Statistical Inference

Instructor: Mohammadreza A. Dehaqani

Pouya HajiMohammadi Gohari, Parmis Bathayan

Fall 2024

Homework 3

Question 1: Say Hello to Neyman-Pearson

Let $X^{'} = (X_1, \cdots, X_n)$ denote a random sample from the distribution that has the pdf

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right), \quad -\infty < x < \infty$$

1. Assume null and alternative hypothesis as follows:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta = \theta_1$$

   Derive the likelihood ratio and identify the test statistic.

2. Express the rejection region in terms of the test statistic.
3. Compute the power under specified parameters:

$$\theta_0 = 0,\ \theta_1 = 1,\ \alpha = 0.05$$

Answer:

$$L_0 = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_i-\theta_0)^2}{2}\right)$$
$$L_1 = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x_i-\theta_1)^2}{2}\right)$$
$$\frac{L_0}{L_1} = \frac{(\frac{1}{\sqrt{2\pi}})^n \exp\left(-\frac{(x_i-\theta_0)^2}{2}\right)}{(\frac{1}{\sqrt{2\pi}})^n \exp\left(\frac{-(x_i-\theta_1)^2}{2}\right)}$$
$$= \exp\left(\frac{-1}{2}\left(\sum_{i=1}^{n}(x_i-\theta_0)^2 - \sum_{i=1}^{n}(x_i-\theta_1)^2\right)\right)$$
$$= \exp\left(\frac{-1}{2}\left(n\theta_0^2 - n\theta_1^2 - 2n\overline{X}\theta_0 + 2n\overline{X}\theta_1\right)\right)$$

The likelihood ratio is a function of $\overline{X}$ thus it is a test statistic.
The rejection area can be formulated as follows and under H_0 the null distribution of $\overline{X}$ is a normal distribution with mean θ_0 and variance $\frac{1}{n}$.

$$P_{H_0}(\overline{X} > x_0) = P\left(\frac{\overline{X}-\theta_0}{\frac{1}{\sqrt{n}}} > \frac{x_0-\theta_0}{\frac{1}{\sqrt{n}}}\right)$$

Therefore:

$$\frac{x_0-\theta_0}{\frac{1}{\sqrt{n}}} = z_\alpha$$
$$x_0 = \theta_0 + \frac{z_\alpha}{\frac{1}{\sqrt{n}}}$$

Then the rejection area is simply:

$$\overline{X} > \theta_0 + \frac{z_\alpha}{\sqrt{n}}$$

For calculating the power:

$$\begin{aligned} P_{H_1}\left(\overline{X} \geq \theta_0 + \frac{z_\alpha}{\sqrt{n}}\right) &= P\left(\frac{\overline{X}-\theta_1}{\frac{1}{\sqrt{n}}} \geq \frac{\theta_0-\theta_1+\frac{z_\alpha}{\frac{1}{\sqrt{n}}}}{\frac{1}{\sqrt{n}}}\right) \\ &= 1-\phi\left(\frac{\theta_0-\theta_1}{\frac{1}{\sqrt{n}}}+z_\alpha\right) \end{aligned}$$

Question 2: Wald Test

Let $X_1, \cdots, X_n \sim Poisson(\lambda)$.

1. From MLE, find $\hat{\lambda}$.
2. What is the standard deviation of the estimated point($\hat{\lambda}$)?
3. Suppose we test the following hypothesis:

$$\begin{aligned} H_0 &: \lambda = \lambda_0 \\ H_1 &: \lambda \neq \lambda_1 \end{aligned}$$

   If we use the statistic:

$$\frac{\hat{\lambda}-\lambda_0}{\hat{se}}$$

   What is the distribution of this statistic under H_0? Express the rejection area in terms of the test statistic.

4. (Programming) Let λ_0, $n = 20$ and $\alpha = 0.05$. Simulate $X_1, \cdots, X_n \sim Possion(\lambda_0)$ and perform above test. Repeat many times and count how often you reject the null. How close is the Type I error rate to 0.05?

Answer: The Poisson distribution is:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

With MLE:

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda_i^x}{x_i!} \\ &= \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \end{aligned}$$

Log-likelihood:

$$LL(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \ln\lambda - \ln\left(\prod_{i=1}^{n} x_i\right)$$

derivative with respect to λ:

$$\begin{aligned} \frac{\partial LL(\lambda)}{\partial\lambda} = -n + \frac{\sum_{i=1}^n x_i}{\hat{\lambda}} &= 0 \\ \hat{\lambda} &= \frac{\sum_{i=1}^n x_i}{n} \\ &= \overline{X} \end{aligned}$$

The variance of the estimate:

$$
\begin{aligned}
V\left(\hat{\lambda}\right) &= V\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \\
&= \frac{1}{n^2} V\left(\sum_{i=1}^{n} x_i\right) \\
&= \frac{\lambda}{n}
\end{aligned}
$$

standard deviation:

$$
\begin{aligned}
se &= \sqrt{V\left(\hat{\lambda}\right)} \\
&= \sqrt{\frac{\lambda}{n}}
\end{aligned}
$$

Under the null distribution, the rejection area is:

$$
P_{H_0}\left(\frac{\hat{\lambda}-\lambda_0}{\sqrt{\frac{\hat{\lambda}}{n}}} > x_0\right) = \alpha
$$

Thus the rejection area can be obtained from the equation above:

$$
\begin{aligned}
\hat{\lambda} &> \lambda_0 + \frac{\sqrt{\hat{\lambda}}}{n} z_{\alpha} \\
\overline{X} &> \lambda_0 + \frac{\sqrt{\overline{X}}}{n} z_{\alpha}
\end{aligned}
$$

If two tail:

$$
\left|\frac{\hat{\lambda}-\lambda_0}{\sqrt{\frac{\hat{\lambda}}{n}}}\right| > z_{\frac{\alpha}{2}} \tag{2.1}
$$

Question 3: $f_0(x)$ vs $f_1(x)$

Consider two p.d.f's $f_0(x)$ and $f_1(x)$ that are defined as follows:

$$
f_0(x) = \begin{cases} \frac{3}{2} & \text{for } 0 \le x \le \frac{2}{3} \\ 0 & \text{otherwise} \end{cases}
$$

and

$$
f_1(x) = \begin{cases} \frac{9}{2}x & \text{for } 0 \le x \le \frac{2}{3} \\ 0 & \text{otherwise} \end{cases}
$$

Suppose that a single observation X is taken from a distribution for which the p.d.f $f(x)$ is either $f_0(x)$ or $f_1(x)$, and the following simple hypothesis are to be tested:

$$
\begin{aligned}
H_0 &: \; f(x) = f_0(x), \\
H_1 &: \; f(x) = f_1(x)
\end{aligned}
$$

1. Describe a test procedure for which the following equation would be minimum:

$$
\alpha + \beta
$$

where α and β are the probability of type I and II errors, respectively.

2. (Programming) Consider 1000 equally spaced decision boundaries c in the range $[0.\frac{3}{2}]$, for each decision boundary c compute the α and β. Plot the power of the test as a function of the decision boundary c and compare it to the theoretical part.

3. A medical diagnostic test is performed to detect a rare disease. The disease is serious, and a missed diagnosis (α) can result in severe consequences, while a false alarm (β) may lead to unnecessary but harmless additional tests. The test result is based on an observed value X, drawn from a distribution $f(x)$, which is either:

$$H_0: \ f(x) = f_0(x) \text{ (patient is healthy)},$$
$$H_1: \ f(x) = f_1(x) \text{ (patient has the disease)}$$

In medical diagnostics, minimizing false negatives (β) is prioritized, as the consequences of missing a diagnosis are far more severe than unnecessary follow-up tests. To address this, we assign risks to each type of error:

- R_1: Risk associated with α
- R_2: Risk associated with β

The expected risk is defined as:

$$\text{Expected Risk} = R_1\alpha + R_2\beta$$

Provide a procedure to minimize the expected risk by selecting an optimal decision boundary c^*(Derive the expression for the optimal decision boundary c^* based on R_1 and R_2).

Answer: Let's plot them first:

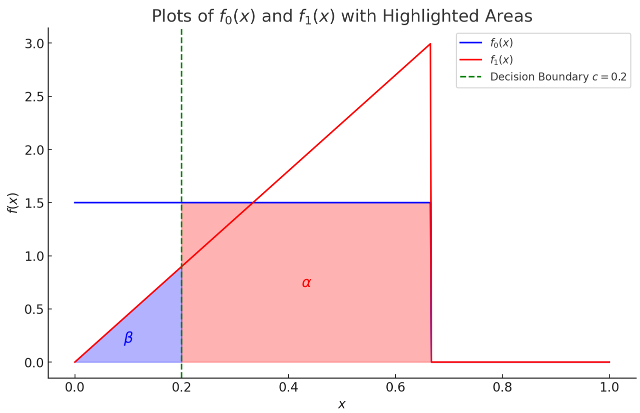


Figure 1: Plot for each distribution

Assume we have set the decision boundary; therefore, the α and β errors would be like in the figure1. The objective is:

$$f(c) = \int_c^{\frac{2}{3}} \frac{3}{2} dx + \int_0^c \frac{9}{2} x dx$$
$$= 1 - \frac{3}{2}c + \frac{9}{4}c^2$$

Derivative with respect to c:

$$\frac{\partial f(c)}{\partial c} = 0$$
$$\frac{9}{2}c - \frac{3}{2} = 0$$
$$c = \frac{1}{3}$$

If we have risk for each error:

$$f(c) = R_1 - \frac{3}{2}R_1 c + \frac{9}{4}R_2 c^2$$

Derivative with respect to c:

$$\begin{aligned}\frac{\partial f(c)}{\partial c} &= 0\\ \frac{9}{2}R_2 c - \frac{3}{2}R_1 &= 0\\ c &= \frac{R_1}{3R_2}\end{aligned}$$

Question 4: Uniformly Most Powerful

Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution with pdf $f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1$, zero elsewhere, where $\theta > 0$. Show the likelihood ratio has the statistic $\prod_{i=1}^n X_i$. Use this to determine the UMP[1] test for $H_0 : \theta = \theta^{'}$ against $H_1 : \theta < \theta^{'}$, for fixed $\theta^{'}$.

Answer: Assume first we want to $H_0 : \theta = \theta^{'}$ against $H_1 : \theta = \theta^{"}$ where $\theta^{"} < \theta'$. These two hypotheses are simple and thus can use Neyman-Pearson as follows:
likelihood ratio:

$$L_0 = \prod_{i=1}^n \theta' x_i^{\theta'-1}$$

$$L_1 = \prod_{i=1}^n \theta^{"} x_i^{\theta^{"}-1}$$

$$\frac{L_0}{L_1} = (\frac{\theta'}{\theta^{"}})^n (\prod_{i=1}^n x_i)^{\theta'-\theta^{"}} \leq k$$

$$\log \prod_{i=1}^n x_i \leq \frac{\log k - n\log\theta' + nlog\theta^{"}}{\theta' - \theta^{"}}$$

The likelihood ratio depends on the product $\prod_{i=1}^n x_i$. When θ'' decreases, the decision boundary increases. Consequently, for every θ'' smaller than θ', this statistic holds true, establishing it as uniformly most powerful.

Question 5: Generalized Likelihood Ratio(Optional)

1. Let $X_1, \cdots, X_n$ and $Y_1, \cdots, Y_n$ be independent random samples from the distributions $N(\theta_1, \theta_3)$ and $N(\theta_2, \theta_4)$, respectively. Show that the likelihood ratio for testing $H_0 : \theta_1 = \theta_2, \theta_3 = \theta_4$ against all alternatives is given by

$$\frac{\left[\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{n}\right]^{n/2} \left[\frac{\sum_{j=1}^m (y_j-\bar{y})^2}{m}\right]^{m/2}}{\left\{\left[\sum_{i=1}^n (x_i-u)^2 + \sum_{j=1}^m (y_j-u)^2\right]/(n+m)\right\}^{(n+m)/2}},$$

   where $\mu = \frac{n\bar{x}+m\bar{y}}{m+n}$.

2. Let the independent random variables X and Y have distributions that are $N(\theta_1, \theta_3)$ and $N(\theta_2, \theta_4)$, where the means θ_1 and θ_2 and common variance θ_3 are unknown. If $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_m$ denote independent random samples from these distribution. Show that the likelihood ratio for testing $H_0 : \theta_1 = \theta_2$, unspecified, and θ_3 unspecified, can be based on the test statistic T(t-distribution) with n+m-2 degrees of freedom.

Answer: With generalized likelihood ratio, we have:

$$\Lambda^* = \frac{\max_{\theta\in\Theta_0} l(\theta)}{\max_{\theta\in\Theta} l(\theta)}$$

Thus we should find the best parameters to maximize the denominator and numerator.

$$l(\theta_1 = \theta_2, \theta_3 = \theta_4) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_3}} \exp\left(-\frac{(x_i-\theta_1)^2}{2\theta_3^2}\right) \prod_{j=1}^m \frac{1}{\sqrt{2\pi\theta_3}} \exp\left(-\frac{(y_j-\theta_1)^2}{2\theta_3^2}\right)$$

Log-likelihood:

$$ll(\theta_1 = \theta_2, \theta_3 = \theta_4) = -\frac{1}{2}\sum_{i=1}^n \ln(2\pi) - \frac{1}{2}\sum_{i=1}^n \ln(\theta_3) - \frac{1}{2}\sum_{j=1}^m \ln(2\pi) - \frac{1}{2}\sum_{j=1}^m \ln(\theta_3) - \sum_{i=1}^n \frac{(x_i-\theta_1)^2}{2\theta_3^2} - \sum_{j=1}^m \frac{(y_j-\theta_1)^2}{2\theta_3^2}$$

[1] Uniformly Most Powerful

Now we will find the optimal θ_1, θ_3:

$$\frac{\partial ll(\theta_1 = \theta_2, \theta_3 = \theta_4)}{\theta_1} = 0$$

$$\frac{1}{2\theta_3^2}\left(\sum_{i=1}^{n}(x_i - \hat{\theta}_1) + \sum_{j=1}^{m}(y_j - \hat{\theta}_1)\right) = 0$$

$$\frac{1}{2\theta_3^2}\left(n\overline{X} - n\hat{\theta}_1 + m\overline{Y} - m\hat{\theta}_1\right) = 0$$

$$\hat{\theta}_1 = \frac{n\overline{X} + m\overline{Y}}{n+m}$$

$$\frac{\partial ll(\theta_1 = \theta_2, \theta_3 = \theta_4)}{\theta_3} = 0$$

$$-\frac{n}{2\theta_3} - \frac{m}{2\theta_3} + \sum_{i=1}^{n}\frac{(x_i - \theta_1)^2}{\theta_3^3} + \sum_{i=1}^{n}\frac{(x_i - \theta_1)^2}{\theta_3^3} = 0$$

$$\hat{\theta}_3 = \frac{\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2 + \sum_{j=1}^{m}(y_j - \hat{\theta}_1)^2}{n+m}$$

Replacing to likelihood of nominator:

$$\max_{\theta \in \Theta_0} l(\theta_1 = \theta_2, \theta_3 = \theta_4) = (\frac{1}{2\pi})^{\frac{n}{2}}(\frac{1}{\hat{\theta}_3})^{\frac{n}{2}}(\frac{1}{2\pi})^{\frac{m}{2}}(\frac{1}{\hat{\theta}_3})^{\frac{m}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2 + \sum_{j=1}^{m}(y_j - \hat{\theta}_1)^2}{2\hat{\theta}_3{}^2}\right)$$

$$= (\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{1}{\hat{\theta}_3})^{\frac{n+m}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2 + \sum_{j=1}^{m}(y_j - \hat{\theta}_1)^2}{2\hat{\theta}_3{}^2}\right)$$

$$= (\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{n+m}{\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2 + \sum_{j=1}^{m}(y_j - \hat{\theta}_1)^2})^{\frac{n+m}{2}} \exp\left(-\frac{n+m}{2}\right)$$

Now for denominator:

$$l(\theta_1, \theta_2, \theta_3, \theta_4) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\theta_3}}\exp\left(-\frac{(x_i - \theta_1)^2}{2\theta_3^2}\right)\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi\theta_4}}\exp\left(-\frac{(y_j - \theta_2)^2}{2\theta_3^2}\right)$$

Log-likelihood:

$$ll(\theta_1, \theta_2, \theta_3, \theta_4) = -\frac{1}{2}\sum_{i=1}^{n}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\ln(\theta_3) - \frac{1}{2}\sum_{j=1}^{m}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{m}\ln(\theta_4) - \sum_{i=1}^{n}\frac{(x_i - \theta_1)^2}{2\theta_3^2} - \sum_{j=1}^{m}\frac{(y_j - \theta_2)^2}{2\theta_4^2}$$

Now we will find the optimal $\theta_1, \theta_2, \theta_3, \theta_4$:

$$\frac{\partial ll(\theta_1, \theta_2, \theta_3, \theta_4)}{\theta_1} = 0$$

$$\frac{1}{2\theta_3^2}\left(\sum_{i=1}^{n}(x_i - \hat{\theta}_1)\right) = 0$$

$$\frac{1}{2\theta_3^2}\left(n\overline{X} - n\hat{\theta}_1\right) = 0$$

$$\hat{\theta}_1 = \overline{X}$$

$$\frac{\partial ll(\theta_1, \theta_2, \theta_3, \theta_4)}{\theta_2} = 0$$

$$\frac{1}{2\theta_4^2}\left(\sum_{j=1}^{m}(y_j - \hat{\theta}_2)\right) = 0$$

$$\frac{1}{2\theta_4^2}\left(n\overline{Y} - n\hat{\theta}_2\right) = 0$$

$$\hat{\theta}_2 = \overline{Y}$$

$$\begin{aligned}
\frac{\partial ll(\theta_1,\theta_2,\theta_3,\theta_4)}{\theta_3} &= 0\\
-\frac{n}{2\hat{\theta_3}}+\frac{\sum_{i=1}^{n}(x_i-\hat{\theta_1})^2}{\hat{\theta_3}^3} &= 0\\
\hat{\theta_3} &= \frac{\sum_{i=1}^{n}(x_i-\hat{\theta_1})^2}{n}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial ll(\theta_1,\theta_2,\theta_3,\theta_4)}{\theta_4} &= 0\\
-\frac{m}{2\hat{\theta_4}}+\frac{\sum_{j=1}^{m}(y_j-\hat{\theta_2})^2}{\hat{\theta_4}^3} &= 0\\
\hat{\theta_4} &= \frac{\sum_{j=1}^{m}(y_j-\hat{\theta_2})^2}{m}
\end{aligned}$$

Replacing to likelihood of denominator:

$$\begin{aligned}
\max_{\theta\in\Theta} l(\theta_1,\theta_2,\theta_3,\theta_4) &= (\frac{1}{2\pi})^{\frac{n}{2}}(\frac{1}{\hat{\theta_3}})^{\frac{n}{2}}(\frac{1}{2\pi})^{\frac{m}{2}}(\frac{1}{\hat{\theta_4}})^{\frac{m}{2}}\exp\left(-\frac{\sum_{i=1}^{n}(x_i-\hat{\theta_1})^2}{2\hat{\theta_3}^2}-\frac{+\sum_{j=1}^{m}(y_j-\hat{\theta_2})^2}{2\hat{\theta_4}^2}\right)\\
&= (\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{1}{\hat{\theta_3}})^{\frac{n}{2}}(\frac{1}{\hat{\theta_4}})^{\frac{m}{2}}\exp\left(-\frac{n}{2}-\frac{m}{2}\right)\\
&= (\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{1}{\hat{\theta_3}})^{\frac{n}{2}}(\frac{1}{\hat{\theta_4}})^{\frac{m}{2}}\exp\left(-\frac{n+m}{2}\right)\\
&= (\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{n}{\sum_{i=1}^{n}(x_i-\hat{\theta_1})^2})^{\frac{n}{2}}(\frac{m}{\sum_{j=1}^{m}(y_j-\hat{\theta_2})^2})^{\frac{m}{2}}\exp\left(-\frac{n+m}{2}\right)
\end{aligned}$$

Now from generalized likelihood ratio, we have:

$$\begin{aligned}
\Lambda^* &= \frac{\max_{\theta\in\Theta_0} l(\theta)}{\max_{\theta\in\Theta} l(\theta)}\\
&= \frac{(\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{n+m}{\sum_{i=1}^{n}(x_i-\mu)^2+\sum_{j=1}^{m}(y_j-\mu)^2})^{\frac{n+m}{2}}\exp\left(-\frac{n+m}{2}\right)}{(\frac{1}{2\pi})^{\frac{n+m}{2}}(\frac{n}{\sum_{i=1}^{n}(x_i-\overline{x})^2})^{\frac{n}{2}}(\frac{m}{\sum_{j=1}^{m}(y_j-\overline{y})^2})^{\frac{m}{2}}\exp\left(-\frac{n+m}{2}\right)}\\
&= \frac{(\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n})^{\frac{n}{2}}(\frac{\sum_{j=1}^{m}(y_j-\overline{y})^2}{m})^{\frac{m}{2}}}{(\frac{\sum_{i=1}^{n}(x_i-\mu)^2+\sum_{j=1}^{m}(y_j-\mu)^2}{n+m})^{\frac{n+m}{2}}}
\end{aligned}$$

Where $\mu=\frac{n\overline{x}+m\overline{y}}{n+m}$
For the second question: Then $\Omega=\{(\theta_1,\theta_2,\theta_3):-\infty<\theta_1<\infty,-\infty<\theta_2<\infty,0<\theta_3<\infty\}$. Let $X_1,X_2,\ldots,X_n$ and $Y_1,Y_2,\ldots,Y_m$ denote independent random samples from these distributions. The hypothesis $H_0:\theta_1=\theta_2$, unspecified, and θ_3 unspecified, is to be tested against all alternatives. Then $\omega=\{(\theta_1,\theta_2,\theta_3):-\infty<\theta_1=\theta_2<\infty,0<\theta_3<\infty\}$. Here $X_1,X_2,\ldots,X_n$, $Y_1,Y_2,\ldots,Y_m$ are $n+m>2$ mutually independent random variables having the likelihood functions

$$L(\omega)=\left(\frac{1}{2\pi\theta_3}\right)^{(n+m)/2}\exp\left\{-\frac{1}{2\theta_3}\left[\sum_{i=1}^{n}(x_i-\theta_1)^2+\sum_{j=1}^{m}(y_j-\theta_1)^2\right]\right\}$$

and

$$L(\Omega)=\left(\frac{1}{2\pi\theta_3}\right)^{(n+m)/2}\exp\left\{-\frac{1}{2\theta_3}\left[\sum_{i=1}^{n}(x_i-\theta_1)^2+\sum_{j=1}^{m}(y_j-\theta_2)^2\right]\right\}$$

If $\frac{\partial \log L(\omega)}{\partial \theta_1}$ and $\frac{\partial \log L(\omega)}{\partial \theta_3}$ are equated to zero, then

$$\sum_{i=1}^{n}(x_i - \theta_1) + \sum_{j=1}^{m}(y_j - \theta_1) = 0$$

$$\frac{1}{\theta_3}\left[\sum_{i=1}^{n}(x_i - \theta_1)^2 + \sum_{j=1}^{m}(y_j - \theta_1)^2\right] = n + m$$

The solutions for θ_1 and θ_3 are, respectively,

$$u = (n+m)^{-1}\left\{\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j\right\}$$

$$w = (n+m)^{-1}\left\{\sum_{i=1}^{n}(x_i - u)^2 + \sum_{j=1}^{m}(y_j - u)^2\right\}$$

Further, u and w maximize $L()$. The maximum is

$$L(\hat{\omega}) = \left(\frac{e^{-1}}{2\pi w}\right)^{(n+m)/2}$$

In a like manner, if

$$\frac{\partial \log L(\Omega)}{\partial \theta_1}, \quad \frac{\partial \log L(\Omega)}{\partial \theta_2}, \quad \frac{\partial \log L(\Omega)}{\partial \theta_3}$$

are equated to zero, then

$$\sum_{i=1}^{n}(x_i - \theta_1) = 0$$

$$\sum_{j=1}^{m}(y_j - \theta_2) = 0$$

$$-(n+m) + \frac{1}{\theta_3}\left[\sum_{i=1}^{n}(x_i - \theta_1)^2 + \sum_{j=1}^{m}(y_j - \theta_2)^2\right] = 0$$

The solutions for θ_1, θ_2, and θ_3 are, respectively,

$$u_1 = n^{-1}\sum_{i=1}^{n} x_i$$

$$u_2 = m^{-1}\sum_{j=1}^{m} y_j$$

$$w' = (n+m)^{-1}\left[\sum_{i=1}^{n}(x_i - u_1)^2 + \sum_{j=1}^{m}(y_j - u_2)^2\right]$$

and, further, u_1, u_2, and w maximize $L(\Omega)$. The maximum is

$$L(\hat{\Omega}) = \left(\frac{e^{-1}}{2\pi w'}\right)^{(n+m)/2}$$

so that

$$\Lambda(x_1, \ldots, x_n, y_1, \ldots, y_m) = \Lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \left(\frac{w'}{w}\right)^{(n+m)/2}$$

The random variable defined by $\Lambda^{2/(n+m)}$ is

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{j=1}^{m}(Y_j-\overline{Y})^2}{\sum_{i=1}^{n}\left\{X_i-\frac{n\overline{X}+m\overline{Y}}{n+m}\right\}^2+\sum_{j=1}^{m}\left\{Y_j-\frac{n\overline{X}+m\overline{Y}}{n+m}\right\}^2}$$

Now

$$\begin{aligned}\sum_{i=1}^{n}\left(X_i-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)^2 &= \sum_{i=1}^{n}\left[(X_i-\overline{X})+\left(\overline{X}-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)\right]^2\\ &= \sum_{i=1}^{n}(X_i-\overline{X})^2+n\left(\overline{X}-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)^2\end{aligned}$$

and

$$\begin{aligned}\sum_{j=1}^{m}\left(Y_j-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)^2 &= \sum_{j=1}^{m}\left[(Y_j-\overline{Y})+\left(\overline{Y}-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)\right]^2\\ &= \sum_{j=1}^{m}(Y_j-\overline{Y})^2+m\left(\overline{Y}-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)^2\end{aligned}$$

But

$$n\left(\overline{X}-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)^2=\frac{m^2n}{(n+m)^2}(\overline{X}-\overline{Y})^2$$

and

$$m\left(\overline{Y}-\frac{n\overline{X}+m\overline{Y}}{n+m}\right)^2=\frac{n^2m}{(n+m)^2}(\overline{X}-\overline{Y})^2$$

Hence the random variable defined by $\Lambda^{2/(n+m)}$ may be written

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{j=1}^{m}(Y_j-\overline{Y})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{j=1}^{m}(Y_j-\overline{Y})^2+\left[\frac{nm}{n+m}\right](\overline{X}-\overline{Y})^2}$$

$$=\frac{1}{1+\frac{\left[\frac{nm}{n+m}\right](\overline{X}-\overline{Y})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{j=1}^{m}(Y_j-\overline{Y})^2}}$$

If the hypothesis $H_0:\theta_1=\theta_2$ is true, the random variable

$$T=\sqrt{\frac{nm}{n+m}}(\overline{X}-\overline{Y})\left\{(n+m-2)^{-1}\left[\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{j=1}^{m}(Y_j-\overline{Y})^2\right]\right\}^{-1/2}$$

has a t-distribution with $n+m-2$ degrees of freedom.Thus the random variable defined by $\Lambda^{2/(n+m)}$ is

$$\frac{n+m-2}{(n+m-2)+T^2}$$

The test of H_0 against all alternatives may then be based on a t-distribution with n+m-2 degrees of freedom.

Question 6: Error Types

One generates a number X from a uniform distribution on the interval $[0,\theta]$. A hypothesis test is performed to test $H_0:\theta=2$ against $H_A:\theta\neq 2$ by rejecting H_0, if $X\leq 0.1$ or $X\geq 1.9$.

1. Compute the probability of a type I error. Illustrate the rejection region on a plot of the uniform probability density function when $\theta=2$.

2. Compute the probability of a type II error if the true value of θ is 2.5. Show the calculations and illustrate how the shift in θ affects the overlap between the rejection and acceptance regions.
3. Suppose we change the decision rule to reject H_0 if $X \leq c_1$ or $X \geq c_2$, where c_1 and c_2 are adjustable thresholds. Derive appropriate values for c_1 and c_2 such that the type I error probability is exactly $\alpha = 0.05$. Explain how the new α affects the power of the test.
4. Discuss the impact of increasing θ on the type II error probability. Provide a mathematical argument to support your answer.
5. Derive an upper bound for the type I error probability α using an appropriate inequality studied in the course (e.g., for distributions with finite variance). Compare this bound to the exact value of α from part 3 and discuss its significance.
6. Identify and use an appropriate inequality to derive an upper bound for the type II error probability β when $\theta = 2.5$. If needed, feel free to search the web for relevant inequalities, and justify your choice. Discuss how such bounds are useful in hypothesis testing, especially when exact calculations are impractical. **(Optional)**

Answer:

1. Since $X \sim U(0, 2)$ under $H_0 : \theta = 2$, the probability density function (PDF) is:

$$f_X(x) = \frac{1}{2}, \quad 0 \leq x \leq 2.$$

The probability of a Type I error, denoted as α, is the probability of rejecting H_0 when it is true. This is given by:

$$P(\text{Type I Error}) = P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

$$P(X \leq 0.1 \text{ or } X \geq 1.9 \mid \theta = 2) = P(X \leq 0.1) + P(X \geq 1.9).$$

$$P(X \leq 0.1) = \frac{0.1}{2}, \quad P(X \geq 1.9) = \frac{0.1}{2}.$$

$$\alpha = \frac{0.1}{2} + \frac{0.1}{2} = \frac{0.2}{2} = 0.1.$$

Illustration of the Rejection Region: The rejection region consists of two shaded areas in the PDF of X, one for $X \leq 0.1$ and another for $X \geq 1.9$. The uniform density is constant at $\frac{1}{2}$, and the total shaded probability sums to 0.1.

2. Since $X \sim U(0, 2.5)$ under $H_A : \theta = 2.5$, the probability density function (PDF) is:

$$f_X(x) = \frac{1}{2.5}, \quad 0 \leq x \leq 2.5.$$

The probability of a Type II error, denoted as β, is the probability of failing to reject H_0 when H_A is true. This is given by:

$$P(\text{Type II Error}) = P(\text{Fail to Reject } H_0 \mid H_A \text{ is true}).$$

$$P(0.1 \leq X \leq 1.9 \mid \theta = 2.5) = P(0.1 \leq X \leq 1.9).$$

$$P(0.1 \leq X \leq 1.9) = \frac{1.9 - 0.1}{2.5} = \frac{1.8}{2.5} = 0.72.$$

$$\beta = 0.72.$$

Illustration of the Shift in θ: The new distribution $U(0, 2.5)$ has a lower density than $U(0, 2)$, and the rejection region remains fixed at $X \leq 0.1$ and $X \geq 1.9$. This results in smaller overlap between the acceptance region and the new distribution, decreasing β.

3. The probability of a Type I error is:

$$P(\text{Type I Error}) = P(X \leq c_1) + P(X \geq c_2).$$

To ensure symmetry in the rejection region, we set:

$$P(X \leq c_1) = P(X \geq c_2) = \frac{\alpha}{2} = 0.025.$$

Since X follows a uniform distribution:

$$P(X \leq c_1) = \frac{c_1}{2}, \quad P(X \geq c_2) = \frac{2 - c_2}{2}.$$

$$c_1 = 0.05, \quad c_2 = 1.95.$$

Effect on Power: A lower α (stricter criteria for rejecting the null hypothesis) usually leads to higher β, hence a lower power.

4. As θ increases, the uniform distribution becomes wider, causing the density of X to decrease (since $f_X(x) = \frac{1}{\theta}$). Consequently, the overlap between the rejection region and the density of X under H_A increases, leading to a higher probability of rejecting H_0. Thus, the type II error probability β decreases as θ increases.

5. Chebyshevs Inequality:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

For $X \sim \text{Uniform}(0, 2)$:

$$\mu = 1, \quad \sigma^2 = \frac{(2-0)^2}{12} = \frac{1}{3}, \quad \sigma = \sqrt{\frac{1}{3}} \approx 0.577.$$

For $|X - \mu| \geq 0.9$ ($X \leq 0.1$ or $X \geq 1.9$):

$$k = \frac{0.9}{\sigma} = \frac{0.9}{0.577} \approx 1.56.$$

$$P(|X - \mu| \geq 0.9) \leq \frac{1}{k^2} = \frac{1}{(1.56)^2} \approx 0.411.$$

Thus, $\alpha \leq 0.411$.

Comparison:

- Exact Value: $\alpha = 0.1$.
- Chebyshev Bound: $\alpha \leq 0.411$.
- While the bound is looser than the exact value, its still a reasonable estimate and useful when exact computation isnt feasible.

6. In this case, inequalities do not provide a good upper bound, as they significantly overestimates the actual Type II error probability. For the bound to be useful, the number of samples should be large. When dealing with a single sample, the bound is too loose to offer meaningful insight.

Question 7: Warm Up t-test (Optional)

Given a two-sample t-test to determine if there is a significant difference between the means of two populations:

$$\text{Null Hypothesis: } H_0 : \mu_1 = \mu_2, \quad \text{Alternative Hypothesis: } H_1 : \mu_1 \neq \mu_2$$

Follow the instructions below to prove that H_0 is rejected if and only if the confidence interval for the difference between the sample means does not include zero.

(a) Derive the test statistic assuming $H_0 : \mu_1 = \mu_2$, and identify the conditions under which H_0 is rejected.

(b) Derive the confidence interval formula for the difference $\mu_1 - \mu_2$, and explain the role of the critical value in determining the interval's endpoints.

(c) Prove that the confidence interval excludes zero if and only if the t-test rejects H_0. Provide a clear explanation of the equivalence between these two approaches and interpret the result in the context of hypothesis testing.

Answer:

In two-sample t-test, we have t-distribution as follows:

$$\frac{\overline{x}_1 - \overline{x}_2 - \delta}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

where

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} \tag{7.1}$$

We can obtain confidence interval for δ from t-distribution:

$$P\left(-t_{\frac{\alpha}{2},n_1+n_2-2} \le \frac{\overline{x}_1 - \overline{x}_2 - \delta}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \le t_{\frac{\alpha}{2},n_1+n_2-2}\right) = 1 - \alpha$$

$$P\left(-t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \overline{x}_1 - \overline{x}_2 - \delta \le t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) = 1 - \alpha$$

$$P\left(-t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} + \overline{x}_1 - \overline{x}_2 \le \delta \le t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} + \overline{x}_1 - \overline{x}_2\right)$$

So the confidence interval for δ or $\mu_{x_1} - \mu_{x_2}$ is:

$$(\overline{x}_1 - \overline{x}_2) \pm t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Now if $\mu_{x_1} - \mu_{x_2}$ does not include zero, two different conditions can be case:

- $(\overline{x}_1 - \overline{x}_2) + t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < 0$
- $(\overline{x}_1 - \overline{x}_2) - t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} > 0$

We know that $\mu_{x_1} - \mu_{x_2}$ is less than zero:

$$\begin{aligned}
\mu_{x_1} - \mu_{x_2} &\le (\overline{x}_1 - \overline{x}_2) + t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < 0 \\
\mu_{x_1} - \mu_{x_2} &< 0 \\
\mu_{x_1} &< \mu_{x_2}
\end{aligned}$$

So if the first condition holds, $\mu_{x_1} < \mu_{x_2}$ and the null hypothesis will be rejected.
In the second condition $\mu_{x_1} - \mu_{x_2}$ is bigger than zero:

$$\begin{aligned}
0 < (\overline{x}_1 - \overline{x}_2) + t_{\frac{\alpha}{2},n_1+n_2-2} * s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} &< \mu_{x_1} - \mu_{x_2} \\
0 &< \mu_{x_1} - \mu_{x_2} \\
\mu_{x_2} &< \mu_{x_1}
\end{aligned}$$

So if second condition holds, $\mu_{x_2} < \mu_{x_1}$ and the null hypothesis will be rejected.

Question 8: t-test

We perform a t-test for the null hypothesis $H_0 : \mu = 10$ at significance level $\alpha = 0.05$ by means of a dataset consisting of $n = 16$ elements with sample mean 11 and sample variance 4.

1. Should we reject the null hypothesis in favor of $H_A : \mu \neq 10$?
2. What if we test against $H_A : \mu > 10$?
3. Explain why the results of the two tests in parts 1 and 2 differ, even though the null hypothesis is the same. Discuss the relationship between the critical regions for the one-tailed and two-tailed tests by providing illustrations of the t-distribution.

Answer:

1. Two-Tailed Test ($H_A : \mu \neq 10$)

   The t-statistic is given by:

   $$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{11 - 10}{2/\sqrt{16}} = 2.$$

   With $n = 16$, the degrees of freedom for this test is:

   $$df = n - 1 = 16 - 1 = 15.$$

   We calculate the two-sided p-value:

   $$p = P(|t| > 2 \mid H_0) = 2 \times (1 - \text{pt}(2, 15)) = 0.063945.$$

   Since $p > \alpha = 0.05$, we fail to reject the null hypothesis.

   Alternatively, we could have done the problem in terms of rejection regions:

   $$(-\infty, 10 - t_{15,0.025}\frac{s}{\sqrt{n}}) \cup [10 + t_{15,0.025}\frac{s}{\sqrt{n}}, \infty) = (-\infty, 8.93] \cup [11.07, \infty).$$

   Since $\bar{x} = 11$ lies outside the rejection region, we again fail to reject the null hypothesis.

2. One-Tailed Test ($H_A : \mu > 10$)

   We have the same t-statistic:

   $$t = 2.$$

   We calculate the one-sided p-value:

   $$p = P(t > 2 \mid H_0) = 1 - \text{pt}(2, 15) = 0.031973.$$

   Since $p < \alpha = 0.05$, we reject the null hypothesis in favor of the alternative.

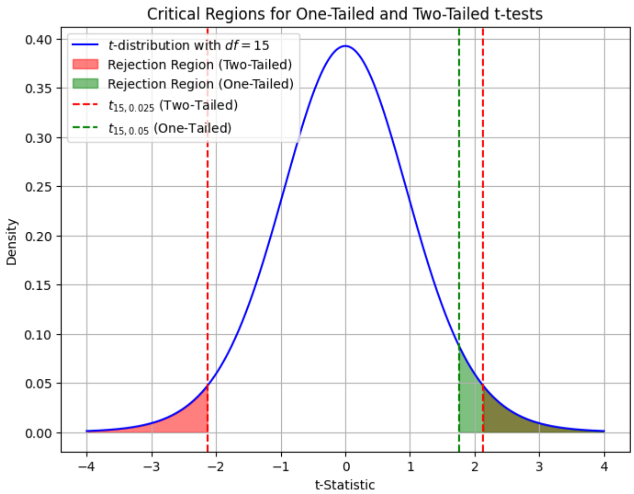
   Again looking at rejection regions:

   $$[10 + t_{15,0.05}\frac{s}{\sqrt{n}}, \infty) = [10.876, \infty).$$

   Since $\bar{x} = 11$ lies inside the rejection region, we reject the null hypothesis in favor of the alternative.

3. The two-tailed test splits the significance level ($\alpha = 0.05$) between the upper and lower tails, resulting in higher critical values and stricter rejection criteria. In contrast, the one-tailed test places all $\alpha = 0.05$ in the upper tail, making rejection easier in that direction. Since the two-tailed test checks for deviations in both directions ($\mu \neq 10$) while the one-tailed test focuses only on increases ($\mu > 10$), the stricter two-tailed test fails to reject H_0, while the one-tailed test does.

   Here is an illustration of the critical regions for both tests:



Question 9: An Apple a Day... (Programming Question)

An agricultural researcher is studying the average number of apples consumed per day in a population. They have conducted preliminary research and found that the populations mean daily consumption is approximately 2 apples, with a standard deviation of 0.5 apples. To explore how confidence intervals behave under repeated sampling, the researcher designs an experiment as follows:

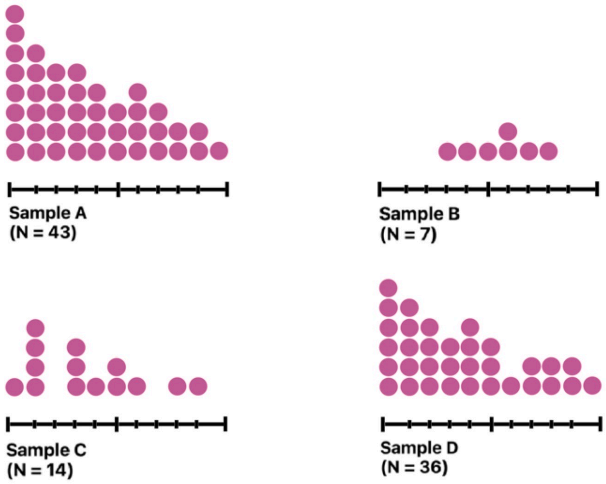
1. **Define the Population and Sample:** Assume the population distribution of daily apple consumption is normal with a mean (μ) of 2 apples per day and a standard deviation (σ) of 0.5 apples. Imagine that the researcher draws a random sample of 12 individuals from this population.
2. **Estimate the Confidence Interval (CI):** Calculate a 95% confidence interval for the sample mean of daily apple consumption for each sample of 12 individuals. Repeat this sampling process 100 times, each time recalculating the 95% confidence interval.
3. **Analyze the Confidence Interval Coverage:** Determine the percentage of times (out of 100) that the calculated confidence interval actually contains the true population mean of 2 apples per day.
4. **Compare Different Methods for CI Estimation:**
   (a) Assume two cases:
      i. The population standard deviation (σ) is known.
      ii. The population standard deviation is unknown.
   (b) For each case, calculate the 95% confidence intervals using the following approaches:
      i. Known population variance (σ): Use the Z-distribution to construct the confidence interval.
      ii. Unknown population variance: Use both the Z-distribution (with sample standard deviation s) and the T-distribution (with sample standard deviation s).
   (c) Compare the results. Observe and comment on whether and why the confidence intervals constructed with the Z-distribution, using s instead of σ, are narrower than expected.

Answer:
You can find the explanation of this question in this link.

Question 10: Short Answers - Big Concepts

Instructions: For each scenario below, analyze the situation, consider the statistical goals, and select the best testing approach. Use your understanding of hypothesis testing, sample conditions, and t-test types to choose the most appropriate answers.

1. Researchers have gathered four samples with varying characteristics. Assess each sample and determine which ones meet the conditions required for performing a t-test.



2. A restaurant advertises that its burgers weigh 250 grams. To verify this claim, the manager randomly selects 16 burgers, finding that their weights are roughly symmetric with no outliers. The manager wants to use these sample data to conduct a t-test about the mean. As a manager, based on the business goals and implications of customer satisfaction, would you test a simple or a composite hypothesis?

3. A fitness scientist wants to test the effectiveness of a new program designed to improve flexibility in senior citizens. Participants were evaluated on flexibility using a standard scale both before and after completing the program. What kind of test, null, and alternative hypothesis are most appropriate for this scenario?

Answer:

1. Since the number of samples in A and D are above 30, and sample B is roughly symmetrical with no outliers, you can perform t-tests on samples A, B, and D as they satisfy the normal condition.

2. To ensure the burgers are big enough, the manager could test a one-sided composite hypothesis of $H_0 : \mu \leq 250$. However, failing to reject this hypothesis could mean the burgers are unnecessarily big. So for a manager the best choice would be to test a two-sided simple hypothesis of $H_A : \mu = 250$ to ensure satisfaction on both ends.

3. In this scenario, a paired t-test is appropriate since each participant is measured twicebefore and after the programcreating paired data. The null hypothesis states that the program has no effect on flexibility:

$$H_0 : \mu_{\text{before}} = \mu_{\text{after}}$$

which simplifies to

$$H_0 : \mu_d = 0$$

where μ_d represents the mean difference in flexibility scores.

For the alternative hypothesis, as the scientist expects an improvement in flexibility, a one-tailed test is appropriate:

$$H_A : \mu_d > 0$$

Question 11: Speed Camera Error Rate

Suppose three speed cameras are set up along a stretch of road to catch people driving over the speed limit of 40 miles per hour. Each speed camera is known to have a normal measurement error modeled on $N(0, 5^2)$. For a passing car, let $\bar{x}$ be the average of the three readings. Our default assumption for a car is that it is not speeding.

1. Find the distribution of $\bar{x}$. Describe the above story in the context of hypothesis testing (write down H_0 and H_A).

2. The police would like to set a threshold on $\bar{x}$ for issuing tickets so that no more than 4% of law-abiding, non-speeders are mistakenly given tickets. This means that they set the threshold conservatively so that no more than 4% of drivers going exactly 40 mph get a ticket.

   (a) Use the hypothesis testing description in part 1 to help determine what threshold they should set.

   (b) Illustrate the distribution and the rejection area.

   (c) If possible, calculate the probability that a person getting a ticket was not speeding.

   (d) Suppose word gets out about the speed trap, and no one speeds along it anymore. What percentage of tickets are given in error?

3. What is the power of this test with the alternative hypothesis that the car is traveling at 45 miles per hour? How many cameras are needed to achieve a power of 0.9 with $\alpha = 0.04$? (Optional)

Answer:

1. Let μ be the actual speed of a given driver.

$$x_i \sim N(\mu, 5^2) \implies \bar{x} \sim N(\mu, 5^2/3).$$

   The hypotheses are:

$$H_0 : \text{the driver is not speeding}, \mu \leq 40, \quad H_A : \text{the driver is speeding}, \mu > 40.$$

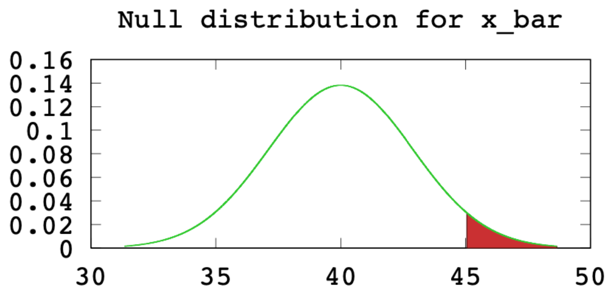
   Note: H_0 is composite, but we can do all our computations with the most extreme value $H_0 = 40$ because the one-sided rejection region will have its largest significance level when $H_0 = 40$.

2. (a) Giving a ticket to a non-speeder is a Type I error (rejecting H_0 when it is true).
   The null distribution is $\bar{x} \sim N(40, 5^2/3)$. The critical value is:

$$c_{0.04} = 40 + z_{0.04}\frac{5}{\sqrt{3}} = 45.054$$

   That is, they should issue a ticket if the average of the three guns is more than 45.054.

   (b) This is a plot of the null distribution $N(40, 5^2/3)$. The rejection region with probability 0.04 is shown.



   (c) We don't know this probability. To find it out would require a prior probability that a random driver is speeding.

   (d) If no one is speeding, then 100% of tickets are given in error.

3. Power $= P(\text{rejection}|H_A)$. So to find the power, we first must find the rejection region. For $n = 3$, this was found to be: rejection region $= [45.054, \infty)$. So,

$$\text{power} = P(\text{rejection} \mid \mu = 45) = 1 - \phi\left(\frac{45.054 - 45}{5/\sqrt{3}}\right) = 0.493.$$

   With n cameras, let's write $\bar{x}_n$ for the sample mean. The null distribution is

$$\bar{x}_n \sim N(40, 5^2/n).$$

We will need to write everything in terms of standard normal values:

$$c_{0.04} = 40 + z_{0.04}\frac{5}{\sqrt{n}},$$

$$\text{power} = P\left(\bar{x} \geq c_{0.04} \mid \mu = 45\right) = 0.9.$$

$$P\left(\frac{\bar{x}-45}{5/\sqrt{n}} \geq \frac{c_{0.04}-45}{5/\sqrt{n}}\right) = 0.9 \implies P\left(z \geq -\frac{5}{5/\sqrt{n}} + z_{0.04}\right) = 0.9.$$

$$z_{0.9} = -\frac{5}{5/\sqrt{n}} + z_{0.04} \implies n = \left(z_{0.04} - z_{0.9}\right)^2 = \left(1.2816 - (-1.2816)\right)^2 = 9.1945.$$

Rounding up, we find $n = 10$.

Question 12: Statistically Delicious Coffee

Imagine you're a data scientist hired by a coffee shop chain that wants to optimize customer satisfaction by improving coffee quality and drink preparation speed. The coffee shop wants to test two things:

- Whether there is a difference in preparation times between the coffee shop branches A and B.
- Whether a new coffee recipe B tastes better than the original coffee recipe A among customers.

1. You design an experiment to collect data for both objectives. Based on the nature of each comparison, what statistical tests would be appropriate to determine if there are significant differences for each case?
2. The coffee shop suspects Branch A might be faster at preparing drinks than Branch B. You measure the preparation time (in minutes) of a random sample of 12 orders from each branch. Based on your answer in part 1, test if there is a statistically significant difference in preparation times between Branch A and Branch B. (Assume equal variances, and let $\alpha = 0.05$). The data is shown in table 1.

Order	Branch A	Branch B
1	6	7
2	7	6
3	7	7
4	8	8
5	6	9
6	9	7
7	5	10
8	7	9
9	8	9
10	7	9
11	5	7
12	7	9

Table 1: Preparation times for Branch A and Branch B.

3. The coffee shop also wants to test whether the new coffee recipe (Recipe B) tastes better than the original recipe (Recipe A). They conduct a taste test with 12 customers, each of whom rates both Recipe A and Recipe B on a scale of 1-10. Based on your answer in part 1, test if there is a statistically significant improvement in taste ratings with Recipe B compared to Recipe A. (Let $\alpha = 0.05$). The data is shown in table 2.

Answer:

Customer	Recipe A	Recipe B
1	8	9
2	7	6
3	7	7
4	6	8
5	9	7
6	6	7
7	5	9
8	8	9
9	8	10
10	7	9
11	5	9
12	7	6

Table 2: Taste test ratings for Recipe A and Recipe B.

1. To compare preparation times between two independent samples, we use a two-sample t-test. To analyze taste preference, we would use a paired t-test.
2. Hypotheses:

$$H_0 : \mu_A = \mu_B$$
$$H_1 : \mu_A \neq \mu_B \quad (or \quad \mu_A \leq \mu_B)$$

Sample statistics:

$$\bar{X}_A = 6.83, \quad s_A = 1.19, \quad n_A = 12$$
$$\bar{X}_B = 8.08, \quad s_B = 1.24, \quad n_B = 12$$

Test Statistics:

$$s_p = 1.22$$
$$t = \frac{\bar{X}_A - \bar{X}_B}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = -2.52$$
$$df = n_A + n_B - 2 = 22$$
$$t_{\text{critical}} = \pm 2.07$$
$$pvalue = 0.0197$$

Since $|t| = 2.52 > 2.07$ and $p < 0.05$, we reject H_0. So there is a significant difference in preparation times. For the one-tailed test we would also reject H_0 since $t = -2.52 < -1.72$ and $p = 0.0098 < 0.05$.

3. Hypothesis:

$$H_0 : \mu_d = 0$$
$$H_1 : \mu_d > 0$$

Sample statistics:

$$s_d = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n-1}} = 1.881$$

Test Statistics:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = 1.995$$

$$t_{crit} = 1.796$$

Since $t = 1.995 > 1.796$, we reject the null hypothesis. So there is statistically significant evidence ($p < 0.05$) that Recipe B tastes better than Recipe A.