

Spørsmål 1

Irene deltek jevnleg i spydkastkonkurranser og vil undersøkje resultatane sine. Anta at lengda, i meter, som Irene kastar i spydkast er normalfordelt med ukjent forventningsverdi μ og kjend standardavvik $\sigma = 8.0$.

Irene ynskjer å teste nullhypotesen at ho gjennomsnittleg kastar 51.0 meter mot den alternative hypotesen at ho kastar lengre enn 51.0 meter. Det vil seie, Irene vil teste

$$H_0 : \mu = 51.0 \text{ mot } H_1 : \mu > 51.0.$$

Irene går fram ved å se på 25 spydkast og målar lengda til kvar av dei. Vidare bestemmer ho seg for å forkaste nullhypotesen dersom gjennomsnittleg lengde av dei 25 kasta er større enn 52.5.

(a) Kva er sannsynet for type I feil? Oppgi svaret som eit desimaltal med tre desimalar, for eksempel 0.053 eller 0.125.

$$① \quad H_0 : \mu = 51$$

$$H_1 : \mu > 51$$

$$② \quad \hat{\mu} = \bar{x} \quad \hat{\mu} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$③ \quad Z = \frac{\hat{\mu} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

$$④ \quad \text{Forkastningsregelen er gitt: } \hat{\mu} > 52.5$$

$$\Rightarrow \text{Vi ønsker } P(\hat{\mu} > 52.5 \mid \mu = 51.0)$$

$$\text{La } Z = \frac{52.5 - 51}{\sqrt{8^2/25}} = \frac{15}{16} = 0.9375$$

$$\Rightarrow P(Z > 0.9375)$$

$$\begin{aligned} ⑤ \quad \text{Tabell } P(Z > 0.9375) &= 1 - P(Z \leq 0.9375) \\ &= 1 - 0.8264 \leftarrow \text{tabell} \\ &= \underline{0.1736} \end{aligned}$$

(b) Dersom $\mu = 52.0$, kva er sannsynet for type II feil? Oppgi svaret som eit desimaltall med tre desimalar, for eksempel 0.053 eller 0.125.

Ser etter $P(\hat{\mu} < 52,5 \mid \underbrace{H_0 \text{ feil}}_{\text{gitt, } \mu=52 \Rightarrow H_0 \text{ er feil}})$

$$\Rightarrow P(\hat{\mu} < 52,5) = P(Z < 0,3125) = \underline{\underline{0,6217}}$$

$$Z = \frac{52,5 - 52}{\sqrt{8^2/25}} = 0,3125$$

Spørsmål 2

Syklisten Arthur ønsker å undersøke sin forventede verdi μ av hemoglobin i blodet når han ikke har dopet seg. I denne forbindelse har han hver morgen i seks dager målt hemoglobinverdien, i g/dl, i blodet sitt. Følgende verdier ble observert: 15.61, 15.89, 16.18, 16.48, 15.66 og 16.35. Vi antar i denne oppgaven at de seks målingene kan betraktes som uavhengige realisasjoner trukket fra en normalfordeling med (ukjent) forventning μ g/dl og (ukjent) standardavvik σ g/dl.

Arthur ønsker å teste om de seks målingene gir grunnlag for å tro at hans forventede hemoglobinverdi (i g/dl) er større enn $\mu_0 = 16$. Han setter opp en hypotesetest med nullhypotese H_0 og alternativ hypotese H_1 som følger:

$$H_0 : \mu = 16 \quad \text{og} \quad H_1 : \mu > 16.$$

En passende testobservator for denne hypotesetesten er

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{6}},$$

der $\bar{X}$ betegner gjennomsnittet av måleverdiene og S det empiriske standardavviket. Testens forkastningsområde ved signifikansnivå 5% kan da defineres som $\{T > k\}$, der k er et tall slik at $P(T > k) = 0.05$ når $\mu = \mu_0$.

(a) Finn verdien til k . Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.053 og 12.125.

Vi har student t-fordeling $\times$ tabell

$$\Rightarrow k = t_{n-1, \alpha} = t_{5, 0,05} = \underline{\underline{2,015}}$$

(b) Arthurs hypotesetest resulterer i forkastning av nullhypotesen. True eller False?

$$\bar{X} = \frac{15,61 + 15,89 + 16,18 + 16,48 + 15,66 + 16,35}{6}$$

$$\bar{X} = 16,028$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$S^2 = \frac{1}{5} \sum_{i=1}^6 (X_i - 16,028)^2 = 0,132$$

$$T = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} = \frac{16,028 - 16}{\sqrt{0,132/6}} = 0,189 < k \Rightarrow \underline{\text{Forkaster ikke}}$$

Spørsmål 3

En forskningsgruppe på et sykehus vil undersøke om det er forskjell i gjennomsnittlig konsentrasjon av et visst stoff i blodet hos kvinner og menn ved hjelp av en hypotesetest. De tar blodprøver av 64 menn og 69 kvinner og måler konsentrasjonen av det aktuelle stoffet i ppm (parts per million). Gjennomsnittlig konsentrasjon av stoffet ble målt til å være 21.1 ppm for mennene og 25.1 ppm for kvinnene. Standardavviket for de to kjønnene antas kjent og er 10.3 ppm for menn og 14.6 ppm for kvinner.

Forskningsgruppen setter opp følgende nullhypotese $H_0 : \mu_m = \mu_k$, og følgende alternativ hypotese $H_1 : \mu_m \neq \mu_k$, der μ_m og μ_k er forventet konsentrasjon for henholdsvis kvinner og menn.

(a) Hva blir absoluttverdien av (den standardiserte) testobservatoren i denne hypotesetesten? Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.239 og 10.933.

$$\bar{X}_m = 21,1 \quad , \quad \bar{X}_k = 25,1$$

$$\sigma_m = 10,3 \quad , \quad \sigma_k = 14,6$$

$$L_a \quad \hat{\theta} = \bar{X}_m - \bar{X}_k$$

$$E[\hat{\theta}] = E[\bar{X}_m - \bar{X}_k] = E[\bar{X}_m] - E[\bar{X}_k] = \mu_m - \mu_k$$

$$\begin{aligned} \text{Var}[\hat{\theta}] &= \text{Var}[\bar{X}_m - \bar{X}_k] = \text{Var}[\bar{X}_m] + \text{Var}[\bar{X}_k] = \text{Var}[\bar{X}_m] + \text{Var}[\bar{X}_k] \\ &= \text{Var}\left[\frac{1}{n} \sum x_i\right] + \left[\frac{1}{m} \sum x_j\right] = \frac{1}{n^2} \cdot n \cdot \sigma_m^2 + \frac{1}{m^2} \cdot m \cdot \sigma_k^2 = \frac{\sigma_m^2}{n} + \frac{\sigma_k^2}{m} \end{aligned}$$

Testobservator:

$$Z = \frac{\hat{\theta} - E[\hat{\theta}]}{\sqrt{\text{Var}[\hat{\theta}]}} = \frac{\hat{\theta} - (\mu_m - \mu_k)}{\sqrt{\frac{\sigma_m^2}{n} + \frac{\sigma_k^2}{m}}} = \frac{(21,1 - 25,1)}{\sqrt{\frac{10,3^2}{64} + \frac{14,6^2}{69}}} = -1,836$$

$$\underline{|Z| = 1,836}$$

(b) Hypotesetesten forkastes dersom testobservatoren har en absoluttverdi større enn en kritisk verdi som bestemmes av fordelingen til testobservatoren og signifikansnivået for testen. Hva blir den kritiske verdien dersom forskningsgruppen bruker et signifikansnivå på 0.05? Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.325 og 2.442.

$$\text{Ønsker } P(|Z| > k | H_0) = P(Z < -k | H_0) + P(Z > k | H_0) = 0,05$$

Z er normalfordelt fra sentralgrenseteoremet

$$\text{Toedeling} \Rightarrow k = Z_{\alpha/2} \text{ (eller } -Z_{1-\alpha/2})$$

$$\Rightarrow \text{Finn } \underline{Z_{0,025} = 1,960}$$

(c) Forskningsgruppen forkaster nullhypotesen. True eller False?

$$|Z| = 1,836 < k$$

$\Rightarrow$ Den forkastes ikke, False

Spørsmål 4

Vi ønsker å undersøke den gjennomsnittlige vekten til egg verpet fra en gruppe høns på et stort gårdsbruk. Vi har et tilfeldig utvalg (random sample) bestående av 5 observasjoner, $x_1, \dots, x_5$, der hver observasjon x_i betegner den målte vekten, i antall gram, til egg nr. i . Eggene er verpet uavhengig fra hver sin høne. Følgende tabell gir oversikt over observasjonene:

64.21	64.09	65.78	66.54	64.99
-------	-------	-------	-------	-------

Vi antar at vekten til hvert egg er normalfordelt med ukjent forventning μ og kjent standardavvik $\sigma=1.0$. Bonden ved gårdsbruket påstår at $\mu > 65$ g. Vi ønsker å utføre en hypotesetest av denne påstanden, og tester nullhypotesen $\mu=65$ g mot bondens alternative hypotese. Hva blir P-verdien? Oppgi svaret som et desimaltall med tre desimaler, som f.eks. 0.053 og 0.125.

$$\textcircled{1} H_0: \mu = 65$$

$$H_1: \mu > 65$$

$$\textcircled{2} \quad \hat{\mu} = \bar{X} \sim n(\mu, \frac{\sigma^2}{n})$$

$$\bar{X} = \frac{64,21 + 64,09 + 65,78 + 66,54 + 64,99}{5} = 65,122$$

$$\textcircled{3} \quad T = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$$

$$\begin{aligned} \Rightarrow P(T > t_{obs} | H_0) &= P(T > \frac{65,122 - 65}{\sqrt{1^2/5}} = 0,2728) \\ &= P(T > 0,2728) \\ &= 1 - P(T \leq 0,2728) \quad \leftarrow \text{tabel} \\ &= 1 - 0,6064 \\ &= \underline{\underline{0,3936}} \end{aligned}$$

Spørsmål 5

I Trondheim er det mange lyskryss. I et av disse kryssene lyser trafikklyset rødt i 2 minutter, grønt i 2 minutter, rødt i 2 minutter, osv (for enkelhets skyld anser vi tiden der lyset er gult som neglisjerbar). Det er dermed rimelig å anta at sannsynligheten for å bli møtt av rødt lys er $p = 0.5$.

Petter kjører gjennom dette krysset til og fra jobb hver arbeidsdag. Han vet imidlertid ikke om at det røde og grønne lyset varer like lenge. Petter synes til stor frustrasjon at han mye oftere møter rødt lys enn grønt lys. Han er derfor interessert i å estimere sannsynligheten p for å bli møtt av rødt lys.

Hver arbeidsdag over en periode på 6 uker, som tilsvarer 30 arbeidsdager og dermed 60 kjøring gjennom lyskrysset, noterer Petter om han blir møtt av rødt eller grønt lys. Petter vil estimere p med estimatoren $\hat{p} = \frac{X}{60}$, der X er antall ganger han møter rødt lys. Resultatet blir at han møter rødt lys 39 ganger.

(a) Hva blir Petters estimat for p ? Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.053 og 0.125.

$$\hat{p} = \frac{X}{60} = \frac{39}{60} = \underline{\underline{0,650}}$$

Petter forteller en av kollegaene sine på jobb om eksperimentet sitt. Kollegaen vet imidlertid om at trafikklyset i virkeligheten lyser rødt og grønt like lenge. Han opplyser Petter om dette, men Petter blir ikke overbevist. Petter setter opp følgende hypotesetest:

$$H_0 : p = 0.5 \text{ vs } H_1 : p > 0.5$$

(b) Benytt sentralgrenseteoremet for å (tilnærmet) regne ut p-verdien til hypotesetesten. Oppgi svaret som et desimaltall med tre desimaler, for eksempel 0.053 og 0.125.

X er binomisk fordelt

$$E[X] = np$$

$$\text{Var}[X] = np(1-p)$$

$$\hat{p} = \frac{X}{60} = \frac{X}{n}$$

$$E[\hat{p}] = \frac{1}{n} E[X] = \frac{1}{n} n \cdot p = p$$

$$\text{Var}[\hat{p}] = \frac{1}{n^2} \text{Var}[X] = \frac{p(1-p)}{n}$$

$$T = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

$$P(T > t_{obs} | H_0) = P\left(T > \frac{0,650 - 0,5}{\sqrt{\frac{0,5(1-0,5)}{60}}}\right) = P(T > 2,32379)$$

$$= 1 - P(T < 2,32)$$

$$= 1 - 0,9898$$

$$= \underline{\underline{0,0102}}$$

(c) Hvis Petter benytter signifikansnivå $\alpha = 0.05$, vil han forkaste nullhypotesen H_0 ? Skriv inn svaret som tallet 1 hvis han forkaster, og tallet 0 hvis han ikke forkaster.

$$0.01 < 0.05$$

$$\Rightarrow p\text{-verdi} \leq \alpha$$

$$\Rightarrow \underline{\text{Forkast } H_0, \text{ Svar} = 1}$$

Spørsmål 6

Petter har generert 10 tilfeldige tall fra en uniformfordeling på intervallet $[0, \theta]$, med en verdi for θ som bare han kjenner. Han viser tallene til Olav og ber ham gjette hva θ er. Olav baserer seg naturlig nok på estimatoren $\hat{\theta} = \max\{X_1, \dots, X_{10}\}$, og observerer at $\hat{\theta} = 1517.2$, som følgelig er hans gjetting. Petter genererer så 10 nye tall fra samme uniformfordeling. Han viser dem til Olav og spør om han beholder gjettingen sin. Olav ønsker å avgjøre dette ved å basere seg på følgende hypotesetest:

$$H_0 : \theta = 1517.2 \quad \text{mot} \quad H_1 : \theta > 1517.2$$

Olav ønsker å utføre testen med signifikansnivå $\alpha = 0.05$ basert på estimatoren $\hat{\theta}$. Han ser at maksimum av de nye tallene er 1512.7.

(a) Hva er p-verdien til det nye estimatet? Oppgi svaret som et desimaltall med tre desimaler, som f.eks. 0.234 og 0.611.

$$P(\theta \leq X) = \underset{\substack{\text{kumulert} \\ \text{for maks}}}{F_{\hat{\theta}}(X)} = \underset{\substack{\text{uniform} \\ \text{fordeling}}}{F_X(X)} = \left(\frac{X - 0}{\theta - 0}\right)^n = \left(\frac{X}{\theta}\right)^n$$

$$\text{La } T = \hat{\theta}$$

$$\begin{aligned} P(T > t_{\text{as}} | H_0) &= P(T > 1512.7 \mid \theta = 1517.2) \\ &= 1 - P(T \leq 1512.7 \mid \theta = 1517.2) \\ &= 1 - \left(\frac{1512.7}{1517.2}\right)^{10} \\ &= \underline{\underline{0.029}} \end{aligned}$$

(b) Hva er forkastningsgrensen til testen? Oppgi svaret som et desimaltall med to desimaler, som f.eks. 0.23 og 0.61. (OBS: Svaret kan være litt "uintuitivt")

Skal ha

$$P(T > k | H_0) = 0,05$$

$$1 - P(T \leq k | H_0) = 0,05$$

$$1 - \left(\frac{k}{1517,2}\right)^{10} = 0,05$$

$$10 \cdot \ln\left(\frac{k}{1517,2}\right) = \ln 0,95$$

$$\ln(k) - \ln 1517,2 = \frac{\ln 0,95}{10}$$

$$\ln k = \frac{\ln 0,95}{10} + \ln 1517,2$$

$$k = \exp\left(\frac{\ln 0,95}{10} + \ln 1517,2\right)$$

$$\underline{\underline{k = 1509,44}}$$