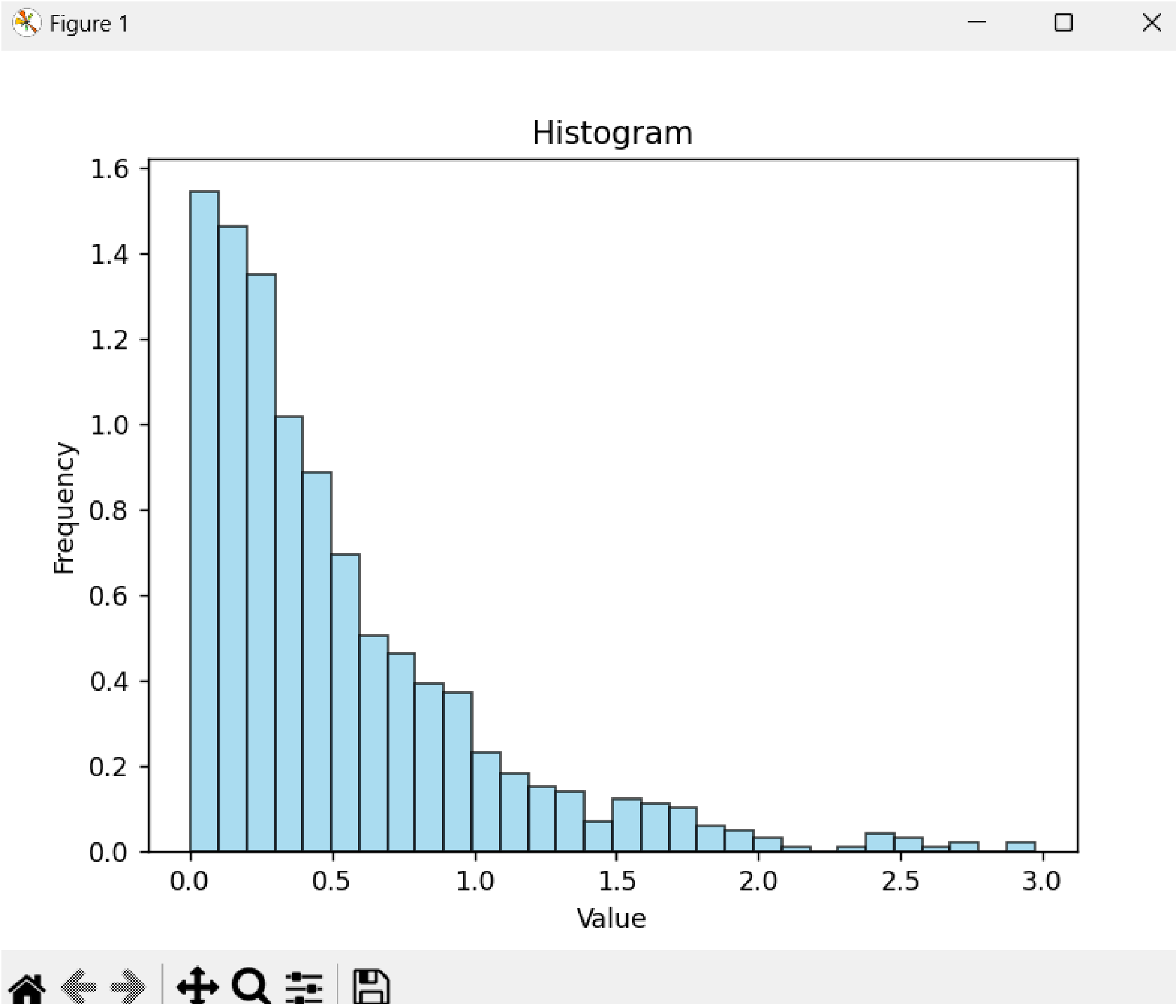
