Stats 210A, Fall 2024
Homework 8

Due on: Wednesday, Oct. 30

Instructions: See the standing homework instructions on the course web page

Problem 1 (Some two-tailed tests). Consider testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ in a one-parameter exponential family of the form $p_\theta(x) = e^{\theta T(x) - A(\theta)} h(x)$. We stated in class that among all *unbiased, level-α* tests, the one that rejects for extreme (i.e., large or small) values of $T(X)$ is uniformly most powerful (simultaneously maximizes power for all alternatives).

The equal-tailed level-α test that rejects for extreme values of $T(X)$ does not satisfy as interesting an optimality property but it is also a competitive test. Depending on the distribution, the equal-tailed test and the UMPU test may or may not coincide.

Numerically find the equal-tailed and UMPU test for the following hypothesis testing problems at level $\alpha = 0.05$. For each problem,

(i) derive the appropriate tests (leaving the cutoff values abstract),

(ii) numerically compute the cutoff values c (no γ necessary since these are continuous problems), and

(iii) invert the equal-tailed test to give an interval for the data value specified (no need to invert the unbiased test).

(a) $X_i \overset{\text{ind.}}{\sim} N(\theta, \sigma_i^2)$ for $i = 1, \ldots, n$, where σ_i^2 are known positive constants and $\theta \in \mathbb{R}$ is unknown. Test $H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$, with $n = 20$ and $\sigma_i^2 = i$. On your power plot, also plot the power function of the (sub-optimal) test that rejects for extreme values of $\sum_i X_i$.

(b) $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \text{Pareto}(\theta) = \theta x^{-(1+\theta)}$, for $\theta > 0$ and $x > 1$ (also called a power law distribution). Test $H_0 : \theta = 1$ vs. $H_1 : \theta \neq 1$, for $n = 100$. On your power plot, also plot the power function of the (sub-optimal) test that rejects for large $\sum_i X_i$.

Problem 2 (Testing in trinomial subfamilies). Recall the problem on a previous homework on subfamilies of the trinomial distribution
$$p_\pi(x) = \pi_1^{x_1} \pi_2^{x_2} \pi_3^{x_3} \cdot \frac{n!}{x_1! x_2! x_3!},$$
and recall that the probability of genotypes **aa**, **Aa**, and **AA** were θ^2, $2\theta(1 - \theta)$, and $(1 - \theta)^2$, respectively, for an individual drawn from a well-mixed population with prevalence $\theta \in (0, 1)$ for allele **a**.

You may appeal to any results from the previous homework problem that are helpful, without re-deriving them.

(a) Suppose we observe a sample of size $n = 1000$ from a well-mixed population with prevalence θ, and we want to test $H_0 : \theta \leq 0.1$ vs $H_1 : \theta > 0.1$ at level $\alpha = 0.1$. Show that there is a UMP test for this problem, and say what the test statistic is. Give the cutoffs c, γ and plot the power as a function of θ, for an appropriate range of values.

(b) Now suppose our population is a mixture of two populations with different prevalences θ_1 and θ_2 for allele **a**. Define $\gamma \in [0, 1]$ as the proportion of individuals from population 1, so the proportion of genotype $i = 1, 2, 3$ is $\pi_i(\gamma) = \gamma\pi_i(1) + (1 - \gamma)\pi_i(0)$. Assume that θ_1, θ_2 are known and only γ is unknown.

Consider testing $H_0 : \gamma = 0$ vs $H_1 : \gamma = \gamma_1$, for some fixed value $\gamma_1 > 0$. That is, we are testing for whether the individuals we are sampling from are coming only from population 2, or whether a nonzero fraction of them are coming from population 1. Show that the likelihood ratio test rejects for large values of the statistic $\sum_{i=1}^3 w_i X_i$, where

$$w_i(\gamma_1) = \log\left(1 + \gamma_1 \left(\frac{\pi_i(1) - \pi_i(0)}{\pi_i(0)}\right)\right).$$

(c) Next consider testing $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$, when n is large. Explain why no UMP test exists, and suggest instead an appropriate test for $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$ that prioritizes small alternative values of γ.

