Stats 210A, Fall 2023
Homework 8

Due date: Wednesday, Oct. 25

1. Directional error claims

Suppose $\mathcal{P} = \{P_\theta : \theta \in \mathbb{R}\}$, and $T(X)$ is a continuous test statistic that is stochastically increasing in θ, meaning

$$\mathbb{P}_{\theta_1}(T(X) \le t) \le \mathbb{P}_{\theta_0}(T(X) \le t), \quad \text{for all } t \in \mathbb{R} \text{ and } \theta_1 > \theta_0.$$

As we have discussed in class, this property guarantees that a one-tailed test of $H_0 : \theta \le 0$ vs. $H_1 : \theta > 0$ that rejects for large values of $T(X)$, with cutoff c chosen to solve $\mathbb{P}_0(T(X) > c) = \alpha$, will give a valid level-α test over the whole null distribution.

Now let $\phi(X)$ represent any level-α *two-tailed* test of $H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$ that rejects for extreme values of $T(X)$.

Assume that we also always make a *directional claim* about $\text{sign}(\theta)$ whenever we reject H_0. That is, we make one of *three* decisions: if $T(X) \in [c_1, c_2]$ then we do not reject H_0 (and we make no claim about the sign of θ); if $T(X) > c_2$, then we reject H_0 and claim further that $\theta > 0$; and if $T(X) < c_1$, then we reject H_0 and claim further that $\theta < 0$. So now there are two kinds of Type I errors we could make: we could reject H_0 when it is really true, *or* we could reject H_0 when it is really false but call the sign wrong. Show that we have control of the *directional error rate*:

$$\sup_{\theta \in \mathbb{R}} \mathbb{P}_\theta(\text{False rejection or wrong sign call}) \le \alpha$$

Moral: Having MLR or exponential families are nice to be able to talk about optimal tests, but stochastically increasing is a useful condition for getting valid tests in various contexts.

In particular, people often complain that we do not learn anything about θ by rejecting $H_0 : \theta = 0$, because we should have already known θ was not exactly zero. This line of argument ignores the fact that (in most testing settings) we can also draw a definite conclusion about the sign of θ whenever we reject $H_0 : \theta = 0$, without inflating the error rate.

2. 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 $\sum_i X_i$.

3. Maximizing average power

In situations where there is not a UMP test, we cannot simultaneously maximize power for all alternatives. However, if the null is simple ($\Theta_0 = \{\theta_0\}$) and we have a prior Λ_1 over the alternative parameter space Θ_1, we can maximize average power by rejecting for large values of:

$$T(x) = \frac{\int_{\Theta_1} p_\theta(x)\, \mathrm{d}\Lambda_1(\theta)}{p_{\theta_0}(x)}$$

Show that this test maximizes the average-case power $\int_{\Theta_1} \mathbb{E}_\theta \phi(X)\, \mathrm{d}\Lambda_1(\theta)$ among all tests with level α.

Hint: Show that it can be viewed as a Neyman-Pearson test for a particular simple alternative.

4. 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 some ambiguity in how we could define the density; to resolve this ambiguity note it is equivalent to show that the CDF is linear, concave, or convex, or you can define the density unambiguously as the derivative of the CDF.