



Toetsen van gemiddelden

Zie ook:
Cursusboek hoofdstuk 26

Inhoud

1. De t -toets voor één steekproef
2. De t -toets voor twee onafhankelijke steekproeven
3. De t -toets voor gepaarde steekproeven
4. De statistische modellen van de t -toetsen

Inhoud

1. De t -toets voor één steekproef
2. De t -toets voor twee onafhankelijke steekproeven
3. De t -toets voor gepaarde steekproeven
4. De statistische modellen van de t -toetsen

Is 37°C de normale lichaamstemperatuur?



- De waarde van 37°C komt uit klassiek onderzoek door Carl Reinhold August Wunderlich gepubliceerd in 1868.
- Klopt die waarde eigenlijk wel?

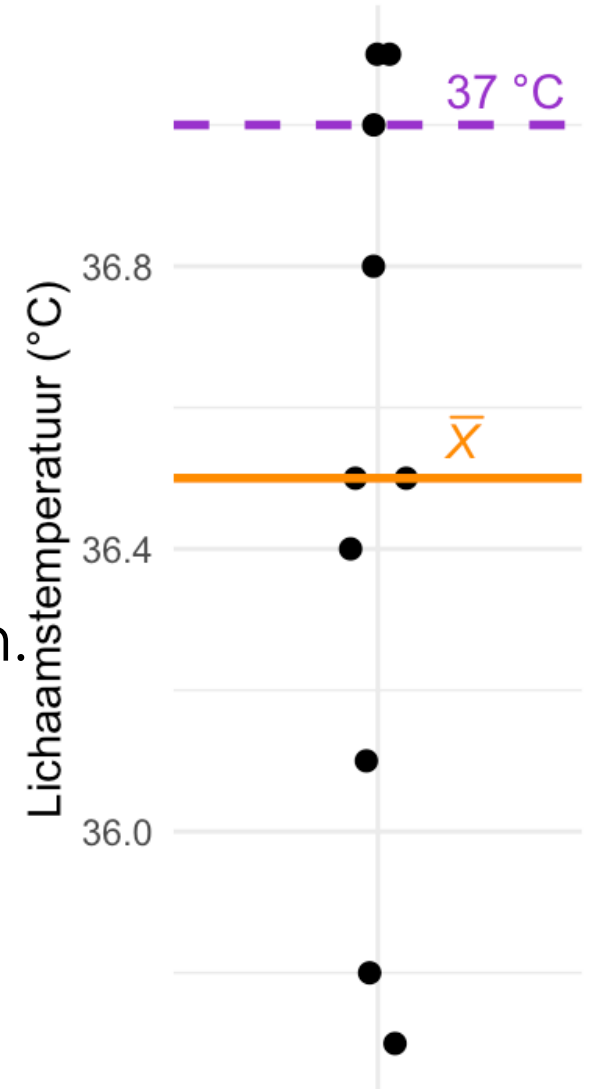
Onderzoek:

Meet de temperatuur van 10 personen.

Gemiddelde: 36,5°C.

Hoe sterk is dit bewijs tegen de waarde van Wunderlich?

Voor t-toets voor één steekproef uit!



Hypothesetoetsen in het algemeen

Structuur in 6 stappen

1. Kies een **statistisch model** dat past bij de casus.
2. Stel **hypotheses** H_0 en H_A op in termen van een parameter van het statistische model
3. Kies een **toetsingsgrootheid** en bepaal de **kansverdeling** van die grootheid “onder H_0 ”
4. **Bereken** de toetsingsgrootheid op basis van de waarnemingen.
5. Bereken de **P-waarde**.
6. Trek je **conclusies**: hoe sterk is het bewijs tegen H_0 ?

Stap 1: Het statistische model

We gebruiken weer het Normale Model:

1. De variabele is normaal verdeeld in de populatie.

Vaak redelijk omdat veel variabelen een klokvormige verdeling hebben.

2. Iedere waarde X_i van de steekproef is *onafhankelijk* uit deze normale verdeling getrokken

Redelijk als

- de populatie veel groter is dan de steekproef
- de steekproef ongeveer eenvoudig aselekt is.

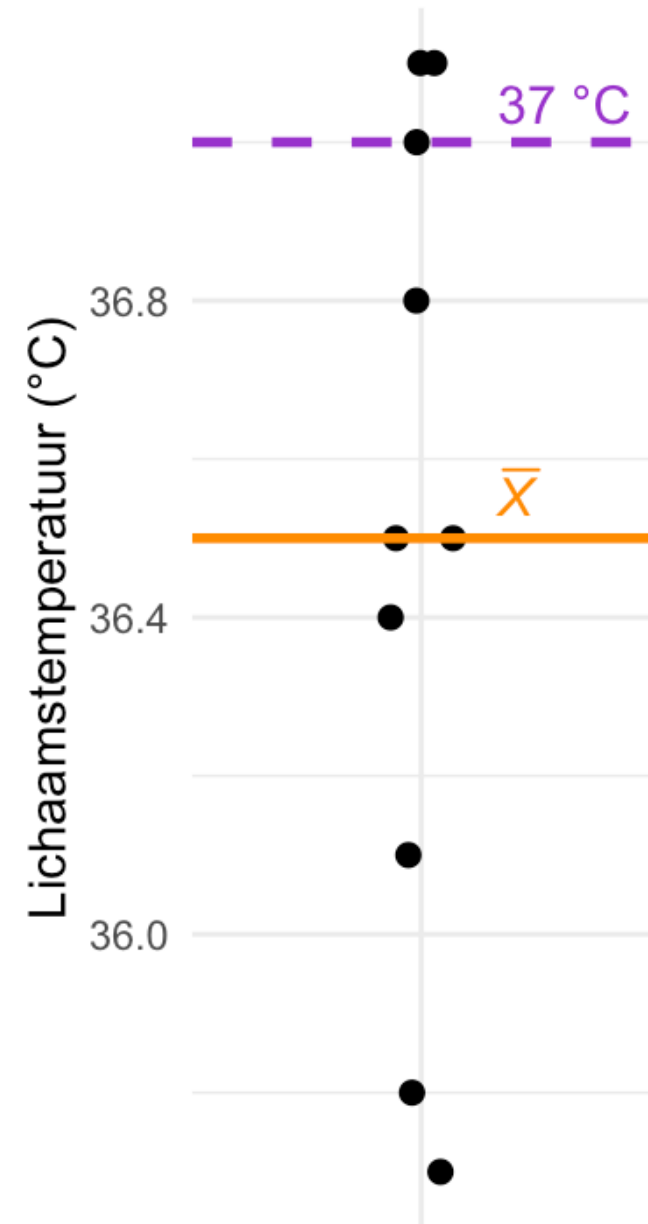
Stap 2: Stel de hypothesen op

Nulhypothese: In de populatie is lichaamstemperatuur gemiddeld 37°C.

Oftewel:

$$H_0: \mu_X = 37^\circ\text{C},$$

$$H_A: \mu_X \neq 37^\circ\text{C}.$$



Stap 3: De toetsingsgrootte en zijn verdeling

We kiezen als toetsingsgrootte:

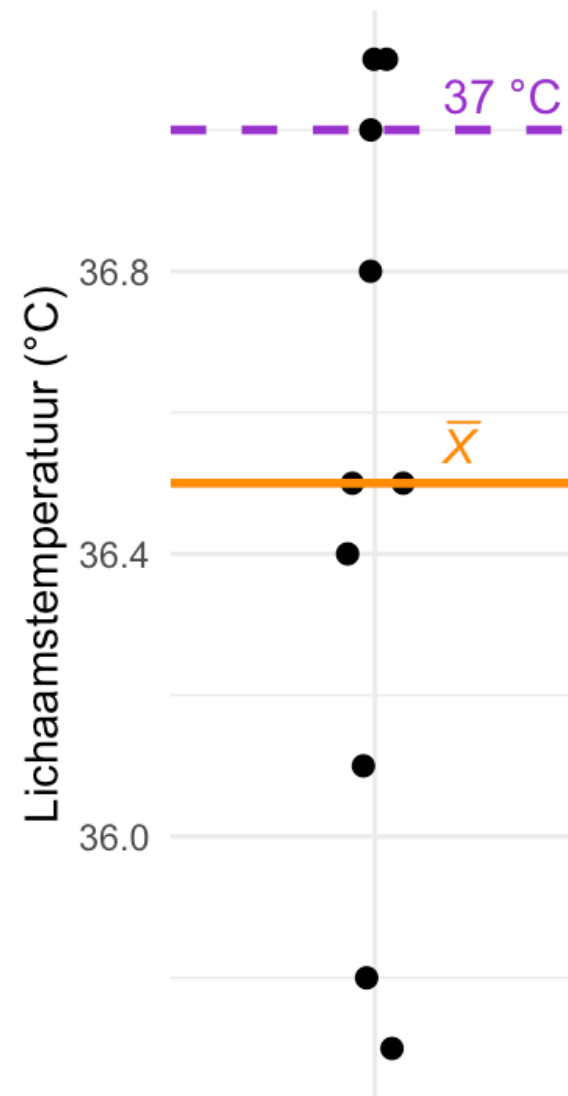
$$t = \frac{\bar{X} - \mu_0}{SE_{\bar{X}}}$$

37 °C

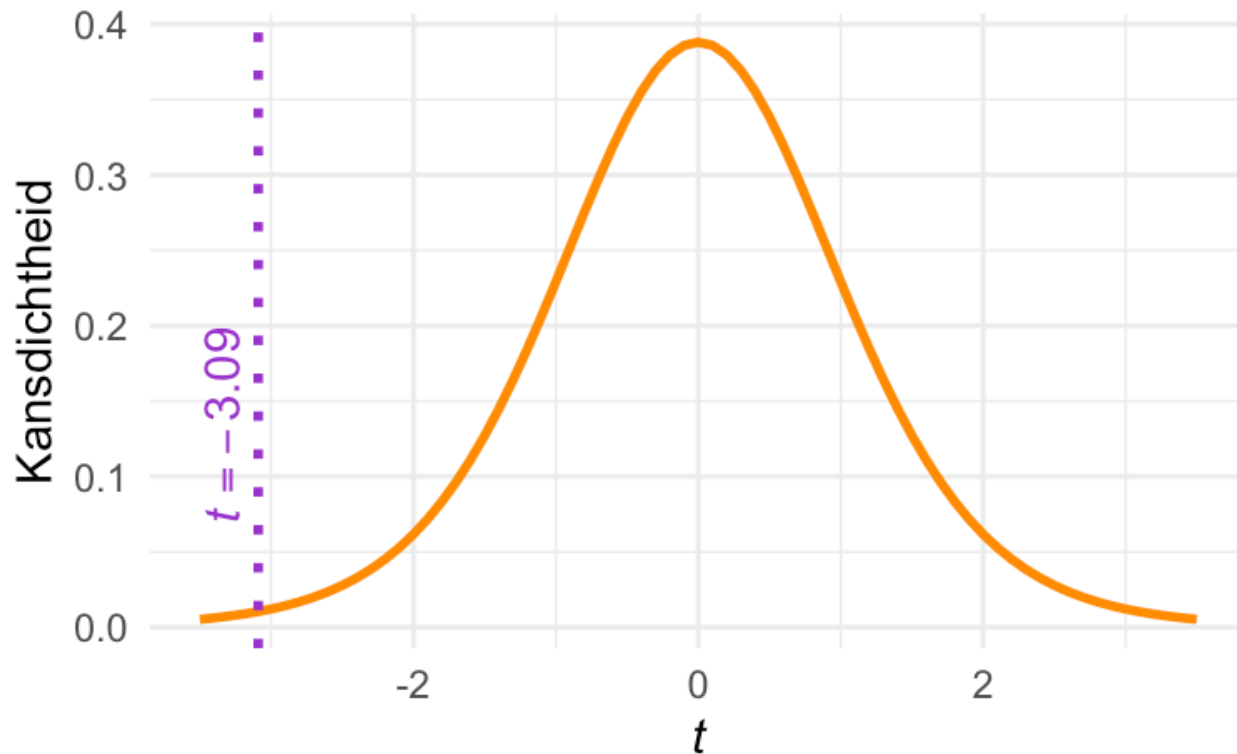


Verdeling:

t -verdeling met $df = n - 1$.



Stap 4: Bereken de toetsingsgrootheid.



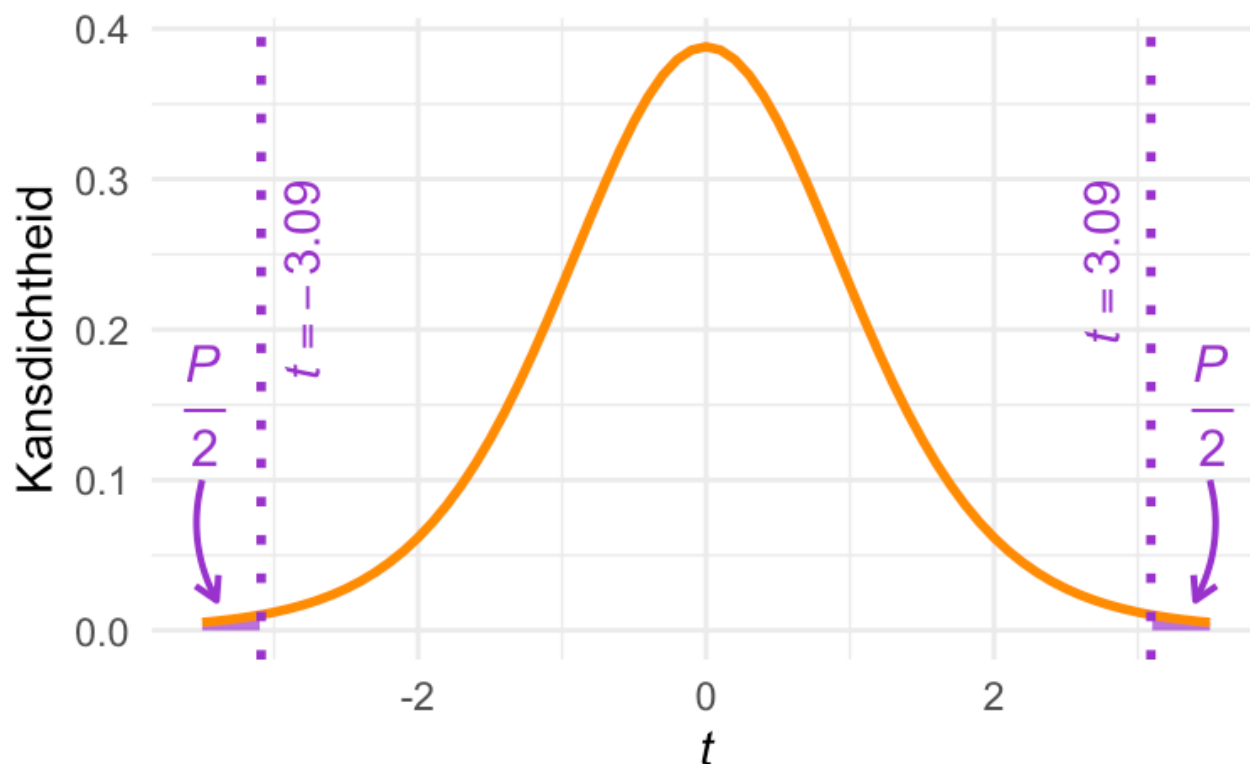
Bereken: $\bar{X} = 36,5^{\circ}\text{C}$

Bereken $s_X = 0,51^{\circ}\text{C}$

Bereken: $SE_{\bar{X}} = \frac{s_X}{\sqrt{n}} = \frac{0,51}{\sqrt{10}} = 0,16^{\circ}\text{C}.$

Bereken: $t = \frac{\bar{X} - \mu_0}{SE_{\bar{X}}} = \frac{36,5 - 37}{0,16} = -3,09.$

Stap 5: Bereken de P -waarde



Gebruik R:

```
t <- -3.087701  
n <- 10  
2*pt(-abs(t), df = n - 1)
```

Cumulative verdeling
van de t-verdeling.

```
[1] 0.01297755
```

Stap 6: Conclusie

- Een P -waarde van 0,013 is best klein (“significant”).
- Het steekproefgemiddelde van $36,5^{\circ}\text{C}$ is dus vrij opmerkelijk als in werkelijkheid het populatiegemiddelde 37°C is.
- Dit is dus een aanwijzing dat er met dat getal iets mis is.

Wanneer gebruik je deze toets?

- Continue variabele X
- Hypothese over populatiegemiddelde:
 $\mu_X = \mu_0$.
- Normale Model is redelijk.

! De t -toets voor één steekproef in 6 stappen

We zijn geïnteresseerd in een continue variabele X en willen de hypothese toetsen dat het populatiegemiddelde van X gelijk is aan een bepaalde waarde μ_0 . We nemen daarom een steekproef van n eenheden.

Dan zijn dit de stappen van de toets:

1. Kies het statistische model: het Normale Model.

2. Formuleer de hypotheses:

$$H_0 : \mu_X = \mu_0,$$

$$H_A : \mu_X \neq \mu_0.$$

3. Kies de toetsingsgrootheid:

$$t = \frac{\bar{X} - \mu_0}{SE_X}.$$

Verdeling van de toetsingsgrootheid onder H_0 :

De t -verdeling met $df = n - 1$.

4. Bereken de toetsingsgrootheid op basis van de steekproef.

Achtereenvolgens: $\bar{X}$, s_X , SE_X , en dan t .

5. Bereken de P -waarde:

Gebruik R:

```
P <- 2 * pt(-abs(t), df = n - 1)
```

6. Trek je conclusies.

t-toets helemaal met R uitvoeren:



```
# gegevens:
temperaturen <- c(35.7, 35.8, 37.0, 36.5, 37.1, 36.4, 36.1, 36.8, 36.5, 37.1)
mu0 <- 37

# berekening:
n <- length(temperaturen)
m <- mean(temperaturen)
s <- sd(temperaturen)
sem <- s/sqrt(n)
t <- (m - mu0) / sem

# bereken en print P-waarde:
(P <- 2*pt(-abs(t), df = n - 1))
```

```
[1] 0.01297755
```

Veel gemakkelijker: de functie `t.test()`



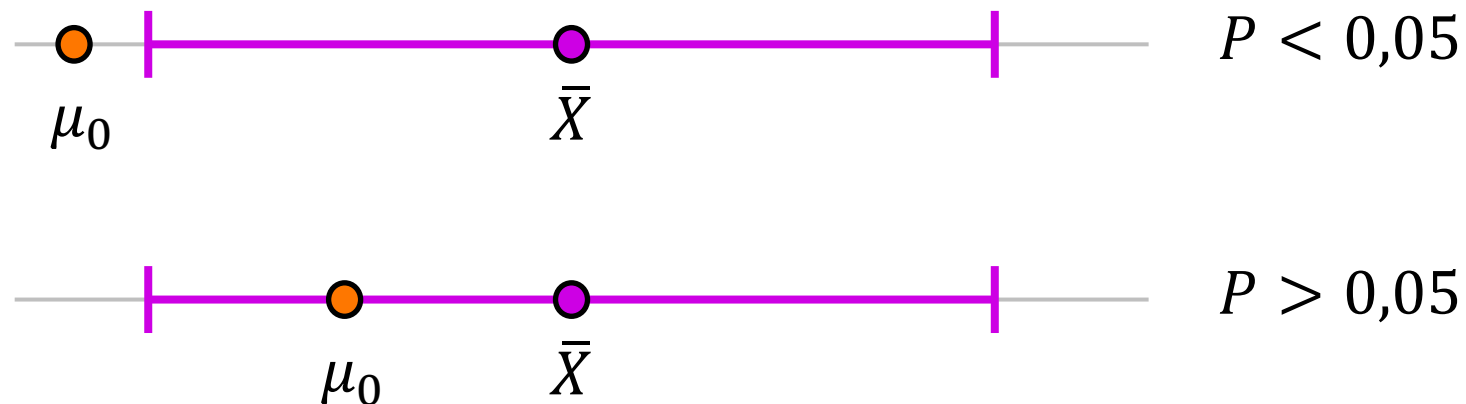
```
temperaturen <- c(35.7, 35.8, 37.0, 36.5, 37.1, 36.4, 36.1, 36.8, 36.5, 37.1)
mu0 <- 37
t.test(temperaturen, mu = mu0)
```

One Sample t-test

```
data:  temperaturen
t = -3.0877, df = 9, p-value = 0.01298
alternative hypothesis: true mean is not equal to 37
95 percent confidence interval:
 36.13368 36.86632
sample estimates:
mean of x
 36.5
```

Verband tussen 95%-BHI en P -waarde

- 95%-BHI is een interval is waarbinnen de ware parameter zich met 95% zekerheid bevindt.
- als H_0 waar is, dan zou 95% van de keren dat we een steekproef trekken μ_0 binnen het berekende 95%-BHI moeten liggen, en maar 5% van de keren daarbuiten
- Daaruit volgt:



Inhoud

1. De t -toets voor één steekproef
2. De t -toets voor twee onafhankelijke steekproeven
3. De t -toets voor gepaarde steekproeven
4. De statistische modellen van de t -toetsen

t-toets voor 2 onafhankelijke steekproeven

Voorbeeld:

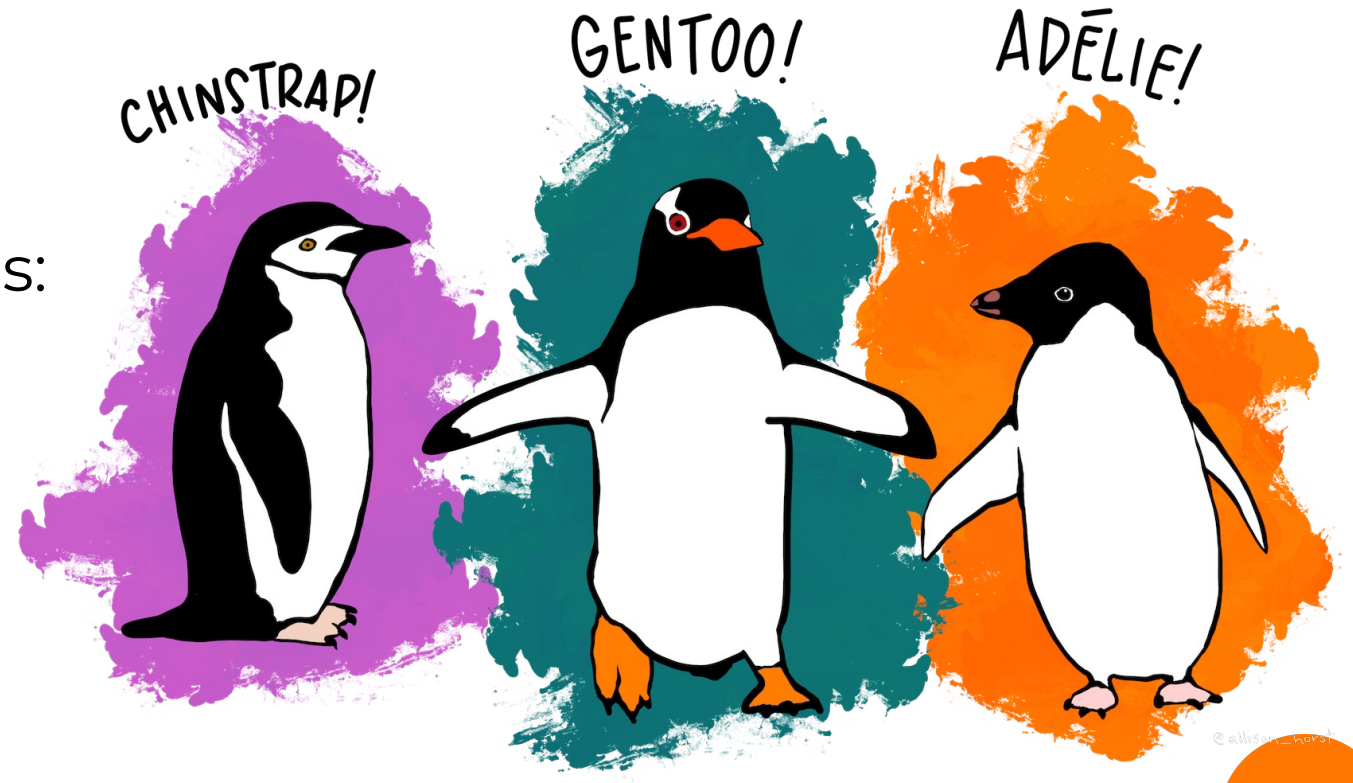
Verschillen de snavellengtes tussen pinguins van de soorten Chinstrap en Gentoo?

Twee steekproeven

uit twee verschillende populaties:

Chinstrap: $n_1 = 68$

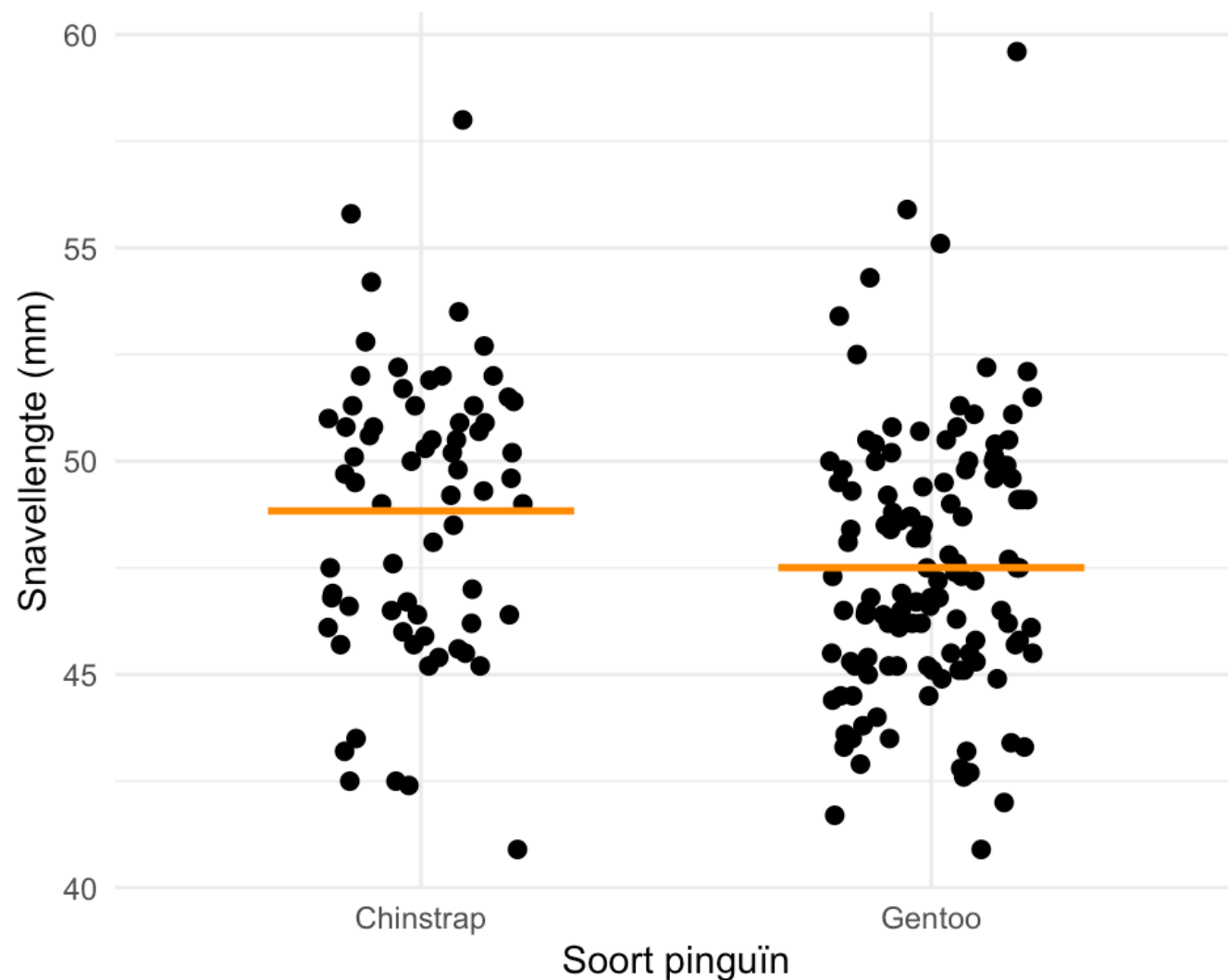
Gentoo: $n_2 = 123$



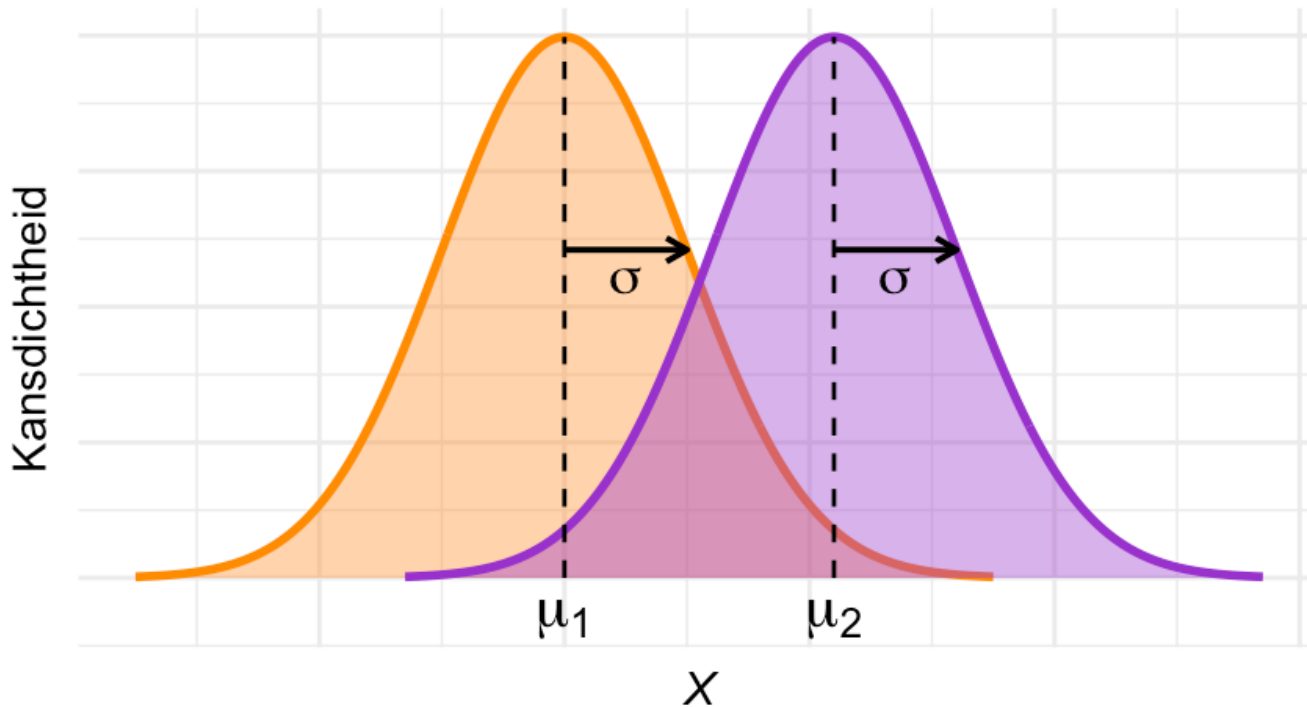
De gegevens:

Kan dit verschil het gevolg zijn van steekproevenvariatie?

Voer een t-toets voor 2 onafhankelijke steekproeven uit.

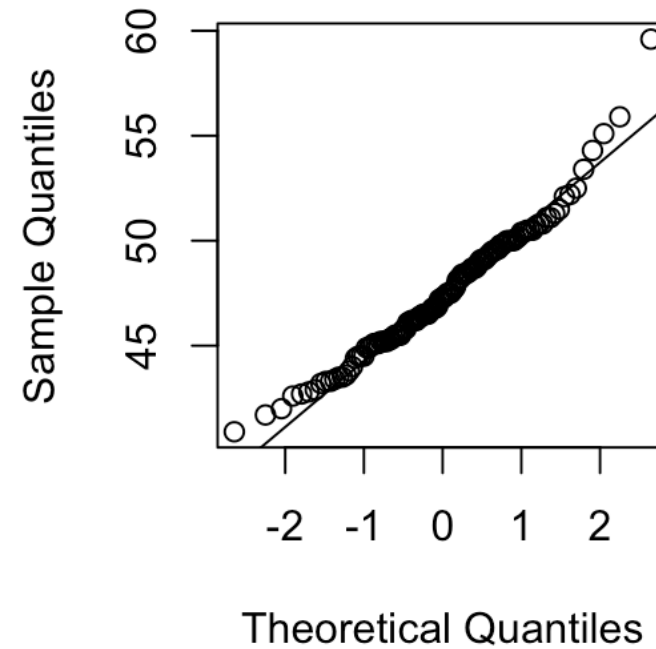
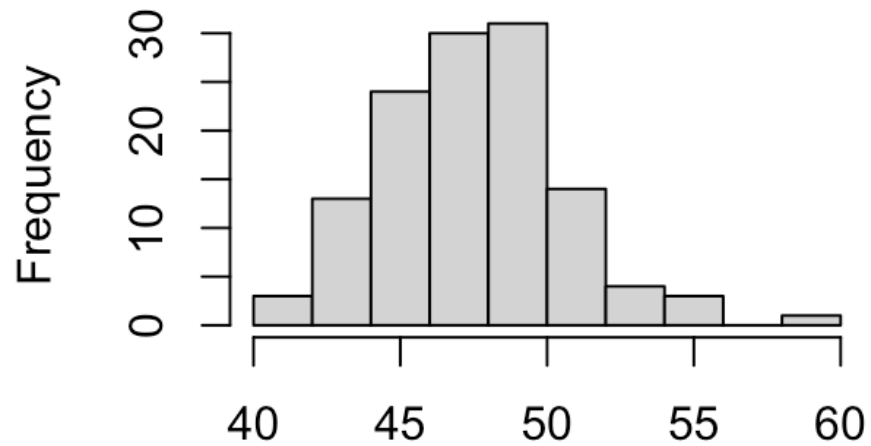
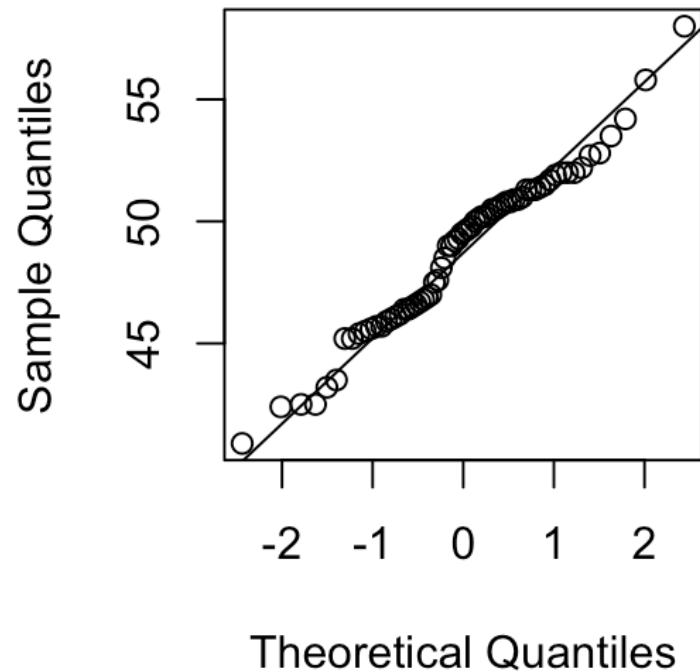
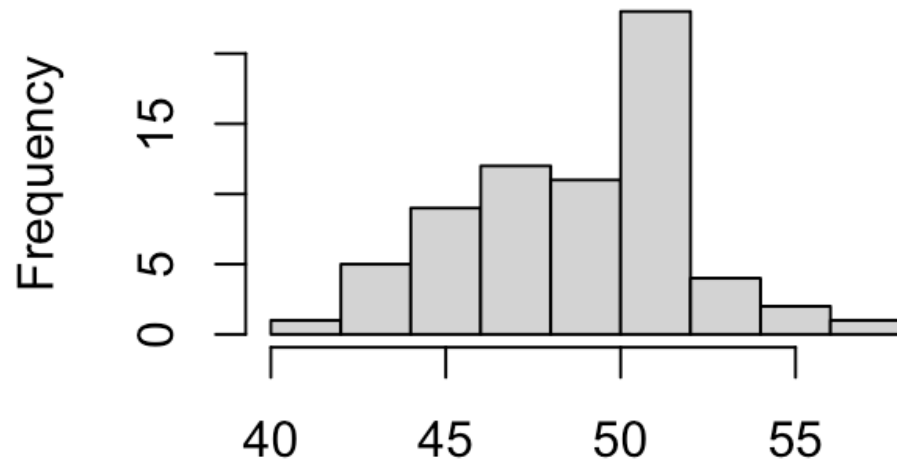


Stap 1: Het statistisch model



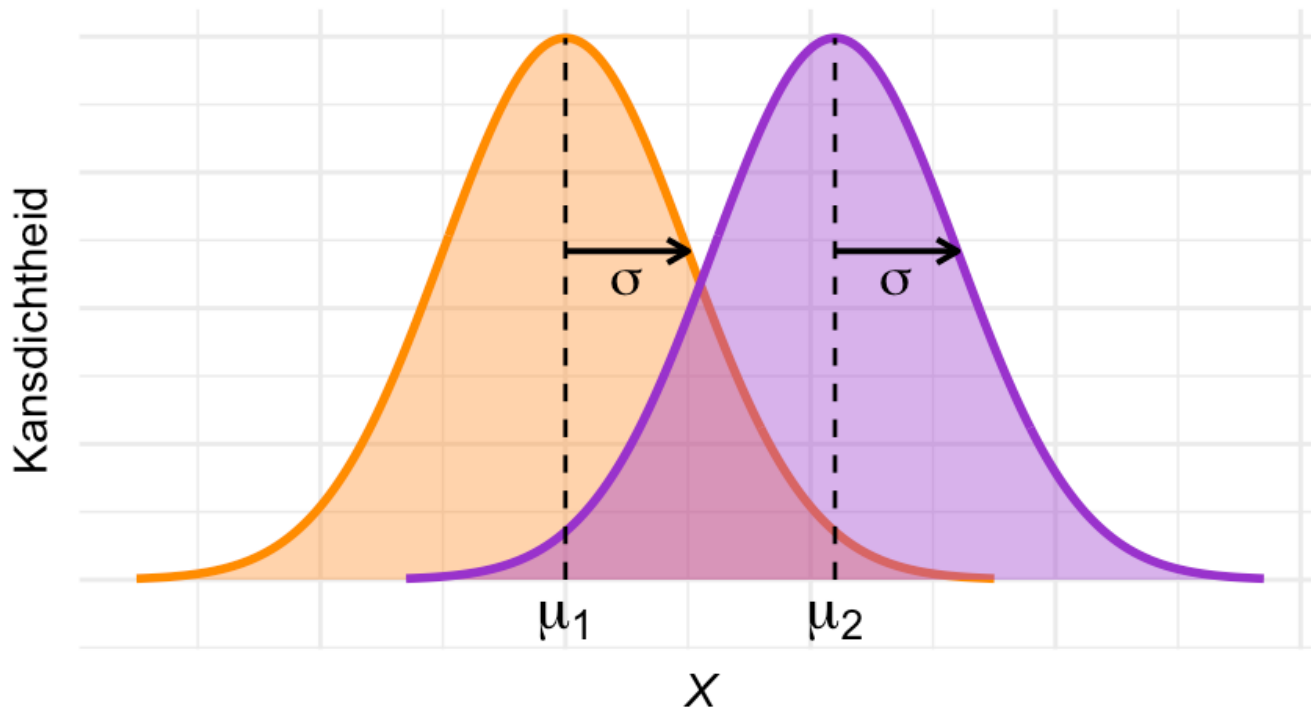
1. Beide populaties en steekproeven voldoen aan het Normale Model.
2. De standaarddeviatie van beide populaties zijn gelijk.

(Altijd controleren of het model redelijk is!)



Stap 2: Stel de hypotheses op

Nulhypothese: De gemiddelden van beide populaties zijn gelijk.



Oftewel:

$$H_0: \mu_1 = \mu_2,$$

$$H_A: \mu_1 \neq \mu_2.$$

Stap 3: De toetsingsgrootheid en zijn verdeling

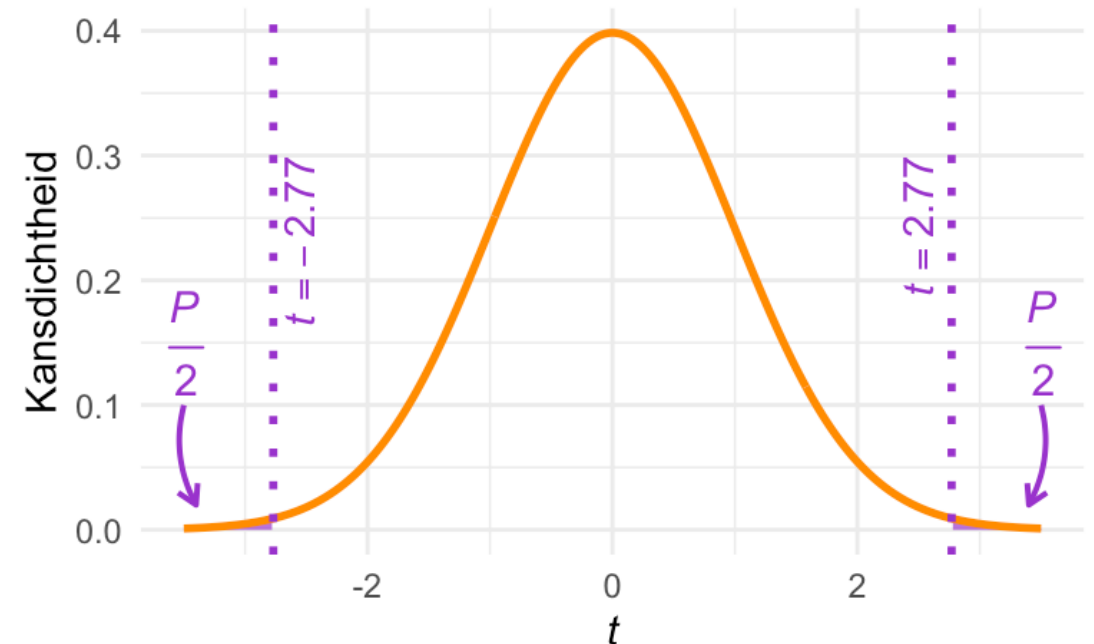
Toetsingsgrootheid:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE_{\bar{X}_1 - \bar{X}_2}}$$

(Formule voor $SE_{\bar{X}_1 - \bar{X}_2}$ staat in het boek.)

Verdeling:

t -verdeling met $df = n_1 + n_2 - 2$.

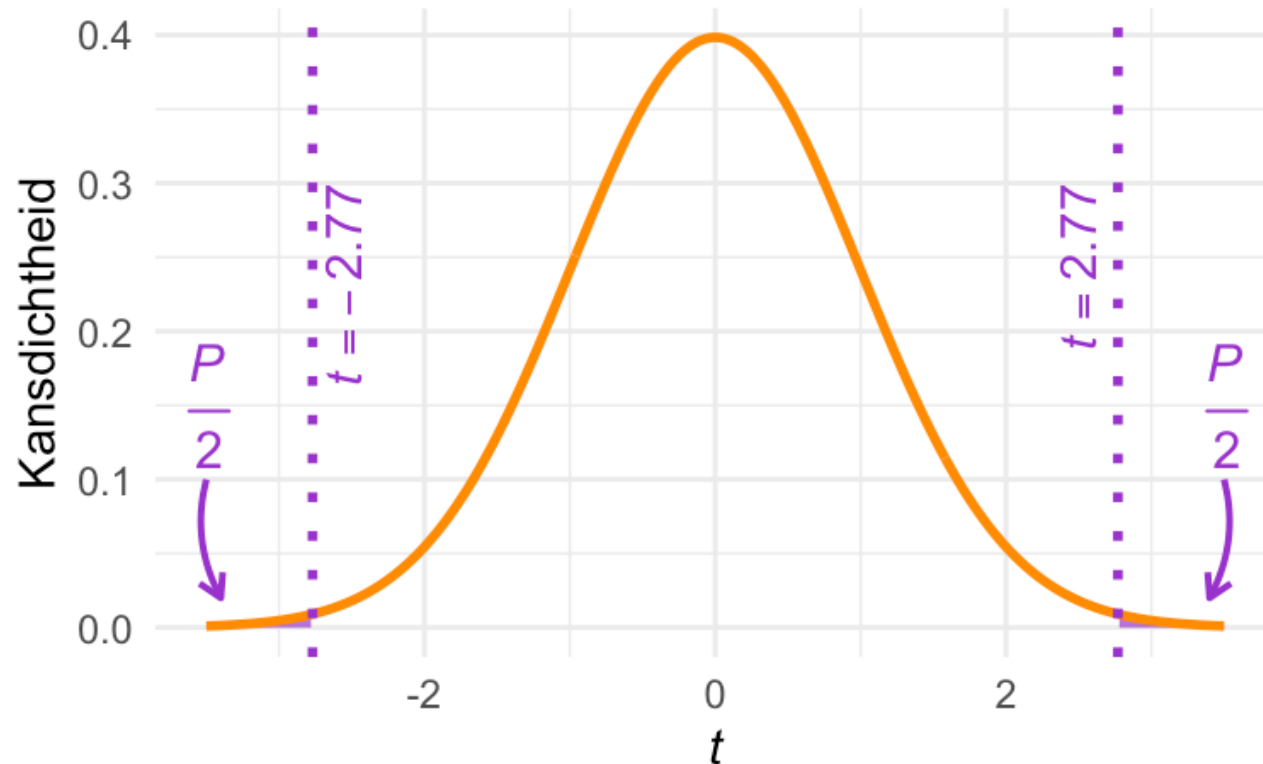


Stap 4: Bereken de toetsingsgrootheid.

Stap 5: Bereken de P -waarde

```
t = 2.7694  
(P <- 2 * pt(-abs(t), df = 189))
```

```
[1] 0.006175896
```



Snellere methode: gebruik `t.test()`!



```
t.test(snavel_Chinstrap, snavel_Gentoo, var.equal = TRUE)
```

Two Sample t-test

data: snavel_Chinstrap and snavel_Gentoo

t = 2.7694, df = 189, p-value = 0.006176

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

0.3823625 2.2755285

sample estimates:

mean of x mean of y

48.83382 47.50488

95%-BHI voor het verschil in
de populatiegemiddelden.

0 valt buiten het BHI.

