

## Statistics 582, Final Exam

Wellner; 3/15/2010

**Instructions:** This is an “in class” and “closed-book” exam. Please do it completely on your own with no books or notes.

1. (30 points) **Define** any *three* of the following terms. In each case, provide an appropriate context for your definition.
  - (a) An *unbiased test* of  $H : \theta \in \Theta_0$  versus  $K : \theta \in \Theta_1$ .
  - (b) A *similar on the boundary test* of  $H : \theta \in \Theta_0$  versus  $K : \theta \in \Theta_1$ .
  - (c) A *uniformly most powerful unbiased level  $\alpha$  test*.
  - (d) A family of distributions  $\mathcal{P} = \{P_\theta : \theta \in \Theta \subset \mathbb{R}\}$  (with densities  $p_\theta = dP_\theta/d\mu$  with respect to a dominating measure  $\mu$ ) with *monotone likelihood ratio*.
  - (e) A *test with Neyman structure* with respect to a statistic  $T$  sufficient for  $\mathcal{P}_B \equiv \{P_\theta : \theta \in \Theta_B\}$ .
2. (30 points) **State** any *three* of the following results:
  - (a) A theorem about admissibility properties of the sample mean  $\bar{X}$  when sampling from a normal distribution on  $\mathbb{R}$  and a contrasting theorem for sampling from a normal distribution on  $\mathbb{R}^d$  with  $d \geq 3$ .
  - (b) Stein's identity for  $Eg'(X)$  where  $X \sim N_1(0, \sigma^2)$  and  $E|g'(X)| < \infty$ .
  - (c) A theorem relating Bayes rules to minimax rules and least favorable prior distributions.
  - (d) The Wald-Wolfowitz-Noether-Hájek finite sampling central limit theorem.
  - (e) The generalized Neyman - Pearson lemma (in the “short form” stated in the notes).

Do one of problems 3, 4, and 5.

3. (40 points) Suppose that  $X$  has distribution function  $F$  on  $[0, \infty)$ .
- (a) Define the cumulative hazard function  $\Lambda$  in terms of  $F$ .
  - (b) Give a general expression for  $1 - F$  in terms of the cumulative hazard function  $\Lambda$ , and show that it holds in the two cases (i)  $F$  is continuous; (ii)  $F$  is discrete with jumps  $\Delta F_j \equiv F(x_j) - F(x_j^-)$  at  $x_1, \dots, x_J$ .
4. (40 points) **State and prove** the short form of the generalized Neyman - Pearson lemma.
5. (40 points) (a) Suppose that  $Z \sim N(0, 1)$  and  $h : \mathbb{R} \rightarrow \mathbb{R}$  is a differentiable function with  $E|h'(Z)| < \infty$ . Give a heuristic proof of the identity  $E(Zh(Z)) = Eh'(Z)$  using integration by parts.
- (b) Suppose that  $X \sim N(\theta, \sigma^2)$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$  is a differentiable function satisfying  $E|g'(X)| < \infty$ . Use the result of (a) to prove Stein's identity.

Do **either** problem 6 **or** problem 7.

6. (40 points) Consider the Beta( $a, b$ ) family of densities given by

$$p(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} 1_{(0,1)}(x), \quad a, b > 0. \quad (1)$$

Suppose that  $X \sim p(\cdot; a, b)$ . Consider testing:

$$\begin{aligned} H_1 : a = b = 1 \quad \text{versus} \quad K_1 : a = b = 2; \\ H_2 : a = b \text{ (unspecified)} \quad \text{versus} \quad K_2 : b > a. \end{aligned}$$

- (a) Is the family of densities (1) an exponential family? Why or why not?
- (b) For  $n = 1$  (or just one  $X$  with the distribution given in (1)), what is the most powerful test of  $H_1$  versus  $K_1$  of size  $\alpha = .1$ ?
- (c) For  $n = 1$  (or just one  $X$  with the distribution given in (1)), what theory do we have to yield tests of  $H_2$  versus  $K_2$  which are unbiased? Construct a uniformly most powerful test of  $H_2$  versus  $K_2$  of size  $\alpha = .2$ .
- (d) Now suppose that  $X_1, \dots, X_n$  are i.i.d. with density given by (1). Give a reasonable test of  $H_2$  versus  $K_2$  based on  $X_1, \dots, X_n$ .

