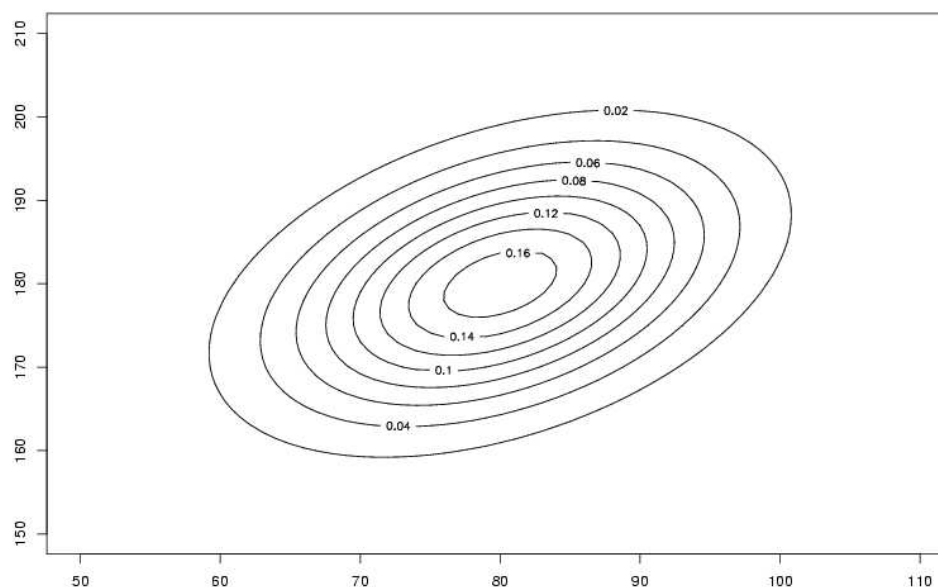


Kapittel 4: Matematisk forventning

TMA4240 Statistikk (F2 og E7)

Multivariate tilfeller foreleses mandag 6.september, 2004



Forventing til funksjon av flere stokastiske variable

DEF 4.2: La X og Y være stokastisk variable med simultan sannsynlighetsfordeling $f(x, y)$. Forventningsverdien til den stokastiske variabelen $g(X, Y)$ er

$$\mu_{g(X,Y)} = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

hvis X og Y er diskrete, og

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

hvis X og Y er kontinuerlige.

Prosjektstyring (forts.)

● X = tid for datainnsamling, Y = tid for dataanalyse.

		x			
		1	2	3	$f_Y(y)$
y	1	0.03	0.05	0.02	0.10
	2	0.03	0.14	0.03	0.20
	3	0.03	0.17	0.10	0.30
	4	0.01	0.24	0.15	0.40
$f_X(x)$		0.10	0.60	0.30	1.00

● Er interessert i forholdet $g(X, Y) = \frac{Y}{X}$ mellom varigheten av datainnsamling og dataanalyse.

$$\begin{aligned}
 E\left[\frac{Y}{X}\right] &= \sum_x \sum_y \frac{y}{x} f(x, y) \\
 &= 1 \cdot 0.03 + (1/2) \cdot 0.05 + (1/3) \cdot 0.02 + 2 \cdot 0.03 + 1 \cdot 0.14 + (2/3) \cdot 0.03 \\
 &\quad + 3 \cdot 0.03 + (3/2) \cdot 0.17 + 1 \cdot 0.10 + 4 \cdot 0.01 + 2 \cdot 0.24 + (4/3) \cdot 0.15 \\
 &= 1.32
 \end{aligned}$$

$$NB : E\left[\frac{Y}{X}\right] \neq \frac{E(X)}{E(Y)}.$$

E (funksjoner av flere stokastiske variabler)

TEO 4.7: Forventningsverdien til summen eller differansen av to eller flere funksjoner av de stokastiske variablene X og Y , er summen eller differansen til forventningsverdiene til funksjonene. Det vil si,

$$E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)].$$

COR 1: Setter vi $g(X, Y) = g(X)$ og $h(X, Y) = h(Y)$

$$E[g(X) \pm h(Y)] = E[g(X)] \pm E[h(Y)].$$

COR 2: Setter vi $g(X, Y) = X$ og $h(X, Y) = Y$

$$E[X \pm Y] = E[X] \pm E[Y].$$

Generalisering

$$Y = \sum_{i=1}^n a_i X_i + b$$

$$E(Y) = \sum_{i=1}^n a_i E(X_i) + b$$

- Formelsamlingen s 34.

Forventing til produktet av uavhengige stokastiske variable

TEO 4.8: La X og Y være to *uavhengige* stokastiske variable. Da er

$$E(X \cdot Y) = E(X) \cdot E(Y).$$

4.2 Varians og kovarians

DEF 4.4: La X og Y være to stokastisk variable med simultan sannsynlighetsfordeling $f(x, y)$, og forventninger hhv. $\mu_X = E(X)$ og $\mu_Y = E(Y)$. *Kovariansen til X og Y er*

$$\begin{aligned}\sigma_{XY} &= \text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \\ &= \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)\end{aligned}$$

hvis X og Y er diskrete, og

$$\begin{aligned}\sigma_{XY} &= \text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy\end{aligned}$$

hvis X og Y er kontinuerlige.

Kovarians og korrelasjon

TEO 4.4: Kovariansen til to stokastiske variable X og Y med forventninger hhv. $\mu_X = E(X)$ og $\mu_Y = E(Y)$, er gitt ved

$$\begin{aligned}\sigma_{XY} &= \text{Cov}(X, Y) \\ &= E(X \cdot Y) - E(X) \cdot E(Y) \\ &= E(X \cdot Y) - \mu_X \mu_Y\end{aligned}$$

DEF 4.5: La X og Y være to stokastisk variable med kovarians σ_{XY} og varianser hhv. σ_X^2 og σ_Y^2 . Korrelasjonskoeffisienten til X og Y er

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Tolkning av Cov og ρ_{XY}

- Kovarians $\text{Cov}(X, Y) = E(X \cdot Y) - \mu_X \mu_Y$.
 - Når X og Y er uavhengige er $E(X \cdot Y) = E(X) \cdot E(Y) = \mu_X \mu_Y$.
 - Dermed når X og Y er uavhengige er $\text{Cov}(X, Y) = \mu_X \mu_Y - \mu_X \mu_Y = 0$.
 - Men, hvis $\text{Cov}(X, Y) = 0$ betyr det nødvendigvis IKKE at X og Y er uavhengige.
- Korrelasjonskoeffisienten:
 - Hvis $Y = aX + b$ og $a > 0 \rightarrow \rho_{XY} = 1$
 - Hvis $Y = aX + b$ og $a < 0 \rightarrow \rho_{XY} = -1$
 - Hvis X og Y er uavhengige $\rightarrow \rho_{XY} = 0$
 - $-1 \leq \rho_{XY} \leq 1$.

Varians til lineærkombinasjon av to stokastiske variable

TEO 4.10: La X og Y være to stokastisk variable med simultan sannsynlighetsfordeling $f(x, y)$, da er

$$\begin{aligned}\sigma_{aX+bY}^2 &= \text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) + 2ab \text{Cov}(X, Y) \\ &= a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab \cdot \sigma_{XY}\end{aligned}$$

COR 1: Hvis X og Y er uavhengige stokastiske variable, så er $\text{Cov}(X, Y) = 0$ og

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) = a^2 \sigma_X^2 + b^2 \sigma_Y^2$$

COR 2: Hvis X og Y er uavhengige stokastiske variable, så er $\text{Cov}(X, Y) = 0$ og

$$\text{Var}(aX - bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) = a^2 \sigma_X^2 + b^2 \sigma_Y^2$$

COR 3: Hvis $X_1, X_2, \dots, X_n$ er uavhengige stokastiske variable, så er

$$\begin{aligned}\text{Var}(a_1 X_1 + a_2 X_2 + \dots + a_n X_n) &= a_1^2 \text{Var}(X_1) + a_2^2 \text{Var}(X_2) + \dots + a_n^2 \text{Var}(X_n) \\ &= a_1^2 \sigma_{X_1}^2 + a_2^2 \sigma_{X_2}^2 + \dots + a_n^2 \sigma_{X_n}^2\end{aligned}$$

