Sheil Sarda
ESE402 Assignment 3
Prof. Hassani
Oct. 25, 2021

Problem 1

PROBLEM 1 A

Given: $\bar{X} = 12$, $s^2 = 5$, $n = 5$, $\sigma^2 = 9$

Hypothesis Test: $H_0: \mu = 10$, $H_a: \mu \neq 10$, $\alpha = 0.05$

~~Cut off~~ Sample Z-score to evaluate against threshold ~~make decision rule~~: $\dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = \dfrac{12-10}{\sqrt{5}/\sqrt{5}} = 2$

$\Rightarrow$ Decision Rule: accept / reject if $Z \geq \underline{?}$ $\Leftarrow$ determined based on 2-tailed value for $\alpha = 0.05$
$= 1.960$

Corresponding p-value to $Z = 2$ in a 2-tailed scenario $= 0.0455$

Since $0.0455 < \alpha\ (0.05)$, we can reject H_0

i.e. we can reject the hypothesis that $\mu = 10$

PROBLEM 1 B

Given: $n = 5 \Rightarrow \#DOF = n - 1 = 4$

$\alpha = 0.05$, 2-tailed ($\because H_a: \mu \neq 10$)

Using a t-value lookup table, corresponding t-statistic is 2.776

Thus, critical rejection regions are where $|Z| > 2.776$

$\Rightarrow -2.776 \leq t \leq 2.776$ $\Leftarrow$ This is a 95% Confidence Interval in terms of T-value

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{12 - \mu}{\sqrt{5}/\sqrt{5}}$$

$\Rightarrow$ 95% Confidence Interval for μ:

$$12 - 2.776 \leq \mu \leq 12 + 2.776$$

$$9.224 \leq \mu \leq 14.776$$ $\Leftarrow$ CI for population mean

Problem 2

PROBLEM 2A

Given:

$n = 51$, $\hat{p} = \frac{41}{51}$, $\alpha = 0.01$

Hypothesis Test: $H_0 : p > 50\%$; $H_a : p \leq 50\%$

We know 1-sided z-score for $\alpha = 0.01 = 2.326$; rejection region is $z \leq -z_\alpha$

Test Statistic Value $= \dfrac{0.80392 - 0.5}{\sqrt{0.5(1-0.5)/51}} = 4.3408$

Since $z \not\leq -z_\alpha$, we accept H_0

PROBLEM 2B

Confidence Interval for true probability P

$$-2.326 \leq t \leq 2.326$$

$$\Rightarrow -2.326 \leq \frac{0.80392 - p}{\sqrt{0.5(1-0.5)/51}}$$ ⇐ Upper ~~Lower~~ Bound

$$\Rightarrow p \leq \left(\frac{(-2.326)\,0.5}{\sqrt{51}} - 0.80392\right)(-1)$$

$$\Rightarrow P \leq 0.967$$

Similarly, lower bound:

$$\frac{0.80392 - p}{\sqrt{0.5(1-0.5)/51}} \leq 2.326 \Rightarrow P \geq \left(\frac{2.326(0.5)}{\sqrt{51}} - 0.80392\right)(-1)$$

$$\Rightarrow P \geq 0.641$$

99% 1-sided confidence Interval:

CI: $[0.641, 0.967]$

Problem 3

PROBLEM 3A

$H_0 : p = 0.25$; $H_a : p < 0.25$

$n = 100$, $\hat{p} = 0.2$

i) $\alpha = 0.05$, left tailed z-value: $z_\alpha = 1.65$

Test statistic value $= \dfrac{0.2 - 0.25}{\sqrt{0.25(1-0.25)/100}} \Rightarrow z = -1.155$

Rejection region : $z \leq -z_\alpha$

Since $-1.155 \nleq -1.65$, we accept H_0

ii) $\alpha = 0.01$, left-tailed z-value : $z_\alpha = 2.33$; we use same z value computed above

