Zhangsheng Lai
1002554

Statistics: Homework 3

10.5 Given $X_1, \dots, X_n \sim \text{Uniform}(0, \theta)$ and $Y = \max\{X_1, \dots, X_n\}$, we have the cdf of Y to be $F_Y(y) = (y/\theta)^n$ for $y \in [0, 1/2]$.

(a) When we choose to reject H_0 when $Y > c$, the power function is $\beta(\theta) = 1 - (c/\theta)^n$, $c \in [0, 1/2]$.

(b) Given size of the test to be .05, we need to solve,

$$1 - (2c)^n = .05$$

which gives us a solution of $c = 1/2(.95)^{1/n}$

(c) The size, $\alpha = \beta(1/2) = 1 - (2c)^n$, $c \in [0, 1/2]$. Thus, when $n = 20, Y = .48$, the p-value is

$$\inf\{\alpha : X^n \in R_\alpha\} = 1 - (2 \times .48)^{20} = 0.557997566$$

We would conclude that we do not reject H_0 with an approximate probability of 0.56, which does not give a strong evidence to reject H_0

(d) When $n = 20, Y = .52$, using the α formula in (c) gives us $1 - (2 \times .52)^{20} = -1.19112314$. But the given $Y = .52 > 1/2$ which is out of the defined boundaries of the size, i.e. $F_Y(0.52; \theta = 1/2) = 0$. Hence the p-value is 0. This allows us to conclude that H_0 is to be rejected as the p-value always lies in the criteria region; a very strong reason to reject H_0.

10.7b Let $H_0 : F_T = F_S$ and $H_1 : F_T \neq F_S$, where the subscripts denote Twain and Snodgrass respectively. The observed value of the test statistic given by the absolute difference of their means, $|\overline{T} - \overline{S}|$ is

$$|0.231875 - 0.2097| = 0.022175$$

```
Have to do some simulation here.
```

Under this p-value, do we reject H_0 at a 5 percent level? How about 2.5 percent level?

10.8 (a) The size of this test with rejection region R is

$$\begin{aligned}\mathbb{P}(T(X^n) > c | \theta = 0) &= \mathbb{P}(\overline{X}_n > c) \\ &= \mathbb{P}\left(Z > \sqrt{n}c\right), \; Z \text{ is the standard normal distribution} \\ &= 1 - \Phi(\sqrt{n}c), \;\; \Phi \text{ is the cdf of the standard normal}\end{aligned}$$

where by Central Limit Theorem, $\overline{X}_n \sim N(0, 1/\sqrt{n})$. Thus given size α, the c is $\Phi^{-1}(1 - \alpha)/\sqrt{n}$

(b) Under $H_1 : \theta = 1$, the power is $\beta(1) = \mathbb{P}(T(X^n) > c | \theta = 1) = 1 - \Phi\left(\sqrt{n}(c - 1)\right)$. Thus when $n \to \infty$, $\sqrt{n}(c - 1) \to \infty$ for $c \neq 1$ which then $1 - \Phi\left(\sqrt{n}(c - 1)\right) \to 1$.

(c)

10.12 (a) We known that the MLE for λ is $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$. The Fisher information $I_n(\lambda)$ is

$$I_n(\lambda) = nI(\lambda) = -n\mathbb{E}_\lambda \left(\frac{\partial^2 f_X(X; \lambda)}{\partial \lambda^2}\right) = -n\mathbb{E}_\lambda \left(-\frac{X}{\lambda^2}\right) = \frac{n}{\lambda}$$

thus by the property of MLE,

$$\frac{\overline{X}_n - \lambda}{\hat{\mathsf{se}}} \rightsquigarrow N(0, 1)$$

thus the size of of the Wald test

$$\mathbb{P}\left(\left|\frac{\overline{X}_n - \lambda_0}{\sqrt{\lambda_0/n}}\right| > z_{\alpha/2}\right)$$

(b)

```python
import numpy as np
from scipy.stats import norm
def poisson_sample(l, n):
    """
    Generates n Poisson distributed samples with parameter l.
    """
    return np.random.poisson(lam = 1, size = n)
def wald_test(sample, n = 20, alpha = .05, null_lambda = 1):
    """
    Perfoms Wald test and returns p-value.
    """
    xbar = np.mean(sample)
    test_statistic = np.absolute((xbar - null_lambda)/ (null_lambda / n) ** 0.5)
    return 2 * (1 - norm.cdf(test_statistic))
def multwald(l = 1, n = 20, alpha = .05, null_lambda = 1, B = 10000):
    """
    Performs Wald test B times and return proportion of test where null hypothesis is
        rejected.
    """
    count = 0
    for i in np.arange(B):
        sample = poisson_sample(l, n)
        if wald_test(sample) < alpha:
            count += 1

    return count/B
multwald()
```

From performing the simulation of Wald 10000 times, the proportion of null rejected is 0.0564 which is very close to the type I error rate of α.

11.3

11.4