3. \( X_1, X_2, \ldots, X_n \) are i.i.d Uniform \((0, \theta)\)

\[ X(n) = \max \left( X_1, X_2, \ldots, X_n \right). \]

(a) Show that \( P \left( \frac{X(n)}{\theta} \leq x \right) = x^n, \quad 0 \leq x \leq 1. \)

The distribution function of \( X_1 \) say \( F \) is:

\[ F(y) = \frac{y}{\theta}, \quad 0 \leq y \leq \theta. \]

Now:

\[ P \left( \frac{X(n)}{\theta} \leq x \right) = P \left( X(n) \leq \theta x \right) \]

= \[ P \left( X_1 \leq \theta x, X_2 \leq \theta x, \ldots, X_n \leq \theta x \right) \]

= \[ P \left( X_1 \leq \theta x \right) P \left( X_2 \leq \theta x \right) \ldots P \left( X_n \leq \theta x \right) \]

= \[ \frac{\theta x}{\theta} \times \frac{\theta x}{\theta} \times \cdots \times \frac{\theta x}{\theta} = x^n = G(x) \]

This shows that \( \frac{X(n)}{\theta} \) is indeed a pivot.

To find a 90% C.I. for \( \theta \):

Here \( \alpha = .1 \). Let \( q_{\alpha} \) denote the \( 100\alpha \)th quantile of the pivotal distribution \( G \).

Then \( G(q) = \beta \) i.e. \( q^n = \beta \) \Rightarrow \( q = \beta^{1/n} \)
We have:
\[ P \left( \frac{q x/2}{\theta} \leq \frac{X(n)}{\theta} \leq q (1 - x)^{1/n} \right) = 1 - \alpha \]

i.e.
\[ P \left( \left( \frac{x}{2} \right)^{1/n} \leq \frac{X(n)}{\theta} \leq \left( 1 - \frac{x}{2} \right)^{1/n} \right) = 1 - \alpha \]

\[ \Rightarrow P \left[ \frac{X(n)}{(1 - x)^{1/n}} \leq 0 \leq \frac{X(n)}{(\frac{x}{2})^{1/n}} \right] = 1 - \alpha \]

Thus a 90% C.I for \( \theta \) is:
\[
\left[ \frac{X(n)}{(0.95)^{1/n}}, \frac{X(n)}{(0.05)^{1/n}} \right]
\]

To test \( \theta = 3 \) at level 0.1.

The pivot computed under \( H_0 : \theta = 3 \) is:
\[ \frac{X(n)}{3} \]

Our level \( \alpha = 0.10 \) test is:
\[ \phi(x) = 0 \text{ if } (0.95)^{1/n} \leq \frac{X(n)}{3} \leq (0.95)^{1/n} \]
\[ = 1 \text{ otherwise } \]

Check that this indeed has level (Probability of Type 1 error) 0.10.
4. \( X_1, X_2, \ldots, X_n \) i.i.d. \( \text{Exp}(\theta) \)

(a) To show that \( \sum_{i=1}^{n} 20X_i \sim \chi^2_{2n} \).

Note that \( 20X_1, 20X_2, \ldots, 20X_n \) are i.i.d random variables.

Now: \( P(20X_1 \leq y) = P\left(X_1 \leq \frac{y}{20}\right) \)

\[ = 1 - e^{-\theta \cdot \frac{y}{20}} \] (\( \sim X_1 \) follows \( \text{Exp}(\theta) \))

\[ = 1 - e^{-\frac{y}{2}} \]

\[
\therefore 20X_1 \sim \text{Exp}\left(\frac{1}{2}\right) \text{ i.e. Gamma}\left(1, \frac{1}{2}\right).
\]

Thus \( 20X_1, 20X_2, \ldots, 20X_n \) are i.i.d \( \text{Gamma}\left(1, \frac{1}{2}\right) \).

\[
\therefore \sum_{i=1}^{n} 20X_i \sim \text{Gamma}\left(n, \frac{1}{2}\right), \text{ by the reproductive property of the Gamma distribution.}
\]

But \( \text{Gamma}\left(n, \frac{1}{2}\right) = \text{Gamma}\left(\frac{2n}{2}, \frac{1}{2}\right) = \chi^2_{2n} \).

Let \( q_{\alpha}(2n) \) denote the \( \alpha \)-th quantile of the \( \chi^2_{2n} \) distribution.
Then, \( P \left[ \frac{q(\frac{2n}{2})}{\sum 2x_i} \leq 0 \leq \frac{q(\frac{2n}{1 - \frac{x}{n}})}{\sum 2x_i} \right] = 1 - \alpha \)

i.e., \( P \left[ \frac{q(\frac{2n}{2})}{\sum 2x_i} \leq 0 \leq \frac{q(\frac{2n}{1 - \frac{x}{n}})}{\sum 2x_i} \right] = 1 - \alpha \)

Thus a level \( 1 - \alpha \) C.I. is:

\[
\left[ \frac{q(\frac{2n}{2})}{2n \bar{x}}, \frac{q(\frac{2n}{1 - \frac{x}{n}})}{2n \bar{x}} \right]
\]

(b) Note that the \( z_i \)'s are i.i.d. Bernoulli(p)

where \( p = P(z_i = 1) = P(x_i > 3) = e^{-3\theta} \).

Note that \( \hat{p} = \frac{1}{n} \sum z_i = \bar{z} \) is a reasonable estimate of \( e^{-3\theta} \).

We can solve: \( \hat{p} = e^{-3\theta} \) to get a reasonable estimate of \( \theta \).

We get: \(- \log \hat{p} = 3\theta\)

\( \therefore \hat{\theta} = -\frac{1}{3} \log \bar{z} \)

So: \( \hat{\theta} = -\frac{1}{3} \log \bar{z} \).
Constructing an approximate level $\alpha = 0.05$ test for testing $H_0: \theta = 0.5$ based on $z_1, \ldots, z_n$.

By the Central Limit Theorem, we get:

$$\frac{\sqrt{n} (\bar{z} - \theta)}{\sqrt{\theta (1 - \theta)}} \sim N(0,1) \text{ approximately,}$$

for large $n$, and is therefore an approximate pivot. Here $\theta = e^{-30}$.

Thus, we get:

$$P(0) \left[ - 3 \alpha/2 \leq \frac{\sqrt{n} (\bar{z} - \theta)}{\sqrt{\theta (1 - \theta)}} \leq 3 \alpha/2 \right] \approx 1 - \alpha$$

for large $n$.

Under $H_0: \theta = \theta_0 = e^{-3 \times 0.5} = e^{-1.5}$ and the value of the approximate pivot is $\frac{\sqrt{n} (\bar{z} - \theta_0)}{\sqrt{\theta_0 (1 - \theta_0)}}$.

Our approximate level $\alpha = 0.05$ test is:

$$\phi(x) = 0 \text{ if } -3 \alpha/2 \leq \frac{\sqrt{n} (\bar{z} - \theta_0)}{\sqrt{\theta_0 (1 - \theta_0)}} \leq 3 \alpha/2$$

$$= 1 \text{ otherwise.}$$
Here $X_{1/2} = 8.025$ can be read off from normal tables.

5. Model: $Y_i = \beta X_i + \epsilon_i$

$\epsilon_i$'s are i.i.d. $N(0, \sigma^2)$ and $X_i$'s are fixed.

(a) Minimize $\sum_{i=1}^{n} (Y_i - \beta X_i)^2$ with respect to $\beta$.

Set $\frac{d}{d\beta} \left[ \sum_{i=1}^{n} (Y_i - \beta X_i)^2 \right] = 0$, to get:

$$\sum_{i=1}^{n} 2 (Y_i - \beta X_i)(-X_i) = 0$$

i.e. $\sum_{i=1}^{n} (X_i Y_i - \beta X_i^2) = 0$

i.e. $\beta \sum X_i^2 = \sum X_i Y_i$

$\therefore \beta = \frac{\sum X_i Y_i}{\sum X_i^2}$

Thus: $\hat{\beta} = \frac{\sum X_i Y_i}{\sum X_i^2}$

(b) Since $\hat{\beta}$ is a linear combination of the $Y_i$'s which are normal, it is also normally distributed.
\[ E(\beta) = E \left[ \frac{\sum X_i Y_i}{\sum X_i^2} \right] \]

\[ = \frac{1}{\sum X_i^2} \sum_{i=1}^{n} X_i \cdot E(Y_i) \]

\[ = \frac{1}{\sum X_i^2} \sum_{i=1}^{n} X_i \cdot \beta \cdot X_i = \beta \frac{1}{\sum X_i^2} \sum X_i^2 \]

\[ = \beta \]

