Cooper Golemme

Question 1: Uniform

Let $X_1, \ldots, X_n \sim \text{Uniform}(0, \theta)$ and let $Y = \max\{X_1, \ldots, X_n\}$. We want to test

$$H_0 : \theta = \frac{1}{2} \quad \text{versus} \quad H_1 : \theta > \frac{1}{2}.$$

The Wald test is not appropriate since Y does not converge to a Normal distribution. Suppose we decide to test this hypothesis by rejecting H_0 when $Y > c$.

(a) Find the power function.

(b) What choice of c will make the size of the test 0.05?

(c) In a sample of size $n = 20$ with $Y = 0.48$, what is the p-value? What conclusion about H_0 would you make?

(d) In a sample of size $n = 20$ with $Y = 0.52$, what is the p-value? What conclusion about H_0 would you make?

1.1 Power Function

The power function is define as the probability of rejecting H_0 given that H_1 is true, or more formally:

$$\beta(\theta) = \mathbb{P}(x \in R | \theta)$$

where R is the rejection region.

Consider Y the statistic that is the maximum of all the X_i's. We reject H_0 when $Y > c$. So the power function is given as

$$\beta(\theta) = \mathbb{P}(Y > c | \theta)$$

As we assumed that Y is a statistic of θ, the power function probability is already implicity conditioned on θ, so we get that

$$\beta(\theta) = \mathbb{P}(Y > c)$$

$$
\begin{aligned}
\beta(\theta) &= \mathbb{P}(Y > c) && (1)\\
&= 1 - \mathbb{P}(Y \leq c) && (2)\\
&= 1 - \mathbb{P}(\max\{X_1, \ldots, X_n\} \leq c) && (3)\\
&= 1 - \mathbb{P}(X_1 \leq c)\mathbb{P}(X_2 \leq c)\ldots\mathbb{P}(X_n \leq c) && (4)\\
&= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i \leq c) && (5)\\
&= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i \leq c) && (6)
\end{aligned}
$$

Each X_i are uniform so the density function is

$$
f_{X_i}(x) = \begin{cases} \frac{1}{\theta} & \textbf{if } 0 \leq x \leq \theta \\ 1 & \textbf{if } x > \theta \\ 0 & \textbf{if } x < 0 \end{cases}
$$

Thus the CDF for X_i is given as:

$$
\mathbb{P}(X_i \leq x) = F_{X_i}(x) = \begin{cases} \frac{x}{\theta} & \textbf{if } 0 \leq x \leq \theta \\ 1 & \textbf{if } x > \theta \\ 0 & \textbf{if } x < 0 \end{cases}
$$

So returning to the calculation for $\beta(\theta)$, we have that

$$
\begin{aligned}
\beta(\theta) &= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i \leq c) && (7)\\
\beta(\theta) &= 1 - \prod_{i=1}^{n} \begin{cases} \frac{c}{\theta} & \textbf{if } 0 \leq c \leq \theta \\ 1 & \textbf{if } c > \theta \\ 0 & \textbf{if } c < 0 \end{cases} && (8)
\end{aligned}
$$

Lets reclassify this function by analysis.

- If $c < 0$, then $\beta(\theta) = 1 - 0 = 1$
- If $c > \theta$, then $\beta(\theta) = 1 - 1 = 0$

- Else then $\beta(\theta) = 1 - \prod_{i=1}^{n} \frac{c}{\theta} = 1 - (\frac{c}{\theta})^n$

Thus, $\beta(\theta)$ is as follows:

$$\beta(\theta) = \begin{cases} 1 \ \ \textbf{if } c < 0 \\ 0 \ \ \textbf{if } c > \theta \\ 1 - (\frac{c}{\theta})^n \ \ \textbf{else} \end{cases}$$

1.2 Size

The size of a test is defined as

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$$

H_0 in this case has only 1 parameter $\theta = 1/2$ so we can use this to solve for α as follows:

$$\sup_{\theta \in \{1/2\}} \beta(\theta) = \alpha \tag{9}$$

$$\beta(1/2) = \alpha \qquad \text{as } \Theta_0 = \{1/2\} \tag{10}$$

$$\alpha = \beta(1/2) = \begin{cases} 1 \ \ \textbf{if } c < 0 \\ 0 \ \ \textbf{if } c > 1/2 \\ 1 - (\frac{c}{1/2})^n \ \ \textbf{else} \end{cases} \tag{11}$$

