TMA4240 Statistikk H2016 [21]

Hypotesetesting
Generelt om statistiske hypoteser [10.1-10.3]
Ett normalfordelt utvalg [10.5]

Mette Langaas

Institutt for matematiske fag, NTNU

wiki.math.ntnu.no/emner/tma4240/2016h/start/

Estimering og hypotesetesting

Fenomen	**Bilkjøring**
Spørsmål	Hvor stor andel av studentene ved NTNU synes de er flinkere enn den gjennomsnittlige bilfører til å kjøre bil?
Populasjon	Alle studenter ved NTNU
Parameter	Andelen p som synes de er flinkere enn gjennomsnittet.
Utvalg	Alle studenter som svarte på spørreundersøkelse.
Data, u.i.f og representative?	Flinkere eller ikke enn gjennomsnittet.

Estimering og hypotesetesting

Fenomen	**Bilkjøring**
Estimator	$\hat{p} = \frac{X}{n}$, X antall som synes de er flinkere enn gjennomsnittet av n spurte.
Størrelse med kjent fordeling	For store n, og p ikke for nært 0 eller 1, så er $Z = \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}}$ *tilnærmet* normalfordelt

Estimering og hypotesetesting

Fenomen	**Bilkjøring**
Konfidens intervall	$[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}]$
Hypotesetesting:	Er andelen av studenter som synes de er flinkere enn gjennomsnittet til å kjøre bil større enn 0.5?
Konklusjon	For hvilke verdier av $\hat{p}$ kan vi konkludere med at andelen i populasjonen er større enn 0.5? Vi observerte 0.51.

Kvalitetskontroll av skruer

- Produksjon av skruer.
- Lengden på produsert skrue skal være 15 mm.
- Tar jevnlig stikkprøve fra prosessen, for å sjekke om skruene som produseres er 15 mm lange.
- Hvis stikkprøven tyder på at de produserte skruene ikke er 15 mm, må maskinen som lager skruene kalibreres på nytt.
- Hvordan skal vi basert på et utvalg fra prosessen bestemme om maskinen må rekalibreres?

Hypotese

- En **statistisk hypotese** er et utsagn om egenskaper til en (eller flere) populasjone(r).
- **Nullhypotese:** Hypotesen vi vil undersøke om vi har grunnlag fra data til å forkaste.
- **Alternativ hypotese:** Reflekterer spørsmålet vi stiller (eller påstanden vi kommer med).

Hypotesetesting og rettsak

- Spørsmål: Er grunn til å tro at skruene som produseres ikke er tatt fra en populasjon med $\mu = 15$ mm?
- Statistisk hypotesetesting: Undersøke om det er nok bevis som underbygger at skruene ikke er fra en populasjon med lengde $\mu = 15$ mm. Som i rettssak: tiltalte er antatt uskyldig til han er bevist skyldig.
- Null hypotesen: skruene som produseres er fra en populasjon med $\mu = 15$ mm.
- Alternativ hypotese: skuene som produseres er ikke fra en populasjon med $\mu = 15$ mm.

$$H_0 : \mu = 15\text{mm} \quad \text{vs.} \quad H_1 : \mu \neq 15\text{mm}$$

To typer feil

- DEF 10.2: Forkasting av nullhypotesen når denne er sann, kalles en type-I-feil.
  - Vi vil være sikre på at skruene ikke er fra en populasjon med $\mu = 15$ mm før vi bestemmer oss for å stoppe produksjonen for å kalibrere. Produksjonsstopp for kalibrering av maskin gjør at produsenten taper penger pga. forsinket produksjon.
- DEF 10.3: Å ikke forkaste nullhypotesen når den er gal, kalles en type-II-feil.
  - Vi vil gjerne kalibrere maskinen på nytt hvis skruene som produseres ikke er fra en populasjon med $\mu = 15$ mm. For lange og for korte skruer påfører kjøper problemer.

Type-I og type-II-feil

	H_0 sann	H_0 falsk
Aksepter H_0	Korrekt	Type-II feil
Forkast H_0	Type-I feil	Korrekt

Ett utvalg: test for μ med σ kjent

	Generell fremgangsmåte	**Kvalitetskontroll av skruer**
0	$X_1, X_2, ..., X_n$ u.i.f. normal(μ, σ) der σ er kjent.	Stikkprøve (utvalg) av $n = 10$ skruer, antar normalfordeling og kjenner σ =0.1mm.
1	To-sidig test $H_0 : \mu = \mu_0$ *vs.* $H_1 : \mu \neq \mu_0$	Er grunn til å tro at skruene som produseres ikke er fra en populasjon med lengde $\mu = 15$ mm? $H_0 : \mu = 15$ *vs.* $H_1 : \mu \neq 15$
2	Signifikansnivå α bestemmes.	Velger $\alpha = 0.05$
3	Testobservator $Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ er under H_0 standard normalfordelt Forkast H_0 hvis $z_0 > z_{\frac{\alpha}{2}}$ eller $z_0 < -z_{\frac{\alpha}{2}}$.	

Ett utvalg: test for μ med σ kjent

	Generell fremgangsmåte	**Kvalitetskontroll av skruer**
4	$z_{\frac{\alpha}{2}}$	$z_{\frac{0.05}{2}} = 1.96$
	Observerer $\bar{x}$ fra utvalget (stikkprøven)	$\bar{x} = 15.05$ mm.
	Beregner $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$z_0 = \frac{15.05-15}{0.1/\sqrt{10}} = 1.58$
	Sammenligner $-z_{\frac{\alpha}{2}}$, z_0 og $z_{\frac{\alpha}{2}}$	$-1.96 < 1.58 < 1.96$
	Forkast H_0 og konkluder med H_1, eller behold H_0.	Beholder H_0. Har ikke sterke nok bevis for at $\mu \neq 15$mm.

Ett utvalg: tosidig test for μ med σ kjent

- $X_1, X_2, ..., X_n$ u.i.f. normal(μ, σ) der σ er kjent.
- To-sidig test:
  1. $H_0 : \mu = \mu_0 \qquad vs. \qquad H_1 : \mu \neq \mu_0$
  2. Signifikansnivå α bestemmes.
  3. Testobservator under H_0 er Z_0. $Z_0 = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ er under H_0 standard normalfordelt.
     Forkast H_0 hvis $|Z_0| > z_{\frac{\alpha}{2}}$.
  4. Beregn $\bar{x}$ fra utvalget, og videre $z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$.
     Sammenlign z_0 og $z_{\frac{\alpha}{2}}$, og forkast H_0 hvis $|z_0| > z_{\frac{\alpha}{2}}$.

Læringsmål

- Gjøre rede for hvordan bestemme null- og alternativ hypotese.
- Gjøre rede for hva Type I og Type II feil er, og hvordan man kan regne ut sannsynligheten for å begå disse feilene.
- Være fortrolig med begrepene forkastningsområde og signifikansnivå.
- Basert på forkastningsområdemetoden kunne spesifisere en tosidig test for forventningsverdien i en normal populasjon med kjent varians.