

Statistiek voor A.I.

College 14

Dinsdag 30 Oktober

2 Deductieve statistiek

Orthodoxe statistiek

Voorbeeld NU.nl 26 Oktober 2012: Helft broodjes döner kebab vol bacteriën.

DEN HAAG - Mensen met een verminderde weerstand, 65-plussers en zwangeren kunnen beter geen broodje döner kebab eten. Dat advies geeft de Consumentenbond vrijdag na onderzoek.



Van de 49 geteste broodjes bleek een meerderheid onder de bacteriën te zitten. Op vijf broodjes werd zelfs de zogenoemde 'poepbacterie' gevonden. De broodjes werden gekocht bij tien willekeurige verkooppunten in Groningen, Utrecht, Amsterdam, Rotterdam en Breda.

2 Deductieve statistiek

Vandaag :

- Kracht
- Grootte steekproef

Kracht

Definitie De *power* (*kracht*) van een hypothese toets is $(1 - \beta)$, waarbij β de kans op een fout van de tweede soort is.

Over het algemeen wil de onderzoeker dat de power van een toets groot is.

De *grootte van het effect* (*ES*) is het verschil tussen de waarde van de nulhypothese en die waarmee gewerkt wordt voor de alternatieve hypothese.

Kracht

Voorbeeld Van een medicijn dat de lichaamstemperatuur verhoogt wil men toetsen of de gemiddelde temperatuur van de gebruikers significant meer dan 38 graden is. μ is de gemiddelde temperatuur van mensen die het medicijn slikken. $\alpha = 0.05$, $\sigma = 1.5$ en de grootte van de steekproef is 25.

$$H_0 : \mu \leq 38 \quad H_a : \mu > 38.$$

$$x_{kw} = 38 + 1.645 \times \frac{1.5}{\sqrt{25}} = 38 + 1.645 \times 0.3 = 38.49.$$

Men is bezorgd over een fout van de tweede soort als het gemiddelde 38.3 of meer is.

$$H_0 : \mu \leq 38 \quad H_a : \mu = 38.3.$$

De fout van de eerste soort is

$$P(\bar{X} \geq x_{kw} \mid \mu = 38) = 0.05.$$

De fout van de tweede soort is

$$P(\bar{X} < x_{kw} \mid \mu = 38.3) = P_s\left(\frac{\bar{X} - 38.3}{0.3} < \frac{38.49 - 38.3}{0.3}\right) = 0.74.$$

Dus de kracht van de toets is $1 - 0.74 = 0.26$.

Voorbeeld Van een medicijn dat de lichaamstemperatuur verhoogt wil men toetsen of de gemiddelde temperatuur van de gebruikers significant afwijkt van 38 graden. μ is de gemiddelde temperatuur van mensen die het medicijn slikken. $\alpha = 0.05$, $\sigma = 1.5$ en de grootte van de steekproef is 25.

$$H_0 : \mu = 38 \quad H_a : \mu \neq 38.$$

$$x_{kw} = 38 + 1.96 \times \frac{1.5}{\sqrt{25}} = 38 + 1.96 \times 0.3 = 38.59.$$

Men is bezorgd over een fout van de tweede soort als het gemiddelde 38.3 of meer is.

$$H_0 : \mu = 38 \quad H_a : \mu = 38.3.$$

De fout van de eerste soort is

$$P(\bar{X} \geq x_{kw} \mid \mu = 38) = 0.05.$$

De fout van de tweede soort is

$$P(\bar{X} < x_{kw} \mid \mu = 38.3) = P_s\left(\frac{\bar{X} - 38.3}{0.3} < \frac{38.59 - 38.3}{0.3}\right) = 0.83.$$

Dus de power van de toets is $1 - 0.83 = 0.17$.

Kracht: factoren die de kracht van een toets beïnvloeden

Wanneer alle andere factoren onveranderd blijven geldt:

- Het vergroten van het significantieniveau maakt een toets krachtiger.
- Het vergroten van de steekproef maakt een toets krachtiger.
- Het vergroten van het effect maakt een toets krachtiger.
- Eenzijdige toetsen zijn krachtiger dan tweezijdige toetsen.

Grootte van een steekproef

Grootte van een steekproef

Stelling Gegeven een hypothese toets met significantieniveau α , kracht $(1 - \beta)$, effectgrootte ES en variantie σ^2 , is de juiste steekproefgrootte:

$$n = \frac{\sigma^2(z_\beta - z_\alpha)^2}{(ES)^2} \text{ (eenzijdig)} \quad n = \frac{\sigma^2(z_\beta - z_{\frac{\alpha}{2}})^2}{(ES)^2} \text{ (tweezijdig)}.$$

Hierbij is z_α de standaard kritische waarde bij significantieniveau α , en z_β de corresponderende standaard kritische waarde in de verdeling van H_a bij de gegeven grootte van het effect.

Bewijs voor een eenzijdige toets. Stel dat H_0 van de vorm $\mu \leq a$, $\mu \geq a$ of $\mu = a$ is, en $H_a : \mu = b$. Dan $ES = (b - a)$ en er geldt

$$z_\alpha = \frac{x_{kw} - a}{\frac{\sigma}{\sqrt{n}}} \quad z_\beta = \frac{x_{kw} - b}{\frac{\sigma}{\sqrt{n}}}$$

$$z_\beta - z_\alpha = \frac{a - b}{\frac{\sigma}{\sqrt{n}}} = \frac{(a - b)\sqrt{n}}{\sigma}$$

$$\sqrt{n} = \frac{\sigma(z_\beta - z_\alpha)}{-(b - a)}.$$



Stelling Gegeven een hypothese toets met twee gemiddelden, significantieniveau α , kracht $(1 - \beta)$, effectgrootte ES en gelijke variantie σ^2 , is de juiste steekproefgrootte by twee steekproeven van grootte n :

$$n = \frac{2\sigma^2(z_\beta - z_\alpha)^2}{(ES)^2} \text{ (eenzijdig)} \quad n = \frac{2\sigma^2(z_\beta - z_{\frac{\alpha}{2}})^2}{(ES)^2} \text{ (tweezijdig)}.$$

Hierbij is z_α de standaard kritische waarde bij significantieniveau α , en z_β de corresponderende standaard kritische waarde in de verdeling van H_a bij de gegeven grootte van het effect.

Voorbeeld Broodjes döner kebab.

$$H_0 : \mu \leq 0.01 \quad H_a : \mu > 0.01$$

Gewenste $\alpha = 0.1$ en $\beta = 0.01$. Dus $z_\alpha = 1.282$ en $z_\beta = -2.326$. Stel de variantie van de populatie is $\sigma^2 = 20$. Bij effectgrootte 5 moet de grootte van de steekproef minstens zijn:

$$n = \frac{\sigma^2(z_\beta - z_\alpha)^2}{ES^2} = \frac{20(-2.326 - 1.282)^2}{5^2} = 10.414.$$

Dus een steekproef van 49 broodjes volstaat.

Grootte van een steekproef

Voorbeeld Een ondernemer die een krant wil oprichten laat onderzoeken hoeveel mensen gemiddeld dagelijks een krant lezen. Er wordt een hypothese toets gedaan en de ondernemer wil: $\alpha = 0.1$ en $\beta = 0.01$. De hypothesen zijn:

$$H_0 : \mu \geq 0.7 \quad H_a : \mu < 0.7.$$

Als grootte van het effect wordt 0.3 genomen en de standaardafwijking van de parameter in de populatie is 0.2.

Hoe groot moet de grootte van de steekproef dan zijn?

Omdat $P_s(z \leq -1.28) = \alpha$ en $P_s(z \geq 2.326) = \beta$ geldt

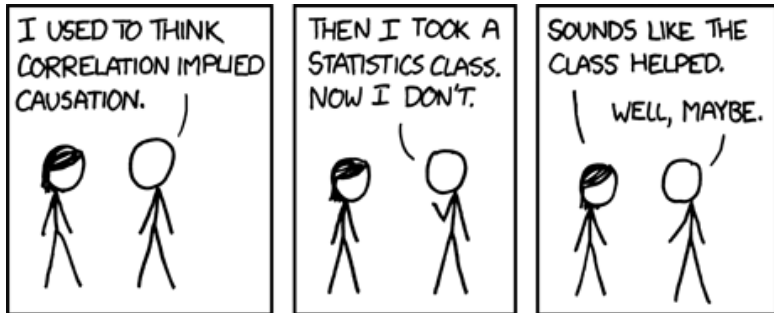
$$z_\alpha = -1.28 \quad z_\beta = 2.326.$$

Daaruit volgt dat

$$n = \frac{(0.4)^2(2.326 + 1.28)^2}{(0.3)^2} = 23.117.$$

Een steekproef ter grootte 24 volstaat dus.

Context



Finis