Kapitel 5

5.4) Den diskreta gemensamma fördelningen för (X,Y) ges av

$$f_{XY}(x,y) = \frac{2}{n(n+1)}, \quad 1 \leq y \leq x \leq n, \quad n \in \mathbb{N}^+$$

a) Verifiera att f är en täthet.

1. $f_{XY}(x,y) \geq 0$ ok!

2. $\sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f_{XY}(x,y) = 1$ Summan över hela utfallsrummet ska vara 1.

$$\sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f_{XY}(x,y) = \sum_{x=1}^{n} \sum_{y=1}^{x} \frac{2}{n(n+1)} =$$

$$= \frac{2}{n(n+1)} \sum_{x=1}^{n} \sum_{y=1}^{x} 1 = \frac{2}{n(n+1)} \sum_{x=1}^{n} x = \frac{2}{n(n+1)} \frac{n(n+1)}{2} = 1$$

Ok!

b) Hitta marginal fördelningarna för X och Y.

summera över Y

$$f_X = \sum_{x=-\infty}^{\infty} f_{XY}(x,y) = \sum_{y=1}^{x} \frac{2}{n(n+1)} = \underline{\underline{\frac{2x}{n(n+1)}}}$$

$$f_Y = \sum_{x=-\infty}^{\infty} f_{XY}(x,y) = \sum_{x=y}^{n} \frac{2}{n(n+1)} = \frac{2}{n(n+1)} \sum_{x=1}^{n-y+1} 1 =$$

$$= \underline{\underline{\frac{2}{n(n+1)}(n-y+1)}}$$

c) Är X och Y oberoende?

X och Y oberoende $\Longleftrightarrow f_{XY} = f_X \cdot f_Y$

$$f_X \cdot f_Y = \frac{2x}{n(n+1)} \frac{2}{n(n+1)}(n-y+1) \neq \frac{2}{n(n+1)}$$

X och Y är alltså inte oberoende.

d) Antag att n=5. Hitta

$$P(X \leq 3 \cap Y \leq 2) \qquad n=5$$

$$P(X \leq 3 \cap Y \leq 2) = \sum_{x=1}^{3} \sum_{y=1}^{\min(2,x)} \frac{2}{n(n+1)} =$$

$$= \frac{2}{n(n+1)} \left(\underset{x=1}{1} + \underset{x=2}{2} + \underset{x=3}{2} \right) \underset{n=5}{=} \frac{2}{\cancel{5}(5+1)} \cancel{5} = \frac{2}{6} = \frac{1}{3}$$

Hitta

$$P(X \leq 3) = \sum_{x=1}^{3} f_X(x) = \frac{2}{n(n+1)} \sum_{x=1}^{3} x =$$

$$= \frac{2}{5 \cdot 6} \cdot 6 = \frac{2}{5}$$

Hitta

$$P(Y \leq 2) = \sum_{y=1}^{2} \frac{2(n-(y-1))}{n(n+1)} = \frac{2n + 2(n-1)}{n(n+1)} =$$

$$= \frac{4n-2}{n(n+1)} = \frac{18}{30} = 3/5$$

5.8

Låt X vara temperatur i grader celsius och Y tiden i minuter det tar att starta en dieselmotor.

Antag också

$$f_{XY}(x,y) = c(4x+2y+1) \qquad \begin{array}{l} 0 \le x \le 40 \\ 0 \le y \le 2 \end{array}$$

a) Hitta värdet på c som gör detta till en täthet.

Vi vet att

$$\underline{1} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = c\int_0^2\int_0^{40}(4x+2y+1)\,dx\,dy =$$

$$= c\int_0^2 4\left[\frac{x^2}{2} + (2y+1)x\right]_{x=0}^{x=40} dy =$$

$$= c\int_0^2 3200 + 80y + 40\,dy = c\left[80\frac{y^2}{2} + 3240y\right]_{y=0}^{2} =$$

$$= c(160 + 6480) = c \cdot 6640 \Rightarrow \underline{\underline{c = \frac{1}{6640}}}$$

b) Hitta sannolikheten att en dag är temperaturen över 20 C och det tar minst **en minut att starta.**

Detta kan alltså skrivas som

$P(X > 20 \cap Y \geq 1)$

$$P(X > 20 \cap Y \geq 1) = \int_{20}^{\infty}\int_{1}^{\infty} f_{XY}(x,y)\,dx\,dy =$$

$$= \frac{1}{6640}\int_{20}^{40}\int_{1}^{2}(4x+2y+1)\,dy\,dx = \frac{1}{6640}\int_{20}^{40}\left[y^2+(4x+1)y\right]_1^2 dx =$$

$$= \frac{1}{6640}\int_{20}^{40} 3+4x+1\,dx = \frac{1}{6640}\int_{20}^{40} 4x+4\,dx =$$

$$= \frac{1}{6640}\left[2x^2+4x\right]_{20}^{40} = \frac{1}{6640}\left(2\cdot 1200+4\cdot 20\right) =$$

$$= \frac{2480}{6640} = \underline{0.3735}$$

c) Hitta marginalfödelningen för X och Y.

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \frac{1}{6640}\int_0^2 (4x+2y+1)\,dy =$$

$$= \frac{1}{6640}\left[y^2+(4x+1)y\right]_0^2 = \frac{1}{6640}\left(4+2(4x+1)\right) =$$

$$= \underline{\underline{\frac{8x+6}{6640}}}$$

$$f_Y(y) = \frac{1}{6640}\int_0^{40}(4x+2y+1)\,dx = \frac{1}{6640}\left[2x^2+(2y+1)x\right]_0^{40} =$$

$$= \frac{1}{6640}\left(2\cdot 40^2+80y+40\right) = \underline{\underline{\frac{80y+3240}{6640}}}$$

d) Hitta sannolikheten att det tar minst **1 minut att starta.**

Vi letar alltså efter $P(Y \geq 1)$

$$P(Y \geq 1) = \int_1^{\infty} f_Y(y)\,dy = \int_1^2 \frac{80y+3240}{6640}\,dy =$$

$$= \frac{1}{6640}\left[40y^2+3240y\right]_1^2 = \frac{1}{6640}\left(40(4-1)+3240\cdot 1\right) =$$

$$= \underline{\underline{0.506\ldots}}$$

e) Hitta sannolikheten att det är varmare än 20 C.

Alltså hitta

$$P(X > 20) = \int_{20}^{\infty} f_X(x)\,dx = \frac{1}{6640}\int_{20}^{40} 8x + 6\,dx =$$

$$= \frac{1}{6640}\left[8\frac{x^2}{2} + 6x\right]_{20}^{40} = \frac{1}{6640}\left(\frac{8}{2}(40^2 - 20^2) + 6 \cdot 20\right) =$$

$$= \frac{1}{6640}\left(4(1600 - 400) + 120\right) = \frac{4920}{6640} = \underline{0.741}$$

f) är X och Y oberoende?

Om oberoende så gäller $f_{XY} = f_X \cdot f_Y$

$$f_X \cdot f_Y = \frac{1}{6640^2}(80y + 3240)(8x + 6) \neq f_{XY}$$

Alltså INTE oberoende.

5.10

Gemensam fördelning för X och Y ges av:

$$f_{XY}(x,y) = \begin{cases} \frac{x^3y^3}{16}, & 0 \le x \le 2, 0 \le y \le 2 \\ 0, & \text{annars} \end{cases}$$

a) Hitta marginal fördelningen för X och Y

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \frac{x^3}{16}\int_0^2 y^3 dy = \frac{x^3}{16}\left[\frac{y^4}{4}\right]_0^2 =$$

$$= \frac{x^3}{16}\frac{16}{4} = \underline{\underline{\frac{x^3}{4}}}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \left\{\begin{matrix}\text{symetri} \\ \text{(Samma som X)}\end{matrix}\right\} = \underline{\underline{\frac{y^3}{4}}}$$

b) Är X och Y oberoende?

