MVE090 LP4 2021 Föreläsning 7-8

Författare: examinator Patrik Albin (=jag)

Kapitel 5. "Joint Distributions"

I kapitel 3-4 räknade jag med stokastiska variabler, antingen en i taget eller annars två eller flera oberoende variabler. Det nya i detta kapitel är att stokastiska variabler kan bero av varandra. Och då räcker det inte veta allt om varje variabel var för sig utan man måste även veta hur de beror av varandra vilket är en ny svårighet.

Definition 5.1. Om X och Y är diskreta stokastiska variabler definieras deras gemensamma <u>frekvensfunktion</u> (engelska "joint probability density function") som

$$f_{X,Y}(x,y) = \mathbf{P}(X=x, Y=y) = \mathbf{P}(\{X=x\} \cap \{Y=y\}).$$

Mao. är $f_{X,Y}(x,y)$ sannolikheten att utfallet $\omega \in S$ av slumpförsöket är sådant att både $X(\omega)=x$ och $Y(\omega)=y$ för möjliga värden x och y för X respektive Y.

Sats 5.2. För diskreta variabler X och Y med frekvensfunktion $f_{X,Y}(x,y)$ är

1. $f_{X,Y}(x,y) \geq 0$ för alla x, y,
2. $\sum_{\text{alla } x} \sum_{\text{alla } y} f_{X,Y}(x,y) = 1$,
3. $\mathbf{P}((X,Y) \in C) = \sum\sum_{\text{alla } (x,y) \in C} f_{X,Y}(x,y)$ för $C \subseteq \mathbb{R}^2$,
4. $f_X(x) = \sum_{\text{alla } y} f_{X,Y}(x,y)$ för alla x och $f_Y(y) = \sum_{\text{alla } x} f_{X,Y}(x,y)$ för alla y.

Även om $f_X(x)$ och $f_Y(y)$ är bestämda av $f_{X,Y}(x,y)$ enligt 4 så är det ej tvärtom: man kan i allmänhet ej beräkna $f_{X,Y}(x,y)$ med kännedom om endast $f_X(x)$ och $f_Y(y)$ ty då saknar man informationen om hur X och Y beror av (påverkar) varandra.

Bevis. 1-3. Helt analogt med beviset av 1-3 i sats 3.5.

4. Överkurs. Följer av ta $C = \{x\} \times (-\infty, \infty)$ respektive $C = (-\infty, \infty) \times \{y\}$ i 3. $\square$

Definition 5.3. Om X och Y är stokastiska variabler definieras deras gemensamma <u>fördelningsfunktion</u> (engelska "joint cummulative distribution function") som

$$F_{X,Y}(x,y) = \mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(\{X \leq x\} \cap \{Y \leq y\}) \quad \text{för alla } x, y.$$

Definition 5.4. Om X och Y är kontinuerliga stokastiska variabler definieras deras gemensamma <u>frekvensfunktion</u> (engelska "joint probability density function")

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) \quad \text{för alla } x,y.$$

Sats 5.5. För två kontinuerliga variabler X och Y är

1. $f_{X,Y}(x,y) \geq 0$ för alla x,y,
2. $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$,
3. $F_{X,Y}(x,y) = \int\limits_{u=-\infty}^{u=x} \int\limits_{v=-\infty}^{v=y} f_{X,Y}(u,v)\,dudv$ för alla x,y,
4. $\mathbf{P}(a < X \leq b, c < Y \leq d) = F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + F_{X,Y}(a,c)$,
5. $\mathbf{P}((X,Y) \in C) = \iint\limits_{(x,y)\in C} f_{X,Y}(x,y)\,dxdy$ för $C \subseteq \mathbb{R}^2$,
6. $F_{X,Y}(-\infty,y) = F_{X,Y}(x,-\infty) = 0$, $F_{X,Y}(x,\infty) = F_X(x)$ och $F_{X,Y}(\infty,y) = F_Y(y)$,
7. $f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$ för alla x och $f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$ för alla y.

Bevis. 1-3. Helt analogt med beviset av 1-3 i sats 4.3.

4. Överkurs. Använd

$$\mathbf{P}(X \leq b, c < Y \leq d) = \mathbf{P}(X \leq b, Y \leq d) - \mathbf{P}(X \leq b, Y \leq c) = F_{X,Y}(b,d) - F_{X,Y}(b,c)$$

först som den är och sedan med b utbytt mot a insatt i

$$\mathbf{P}(a < X \leq b, c < Y \leq d) = \mathbf{P}(X \leq b, c < Y \leq d) - \mathbf{P}(X \leq a, c < Y \leq d).$$

5. Överkurs. Använd 4 till visa 5 analogt med beviset av 5 i sats 4.3.

6. Följer direkt av definition 5.3.

7. Med utnyttjande av de två sista likheterna i 6 följer 7 av först ta $y = \infty$ och derivera map. x och sedan ta $x = \infty$ och derivera map. y i 3. $\square$

Definition 5.6. Två stokastiska variabler X och Y är oberoende om

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\,\mathbf{P}(Y \in B) \quad \text{för alla } A,B \subseteq \mathbb{R}.$$

Sats 5.7. Två stokastiska variabler X och Y är oberoende om och endast om

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \text{för alla } x,y.$$

Bevis. (KONTINUERLIGT) Om $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ ger 5 i satserna 4.3 och 5.5 att

$$\mathbf{P}(X\in A, Y\in B) = \int_{x\in A}\int_{y\in B} f_{X,Y}(x,y)\,dxdy = \int_{x\in A}\int_{y\in B} f_X(x)f_Y(y)\,dxdy = \mathbf{P}(X\in A)\,\mathbf{P}(Y\in B).$$

Om omvänt X och Y är oberoende är enligt definitionerna 5.3-5.4, 5.6 och 4.2

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x\partial y}\,\mathbf{P}(X\le x, Y\le y) = \frac{\partial^2}{\partial x\partial y}\,\mathbf{P}(X\le x)\,\mathbf{P}(Y\le y) = f_X(x)f_Y(y). \qquad \square$$

Exempel 5.1. (Example 5.1.5 M&A) Låt X och Y vara kontinuerliga stokastiska variabler med $f_{X,Y}(x,y) = c/x$ för $27 \le y \le x \le 33$[1] för en konstant $c > 0$. Bestäm c, $f_X(x)$ och $f_Y(y)$. Är X och Y oberoende?

Lösning. Jag bestämmer c mha. 2 i sats 5.5:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = \int_{x=27}^{x=33}\int_{y=27}^{y=x} \tfrac{c}{x}\,dxdy = c\int_{x=27}^{x=33}(1-\tfrac{27}{x})\,dx = c\,(6-27\ln(\tfrac{33}{27})).$$

Enligt 7 i sats 5.5 är vidare

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{y=27}^{y=x} \tfrac{c}{x}\,dy = c\,(1-\tfrac{27}{x}) \quad \text{för } x\in[27,33],$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_{x=y}^{x=33} \tfrac{c}{x}\,dx = c\ln(\tfrac{33}{y}) \quad \text{för } y\in[27,33].$$

Eftersom $f_{X,Y}(x,y) \ne f_X(x)f_Y(y)$ är X och Y enligt sats 5.7 ej oberoende.

Exempel 5.2. (Tal 3 20160405) Beräkna $\mathbf{P}(X > Y)$ för X och Y oberoende med X standard normalfördelad och Y exponentialfördelad med parameter 1.

Lösning. Eftersom enligt sats 5.7 och definitionerna 4.9-4.11

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}\,\mathrm{e}^{-y} \quad \text{för } x\in\mathbb{R} \text{ och } y>0$$

och händelsen $\{X > Y\}$ betyder att $(X,Y)\in\{(x,y)\in\mathbb{R}^2 : x>y\}$ ger 5 i sats 5.5

$$\begin{aligned}
P(X>Y) &= \iint_{\{(x,y)\in\mathbb{R}^2\,:\,x>y\}} f_{X,Y}(x,y)\,dxdy \\
&= \int_{x=0}^{x=\infty}\int_{y=0}^{y=x} \tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}\,\mathrm{e}^{-y}\,dydx \\
&= \int_0^{\infty} \tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}\,(1-\mathrm{e}^{-x})\,dx = \tfrac{1}{2} - \sqrt{\mathrm{e}}\int_0^{\infty}\tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-(x+1)^2/2}\,dx.
\end{aligned}$$

Sats 5.8. För två stokastiska variabler X och Y och en funktion $h:\mathbb{R}^2\to\mathbb{R}$ är

$$\mathbf{E}(h(X,Y)) = \begin{cases} \sum_{\text{alla } x}\sum_{\text{alla } y} h(x,y)f_{X,Y}(x,y) & \text{för } X \text{ och } Y \text{ diskreta} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f_{X,Y}(x,y)\,dxdy & \text{för } X \text{ och } Y \text{ kontinuerliga} \end{cases}.$$

[1] Notera de konstiga siffrorna – en matematiker hade valt $0 \le y \le x \le 1$.

Bevis. Dubbel överkurs. I diskreta fallet är $h(X,Y)$ en vanlig diskret variabel och beviset ekvivalent med det av sats 3.7. I kontinuerliga fallet approximeras X och Y uppifrån med diskreta variabler som i beviset av sats 4.5 och sedan resoneras som där. $\square$

Exempel 5.1. (FORTS.) För beräkna $\mathbf{E}(X)$ kan jag utnyttja det tidigare gjorda

$$\mathbf{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_{x=27}^{x=33} x\,c\,(1-\tfrac{27}{x})\,dx = \ldots = 18\,c.$$

Men om detta varit ogjort kunde jag istället använt sats 5.8 med $h(x,y)=x$:

$$\mathbf{E}(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f_{X,Y}(x,y)\,dxdy = \int_{y=27}^{y=33}\int_{x=y}^{x=33} x\,\tfrac{c}{x}\,dxdy = \ldots = 18\,c.$$

Jag kan nu göra de tidigare utelämnade bevisen av 3 i sats 3.8 och 2 i sats 4.6:

Sats 5.9. För stokastiska variabler $X_1,\ldots,X_n$ och reella tal $a_1,\ldots,a_n$ är

$$\mathbf{E}\Big(\sum_{k=1}^{n} a_k X_k\Big) = \sum_{k=1}^{n} a_k\,\mathbf{E}(X_k)$$

Bevis. Överkurs. Det räcker visa 3 i sats 3.8 och 2 i sats 4.6 ty sedan kan dessa användas upprepat tillsammans med 2 i sats 3.8 och 1 i sats 4.6 till att plocka ut en term $\mathbf{E}(a_k X_k) = a_k\,\mathbf{E}(X_k)$ i taget ur vänsterledets väntevärde till högerledets summa. För visa 2 i sats 4.6 tar jag $h(x,y)=x+y$ i sats 5.8 och får mha. 7 i sats 5.5

$$\begin{aligned}\mathbf{E}(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\,f_{X,Y}(x,y)\,dxdy\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\,f_{X,Y}(x,y)\,dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\,f_{X,Y}(x,y)\,dxdy\\ &= \int_{-\infty}^{\infty} x\,f_X(x)\,dx + \int_{-\infty}^{\infty} y\,f_Y(y)\,dy = \mathbf{E}(X)+\mathbf{E}(Y).\end{aligned}$$ $\square$

Definition 5.10. Kovariansen mellan två stokastiska variabler X och Y definieras

$$\mathbf{Cov}(X,Y) = \mathbf{E}((X-\mu_X)(Y-\mu_Y)).$$

Kovariansen är ett enkelt men mycket använt (speciellt i teknik etc.) mått på hur mycket X och Y beror på varandra. För avgöra vad som är stort eller litet beroende normaliseras kovariansen till korrelationskoefficient, se definition 5.14 nedan.

Notera det triviala men ändå viktiga faktumet att $\mathbf{Var}(X) = \mathbf{Cov}(X,X)$.

Definition 5.11. Om $\mathbf{Cov}(X,Y) = 0$ säges X och Y vara okorrelerade.

Sats 5.12. Om X och Y är oberoende är de okorrelerade.

Omvändningen till sats 5.11 är inte sann (alls!): det är lätt konstruera stokastiska variabler X och Y som är fullständigt beroende men ändå har $\mathbf{Cov}(X,Y)=0$.

Överkurs. Visa att om X är Uni$[-\pi,\pi]$ och $Y=\cos(X)$ så är $\mathbf{Cov}(X,Y)=0$.

Bevis. (Kontinuerligt) Enligt sats 5.7 och 5.8 (samt definition 4.4) är

$$\begin{aligned}\mathbf{Cov}(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)\,(y-\mu_Y)f_{X,Y}(x,y)\,dxdy\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)\,(y-\mu_Y)f_X(x)f_Y(y)\,dxdy\\ &= \Big(\int_{-\infty}^{\infty}(x-\mu_X)f_X(x)\,dx\Big)\Big(\int_{-\infty}^{\infty}(y-\mu_Y)f_Y(y)\,dy\Big)\\ &= (\mathbf{E}(X)-\mu_X)\,(\mathbf{E}(Y)-\mu_Y)=0\cdot 0=0.\end{aligned}$$ □

Sats 5.9 har en viktig motsvarighet för kovarians och varians:

Sats 5.13. För variabler $X_1,\ldots,X_m,Y_1,\ldots,Y_n$ och tal $a_1,\ldots,a_m,b_1,\ldots,b_n$ är

1. $\mathbf{Cov}\Big(\sum_{i=1}^m a_iX_i,\sum_{j=1}^n b_jY_j\Big)=\sum_{i=1}^m\sum_{j=1}^n a_ib_j\,\mathbf{Cov}(X_i,Y_j)$,
2. $\mathbf{Var}\Big(\sum_{i=1}^m a_iX_i\Big)=\sum_{i=1}^m\sum_{j=1}^m a_ia_j\,\mathbf{Cov}(X_i,X_j)$,
3. $\mathbf{Var}\Big(\sum_{i=1}^m a_iX_i\Big)=\sum_{i=1}^m a_i^2\,\mathbf{Var}(X_i)$ om $X_1,\ldots,X_m$ är okorrelerade.

Bevis. Överkurs. 1. Inse det räcker visa

$$\mathbf{Cov}(a\,X,Y)=a\,\mathbf{Cov}(X,Y)\quad\text{och}\quad \mathbf{Cov}(X+Y,Z)=\mathbf{Cov}(X,Z)+\mathbf{Cov}(Y,Z)\qquad(1)$$

för att sedan konkludera 1 mha. upprepat användande av (1) och symmetrin $\mathbf{Cov}(X,Y)$ $=\mathbf{Cov}(Y,X)$. Men (1) följer av sats 5.9 och en inspektion av definition 5.10.

2. Följer direkt av 1 och faktumet att $\mathbf{Var}(\sum_{i=1}^m a_iX_i)=\mathbf{Cov}(\sum_{i=1}^m a_iX_i,\sum_{j=1}^m a_jX_j)$.

3. Följer direkt av 2 och definition 5.11 av okorrelerad. □

Definition 5.14. Korrelationskoefficienten för två variabler X och Y definieras

$$\rho_{X,Y}=\frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{Var}(X)}\sqrt{\mathbf{Var}(Y)}}.$$

Sats 5.15. $|\rho_{X,Y}|\le 1$ med likhet om och endast om X och Y är linjärt beroende, dvs. $a\,X+b=c\,Y+d$ för några konstanter $a,b,c,d\in\mathbb{R}$.

Bevis. Överkurs. Enligt sats 5.13 gäller att

$$0 \leq \mathbf{Var}\Big(\tfrac{X}{\sqrt{\mathbf{Var}(X)}} \pm \tfrac{Y}{\sqrt{\mathbf{Var}(Y)}}\Big) = \tfrac{\mathbf{Var}(X)}{(\sqrt{\mathbf{Var}(X)}\,)^2} \pm 2\tfrac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{Var}(X)}\sqrt{\mathbf{Var}(Y)}} + \tfrac{\mathbf{Var}(Y)}{(\sqrt{\mathbf{Var}(Y)}\,)^2}$$
$$= 1 \pm 2\,\rho_{X,Y} + 1$$

som enligt 3 i sats 3.10 är lika noll om och endast om $X/\sqrt{\mathbf{Var}(X)} \pm Y/\sqrt{\mathbf{Var}(Y)}$ är en konstant, vilket i sin tur är samma som att X och Y är linjärt beroende. $\square$

Definition 5.16. Två diskreta variabler X och Y har betingad frekvensfunktion

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

Den första beteckningen $f_{X|Y}(x|y)$ är den gängse använda medan M&A använder $f_{X|y}(x)$. Jag kommer använda den första.

Den betingade frekvensfunktionen $f_{X|Y}(x|y)$ är en frekvensfunktion som uppfyller reglerna 1-2 i sats 3.5. Det är en annan frekvensfunktion för X än den ursprungliga $f_X(x)$ där man tagit hänsyn till man har informationen att $Y = y$.

Sats 5.17. (Total sannolikhet) För diskreta variabler X och Y och $C \subseteq \mathbb{R}$ är

$$\mathbf{P}(X \in C|Y = y) = \sum_{x \in C} f_{X|Y}(x|y) \quad \text{och} \quad \mathbf{P}(X \in C) = \sum_{\text{alla } y} \mathbf{P}(X \in C|Y = y) f_Y(y).$$

Bevis. Då händelsen $\{X \in C, Y = y\}$ betyder $(X,Y) \in C \times \{y\}$ är enligt 5 i sats 5.5

$$\mathbf{P}(X \in C|Y = y) = \frac{\mathbf{P}(X \in C, Y = y)}{\mathbf{P}(Y = y)} = \frac{\sum_{x \in C} f_{X,Y}(x,y)}{f_Y(y)} = \sum_{x \in C} f_{X|Y}(x|y)$$

som i sin tur mha. definition 5.16 och 4 i sats 5.2 (samt 3 i sats 3.5) ger att

$$\sum_{\text{alla } y} \mathbf{P}(X \in C|Y = y) f_Y(y) = \sum_{\text{alla } y} \Big(\sum_{x \in C} f_{X|Y}(x|y)\Big) f_Y(y)$$
$$= \sum_{x \in C} \Big(\sum_{\text{alla } y} f_{X,Y}(x,y)\Big) = \sum_{x \in C} f_X(x) = \mathbf{P}(X \in C). \qquad \square$$

Definition 5.18. För diskreta variabler X och Y definieras betingat väntevärde

$$\mathbf{E}(X|Y = y) = \sum_{\text{alla } x} x\, f_{X|Y}(x|y).$$

Sats 5.19. (Totalt väntevärde) För diskreta variabler X och Y är

$$\mathbf{E}(X) = \sum_{\text{alla } y} \mathbf{E}(X|Y = y) f_Y(y).$$

Bevis. Enligt definitionerna 5.16 och 5.18 samt 4 i sats 5.2 är

$$\sum_{\text{alla } y} \mathbf{E}(X|Y=y) f_Y(y) = \sum_{\text{alla } y} \Big(\sum_{\text{alla } x} x\, f_{X|Y}(x|y) \Big) f_Y(y)$$

$$= \sum_{\text{alla } x} x \Big(\sum_{\text{alla } y} f_{X,Y}(x,y) \Big) = \sum_{\text{alla } x} x\, f_X(x) = \mathbf{E}(X). \qquad \square$$

Exempel 5.3. (Tal 3 20150605) Givet ett tal $\lambda > 0$, bestäm $\mathbf{E}(X|Y=\ell)$ för en tvådimensionell diskret stokastisk variabel (X,Y) med frekvensfunktion

$$f_{X,Y}(k,\ell) = \frac{(k+\ell)\,\lambda^{k+\ell}}{2\,\lambda\,(k!)\,(\ell!)}\,\mathrm{e}^{-2\lambda} \quad \text{för } k,\ell = 0,1,2,\ldots \quad \text{och} \quad f_{X,Y}(k,\ell) = 0 \quad \text{fö.}$$

Lösning. Jag kollar först att $f_{X,Y}(k,\ell)$ verkligen är en frekvensfunktion:

```
In[1]:= fXY[k,l]=(k+l)*lambda^(k+l)*Exp[-2*lambda]/(2*lambda*(k!)*(l!));
In[2]:= Sum[Sum[fXY[k,l],{k,0,Infinity}],{l,0,Infinity}]
Out[2]= 1
```

Så räknar jag ut $f_Y(l)$ mha. 4 i sats 5.2

```
In[3]:= fY[l]=Sum[fXY[k,l],{k,0,Infinity}]
```

Out[3]= $\frac{\mathrm{e}^{-lambda}\, lambda^{l-1}\,(l+lambda)}{2\, l!}$

Också räknar jag slutligen ut $\mathbf{E}(X|Y=\ell)$ enligt definition 5.18:

```
In[4]:= Sum[k*fXY[k,l]/fY[l],{k,0,Infinity}]
```

Out[4]= $\frac{lambda + l\, lambda + lambda^2}{l + lambda}$

Definition 5.20. Kontinuerliga variabler X och Y har betingad frekvensfunktion

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

Överkurs. Definition 5.20 kan motiveras antingen genom tänka att $f_{X|Y}(x|y)$ bör vara proportionell mot $f_{X,Y}(x,y)$ och för att x-arean under $f_{X|Y}(x|y)$ skall vara 1 måste då proportionalitetskonstanten enligt 7 i sats 5.5 vara $1/f_X(x)$. Eller låter man $\varepsilon \downarrow 0$ i

$$\frac{d}{dx}\mathbf{P}(X \le x | Y \in (y-\varepsilon, y]) = \frac{d}{dx}\frac{F_{X,Y}(x,y) - F_{X,Y}(x,y-\varepsilon)}{\varepsilon}\,\frac{\varepsilon}{F_Y(y) - F_Y(y-\varepsilon)} \to \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Den betingade frekvensfunktionen $f_{X|Y}(x|y)$ är en annan frekvensfunktionen för X än den ursprungliga $f_X(x)$ där man tagit hänsyn till man har informationen att $Y = y$.

Definition 5.21. För kontinuerliga X och Y och $C \subseteq \mathbb{R}$ definieras

$$\mathbf{P}(X \in C | Y = y) = \int_{x \in C} f_{X|Y}(x|y)\,dx \quad \text{och} \quad \mathbf{E}(X|Y=y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx.$$

$\mathbf{E}(X|Y=y)$ betecknas även $\mu_{X|y}$ och kallas regressionskurva av X på Y.

Sats 5.22. (Total sannolikhet och Totalt väntevärde) För kontinuerliga X och Y och $C \subseteq \mathbb{R}$ är

$$\mathbf{P}(X \in C) = \int_{-\infty}^{\infty} \mathbf{P}(X \in C|Y=y) f_Y(y)\, dy \quad \text{och} \quad \mathbf{E}(X) = \int_{-\infty}^{\infty} \mathbf{E}(X|Y=y) f_Y(y)\, dy.$$

Bevis. Helt analogt med de diskreta bevisen av sats 5.17 och 5.19. □

Exempel 5.4. (Tal 3 20161007) Beräkna $\mathbf{E}(X\,|\,Y=y)$ för en tvådimensionell kontinuerlig stokastisk variabel (X, Y) med frekvensfunktion

$$f_{XY}(x, y) = 1/x \quad \text{för } 0 \leq y \leq x \leq 1 \quad \text{och} \quad f_{X,Y}(x, y) = 0 \quad \text{fö.}$$

Lösning. Jag kollar först att $f_{X,Y}(x, y)$ är en frekvensfunktion:

```
In[6]:= fXY[x,y]=1/x; Integrate[Integrate[f[x,y],{x,y,1}],{y,0,1}]
Out[6]= 1
```

Så räknar jag ut $f_Y(y)$ mha. 7 i sats 5.5

```
In[7]:= fY[y]=Integrate[f[x,y],{x,y,1}]
Out[7]= −Log[y]
```

Också räknar jag slutligen ut $\mathbf{E}(X|Y=y)$ enligt definition 5.21:

```
In[8]:= Integrate[x*fXY[x,y]/fY[y],{x,y,1}]
```
Out[8]= $\frac{1-y}{-\mathrm{Log}[y]}$

Sats 5.23. Låt X och Y vara kontinuerliga stokastiska variabler med frekvensfunktion $f_{X,Y}(x, y)$ och $g_1, g_2, h_1, h_2 : \mathbb{R}^2 \to \mathbb{R}$ funktioner sådana att

$$(u, v) = (g_1(x, y), g_2(x, y)) \;\Leftrightarrow\; (x, y) = (h_1(u, v), h_2(u, v)) \quad \text{för } x, y, u, v \in \mathbb{R}.$$

De kontinuerliga stokastiska variablerna $U = g_1(X, Y)$ och $V = g_2(X, Y)$ har då tvådimensionell frekvensfunktion

$$f_{U,V}(u, v) = f_{X,Y}(h_1(u, v), h_2(u, v)) \begin{vmatrix} \frac{\partial h_1(u,v)}{\partial u} & \frac{\partial h_1(u,v)}{\partial v} \\ \frac{\partial h_2(u,v)}{\partial u} & \frac{\partial h_2(u,v)}{\partial v} \end{vmatrix} \quad \text{för } (u, v) \in \mathbb{R}^2.$$

Determinanten $|\cdot|$ ovan är den sk. Jacobianen eller funktionaldeterminanten.

Bevis. Följer av formeln för variabelbyte i dubbelintegral:

$$
\begin{aligned}
\mathbf{P}((U,V)\in C) &= \mathbf{P}\big((g_1(X,Y),g_2(X,Y))\in C\big)\\
&= \int\int_{\{(x,y)\in\mathbb{R}^2:(g_1(x,y),g_2(x,y))\in C\}} f_{X,Y}(x,y)\,dxdy\\
&= \left[\begin{cases} x=h_1(u,v)\\ y=h_2(u,v)\end{cases},\ dxdy=\begin{vmatrix}\cdot & \cdot\\ \cdot & \cdot\end{vmatrix}dudv\right]\\
&= \int\int_{\{(u,v)\in\mathbb{R}^2:(u,v)\in C\}} f_{X,Y}(h_1(u,v),h_2(u,v))\begin{vmatrix}\frac{\partial h_1(u,v)}{\partial u} & \frac{\partial h_1(u,v)}{\partial v}\\ \frac{\partial h_2(u,v)}{\partial u} & \frac{\partial h_2(u,v)}{\partial v}\end{vmatrix}dudv. \quad \square
\end{aligned}
$$

Exempel 5.5. (EXAMPLE 5.5.2 M&A) Om X och Y är obereonde och $\mathrm{Uni}[0,2]$ respektive $\mathrm{Uni}[0,3]$ är enligt definition 4.8 och sats 5.7 $f_{X,Y}(x,y)=1/6$. Om

$$(U,V)=(g_1(X,Y),g_2(X,Y))=(X-Y,X+Y)$$

blir

$$(X,Y)=((U+V)/2,(V-U)/2)=(h_1(U,V),h_2(U,V))$$

så att sats 5.23 enkelt ger

$$f_{U,V}(u,v)=f_{X,Y}\left(\tfrac{u+v}{2},\tfrac{v-u}{2}\right)\begin{vmatrix}\frac{\partial}{\partial u}\frac{u+v}{2} & \frac{\partial}{\partial v}\frac{u+v}{2}\\ \frac{\partial}{\partial u}\frac{v-u}{2} & \frac{\partial}{\partial v}\frac{v-u}{2}\end{vmatrix}=\frac{1}{6}\begin{vmatrix}\frac{1}{2} & \frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{vmatrix}=\frac{1}{12}.$$

VB definitionsområdet för $f_{U,V}(u,v)$, dvs. för vilka (u,v) formeln gäller och för vilka (u,v) istället $f_{U,V}(u,v)=0$ visas det i figuren nedan stulen från M&A där området till vänster $0\le x\le 2$, $0\le y\le 3$ är definitionsområdet för $f_{X,Y}(x,y)$ och området till höger $0\le u+v\le 4$, $0\le v-u\le 6$ är definitionsområdet för $f_{U,V}(u,v)$: