

Forelesning 25.april

Intro til hypotesetesting (max 45 min)

Utforming av Prosjekt idé (45 min)

- Sitte sammen i prosjekt-2 gruppene
 - Jeg går rundt til hver av gruppene og hører hva de ser for seg
-

Hypotesetesting

- Hypotese er?
 - Et verktøy for å bestemme om en formulert hypotese stemmer eller ikke
 - Testen gjøres basert på data
-

Nullhypotesen og alternativhypotesen

- Nullhypotesen er det motsatte av det vi vil teste om er stemmer
 - I bøker vil denne ofte noteres med H_0
 - Alternativhypotesen er det vi ønsker å teste
 - H_1
-

Finne nullhypotesen

- Eks.1: Et firma vil sjekke om elektriske biler er mer energi effektive enn diesel-biler
 - Nullhypotese og alternativhypotese:
 - Eks.2: Ny medisin skal vurderes om endrer blodtrykket (en gruppe på medisin og en placebo)
 - Nullhypotese og alternativhypotese:
 - Eks.3 : Vi vil undersøke om elever som tar Exphil på 1-2-3 får dårligere karakterer enn de som ikke gjør det:
 - Nullhypotese og alternativhypotese:
-

- Eks.1: Et firma vil sjekke om elektriske biler er mer energi effektive enn diesel-biler
 - Nullhypotese: Effektiviteten til el-biler er lavere eller lik enn effektiviteten til diesel-biler
 - Alternativhypotese: Effektiviteten til el-biler er høyere enn effektiviteten til diesel-biler
-

- Eks.2: Ny medisin skal vurderes om endrer blodtrykket (en gruppe på medisin og en placebo)
 - Nullhypotese: Gjennomsnittlig blodtrykk til de på medisin er lik gjennomsnittlig blodtrykk for de uten medisin
 - Alternativhypotese: Gjennomsnittlig blodtrykk for de på og uten er ulik
-

- Eks.3 : Vi vil undersøke om elever som tar Exphil på 1-2-3 får dårligere karakterer enn de som ikke gjør det:
 - Nullhypotese: Karakterer for de med Exphil-123 er bedre eller lik karakterer for de uten Exphil 123

- Alternativhypotese: Karakterer for de med Exphil-123 er lavere enn karakterer for de uten Exphil 123
-

- Eks.1:
 - H_0 : Effektivitet el-bil \leq Effektivitet diesel-bil
 - H_1 : Effektivitet el-bil $>$ Effektivitet diesel-bil
 - Eks.2:
 - H_0 : Gjennomsnitt blodtrykk på medisin = Gjennomsnittlig blodtrykk uten medisin
 - H_1 : Gjennomsnitt blodtrykk på medisin \neq Gjennomsnittlig blodtrykk uten medisin
 - Eks.3:
 - H_0 : Karakter med Ex-123 \geq Karakter uten Exphil 1-2-3
 - H_1 : Karakterer med Ex-123 $<$ Karakter uten Exphil 1-2-3
-

Type en og type to feil

- Type en
 - Vi forkaster H_0 , men den er sann
 - Type to
 - Vi forkaster ikke H_0 , men den er ikke sann
-

Forkastningsområdet

Før vi gjør en hypotesetest må vi bestemme hvor sikre vi ønsker å være, før vi forkaster H_0
Altså hvilke resultater/data skal gjøre at vi forkaster nullhypotesen

- Bestemmes ut i fra signifikansnivå α vi bestemmer
 - Signifikansnivå er sannsynligheten vi tillater av type en feil kan inntreffe
 - Gjennomsnittlig karakter for de med Ex-123 er 0.4 svakere eller mer
-

P-verdi

- Hvor sikre er vi på at vi har tatt riktig beslutning fra hypotesetesten
 - Om forkastnings området er 0.4 eller mer, vil vi være enda sikrere om den gjennomsnittlige karakterforskjellen er 0.9 eller igjen 1.5
 - P-verdien forteller oss noe om dette
 - Definert som "Det laveste signifikansnivået vi kan bruke for å forkaste nullhypotesen"
-

Eksempel

Vi vil teste om ny type salve fungerer bedre mot eksem
Finner $n = 100$ mennesker som tester de to salvene på hver av hendene
Resultatet er f.eks at for $X = 65$ stykker var den nye salven bedre

Om vi antar binomisk fordeling med parameter p , får vi at vi kan lage følgende hypotesetest:
 $p > 0.5 \Rightarrow$ den nye salven bedre

$p \leq 0.5 \Rightarrow$ den nye salven ikke bedre

$H_0 : p \leq 0.5$

$H_1 : p > 0.5$

La oss ta en kalkulert gjetning på forkastingsområdet med $X \geq 62$

Så finner sansynligheten for at vi har gjort en type 1 feil:

$$P(\text{feil type 1}) = P(\text{forkaster } H_0 | H_0 \text{ sann}) = P(X \geq 62 | p \leq 0.5)$$

$$P(\text{feil type 1}) \leq P(X \geq 62 | p = 0.5) = 1\%$$

Altså hvis vi har $X \geq 62$ i forsøket vårt, kan vi med 99% sikkerhet si at vi ikke har gjort en type en feil.

Om vi hadde gått motsatt vei, som er vanlig, er 1% eller $0.01 = \alpha$.

Det betyr at om vi gjør salve testen vår og krever signifikansnivå på 0.01 er forkasningsområdet vårt $X \geq 62$

I eksemplet med salvene var det $X = 65$ som den nye salven fungerte bedre for.

Ut i fra signifikansnivået på $\alpha = 0.01$ skal vi da forkaste nullhypotesen.

P-verdien kan vi da regne ut som "hva er det minste signifikansnivået vi kunne valgt oss, men fortsatt forkastet nullhypotesen" :

$$P(X \geq 65 | p = 0.5) = 0.002$$

Altså en p-verdi på 0.002, (merk: Lavere jo bedre)

Kommentarer:

$$P(X \geq 62 | p = 0.5) = 0.010$$

- Dette er utregnet på bakgrunn av at vi vet hvilken fordeling resultatene våre kommer fra (den binomiske), og kan derfor regne sansynligheten for at vi skal få resultater lik eller større enn det vi fikk.
Hvordan går vi baklengs fra signifikansnivå og bestemmer forkasningsområdet?
 - Slår opp i tabell eller bruker pakker i python eller R
Det finner ulike hypotesetester (z-test og t-test er de mest kjente), utvalget og fordelingen angir hvilken man burde bruke.
-

Følgende skal oppgis og leveres inn som README.md på github-repoet for prosjekt 2 innen 2. mai:

- Beskrivelse av datasettet gruppa skal bruke. Dette kan enten være et datasett som finnes på nettet allerede, eller et datasett dere samler inn på en eller annen måte. F. eks. ved å skrape nettet eller snakke med ChatGPT.
- Problemstilling/forskningsspørsmål/prosjektbeskrivelse. Denne inneholde delproblemstillinger fra alle studieretninger som deltar i gruppa, samt en tverrfaglig problemstilling.