Oberoende om $f_{XY} = f_X \cdot f_Y$

$$f_{XY} = \frac{x^3y^3}{16}$$

$$f_X \cdot f_Y = \frac{x^3}{4} \cdot \frac{y^3}{4} = \frac{x^3y^3}{16}$$

Ja dessa är oberoende!

c) Hitta P(X=<1)

$$P(X \le 1) = \int_{-\infty}^{1} f_X(x)\,dx = \frac{1}{4}\int_0^1 x^3 dx = \underline{\underline{\frac{1}{16}}}$$

d) Vad är P(X=<1) om vi vet att y=1?

Oberoende ger oss

$P(X \leq 1 \mid Y=1) = P(X \leq 1) = \frac{1}{16}$

5.16

Låt X=# syntaxfel och Y=# logik-fel som fås från ett program. Den gemensamma fördelningen fås som

X/Y	0	1	2	3
0	0.400	0.100	0.020	0.005
1	0.300	0.040	0.010	0.004
2	0.040	0.010	0.009	0.003
3	0.009	0.008	0.007	0.003
4	0.008	0.007	0.005	0.002
5	0.005	0.002	0.002	0.001

a) X och Y är inte oberoende. Säger detta något om kovariancen?

Nej!

Oberoende $\Rightarrow$ Kovarans 0.

Doch

Inte Oberoende $\not\Rightarrow$ Kovarians skild från 0

Kan alltså inte säga något i detta fall.

b) Hitta Väntevärde för X, Y och XY. Hitta även kovariansen och ge en kort förklaring.

Väntevärde

Y

$P(Y=k)$ ges av summan av kolumnen.

$$f_Y(y) = \sum_x f(x,y)$$

y	0	1	2	3
$f(y)$	0.762	0.167	0.053	0.018

$$E[Y] = \sum y f(y) = 0 \cdot 0.762 + 1 \cdot 0.167 + 2 \cdot 0.053 + 3 \cdot 0.018 =$$

$$= 0.327$$

X

$P(X=k)$ ges av summan av raderna

$$f_X(x) = \sum_y f(x,y)$$

x	0	1	2	3	4	5
$f(x)$	0.525	0.354	0.062	0.027	0.022	0.01

$$E[X] = \sum x f(x) = 0 \cdot 0.525 + 1 \cdot 0.354 + 2 \cdot 0.062 + 3 \cdot 0.027 +$$

$$+ 4 \cdot 0.022 + 5 \cdot 0.01 = 0.697$$

XY

$$E[XY] = \sum_{x,y} xy\, f(x,y) =$$

produkten viktad mot sannolikheten för den produkten för alla x och Y.

$$= \left\{ \begin{array}{c|cccc} x\backslash y & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 0 & 2 & 4 & 6 \\ 3 & 0 & 3 & 6 & 9 \\ 4 & 0 & 4 & 8 & 12 \\ 5 & 0 & 5 & 10 & 15 \end{array} \leftarrow X \cdot Y \right\} = \ldots = \underline{\underline{0.376}}$$

Kovarians

$$\operatorname{cov}(X, Y) = E[XY] - E[X] \cdot E[Y] =$$

$$= 0.376 - 0.697 \cdot 0.327 \approx \underline{\underline{0.148}} > 0$$

En positiv kovarians betyder att variablerna är positivt korrelerade dvs. om den ena ökar tenderar den andra att också öka. I vårat fall betyder det om ena typen av fel ökar tenderar den andra att göra det samma.

öka/minska (De följer varandra)

c) Hitta väntevärdet för X+Y och tolka detta.

Vänte värdet är linjärt dvs.

$$E[aX+bY]=aE[X]+bE[Y]$$

$$\Rightarrow E[X+Y]=E[X]+E[Y]=0.697+0.327=$$

$$=1.024$$

Tolkning

Förväntat antal fel i kod.

5.20

Beteckna återigen X som temperatur för en motor och Y som starttiden för motorn.

$$f_{XY}(x,y) = \frac{1}{6640}(4x + 2y + 1), \quad 0 \le x \le 40; \; 0 \le y \le 2$$

som tidigare

$$f_X(x) = \frac{1}{6640}(8x + 6), \quad f_Y(y) = \frac{1}{6640}(80y + 3240)$$

a) Från ett fysikaliskt perspektiv vilket värde borde Cov(X,Y) ha?

Lägre temp. $\Rightarrow$ Större starttid.

Högre temp. $\Rightarrow$ mindre startid.

Dvs. de går i motsatt riktning alltså negativt korolerade.

$$\Rightarrow Cov(X,Y) < 0$$

b) hitta E[X], E[Y], E[XY] samt Cov(X,Y).

som i uppgif 5.8

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \frac{1}{6640}\int_0^{40}(8x^2+6x)\,dx =$$

$$= \frac{1}{6640}\left[8\frac{x^3}{3}+6\frac{x^2}{2}\right]_0^{40} = \frac{1}{6640}\left(\frac{8\cdot 40^3}{3}+\frac{6\cdot 40^2}{2}\right) = \underline{\underline{26.43}}$$

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy = \frac{1}{6640}\int_0^{2}(80y^2+3240y)\,dy =$$

$$= \frac{1}{6640}\left(80\frac{2^3}{3}+3240\frac{y^2}{2}\right) = \underline{\underline{1.008}}$$

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)\,dx\,dy =$$

$$= \frac{1}{6640}\int_0^{40} x \int_0^{2} y(4x+2y+1)\,dy\,dx =$$

$$= \frac{1}{6640}\int_0^{40} x \int_0^{2} 4xy+2y^2+y\,dy\,dx =$$

$$= \frac{1}{6640}\int_0^{40} x\left[2xy^2+\frac{2}{3}y^3+\frac{y^2}{2}\right]_0^{2} dx =$$

$$= \frac{1}{6640}\int_0^{40} 8x^2+\frac{16}{3}x+2x\,dx = \frac{1}{6640}\left[\frac{8}{3}x^3+\frac{8}{3}x^2+x^2\right]_0^{40} =$$

$= \underline{\underline{26{,}586}}$

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

$= 26{,}586 - 26{,}43 \cdot 1{,}008 = -0.052 < 0$

Som vi väntade oss.

5.22

Låt (X,Y) ha en gemensam fördelning

$$f_{XY}(x,y) = \begin{cases} \frac{x^3 y^3}{16}, & 0 \leq x \leq 2,\ 0 \leq y \leq 2 \\ 0, & \text{annars} \end{cases}$$

$$f_X(x) = \frac{x^3}{4} \quad f_Y(y) = \frac{y^3}{4}$$

Hitta kovariansen

Hitta $Cov(X,Y)$

X och Y är oberoende som visat i 5.10

$$\Rightarrow \quad Cov(X,Y) = E[X,Y] - E[X] \cdot E[Y] = (\text{sats } 5.2.2) =$$

$$= E[X] \cdot E[Y] - E[X] \cdot E[Y] = 0$$

5.24
Visa att Cov(X,Y)=E[XY]-E[X]E[Y].

$$\mathrm{Cov}(X,Y) \underset{\text{Def}}{:=} E\left[(X-E(X))(Y-E(Y))\right] = \left\{\begin{matrix}\text{utveckla}\\ \text{produkt}\end{matrix}\right\} =$$

$$= E\left[XY - XE[Y] - YE[X] + E[X]E[Y]\right] = \{\text{Linjäritet}\} =$$

$$= E[XY] - E\left[XE[Y]\right] - E\left[YE[X]\right] + E\left[E[X]E[Y]\right] =$$

$$= \left\{\begin{matrix}\text{Notera här} \\ \text{att } E[X] \text{ o } E[Y] \\ \text{är konstanter} \\ \text{kan alltså flyttas ut.}\end{matrix} \quad \begin{matrix} E[X]=\mu_X \\ E[Y]=\mu_Y\end{matrix}\right\} = E[XY] - E[Y]E[X] - E[X]E[Y] + E[X]E[Y] =$$

$$= E[XY] - 2E[Y]E[X] + E[X]E[Y] =$$

$$= \underline{\underline{E[XY] - E[X]\cdot E[Y]}} \quad \text{V.S.V}$$

5.26

Visa att om X och Y är oberoende så gäller att Var(X+Y) = Var(X) + Var(Y).

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

$$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 =$$

$$= E[X^2] + 2E[XY] + E[Y^2] - (E[X]^2 + 2E[X]E[Y] + E[Y]^2) =$$

$$= \underbrace{E[X^2] - E[X]^2}_{Var[X]} + \underbrace{E[Y^2] - E[Y]^2}_{Var[Y]} + 2\underbrace{(E[XY] - E[X] \cdot E[Y])}_{Cov(X,Y)} =$$

oberoende!

$$= Var(X) + Var(Y) + 2(\cancel{E[X]E[Y]} - \cancel{E[X] \cdot E[Y]}) =$$

$$= Var(X) + Var(Y).$$

V.S.V.

5.30

Hitta E[X^2], E[Y^2], Var(X), Var(Y) samt "pearson coefficient of correlation) rho_XY.

X och Y har samma fördelning som i 5.16. Förväntar vi oss en tydlig linjär trend mellan X och Y?

från 5.16 kan vi hämta fördelningen för X och Y och då räkna väntevärdena.

$$E[X^2] = \sum_x x^2 f(x) = 1.447$$

$$E[Y^2] = \sum_y y^2 f(y) = 0.541$$

<u>5.16</u> $\Rightarrow E[X] = 0.697,\ E[Y] = 0.327$

$$\mathrm{cov}(X, Y) = 0.148$$

$$\mathrm{Var}(X) = E[X^2] - E[X]^2 = 1.447 - 0.697^2 = 0.96119$$

$$\mathrm{Var}(Y) = E[Y^2] - E[Y]^2 = 0.541 - 0.327^2 = 0.434$$

<u>Pearson coefficient of correlation ρ_{XY}</u>

$$\rho_{XY} = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}} = \frac{0.148}{\sqrt{0.96 \cdot 0.43}} = 0.229$$

Correlation coefficient ρ_{XY} mäter linjäritet

$\rho_{XY} \in [-1, 1]$

Ett ρ med $|\rho|$ nära 1 indikerar ett linjärt beroende. I vårat fall $|\rho| = 0.23$ ger det oss indikation på linjärt samband.

5.32

Hitta E[X^2], E[Y^2], Var(X), Var(Y) samt correlation coefficent för X och Y med fördelning från uppg. 20 motor problemet.

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \frac{1}{6640}\int_0^{40}(8x^3+6x^2)\,dx =$$

$$= \frac{1}{6640}\left[2x^4+2x^3\right]_0^{40} = \frac{1}{6640}(5248000) = \underline{\underline{790,4}}$$

$$E[Y^2] = \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy = \frac{1}{6640}\int_0^{2}(80y^3+3240y^2)\,dy =$$

$$= \frac{1}{6640}\left[20y^4+\frac{3240}{3}y^3\right]_0^{2} = \underline{\underline{1,35}}$$

$$Var(X) = E[X^2]-E[X]^2 = 790.3-26.4^2 = 92,04$$

$$Var(Y) = E[Y^2]-E[Y]^2 = 1.35-1.008^2 = 0.33$$

Correlation Coefficient

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)\,var(Y)}} = \frac{-0.052}{\sqrt{92,04\cdot 0.33}} =$$

$$= \underline{\underline{-0.00932}}$$

$|\rho_{XY}|$ mycket litet $\Rightarrow$ "$\overbrace{\text{Inget}}^{\text{näterbart}}$" samband.

5.40

Betrakta exempel 5.1.4. I den uppgiften är X = kalcium nivå i blodet och Y = kolesterol nivå i blodet. X och Y har den gemensamma fördelningen

$$f_{XY}(x,y) = \frac{1}{240} \qquad \begin{matrix} 8.5 \leq x \leq 10.5 \\ 120 \leq y \leq 240 \end{matrix}$$

Visar

$$f_X(x) = \int_{120}^{240} \frac{1}{240}\,dy = \frac{1}{2}$$

$$f_Y(y) = \int_{8.5}^{10.5} \frac{1}{240}\,dx = \frac{1}{120}$$

a) Hitta f_X|y(x)

Vi letar alltså efter fördelningen för X givet ett Y. Detta kan skrivas som

$$f_{X|Y}(x) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{1/240}{1/120} = \frac{1}{2}$$

Kan Notera att $f_{X|Y} = f_X$. Detta kan tolkas som att kolestorol nivån säger inget om kalcium nivån.

b) Är f_Y|x =f_Y?

$$f_{Y|X}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1/240}{1/2} = \frac{1}{120} = f_Y$$

<u>Ja.</u>

c) Hitta regressions kurvan av X på Y och Y på X. Är dessa linjära?

Regressions kurvan av X på Y är väntevärdet av X givet ett $Y=y$.
Låt oss beteckna det.

$\mu_{X|y}$

Kan räknas ut som

$$\mu_{X|y} = E[X|y] = \int_{-\infty}^{\infty} x f_{X|y}(x)dx =$$

$$= \int_{8.5}^{10.5} x \cdot \frac{1}{2} dx = \left[\frac{x^2}{4}\right]_{8.5}^{10.5} = \frac{1}{4}(10.5^2 - 8.5^2) = \underline{\underline{9.5}}$$

$$\mu_{Y|x} = E[Y|x] = \int_{-\infty}^{\infty} y f_{Y|x}(y)dy = \dots = \frac{240+120}{2} = \underline{\underline{180}}$$

(Notera att generelt så är $\mu_{X|y}(y)$ funktion av y och $\mu_{Y|x}$ funktion av x.)

Är dessa linjära.
De är kostanta funktioner och konstanta funktioner är linjära. så Ja.

$ax+b$

$a=0$

5.41

Låt (X,Y) vara slumpvariabeler som i övning 5.4. Det vill säga ha den gemensamma fördelningen

$$f_{XY}(x,y) = \frac{2}{n(n+1)} \quad 1 \leq y \leq x \leq n$$

$$f_X(x) = \frac{2x}{n(n+1)}, \quad f_Y(y) = \frac{2}{n(n+1)}(n-(y-1))$$

a) Hitta regressionen av X på Y. Är denna linjär?

$$f_{X|Y} = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{\frac{\cancel{2}}{\cancel{n(n+1)}}}{\frac{\cancel{2}}{\cancel{n(n+1)}}(n-(y-1))} = \underline{\underline{\frac{1}{n-(y-1)}}}$$

$, \; y \leq x \leq n$

$$E[X|Y] = \sum_{x=y}^{n} x \cdot f_{X|Y}(x) = \sum_{x=y}^{n} \frac{x}{n-(y-1)} =$$

$$= \frac{1}{n-(y-1)}\left(\sum_{x=1}^{n} x - \sum_{x=1}^{y-1} x\right) = \frac{1}{n-(y-1)}\left(\frac{n(n+1)}{2} - \frac{(y-1)y}{2}\right) =$$

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

$$= \frac{(n+y)\cancel{(n-(y-1))}}{2\cancel{(n-(y-1))}} = \frac{n+y}{2}.$$ Den är linjär.

b) Antag att n=10 och hitta väntevärdet av X givet att Y=4.

Använder vad vi vet från a).

$$E[X|4] = \frac{10+4}{2} = \underline{\underline{7}}$$

c) Hitta regressionen av Y på X och avgör om den är linjär.

$$f_{Y|X} = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\frac{2}{n(n+1)}}{\frac{2x}{n(n+1)}} = \frac{1}{x}, \ 1 \le y \le x$$

$$\mu_{Y|x} = E[Y|X=x] = \sum_{y=1}^{x} y \cdot \frac{1}{x} = \frac{1}{x} \sum_{y=1}^{x} y = \frac{1}{x} \frac{x(x+1)}{2} =$$

$$= \frac{x+1}{2}.$$ Den är linjär.

d) Antag att n=10 och hitta väntevärdet av Y givet att X=4.

$$E[Y|X=4] = \frac{x+1}{2} = \frac{4+1}{2} = \underline{\underline{\frac{5}{2}}}$$

5.42

Låt (X,Y) ha den gemensamma fördelningen

$$f_{XY}(x,y) = \frac{1}{x}, \quad 0<y<x<1$$

a) Hitta regressionen av X på Y.

Kommer behöva föreIningen av X givet Y=y.

$$f_{X|Y} = \frac{f_{XY}(x,y)}{f_Y(y)}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_y^1 \frac{1}{x}\,dx = \left[\ln x\right]_y^1 =$$

$$= \ln(1) - \ln(y) = -\ln(y)$$

$$\Rightarrow f_{X|Y} = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{1/x}{-\ln(y)} = -\frac{1}{x\ln(y)}$$

$$\Rightarrow \mu_{X|Y} = \int_y^1 x\, f_{X|Y}\,dx = -\int_y^1 \frac{1}{\ln(y)}\,dx = \frac{-(1-y)}{\ln(x)} =$$

$$= \frac{y-1}{\ln(y)}.$$ INTE linjärt.

b) Hitta väntevärdet av X givet Y=0.5.

$$\mu_{X|y=0.5} = E(X \mid y=0.5) \overset{a}{=} \frac{0.5-1}{\ln(0.5)} =$$

$$= \frac{-1}{2\ln(1/2)} \approx 0.7213$$

c) Hitta regressionen av Y på X.

$$\mu_{Y|x} = E[Y|x]$$

$$f_{Y|x} = \frac{f_{XY}(x,y)}{f_X(x)}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \{y<x\} = \int_0^x \frac{1}{x}\,dy = \underline{1}$$

$$f_{Y|x} = \frac{1}{x}$$

$$E[Y|x] = \int_0^x y \cdot \frac{1}{x}\,dy = \frac{1}{x}\left[\frac{y^2}{2}\right]_0^x = \underline{\underline{\frac{x}{2}}}.$$ Den är linjär.

d) Hitta väntevärdet av Y givet x=0.75.

$$\mu_{Y|x=0.75} \underset{c)}{=} \frac{0.75}{2} = \underline{\underline{0.375}}$$

5.48

Låt X och Y vara kontinuerliga slump variabler med den gemensamma fördelningen f_XY. Låt sedan U = X+Y. Visa att f_U ges av

$$f_U(z) = f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_{X,Y}(z-y, y)\,dy$$

<u>Lösning</u>

$$f_{X+Y}(z) = \frac{d}{dz} P(X+Y \leq z)$$

$$= \frac{d}{dz} \iint_{\{(x,y)\in \mathbb{R}^2;\ x+y\leq z\}} f_{X,Y}(x,y)\,dx\,dy$$

$$= \frac{d}{dz} \int_{y=-\infty}^{y=+\infty} \left[\int_{x=-\infty}^{x=z-y} f_{X,Y}(x,y)\,dx \right] dy$$

$$= \int_{y=-\infty}^{y=+\infty} \frac{d}{dz} \left[\text{———— } '' \text{ ————} \right] dy$$

$$= \int_{y=-\infty}^{+\infty} f_{X,Y}(z-y, y)\,dy \qquad \text{V.S.V.}$$