Since $-1.155 \nleq -2.33$, we also accept H_0

PROBLEM 3B

$H_0 : p = 0.25$, $H_a : p \neq 0.25$, $\alpha = 0.05$ & $\alpha = 0.01$

Rejection region: $z \geq z_{\alpha/2}$ OR $z \leq -z_{\alpha/2}$

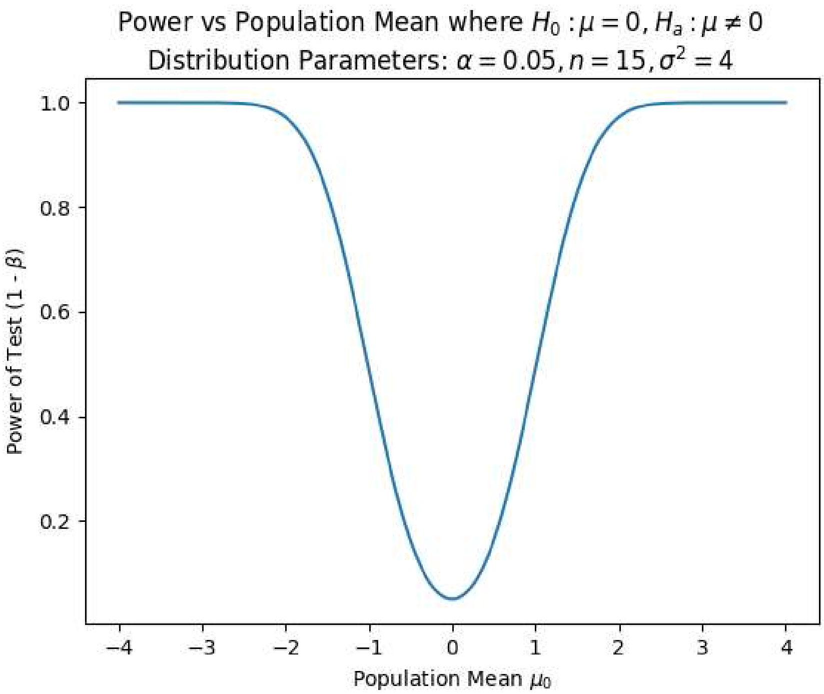
i) $\alpha = 0.05$ 2-tailed z value $= 1.960$

Since $-1.155 \ngeq 1.960$ & $-1.155 \nleq -1.960$, we accept H_0

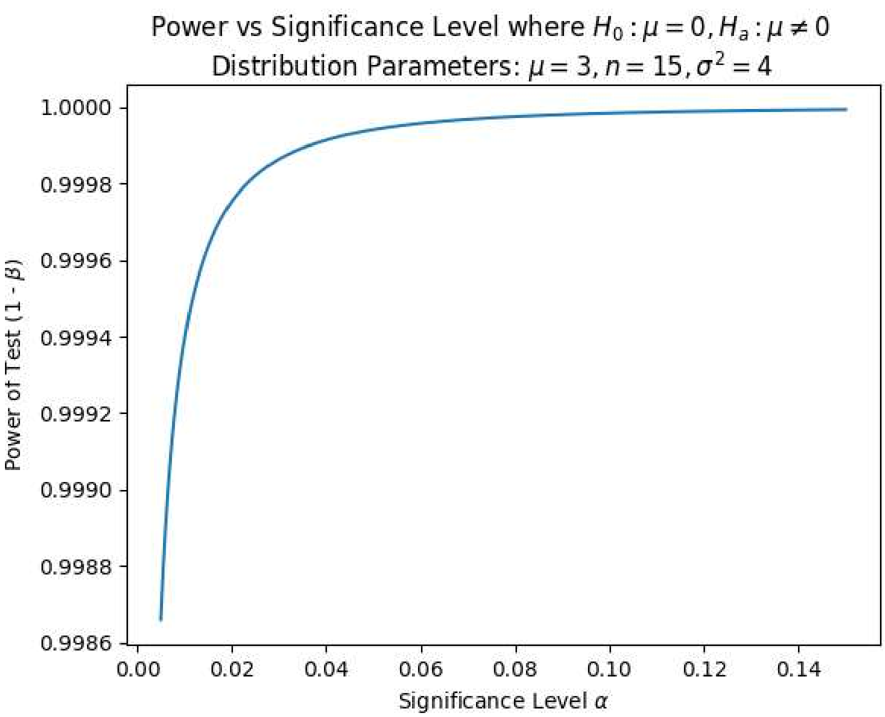
ii) $\alpha = 0.01$, $z_{\alpha/2} = 2.576$

