MVE090 LP4 2021 Föreläsning 7-8

Kapitel 5. "Joint Distributions"

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

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

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

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

Definition 5.3. Om X och Y är stokastiska variabler definieras deras gemensamma fördelningsfunktion (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 frekvensfunktion (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. $f_{X,Y}(x,y) \geq 0$,
2. $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\, dxdy = 1$,
3. $F_{X,Y}(x,y) = \int\limits_{u=-\infty}^{u=x} \int\limits_{v=-\infty}^{v=y} f_{X,Y}(u,v)\, dudv$,
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. $f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$ och $f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$.

Man kan i allmänhet ej beräkna $f_{X,Y}(x,y)$ med kännedom om $f_X(x)$ och $f_Y(y)$!

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. Låt X och Y vara kontinuerliga med $f_{X,Y}(x,y) = c/x$ för $27 \le y \le x \le 33$[1] för ett $c > 0$. Bestäm c, $f_X(x)$ och $f_Y(y)$. Är X och Y oberoende?

Lösning. Konstanten c bestäms 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. 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

$$P(X>Y) = \int\int_{\{(x,y)\in\mathbb{R}^2\,:\,x>y>0\}} \tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}\,\mathrm{e}^{-y}\,dydx = \int_0^{\infty}\tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}\,(1-\mathrm{e}^{-x})\,dx.$$

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

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

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

Exempel 5.1. (Forts.) För beräkna $\mathbf{E}(X)$ kan det tidigare gjorda utnyttjas:

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

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

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

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

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

Definition 5.10. <u>Kovariansen</u> mellan två stokastiska variabler X och Y definieras

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

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

Kovariansen är ett basalt mått på hur mycket X och Y beror av varandra. För veta vad som är stort eller litet beroende normaliseras kovariansen till korrelationskoefficient:

Definition 5.11. <u>Korrelationskoefficienten</u> för två variabler X och Y definieras

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

Sats 5.12. $|\rho_{X,Y}| \leq 1$ med likhet om och endast om X och Y är linjärt beroende.

Definition 5.13. Om $\mathbf{Cov}(X,Y) = 0$ säges X och Y vara <u>okorrelerade</u>.

Sats 5.14. Om X och Y är oberoende är de okorrelerade.

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

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

$\square$

Sats 5.15. För stokastiska variabler $\{X_i\}_{i=1}^m$, $\{Y_j\}_{j=1}^n$ och tal $\{a_i\}_{i=1}^m$, $\{b_j\}_{j=1}^n$ är

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

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

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

Beteckningen $f_{X|Y}(x|y)$ är den gängse använda medan M&A använder $f_{X|y}(x)$.

Betingade frekvensfunktionen $f_{X|Y}(x|y)$ uppfyller reglerna 1-2 i sats 3.5 och är en annan frekvensfunktion för X än $f_X(x)$ där hänsyn tagits till informationen $Y = y$.

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

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

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

$$\begin{aligned}\sum_{\text{alla } y} \mathbf{E}(X|Y = y) f_Y(y) &= \sum_{\text{alla } y} \Big(\sum_{\text{alla } x} x\, f_{X|Y}(x|y)\Big) f_Y(y) \\ &= \sum_{\text{alla } x} x \Big(\sum_{\text{alla } y} f_{X,Y}(x, y)\Big) = \sum_{\text{alla } x} x\, f_X(x) = \mathbf{E}(X). \qquad \square\end{aligned}$$

Exempel 5.3. Finn $\mathbf{E}(X|Y=\ell)$ för en diskret stokastisk variabel (X,Y) med

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

Lösning. Beräkna $f_Y(l)$ mha. 4 i sats 5.2:

```
In[1]:= fY[l]=Sum[fXY[k,l],{k,0,Infinity}]
```

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

Räkna nu ut $\mathbf{E}(X|Y=\ell)$ enligt definition 5.18:

```
In[2]:= Sum[k*fXY[k,l]/fY[l],{k,0,Infinity}]
```

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

Definition 5.20. Kontinuerliga variabler X och Y har betingad frekvensfunktion

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

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 regressionskurvan av X på Y.

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

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

Exempel 5.4. Finn $\mathbf{E}(X|Y=y)$ för en kontinuerlig stokastisk variabel (X,Y) med

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

Lösning. Beräkna $f_Y(y)$ mha. 7 i sats 5.5:

```
In[3]:= fY[y]=Integrate[f[x,y],{x,y,1}]
```

Out[3]= $-\mathrm{Log}[y]$

Räkna nu ut $\mathbf{E}(X|Y=y)$ enligt definition 5.21:

```
In[4]:= Integrate[x*fXY[x,y]/fY[y],{x,y,1}]
```

Out[4]= $\frac{1-y}{-\mathrm{Log}[y]}$

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

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

har de stokastiska variablerna $U = g_1(X,Y)$ och $V = g_2(X,Y)$ frekvensfunktion

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

Bevis. Följer av formeln för variabelbyte i dubbelintegral:

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

Exempel 5.5. Om X och Y är obereonde och Uni$[0,2]$ respektive 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

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

I figurerna nedan stulna från M&A visas definitionsområdena $0 \le x \le 2$, $0 \le y \le 3$ för $f_{X,Y}(x,y)$ och $0 \le u+v \le 4$, $0 \le v-u \le 6$ för $f_{U,V}(u,v)$: