Page No.
Date

ASSIGNMENT - 3

REPORT

- Nimish Manware (22B0944)
- Prasanna Nage (22B0953)
- Jayesh Vinod Jadhav (22B1056)

Page No.
Date

1. Solution:

a) $x_1 = 1$ since, any picked book will have a different colour than already picked books, which are none.
There are $n-(i-1)$ books which are not picked yet, so the probability that we will get a new coloured book will be

$$\cancel{P(X_i \mid X_{i-1}) = \frac{n-(i-1)}{n}}$$

$$\text{Probability} = \frac{n-(i-1)}{n} = P(X_i = 1)$$

b) For any X_i, we have seen, that,

$$P(X_i = 1) = \frac{n-i+1}{n}$$

Now, $P(X_i = 2) = \left(\frac{i-1}{n}\right)\left(\frac{n-i+1}{n}\right)$ Because, we will have to choose, already selected books in the first trial and new one in the second.

similarly, $P(X_i = K) = \left(\frac{i-1}{n}\right)^{K-1}\left(\frac{n-i+1}{n}\right)$

comparing with geometric random variable,

$$P(X = K) = p(1-p)^{K-1}$$

so, $p = \frac{n-i+1}{n}$

c) $$E(Z) = \sum_{K=1}^{\infty} K P(Z=K) = \sum_{K=1}^{\infty} K \cdot p(1-p)^{K-1}$$

$$= p \sum_{K=1}^{\infty} K(1-p)^{K-1}$$

Now, we know that,

$$\sum_{K=0}^{\infty} (1-p)^K = \frac{1}{1-(1-p)} = \frac{1}{p}$$

Page No.
Date

Differentiating with respect to p on both sides,

$$-\sum_{k=0}^{\infty} k\,(1-p)^{k-1} = \frac{-1}{p^2}$$

$$\therefore \sum_{k=0}^{\infty} k\,(1-p)^{k-1} = \frac{1}{p^2}.$$

$$\therefore 0.(1-p)^{-1} + \sum_{k=1}^{\infty} k\,(1-p)^{k-1} = \frac{1}{p^2} \quad -(1)$$

$\therefore$ we get, $E(z) = p\sum_{k=1}^{\infty} k\,(1-p)^{k-1}$

$$= p \cdot \frac{1}{p^2} = \frac{1}{p}$$

d) $\mathrm{Var}(z) = E[z^2] - (E[z])^2$

$$(E[z])^2 = \frac{1}{p^2}$$

$$E[z^2] = \sum_{k=1}^{\infty} k^2\, p\,(1-p)^{k-1}$$

$$= p\sum_{k=1}^{\infty} k^2 (1-p)^{k-1}$$

Differentiating equation (1) on both sides, with respect to p,

$$-\sum_{k=1}^{\infty} (k^2-k)\,(1-p)^{k-2} = \frac{-2}{p^3}$$

$$\sum_{k=1}^{\infty} k^2 (1-p)^{k-2} = \frac{2}{p^3} + \sum_{k=1}^{\infty} k\,(1-p)^{k-2}$$

$$\frac{E[z^2]}{p(1-p)} = \frac{2}{p^3} + \frac{1}{p^2(1-p)}$$

$$E[z^2] = \frac{2}{p^2}(1-p) + \frac{1}{p} = \frac{2-p}{p^2}$$

$$\therefore \mathrm{Var}(z) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Page No.
Date

(d) $E(X^{(n)}) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{1}{p_i} = \sum_{i=1}^{n} \frac{n}{n+1-i}$

(e) $Var(X^{(n)}) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$ [$\because X_i$'s are independent random variables].

$= \sum_{i=1}^{n} \left(\frac{1-p_i}{p_i^2}\right) = \sum_{i=1}^{n} \left(\frac{i-1}{n}\right) \cdot \frac{n^2}{(n+1-i)^2}$

$= n \sum_{i=1}^{n} \frac{i-1}{(n+1-i)^2}$

$= n \cdot \left[\frac{1}{(n-1)^2} + \frac{2}{(n-2)^2} + \cdots + \frac{n-1}{1^2}\right]$

