MNFSIB1 STATISTIKK FOR SAMFUNNSVITERE
OBLIGATORISK ØVING 1

LØSNINGSFORSLAG

4. mars 2003

KAPITTEL 2

150

C:	20	60	60	70	90
D:	20	30	70	90	90

Ser på hvilken effekt en endring av kolonne 2 og 3 i sample C fra $(60, 60)$ til $(30, 90)$ vil ha på gjennomsnitt, median, "midrange", "range", varians og standardavvik.

a. $\bar{\text{C}} = 60$, $\bar{\text{D}} = 60 \Rightarrow$ Ingen effekt.

b. $\tilde{\text{C}} = 60$, $\tilde{\text{D}} = 70 \Rightarrow$ Medianen blir større.

c. Mode(C) = 60, n=2; Mode(D) = 90, n=2 $\Rightarrow$ Øker fra 60 til 90.

d. Midrange(C) $= \frac{20+90}{2} =$ Midrange(D) $= \frac{20+90}{2} = 55 \Rightarrow$ Ingen effekt.

e. Range(C) = Range(D) = 90-20 = 70 $\Rightarrow$ Ingen effekt.

f. $s^2 = \frac{1}{n-1}\sum(x-\bar{x})^2$

$s_C^2 = \frac{1}{4}[(-40)^2 + 0^2 + 0^2 + 10^2 + 30^2] = 650$
$s_D^2 = \frac{1}{4}[(-40)^2 + (-30)^2 + 30^2 + 10^2 + 30^2] = 1100$
$\Rightarrow$ Øker.

g. Standardavvik $= \sqrt{Varians} = s$
$s_C = \sqrt{650} = 25.5$
$s_D = \sqrt{1100} = 33.17$
$\Rightarrow$ Øker.

161

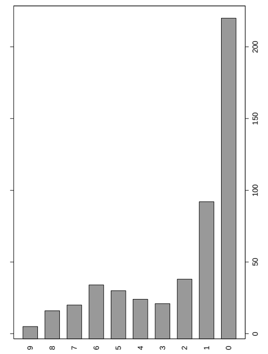
x	0	1	2	3	4	5	6	7	8	9
f	220	92	38	21	24	30	34	20	16	5

a. Konstruerer et histogram, se figur 1.
b. Gjennomsnitt: $\bar{x} = \frac{\sum xf}{\sum f} = \frac{994}{500} = \underline{\underline{1.988}}$.
Median: $\tilde{\text{f}} = \frac{1+\sum f}{2} = \frac{501}{2} = 250.5 \Leftrightarrow \tilde{\text{x}} = \underline{\underline{1}}$.
Mode: $\underline{\underline{0}}$ (har høyest frekvens).
Midrange: $\frac{L+H}{2} = \frac{0+9}{2} = \underline{\underline{4.5}}$.

c. Varians: $s^2 = \frac{1}{n-1}\sum(x-\bar{x})^2 = \underline{\underline{6.461}}$.
Standardavvik: $s = \sqrt{6.461} = \underline{\underline{2.542}}$.

d. $\frac{nk}{100} = \frac{500 \cdot 25}{100} = 125 \Rightarrow Q_1 = \underline{\underline{0}}$
$\frac{nk}{100} = \frac{500 \cdot 75}{100} = 375 \Rightarrow Q_3 = \underline{\underline{4}}$
$\frac{nk}{100} = \frac{500 \cdot 90}{100} = 450 \Rightarrow P_{90} = \underline{\underline{6}}$.

e. Midtkvartil: $\frac{Q_1+Q_3}{2} = \frac{4}{2} = \underline{\underline{2}}$.



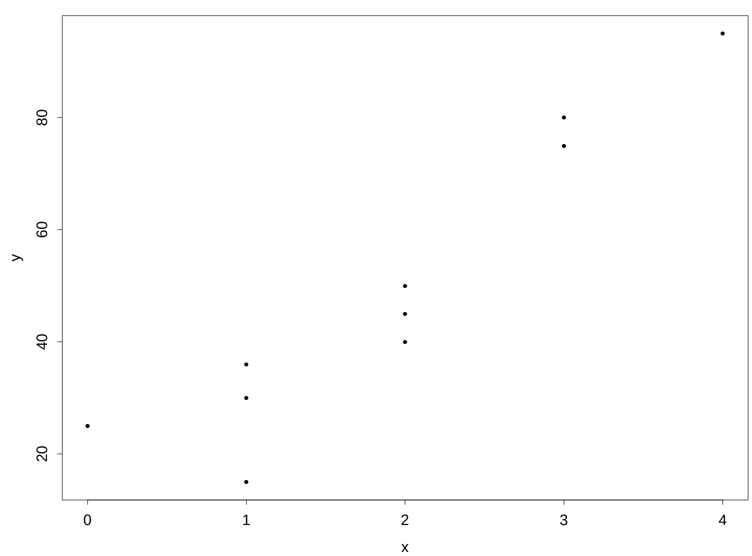
Figur 1: Histogram av x.

KAPITTEL 3

72

x	0	3	2	2	1	3	2	4	1	1
y	25	80	45	40	36	75	50	95	30	15

a. Scatterplot, se figur 2.



Figur 2: Scatterplot av x mot y.

b. Korrelasjonskoeffisient: $r = \frac{SS(xy)}{\sqrt{SS(x) \cdot SS(y)}}$
$SS(x) = \sum x^2 - \frac{1}{n}(\sum x)^2 = 49 - 36.1 = 12.9$
$SS(y) = \sum y^2 - \frac{1}{n}(\sum y)^2 = 30221 - 24108.1 = 6112.9$
$SS(xy) = \sum xy - \frac{1}{n}(\sum x \sum y) = 1196 - 932.9 = 263.1$
$r = \frac{263.1}{\sqrt{12.9 \cdot 6112.9}} = \underline{\underline{0.937}}$.

c. $\hat{Y} = b_0 + b_1 x, b = \frac{SS(xy)}{SS(x)} = \frac{263.1}{12.9} = 20.395$
$b_0 = \frac{1}{n}(\sum y - b_1 \sum x) = \frac{491 - 20.395 \cdot 19}{10} = 10.349$
$\Rightarrow \underline{\underline{\hat{Y} = 10.3 + 20.4x}}$.

d. En så høy korrelasjonskoeffisient med positivt fortegn antyder en sterk, positiv sammenheng mellom x og y, dvs. at en høy x-verdi vil medføre en høy y-verdi. Prediksjonsligningen er den linja som beskriver dataene best. b_0 er y-verdien dersom x settes lik null, altså skjæringspunktet med y-aksen, mens b_1 er stignigstallet til linja, dvs. den sier noe om hvor mye y øker når x øker.

KAPITTEL 4

126

		Opinion		
Employee	**Favor**	**Neutral**	**Opposed**	**Totals**
Male	800	200	500	1500
Female	400	100	500	1000
Totals	1200	300	1000	2500

a. $P(\text{mot}) = \frac{1000}{2500} = \underline{\underline{0.4}}$

b. $P(\text{kvinne}) = \frac{1000}{2500} = \underline{\underline{0.4}}$

c. $P(\text{mot}|\text{mann}) = \frac{500}{1500} = \underline{\underline{0.33}}$

d. $P(\text{mot og kvinne}) = \frac{500}{2500} = 0.2$
$P(\text{mot}) \cdot P(\text{kvinne}) = \frac{1000}{2500}\frac{1000}{2500} = 0.16$
$\Rightarrow$ "Mot" og "Kvinne" er avhengige siden $P(\text{mot og kvinne}) \neq P(\text{mot}) \cdot P(\text{kvinne})$.

128

			Jones		
Adams	**0**	**1**	**2**	**3**	**4**
0	0.0	1.0	2.0	3.0	4.0
1	0.1	1.1	2.1	3.1	4.1
2	0.2	1.2	2.2	3.2	4.2
3	0.3	1.3	2.3	3.3	4.3

A = minst én av selgerne får ikke solgt noe.

B = tilsammen eksakt tre salg.
C = Alle selgerne solgte like mange.
D = Adams solgte akkurat én.

a. $P(A) = \frac{4}{20} + \frac{4}{20} = \underline{\underline{\frac{2}{5}}}$

b. $P(B) = \frac{4}{20} = \underline{\underline{\frac{1}{5}}}$

c. $P(C) = \frac{4}{20} = \underline{\underline{\frac{1}{5}}}$

d. $P(D) = \frac{5}{20} = \underline{\underline{\frac{1}{4}}}$

e. $P(A \text{ og } B) = \frac{2}{20} = \underline{\underline{\frac{1}{10}}}$

f. $P(B \text{ og } C) = \underline{\underline{0}}$, summen 3 kan ikke fremkomme av likt antall hos begge.

g. $P(A \text{ eller } B) = P(A) + P(B) - P(A \text{ og } B) = \frac{2}{5} + \frac{1}{5} - \frac{1}{10} = \underline{\underline{\frac{1}{2}}}$

h. $P(B \text{ eller } C) = P(B) + P(C) - P(B \text{ og } C) = \frac{1}{5} + \frac{1}{5} - 0 = \underline{\underline{\frac{2}{5}}}$

i. $P(A|B) = \frac{P(A \text{ og } B)}{P(B)} = \frac{\frac{1}{10}}{\frac{1}{5}} = \frac{1}{10} \cdot \frac{5}{1} = \underline{\underline{\frac{1}{2}}}$

j. $P(B|D) = \frac{P(B \text{ og } D)}{P(D)} = \underline{\underline{\frac{1}{5}}}$

k. $P(C|B) = \frac{P(B \text{ og } C)}{P(B)} = \underline{\underline{0}}$

l. $P(\bar{A}) = 1 - P(A) = 1 - \frac{2}{5} = \frac{3}{5}$
$P(\bar{A} \text{ og } B) = \frac{2}{20}$
$P(B|\bar{A}) = \frac{P(\bar{A} \text{ og } B)}{P(\bar{A})} = \frac{\frac{1}{10}}{\frac{3}{5}} = \underline{\underline{\frac{1}{6}}}$

m. $P(\bar{A} \text{ og } C) = \frac{3}{20}$
$P(C|\bar{A}) = \frac{P(\bar{A} \text{ og } C)}{P(\bar{A})} = \frac{\frac{3}{20}}{\frac{3}{5}} = \underline{\underline{\frac{1}{4}}}$

n. $P(A \text{ eller } B \text{ eller } C) = P(A) + P(B) + P(C) - P(A \text{ og } B) - P(A \text{ og } C) = \frac{2}{5} + \frac{1}{5} + \frac{1}{5} - \frac{1}{10} - \frac{1}{20} = \underline{\underline{\frac{13}{20}}}$

o. $P(A) + P(B) = \frac{2}{5} + \frac{1}{5} = \frac{3}{5}$ mens $P(A \text{ eller } B) = \frac{1}{2}$.
At den ene ikke gjør noe salg utelukker ikke at summen kan bli 3 og omvendt $\Rightarrow$ ikke disjunkte.

p. $P(B) + P(C) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$
$P(B \text{ eller } C) = \frac{2}{5}$, se f. og h.
$\Rightarrow$ Hendelsene er disjunkte.

q. $P(B) + P(D) = \frac{1}{5} + \frac{1}{4} = \frac{9}{20}$
$P(B \text{ og } D) = \frac{1}{20}$
$P(B \text{ eller } D) = P(B) + P(D) - P(b \text{ og } D) = \frac{9}{20} - \frac{1}{20} = \frac{2}{5}$
$\Rightarrow$ Summen kan bli 3 også når Adams gjør ett salg, hendelsene er ikke disjunkte.

r. $P(A \text{ og } B) = \frac{1}{10}$
$P(A) \cdot P(B) = \frac{2}{25}$
$\Rightarrow$ Hendelsene er ikke uavhengige da $P(A) \cdot P(B) \neq P(A \text{ og } B)$.

s. $P(B \text{ og } C) = 0$
$P(B) \cdot P(C) = \frac{1}{25}$

$\Rightarrow$ Hendelsene er ikke uavhengige da $P(B) \cdot P(C) \neq P(B \text{ og } C)$.

t. $P(B \text{ og } D) = \frac{1}{20}$
$P(B) \cdot P(D) = \frac{1}{20}$
$\Rightarrow$ Hendelsene er uavhengige da $P(B) \cdot P(D) = P(B \text{ og } D)$.

