Stat 8003, Homework 6

Group G: `sample( c( "David" , "Andrew", "Salam" ))`

October 8, 2014

Question 6.1. A coin is thrown independently 10 times to test the hypothesis that the probability of heads is 1/2 versus the alternative that the probability is not 1/2. The test rejects if either 0 or 10 heads are observed.

(a) What is the significance level of the test?

(b) If in fact the probability of heads is .1, what is the power of the test?

Answer:

(a) Let X be the number of heads showing up in 10 tosses of the coin, i.e. $X \sim Binomial(10, p)$ where p is the probability of landing heads. The hypotheses are

$$H_0 : \ p = \frac{1}{2}, \quad vs \quad H_a : \ p \neq \frac{1}{2}.$$

By definition,

$$\begin{aligned}\alpha &= P(X = 0 \mid H_0) + P(X = 10 \mid H_0)\\ &= \binom{10}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{10} + \binom{10}{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^0\\ &= \frac{2}{2^{10}} = \frac{1}{512}\\ &= 0.002\end{aligned}$$

(b) If $p = 0.1$, then the probability of rejecting the null hypothesis is

$$\begin{aligned}1 - \beta &= \binom{10}{0}(0.1)^0(0.9)^{10} + \binom{10}{10}(0.1)^{10}(0.9)^0\\ &= 0.3486784\end{aligned}$$

Thus the significance level of this test is 0.002 and its power is 0.3487.

Question 6.2. Let $X_1, , X_n$ be a random sample from an exponential distribution with the density function

$$f(x \mid \theta) = \theta e^{-\theta x}.$$

We want to test the hypothesis

$$H_0: \ \theta = 1, \quad vs \quad H_a: \ \theta \neq 1.$$

Set the desired level of significance as $\alpha = 5\%$.

(a) Derive a generalized likelihood ratio test and show that the rejection region is of the form $\mathcal{R} = \{\bar{X} e^{-\bar{X}} \leq c\}$;

(b) Suppose n = 10. Show that the rejection region in (a) is of the form $\mathcal{R} = \{\bar{X} \leq x_0\} \cup \{\bar{X} \geq x_1\}$, where x_0 and x_1 are determined by c;

(c) When $\theta = 1$, it is known that $\sum_i X_i$ follows $Gamma(n, \frac{1}{\theta})$. How could this knowledge be used to choose c?

Answer:

(a) Since the `MLE` of the parameter θ for the exponential distribuiton is given by:

$$\hat{\theta} = 1/\bar{X},$$

We have that the `GLRT` is given by:

$$\begin{aligned}
\Lambda &= \frac{max_{\theta \in \Theta_0 \cup \Theta_1} L(\theta; x)}{max_{\theta \in \Theta_0} L(\theta; x)} = \frac{max_{\theta \in \Theta_0 \cup \Theta_1} \theta^n e^{-\theta \sum_i x_i}}{max_{\theta \in \Theta_0} \theta^n e^{-\theta \sum_i x_i}} \\
&= \frac{(1/\bar{X})^n e^{-1/\bar{X} \sum_i x_i}}{\theta_0^n e^{-\theta_0 \sum_i x_i}} \\
&= \frac{(1/\bar{X})^n e^{(-1/\bar{X}) \ n \ (\sum_i x_i/n)}}{1^n e^{-1 \ \sum_i x_i}} \\
&= \frac{(1/\bar{X})^n e^{-n}}{e^{-n \ (\sum_i x_i)/n}} \\
&= (1/\bar{X})^n e^{-n} e^{n\bar{X}}
\end{aligned}$$

Our criterion for rejecting the Null Hypothesis is that this ratio be greater than some constant:

$$\Lambda \geq k^*$$

We can simplify the above expression for Λ further by taking the n-th root of both sides of this expression:

$$\begin{aligned}
\Lambda^{\frac{1}{n}} &= \left((1/\bar{X})^n e^{-n} e^{n\bar{X}} \right)^{\frac{1}{n}} \\
&\geq (k^*)^{\frac{1}{n}}
\end{aligned}$$

