Math 541b: Introduction to Mathematical Statistics
Homework 3 (due on Fri 12/02)

Disclaimer: problem 1 is an important exercise in theory; problem 2 includes a calculus exercise; finally, problems 3 to 6 are especially relevant for the final exam.

Please submit electronically directly to Blackboard in a PDF file.

1^o (Algebraic structure of exponential families). Consider a generic one-parameter exponential family, with density[1]

$$p_\theta(x) = c(\theta) \exp\left(T(x)\theta\right)$$

with respect to some dominating measure $\mu(dx)$ supported on $\mathcal{X} \subseteq \mathbb{R}$. Prove the following results:

1. The *canonical domain* $\Theta := \{\theta \in \mathbb{R} : c(\theta) > 0\}$ is a convex set (i.e., a segment—since $\Theta \subseteq \mathbb{R}$). In other words, if $\theta_0, \theta_1 \in \Theta$, then also $\frac{1}{2}(\theta_0 + \theta_1) \in \Theta$.

2. The *log-cumulant* function $a(\theta) = -\log(c(\theta))$ is convex on its its domain Θ, i.e.: $\forall \theta_0, \theta_1 \in \Theta$,

$$a(\tfrac{1}{2}(\theta_0 + \theta_1)) \le \tfrac{1}{2}(a(\theta_0) + a(\theta_1)).$$

   You can either prove this directly, or use that convexity of $a(\cdot)$ on Θ is equivalent to having $a''(\theta) \ge 0 \ \forall \theta \in \Theta$, assuming $c(\theta)$ is smooth enough, and differentiate in θ under the integral.

3. Show that $a'(\theta) = \mathbb{E}_\theta[T(X)]$ and $a''(\theta) = \mathrm{Var}_\theta[T(X)]$ where $\mathbb{E}_\theta[\cdot]$ and $\mathrm{Var}_\theta[\cdot]$ are the expectation and variance over the distribution with density p_θ. *You can skip item 2 if you do this—why?*

2^o (UMP test for the ratio of independent exponentially distributed r.v.'s). Revisiting the setup of problem **3^o** in Homework 1, consider $Z = X/Y$ where $X \sim \mathsf{Exp}(\lambda)$ and $Y \sim \mathsf{Exp}(\mu)$ are independent. Recall that the p.d.f. of Z is

$$f_\theta(z) = \frac{\theta}{(z+\theta)^2}$$

where $\theta = \mu/\lambda$, and the corresponding c.d.f. is $F_\theta(z) = 1 - \frac{\theta}{\theta + z}$.

1. Is this family (parametrized by $\theta \in \mathbb{R}^+$) an exponential family? (Justify your answer.)

[1] We assume that the initial parametrization is already a canonical one, and the factor $h(x)$ is merged into $\mu(dx)$.

2.* Consider testing $H_0 : \theta \le 1$ vs. $H_1 : \theta > 1$ from the sample $Z_{1:n} = (Z_1, ..., Z_n)$ with the test

$$\varphi(z_{1:n}) = \begin{cases} 1 & \text{if } \prod_{i=1}^{n}(1+z_i) > e^s, \\ 0 & \text{otherwise,} \end{cases}$$

where $s \ge 0$ is the threshold parameter. Show that for any $s \ge 0$, this test is unbiased.

Hint: try to show that the power function $\beta_\varphi(\theta)$ is increasing. This is a nice exercise!

In what follows, let $s = s(\alpha)$ be such that the corresponding test has size α. Next we study it.

3. Show that $s = s(\alpha)$ is the α-quantile of the law $\mathsf{Gamma}(n, 1)$ where the pdf of $\mathsf{Gamma}(k, \lambda)$ is

$$f_{k,\lambda}(u) = \frac{\lambda^k}{(k-1)!} u^{k-1} e^{-\lambda u}, \quad u \ge 0.$$

   (i) Derive the distribution of $U_i = \log(1 + Z_i)$ under $\theta = 1$. *(Hint: when transforming a random variable, it is often more convenient to think in terms of cdf's rather than pdf's.)*

   (ii)* Show that the sum of n independent r.v.'s, each distributed as $\mathsf{Exp}(\lambda)$, is $\mathsf{Gamma}(n, \lambda)$.

   (iii) Conclude by rewriting $\varphi(Z_{1:n})$ appropriately and using the previous results.

