04b Sample Examination Problems Chapter 4 SOLUTIONS

1. X and Y are random variables with Normal distributions with mean 0, variance 1 and correlation coefficient 0.5. What is $P(X+Y>2)$? Assume that $X+Y$ has a normal distribution.

X ~ N(0,1) , Y ~ N(0,1) , $\rho = 0.5$, X + Y ~ N(?,?)
E(X+Y) = E(x) + E(Y) = 0 + 0 = 0
Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)
Note If X and Y were independent then Cov(X,Y)= 0
And Var(X+Y) = Var(X) + Var(Y)which is not the case here since $\rho = 0.5 \neq 0$

$$Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\sigma_{XY}}{\sqrt{\sigma_X{}^2\sigma_Y{}^2}}$$

$$\Rightarrow 0.5 = \frac{\sigma_{XY}}{\sqrt{1^2 \times 1^2}} \Rightarrow \sigma_{XY} = 0.5$$

Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)= 1 + 1+ 2(0.5) = 3

$\Rightarrow$ X + Y ~ N(0,3)
P(X+Y > 2) = P (W > 2) by letting W = X + Y

Standardize : P (W > 2) = $P(\frac{W-0}{\sqrt{3}} > \frac{2-0}{\sqrt{3}}) = P(Z > 1.15)$

= 1 – P(Z≤ 1.15) = 1 - $\phi(1.15)$= 1 – 0.8749 = 0.1251

---

2. X and Y are independent random variables with Normal distributions with mean 0 and variance 1. For some choice of $c>0$ $P(X+cY>4.2732)=0.15$. What is c?

X ~ N(0,1) , Y ~ N(0,1) , X and Y are independent

P(X+cY > 4.2732) = 0.15 , X + cY ~ N(?,?) , <u>c > 0</u>

E(X+cY) = E(x) + E(cY) = E(x) + cE(Y) = 0 + c(0) = 0

Var(X+cY) = Var(X) + Var(cY)+ + 2Cov(X,cY) but Cov(X,cY)= 0
= 1 + c^2(1) = 1 + c^2

X + cY ~ N(0,1+c^2) , standardize : Let X + cY = W

P(W > 4.2732) = 0.15 =

$$P(\frac{W-0}{\sqrt{1+c^2}} > \frac{4.2732-0}{\sqrt{1+c^2}}) = P(Z > \frac{4.2732}{\sqrt{1+c^2}}) = 0.15$$

$$\Rightarrow 1 - P(Z \le \frac{4.2732}{\sqrt{1+c^2}}) = 0.15 \Rightarrow 1-\phi(\frac{4.2732}{\sqrt{1+c^2}}) = 0.15 \Rightarrow \phi(\frac{4.2732}{\sqrt{1+c^2}}) = 0.85$$

From Table 4 : $\phi(1.03) = 0.8485$ **and** $\phi(1.04) = 0.8508$

We use Linear Interpolation :

Steps :

-take the difference : $\phi(1.04) - \phi(1.03)$ **=0.8508 – 0.8485 =0.0023**

-take the difference : 0.85 – 0.8485 = 0.0015

Now :

$$\frac{4.2732}{\sqrt{1+c^2}} = 1.03 + (\frac{0.0015}{0.0023})(0.01) = 1.0365$$

$$\sqrt{1+c^2} = \frac{4.2732}{1.0365} \Rightarrow \sqrt{1+c^2} = 4.1227 \Rightarrow 1+c^2 = 16.9968 \Rightarrow c^2 = 15.9968$$

$\Rightarrow c = \pm\sqrt{15.9968} = \pm 3.9996 \approx \pm 4$ **but c > 0** $\Rightarrow$ **c = 4**

---

3. Prove that

$$\text{var}(X+Y) = \text{var}\ (X) + \text{var}\ (Y) + 2\text{cov}(X,Y).$$

Remember Var(X) = E(X^2) – E^2(X)= E[X-E(X)]^2

$$\begin{aligned}\text{Var(X+Y)} &= \text{E[(X+Y)-E(X+Y)]}^2 = \text{E[X+Y-EX-EY]}^2 \\ &= \text{E[(X-EX)+(Y-EY)]}^2 \\ &= \text{E[(X-EX)}^2 + \text{(Y-EY)}^2 + 2\text{(X-EX)(Y-EY)]} \\ &= \text{E[(X-EX)}^2\text{]} + \text{E[(Y-EY)}^2\text{]} + 2\text{E[(X-EX)(Y-EY)]} \\ &= \text{Var(X)} + \text{Var(Y)} + 2\text{Cov(X,Y)}\end{aligned}$$

---

4. The distribution of (X,Y) is specified in the following table:

X	Y	Probability
1	6	1/3
2	5	1/3
3	4	1/3

Find the correlation coefficient of X, Y.

$$Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

	x	1	2	3	
Y					
4		0	0	1/3	1/3
5		0	1/3	0	1/3
6		1/3	0	0	1/3
		1/3	1/3	1/3	1

Var(X) = E(X^2) – E^2(X)

E(X) = $\sum_{x=1}^{3} xp_X(x) = 1\times\frac{1}{3} + 2\times\frac{1}{3} + 3\times\frac{1}{3} = 2$

E(X^2) = $\sum_{x=1}^{3} x^2 p_X(x) = 1\times\frac{1}{3} + 4\times\frac{1}{3} + 9\times\frac{1}{3} = \frac{14}{3}$

Var(X) = E(X^2) – E^2(X) = $\frac{14}{3} - 4 = \frac{2}{3}$

Similarly E(Y) = 5, E(Y^2)= $\frac{77}{3}$ **, Var(Y) =** $\frac{2}{3}$

Cov(X,Y) = E(XY) – E(X)E(Y)

E(XY) = $\sum_{x,y} xyp_X(x) = 1\times 6\times\frac{1}{3} + 2\times 5\times\frac{1}{3} + 3\times 4\times\frac{1}{3} = \frac{28}{3}$

Cov(X,Y) = E(XY) – E(X)E(Y) = $\frac{28}{3} - (2)(5) = \frac{-2}{3}$

$$Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{-2/3}{\sqrt{2/3\times 2/3}} = \frac{-2/3}{2/3} = -1$$

---

5. Why is a correlation coefficient used to measure linear association rather than covariance?

Cov(X,Y) can give a measure of linear relationship but we use ρ because :
-It is unaffected by either location or scale of measurement of X and Y.
-It is always in [-1 ,1]

---

6. Find the correlation coefficient of X and X^2 where X is a binomial random variable from 3 trials with probability of success 0.5.

$$Corr(X, X^2) = \rho_{XY} = \frac{Cov(X, X^2)}{\sqrt{Var(X)Var(X^2)}}$$

X ~ Bin(3 , 0.5) , n = 3 , p = 0.5

E(X) = np = 3(0.5)= 1.5
Var(X) = np(1-p)= 3(0.5)(0.5) = 0.75
Var(X^2) = E[(X^2)^2] – E^2(X^2)= E[X^4] – E^2(X^2)

X ~ Bin(3 , 0.5) : P(X = x)=$C_x^n (1/2)^x (1/2)^{n-x}$

P(X = 0)=$C_0^3 (1/2)^0 (1/2)^3 = 1/8$**, P(X = 1)=**$C_1^3 (1/2)^1 (1/2)^2 = 3/8$

X	0	1	2	3
P(X=x)	1/8	3/8	3/8	1/8

$\sum P(X = x) = 1$

Y=X^2	0	1	4	9
P(Y=y)	1/8	3/8	3/8	1/8

$\sum P(Y = y) = 1$

E(X^2)=E(Y)= $\sum_y y p_Y(y) = (0)(1/8) + (1)(3/8) + (4)(3/8) + 9(1/8) = 3$

Note this could have been obtained from :
Var(X) = E(X^2) – E^2(X)
$\Rightarrow$**E(X^2) = Var(X) + E^2(X) = 0.75 + (1.5)^2 = 3**

Var(X^2)= Var(Y) = E(Y^2) – E^2(Y)
E(Y^2) = $(0^2)(1/8) + (1^2)(3/8) + (4^2)(3/8) + (9^2)(1/8) = 33/2$
Or E(X^4) = $(0^4)(1/8) + (1^4)(3/8) + (2^4)(3/8) + (3^4)(1/8) = 33/2$
Var(X^2)= 33/2 – 9 = 15/2 = 7.5
Cov(X,X^2) = E(X.X^2) – E(X)E(X^2)= E(X^3) – E(X)E(X^2)
E(X^3)= $(0^3)(1/8) + (1^3)(3/8) + (2^3)(3/8) + (3^3)(1/8) = 54/8 = 6.75$
Cov(X,X^2)= 6.75 – (1.5)(3)= 2.25

$$Corr(X, X^2) = \rho_{XY} = \frac{Cov(X, X^2)}{\sqrt{Var(X)Var(X^2)}} = \frac{2.25}{\sqrt{0.75 \times 7.5}} = 0.9487$$