Math 541b: Introduction to Mathematical Statistics
Homework 4 (due on Mon 12/12)

Please submit electronically directly to Blackboard in a PDF file.

1^o (Continuation of 6^o from Homework 3: confidence intervals for the Bernoulli parameter via test inversion).

6. Inverting the just constructed size α test would give us a $(1-\alpha)$-level confidence set (using a theorem proved in class). Unfortunately, Λ_n, as a function of p_0 with given $\hat{p}_n$ cannot be inverted *explicitly* (its inverse function can be obtained by solving a non-linear equation). Hence, the corresponding confidence sets are not explicit.
   - Check that the function $h_n(p) := \Lambda_n(p|\hat{p}_n)$ is strictly convex on $(0,1)$, and $h_n(\hat{p}_n) = 0$. Conclude that the corresponding confidence set $C(X_{1:n})$ is a segment that contains $\hat{p}_n$ (level sets of a convex function are convex). Which two equations define its borders?
   - Use a suitable Taylor expansion of $h(p)$ around $\hat{p}_n$, along with the result of item 4, to obtain an approximate $(1-\alpha)$-level confidence interval in the form $[\hat{p}_n - \delta_n, \hat{p}_n + \delta_n]$ with
   $$\delta_n = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}},$$
   $z_{1-\alpha/2}$ being a quantile of $\mathcal{N}(0,1)$. Explain why this approximation is satisfactory (cf. 1).

2^o (Two-sample testing for the exponential distribution). Let $X_{1:n} := (X_1, ..., X_n)$ and $Y_{1:n} := (Y_1, ..., Y_n)$ be sampled i.i.d. from $\mathsf{Exp}(\lambda)$ and $\mathsf{Exp}(\mu)$ respectively, and independent from each other; recall that $\mathsf{Exp}(\lambda)$ has density $f_\lambda(z)\lambda e^{-\lambda z}\mathbb{1}\{z \ge 0\}$ with $\lambda > 0$. Assuming λ, μ unknown, consider testing
$$H_0 : \lambda = \mu \quad \text{vs.} \quad H_1 : \lambda \neq \mu.$$
Define $N := 2n$ and $Z_{1:N} := (X_{1:n}, Y_{1:n})$. Our final goal is to find a UMP unbiased size α test.

1. Write the density of $Z_{1:N}$ explicitly to show that this is an exponential family parametrized by $\theta = (\lambda, \mu) \in \mathbb{R}_+^2$. Is this parametrization a canonical one? Justify your answer, and write down some complete sufficient statistic $T(Z_{1:N}) = (T_1(Z_{1:N}), T_2(Z_{1:N}))$.

2. Now, pass to a canonical parameterization $\eta(\theta) = (\tau(\theta), \nu(\theta))$ where $\tau = \tau(\theta)$ and $\nu = \nu(\theta)$ are such that the hypotheses depend only on τ, but not on ν. (If the previously found parametrization already satisfies this requirement, I do *not* ask you to find yet another one.)

3. Using the parametrization $\eta(\lambda, \mu)$ you have just found:

(*i*) Write down $V(Z_{1:N})$ and $U(Z_{1:N})$—respectively, (complete) sufficient statistics for the boundary family and for the "conditioned" family $\mathbb{P}_\tau(Z_{1:N}|V = v)$ parametrized by τ.

(*ii*)* Find the general form of a UMP unbiased test in each conditioned problem. To this end, show that the asymptotic distribution of $U|V = v$ under H_0 has density

$$f_{U|V}(u|v) = \frac{1}{c_n v}\left(1 - \frac{u^2}{v^2}\right)^{n-1} \mathbb{1}\{|u| \le v\} \quad \text{with } c_n = \int_{-1}^{1} \left(1 - t^2\right)^{n-1} dt.$$

and use this to specify, for any given α and v, the test with size α.

Hint. *You may find useful the following facts (no need to prove them):*

(a) $\mathsf{Exp}(\lambda)$ *is the same as* $\mathsf{Gamma}(1, \alpha)$, *where* $\mathsf{Gamma}(n, \alpha)$ *has density*

$$f_{n,\lambda}(z) = \lambda^n z^{n-1} e^{-\lambda z} \mathbb{1}\{z \ge 0\}$$

for $n \in \mathbb{N}$ *and* $\lambda > 0$.

(b) If $T_1 \sim \mathsf{Gamma}(n, \lambda)$ *and* $T_2 \sim \mathsf{Gamma}(m, \lambda)$ *are independent, then*

$$T_1 + T_2 \sim \mathsf{Gamma}(n + m, \lambda).$$

(c) The joint distribution of $W = (T_1 - T_2, T_1 + T_2)$ *can be derived from that of* $T = (T_1, T_2)$, *since* $T \mapsto W(T)$ *is a bijective linear map.*

(*iii*) Argue, as we did in the class, that the test in which v is replaced with the boundary-family sufficient statistic $V = V(Z_{1:N})$ is UMP unbiased (of its size—which depends on v) for the initial problem. You don't have to write a lot here—just give a brief explanation.

3^o (Local behavior of f-divergences). In this exercise you are invited to show that f-divergence with a *strictly convex* function f locally behaves as the χ^2-divergence. Namely, assume that $f : \mathbb{R}_{++} \to \mathbb{R}$ (where $\mathbb{R}_{++}$ is the set of all positive reals) satisfies the following assumptions:

- $f(1) = 0$;
- uniformly bounded third derivative on $\mathbb{R}_{++}$, that is f''' exists on $\mathbb{R}_{++}$ and $\sup_{r>0} |f'''(r)| < \infty$;
- f is strictly convex (and thus by the previous assumption $f''(r) > 0$ for any $r > 0$).

In fact, all common f-divergences, except the TV distance, satisfy these assumptions (including Hellinger, chi-squared, and Kullback-Leibler). Recall that the associated f-divergence between two distributions P, Q on the same space, with densities p, q with respest to a dominating measure μ, is

$$D_f(P||Q) := \mathbb{E}_Q\left[f\left(\frac{dP}{dQ}\right)\right] = \int_{\mathcal{X}} f\left(r(x)\right) q(x) d\mu(x),$$

where $r(x) := \frac{p(x)}{q(x)}$ is the likelihood ratio and $\mathcal{X}$ is the support of μ. Fix P and Q, and consider the "segment" between them, that is, the family of distributions $P_t := (1-t)Q + tP$ for $t \in [0, 1]$ (in particular, $P_1 = P$ and $P_0 = Q$).

- Show that as $t \to 0$,

$$D_f(P_t||Q) = (1 + o(1))\frac{f''(1)}{2}\chi^2(P_t||Q)$$

where $o(1) \to 0$ and $\chi^2(P||Q)$ is the chi-square divergence, i.e. $D_h(P||Q)$ with $h(r) = (1-r)^2$.

- Check that $\chi^2(P_t||Q) = t^2\chi^2(P||Q)$ and conclude that $D_f(P_t||Q)$ is locally quadratic in t.

Hint: *Consider the 3rd-order Taylor expansion of $f(r)$ at $r = 1$. The 1st-order term must vanish.*