Since $-1.155 \ngeq 2.576$ & $-1.155 \nleq -2.576$, we accept H_0

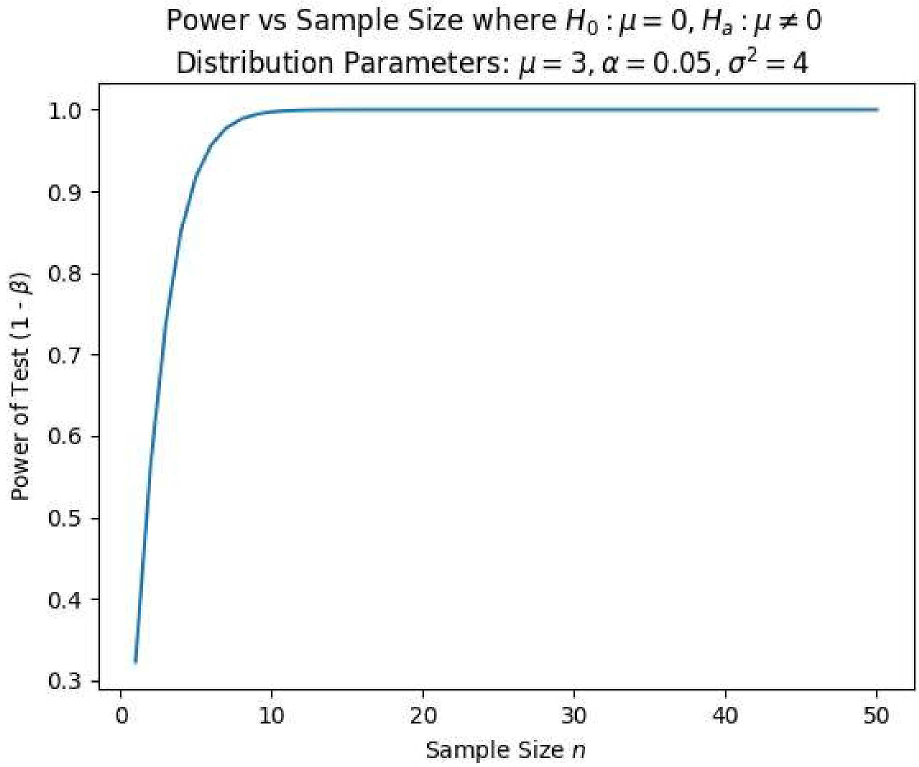
Problem 4A



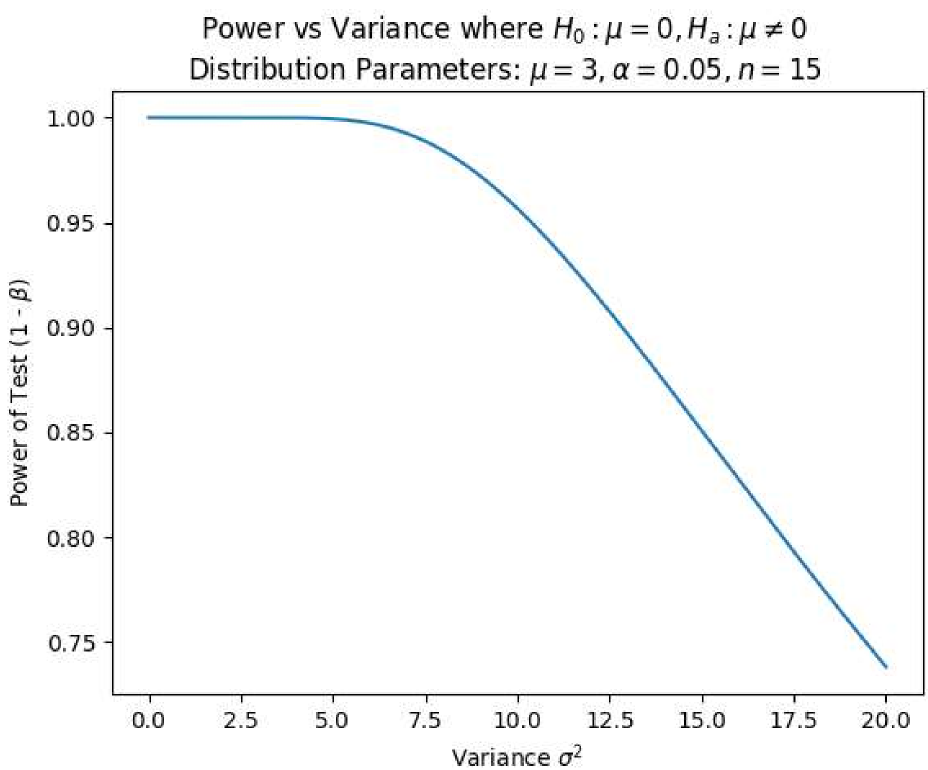
Problem 4B



Problem 4C



Problem 4D



```python
import numpy as np
import scipy.special as sp
import scipy.stats as st
from matplotlib import pyplot as plt
from time import sleep


def generate_mu():
    # Simulate True Mean of Normal Distribution
    mu_array = np.linspace(-4, 4, 1000)
    return mu_array


def generate_alpha():
    # Simulate Probability of Type 1 Errors
    alpha_error = np.linspace(0.005, 0.15, 1000)
    return alpha_error


def generate_n():
    # Simulate Various Sample Sizes
    max_n = 50
    n_array = np.linspace(1, max_n, max_n)
    return n_array


def generate_var():
    # Simulate Different Variances
    sigma_squared_array = np.linspace(0.001, 20, 1000)
    return sigma_squared_array


def compute_power(mu, z_alpha, n, sigma):
    # Power is probability of rejecting the null hypothesis
    # when it is wrong

    # Find acceptance and rejection regions in terms of sample means
    Xbar_lower_bound = -z_alpha * (sigma / n ** 0.5)
    Xbar_upper_bound = z_alpha * (sigma / n ** 0.5)

    # Find power in terms of Z value
    Z_left_tail = (Xbar_lower_bound - mu) / (sigma / n ** 0.5)
    Z_right_tail = (Xbar_upper_bound - mu) / (sigma / n ** 0.5)

    # Compute and return power
    power = abs(sp.ndtr(Z_left_tail)) + (1 - abs(sp.ndtr(Z_right_tail)))
    return power


def plot_power(scenario):
```

```
if (scenario == 'vary_mu'):
    mu_array = generate_mu()
    z_alpha = 1.96  # two-tailed test value for alpha = 0.05
    n = 15  # number of observations
    sigma = 2  # standard deviation
    power_array = []
    for mu in mu_array:
        power = compute_power(mu, z_alpha, n, sigma)
        power_array.append(power)

    plt.plot(mu_array, power_array)
    plt.suptitle(
        r'Power vs Population Mean where $H_0 : \mu = 0, H_a : \mu \neq 0$' + '\n' +
        r'Distribution Parameters: $\alpha = 0.05, n = 15, \sigma^2 = 4$')
    plt.xlabel(r'Population Mean $\mu_0$')

elif (scenario == 'vary_alpha'):
    alpha_array = generate_alpha()
    mu = 3.0  # population mean
