DEPARTMENT OF STATISTICS, Penn State University
Statistics 514 (Spring 2010)

Practice Problems

1. Suppose someone has a coin they claim has $\theta = P(H) = 1/4$. You conduct the trial: toss a coin until a head appears and record X = number of tails until the first head appears. Then X has the probability mass function: $f(x, \theta) = \theta(1-\theta)^x$, $x = 0, 1, 2, \ldots$, $\theta > 0$. Now carry out 10 trials to get $x_1, \ldots, x_{10}$. You wish to test $H_0 : \theta = 1/2$ vs. $H_A : \theta = 1/4$. Find the most powerful size α critical region. Explain in detail how you would find the critical value for the test.

2. Suppose $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ with μ and σ^2 both unknown. Show that the size α likelihood ratio test of $H_0 : \mu = 0$, σ^2 unspecified vs. $H_A : \mu \neq 0$, σ^2 unspecified can be based on a t-statistic. Provide statements to support the arguments you use in the derivation.

3. Let $X_1, \ldots, X_n$ be iid with pmf $f(x; p) = p^x(1-p)^{1-x}$, $x = 0, 1$, zero elsewhere. Find the most powerful critical region for testing $H_0 : p = 1/2$ vs. $H_A : p = 1/3$. Use the CLT to find n and c (critical value of the test) so that you obtain a size $\alpha = 0.10$ test with power 0.80.

4. Consider a random sample $X_1, \ldots, X_n$ from a distribution with pdf $f(x; \theta) = \theta(1-x)^{\theta-1}$, $0 \leq x \leq 1$, zero elsewhere, where $\theta > 0$. Find the form of the uniformly most powerful test of $H_0 : \theta = 1$ vs. $H_A : \theta > 1$. Argue carefully why it is the uniformly most powerful test.

5. You are interested in estimating the number of times adult female fruit flies mate during their brief lifespan. You set up an experiment where you observe 6 female flies, and find that they mate $3, 5, 1, 2, 2$ and 4 times. You decide to model the number of times they mate as a Poisson random variable with mean λ. A literature search reveals that closely related species mate 3 times on average, so you put a Gamma prior on λ, with parameters $\alpha = 3$, $\beta = 1$.

   a. What is the posterior distribution of λ?

   b. If we are using squared error loss, what is the Bayes estimate for λ? (You should give a number here.)

   c. What is the MLE for λ using this data? Explain any differences from the Bayes estimate.

6. Consider Bayes estimation of p, the proportion of turkeys at Giant exceeding 20 pounds. Our prior on the proportion is $beta(0.5, 1)$. We go to Giant and observe 100 randomly selected turkeys. 32 exceed 20 pounds.

   a. What is the posterior of p?

   b. If we are using squared error loss, what is the Bayes estimate for p? (You should give a number here.)

7. Assume that the number of plants of a certain species in a region R has a Poisson distribution with parameter $a\lambda$ where a is the area of the region R. To make inferences about the unknown parameter λ, an ecologist selects n points in the entire region of interest and measures the random variable $Y_i =$ distance from the ith point to the nearest plant.

   a. Find the pdf of the random variable $Y =$ distance from a given point to the nearest plant?
   Hint: $P(Y > y) = P(\text{no plants in a circle of area } \pi y^2)$.

   b. Is there a complete and sufficient statistic for λ? Justify your answer.

   c. Find the maximum likelihood estimator of λ.

8. Let $\bar{X}$ denote the mean of the random sample $X_1, \ldots, X_n$ from a gamma distribution with parameters $\lambda = 3$ and $\beta = \theta > 0$. Compute $E(X_1|\bar{x})$.
   Hint: Can you find a function $\psi(\bar{X})$ of $\bar{X}$ such that $E(\psi(\bar{X})) = \theta$? Is $E(X_1|\bar{x}) = \psi(\bar{x})$? Why?

9. Let $X_1, \ldots X_n$, $n > 2$, be a random sample from the binomial distribution $b(1, \theta)$.

   a. Find the complete sufficient statistic for θ, and call it Y_1.

   b. Find the UMVUE estimator of θ.

   c. Find the UMVUE estimator of θ^2.

   d. Let $Y_2 = (X_1 + X_2)/2$ and compute $E(Y_2)$.

   e. Determine $E(Y_2|Y_1 = y_1)$.

