STK1110 Høsten 2021

Innledning til hypotesetesting

Tilsvarer Avsnitt 9.1

Ingrid Hobæk Haff
Matematisk institutt
Universitetet i Oslo

Eksempel

- Et farmasøytisk firma vil undersøke om en ny salve mot eksem er bedre enn den gamle.
- 100 pasienter med eksem på begge hendene blir med på et forsøk.
- Hver pasient får ved loddtrekning den nye salven på en hånd og den gamle salven på den andre hånden.
- Etter 3 uker avgjør legen hvilken av de to hendene som er best, og en teller opp antall pasienter for hvem den nye salven ga best resultat.

Eksempel (forts.)

- La oss si at den nye salven ga best resultat for 60 av pasientene.
- Kan det farmasøytiske firmaet med **rimelig grad av sikkerhet** konkludere med at den nye salven er bedre enn den gamle?
- For å kunne avgjøre det, må vi:
  - anta en statistisk modell
  - formulere problemet som utsagn/hypoteser om en parameter i modellen
- Vi gjør det da til et **hypotesetestingsproblem**.

Hypotesetesting

- I en statistisk hypotesetest tester en én hypotese mot en annen.
- **Nullhypotesen** H_0 er det utsagnet som i utgangspunktet antas å være sant, mens den **alternative hypotesen** H_a motsier H_0.
- Dersom dataene gir sterke indikasjoner om at H_0 ikke er sann, forkaster en denne til fordel for alternativet H_a.
- En kan da med rimelig grad av sikkerhet konkludere med at H_a er sann.

Hypotesetesting

- En **statistisk test** er en regel for når vi skal forkaste H_0.
- Den består i å spesifisere:
    1. en **testobservator**, som er en stokastisk variabel en baserer testen på.
    2. et **forkastningsområde**, som er verdiene av testobservatoren en skal forkaste H_0 for.

Eksempel (forts.)

- La X være antall pasienter som den nye salven har fungert best for.
- I det konkrete forsøket fikk X verdien 60, men hvis en hadde gjentatt forsøket ville nok X ha fått en annen verdi.
- Det er rimelig å anta at $X \sim Binomisk(n, p)$, der $n = 100$ og p er sannsynligheten for at den nye salven er best for en tilfeldig valgt pasient.
- Problemstillingen til det farmasøytiske firmaet kan nå formuleres som hypoteser om parameteren p:
  - hvis den nye salven er bedre enn den gamle, er $p > 0.50$
  - hvis den nye salven ikke er bedre enn den gamle er $p \leq 0.5$.

Eksempel (forts.)

- I dette eksempelet blir hypotesene som følger:
  - Nullhypotese $H_0 : p \leq 0.50$, altså den nye salven **er ikke** bedre enn den gamle.
  - Alternativ hypotese $H_a : p > 0.50$, altså den nye salven **er** bedre enn den gamle.
- Da $\mathrm{E}(X) = np = 100p$, er det naturlig å forkaste H_0 dersom X er tilstrekkelig stor, altså $X \geq k$ for en passende k.
- Hva bør k være?

Type I- og type II-feil

- I en hypotesetest har vi følgende mulige utfall:

	H_0 er sann	*H_a er sann*
Ikke forkaste H_0	Riktig	Feil
Forkaste H_0	Feil	Riktig

- En kan altså gjøre to typer feil:
  - **type I-feil**: forkaste H_0 når H_0 er sann.
  - **type II-feil**: ikke forkaste H_0 når H_a er sann.
- I hypotesetesting formulerer en problemet slik at feil av type I er verre enn feil av type II.

Type I- og type II-feil (forts.)

- Valget av forkastningsområde styres av hensynet til type I- og type II-feil.
- For en gitt testobservator og en gitt størrelse *n* på utvalget en bruker i testen vil en reduksjon av forkastningsområdet gi lavere sannsynlighet for type I-feil, men samtidig øke sannsynligheten for type II-feil.
- Valget av forkatningsområde må altså være en avveining mellom de to typene feil, men med størst hensyn til type I-feil.
- **Signifikansnivået** α til en test er (den maksimale) sannsynligheten for feil av type I.
- Vanlige verdier for α er 5% og 1%.

Styrkefunksjonen og type II-feil

- Anta at en har en test vedrørende verdien av parameteren θ, f.eks $H_0 : \theta \leq \theta_0$ mot $H_a : \theta > \theta_0$.
- Da er $\beta(\theta) = \mathrm{P}(\mathit{TypeII} - \mathit{feil}|\theta)$, altså sannsynligheten for ikke å forkaste H_0 for en gitt verdi av θ i samsvar med H_a.
- **Styrkefunksjonen** til testen er definert som

$$\begin{aligned}\gamma(\theta) =& \mathrm{P}(\text{Forkaste } H_0|\theta) \\ =& \begin{cases} \mathrm{P}(\text{Type I-feil}|\theta), & \theta \text{ i samsvar med } H_0 \\ 1 - \mathrm{P}(\text{Type II-feil}|\theta) = 1 - \beta(\theta), & \theta \text{ i samsvar med } H_a \end{cases}.\end{aligned}$$

- Styrkefunksjonen oppsummerer egenskapene til testen.

Eksempler

Eksempel

Forsøk med salve. Vil vil utføre testen og beregne styrkefunksjonen.

Eksempel

Eks. 9.2 fra boka. La X_i være tørketida på flate i, $i = 1, \dots, 25$. Det antas at $X_1, \dots, X_{25} \overset{uif}{\sim} N(\mu, 9^2)$. Vi vil teste

$$H_0 : \mu \geq 75 \text{ mot } H_a : \mu < 75.$$