Generalisering

$$Y = \sum_{i=1}^n a_i X_i + b$$

$$E(Y) = \sum_{i=1}^n a_i E(X_i) + b$$

$$\begin{aligned} \text{Var}(Y) &= \sum_{i=1}^n a_i^2 \text{Var}(X_i) \\ &+ 2 \sum_{i=1}^n \sum_{j=1}^{i-1} a_i a_j \text{Cov}(X_i, X_j) \end{aligned}$$



Prosjektstyring



Total varighet av aktiviteter er $X + Y$, hva er $\text{Var}(X + Y)$?

$$\mu_X = E(X) = 1 \cdot 0.10 + 2 \cdot 0.60 + 3 \cdot 0.30 = 2.2$$

$$\mu_Y = E(Y) = 1 \cdot 0.10 + 2 \cdot 0.20 + 3 \cdot 0.30 + 4 \cdot 0.40 = 3.0$$

$$\begin{aligned} \sigma_X^2 = \text{Var}(X) &= E(X^2) - \mu_X^2 = 1^2 \cdot 0.10 + 2^2 \cdot 0.60 + 3^2 \cdot 0.30 - 2.2^2 \\ &= 5.2 - 4.84 = 0.36 \end{aligned}$$

$$\begin{aligned} \sigma_Y^2 = \text{Var}(Y) &= E(Y^2) - \mu_Y^2 \\ &= 1^2 \cdot 0.10 + 2^2 \cdot 0.20 + 3^2 \cdot 0.30 + 4^2 \cdot 0.40 - 3.0^2 \\ &= 10.00 - 9.00 = 1.00 \end{aligned}$$

$$\begin{aligned} \sigma_{XY} = \text{Cov}(X, Y) &= E(XY) - \mu_X \mu_Y = 1 \cdot 1 \cdot 0.03 + 2 \cdot 1 \cdot 0.05 + \dots + 3 \cdot 4 \cdot 0.15 \\ &\quad - 2.2 \cdot 3.0 = 6.76 - 6.6 = 0.16 \end{aligned}$$

$$E(X + Y) = E(X) + E(Y) = 2.2 + 3.0 = 5.2$$

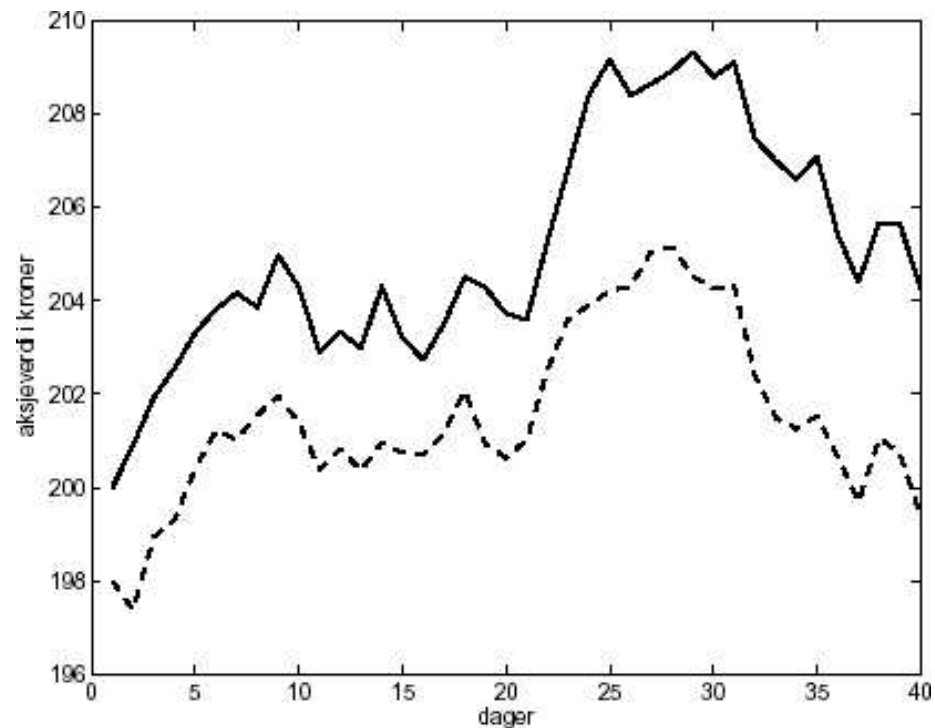
$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) + 2ab \text{Cov}(X, Y)$$

$$\begin{aligned} \text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y) \\ &= 0.36 + 1.00 + 2 \cdot 0.16 = 1.68 \end{aligned}$$

Aksjekurs

- Modifisert versjon av eksamen Juni 2004, oppgave 2b+c
- Selskapet Agderfrukt er notert på børsen. Vi antar at endringen X i verdien på en Agderfrukt-aksje i løpet av en dag har forventningsverdi $\mu_X = 0.15$ kroner og standardavvik $\sigma_X = 0.60$ kroner. Har du en aksje i Agderfrukt vil $X > 0$ bety fortjeneste, mens $X < 0$ er tap.
- Selskapet Trønderfrukt er også notert på børsen. Vi kaller endringen på en Trønderfrukt-aksje i løpet av en dag for Y , der Y har forventningsverdi $\mu_Y = 0.15$ kroner og standardavvik $\sigma_Y = 0.80$ kroner.
- Vi ser på aksjekursendringen på den samme dagen for Agderfrukt og Trønderfrukt og antar i dette punktet at aksjekursendringene X og Y er uavhengige.
- Idag er verdien til en Agderfrukt-aksje den samme som verdien til en Trønderfrukt-aksje. Vi ønsker å undersøke tre mulige strategier for aksjekjøp, der vi kjøper aksjer idag og selger imorgen.
 - i) Kjøp to aksjer i Agderfrukt.
 - ii) Kjøp en aksje i Agderfrukt og en aksje i Trønderfrukt.
 - iii) Kjøp to aksjer i Trønderfrukt.
- Dersom du vil ha minst mulig risiko for investeringen din, hvilken av de tre investeringsstrategiene over vil du velge? Begrunn svaret.

Aksjekurs, forts.



Figuren viser utviklingen av aksjekursen til Agderfrukt (stiplet) sammen med aksjekursen til Trønderfrukt (heltrukket).

Aksjekurs, forts.

- Kursendringen dag i for Agderfrukt kaller vi X_i , og vi antar at X_i har forventning $\mu_X = 0.15$ kroner og standardavvik $\sigma_X = 0.60$ kroner.
- Kursendringen dag i for Trønderfrukt kaller vi Y_i , og vi antar at Y_i har forventning $\mu_Y = 0.15$ kroner og standardavvik $\sigma_Y = 0.80$ kroner.
- Kursendringer for ulike dager antas å være uavhengige.
- Vi sammenlikner de to selskapene ved å måle differansen mellom de daglige kursendringene, $D_i = X_i - Y_i$, og ta gjennomsnitt. Vi ser på 10 dager og får $\bar{D} = \frac{1}{10} \sum_{i=1}^{10} D_i = \frac{1}{10} \sum_{i=1}^{10} (X_i - Y_i)$.
- Gir fi guren grunn til å tro at endringene i de to aksjekursene samme dag, X_i og Y_i , er uavhengige?
- Korrelasjonen mellom X_i og Y_i for disse to selskapene, $\rho(X_i, Y_i)$, er enten -0.5, 0.0 eller 0.5. Hvilken av disse verdiene virker mest rimelig fra fi guren? Begrunn kort.
- Hva blir forventningsverdi og varians for $\bar{D}$? Benytt verdien for korrelasjonen, $\rho(X_i, Y_i)$, som du valgte over.

Vinduskarmen

- Andersen skal bytte ut nedre del av en vinduskarm som er 2 meter lang. Han kutter to planker slik at de er 1 meter hver.
- Det viktigste er at den totale lengden på de to plankene er 2 meter (passer i vinduskarmen).
- Andersen ser for seg to måter å gjøre dette på:
 - a) Uavhengig: måle planke 1 og kutte, deretter måle planke 2 og kutte.
 - b) Avhengig: legge de to plankene oppå hverandre, måle og kutte begge samtidig (antar de to plankene da blir akkurat like lange).
- Hvilken av disse to måtene bør Andersen velge?