Statistikk:

Matematiske fag, NTNU

Hypotesetesting - tabell

Decision

Truth	H 0	H 1
H 0	Riktig	Type I feil
H 1	Type II feil	Riktig

Hypotesetesting

H_0 : etablert hypotese.
H_1 : ny, annerledes hypotese.
Har vi evidens til å endre syn? Forlate det etablerte (H_0) og gå til det nye paradigmet (H_1)?

Riktig og feil beslutning

1. Type I feil: H_0 forkastes, når H_0 er riktig.
2. Type II feil: H_0 forkastes ikke, når H_0 er gal.

Vi vil ha liten sannsynlighet for Type I feil.
Vi ønsker ikke å forlate det etablerte uten at det er tilstrekkelig evidens for dette.
$P(\text{Type I feil}) = P(\text{forkast} H_0 | H_0 \text{ er riktig}) = \alpha$.
α er signifikansnivå.
Typisk $\alpha = 0.05, 0.01$ eller lignende.
Avhengig av kostnad ved å skifte til H_1.

Hypotesetesting for to-utvalg

- Populasjon 1: $X_1, \ldots, X_{n_1}$
- Populasjon 2: $Y_1, \ldots, Y_{n_2}$.

H_0 : De to populasjonene kommer fra en fordeling med samme parametre.
H_1 : De to populasjonene kommer fra like fordelingsklasser men med ulike parametre.

Eksempel binomisk

- Populasjon 1: n_1 forsøk, $X_i \sim Bin(p_1, 1)$. $\hat{p}_1 = \frac{\sum_{i=1}^{n_1} X_i}{n_1}$,
- Populasjon 2: n_2 forsøk, $Y_j \sim Bin(p_2, 1)$. $\hat{p}_2 = \frac{\sum_{j=1}^{n_2} Y_j}{n_2}$

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \approx N(0,1)$$

Under H_0: $p_1 = p_2$.

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \approx N(0,1)$$

Eksempel: Kirurgi, infeksjoner (Joseph Lister, 1867)

Arbeid med amputasjoner. Hjelper desinfeksjonsmiddel? Lister tror det.

- Populasjon 1: Med desinfeksjonsmiddel. $p_1 = P(\text{ overlevelse })$
- Populasjon 2: Uten desinfeksjonsmiddel. $p_2 = P(\text{ overlevelse }),$

H_0: $p_1 = p_2$. H_1: $p_1 > p_2$.

Signifikansnivå $\alpha = 0.01$

Eksempel: Kirurgi, infeksjoner (Joseph Lister, 1867)

$n_1 = 40 =$ antall amputasjoner utført med desinfeksjonsmiddel. Andel overlevende $\hat{p}_1 = 34/40 = 0.85$.
$n_2 = 35 =$ antall amputasjoner utført uten desinfeksjonsmiddel. Andel overlevende $\hat{p}_2 = 19/35 = 0.54$.

Eksempel: Kirurgi, infeksjoner (Joseph Lister, 1867)

Under H_0: Forkast dersom $Z > z_{0.01} = 2.33$. Vi får:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{0.85 - 0.54}{\sqrt{\frac{0.85 \cdot 0.15}{40} + \frac{0.54 \cdot 0.46}{35}}} = 3.05$$

Forkast H_0. Desinfeksjonsmiddel hjelper under kirurgiske inngrep.

P-verdi: $P(N(0,1) > 3.05) = 0.0011$. Liten. Forkast H_0.

Ensidige og tosidige hypotesetester

Vi kan ha ensidige tester: $H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_0$ eller $H_0 : \mu = \mu_0$, $H_1 : \mu < \mu_0$
Vi kan ha tosidige test$erH_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$. Her vil testen H_0 ikke forkastes dersom μ_0 er innen konfidensintervallet for μ.

Hypotesetesting for ett-utvalg

Data: $X_1, \ldots, X_n$

- Binomisk fordeling: Eksakt test, Normalapproksimasjon, med eller uten halvkorreksjon. $Z = \frac{\hat{p}-p_0}{\sqrt{\hat{p}(1-\hat{p})}/\sqrt{n}} \sim N(0,1)$
- Normalfordelt, kjent varians : Normaltest: $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$
- Normalfordelt, ukjent varians : T-test: $T = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$
- Normalfordelt og test på varians σ: $\frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$

Eksempel hypotesetest - eksempel (H2014,3)

Oppgave 3

To El-bil produsenter konkurrerer om det samme kundemarkedet med hver sine modelltyper. La oss kalle de to modelltypene for modell A og modell B. En viktig faktor for valg av modelltype er rekkevidden som bilen har før en må lade batteriet på nytt. Produsenten av modell B beskylder produsenten av modell A for å oppgi for lang rekkevidde. De hevder at rekkevidden som er oppgitt ikke er realistisk å oppnå under normale kjøreforhold. For å undersøke dette nærmere bestemmer produsenten av modell A seg for å gjøre et lite forsøk. 10 sjåfører får i oppdrag å kjøre hver sin bil i noe som ble definert som normalterreng og med en snittfart på 60 km/t. La X_i være rekkevidden for sjåfør i, $i = 1, 2, \cdots, 10$. De observerte rekkeviddene i km er gitt nedenfor:

x_i: 123.0, 119.3, 119.4, 118.2, 119.5, 118.5, 117.7, 118.1, 119.6, 119.4

La $E(X_i) = \mu_A$ og $Var(X_i) = \sigma^2$, $i = 1, 2, \cdots, 10$

a) Sett opp forventningsrette estimatorer for μ_A og σ^2 basert på hele utvalget. Hva blir estimatene når $\sum_{i=1}^{10} x_i = 1192.70$ og $\sum_{i=1}^{10} (x_i - \bar{x})^2 = 19.68$?

Eksempel hypotesetest - eksempel (H2014,3)

Produsenten av modell A oppgir at forventet rekkevidde under slike forhold er 125 km. Anta nå at $X_i \sim N(\mu_A, \sigma^2)$ og uavhengige, $i = 1, 2, \cdots, 10$.

b) Kan en ut i fra dataene konkludere med at produsenten har satt forventet rekkevidde for høyt? Formuler dette som ett hypotesetestingsproblem og utfør testen. Bruk signifikansnivå $\alpha = 0.01$.

Eksempel - gammel eksamen

Oppgave 3 Vinkorker

Kraften som er nødvendig for å trekke ut en kork fra en vinflaske er en viktig egenskap ved korken. Dersom kraften er for liten gir ikke korken nok beskyttelse mot innsig av luft for vinen inni flasken. Dersom kraften er for stor, vil korken være vanskelig å fjerne. Kraften (Newton) for en bestemt korktype kan antas normalfordelt (Gaussisk fordelt) med forventning μ og standardavvik σ.

a) Anta i dette punktet at kraften har forventning 310.0 og standardavvik 36.0.

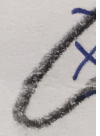
Hva er sannsynligheten for at kraften som er nødvendig for å trekke ut en kork er mellom 300.0 og 310.0?

En kork ble utsatt for en kraft på 330 uten at korken gikk ut. Hva er sannsynligheten for at en kraft større enn 360 er nødvendig for å trekke ut korken?

For et utvalg på 8 korker, hva er sannsynligheten for at utvalgsgjennomsnittet av kreftene er større enn 320?

Et utvalg av 8 tilfeldig valgte flasker med samme korktype ble plukket ut og kreftene som var nødvendig for å fjerne korkene var:

305.98 205.48 322.97 198.58 191.76 288.50 341.18 222.62



En ønsker at kraften som er nødvendig for å fjerne korken har forventning 310. Da gir korken god beskyttelse og er samtidig enkel å ta ut.

Eksempel - gammel eksamen

Generelt ønsker en lavest mulig standardavvik for korkene, og det er et krav at standarda
for kraften ikke er større enn 36.0. Produsenten av korkene hevder at dette kravet er op

c) Test $H_0 : \sigma = 36$ mot $H_1 : \sigma > 36$ ved signifikansnivå 5%.

Hva menes med p-verdi?

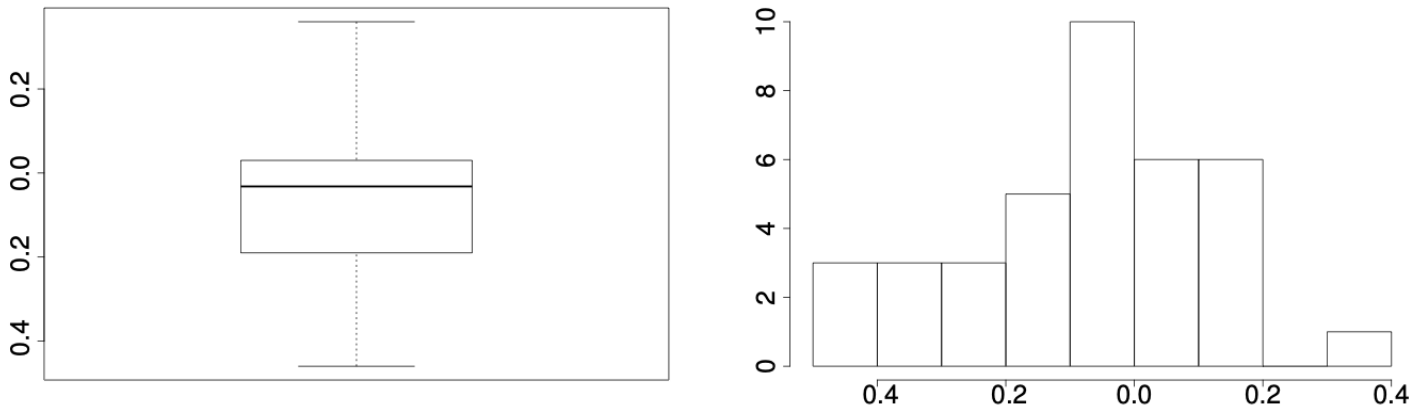
Finn p-verdien for testen over.

Eksempel hypotesetest - eksempel (S2015, 2)

Oppgave 2 Siste indre på 500m skøyteløp

Ved verdensmesterskap på enkeltdistanser på skøyter går hver deltager to 500 meters løp, ett løp hvor deltageren har indre bane i siste sving og ett løp hvor deltageren har ytre bane i siste sving. Rekkefølgen av deltagerne baserer seg på summen av tidene på de to løpene. Tilsvarende regel benyttes også i olympiske leker. Denne regelen ble innført fra og med verdensmesterskapet på Hamar i 1995. Tidligere ble rekkefølgen av deltagerne basert på kun et løp for hver deltager. Bakgrunnen for regelen om at hver deltager skal gå to løp er at det kan være en fordel å ha siste ytre siden løperne har stor fart i siste sving og i indre bane er krumningen større enn i ytre bane.

Eksempel hypotesetest - eksempel (S2015,2)



Figur 1: Boxplott og histogram over differansene d_i for de 37 løperne i 500 meter for menn i Sochi som ikke falt. I boxplottet er differansene langs y-aksen, mens i histogrammet er differansene langs x-aksen.

d) Formuler en hypotesetest for situasjonen med utgangspunkt i $\bar{D}$. Spesifiser H_0 og H_1, velg en passende testobservator og utled en beslutningsregel når signifikansnivået er α.

Av de 39 løperne i Sochi var det to som falt i et av sine løp. Dersom vi ser bort fra disse to løperne gir resultatene i Sochi $n = 37$, $\sum_{i=1}^{37} d_i = -2.654$ og $\sum_{i=1}^{37} d_i^2 = 1.552$. Hva blir da konklusjonen av Test 2 med resultatene fra Sochi når $\alpha = 0.05$? Avrund om nødvendig til nærmeste frihetsgrad oppgitt i tabell.

Eksempel hypotesetest - eksempel (H2022,3)

TMA4240 Statistikk, 3. desember 2022 – bokmål *Side 3 av 3*

Oppgave 3 Bilverkstedet

Et verksted klassifiserer feil i det elektriske anlegget til en bil til å være av 2 typer, A-feil og B-feil, avhengig av hvilken feilkode som blir gitt i feilsøkingen. Tidene målt i timer som går med til å finne en A-feil, X, og en B-feil, Y, er gitt ved sannsynsynlighetstetthetene $f_X(x) = \frac{1}{\beta}e^{-x/\beta}$, $x > 0$, $\beta > 0$ og $f_Y(y) = \frac{y}{\beta^2}e^{-y/\beta}$, $y > 0$, $\beta > 0$. Parameteren β er den samme i begge fordelingene.

Eksempel hypotesetest - eksempel (H2022,3)

$$\hat{\beta} = \frac{\sum_{i=1}^{m_1} X_i + \sum_{j=1}^{m_2} Y_j}{m_1 + 2m_2}$$

Finn forventning og varians til denne estimatoren.

d) En kunde har ventet i mange timer på at feilen skal bli funnet og mistenker at β er større enn 1 som er den verdien verkstedet hevder er riktig. Anta vi har $m_1 = m_2 = 100$ og at $\bar{x} = 1.43$ og $\bar{y} = 2.80$.

Forklar hvorfor $\hat{\beta}$ tilnærmet er normalfordelt.

Formuler en hypotesetest for denne situasjonen basert på mistanken til kunden. Foreslå en egnet testobservator (teststatistikk) og regn ut p-verdien til denne hypotesetesten med de gitte dataene. Kommenter resultatet.