Statistics 210A Lecture 17 Notes

Daniel Raban

October 21, 2021

1 Nuisance Parameters, Tests for Multiparameter Exponential Families, and Permutation Tests

1.1 Nuisance parameters

We have been looking at tests with one real parameter $\theta \in \Theta \subseteq \mathbb{R}$. In a one sided test, $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$, we reject for large $T(X)$. This is valid if $T(X)$ is stochasticly increasing in θ and UMP if we the density has MLR in $T(X)$.

For two-sided tests, $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, we reject for extreme values of $T(X)$. We get valid **directional inference** if $T(X)$ is tochastically increasing, and this is UMPU if we have an exponential family and calibrate $c_1, c_2(x_1, x_2)$. So

$$\frac{d\text{Power}}{d\theta} = 0$$

at θ_0.

What about tests with multiple parameters?

Now, our model is $\mathcal{P} = \{P_{\theta,\lambda} : (\theta, \lambda) \in \Omega\}$, and we want to test $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$. We call θ the **parameter of interest** and λ the **nuisance paramater**. λ can affect our hypothesis test, even if we are only interested in estimating θ.

Example 1.1. Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ and $Y_1, \dots, Y_n \overset{\text{iid}}{\sim} N(\nu, \sigma^2)$, where μ, ν, σ^2 are unknown. We want to test $H_0 : \mu = \nu$ vs $H_i : \mu - \nu$. Here, we only care about $\theta = \mu - \nu$, so $\lambda = (\mu + \nu, \sigma)$ or $\lambda = (\mu, \sigma)$.

Example 1.2. Let $X_0 \sim \text{Binom}(n_0, \pi_0)$ and $X_1 \sim \text{Binom}(n_1, \pi_1)$ with $X_0 \perp\!\!\!\perp X_1$. Here, n_0, n_1 are known (not nuisance parameters). We want to test $H_0 : \pi_1 \leq \pi_0$ vs $H_1 : \pi_1 > \pi_0$. A nice choice of θ is the log **odds ratio** $\theta = \log \frac{\pi_1/(1-\pi_1)}{\pi_0/(1-\pi_0)}$, which we can use to write the null hypothesis as $\theta \leq 0$. In this case, $\lambda = \pi_0$.

1.2 Dealing with nuisance parameters in hypothesis tests for multiparameter exponential families

Suppose we have an exponential family $X \sim p_{\theta,\lambda}(x) = e^{\theta^\top T(x)+\lambda^\top U(x)-A(\theta,\lambda)}h(x)$ with $\theta \in \mathbb{R}^s$ and $\lambda \in \mathbb{R}^r$ both unknown. The distribution of $X \mid U(X)$ only depends on θ. This blocks the dependence on λ. Proceed in steps:

1. Make a sufficiency reduction to T, U:

$$(T(X), U(X)) \sim q_{\theta,\lambda}(t,u)e^{\theta^\top t+\lambda^\top u-A(\theta,\lambda)}g(t,u)$$

2. Condition on U to get

$$q_\theta(t \mid u) = \frac{q_{\theta,\lambda}(t,u)}{\int q_{\theta,\lambda}(z,u)\,dz} = e^{\theta^\top t-B_u(\theta)}g(t,u)$$

3. Perform the **conditional test** $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$ in the s-parameter model $\mathcal{Q}_u = \{q_\theta(t \mid u) : \theta \in \Theta\}$.

If $H_0 : \theta \le \theta_0$, then

$$\phi(x) = \mathbb{1}_{\{T(X)>c_\alpha(u(x))\}},$$

where

$$\mathbb{E}_{\theta,\lambda}[\phi(X) \mid U(X)] \le \alpha, \qquad \forall \theta \in \Theta_0.$$

Conditional control of the Type I error rate is *stronger* than marginal control of the Type I error rate.

Remark 1.1. We may not want conditional control of the Type I error if we don't need to get rid of a nuisance parameter because requiring this may give a less powerful test.

Theorem 1.1. *Assume $\mathcal{P}$ is a full-rank exponential family with densities*

$$p_{\theta,\lambda}(x) = e^{\theta^\top T(x)+\lambda^\top U(x)-A(\theta,\lambda)}h(x),$$

where $\theta \in \mathbb{R}$, $\lambda \in \mathbb{R}^r$, $(\theta, \lambda) \in \Omega$ is open, and θ_0 is possible.

(a) To test $H_0 : \theta \le \theta_0$ vs $H_1 : \theta > \theta_0$, there is a UMPU test $\phi^*(x) = \psi(T(x), U(x))$, where

$$\psi(t,u) = \begin{cases} 1 & t > c(u) \\ \gamma(u) & t = c(u) \\ t < c(u), & \end{cases}$$

where $c(u)$ and $\gamma(u)$ are chosen so that

$$\mathbb{E}_{\theta_0}[\phi^*(X) \mid U(X) = u] = \alpha.$$

(b) *To test $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, there is a UMPU test $\phi^*(x) = \psi(T(x), U(x))$, where*

$$\psi(t,u) = \begin{cases} 1 & t < c_1(u) \text{ or } t > c_1(u) \\ \gamma_i(u) & t = c_i(u) \\ 0 & t \in (c_1(u), c_2(u)) \end{cases}$$

with $c_i(u), \gamma_i(u)$ chosen to make

$$\mathbb{E}_{\theta_0}[\phi^*(X) \mid U(X) = u] = \alpha,$$

$$\mathbb{E}_{\theta_0}[T(X)(\phi^*(X) - \alpha) \mid U(X) = u] = 0.$$

We will prove this theorem next time.

Example 1.3. Suppose $X_i \overset{\text{ind}}{\sim} \text{Pois}(\mu_i)$ $i = 1, 2$, where we want to test $H_0 : \mu_1 \leq \mu_2$ vs $H_1 : \mu_1 > \mu_2$.[1] If we let $\eta_i = \log \mu_i$, then

$$p_\mu(x) = \prod_{i=1,2} \frac{\mu_i^{x_i} e^{-\mu_i}}{x_i!} = e^{x_1\eta_1 + x_2\eta_2 - (e^{\eta_1} e^{\eta_2})} \frac{1}{x_1! x_2!} \quad = e^{x_1(\mu_1 - \mu_2) + (x_1 + x_2)\eta_2 - (\cdots)} \frac{1}{x_1! x_2!}.$$

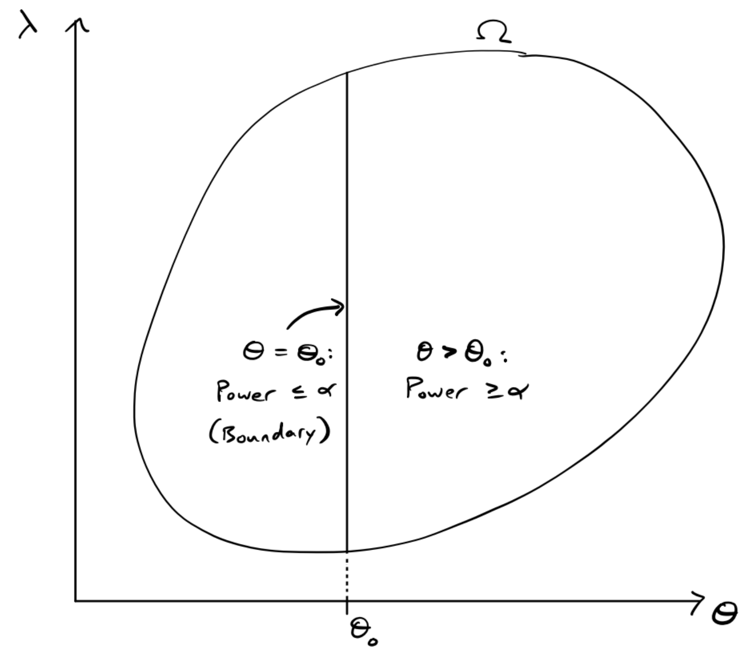
The null hypothesis is $H_0 : \mu_1 \leq \mu_2 \iff \eta_1 \leq \eta_2 \iff \eta_1 - \eta_2 \leq 0$. So our test is to reject when X_1 is *conditionally* large given $X_1 + X_2$.

