Sample/practice exam Fall 2016, questions

Intermediate Statistics (Carnegie Mellon University)

Practice Final Exam

1. Let $X_1, \ldots, X_n \sim \text{Uniform}(\theta, 2\theta)$ where $\theta > 0$.

   (a) Find a minimal sufficient statistic.

   (b) Find the maximum likelihood estimator $\widehat{\theta}_n$. Show that $\widehat{\theta}_n$ is a consistent estimator.

   (c) Find the method of moments estimator $\widetilde{\theta}_n$ for θ. Find the limiting distribution of $\sqrt{n}(\widetilde{\theta}_n - \theta)$.

2. Let $X_1, \ldots, X_n \sim N(\theta, 1)$.

   (a) Find the Neyman-Pearson test for testing

   $$H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \theta = a$$

   where $a > 0$. Be precise about the critical value. Show that the power tends to 1 as $n \to \infty$ when H_1 is true.

   (b) Find the likelihood ratio test for testing

   $$H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \theta \neq 0.$$

3. Let $X_1, \ldots, X_n$ be iid Multinomial(1,p) random variables where $p = (p_1, p_2, p_3)$. In other words, $X_i \in \{1, 2, 3\}$ and

   $$\mathbb{P}(X_i = 1) = p_1, \quad \mathbb{P}(X_i = 2) = p_2, \quad \mathbb{P}(X_i = 3) = p_3$$

   where $p_1 + p_2 + p_3 = 1$.

   (a) Find the maximum likelihood estimator of p.

   (b) Find the Fisher information matrix.

   Hint: there are really only two parmeters since $p_3 = 1 - p_1 - p_2$. Write the log-likelihood in terms of p_1, p_2. The Fisher information matrix will be a 2-by-2 matrix.

   (c) Use the delta method to find an asymptotic $1 - \alpha$ confidence interval for ψ where

   $$\psi = \log p_1.$$

   (d) Explain how to use the bootstrap to get a confidence interval for ψ. List the steps clearly.

4. Suppose that (Y, X) are random variables where $Y \in \{0, 1\}$ and $X \in \mathbb{R}$. Suppose that

   $$X | Y = 0 \sim \text{Uniform}(0, 5)$$

   and that

   $$X | Y = 1 \sim \text{Uniform}(4, 9)$$

Further suppose that $\mathbb{P}(Y=0)=1/3$ and $\mathbb{P}(Y=1)=2/3$.

(a) Find $m(x)=\mathbb{P}(Y=1|X=x)$.

(b) Find the Bayes classification rule $h_*(x)$.

(c) Find the risk of h_*. (Recall that the risk is $P(Y \neq h_*(X))$.)

5. Recall that the exponential distribution has density

$$p(x;\theta) = \begin{cases} \theta e^{-\theta x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Here, $\theta > 0$ is an unknown parameter. Let $X_1, \ldots, X_n$ be iid from $p(x;\theta)$.

(a) Show that $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is a minimal sufficient statistic.

(b) Find the maximum likelihood estimator $\widehat{\theta}_n$, the score function, the Fisher information, and the limiting distribution of $\sqrt{n}(\widehat{\theta}_n - \theta)$.

6. Let $X \sim \text{Binomial}(n,p)$.

(a) Find the Wald test for testing

$$H_0 : p = \frac{1}{3} \quad \text{versus} \quad H_1 : p \neq \frac{1}{3}.$$

(b) Show how to use BIC to choose between these two models:

Model I: $X \sim \text{Binomial}(n, 1/3)$.
Model II: $X \sim \text{Binomial}(n,p)$ for some $p \in (0,1)$.

(c) Find the large sample, approximate, $1-\alpha$ likelihood ratio confidence interval for e^p.

7. Let $X_1, \ldots, X_n \sim \text{Normal}(\theta, 1)$. Let $\widehat{\theta}_n = a\overline{X}_n$ where $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $0 \leq a \leq 1$ is a fixed constant. Find the mean squared error for $\widehat{\theta}_n$. What value of a gives the smallest mean squared error?

8. Suppose that (Y,X) are random variables where $Y \in \{0,1\}$ and $X \in \mathbb{R}$. Suppose that

$$X|Y=0 \sim \text{Normal}(0,1)$$

and that

$$X|Y=1 \sim \text{Normal}(2,1).$$

Further suppose that $\mathbb{P}(Y=0)=\mathbb{P}(Y=1)=1/2$.

(a) Find $m(x)=\mathbb{P}(Y=1|X=x)$.

(b) Let $\mathcal{A} = \mathcal{A}_1 \bigcup \mathcal{A}_2$ where $\mathcal{A}_1 = \{(\infty, a) :\ a \in \mathbb{R}\}$ and $\mathcal{A}_2 = \{(a,\infty) :\ a \in \mathbb{R}\}$. Find the VC dimension of $\mathcal{A}$.

(c) Let $\mathcal{H} = \{h_A :\ A \in \mathcal{A}\}$ where $h_A(x)=1$ if $x \in A$ and $h_A(x)=0$ if $x \notin A$. Show that the Bayes rule h_* is in $\mathcal{H}$.

9. Recall that the Beta(a, b) density has the form

$$p(y; a, b) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{a-1}(1-y)^{b-1} \quad 0 \leq y \leq 1$$

where

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt$$

is the Gamma function. The mean of the Beta distribution is $a/(a+b)$. Let $Y \sim$ Binomial(n, p). Let p have a Beta(a, b) prior.

(a) Find the posterior distibution.

(b) Find the posterior mean $\overline{p}_n$.