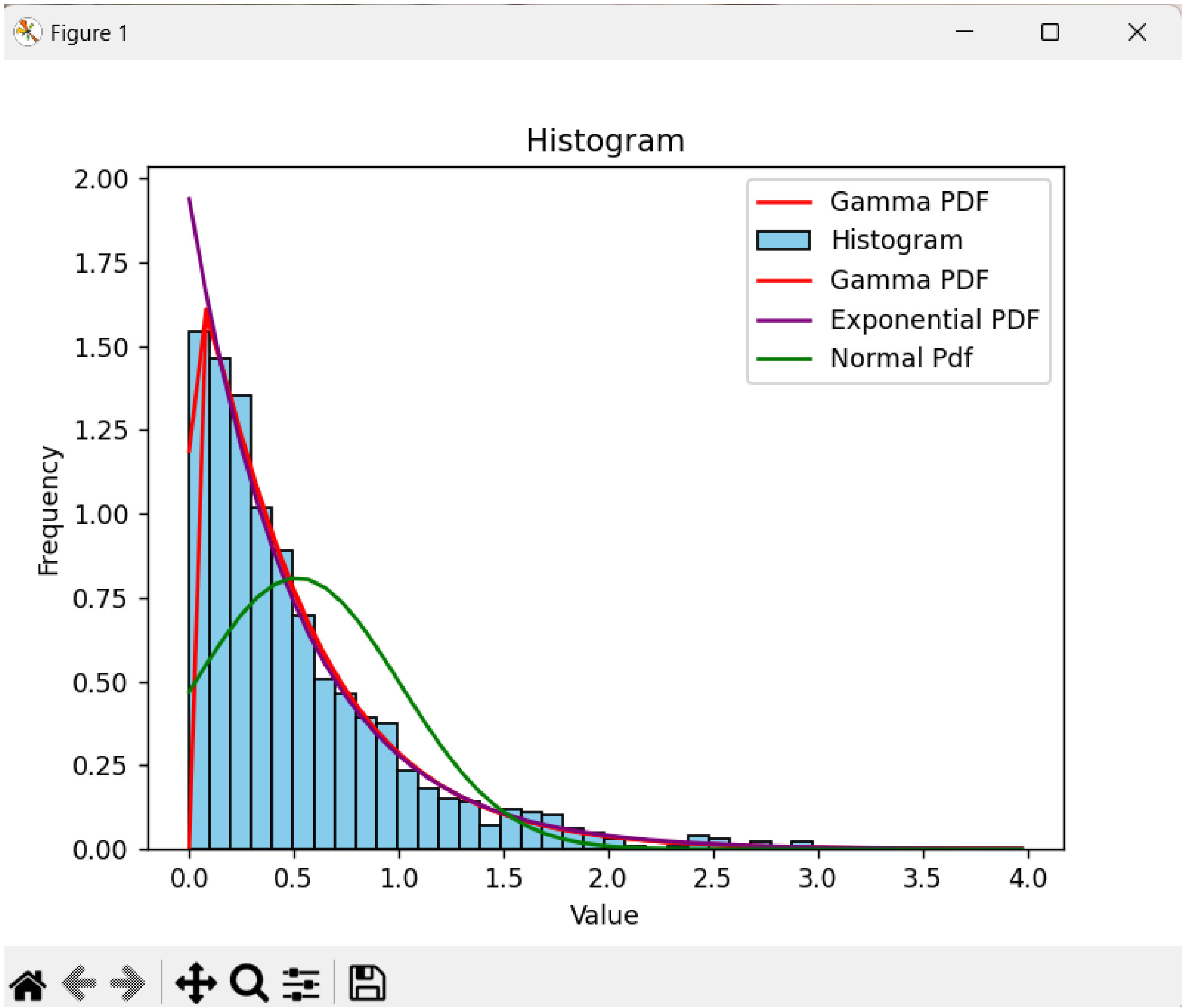
Assignment 1

