Homework #5

(due Wednesday, February 15)

1. Larsen and Marx, Problem 5.5.2, p. 323.

2. Larsen and Marx, Problem 5.5.6, p. 323. (Hint: to compute the mean and variance of $\hat{\theta}$, make the substitution $x = y/\theta$ and use the fact that $\Gamma(m) = \int_0^\infty x^{m-1}e^{-x}\,dx = (m-1)!$ for any positive integer m.)

3. Larsen and Marx, Problem 5.7.2, p. 333.

4. Larsen and Marx, Problem 5.7.4, part (b), p. 574. (Hint: use Markov's Inequality.)

5. Suppose $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables having the Poisson distribution with mean λ. Let $\hat{\lambda}_n = X_n$.

   (a) Is $\hat{\lambda}_n$ an unbiased estimator of λ? Explain your answer.

   (b) Is $\hat{\lambda}_n$ a consistent estimator of λ? Explain your answer.

6. Suppose $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables having a normal distribution with mean μ and variance σ^2. Let

$$\hat{\mu}_n = \frac{X_1 + \cdots + X_{\lfloor\sqrt{n}\rfloor}}{\lfloor\sqrt{n}\rfloor},$$

   where $\lfloor\sqrt{n}\rfloor$ denotes the largest integer less than or equal to $\sqrt{n}$.

   (a) Is $\hat{\mu}_n$ an unbiased estimator of μ? Explain your answer.

   (b) Is $\hat{\mu}_n$ an efficient estimator of μ? Explain your answer.

   (c) Is $\hat{\mu}_n$ a consistent estimator of μ? Explain your answer.

7. Suppose $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables with $E[X_i] = \theta$ and $\mathrm{Var}(X_i) = \sigma_1^2 < \infty$. Suppose also that $Y_1, Y_2, \ldots$ is a sequence of i.i.d. random variables, independent of $X_1, X_2, \ldots$, with $E[Y_i] = \mu$ and $\mathrm{Var}(Y_i) = \sigma_2^2 < \infty$. Suppose $(\alpha_n)_{n=1}^\infty$ is a sequence of numbers such that $0 < \alpha_n < 1$ for all n and

$$\lim_{n\to\infty} \alpha_n = 1.$$

   Let

$$\hat{\theta}_n = \alpha_n\left(\frac{X_1 + \cdots + X_n}{n}\right) + (1-\alpha_n)\left(\frac{Y_1 + \cdots + Y_n}{n}\right).$$

   Show that $\hat{\theta}_n$ is a consistent estimator of θ. (Hint: compute the mean squared error of $\hat{\theta}_n$ and use the result of Problem 4.)

Problems 8, 9, and 10 pertain to asymptotic properties of maximum likelihood estimators, a topic which is not covered in Larsen and Marx. If $X_1, \ldots, X_n$ is a random sample from the pdf $f_X(x;\theta)$, which satisfies certain smoothness conditions, then the maximum likelihood estimator $\hat{\theta}$ has approximately a normal distribution with mean θ and asymptotic variance $1/[nI(\theta)]$, where

$$I(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\log f_X(X;\theta)\right)^2\right] = -E\left[\frac{\partial^2}{\partial\theta^2}\log f_X(X;\theta)\right].$$

An approximate $100(1-\alpha)$ percent confidence interval for θ is given by

$$\left(\hat{\theta} - z_{\alpha/2}\sqrt{\frac{1}{nI(\hat{\theta})}},\ \hat{\theta} + z_{\alpha/2}\sqrt{\frac{1}{nI(\hat{\theta})}}\right).$$

The same results hold in the discrete case with $p_X(k;\theta)$ in place of $f_X(x;\theta)$. More details can be found in sections 8.5.2 and 8.5.3 of Rice, *Mathematical Statistics and Data Analysis*, which is on reserve in the library.

8. Between 1951 and 2000, there were 295 hurricanes in the Atlantic Ocean. Assume that the numbers of hurricanes in different years can be accurately modeled as i.i.d. random variables having a Poisson distribution with mean λ. Find a 90 percent confidence interval for λ.

9. Let $X_1, \ldots, X_n$ be a random sample from the pdf $f_X(x;\theta) = (\theta+1)x^\theta$ for $0 < x < 1$, and $f_X(x;\theta) = 0$ otherwise, where $\theta > -1$. Recall that in Problem 8 of Homework 2, you found that the MLE for θ is given by

$$\hat{\theta} = -\left(1 + \frac{n}{\sum_{i=1}^n \log X_i}\right).$$

   Find the asymptotic variance of $\hat{\theta}$.

10. Let $X_1, \ldots, X_n$ be i.i.d. random variables having a geometric distribution with parameter $p \in [0,1]$, which means that $P(X_i = k) = (1-p)^{k-1}p$ for $k = 1, 2, \ldots$ and $E[X_i] = 1/p$.

    (a) Find the maximum likelihood estimator of p.

    (b) Find the asymptotic variance of the maximum likelihood estimator.

    (c) If $n = 100$ and you observe the the following data:

Number	Frequency
1	60
2	21
3	12
4	4
5	3

    (meaning that 60 of the random variables $X_1, \ldots, X_{100}$ equal 1, and so on), find a 95 percent confidence interval for p.