130

Mynt A er en skev mynt, dvs at det ikke er lik sannsynlighet for å få hode og mynt når man kaster mynten. Mynt B er en vanlig mynt.

a. $P(head_A) = P(A) = 0.6$
$P(\bar{A}) = 0.4$
$P(head_B) = P(B) = P(\bar{B}) = 0.5$
Utfallsrommet blir $S = (A, B), (\bar{A}, B), (A, \bar{B}), (\bar{A}, \bar{B})$ og $n(S) = 4$.

b. $P(A \text{ og } B) = P(A) \cdot P(B) = 0.6 \cdot 0.5 = \underline{\underline{0.3}}$

c. To muligheter av i alt fire: $(A, \bar{B})$ eller $(\bar{A}, B)$.
$P(A \text{ og } \bar{B}) = P(A) \cdot P(\bar{B}) = 0.6 \cdot 0.5 = 0.3$
$P(\bar{A} \text{ og } B) = P(\bar{A}) \cdot P(B) = 0.4 \cdot 0.5 = 0.2$
$P[(A \text{ og } \bar{B}) \text{ eller } (\bar{A} \text{ og } B)] = P(A \text{ og } B) + P(\bar{A} \text{ og } B) = 0.3 + 0.2 = \underline{\underline{0.5}}$

d. $P(\bar{A} \text{ og } \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = 0.4 \cdot 0.5 = \underline{\underline{0.2}}$

e. $P[(A \text{ og } B)|A] = P(B|A)$
Utfallet av terning B uavhengig av terning A.
$P(B|A) = P(B) = \underline{\underline{0.5}}$

f. $P[(A \text{ og } B)|B] = P(A|B) = P(A) = \underline{\underline{0.6}}$

g. Her har vi gitt at eksakt én terning viser "head"(krone) også skal vi finne sannsynligheten for at det er terning A som viser krone. De to mulige utfallene er $S = (A, \bar{B}), (\bar{A}, B)$.
$P(A) = 0.6, P(\bar{A}) = 0.4, P(B) = P(\bar{B}) = 0.5, P(A \text{ og } \bar{B}) = 0.3, P(\bar{A} \text{ og } B) = 0.2$, øvrige utfall er utelukket.
$P(A|\text{én krone}) = \frac{P(A \text{ og } \bar{B})}{P(\text{én krone})} = \frac{P(A \text{ og } \bar{B})}{P(A \text{ og } \bar{B}) + P(\bar{A} \text{ og } B)} = \frac{0.3}{0.3+0.2} = \underline{\underline{0.6}}(= P(A))$.

KAPITTEL 5

104

Vi har en undersøkelse av kvinners lesevaner der 90% sa de aldri hadde lest en utgave av "Vogue". Vi antar at denne informasjonen er riktig og vil finne sannsynligheten for at det i et tilfeldig utvalg på tre kvinner vil være mindre enn to kvinner som har lest bladet. Dvs. at ingen eller en av de tre kvinnene har lest bladet=hendelse A. Vi må finne $P(A)$.
Vi har altså en binomisk fordeling siden en kvinne enten har lest bladet eller ikke, $p = 0.1, (1-p) = q = 0.9, n = 3$.
$p(x) = \begin{pmatrix} n \\ x \end{pmatrix} \cdot p^x \cdot q^{n-x}$
$p(1) = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot (0.1) \cdot (0.9)^2 = 0.243$

$p(0) = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \cdot (0.1)^0 \cdot (0.9)^3 = 0.729$
$P(A) = p(0) + p(1) = 0.243 + 0.729 = \underline{\underline{0.972}}$

118

a. x = antall riktige svar, binomisk fordelt:
$b(x;n,p) = P(x) = \begin{pmatrix} n \\ x \end{pmatrix} \cdot p^x \cdot q^{n-x}$, der $n = 10, p = \frac{1}{5} = 0.2, q = 1 - p = 0.8$.

b. $P(7) = \begin{pmatrix} 10 \\ 7 \end{pmatrix} \cdot p^7 \cdot q^3 = \underline{\underline{0.0008}}$. Dvs. 0.1% sjanse for sju rette.

c. Seks eller flere rette svar: $P(x \geq 6) = P(6) + P(7) + P(8) + P(9) + P(10) \approx \underline{\underline{0.007}}$.

d. Sannsynligheten for å få sju rette av ti ved helt tilfeldig gjetning er svært liten så han har nok ikke tippet helt tilfeldig.

KAPITTEL 6

100

Vi har en x som er normalfordelt med $\mu = 900$ og $\sigma = 75$. Dersom ikke mer enn 10% av lyspærene skal ryke, når må lyspærene skiftes ut?

Vi vil finne x slik at maksimalt 10% av lyspærene ryker. Dvs. at vi skal finne en k slik at $P(x < k) = 0.1$. Vi standardiserer slik at vi får $P(\frac{x-900}{75} < \frac{k-900}{75}) = P(Z < \frac{k-900}{75}) = 0.1$. Da får vi fra tabell at

$$\begin{aligned} \frac{k-900}{75} &= -1.28 \\ \Rightarrow k &= \underline{\underline{804}} \end{aligned}$$

Så lyspærene må skiftes ut innen 804 timers bruk dersom maks 10% av pærene skal ryke.

108

80% av kundene til et firma vil ikke trenge reparasjoner i løpet av de første to åra etter kjøpet. En studie viser at 70 av totalt 100 holdt i to år uten behov for reparasjoner. Vi skal finne sannsynligheten for at 70 eller færre trenger reparasjoner. Vi har en binomisk fordeling, maskina trenger enten reparasjoner eller ikke.

x = antall maskiner av våre 100 som ikke trenger reparasjoner i løpet av de første to åra. x er binomisk fordelt, enten trenger maskina reparasjon eller så gjør den det ikke.

$P(x) = \begin{pmatrix} n \\ x \end{pmatrix} \cdot p^x \cdot q^{n-x}$, der $n = 100, p = 0.8, q = 1 - p = 0.2$.

Vi vil finne $P(x \leq 70)$. Det vil her bli veldig mye arbeid å regne ut vha. den binomiske fordelinga, da må vi summere for $x = 0, 1, \ldots, 70$. Vi tilnærmer istedet den binomiske fordelinga med en normalfordeling. Det kan vi gjøre dersom np og npq er større enn 5. Her har vi $np = 100 \cdot 0.8 = 80$ og $npq = 1000 \cdot 0.8 \cdot 0.2 = 16$, så vi kan godt tilnærme med normalfordeling, vi kaller den normalfordelte variabelen for y. y er da normalfordelt med $\mu = 80$ og $\sigma = 4$. Siden vi tilnærmer den binomiske fordelinga med en normalfordeling så må vi finne $P(y \leq 70.5)$, og ikke $P(y \leq 70)$,se kapittel 6.5.

$P(x \leq 70) = P(y \leq 70.5) = P(z \leq \frac{70.5-80}{4}) = P(z \leq -2.375)$, slår opp i tabell og får
$P(z \leq -2.375) = 0.5 - 0.4912 = \underline{\underline{0.0088}}$

Dvs. at sannsynligheten for at 70 maskiner eller færre ikke trenger reparasjoner er lik $8.8 \cdot 10^{-3}$.

KAPITTEL 7

48

Baggasjevekten til passasjerer på et fly er normalfordelt med forventningsverdi $20lb$ og varians 4^2lb. Dersom grensen til flyselskapet er på $2125lb$, hva er sannsynligheten for at grensen overskrides når det er 100 passasjerer?

x=vekten til en passasjers bagasje. Vi må finne $P(\sum x \geq 2125) = P(\bar{x} \geq 21.25)$. Vi vil derfor finne sannsynlighetsfordelinga til $\bar{x}$.
Gjennomsnittet er summen av normalfordelte variable og er derfor selv normalfordelt.

$\bar{x}$ er normalfordelt med $\mu_{\bar{x}}$ og $\sigma_{\bar{x}}^2$ der $\mu_{\bar{x}} = \mu = 20$ og $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{10} = 0.4$.

$P(\bar{x} \geq 21.25) = P(z \geq \frac{21.25-20}{0.4}) = P(z \geq 3.125)$, som vha. tabell blir:
$P(z \geq 3.125) = 0.5 - 0.4991 = \underline{\underline{0.0009}}$.

Dvs. at sannsynligheten for at grensa til flyselskapet overskrides ved 100 passasjerer er 0.09%.

SPØRSMÅL RETTES TIL danielka@stud.math.ntnu.no, olefridt@stud.math.ntnu.no eller oyvinert@stud.math.ntnu.no