STAT 461_Hw 4

Julia Andrus

4.1-2a) Show space of X and Y.

X can take on the values {1,2,3,4} and Y can take the values of {2,3,4,5,6,7,8}

a) 4 sided die {1, 2, 3, 4}

X = red die, Z = black die outcome

$Y = X + Z$

$P(X=1) = 1/4$
$P(X=2) = 1/4$
$P(X=3) = 1/4$
$P(X=4) = 1/4$

For the black, we have equal outcomes

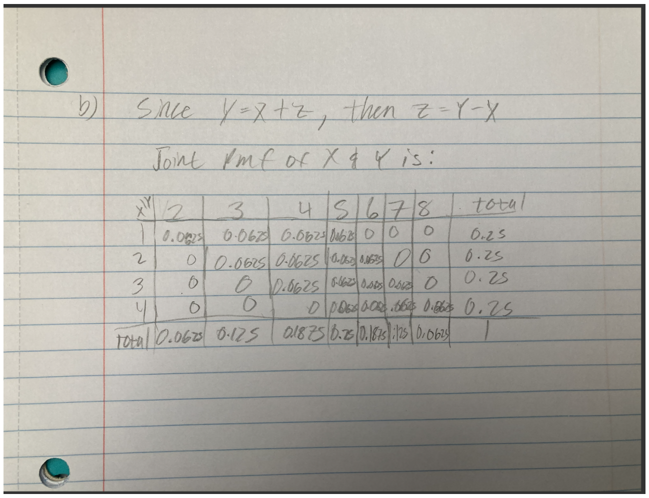
$P(Z=1) = 1/4$
$P(Z=1) = 1/4$
$P(Z=1) = 1/4$
$P(Z=1) = 1/4$

X & Z are independent

$P(X=x, Z=z) = P(X=x)P(Z=z)$
$0.0625 = (0.25)(0.25)$

X \ Z	1	2	3	4
1	0.0625	"	"	"
2	"	"	"	"
3	"	"	"	"
4	"	"	"	"

4.1-2b) Define the joint pmf on the space.



b) Since $Y = X + Z$, then $Z = Y - X$

Joint pmf of X & Y is:

X \ Y	2	3	4	5	6	7	8	total
1	0.0625	0.0625	0.0625	0.0625	0	0	0	0.25
2	0	0.0625	0.0625	0.0625	0.0625	0	0	0.25
3	0	0	0.0625	0.0625	0.0625	0.0625	0	0.25
4	0	0	0	0.0625	0.0625	.0625	0.0625	0.25
total	0.0625	0.125	0.1875	0.25	0.1875	.125	0.0625	1

Answer: $f(x,y) = \frac{1}{16}$, where x and y=1,2,3,4,

4.1-2c) Give the marginal pmf of X in the margin.

X is the outcome of the red 4 sided die, each outcome has an equal chance of occurring:

Answer: $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = \frac{1}{4}$. So $f_X(x) = \frac{1}{4}$, where x=1,2,3,4.

4.1-2d) Give the marginal pmf of Y in the margin

Shown in table in part(a):

P(Y=2)=P(X=1, Z=1)=0.0625
P(Y=3)=P(X=1, Z=2)+P(X=2, Z=1)=2(0.0625)=0.125
P(Y=4)=P(X=1, Z=3)+P(X=2, Z=2)+P(X=3, Z=1)=3(0.0625)=0.1875
P(Y=5)=P(X=1, Z=4)+P(X=2, Z=3)+P(X=3, Z=2)+P(X=4, Z=1)=4(0.0625)=0.25
P(Y=6)=P(X=2, Z=4)+P(X=3, Z=3)=3(0.0625)=0.1875
P(Y=7)=P(X=3, Z=4)+P(X=4, Z=3)=2(0.0625)=0.125
P(Y=8)=P(X=4, Z=4)=0.0625
Answer: $P(Y = 1) = P(Y = 2) = P(Y = 3) = P(Y = 4) = \frac{1}{4}$. So $f_Y(y) = \frac{1}{4}$, where y=1,2,3,4.