Ans 1) Best fit is Gamma. How i found it is very simple

First Fiited some graphs / basic common distributions on histogram formed . Conflict arised between exponential and gamma but gamma wins as for exponential mean nearly equals variance but here it failed.



(distribution)



(fitted)
Answer BEST FIT ESTIMATE IS GAMMA FUNCTION

Some basic things
Shape parameter (alpha): 1.0670699752982742
Scale parameter (beta): 0.4826855809531399

(Confirmed answer using inbuilt scipy script)

```
ans_to_hold = scipy.stats.gamma.fit(data)
print(ans_to_hold)
print("Location parameter (theta):", ans_to_hold[1])
print("Shape parameter (alpha):", ans_to_hold[0])
print("Scale parameter (beta):", ans_to_hold[2])

# # yeh hm inbuilt libnrary se bhi kr skte hai but uske bina bhi ho jayega
```

Ans 2) Four estimators pair for alpha and beta found

Assignment Report

Q2] For two estimators, of above distribution.

We shown above that it is gamma.

Let our Distribution be $G(\alpha, \beta)$

W.K.T for Gamma,

$$E[X] = \alpha\beta$$

Using MOM

$$\cancel{E[X] = \bar{X}}$$

$$E[X] = \frac{1}{n} \Sigma X_i$$

$$\alpha\beta = \frac{1}{n} \Sigma X_i \quad \left(\text{As } \frac{1}{n} \Sigma X_i = \bar{X} \right)$$

$$\Rightarrow \alpha\beta = \bar{X} \quad \text{— (1)}$$

for $E[X^2] = \frac{\beta^2 \Gamma(\alpha+2)}{\Gamma\alpha}$

Using MOM

$$E[X^2] = \frac{1}{n} \Sigma x_i^2$$

$$\frac{\beta^2 \Gamma\alpha+2}{\Gamma\alpha} = \frac{1}{n} \Sigma x_i^2$$

using Gamma function property

$$\boxed{\Gamma\alpha+2 = (\alpha)(\alpha+1)\Gamma\alpha}$$

$$\frac{\beta^2 (\alpha)(\alpha+1)\cancel{\Gamma\alpha}}{\cancel{\Gamma\alpha}} = \frac{1}{n} \Sigma x_i^2 \quad (2)$$

from eqn (1) $\alpha = \frac{\bar{X}}{\beta}$

using it in eqn 2

$$\beta^2 \left[\frac{\bar{X}}{\cancel{\beta}}\right] \left[\frac{\bar{X}}{\beta}+1\right] = \frac{1}{n} \Sigma x_i^2$$

$$\frac{\cancel{\beta} [\bar{X}] [\bar{X}+\beta]}{\cancel{\beta}} = \frac{1}{n} \Sigma x_i^2$$

$$(\bar{X})^2 + \bar{X} \cdot \beta = \frac{1}{n} \Sigma x_i^2$$

Date
DELTA Pg No

$$\bar{x}\beta = \frac{1}{n}\sum x_i^2 - (\bar{x})^2$$

$$\bar{x}\beta = \frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2$$

$$= \frac{1}{n}\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)$$

$$= \frac{1}{n}\sum x_i^2 - (\bar{x})^2$$

$$= \frac{1}{n}\left(\sum x_i^2 - n\bar{x}^2\right)$$

$$= \frac{1}{n}\left(\sum x_i^2 + \bar{x}^2 n - 2n\bar{x}^2\right)$$

$$= \frac{1}{n}\sum (x_i - \bar{x})^2$$

$$\beta = \frac{1}{\bar{x}}\left[\frac{1}{n}\sum (x_i - \bar{x})^2\right]$$

for α, $\alpha = \frac{\bar{x}}{\beta}$

$$\alpha = \frac{\bar{x}}{\frac{1}{\bar{x}}\left[\frac{1}{n}\sum (x_i - \bar{x})^2\right]}$$

$$\alpha = \frac{(\bar{X})^2}{\frac{1}{n}\Sigma(x_i-\bar{x})^2}$$

for other 100 estimators

consider any $a>0$ multiply to both

that 100 is a estimator

let $a = 2$

for $\hat{\alpha} = \dfrac{(\bar{X})^2}{\frac{1}{n}\left[\Sigma(x_i-\bar{x})^2\right]} \times 2$

$\hat{p}$ $\dfrac{1}{\bar{x}}\left[\frac{1}{n}\Sigma(x_i-\bar{x})^2\right] \times 2$

Found a estimator for each

DELTA

3) i) for unbiasness

we need

$$E\left[\frac{(\bar{X})^2}{\frac{1}{n}\Sigma(x_i-\bar{x})^2}\right] = \alpha$$

$$E\left[\frac{\Sigma x_i/n}{\Sigma(x_i-\bar{x})^2}\right] \Rightarrow$$

$$= E\Big[$$

Now as, law this n increase or simply $n \to \infty$ (by city

Law of large numbers (Taught in Tut by 'Anupam Sir' + cited)

the demonstrations will convey

$$E[\hat{\alpha}] = \alpha$$

Hence, it is unbiased

for consistency

+ Consistency ensures that the estimator converge in probability to the true parameter value as the sample size increase

$Var(\hat{\alpha})$

$$= \frac{1}{n^2} var \left(\frac{\Sigma w}{\Sigma (x_i - \bar{x})^2} \right)$$

as $n \to \infty$ this consequently

the denominator will converge to

∞

$Var(\hat{\alpha}) \longrightarrow 0$

Hence, it is consistent

For efficiency, we first need to

know minima variance achievable

```
print("Found on Notebook values are \n alpha = X(bar)**2/((1/n)*(summ(Xi - X(bar))**2)) ")
```

```
print("Found on Notebook values are \n alpha = 2*X(bar)**2/((1/n)*(summ(Xi - X(bar))**2)) ")

print("Beta = (1/n)*summ(Xi - X(bar))**2/X(bar)")

print("Beta = 2*(1/n)*summ(Xi - X(bar))**2/X(bar)")
```

```
Found on Notebook values are
 alpha = X(bar)**2/((1/n)*(summ(Xi - X(bar))**2))
Found on Notebook values are
 alpha = 2*X(bar)**2/((1/n)*(summ(Xi - X(bar))**2))
Beta = (1/n)*summ(Xi - X(bar))**2/X(bar)
Beta = 2*(1/n)*summ(Xi - X(bar))**2/X(bar)

Shape parameter (alpha): 1.0670699752982742
Scale parameter (beta): 0.4826855809531399
```

Ans 3)

Need to find UMVUE

$$\operatorname{Var}_\theta(W(X)) \geq \frac{\left[\frac{\partial}{\partial\theta} E_\theta(W(X))\right]^2}{E\left[\frac{\partial}{\partial\theta} f(X|\theta)\right]}$$

$$-E\left[\frac{\partial^2}{(\partial\beta^2)^2}\left(\log\left(\frac{e^{-\Sigma x_i/\beta}}{\beta^{n\alpha}} \frac{\prod (x_i^{(\alpha-1)})}{(\Gamma\alpha)^n}\right)\right)\right]$$

$$= -E\left[\frac{\partial^2}{(\partial\beta^2)^2}\left[-\frac{\Sigma x_i}{\beta} + (\alpha-1)\Sigma \ln x_i - n\alpha \ln\beta - n\ln\Gamma\alpha\right]\right]$$

$$= -E\left[\frac{\partial}{\partial\beta}\left(\frac{\Sigma x_i}{\beta^2} - \frac{n\alpha}{\beta}\right)\right]$$

$$-E\left[\frac{-2\Sigma x_i}{\beta^3} + \frac{n\alpha}{\beta^2}\right]$$

$$= \frac{2\alpha\beta n}{\beta^3} - \frac{n\alpha}{\beta^2}$$

$$= \frac{n\alpha}{\beta^2}$$

$$= \operatorname{Var}_\alpha(W(X)) \geq \frac{\beta^2}{n\alpha}$$

Now we need to find for α

Some cramer & rao lower bound

$$Var_\theta(W(X)) \geq \frac{[\frac{\partial}{\partial\theta} E(W(X))]}{E\left(\frac{\partial}{\partial\theta} \log f(x|\theta)\right)}$$

