Høgskolen i Østfold

4. mai 2017

# EKSAMEN – løsningsforslag

| **Emnekode:** | **Emnenavn:** |
|---|---|
| ITD20106 | Statistikk og økonomi |
| **Dato:**<br>2. mai 2017 | **Eksamenstid:**<br>09.00 – 13.00 |
| **Hjelpemidler:**<br>- Alle trykte og skrevne.<br>- Kalkulator. | **Faglærer:**<br>Christian F Heide |

**Om eksamensoppgaven og poengberegning:**

Oppgavesettet består av 12 sider inklusiv denne forsiden og seks sider vedlegg. Kontroller at oppgavesettet er komplett.

Oppgavesettet består av 8 oppgaver med i alt 18 deloppgaver. Det er i oppgavesettet angitt hvor mye hver oppgave teller ved sensuren.

Der det er mulig skal du:

- vise utregninger og hvordan du kommer fram til svarene
- begrunne dine svar, selv om dette ikke er eksplisitt sagt i hvert spørsmål

Om noe er uklart eller mangelfullt i oppgaven, gjør selv de nødvendige forutsetninger.

**Sensurfrist:** 26. mai 2017

Karakterene er tilgjengelige for studenter på Studentweb senest 2 virkedager etter oppgitt sensurfrist. www.hiof.no/studentweb

## Oppgave 1 (15 %)

*Sølvsmia AS er en produksjonsbedrift i gull- og sølvvarebransjen. I november 2016 starter bedriften med budsjettarbeider for 2017. Regnskapet for 2016 er ikke ferdig, men økonomisjefen mener at regnskapet ikke vil avvike vesentlig fra budsjettet. Resultatbudsjettet for 2016 ser slik ut*:

| | |
|---|---|
| Salgsinntekt | 26 700 000 |
| Annen driftsinntekt | 852 000 |
| **Sum driftsinntekter** | **27 552 000** |
| Vareforbruk | 8 758 000 |
| Lønns-og personalkostnader | 9 358 000 |
| Avskrivninger | 458 000 |
| Annen driftskostnad | 7 897 000 |
| **Sum driftskostnader** | **26 471 000** |
| **Driftsresultat** | **1 081 000** |
| Renteinntekt | 130 000 |
| Rentekostnad | -876 000 |
| Resultat før skattekostnad | 335 000 |
| Skattekostnad | -90 000 |
| **Årsresultat** | **245 000** |

*Ledelsen i Sølvsmia AS er optimistisk og tror på et høyere salg i tiden fremover. De tror at omsetningen vil øke med 5% i 2017, og at de fleste kostnadene vil øke tilsvarende. Lønninger og personalkostnader kommer trolig til å øke med 6%. Bedriften budsjetterer med en nedgang i rentenivået på ca. 10%. Avskrivningene blir uendret. Annen driftsinntekt antar de blir redusert til kr. 330,000.*

*Sett opp resultatbudsjettet for 2017. Du kan ta egne forutsetninger der du mener det er nødvendig. Egne forutsetninger skal begrunnes.*

*Foreta en vurdering av resultatbudsjettet for 2017, og foreslå tiltak his du mener det er nødvendig.*

| | Regnskap År 2016 | | Budsjett År 2017 | |
|---|---|---|---|---|
| | Kroner | % | Kroner | % |
| Salgsinntekter | 26 700 000 | 96,9 % | 28 035 000 | 98,8 % |
| Annen driftsinntekt | 852 000 | 3,1 % | 330 000 | 1,2 % |
| Sum driftsinntekter | 27 552 000 | 100,0 % | 28 365 000 | 100,0 % |
| Varekostnader | 8 758 000 | 31,8 % | 9 195 900 | 32,4 % |
| Lønnskostnader | 9 358 000 | 34,0 % | 9 919 480 | 35,0 % |
| Avskrivninger | 458 000 | 1,7 % | 458 000 | 1,6 % |
| Andre driftskostnader | 7 897 000 | 28,7 % | 8 291 850 | 29,2 % |
| Sum driftskostader | 26 471 000 | 96,1 % | 27 865 230 | 98,2 % |
| Driftsresultat | 1 081 000 | 3,9 % | 499 770 | 1,8 % |
| Renteinntekter | 130 000 | 0,5 % | 117 000 | 0,4 % |
| Rentekostnader | 876 000 | 3,2 % | 788 400 | 2,8 % |
| Resultat før skattekostnad | 335 000 | 1,2 % | -171 630 | -0,6 % |
| Skattekostnad | 90 000 | 0,3 % | 0 | 0,0 % |
| Årsresultat | 245 000 | 0,9 % | -171 630 | -0,6 % |

Det budsjetterte underskuddet har sin hovedforklaring i de reduserte «andre driftsinntekter».