4.1-2e) Are X and Y independent?

Answer: X and Y are not independent because $P(X = x, Y = y) \neq P(X = x)P(Y = y)$.

4.1-6) n=25, very high=30%, high=40%, average=20%, and low=10%. P of 7 very high, P of 8 high, P of 6 average, P of 4 low.

For this, we have to use the Trinomial to Quadnomial pmf:

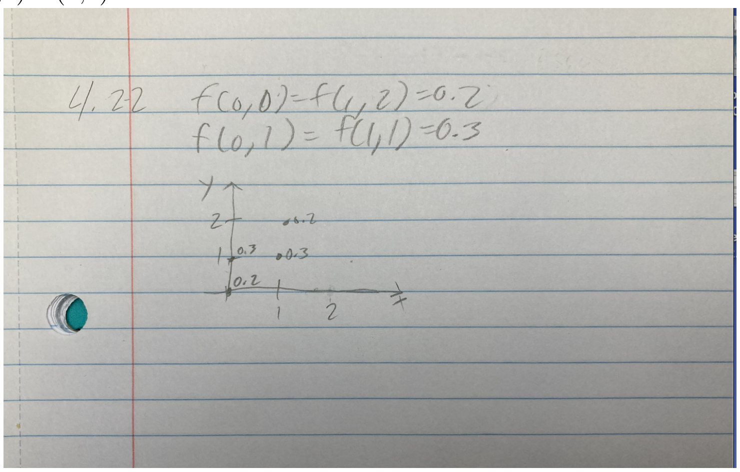
$f(x, y, z) = \frac{n!}{x!y!z!(n-x-y-z)}(P(X)^x)(P(Y)^y)(P(Z)^z)(1 - P(X) - P(Y) - P(Z)^{n-x-y-z}$

$f(7, 8, 6) = \frac{25!}{7!8!6!(25-7-8-6)}(0.3^7)(0.4^8)(0.2^6)(1 - 0.3 - 0.4 - 0.2)^{25-7-8-6}$

$\approx 0.004052 => 0.4052\%$

Answer: 0.4052%

4.2-2a) X and Y have the joint pmf defined by f(0,0)=f(1,2)=0.2 and f(0,1)=f(1,1)=0.3.



4.2-2b) Give the marginal pmfs in the margins.

Answer: For x:
$f_x(0) = f(0,0) + f(0,1) = 0.2 + 0.3 = 0.5$

$f_x(1) = f(1,1) + f(1,2) = 0.3 + 0.2 = 0.5$

Answer: For y:

$f_y(0) = f(0,0) = 0.2$

$f_y(1) = f(0,1) + f(1,1) = 0.3 + 0.3 = 0.6$

$f_y(2) = f(1,2) = 0.2$

4.2-2c) Compute $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$, COV(X,Y) and P.

1) μ_X and μ_Y

Answer: $\mu_X = E(X) = \sum x f(x) = 0(0.5) + 1(0.5) = \frac{1}{2}$

Answer: $\mu_Y = E(Y) = \sum y f(y) = 0(0.2) + 1(0.6) = 1$

2) COV(X,Y)

$E(XY) = \sum xyf(x,y) = 0(0)(0.2) + 0(1)(0.3) + 1(1)(0.3) + 1(2)(0.2) = \frac{7}{10}$
Answer: $COV(X,Y) = E(XY) - \mu_X \mu_Y = \frac{7}{10} - (\frac{1}{2})(1) = \frac{1}{5}$

3) $P = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$

$\sigma_X^2 = \sum (x - \mu_X)^2 f(x,y) = (0 - .5)^2(0.2) + (0 - .5)^2(0.3) + (1 - .5)^2(0.3) + (1 - .5)^2(0.3) = \frac{1}{4}$

