Statistics 210A Lecture 18 Notes

Daniel Raban

October 26, 2021

1 Hypothesis Tests for Gaussian Models

1.1 Recap: hypothesis testing with nuisance parameters

Last time, we discussed hypothesis testing with nuisance parameters. If we have an exponential family $X \sim e^{\theta T(x)+\lambda^\top U(x)-A(\theta,\lambda)}h(x)$ with the one-sided test $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$, then the UMPU test rejects for conditionally large $T \mid U$. If we have the two-sided test $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, the UMPU test rejects for conditionally extreme $T \mid U$. Here, $P_\theta(T \mid U)$ depends only on θ (U is sufficient after fixing θ).

We saw the nonparametric test where if $X_i \stackrel{\text{iid}}{\sim} P$, $Y_i \stackrel{\text{iid}}{\sim} Q$, then we can test $H_0 : P = Q$ vs $H_1 : P \neq Q$ by conditioning on the pooled order statistics. There are various choices of test statistics to use for permutation tests with various properties.

1.2 Distributions related to Gaussians

Example 1.1 (χ^2 distribution)**.** If $Z_1, \dots, Z_d \stackrel{\text{iid}}{\sim} N(0, 1)$, then

$$V = \sum_{i=1}^d Z_i^2 \sim \chi_d^2 = \text{Gamma}(d/2, 2)$$

with

$$\mathbb{E}[V] = d, \qquad \text{Var}(V) = 2d.$$

Note that the standard deviation grows slower than the mean, so as $d \to \infty$,

$$\frac{V}{d} \xrightarrow{p} 1.$$

That is,

$$\mathbb{P}(|\tfrac{V}{d} - 1| \geq \varepsilon) \to 0$$

for all $\varepsilon > 0$. This is what we would expect from the law of large numbers. The central limit theorem tells us that $V \approx N(d, 2d)$ because $\sqrt{d}(\frac{V}{d} - 1) \xrightarrow{d} N(0, 2)$.

Example 1.2 (t-distribution)**.** If $Z \sim N(0,1)$ and $V \sim \chi^2_d$ with $Z \amalg V$, then

$$\frac{Z}{\sqrt{V}/d} \sim t_d,$$

the **Student's T-distribution**, where $t_d \approx N(0,1)$ for large d.

Example 1.3 (F-distribution)**.** If $V_1 \sim \chi^2_{d_1}$ and $V_2 \sim X^2_{d_2}$ with $V_1 \amalg V_2$, then

$$\frac{V_1/d_1}{V_2/d_2} \sim F_{d_1,d_2},$$

the **F-distribution**, which has 2 degrees of freedom. $F_{d_1,d_2} \approx \chi^2_{d_1}$ if $d_2 \to \infty$. If $t \sim t_d$, then

$$T^2 \sim \frac{Z^2}{V/d} \sim F_{1,d}.$$

Example 1.4. If $Z \sim N_d(\mu, \Sigma)$ with $A \in \mathbb{R}^{k \times d}$ and $b \in \mathbb{R}^k$, then

$$AZ + b \sim N(A\mu + b, A\Sigma A^\top).$$

1.3 Analysis of the one-sample t-test

We saw earlier that if $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ with both $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ unknown and we test $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$, then the UMPU test says to reject for extreme $\overline{X}$ given $\|X\|^2$. We could also say to reject for large $|\overline{X}|/\|X\|$; this gets rid of the conditioning given $\|X\|^2$, since under the null, $|\overline{X}|/\|X\| \amalg \|X\|^2$. Equivalently, we can reject for large values of

$$\frac{n\overline{X}}{\|X\|^2 - n\overline{X}^2} = \frac{n\overline{X}^2/\|X\|^2}{1 - \overline{X}^2/\|X\|^2}.$$

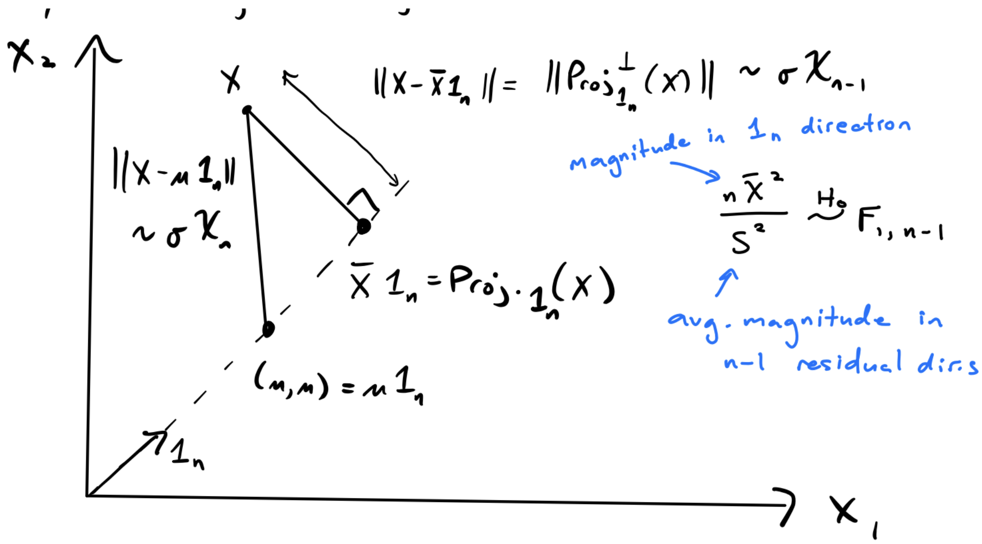
If we didn't want to square this, we could equivalently reject for extreme

$$\frac{\sqrt{n}\overline{X}}{\sqrt{S}^2}, \overset{\text{H}_0}{\sim} t_{n-1},$$

where $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$. This has a T-distirbution because

$$S^2 = \frac{1}{n-1}\| \operatorname{Proj}^\perp_{\mathbf{1}_n} X\|^2 \sim \sigma^2 \chi_{n-1}.$$

Here is a picture for $n = 2$:



What's happening geometrically is that

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

is a ratio of squared magnitudes in different directions. If $\mu = 0$, then no direction should be special. Let's make this more precise with linear algebra.

Here is a change of basis: Let

$$Q = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & \\ q_1 & Q_r \\ | & \end{bmatrix}$$

with $q_1 = \frac{1}{\sqrt{n}}\mathbf{1}_n$ and $q_2, \ldots, q_n$ completing this to an orthonormal basis. Then $Q^\top Q = QQ^\top = I_n$, and $X \sim N_n(\mu\mathbf{1}_n, \sigma^2 I_n)$. Let

$$Z = Q^\top X = \begin{bmatrix} q_1^\top X \\ Q_r^\top X \end{bmatrix} = \begin{bmatrix} \sqrt{n}\overline{X} \\ Q_r^\top X \end{bmatrix}.$$

Then

$$Q^\top X \sim N_n\left(\begin{bmatrix} \sqrt{n}\mu \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \sigma^2 I_n\right),$$

so

$$\|Q_r^\top X\|^2 = \|Q^\top X\|^2 - n\overline{X}^2 = \|X\|^2 - n\overline{X}^2 = (n-1)S^2.$$

This tells us that

$$Q_r^\top X \sim N_{n-1}(0, \sigma^2 I_{n-1}).$$

Here, we have

$$(n-1)S^2 \sim \sigma^2\chi^2_{n-1}, \qquad \sqrt{n}\overline{X} \sim N(\sqrt{n}\mu, \sigma^2).$$

1.4 Canonical linear model

Assume

$$Z = \begin{bmatrix} Z_0 \\ Z_1 \\ Z_r \end{bmatrix} \sim N_n\left(\begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_r \end{bmatrix}, \sigma^2 I_n\right)$$

where Z_i has dimension d_i with $n - d_0 - d_1 = d_r$. Here, $\mu_0 \in \mathbb{R}^{d_0}$, $\mu_1 \in \mathbb{R}^{d_1}$, $\mu_r \in \mathbb{R}^{d_r}$. We want to test $H_0 : \mu_1 = 0$ vs $H_1 : \mu_1 \neq 0$.

This is an exponential family, with

$$p(z) \propto e^{\frac{\mu_1^\top}{\sigma^2}Z_1 + \frac{\mu_0^\top}{\sigma^2}Z_0 - \frac{1}{2\sigma^2}\|z\|^2} h(z).$$

We want to condition on (i.e. ignore) the nuisance parameter Z_0.

- If σ^2 is known and $d_1 = 1$, then the UMPU test rejects for large/small/extreme values of
$$\frac{Z_1}{\sigma} \overset{H_0}{\sim} N(0,1), \qquad (Z\text{-test}).$$

- For known σ^2 and $d_1 \geq 1$, reject for large values of
$$\frac{|Z_1||^2}{\sigma^2} \overset{H_0}{\sim} \chi^2_{d_1} \qquad (\chi^2\text{-test}[1]).$$

- For $d_1 = 1$ and unknown σ^2, condition on Z_0, $\|Z\|^2 = \|Z_0\|^2 + Z_1^2 + \|Z_r\|^2$. We reject for conditionally large/small/extreme Z_1, which is the same as conditioning on extreme values of $\frac{Z_1}{\sqrt{Z_1^2 + \|Z_r\|^2}}$. This is equivalent to rejecting for extreme values of
$$\frac{Z_1}{\sqrt{\|Z_1\|^2/d_1}} \overset{H_0}{\sim} t_{d_1} \qquad (t\text{-test}).$$

- If $d_1 \geq 1$ and σ^2 is unknown, condition on $Z_0, \|Z_1\|^2 + \|Z_r\|^2$. Reject for large
$$\frac{\|Z_1\|^2/d_1}{\|Z_r\|^2/d_r} \overset{H_0}{\sim} F_{d_1,d_r} \qquad (F\text{-test}).$$

[1]There are a number of hypothesis tests refered to as a "χ^2-test."

Remark 1.1. This is related to the Beta distribution.

$$\|X_1\|^2 \sim \sigma^2 \,\mathrm{Gamma}(d_1/2, 2\sigma^2)$$

and

$$\|Z_r\|^2 \sim \sigma^2 \chi^2_{d_r} = \mathrm{Gamma}(d_r/2, 2\sigma^2).$$

so

$$\frac{\|Z_1\|^2}{\|Z_1\|^2 + \|Z_r\|^2} \sim \mathrm{Beta}(d_1/2, d_r/2).$$

More generally, if $U \sim \mathrm{Beta}(d_1/2, d_r/2)$, then

$$\frac{U/d_1}{(1-U)/d_r} \sim F_{d_1,,d_2}.$$

The normalized residual vector is

$$\frac{1}{d_r}\|Z_r\|^2 \sim \frac{\sigma^2}{d_r}\chi^2_{d_r} \approx \sigma^2.$$

If we write

$$\widehat{\sigma}^2 = \frac{\|Z_r\|^2}{d_r},$$

then the t-statistic is

$$\frac{Z_1}{\widehat{\sigma}},$$

and the F-statistic is

$$\frac{1}{d_1}\frac{\|Z_1\|^2}{\widehat{\sigma}^2}.$$

This is a solution to a problem presented in a nice form. Next time, we will talk about how to use a change of basis to solve more general Gaussian model problems.