Chandan Suri
CS4090

1. X = Number of points earned on the first part.
Y = Number of points earned on the second part.

(a) for calculating the expected recorded score

$E(X+Y) \Rightarrow$

$$E[X+Y] \Rightarrow \sum_x \sum_y (x+y)\, p(x,y)$$

$$\Rightarrow \Big[(0+0)0.02 + (0+5)0.06 + (0+10)0.02 + (0+15)0.10 + (5+0)0.04 + (5+5)0.15 + (5+10)0.20 + (5+15)0.1 + (10+0)0.01 + (10+5)0.15 + (10+10)0.14 + (10+15)0.01\Big]$$

$$\Rightarrow \Big[0 + 0.3 + 0.2 + 1.5 + 0.2 + 1.5 + 3.0 + 2.0 + 0.1 + 2.25 + 2.8 + 0.25\Big]$$

$$\Rightarrow \underline{\underline{14.1}}$$

$\therefore$ $\underline{\underline{E[X+Y] = 14.1}}$ Ans.

(b) If maximum of the two scores was recorded and this is a discrete distribution (rv) $\Rightarrow$

$$\therefore E[\max(x,y)] = \sum_x \sum_y \max(x,y)\, p(x,y)$$

$$\Rightarrow \Big[\max(0,0)0.02 + \max(0,5)0.06 + \max(0,10)0.02 + \max(0,15)0.10 + \max(5,0)0.04 + \max(5,5)0.15 + \max(5,10)0.20 + \max(5,15)0.10 + \max(10,0)0.01 + \max(10,5)0.15 + \max(10,10)0.14 + \max(10,15)0.01\Big]$$

$\Rightarrow [0 + 0.3 + 0.2 + 1.5 + 0.2 + 0.75 + 2.0 + 1.5$
$+ 0.1 + 1.5 + 1.4 + 0.15]$

$\Rightarrow \underline{\underline{9.6}}$.

$\therefore \underline{\underline{E[\max(x,y)] = 9.6}}$

Ans.

30.a To compute the covariance for X and Y $\Rightarrow$

$$\boxed{Cov(X,Y) = E[XY] - E[X]E[Y]}$$

So, we first need to compute the expected values $\Rightarrow$

for expected values, we need to compute the marginal PMF $\Rightarrow$

Marginal PMF of X is computed as $\Rightarrow$

$$\Rightarrow \boxed{P_X(x) = \sum_y p(x,y)}$$

So,

$P_X(0) = 0.02 + 0.06 + 0.02 + 0.1 = 0.2$

$P_X(5) = 0.04 + 0.15 + 0.2 + 0.1 = 0.49$

$P_X(10) = 0.01 + 0.15 + 0.14 + 0.01 = 0.31$

Marginal PMF of Y is computed as $\Rightarrow$

$$\Rightarrow \boxed{P_Y(y) = \sum_x p(x,y)}$$

So, $P_Y(0) = 0.02 + 0.04 + 0.01 = 0.07$

$P_Y(5) = 0.06 + 0.15 + 0.15 = 0.36$

$P_Y(10) = 0.02 + 0.2 + 0.14 = 0.36$

$P_Y(15) = 0.1 + 0.1 + 0.01 = 0.21$

Now,

Expected value of X is $\Rightarrow$

$$\boxed{E[X] = \sum_x x\, P_X(x)}$$

$$= 0(0.2) + 5(0.49) + 10(0.31)$$

$$= 0 + 2.45 + 3.1 = \underline{\underline{5.55}}$$

Now,

Expected value of Y is $\Rightarrow$

$$\boxed{E[Y] = \sum_y y\, P_Y(y)}$$

$$= 0(0.07) + 5(0.36) + 10(0.36) + 15(0.21)$$

$$= 0 + 1.8 + 3.6 + 3.15$$

$$= \underline{\underline{8.55}}$$

for calculating the expected value of $XY \Rightarrow$

$$\boxed{E[XY] = \sum_x \sum_y xy\, p(x,y)}$$

$$= \Big[(0\times0)0.02 + (0\times5)0.06 + (0\times10)0.02 + (0\times15)0.1$$
$$+ (5\times0)0.04 + (5\times5)0.15 + (5\times10)0.2 + (5\times15)0.1$$
$$+ (10\times0)0.01 + (10\times5)0.15 + (10\times10)0.14 + (10\times15)0.01\Big]$$

$$= \Big[0 + 0 + 0 + 0 + 0 + 3.75 + 10 + 7.5 + 0 + 7.5$$
$$+ 14 + 1.5\Big]$$

$$= \underline{\underline{44.25}}$$

for covariance,

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

$$= 44.25 - (5.55)(8.55)$$

$$= 44.25 - 47.45$$

$$= -3.2$$

$$\therefore \quad Cov(X,Y) = -3.2$$

Ans.

30.b To compute the correlation coefficient of X and Y =>

$$\boxed{\rho(x,y) \Rightarrow \frac{Cov(X,Y)}{\sigma_X \sigma_Y}}$$

We already know $Cov(X,Y) = -3.2$, So, to compute the standard deviation of X and Y =>

$$\sigma_X^2 = E[X^2] - (E[X])^2$$

(Variance of X)

Now,

$$E[X^2] = \sum_x x^2 p_X(x)$$

$$= 0^2 \times 0.2 + 5^2 \times 0.49 + 10^2 \times 0.31$$

$$= 0 + 12.25 + 31 = 43.25$$

for, $\sigma_Y^2 = E[Y^2] - (E[Y])^2$

So,

$$E[Y^2] = \sum_y y^2 P_Y(y)$$

$$= 0^2 \times 0.07 + 5^2 \times 0.36 + 10^2 \times 0.36 + 15^2 \times 0.21$$

$$= 0 + 9 + 36 + 47.25$$

$$= \underline{\underline{92.25}}$$

for Variance of X ⇒

$$\sigma_X^2 = 43.25 - (5.55)^2$$

$$= 43.25 - 30.8025$$

$$= \underline{\underline{12.4475}} = \underline{\underline{12.45}}$$

for Variance of Y ⇒

$$\sigma_Y^2 = 92.25 - (8.55)^2$$

$$= 92.25 - 73.1025$$

$$= \underline{\underline{19.1475}} = \underline{\underline{19.15}}$$

Thus,
the correlation coefficient of X and Y ⇒

$$\rho(x,y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{-3.2}{\sqrt{12.4475}\sqrt{19.1475}}$$

$$= \frac{-3.2}{3.5281 \times 4.3758} = -0.2073$$

$$= \underline{\underline{-0.21}}$$ Ans.

(2.)

joint pdf is given by =>

$$f(x,y) = \begin{cases} K(x^2+y^2) & 20 \le x \le 30,\ 20 \le y \le 30 \\ 0 & \text{Otherwise} \end{cases}$$

$$\boxed{K = 3/380000}$$

(a)

Referring to Exercise 7 from Homework 5, we have computed the marginal PDFs as =>

$$f_X(x) = \begin{cases} \frac{3x^2}{38000} + 0.05 & 20 \le x \le 30 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3y^2}{38000} + 0.05 & 20 \le y \le 30 \\ 0 & \text{otherwise} \end{cases}$$

Therefore,

Covariance between X & Y can be computed as

=>

$$\boxed{Cov(X,Y) = E[XY] - E[X]E[Y]}$$

Now, we need to compute the expected values,

$$\therefore E[X] = \int_{20}^{30} x f_X(x)\,.dx$$

$$= \int_{20}^{30} \left(\frac{3x^3}{38000} + 0.05x\right).dx$$

$$= \left[\frac{3}{38000} \cdot \frac{x^4}{4} + \frac{0.05x^2}{2}\right]\Bigg|_{20}^{30}$$

$$= \frac{3}{38000}\left[\frac{30^4}{4} - \frac{20^4}{4}\right] + 0.05\left[\frac{30^2}{2} - \frac{20^2}{2}\right]$$

$$= \frac{3}{38\cancel{000}} \times \frac{6500\cancel{00}}{4} + \frac{0.05}{1\cancel{00}} \times \frac{5\cancel{00}}{2}$$

$$= 12.8289 + 12.5 = \underline{\underline{25.3289}} = \underline{\underline{25.33}}$$

$\therefore$

$$E[Y] = \int_{20}^{30} y f_Y(y).dy$$

$$= \int_{20}^{30}\left(\frac{3y^3}{38000} + 0.05y\right).dy$$

$$= \left[\frac{3y^4}{38000 \times 4} + \frac{0.05y^2}{2}\right]\Bigg|_{20}^{30}$$

$$= \frac{3}{38000}\left[\frac{30^4 - 20^4}{4}\right] + 0.05\left[\frac{30^2 - 20^2}{2}\right]$$

$$= \frac{3}{38000} \times \frac{650000}{4} + \frac{0.05}{100} \times \frac{500}{2}$$

$$= 12.8289 + 12.5 = \underline{\underline{25.3289}} = \underline{\underline{25.33}}$$

Now,

$$E[XY] = \int_{20}^{30}\int_{20}^{30} xy f(x,y).\,dx\,dy$$

$\Rightarrow$

$$\Rightarrow \int_{20}^{30}\int_{20}^{30} xy\left(\frac{3}{380000}(x^2+y^2)\right) dx\, dy$$

$$\Rightarrow \frac{3}{380000}\int_{20}^{30}\int_{20}^{30}(x^3y + xy^3)\, dx\, dy$$

$$\Rightarrow \frac{3}{380000}\int_{20}^{30}\left[\frac{x^4y}{4} + \frac{x^2y^3}{2}\right]\Bigg|_{20}^{30} dy$$

$$\Rightarrow \frac{3}{380000}\int_{20}^{30}\left[y\left(\frac{30^4-20^4}{4}\right) + y^3\left(\frac{30^2-20^2}{2}\right)\right]dy$$

$$\Rightarrow \frac{3}{380000}\int_{20}^{30}\left(\frac{650000y}{4} + \frac{500y^3}{2}\right)dy$$

$$\Rightarrow \frac{3}{380000}\left[\frac{650000}{4}\cdot\frac{y^2}{2} + \frac{500}{2}\cdot\frac{y^4}{4}\right]\Bigg|_{20}^{30}$$

$$\Rightarrow \frac{3}{380000}\left[\left(\frac{650000}{4}\cdot\frac{30^2}{2} + \frac{500}{2}\cdot\frac{30^4}{4}\right) - \left(\frac{650000}{4}\cdot\frac{20^2}{2} + \frac{500}{2}\cdot\frac{20^4}{4}\right)\right]$$

$$\Rightarrow \frac{3}{380000}\left[(73125000 + 50625000) - (32500000 + 10000000)\right]$$

$$\Rightarrow \frac{3}{380000}\left[81250000\right] = \underline{\underline{641.4474}} = \underline{\underline{641.45}}$$

for covariance of X and Y =>

$$Cov(X,Y) \Rightarrow 641.4474 - (25.3289)(25.3289)$$

$$\Rightarrow \underline{\underline{-0.1058}} \quad \Rightarrow \underline{\underline{-0.11}}$$

Ans. Ans.

(b) for calculating the correlation coefficient of X and Y =>

$$\boxed{\rho(X,Y) \Rightarrow \frac{Cov(X,Y)}{\sigma_X \sigma_Y}}$$

To calculate the standard deviation of X and Y =>

we need to compute some expected values as follows =>

$$\therefore E[X^2] = \int_{20}^{30} x^2 f_X(x) \cdot dx$$

$$= \int_{20}^{30} x^2 \left(\frac{3x^2}{38000} + 0.05 \right) \cdot dx = \int_{20}^{30} \left(\frac{3x^4}{38000} + 0.05x^2 \right) \cdot dx$$

$$= \left[\frac{3}{38000} \cdot \frac{x^5}{5} + 0.05 \cdot \frac{x^3}{3} \right] \Bigg|_{20}^{30}$$

$$= \left[\left(\frac{3}{38000} \cdot \frac{30^5}{5} + 0.05 \cdot \frac{30^3}{3} \right) - \left(\frac{3}{38000} \cdot \frac{20^5}{5} + 0.05 \cdot \frac{20^3}{3} \right) \right]$$

$\Rightarrow \left(333.1579 + 316.6667\right)$

$\Rightarrow 649.8246 \approx \underline{\underline{649.83}}$ Ans.

$\therefore E[Y^2] = \int_{20}^{30} y^2 f_Y(y) . dy$

$= \int_{20}^{30} y^2 \left(\frac{3y^2}{38000} + 0.05\right) . dy = \int_{20}^{30} \left(\frac{3y^4}{38000} + 0.05y^2\right) . dy$

$= \left[\frac{3}{38000} \cdot \frac{y^5}{5} + 0.05 . \frac{y^3}{3}\right]\Bigg|_{20}^{30}$

$= \left[\left(\frac{3}{38000} \cdot \frac{30^5}{5} + 0.05 . \frac{30^3}{3}\right) - \left(\frac{3}{38000} \cdot \frac{20^5}{5} + 0.05 . \frac{20^3}{3}\right)\right]$

$= \left[333.1579 + 316.6667\right] = 649.8246$

$\approx \underline{\underline{649.83}}$ Ans

Now, variance of $X \Rightarrow$

$\sigma_x^2 = E[X^2] - (E[X])^2$

$= 649.83 - (25.\cancel{3289}33)^2$

$= \underline{\underline{8.2\cancel{7}}}$

and variance of $Y \Rightarrow$

$$\sigma_Y^2 = E[Y^2] - (E[Y])^2$$

$$= 649.83 - (25.3289)^2$$

$$= \underline{\underline{8.2\cancel{7}}}$$

To compute the correlation coefficient of X & Y =>

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$= \frac{-0.1058}{\sqrt{8.2\cancel{2}}\sqrt{8.2\cancel{2}}}$$

$$= \cancel{-0.01279} \approx \cancel{-0.0128}\ \underline{\underline{-0.0134}}$$

$$-0.01338$$

Ans. $\underline{\underline{(-0.01)}}$

③

Let's consider the case of 5 working days in a week. For those 5 days, the person would take the bus twice in morning and evening.

Hence, the person would take the bus 10 times during the week.

Considering the hint, we will assume,

X_1, X_2, X_3, X_4, X_5 as the waiting times of the person in morning for the 5 days.

&

$X_6, X_7, X_8, X_9, X_{10}$ as the waiting times of the person in the evening for the 5 days.

Let T denotes the total waiting time.

$$\boxed{\therefore\ T = \sum_{i=1}^{10} X_i}$$

ⓐ To compute the total expected waiting time ⇒

$$E(T) = E\left(\sum_{i=1}^{10} X_i\right) = \underline{E\left(\sum_{i=1}^{5} X_i\right) + E\left(\sum_{i=6}^{10} X_i\right)}$$

By the property of linearity,

$$= \sum_{i=1}^{5} E(X_i) + \sum_{i=6}^{10} E(X_i)$$

for $i=1$ to 5, as X_i follows a uniform distribution over $[0, 8]$.

⇒

$E(X_1) = E(X_2) = E(X_3) = E(X_4) = E(X_5)$

Referring to homework 5, question 4, the expectation of uniformly distributed random variable in the range [A, B] is $\frac{A+B}{2}$.

$\therefore \quad E(X_1) = \frac{0+8}{2} = 4$

$\therefore \quad \underline{E(X_1) = E(X_2) = \dots = E(X_5) = 4}$

Similarly for $i = 6$ to 10, X_i follows a uniform distribution over $[0, 10]$.

So, $\quad E(X_6) = \frac{0+10}{2} = 5$

$\therefore \quad \underline{E(X_6) = E(X_7) = \dots = E(X_{10}) = 5}$

Thus, the total expected time can be computed as $\Rightarrow$

$$E(T) = \sum_{i=1}^{5} E(X_i) + \sum_{i=6}^{10} E(X_i)$$

$$= 5E(X_1) + 5E(X_4)$$

$$= 5[4 + 5] = 45$$

$\therefore \quad \underline{E(T) = 45}$

Ans.

(b.) To compute the variance of total waiting time ⇒

$$V(T) = V\left(\sum_{i=1}^{10} X_i\right)$$

$$= V\left(\sum_{i=1}^{5} X_i\right) + V\left(\sum_{i=6}^{10} X_i\right)$$

By the property of linearity,

$$= \sum_{i=1}^{5} V(X_i) + \sum_{i=6}^{10} V(X_i)$$

Referring to question 4 in homework 5,

Variance of uniformly distributed rv in range [A, B] is ⇒ $\boxed{\dfrac{(B-A)^2}{12}}$

Also,

$\because$ $X_1 \ldots X_5$ follow an identical uniform distribution in the range [0, 8],

$$\therefore V(X_1) = V(X_2) = \ldots = V(X_5)$$

$$\therefore V(X_1) = \frac{(8-0)^2}{12} = \frac{64}{12}$$

$$\Rightarrow V(X_1) = V(X_2) = \ldots = V(X_5) = \frac{64}{12}$$

Similarly, $X_6 \ldots X_{10}$ follow an identical uniform distribution in the range [0, 10],

$$\therefore V(X_6) = V(X_7) = \ldots = V(X_{10})$$

$$\therefore\ V(X_6) = \cancel{\quad} \frac{(10-0)^2}{12} = \frac{100}{12}$$

$$\therefore\ V(X_6) = V(X_7) = \dots = V(X_{10}) = \frac{100}{12}$$

So, the variance of total waiting time becomes

$$= 5V(X_1) + 5V(X_6)$$

$$= 5\left[\frac{64}{12} + \frac{100}{12}\right] = 5 \times \frac{164}{12}$$

$$= \frac{820}{12} = \underline{\underline{68.33}}$$ Ans.

(c.)

We can compute the expected value and variance of the difference between morning and evening waiting times on a given day by taking into consideration the first day of the work week as rv's have identical uniform distributions for the morning and evening waiting times.

$\therefore$

Difference between morning and evening waiting time on the first day can be written as $X_1 - X_6$.

$$\therefore\ E(X_1 - X_6) = E(X_1) - E(X_6)$$ [By linearity property]

$$= 4 - 5 = \underline{\underline{-1}}$$ Ans

Similarly, variance of the difference will be ⇒

$$V(X_1 - X_6) = V(X_1) + V(X_6)$$

$$= \frac{64}{12} + \frac{100}{12} = \frac{164}{12}$$

$$= \underline{\underline{13.67}}$$

Ans

And, as stated above,
the expectation and variance of the difference between morning and evening wait times for any other day of the week will be the same.

Ans

(d.) Let,

T_1 ⇒ Total morning waiting time for a week.

T_2 ⇒ Total evening waiting time for a week.

So, $T_1 - T_2$ ⇒ Difference between total morning waiting time and total evening waiting time.

&

$$T_1 = X_1 + X_2 + X_3 + X_4 + X_5$$

$$T_2 = X_6 + X_7 + X_8 + X_9 + X_{10}$$

for calculating $E(T_1 - T_2)$ ⇒

$$= \sum_{i=1}^{5} E(X_1) - \sum_{i=6}^{6} E(X_6)$$ [By linearity property]

$$= 5 \times 4 - 5 \times 5 = \underline{\underline{-5}}$$ Ans

for calculating $V(T_1 - T_2) \Rightarrow$

$$\Rightarrow \sum_{i=1}^{5} V(X_i) + \sum_{i=6}^{10} V(X_i)$$

[Variances will be added]

$$\Rightarrow 5\,V(X_1) + 5\,V(X_6)$$

$$= 5\left[\frac{64}{12} + \frac{100}{12}\right]$$

$$= \frac{820}{12} \approx \underline{\underline{68.33}}$$

Ans.

4. Number of cars coming from each road onto the freeway is a random variable.

So, let,

$X_1 \rightarrow$ denotes (rv) the cars on Road 1 with expected value as 800 and standard deviation (σ_1) as 16.

$X_2 \rightarrow$ denotes (rv) the cars on Road 2 with expected value as 1000 and standard deviation (σ_2) as 25.

$X_3 \rightarrow$ denotes (rv) the cars on Road 3 with expected value as 600 and standard deviation (σ_3) as 18.

a) So,

expected total number of cars entering the freeway at this point during the period can be computed as $\Rightarrow$

$$= E[X_1 + X_2 + X_3]$$

By the property of linearity,

$$= E[X_1] + E[X_2] + E[X_3]$$

$$= 800 + 1000 + 600$$

$$= 2400$$

Ans.

b) To find the variance of the total number of entering cars $\Rightarrow$

In the question, the standard deviation of the number of cars entering the freeway is given but the information regarding covariance between the number of cars entering the freeway is not provided.

