Problem 12

Let $X_1 \dots X_n$ be iid from an exponential distribution with $f(x|\theta) = \theta e^{-\theta x}, x \geq 0$. Derive a likelihood test for $H_0 : \theta = \theta_0$ versus $H_A : \theta \neq \theta_0$ and show that the rejection region is of the form $\bar{X}e^{-\theta_0 \bar{X}} \leq c$.

The likelihood ratio is

$$\Lambda = \frac{\max_{\theta=\theta_0} lik(\theta)}{\max_{\theta \in \mathbb{R}^+} lik(\theta)}$$

where $lik(\theta|x) = \theta^n e^{-\theta \sum X_i}$. The maximum likelihood estimator of an exponential distribution is $\theta^* = \frac{1}{\bar{X}}$. Since small values of Λ are evidence against the null, the test rejects the null when

$$\Lambda = \frac{\theta_0^n e^{-\theta_0 n \bar{X}}}{(\frac{1}{\bar{X}})^n e^n} < c$$

θ_0^n is a constant and can be immediately absorbed into c. For the exponent, $\bar{X}^n e^{-\theta_0 n \bar{X} - n} = (\bar{X}e^{-\theta_0 \bar{X} - 1})^n$, and since Λ is positive, taking the nth root won't introduce more inequalities. Thus by taking the nth root of both sides, e^{-1} can be absorbed into c, and so the rejection region has the form

$$\bar{X}e^{-\theta_0 \bar{X}} < c$$

as desired.

Problem 13

Continuing from the previous problem, suppose that $\theta_0 = 1$, $n = 10$, and $\alpha = .05$.

Part a

Show that the rejection region is of the form $\{\bar{X} \leq x_0\} \cup \{\bar{X} \geq x_1\}$, where x_0, x_1 are determined by c.

First note that $f(\bar{X}) = \bar{X}e^{-\theta_0 \bar{X}}$ is continuously differentiable. Note that $f(\bar{X})$ is always nonnegative, $f(0) = 0$, and $\lim_{\bar{X} \to \infty} f(\bar{X}) = 0$. Since $f'(\bar{X})$ is continuous, $f'(\bar{X}) = 0$ is a necessary condition for a local extrema. $f(\bar{X}) = 0$ when $\bar{X}^* = \frac{1}{\theta_0}$, and since this is the only inflection point of f, $f(\bar{X})$ must have a local extrema at $\bar{X}^*$. It's obvious that this is a maximum, since the endpoint limits go to zero.

Therefore, small values of $f(\bar{X})$ correspond to areas towards 0 and infinity, and away from the central hump. This corresponds to $\{\bar{X} \leq x_0\} \cup \{\bar{X} \geq x_1\}$.

Part b

Explain why c should be chosen so that $P(\bar{X}e^{-\bar{X}} \leq c) = .05$ when $\theta_0 = 1$.

Under the null hypothesis of $\theta_0 = 1$, the test statistic $\bar{X}e^{-\bar{X}}$ being small is evidence against the null. We want a 95 percent confidence set, so under the null, $P(rejectnull) = .05$. We reject the null when the test statistic is below a level c, so we need to choose c such that $P(rejectnull) = P(\bar{X}e^{-\bar{X}} \leq c) = .05$.

Part c

Explain why $\sum_{i=1}^{10} X_i$ and $\bar{X}$ follow gamma distributions. Use this information to choose c.

It's well established that when X and Y are iid Gamma with the same rate parameters, their sum is a Gamma random variable with its rate parameter equal to X and Y's common rate parameter, and its shape parameter equal to the sum of X and Y's shape parameters. Since an exponential random variable is a $Gamma(1, \theta_0)$ random variable, $\sum_{i=1}^{10} X_i$ is distributed $Gamma(10, \theta_0)$.

Let X be a Gamma random variable, and $Y = \frac{X}{m}$, with m a positive constant. Then $F_Y y = F_X cy$, and

$$f_Y(y) = cf_X cy = m\frac{\theta^\alpha}{\Gamma(\alpha)}(my)^{\alpha-1}e^{-\theta(my)} = \frac{(\theta m)^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-(\theta m)y}, y \geq 0$$

which is a $Gamma(\alpha, \theta m)$ random variable. Thus, since $\sum_{i=1}^{1} 0 \sim Gamma(10, \theta_0)$, $\bar{X} = \frac{1}{10}\sum_{i=1}^{1} 0 \sim Gamma(10, 10\theta_0)$. Under the assumption that $\theta_0 = 1$, $\bar{X} \sim Gamma(10, 10)$.

Using this knowledge, we can use a transformation to find the distribution of $\bar{X}e^{-\bar{X}}$. However, I do not know how to do this, as the inverse of $\bar{X}e^{-\bar{X}}$ is $-W(-\bar{X})$ according to Wolfram Alpha, where $W(x)$ is the Lambert W function. I do not know enough about this function at the moment to numerically determine c.

Part d

Explain how to approximate c by simulation.

Since we know the distribution of $\bar{X}$, we can approximate the distribution of $\bar{X}e^{-\bar{X}}$. First we generate n $Gamma(10, 10)$ variables, apply $f(x) = xe^{-x}$ to the generated numbers, and rank them from smallest to largest. These numbers estimate the distribution of $\bar{X}e^{-\bar{X}}$, and so the $.05n$th number from the bottom is our estimate for c such that $P(\bar{X}e^{-\bar{X}} < c)$.