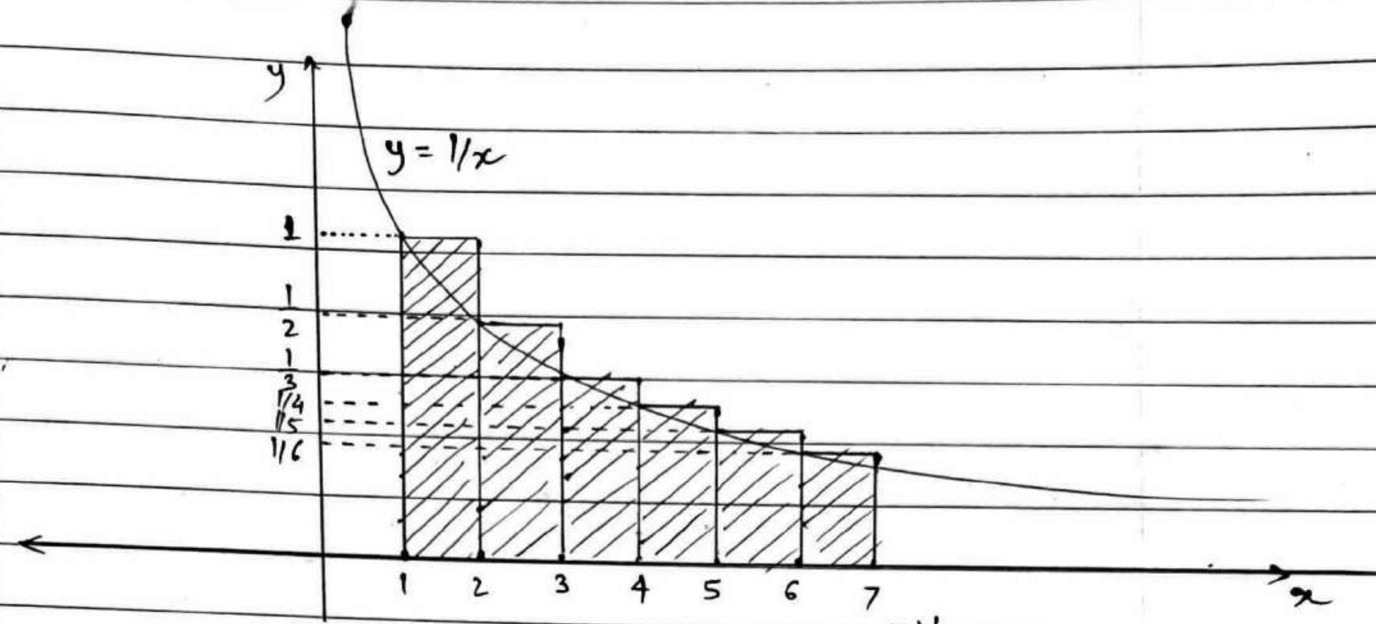
$\leq n \cdot \left[\frac{n-1}{(n-1)^2} + \frac{n-1}{(n-2)^2} + \cdots + \frac{n-1}{1^2}\right]$

$\leq n(n-1) \cdot \left[\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{(n-1)^2}\right]$

$\leq n(n-1) \left(\sum_{i=1}^{\infty} \frac{1}{i^2}\right) \leq n(n-1) \frac{\pi^2}{6}$

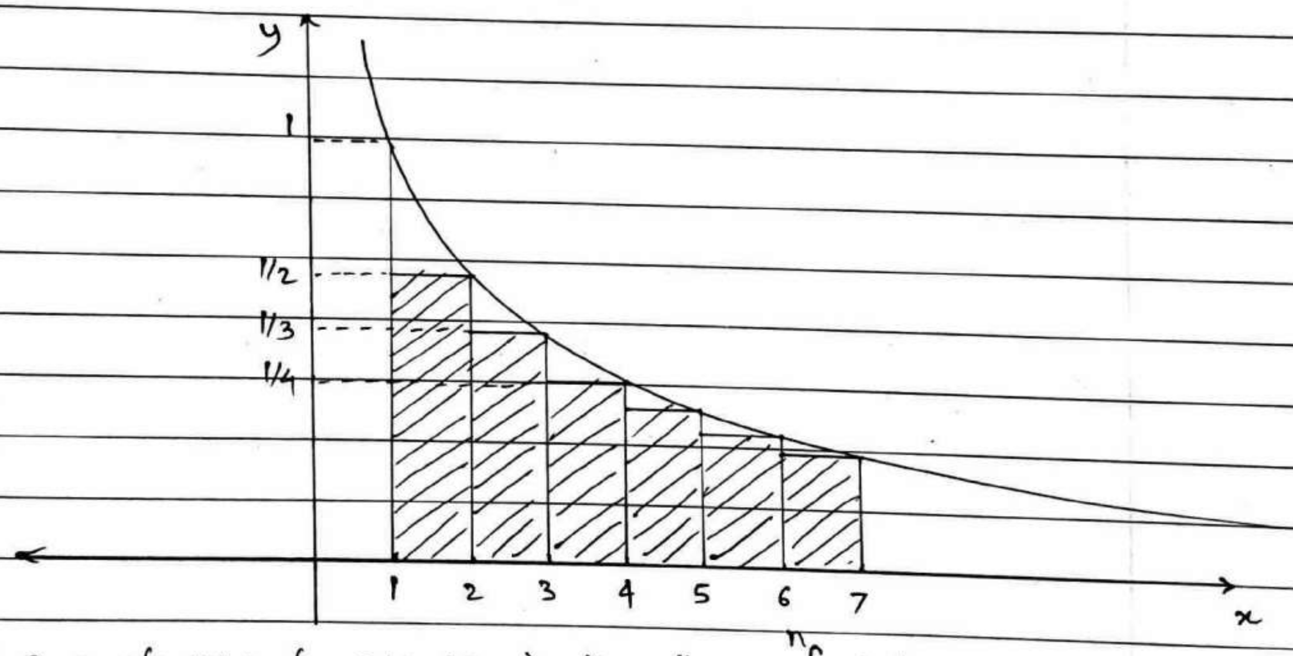
$\therefore Var(X^{(n)}) = \frac{\pi^2}{6}(n^2-n)$

(f)



sum of areas of rectangles is greater than $\int_1^{n+1} \frac{1}{x} dx$,

so, $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > \log(n+1) > \log(n) \quad -(1)$



sum of areas of rectangles is less than $\int_1^{n} \frac{1}{x} dx$

$$\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} < \log(n)$$

so, $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} < 1 + \log(n) < 2\log(n) \quad -(2)$

[For large n]

From (1) and (2),

$$\log(n) < \sum_{i=1}^{n} \frac{1}{i} < 2\log(n)$$

so, $n\log(n) < n\sum_{i=1}^{n} \frac{1}{i} < 2n\log(n)$

so, $n\log(n) < E[X^{(n)}] < 2n\log(n)$

so, $E[X^{(n)}] = \Theta(\log(n))$.

Thus, $f(n) = \log(n)$

Page No.
Date

2. Solution:

(a) $\{v_i\}_{i=1}^{n}$ where each $v_i = F^{-1}(u_i)$ and u_i is a $[0,1]$ uniform distribution.

So,

$$P(V < x) = P(F^{-1}(u_i) < x) \text{ for some } i \in [1,n]$$
$$= P(u_i < F(x))$$
$$= F(x) \quad [\because u_i \text{ belongs to uniform distribution in } [0,1]]$$

Hence, V follows distribution $F(x)$.

(b) We have

$$P(D \geq d) = P\left(\max_x \left| \frac{\sum_{i=1}^{n} 1.(Y_i \leq x)}{n} - F(x) \right| \geq d \right)$$

As, F is a CDF, it is non decreasing

so, $Y_i \leq x \Rightarrow F(Y_i) \leq F(x)$

$$P(D \geq d) = P\left(\max_x \left| \frac{\sum_{i=1}^{n} 1(F(Y_i) \leq F(x))}{n} - F(x) \right| \geq d \right)$$

Putting $y = F(x)$, we get $y \in [0,1]$.

$$P(D \geq d) = P\left(\max_{0 \leq y \leq 1} \left| \frac{\sum_{i=1}^{n} (F(Y_i) \leq y)}{n} - y \right| \geq d \right)$$
$$= P\left(\max_{y \in [0,1]} \left| \frac{\sum_{i=1}^{n} 1.(U_i \leq y)}{n} - y \right| \geq d \right)$$

[since, from the last subpart, we know that $F(Y_i) \sim U_i$, where U_i is a random variable from Uniform distribution on $[0,1]$

Hence, $P(D \geq d) = P\left(\max_{0 \leq y \leq 1} \left| \frac{\sum_{i=1}^{n} 1(U_i \leq y)}{n} - y \right| \geq d \right)$

So, $P(D \geq d) = P(E \geq d)$

The significance of this result is that we can use uniform distribution to pick samples from a particular distribution function and, thus it will be closer to the real values as the samples from uniform random variables will increase.