7. (40 points) Suppose that  $X$  is a random variable with density  $p(\cdot; \theta)$  given by

$$p(x; \theta) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}} \cdot \frac{\theta}{1 - (1 - \theta^2)x} 1_{(0,1)}(x), \quad \theta \in (0, \infty). \quad (2)$$

This is the distribution of the total time spent in  $(0, \infty)$  up to time  $t = 1$  by a skew Brownian motion process with skewing parameter  $\theta = \sigma_+/\sigma_-$  where  $\sigma_+^2$  is the variance parameter for the positive space axis and  $\sigma_-^2$  is the variance parameter for the negative space axis. Note that for  $\theta = 1$ ,

$$p(x, 1) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}} 1_{(0,1)}(x)$$

is the Beta(1/2, 1/2) density corresponding to the arcsin distribution

$$F_1(x) = P_1(X \leq x) = \frac{2}{\pi} \arcsin(\sqrt{x}).$$

The mean and variance of  $X$  with density  $p(\cdot; \theta)$  are  $E_\theta X = 1/(1 + \theta)$  and  $Var_\theta(X) = 2^{-1} \frac{1}{1+\theta} \cdot \frac{\theta}{1+\theta}$ .

(a) Show that the family  $\mathcal{P} = \{p(\cdot; \theta) : \theta \in (0, \infty)\}$  has monotone likelihood ratio in  $T(x) = 1/x$ . [Hint: rewrite the last factor in (2) in terms of  $1/x$ .]

(b) Find the UMP size .05 test of  $H : \theta \leq 1 \equiv \theta_0$  versus  $K : \theta > 1$ ? Specify your test completely, including the constant(s).

(c) Compute the power function of the UMP test in (b) as explicitly as possible.

(d) Now suppose that  $X_1, \dots, X_n$  are i.i.d. with density  $p(\cdot; \theta)$ . Find the form of the locally most powerful test of  $H : \theta \leq 1$  versus  $K : \theta > 1$  based on  $X_1, \dots, X_n$ , and use the central limit theorem to find appropriate constants so that your test has approximate size  $\alpha = .05$ .

Do **either** problem 8 **or** problem 9.

8. (40 points). Suppose that conditional on  $\boldsymbol{\theta} = \theta$ ,  $X \sim \text{Uniform}(0, \theta)$  and that  $\boldsymbol{\theta} \sim \text{Gamma}(2, \mu)$ ; i.e.  $\boldsymbol{\theta}$  has density  $\lambda(\theta) = \mu^2 \theta \exp(-\mu\theta) 1_{(0, \infty)}(\theta)$  for  $\theta > 0$ .
- Find the marginal density of  $X$  and the conditional density of  $\boldsymbol{\theta}$  given  $X = x$ .
  - What is the Bayes estimator of  $\theta$  (for squared error loss),  $d_\mu(X) \equiv d_{\lambda_\mu}(X)$ ?
  - Find the maximum likelihood estimator of  $\mu$  based on the marginal distribution of  $X$ , and use this to find the empirical Bayes estimator  $d_{EB}(X) \equiv d_{\hat{\lambda}}(X)$  of  $\theta$ .
  - Compute and compare the risks of the maximum likelihood estimator  $d_{ML}$ , the Bayes estimator  $d_B$ , and the empirical Bayes estimator  $d_{EB}$ .
9. (40 points). Suppose that  $X_1, \dots, X_m$  are i.i.d.  $\text{Poisson}(\mu)$  and  $Y_1, \dots, Y_n$  are i.i.d.  $\text{Poisson}(\nu)$  random variables independent of the  $X_i$ 's. Consider testing  $H : \mu \geq \nu$  versus  $K : \nu > \mu$ .
- Find the joint distribution of  $(\underline{X}, \underline{Y})$  and put it in exponential family form. What are the marginal distributions of  $R \equiv \sum_1^m X_i$ ,  $S \equiv \sum_1^n Y_j$ , and  $T = R + S = \sum_i X_i + \sum_j Y_j$ ?
  - Find the conditional distribution of  $S = \sum_1^n Y_j$  given  $T = t$  and specialize this to the case when  $\mu = \nu$ .
  - Find the UMP unbiased test of  $H$  versus  $K$  of size  $\alpha$  and describe how you would find the constants involved as explicitly as possible.
  - Show that you can rewrite the test you derived in (c) as

$$\phi(\underline{X}, \underline{Y}) = 1\{V(\underline{X}, \underline{Y}) > \tilde{c}(T)\} + \gamma(T)1\{V(\underline{X}, \underline{Y}) = \tilde{c}(T)\}$$

where

$$V(\underline{X}, \underline{Y}) \equiv \frac{\sum_{j=1}^n Y_j - T \frac{n}{N}}{\sqrt{T \frac{n}{N} \frac{m}{N}}} = \frac{\sqrt{\frac{mn}{N}}(\bar{Y}_n - \bar{X}_m)}{\sqrt{T/n}}.$$

- Use (d), the CLT, and the WLLN to show that if the constant in (d) is taken to be  $z_\alpha \equiv \Phi^{-1}(1 - \alpha)$ , then the test in (d) has approximate size  $\alpha$  for large  $N = m + n$ .