Bedriften må gjennomgå sine kostnader for å finne eventuelle besparelser. Det er også aktuelt å finne en erstatning for driftsinntektene som er budsjettert med en reduksjon.

Forutsetninger for budsjettoppsettet:
Vi har redusert både rentekostnadene og renteinntektene med 10 %. Det er budsjettert med et underskudd, og vi har derfor satt skattekostnaden til kr. 0.

## Oppgave 2 (15 %)

*Høknes stolmontering har en kapasitet på 70 stoler per uke. Bedriften har funnet denne sammenhengen mellom produksjon og variable kostnader*:

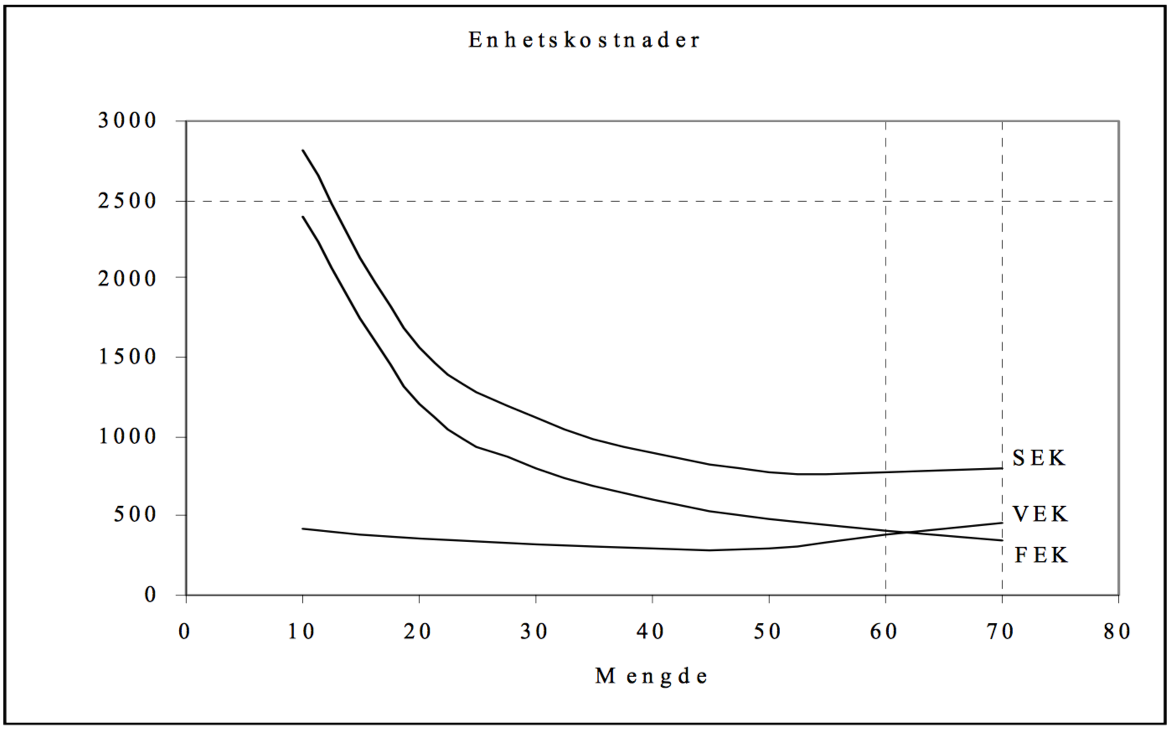
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Mengde | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| Variable kostnader | 4 200 | 7 200 | 9 600 | 12 000 | 15 000 | 22 800 | 32 200 |

*De faste kostnadene er driftsuavhengige og utgjør kr 24 000 per uke.*

**a)** *Lag en kostnadstabell hvor du blant annet finner sum enhetskostnader, variable enhetskostnader og faste enhetskostnader.*

| Antall stoler | Totale kostnader | | | Enhetskostnader | | |
|---|---|---|---|---|---|---|
| | Variable | Faste | Sum | Variable | Faste | Sum |
| 0 | 0 | 24 000 | 24 000 | - | - | - |
| 10 | 4 200 | 24 000 | 28 200 | 420 | 2 400 | 2 820 |
| 20 | 7 200 | 24 000 | 31 200 | 360 | 1 200 | 1 560 |
| 30 | 9 600 | 24 000 | 33 600 | 320 | 800 | 1 120 |
| 40 | 12 000 | 24 000 | 36 000 | 300 | 600 | 900 |
| 50 | 15 000 | 24 000 | 39 000 | 300 | 480 | 780 |
| 60 | 22 800 | 24 000 | 46 800 | 380 | 400 | 780 |
| 70 | 32 200 | 24 000 | 56 200 | 460 | 343 | 803 |

**b)** *Fremstill de tre enhetskostnadene grafisk.*




*Høknes stolmontering har hittil fremstilt stolene mer eller mindre manuelt. Gå ut fra at bedriften vurderer å kjøpe en avansert maskin som helt endrer fremstillingsprosessen. Den daglige lederen i bedriften er gift med en lærer i bedriftsøkonomi. Læreren ber studentene drøfte hvilke konsekvenser overgang til en slik maskin vil få for bedriften. Et sammendrag av påstandene til studentene er gitt nedenfor.*

**c)** *Du skal kommentere hver påstand. Det bør fremgå om du mener påstanden er rett eller gal. Gi begrunnelse og presiser hvilke forutsetninger du eventuelt bygger svaret ditt på.*

1. *Det blir færre arbeidsplasser i bedriften.*

   Dersom bedriften ikke øker kapasiteten som følge av nyanskaffelsen, er nok påstanden korrekt. En maskin reduserer behovet for manuell arbeidskraft. Økt kapasitet kan imidlertid gi rom for alle ansatte og endog øke antall arbeidsplasser.

2. *De faste kostnadene går ned, og de variable øker.*

   Feil. Overgangen til mer maskinell drift, fører til at faste kostnader øker, mens de variable kostnadene per enhet synker (for eksempel lønn til produksjonsarbeidere). Ved økning av kapasiteten vil nok også de variable kostnadene øke (totalt sett).

3. *De faste kostnadene stiger, og de variable går ned.*

   Korrekt. Se begrunnelsen under påstand 2.

4. *Kapasiteten øker slik at bedriften kan fremstille langt flere stoler.*

   Oppgaveteksten sier ikke noe eksplisitt om kapasiteten. Men normalt sett vil en ny maskin føre til økt kapasitet. Påstanden er derfor korrekt.

5. *Bedriften blir mer sårbar dersom etterspørselen etter stolene blir mindre.*

   Korrekt. Ved å pådra seg ekstra faste kostnader blir bedriften mer sårbar fordi disse kostnadene ikke uten videre faller bort dersom aktiviteten reduseres på grunn av redusert etterspørsel.

## Oppgave 3 (15 %)

*Et kaffebrenneri benytter en maskin som fyller kaffe i poser. Hver pose skal inneholde en kvart kilo (250 g) kaffe. Mengden kaffe som fylles i hver pose kan oppfattes som uavhengig og normalfordelt. Man vet at standardavviket for fyllingsmaskinen er* $\sigma = 9\,g$.

*I det siste har det kommet klager fra brenneriets kunder på at kaffeposene inneholder for lite kaffe, altså mindre 250 g.*

*For å sjekke om det er hold i disse påstandene, plukker man tilfeldig ut 10 poser kaffe og veier disse. Vekten for disse kaffeposene er som følger*:

242, 245, 240, 249, 239, 250, 241, 242, 262, 241

**a)** *Sett opp hypoteser, og utfør en hypotesetest for å undersøke om vekten på de ti utplukkede kaffeposene gir grunnlag for å hevde at forventet vekt på kaffeposene er lavere enn 250 g. Benytt et signifikansnivå på* α = 0.05.

