Math Statistics
Semiweekly HW 15

Ozaner Hansha

November 24, 2020

Question 1

Problem: Suppose we have a Poisson population $X_i \sim \text{Pois}(\lambda)$. Our null hypothesis H_0 is that $\lambda = 2$ and our alternative hypothesis H_1 is that $\lambda = 1$. Given an i.i.d. sample X of $n = 7$, use the Neyman-Pearson lemma to produce a most powerful critical region C for $\alpha = .01$ to test these hypotheses.

Solution: Note that the ratio of the likelihood of the null and alternative hypotheses must be less than some finite constant k (since we can only observe items in the support):

$$\begin{aligned}
k &\geq \frac{L(\lambda_0)}{L(\lambda_1)} \\
&= \frac{\prod_{i=1}^n P(X_i; \lambda = \lambda_0)}{\prod_{i=1}^n P(X_i; \lambda = \lambda_1)} && \text{(def. of likelihood)} \\
&= \prod_{i=1}^n \frac{P(X_i; \lambda = \lambda_0)}{P(X_i; \lambda = \lambda_1)} && \text{(associativity)} \\
&= \prod_{i=1}^n \frac{\lambda_0^{X_i} e^{-\lambda_0}/X_i!}{\lambda_1^{X_i} e^{-\lambda_1}/X_i!} && \text{(pmf of Poisson RV)} \\
&= \prod_{i=1}^n \left(\frac{\lambda_0}{\lambda_1}\right)^{X_i} e^{\lambda_1 - \lambda_0} \\
\ln k &\geq \ln \prod_{i=1}^n \left(\frac{\lambda_0}{\lambda_1}\right)^{X_i} e^{\lambda_1 - \lambda_0} && \text{(ln is monotone increasing)} \\
&= \sum_{i=1}^n X_i(\ln \lambda_0 - \ln \lambda_1) + \lambda_1 - \lambda_0 \\
&= n(\lambda_1 - \lambda_0) + (\ln \lambda_0 - \ln \lambda_1) \sum_{i=1}^n X_i \\
\frac{\ln k - n(\lambda_1 - \lambda_0)}{\ln \lambda_0 - \ln \lambda_1} &\geq \sum_{i=1} X_i \\
\frac{\ln k + 7}{\ln 2} &\geq \sum_{i=1} X_i \\
k' &\geq \sum_{i=1} X_i && \text{(let LHS} = k')
\end{aligned}$$

Note that since both our hypotheses are simple, the Neyman-Pearson lemma tells us that the most powerful test for testing them is given by the above inequality for some constant k':

$$\sum_{i=1}^n X_i \leq k'$$

What that k' is depends on our α which, in this case, is equal to .01. Below we will solve for k':

$$\begin{aligned}
\alpha &= P(\text{Type I error}) && \text{(def. of } \alpha\text{)} \\
&= P(\neg \hat{H}_0 \mid H_0) && \text{(def. of Type I error)} \\
&= P\left(\sum_{i=1}^{n} X_i > k' \mid H_0\right) && \text{(Neyman-Pearson lemma)} \\
&= P\left(\sum_{i=1}^{n} X_i > k' \mid \lambda = \lambda_0\right) && \text{(given null hypothesis)} \\
&= 1 - p_Y(k') && (Y = \textstyle\sum_{i=1}^{n} X_i \sim \text{Pois}(n\lambda_0), \text{ i.i.d } X_i) \\
.01 &= 1 - p_Y(k') && \text{(desired } \alpha\text{)} \\
.99 &= p_Y(k')
\end{aligned}$$

And so to solve for k' we must find the 99th quantile of the Poisson cdf with $\lambda = n\lambda_0 = 14$. Using a calculator, we find that:

$$k' = q(.99; \lambda = 14) \approx 22.88387$$

And so our most powerful test of the hypotheses that splits $\mathbb{R}$ into regions R_0 and R_1 is given by:

$$\begin{aligned}
\sum_{i=1}^{7} X_i &\leq 22.88387 \\
\bar{X} &\leq 3.26912 && \text{(divide both sides by } n = 7\text{)}
\end{aligned}$$

If the sample passes this test we accept the alternative hypothesis, else we reject it.