Rubric for End Sem

1 (a) The random variable X takes the values 0, 1 and 2

$$P[X=0] = \frac{2}{3} \times \frac{4}{5} = \frac{8}{15}$$

$$P[X=1] = \frac{1}{3} \times \frac{4}{5} + \frac{2}{3} \times \frac{1}{5} = \frac{4}{15} + \frac{2}{15} = \frac{6}{15}$$

+2

$$P[X=2] = \frac{1}{3} \times \frac{1}{5} = \frac{1}{15}$$

Then $E[X] = 0 \times \frac{8}{15} + 1 \times \frac{6}{15} + 2 \times \frac{1}{15}$

+2

$$= \boxed{\frac{8}{15}}$$

+4

Now $E[X^2] = 0^2 \times \frac{8}{15} + 1^2 \times \frac{6}{15} + 2^2 \times \frac{1}{15}$

+2

$$= \frac{6+4}{15} = \frac{10}{15} = \frac{2}{3}$$

Then $Var[X] = E[X^2] - (E[X])^2$

$$= \frac{2}{3} - \left(\frac{8}{15}\right)^2 = \frac{2}{3} - \frac{64}{225}$$

+4 +2

$$= \frac{2 \times 75 - 64}{225} = \frac{150-64}{225}$$

$$= \boxed{\frac{86}{225}}$$

(Total = 8 points)

(b) Let X be the gain of the player.

+1 Then X can take two values -15 and 1.

$P[X=-15] = P[\text{All four tosses result in tails}]$

+3 $= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$

$P[X=1] = P[\text{The first toss is Head}]$

$+ P[\text{The first toss is tail, the second toss is head}]$

$+ P[\text{The first two tosses are tails, the third toss is head}]$

+4 $+ P[\text{The first three tosses are tails, the fourth toss is head}]$

$= \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$

$= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{8+4+2+1}{16}$

$= \frac{15}{16}$

Therefore $E[X] = (-15) \times \frac{1}{16} + 1 \times \frac{15}{16}$

$= -\frac{15}{16} + \frac{15}{16} = \boxed{0}$

+4

(Total = 12 points)

2. Suppose the professor makes n measurements.

If $X_1, X_2, \dots, X_n$ are n measurements, then by central limit theorem

$$Z_n = \frac{\sum_{i=1}^{n} X_i - nd}{\sqrt{4}\sqrt{n}} \qquad \left(\begin{array}{l}\text{since } E[X_i] = d \\ \text{and } Var[X_i] = 4\end{array}\right)$$

$$= \frac{\sum_{i=1}^{n} X_i - nd}{2\sqrt{n}} \quad \text{is approximately } N(0,1)$$

+3

Now, we want

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - d\right| \le .5\right] \geq .96$$

(Equality will be o.k.)

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - d\right| \le .5\right]$$

$$= P\left[-.5 \le \frac{\sum_{i=1}^{n} X_i}{n} - d \le .5\right]$$

$$= P\left[-.5 \le \frac{\sum_{i=1}^{n} X_i - nd}{n} \le .5\right]$$

$$= P\left[-.5n \le \sum_{i=1}^{n} X_i - nd \le .5n\right]$$

$$= P\left[\frac{-.5n}{2\sqrt{n}} \le \frac{\sum_{i=1}^{n} X_i - nd}{2\sqrt{n}} \le \frac{.5n}{2\sqrt{n}}\right]$$

$$= P\left[-\frac{\sqrt{n}}{4} \le Z_n \le \frac{\sqrt{n}}{4}\right]$$

$$= \Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right)$$

$$= 2\,\Phi\left(\frac{\sqrt{n}}{4}\right) - 1$$

+7

We want $2\,\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 \geq .96$

Now $2\,\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 = .96$

$$\Rightarrow 2\,\Phi\left(\frac{\sqrt{n}}{4}\right) = 1.96 \Rightarrow \Phi\left(\frac{\sqrt{n}}{4}\right) = .98$$

+3

$$\Rightarrow \frac{\sqrt{n}}{4} = 2.06$$

$$\Rightarrow \sqrt{n} = 8.24 \Rightarrow n = 67.89$$

+2

So, n will be atleast 68

$\boxed{\text{Thus the professor will need 68 measurements}}$

(Total = 15 points)

Note: Some student may use a related theorem to solve this problem.

Theorem: Let X be a Gaussian (μ, σ) random variable.