    n = 15  # number of observations
    sigma = 2  # standard deviation
    power_array = []
    for alpha in alpha_array:
        z_alpha = abs(st.norm.ppf(alpha / 2.0))
        power = compute_power(mu, z_alpha, n, sigma)
        power_array.append(power)

    plt.plot(alpha_array, power_array)
    plt.suptitle(
        r'Power vs Significance Level where $H_0 : \mu = 0, H_a : \mu \neq 0$' + '\n' +
        r'Distribution Parameters: $\mu = 3, n = 15, \sigma^2 = 4$')
    plt.xlabel(r'Significance Level $\alpha$')

elif (scenario == 'vary_n'):
    n_array = generate_n()
    mu = 3.0  # population mean
    z_alpha = 1.96  # two-tailed test value for alpha = 0.05
    sigma = 2  # standard deviation
    power_array = []
    for n in n_array:
        power = compute_power(mu, z_alpha, n, sigma)
        power_array.append(power)

    plt.plot(n_array, power_array)
    plt.suptitle(
        r'Power vs Sample Size where $H_0 : \mu = 0, H_a : \mu \neq 0$' + '\n' +
        r'Distribution Parameters: $\mu = 3, \alpha = 0.05, \sigma^2 = 4$')
    plt.xlabel(r'Sample Size $n$')
```

```
    elif (scenario == 'vary_sigma_squared'):
        sigma_squared_array = generate_var()
        z_alpha = 1.96  # two-tailed test value for alpha = 0.05
        n = 15  # number of observations
        mu = 3.0  # population mean
        power_array = []
        for sigma_squared in sigma_squared_array:
            sigma = sigma_squared ** 0.5
            power = compute_power(mu, z_alpha, n, sigma)
            power_array.append(power)

        plt.plot(sigma_squared_array, power_array)
        plt.suptitle(
            r'Power vs Variance where $H_0 : \mu = 0, H_a : \mu \neq 0$' + '\n' +
            r'Distribution Parameters: $\mu = 3, \alpha = 0.05, n = 15$')
        plt.xlabel(r'Variance $\sigma^2$')

    plt.ylabel(r'Power of Test (1 - $\beta$)')
    plt.show()


def main():
    scenario_array = ['vary_mu',
                      'vary_alpha',
                      'vary_n',
                      'vary_sigma_squared'
                      ]
    for scenario in scenario_array:
        plot_power(scenario)
    return


if __name__ == '__main__':
    main()
```

