

Statistiek voor A.I.

College 12

Dinsdag 23 Oktober

2 Deductieve statistiek

Orthodoxe statistiek

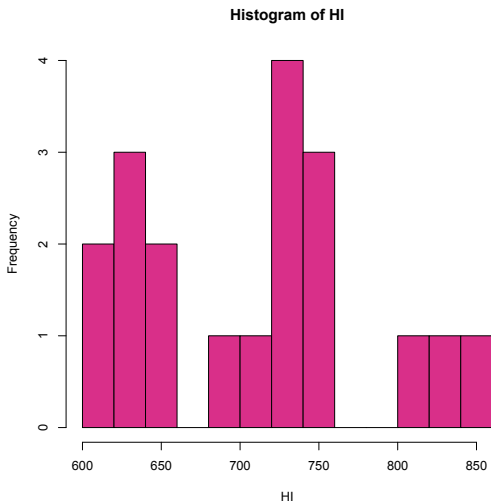
The Six Most Confusing Words in Statistics

failed to reject the null hypothesis

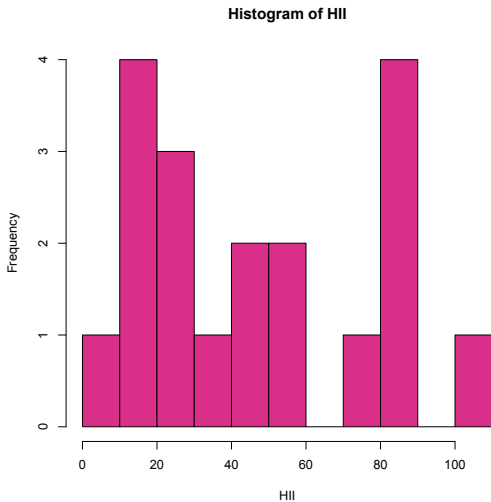
TRIPLE NEGATIVE!
my brain hurts!



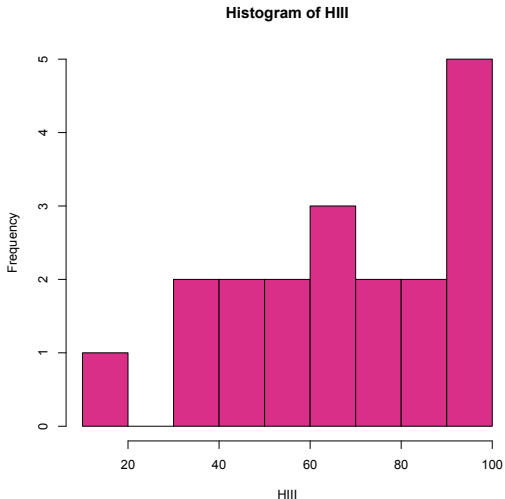
Tijd waarop je opstaat (uu:mm wordt weergegeven als uumm).



Reistijd van huis naar college in minuten.



Hoe vaak je op tijd bent (in %).

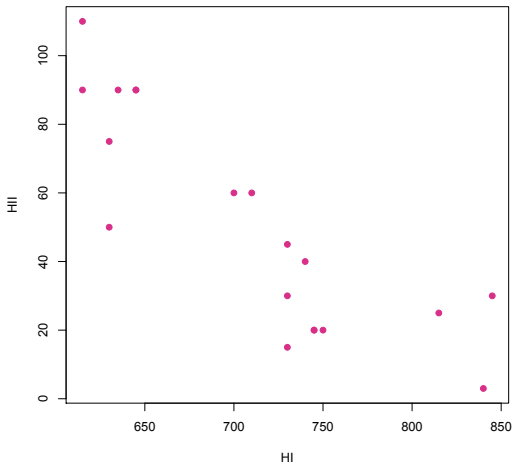


Geen noemenswaardige correlatie tussen reistijd en op tijd komen, noch tussen uur van opstaan en op tijd komen.

Jullie - onderzoek Wivine

Horizontaal: Tijd van opstaan.

Verticaal: Reistijd.



Correlatiecoëfficiënt: -0.85

2 Deductieve statistiek

Vandaag :

- Hypothese toetsen

Hypothese toetsen

Hypothese toetsen: criterium voor verwerpen H_0

Definitie De hypothesen die we gaan beschouwen zijn meestal van de vorm:

<i>tweezijdig</i>	<i>rechtszijdig</i>	<i>linkszijdig</i>
$H_0 : \mu = a$	$H_0 : \mu \leq a$	$H_0 : \mu \geq a$
$H_a : \mu \neq a$	$H_a : \mu > a$	$H_a : \mu < a$

- H_0 is waar en wordt aangenomen: correcte beslissing.
- H_0 is waar en wordt verworpen: fout van de 1^e soort.
- H_a is waar en wordt aangenomen: correcte beslissing.
- H_a is waar en wordt verworpen: fout van de 2^e soort.

Criterium voor het verwerpen van de nulhypothese:

Voor een zekere constante x_k (of x_k^l en x_k^r) wordt op grond van een steekproef met gemiddelde $\bar{X}$

H_0 verworpen als $\bar{X} \geq x_k^r$ of $\bar{X} \leq x_k^l$	(tweezijdige toets)
H_0 verworpen als $\bar{X} \geq x_k$	(rechtszijdige toets)
H_0 verworpen als $\bar{X} \leq x_k$	(linkszijdige toets).

Hypothese toetsen: significantieniveau

Definitie Het *significantieniveau* is de kans op een fout van de 1^e soort en wordt aangeduid met α .

Het criterium voor het verwerpen van H_0 wordt zo gekozen dat de kans op een fout van de 1^e soort het significantieniveau α is:

$$P(\bar{X} \geq x_k^r \text{ of } \bar{X} \leq x_k^l \mid H_0) = \alpha \quad (\text{tweezijdige toets})$$

$$P(\bar{X} \geq x_k \mid H_0) = \alpha \quad (\text{rechtszijdige toets})$$

$$P(\bar{X} \leq x_k \mid H_0) = \alpha \quad (\text{linkszijdige toets}).$$

Meestal wordt het “ $\mid H_0$ ” weggelaten.

Hypothese toetsen: kritisch gebied

Definitie De *kritische waarde* is de waarde x_k waarvoor

$$\begin{aligned}P(\bar{X} \geq x_k^r \text{ of } \bar{X} \leq x_k^l \mid \mu = a) &= \alpha && \text{(tweezijdige toets met } H_0 : \mu = a) \\P(\bar{X} \geq x_k \mid \mu = a) &= \alpha && \text{(rechtszijdige toets met } H_0 : \mu \leq a) \\P(\bar{X} \leq x_k \mid \mu = a) &= \alpha && \text{(linkszijdige toets met } H_0 : \mu \geq a).\end{aligned}$$

Het *kritische gebied* (*region of rejection*) zijn die waardes x die liggen in de intervallen

$$\begin{aligned}(-\infty, x_k^l] \text{ of } [x_k^r, \infty) &&& \text{(tweezijdige toets)} \\[x_k, \infty) &&& \text{(rechtszijdige toets)} \\(-\infty, x_k] &&& \text{(linkszijdige toets)}.\end{aligned}$$

Als $\bar{X}$ in the kritisch gebied ligt, dan geldt dat onder de aanname H_0 de kans dat het steekproefgemiddelde gelijk aan $\bar{X}$ is, $\leq \alpha$ is.

Daarom wordt H_0 *verworpen* als $\bar{X}$ in het kritische gebied ligt. Als $\bar{X}$ niet in het kritische gebied ligt wordt H_0 *aangenomen*.

Hypothese toetsen: verdeling van het steekproefgemiddelde

Definitie Wanneer het significantieniveau eenmaal gekozen is, kan de kritische waarde berekend worden als de verdeling P van het steekproefgemiddelde $\bar{X}$ bekend is.

Als de standaardafwijking van de populatie bekend en gelijk aan σ is, dan wordt vaak aangenomen dat P de normale verdeling is met gemiddelde μ en standaardafwijking $\frac{\sigma}{\sqrt{n}}$, waarbij n de grootte van de steekproef is.

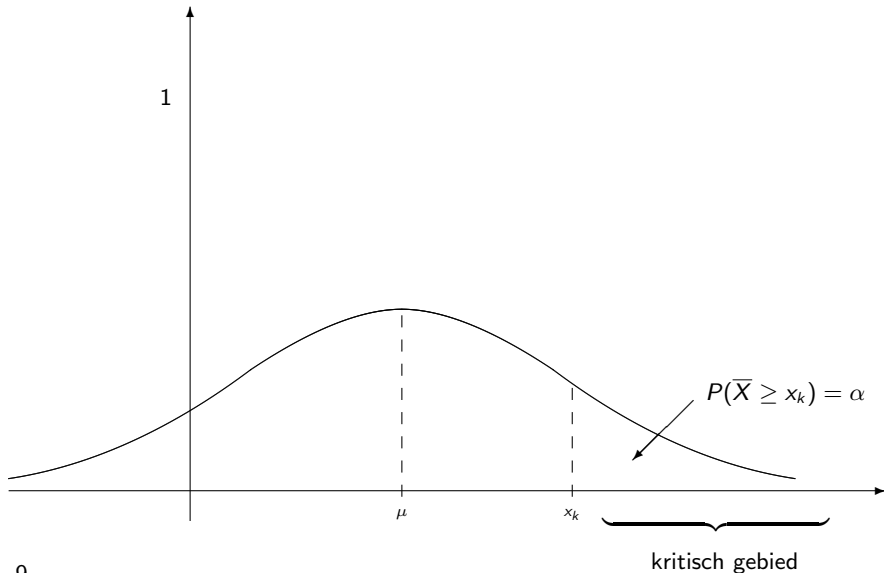
De Centrale Limietstelling wordt beschouwd als rechtvaardiging voor die keuze.

Als de standaardafwijking van de populatie niet bekend is, dan wordt vaak aangenomen dat P de t -verdeling P_t is bij $n - 1$ vrijheidsgraden, met gemiddelde μ en standaardafwijking $\frac{s}{\sqrt{n}}$, waarbij s de standaardafwijking van de steekproef is en n de grootte van de steekproef.

Deze keuze kan gerechtvaardigd worden op dezelfde wijze als voor de normale verdeling.

Hypothese toetsen: kritisch gebied

Bij een rechtszijdige toets:



Hypothese toetsen

Voorbeeld Het gemiddelde aantal patiënten dat met 200 mg. van medicijn M geneest van een zekere ziekte is 30%. Men wil testen of een dosis van 300 mg. significant meer patiënten geneest. Men dient aan een steekproef van 25 patiënten 300 mg. M toe, waarvan 40% geneest.

Laat μ het gemiddelde aantal patiënten in de populatie zijn dat bij 300 mg. M geneest. Als significantieniveau wordt 0.05 genomen. Er wordt aangenomen dat σ bekend en gelijk aan 4% is.

$$H_0 : \mu \leq 30\%$$

$$H_a : \mu > 30\%.$$

Fout van de 1^e soort: $\mu = 30\%$, maar H_0 wordt verworpen.

Voor de kritische waarde x_k moet gelden dat

$$P(\bar{X} \geq x_k \mid \mu = 30) = \alpha = 0.05.$$

Aangenomen wordt dat $\bar{X}$ normaal verdeeld is. Dus moet gelden voor x_k :

$$P(\bar{X} \geq x_k) = P_s\left(\frac{\bar{X} - 30}{\frac{\sigma}{\sqrt{25}}} \geq \frac{x_k - 30}{\frac{\sigma}{\sqrt{25}}}\right) = P_s\left(\frac{\bar{X} - 30}{0.8} \geq \frac{x_k - 30}{0.8}\right) = 0.05.$$

$P_s(z \geq 1.645) = 0.05$. Dus $x_k = \mu + 1.645 \frac{\sigma}{\sqrt{25}} = 30 + (1.645)(0.8) = 31.316\%$.

$\bar{X} = 40\% > 31.316\%$. $\bar{X}$ ligt in the kritische gebied, en H_0 wordt verworpen.

Conclusie: bij een dosis van 300 mg. genezen significant meer patiënten.

Hypothese toetsen

Voorbeeld De UU wil kijken of het een anti-alcohol campagne moet beginnen en wil daartoe een toets laten uitvoeren waarbij wordt bekeken of het aantal glazen alcohol dat studenten per week drinken significant groter is dan 5.

Hypothesen $H_0 : \mu \leq 5$ en $H_a : \mu > 5$.

Er wordt een steekproef ter grootte 22 genomen door Peter, Huub en Rens: $\bar{X} = 6.9$ en $s = 6.7$. Op grond daarvan wordt bepaald welke hypothese verworpen moet worden bij significantieniveau 0.01.

Voor de kritische waarde x_k moet gelden:

$$P(\bar{X} \geq x_k \mid \mu = 5) = P_t\left(\frac{\bar{X} - 5}{\frac{6.7}{\sqrt{22}}} \geq \frac{x_k - 5}{\frac{6.7}{\sqrt{22}}}\right) = P_t\left(\frac{\bar{X} - 5}{1.4} \geq \frac{x_k - 5}{1.4}\right) = 0.01.$$

Omdat de variantie van de populatie niet bekend is wordt met de t -verdeling P_t getoetst bij 21 vrijheidsgraden, met gemiddelde 6.9 en de standaardafwijking van het gemiddelde 1.4. Uit C.3: $P_t(z \geq 2.518) = 0.01$.

Daarmee is de kritische waarde $x_k = 5 + (2.518)(1.4) = 8.5$ en het kritische gebied $[8.5, \infty)$.

$\bar{X}$ ligt niet in het kritische gebied: H_0 wordt niet verworpen. En daarom onderneemt de UU geen actie.

Hypothese toetsen: p -waarde

Definitie De p -waarde van een statistiek (mogelijke waarde) c is de kans dat onder de nulhypothese het steekproefgemiddelde een waarde minstens zo extreem als c heeft.

	<i>tweezijdig</i>	<i>rechtszijdig</i>	<i>linkszijdig</i>
p -waarde	$2P(\bar{X} \geq c \mid H_0)$	$P(\bar{X} \geq c \mid H_0)$	$P(\bar{X} \leq c \mid H_0)$

Stelling Als bij een rechts- of linkszijdige toets de p -waarde van het gevonden steekproefgemiddelde kleiner is dan het significantieniveau, dan wordt de nulhypothese verworpen.

Bij een tweezijdige toets moet de p -waarde daarvoor kleiner dan de helft van het significantieniveau zijn.

Hypothese toetsen: binomiale verdeling

Voorbeeld In opdracht van de regering wordt onderzocht wat het percentage mensen is dat illegaal films downloadt. De regering zal in actie komen als dat percentage meer dan 30% is.

Een onderzoeksbureau voert een hypothesetoets uit met hypothesen (in %)

$H_0 : \mu \leq 30$ en $H_a : \mu > 30$.

In dit geval wordt aangenomen dat $\bar{X}$ een binomiale verdeling heeft ($X = 1$ als iemand illegaal films downloadt, $X = 0$ anders). Als significantieniveau wordt 0.02 genomen.

Er een steekproef ter grootte 1000 genomen uit de populatie Nederlanders en op grond daarvan wordt bepaald welke hypothese verworpen moet worden. Van die 1000 Nederlanders downloadt 37% illegaal films.

De kans op een fout van de eerste soort is in dit geval

$$P(\bar{X} \geq x_k \mid \mu = 30) = \sum_{i=x_k}^{1000} \binom{1000}{i} (0.3)^i (0.7)^{1000-i}.$$

Er geldt dat $P(\bar{X} \geq 330) = 0.0215$ en $P(\bar{X} \geq 331) = 0.018$, dus $x_k = 331$ en het kritische gebied is $\{331, 332, \dots, 1000\}$.

$\bar{X}$ ligt in het kritische gebied: H_0 wordt verworpen. De regering onderneemt actie.

Hypothese toetsen: kritiek

Bij toenemende n wordt het kritische gebied groter en neemt de kans dat de nulhypothese verworpen wordt toe. “Met een voldoende grote steekproef kan elke hypothese verworpen worden”.

Bij hypothese toetsen wordt de uitkomst van de toets niet gebruikt om de test aan te passen. Bij Bayesiaans leren wel.

Finis