Page No.			
Date			

3. ~~2~~. Solution :

(a) We have

$$z_i = a x_i + b y_i + c + \varepsilon_i$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$

$$\therefore \quad z_i \sim \mathcal{N}(a x_i + b y_i + c, \sigma^2)$$

$$\therefore \quad P\left(\{z_i\}_{i=1}^{n}; \{x_i, y_i\}_{i=1}^{n}, a, b, c\right)$$

$$= \text{likelihood} = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z_i - a x_i - b y_i - c)^2}{2\sigma^2}\right)$$

$$\mathcal{L} = -\sum_{i=1}^{n} \frac{(z_i - a x_i - b y_i - c)^2}{2\sigma^2} - n \log \sigma - n \log \sqrt{2\pi}$$

$$= \log(\text{likelihood})$$

Differentiating partially w.r.t. a :

$$\frac{\partial \mathcal{L}}{\partial a} = -\sum_{i=1}^{n} \frac{2(z_i - a x_i - b y_i - c)(-x_i)}{2\sigma^2}$$

setting it to 0

$$\sum_{i=1}^{n} \frac{(z_i - a x_i - b y_i - c)(x_i)}{\sigma^2} = 0$$

$$\Rightarrow \sum_{i=1}^{n} (z_i - a x_i - b y_i - c)(x_i) = 0 \quad [\sigma \neq 0]$$

$$\Rightarrow a \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i y_i + c \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} z_i x_i \quad -(1)$$

Similarly after differentiating w.r.t. b and c, separately we get,

$$a \sum_{i=1}^{n} x_i y_i + b \sum_{i=1}^{n} y_i^2 + c \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} z_i y_i$$

$$a \sum x_i + b \sum y_i + nc = \sum_{i=1}^{n} z_i$$

Write, $S_x = \sum_{i=1}^{n} x_i$ $S_{xy} = \sum_{i=1}^{n} x_i y_i$ and $S_{x^2} = \sum_{i=1}^{n} x_i^2$

and similarly for others

$$\therefore \quad aS_{x^2} + bS_{xy} + cS_x = S_{xz}$$
$$aS_{xy} + bS_{y^2} + cS_y = S_{yz}$$
$$aS_x + bS_y + cn = S_z$$

Therefore, the matrix form will be :-

$$\begin{bmatrix} S_{x^2} & S_{xy} & S_x \\ S_{xy} & S_{y^2} & S_y \\ S_x & S_y & n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} S_{xz} \\ S_{yz} \\ S_z \end{bmatrix}$$

For vector form, let $\vec{v} = (a, b, c)$

$$\therefore \quad (S_{x^2}, S_{xy}, S_x) \cdot \vec{v} = S_{xz}$$
$$(S_{xy}, S_{y^2}, S_y) \cdot \vec{v} = S_{yz}$$
$$(S_x, S_y, n) \cdot \vec{v} = S_z$$

b) similar to that in the previous case,

$$\text{Likelihood} = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(z_i - a_1x^2 - a_2y^2 - a_3xy - a_4x - a_5y - a_c)^2}{2\sigma^2}\right\}$$

taking logarithm both sides

$$\cancel{\log(\text{likelihood}) = \mathcal{L} = -\sum_{i=1}^{n} \frac{(z_i - ax_i - by_i - c)^2}{2\sigma^2} - n\log\sigma - n\log\sqrt{2\pi}}$$

Similar to the previous subpart, the linear equations will be,

$$\log(\text{likelihood}) = \mathcal{L} = -\sum_{i=1}^{m} \frac{(z_i - a_1x_i^2 - a_2y_i^2 - a_3x_iy_i - a_4x_i - a_5y_i - a_c)^2}{2\sigma^2}$$
$$- n\log\sigma - n\sqrt{2\pi}$$

$$a_1\sum x_i^4 + a_2\sum x_i^2y_i^2 + a_3\sum x_i^3y_i + a_4\sum x_i^3 + a_5\sum x_i^2y_i + a_c\sum x_i^2 = \sum z_ix_i^2 \quad -(1)$$
$$a_1\sum x_i^2y_i^2 + a_2\sum y_i^4 + a_3\sum x_iy_i^3 + a_4\sum x_iy_i^2 + a_5\sum y_i^3 + a_c\sum y_i^2 = \sum z_iy_i^2 \quad -(2)$$
$$a_1\sum x_i^3y_i + a_2\sum x_iy_i^3 + a_3\sum x_i^2y_i^2 + a_4\sum x_i^2y_i + a_5\sum x_iy_i^2 + a_c\sum x_iy_i = \sum z_ix_iy_i \quad -(3)$$

4. Solution:

(b) Joint likelihood = $\hat{p}(\{x_i\}_V ; \sigma)$

$$= \prod_{x_i \in V} \left(\frac{\sum_{x_j \in T} \exp\left\{ \frac{-(x_i - x_j)^2}{2\sigma^2} \right\}}{n \sigma \sqrt{2\pi}} \right)$$

(c) The best sigma value for the LL comes out to be $\sigma = 1$. (it was observed to fluctuate on every run, but mostly it came out to be 1)

(d) The best sigma value for the value of D came out to be $\sigma = 1$, and the D value was which gave the best LL at this sigma is 8.305863×10^{-3}. This was fluctuating with every run.

(e) ~~If T and V~~ If we use the same data for both training and validation, that is T and V are taken equal, the cross-validation method becomes ineffective. Cross-validation is meant to check how a model works on a new data it has not seen before. This leads to a spiked graph because the cross validation procedure indicates that the model with smallest σ is the best choice since it fits the data perfectly. Also, for the smallest σ, the joint likelihood will be maximum.

The graph for this procedure is also given in the file.

Page No.			
Date			

5. Solution :

we know by IR,

$$\phi_{X-E[X]}(s) \leq e^{s^2(b-a)^2/8}$$

$$\therefore \phi_{S_n - E[S_n]}(s) = \prod_{\forall x_i} \phi_{x_i - E[x_i]}(s)$$

$$\leq \prod_{\forall i} e^{s^2(b_i-a_i)^2/8}$$

$$= \exp\left\{ \sum_{\forall i} \frac{s^2}{8} (b_i - a_i)^2 \right\}$$

$$= \exp\left\{ \frac{s^2}{8} \sum_{\forall i} (b_i - a_i)^2 \right\} \quad -(1)$$

Now,

$$P(S_n - E(S_n) > t) = P\left(e^{s[S_n - E(S_n)]} > e^{st}\right)$$

$$\leq \frac{E\left(e^{s[S_n - E(S_n)]}\right)}{e^{st}} = \phi_{S_n - E(S_n)}(s)$$

$$\leq \exp\left\{ \frac{s^2}{8} \sum_i (b_i - a_i)^2 - st \right\} \quad \forall s > 0$$

Now, $\because$ this is true for all s, $P(S_n - E(S_n) < t)$ will be less than the minimum possible value of the RHS of inequality.

$$\therefore P(S_n - E(S_n) > t) \leq \min\left\{ \exp\left(\frac{s^2}{8} \sum_i (b_i - a_i)^2 - st \right) \right\}$$

differentiating RHS of the inequality w.r.t. s and setting it to 0, we get

$$\exp\left\{ \frac{s^2}{8} \sum_i (b_i - a_i)^2 - st \right\} \left(\frac{s}{4} \sum_i (b_i - a_i)^2 - t \right) = 0$$

so,
$$s = \frac{4t}{\sum_i (b_i - a_i)^2}$$

Page No.
Date

$a_1 \Sigma x_i^3 + a_2 \Sigma x_i y_i^2 + a_3 \Sigma x_i^2 y_i + a_4 \Sigma x_i^2 + a_5 \Sigma x_i y_i + a_6 \Sigma x_i = \Sigma z_i x_i$ -(4)

$a_1 \Sigma x_i^2 y_i + a_2 \Sigma y_i^3 + a_3 \Sigma x_i y_i^2 + a_4 \Sigma x_i y_i + a_5 \Sigma y_i^2 + a_6 \Sigma y_i = \Sigma z_i y_i$ -(5)

$a_1 \Sigma x_i^2 + a_2 \Sigma y_i^2 + a_3 \Sigma x_i y_i + a_4 \Sigma x_i + a_5 \Sigma y_i + a_6 n = \Sigma z_i$ -(6)

These will be the equations obtained by differentiating L and setting it to zero, w.r.t. $a_1, a_2, \ldots, a_6$

