Oppgave 1:

a) Vi har at $\bar{X} \sim N(5, \frac{16}{4})$

Finn $P(6 < \bar{X} < 8)$:

$$P(6 < \bar{X} < 8) = P\left(\frac{6-5}{2} < Z < \frac{8-5}{2}\right)$$

$$= P(0,5 < Z < 1.5)$$

$$= P(0 < Z < 1.5) - P(0 < Z < 0,5)$$

$$= 0.4332 - 0.1915$$

$$= 0,2417 \approx 0,24$$

Riktig svar: D

b) Har fått oppgitt følgende:

A – har diabetes
B – har fått diabetes-diagnose

$P(A) = 0,10 \Rightarrow P(\bar{A}) = 0.90$
$P(B \mid A) = 0,95$
$P(B \mid \bar{A}) = 0,05$

Finn $P(A \mid B)$:

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

$\left[\dfrac{\text{ant. av de som har diabetes som har fått diagnose}}{\text{tot. ant. som har fått diagnose}} \right]$

Må først finne $P(B)$:

$$P(B) = \underbrace{0,95 \cdot 0,10}_{\text{du har diabetes og får diagnose}} + \underbrace{0,05 \cdot 0,90}_{\text{du har ikke diabetes og får diagnose}}$$

$$= 0,14$$

$$\Rightarrow P(A|B) = \frac{0,95 \cdot 0,10}{0,14}$$

$$\approx 0,68$$

Riktig svar: $\underline{\underline{D}}$

c) La X være antall skårede mål. Da er X binomisk fordelt med $n=4$ og $p=0{,}8$. Finn $P(X=3$ eller $X=4)$:

$\Rightarrow P(X=3) + P(X=4) =$

$$\frac{4!}{(4-3)!\,3!}\, 0{,}8^3 \cdot 0{,}2^{4-3} + \frac{4!}{(4-4)!\,4!}\, 0{,}8^4 \cdot 0{,}2^{4-4}$$

$= 4 \cdot 0{,}8^3 \cdot 0{,}2 + 0{,}8^4$

$= 0{,}819 \approx 0{,}82$

Riktig svar : $\underline{\underline{C}}$

d) Gitt at personen ikke fikk hodeskader, hva er sanns. for at vedkommende brukte hjelm ?

$P(\text{hjelm} \mid \text{ikke hodeskade}) = \frac{83}{319} = 0{,}26$

Riktige svar : $\underline{\underline{E}}$

e) $\mu = \sum X \cdot p(X)$

$= 1 \cdot 0{,}4 + 2 \cdot 0{,}2 + 3 \cdot 0{,}1 + 4 \cdot 0{,}3$

$= 2{,}3$

$\sigma^2 = \sum X^2 p(X) - \mu^2$

$= 1 \cdot 0{,}4 + 4 \cdot 0{,}2 + 9 \cdot 0{,}1 + 16 \cdot 0{,}3 - 2{,}3^2$

$= 6{,}9 - 5{,}29$

$= 1{,}61$

Riktig svar: $\underline{\underline{C}}$

Oppgave 2:

Finn forventningsverdien til Z:

$$
\begin{aligned}
E(Z) &= E(2X - Y) \\
&= 2E(X) - E(Y) \\
&= 2 \cdot 1{,}5 - 1{,}5 \\
&= \underline{\underline{1{,}5}}
\end{aligned}
$$

Finn variansen til Z:

$$
\begin{aligned}
Var(Z) &= Var(2X - Y) \\
&= 2^2 Var(X) + Var(Y) - 2 \cdot 2 cov(X,Y) \\
&= 4 \cdot 0{,}75 + 0{,}75 - 4(-0{,}5) \\
&= \underline{\underline{5{,}75}}
\end{aligned}
$$

Siden Z er en lineær kombinasjon av to normalfordelte variabler så er Z selv normalfordelt. Har altså følgende fordeling:

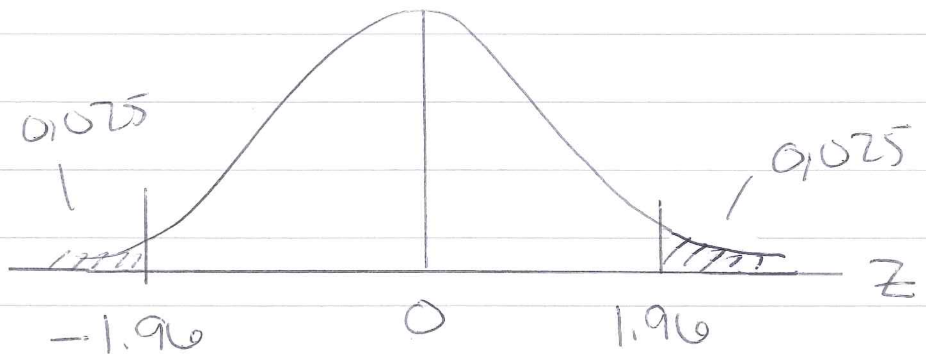
$$\underline{\underline{Z \sim N(1{,}5\ ,\ 5{,}75)}}$$

Oppgave 3

a) 95 % konfidensintervall for μ:

$$\bar{X} \pm z_{\alpha/2} \frac{S}{\sqrt{n}}$$

Kritisk verdi for $\alpha = 0{,}05$ for en tosidig test finner vi fra tabell A.7:



Har altså at $z_{\alpha/2} = 1.96$

Regner ut konfidensintervallet:

$$455\,000 \pm 1.96 \cdot \frac{27\,000}{\sqrt{36}}$$

$$455\,000 \pm 8\,820$$

$$\Rightarrow \underline{\underline{[446\,180\ ,\ 463\,820]}}$$

b) Vi ønsker å finne et 95% konfidens-intervall slik at avstanden mellom $\bar{X}$ og μ er max 5000 kr.

Bruker uttrykket fra forrige oppgave:

$$z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 1.96 \cdot \frac{27000}{\sqrt{n}} \leq 5000$$

Løser uttrykket for n:

$$\left(\frac{1.96 \cdot 27000}{5000}\right)^2 \leq n$$

Setter likhetstegn og regner ut:

$$\underline{\underline{n \approx 112}}$$

Utvalget måtte ha vært på 112 lærere for at avstanden mellom $\bar{X}$ og μ skulle ha vært mindre enn 5000 kr med 95% sikkerhet.

c) Bruk 5 - stegs metoden:

Steg 1:

μ - gj.snittlig årsinntekt blant nyutdannede kokker.

H_0: $\mu = 448\,000$ kr
H_A: $\mu > 448\,000$ kr

Steg 2:

$n > 30$ => sentralgrenseteoremet sier at $\bar{X}$ er normalfordelt:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Får følgende testobservator:

$$TS = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0,1)$$

Velger $\alpha = 0{,}05$

Steg 3:

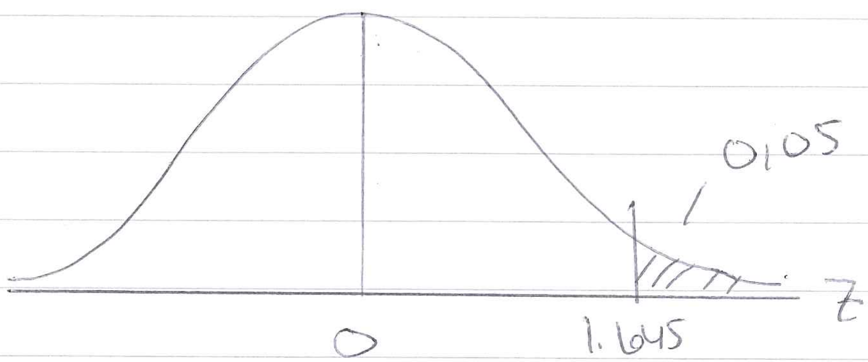
$n = 36$, $\bar{X} = 455\,000$, $S = 27\,000$

$$TS = \frac{455\,000 - 448\,000}{27\,000 / \sqrt{36}} = \frac{7000}{4500} \approx 1.56$$

Steg 4:

Ensidig test, kun interessert i å forkaste H_0 i høyre-halen av z-fordelinga.

Forkast H_0 dersom $TS > z_\alpha$
Fra tabell A.1: $z_{0,05} = 1.645$



Steg 5:

Siden $TS = 1.56 < 1.645$

$\Rightarrow$ Vi kan ikke forkaste H_0 på et 5% signifikansnivå. Årsinntekt til nyutdannede lærere har ikke økt.

d) Signifikansnivået gir oss sanns. for å gjøre en type 1-feil som er å forkaste en sann nullhypotese.

e) $p\text{-verdi} = P(Z > 1.56)$

$= 0{,}5 - P(0 < Z < 1.56)$

$= 0.5 - 0{,}4406$

$= \underline{\underline{0{,}0594}}$

Det er en 5.94% sanns. for å observere TS = 1.56 når H_0 er sann. Det betyr at 5.94% er det laveste sig. nivået som H_0 kan forkastes på. Vi var altså veldig nærme å forkaste H_0 i opg. b).

Oppgave 4

Bruk 5-stegs metoden:

Steg 1:

π_i – andelen som pådrar seg hodeskade i gruppe i hvor:
1 – de som brukte hjelm
2 – de som ikke brukte hjelm

$H_0: \pi_1 = \pi_2$
$H_A: \pi_1 < \pi_2$

Steg 2:

Utvalg 1 (de med hjelm):
Siden $n_1 \cdot p_1 = 113 \cdot \frac{30}{113} > 5$ og
$n_1 \cdot (1 - p_1) = 113 \cdot (1 - \frac{30}{113}) > 5$
$\Rightarrow p_1$ er normalfordelt

Utvalg 2 (de uten hjelm):
Siden $n_2 \cdot p_2 = 418 \cdot \frac{182}{418} > 5$ og

$n_2 \cdot (1 - p_2) = 418 \cdot (1 - \frac{182}{418}) > 5$

$\Rightarrow p_2$ er normalfordelt

Antar at utvalgene kommer fra to uavhengige populasjoner.

Har følgende testobservator:

$$TS = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

Felles estimat på π er gitt ved:

$$p = \frac{n_1 \cdot p_1 + n_2 \cdot p_2}{n_1 + n_2}$$

Steg 3:

$n_1 = 113$, $p_1 = \frac{30}{113} = 0,265$

$n_2 = 418$, $p_2 = \frac{182}{418} = 0,435$

$$p = \frac{113 \cdot 0,265 + 418 \cdot 0,435}{113 + 418}$$

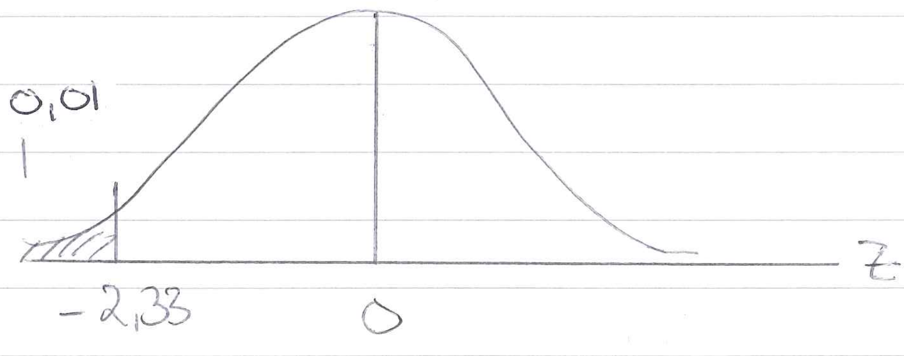
$\approx 0,40$

$$\Rightarrow TS = \frac{0,265 - 0,435}{\sqrt{0,40(1-0,40)\left(\frac{1}{113}+\frac{1}{418}\right)}}$$

$$\approx -4$$

Steg 4:

Forkast H_0 dersom $TS < -z_\alpha$. Har fått oppgitt at $\alpha = 0.01$. Fra tabell A.1 finner vi at $z_{0,01} = 2.33$



Steg 5:

Siden $TS \approx -4 < -2,33$

$\Rightarrow$ H_0 kan forkastes til et 1% sig.-nivå. Vi kan altså konkludere med at andelen som pådrar seg hodeskader er lavere blant de som bruker hjelm.

Oppgave 5

a) Beregn korrelasjonskoeffisienten:

$$R = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2}\sqrt{\Sigma(y_i - \bar{y})^2}}$$

$$= \frac{-989.75}{\sqrt{3462.5} \cdot \sqrt{323.8}}$$

$$= \underline{\underline{-0.93}}$$

Siden $|R|$ er såpass nær 1 så virker det som en fornuftig antakelse å anta en lineær sammenheng mellom x og y.

b) Estimer α og β i populasjonens regresjonslinje ved bruk av OLS:

$$b = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2}$$

$$= \frac{-989.75}{3462.5} \approx \underline{\underline{-0.29}}$$

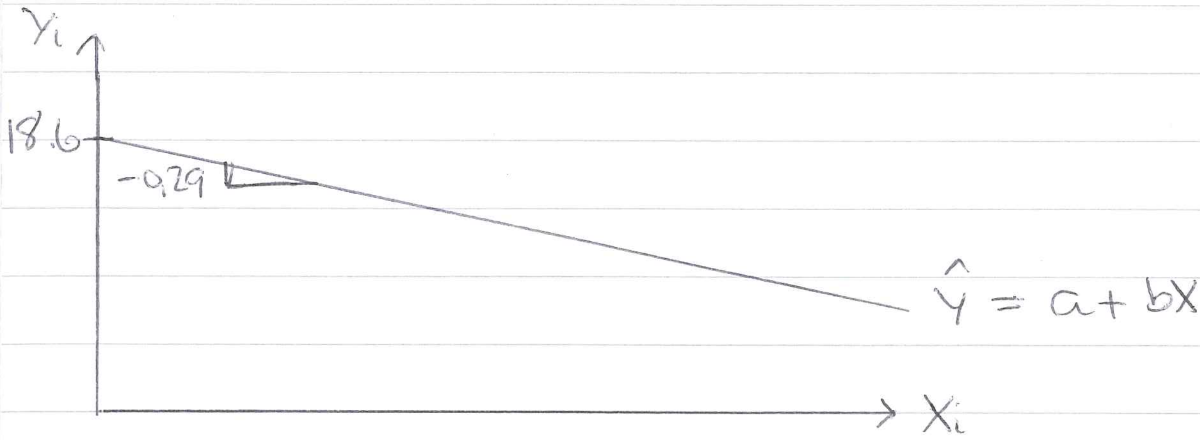
$$a = \bar{Y} - b\bar{X}$$

$$= \frac{77{,}1}{10} - (-0{,}29) \cdot \frac{375}{10}$$

$$\approx \underline{\underline{18.6}}$$

Predikert reg. linje for populasjonen:

$$\hat{Y}_i = 18.6 - 0{,}29 \cdot X_i$$



b gir oss endringen i nivået på mellasyre for hver enhet dopmiddel.

c) For $x = 15$ følger det så at:

$$y(15) = 18.6 - 0{,}29 \cdot 15$$
$$= \underline{\underline{14{,}25}}$$

Melkesyren for denne personen antas

å bli 14.25 mg.

d) Modellens forklaringskraft:

$$R^2 = (-0.93)^2 \approx \underline{\underline{0,86}}$$

$\Rightarrow$ 86% av variasjonen i melkesyre forklares av variasjonen i dopinginntaket.