$$= -E\left[\frac{\partial^2}{\partial\alpha^2}\left(\log\left(\frac{e^{-\Sigma x_i/\beta}\cdot\prod x_i^{\alpha-1}}{\beta^{n\alpha}\cdot(\Gamma\alpha)^n}\right)\right)\right]$$

$$= -E\left[\frac{\partial^2}{\partial\alpha^2}\left(\frac{-\Sigma x_i}{\beta} + (\alpha-1)\sum_{i=1}^{n}\ln x_i - n\alpha\ln\beta - n\ln\Gamma\alpha\right)\right]$$

$$\left(\text{please note } \frac{\partial}{\partial\alpha}\log\Gamma\alpha = \ln\alpha - \frac{1}{2\alpha}\right)$$

calculator found using differentials

$$= -E\left[\frac{\partial}{\partial\alpha}\left(\Sigma\ln x_i - n\ln\beta - n\left(\ln\alpha - \frac{1}{2\alpha}\right)\right)\right]$$

$$-E\left[-n\left[\frac{1}{\alpha} + \frac{1}{2\alpha^2}\right]\right]$$

$$nE\left[\frac{1}{\alpha}+\frac{1}{2\alpha^2}\right]$$

$$nE\left[\frac{2\alpha+1}{2\alpha^2}\right]$$

$$n\times\left(\frac{2\alpha+1}{2\alpha^2}\right]$$

$$n\left[\cancel{\frac{1}{\alpha}+\frac{1}{2\alpha^2}}\right]=\frac{n}{2\alpha^2}(2\alpha+1)$$

$$\frac{2\alpha^2}{n(2\alpha+1)}$$ for large val of α $\approx \frac{2\alpha^2}{n\times 2\alpha} = \alpha/n$

Now, to check effiency

for β. $V\left(\frac{\bar{X}}{\alpha}\right)=\frac{1}{\alpha^2}\times\frac{1}{n}\Sigma V(x_i)$

β's estimator

$\cdot V(x_i)=\alpha\beta^2$

$$=\frac{1}{n^2\alpha^2}\times\alpha\beta^2\times n=\frac{\beta^2}{n\alpha}$$

It is efficient

for α, $V\left(\frac{\bar{X}}{\beta}\right)=\frac{1}{\beta^2 n^2}\times\Sigma V(x_i)$

$$=\frac{1}{n^2\beta^2}\times n\times\alpha\beta^2=\frac{\alpha}{n}$$

Not efficient, efficient only for large val of α

for other roo

$V\left(\frac{2\times 1}{\bar{x}}\left[\frac{1}{n}\sum(x_i-\bar{x})^2\right]^{1}\right]$

$= 4V$ (same old estimator)

$= 4\frac{\beta}{n\alpha}$ hence not efficient

the other one is not efficient as, same reason

For consistency check for other roo, clearly they didn't give same expection as, $E[\alpha+\beta] = E[\alpha]+E[\beta]$

for any $\alpha, \beta \in R$ (Not common one)

so, a factor of 1 as bias is added

for ~~other~~ parameter of x

checking for unbiasness

$$E\left[\frac{\bar{X}}{\alpha}\right] = \frac{1}{\alpha} \times E\left[\frac{\sum x_i}{n}\right]$$

$$\frac{1}{\cancel{\alpha} \times \cancel{n}} \cancel{n} \times \cancel{\alpha} \beta$$

$$= \beta$$

(shortest process)

Hence, it is unbiased estimator

for $\hat{\beta} = \frac{1}{\bar{x}}\left[\frac{1}{n} \sum (x_i - \bar{x})^2\right]$

we can show a similar approach by law of large numbers that it will converge ~~of~~ to β; a similar approach used in α also for consistency

we can show

$\cancel{V(\hat{\beta}) = \frac{\beta^2}{n\alpha}}$ ~~(proof shown in next page of consistency)~~

$\cancel{\lim_{n\to\infty} \frac{\beta^2}{n\alpha}}$

$\cancel{= 0}$

$E[\hat{\beta}] =$

$E\left[\frac{1}{\bar{x}} \times \frac{1}{n} \Sigma(x_i - \bar{x})^2\right]$

$E\left[\frac{1}{\Sigma x_i} \times \Sigma(x_i)^2 - n\bar{x}^2\right]$

$E\left[\frac{\Sigma(x_i)^2}{\Sigma x_i} - \frac{n(\Sigma x_i)^2}{n^2 \times \Sigma x_i}\right]$

$E\left[\frac{\Sigma(x_i)^2}{\Sigma x_i} - \bar{x}\right]$

$\cancel{\alpha\beta}$ — $\alpha\beta$ $t = \alpha\alpha = 1$

$t = \alpha t \alpha$

$\alpha(\beta + \alpha)$

$= E\left[\frac{\cancel{(\Sigma x_i)^2}}{\Sigma x_i}\right] - \cancel{\alpha\beta}$

$E\left[\frac{\Sigma x_i^2}{\Sigma x_i}\right] - \alpha\beta$

$E(\cancel{x_i}^2)$

$= E[x]^2 -$

$V(x) = E[x^2] - E(x)^2$

Clearly it is independent of n for large n, it will give a finite result

$= \beta(1+\alpha) - \alpha\beta$

$\beta + \alpha\beta - \cancel{\alpha\beta}$

$= \beta$

$V(\hat{\beta}) = \frac{\beta^2}{n\alpha}$

$\lim_{n\to\infty} \frac{\beta^2}{n\alpha} = 0$

Hence it is consistent

(6) Other two are biased ~~but not~~

Becaus .

$E[2\hat{\alpha}]$ where $\hat{\alpha}$ is other estimator

$2E[\hat{\alpha}] = 2\hat{\alpha} \rightarrow$ biased

Similarly for one

Ans 4) Shown in above pictures

Ans 5) For alpha —> learge it makes estimator efficient

DELTA Pg No.

Need to find UMVUE

$$\mathrm{Var}_\theta(\omega(X)) \geqslant \frac{\left[\frac{\partial}{\partial\theta} E_\theta(\omega(X))\right]^2}{E\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right]}$$

$$-E\left[\frac{\partial^2}{(\partial\beta^2)^2}\left(\log\left(\frac{e^{-\Sigma x_i/\beta}}{\beta^{n\alpha}}\frac{\prod(x_i^{(\alpha-1)})}{(\Gamma\alpha)^n}\right)\right)\right]$$

$$= -E\left[\frac{\partial^2}{(\partial\beta^2)^2}\left[-\frac{\Sigma x_i}{\beta} + (\alpha-1)\Sigma \ln x_i - n\alpha\ln\beta - n\ln\Gamma\alpha\right]\right]$$

$$= -E\left[\frac{\partial}{\partial\beta}\left(\frac{\Sigma x_i}{\beta^2} - \frac{n\alpha}{\beta}\right)\right]$$

$$-E\left[\frac{-2\Sigma x_i}{\beta^3} + \frac{n\alpha}{\beta^2}\right]$$

$$= \frac{2\alpha\beta n}{\beta^3} - \frac{n\alpha}{\beta^2}$$

$$= \frac{n\alpha}{\beta^2}$$

$$= \mathrm{Var}_\alpha(\omega(X)) \geqslant \frac{\beta^2}{n\alpha}$$

Date
DELTA Pg No.

Now we need to find for α

Same cramer & rao lower bound

$$Var_\theta(W(X)) \geq \frac{\left[\frac{\partial}{\partial\theta} E(W(X))\right]}{E\left(\frac{\partial}{\partial\theta} \log f(x|\theta)\right)}$$

$$= -E\left[\frac{\partial^2}{\partial\alpha^2}\left(\log\left(\frac{e^{-\Sigma x_i/\beta} \cdot \prod x_i^{\alpha-1}}{\beta^{n\alpha}\cdot(\Gamma\alpha)^n}\right)\right)\right]$$

$$= -E\left[\frac{\partial^2}{\partial\alpha^2}\left(\frac{-\Sigma x_i}{\beta} + (\alpha-1)\sum_{i=1}^{n} \ln x_i - n\alpha \ln\beta - n\ln\Gamma\alpha\right)\right]$$

(please note $\frac{\partial}{\partial\alpha}\log\Gamma\alpha = \ln\alpha - \frac{1}{2\alpha}$)

found using differential calculator

$$= -E\left[\frac{\partial}{\partial\alpha}\left(\Sigma \ln x_i - n\ln\beta - n\left(\ln\alpha - \frac{1}{2\alpha}\right)\right)\right]$$

$$-E\left[-n\left[\frac{1}{\alpha} + \frac{1}{2\alpha^2}\right]\right]$$

$$nE\left[\frac{1}{\alpha}+\frac{1}{2\alpha^2}\right]$$

$$nE\left[\frac{2\alpha+1}{2\alpha^2}\right]$$

$$n\times\left(\frac{2\alpha+1}{2\alpha^2}\right]$$

$$\cancel{n\left[\frac{1}{\alpha}+\frac{1}{2\alpha^2}\right]} = \frac{n}{2\alpha^2}(2\alpha+1)$$

$$= \frac{2\alpha^2}{n(2\alpha+1)} \text{ for large val of } \alpha \approx \frac{2\alpha^2}{n\times 2\alpha} = \alpha/n$$

Now, to check efficiency

for β. $V\left(\frac{\overline{X}}{\alpha}\right) = \frac{1}{\alpha^2}\times\frac{1}{n}\Sigma V(x_i)$

β's estimator

$V(x_i) = \alpha\beta^2$

$$= \frac{1}{n^2\alpha^2}\times\alpha\beta^2\times n = \frac{\beta^2}{n\alpha}$$

It is efficient

for α, $V\left(\frac{\overline{X}}{\beta}\right) = \frac{1}{\beta^2 n^2}\times\Sigma V(x_i)$

$$= \frac{1}{n^2\beta^2}\times n\times\alpha\beta^2 = \frac{\alpha}{n}$$

Not efficient, efficient only for large val of α

for ~~scale~~ parameter of x

Checking for unbiasedness

$$E\left[\frac{\bar{X}}{\alpha}\right] = \frac{1}{\alpha} \times E\left[\frac{\Sigma x_i}{n}\right]$$

$$\frac{1}{\cancel{\alpha} \times \cancel{n}} \cancel{n} \times \cancel{\alpha} \beta$$

$$= \beta$$

(shorter proof)

Hence, it is unbiased estimator

for $\hat{\beta} = \frac{1}{\bar{x}}\left[\frac{1}{n} \Sigma(x_i - \bar{x})^2\right]$

we can show a similar approach by law of large numbers that it will converge ~~of~~ to β; a similar approach used in α also

Ans 6)

Question 6 Interval Estimation

alpha hat is 1.0906280438082017 beta hat is 0.4725638154562607

0.01 : (1.0017910732316269, 1.1794650143847765)

0.05 : (1.023031459443765, 1.1582246281726383)

0.1 : (1.0338992019323308, 1.1473568856840726)

How we found

```
# Question 6
print("Question 6 Interval Estimation")

# Recalculate alpha and beta using method of moments
mean_x = data.mean()
variance_x = data.var()

alpha_hat = mean_x**2 / variance_x
beta_hat = mean_x / alpha_hat
```

```
se_alpha = alpha_hat / np.sqrt(data.shape[0])


confidence_levels = [0.01, 0.05, 0.1]
z_values = {conf: stats.norm.ppf(1 - conf/2) for conf in
confidence_levels}


ci = {conf: (alpha_hat - z * se_alpha, alpha_hat + z * se_alpha) for conf,
z in z_values.items()}

print("alpha hat is ",alpha_hat,"beta hat is ",beta_hat)

for i in ci:
    print(i,":",ci[i])
```

Simply utilized a simple method first took vals in form from of list then for each used stats to find Z val for that alpha in its parameter
To comply the error

Ans 7) Ok now we can directly check For hypothesis by normal calculations and T test used as
n>=30 also simple random sampling test
Why T test not Z

Reason :- here we dont know population variance we are using sample variance thats why we are using
T test.

```
mean_0 = 0.525  # Hypothesized variance
real_mn = 0.515
```

```
# Calculating the t-test statistic
t_statistic = (real_mn - mean_0) / (math.sqrt(s_squared / n))

# Degrees of freedom
df = n - 1
```

Bounds = -1.962341461133449
1.9623414611334487
My val = -0.6415002990995847

Result : - Failed to reject the null hypothesis

Q7) Ho: $\mu = 0.525$

$\mu \neq 0.525$

Step 2 → Assumption check

1) Sample as Random Sampling

2) $n \geq 30$

3) but σ is unknown

so, we can use our good old t test

$\bar{x} = 0.515$

Test stat, $\dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ $s =$

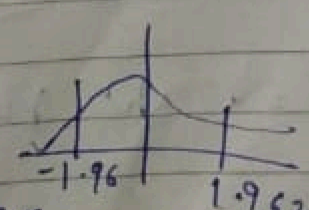
$\bar{x} = 0.515$.

$$= \frac{0.515 - 0.525}{\frac{\sqrt{0.243}}{\sqrt{10^3}}}$$

$$= \frac{-(0.010)\sqrt{10^3}}{\sqrt{0.243}}$$

$$= -0.64207$$

$t_{999, \alpha/2} = 1.9623$



Accept Ho, Not in RR

Ans 8) Ok now we can directly check For hypothesis by normal calculations and using chi square test

All assumptions are accepted

```
# Given values
n = 1000  # Sample size
s_squared = 0.24  # Sample variance
sigma_squared_0 = 0.25  # Hypothesized variance

print("H0 : s_squared = 0.24")
print("H0 : s_squared ≠ 0.24")
# Calculating the chi-squared test statistic
chi_squared_statistic = (n - 1) * s_squared / sigma_squared_0
print(chi_squared_statistic)

from scipy.stats import chi2

# Significance level for two-tailed test
alpha = 0.05
df = n - 1  # Degrees of freedom

# Critical values for both tails
lower_critical_value = chi2.ppf(alpha / 2, df)  # Lower tail
upper_critical_value = chi2.ppf(1 - alpha / 2, df)  # Upper tail

print(lower_critical_value, upper_critical_value)  # hypothesis test

print("accept") if chi_squared_statistic > lower_critical_value and
chi_squared_statistic < upper_critical_value else print("reject")
```

Result : - Failed to reject the null hypothesis
My val : 959.04
Bound : 913.3009983021134
1088.4870677259353

Result : - Failed to reject the null hypothesis

Date
DELTA Pg No.

(8) $H_0 : \sigma^2 = \sigma_0^2$ $\sigma_0^2 = 0.25$
$\sigma^2 \neq \sigma_0^2$

$s^2 = 0.24$

Steps Assumption

step 1) Normal sampling. Or
n is large

2) Simple random sampling

we can test using chi square test

$\frac{(n-1)s^2}{\sigma^2}$
test stat

$$\frac{999 \times 0.24}{0.25} = 959.07$$

$\chi^2_{999;0.025} = 1080.531$

$\chi^2_{Cal} < \chi^2_{Tab}$

Accept H_0.

Ans 9) Good ness of test

Can be computed using chi square

```
print("Goodness of test")
```

```
test_stat = "chi_squared_statistic"
alpha = 0.05
print("start kre")
rep = 0
mu = np.mean(data)
for i in data:
    rep +=(i-mu)**2
rep/=mu

# print(rep)

ans = chi2.ppf(1-alpha, df)
print("my calculated answer is ",rep)
print("table value is ",ans)
print("accept") if rep < ans else print("reject")
```

Results

my calculated answer is 472.5638154562601

table value is 1073.6426506574246

Failed to reject the null hypothesis

Result : - Failed to reject the null hypothesis