Opgave 1

Anita har været på ferie i Paris. Da hun ankommer til lufthavnen er hendes kufferter ikke med hjem i flyet. Hun går straks til flyselskabets skranke og spørger til sin bagage. Den unge ekspedient ved skranken fortæller hende, at bagagen skulle overføres 2 gange for at komme med hjem. Sandsynlighederne for at bagagen ikke blev korrekt overført ved de to overførsler er henholdsvis $\frac{1}{97}$ og $\frac{3}{100}$.

1. Hvad er sandsynligheden for at kufferterne ikke er med hjem?

Lad A og B være hændelsen at bagagen ikke blive overført. Lad C være en hændelsen at bagagen ikke er rmed hjem. Vi kender følgende ssh. $P(A) = \frac{1}{97}$ og $P(B|A^c) = \frac{3}{100}$.$P(C) = P(A) + P(B|A^c) \cdot P(A^c) = \frac{1}{97} + \frac{96}{97}\frac{3}{100} = 0,04$. Alternativt kan den beregnes ved at at finde den komplementære hændelse C^c.

De 179 andre ombord på flyet havde også kufferter, som skulle igennem de samme overførsler, med. Antag uafhængighed mellem passagernes overførsler af kufferter.

2. Opstil en sandsynlighedsmodel for antallet af passagerer, der ikke får deres kufferter med hjem. Er uafhængighed en realistisk antagelse? Hvad er det forventede antal passagerer, der ikke får deres kufferter med hjem? Hvad er sandsynligheden for, at højest 7 ikke får deres kufferter med hjem?

Lad $n = 180$ og $p = 0,04$. Pga uafhængighed og konstant ssh for sandsynlighed er antallet af passagerer, der ikke får deres bagage med hjem, binomialt fordelt. $X \sim Bin(180; 0,04)$

Uafhængighed er en tvivlsom antagelse, da overførslerne formentlig sker nogenlunde samtidigt. Enhver diskussion af uafhængighed for eller imod giver point. $E[X] = 180 \cdot 0,04 = 7,2$.$P(X \leq 7) = 0,568$.Hvis Excel ikke anvendes kan ssh. approksimeres med en normalfordeling.

Efter Anitas flyvning er der 7 passagerer inkl. Anita, der ikke har fået deres kufferter med hjem.

3. Opstil en model for antallet af passagerer, der er med flyet, når 7 passagerer ikke får deres kufferter med hjem (Det kan ignoreres at der højest kan være 320 passagerer med i flyet). Hvad er det forventede antal passagerer? Find det antal passagerer, som giver den største sandsynlighed for at præcis 7 passagerer har mistet deres bagage.

Lad Y være antallet af passagerer med i flyet. Vi ved at der er uafhængighed og konstant ssh $p = 0,04$. Vi 'venter' på 7 successer og $y = 7, ..., 320$, men de 320 kan ignoreres. Derfor kan vi bruge en negativ binomialfordeling . $Y \sim NB(0,04; 7)$. $E[Y] = \frac{7}{0,04} = 175$.Maksimum eller mode af sandsynlighedsfunktionen er $P(X = 150)$ og $P(X = 151)$.

Opgave 2

En virksomhed sender sine medarbejdere på efteruddannelse for at omskole dem til et nyt IT-system, der skal implementeres. Der er to typer af medarbejdere. Type 1 har let ved at blive omskolet. og type 2 er sværere at omskole. Type 1 omskoles med sandsynligheden $0,9$ og type 2 med sandsynligheden $0,75$.

1. Hvad er sandsynligheden for at en tilfældigt udvalgt medarbejder bliver omskolet, når vi antager at 60 pct er af type 1 i virksomheden og 40 pct. er type 2?

Lad $X = 1$ ved omskoling og 0 ellers. Lad T være en indikator for om det er type 1 ($T = 1$) eller type 2 (T=2). Find $P(X = 1) = P(X = 1|T = 1) \cdot P(T = 1) + P(X = 1|T = 2) \cdot P(T = 2) = 0,84$

2. En medarbejder er blevet omskolet, hvad er sandsynligheden for at han er type 1?

$P(T = 1|X = 1) = \frac{P(X=1|T=1)P(T=1)}{P(X=1)} = \frac{0,9 \cdot 0,6}{0,84} = 0,643$. Ssh for type 1 er $0,6$ når vi ikke ved noget om medarbejderen er blevet omskolet. Denne ssh er steget til 0,643 når vi får oplyst at medarbejderen er omskolet, da type 1 lader sig lettere omskole.

Set fra virksomheden er det en fordel at sende medarbejdere af type 1 på kursus, da de er nemmest at omskole. Men virksomheden kan ikke se forskel på typerne. Derimod kender virksomheden medarbejdernes uddannelsesniveau. Fordelingen af uddannelse og type er givet ved denne tabel:

	Type 1	Type 2	I alt
Ingen uddannelse	$0,1$	$0,2$	$0,3$
Noget uddannelse	$0,2$	$0,2$	$0,4$
Meget uddannelse	$0,3$	0	$0,3$
I alt	$0,6$	$0,4$	

3. Hvad er sandsynligheden for at blivet omskolet for de forskellige uddannelsestyper? Kan virksomheden anvende uddannelse til at vælge mellem medarbejdere, der skal på kursus?

gruppe	ssh
ingen uddannelse	0,8
noget uddannelse	0,825
meget uddannelse	0,9

Det er meget tydeligt at mere uddannelse øger ssh for omskoling. Det kan virksomheden klart bruge.

Opgave 3

Gennem længere tids anvendelse af en bestemt vægt har det vist sig, at vægtaflæsninger ved hjælp af vægten kan beskrives ved en normalfordeling med standardafvigelse 0,18 g. Man ønsker nu at kontrollere, om vægten har denne egenskab, da man har en bekymring for, at vægtaflæsningerne har fået en forøget standardafvigelse. Der er derfor foretaget 20 aflæsninger af et bestemt legemes vægt i gram.

$10,9$	$11,2$
$11,3$	$10,7$
$11,0$	$10,8$
$11,0$	$10,9$
$10,7$	$10,8$
$11,0$	$10,8$
$11,1$	$10,8$
$10,9$	$11,1$
$11,4$	$10,9$
$11,0$	$11,2$

Det kan oplyses at $\sum_{i=1}^{20} x_i = 219,5$ og $\sum_{i=1}^{20} (x_i - \overline{x})^2 = 0,7175$

Der opstilles nu den statistiske model, at X_i er uafhængige normalfordelte med middelværdi $= \mu$ og varians $= \sigma^2$.

1. Estimer de to parametre μ og σ^2 og angiv deres fordelingsegenskaber

$\widehat{\mu} = \frac{1}{20}\sum_{i=1}^{20} x_i = \overline{x} = \frac{219,5}{20} = 10,975$ $\widehat{\sigma}^2 = s^2 = \frac{1}{20-1}\sum_{i=1}^{20}(x_i - \overline{x})^2 = \frac{0,7175}{19} = 0,038$

E$(\widehat{\mu}) = \mu$ og E$(\widehat{\sigma}^2) = \sigma^2$ og dermed er de begge centrale.

V$(\widehat{\mu}) = \frac{\sigma^2}{n}$ så variansen gåer mod nul når n går mod uendelig og dermed er $\widehat{\mu}$ konsistent

Det forventes ikke at man kommentere yderligere vedr. σ^2. Men der gælder at V$(\widehat{\sigma}^2) = \frac{2\sigma^4}{n-1}$

2. Test hypotesen om at $\sigma = 0,18$ (dvs $\sigma^2 = (0,18)^2$) mod alternativet $\sigma > 0,18$

Testatørrelse q=(n-1)$\frac{s^2}{\sigma^2} = (20-1)\frac{0,038}{0,18^2} = 22,2$ som er $\chi^2(19)$

Sss=p-værdi=P$(\chi^2(19) > 22,2) = 27\%$ og altså større end de berømte 5%.

3. Udarbejd et 95% konfidensinterval for μ og redegør for de nødvendige forudsætninger.

$\overline{x}$+-$t_{0.975}(n-1)*\sqrt{\frac{s^2}{n}}$ hvilket giver 10,975 +-2,09*$\sqrt{\frac{0{,}038}{20}}$ [10,88 - 11,07]

4. Formuler og test hypotesen om, at vægten af det målte legeme er 11 g.

$H_0 : \mu = 11$ $H_A : \mu <> 11$
teststørrelse t=$\sqrt{20}$*$\frac{\overline{x}-11}{\sqrt{0{,}038}}$ = -0,57   som er t-fordelt med 19 frihedsgrader
Sss=p-værdi=2*P(t(19)<-0,57)= 57% som igen er større end de 5%

For at afgøre om et bestemt tilsætningsstof til motorolien formindsker sandsynligheden for, at en bil skal til hovedreparation indenfor to år, udvælger et biludlejningsselskab, der disponerer over 25 nye biler, tilfældigt 10 biler, der får olie med den omtalte tilsætning. Efter to år var der 17 biler, der havde fået en hovedreparation, mens 8 ikke have haft behov for en hovedreparation. Sammenhængen mellem nødvendigheden af en hovedreparation og anvendelsen af tilsætningsstof fremgår af nedenstående tabel

	hoved	ingen	I alt
tilsætning	5	5	10
ingen	12	3	15
I alt	17	8	25

5. Opstil en statistisk model for ovenstående

X=#biler der skal have hovedreparation blandt biler med tilsætning
Y=#blier der skal have hovedreparation blandt biler uden tilsætning
X er BIN(10,p_1) og Y er BIN(15,p_2) og uafhængige af hinanden.

forudsætningerne for binomialfordelingen: n forsøg, to udfald, samme sandsynlighed og uafhængighed må siges at være opfyldt

6. Test om tilsætningsstoffet har virket.

H_o : p_1 =p_2 H_A både et ensidet og et tosidet alternativ bør accepteres her.

$\widehat{p}_1 = \frac{5}{10} = 0,5$ $\widehat{p}_2 = \frac{12}{15} = 0,8$ under antagelse af at H_o er sand så bliver den fælles p estimeret ved
$\widehat{p}_{ens} = \frac{17}{25} = 0,68$
teststørrelse=$\frac{\widehat{p}_1-\widehat{p}_2}{\sqrt{\widehat{p}_{ens}*(1-\widehat{p}_{ens})*(\frac{1}{10}+\frac{1}{15})}} = \frac{0,5-0,8}{\sqrt{0,68*(1-0,68)*(\frac{1}{10}+\frac{1}{15}}} = -1,57$
ensidet Sss=p-værdi=P(U<-1,57)=$\phi(-1,57) = 5,8\%$
tosidet Sss=p-værdi=P(U<-1,57)=$\phi(-1,57) = 11,5\%$

havde man brugt det betingede test (også kaldet Fischers eksakte test) ville man have fået en Sss på ca 12%. Her er det nok mest korrekt at bruge det betingede. Men alle beregninger af test accepteres her. Så man kan ikke konkludere at stffet virker, men man ville jo anbefale en test med flere observationer.