Førelesing TMA4240

13.09.12

4.2 Varians og kovarians

Eks. Kast mynt og krone

X = gevinst

$$X = \begin{cases} 1 & \text{dersom krone} \\ -1 & \text{dersom mynt} \end{cases}$$

$$E[X] = 1 \cdot \frac{1}{2} + -1 \cdot \frac{1}{2} = 0$$

x	-1	1
$P(X=x)$	$\frac{1}{2}$	$\frac{1}{2}$

x	-10^6	10^6
$P(X=x)$	$\frac{1}{2}$	$\frac{1}{2}$

$$E[X] = 0$$

Har definert $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$, utvalsvarians

For stor n er $\frac{1}{n-1} \approx \frac{1}{n}$ s.a

$$s^2 \approx \sum_i (x_i - \bar{x})^2 \cdot \text{rel.frekvens}(x_i) \xrightarrow{n \to \infty} \sum_x (x-\mu_X)^2 f(x)$$

<u>Definisjon 4.3</u>

X er stokastisk variabel, $E[X] = \mu_X$. Variansen til X er definert som

$$\sigma^2 \triangleq \begin{cases} \sum_x (x-\mu_X)^2 f(x) = E[(X-\mu_X)^2], & X \text{ diskret} \\ \int_{-\infty}^{\infty} (x-\mu_X)^2 f(x)\,dx = E[(X-\mu_X)^2], & X \text{ kont} \end{cases}$$

Standardavviket $\sigma = \sqrt{Var[X]}$

Varians (standardavvik) er eit mål på variasjon om forventingsverdien

Eks.

x	-20	5	10
$P(X=x)$	0.2	0.3	0.5

$E[X] = 2.5$

$$Var[X] = \sigma_X^2 = (-20-2.5)^2 \cdot 0.2 + (5-2.5)^2 \cdot 0.3 + (10-2.5)^2 \cdot 0.5$$

$$= 131.25$$

$$\sigma_X = \sqrt{131.25} = 11.46$$

Kovarians

La $g(x,y) = (X-\mu_X)(Y-\mu_Y)$

Definisjon 4.4

Kovariansen mellom X og Y er definert som

$$\sigma_{XY} = E[g(X,Y)] = \begin{cases} \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f(x,y), & X, Y \text{ er diskret} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f(x,y)\,dx\,dy, & X, Y \text{ er kont} \end{cases}$$

Kovarians er eit mål på korleis X og Y varierer saman. Eks høgde og vekt

$X > \mu_X$ når $Y > \mu_Y$ eller $X < \mu_X$ når $Y < \mu_Y$ gjev positiv σ_{XY}

$X > \mu_X$ når $Y < \mu_Y$ eller $X < \mu_X$ når $Y > \mu_Y$ — " — negativ σ_{XY}

Vi har $E[g(X,Y)] = E[(X-\mu_X)(Y-\mu_Y)]$

$$= E[XY] - \mu_X E[Y] - \overbrace{\mu_Y \underbrace{E[X]}_{= \mu_X} + \mu_X \cdot \mu_Y}^{0} = E[XY] - E[X] \cdot E[Y]$$

x y	1	2	3	4	P(X=x)
0	0.11	0.09	0.07	0.01	0.28
1	0.07	0.12	0.12	0.02	0.33
2	0.02	0.05	0.17	0.05	0.29
3	0.00	0.02	0.04	0.02	0.08
4	0.00	0.00	0.01	0.01	0.02
P(Y=y)	0.20	0.28	0.01	0.01	

X = talet på barn

Y = talet på rom

$$E[XY] = 1\cdot 1\cdot 0{,}07 + 1\cdot 2\cdot 0{,}12 + 1\cdot 3\cdot 0{,}12 + 1\cdot 4\cdot 0{,}02$$
$$+ 2\cdot 1\cdot 0{,}02 + 2\cdot 2\cdot 0{,}05 + 2\cdot 3\cdot 0{,}17 + 2\cdot 4\cdot 0{,}05$$
$$+ 3\cdot 2\cdot 0{,}02 + 3\cdot 3\cdot 0{,}04 + 3\cdot 4\cdot 0{,}02$$
$$+ 4\cdot 3\cdot 0{,}01 + 4\cdot 4\cdot 0{,}01 = 3{,}41$$

$\sigma_{XY} = 3.41 - 1{,}23\cdot 2{,}43 = 0.43$. NB! σ_{XY} er skalaavhengig

Dersom X og Y er uavh, er $\sigma_{XY} = 0$

Bevis. X og Y uavh $\Rightarrow$ $\sigma_{XY} = \sum_x \sum_y xy\, g(x) h(y) - \sum_x x g(x) \cdot \sum_y y h(y)$

$$= \sum_x x g(x) \sum_y y h(y) - \sum_x x g(x) \sum_y y h(y) = 0$$

For å ha eit mål på samvariasjon som er uavh. av benemning innfører ein korrelasjonskoeffisienten

Def 4.5. X, Y stokastiske variable med standardavvik σ_X og σ_Y. Korrelasjonskoeffisienten er definert som

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}, \quad -1 \leq \rho_{XY} \leq 1$$

4.3 Forventning og varians for lineær kombinasjonar av tilfeldige variable

Teorem 4.7

$$E[g(X,Y)+h(X,Y)] = E[g(X,Y)] + E[h(X,Y)]$$

La $Z(X,Y) = g(X,Y) + h(X,Y)$

$$E[Z(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (g(x,y)+h(x,y))\, f(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f(x,y)\, dx\, dy$$

$$= E[g(X,Y)] + E[h(X,Y)]$$

Korrolar 4.6

$$g(X,Y) = g(X), \quad h(X,Y) = h(Y)$$

$$\Rightarrow \quad E[g(X) + h(Y)] = E[g(X)] + E[h(Y)]$$

Dersom $h(X,Y) = h(X)$

$$\Rightarrow \quad E[g(X) + h(X)] = E[g(X)] + E[h(X)]$$

Eks. $E[X(X+1)] = E[X^2 + X] = E[X^2] + E[X]$

Spesielt får vi: La $g(x) = aX$ og $h(x) = b$

$$E[aX + b] = E[aX] + E[b] = \int_{-\infty}^{\infty} ax f(x)\,dx + \int_{-\infty}^{\infty} b f(x)\,dx$$

$$= aE[X] + b \quad \text{(Teorem 4.5)}$$

Teorem 4.9.

La $Y = aX + b \Rightarrow E[Y] = aE[X] + b = a\mu + b$

$$Var[Y] = \sigma_Y^2 = E[(aX + b - (a\mu + b))^2] = E[(aX - a\mu)^2]$$

$$= E[a^2(X-\mu)^2] = a^2 E[(X-\mu)^2] = a^2 Var[X]$$

Teorem 4.10

La $Z = g(X,Y) = aX + bY \Rightarrow E[Z] = a\mu_X + b\mu_Y$

$$\sigma_Z^2 = E[(aX + bY - (a\mu_X + b\mu_Y))^2] = E[(a(X-\mu_X) + b(Y-\mu_Y))^2]$$

$$= a^2\sigma_X^2 + 2ab\,\sigma_{XY} + b^2\sigma_Y^2$$