Sensorveiledning SØK 1004 Statistikk for økonomer vår 20.

Oppgave 1

Definerer hendelsene

S- Har Sol

$\underline{S}$ – Har ikke Sol

D- Positiv diagnose

$\underline{D}$ – Negativ diagnose

a. P(D) = P(D|S)*P(S) + P(D|$\underline{S}$)*P($\underline{S}$) = 0.0725
b. Skal finne P(S|D)
   Kjenner P(D|S) Bruker da Bayes regel

   P(S|D)= P(S)*P(D|S) / P(D) = (0.03*0.8)/0.0725 =0.331

c. To tilfeller: Positiv diagnose for person med Sol og negativ diagnose for person uten Sol.

$$P(D \cap S) + P(D' \cap S') = \mathrm{P(D)} * \mathrm{P(S|D)} + \mathrm{P(D')} * \mathrm{P(S'|D')}$$

   =0.03*0.8+0.97*0.95=0.9455

Oppgave 2

a. Binomisk fordeling

$$P(5) = \frac{8!}{(8-5)!\,5!} 0.2^5 0.8^3 = 0.009$$

b. Beregner med normaltilnærming
   N=750
   π=0.2
   πn=150

$$\sigma^2 = 750 * 0.2 * 0.8 = 120$$

$$X \sim N(150,120)$$

P(X≥175) =

$$P\left(\frac{175-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(\frac{25-0.5}{10.95} \leq Z\right) = P(2.28 \leq Z) = 0.0125$$

c. H0: p=0.2, H1: p>0.2
d. Bruker test på p for binomisk fordeling (stort utvalg)

Testobservator:

$$Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p0(1-p0)}{n}}}$$

Setter inn for $\widehat{p = 0.25, p_0 = 0.2}, n = 200$

Dette gir Z=1.77 Forkaster H0 siden kritisk verdi er 1,645

e. Starter med å finne kritisk verdi ensidig test 1 % nivå som er 2.33

Testobservatoren skal bli 2.33 og den observerte andelen er ukjent. N=1000

Lar den ukjente andelen være X. Løser vi likningen finner vi at vi forkaster H0 ved X=0.2294

X=x/n =x/1000. Finner da at hvis 230 svarer de vil stemme Høyre vil hullhypotesen forkastes.

Oppgave 3

a. P=22.9 + 1.044T
   En time ekstra forberedelse siste uka gir 1.044 poeng mer på eksamen.
   Antar lineær sammenheng
b. 0.73
c. C. Må først beregne standardavvik som er 10.35

To standardavvik er da 20.7

Økning er da 21.61 (Fra gjennomsnitt på 26.8 øker poeng da til 72,49)

Oppgave 4

a. Følger av teksten at $X \sim N(\mu, \sigma) = N(200{,}40)$

$$P(X < 220) = P\left(Z < \frac{220 - 200}{40}\right) = P(Z < 0.5) = 0.6915$$

b. $P(X \geq 190) = 1 - P(X \leq 190) = 1 - 0.4013 = 0.5987$

P(190<X>210)=0,1974

c. Normalfordelingskurven er symmetrisk om forventningsverdien (200). Dette betyr at arealet til venstre for 180 er lik arealet til høyre for 220. Bruker da svaret for a der vi fant arealet til venstre for 220.

Har da at $P(X < 180) = P(X \geq 220) = 1 - P(X \leq 220) = 1 - 0.6915 = 0.3085$

d. Skal her finne 5 % av personer som ser mest TV. Sannsynligheten for at en standard normalfordelt variabel er større enn 1.64 er 5 %.

X-200/40=1,645 som gir x=265,8

e. Dette er et binomisk forsøk. P(Y<3) =P(Y<=2) =P(Y=0)+P(Y=1)+P(Y=2) =0,001948

f. T-fordeling, alfa=0.05

$$TS = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{189{,}55 - 200}{23{,}8/\sqrt{9}} - 1{,}315$$

N=9, 8 df gir $t_{0,05}(8) = 1{,}86$

Beholder H0

g. Flere observasjoner, bruker normalfordeling

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

Setter inn og får TS=-2,66 Forkaster H0

Forskjellen er at her forkaster vi null selv om utvalgssnittet er mindre redusert. Men større utvalg gir mindre standadrdavvik. Vi kan også anta normalfordeling slik at halene er mindre fete.