Rettevejledning

Opgave 1

1. Det regner ofte flere dage i træk, derfor vil uafhængighed være en tvivlsom antagelse. Andre forklaringer godtages. Der er tre udfald: sne, regn, og tørvejr. Lad X_1 være antallet af snedage på en uge, X_2 antallet af regnsvejrsdage, X_3 antallet af tørvejsdage. Så er $X_1, X_2 \sim MULT(7, \frac{2}{7}, \frac{2}{7})$. Antallet af tørvejrsdage følger af $n = x_1 + x_2 + x_3$.

$$P(X_1 = 4, X_2 = 1) = \binom{7}{4,1,2} \cdot \left(\frac{2}{7}\right)^4 \cdot \frac{2}{7} \cdot (1 - \frac{2}{7} - \frac{2}{7})^2 = 0,037$$

2. Y er antallet snedage i løbet af en uge. $Y\tilde{\ }BIN(7, \frac{2}{7})$ fordi den marginale fordeling i multinomialfordelingen er binomialfordelt. $P(Y \geq 2) = 1 - P(Y \leq 1) = 0,640$.

$$\begin{aligned} E[Y] &= 7\frac{2}{7} = 2 \\ Var(Y) &= 7\frac{2}{7}(1 - \frac{2}{7}) = \frac{10}{7}. \end{aligned}$$

Sandsynligheden for at Y ligger indenfor 2 std. afv. er givet ved $P(2 - \frac{20}{7} \leq Y \leq 2 + \frac{20}{7}) = P(Y \leq 4) = 0,977$

3. Lad Y_1 være antallet af snedage i uge 1 og Y_2 være antallet af snedage i uge 2. $Y_1 + Y_2 \sim BIN(14, \frac{2}{7})$. Find $P(Y_1 = 1, Y_2 = 6 | Y_1 + Y_2 = 7) = \frac{P(Y_1=1) \cdot P(Y_2=6)}{P(Y_1+Y_2=7)} = \frac{0.000722}{0.1045} = 0.0069$.

Opgave 2

1. Uden tilbagelægning skal vi anvende en hypergeometrisk fordeling. $X \sim HYPGEO(942, 700, 100).P(X \geq 75) = 1 - P(X \leq 74) = 0,488$. Binomialfordelingen anvendes når n er lille i fht. N. I dette tilfælde blæser svaret i vinden, da lærebogen ikke opstiller en tommelfingerregel. På forelæsningsslides er anvendt $\frac{n}{N} \leq 0.05$, hvilket også kan anvendes. Med binomialfordelingen, $X\tilde{\ }BIN(100; \frac{700}{942})$ er $P(X \geq 75) = 1 - P(X \leq 74) = 0,490$.

2. Lad Y være en sum af 100 uafhængige stokastiske variable, der hver især er Bernouilli fordelte, og Y vil derfor være normalfordelt. Der skal approksimeres med middelværdi og varians i binomialfordeling fra 1. Hvis der vælges middelværdi og varians fra den hypergeometrisk fordeling er

det ok. Så $Y\tilde{}N(100 \cdot \frac{700}{942}; 100 \cdot \frac{700}{942} \cdot (1 - \frac{700}{942})) = N(74.3; 19.1)$. Bemærk at sandsynligheden, som skal findes og approksimeres er $P(Y \leq 74)$.

$$P(Y \leq 74) = \Phi(\frac{74 + 0.5 - 74.3}{\sqrt{19.1}}) = 0.518$$

3. Uafhængighed vil være en fin antagelse hvis sygdom ikke smitter mellem eleverne. Ellers vil den være tvivlsom.

$$D\tilde{}N(\frac{1}{56}\sum \mu_{Y_1} - \frac{1}{44}\sum \mu_{Y_2}, \frac{1}{56^2}\sum \sigma^2_{Y_1} + \frac{1}{44^2}\sum \sigma^2_{Y_2}) = N(4-3, \frac{1}{56}\cdot 2 + \frac{1}{44}\cdot 1)$$

$$\begin{aligned} P(D &> 0) = 1 - P(D < 0) \\ &= 1 - \Phi(\frac{0-1}{\sqrt{\frac{2}{56} + \frac{1}{44}}}) = 0.999 \end{aligned}$$

Opgave 3

1. Lad Y være resultatet af et terningslag. Så er $E(Y) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = 3,5 = \frac{21}{6}$.$E(Y^2) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6}$.$Var(Y) = E(Y^2) - E(Y)^2 = \frac{91}{6} - (\frac{21}{6})^2 = \frac{105}{36}$. Lad $X = \frac{1}{5}(Y_1 + Y_2 + Y_3 + Y_4 + Y_5)$. Dermed er $E(X) = 3,5$ og $Var(X) = \frac{1}{25}5Var(Y) = \frac{105}{36\cdot 5} = \frac{105}{180} = 0,58$

2. $H_0 : \sigma^2 = 0,58$. $H_A : \sigma^2 > 0,58$. Lad $Q = \frac{(n-1)s^2}{\sigma^2} = (20-1) \cdot \frac{0,60}{0,58} = 19,7$ som er χ^2 med 19 frihedsgrader. Sss $= P(\chi^2(19) > 19,7) = 41,6\%$ dvs. stor Sss og dermed opretholdes H_0. Så vi opretholder, at variansen er $0,58$.

3. $H_0 : \mu = 3,5$, $H_A : \mu \neq 3,5$. Lad $t = \sqrt{n}\ \frac{(\mu-3,5)}{s} = \sqrt{20}\ \frac{(3,66-3,5)}{\sqrt{0,6}} = 0,92$ som er t-fordelt med 19 frihedsgrader. Sss=p-værdi= $2P$(t-fordeling med 19 frihedsgrader $> 1,19) = 0,37$ så vi opretholder H_0 og dermed opretholder vi at middelværdien er lig $3,5$. (Bruges et U-test er dette også OK, så skal $\sigma^2 = 0,58$. Vi har jo lige testet dette)

4. $(X_1, X_2, X_3, X_4, X_5, X_6)$ bliver multinomisk fordelt, hvor $N = 100$ og med sandsynligheden $1/6$ hvis terningen er fair.

5.

værdi	antal kast	Forventet	O-F	$\frac{(O-F)^2}{F}$
1	12	$16,67$	$-4,76$	$1,31$
2	14	$16,67$	$-2,67$	$0,43$
3	16	$16,67$	$-0,67$	$0,03$
4	17	$16,67$	$0,33$	$0,01$
5	28	$16,67$	$11,33$	$7,71$
6	13	$16,67$	$-0,67$	$0,81$
I alt	100	100	0	$10,28$

Teststørrelsen bliver $10,28$ som er χ^2 fordelt med $df = 6-1$.

Sss bliver da 7% som jo altså er større end de "berømte" 5%. Så vi opretholder, at terningen er fair.

6. Estimatoren bliver $\widehat{p} = \frac{28}{100} = 0,28$ og det tilhørende 95% cf interval bliver $\widehat{p} \pm 1,96\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$. $0,28 \pm 1,96\sqrt{\frac{0,28(1-0,28)}{100}} = [0,19 - 0,37]$. Et interval der ikke indeholder $\frac{1}{6} = 0,16$. Så noget kunne tyde på at terningen er skæv.