Problem 5

PROBLEM 5

$H_0: \lambda = \lambda_0$; $H_a: \lambda = \lambda_1$ where $\lambda_1 > \lambda_0$; Given α_0, determine rejection region

Likelihood Ratio Test:

a) Likelihood of λ_0 : $\prod_{i=1}^{n} \left(\frac{e^{-\lambda_0} \lambda_0^{x_i}}{x_i!} \right)$

b) Likelihood of λ_1 : $\prod_{i=1}^{n} \left(\frac{e^{-\lambda_1} \lambda_1^{x_i}}{x_i!} \right)$

Likelihood ratio test: $\frac{(a)}{(b)} = \frac{e^{-n\lambda_0} (\lambda_0)^{\sum_{i=1}^{n} x_i}}{e^{-n\lambda_1} (\lambda_1)^{\sum_{i=1}^{n} x_i}} < c$

/100

$$= \left(\frac{\lambda_0}{\lambda_1}\right)^{\sum_{i=1}^{n} x_i} < c \cdot e^{n(\lambda_0 - \lambda_1)}$$

/100 = 1.94

Taking log of both sides $\Rightarrow$

$= \sum_{i=1}^{n}$ /100 = 1.97 $< n(\lambda_0 - \lambda_1) + \log c$

We flip inequality because $\log(\lambda_0/\lambda_1)$ is negative $\because \lambda_1 > \lambda_0$ $\Rightarrow$

$$\sum_{i=1}^{n} x_i > \frac{n(\lambda_0 - \lambda_1) + \log c}{\log(\lambda_0/\lambda_1)}$$

$\Downarrow$

We know sum of independent Poisson RVs w/ param. λ_0 is another Poisson$(n\lambda_0)$

constant $= C_1$

$\Rightarrow \quad P_{\lambda=\lambda_0}\left(\sum_{i=1}^{n} x_i > C_1\right) = P(Y > C_1)$ where $Y \sim$ Poisson$(n\lambda_0)$

$\Rightarrow$ Rejection region should be inverse CDF of the Poisson$(n\lambda_0)$ distribution for α_0

Problem 6

PROBLEM 6

$$L(\theta) = \prod_{i=1}^{n} \theta \exp\{-\theta x_i\}$$

$$\Rightarrow \ell(\theta) = n \log\theta - \theta \sum_{i=1}^{n} x_i$$

$$\Rightarrow \ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \hat{\theta} = \frac{1}{\bar{X}}$$

Likelihood Ratio test: $H_0 : \theta = \theta_0$, $H_a : \theta \neq \theta_0$

$$\frac{L(\theta_0)}{L(\hat{\theta})} = \frac{\theta_0^n \exp\{-\theta_0 \sum_{i=1}^{n} x_i\}}{\hat{\theta}^n \exp\{-\hat{\theta} \sum_{i=1}^{n} x_i\}}$$

$$= \left(\frac{\theta_0^{\,n}}{\hat{\theta}^{\,n}}\right) \exp\left\{-n\bar{X}(\theta_0 - \hat{\theta})\right\}$$

$$= \left(\theta_0 \bar{X} \exp\{-\theta_0 \bar{X} + 1\}\right)^n$$

Now, we wish to determine rejection region:

$$\left(\theta_0 e^{-1} \bar{X} \exp\{-\theta_0 \bar{X}\}\right)^n \leq c_1$$

$$\theta_0 e^{-1} \bar{X} \exp\{-\theta_0 \bar{X}\} \leq c_1^{1/n}$$

$$\bar{X} \exp\{-\theta_0 \bar{X}\} \leq \underbrace{e\theta_0^{-1} c_1^{1/n}}_{=c}$$

$\Rightarrow$ rejection region corresponds to $\bar{X} \exp\{-\theta_0 \bar{X}\} \leq c$ where c is defined above