$\sigma_Y^2 = \sum(y-\mu_Y)^2 f(x,y) = (0-1)^2(0.2) + (1-1)^2(0.3) + (1-1)^2(0.3) + (2-1)^2(0.2) = \frac{2}{5}$

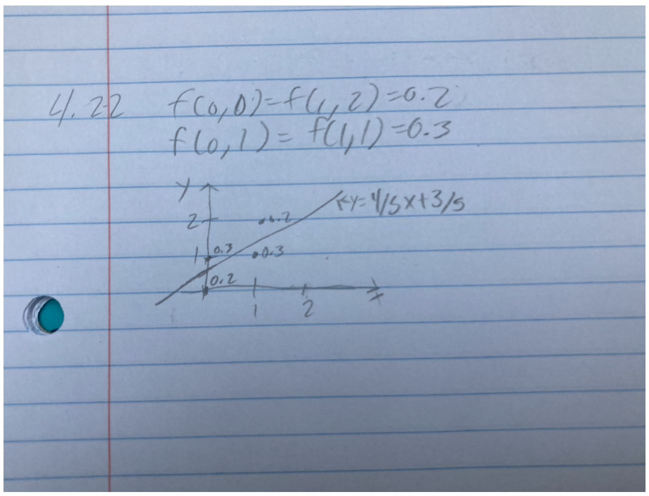
Now,

Answer: $P = \frac{COV(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{1}{5}}{\sqrt{\frac{1}{4}}\sqrt{\frac{2}{5}}} = \frac{\sqrt{10}}{5}$

4.2-2d) find the equation of the least squares regression line and draw on graph. Does the line make sense?

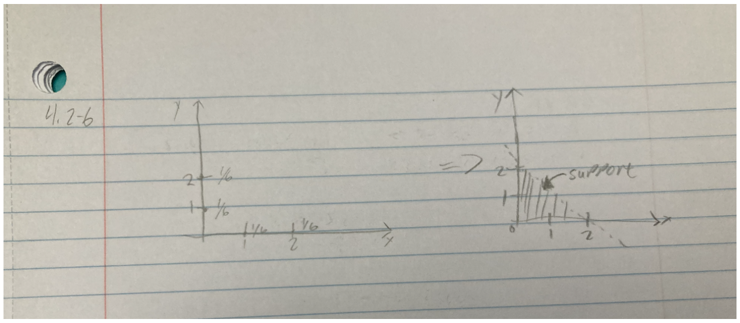
$y = \mu_Y + P(\frac{\sigma_Y}{\sigma_X})(x - \mu_X$

$y = 1 + \frac{\sqrt{10}}{5}(\frac{\frac{2}{5}}{\frac{1}{4}}(x - \frac{1}{2})$
$y = \frac{4}{5} + \frac{3}{5}$



Answer: Line does not make sense because none of the point lie on the least squares regression line.

4.2-6a) $f(x, y) = \frac{1}{6}, 0 \leq x + y \leq 2$, where x and y are nonnegative integers. Sketch the support of X and Y.



4.2-6b) Marginal pmfs $f_x(x)$ and $f_y(y)$.

Marginal distribution of X at x is the sum of the joint pmf at x overall possible y-values:

$f_X(0) = f(0, 0) + f(0, 1) + f(0, 2) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

$f_X(1) = f(1, 0) + f(1, 1) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

$f_X(2) = f(2, 0) = \frac{1}{6}$

Marginal distribution of Y at y is the sum of the joint pmf at y overall possible x-values:

$f_Y(0) = f(0, 0) + f(1, 0) + f(2, 0) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

$f_Y(1) = f(0, 1) + f(1, 1) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

$f_Y(2) = f(0, 2) = \frac{1}{6}$

4.2-6c) Compute COV(X,Y)

1) $\mu_x = E(X) = \sum xf(x) = 0(\frac{1}{2}) + 1(\frac{1}{3}) + 2(\frac{1}{6}) = \frac{2}{3}$
$\mu_y = E(Y) = \sum yf(y) = 0(\frac{1}{2}) + 1(\frac{1}{3}) + 2(\frac{1}{6}) = \frac{2}{3}$

