E6 opgavesæt 1

Jogvan M. Poulsen

9. oktober 2000

Opgave 1

Vi har to kasser med kugler, nummereret $1, \cdots, 5$ i den ene og $1, \cdots, 7$ i den anden kasse. En kugle trækkes tilfældigt fra hver kasse. Vi lader X_1 hhv. X_2 betegne nummeret på kuglen trukket fra kasse 1 hhv. 2. Det antages, at X_1 og X_2 er stokastisk uafhængige.

Opgave a Anfør udfaldsrum og den simultane sandsynlighedsfordeling for (X_1, X_2).

Vi betragter systemet som om der var tale om to terninger med hhv. fem og syv sider. Udfaldsrummet E må derfor være

$$
\begin{array}{c|ccccccc}
5 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
4 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
3 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
 & 1 & 2 & 3 & 4 & 5 & 6 & 7
\end{array}
$$

$$E = \{1, \cdots, 5\} \times \{1, \cdots, 7\} \qquad , |E| = 35$$

hvor alle udfald af (X_1, X_2) er lige sandsynlige. Den simultane sandsynlighedsfordeling er derfor en ligefordeling over hele udfaldsrummet E

$$P(X_1, X_2) = \frac{1}{|E|} = \frac{1}{35}$$

Opgave b Angiv sandsynlighedsfordelingen (ved sin sandsynlighedsfunktion) for de afledte stokastiske variable $M = min(X_1, X_2)$ og $S = max(X_1, X_2)$

Sandsynlighedsfordelingen af M, som er $P(min(X_1, X_2))$, er alle de mulige hendelser beskrevet før, men hvor det endelige udfald registeres som den kugle der viser mindst øjne. Udfaldene er fordelt på følgende måde

$$
\begin{array}{ll}
P(1) = {}^{11}\!/_{35} & P(2) = {}^{9}\!/_{35} \\
P(3) = {}^{7}\!/_{35} & P(4) = {}^{5}\!/_{35} \\
P(5) = {}^{3}\!/_{35} &
\end{array}
$$

$$P(M = m) = \frac{7 - m + 5 - m + 1}{35} = \frac{13 - 2m}{35} \qquad , m = (1, \cdot, 5)$$

Tilsvarende er det for S, hvor det bare den kugle der viser højst øjne, der angiver S's værdi

$$
\begin{array}{ll}
P(1) = {}^{1}\!/_{35} & P(2) = {}^{3}\!/_{35} \\
P(3) = {}^{5}\!/_{35} & P(4) = {}^{7}\!/_{35} \\
P(5) = {}^{9}\!/_{35} & P(6) = {}^{5}\!/_{35} \\
P(7) = {}^{5}\!/_{35} &
\end{array}
$$

$$P(S = s) = \begin{cases} \frac{2s-1}{35} & s = 1, \cdots, 5 \\ \frac{5}{35} & s = 6, 7 \end{cases}$$

Opgave c Udfaldet af forsøget gav m som den mindste værdi. Hvad er sandsynligheden for at dette nummer blev trukket fra kasse 1? Anfør denne sandsynlighed for alle værdier af m eller som funktion af m.

Sandsynligheden for at m er værdien på kuglen fra kasse 1 er den betingede sandsynlighed for

$$P(X_1 = m|M = m) = \frac{P(X_1 = m \cap M = m))}{P(M = m)}$$

Her udregner vi den betingede sandsynlighed ud

$$\begin{aligned}
P(X_1 = 1|M = 1) &= \frac{7/35}{11/35} = \frac{7}{11} \\
P(X_1 = 2|M = 2) &= \frac{6/35}{9/35} = \frac{6}{9} \\
P(X_1 = 3|M = 3) &= \frac{5/35}{7/35} = \frac{5}{7} \\
P(X_1 = 4|M = 4) &= \frac{4/35}{5/35} = \frac{4}{5} \\
P(X_1 = 5|M = 5) &= \frac{3/35}{3/35} = \frac{3}{3}
\end{aligned}$$

Opgave d Find sandsynligheden for at kuglen med det mindste nummer blev trukket fra kasse 1. Vi tæller alle punktsandsynlighederne $p(M)$ hvor $X_2 \geq X_1$, og det er

$$P(X_2 \geq X_1) = \sum_{n=1}^{5} \frac{n+2}{35} = \frac{25}{35} = \frac{5}{7}$$

Opgave e To personer, A og B, vil spille følgende spil: De vælger hver en kasse, trækker en kugle fra den og det højeste nummer vinder. Er begge numre ens, er spillet uafgjort.

Angiv sandsynlighederne for, i) at apilleren med kasse 1 vinder, ii) at spilleren med kasse 2 vinder.

Der er tre muligheder. Enten vinder A,B eller uafgjort. De tilfælde hvor spillet er uafgjort, er hvor $X_1 = X_2$, som i alt er i 5 udaf de ialt 35 tilfælde. Tilbage er der $X_1 > X_2$, som giver 10, og ræsten $X_1 < X_2$ som giver 20 tilfælde. Det betyder mere formalistisk, at

$$\begin{aligned}
P(X_1 > X_2) &= \sum_{n=1}^{5} \frac{n-1}{35} = \frac{10}{35} \\
P(X_1 < X_2) &= \sum_{n=1}^{5} \frac{n+1}{35} = \frac{20}{35}
\end{aligned}$$

Opgave f Det er altrå en fordel at trække fra kasse 2. Derfor bliver A og B enige om at skiftes til at vælge kasse først. Den der trækker først, kaster en terning. Er udfaldet 5 eller 6 øjne, vælges kasse 2, ellers vælger kasse 1. Hvad er sandsynligheden for at have trukket først, når man vinder?

Hvis man trækker først, har man $^1/_3$ mulighed for at kunne trække fra kasse 2 og $^2/_3$ for at trække fra kasse 1. Det vil sige, at sandsynligheden for at vinde spillet (vi kalder heldelse A) når man trækker først er

$$P(A) = \frac{1}{3} \cdot \frac{20}{35} + \frac{2}{3} \cdot \frac{10}{35} = \frac{20}{105} + \frac{20}{105} = \frac{8}{21}$$

Opgave 2 (eksanem, vinter 1994-95, opgave 1)

På en ø i Finland fanges der rødspætter og aborrer hver nat med garn som overføres til et hyttefad hver morgen. Antallet af rødspætter, X_r, der fanges i løbet af en nat er Poissonfordelt med parameter λ_r, og antallet af aborrer, X_a, tilsvarende poissonfordelt med parameter λ_a. Det antager, at X_r og X_a er stokastisk uafhængige.

Opgave a Gennemfør beviset for, at antallet af fisk $X = X_a + X_r$, er Poissonfordelt med parameter $\lambda = \lambda_a + \lambda_r$

Beviset skal vi ikke gennemgå her, men det står i sætning 3.6.1TT

$$P(X_a + X_r = x) = e^{-(\lambda_a+\lambda_r)} \frac{(\lambda_a + \lambda_r)^x}{x!} = e^{-\lambda} \frac{\lambda^x}{x!} \quad , \lambda = \lambda_a + \lambda_r$$

Opgave b Find fordelingen af X_r, givet at $X = x$ Dette findes ved at udregne den betingede fordeling.

$$P(X_r = x_r | X = x) = \frac{P((X_r = x_r) \cap (X = x))}{P(X = x)}$$

Vi skal omskrive fællesmængden så den kun indeholder stokastisk uafhængige variable. Vi ved at X_r og X_a er stokastisk uafhængige, og at hvis $X = x$ og $X_r = x_r$ må $X_a = x - x_r$. Fællesmængden som dannes af X_r og X er selve X_r. Det betyder, at vi kan omskrive ligningen til følgende

$$\begin{aligned}
P(X_r = x_r | X = x) &= \frac{P(X_r = x_r) \cdot P(X_a = x - x_r)}{P(X = x)} \\
&= \frac{e^{-\lambda_r} \frac{\lambda_r^{x_r}}{x_r!} \cdot e^{-\lambda_a} \frac{\lambda_a^{x-x_r}}{(x-x_r)!}}{e^{-(\lambda_r+\lambda_a)} \frac{(\lambda_r+\lambda_a)^x}{x!}} \\
&= \frac{e^{-(\lambda_r+\lambda_a)}}{e^{-(\lambda_r+\lambda_a)}} \cdot \frac{\frac{\lambda_r^{x_r}}{x_r!} \cdot \frac{\lambda_a^{x-x_r}}{(x-x_r)!}}{\frac{(\lambda_r+\lambda_a)^x}{x!}} \\
&= \frac{x!}{x_r!(x - x_r)!} \cdot \frac{\lambda_r^{x_r} \cdot \lambda_a^{(x-x_r)}}{(\lambda_r + \lambda_a)^x} \\
&= \binom{x}{x_r} \left(\frac{\lambda_r}{\lambda_r + \lambda_a}\right)^{x_r} \left(\frac{\lambda_a}{\lambda_r + \lambda_a}\right)^{x-x_r}
\end{aligned}$$

X_r er hermed binomialfordelt med sandsynlighedsparameteren $p = {}^{\lambda_r}/_{\lambda_r+\lambda_a}$.

Opgave c Når fiskene overføres til hyttefadet, sker det, at de undslipper. Dette kan beskrives på følgende måde: Lad $F_1, F_2, \cdots$ være indbyrdis uafhængige $0-1$-variable med

$$P(F_i = 1) = 1 - P(F_i = 0) = p \qquad , 0 < p < 1$$

Hændelsen $P(F_i = 1)$ angiver, at fisk nr. i blev overført til hyttefadet, og det antages, at $F_1, F_2, \cdots$ er uafhængige af X_a og X_r. Hvis der i alt er fanget X fisk, vil

$$Y = \sum_{i=0}^{X} F_i$$

angive antallet af fisk, der faktisk blev anbragt i hyttefadet.

Givet, at der er fanget $X = x$ fisk, vis da, at Y er binomialfordelt med antalsparemeter x og sandsynlighedsparameter p

Udfra definitionen på en binomialfordeling s.27TT, er Y binomialfordelt (når vi kender $X = x$) få følgende måde

$$P(Y = y|X = x) = \binom{x}{y} p^y (1-p)^{x-y}$$

hvor $p = P(F_i = 1) = 1 - P(F_i = 0)$.

Opgave d Vis, at Y er Poissonfordelt med parameter $p\lambda$

Da F_i'erne er parvis disjunkte, kan vi med fordel se på sætning 2.3.2TT, som er

$$\begin{aligned} P(Y = y) &= P(X = 1)P(Y = y|X = 1) + \cdots + P(X = x)P(Y = y|X = x) + \cdots \\ &= \sum_{x=1}^{\infty} P(Y = y|X = x)P(X = x) \\ &= \sum_{x=1}^{\infty} \binom{x}{y} p^y (1-p)^{x-y} e^{-\lambda} \frac{\lambda^x}{x!} \end{aligned}$$

Opgave 3

Lad X_1 og X_2, være stokastiske variable på $\mathbf{N}_0 = \{0, 1, \cdots\}$, stokastisk uafhængige med samme geometriske fordeling

$$P(X_i = x) = p(1-p)^x \qquad , 0 < p < 1 \quad , i = 1, 2$$

og betragt de afledte stokastiske variable $S = X_1 + X_2$ og $D = X_1 - X_2$

Opgave c Opskriv sandsynlighedsfunktionen for S

Her genkender vi problemet fra opgave 3.5.4TT.

$$\begin{aligned} P(S = s) &= P(X_1 + X_2 = s) \qquad , \text{hvor } s = x_1 + x_2 \\ &= P(X_1 = 0)P(X_2 = s) + P(X_1 = 1)P(X_2 = s-1) + \cdots \\ &\quad + P(X_1 = s)P(X_2 = 0) \end{aligned}$$

$$\begin{aligned}
&= \sum_{x_1=0}^{s} P(X_1 = x_1)P(X_2 = s - x_1) \\
&= \sum_{x_1=0}^{s} p(1-p)^{x_1} \cdot p(1-p)^{s-x_1} \\
&= \sum_{x_1=0}^{s} p^2(1-p)^s \\
&= (s+1)p^2(1-p)^s
\end{aligned}$$

Dette genkender vi som en negativ binomialfordeling med sandsynlighedsparameter p og antalsparameter $n = 2$ jvnf. TT. side 60

Opgave b Opskriv sandsynlighedsfunktionen for den betingede fordeling af (X_1, X_2) givet $S = s$

Den kan i første omgang opskrives som

$$P(X_1, X_2 | S = s) = \frac{P((X_1, X_2) = (x_1, x_2) \cap S = s)}{P(S = s)}$$

Først kan vi omskrive $P((X_1, X_2) = (x_1, x_2) \cap S = s)$ til $P((X_1, X_2) = (X_1 = x_1, X_2 = s - x_2))$. Da disse er stokastisk uafhænging, kan vi skrive dem som et produkt, nemlig $P(X_1 = x_1)P(X_2 = s - x_2)$. Vi kan nu indsætte de ting vi fandt i a)'eren

$$\begin{aligned}
P(X_1, X_2 | S = s) &= \frac{P(X_1 = x_1)P(X_2 = s - x_2)}{(s+1)p^2(1-p)^s} \\
&= \frac{p(1-p)^{x_1} \cdot p(1-p)^{s-x_2}}{(s+1)p^2(1-p)^s} \\
&= \frac{1}{s+1}
\end{aligned}$$

Jeg kan se, at resultatet ikke er helt hen i skoven, da det hverken bliver negativt, eller større end 1, men hvad resultatet kan bruges til, eller hvad det er for en fordeling ved jeg ikke, men har det ikke noget at gøre med en ventetid?

Opgave c Opskriv sandsynlighedsfunktionen for D. Da vi ikke ved endnu hvad vi skal forstå ved en negativ stokastisk variabel, skal vi ikke besvare dette såørgsmål.

Opgave d Udregn middelværdi og varians for S og D

Middelværdien for S Det var her det var smart at vi i a'eren kunne genkende fordelingen for S som den negative binomialfordeling. Jvnf. TT. side 60, så er middelværdien for geometrisk fordeling

$$\mathrm{E}X = \frac{1-p}{p}$$

Da summen at to uafhænginge geometriske fordelinger med sandsynlighedsparameter p giver en negativ binomialfordeling, følger det, at middelværdien for den negative binomialfordeling også er summen af de to middelværdier for de geometriske fordelinger. Dermed er

$$\mathrm{E}S = \mathrm{E}X_1 + \mathrm{E}X_2 = 2\frac{1-p}{p}$$

Variansen for S var(S) er givet ved (jvmf. TT side 62)

$$\mathrm{var}(S) = \mathrm{E}(S^2) - (\mathrm{E}S)^2$$

I opgave 4.3.3 får vi at vide, at variansen for den negative binomialfordeling er givet ved

$$\mathrm{var}(S) = n\frac{1-p}{p^2} = 2\frac{1-p}{p^2}$$

hvor vi har antalsparameter $n = 2$. (Jeg skulle nok have regnet det hele ud, men jeg kunne ikke komme igennem opgaven, så jeg har 'snydt' lidt ved at lede i TT om ikke der var noget der lignede, og jeg har brugt det jeg fandt)

Middelværdien for D her kan vi finde middelværdien udfra $\mathrm{E}X_1$ og $\mathrm{E}X_2$. Middelværdien for en geometrisk fordeling er ifølge eksempel 4.2.6TT givet ved

$$\mathrm{E}X = \frac{1-p}{p}$$

Vi har nu middelværdien for de to stokastisk uafhængige variable X_1 og X_2. Det vi nu skal, er at finde middelværdien af en linarkombination af disse nemlig

$$\begin{aligned} \mathrm{E}D &= a\mathrm{E}X_1 + b\mathrm{E}X_2 \qquad , a = 1 \text{ og } b = -1 \\ &= \frac{1-p}{p} + (-1)\frac{1-p}{p} \\ &= 0 \end{aligned}$$

Hvis vi tænker intuitivt, så skulle middelværdien for $D = X_1 - X_2$ også gerne være 0

Variansen for D For at udregne $\mathrm{var}(D)$, følger vi proceduren side 66TT. hvor der står

$$\begin{aligned} \mathrm{var}(D) &= \mathrm{var}(aX_1 + bX_2) \qquad , a = 1 \text{ og } b = -1 \\ &= \mathrm{var}(X_1) + \mathrm{var}(X_2) - 2\mathrm{cov}(X_1, X_2) \end{aligned}$$

Siden variablene er stokastisk uafhængige er $\mathrm{cov}(X_1, X_2) = 0$, da

$$\mathrm{cov}(X_1, X_2) = \mathrm{E}(X_1X_2) - (\mathrm{E}X_1)(\mathrm{E}X_2) = 0$$

hvor $\mathrm{E}(X_1X_2) = (\mathrm{E}X_1)(\mathrm{E}X_2)$ pga den stokastiske uafhængighed. Vi har nu tilbage kun at udregne

$$\begin{aligned} \mathrm{var}(D) &= \mathrm{var}(X_1) + \mathrm{var}(X_2) \\ &= 2\frac{1-p}{p^2} \end{aligned}$$

Vi ser, at variansen for D er den samme som for S. Det ved jeg ikke om er rigtigt, men vi ser udfra regneproceduren, at om b er 1 eller -1, ikke indvirker på resultatet når covarianser er 0.