10. Let $X_1, \ldots, X_n$ denote a random sample from a distribution that is $N(\theta, 1)$, where the mean θ is an unknown real number.

    a. Find the form of the uniformly most powerful test of size α of $H_0 : \theta = 0$ vs. $H_A : \theta > 0$. Argue carefully why it is the uniformly most powerful test. Explain in detail how you would find the critical value for the test referring to specific distributions.

b. Find the critical value of the test and the power at $\theta = 1$ if $n = 25$ and $\alpha = 0.05$.

11. Let $X_1, \dots, X_n$ be iid Uniform$(0, \theta)$ for a real number θ.

    a. Find the sufficient statistic. Justify your answer.

    b. Find the pdf of the sufficient statistic you found in part (a).

    c. Find the minimal sufficient statistic if $X_1, \dots, X_n$ were iid Uniform$(\theta, \theta+1)$. Justify your answer.

    d. Derive at the joint pdf of the minimal sufficient statistic you have found in part (c).

12. Suppose we observe $X_1, \dots, X_n$ independent Bernoulli(p) random variables. A popular parameter of interest is $\theta = p/(1-p)$, the odds.

    a. Find the mle estimator of θ.

    b. Find the asymptotic distribution of the mle estimator you have found in part (a). Fully justify your answer.

13. Observations (Y_i, x_i), $i = 1, \dots, n$ are collected according to the model

    $$Y_i = \beta x_i + \epsilon_i,$$

    where ϵ_i are independent and normally distributed with mean zero and variance σ^2. The goal is to estimate the unknown parameter β, where $x_1, \dots, x_n$ are fixed constants, not all equal to zero.

    a. Find the maximum likelihood estimator of β. Compute its mean and variance.

    b. Compute the Cramer-Rao lower bound for the variance of an unbiased estimator of β.

    c. Find the best unbiased estimator (i.e. the UMVUE) of β.

14. Suppose we are interested in estimating $\theta = p^2$, $p \in (0, 1)$, and that we have two different ways of designing an experiment. The first option is to observe $X_1, \dots, X_n$, n independent and identically distributed Bernoulli(p) random variables and then to estimate θ by the maximum likelihood estimate $\hat{\theta}_1 = (\hat{p}_n)^2 = (\bar{X}_n)^2$. The second possibility is to observe $Y_1, \dots, Y_n$, n independent and identically distributed Bernoulli(p^2) random variables and then to estimate θ by $\hat{\theta}_2 = \bar{Y}_n$. Which of the two different methods of estimating θ would you prefer if n is large.

    Hint: Look at the asymptotic distributions of $\sqrt{n}(\hat{\theta}_1 - \theta)$ and $\sqrt{n}(\hat{\theta}_2 - \theta)$.

15. Let X_1 and X_2 be independent and identically distributed from a geometric distribution with pmf $P(X = x) = p(1-p)^{x-1}$ where $x = 1, 2, 3, \ldots$. Find the UMVUE for $1/p$.

16. Let $X_1, \ldots, X_n$ be a random sample from the density $f(x, \theta) = \theta^2 x^{\theta^2 - 1}$, $0 < x < 1$.

    a. Derive the form of the cut region for the UMP test for testing $H_0 : \theta = 1$ versus $H_1 : \theta > 1$. You don't have to find the constant involved.

    b. Does the UMP test for testing $H_0 : \theta = 1$ vs. $H_1 : \theta < 1$ exist? Why?

17. Let $X_1, \ldots, X_n$ denote the observations of a random sample of size $n > 1$ from a binomial distribution, $b(1, \theta)$, $0 < \theta < 1$. Give full justification in the following parts.

    a. (5 points) Find the UMVUE estimator of θ.

    b. (5 points) Suppose $\delta = \theta(1 - \theta)$. Find the mle estimator of δ.

    c. (10 points) Find the asymptotic distribution of the mle estimator of δ.

    d. (5 points) Using the asymptotic distribution you found in part c, form a $(1 - \alpha)100\%$ asymptotic confidence interval for δ.

    e. (15 points) Find the UMVUE estimator of δ.

    f. (5 points) Find the asymptotic distribution of the UMVUE estimator of δ.