2) $E(XY) = \sum xyf(x, y) = 0(0)(\frac{1}{6}) + 0(1)(\frac{1}{6}) + 1(1)(\frac{1}{6}) + 1(0)(\frac{1}{6}) + 2(0)(\frac{1}{2}) + 0(2)(\frac{1}{6}) = \frac{1}{6}$
$Cov(X, Y) = E(XY) - \mu_x\mu_y = \frac{1}{6} - (\frac{2}{3})(\frac{2}{3}) = -\frac{5}{18} \approx -0.2778$

Answer: ≈ -0.2778

4.2-6d) Determine P, the correlation coefficient.

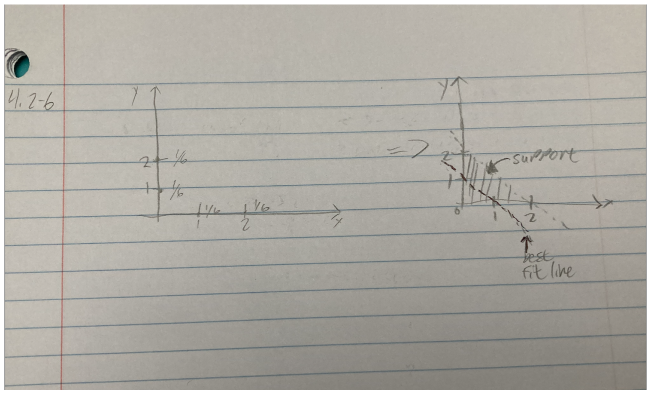
1) $\sigma_X^2 = \sum(x - \mu_x)^2 f(x, y) = (0 - \frac{2}{3})^2(\frac{1}{2}) + (1 - \frac{2}{3})^2(\frac{1}{3}) + (2 - \frac{2}{3})^2(\frac{1}{6}) = \frac{5}{9}$

2) $\sigma_Y^2 = \sum(x - \mu_y)^2 f(x, y) = (0 - \frac{2}{3})^2(\frac{1}{2}) + (1 - \frac{2}{3})^2(\frac{1}{3}) + (2 - \frac{2}{3})^2(\frac{1}{6}) = \frac{5}{9}$

$P = \frac{COV(X,Y)}{\sigma_x}\sigma_y = \frac{-\frac{5}{18}}{\sqrt{\frac{5}{9}}\sqrt{\frac{5}{9}}} = -\frac{1}{2} = -0.5$

Answer: -0.5

4.2-6e) Find the best fitting line and draw it.



4.3-4a) If the total number of offspring is n=400, how is X distributed?

Equally outcomes (R,R)=(R,W)=(W,W)=(W,R)-$\frac{1}{4}$

The offspring has white outcomes: (R,W), (W,W), or (W,R). (R,W)=(W,W)=(W,R)=$\frac{3}{4}$=0.75.

X=number of successes, n=400, p=0.75, then the binomial distribution is $X \sim B(400, 0.75)$

Answer: $X \sim B(400, 0.75)$

4.3-4b) Give values of E(X) and Var(X).

Answer: E(X)=np=400(0.75)=300

Answer: Var(X)=npq=400(0.75)(1-0.75)=75

4.3-4c) Given X=300, how is Y distributed?

Y= number of red-eyed offspring: (R,W) or (W,R) genes with white eyes. With n=300, p=$\frac{2}{3}$, then $g(y|300) \sim B(300, \frac{2}{3}$
Answer: $g(y|300) \sim B(300, \frac{2}{3}$

4.3-4d) Give values of E(Y — X=300) and Var(Y — X=300).

