Mathematical Statistics Recitation 8

TA: Sonia Reilly

March 27, 2026

Lecture Review

Lecture 15: Generalized LR (Last of Midterm 2 material!)

Generalized likelihood ratio test Composite vs. composite testing, hypotheses are sets of values of the parameter θ:

$$H_0 : \theta \in \Theta_0, \qquad H_1 : \theta \in \Theta_1, \qquad \Theta = \Theta_0 \cup \Theta_1$$

$$LR = \frac{\max\limits_{\theta \in \Theta_0} \mathcal{L}_n(\theta)}{\max\limits_{\theta \in \Theta} \mathcal{L}_n(\theta)}$$

Reject H_0 if $LR \leq c$.

Wilk's Theorem:

$$-2 \log LR \xrightarrow{d} \chi^2_d$$

under H_0, with $d = \dim \Theta - \dim \Theta_0$, given some regularity conditions on pdf's.

Gives us the ability to build confidence intervals / level α rejection regions.

Goodness of fit for Multinomial Data Model: m categories ("cells"), each random variable from the model belongs to the ith category with probability $p_i(\theta)$, θ some unknown parameter.

$$H_0 = \text{data comes from this model}, \qquad H_1 = \text{data does not come from this model}$$

Pearson's test statistic:

$$\chi^2 = \sum_{i=1}^{m} \frac{(X_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})} = \sum_{i=1}^{m} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{m-k-1}$$

where:

- $X_i = O_i =$ observed count in cell i
- $E_i =$ expected count in cell i
- $\hat{\theta} =$ MLE value of θ assuming data follows model ($E_i = np_i(\hat{\theta})$)
- $k =$ dimension of θ

Comes from applying generalized LR test — for large n, $-2 \log \Lambda \approx \chi^2$.

Pearson's Chi-squared test: reject H_0 if $\chi^2 \geq c$

Lecture 16: Nonparametric methods

Nonparametric: $X_1, \ldots, X_n$ iid from some distribution with CDF $F(x)$, unknown. No specific type of distribution assumed!

Empirical CDF

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq x) = \frac{1}{n}\ (\#\text{ of } X_i \text{ less than } x)$$

For all x,

$$\mathbb{E}[F_n(x)] = F(x), \qquad \text{Var}(F_n(x)) = \frac{F(x)(1-F(x))}{n} \quad \text{(pointwise)}$$

Kolmogorov-Smirnov Test The Glivenko-Cantelli Theorem says that

$$\sup_x |F(x) - F_n(x)| \xrightarrow{p} 0.$$

It turns out that

$$\sqrt{n}\sup_x |F_n(x) - F(x)| \xrightarrow{d} K$$

where K has a distribution called the Kolmogorov distribution. We can use this to form CIs:

$$\mathbb{P}\left(F_n(x) - \frac{K(\alpha)}{\sqrt{n}} \leq F(x) \leq F_n(x) + \frac{K(\alpha)}{\sqrt{n}}\ \forall x\right) \to 1-\alpha,$$

where $K(\alpha)$ is the number for which $P(K > K(\alpha)) = \alpha$, and to create tests. The test that rejects when

$$\sup_x |F_n(x) - F(x)| > \frac{K(\alpha)}{\sqrt{n}}$$

is a level α test for

$$H_0 : X_1, \ldots, X_n \sim F \qquad \text{vs.} \qquad H_1 : X_1, \ldots, X_n \nsim F$$

called the Kolmogorov–Smirnov test.

Bootstrap Method for Estimating Variance For some estimator $\hat{\theta}$, want approx $\text{Var}(\hat{\theta})$ given only the data, no info about distribution.

1. Generate B samples of size n from F_n iid, i.e., sampling with replacement from original data $X_1, \ldots, X_n$.
2. Compute the estimates $\hat{\theta}_j$ from each sample.
3. Compute sample variance of the estimates:

$$\widehat{\text{Var}}(\hat{\theta}) = \frac{1}{B}\sum_{j=1}^{B}(\hat{\theta}_j - \bar{\theta}_B)^2, \quad \bar{\theta}_B = \frac{1}{B}\sum_{j=1}^{B}\hat{\theta}_j$$

Can do something similar to estimate other quantities, such as bias.

