Shashank. P. 200010048

(1) (a) Let $X \sim \text{Uniform}(0,1)$

Here I considered 500 X_i's.

Let $Y = \dfrac{(\sum X_i) - 500 \cdot \frac{1}{2}}{\sqrt{n}\sqrt{Var[X_i]}}$

$Var[X_i] = \dfrac{(b-a)^2}{12} = \dfrac{1}{12}$

$\therefore \quad Y = \sqrt{\dfrac{12}{500}}\left(\sum X_i - 250\right)$

Here Y is $\sim$ Normal (0,1) by CLT.
[File Q1_a.txt]
Using Y I generated 10000 data points.

(b) Let $Z = Y_1^2 + Y_2^2 + Y_3^2 + Y_4^2$

where Y is same as above.

$Z \sim \chi^2(df = 4)$

Using this Z value I generated 10000 data points [Q1_b.txt].

classmate

Date

Page

2. (a) Let $Y = X + \omega$ where $\omega \sim \mathcal{N}(0,1)$

$X \sim$ Uniform (a, b) and Y is observed.

Consider the noise to be independent.

$E[Y] = E[X]$ — ① & $Var[Y] = Var[X] + Var[W]$
$= Var[X] + 1$ — ②

Let v = sample variance of Y.

m = sample mean of Y

We know both v & m are unbiased and consistent.

Using ① & ② & $E[X] = \frac{b+a}{2}$ & $Var[X] = \frac{(b-a)^2}{12}$

$\therefore m = \frac{b+a}{2}$ & $v = \frac{(b-a)^2}{12} + 1$

$\Rightarrow b+a = 2m$ & $2\sqrt{3(v-1)} = b-a$

$\Rightarrow \boxed{b = m + \sqrt{3(v-1)}}$ $\boxed{a = m - \sqrt{3(v-1)}}$

$\boxed{b = 8}$ $\boxed{a = 2}$

(b) $\boxed{\text{Assuming that noise} = 0}$

The given data represents uniform random variable $X \sim (a, b)$.

Let $\vec{X} = (x_1, x_2, \ldots\ldots x_n)$ given data

Let $\hat{b} = \max(x_1, x_2, \ldots x_n) = \vec{X}_{max}$

$\hat{a} = \min(x_1, x_2, \ldots x_n) = \vec{X}_{min}$

$$P(\vec{X}_{max} < x) = P(X_i < x) = \prod_i P(X_i < x)$$

$$\Rightarrow f_{max}(x; b) = \begin{cases} \frac{n}{b^n} x^{n-1} & 0 \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

$$E[\vec{X}_{max}] = \int_0^b x \frac{n}{b^n} x^{n-1} \, dx = \frac{nb}{n+1}$$

$\therefore$ $\boxed{\hat{b}' = \frac{n+1}{n} \vec{X}_{max}}$ is unbiased and consistent

Illy $\boxed{\hat{a}' = \frac{n+1}{n} \vec{X}_{min}}$ is biased and consistent

$\therefore$ Assuming no noise $\boxed{\hat{b}' = 11 ; \quad \hat{a}' = -1}$

3 a We see $\alpha = 0.01$

$\mu_1 = 0.088$ $\mu_2 = -0.075$

$\sigma_1^2 = 8.539$ $\sigma_2^2 = 8.705$

$\mu_s = -0.457$ $s^2 = 14.971$

$n = 50$

C.I for mean. ~~assuming~~ ~~n is small~~

$$\left[\mu - Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}, \mu + Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$$

$$= [-1.866, 0.952]$$

C.I for variance

$$\left[\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}\right] = [9.377, 26.921]$$

We can deduce that sample belongs to population 2.

(b) Let $\mu_2 = -0.075$ $\mu_1 = -0.457 = \mu_1$

$H_0 : \mu = \mu_2$ $H_1 : \mu \neq \mu_2$.

Let $T = \left| \frac{\mu - \mu_2}{S/\sqrt{n}} \right| = 0.698$.

$t_{\frac{\alpha}{2}} = 2.01$ where $\alpha = 0.05$

We see $T < t_{\alpha/2}$ we accept H_0

i.e $\underline{\mu = \mu_2}$.

(c) F-test performed by $F = \frac{\max(a_1^2, a_2^2)}{\min(a_1^2, a_2^2)}$

Using significance level 0.6, we get an interval for 'm' : 4-18.

4 (a) Generating uniform normal data $N(0,1)$ and appending it to the file.

(a) File Q4_a.txt has 3 columns.

where $F = \frac{MSE}{MSE}$;

Since 'K' can be anything, I have printed p-value for the test.

(b) Spearman Rank correlation coefficient

$$\boxed{\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}}$$

d_i = difference in ranks of data in both columns.

I have printed the 'ρ' value and p-value for the test.