We want to find a choice of c that makes the size $\alpha = 0.05$

$$0.05 = \beta(1/2) = \begin{cases} 1 \ \ \textbf{if } c < 0 \\ 0 \ \ \textbf{if } c > 1/2 \\ 1 - (2c)^n \ \ \textbf{else} \end{cases} \tag{12}$$

$\beta(1/2) = 1$ if $c < 0$ and 0 if $c > 1/2$ so we should choose a c that is $0 < c < 1/2$ to get a value for $\beta(1/2) = 0.05$

$$
\begin{aligned}
0.05 &= \beta(1/2) = 1-(2c)^n && \text{if } 0<c<1/2 && (13)\\
0.05 &= 1-(2c)^n && && (14)\\
(2c)^n &= 0.95 && && (15)\\
2c &= 0.95^{\frac{1}{n}} && && (16)\\
c &= \frac{1}{2}\dot{0}.95^{\frac{1}{n}} && && (17)
\end{aligned}
$$

Thus when $c = \frac{1}{2}\dot{0}.95^{\frac{1}{n}}$, the size of the test is 0.05.

1.3 Sample

In a sample of size $n = 20$ with $Y = 0.48$, what is the p-value? What conclusion about H_0 would you make?

We reject H_0 when Y lies in the rejection region i.e $Y > c$. We showed in the previous part that for a given α, $c = \frac{1}{2}(1-\alpha)^{\frac{1}{n}}$.

So if $Y = 0.48$ and $Y > c$, then $0.48 > \frac{1}{2}(1-\alpha)^{\frac{1}{20}}$

Rearranging this we get $\alpha > 1-(2\cdot 0.48)^{20}$, which simplifies to $\alpha > 0.55799$. This implies that we are 55% confident that H_0 is true, which is basically nothing, almost a 50/50 chance which is only slightly better than guessing randomly.

1.4 Sample 2

In a sample of size $n = 20$ with $Y = 0.52$, what is the p-value? What conclusion about H_0 would you make?

Returning to the power function, which we calculated in the previous part:

$$
\beta(1/2) = \begin{cases} 1 & \textbf{if } c < 0 \\ 0 & \textbf{if } c > 1/2 \\ 1-(2c)^n & \textbf{else} \end{cases}
$$

Even if we set $\alpha = 0$, we get that:

$$
\begin{aligned}
0 &= \beta(1/2) = 1-(2c)^n && \text{if } 0<c<1/2 && (18)\\
c &= \frac{1}{2} && && (19)
\end{aligned}
$$

And we know that $Y = 0.52$, so even with the lowest possible value for α, $Y > c$, so the p-value for this is 0. This means that our test in this case will always reject H_0.

This makes sense intuitively as our hypothesis is that $X_1, X_2, \ldots, X_n$ is sampled from a uniform distribution $0, 1/2$ and $Y = \max(X_1, X_2, \ldots, X_n)$. It is not possible that $Y > 1/2$ because a uniform distribution assigns a probability of 0 to all values that are not within its bounds. Thus, a good test should reject things that are fundementally impossible.

Question 2: Mark Twain

In 1861, 10 essays appeared in the *New Orleans Daily Crescent.* They were signed "Quintus Curtius Snodgrass" and some people suspected they were actually written by Mark Twain. To investigate this, we will consider the proportion of three-letter words found in an author's work. From eight Twain essays we have:

$$0.225,\ 0.262,\ 0.217,\ 0.240,\ 0.230,\ 0.229,\ 0.235,\ 0.217$$

From 10 Snodgrass essays we have:

$$0.209,\ 0.205,\ 0.196,\ 0.210,\ 0.202,\ 0.207,\ 0.224,\ 0.223,\ 0.220,\ 0.201$$

(a) Perform a Wald test for equality of the means. Use the nonparametric plug-in estimator. Report the p-value and a 95% confidence interval for the difference of means. What do you conclude?