~~Ans~~ So, an assumption here would be that the number of cars entering the freeway from different roads are independent of each other.

$\therefore$ Variance can be computed as $\Rightarrow$
(of total number of cars entering freeway)

$$= V(X_1+X_2+X_3)$$

$$= V(X_1)+V(X_2)+V(X_3)$$

$$= \sigma_1^2 + \sigma_2^2 + \sigma_3^2$$

$$= 16^2 + 25^2 + 18^2$$

$$= \underline{1205}$$

Ans.

(c.)

$Cov(X_1, X_2) = 80$

$Cov(X_1, X_3) = 90$

$Cov(X_2, X_3) = 100$

Thus, three streams of traffic aren't independent!

Even though the three streams of traffic are dependent, this wouldn't change the expected value of total number of cars entering the freeway,

$\therefore E[X_1+X_2+X_3] = 2400$ [as computed before]

Now, to compute the standard deviation of the total, we need to first compute the variance of the total now that the 3 streams of traffic are dependent.

$$\Rightarrow V[X_1 + X_2 + X_3]$$

[can be re-written as]

$$\Rightarrow \sum_{i=1}^{3} \sum_{j=1}^{3} Cov(X_i, X_j)$$

[$\because X_1, X_2, X_3$ are dependent]

$$= Cov(X_1, X_1) + Cov(X_1, X_2) + Cov(X_1, X_3) + Cov(X_2, X_1) + Cov(X_2, X_2) + Cov(X_2, X_3) + Cov(X_3, X_1) + Cov(X_3, X_2) + Cov(X_3, X_3)$$

$$= Var(X_1) + Var(X_2) + Var(X_3) + 2Cov(X_1, X_2) + 2Cov(X_2, X_3) + 2Cov(X_1, X_3)$$

[$\because Cov(X, X) = Var(X)$ & $Cov(X, Y) = Cov(Y, X)$]

$$= 16^2 + 25^2 + 18^2 + 2 \times 80 + 2 \times 90 + 2 \times 100$$

$$= 1205 + 160 + 180 + 200$$

$$= \underline{\underline{1745}}$$

$\therefore$ Standard deviation $\Rightarrow \sqrt{V[X_1 + X_2 + X_3]}$

$$= \sqrt{1745}$$

$$= \underline{\underline{41.77}}$$ Ans

5.

Let X denote the truck haul time.

$$\boxed{X \sim N(8.46, 0.913)} \quad (N(\mu, \sigma))$$

(a) Probability that haul time will be atleast 10 minutes

$\Rightarrow$ $P(X \geq 10)$

$= 1 - P(X \leq 10)$ [$\because$ continuous, $P(X < 10) = P(X \leq 10)$]

Let's standardize the rv,

$$= 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{10-\mu}{\sigma}\right)$$

$$= 1 - P\left(\frac{X-8.46}{0.913} \leq \frac{10-8.46}{0.913}\right)$$

$$= 1 - P(Z \leq 1.6867) \quad \left[\frac{X-8.46}{0.913} = Z\right]$$

$$= 1 - P(Z \leq 1.69)$$

[from Appendix] $[\Phi(1.69) = 0.9545]$

$= 1 - 0.9545$

$= \underline{\underline{0.0455}}$ Ans.

Since, this is a continuous variable,

$P(X \geq 10) = P(X > 10) = \underline{\underline{0.0455}}$

Ans.

(b) Probability that haul time will be atleast 15 minutes

$\Rightarrow P(X > 15)$

$\Rightarrow 1 - P(X \leq 15)$ [∵ this a continuous variable and is cdf]

Let's standardize the random variable,

$\Rightarrow 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{15-\mu}{\sigma}\right)$

$\Rightarrow 1 - P\left(Z \leq \frac{15-8.46}{0.913}\right)$

$\Rightarrow 1 - P(Z \leq 7.16)$

$= 1 - \Phi(7.16)$ [from Appendix, this $\Phi(7.16)$ becomes 1.0]

$= 1 - 1 = \underline{\underline{0}}$ Ans.

(c.) Probability that haul time will be between 8 and 10 min $\Rightarrow$

$= P(8 \leq X \leq 10)$

Let's standardize the random variable,

$\Rightarrow P\left(\frac{8-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{10-\mu}{\sigma}\right)$

$\Rightarrow P\left(\frac{8-8.46}{0.913} \leq Z \leq \frac{10-8.46}{0.913}\right)$

$\Rightarrow P(-0.50 \leq Z \leq 1.69) = \Phi(1.69) - \Phi(-0.5)$

$\Rightarrow 0.9545 - 0.3085$

$= \underline{\underline{0.6460}}$

Ans.

(e.) If 4 haul times are independently selected, then,

the probability that atleast one of them exceeds 10 min can be computed as $\Rightarrow$

$\Rightarrow$ 1 − Probability that none of the four haul times exceeds 10 minutes.

$\Rightarrow$ ($\because$ All are iid,)

$\Rightarrow 1 - [P(X \le 10)]^4$

We already computed $P(X \le 10)$ in (a) part after standardizing the random variable.

Therefore,

$= 1 - [0.9545]^4 \quad \leftarrow [\because P(Z \le 1.69) = \Phi(1.69)]$

$\approx 1 - 0.8300$

$= \underline{\underline{0.1700}}$

Ans.

6. Bolt thread length follows a normal distribution, (X)

$$\boxed{X \sim N(\mu, \sigma)}$$ [X → rv denoting bolt thread length]

ⓐ Probability that thread length of a randomly selected bolt is within 1.5 SDs of it's mean value ⇒

$$= P(\mu - 1.5\sigma \leq X \leq \mu + 1.5\sigma)$$

Let's standardize the random variable,

$$\Rightarrow P\left(\frac{\mu - 1.5\sigma - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{\mu + 1.5\sigma - \mu}{\sigma}\right)$$

$$\Rightarrow P(-1.5 \leq Z \leq 1.5)$$

$$\Rightarrow \Phi(1.5) - \Phi(-1.5)$$

$$= 0.9332 - 0.0668$$

$$= \underline{\underline{0.8664}}$$ Ans.

ⓑ Probability that the thread length of a randomly selected bolt is farther than 2.5 SDs from its mean value ⇒

$$= P(X < \mu - 2.5\sigma \text{ or } X > \mu + 2.5\sigma)$$

$$\Rightarrow 1 - P(\mu - 2.5\sigma \leq X \leq \mu + 2.5\sigma)$$

Let's standardize the random variable,

$\Rightarrow$

$$1 - P\left(\frac{\mu - 2.5\sigma - \mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu + 2.5\sigma - \mu}{\sigma}\right)$$

$$= 1 - P(-2.5 \le Z \le 2.5)$$

$$= 1 - [\Phi(2.5) - \Phi(-2.5)]$$

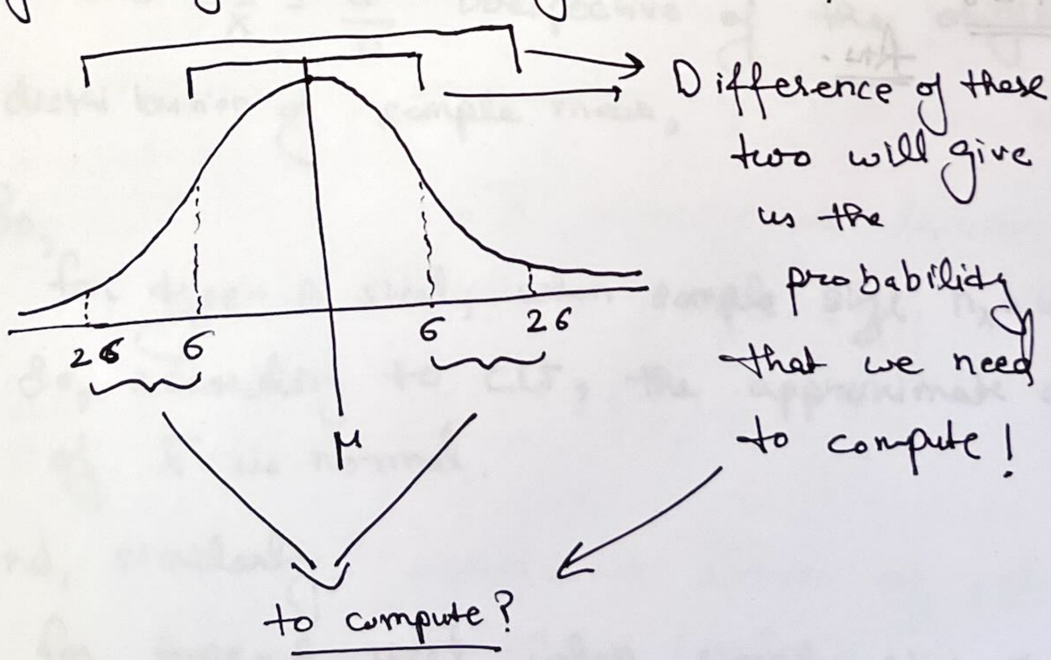
$$= 1 - [0.9938 - 0.0062]$$

$$= 1 - [0.9876] = \underline{\underline{0.0124}}$$

Ans

(c) Probability that the thread length is between 1 and 2 SDs from it's mean value $\Rightarrow$

We can come up with the computation for this by looking at the figure below $\Rightarrow$



Therefore,

$$= P(\mu - 2\sigma \le X \le \mu + 2\sigma) - P(\mu - \sigma \le X \le \mu + \sigma)$$

$\Rightarrow$

Let's standardize the random variable X,

$$\Rightarrow P\left(\frac{\mu-2\sigma-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu+2\sigma-\mu}{\sigma}\right) - P\left(\frac{\mu-\sigma-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu+\sigma-\mu}{\sigma}\right)$$

$$\Rightarrow P(-2 \le Z \le 2) - P(-1 \le Z \le 1)$$

$$\Rightarrow \left[\Phi(2) - \Phi(-2)\right] - \left[\Phi(1) - \Phi(-1)\right]$$

$$\Rightarrow \left[0.9772 - 0.0228\right] - \left[0.8413 - 0.1587\right]$$

$$\Rightarrow 0.9544 - 0.6826$$

$$\Rightarrow \underline{\underline{0.2718}}$$

Ans.

⑦ $\bar{X} \Rightarrow$ sample average tensile strength of a random sample of 40 type-A specimens.

$\bar{Y} \Rightarrow$ sample average tensile strength of a random sample of 35 type-B specimens.

$E[X] = 105$ ksi	$E[Y] = 100$ ksi
$\sigma_X = 8$ ksi	$\sigma_Y = 6$ ksi

ⓐ According to the central limit theorem (CLT), when the sample size is fairly large (or greater than 30 in reference to the book), then the distribution of the sample means is approximated as normal with sample mean $\mu_{\bar{X}} = \mu$ and sample variance $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ irrespective of the original distribution of sample mean,

So,

for type-A steel, when sample size $n_X = 40 > 30$.

So, according to CLT, the approximate distribution of $\bar{X}$ is normal.

And, similarly,

for type-B steel, when sample size $n_Y = 35 > 30$.

So, according to CLT, the approximate distribution of $\bar{Y}$ is normal.

(b) Since, the approximate distribution of both $\bar{X}$ and $\bar{Y}$ is normal and,

$\bar{X} - \bar{Y}$ is just a linear combination of those normally distributed variables,

So, the linear combination of two random variables ($\bar{X}$ and $\bar{Y}$ in this case) is also normal.

Therefore,
the approximate distribution of $\bar{X} - \bar{Y}$ is normal.
Ans

(c) To calculate $P(-1 \leq \bar{X} - \bar{Y} \leq 1) \Rightarrow$

we need to first calculate the expected value of the distribution denoted by $\bar{X} - \bar{Y}$ and the variance for the same distribution.

for normal distribution $\bar{X} \Rightarrow$

$$\mu_{\bar{X}} = E[\bar{X}] = E[X] = 105$$

$$\sigma_{\bar{X}}^2 = V[\bar{X}] = \frac{V[X]}{n_X} = \frac{(\sigma_X)^2}{n_X} = \frac{64}{40} = 1.6$$

Similarly, for normal distribution $\bar{Y} \Rightarrow$

$$\mu_{\bar{Y}} = E[\bar{Y}] = E[Y] = 100$$

$$\sigma_{\bar{Y}}^2 = V[\bar{Y}] = \frac{V[Y]}{n_Y} = \frac{(\sigma_Y)^2}{n_Y} = \frac{36}{35} = 1.03$$

Thus, $\bar{X} - \bar{Y}$ is linear combination of normally distributed vars, Thus, to compute mean =>

$$\mu_{\bar{X}-\bar{Y}} = E(\bar{X}-\bar{Y})$$
$$= E[\bar{X}] - E[\bar{Y}]$$
$$= 105 - 100 = \underline{5}$$

&

$$\sigma^2_{\bar{X}-\bar{Y}} = V[\bar{X}-\bar{Y}]$$
$$= V[\bar{X}] + V[\bar{Y}] = 1.6 + 1.03$$
$$= \underline{\underline{2.63}}$$

Standard deviation can be computed as =>

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\sigma^2_{\bar{X}-\bar{Y}}} = \sqrt{2.63} = 1.62$$ Ans.

Thus, the probability,

$$P(-1 \le \bar{X}-\bar{Y} \le 1) \Rightarrow$$

Let's standardize the random variable for the distribution denoted by $\bar{X}-\bar{Y}$ =>

$$= P\left(\frac{-1-\mu_{\bar{X}-\bar{Y}}}{\sigma_{\bar{X}-\bar{Y}}} \le \frac{\bar{X}-\bar{Y}-\mu_{\bar{X}-\bar{Y}}}{\sigma_{\bar{X}-\bar{Y}}} \le \frac{1-\mu_{\bar{X}-\bar{Y}}}{\sigma_{\bar{X}-\bar{Y}}}\right)$$

$$= P\left(\frac{-1-5}{1.62} \le Z \le \frac{1-5}{1.62}\right)$$

$$= P(-3.70 \le Z \le -2.47)$$

$$= \Phi(-2.47) - \Phi(-3.70)$$ [Values from Appendix]

$$= 0.0068 - 0 = \underline{0.0068}$$

$\therefore P(-1 \le \bar{X} - \bar{Y} \le 1) = 0.0068$

Ans.

(d) To calculate $P(\bar{X}-\bar{Y} \geq 10) \Rightarrow 1 - P(\bar{X}-\bar{Y} \le 10)$

Let's standardize the random variable,

$$= 1 - P\left(\frac{\bar{X}-\bar{Y}-\mu_{\bar{X}-\bar{Y}}}{\sigma_{\bar{X}-\bar{Y}}} \le \frac{10-\mu_{\bar{X}-\bar{Y}}}{\sigma_{\bar{X}-\bar{Y}}}\right)$$

$$= 1 - P\left(Z \le \frac{10-5}{1.62}\right)$$

$$= 1 - P(Z \le 3.09)$$

$$= 1 - \Phi(3.09)$$

$$= 1 - 0.9990$$

$$= 0.001$$

Ans. $\Rightarrow \mu_1 - \mu_2$ is same as $\mu_{\bar{X}} = \mu_{\bar{Y}} = \mu_{\bar{X}-\bar{Y}} = 5$.

Then mean of $\mu_{\bar{X}-\bar{Y}} = 5$ and the probability $P(\bar{X}-\bar{Y} \geq 10) = 0.001$ which is quite small and not likely to occur.

