TMA4245 Statistikk
Vår 2013

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Øving nummer 2, blokk I
Løsningsskisse

Oppgave 1 En mynt kastes 2 ganger. La Z være antall kroner i første kast, og W være antall kroner i to kast.

Utfallsrom: $S = \{\text{KK}, \text{KM}, \text{MK}, \text{MM}\}$. $P(\text{K}) = 0.4, P(\text{M}) = 0.6$

a) $f(z,w) = P(Z = z, W = w)$, f.eks.:

$$f(1,0) = P(Z = 0, W = 1) = P(\text{MK}) = P(\text{M})\,P(\text{K}) = 0.4 \cdot 0.6 = 0.24$$

	$f(z,w)$ \ w	0	1	2
z	0	0.36	0.24	0
	1	0	0.24	0.16

Videre får vi $f_W(w) = \sum_{z=0}^{1} f(z,w)$, f.eks $f_W(0) = 0.36 + 0 = 0.36$

w	0	1	2
$f_W(w)$	0.36	0.48	0.16

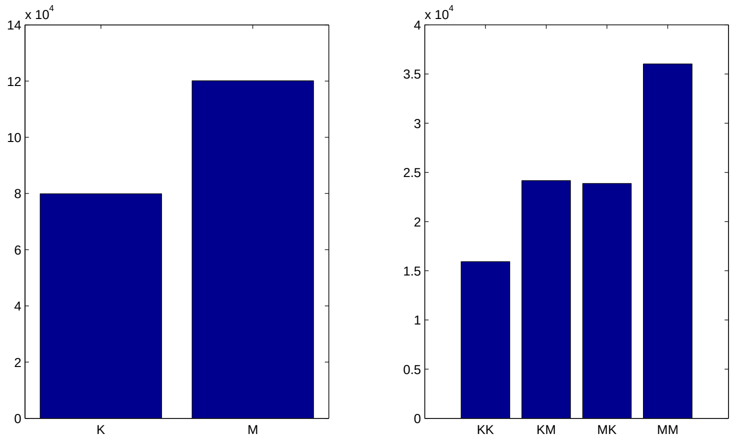
Vi får også $f_Z(z) = \sum_{w=0}^{2} f(z,w)$, f.eks $f_Z(0) = 0.36 + 0.24 + 0 = 0.6$

z	0	1
$f_Z(z)$	0.6	0.4

Dette gir $P(\text{minst 1 krone }) = P(W \geq 1) = \sum_{w=1}^{2} f_W(w) = 0.48 + 0.16 = 0.64$

b) En grei måte å sette sannsynligheten for å kaste K som en input parameter er

- Fjern linje 18 i koden, *P_K = 0.5;* som setter $P(K) = 0.5$ fast
- Sett P_K som input i funksjonen, da vil den f.eks. se slik ut: *throw_coin(throws, P_K, display)*. For f.eks. å kaste myntene 100000 ganger hvor $P(K) = 0.4$ kaller man da *throw_coin(100000, 0.4, 1)* ('display'=1 vil plotte histogram).



Figur 1: Eksempel barplot med 100000 myntkast med 2 mynter.

c) I figur 1 ser vi en realisasjon av 100000 simulerte kast med de to myntene. Tabellene under viser hvor mange ganger vi har kastet K og M, og hvor mange ganger de ulike kombinasjonene KK, KM, MK og MM har blitt kastet.

K	M
79842	120158

KK	KM	MK	MM
15867	24114	23994	36025

Vi kan sammenligne med a) ved å beregner andelen av tilfeller:

- Ser på $f(z,w)$, f.eks. for $z=0$ kroner i første kast og $w=0$ kroner i to kast (bare i tilfellet MM):

$$\frac{\#\text{ ganger MM}}{\text{Totalt }\#\text{ kast av to mynter}} = \frac{36025}{100000} = 0.3603 \approx 0.36 = f(0,0)$$

  Tilsvarende gir

		w		
		0	1	2
z	0	$\frac{36025}{100000} = 0.3603$	$\frac{23994}{100000} = 0.2399$	0
	1	0	$\frac{24114}{100000} = 0.2411$	$\frac{15867}{100000} = 0.1587$

  som er tilnærmet likt tabellen for $f(z,w)$ i Oppgave 4a).

- Ser på $f_W(w)$, f.eks. for $w=1$ kroner i to kast (i tilfellene KM og MK):

$$\frac{\#\text{ ganger KM} + \#\text{ ganger MK}}{\text{Totalt }\#\text{ kast av to mynter}} = \frac{24114+23994}{100000} = 0.4811 \approx 0.48 = f_W(1)$$

Tilsvarende gir

w	0	1	2
Andel	$\frac{36025}{100000} = 0.3603$	$\frac{24114+23994}{100000} = 0.4811$	$\frac{15867}{100000} = 0.1587$

som er tilnærmet likt tabellen for $f_W(w)$ i Oppgave 4b).

- Ser på $f_Z(z)$, f.eks. for $z = 1$ kroner i første kastet (i tilfellene KM og KK):

$$\frac{\text{\# ganger KM + \# ganger KK}}{\text{Totalt \# kast av to mynter}} = \frac{24114 + 15867}{100000} = 0.3998 \approx 0.40 = f_Z(1)$$

Tilsvarende gir

z	0	1
Andel	$\frac{23994+36025}{100000} = 0.6002$	$\frac{24114+15867}{100000} = 0.3998$

som er tilnærmet likt tabellen for $f_Z(z)$ i Oppgave 4c).

- Andelen for minst en krone er (for tilfellene KK, KM, MK)

$$\frac{\text{\# ganger KK + \# ganger KM + \# ganger MK}}{\text{Totalt \# kast av to mynter}} = \frac{15867 + 24114 + 23994}{100000} = 0.6398 \approx 0.64$$

som er tilnærmet likt svaret i Oppgave 4d).

Vi ser at simuleringen gir tilnærmet like svar som i Oppgave 4.

Oppgave 2

Det er 100 lodd: 1 lodd gir gevinst A, 1 lodd gir gevinst B og 98 gir ingen gevinst. Per trekker 4 lodd uten tilbakelegging. La **A** være at Per vinner gevinst A, **B** være at Per vinner gevinst B og **C** være at Per vinner minst en gevinst.

$$P(A) = \frac{\binom{1}{1}\binom{99}{3}}{\binom{100}{4}} = \underline{\underline{0.04}}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{\binom{1}{1}\binom{1}{1}\binom{98}{2}}{\binom{100}{4}}}{\frac{\binom{1}{1}\binom{99}{3}}{\binom{100}{4}}} = \frac{\binom{1}{1}\binom{1}{1}\binom{98}{2}}{\binom{1}{1}\binom{99}{3}} = \underline{\underline{0.0303}}$$

$$P((A \cap B)|C) = \frac{P(A \cap B \cap C)}{P(C)} = \frac{P(A \cap B)}{P(C)}$$

$$= \frac{\binom{1}{1}\binom{1}{1}\binom{98}{2}}{\binom{1}{1}\binom{1}{0}\binom{98}{3} + \binom{1}{0}\binom{1}{1}\binom{98}{3} + \binom{1}{1}\binom{1}{1}\binom{98}{2}} = \underline{\underline{0.0103}}$$

Oppgave 3

Definer

M : Mann
K : Kvinne
F : Fargeblind

med oppgitte sannsynligheter

$P(M) = 0.5$
$P(K) = 0.5$
$P(F|M) = 0.05$
$P(F|K) = 0.0025.$

Vi skal beregne $P(M|F)$, og bruker Bayes regel:

$$
\begin{aligned}
P(M|F) &= \frac{P(M \cap F)}{P(F)} \\
&= \frac{P(M)\,P(F|M)}{P(M)\,P(F|M) + P(K)\,P(F|K)} \\
&= \frac{0.5 \times 0.05}{0.5 \times 0.05 + 0.5 \times 0.0025} \\
&= 0.952.
\end{aligned}
$$

Oppgave 4

Det er enklest å se på de komplementære hendelsene A' og B':

$$
\begin{aligned}
P(A') &= P(\text{mynt}|\text{falsk}) \cdot P(\text{falsk}) + P(\text{mynt}|\text{ekte}) \cdot (\text{ekte}) \\
&= 0 + \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{1}{4}
\end{aligned}
$$

Samme verdi gjelder også for $P(B')$, dvs. $P(B') = \frac{1}{4}$.

Videre har vi

$$P(A' \cap B') = P(B'|A')P(A') = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

(Hvis 1. kast er mynt, er den ekte mynten trukket først. Dermed er første tallet her $\frac{1}{2}$).

Dermed er $P(A' \cap B')$ ikke lik $P(A') \cdot P(B')$, slik at A' og B' er avhengige. Da er også A og B avhengige.

Hvis ikke en ser på komplementære hendelser, blir regningen som følger.

$$
\begin{aligned}
P(A) &= P(\text{kron} \cap \text{falsk}) + P(\text{kron} \cap \text{ekte}) \\
&= P(\text{kron}|\text{falsk})P(\text{falsk}) + P(\text{kron}|\text{ekte})P(\text{ekte}) \\
&= 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{3}{4}.
\end{aligned}
$$

De to kastene er like, så $P(B) = P(A)$.

$$
\begin{aligned}
P(A \cap B) &= P(A \cap B \cap \text{ekte}) + P(A \cap B \cap \text{falsk}) \\
&= P((A \cap B)|\text{ekte})P(\text{ekte}) + P((A \cap B)|\text{falsk})P(\text{falsk}) \\
&= \left(\frac{1}{2}\right)^2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \\
&= \frac{5}{8}.
\end{aligned}
$$

Konklusjonen er den samme, $P(A)P(B) \neq P(A \cap B)$, dermed er hendelsene avhengige.

Oppgave 5

a) Vi lar X være antall normenn som har sykdommen. X vil da være binomisk fordelt med parametere $n = 3 \cdot 10^6$ og $p = 0.003$. Dermed er

$$E(X) = pn = 0.003 \cdot 3 \cdot 10^6$$

.

Vi lar Y være antall av de syke som ikke blir avslørt i undersøkelsen. Gitt at det er X som er syke vil Y være binomisk fordelt med parametere $n = X$ og $p = 0.01$. Dermed er

$$E(Y) = pX$$

. Hvis $X = \mathrm{E}(X) = 9000$ er

$$E(Y) = 9000 \cdot 0.01 = 90$$

.

b)

$$
\begin{aligned}
P(D^c|A) &= \frac{P(A|D^c)P(D^c)}{P(A|D)P(D) + P(A|D^c)P(D^c)} \\
&= \frac{0.005 \cdot 0.997}{0.99 \cdot 0.003 + 0.005 \cdot 0.997}
\end{aligned}
$$

$$P(D|A^c) = \frac{P(A^c|D)P(D)}{P(A^c)}$$
$$= \frac{P(A^c|D)P(D)}{P(A^c|D)P(D) + P(A^c|D^c)P(D^c)}$$

c) Sannsynligheten for B gitt at en person ikke har sykdommen kaller vi p_B. Da er sannsynlighet for B gitt at en person har sykdommen $10p_B$

$$P(D^c|A \cap B) = 1 - P(D|A \cap B)$$

$$P(D|A \cap B) = \frac{P(B|D \cap A)P(A|D)P(D)}{P(B|D \cap A)P(A|D)P(D) + P(B|D^c \cap A)P(A|D^c)P(D^c)}$$
$$\frac{10p_B P(A|D)P(D)}{10p_B P(A|D)P(D) + p_B P(A|D^c)P(D^c)}$$
$$\frac{10p_B P(A|D)P(D)}{p_B(10P(A|D)P(D) + P(A|D^c)P(D^c))}$$
$$\frac{10P(A|D)P(D)}{(10P(A|D)P(D) + P(A|D^c)P(D^c))}$$