Let

$$H_0 : \mu_T - \mu_S = 0 \quad \text{versus} \quad H_1 : \mu_T - \mu_S \neq 0$$

In order to perform a Wald test, we have to have an estimator that is asymtoically normal. As we do not know the parameter for the underlying random process, we can use the difference in sample means as the estimator for our Wald test.

This works because we know from the central limit theorem that, the sample mean is asymtotically normal under the right circumstances. Specifically, for non-parametric estimator $\hat{\mu}$:

$$\frac{\hat{\mu} - \mu}{\hat{SE}} \to N(0, 1)$$

So in our context we have it that the Wald Statistic w is:

$$w = \frac{\mu_T - \mu_S - 0}{\hat{SE}}$$

We need to find each part of this in order to continue with the analysis. First, we will compute the sample means. The sample mean for each of the authors three-letter word frequencies is:

$$\hat{\mu_T} = 0.231875 \quad \text{and} \quad \hat{\mu_S} = 0.2097$$

Thus the sample difference between the two is:

$$\hat{\mu_T} - \hat{\mu_S} = 0.022175$$

Now we must computer the standard error for this estimator

$$SE(\hat{\mu_T} - \hat{\mu_S}) = \sqrt{Var(\hat{\mu_T} - \hat{\mu_S})} \tag{20}$$

$$= \sqrt{Var(\hat{\mu_T}) + Var(\hat{\mu_S})} \tag{21}$$

The variance for the empirical mean is defined as

$$Var(\hat{\mu}) = \frac{1}{n-1}\sum_{i=1}^{n}[X_i - \hat{\mu}]^2$$

The $\frac{1}{n-1}$ rather than $\frac{1}{n}$ is to make this unbiased. If we do this calculation for both of these estimators, we get:

$$Var(\hat{\mu_T}) \approx 0.0001856 \quad Var(\hat{\mu_S}) \approx 0.000084009$$

Plugging this back into the formula for $\hat{SE}$

$$\hat{SE} \approx \sqrt{0.0001856 + 0.000084009} \approx 0.00598$$

Plugging this in to the formula for the w statistic:

$$w = \frac{\mu_T - \mu_S - 0}{\hat{SE}} \approx \frac{0.0221}{0.00598} \approx 3.70$$

2.1 95 Confidence Interval

To computer 95% confidence interval we take:

$$C = (\hat{\mu_T} - \hat{\mu_S} - Z_{0.025} * SE(\hat{\mu_T} - \hat{\mu_S}), \hat{\mu_T} - \hat{\mu_S} + Z_{0.025} * SE(\hat{\mu_T} - \hat{\mu_S}))$$

Computing this we get

$$C = (0.01232, 0.03202)$$

2.2 p-value

We want to find the value for α such that

$$w = Z_{\frac{\alpha}{2}}$$

$$3.70 = Z_{\frac{\alpha}{2}}$$

$$\alpha = 0.000212$$

Thus we have a p-value of 0.000212.

2.3 Conclusion

With such a small p-value, there is very high confidence that we can reject the null hypothesis. In this case, the null hypothesis was that the samples were written by the same author, so with great certainty, we can refute this claim.

Question 3: Poisson Wald Test

Let $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$.

(a) Let $\lambda_0 > 0$. Find the size α Wald test for

$$H_0 : \lambda = \lambda_0 \quad \text{versus} \quad H_1 : \lambda \neq \lambda_0.$$

(b) (*Computer Experiment.*) Let $\lambda_0 = 1$, $n = 20$, and $\alpha = 0.05$. Simulate $X_1, \ldots, X_n \sim \text{Poisson}(\lambda_0)$ and perform the Wald test. Repeat many times and count how often you reject the null. How close is the type I error rate to 0.05?

