

# Førelesing TMA4240

13.09.12

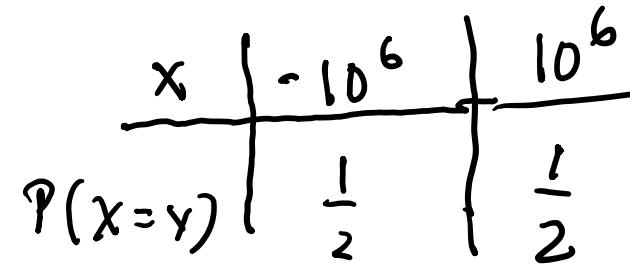
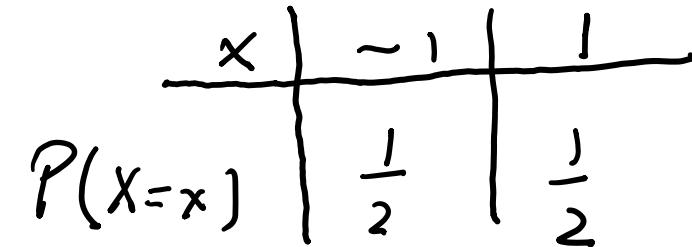
## 4.2 Varians og kovarians

Eks. Kast mynt og krone

$$X = \text{ gevinst}$$

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

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



$$\mathbb{E}[X] = 0$$

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

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

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

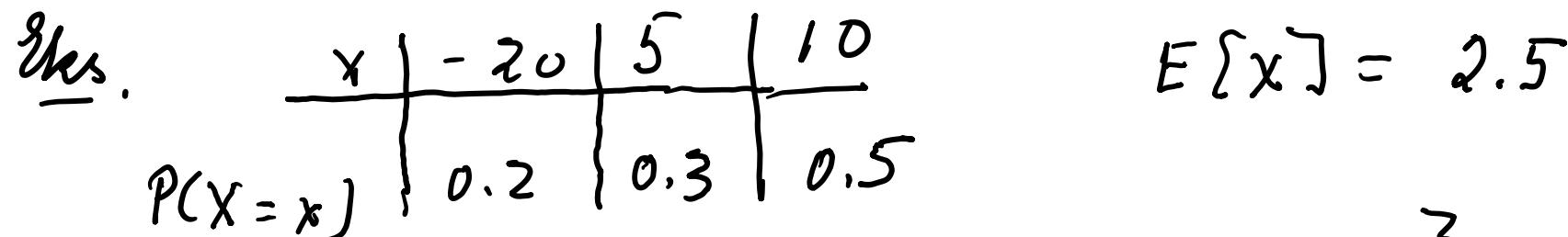
### Definisjon 4.3

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

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

$$\text{Standardavviket} \quad \sigma = \sqrt{\text{Var}[x]}$$

Varians (standardavvik) er eit mål på variasjon om forventingsverdien



$$\begin{aligned}\text{Var}[x] &= \sigma_x^2 = (-20 - 2.5)^2 \cdot 0.2 + (5 - 2.5)^2 \cdot 0.3 + (10 - 2.5)^2 \cdot 0.5 \\ &= 131.25\end{aligned}$$

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

## Kovarians

$$\text{La } g(x, y) = (x - \mu_x)(y - \mu_y)$$

### 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  $y < \mu_y$  gir positiv  $\sigma_{xy}$

$x > \mu_x$  når  $y < \mu_y$  eller  $x < \mu_x$  når  $y > \mu_y$  gir negativ  $\sigma_{xy}$

Vi har  $E[g(x, y)] = E\left[\overset{\circ}{(x - \mu_x)(y - \mu_y)}\right]$

$$= E[x y] - \mu_x E[y] - \underbrace{\mu_y E[x] + \mu_x \cdot \mu_y}_{\mu_X} = E[x y] - E[x] \cdot E[y]$$

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	

$X = \text{talst} \overset{\circ}{\text{pa}} \text{ borm}$

$Y = \text{talst} \overset{\circ}{\text{pa}} \text{ roun}$

$$\begin{aligned}
 E[XY] &= 1 \cdot 1 \cdot 0.07 + 1 \cdot 2 \cdot 0.12 + 1 \cdot 3 \cdot 0.12 + 1 \cdot 4 \cdot 0.02 \\
 &\quad + 2 \cdot 1 \cdot 0.02 + 2 \cdot 2 \cdot 0.05 + 2 \cdot 3 \cdot 0.17 + 2 \cdot 4 \cdot 0.05 \\
 &\quad + 3 \cdot 2 \cdot 0.02 + 3 \cdot 3 \cdot 0.04 + 3 \cdot 4 \cdot 0.02 \\
 &\quad + 4 \cdot 3 \cdot 0.01 + 4 \cdot 4 \cdot 0.01 = 3.41
 \end{aligned}$$

$$G_{XY} = 3.41 - 1.23 \cdot 2.43 = 0.43. \text{ NB! } G_{XY} \text{ vs Skalaavhengig}$$

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

Bewis.

$$\begin{aligned} X \text{ og } Y \text{ uavh} \Rightarrow \sigma_{XY} &= \sum_x \sum_y xy g(x) h(y) - \sum_x x g(x) \cdot \sum_y y h(y) \\ &= \sum_x x g(x) \sum_y y h(y) - \sum_x x g(x) \sum_y y h(y) = 0 \end{aligned}$$

For å ha eit mål på sannvariasjon som er uavh.  
av betrekning innføres ein korrelasjonskoefisiens

Dif 4.5.  $X, Y$  stokastiske variable med standardavvik  $\sigma_X$   
og  $\sigma_Y$ . Korrelasjonskoefisiensen er definert som

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

4.3 Forventning og varians for lineær kombinasjoner  
av tilfeldige variable

Teorem 4.7

$$E[g(x, y) + h(x, y)] = E[g(x, y)] + E[h(x, y)]$$

La  $Z(x, y) = g(x, y) + h(x, y)$

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

## Korrolas 4.6

$$g(x, y) = g(x), \quad h(x, y) = h(y)$$

$$\Rightarrow E[g(x) + h(y)] = E[g(x)] + E[h(y)]$$

Dersom  $h(x, y) = h(x)$

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

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

Specielt fås vi: La  $g(x) = ax$  og  $h(x) = b$

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

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

Theorem 4.9.

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

$$\begin{aligned} \text{Var}[Y] &= \sigma_Y^2 = E[(aX + b - (a\mu + b))^2] = E[(aX - a\mu)^2] \\ &= E[a^2(X - \mu)^2] = a^2 E[(X - \mu)^2] = a^2 \text{Var}[X] \end{aligned}$$

Theorem 4.10

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

$$\begin{aligned} \sigma_Z^2 &= E[(aX + bY - (a\mu_X + b\mu_Y))^2] = E[(a(X - \mu_X) + b(Y - \mu_Y))^2] \\ &= a^2 \sigma_X^2 + 2ab \sigma_{XY} + b^2 \sigma_Y^2 \end{aligned}$$