ISBN 9788770668699

Projekt 9.1 Regneregler for stokastiske variable – middelværdi, varians og spredning

Sætning 1: Regneregler for middelværdi, varians og spredning for diskrete stokastiske variable

Lad X være en stokastisk variabel med sandsynlighedsfordelingen:

Stokastisk variabel X	x_1	x_2	...	...	x_n
Sandsynlighed $P(X=x_i)$	p_1	p_2	...	...	p_n

Og lad Y være en stokastisk variabel med sandsynlighedsfordelingen:

Stokastisk variabel Y	y_1	y_2	...	...	y_m
Sandsynlighed $P(Y=y_i)$	q_1	q_2	...	...	q_m

og lad k være et vilkårligt reelt tal, der ikke varierer.
Da gælder følgende regneregler:

a) $E(k)=k$	f) $V(k)=0$
b) $E(X+k)=E(X)+k$	g) $V(X+k)=V(X)$
c) $E(k\cdot X)=k\cdot E(X)$	h) $V(aX)=a^2V(X)$
d) $E(X+Y)=E(X)+E(Y)$ og $E(X\cdot Y)=E(X)\cdot E(Y)$ når X og Y er uafhængige.	i) $V(X+Y)=V(X)+V(Y)$, og $V(X\cdot Y)=E(X^2)\cdot E(Y^2)-\left(E(X)\right)^2\cdot\left(E(Y)\right)^2$ når X og Y er uafhængige.
e) $V(X)=E(X^2)-\left(E(X)\right)^2$	

Forudsætninger der anvendes i beviset:
Vi anvender definitionerne:
Middelværdien $\mu=E(X)$ defineres som

$$\mu=E(X)=x_1\cdot p_1+x_2\cdot p_2+x_3\cdot p_3+\ldots+x_{n-1}\cdot p_{n-1}+x_n\cdot p_n\,.$$

Variansen $Var(X)$ defineres som

$$Var(X)=\left(x_1-\mu\right)^2\cdot p_1+\left(x_2-\mu\right)^2\cdot p_2+\left(x_3-\mu\right)^2\cdot p_3+\ldots+\left(x_{n-1}-\mu\right)^2\cdot p_{n-1}+\left(x_n-\mu\right)^2\cdot p_n\,.$$

Spredningen $\sigma\left(X\right)$ defineres som

$$\sigma\left(X\right)=\sqrt{Var(X)}\,.$$

I beviset anvendes flere steder, at summen af sandsynlighederne er 1: $p_1+p_2+\ldots+p_n=1$

Bevis for a):
Når den stokastiske variabel er en konstant k, så er sandsynlighedsfordelingen simpel

$X=x_i$	k
$P\left(X=x_i\right)$	1

Dvs. $E(X)=k\times 1=k$.

Bevis for b)
For en stokastisk variabel X har vi

$X=x_i$	x_1	x_2	x_3	...	x_{n-1}	x_n
$P\left(X=x_i\right)$	p_1	p_2	p_3	...	p_{n-1}	p_n

Dvs. der må gælde for

$X+k=x_i+k$	x_1+k	x_2+k	x_3+k	...	$x_{n-1}+k$	x_n+k
$P\left(X+k=x_i+k\right)$	p_1	p_2	p_3	...	p_{n-1}	p_n

Vi får derfor

$$E(X+k)=(x_1+k)\cdot p_1+(x_2+k)\cdot p_2+(x_3+k)\cdot p_3+\ldots+(x_{n-1}+k)\cdot p_{n-1}+(x_n+k)\cdot p_n=$$
$$x_1\cdot p_1+x_2\cdot p_2+x_3\cdot p_3+\ldots+x_{n-1}\cdot p_{n-1}+x_n\cdot p_n+k\cdot p_1+k\cdot p_2+k\cdot p_3+\ldots+k\cdot p_{n-1}+k\cdot p_n=$$
$$x_1\cdot p_1+x_2\cdot p_2+x_3\cdot p_3+\ldots+x_{n-1}\cdot p_{n-1}+x_n\cdot p_n+k\cdot(p_1+p_2+p_3+\ldots+p_{n-1}+p_n)=$$
$$x_1\cdot p_1+x_2\cdot p_2+x_3\cdot p_3+\ldots+x_{n-1}\cdot p_{n-1}+x_n\cdot p_n+k\cdot 1=$$
$$E(X)+k$$

Bevis for c)
Som i b) må der gælde for

$k\cdot X=k\cdot x_i$	$k\cdot x_1$	$k\cdot x_2$	$k\cdot x_3$	...	$k\cdot x_{n-1}$	$k\cdot x_n$
$P(k\cdot X=k\cdot x_i)$	p_1	p_2	p_3	...	p_{n-1}	p_n

Vi får derfor

$$\begin{aligned}E(k\cdot X)&=(k\cdot x_1)\cdot p_1+(k\cdot x_2)\cdot p_2+\ldots+(k\cdot x_n)\cdot p_n\\&=k\cdot(x_1\cdot p_1+x_2\cdot p_2+\ldots+x_n\cdot p_n)\\&=k\cdot E(X)\end{aligned}$$

Bevis for d)
Ved kast med to terninger, kunne $X+Y$ angive summen af øjnene. Men der behøver ikke at være lige mange udfald for *X* og *Y*. Forestil dig fx, at *X* angiver øjnene ved kast med en terning og *Y* angiver summen af øjnene ved kast med to terninger. Vi kan her få alle værdier fra 3 til 18. Værdien 12 kan her fremkomme som: 1+11, 2+10, 3+9, 4+8, 5+7, 6+6. Men de enkelte summer her harv forskellig sandsynlighed. Skal vi få resultatet 4+8 kræver det, at *X* giver 4, hvilket har sandsynligheden $\frac{1}{6}$, og at *Y* giver 8, hvilket har sandsynligheden $\frac{4}{36}$. Da det skal indtræffe samtidig, er sandsynligheden for at vi får udfaldet 4+8 lig med $\frac{1}{6}\cdot\frac{4}{36}=\frac{1}{54}$. Her udnyttes , at de to variable er uafhængige, så vi kan gange sandsynlighederne. Det virker som om, der er meget at holde styr på. Derfor er det godt, at vi har en så simpel formel! Vi gennemfører nu beviset med symbol-regning:

For den anden stokastiske variabel *Y* har vi

$Y=y_i$	y_1	y_2	y_3	...	y_{m-1}	y_m
$P(Y=y_i)$	q_1	q_2	q_3	...	q_{m-1}	q_m

Middelværdien af $X+Y$ beregnes ud fra formlen: Summen af alle led $(x_i+y_j)\cdot p_i\cdot q_j$ (jfr argumentet oevnfor)
Vi skaver os overblik: For summen $X+Y$ har vi,

$X+Y=x_i+y_j$	x_1	x_2	x_3	...	x_{n-1}	x_n
y_1	x_1+y_1					
y_2						
y_3						
...						
y_{m-1}						
y_m						x_n+y_m

ISBN 9788770668699

Vi summerer nu leddene rækkevis:
Dvs. at vi for hver af værdierne for *Y* får

$$(y_1+x_1)\cdot p_1\cdot q_1+(y_1+x_2)\cdot p_2\cdot q_1+\ldots+(y_1+x_n)\cdot p_n\cdot q_1+$$
$$(y_2+x_1)\cdot p_1\cdot q_2+(y_2+x_2)\cdot p_2\cdot q_2+\ldots+(y_2+x_n)\cdot p_n\cdot q_2+$$
$$\vdots$$
$$(y_m+x_1)\cdot p_1\cdot q_m+(y_m+x_2)\cdot p_2\cdot q_m+\ldots+(y_m+x_n)\cdot p_n\cdot q_m$$

multipliceres nu alle parenteserne ud for hver af værdierne for *Y*, og omarrangeres leddene får vi fra første række:

$$\left(x_1\cdot p_1\cdot q_1+x_2\cdot p_2\cdot q_1+\ldots+x_n\cdot p_n\cdot q_1\right)+\left(y_1\cdot p_1\cdot q_1+y_1\cdot p_2\cdot q_1+\ldots+y_1\cdot p_n\cdot q_1\right)=$$
$$\left(x_1\cdot p_1+x_2\cdot p_2+\ldots+x_n\cdot p_n\right)\cdot q_1+\left(y_1\cdot p_1\cdot q_1+y_1\cdot p_2\cdot q_1+\ldots+y_1\cdot p_n\cdot q_1\right)=$$
$$E(X)\cdot q_1+\left(y_1\cdot p_1\cdot q_1+y_1\cdot p_2\cdot q_1+\ldots+y_1\cdot p_n\cdot q_1\right)$$

Udfører vi det samme for næste række

$$\left(x_1\cdot p_1\cdot q_2+x_2\cdot p_2\cdot q_2+\ldots+x_n\cdot p_n\cdot q_2\right)+\left(y_2\cdot p_1\cdot q_2+y_2\cdot p_2\cdot q_2+\ldots+y_2\cdot p_n\cdot q_2\right)$$

Vi ser, at det giver

$$E(X)\cdot q_2+\left(y_2\cdot p_1\cdot q_2+y_2\cdot p_2\cdot q_2+\ldots+y_2\cdot p_n\cdot q_2\right)$$

Samlet får vi for alle rækkerne:

$$E(X)\cdot q_1+\left(y_1\cdot p_1\cdot q_1+y_1\cdot p_2\cdot q_1+\ldots+y_1\cdot p_n\cdot q_1\right)$$
$$E(X)\cdot q_2+\left(y_2\cdot p_1\cdot q_2+y_2\cdot p_2\cdot q_2+\ldots+y_2\cdot p_n\cdot q_2\right)$$
$$\ldots$$
$$E(X)\cdot q_m+\left(y_m\cdot p_1\cdot q_m+y_m\cdot p_2\cdot q_m+\ldots+y_m\cdot p_n\cdot q_m\right)$$

Nu summerer vi det hele, ved at summere lodret. De første led giver:

$$E(X)\cdot(q_1+q_2+\ldots+q_m)$$

Den lodrette sum af de første led i y-parenteserne giver:

$$y_1\cdot p_1\cdot q_1+y_2\cdot p_1\cdot q_2+\ldots+y_m\cdot p_1\cdot q_m=E(Y)\cdot p_1$$

Overvej selv dette!

dvs alle led giver:

$$E(X)\cdot(q_1+q_2+\ldots+q_m)+E(Y)\cdot(p_1+p_2+\ldots+p_n)=$$
$$E(X)\cdot 1+E(Y)\cdot 1=E(X)+E(Y)$$

slut på beviset for d i tilfældet med sum af to stokastiske variable.

Øvelse 1.

Gennemfør selv beviset vi tilfældet med produkt af to stokastiske variable. Vi anvender igen skemaet ovenfor, og at sandsynligheden for at få et produkt $x_i\cdot y_j$ som før er produktet af sandsynlighederne, når *X* og *Y* er uafhængige.

Anvend samme ide ved først at se på rækkerne, dernæst summere lodret.

ISBN 9788770668699

Bevis for e)

Vi tager udgangspunkt i definition på $V(X)$, og anvender også definitionen på $E(X)$:

$V(X) = (x_1-\mu)^2 \cdot p_1 + (x_2-\mu)^2 \cdot p_2 + (x_3-\mu)^2 \cdot p_3 + \ldots + (x_{n-1}-\mu)^2 \cdot p_{n-1} + (x_n-\mu)^2 \cdot p_n$.

I hvert led indgår der ”kvadratet på en toleddet størrelse”, som vi kan gange ud.

$$\begin{aligned}
V(X) &= (x_1-\mu)^2 \cdot p_1 + (x_2-\mu)^2 \cdot p_2 + (x_3-\mu)^2 \cdot p_3 + \ldots + (x_{n-1}-\mu)^2 \cdot p_{n-1} + (x_n-\mu)^2 \cdot p_n \\
&= \left(x_1^2 \cdot p_1 + \mu^2 \cdot p_1 - 2x_1 \cdot \mu \cdot p_1\right) + \left(x_2^2 \cdot p_2 + \mu^2 \cdot p_2 - 2x_2 \cdot \mu \cdot p_2\right) + \ldots + \left(x_n^2 \cdot p_n + \mu^2 \cdot p_n - 2x_n \cdot \mu \cdot p_n\right) \\
&= \left(x_1^2 \cdot p_1 + x_2^2 \cdot p_2 + \ldots + x_n^2 \cdot p_n\right) + \left(\mu^2 \cdot p_1 + \mu^2 \cdot p_2 + \ldots + \mu^2 \cdot p_n\right) - 2x_1 \cdot \mu \cdot p_1 - 2x_2 \cdot \mu \cdot p_2 - \ldots - 2x_n \cdot \mu \cdot p_n \\
&= \left(x_1^2 \cdot p_1 + x_2^2 \cdot p_2 + \ldots + x_n^2 \cdot p_n\right) + \mu^2 \cdot (p_1 + p_2 + \ldots + p_n) - 2 \cdot \mu \cdot (x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_n \cdot p_n) \\
&= E(X^2) + \mu^2 \cdot 1 - 2 \cdot \mu \cdot E(X) \\
&= E(X^2) + \mu^2 - 2 \cdot \mu \cdot \mu \\
&= E(X^2) - \mu^2 \\
&= E(X^2) - (E(X))^2
\end{aligned}$$

Gør selv nøje rede for hvert trin i omskrivningen!
Hermed er formel e) bevist.

Bevis for f)

Anvend punkt e):

$V(k) = E(k^2) - (E(k))^2 = k^2 - (k)^2 = 0$

hvor vi også har udnyttet punkt a) to gange.

Bevis for g)

Vi har:

$a \times X = a \times x_i$	$a \times x_1$	$a \times x_2$	$a \times x_3$	…	$a \times x_{n-1}$	$a \times x_n$
$P(a \times X = a \times x_i)$	p_1	p_2	p_3	…	p_{n-1}	p_n

Vi kunne sætte ind i definitionen på variansen og regne det igennem, men det er lettere nu at udnytte e):

$V(a \cdot X) = E\left((a \cdot X)^2\right) - (E(a \cdot X))^2 = E\left(a^2 \cdot X^2\right) - (E(a \cdot X))^2$

Udnyt nu punkt c) og gør udregningerne færdig.

ISBN 9788770668699

Bevis for h)

Vi udnytter igen punkt e):

$$V(X+b)=E\left((X+b)^2\right)-\left(E(X+b)\right)^2$$
$$=E(X^2+b^2+2\cdot b\cdot X)-\left(E(X)+b\right)^2$$
$$=E(X^2)+b^2+2\cdot b\cdot E(X)-\left(\left(E(X)\right)^2+b^2+2\cdot b\cdot E(X)\right)$$
$$=E(X^2)+b^2+2\cdot b\cdot E(X)-\left(E(X)\right)^2-b^2-2\cdot b\cdot E(X)$$
$$=E(X^2)-\left(E(X)\right)^2=V(X)$$

Bevis for i)

Vi skal vise, at $V(X+Y)=V(X)+V(Y)$, når X og Y er uafhængige**.**

Vi udnytter igen punkt e):

$$V(X+Y)=E\left(\left(X+Y\right)^2\right)-\left(E(X+Y)\right)^2$$
$$=E\left(X^2+Y^2+2\cdot X\cdot Y\right)-\left(E(X)+E(Y)\right)^2$$
$$=E(X^2)+E(Y^2)+E(2\cdot X\cdot Y)-\left(\left(E(X)\right)^2+\left(E(Y)\right)^2+2\cdot E(X)\cdot E(Y)\right)$$
$$=E(X^2)+E(Y^2)+2\cdot E(X)\cdot E(Y)-\left(E(X)\right)^2-\left(E(Y)\right)^2-2\cdot E(X)\cdot E(Y)$$
$$=E(X^2)-\left(E(X)\right)^2+E(Y^2)-\left(E(Y)\right)^2$$
$$=V(X)+V(Y)$$

hvilket var påstanden i i)

Øvelse 2

Vis reglen for varians af et produkt ud fra punkt e) og reglerne for middelværdi af et produkt.

Hermed er sætningen bevist!