3.1 Wald Test

To use a Wald Test, we need a statistic that is asymptotically normal. For this we can use the MLE for the Poisson distribution, which is given as

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

We can compute the W statistic as

$$W = \frac{\hat{\lambda} - \lambda_0}{\hat{SE}}$$

The standard error for the MLE estimator for a Poisson random variable is given as

$$SE(\hat{\lambda}) = \frac{\sqrt{\sum_{i=1}^{n} X_i}}{n}$$

Proof:

$$\begin{aligned}
Var(\hat{\lambda}) &= Var(\frac{1}{n}\sum_{i=1}^{n} X_i) && (22)\\
&= \frac{1}{n^2}Var(\sum_{i=1}^{n} X_i) && (23)\\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \qquad \text{because each } X_i \text{ are independent} && (24)\\
&= \frac{1}{n^2}\sum_{i=1}^{n} \lambda \qquad \text{variance of possion}(\lambda) && (25)\\
&= \frac{\lambda}{n} && (26)\\
Se(\hat{\lambda}) &= \sqrt{\frac{\lambda}{n}} && (27)
\end{aligned}$$

Because MLE is equivarient, we can replace λ in this expression with $\hat{\lambda}$ and we get that $\hat{SE} = \sqrt{\frac{\hat{\lambda}}{n}}$. We can simplify this knowing that $\hat{\lambda} = \sum_{i=1}^{n} X_i$ to $\hat{SE} = \sqrt{\frac{\sum_{i=1}^{n} X_i}{n^2}}$

Thus, the Wald statistic is

$$W = \frac{n((\frac{1}{n}\sum_{i=1}^{n} X_i) - \lambda_0)}{\sqrt{\sum_{i=1}^{n} X_i}} = \frac{\sum_{i=1}^{n} X_i - n\lambda_0}{\sqrt{\sum_{i=1}^{n} X_i}}$$

If we want to create a size α test, we need to use the power function as follows

$$\beta(\lambda_0) = \mathbb{P}\left(\left|\frac{\sum_{i=1}^{n} X_i - n\lambda_0}{\sqrt{\sum_{i=1}^{n} X_i}}\right| > z_{\alpha/2}\right) \tag{28}$$

3.2 Computer Experiment

```python
import numpy as np
from scipy.stats import norm

lambda0 = 1
n = 20
alpha = 0.05
num_simulations = 10000
reject_count = 0

for _ in range(num_simulations):
    X = np.random.poisson(lambda0, n)
    sum_data = np.sum(X)
    sample_mean = np.mean(X)

    wald_statistic = (n * (sample_mean - lambda0)) / np.sqrt(sum_data)

    p_value = 2 * (1 - norm.cdf(abs(wald_statistic)))  # Two-tailed test

    if p_value < alpha:
        reject_count += 1

type_I_error_rate = reject_count / num_simulations
print(f"Estimated Type I Error Rate: {type_I_error_rate:.4f}")
```

```
Estimated Type I Error Rate: 0.0509
```

The type 1 error rate from this simulation is 0.0509 whcih is very close to 0.05.

Question 4: Supplemental

Suppose we sample $x_1, \ldots, x_n$ i.i.d. from $N(0, 1)$. Let

$$H_0 : \mu = 0 \quad \text{versus} \quad H_1 : \mu \neq 0.$$

Let $n = 100$. Simulate $T = 10,000$ trials of sampling and running a Wald test at the $\alpha = 0.01$ level. How many times do you reject H_0? Interpret this in terms of Type I errors and the meaning of the significance level.

```python
import numpy as np
from scipy.stats import norm

n = 100
alpha = 0.01
T = 10000
reject_count = 0

for _ in range(T):
    X = np.random.normal(0, 1, n)

    sample_mean = np.mean(X)

    # SE = 1/sqrt(n) for N(0,1)
    wald_statistic = sample_mean / (1 / np.sqrt(n))

    p_value = 2 * (1 - norm.cdf(abs(wald_statistic)))

    if p_value < alpha:
        reject_count += 1

print(f"Number of rejections of H0: {reject_count}")
```

```
Number of rejections of H0: 112
```

In this experiment, we fix the mean to be 0 and are testing whether the test that we create goes along with what we would expect. We expect that we will reject H_0 around $\alpha \times T$ times as the Wald Test level can be interpreted as the amount of error that we allow for the test to have. Specifically, it is the amount of Type 1 errors that we are allowed to have (when we reject the null hypothesis when it is indeed true). In this case we know the null hypothesis is true, as we are indeed sampling from a normal distribution with mean 0

When we run this experiment, we rejected H_0 112 times out of 10000. This means that we rejected H_0 about 0.96% of the time. This is very close to the 1% Type 1 error rate that we are allowing when we set $\alpha = 0.01$