Hold 09 mach: Opg3

Ali Kadum Hassan, Frederik Henriques Altmann, Gustav Mark-Hansen

December 17, 2021

1

a

Pmf'en er geometrisk fordelt, da vi tæller antal forsøg indtil succes [table 1. EPT]. Dvs. X antager værdierne $Support\{1, 2...\infty\}$, og kan derfor opskrive vores pmf:

$$p(x) = \frac{1}{2^x}$$

b

Hvis det koster C at spille, findes f for gevinsten :

$$f_x(x) = 2^x - C$$

c

Bevis [EPT Def: 9]:

$$\sum_{x=supp(x)}^{\infty} |f_x(x)| P(X = x) = \infty$$

$$\sum_{x=supp(x)}^{\infty} |f_x(x)| P(X = x) \geq \sum_{x=supp(x)}^{\infty} f_x(x) P(X = x)$$

$$\sum_{x=supp(x)}^{\infty} |2^x - C| \frac{1}{2^x} \geq \sum_{x=supp(x)}^{\infty} (2^x - C) \frac{1}{2^x}$$

$$\sum_{x=supp(x)}^{\infty} (2^x - C) \frac{1}{2^x} \Rightarrow \sum_{x=supp(x)}^{\infty} 1 - \frac{C}{2^x} = \infty$$

Da dette er opfyldt og $\sum_{x=supp(x)}^{\infty}(2^x - C)\frac{1}{\frac{1}{2}^x} = \infty$ og man kan ikke finde en middelværdi.

2

a

p(x, y)	y = 1	y = 1	y = 3
x = 1	0.1	0.1	z
x = 1	0.3	0.05	0.05

Da $P(X = x, Y = y) = 1$ [EPT def. 1], kan vi udregne z ved:

$$1 = 0.1 + 0.1 + 0.3 + 0.05 + 0.05 + z = 0.6 + z \quad (1)$$
$$z = 1 - 0.6 = 0.4 \quad (2)$$

b

Beregn den forventede gevinst [EPT Rem: 18]

$$E[X + Y] = E[X] + E[Y] \quad (3)$$
$$E[X] = -1 * 0.6 + 1 * 0.4 = -0.2 \quad (4)$$
$$E[Y] = -1 * 0.4 + 1 * 0.15 + 3 * 0.45 = 1.1 \quad (5)$$
$$E[X + Y] = 1.1 - 0.2 = 0.9 \quad (6)$$

c

beregn kovariansen mellem X og Y. Er X og Y uafhængige? \

$$EPTDef : 19]$$

COV(X,Y) := E(XY)-E(x)E(y)

P(XY=-3)=0.4, P(XY=-1)=0.4, P(XY=1)=0.15, P(XY=3)=0.05

E(XY)=(-3)*0.4+(-1)*0.4+1*0.15+3*0.05 $\Rightarrow -1.03 E(x)$ og $E(y)$ kender vi fra sidste opgave: $E(x) = -0.2, E(y) = 1.1$

$$COV(X,Y) = (-1.03) - (-0.2) * 1.1 \Rightarrow -1.08$$

Hvilket betyder at de er afhængige af hinanden, da $cov \neq 0$ [Rem 22 EPT]

3

a

Vi ved at hvis A har et udfald som er givet, har den altid en sandsynlighed 1 betinget at det er udfaldet.

CPT	A=1	A=0
P(T=1 givet A)	0.998	a
P(T=0 givet A)	b	0.993
sum	1	1

Derfor må $a = 1 - 0.993$ og $b = 1 - 0.998$.

CPT	A=1	A=0
P(T=1 givet A)	0.998	0.007
P(T=0 givet A)	0.002	0.993
sum	1	1

Sandsynligheden for at have en peanut allergi er $P(A = 1) = 0.01$, 1%, samt $P(A = 0) = 0.99$, 99% for ikke at have allergien.

Da

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

$$P(AB) = P(A|B) \cdot P(B)$$

kan en **pmf** beregnes.

PMF	A=1	A=0	sum
T=1	0.00998	0.00693	0.01691
T=0	0.00002	0.98307	0.98309
sum	0.01	0.99	1

Summen af rækkerne er så hhv. $P(T = 1)$ og $P(T = 0)$.

b

Sandsynligheden for ikke at have allergien givet en negativ test er:

$$P(A = 0|T = 0) = \frac{P(A = 0, T = 0)}{P(T = 0)} = \frac{0.98307}{0.98309} = 0.9999797$$

c

Sandsynligheden for at have allergien givet en positiv test er:

$$P(A = 1|T = 1) = \frac{P(A = 1, T = 1)}{P(T = 1)} = \frac{0.00998}{0.01} = 0.998$$

4

a

$$p_i(x) = \begin{cases} \frac{1}{365} & if \quad x \in \{1, \cdots, 365\} \\ 0 & if \quad x \notin \{1, \cdots, 365\} \end{cases} \tag{7}$$

b

`ER DET GODT NOK FORKLARET?`
Da $x_1, x_2, \cdots, x_n$ er disjunkte kan $P(x_1, x_2, \cdots, x_n)$ udtrykkes som produktet af sandsyndligheden for hvert udfald. **mangler citation**.

$$\begin{aligned} P(x_1, x_2, \cdots, x_n) &= P(X_i = x_1) \cdot P(X_i = x_2) \cdots P(X_i = x_n) && (8) \\ &= p_i(X_i = x_1) \cdot p_i(X_i = x_2) \cdots p_i(X_i = x_n) && (9) \\ p_i(x) &= \begin{cases} \frac{1}{365} & if \quad x \in \{1, \cdots, 365\} \\ 0 & ellers \end{cases} && (10) \\ P(x_1, x_2, \cdots, x_n) &= \begin{cases} \frac{1}{365^n} & hvis \quad (x_1, x_2, \cdots, x_n) \in \{1, \cdots, 365\} \\ 0 & ellers \end{cases} && (11) \end{aligned}$$

c

$$\begin{aligned} \forall x, \quad p(x) \geq 0 && (12) \\ \sum_x p(x) = 1 && (13) \\ && (14) \end{aligned}$$

$p_2(x)$ er et produkt af to muglige faktorer 0 og $\frac{1}{365}$.

$$0 \geq 0 \quad \frac{1}{365} \geq 0$$

x_1 kan antage 365 udfald, x_2 365, og samme for resten op til x_n. Derfor må pmf; $P(x_1, x_2, \cdots, x_n)$ have $365 \cdot \overset{n}{\cdots} \cdot 365 = 365^n$ udfald.

$$\sum_{(x_1,x_2,\cdots,x_n)\in\{1,\cdots,365\}} p(x_1, x_2, \cdots, x_n) = \sum_x \frac{1}{365^n} = 365^n \cdot \frac{1}{365^n} = 1$$

d

Sandsynligheden for nogen i en gruppe, dvs. 2 eller mere, har fødseldag på samme dag er det omvendte af at ingen i gruppen har fødseldag på samme dag, dvs. 1 eller mindre.

$$B = \text{"}\geq 2 \text{ har fødselsdag på samme dag"} \tag{15}$$

$$B^C = \text{"}\leq 1 \text{ har fødselsdag på samme dag"} = \text{"Alle har en unik fødselsdag"} \tag{16}$$

Ethvert udfald vil ligge i enten B eller B^C, `ER DET TRIVIELT AT B, BC ER DISJUNKTE?`

$$P(x_1, x_2, \cdots, x_n \in B \cup B^C) = 1$$

$$P(x_1, x_2, \cdots, x_n \in B) = 1 - P(x_1, x_2, \cdots, x_n \in B^C)$$

Mængden af dage hvor man kan have en unik fødselsdag er 365 hvis der blot er én person. Hvis der er to personer er der 365 dage for den første og 364 for den anden, da den anden person ikke kan have fødselsdag på samme dag som den første. Det giver $365 \cdot 364$ kombinationer. Det gælder generelt at antalet af kombinationer for n personer er:

$$365 \cdot (365 - 1) \cdots (365 - n + 1)$$

Eller omskrevet:

$$\frac{365!}{(365 - n)!}$$

Vi kender sandsyndligheden for et specifikt udfald for n personer, ved \$P(x1,x2,···,xn). Antages det at alle har rigtige fødselsdage, kan sandsynligheden for at n personer alle har unikke fødselsdage ift. hindanden.

$$P(x_1, x_2, \cdots, x_n \in B^C) = \frac{365!}{(365 - n)! \cdot 365^n}$$

Dvs.

$$P(x_1, x_2, \cdots, x_n \in B) = 1 - P(x_1, x_2, \cdots, x_n \in B^C)$$

$$P(x_1, x_2, \cdots, x_n \in B) = 1 - \frac{365!}{(365-n)! \cdot 365^n}$$

Skrevet i R (muligvis med forstærkede afrundingsfejl):

```
p = \(n) 1 - prod(c((365-n):365)/365)
p(10) # 0.1411414
p(20) # 0.4436883
p(50) # 0.974432
```