Het toetsen van hypothesen

Sandra Van Aert

10 november 2011

Tweezijdige hypothesetoets

- algemeen:
  $H_0 : \mu = \mu_0 \qquad H_a : \mu \neq \mu_0$
- voorbeeld:
  $H_0 : \mu = 34 \text{ cl} \qquad H_a : \mu \neq 34 \text{ cl}$
- H_0 verwerpen bij $\overline{x} \gg 34$ cl of $\overline{x} \ll 34$ cl
- kritieke waarden:

$$c_L = \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$c_U = \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Beslissingsregels

- benadering 1:
  H_0 verwerpen als $\overline{x} < c_L$ of $\overline{x} > c_U$
  H_0 aanvaarden als $c_L \leq \overline{x} \leq c_U$

- benadering 2:

  toetsingsgrootheid $z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$

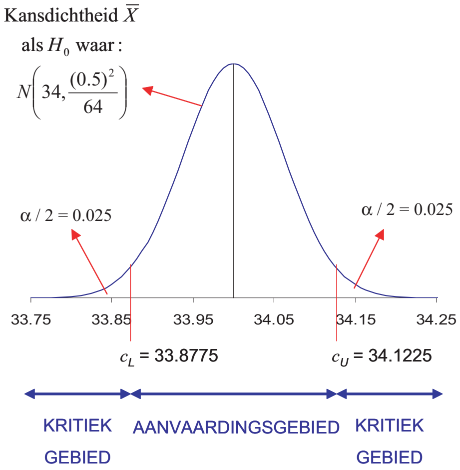
  H_0 verwerpen als $z < -z_{\alpha/2}$ of $z > z_{\alpha/2}$

  H_0 aanvaarden als $-z_{\alpha/2} \leq z \leq z_{\alpha/2}$

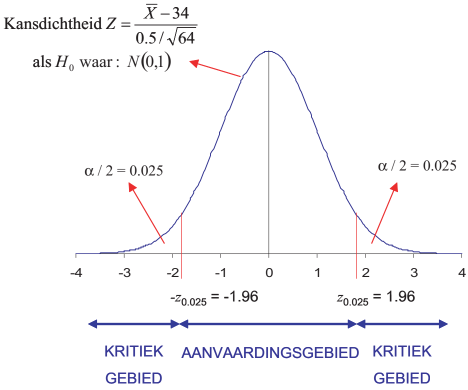
Voorbeeld

- gegeven: $n = 64$
  $\overline{x} = 33.89$ cl
  $\sigma = 0.5$ cl
- $H_0 : \mu = 34$ cl
  $H_a : \mu \neq 34$ cl
- significantieniveau $\alpha = 5\%$

Voorbeeld: benadering 1



Voorbeeld: benadering 2



p-waarde bij tweezijdige toets

- $\overline{x} < \mu_0 : p = 2P(\overline{X} < \overline{x} \mid \mu = \mu_0)$
  $\overline{x} > \mu_0 : p = 2P(\overline{X} > \overline{x} \mid \mu = \mu_0)$

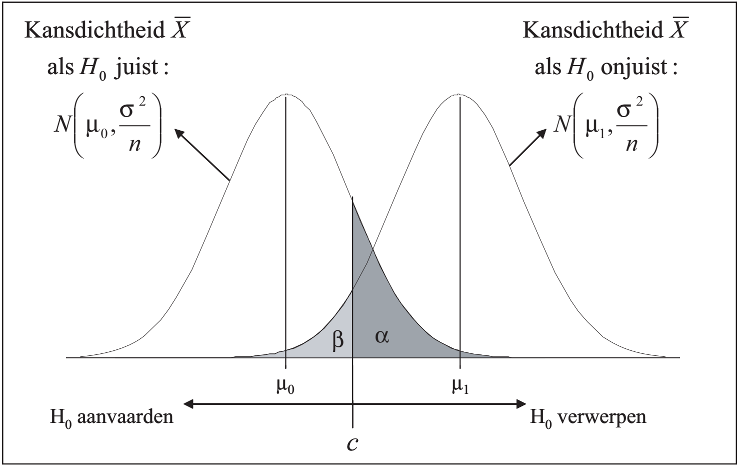
- komt op hetzelfde neer als
  $p = 2P(Z > |z| \mid \mu = \mu_0)$

- beslissingsregel (benadering 3):
  H_0 verwerpen als $p < \alpha$
  H_0 aanvaarden als $p \geq \alpha$