(We have used the fact that \( Y_i \sim N(\alpha X_i \beta, \sigma^2) \), so \( E Y_i = X_i \beta \)).

\[ \text{VAR}(\beta) = \text{VAR} \left[ \frac{\sum X_i Y_i}{\sum X_i^2} \right] \]

\[ = \frac{1}{(\sum X_i^2)^2} \sum_{i=1}^{n} \text{VAR} \left( \frac{X_i Y_i}{\sum X_i^2} \right) \]

\[ = \frac{1}{(\sum X_i^2)^2} \sum_{i=1}^{n} \text{VAR}(X_i Y_i) \]

\[ = \frac{1}{(\sum X_i^2)^2} \sum_{i=1}^{n} X_i^2 \text{VAR}(Y_i) = \frac{\sigma^2}{\sum X_i^2} \]
We assume $\sigma^2$ is known.

\[ \hat{\beta} \sim N(\beta, \frac{\sigma^2}{\sum x_i^2}) \]

\[ \frac{\hat{\beta} - \beta}{\sigma / \sqrt{\sum x_i^2}} \sim N(0, 1) \]

\[ \Pr \left[ -3\alpha/2 \leq \frac{\hat{\beta} - \beta}{\sigma / \sqrt{\sum x_i^2}} \leq 3\alpha/2 \right] = 1 - \alpha \]

\[ \Pr \left[ \hat{\beta} - \frac{3\alpha/2}{\sqrt{\sum x_i^2}} \leq \beta \leq \hat{\beta} + \frac{3\alpha/2}{\sqrt{\sum x_i^2}} \right] = 1 - \alpha \]

Level 1 - $\alpha$ c.i. for $\beta$ is given by:

\[ \left[ \hat{\beta} - \frac{3\alpha/2}{\sqrt{\sum x_i^2}}, \hat{\beta} + \frac{3\alpha/2}{\sqrt{\sum x_i^2}} \right] \]

1. We follow the hint. It suffices to show that $\hat{\beta}$ is independent of $y_i - \hat{\beta}x_i$.

Since $\hat{\beta}$ and $y_i - \hat{\beta}x_i$ are both normal, it suffices to show that the covariance between $\hat{\beta}$ and $y_i - \hat{\beta}x_i$ is equal to 0.
\[ \text{Cov} (\widehat{\beta}, Y_i - \widehat{\beta}x_i) \]
\[ = \text{Cov} (\widehat{\beta}, Y_i) - \text{Cov} (\widehat{\beta}, \widehat{\beta}x_i) \]
\[ = \text{Cov} (\widehat{\beta}, Y_i) - x_i \text{ Cov} (\widehat{\beta}, \widehat{\beta}) \]
\[ = \text{Cov} (\widehat{\beta}, Y_i) - x_i \text{ Var} (\widehat{\beta}) \]

Now, \( \text{Cov} (\widehat{\beta}, Y_i) = \text{Cov} \left( \frac{1}{\sum x_i^2} \sum x_j y_j, Y_i \right) \)
\[ = \frac{1}{\sum x_j^2} \text{Cov} \left( \sum x_j y_j, Y_i \right) \]
\[ = \frac{1}{\sum x_j^2} \sum_{j=1}^{n} \text{Cov} (x_j y_j, Y_i) \]
\[ = \frac{1}{\sum x_j^2} \text{Cov} (x_i y_i, Y_i) \text{ (since for } j \neq i, \text{ Y}_j \text{ and } Y_i \text{ are independent and so } \text{Cov} (x_j y_j, Y_i) = 0) \]
\[ = \frac{1}{\sum x_j^2} x_i \text{ Var} (Y_i) \]
\[ = \frac{s^2 x_i}{\sum x_j^2} = x_i \text{ Var} (\widehat{\beta}) \]

Therefore, \( \text{Cov} (\widehat{\beta}, Y_i - \widehat{\beta}x_i) = 0 \)
1. (a) Here \( H_0 : \mu_x = \mu_y \) i.e. \( \mu_x - \mu_y = 0 \)

Refer to the solution of Problem 2(i).