Problems

1. True or False:

   (a) The generalized likelihood ratio statistic LR is always less than or equal to 1.

   (b) Since $\text{Var}(F_n(x)) = S_n^2(x) = \frac{F(x)(1-F(x))}{n}$,

   $$\left(F_n - 1.96\frac{1}{\sqrt{n}}S_n(x), F_n + 1.96\frac{1}{\sqrt{n}}S_n(x)\right)$$

   is a 0.95 confidence interval for the true CDF $F(x)$.

   Solution:

   (a) True — numerator is max over a smaller set than the denominator, so numerator $\leq$ denominator.

   (b) False — that is a pointwise variance.

   $$\mathbb{P}(f_1(x) \leq F(x) \leq f_2(x)\ \forall x) \neq \mathbb{P}(f_1(x) \leq F(x) \leq f_2(x) \text{ for a given } x)$$

   For specific x, this would be a 0.95 confidence interval, but it's not a 0.95 CI for the whole function.

2. [9.35] Under a standard genetic model, the genotypes AA, Aa, and aa occur with probabilities $(1-\theta)^2$, $2\theta(1-\theta)$, and θ^2, respectively, for some $0 \leq \theta \leq 1$. A sample of 190 people reveals that 10 have type AA, 68 have type Aa, and 112 have type aa. [You may use the fact that the MLE is $\hat{\theta} = 0.768$. You should be able to derive this yourself to check it.]

   (a) What is the Pearson chi-squared statistic?

   (b) Use the table of the CDF of the χ^2 distribution on the other side of the page to estimate the p-value of the data. At a significance level of 0.05, would we accept or reject the null hypothesis that the data comes from our model?

   Solution:

   To find the MLE, see Example A in section 8.5.1, where they derive the MLE $\hat{\theta}$ for this model.

   $$\hat{\theta} = \frac{2X_3 + X_2}{2n} = \frac{2(112) + 68}{2(190)} = 0.768.$$

   (a)

   $$\chi^2 = \sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} = \frac{(10 - 190(1-\hat{\theta})^2)^2}{190(1-\hat{\theta})^2} + \frac{(68 - 190(2\hat{\theta}(1-\hat{\theta}))^2}{2\hat{\theta}(1-\hat{\theta})} + \frac{(112 - 190\hat{\theta}^2)^2}{190\hat{\theta}^2} = 0.0067$$

   (b)

   $$\chi^2 \sim \chi^2_{m-k-1} = \chi^2_1 \quad \text{since } m = 3,\ k = 1$$

   So p-value is

   $$P(\chi^2 > 0.0067) \approx \text{between 0.90 and 0.95}$$

   according to the table.

   Since p-value $> 0.90 > 0.05$, the data fits the model well and we should not reject H_0 at a significance level of 0.05.

3. Suppose we have two samples from a distribution, $X_1 = 2$ and $X_2 = 4$.

   (a) Write out all the possible samples of size $n = 2$ with replacement from this data.
   (b) Letting B be all possible samples, compute the B sample medians.
   (c) What is the bootstrap estimate of the variance of the sample median for this tiny dataset?

   Solution:

   (a) Possible samples of size $n = 2$ with replacement:

   $$\{2, 2\}, \{2, 4\}, \{4, 2\}, \{4, 4\}$$

   (b) Sample medians:

   $$2, 3, 3, 4$$

   This is the bootstrap distribution of medians.

   (c) Bootstrap estimate of variance:

   $$\widehat{\text{Var}}(\hat{\theta}) = \frac{1}{4}\Big[(2-3)^2 + (3-3)^2 + (3-3)^2 + (4-3)^2\Big] = \boxed{\frac{1}{2}}.$$

4. A researcher claims that a specific random number generator produces values following a Uniform(0, 1) distribution. You generate a sample of $n = 4$ independent observations to test this claim:

   $$X = \{0.75, 0.40, 0.15, 0.90\}$$

   (a) State the null hypothesis H_0 and the alternative hypothesis H_1.
   (b) Calculate and sketch the empirical CDF, $F_n(x)$, for this sample.
   (c) Calculate the K-S test statistic $d_{KS} = \sup_x |F_n(x) - F(x)|$.
   (d) At the $\alpha = 0.05$ level, should you reject H_0? Use the table on the other side of this page.
   (e) At the same level, should we reject H_0 if $X = \{0.82, 0.71, 0.96, 0.63\}$?

   Solution:

   (a) $H_0 : X \sim$ Uniform(0, 1), or equivalently, X has CDF $F(x) = x$ for $x \in [0, 1]$.
   $H_1 : X \not\sim$ Uniform(0, 1).

   (b) Order the observations: $\{0.15, 0.40, 0.75, 0.90\}$. The empirical CDF increases by $1/n = 1/4$ at each ordered observation. See figure below (the line in blue).

   (c) We want the max difference between the hypothesized and empirical CDFs. Check either side of each observation point:

   $$\begin{aligned} d_{KS} &= \max(|0.15-0|, |0.15-0.25|, |0.40-0.25|, |0.40-0.5|, \\ &\qquad |0.75-0.5|, |0.75-0.75|, |0.90-0.75|, |0.90-1|) \\ &= \max(0.15, 0.10, 0.15, 0.10, 0.25, 0.00, 0.15, 0.10) \\ &= \boxed{0.25} \end{aligned}$$

(d) Looking at the table, for $n = 4$ and $\alpha = 0.05$, the critical value is 0.624. (For large n, the critical value is $K(\alpha)/\sqrt{n}$ for some $K(\alpha)$, but for small n we have to look in a table because the dependence on n is not as simple.) Since $d_{KS} = 0.25 < 0.624$, we fail to reject the null hypothesis. There is no significant evidence at the 5% level to suggest the generator is not Uniform(0, 1).

(e) For $X = \{0.82, 0.71, 0.96, 0.63\}$, $X_{(1)} = 0.63$, so $d_{KS} \geq |0.63 - 0| > 0.624$. Since we are interested in the max (supremum) difference, we don't need to check any other points if there is one point where the difference is larger than $K(\alpha)$. We should reject at this level.

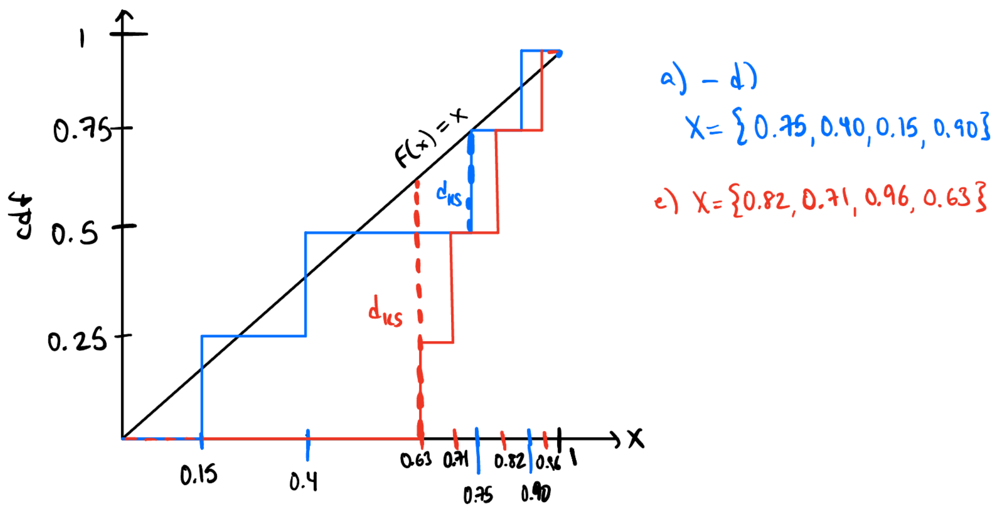

Figure 1: Sketch for Problem 4.