Ályktanir fyrir talnabreytur

Kennarar diffra 2024

Anna Helga Jónsdóttir

Helstu atriði:

1. Ályktanir um meðaltal þýðis
2. Ályktanir um meðaltöl tveggja þýða
3. Ályktanir fyrir paraðar mælingar
4. Ályktanir um dreifni þýðis
5. Ályktanir um dreifni tveggja þýða

Ályktanir um talnabreytur

- Byrjum á að skoða ályktanir fyrir meðaltal/meðaltöl þýðis/þýða:
  - Meðaltal eins þýðis.
  - Mismun meðaltala tveggja þýða.
  - Mismun meðaltals paraðra mælinga.
- Skoðum svo algengustu ályktanir fyrir meðaltal/meðaltöl:
  - Dreifni í einu þýði.
  - Dreifni í tveimur þýðum.
  - Dreifni í fleiri þýðum - bara í R.

Framkvæmd tilgátuprófa

Framkvæmd tilgátuprófa

1. Ákveða hvaða tilgátupróf er viðeigandi fyrir gögnin okkar.
2. Ákveða hæstu ásættanlegu villulíkur.
3. Setja fram núlltilgátu og ákveða um leið áttun prófsins (einhliða/tvíhliða).
4. Reikna prófstærðina sem svarar til tilgátuprófsins.

5a Kanna hvort prófstærðin falli á höfnunarsvæði tilgátuprófsins.

5b Kanna p-gildi tilgátuprófsins.

6 Draga ályktun.

Hvert erum við komin...

Ályktanir um meðaltal þýðis

Algeng leið til að draga ályktanir um talnabreytu er að skoða meðaltal hennar.

- Við látum μ lýsa þessu meðaltali fyrir allt **þýðið**.
- Við látum $\bar{x}$ lýsa þessu meðaltali fyrir **úrtakið okkar**.
- Við viljum nota mælingarnar okkar til að draga ályktanir um **þýðismeðaltalið** μ:
    - Með því að reikna **öryggisbil** fyrir μ.
    - Með því að framkvæma **tilgátupróf** um μ.

Meðaltal margra mælinga

- **Höfuðsetning tölfræðinnar** gefur okkur að líkindadreifing meðaltals breytu fylgir **normaldreifingu** ef það byggir á nógu mörgum mælingum af breytunni.
- Staðalfrávik meðaltalsins er staðalskekkjan, $\frac{\sigma}{\sqrt{\mathbf{n}}}$.
- Í um 95% tilvika mun meðaltal úrtaksins lenda innan tveggja staðalskekkja frá þýðismeðaltalinu.
- Við þekkjum nánast aldrei $\frac{\sigma}{\sqrt{\mathbf{n}}}$, en sé n mjög stórt er $\frac{\mathbf{s}}{\sqrt{\mathbf{n}}}$ góð nálgun.

Öryggisbil fyrir μ, þýði fylgir normaldreifingu og σ er þekkt (mjög óalgengt...)

Þýði fylgir normaldreifingu og σ er þekkt

Neðra mark $1-\alpha$ öryggisbils:

$$\bar{x} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$$

Efra mark $1-\alpha$ öryggisbils:

$$\bar{x} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$$

Öryggisbilið má því skrifa:

$$\bar{x} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Öryggisbil fyrir μ þegar n er mjög stórt

n er mjög stórt

Neðra mark $1-\alpha$ öryggisbils:

$$\bar{x} - z_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}.$$

Efra mark $1-\alpha$ öryggisbils:

$$\bar{x} + z_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}.$$

Öryggisbilið má því skrifa:

$$\bar{x} - z_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

þar sem $\bar{x}$ og s eru meðaltal og staðalfrávik úrtaksins.

Tilgátupróf fyrir μ, þýði fylgir normaldreifingu og σ er þekkt

Þýði fylgir normaldreifingu og σ er þekkt

Núlltilgátan er:

$$H_0 : \mu = \mu_0$$

Prófstærðin er:

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

Ef núlltilgátan er sönn fylgir prófstærðin stöðluðu normaldreifingunni, eða $Z \sim N(0,1)$.

Höfnunarsvæðin eru:

Gagntilgáta	Hafna H_0 ef:
$\mathrm{H}_1 : \mu < \mu_0$	$Z < -z_{1-\alpha}$
$\mathrm{H}_1 : \mu > \mu_0$	$Z > z_{1-\alpha}$
$\mathrm{H}_1 : \mu \neq \mu_0$	$Z < -z_{1-\alpha/2}$ eða $Z > z_{1-\alpha/2}$

Tilgátupróf fyrir μ þegar n er mjög stórt

n er mjög stórt

Núlltilgátan er:

$$H_0 : \mu = \mu_0$$

Prófstærðin er:

$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

Ef núlltilgátan er sönn fylgir prófstærðin stöðluðu normaldreifingunni, eða $Z \sim N(0,1)$.

Höfnunarsvæðin eru:

Gagntilgáta	Hafna H_0 ef:
$\mathrm{H}_1 : \mu < \mu_0$	$Z < -z_{1-\alpha}$
$\mathrm{H}_1 : \mu > \mu_0$	$Z > z_{1-\alpha}$
$\mathrm{H}_1 : \mu \neq \mu_0$	$Z < -z_{1-\alpha/2}$ eða $Z > z_{1-\alpha/2}$

Hvað ef úrtakið er ekki stórt?

- Meðaltöl fylgja normaldreifingu ef að upprunalegu mælingarnar sem þau byggja á fylgja normaldreifingu.
- Í þeim tilvikum getum við smíðað öryggisbil og tilgátupróf með svipuðum hætti.
- Nú er hins vegar talsverð óvissa í matinu á staðalskekkjunni, þ.e.a.s $\frac{s}{\sqrt{n}}$ getur orðið talsvert frábrugðið $\frac{\sigma}{\sqrt{n}}$.
- Ef að X fylgir normaldreifingu með meðaltal μ og staðalfrávik σ fylgir lýsistærðin:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

t-dreifingu með $\mathbf{n-1}$ **frígráðu**.

Öryggisbil með t-dreifingu

Normaldreift þýði eða n stórt

Neðra mark $1-\alpha$ öryggisbils:

$$\bar{x} - t_{1-\alpha/2,(n-1)} \cdot \frac{s}{\sqrt{n}}.$$

Efra mark $1-\alpha$ öryggisbils:

$$\bar{x} + t_{1-\alpha/2,(n-1)} \cdot \frac{s}{\sqrt{n}}.$$

Öryggisbilið má því skrifa:

$$\bar{x} - t_{1-\alpha/2,(n-1)} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{1-\alpha/2,(n-1)} \cdot \frac{s}{\sqrt{n}}$$

þar sem $\bar{x}$ og s eru meðaltal og staðalfrávik úrtaksins.

t-próf fyrir eitt meðaltal

t-próf: normaldreift þýði eða n stórt

Núlltilgátan er:

$$H_0 : \mu = \mu_0$$

Prófstærðin er:

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

Ef núlltilgátan er sönn fylgir prófstærðin t-dreifingu með $(n-1)$ frígráðu, eða $T \sim t_{1-\alpha,(n-1)}$.
Höfnunarsvæðin eru:

Gagntilgáta	Hafna H_0 ef:
$\mathrm{H}_1 : \mu < \mu_0$	$T < -t_{1-\alpha,(n-1)}$
$\mathrm{H}_1 : \mu > \mu_0$	$T > t_{1-\alpha,(n-1)}$
$\mathrm{H}_1 : \mu \neq \mu_0$	$T < -t_{1-\alpha/2,(n-1)}$ eða $T > t_{1-\alpha/2,(n-1)}$

Hvert erum við komin...

Ályktanir um meðaltöl tveggja þýða

- Skoðum nú tilgátupróf og öryggisbil sem eiga við þegar bera á saman meðaltöl tveggja þýða.
- Við köllum meðaltöl þýðanna μ_1 og μ_2.
- Við köllum meðaltöl úrtakanna $\bar{x}_1$ og $\bar{x}_2$.
- Við viljum framkvæma tilgátupróf sem prófar núlltilgátuna:

$$H_0 : \mu_1 - \mu_2 \ = \delta$$

- Við viljum reikna öryggisbil fyrir mismuninn:

$$\mu_1 - \mu_2$$

Ályktanir um meðaltöl tveggja þýða

- Við notum z-próf/öryggisbil eða t-próf/öryggisbil líkt þegar unnið er með eitt meðaltal þýðis.
- z-próf/öryggisbil eiga við þegar þýðin fylgja normaldreifingu og dreifnin í þýðunum er þekkt (mjög óalgengt...) eða þegar úrtökin eru mjög stór.
- t-próf/öryggisbil eiga við þegar þýðin fylgja normaldreifingu (óháð stærð úrtakanna) eða þegar úrtökin eru mjög stór (óhað dreifingu þýðanna).
- Notum mismunandi t-próf/öryggisbil eftir því hvort gera megi ráð fyrir að dreifni þýðanna sé svipuð eða ekki.
- Byrjum á að kanna hvor dreifnin sé ólík með viðeigandi tilgátuprófi - sjá tilgátupróf fyrir dreifni þýða.

tveggja hópa t-próf: svipuð dreifni

Ef gera megi ráð fyrir að dreifnin sé svipuð byrjum við á að reikna **vegna dreifni**, sem við táknum s_p^2:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

t-öryggisbil fyrir mismun meðaltala tveggja þýða - svipuð dreifni

t-öryggisbil fyrir mismun meðaltala tveggja þýða - svipuð dreifni

Neðra mark $1-\alpha$ öryggisbils er:

$$\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2,(n_1+n_2-2)} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Efra mark $1-\alpha$ öryggisbils er:

$$\bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2,(n_1+n_2-2)} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

t-próf fyrir meðaltöl tveggja þýða - svipuð dreifni

t-próf fyrir meðaltöl tveggja þýða - svipuð dreifni

Núlltilgátan er:

$$H_0 : \mu_1 - \mu_2 = \delta$$

Prófstærðin er:

$$T = \frac{\overline{X}_1 - \overline{X}_2 - \delta}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Ef núlltilgátan er sönn fylgir prófstærðin t dreifingu með $(n_1 + n_2 - 2)$ frígráður, eða $T \sim t_{(n_1+n_2-2)}$.

Athugið að δ getur verið hvaða fasti sem er en í langflestum tilvikum er $\delta = 0$. Höfnunarsvæðin eru:

Gagntilgáta	Hafna H_0 ef:
$H_1 : \mu < \mu_0$	$T < -t_{1-\alpha,(n_1+n_2-2)}$
$H_1 : \mu > \mu_0$	$T > t_{1-\alpha,(n_1+n_2-2)}$
$H_1 : \mu \neq \mu_0$	$T < -t_{1-\alpha/2,(n_1+n_2-2)}$ eða $T > t_{1-\alpha/2,(n_1+n_2-2)}$

Tveggja hópa t-próf: ólík dreifni

- Ef ekki er hægt að gera ráð fyrir að dreifnin sé sú sama reiknum við ekki vegna dreifni.
- Þá þurfum við hins vegar að reikna fjölda frígráða fyrir tilgátuprófið með formúlunni:

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

Öryggisbil fyrir meðaltöl tveggja þýða - ólík dreifni

Öryggisbil fyrir meðaltöl tveggja þýða - ólík dreifni

Neðra mark $1-\alpha$ öryggisbils er:

$$\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2,(\nu)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Efra mark $1-\alpha$ öryggisbils er:

$$\bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2,(\nu)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Öryggisbilið má því skrifa:

$$\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2,(\nu)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < \bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2,(\nu)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

þar sem $\bar{x}_1$, $\bar{x}_2$ eru meðaltöl úrtakanna og s_1^2, s_2^2 eru dreifni úrtakanna.

Tilgátupróf fyrir mismun meðaltala tveggja þýða - ólík dreifni

Tilgátupróf fyrir mismun meðaltala tveggja þýða - ólík dreifni

Núlltilgátan er:

$$H_0 : \mu_1 - \mu_2 = \delta$$

Prófstærðin er:

$$T = \frac{\overline{X}_1 - \overline{X}_2 - \delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Ef núlltilgátan er sönn fylgir prófstærðin t-dreifingu með ν frígráðum

Athugið að δ getur verið hvaða fasti sem er en í langflestum tilvikum er $\delta = 0$.
Höfnunarsvæðin eru:

Gagntilgáta	Hafna H_0 ef:
$\mathrm{H}_1 : \mu < \mu_0$	$T < -t_{1-\alpha,(n_1+n_2-2)}$
$\mathrm{H}_1 : \mu > \mu_0$	$T > t_{1-\alpha,(n_1+n_2-2)}$
$\mathrm{H}_1 : \mu \neq \mu_0$	$T < -t_{1-\alpha/2,(n_1+n_2-2)}$ eða $T > t_{1-\alpha/2,(n_1+n_2-2)}$

Hvert erum við komin...

Paraðar mælingar

- Margar rannsóknir eru gerðar á pöruðum slembiúrtökum.
- Þær rannsóknir eru oftar en ekki þannig að gögnum er aflað fyrir og eftir eitthvert inngrip.
- Tilgáturnar ganga þá iðulega út á að kanna hvort inngripið hafi borið árangur.
- Við notum parað próf til að kanna þessar tilgátur.

Paraðar mælingar

- Gerum ráð fyrir að við höfum n pör mælinga (X_i, Y_i), $i = 1, 2, 3, ..., n$.
- Þurfum að finna mismun þessara pöruðu mælinga,

$$D_i = X_i - Y_i.$$

  Við lítum á D_i sem slembiúrtak af stærð n úr þýði með meðaltal μ_D.
- Tilgátuprófin ganga út á að kanna μ_D.

Paraðar mælingar

Áður en við getum hafist handa við tilgátuprófin þurfum við að reikna eftirfarandi stærðir:

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}$$

sem er meðaltal mismunanna og

$$S_D{}^2 = \frac{\sum_{i=1}^{n} (D_i - \overline{D})^2}{n-1}$$

sem er dreifni mismunanna.

Ályktanir um paraðar mælingar

Ályktanir um paraðar mælingar

Núlltilgátan er:

$$H_0 : \mu_D = \mu_{D,0}$$

Prófstærðin er:

$$T = \frac{\overline{D} - \mu_{D,0}}{S_D/\sqrt{n}}$$

Ef núlltilgátan er sönn fylgir prófstærðin t-dreifingu með $(n-1)$ frígráðu, eða $T \sim t_{(n-1)}$ þar sem n er fjöldi para.
Höfnunarsvæðin eru:

Gagntilgáta	Hafna H_0 ef:
$\mathrm{H}_1 : \mu < \mu_0$	$T < -t_{1-\alpha,(n-1)}$
$\mathrm{H}_1 : \mu > \mu_0$	$T > t_{1-\alpha,(n-1)}$
$\mathrm{H}_1 : \mu \neq \mu_0$	$T < -t_{1-\alpha/2,(n-1)}$ eða $T > t_{1-\alpha/2,(n-1)}$

Dæmi um paraðar mælingar

Lyfjafyrirtæki hefur þróað blóðþrýstingslyf sem lækka á blóðþrýsting fólks. Fyrirtækið stóð fyrir rannsókn þar sem SBP 6 einstaklinga var mældur fyrir inntöku lyfsins og aftur eftir inntöku lyfsins. Niðurstöðurnar má sjá hér að neðan:

Einstaklingur	Blóðþrýstingur fyrir inntöku	Blóðþrýstingur eftir inntöku
1	125	119
2	119	117
3	136	121
4	127	122
5	119	113
6	115	116

Hvert erum við komin...

Ályktanir um dreifni þýðis

- Tilgátuprófin og öryggisbilin sem við munum skoða í þessum hluta eiga við þegar draga á ályktun um dreifni normaldreifðs þýðis, σ^2.
- Þegar reikna á öryggisbil og prófa tilgátur um dreifni þýðis er notast við χ^2-dreifinguna.
- Núlltilgátan í þessum hluta er að dreifni þýðisins sé jöfn einhverju ákveðnu gildi sem við köllum σ_0^2.
- Núlltilgátuna ritum við $H_0 : \sigma^2 = \sigma_0^2$.

Öryggisbil fyrir dreifni þýðis

Öryggisbil fyrir dreifni þýðis

Neðra mark $1-\alpha$ öryggisbils er:
$$\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2,(n-1)}}$$

Efra mark $1-\alpha$ öryggisbils er:
$$\frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2,(n-1)}}$$

Öryggisbilið má því skrifa:
$$\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2,(n-1)}} < \sigma^2 < \frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2,(n-1)}}$$

þar sem n er fjöldi mælinga í úrtakinu og s^2 er dreifni úrtaksins. $\chi^2_{1-\alpha/2,(n-1)}$ og $\chi^2_{\alpha/2,(n-1)}$ má finna í χ^2-töflu.

Tilgátupróf fyrir dreifni þýðis

Tilgátupróf fyrir dreifni þýðis

Núlltilgátan er:

$$H_0 : \sigma^2 = \sigma_0^2$$

Prófstærðin er:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

Gagntilgáturnar og höfnunarsvæðin eru

Gagntilgáta	Hafna H_0 ef:
$\mathrm{H}_1 : \sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{\alpha,(n-1)}$
$\mathrm{H}_1 : \sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{1-\alpha,(n-1)}$
$\mathrm{H}_1 : \sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{\alpha/2,(n-1)}$ eða $\chi^2 > \chi^2_{1-\alpha/2,(n-1)}$

Hvert erum við komin...

Ályktanir um dreifni tveggja þýða

- Tilgátuprófin sem við skoðum í þessum hluta eru notuð til að bera saman dreifni tveggja þýða sem fylgja normaldreifingu.
- Próf af þessu tagi þarf meðal annars oft að framkvæma áður en tilgátupróf þar sem meðaltöl tveggja þýða eru borin saman.
- Núlltilgátan í þessum hluta spyr hvort dreifni þýðanna tveggja sé jöfn, ritað $H_0 : \sigma_1^2 = \sigma_2^2$.
- Sé tilgátuprófið tvíhliða getum við dregið þær ályktanir að dreifnin sé ólík, sé það einhliða getum við fullyrt að dreifni annars þýðisins sé stærri en dreifni hins þýðisins.

Tilgátupróf fyrir dreifni tveggja normaldreifðra þýða

Tilgátupróf fyrir dreifni tveggja normaldreifðra þýða

Núlltilgátan er:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

Gagntilgátan getur verið einhliða eða tvíhliða og er prófstærðin mismunandi eftir því hvernig gagntilgátan er sett upp. Mögulegar gagntilgátur, prófstærðir og höfnunarsvæði þeirra má sjá hér að neðan.

Gagntilgáta	Prófstærð	Hafna H_0 ef:
$\mathrm{H}_1 : \sigma_1^2 < \sigma_2^2$	$F = \frac{S_2^2}{S_1^2}$	$F > F_{1-\alpha,(n_2-1,n_1-1)}$
$\mathrm{H}_1 : \sigma_1^2 > \sigma_2^2$	$F = \frac{S_1^2}{S_2^2}$	$F > F_{1-\alpha,(n_1-1,n_2-1)}$
$\mathrm{H}_1 : \sigma_1^2 \neq \sigma_2^2$	$F = \frac{S_M^2}{S_m^2}$	$F > F_{1-\alpha/2,(n_M-1,n_m-1)}$

Í tvíhliða prófinu skal ávalt velja úrtakið með hærri dreifni sem úrtak M og úrtakið með lægri dreifni sem úrtak m.

Framkvæmd tilgátuprófa

Framkvæmd tilgátuprófa

1 Ákveða hvaða tilgátupróf er viðeigandi fyrir gögnin okkar.

2 Ákveða hæstu ásættanlegu villulíkur.

3 Setja fram núlltilgátu og ákveða um leið áttun prófsins (einhliða/tvíhliða).

4 Reikna prófstærðina sem svarar til tilgátuprófsins.

5a Kanna hvort prófstærðin falli á höfnunarsvæði tilgátuprófsins.

5b Kanna p-gildi tilgátuprófsins.

6 Draga ályktun.