$$\begin{aligned} \mathbb{P}_\theta(X_1 = x_1 \mid X_0 + X_1 = u) &= \frac{e^{x_1\theta + u\lambda - A(\theta,\lambda)} \frac{1}{x_1!} (u - x_1)!}{\sum_{z=0}^{u} (\cdots)} \\ &\propto_\theta e^{x_1\theta} \frac{1}{x_1!(u - x_1)!} \\ &\propto \text{Binom}(u, \tfrac{e^\theta}{1+e^\theta}) \\ &= \text{Binom}(X_0 + X_1, \tfrac{\mu_1}{\mu_1 + \mu_2}). \end{aligned}$$

Here is a sketch of the proof of the theorem:

[1] For example, Professor Fithian's wife has gotten pooped on by a bird 4 times in her life, and Professor Fithian has only gotten pooped on once. We can test if Professor Fithian's wife is more unlucky than average. This is the real reason to learn statistics.

Proof. We can think of the power function as a function on the set Ω:



1. We must have power $= \alpha$ on the boundary $\theta = \theta_0$.
2. On the boundary, U is complete sufficient. Then $\mathbb{E}[_{\theta_0,\lambda}[\phi(X)] = \alpha$ for all λ, so $\mathbb{E}_{\theta_0,\lambda}[\phi(X) \mid U(X)] \overset{a.s.}{=} \alpha$.
3. ϕ^* must then be the best among conditional tests.

The idea is the same for the two-sided tests, where we have constant power on the null hypothesis. $\square$

Example 1.4. Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$, where μ, ν, σ^2 are unknown. Test $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$. We have

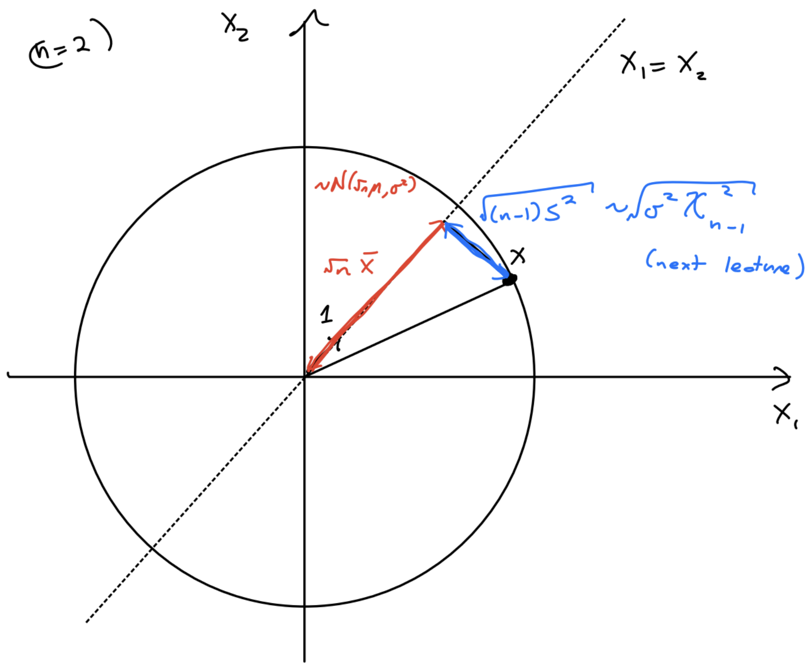
$$\rho_{\mu,\sigma^2}(x) = e^{\frac{\mu}{\sigma^2}\sum_i x_i - \frac{1}{2\sigma^2}\sum_i x_i^2 - \frac{n\mu}{\sigma^2}} \left(\frac{1}{2\pi\sigma^2}\right)^{n/2}.$$

Condition on $\sum_i X_i^2 = \|X\|^2 = U(X)$. Then under H_0, $X \mid \|X\|^2 = u \overset{H_0}{\sim} \text{Unif}(\sqrt{n}\mathbb{S}^{n-1})$ is uniform on the sphere. This is equivalent to $\frac{X}{\|X\|} \overset{H_0}{\sim} \text{Unif}(\mathbb{S}^{n-1})$, where $\frac{X}{\|X\|} \amalg \|X\|^2$. The UMPU test rejects when $\sum_i X_i$ is extreme given $\|X\|^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, so

$$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$$

$$
\begin{aligned}
&= \frac{1}{n-1}\left(\sum_{i=1}^n X_i^2 - 2\overline{X}\sum_{i=1}^n X_i + n\overline{X}^2\right) \\
&= \frac{1}{n-1}(\|X\|^2 - n\overline{X}^2) \\
&= \frac{1}{n-1}\left(\|X\|^2 - \left(\frac{1}{\sqrt{n}}\mathbf{1}_n^\top X\right)^2\right).
\end{aligned}
$$

This means $(n-1)S^2 = \|\operatorname{Proj}_{X_0}^\perp X\|$. Here is the picture when $n = 2$:



Reject for extreme $\frac{\sqrt{n}\overline{X}}{\sqrt{\|X\|^2 - n\overline{X}^2}}$ (note that this is monotone in $\overline{X}$). This is

$$
\frac{\sqrt{n}\overline{X}/\|X\|}{\sqrt{1 - n\overline{X}^2/\|X\|^2}},
$$

which is a function of $X/\|X\|$. So this is independent of $U = \|X\|^2$. This statistic is a scaled version of the T**-statistic**:

$$
\frac{1}{\sqrt{n-1}}\frac{\sqrt{n}\overline{X}}{\sqrt{S^2}},
$$

where $\frac{\sqrt{n}}{\overline{X}}\sqrt{S}^2 \overset{H_0}{\sim} t_{n-1}$.

1.3 Permutation tests

Example 1.5. Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} P$ and $Y_1, \dots, Y_n \overset{\text{iid}}{\sim} Q$ be independent, where we want to test $H_0 : P = Q$ vs $H_1 : P \neq Q$. Under H_0, $X_1, \dots, X_n, Y_1, \dots, X_n \overset{\text{iid}}{\sim} P$. So condition on the complete sufficient statistic for the null hypothesis (which is not complete for the alternative!): Define $(Z_1, \dots, Z_{n+m}) = (X_1, \dots, X_n, Y_1, \dots, Y_m)$. Under H_0, the order statistics $U(Z) = (Z_{(1)}, \dots, Z_{(n+m)})$ is complete sufficient. So let $S_{n+m} = \{\pi : \text{permutation on } n+m \text{ elements}\}$. Then

$$Z = (X, Y) \mid U \overset{H_0}{\sim} \text{Unif}(\{\pi U : \pi \in S_{n+m}\}).$$

For *any* test statistic Y, if $P = Q$, then

$$\mathbb{P}_{P,Q}(T(Z) \geq t \mid U) = \frac{1}{(n+m)!} \sum_{\pi \in S_{n+m}} \mathbb{1}_{\{T(\pi Z) \geq t\}}.$$

In practice, we can do a Monte Carlo version of this; sample $\pi_1, \dots, \pi_B \overset{\text{iid}}{\sim} \text{Unif}(S_{n+m})$. Then $Z, \pi_1 Z, \pi_2 Z, \dots, \pi_B Z \overset{\text{iid}}{\sim} \text{Unif}(S_{n+m} U)$. Let the p-value be

$$\begin{aligned} p &= \frac{1}{1+B} \sum_{b=1}^{B} \mathbb{1}_{\{T(Z) \leq T(\pi_b Z)\}} \\ &\overset{H_0}{\sim} \text{Unif}(\{\tfrac{1}{1+B}, \dots, \tfrac{B+1}{1+B}\}). \end{aligned}$$

So if we take a test statistic and apply it to all permutations of the data, if the original data looks special, then we should reject.