4. Using CLT, find an approximation of $s(\alpha)$ in terms of an $\mathcal{N}(0, 1)$ quantile when $n \to \infty$.

3^o (Student's t-test is UMP unbiased for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \ne \mu_0$ in the full Gaussian family.) In this problem we shall derive the UMP unbiased test for testing

$$H_0 : \mu = \mu_0 \text{ against } H_1 : \mu \ne \mu_0$$

observing an i.i.d. sample $X_{1:n} := (X_1, ..., X_n)$ from $\mathcal{N}(\mu, \sigma^2)$ with unknown variance.[2] The test will be derived via the conditioning method.

1. Without loss of generality we can assume that $\mu_0 = 0$ (*why?*). Show that (U, V) with $U := \sum_{i=1}^n X_i$ and $V := \sum_{i=1}^n X_i^2$ is a complete sufficient statistic for the whole family.

2. Explain why it suffices to consider tests depending only on (U, V) (rather than on $X_{1:n}$).

3. Find the boundary family and show that V is a complete sufficient statistic for that family. (You can use the general result for exponential families stated in the class without a proof.)

4. Next, we shall consider a new, possibly simpler, problem by conditioning on the value of V:

$$\begin{aligned} H_0'(v) &: U|\{V = v\} \sim \mathbb{P}_\tau(v) \quad \tau \in \mathcal{T}_0 \;\; [\subseteq \mathbb{R}] \\ H_1'(v) &: U|\{V = v\} \sim \mathbb{P}_\tau(v) \quad \tau \in \mathcal{T}_0^c \;\; [= \mathbb{R} \setminus \mathcal{T}_0]. \end{aligned} \tag{1}$$

   Here $(\mathbb{P}_\tau)_{\tau \in \mathbb{R}}$ is the "induced" family of conditional distributions of $U|\{V = v\}$ with a *one-dimensional* parameter $\tau = \tau(\mu, \sigma^2) \in \mathbb{R}$; the set $\mathcal{T}_0$ corresponds to the null H_0 as follows:

$$\mathcal{T}_0 := \{\tau(0, \sigma^2) \;\text{ for }\; \sigma^2 \in \mathbb{R}_{++}\}.$$

[2] This is a companion problem to the one considered in the class, where we tested (one-sided) hypotheses about σ^2.

As we shall see next, this "conditioned" problem is easy because the map $\tau(\mu, \sigma^2)$ is such that $\tau(0, \sigma^2)$ *does not* depend on σ^2, so $\mathcal{T}_0$ is a singleton, and $H_0'(v)$ is a *simple* hypothesis.

Before delving into analysis, explain why finding a UMP unbiased test of size α for every such problem (for each $v > 0$) would give a solution to the initial problem. *(Appeal to the general results we proved in class about the connections between the UMP unbiased, UMP similar on the boundary (SOB), and UMP Neyman-structured (NS) size α tests.)*

5. Now we analyze the conditioned problem. Show that the joint density of (U, V) is of the form

$$p_{U,V}(u, v) = g(\tau, \nu) \exp(\tau u + \nu v)\, h(u, v)$$

for some functions $g(\tau, \nu)$ and $h(u, v)$ (that are not important), where $\tau = \tau(\mu, \sigma^2)$, $\nu = \nu(\mu, \sigma^2)$ is the canonical parametrization of the corresponding (2-parameter) exponential family. Write down τ and ν explicitly in terms of μ, σ^2.

6. Derive, from the above, that the corresponding family of conditional densities has the form

$$p_{U|V}(u|v) = \widetilde{g}_v(\tau) \exp(\tau u)\, \widetilde{h}_v(u)$$

for some functions $\widetilde{g}_v(\cdot), \widetilde{h}_v(\cdot)$ that might depend on v (their form is not important), i.e. they form a 1-parameter exponential family with canonical parameter $\tau = \tau(\mu, \sigma^2)$. From the explicit form of $\tau(\mu, \sigma^2)$ conclude that the conditioned null hypothesis $H_0'(v)$ is simple, i.e., $\mathcal{T}_0 = \{\tau_0\}$ for some τ_0.

7. Conclude, from the general result on the form of a UMP unbiased test for a 1-parameter exponential family, that a UMP unbiased test of level α for the conditioned problem (1) reads

$$\phi^*(U, V) = \begin{cases} 1, & |U| > k(V), \\ 0, & |U| < k(V) \end{cases} \tag{2}$$

for $k(V)$ defined by the condition $\mathbb{E}_{\tau=\tau_0}[\phi^*(U, V)|V] = \alpha$ (*a.s.* over V).