(d) Implement your test from part (c) when $\theta_1 = 0.7$, $\theta_2 = 0.05$, and $n = 1000$, at level $\alpha = 0.1$, and compare its power to the Neyman–Pearson test for $H_0 : \gamma = 0$ vs $H_1 : \gamma = 1$. Give an explicit expression for both test statistics and comment qualitatively on how they are examining the data differently. Plot their power curves for an appropriate range of γ values.

Moral: We saw before that, while a one-parameter exponential family like the one in part (a) has a complete sufficient statistic that allows for optimal inference throughout the parameter space, a curved exponential family like the mixture family in parts (b-d) has no complete sufficient statistic. Instead, the score acts like a complete sufficient statistic in a local neighborhood of the parameter space, but the score is different in different parts of the parameter space. Hence, the structure of the family has important ramifications for how we should think about statistically efficient inference.

Problem 3 (Confidence intervals for quantiles). Assume we observe $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} F$, where the cdf F is assumed to be strictly increasing and continuous. Consider inference on the quantile $q(F) = F^{-1}(p)$, for a fixed $p \in (0, 1)$.

(a) Suggest a nonparametric level-α test for $H_0 : q(F) = q_0$ vs $H_1 : q(F) \neq q_0$. Since α and p are unspecified, you don't need to solve for any cutoff values, but describe how you might do so.

(b) For $n = 1000$ and $p = 0.9$, invert your test from part (a) to find an explicit confidence interval for $q(F)$, as a function of $X_1, \ldots, X_n$.

Problem 4 (Testing with multidirectional alternatives). Suppose $X \sim N_d(\mu, I_d)$ for unknown $\mu \in \mathbb{R}^d$. Consider testing $H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$.

(a) If $d = 1$, elaborate on the UMPU proof from class to show that the usual two-sided test $\phi^*(X) = 1\{|X| > z_{\alpha/2}\}$ is the only UMPU test up to almost sure equality.

(b) Show that for any $d > 1$ and $\alpha \in (0, 1)$, there exists no UMP or UMPU level-α test.

Hint: what would we do if we knew $\mu = (\theta, 0, 0, \ldots, 0)$ for an unknown $\theta \in \mathbb{R}$?

(c) Suppose we have a prior Λ_1 for the value that μ takes under the alternative; that is, $\mu \sim \Lambda_1$ if H_1 is true and $\mu = 0$ if H_0 is true. Define the average power as

$$\int_{\mathbb{R}^d} \mathbb{E}_\mu[\phi(X)] \, d\Lambda_1(\mu).$$

If $\Lambda_1 = N(\nu, \Sigma)$, with positive definite covariance matrix Σ, find the level-α test that maximizes the average power. Show that the acceptance region is an ellipse centered at 0 if $\nu = 0$.

(d) **Optional:** (Not graded, no extra points) Suppose the prior Λ_1 (not necessarily multivariate Gaussian) is rotationally invariant, meaning $\Lambda_1(QA) = \Lambda_1(A)$ where $A \subseteq \mathbb{R}^d$, Q is any rotation matrix and $QA = \{Qa : a \in A\}$. Show that the χ^2 test that rejects for large $\|X\|^2$ maximizes the average power.

Moral: Choosing a test in higher dimensions requires us to think harder about how to compromise across different alternative directions, and Bayesian thinking can give us some guidance.

Problem 5 (p-value densities)**.** Suppose $\mathcal{P}$ is a family with monotone likelihood ratio in $T(X)$, and the distribution of $T(X)$ is continuous with common support for all θ. Let ϕ_α denote the UMP level-α test of $H_0 : \theta \leq \theta_0$ vs. $H_0 : \theta > \theta_0$ that rejects when $T(X)$ is large. Let $p(X)$ denote the resulting p-value. Show that $p(X) \sim \text{Unif}[0, 1]$ if $\theta = \theta_0$, has non-increasing density on $[0, 1]$ if $\theta > \theta_0$, and has non-decreasing density on $[0, 1]$ if $\theta < \theta_0$.

Note: As always there is technically some ambiguity in how we could define the density, but the version we want is the derivative of the cdf. Feel free to work more informally and ignore such technical issues.