Therefore, when the actual observation is $\bar{X}-\bar{Y} \geq 10$, the claim $\mu_1 - \mu_2 = 5$ should be doubted.

Ans.

(8.) $X \rightarrow$ denotes the courtship time for a randomly selected female - male pair.

$$\mu_x = 120 \text{ min} \quad ; \quad \sigma_x = 110 \text{ min}$$

(b) To find the probability that the sample mean courtship time is between 100 min and 125 min, for a random sample of 50 such pairs.

As here the sample size is fairly large (>30), we can apply central limit theorem here.

As a result of CLT, sample mean $\bar{x}$ has an approx. normal distribution with mean $\mu_{\bar{x}}$ and variance $\sigma_{\bar{x}}$ as $\Rightarrow$

$$\mu_{\bar{x}} = \mu_x = 120$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n} = \frac{110^2}{50} = 242$$

To compute $P(100 \leq \bar{x} \leq 125)$ for a random sample of 50 such pairs,

let's standardize,

$$= P\left(\frac{100 - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \leq \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \leq \frac{125 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$= P\left(\frac{100-120}{\sqrt{242}} \leq Z \leq \frac{125-120}{\sqrt{242}}\right)$$

$$= P\left(\frac{-20}{\sqrt{242}} \leq Z \leq \frac{5}{\sqrt{242}}\right) = P(-1.29 \leq Z \leq 0.32)$$

$\Rightarrow \Phi(0.32) - \Phi(-1.29)$

Taking values from the Appendix Table,

$\Rightarrow 0.6255 - 0.0985 = \underline{0.5270}$

Therefore, the approx. probability that the sample mean courtship time is between 100 min and 125 min is

$\underline{0.5270}$. Ans.

(c) To compute the probability that the total courtship time exceeds 150 hrs $\Rightarrow$

As sample size (50) is greater than 30,

So, CLT can be applied here,

let, total courtship time is denoted by T.

As a result of CLT, T has an approximate normal distribution with mean μ_T and σ_T given as:-

$$\mu_T = n\mu_x = 50 \times 120$$
$$= 6000$$

&

$$\sigma_T = n\sigma^2 = 50 \times (110)^2$$
$$= 605000$$

Now,

150 hours $= (150 * 60)$ minutes

$= 9000$ minutes.

To find $P(T > 9000) \Rightarrow$

Let's standardize first,

$$\Rightarrow P\left(\frac{T-\mu_T}{\sigma_T} > \frac{9000-\mu_T}{\sigma_T}\right)$$

$$\Rightarrow 1-P\left(\frac{T-\mu_T}{\sigma_T} \leq \frac{9000-\mu_T}{\sigma_T}\right)$$

$$= 1 - P\left(Z \leq (9000-6000)/\sqrt{605000}\right)$$

$$= 1 - P(Z \leq 3.86) = 1 - \Phi(3.86)$$

Using Appendix Tables,

$$\Rightarrow \Phi(3.86) = 1.0$$

$$\Rightarrow 1 - 1 \approx 0$$

Therefore, the approximate probability that the total courtship time exceeds 150 hour is zero.

Ans.

9. $E[XY] = E[X]E[Y]$

Supposing that $X \sim \mathcal{U}(\{-1,0,1\})$ and

$U \sim \mathcal{U}(\{-1,1\})$ are independent &

$Y = XU$.

(a) To show that X and Y are uncorrelated, it would mean that

$$\mathrm{Corr}(X,Y) = 0$$

$\because$ $Corr(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$

and $Cov(X,Y) = E[XY] - E[X]E[Y]$

$\because$ $E[XY] = E[X]E[Y]$.

$\therefore$ $\underline{Cov(X,Y) = 0}$.

Thus, $\underline{Corr(X,Y) = 0}$.

$\therefore$ X and Y are <u>uncorrelated</u>. <u>Ans</u>.

(b). $\because$ X and Y are uncorrelated, that means that the joint probability of X and Y for all possible values in X and Y should be equal to zero.

so, $P_{X,Y}(x,y) = 0$ for $X = \{-1, 0, 1\}$
and $Y = \{-1, 0, 1\}$

$\because$ $Y = XU$

then, possible values of Y with all combinations of X and U will be $\Rightarrow$ (Pairs are shown as (x,y)).

$\Rightarrow \{(-1,1), (-1,-1), (0,0), (0,0), (1,-1), (1,1)\}$

$\therefore$ $P_X(x) = \begin{cases} 2/6 = 1/3 & x=-1 \\ 2/6 = 1/3 & x=0 \\ 2/6 = 1/3 & x=1 \end{cases}$; $P_Y(y) = \begin{cases} 2/6 = 1/3 & y=-1 \\ 2/6 = 1/3 & y=0 \\ 2/6 = 1/3 & y=1 \end{cases}$

[Marginal probabilities]

Now, let's take $(x, y) = (1, 1)$

$\Rightarrow$

$$P_X(1) \cdot P_Y(1) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

[We know that the product of the marginal probabilities should be equal to the joint probability for all possible values in the domain for the rv's to be independent.]

$\because$ $P_X(1) \cdot P_Y(1) = \frac{1}{9} \neq P_{X,Y}(1, 1) = 0$

$\therefore$ X and Y are not independent.

Hence, proved

Ans.