Presentation 5 Problem 5

Luis Park

March 11, 2025

Question 5 part A

Rejecting for small values of the GLR test statistic Λ is equivalent to rejecting for large values of $-2log(\Lambda)$. Theorem A on p. 341 states that for large samples from Exponential family distributions, the distribution of $-2log(\Lambda)$ is approximately $\chi^2_{(\nu)}$, where ν is the difference in the number of parameters that need to be estimated overall, and under H_o. Show (as in Example A on p. 339) that this is exactly true for testing a Normal mean with known σ, and approximately true when σ is unknown (using the result of 4d).

Answer:

Likelihood function for Normal

Normal Mean with known variance σ:

Suppose $x_1, x_2, ..x_n \sim N(\mu, \sigma^2)$ with known σ^2

The likelihood function can be defined as $L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x_i-\mu)^2}{2\sigma^2}) = \frac{1}{(2\pi\sigma^2)^{n/2}} * exp(\frac{-1}{2\sigma^2}\sum(x_i-\mu)^2)$

GLR Test Statistic for known variance

The hypotheses are $H_o : \mu = \mu_o$ vs $H_a : \mu \neq \mu_o$

The likelihood ratio is defined as $\Lambda = \frac{L(\mu_o)}{L(\hat{\mu})}$

$L(\mu_o) = \frac{1}{(2\pi\sigma^2)^{n/2}} * exp(\frac{-1}{2\sigma^2}\sum(x_i - \mu_o)^2)$

$L(\hat{\mu}) = \frac{1}{(2\pi\sigma^2)^{n/2}} * exp(\frac{-1}{2\sigma^2}\sum(x_i - \overline{x})^2)$

So $\Lambda = exp(-\frac{1}{2\sigma^2}(\sum(x_i - \mu_o)^2 - \sum(x_i - \overline{x})^2)$

$-2log\Lambda = \frac{1}{\sigma^2}(\sum(x_i - \mu_o)^2 - \sum(x_i - \overline{x})^2)$

And we also know that $\sum(x_i - \mu_o)^2 = \sum(x_i - \overline{x})^2 + n(\overline{x} - \mu_o)^2$

So then we know $-2log\Lambda = \frac{1}{\sigma^2}(\sum(x_i - \overline{x})^2 + n(\overline{x} - \mu_o)^2 - \sum(x_i - \overline{x})^2) = \frac{n}{\sigma^2}(\overline{x} - \mu_o)^2$

So we reject for large $\frac{n}{\sigma^2}(\overline{x} - \mu_o)^2$

Under H_o, we know that $\overline{x} \sim N(\mu_o, \sigma^2/n)$

And we also know that to standardize the normal distribution, we have $\frac{\sqrt{n}(\overline{x}-\mu_o)}{\sigma} \sim N(0, 1)$

And we also know squaring a standard normal gives us a chi-square distribution

So $\frac{n}{\sigma^2}(\overline{x}-\mu_o)^2 \sim \chi^2_1$

And $\frac{n}{\sigma^2}(\overline{x}-\mu_o)^2$ is the likelihood ratio we computed, so the likelihood ratio of a normal distribution follows a chi-square distribution

GLR Test Statistic for unknown variance

For σ unknown, we know that $\Lambda = (1+\frac{T^2}{n-1})^{-n/2}$ with $T = \frac{\overline{x}-\mu_o}{s/\sqrt{n}}$ (This was proven in question 4 part d)

Then $-2log\Lambda = nlog(1+\frac{T^2}{n-1}) \approx \frac{nT^2}{(n-1)} \approx T^2 \approx Z^2 \sim \chi^2_{(1)}$

Question 5 part B

Problem:

Show how to use Lagrange multipliers to find the maximum likelihood estimates for multinomial cell probabilities.

Answer:

Define the Likelihood Function

Let us we have a multinomial distribution with k categories and conduct n independent trials

Suppose the counts for each category is $x_1, x_2 ... x_k$ where $\sum_{i=1}^{k} x_i = n$

The probability of observing these counts with probabilities $p_1, p_2 ... p_k$ can be given by this likelihood function:

$\frac{n!}{x_1!x_2!..x_k!} p_1^{x_1} p_2^{x_2} .... p_k^{x_k}$

The log likelihood function is then $l(p_1, p_2 ... p_k) = log(\frac{n!}{x_1!x_2!...x_k!}) + \sum_{i=1}^{k} x_i log(p_i) = \sum_{i=1}^{k} x_i log(p_i)$ because the first term is constant

Introducing Lagrange Constraints

Since $\sum_{i=1}^{k} p_i = 1$, we can introduce a lagrange constraint to assure this sum is 1

So $L(p_1, p_2, ..p_k, \lambda) = \sum_{i=1}^{k} x_i log(p_i) + \lambda(1 - \sum_{i=1}^{k} p_i)$

Now to find a maximum likelihood of p_i, we need to take the partial derivative with respect to p_i and λ

So $\frac{\partial L}{\partial p_i} = \frac{x_i}{p_i} - \lambda$ and now set equal to 0 to find critical point: $\frac{x_i}{p_i} - \lambda = 0$

So $p_i = \frac{x_i}{\lambda}$

Now let us find $\frac{\partial L}{\partial \lambda}$ to solve for λ

So $\frac{\partial L}{\partial \lambda} = (1 - \sum_{i=1}^{k} p_i) = 0$

Now substitute $p_i = \frac{x_i}{\lambda}$, so $(1 - \sum_{i=1}^{k} p_i) = 1 - \sum_{i=1}^{k} \frac{x_i}{\lambda}) = 1 - n/\lambda = 0$

Now $\lambda = n$

So finally we get $p_i = \frac{x_i}{n}$, the MLE for each p_i

Question 5 part C

Problem:

Show how to test multinomial probabilities using the approximate Generalized Likelihood Ratio (GLR) test. Show that the GLR test statistic is asymptotically equivalent to the Pearson Chi-square statistic:

$$P = \sum \frac{(O-E)^2}{E}, \tag{1}$$

where O and E are the observed and expected counts. As an example, suppose $n = 30$ rolls of a 6-sided die result in counts of $(10, 5, 5, 5, 5, 0)$ for the outcomes $1, 2, \ldots, 6$. This is what one would expect for a biased die that has two 1's and no 6. Compute the P-value of a test for the null hypothesis:

$$H_0 : p_1 = p_2 = \cdots = p_6 = \frac{1}{6} \tag{2}$$

against a general alternative using the Chi-square approximation. Additionally, use simulation to compute an exact P-value (up to simulation error). Simulate many sets of 30 fair dice rolls and compute the GLR statistic for each replicate dataset. Estimate the P-value as the proportion of times you get a statistic as large or larger than the observed value.

Answer:

GLR of Multionmial Probabilities

Given $H_o : p_1 = p_2 = ...p_6 = 1/6$ and the alternative is that the dice is biased, the likelihood ratio is $\Lambda = \frac{\text{likelihood under null}}{\text{likelihood under alternative}}$

$L(p_i = 1/6) = \frac{n!}{O_1!O_2!..O_6!} \frac{1}{6}^{O_1} \frac{1}{6}^{O_2} .... \frac{1}{6}^{O_6} = \frac{n!}{O_1!O_2!..O_6!} \prod_{i=1}^{6} \frac{1}{6}^{O_i}$ and $L(p_i = \frac{O_i}{n}) = \frac{n!}{O_1!O_2!..O_6!} (\frac{O_1}{n})^{O_1} (\frac{O_2}{n})^{O_2} .... (\frac{O_6}{n})^{O_6} = \frac{n!}{O_1!O_2!..O_6!} \prod_{i=1}^{6} (\frac{O_i}{n})^{O_i}$

So $\Lambda = \frac{L(p_i=1/6)}{L(p_i=\frac{O_i}{n})} = \frac{\prod_{i=1}^{6} (\frac{1}{6})^{O_i}}{\prod_{i=1}^{6} (\frac{O_i}{n})^{O_i}}$

So $log(\Lambda) = \sum_{i=1}^{6} O_i log(1/6) - \sum_{i=1}^{6} O_i log(\frac{O_i}{n}) = \sum_{i=1}^{6} O_i (log(\frac{1}{6}) - log(\frac{O_i}{n})) = \sum_{i=1}^{6} O_i log(\frac{1/6}{O_i/n}) = \sum_{i=1}^{6} O_i log(\frac{n}{6O_i})$

And finally $-2log(\Lambda) = 2 \sum_{i=1}^{6} O_i log(\frac{6O_i}{n})$

And recall that under the null hypothesis, $E_i = np_i = n/6$ so $-2log(\Lambda) = 2 \sum_{i=1}^{6} O_i log(\frac{O_i}{E_i})$

And $-2log(\Lambda) = 2 \sum_{i=1}^{6} O_i log(\frac{O_i}{E_i}) \approx \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$. This is true because of the taylor expansion for $log(\frac{O_i}{E_i})$ around $O_i = E_i$

Thus $-2log(\Lambda) \approx \chi^2$ with k-1=6-1 degrees of freedom

Computing P-Value

Given the observed counts of (10,5,5,5,5,0), $E_i = 30/n = 30/6 = 5$ for all i

So the approximate chi-squared statistic is $-2log(\Lambda) = 2 \sum_{i=1}^{6} O_i log(\frac{O_i}{E_i}) = 2(10log(\frac{10}{5}) + 5log(\frac{5}{5}) + 5log(\frac{5}{5}) + 5log(\frac{5}{5}) + 5log(\frac{5}{5}) + 0log(\frac{0}{5})) = 13.86$

The degrees of freedom are k-1=5, so the p value is $P(\chi^2_5 \geq 13.86) = .016$

However, when using 100,000 runs of a simulation on R(provided by Professor Everson), we estimated an exact p-value of 0.02582. R code will be provided

Question 5 part D

Problem:

Consider a test of $H_o : \theta = \theta_o vs. H_a : \theta \neq q\theta_o$, based on $X_1, ..., X_n \sim Unif(0, \theta)$. Explain why this distribution is not in the Exponential family. Show the null distribution is exactly Chi-square, but with 2 df, not the 1 df prescribed by Theorem A.

Answer:

Suppose $x_1, x_2, ...x_n \sim Unif(0, \theta)$

Our null hypothesis $H_o : \theta = \theta_o$ vs. $H_a : \theta \neq \theta_o$ The likelihood function is defined as $L(\theta) = (\frac{1}{\theta})^n$ with support of $I(\theta > x_{(n)})$

$\Lambda = \frac{L(\theta_o)}{L(x_{(n)})} = \frac{\theta_o^{-n}}{X_{(n)}^{-n}}$ $-2log(\Lambda) = -2nlog(\frac{x_{(n)}}{\theta_o})$