Answer: $E(Y|X = 300) = E(g(y|300)) = 300(\frac{2}{3}) = 200$

Answer: $Var(Y|X = 300) = npq = 300(\frac{2}{3})(1 - \frac{2}{3}) = \frac{200}{3} = 66.67$

4.3-8a) 6-sided die, 30 times rolled independently (6^{30}). X= # ones and Y= # twos What is the joint pdf of X and Y.

f(x,y)=P(X=x, Y=y):

$\binom{30}{x} ways$ => ones falls/chosen

$\binom{30-x}{y} ways$ => ones no falls/ not chosen, but twos falls/chosen

$(6)^{30-x-y} ways$ => where any other numbers other than two and one falls/chosen

$\binom{30}{x}\binom{30-x}{y}4^{30-x-y}$ => total outcomes

Now,

$f(x,y) = P(X = x, Y = y) = \frac{\binom{30}{x}\binom{30-x}{y}}{6^{30}}$

Answer: $f(x,y) = P(X = x, Y = y) = \frac{\binom{30}{x}\binom{30-x}{y}4^{30-x-y}}{6^{30}}$

4.3-8b) Find the conditional pmf of X and Y.

Because $f(x,y) = P(X = x, Y = y) = \frac{\binom{30}{x}\binom{30-x}{y}}{6^{30}}$, then $X, Y \sim B(30, \frac{1}{6})$

So,

$P(Y = y) = \binom{30}{y}(\frac{1}{6})^y(\frac{5}{6})^{30-y} = \binom{30}{y}\frac{5^{30-y}}{6^{30}}$

Now everything together:

$P(X = x, Y = y) = \frac{\binom{30}{x}\binom{30-x}{y}}{\binom{30}{y}}\frac{4^{30-x-y}}{5^{30-y}}$

$= \frac{\frac{30!}{(30-x)!x!}\frac{30-x}{30-x-y}}{\frac{30!}{(30-y)!y!}}(\frac{1}{5})^x(\frac{4}{5})^{30-x-y}$

$= \frac{(30-y)!}{x!(30-y-x)!}(\frac{1}{5}^x(\frac{4}{5})^{30-x-y}$

So X conditioned on Y=y is

$B(30 - y, \frac{1}{5})$ distribution

Answer: $B(30 - y, \frac{1}{5})$ distribution

4.3-8c) Compute $E(X^2 - 4XY + 3Y^2)$.

$E(X^2) = E(Y^2) = \sigma_Y^2 + (E(Y))^2$

$\sigma_Y^2 = 30(\frac{1}{6})(\frac{5}{6}) = \frac{25}{6}$

$(E(Y))^2 = (\frac{30}{6})^2 = 25$

$\sigma_Y^2 + (E(Y))^2 = \frac{25}{6} + 25 = \frac{175}{6}$

Now $P\sigma_X\sigma_Y = COV(X,Y) = E(XY) - E(X)E(Y) = E(XY) - (E(X))^2$, where $\sigma_X\sigma_Y = \sigma_Y^2 = \frac{25}{6}$ and $(E(X))^2 = 25$

So $E(XY) = \frac{25}{6}P + 25$

Using the idea from the book:

$P = -\sqrt{\frac{\frac{1}{6}\frac{1}{6}}{\frac{5}{6}}\frac{5}{6}} = -\frac{1}{5}$

$E(XY) = \frac{25}{6}(-\frac{1}{5}) + 25 = \frac{145}{6}$

Finally, with $E(X^2) = E(Y^2) = \frac{175}{6}$, $E(XY) = \frac{145}{6}$

$E(X^2 - 4XY + 3Y^2) = \frac{175}{6} - 4(\frac{145}{6}) + 3(\frac{175}{6}) = 20$

Answer: 20

4.4-4a) Let $f(x,y) = \frac{3}{2}, X^2 \leq y \leq 1, 0 \leq x \leq 1$ be the joint pdf of X and Y. Find $P(0 \leq X \leq \frac{1}{2})$

First we have to find our marginal of X and Y:

marginal of X:

$f(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{x^2}^{1} \frac{3}{2}dy = \frac{3}{2}(1-x^2), 0 \leq x \leq 1.$

0, otherwise.

marginal of Y:

$f(Y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{0}^{\sqrt{y}} \frac{3}{2}dx = \frac{3}{2}(\sqrt{y}, 0 \leq x \leq 1.$

0, otherwise.

$P(0 \leq X \leq \frac{1}{2}) = \int_{0}^{\frac{1}{2}} f(x)dx = \int_{0}^{\frac{1}{2}} \frac{3}{2}(1-x^2)dx = \frac{11}{16}$

Answer: $\frac{11}{16}$

4.4-4b) Find $P(\frac{1}{2} \leq Y \leq 1)$

$P(\frac{1}{2} \leq Y \leq 1) = \int_{\frac{1}{2}}^{1} f(y)dy = \int_{\frac{1}{2}}^{1} \frac{3}{2}(\sqrt{y})dy = 1 - (\frac{1}{2})^{\frac{3}{2}}$

Answer: $1 - (\frac{1}{2})^{\frac{3}{2}}$

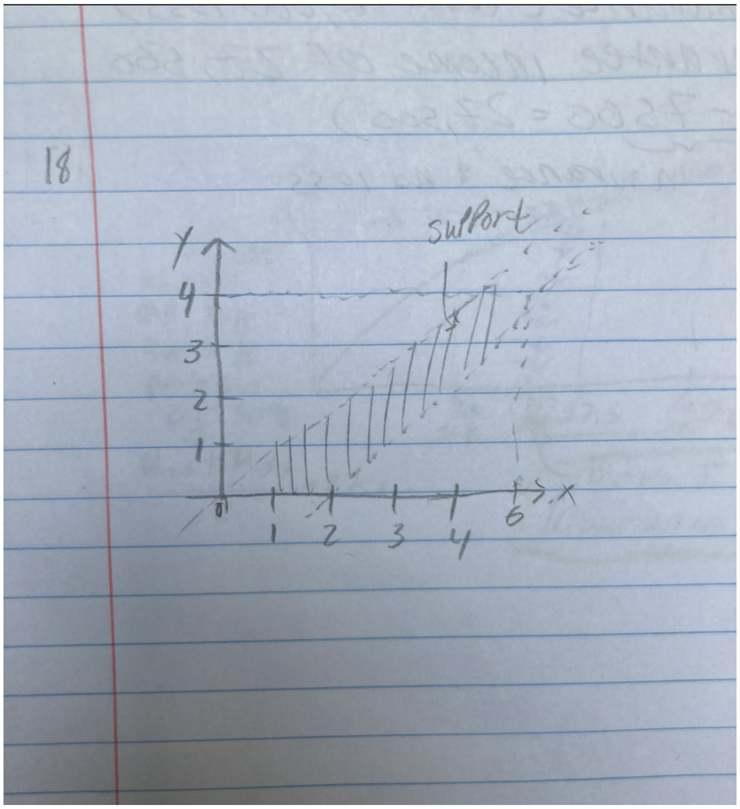
4.4-4c) Find $(X \geq \frac{1}{2}, Y \geq \frac{1}{2})$

$(X \geq \frac{1}{2}, Y \geq \frac{1}{2}) = \frac{5}{8} - -(\frac{1}{2})^{\frac{3}{2}}$

4.4-4d) Are X and Y independent? Why/why not?

Answer: X and Y are not independent, because f(x,y)$\neq$ f(x)f(y).

4.4-18a) $f(x,y) = \frac{1}{8}, 0 \le y \le 4, y \le x \le y+2$ be the joint pdf of X and Y. Sketch the region for which $f(x,y) > 0$.



4.4-18b) Find $f_x(x)$, the marginal pdf of X.

$f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{x-2}^{x} \frac{1}{8}dy = \frac{1}{4}$

Answer: $\frac{1}{4}$

4.4-18c) Find $f_y(y)$, the marginal pdf of X

$f_y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{y}^{y+2} \frac{1}{8}dx = \frac{1}{4}$

Answer: $\frac{1}{4}$

4.4-18d) Determine $h(y|x)$, the conditional of pdf of Y, given that X=x.