So, the matrix formed will be,

$$\begin{bmatrix} S_{x^4} & S_{x^2y^2} & S_{x^3y} & S_{x^3} & S_{x^2y} & S_{x^2} \\ S_{x^2y^2} & S_{y^4} & S_{xy^3} & S_{xy^2} & S_{y^3} & S_{y^2} \\ S_{x^3y} & S_{xy^3} & S_{x^2y^2} & S_{x^2y} & S_{xy^2} & S_{xy} \\ S_{x^3} & S_{xy^2} & S_{x^2y} & S_{x^2} & S_{xy} & S_x \\ S_{x^2y} & S_{y^3} & S_{xy^2} & S_{xy} & S_{y^2} & S_y \\ S_{x^2} & S_{y^2} & S_{xy} & S_x & S_y & n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \begin{bmatrix} S_{zx^2} \\ S_{zy^2} \\ S_{zxy} \\ S_{zx} \\ S_{zy} \\ S_z \end{bmatrix}$$

let $\vec{u} = (a_1, a_2, a_3, a_4, a_5, a_6)$

So vector form will be,

$(S_{x^4}, S_{x^2y^2}, S_{x^3y}, S_{x^3}, S_{x^2y}, S_{x^2}) \cdot \vec{u} = S_{zx^2}$

$(S_{x^2y^2}, S_{y^4}, S_{xy^3}, S_{xy^2}, S_{y^3}, S_{y^2}) \cdot \vec{u} = S_{zy^2}$

$(S_{x^3y}, S_{xy^3}, S_{x^2y^2}, S_{x^2y}, S_{xy^2}, S_{xy}) \cdot \vec{u} = S_{zxy}$

$(S_{x^3}, S_{xy^2}, S_{x^2y}, S_{x^2}, S_{xy}, S_x) \cdot \vec{u} = S_{zx}$

$(S_{x^2y}, S_{y^3}, S_{xy^2}, S_{xy}, S_{y^2}, S_y) \cdot \vec{u} = S_{zy}$

$(S_{x^2}, S_{y^2}, S_{xy}, S_x, S_y, n) \cdot \vec{u} = S_z$

c) Equation of plane is:

$$z = 10.022x + 19.998y + 29.95$$

Noise variance = 23.0685

Page No.
Date

Now,

$$e^{L(s(b-a))} = \exp\left\{ \frac{s(b-a)a}{b-a} + \log\left(1+\frac{(a-ae^{s(b-a)})}{b-a}\right)\right\}$$

$$= \exp.\left\{ sa + \log\left(\frac{b-ae^{s(b-a)}}{b-a}\right)\right\}$$

$$= \left\{\frac{b-ae^{s(b-a)}}{b-a}\right\} e^{sa}$$

$$e^{L(s(b-a))} = \frac{be^{sa}-ae^{sb}}{b-a} \quad -(2)$$

Also,

$$E(e^{sx}) < E\left[\frac{x(e^{sb}-e^{sa})}{b-a} + \frac{be^{sa}-ae^{sb}}{b-a}\right]$$

$$< E\left[x\left(\frac{e^{sb}-e^{sa}}{b-a}\right)\right] + E\left[\frac{be^{sa}-ae^{sb}}{b-a}\right]$$

$$< \frac{e^{sb}-e^{sa}}{b-a} E[x] + \left(\frac{be^{sa}-ae^{sb}}{b-a}\right)$$

$\because E[x]=0$,

so, $$E(e^{sx}) < \frac{be^{sa}-ae^{sb}}{b-a}$$

Thus, by equation (2),

$$E(e^{sx}) < e^{L(s(b-a))}$$

(c) We have,

$$L(h) = \frac{ha}{b-a} + \log\left(1+\frac{(a-ae^{h})}{b-a}\right)$$

$$L(h) = \frac{ha}{b-a} + \log\left(\frac{b-ae^{h}}{b-a}\right)$$

$$= \frac{ha}{b-a} + \log(b-ae^{h}) - \log(b-a)$$

Page No.			
Date			

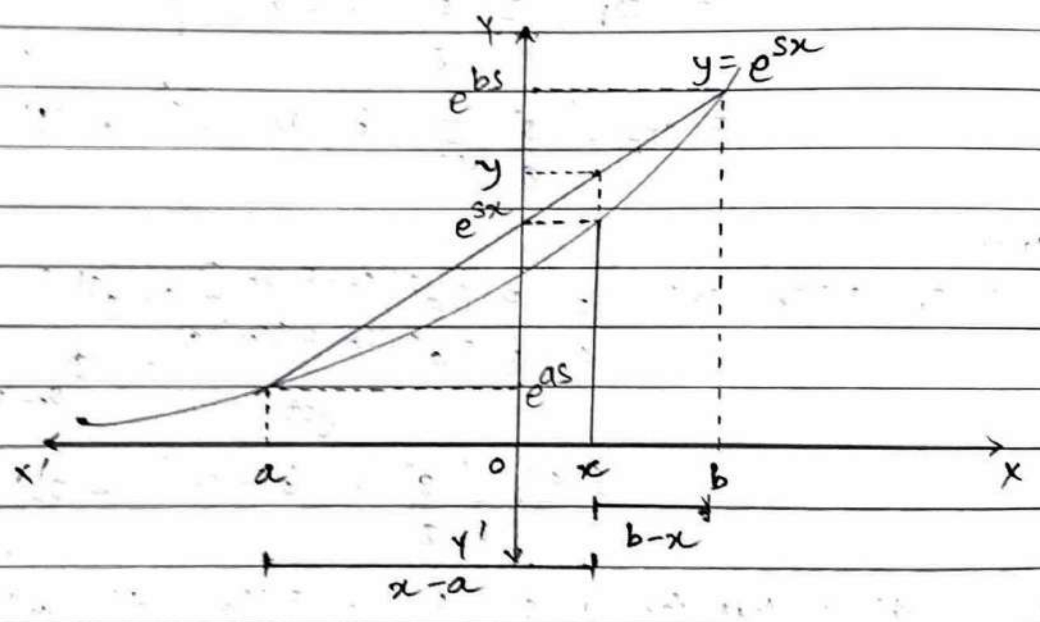
so

$$P(S_n - E(S_n) > t) \le \exp\left\{ \frac{1}{8} \frac{16t^2}{[\Sigma(b_i-a_i)^2]^2} [\Sigma(b_i-a_i)^2] - st \right\}$$

$$\le \exp\left\{ \frac{2t^2}{\Sigma(b_i-a_i)^2} - \frac{4t^2}{\Sigma(b_i-a_i)^2} \right\}$$

$$P(S_n - E(S_n) > t) \le \exp\left\{ \frac{-2t^2}{\Sigma(b_i-a_i)^2} \right\}$$

(a)



From graph it can easily be seen that $e^{sx} < y = \frac{(b-x)e^{sa}}{b-a} + \frac{(x-a)e^{sb}}{b-a}$

[by using section formula]

(b) $$E(e^{sx}) < E\left[\frac{(b-x)e^{sa}}{b-a} + \frac{(x-a)e^{sb}}{b-a} \right] \quad -(1)$$

[taking expectation on both sides].

Differentiating both sides w.r.t. h, we get

$$L'(h) = \frac{a}{b-a} + \frac{-ae^h}{b-ae^h}$$

$$L'(h) = \frac{a}{b-a} - \frac{ae^h}{b-ae^h}$$

Now,

$$L''(h) = \frac{-ae^h(b-ae^h) - (-ae^h)(-ae^h)}{(b-ae^h)^2}$$

$$= \frac{(-ae^h)\left[b-ae^h+ae^h\right]}{[b-ae^h]^2}$$

$$L''(h) = \frac{-abe^h}{(b-ae^h)^2}$$

$$= -\left(\frac{b}{b-ae^h}\right)\left(\frac{ae^h}{b-ae^h}\right) = \underbrace{\left(\frac{-b}{b-ae^h}\right)}_{x}\underbrace{\left(\frac{ae^h}{b-ae^h}\right)}_{y}$$

By AM $\geq$ GM inequality,

$$xy \leq \frac{(x+y)^2}{4}$$

so,

$$L''(h) \leq \frac{1}{4}\left(\frac{ae^h-b}{b-ae^h}\right)^2 = \frac{1}{4}$$

so,

$$L''(h) \leq \frac{1}{4}$$

$L(h)$ by Taylor's expansion theorem $\exists z \in [0, x]$ such that,

$$L(h) = L(0) + hL'(0) + \frac{h^2 L''(z)}{2!}$$

$$L(h) = \frac{h^2}{2} L''(z) \leq \frac{h^2}{2} \cdot \frac{1}{4}$$

$$L(h) \leq \frac{h^2}{8} \qquad -(3)$$

(d) $\therefore \quad e^{sx} \leq e^{L(s(b-a))} \leq e^{s^2(b-a)^2/8}$. [using (3)].