Our level \( \alpha \) test is given by:

\[
\phi(x) = 0 \quad \text{if the hypothesized value of } \mu_x - \mu_y \text{ under } H_0, \text{ which is } 0, \text{ lies in the level } 1 - \alpha \text{ confidence interval }
\]

i.e. \( 0 \in \left[ x - \bar{y} - \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \cdot z_{\alpha/2}, \right. \\
\left. x - \bar{y} + \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \cdot z_{\alpha/2} \right] 
\]

and \( \phi(x) = 1 \) otherwise.

1. (b) Here \( H_0 : \frac{\sigma_x^2}{\sigma_y^2} = 1 \).

Our level \( \alpha \) test is then seen to be given by:

\[
\phi(x) = 0 \quad \text{if } q_{\beta_1} \leq \frac{s_x^2}{s_y^2} \leq q_{1-\beta_2}
\]

\[
= 1 \quad \text{otherwise}
\]

(Refer to Problem 2(ii) in Homework 3)
2. \( X \) is a single observation from a \( N(\mu, 1) \) distribution.

We want to test: \( H_0: \mu = 0 \).

Test: \( \phi(x) = 1 \) if \( |x| > 3\sigma/2 \)

\[ = 0 \text{ otherwise} \]

(a) This is a reasonable test. It rejects for large absolute values of \( |x| \), which is a sensible thing to do, since under \( H_0 \), \( X \) comes from a \( N(0, 1) \) distribution and values of \( X \) with small absolute magnitude are more likely.

Level of the test:

\[ = \text{E}_{\mu = 0} \left[ \phi(x) \right] \]

\[ = P_{\mu = 0} \left[ |x| > 3\sigma/2 \right] \]

\[ = \alpha \]

Since \( X \sim N(0, 1) \) under \( H_0 \).
(b) \[ \beta(\mu) = P(\mu \mid |X| > 3x/2) \]

\[ = P(\mu \mid X > 3x/2) + P(\mu \mid X < -3x/2) \]

\[ = P(\mu \mid X - \mu > 3x/2 - \mu) + P(\mu \mid X - \mu < -3x/2 - \mu) \]

When \( \mu \) is the true mean, \( X - \mu \sim N(0, 1) \)

Thus:

\[ \beta(\mu) = P(\frac{z}{\sigma} > 3x/2 - \mu) + P(\frac{z}{\sigma} < -3x/2 - \mu) \]

where \( z \sim N(0, 1) \). Let \( \Phi \) denote the cumulative distribution function of \( N(0, 1) \). Then:

\[ \beta(\mu) = 1 - \Phi (3x/2 - \mu) + \Phi (-3x/2 - \mu) \]

\[ = \Phi (\mu - 3x/2) + \Phi (-3x/2 - \mu) \]

\[ \beta'(\mu) = \Phi'(\mu - 3x/2) + \Phi' (-3x/2 - \mu) \]

Setting \( \beta'(\mu) = 0 \) gives:

\[ \Phi'(\mu - 3x/2) = \Phi' (-3x/2 - \mu) \]

Since \( \Phi' \) is the normal density which is symmetric about 0.
\[ \mu = 3 \alpha/2 = \pm (-3 \alpha/2 - \mu) \]

The possibility \( \mu = 3 \alpha/2 = -(-3 \alpha/2 - \mu) \) is ruled out since this implies \( 3 \alpha/2 = 0 \), which is false.

So \( \mu = 3 \alpha/2 = -3 \alpha/2 - \mu \Rightarrow \mu = 0 \).

It can be verified that \( \beta''(\mu) > 0 \) so \( \mu = 0 \) indeed gives a minimum.

\[ \beta(\Delta) = \phi(\Delta - 3 \alpha/2) + \Phi(-3 \alpha/2 - \Delta) \]

\[ \beta(-\Delta) = \phi(-\Delta - 3 \alpha/2) + \Phi(-3 \alpha/2 + \Delta) \]

Which is the same expression as that for \( \beta(\Delta) \).

Schematically, the power function looks the same as the one drawn on Page 18 of "Statistics Notes - I".