$h(y|x) = \frac{f(x,y)}{f_x(x)} = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2}$

Answer: $\frac{1}{2}$

4.4-18e) Determine $g(y|x)$, the conditional of pdf of X, given that Y=y.

$g(y|x) = \frac{f(x,y)}{f_y(y)} = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2}$

Answer: $\frac{1}{2}$

4.4-18f) Compute y=E(Y—x), the conditional mean of Y, given that X=x.

$y = E(Y|x) = \int_{x-2}^{x} y(h(y|x))dy = \int_{x-2}^{x} y(\frac{1}{2})dy = x - 1$

Answer: $x - 1$

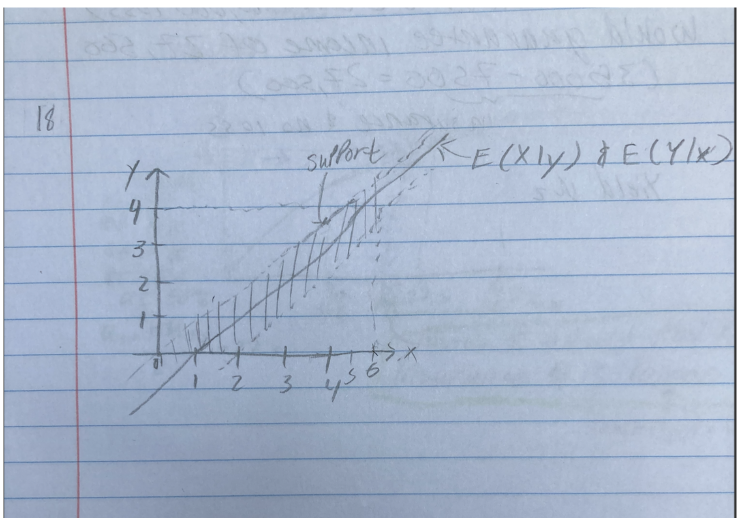
4.4-18f) Compute x=E(X—y), the conditional mean of X, given that Y=y.

$x = E(X|y) = \int_{y}^{y+2} x(g(x|y))dx = \int_{y}^{y+2} x(\frac{1}{2})dx = y + 1$

Answer: $y + 1$

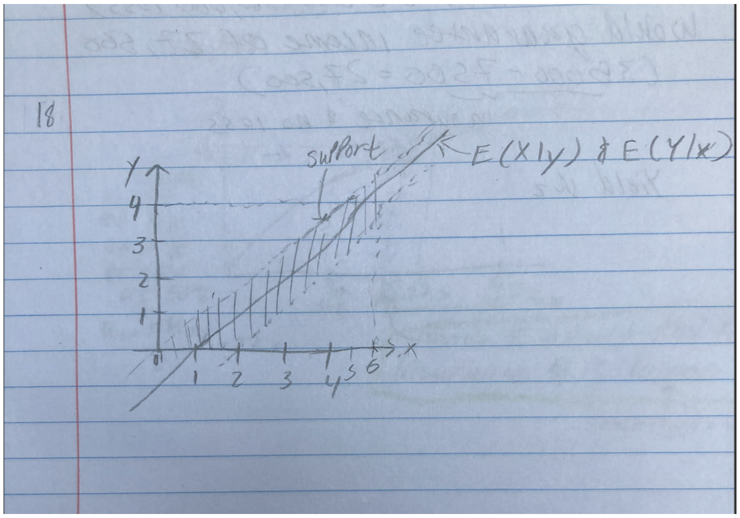
4.4-18h) Graph y=E(Y—x) on sketch in part (a). Is y=E(Y—x) linear?

Answer: y=E(Y—x) is linear.



4.4-18i) Graph x=E(X—y) on sketch in part (a). Is x=E(X—y) linear?