A confidence interval estimate of μ of the form $M_n(X) - C \le \mu \le M_n(X) + C$

has confidence coefficient $1-\alpha$ where

$$\frac{\alpha}{2} = 1 - \Phi\left(\frac{c\sqrt{n}}{\sigma}\right)$$

+5

So, $P\left[\,|M_n(x) - \mu| \le c\,\right] = 1-\alpha$

where $\frac{\alpha}{2} = 1 - \Phi\left(\frac{c\sqrt{n}}{\sigma}\right)$

In our problem $\mu = d$, $c = .5$

and $1-\alpha = .96 \Rightarrow \alpha = .04$

and $\sigma = 2$

+5

So, $$\frac{.04}{2} = 1 - \Phi\left(\frac{.5\sqrt{n}}{2}\right)$$

$$.02 = 1 - \Phi\left(\frac{\sqrt{n}}{4}\right)$$

$$\Rightarrow \Phi\left(\frac{\sqrt{n}}{4}\right) = 1 - .02$$

$$\Rightarrow \Phi\left(\frac{\sqrt{n}}{4}\right) = .98$$

+3

$$\Rightarrow \frac{\sqrt{n}}{4} = 2.06$$

$$\Rightarrow \sqrt{n} = 4 \times 2.06 = 8.24$$

$$\Rightarrow n = 67.89$$

+2

So, n will be atleast 68

Either of the solution is correct but they need to mention the theorem clearly in their solution.

(Total = 15 points)

3. $X_1, X_2, \ldots$ are i.i.d. random variables with $E[X_i] = 75$ and standard deviation of X_i is 15

(a) We want to find the value of n such that $P[74 < M_n(X) < 76] = .99$

Now $P[74 < M_n(X) < 76]$

$= P[74 - 75 < M_n(X) - E(X) < 76 - 75]$

$= P[\,|M_n(X) - E(X)| < 1]$

$= 1 - P[\,|M_n(X) - E(X)| \geq 1]$

$\geq 1 - \frac{Var(X)}{n} = 1 - \frac{225}{n}$

(By chebyshev inequality)

+4

Thus $1 - \frac{225}{n} = .99 \Rightarrow \frac{225}{n} = .01$

$\Rightarrow n = \frac{225}{.01} = \boxed{22500}$

+3

So, $\boxed{n \geq 22500}$

(b) If each X_i is gaussian, the Sample mean $M_n(X)$ will be Gaussian with mean $E[M_n(X)] = E[\cancel{X}] = 75$

and $Var[M_n(X)] = \frac{Var[X]}{n} = \frac{225}{n}$

8

Now $P[74 < M_n(X) < 76]$

$$= P[74-75 < M_n(x) - 75 < 76-75]$$

$$= P\left[-\frac{1}{\sqrt{\frac{225}{n}}} < \frac{M_n(X)-75}{\sqrt{\frac{225}{n}}} < \frac{1}{\sqrt{\frac{225}{n}}}\right]$$

$$= P\left[-\frac{\sqrt{n}}{15} < \frac{M_n(x)-75}{\frac{15}{\sqrt{n}}} < \frac{\sqrt{n}}{15}\right]$$

$$= \Phi\left(\frac{\sqrt{n}}{15}\right) - \Phi\left(-\frac{\sqrt{n}}{15}\right)$$

+4

$$= 2\Phi\left(\frac{\sqrt{n}}{15}\right) - 1$$

+2

Now $2\Phi\left(\frac{\sqrt{n}}{15}\right) - 1 = .99$

$$\Rightarrow 2\Phi\left(\frac{\sqrt{n}}{15}\right) = 1.99 \Rightarrow \Phi\left(\frac{\sqrt{n}}{15}\right) = .995$$

$$\Rightarrow \frac{\sqrt{n}}{15} = 2.6 \Rightarrow n = (2.6 \times 15)^2$$

$$\Rightarrow n = (39)^2 = 1521$$

+2

So, $\boxed{n = 1521}$ or $\boxed{n \geq 1521}$

(Total = 15 points)

4 Let T be the arrival time of the train. Let X be the arrival time of the bus

Let Y be the arrival time of the employer's car.

