Final Review Questions

Note: Many of these questions are drawn from previous years' finals. I will not provide solutions to these questions. The final will cover material from the entire semester, but with more weight on the second half of the course.

1. Suppose that $X_1$ is uniformly distributed on $(0, 1)$ and $X_2$ is uniformly distributed on $(0, 2)$, and $X_1$ and $X_2$ are independent. Let $Y \equiv \max(X_1, X_2)$.
   
   (a) Derive the CDF and PDF of $Y$.
   
   (b) Calculate $E[Y]$.
   
   (c) Suppose that instead of being independent, $X_1$ and $X_2$ are related in the following way: $X_1$ is uniform $(0, 1)$, and $X_2 = 2 \cdot X_1$. Let $Y$ be defined as before. Now calculate the CDF and PDF of $Y$.

2. Suppose $X_n \sim Bin(n, \frac{\alpha}{n})$, where $\alpha > 0$.
   
   (a) $X_n \xrightarrow{d} Y$. What's the probability mass function for $Y$?
   
   (b) Does $\frac{X_n}{\sqrt{n}} \xrightarrow{p} 0$? Prove your answer.

3. Suppose that $Y_1, \ldots, Y_n$ are IID with a discrete distribution with PMF
   
   $$f_Y(y; K) = \begin{cases} \frac{1}{K} & \text{for } y = 1, 2, \ldots, K \\ 0 & \text{otherwise.} \end{cases}$$

   The parameter $K$ is an integer $\geq 1$.
   
   (a) Find the maximum likelihood estimator for the parameter $K$.
   
   (b) Show that if $K_0$ is the true value of the parameter, then the MLE $\hat{K}$ has the following properties:

   $$Pr(\hat{K} = 1) > 0;$$

   $$Pr(\hat{K} > K_0) = 0.$$

   (c) Use your result from (b) to show that the MLE $\hat{K}$ is biased towards 0.

4. Let $X$ be a random variable with probability density function
   
   $$f_X(x; \mu) = \frac{1}{\mu} \exp(-x/\mu),$$

   for $x > 0$ and zero elsewhere.
(a) Calculate the mean and variance of $X$.

(b) Calculate the mean and variance of $X$ conditional on $X < 8$.

(c) Let $x_1, x_2, \ldots, x_N$ be a random sample from this distribution, with $N = 20$, $\sum x = 95$, and $\sum x^2 = 590$. Calculate the maximum likelihood estimate.

(d) Test the hypothesis that $\mu = 4$ at the 10% level using a likelihood ratio test.

(e) Test the same hypothesis using a Lagrange multiplier (score) test.

5. Let the marginal distribution of $X$ be binomial with $N = 1$ and $p = 1/4$. Conditional on $X$, the random variable $Y$ has a normal distribution with mean $\mu \cdot (X + 1)$ and variance 1.

(a) Find the marginal density of $Y$.

(b) Suppose you have a random sample of size $N$ from this joint distribution. What is the maximum likelihood estimator and its large sample variance?

(c) Suppose you only observe $y_1, \ldots, y_N$. Find an unbiased estimator for $\mu$. What is its large sample variance and how does that compare to that of the maximum likelihood estimator derived before?

6. Suppose there is a random sample of size 10 ($X_1, \ldots, X_{10}$) from a Poisson distribution (so $f_{X_i}(x_i|\theta) = \frac{\theta^{x_i}e^{-\theta}}{x_i!}$, when $x_i$ is a nonnegative integer). The sample mean of the random sample is 4.

(a) Using the sample, what is the maximum likelihood estimate for $\theta$?

(b) Provide an approximate 95% confidence interval for $\theta$.

(c) Using a large-sample LR test, test the hypothesis that $\theta = 3$ at the 0.05 level. (Note that if $Z$ is a chi-squared random variable with 1 degree of freedom, $Pr(Z > 3.84) = 0.05$.)

(d) Now suppose that the prior distribution for $\theta$ is $Gamma(3, 5)$:

$$\theta \sim Gamma(3, 5).$$

What is the posterior distribution for $\theta$ given $X_1, \ldots, X_{10}$? (Hint: $Z \sim Gamma(\alpha, \beta)$ means $f_Z(z|\alpha, \beta) = \frac{z^{\alpha-1}e^{-z/\beta}}{\Gamma(\alpha)\beta^\alpha}$)

7. A firm samples machine parts until it finds a defective part; let $X_i$ be the number of samples until a defective part. Assume that each sample is independent with probability $p$ of being defective.

(a) Derive the probability mass function of $X_i$. 

2
(b) Suppose that the firm obtains independent observations $X_1, \ldots, X_n$. The firm has a Beta(1,1) prior distribution for $p$. What is the posterior distribution for $p$?

Note: the Beta distribution with parameters $\alpha$ and $\beta$ has PDF

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}.$$ 

8. Suppose we have a random sample of individuals from a population, and assume that earnings in a given month for the $i$th sampled individual is $X_i$, where $X_i$ has PDF

$$f(x; \theta) = \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{(\log x - \theta)^2}{2}\right), \quad x > 0,$$

where the parameter $\theta \in \mathbb{R}$.

(a) Is the statistical model an exponential family?

(b) Derive the MLE for $\theta$ based on a sample of size $n$. Is the MLE a minimum variance unbiased estimator?

(c) Construct a large sample Wald test for the hypothesis that $\theta = \theta_0$. (Hint: if possible, show that the single-observation score function has variance 1. If you cannot show this, you can just take as given that its variance is 1.)

(d) Suppose in our data set we observe $n = 100$, $\sum_i X_i = 538$, $\sum_i \log X_i = 127$, and $\sum_i X_i^2 = 5628$. Calculate the MLE and provide a large sample 95% confidence interval for $\theta$.

9. Suppose $X_1, \ldots, X_n$ are i.i.d. random variables each with PMF

$$f_X(x|\alpha) = \begin{cases} \frac{\alpha-1}{\alpha^2} & \text{for } x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

for $\alpha > 1$. Note that $E[X] = \frac{\alpha}{\alpha - 1}$.

(a) Does there exist a one-dimensional sufficient statistic (based on $X_1, \ldots, X_n$) for $\alpha$? If so what is it?

(b) Derive the method of moments and maximum likelihood estimators for $\alpha$.

(c) Obtain the limiting distribution of $\sqrt{n}(\hat{\alpha}_{ml} - \alpha)$.

(d) Suppose that the sample size is $n = 2$, and the data are $X_1 = 2$ and $X_2 = 1$. Given a prior distribution on $\alpha$ with PDF

$$p(\alpha) = \frac{\mathbbm{1}(\alpha > 1)}{\alpha^2},$$

what is the mean of the posterior distribution of $\alpha$?
10. Let \( X \) have a Bernoulli distribution with \( Pr(X = 1) = \frac{1}{2} \). Conditional on \( X = x \), the random variable \( Y \) has a Poisson distribution with parameter \( \lambda(1 + x) \):

\[
Pr(Y = y | X = x) = \frac{(\lambda(1 + x))^y \exp(-\lambda(1 + x))}{y!},
\]

for \( x = 0, 1 \), and \( y = 0, 1, \ldots \). We observe a single draw for the pair \((X, Y)\).

(a) Calculate \( Pr(Y = 2) \).

(b) Calculate \( Pr(X = 1 | Y = 2) \).

(c) Find an unbiased estimator for \( \lambda \) that is a function of \( Y \) alone.

(d) Derive the maximum likelihood estimator for \( \lambda \) given the pair \((X, Y)\). Give as simple an expression as possible.

(e) Is the maximum likelihood estimator unbiased?

11. Suppose that \( X_1, X_2, \ldots \) are IID with mean 0 and variance \( \sigma^2 \). The variance \( \sigma^2 \) is not known, but assumed to exist and be finite, and we also assume that the distribution of \( X_i \) has finite higher order moments. We do not know anything else about the distribution of the \( X_i \).

Consider the following estimator of the variance:

\[
\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2.
\]

(a) Show that \( \sqrt{n}(\hat{\sigma}^2 - \sigma^2) \) converges in distribution and give the limiting distribution.

(b) Derive a large-sample 95% confidence interval for \( \sigma^2 \).

(c) Provide (and justify) a large-sample test of the hypothesis that \( \sigma^2 = 1 \) against the alternative that \( \sigma^2 \neq 1 \).

12. Suppose that \( Y_1, Y_2, \ldots, Y_n \) are a random sample (iid) from a Geometric distribution with PMF

\[
f(y; p) = p \cdot (1 - p)^{y-1} \quad \text{for } x = 1, 2, \ldots.
\]

The parameter is \( p \in (0, 1) \). Note that \( E[Y_i] = 1/p \) and \( V[Y_i] = (1 - p)/p^2 \).

(a) Provide a one-dimensional sufficient statistic; be sure to explain your reasoning.

(b) Derive the maximum likelihood estimator \( \hat{p} \) and obtain the large-sample distribution of \( \sqrt{n}(\hat{p} - p) \).

(c) In the sample we observe \( \sum_{i=1}^{n} Y_i = 100 \) for \( n = 40 \). Test the hypothesis that \( p = .5 \) against the alternative that \( p \neq .5 \) using a Wald test at the 5% significance level.
13. Suppose we observe the earnings of pairs of siblings: \((X_i, Y_i)\) where \(X_i\) is the earnings of the younger sibling and \(Y_i\) is the earnings of the older sibling in family \(i\). We have a sample \((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\) where the vectors \((X_i, Y_i)\) are iid with 

\[
\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma^2_x & 0 \\ 0 & \sigma^2_y \end{pmatrix} \right).
\]

We are interested in the quantity \(\tau = \mu_y - \mu_x\).

(a) Explain the practical meaning of the null hypothesis \(H_0: \tau = 0\).

(b) Suppose that \(\sigma^2\) is known. Provide an exact (nonasymptotic) test of the hypothesis that \(\tau = 0\) against the alternative that \(\tau \neq 0\) at the 5% level. Describe the procedure in enough detail so that it could be easily coded by a research assistant familiar with R (but not otherwise trained in statistics).