UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Basic Exam - Statistics
Wednesday, August 30, 2023

Work all problems. 60 points are needed to pass at the Masters Level and 75 to pass at the Ph.D. level. Each answer is worth approximately the same number of points.

1. Let $X_1, \ldots X_n$ be a random sample from a normal distribution with mean 0 and variance σ^2 unknown.
   (a) (7 points) Find the MLE of σ^2. Is this an UMVUE? Why or why not? If not, find the UMVUE.
   (b) (5 points) Find an exact $(1-\alpha)$-level confidence interval for σ^2.
   (c) (5 points) Find an exact $(1-\alpha)$-level confidence interval for σ.

2. Let $X_1, \ldots, X_n$ denote a random sample of size $n > 2$ from a distribution with a probability density function

$$f_\theta(x) = \theta x^{\theta-1}$$

   where $0 < x < 1$ and $\theta \in (0, \infty)$. Note that $E(W^k) = \frac{(n+k-1)!}{\theta^k (n-1)!}$ for $W = \sum_{i=1}^n (-\log X_i)$ and $k > -n$.
   (a) (4 points) Obtain the Maximum Likelihood Estimator (MLE) for θ, denoted as $\hat{\theta}_n$.
   (b) (4 points) Compute the bias of $\hat{\theta}_n$.
   (c) (4 points) Find an unbiased estimator for θ, denoted as $\tilde{\theta}_n$, using the result in (b).
   (d) (4 points) Compute the variance of $\tilde{\theta}_n$ obtained in (c).
   (e) (4 points) Prove or disprove if $\tilde{\theta}_n$ is efficient.

3. Mendel bred peas with round yellow and wrinkled green seeds. The progeny has four possible outcomes:

$$\{\text{Yellow, Green}\} \times \{\text{Wrinkled, Round}\}.$$

   His theory of inheritance implies that P(Yellow) = 3/4, P(Green) = 1/4, P(Round) = 3/4, and P(Wrinkled) = 1/4, and the color and shape traits are inherited independently. He randomly selected $n = 556$ peas and found among them 315 yellow round peas, 101 yellow wrinkled peas, 108 green round peas, and 32 green wrinkled peas.
   (a) (5 points) Let X_1 be the number of yellow round peas, X_2 be the number of yellow wrinkled peas, X_3 be the number of green round peas, and X_4 be the number of green wrinkled peas, among the $n = 556$ peas. What is an appropriate probability model for (X_1, X_2, X_3, X_4)?
   (b) (5 points) What is the maximum likelihood estimate for (p_1, p_2, p_3, p_4) where p_i's are the probability of each type of pea respectively?

(c) (5 points) What are the null and alternative hypotheses in order to test Mendel's inheritance theory?

(d) (5 points) Derive the likelihood ratio test statistics for the above hypothesis test.

(e) (5 points) Does the data Mendel collected support his theory of inheritance?

4. Let $X_1, \ldots, X_n, X_{n+1}$ be a random (i.i.d.) sample where each X_i is defined by

$$X_i = \begin{cases} 1 \ , & \text{it is sunny on the } i\text{th day} \\ 0 \ , & \text{otherwise} \end{cases}$$

where $i = 1, \ldots, n, n+1$ and $\theta = P(\text{it is sunny on the } i\text{th day})$ is unknown. Assume that $X_i \mid \theta \sim \mathit{Bernoulli}(\theta)$ and a prior distribution for θ is $f(\theta) = 1$.

(a) (7 points) Suppose that the sun has risen for the last n days, i.e. $x_1 = 1, \ldots, x_n = 1$. Derive the exact form of the posterior density for θ, conditional on $\underline{x}_n = (x_1, \ldots, x_n) = (1, \ldots, 1)$.

(b) (7 points) Suppose that the sun has risen for the last n days, i.e. $x_1 = 1, \ldots, x_n = 1$. We wish to test $H_0 : \theta \leq 1/2$ versus $H_a : \theta > 1/2$. Compute the Bayes factor of H_0 by computing the posterior odds and the prior odds of H_0.

(c) (7 points) Show that the posterior predictive probability that the sun will rise on the $(n+1)$th day, given that it has risen for the last n days is $\frac{n+1}{n+2}$.

5. Let $\mathbf{X}_n = (X_1, \ldots, X_n)$ denote a random sample from a Poisson distribution with mean λ:

$$g(x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!},$$

where $0 < \lambda < \infty$. Suppose that Λ has the prior probability distribution

$$h(\lambda \mid a) = \frac{1}{a}\exp\left(-\frac{\lambda}{a}\right),$$

where $a > 0$. The goal is to obtain the posterior distribution of λ given the data $\mathbf{X}_n = \mathbf{x}_n$ only. To this end, we need to estimate a in $h(\lambda \mid a)$ based on $\mathbf{X}_n = \mathbf{x}_n$.

(a) (5 points) Derive the probability density function $m(\mathbf{X}_n \mid a)$ using the following equation

$$m(\mathbf{X}_n \mid a) = \int_0^\infty g(\mathbf{X}_n \mid \lambda)h(\lambda \mid a)d\lambda.$$

[Hint] $\int_0^\infty x^{\alpha-1}\exp(-x/\beta)dx = \Gamma(\alpha)\beta^\alpha$ for $\alpha > 0$ and $\beta > 0$.

(b) (4 points) Estimate the Maximum Likelihood Estimator (MLE) for a, denoted as $\hat{a}_n \equiv \hat{a}(\mathbf{x}_n)$.

(c) (5 points) Obtain the posterior distribution of λ given the data $\mathbf{x}_n$, denoted as $k(\lambda \mid \mathbf{x}_n, \hat{a}_n)$.

(d) (3 points) Compute the Bayes estimator for λ under squared-error loss.