Hypoteser:

$H_0: \ \mu \geq 250$ (Forventet vekt av kaffe i posene er større eller lik 250 g.)

$H_1: \ \mu < 250$ (Forventet vekt av kaffe i posene er mindre enn 250 g.)

Siden standardavviket er kjent, kan vi bruke en **Z-test**. Et signifikansnivå på $\alpha = 0.05$ betyr at $z_\alpha = z_{0.05} = 1.645$. Vi forkaster nullhypotesen dersom

$$Z < -z_\alpha = -1.645$$

Testobservatoren $Z$ er gitt ved

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

hvor $\overline{X}$ er gjennomsnittsvekten av de kaffeposene vi veier. Her er

$$\bar{x} = \frac{1}{10}(242 + 245 + 240 + 249 + 239 + 250 + 241 + 242 + 262 + 241)$$

$$= 245.1$$

Videre er $\sigma = 9$ (oppgitt i oppgaven) og $n = 10$ (antall kaffeposer vi veier). $\mu_0$ er forventningsverdien angitt i hypotesene, altså $\mu_0 = 250$ . Setter vi inn disse tallene får vi

$$z = \frac{245.1 - 250}{9 / \sqrt{10}} = \underline{-1.72}$$

Vi ser av dette at

$$z = -1.72 < -z_\alpha = -1.645$$

Vi kan følgelig forkaste nullhypotesen.

Konklusjon:

Forventet vekt av kaffe i posene er mindre enn 250 g.

**b)** *Beregn testens p-verdi.*
Dette angir sannsynligheten for at vi forkaster nullhypotesen gitt at den er sann. Altså: hva er sannsynligheten for at vi kunne fått de måleresultatene vi fikk dersom forventningsverdien til vekten av kaffe i posene faktisk ikke *var* mindre enn 250 g. Denne er

$$p = P(\overline{X} \leq 245.1 \mid \mu = 250) =$$

$$P\left( Z \le \frac{245.1 - 250}{9/\sqrt{10}} \right) = P(Z \le -1.72) =$$

$$P(Z < -1.72) = G(-1.72) = \underline{\underline{0.0427}}$$

Vi ser at $p$-verdien er mindre enn signifikansnivået på 0.05 som vi brukte ved hypotesetesten, og vi kunne også brukt $p$-verdien til vår hypotesetest. At $p$-verdien er mindre enn signifikansnivået innebærer at vi kan forkaste nullhypotesen. $p$-verdien gir imidlertid også informasjon om hvor langt ned i signifikansnivå vi kunne gått før vi ikke kunne forkastet nullhypotesen. Her ser vi at dersom vi hadde valgt et signifikansnivå mindre enn 0.0427 kunne vi ikke forkastet nullhypotesen.

*Bedriften går til anskaffelse av en ny maskin for å fylle kaffe i posene. Man kjenner ikke standardavviket for denne nye maskinen. Det gjøres nå nye stikkprøver hvor det velges ut tre poser kaffe som veies. Resultatet er:*

240, 241, 254

**c)** *Utfør nå en ny hypotesetest for å undersøke om det er grunnlag for å hevde at den nye fyllingsmaskinen fyller poser hvor forventningsverdien til kaffemengden er mindre enn 250 g. Benytt også denne gangen et signifikansnivå på 0.05.*

Siden standardavviket er ukjent, må vi bruke en ***T*-test**. Hypotesene blir som tidligere:

$H_0: \ \mu \ge 250$ (Forventet vekt av kaffe i posene er større eller lik 250 g.)

$H_1: \ \mu < 250$ (Forventet vekt av kaffe i posene er mindre enn 250 g.)

Gjennomsnittet for målingene er

$$\bar{x} = 245.0$$

Vårt estimat for standardavviket, $S$, er gitt ved

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{3-1}\sum_{i=1}^{3}(x_i - \bar{x})^2 =$$

$$\frac{1}{2} \cdot \left( (240 - 245)^2 + (241 - 245)^2 + (254 - 245)^2 \right)$$

$$\frac{1}{2} \cdot (5^2 + 4^2 + 9^2) = \frac{1}{2} \cdot 122 = 61.0$$

og følgelig er

$$s = \sqrt{61} = 7.81$$

Siden vi skal bruke signifikansnivå 0.05 og antall frihetsgrader her er $n - 1 = 2$ (siden vi har tre observasjoner), må vi finne $t$-fordelingens kvantil $t_{\alpha,n-1} = t_{0.05,2}$ som er

$$t_{0.05,2} = 2.920$$

Vi forkaster nullhypotesen dersom den observerte verdien av $T$ er mindre enn –2.920 (fordi vi benytter venstre hale av sannsynlighetsfordelingen).

Vår observerte $t$-verdi er

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{245 - 250}{7.81/\sqrt{3}} = \underline{-1.11}$$

Vi ser at vår verdi ikke er mindre enn kvantilverdien –2.920.

Vi kan derfor **ikke** forkaste nullhypotesen på signifikansnivå 0.05.

Konklusjon:

Det er ikke grunnlag for å hevde at forventningsverdien til maskinen er mindre enn 250 g.

## Oppgave 4 (10 %)

*I en urne er det 3 røde og 4 grønne kuler som er helt like bortsett fra fargen.*

*Vi trekker tilfeldig tre kuler uten tilbakelegging. La den stokastiske variabelen X være antall røde kuler i trekningen.*

**a)** *Finn* $P(X = 2)$ *og* $P(X \leq 2)$.

Siden vi har trekning uten tilbakelegging, er $X$ **hypergeometrisk** fordelt med følgende parametere:
$n$ = 3 er antall enheter vi trekker
$N$ = 7 er totalt antall enheter
$M$ = 3 er antall merkede enheter

Formelen for sannsynlighet i hypergeometrisk fordeling er

$$P(X = x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$$

Av denne finner vi

$$P(X = 2) = \frac{\binom{3}{2} \cdot \binom{7-3}{3-2}}{\binom{7}{3}} = \frac{\binom{3}{2} \cdot \binom{4}{1}}{\binom{7}{3}} = \frac{3 \cdot 4}{\frac{7 \cdot 6 \cdot 5 \cdot 4!}{4! \cdot 3 \cdot 2 \cdot 1}} = \frac{12}{35} = \underline{\underline{0.343}}$$

Vi skal så finne $P(X \leq 2)$. Denne finner vi lettest slik:

$$P(X \leq 2) = 1 - P(X \geq 3)$$

Siden vi trekker tre kuler, vil $P(X \geq 3)$ være det samme som $P(X = 3)$ (vi kan jo ikke få fire eller flere røde kuler når vi trekker tre kuler). Vi får da

$$P(X = 3) = \frac{\binom{3}{3} \cdot \binom{4}{0}}{\binom{7}{3}} = \frac{1 \cdot 1}{35} = 0.029$$

og altså

$$P(X \leq 2) = 1 - P(X \geq 3) = 1 - 0.029 = \underline{\underline{0.971}}$$

Vi kan også finne den slik:

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$P(X = 0) = \frac{\binom{3}{0} \cdot \binom{4}{3}}{\binom{7}{3}} = \frac{1 \cdot 4}{35} = \underline{\underline{0.114}}$$

$$P(X = 1) = \frac{\binom{3}{1} \cdot \binom{4}{2}}{\binom{7}{3}} = \frac{3 \cdot 6}{35} = \underline{\underline{0.514}}$$

Følgelig får vi

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.114 + 0.514 + 0.343 = \underline{\underline{0.971}}$$

**b)** *Finn forventningsverdi og standardavvik for X.*
Det enkleste her er å bruke formlene for forventningsverdi og standardavvik i en hypergeometrisk fordeling:

$$E(X) = np$$

Her er $n$ antall kuler vi trekker, altså 3, og $p$ er sannsynligheten for å trekke en rød kule i første trekning, altså $p = \frac{3}{7} = 0.429$ . Vi får da

$$E(X) = 3 \cdot \frac{3}{7} = \underline{\underline{1.286}}$$

Variansen er gitt ved

$$Var(X) = np(1-p)\frac{N-n}{N-1} = 3 \cdot \frac{3}{7} \cdot \frac{4}{7} \cdot \frac{7-3}{7-1} = 0.4898$$

Standardavviket er derfor

$$\sigma = \sqrt{Var(X)} = \sqrt{0.4898} = \underline{\underline{0.70}}$$

## Oppgave 5 (10 %)

*En turist fra England besøker Frankrike. Turisten oppdager at veldig få franskmenn snakker engelsk. Imidlertid er han ved Eiffeltårnet hvor det i tillegg til mange franskmenn også er mange utenlandske turister, og mange av disse snakker engelsk.*

*Anta følgende:*

- *25 % av franskmenn snakker engelsk.*
- *70 % av personer turisten møter ved Eiffeltårnet er utenlandske turister.*
- *85 % av de utenlandske turistene snakker engelsk.*

**a)** *Finn sannsynligheten for at en tilfeldig person som den engelske turisten møter ved Eiffeltårnet snakker engelsk.*

Det kan her være lurt å definere hendelser, f. eks. følgende:
*F*: personen den engelske turisten møter er en franskmann
*E*: personen den engelske turisten møter snakker engelsk

Vi kunne også definert hendelsen at personen den engelske turisten møter er en utenlandsk turist, men dette er det komplementære av at personen turisten møter er en franskmann.

Oppgaven gir da følgende sannsynligheter:

$$P(\overline{F}) = 0.7$$
$$P(F) = 1 - P(\overline{F}) = 1 - 0.7 = 0.3$$
$$P(E \mid F) = 0.25$$
$$P(E \mid \overline{F}) = 0.85$$

Oppgaven er å finne $P(E)$, og denne kan vi finne ved setningen om **total sannsynlighet**:

$$P(E) = P(F) \cdot P(E \mid F) + P(\overline{F}) \cdot P(E \mid \overline{F}) = 0.3 \cdot 0.25 + 0.7 \cdot 0.85 =$$

$$0.075 + 0.595 = \underline{\underline{0.67}}$$

**b)** *Hva er sannsynligheten for at en tilfeldig person som den engelske turisten møter ved Eiffeltårnet er fransk, gitt at denne personen snakker engelsk*?

Her skal vi finne $P(F\mid E)$. Denne sannsynligheten kan vi finne ved hjelp av **Bayes' setning**:

$$P(F\mid E)=\frac{P(F)\cdot P(E\mid F)}{P(E)}=\frac{0.3\cdot 0.25}{0.67}=\underline{\underline{0.11}}$$

## Oppgave 6 (10 %)

*Ingen operasjoner er uten risiko. Det er imidlertid mistanke om at en hjertekirurg gjør oftere feil enn det som er akseptabelt ved bytte av hjerteklaff. Verdens helseorganisasjon oppgir at man må akseptere at 7 % av slike operasjoner slår feil på en eller annen måte (dette tallet har jeg bare funnet på, og enhver likhet med virkeligheten er nokså tilfeldig).*

*Fylkeslegen innhenter informasjon om den aktuelle legens hjerteklaffoperasjoner, og finner at det i løpet av 149 operasjoner var 16 som gikk galt, altså over 10 %.*

*Vi gjør den forenklede antagelsen at antall operasjoner som går galt er binomisk fordelt.*

***a)*** *Begrunn at vi kan benytte normalfordelingen i dette tilfellet selv om vi i utgangspunktet har en binomisk fordeling. Gjennomfør så en hypotesetest for å undersøke om tallmaterialet gir fylkeslegen grunn til å reagere overfor legen fordi hans feilandel overskrider 7 %. Benytt et signifikansnivå på 0.05.*

En binomisk fordeling kan approksimeres ved en normalfordeling forutsatt at antall observasjoner, *n*, er så stort at

$$n\hat{p}(1-\hat{p})>5$$

(Dette kan begrunnes med sentralgrenseteoremet.)

Her er

$$\hat{p}=\frac{X}{n}$$

punktestimatoren for sannsynligheten i den binomiske fordelingen, som i vårt tilfelle er

$$\hat{p}=\frac{16}{149}=0.107$$

Det betyr at i vårt tilfelle er

$$n\hat{p}(1-\hat{p})=149\cdot 0.107\cdot(1-0.107)=14.3$$

som altså langt overstiger 5. Dette viser at vi kan bruke normalfordelingen i vårt tilfelle.

Til hypotesetesten setter vi opp følgende hypoteser:

$H_0: \ p \le p_0 = 0.07$ (legens feilsannsynlighet er *ikke* større enn grensen gitt av WHO)
$H_1: \ p > p_0 = 0.07$ (legens feilsannsynlighet *er* større enn grensen gitt av WHO)

Vi kan benytte testobservatoren

$$Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$$

som er tilnærmet standardnormalfordelt. Vi skal benytte et signifikansnivå på $\alpha = 0.05$. Vi forkaster altså $H_0$ dersom den $z$-verdien vi regner ut er større enn 95-prosentilen $z_\alpha = z_{0.05} = 1.645$.

Setter vi inn tall finner vi følgende $z$-verdi:

$$z = \frac{16 - 149 \cdot 0.07}{\sqrt{149 \cdot 0.07 \cdot (1 - 0.07)}} = \frac{16 - 10.43}{\sqrt{9.7}} = \underline{1.79}$$

Vi ser at denne $z$-verdien er større enn kvantilverdien, og vi må derfor konkludere med:

Vi kan forkaste $H_0$ på signifikansnivå 0.05.

Eller sagt på en annen måte:

Tallmaterialet gir fylkeslegen grunn til å reagere overfor kirurgen.

**b)** *Forklar begrepene type I-feil og type II-feil med utgangspunkt i denne hypotesetesten. Forklar hvorfor og for hvilke parter det er viktig at sannsynlighetene for å gjøre disse to typene av feil er små, og forklar hvordan vi kan redusere sannsynlighetene for de to feiltypene.*

Type I-feil er feilen vi gjør når vi feilaktig forkaster $H_0$ (forkastningsfeil). Altså når vi basert på våre observasjoner konkluder med at kirurgen gjør for mye feil selv om dette ikke er tilfelle.

Type II-feil er feilen vi gjør når vi feilaktig godtar $H_0$ (godtakingsfeil). Altså når vi basert på våre observasjoner konkluderer med at kirurgen ikke gjør for mye feil selv om han faktisk gjør det.

For kirurgen er det viktig at vi ikke gjør type I-feil, for det betyr at han i så fall får en reaksjon fra fylkeslegen selv om det ikke er grunnlag for det. En slik reaksjon vil kunne gå ut over karrieren hans, og i verste fall (for legen) føre til at han mister retten til å praktisere som lege.

For pasientene (og fylkeslegen) er det viktig at vi ikke gjør type II-feil. For hvis legen faktisk er en dårlig kirurg, er det viktig å få ham fjernet slik at han ikke skader for mange pasienter.

Vi kan redusere sannsynligheten for type I-feil ved å endre signifikansnivået, for eksempel fra 0.05 til 0.01. Dette vil imidlertid øke sannsynligheten for type II-feil.

Sannsynligheten for type II-feil, kunne i prinsippet blitt redusert ved å endre signifikansnivået fra for eksempel 0.05 til 0.1. Dette vil man imidlertid ikke ønske å gjøre på grunn av den økte sannsynligheten for type I-feil dette medfører.

Den eneste måten å redusere sannsynligheten for type II-feil uten å øke sannsynligheten for type I-feil, er å øke antall observasjoner. Dette vil nemlig redusere utvalgsstandardavviket slik at vi får et mer presist estimat for sannsynligheten.

## Oppgave 7 (10 %)

*På et hogstfelt settes det opp barkbillefeller. Antall barkbiller som fanges i en felle er poissonfordelt med* $\lambda = 4.4$ *per time.*

**a)** *Hva er sannsynligheten for at en felle inneholder mer enn fem barkbiller etter en time*?
La $X$ være antall barkbiller som fanges i en felle. Vi skal da finne

$$P(X > 5) = 1 - P(X \le 5)$$

Dette finner vi ved å slå opp i tabellen for poissonfordelingen på den linjen med forventningsverdi $\lambda t = 4.4 \cdot 1 = 4.4$:

$$P(X > 5) = 1 - P(X \le 5) = 1 - 0.720 = \underline{\underline{0.280}}$$

**b)** *En felle tømmes. Hva er sannsynligheten for at det kommer enten 2 eller 3 barkbiller i fella i løpet av det første kvarteret etter at den er tømt*?

Vi skal altså finne

$$P(2 \le X \le 3) = P(X \le 3) - P(X \le 1)$$

Siden $\lambda$ har måleenhet «pr. time», vil 15 minutter tilsvare $t = \frac{1}{4} = 0.25$ timer. Følgelig må vi slå opp på forventningsverdien $\lambda t = 4.4 \cdot 0.25 = 1.1$ i tabellen. Dette gir

$$P(2 \le X \le 3) = P(X \le 3) - P(X \le 1) = 0.974 - 0.699 = \underline{\underline{0.275}}$$

## Oppgave 8 (15 %)

*Sirisser er gresshoppelignende insekter som «synger» ved å gni vingene mot hverandre. Det har vært gjort undersøkelser om hvorvidt det finnes en sammenheng mellom temperatur og hvor ofte sirissene gir fra seg et «gniss». Fra boken «The song of insects» av George W. Pierce (Harvard University Press, 1948) kan vi hente følgende data angående dette for en bestemt siriss-art*:

| | | | | |
|---|---|---|---|---|
| Antall gniss pr. minutt ($x$) | 882 | 1104 | 1200 | 1032 |
| Temperatur i °C ($Y$) | 20.9 | 29.1 | 31.4 | 28.1 |

*Følgende er alt utregnet for dette datasettet*:

$\bar{x} = 1054.5$

$\bar{y} = 27.4$

| $x_i$ | $y_i$ | $(x_i - \bar{x})$ | $(y_i - \bar{y})$ | $(x_i - \bar{x})(y_i - \bar{y})$ | $(x_i - \bar{x})^2$ | $(y_i - \bar{y})^2$ |
|---|---|---|---|---|---|---|
| 882 | 20.9 | –172.5 | –6.5 | 1121.25 | 29756.25 | 42.25 |
| 1104 | 29.1 | 49.5 | 1.7 | 84.15 | 2450.25 | 2.89 |
| 1200 | 31.4 | 145.5 | 4.0 | 582.00 | 21170.25 | 16.00 |
| 1032 | 28.1 | –22.5 | 0.7 | –15.75 | 506.25 | 0.49 |

**a)** *Finn regresjonslinjen* $Y = \hat{\alpha} + \hat{\beta}x$ *som best beskriver den lineære sammenhengen mellom x og Y basert på de foreliggende data.*

Regresjonslinjens koeffisienter finner vi ved uttrykkene

$$\hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} \qquad \hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{r \cdot S_Y}{S_X}$$

Vi får

$$\hat{\beta} = \frac{1771.65}{53883} = \underline{0.033}$$

og

$$\hat{\alpha} = 27.4 - 0.033 \cdot 1054.5 = \underline{-7.3}$$

Den estimerte regresjonslinjen er derfor

$$\underline{\underline{Y = -7.3 + 0.033x}}$$

**b)** *Finn et 95 % konfidensintervall for forventningsverdien til temperaturen når antall gniss pr minutt er 1000. Som en hjelp i utregningen får du opplyst at standardfeilen til residualene er s* = 1.30 *og at* $SE(\hat{\beta}) = 0.0056$ .

Konfidensintervallet for forventningsverdien til *Y*, er gitt ved

$$\hat{\alpha} + \hat{\beta}x \pm t_{\alpha/2, n-2} \cdot s \cdot \sqrt{\frac{1}{n} + \left( \frac{(x - \bar{x})}{s / SE(\hat{\beta})} \right)^2}$$

Her er altså *x* = 1000. Antall observasjoner er *n* = 4. Siden vi skal beregne et 95 % konfidensintervall, så er $\alpha = 0.05$, og vi må finne kvantilet $t_{0.025,2}$ i tabellen. Vi finner

$$t_{0.025,2} = 4.303$$

Setter vi inn tallene, finner vi

$$-7.3 + 0.033 \cdot 1000 \pm 4.303 \cdot 1.30 \cdot \sqrt{\frac{1}{4} + \left( \frac{1000 - 1054.5}{1.30 / 0.0056} \right)^2}$$

som gir

$$25.7 \pm 3.1$$

altså

$$\underline{\underline{[22.6, 28.8]}}$$