$\Rightarrow$

$$(1/\bar{X})\mathrm{e}^{-1}\mathrm{e}^{\bar{X}} \geq (k^*)^{\frac{1}{n}}$$

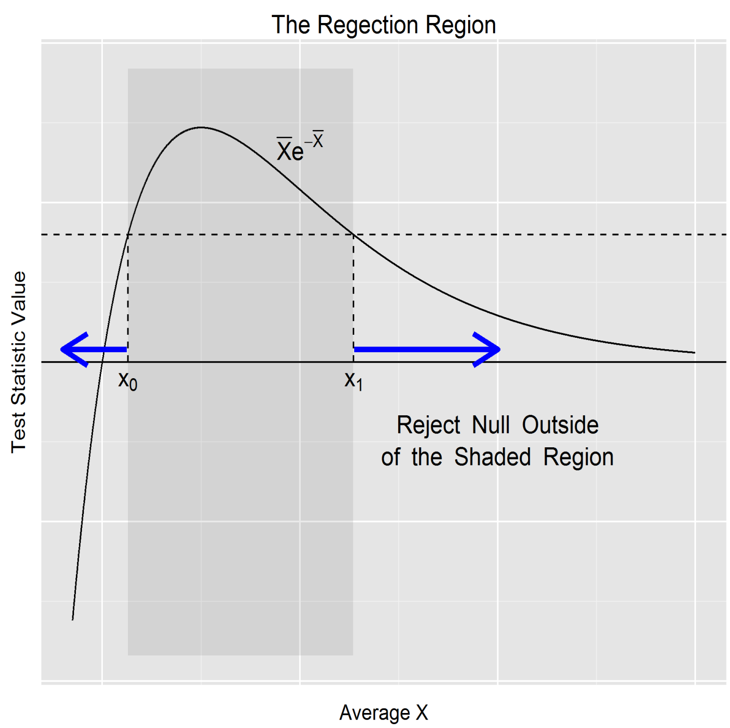
$\Rightarrow$

$$\frac{1}{(1/\bar{X})\mathrm{e}^{\bar{X}}} = \bar{X}\mathrm{e}^{-\bar{X}} \leq \frac{1}{(k^*)^{\frac{1}{n}}\mathrm{e}} = c$$

Thus we have shown that the rejection region $\mathcal{R}$ is given by:

$$\boxed{\bar{X}\mathrm{e}^{-\bar{X}} \leq \text{ some constant } c}$$

(b) Some people like to claim that a picture is worth a few words[1]. Here's what the rejection region from part (a) might look like:



The dashed horizontal line in this picture corresponds to the constant c from part (a). The test statistic has to be below this constant, that is $\bar{X}$ has to fall outside the shaded

[1]One author of this homework has heard of pictures being worth thousands of words, but he's skeptical of this claim.

region, for us to reject H_0. This picture shows that x_0 and x_1 are determined by the constant c: at these points our test statistic $\bar{X}\mathrm{e}^{-\bar{X}}$ intersects the horizontal line at level c.

(c) To simplify the notation, let $Y = \sum_i^{10} X_i$. We know that under the null hypothesis Y follows the $Gamma(10, 1)$ distribution. We rewrite the expression we got in part (a) as follows:

$$\bar{X}\mathrm{e}^{-\bar{X}} \leq c \quad \Rightarrow \quad \frac{Y}{10}(\mathrm{e}^{-Y})^{\frac{1}{10}} \leq c$$

$$\Rightarrow \quad Y^{10}\mathrm{e}^{-Y} \leq (10c)^{10}$$

Thus we are looking for a constant c such that

$$P\left(Y^{10}\mathrm{e}^{-Y} \leq (10c)^{10}\right) = \alpha = 0.05.$$

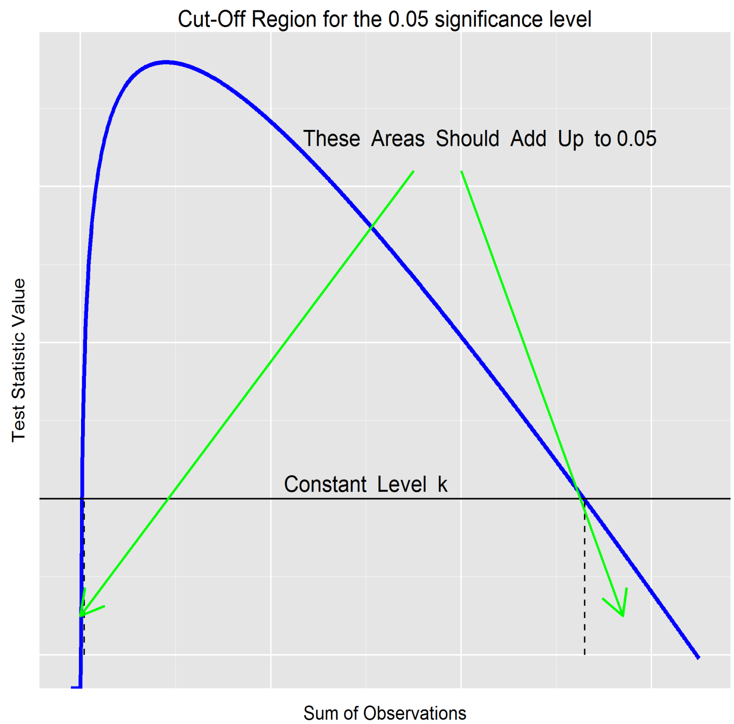
Or, taking the logarithms of both sides of the inequality,

$$P\left(10\log Y - Y \leq \log\left((10c)^{10}\right)\right) = 0.05.$$

Again, for simplicity of notation, denote $\log((10c)^{10})$ by k. Thus, for $Y \sim Gamma(10, 1)$, we are looking for such constant k that satisfies the following inequality:

$$\boxed{P\ (\ 10\log Y - Y \leq k\ ) = 0.05}$$

We can plot $10\log Y - Y$ and find the constant k that makes the area under the curve, and below this constant, equal to $\alpha = 0.05$. (We can then easily recover c from k, if need be).



Question 6.3. Suppose, to be specific, that in Problem 2, the observed data are the following:

`1.07 0.88 0.66 0.55 1.15 0.65 3.45 3.55 3.51 0.48`

(a) Based on the result in Problem 2, will you reject H_0? What's your p-value?

(b) If we start from generalized likelihood ratio test, and use the asymptotic distribution of $2 \log \Lambda$, will you reject H_0? What's you p-value?

Answer: