Föreläsning 8

Flerdimensionella fördelningar (begr. oss till 2-dim)

Ex. Kasta två tärningar, låt antalet prickar betecknas av X & Y. Då är (X,Y) en 2-dim s.v.

Diskreta fallet:

Def: Två-dim simultan slh-funktion

$$p_{XY}(j,k) = P(\{X=j\} \cap \{Y=k\}) =$$
$$= P(X=j \text{ och } Y=k)$$

uppfyller $\sum_{j,k} p_{XY}(j,k) = 1$

Ex. Vi har en urna med 3 svarta och 2 vita bollar. Ta slumpmässigt upp 2 bollar & låt X vara antal svarta & Y antal vita. Bestäm slh-funktionen $p_{XY}(j,k)$

p_{XY}

svart X, j \ vit Y, k	0	1	2
0	0	0	$\frac{2}{5}\cdot\frac{1}{4}=\frac{1}{10}$
1	0	$\frac{2}{5}\cdot\frac{3}{4}+\frac{3}{5}\cdot\frac{2}{4}= \frac{3}{5}$	0
2	$\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10}$	0	0

Om vi vill beräkna $P(A)$ för en händelse A så summerar vi över alla $(j,k) \in A$

$$P(A) = \sum_{(j,k) \in A} p_{XY}(j,k)$$

Vill vi komma åt sannolikhetsfunktionen för en av de två variablerna använder vi:

~~Def~~ Marginell sannolikhetsfunktion

$$\boxed{\begin{aligned} p_X(j) &= P(X=j) = \sum_k p_{XY}(j,k) \\ p_Y(k) &= P(Y=k) = \sum_j p_{XY}(j,k) \end{aligned}}$$

Kontinuerliga fallet:

Kont. 2-dim s.v. (X,Y) har simultan täthetsfunktion $f_{XY}(x,y)$ som uppfyller

$$f_{XY}(x,y) \geq 0 \quad \text{överallt}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1$$

$$P(A) = \iint_{(x,y) \in A} f_{XY}(x,y)\,dx\,dy$$

p.s.s. som för diskreta s.v. kan vi från en simultan täthetsfunktion f_{XY} få marginalfördelningar f_X & f_Y:

$$\boxed{\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx \end{aligned}}$$

Beroende och oberoende osv

<u>Def</u> 2-dim fördelningsfunktion:

$$F_{XY}(x,y) = P(\{X \le x\} \cap \{Y \le y\}) = \begin{cases} \sum\limits_{j \le x} \sum\limits_{k \le y} p_{XY}(j,k) \\ \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(s,t)\,dt\,ds \end{cases}$$

<u>Def</u>. 2-dim <u>marginell</u> fördelningsfunktion:

$$F_X(x) = P(X \le x) = \begin{cases} \sum\limits_{j \le x} p_X(j) \\ \int_{-\infty}^{x} f_X(t)\,dt \end{cases}$$

<u>marginell</u> slh-fkn och täthet.

<u>Oberoende s.v.</u>

Oberoende händelser: om $\{X \le x\}$ och $\{Y \le y\}$ ob så

$$\begin{array}{ccccc} P(\{X \le x\} \cap \{Y \le y\}) & = & P(X \le x) & \cdot & P(Y \le y) \\ \parallel & & \parallel & & \parallel \\ F_{XY}(x,y) & = & F_X(x) & \cdot & F_Y(y) \end{array}$$

→ Def. Om för alla x, y gäller

$$\boxed{F_{XY}(x,y) = F_X(x)F_Y(y)}$$

Så är X & Y oberoende

i kont. fallet: derivera map x & y $\Rightarrow$ $f_{XY}(x,y) = f_X(x)f_Y(y)$

diskr. fallet: $p_{XY}(j,k) = p_X(j)p_Y(k)$.

Ex. (X,Y) har sannolikhetsfunktionen p_{XY} enl.

$p_{XY}(0,0) = 1/20$; $p_{XY}(0,1) = 3/20$, $p_{XY}(1,0) = 1/5$, $p_{XY}(1,1) = 3/5$

Är X & Y oberoende?

Beräknar $p_X(j)$ & $p_Y(k)$:

$$p_X(0) = p_{XY}(0,0) + p_{XY}(0,1) = 4/20$$
$$p_X(1) = p_{XY}(1,0) + p_{XY}(1,1) = 16/20$$
$$p_Y(0) = p_{XY}(0,0) + p_{XY}(1,0) = 5/20$$
$$p_Y(1) = p_{XY}(0,1) + p_{XY}(1,1) = 15/20$$

Vill kolla om $p_{XY}(j,k) = p_X(j)p_Y(k)$ överallt

$p_{XY}(0,0) = 1/20$; $p_X(0)p_Y(0) = \frac{4}{20}\cdot\frac{5}{20} = 1/20$ OK!

$p_{XY}(0,1) = 3/20$; $p_X(0)\cdot p_Y(1) = \frac{4}{20}\cdot\frac{15}{20} = 3/20$ OK!

$p_{XY}(1,0) = 4/20$; $p_X(1)p_Y(0) = \frac{16}{20}\cdot\frac{5}{20} = 4/20$ OK!

$p_{XY}(1,1) = 12/20$; $p_X(1)p_Y(1) = \frac{16}{20}\cdot\frac{15}{20} = 12/20$ OK!

$\Rightarrow$ $p_{XY} = p_X p_Y$

$\Rightarrow$ X & Y oberoende!

Kovarians

mått på (linjärt) beroende mellan två s.v.

Def. $C(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - \mu_X \mu_Y$

Om X & Y är oberoende så är

$$C(X,Y) = E[XY] - \mu_X \mu_Y = E[X]E[Y] - \mu_X \mu_Y$$
$$= \mu_X \mu_Y - \mu_X \mu_Y = 0$$

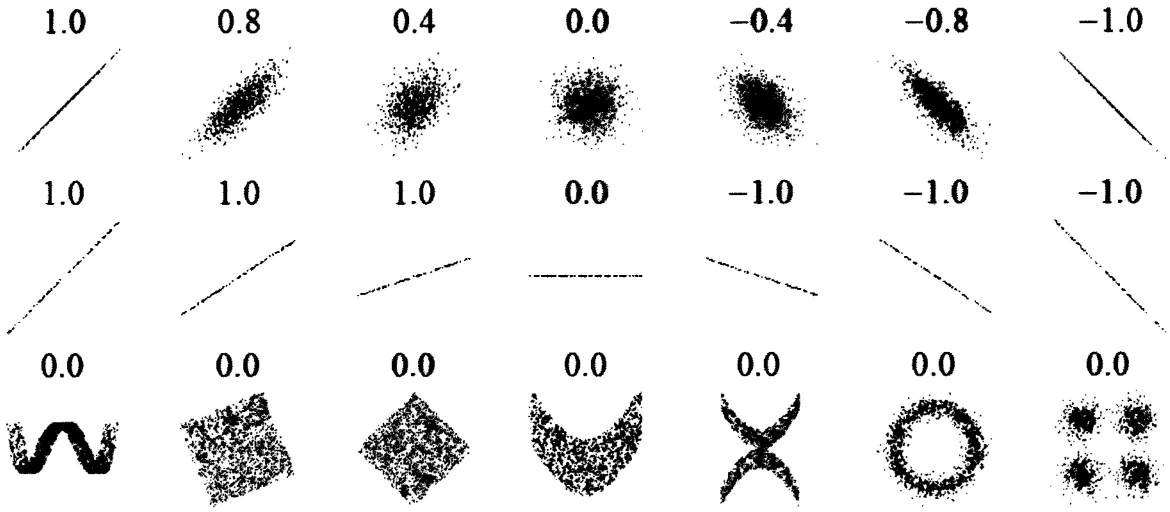
om $C(X,Y) = 0$ så är X & Y okorrelerade; behöver dock inte vara oberoende

Korrelationskoefficienten $\rho(X,Y) = \frac{C(X,Y)}{\sigma_X \sigma_Y}$ är en skalad variant s.a. $-1 \le \rho \le 1$.

$\boxed{\text{SLIDE}}$

Kovarians och korrelationskoefficient mäter bara LINJÄRA beroenden

Värden på korrelationskoefficient



Obs 1: Notera att om vi låter $Y = X$ så får vi

$$C(X,X) = E\left[(X-\mu_X)(X-\mu_X)\right] = E\left[(X-\mu_X)^2\right] = \\ = V(X).$$

Obs 2: Om vi beräknar varians av summa:

$$V(X+Y) = E\left[(X+Y-\mu_X-\mu_Y)^2\right] = \\ = E\left[\left((X-\mu_X)+(Y-\mu_Y)\right)^2\right] = E\left[(X-\mu_X)^2+(Y-\mu_Y)^2+ \\ +2(X-\mu_X)(Y-\mu_Y)\right] =$$

$$= V(X)+V(Y)+2\underbrace{E\left[(X-\mu_X)(Y-\mu_Y)\right]}_{C(X,Y)} =$$

$$= V(X)+V(Y)+\underline{2C(X,Y)}$$

Så additionsregeln för varians gäller för okorrelerade variabler (måste inte vara oberoende)