Then T is $N(8{:}42, 4^2)$

X is $N(8{:}58, 3^2)$

and Y is $N(8{:}57, 2^2)$

Now

P[The clerk is late]

= P[The train arrives after 8:45]

+ P[The train arrives before 8:45]

× P[The bus arrives after 9:00]

$$= 1 - \Phi\left(\tfrac{3}{4}\right) + \Phi\left(\tfrac{3}{4}\right)\left[1 - \Phi\left(\tfrac{2}{3}\right)\right]$$

$$= 1 - \cancel{\Phi\left(\tfrac{3}{4}\right)} + \cancel{\Phi\left(\tfrac{3}{4}\right)} - \Phi\left(\tfrac{3}{4}\right)\Phi\left(\tfrac{2}{3}\right)$$

$$= \boxed{1 - \Phi\left(\tfrac{3}{4}\right)\Phi\left(\tfrac{2}{3}\right)}$$

+2

10

(a) P[Both the clerk and employer arrive late]

$= P[\text{The clerk is late}]\; P[\text{The employer is late}]$

$= \left[1 - \Phi\left(\frac{3}{4}\right)\Phi\left(\frac{2}{3}\right)\right] P[\text{The employer arrives after 9:00}]$

+3

$$= \boxed{\left[1 - \Phi\left(\frac{3}{4}\right)\Phi\left(\frac{2}{3}\right)\right]\left[1 - \Phi\left(\frac{3}{2}\right)\right]}$$

(b) First note that

$X - Y$ is a normal random variable

with $E[X-Y] = E[X] - E[Y] = -1$ minute
(8:57 − 8:58)

and $Var[X-Y] = Var(X) + Var(Y)$

$= 3^2 + 2^2 = 13$

+2

So, standard deviation of $X - Y$
is $\sqrt{13}$

Now P[The bus arrives after the employer's car]

$= P[X > Y] = P[X - Y > 0]$

$= P[X - Y - 1 > -1] = P\left[\frac{X-Y-1}{\sqrt{13}} > -\frac{1}{\sqrt{13}}\right]$

+3

$= 1 - \Phi\left(-\frac{1}{\sqrt{13}}\right) = \Phi\left(\frac{1}{\sqrt{13}}\right)$

Thus

P[The employee arrives before the clerk]

= P[The train arrives before 8:45] ×

P[The bus arrives after the employee's car]

+ P[The train arrives after 8:45]

$$= \boxed{\Phi\left(\frac{3}{4}\right)\Phi\left(\frac{1}{\sqrt{13}}\right) + \left[1-\Phi\left(\frac{3}{4}\right)\right]}$$

+5

(Can also be written as

$$1 - \Phi\left(\frac{3}{4}\right)\left[1-\Phi\left(\frac{1}{\sqrt{13}}\right)\right]$$

$$= 1 - \Phi\left(\frac{3}{4}\right)\Phi\left(-\frac{1}{\sqrt{13}}\right)$$
)

(Total = 15 points)

5 $$P_{N,K}(n,k) = \begin{cases} \dfrac{100^n e^{-100}}{(n+1)!} & k=0,1,\dots,n \\ & n=0,1,\dots \\ = 0 & \text{otherwise} \end{cases}$$

Marginal PMF of N

$$P_N(n) = \sum_{k=0}^{n} \frac{100^n e^{-100}}{(n+1)!} \quad \text{for } n=0,1,\dots$$

$$= \frac{100^n e^{-100}}{(n+1)!}(n+1)$$

+3

So, $$\boxed{P_N(n) = \begin{cases} \dfrac{100^n e^{-100}}{n!} & \text{for } n=0,1,\dots \\ = 0 & \text{otherwise} \end{cases}}$$

(Thus N is a Poisson RV with expected value 100)

Now for $n \geq 0$,
the conditional PMF of K given $N=n$ is

$$P_{K|N}(k|n) = \frac{P_{N,K}(n,k)}{P_N(n)} = \frac{\dfrac{100^n e^{-100}}{(n+1)!}}{\dfrac{100^n e^{-100}}{n!}} \quad \text{for } k=0,1,\dots,n$$

$$= 0 \quad \text{otherwise}$$

$$\Rightarrow \boxed{\begin{aligned} P_{K|N}(k|n) &= \frac{1}{n+1} && \text{for } k = 0, 1, \dots, n \\ &= 0 && \text{otherwise} \end{aligned}}$$

+3

Now $E[K|N=n] = \sum_{k=0}^{n} \frac{k}{(n+1)}$

$$= \frac{1}{(n+1)} \sum_{k=0}^{n} k$$

$$= \frac{1}{(n+1)} \sum_{k=1}^{n} k = \frac{1}{\cancel{(n+1)}} \cdot \frac{n\cancel{(n+1)}}{2}$$

+3

$$\Rightarrow \boxed{E[K|N=n] = \frac{n}{2}}$$

Therefore $\boxed{E[K|N] = \frac{N}{2}}$ +2

Now $E[K] = E\left[E[K|N]\right] = E\left[\frac{N}{2}\right]$

$$= \frac{1}{2} E[N] = \frac{100}{2} = \boxed{50}$$

+4

(because we have found the marginal PMF of N to be a Poisson (100) distribution and so $E[N] = 100$)

(Total = 15 points)

6 Let X be the number of threes in 3600 throws of a fair die.

Then X is Binomial $\left(3600, \frac{1}{6}\right)$ Random Variable.

+3

$$E[X] = 3600 \times \frac{1}{6} = \boxed{600}$$

$$Var[X] = 3600 \times \frac{1}{6} \times \frac{5}{6} = \boxed{500}$$

Now, $P[550 < X < 650]$

+2

$$= P[550 - 600 < X - 600 < 650 - 600]$$

$$= P[-50 < X - 600 < 50]$$

$$= P[\,|X - 600| < 50]$$

+3

$$= 1 - P[\,|X - 600| \geq 50]$$

$$\geq 1 - \frac{Var(X)}{(50)^2}$$

(since by chebyshev's inequality $P[\,|X-600| \geq 50] \leq \frac{Var(X)}{(50)^2}$)

+2

$$= 1 - \frac{\cancel{500}\,10}{\cancel{50} \times \cancel{50}\,5}$$

$$= 1 - \frac{1}{5} = \frac{4}{5}$$

So, $\boxed{P[550 < X < 650] \geq \frac{4}{5}}$

(Total = 10 points)

(7) If B_i is the number of points the Professor earned for the i-th book then B_i can take values 0, 1 or 3.

The PMF of B_i is: $P_{B_i}(x) = \begin{cases} \frac{1}{3} & \text{for } x = 0, 1, 3 \\ = 0 & \text{otherwise} \end{cases}$

Also $T = B_1 + B_2 + \cdots + B_{18}$

Now MGF of B_i: $\phi_{B_i}(s) = E[e^{sB_i}]$

$$= e^{s\times 0} \times \frac{1}{3} + e^{s\times 1} \times \frac{1}{3} + e^{s\times 3} \times \frac{1}{3}$$

$$= \boxed{\frac{1}{3}(1 + e^s + e^{3s})}$$

+2

Now MGF of T:

$$\phi_T(s) = \left[\phi_{B_i}(s)\right]^{18}$$ (Since $B_1, B_2, \ldots, B_{18}$ are i.i.d.)

$$= \boxed{\frac{[1 + e^s + e^{3s}]^{18}}{3^{18}}}$$

+2

$E[T]$ and $Var[T]$ can be calculated from the MGF $\phi_T(s)$

+2

$$\frac{d\left(\phi_T(s)\right)}{ds} = \frac{18}{3^{18}}\left[1+e^{s}+e^{3s}\right]^{17}\left(e^{s}+3e^{3s}\right)$$

$$\left.\frac{d\,\phi_Y(s)}{ds}\right|_{s=0} = \frac{18}{3^{18}}(3)^{17}(4) = \frac{18}{3}\times 4 = \boxed{24}$$

So, $\boxed{E[T] = 24}$

+2

Now
$$\frac{d^2\left(\phi_T(s)\right)}{ds^2} = \frac{18\times 17}{3^{18}}\left(1+e^{s}+e^{3s}\right)^{16}\left(e^{s}+3e^{3s}\right)^2 + \frac{18}{3^{18}}\left(1+e^{s}+e^{3s}\right)^{17}\left(e^{s}+9e^{3s}\right)$$

So,
$$\left.\frac{d^2\left(\phi_T(s)\right)}{ds^2}\right|_{s=0} = \frac{\cancel{18}^{\,2}\times 17\times \cancel{3^{16}}\times 4^2}{\cancel{3^{18}}\ \cancel{3^{2}}} + \frac{\cancel{18}^{\,6}}{\cancel{3^{18}}_{\,3}}\times \cancel{3^{17}}\,(10)$$

$$= 544 + 60 = 604$$

So, $E[T^2] = 604$

+2

Hence
$$Var[T] = E[T^2] - \left(E[T]\right)^2$$
$$= 604 - (24)^2 = 604 - 576 = 28$$

$\Rightarrow \boxed{Var[T] = 28}$

(Total = 10 points)

Note:

It is possible to calculate $E[T]$ and $Var[T]$ directly. But they will not get any Credit unless they use MGF. The question clearly says that using the first result, they have to get the second result.