Kans op type II fout bij rechtseenzijdige toets

$$
\begin{aligned}
\beta &= P(H_0 \text{ aanvaarden} \mid H_0 \text{ is onjuist}) \\
&= P(\overline{X} < c \mid H_0 \text{ is onjuist}) \\
&= P(\overline{X} < c \mid \mu = \mu_1) \qquad \text{waarbij } \mu_1 > \mu_0 \\
&= P\left(\frac{\overline{X} - \mu_1}{\sigma/\sqrt{n}} < \frac{c - \mu_1}{\sigma/\sqrt{n}} \mid \mu = \mu_1\right) \\
&= P\left(Z < \frac{c - \mu_1}{\sigma/\sqrt{n}}\right)
\end{aligned}
$$

Type II fout: grafisch



Voorbeeld

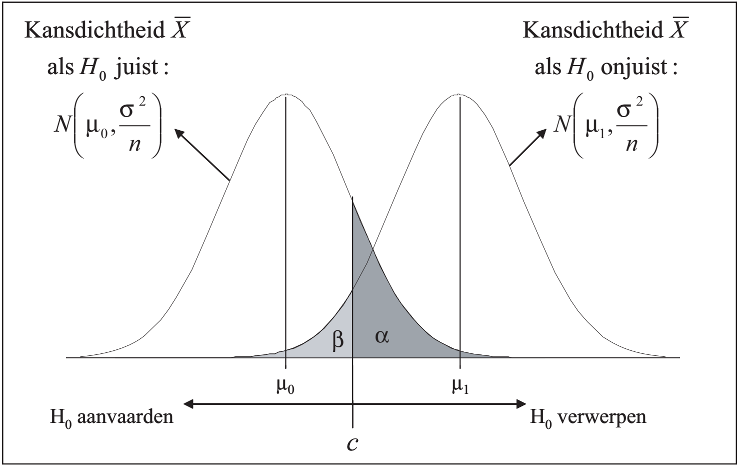
- stel H_0 is fout
  meer bepaald: $\mu = -0.535$

$$\begin{aligned}
\beta &= P(\overline{X} < -0.5391 \mid \mu = -0.535) \\
&= P(Z < \frac{-0.5391 - (-0.535)}{0.008/\sqrt{5}}) \\
&= P(Z < -1.146) \\
&= 0.1259
\end{aligned}$$

$\Rightarrow$ onderscheidingsvermogen van deze toets

$$1 - \beta = 1 - 0.1259 = 0.8741$$

Type II fout: grafisch



Beïnvloedende factoren

- α daalt $\Rightarrow$ β stijgt
- n stijgt $\Rightarrow$ β daalt
- σ^2 daalt $\Rightarrow$ β daalt
- μ_1 stijgt $\Rightarrow$ β daalt

Hypothesetoetsen voor één populatie

Sandra Van Aert

10 november 2011

Populatiegemiddelde μ - Normaal verdeelde populatie

- σ niet gekend: $$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \rightarrow \underbrace{\frac{\overline{X}-\mu_0}{S/\sqrt{n}}}_{t\text{-verdeeld met } n-1 \text{ vrijheidsgraden}}$$
- gevolg:

$$t_{\alpha,n-1} \text{ in plaats van } z_\alpha$$
$$t_{\alpha/2,n-1} \text{ in plaats van } z_{\alpha/2}$$

Linkseenzijdige t-toets

- beslissingsregel volgens benadering 1:

$$\frac{c-\mu_0}{s/\sqrt{n}} = -t_{\alpha;n-1} \Leftrightarrow c = \mu_0 - t_{\alpha;n-1}\frac{s}{\sqrt{n}}$$

H_0 verwerpen als $\overline{x} < \mu_0 - t_{\alpha;n-1}\dfrac{s}{\sqrt{n}}$

H_0 aanvaarden als $\overline{x} \geq \mu_0 - t_{\alpha;n-1}\dfrac{s}{\sqrt{n}}$

Linkseenzijdige t-toets (vervolg)

- beslissingsregel volgens benadering 2:

  toetsingsgrootheid $t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$

  H_0 verwerpen als $t < -t_{\alpha;n-1}$

  H_0 aanvaarden als $t \geq -t_{\alpha;n-1}$

Linkseenzijdige t-toets (vervolg)

- beslissingsregel volgens benadering 3:

  $p = P(t_{n-1} < t)$ met t = toetsingsgrootheid

  als $p < \alpha$, verwerp nulhypothese

  als $p \geq \alpha$, aanvaard nulhypothese

Populatiegemiddelde μ - Niet normaal verdeelde populatie

- kleine steekproeven ($n < 30$)
  → geen hypothesetoets voor μ
  → niet-parametrische toets voor de
  $\begin{cases} \text{mediaan} \\ \text{centrale locatie/ligging} \end{cases}$ (zie § 14.4)

- grote steekproeven ($n \geq 30$)
  → σ^2 bekend: gebruik $N(0,1)$-verdeling
  → σ^2 onbekend: gebruik t-verdeling
  met $n-1$ vrijheidsgraden

Voorbeeld

- wet 1 januari 1980:
  $\rightarrow n = 50$
  $\rightarrow$ beschuldig leverancier indien

$$\overline{x} < Q_n - 0.379s$$

  Q_n = nettogewicht vermeld op verpakking
  s = steekproefstandaarddeviatie

- getoetste hypothesen:

$$H_0 : \mu = Q_n$$
$$H_a : \mu < Q_n$$

Vervolg voorbeeld

- nulhypothese verwerpen indien

$$\overline{x} < \mu_0 - t_{\alpha,n-1}\frac{s}{\sqrt{n}}$$

$$\overline{x} < Q_n - t_{\alpha,49}\frac{s}{\sqrt{50}}$$

- is identiek aan wet als

$$0.379 = \frac{t_{\alpha,49}}{\sqrt{50}} \Leftrightarrow t_{\alpha,49} = 0.379\sqrt{50} = 2.6799$$

- $\alpha = 0.005$
  $\rightarrow$ kleine kans op type I fout
  $\rightarrow$ overheid wil niet al te veel risico nemen om iemand valselijk te beschuldigen

Populatievariantie σ^2 (rechtseenzijdig)

- $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 > \sigma_0^2 \end{cases}$
- beslissingsregel:

  verwerp H_0 indien $s^2 > \dfrac{\sigma_0^2}{n-1}\chi^2_{\alpha;n-1}$

  aanvaard H_0 indien $s^2 \leq \dfrac{\sigma_0^2}{n-1}\chi^2_{\alpha;n-1}$

  of

  verwerp H_0 indien $p < \alpha$

  met $p = P(\chi^2_{n-1} > \dfrac{(n-1)s^2}{\sigma_0^2})$

  aanvaard H_0 indien $p \geq \alpha$

Populatievariantie σ^2 (tweezijdig)

- $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 \neq \sigma_0^2 \end{cases}$
- beslissingsregel:
  verwerp H_0 indien

$$s^2 < \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha/2;n-1} \qquad \text{of} \qquad s^2 > \frac{\sigma_0^2}{n-1}\chi^2_{\alpha/2;n-1}$$

aanvaard H_0 indien

$$\frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha/2;n-1} \leq s^2 \leq \frac{\sigma_0^2}{n-1}\chi^2_{\alpha/2;n-1}$$

p-waarde bij tweezijdige toets σ^2

- χ^2-verdeling is niet symmetrisch
  dus onderscheid tussen $s^2 > \sigma_0^2$ en $s^2 < \sigma_0^2$
- $s^2 > \sigma_0^2$:
  $$p = 2P(\chi^2_{n-1} > \frac{(n-1)s^2}{\sigma_0^2})$$
- $s^2 < \sigma_0^2$:
  $$p = 2P(\chi^2_{n-1} < \frac{(n-1)s^2}{\sigma_0^2})$$

Voorbeeld

- $n = 20$
  $s = 0.4294$
  $\alpha = 0.05$
- getoetste hypothesen:

$$H_0 : \sigma^2 = 0.1$$
$$H_a : \sigma^2 \neq 0.1$$

- berekende toetsingsgrootheid:

$$\frac{(n-1)s^2}{\sigma_0^2} = \frac{(20-1)(0.4294)^2}{0.1} = 35.033$$

Vervolg voorbeeld

- berekende toetsingsgrootheid $x = 35.033$
- kritieke waarden:

$$\chi^2_{1-\alpha/2;n-1} = \chi^2_{0.975;19} = 8.9065 (= \text{qchisq}(0.025, 19))$$
$$\chi^2_{\alpha/2;n-1} = \chi^2_{0.025;19} = 32.8523 (= \text{qchisq}(0.975, 19))$$

- besluit:
  berekende toetsingsgrootheid ligt niet in aanvaardingsgebied
  $\Rightarrow H_0$ verwerpen
  $\Rightarrow \sigma^2$ is significant verschillend van 0.1

Populatieproportie

- $X = \#$ successen in steekproef van grootte n
- $X \sim \text{binomiaal}(n, \pi)$
- voor grote n: $X \sim N(n\pi, n\pi(1-\pi))$
- $\hat{P} = \dfrac{X}{n} =$ proportie successen

$$\hat{P} \sim N(\pi, \frac{\pi(1-\pi)}{n})$$

- voorwaarde in de praktijk:

$$n\hat{p} > 5 \text{ en } n(1-\hat{p}) > 5$$

Rechtseenzijdige toets voor π

- $\begin{cases} H_0 : \pi = \pi_0 \\ H_a : \pi > \pi_0 \end{cases}$
- benadering 1:

  verwerp H_0 indien $\hat{p} > \pi_0 + z_\alpha \sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}$

  aanvaard H_0 indien $\hat{p} \leq \pi_0 + z_\alpha \sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}$
- p-waarde $= P(Z > \dfrac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}})$

  verwerp H_0 indien $p < \alpha$
  aanvaard H_0 indien $p \geq \alpha$

Tekentoets voor populatiemediaan

- niet-parametrische of verdelingsvrije hypothesetoets
  → hangen niet af van veronderstellingen omtrent kansverdeling/kansdichtheid populatie
  → werken altijd
  → klassieke toetsen zijn beter
  als veronderstellingen juist zijn

- $$H_0 : \text{Me} = \text{Me}_0 \text{ versus } \begin{cases} H_a : \text{Me} > \text{Me}_0 \\ H_a : \text{Me} < \text{Me}_0 \\ H_a : \text{Me} \neq \text{Me}_0 \end{cases}$$

Werking tekentoets

- betekenis populatiemediaan:

$$50\% \text{ van populatie } > \text{ populatiemediaan}$$

$$50\% \text{ van populatie } < \text{ populatiemediaan}$$

- als H_0 juist:

$$50\% \text{ populatie } > \mathrm{Me}_0 \text{ en } 50\% \text{ populatie } < \mathrm{Me}_0$$

- als H_0 juist:

$$\text{ongeveer } 50\% \text{ steekproef } > \mathrm{Me}_0$$

$$\text{ongeveer } 50\% \text{ steekproef } < \mathrm{Me}_0$$

Werking tekentoets

- tellen hoeveel waarnemingen groter of kleiner zijn dan Me_0
- vergelijken met $\dfrac{n}{2}$
- hoe?

  aantal waarnemingen > of < Me_0

  = binomiaal verdeeld met parameters

$$\begin{cases} n \\ \pi = 0.5 \quad \text{(als } H_0 \text{ waar is)} \end{cases}$$

Rechtseenzijdige tekentoets

- $H_a : \mathrm{Me} > \mathrm{Me}_0$
- stel s = aantal waarnemingen $> \mathrm{Me}_0$
- p-waarde:

$$p = P(X \geq s) \text{ met } X \sim \text{ binomiaal}(n, \frac{1}{2})$$

- beslissingsregel:
  verwerp H_0 als $p < \alpha$
  aanvaard H_0 als $p \geq \alpha$
- $P(X \geq s)$
  $\rightarrow$ R: `"=1-pbinom(`$s-1$`,`n`,0.5)"`
  $\rightarrow$ Matlab: `"=1-binocdf(`$s-1$`,`n`,0.5)"`

Linkseenzijdige tekentoets

- $H_a : \mathrm{Me} < \mathrm{Me}_0$
- stel s = aantal waarnemingen $< \mathrm{Me}_0$
- p-waarde:

$$p = P(X \geq s) \text{ met } X \sim \text{ binomiaal}(n, \frac{1}{2})$$

- beslissingsregel:
  verwerp H_0 als $p < \alpha$
  aanvaard H_0 als $p \geq \alpha$

Tweezijdige tekentoets

- $H_a : \text{Me} \neq \text{Me}_0$
- stel

$$s = \text{ grootste van } \begin{cases} \#\text{ waarnemingen } > \text{Me}_0 \\ \#\text{ waarnemingen } < \text{Me}_0 \end{cases}$$

- p-waarde:

$$p = 2P(X \geq s) \text{ met } X \sim \text{ binomiaal}(n, \frac{1}{2})$$

- beslissingsregel:

  verwerp H_0 als $p < \alpha$
  aanvaard H_0 als $p \geq \alpha$

Voorbeeld

- fabrikant CD-spelers: mediaan = 5250 uur
- test 20 spelers van concurrent: 14 spelers hebben levensduur > 5250 uur
- H_0 : Me = 5250 versus H_a : Me > 5250 ($\alpha = 0.05$)
- oplossing:
  $\rightarrow s = 14$
  $\rightarrow p = P(X \geq s)$
  $= P(X \geq 14)$ met $X \sim$ binomiaal $(20, \frac{1}{2})$
  $= 0.0577$

  $0.0577 > \alpha$ dus aanvaard H_0 (randgeval)

Tekentoets

- vereist ordinale gegevens
- wordt vaak gebruikt voor kwantitatieve, niet-normaal verdeelde gegevens
- beslissingsboom voor hypothesetoets centrale ligging