8. Finally, we would like to reformulate test (2) such as the threshold does *not* depend on V.

   (i) First, observe that

   $$|U| > k(V) \iff \frac{|U|}{\sqrt{V}} > k' := \frac{k(V, \alpha)}{\sqrt{V}},$$

   and $\frac{U}{\sqrt{V}}$ is independent from V when $\mu = 0$ (i.e., under the null). (Justify this!). Conclude that we can select $k' = k'(\alpha)$ independently of V.

   (ii) Show that the just obtained test

   $$\phi(U, V) = \mathbb{1}\left\{\frac{|U|}{\sqrt{V}} > k'(\alpha)\right\}$$

   can, in fact, be expressed in the form

   $$\phi(U, V) = \mathbb{1}\left\{\frac{|U|}{\sqrt{nS^2}} > \widetilde{k}(\alpha)\right\}$$

   where $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ is the familiar unbiased estimate of σ^2. *Hint:* express $\frac{|U|}{\sqrt{nS^2}}$ as $\psi(\frac{|U|}{\sqrt{V}})$ for some one-to-one function $\psi(\cdot)$; then, clearly, we can take $\widetilde{k}(\alpha) := \psi(k'(\alpha))$.

(iii) Using the properties of the standard (multivariate) Gaussian distribution $\mathcal{N}(0, I)$, show that $W = \frac{U}{\sqrt{nS^2}}$ has the Student distribution with $n-1$ degrees of freedom under H_0. (By definition, this is the law of $\frac{Z}{Y}$ where $Z \sim \mathcal{N}(0,1)$, $Y \sim \chi^2_{n-1}$, and Z, Y are *independent.*) Express $\widetilde{k}(\alpha)$ in terms of the corresponding quantile (don't forget about the abs. value).

4^o (Permutation test). Let $X_1, ..., X_n, Y_1, ..., Y_m$ be a sequence of random variables. Find the permutation test for testing

$H_0 : X_1, ..., X_n, Y_1, ..., Y_m$ are i.i.d.

against the alternative

$H_1 : X_1, ..., X_n, Y_1, ..., Y_m$ are independent; $X_1, ..., X_n \sim \mathsf{Poisson}(\lambda)$, $Y_1, ..., Y_m \sim \mathsf{Poisson}(\mu)$ with $\lambda > \mu > 0$.

5^o (Likelihood ratio test for $\mathsf{Unif}([0, \theta])$). Let $X_1, ..., X_n$ be an i.i.d. sample from the uniform distribution $\mathsf{Unif}([0, \theta])$ with $\theta > 0$. We test $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ for some (known) θ_0.

1. Derive the likelihood ratio test for this problem.
2. Find the distribution of the LRT statistic.
3. In few words, explain why the χ^2 asymptotics derived in class does not apply here.

6^o (Likelihood ratio test for the Bernoulli parameter). Let $X_i \sim \mathsf{B}(p)$ (Bernoulli distribution) with success probability $p \in (0, 1)$ and let $X_{1:n} = (X_1, ..., X_n)$ be an i.i.d. sample. Denote $f_p(x)$ the corresponding p.m.f. (with $x \in \{0, 1\}$). The sample average $\hat{p}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is a reasonable point estimate of p, converging to p in almost surely by the strong law of large numbers.

1. Derive $\hat{p}_n$ as the MLE of p. What can be said about the *rate* of convergence of $\hat{p}_n$ to p in terms of n? (Appeal to the asymptotic result about the asymptotics of MLE in regular families.)
2. Express the family of distributions $\{B(p),\ p \in (0, 1)\}$ as a full exponential family with $\Theta \subseteq \mathbb{R}^1$.
3. Now, consider testing $H_0(p_0) : p = p_0$ against $H_1(p_0) : p \neq p_0$ for some fixed $p_0 \in (0, 1)$. Derive an explicit expression for the likelihood ratio test (LRT) statistic

$$\Lambda_n(p_0|\hat{p}_n) := n(L_n(\hat{p}_n) - L_n(p_0)),$$

where $L_n(p) = \frac{1}{n}\sum_{i=1}^n \log f_p(X_i)$ is the empirical log-likelihood, in terms of $\hat{p}_n$ and p_0.
4. Use the theorem about the asymptotic distribution of $2\Lambda_n$ stated in class to select the threshold that gives a test with asymptotic size α. Express the threshold in terms of a quantile of $\mathcal{N}(0, 1)$.