Answer: x=E(X—y) is linear.


18
support
$E(X|y)$ & $E(Y|x)$
Y
4
3
2
1
0
1
2
3
4
5
6
X

4.5-6a) Find $P(19.0 < 26.9)$.

We are given $\mu_X = 22.7, \sigma_X^2 = 17.64$ and $\mu_Y = 22.7, \sigma_Y^2 = 12.25$ and p=0.78.

$z = \frac{x-\mu}{\sigma} = \frac{19.9-22.7}{\sqrt{12.25}} = -.8$

$z = \frac{x-\mu}{\sigma} = \frac{26.9-22.7}{\sqrt{12.25}} = 1.2$

$P(19.0 < 26.9) = P(-.8 < Z < 1.2) = P(Z < 1.2) - P(Z < -.8)$

$= P(Z < 1.2) - P(Z > .8) = P(Z < 1.2) - (1 - P(Z < .8)) = 0.8849 - (1 - .7881) = .673 = 67.3\%$

Answer: 67.3%

4.5-6b) Find E(Y — x).

$E(Y|x) = E(Y|X = x) = \mu_y + p(\frac{\sigma_y}{\sigma_x})(x - \mu_x)$

Answer: $= 22.7 + 0.78(\frac{3.5}{4.2}(x - 22.7) = 22.7 + .65(x - 22.7) = 22.7 + 0.65x - 14.755 = 0.65x + 7.945$

4.5-6c) Find Var(Y—x).

$Var(Y|x) = Var(Y|X = x) = \sigma^2_{Y|x=x} = \sigma_Y^2 - p^2(\sigma_Y^2)$

$= 12.25 - 0.78^2(12.25) = 4.7971$

Answer: Var(Y—X=x)=4.7971

4.5-6d) Find $P(19.9 < Y < 26.9|X = 23)$

Now x=23

$E(Y|X = x) = 0.65x + 7.945$

$Var(Y|X = x) = 4.7971$

$z = \frac{x-\mu}{\sigma} = \frac{19.9-(7.945+(0.65(23))}{\sqrt{4.7971}} \approx -1.37$

$z = \frac{x-\mu}{\sigma} = \frac{26.9-(7.945+(0.65(23))}{\sqrt{4.7971}} \approx 1.83$

$P(19.9 < Y < 26.9|X = 23) = P(-1.37 < Z < 1.83) = P(Z < 1.83) - P(Z < -1.37)$

$= P(Z < 1.83) - P(Z > 1.37) = P(Z < 1.83) - (1 - P(Z < 1.37))$

$= 0.966 - (1 - .915) = 0.881 = 88.1\%$

Answer: 88.1%

4.5-6e) Find $P(19.9 < Y < 26.9|X = 25)$

Now x=25

$E(Y|X = x) = 0.65x + 7.945$

$Var(Y|X = x) = 4.7971$

$z = \frac{x-\mu}{\sigma} = \frac{19.9-(7.945+(0.65(25))}{\sqrt{4.7971}} \approx -1.96$

$z = \frac{x-\mu}{\sigma} = \frac{26.9-(7.945+(0.65(25))}{\sqrt{4.7971}} \approx 1.24$

$P(19.9 < Y < 26.9|X = 25) = P(-1.96 < Z < 1.24) = P(Z < 1.24) - P(Z < -1.96)$

$= P(Z < 1.24) - P(Z > 1.96) = P(Z < 1.24) - (1 - P(Z < 1.96))$

$= 0.8925123 - (1 - 0.9750021) = .8675 = 86.75\%$

Answer: 86.75%

4.5-6f) For x=21, 23, and 25 draw a graph of z=h(y — x).

x=21:
E(Y — X=21)=7.945 +0.65(21)= 21.595

x=23:
E(Y — X=21)=7.945 +0.65(23)= 22.895

x=25:
E(Y — X=21)=7.945 +0.65(25)= 24.195

Graph: