameters and $c > 0$ is a fixed constant. Let $\theta = (\mu, \sigma)$, $\bar{X}$ be the sample mean, and $S^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$ be the sample variance.

(a) Show that $c = 2/\Gamma(1/4)$, where $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$ is the gamma function, $t > 0$.

(b) Calculate $E(X_1 - \mu)^2$ and find a constant a such that aS^2 is unbiased for σ^2.

(c) Obtain the covariance between $\bar{X}$ and S^2.

(d) Obtain the likelihood given observed $X_1, ..., X_n$, $L(\theta)$, and show that $\partial \log L(\theta)/\partial\mu$ is a monotone function of μ and that an MLE $\hat{\mu}$ of μ exists and is unique.

(e) Calculate the asymptotic relative efficiency of the sample mean $\bar{X}$ with respect to the MLE $\hat{\mu}$.

(f) Define $n_1 = \sum_{i=1}^n I(X_i \leq 0)$ and $n_2 = \sum_{i=1}^n I(X_i > 0)$, where $I(\cdot)$ is the indicator function. Let $p_1 = P(X_1 \leq 0)$, $p_2 = 1 - p_1$, and

$$T_n = \sum_{j=1}^{2} \frac{(n_j - np_j)^2}{np_j}$$

Derive and identify the limit distribution of T_n as $n \to \infty$.

(g) For testing $H_0 : p_1 = p_{10}$ versus $H_1 : p_1 \neq p_{10}$, where $p_{10} \in (0, 1)$ is a known value, develop an asymptotically level α test procedure.

3. A rocket is launched from ground level (altitude $y = 0$) at time $t = 0$. Its upward acceleration is a. The altitude of the rocket is measured by two separate instruments at times $t_1, t_2, \cdots, t_n$. At time t_i, the altitudes $y_{i,1}$ and $y_{i,2}$ measured by the respective instruments satisfy the model

$$y_{i,j} = \frac{1}{2}at_i^2 + e_{i,j}, \quad 1 \le i \le n \quad j = 1, 2.$$

The measurement errors $\{e_{i,j} : 1 \le i \le n, j = 1, 2\}$ are taken to be independent, identically distributed, each having a $N(0, \sigma^2)$ distribution. Both the acceleration a and the variance σ^2 are unknown.

(a) Let $y = (y_{1,1}, y_{1,2}, y_{2,1}, y_{2,2}, \cdots, y_{n,1}, y_{n,2})'$, a vector of dimension $2n \times 1$, and let $\beta = a$. Find covariate matrix X such that the model discussed above is equivalent to the assertion that the y has a $N(X\beta, \sigma^2 I_{2n})$ distribution.

(b) Using general theory or other argument, find and simplify the least squares estimate $\hat{a}$ for the acceleration a. Find the distribution of $\hat{a}$.

(c) Consider the following acceleration estimate $\tilde{a} = \sum_{i=1}^n (y_{i,1} + y_{i,2}) / \sum_{i=1}^n t_i^2$. Find the distribution of $\tilde{a}$.

(d) Find the mean squared errors $E(\hat{a} - a)^2$ and $E(\tilde{a} - a)^2$. Which of the two estimators $\hat{a}$ and $\tilde{a}$ has smaller mean squared error?

(e) Construct a $100 \times (1 - \alpha)\%$ confidence interval for the acceleration a that is based on the least squares estimate $\hat{a}$.

(f) Construct a $100 \times (1 - \alpha)\%$ confidence interval for $\frac{1}{2}at_0^2$, the expected altitude of the rocket at time t_0.

4. A split plot experiment was conducted to study the effect of factors that might affect the sweetness of wine measured in grams of sugar per liter of juice. Two experimental factors were of interest: the fertilization of the plants, and type of pruning method used. The fertilizer treatments consisted of: (1) amendment with nitrogen only (NO); or (2) amendment with nitrogen and phosphorus (NP). The pruning methods are defined in terms of the number of buds left. In this experiment pruning was done to leave either 50, 45, 40, or 30 buds. Pruning to 50 buds is least aggressive; pruning to 30 is most aggressive.

   The experiment was conducted in three vineyards at a large winery. We will simply label these as the East, West, and Central vineyard. Within each vineyard, two rows of vines were chosen at random, and each one was randomly assigned to receive one of the fertilizer treatments. Within each of the rows, four trunks were chosen, and the each of the trunks was randomly assigned to receive one of the four pruning methods. Finally, at harvest time, two grapes were sampled from each trunk (call them Grape 1 and Grape 2) and the sweetness of each grape was determined.

   The data are shown below. The columns are, in order, the sweetness, the vineyard, the fertilizer, the amount of pruning, and the grape sampled for each trunk.

```
45 E NO 50 1
53 E NO 50 2
51 E NO 45 1
61 E NO 45 2
53 E NO 40 1
66 E NO 40 2
54 E NO 30 1
66 E NO 30 2
76 E NP 50 1
86 E NP 50 2
71 E NP 45 1
81 E NP 45 2
71 E NP 40 1
82 E NP 40 2
62 E NP 30 1
73 E NP 30 2
31 C NO 50 1
43 C NO 50 2
27 C NO 45 1
40 C NO 45 2
35 C NO 40 1
44 C NO 40 2
49 C NO 30 1
57 C NO 30 2
61 C NP 50 1
73 C NP 50 2
61 C NP 45 1
71 C NP 45 2
73 C NP 40 1
82 C NP 40 2
81 C NP 30 1
88 C NP 30 2
66 W NO 50 1
76 W NO 50 2
53 W NO 45 1
```

```
66 W NO 45 2
79 W NO 40 1
96 W NO 40 2
83 W NO 30 1
98 W NO 30 2
79 W NP 50 1
88 W NP 50 2
75 W NP 45 1
83 W NP 45 2
86 W NP 40 1
100 W NP 40 2
101 W NP 30 1
113 W NP 30 2
```

These data were also analyzed in SAS as follows:

```
options ls=80;

data a;
    infile "grapes.txt" ;
    input sweet yard$ fert$ prune grape;

proc glm;
    class yard fert prune grape;
    model sweet = yard|fert|prune|grape;

    lsmeans yard fert prune grape;
```

Here is some edited output.

```
                              The GLM Procedure

                           Class Level Information

                       Class         Levels    Values

                       yard               3    C E W
                       fert               2    NO NP
                       prune              4    30 40 45 50
                       grape              2    1 2

Dependent Variable: sweet

                                        Sum of
 Source                      DF        Squares     Mean Square    F Value    Pr > F

 Model                       47    17422.81250       370.69814        .        .
 Error                        0        0.00000         .
 Corrected Total             47    17422.81250

 Source                      DF      Type I SS     Mean Square    F Value    Pr > F
```

yard	2	5924.625000	2962.312500	.	.
fert	1	5742.187500	5742.187500	.	.
yard*fert	2	805.875000	402.937500	.	.
prune	3	1772.729167	590.909722	.	.
yard*prune	6	1215.208333	202.534722	.	.
fert*prune	3	82.729167	27.576389	.	.
yard*fert*prune	6	369.958333	61.659722	.	.
grape	1	1441.020833	1441.020833	.	.
yard*grape	2	10.791667	5.395833	.	.
fert*grape	1	6.020833	6.020833	.	.
yard*fert*grape	2	4.041667	2.020833	.	.
prune*grape	3	6.562500	2.187500	.	.
yard*prune*grape	6	35.375000	5.895833	.	.
fert*prune*grape	3	3.562500	1.187500	.	.
yard*fert*prun*grape	6	2.125000	0.354167	.	.

Least Squares Means

yard	sweet LSMEAN
C	57.2500000
E	65.6875000
W	83.8750000

fert	sweet LSMEAN
NO	58.0000000
NP	79.8750000

prune	sweet LSMEAN
30	77.0833333
40	72.2500000
45	61.6666667
50	64.7500000

grape	sweet LSMEAN
1	63.4583333
2	74.4166667

(a) Explain why this experiment is a split plot experiment.

(b) Make a suitable plot to determine whether there is evidence of an interaction between fertilization and pruning method.

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

(d) Perform a formal test to determine whether there is a main effect for fertilization.

(e) Perform a formal test to determine whether there is a main effect for pruning.

(f) Suppose that the experimenters had instead decided to focus on the NP fertilizer treatment, and the use of a 40 bud pruning regime. Further, suppose they observed 8 trunks, and 3 grapes per trunk. Imagine calculating the mean sweetness from the resulting 24 grapes. Estimate the variance of this mean.