Stap 6: Conclusies

- We vonden $P = 0,0062$.
- Dat is een erg kleine waarde (“highly significant”).
- Deze gegevens leveren dus sterk bewijs voor een verschil in de gemiddelde lengte van de snavels van Chinstrap-pinguïns en Gentoo-pinguïns.
- 95%-BHI voor het verschil: [0,38 mm; 2,28 mm]
Op een lengte van ~48mm is dat dus een klein verschil,

Inhoud

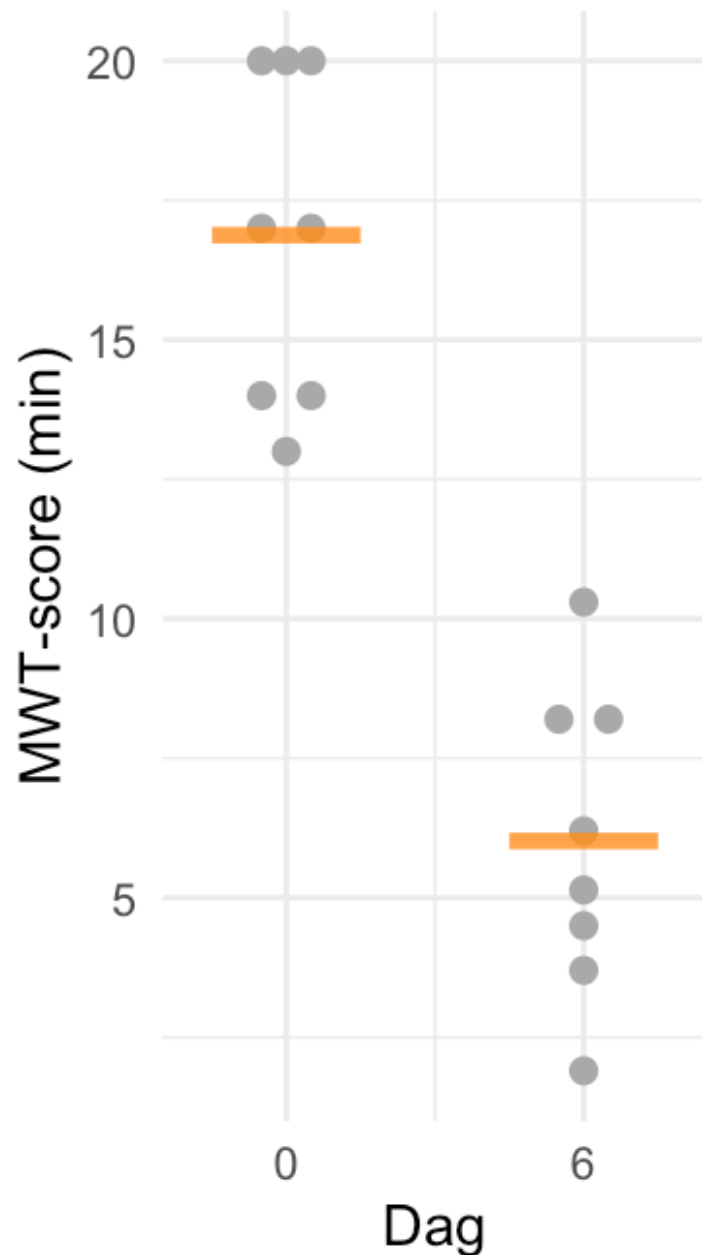
1. De t -toets voor één steekproef
2. De t -toets voor twee onafhankelijke steekproeven
3. De t -toets voor gepaarde steekproeven
4. De statistische modellen van de t -toetsen

De effecten van slaaptekort



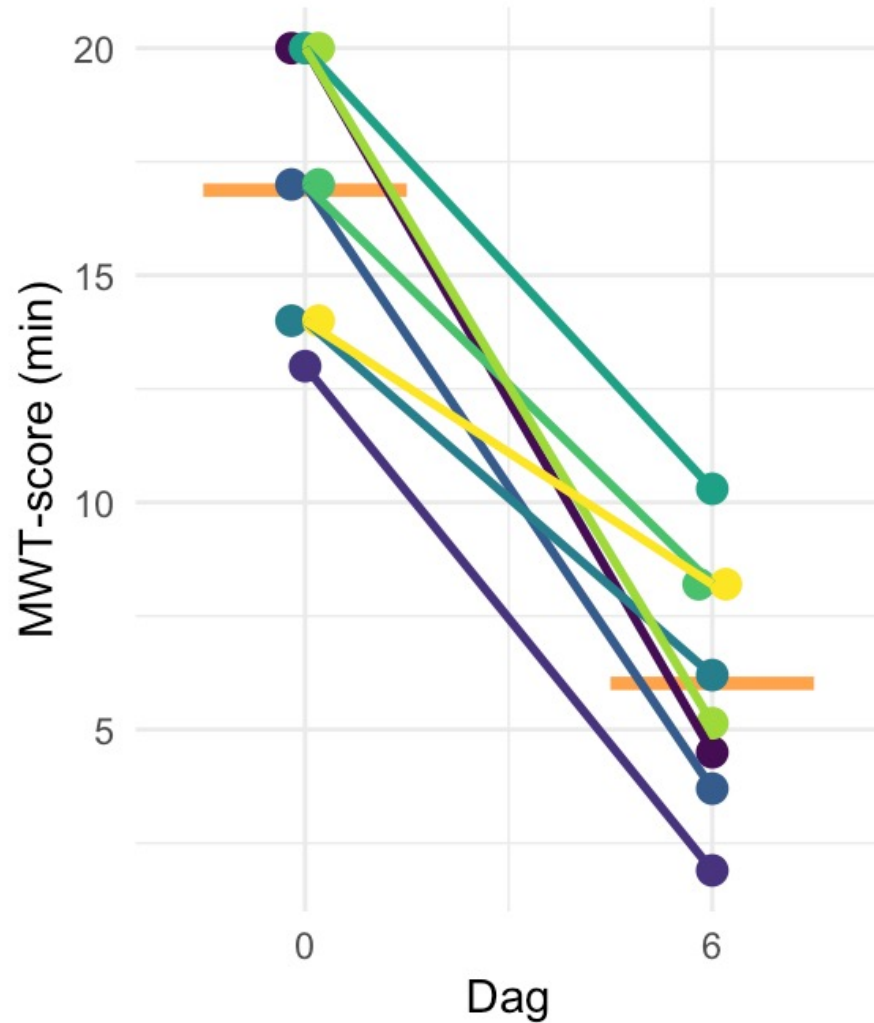
- Gedurende één week mochten acht proefpersonen slechts 4 uur per nacht in bed doorbrengen.
- Vervolgens werd elke dag gemeten hoe slaperig de proefpersonen waren door middel van de *Maintenance of Wakefulness test* (MWT).

Resultaten op dag 0 en dag 6



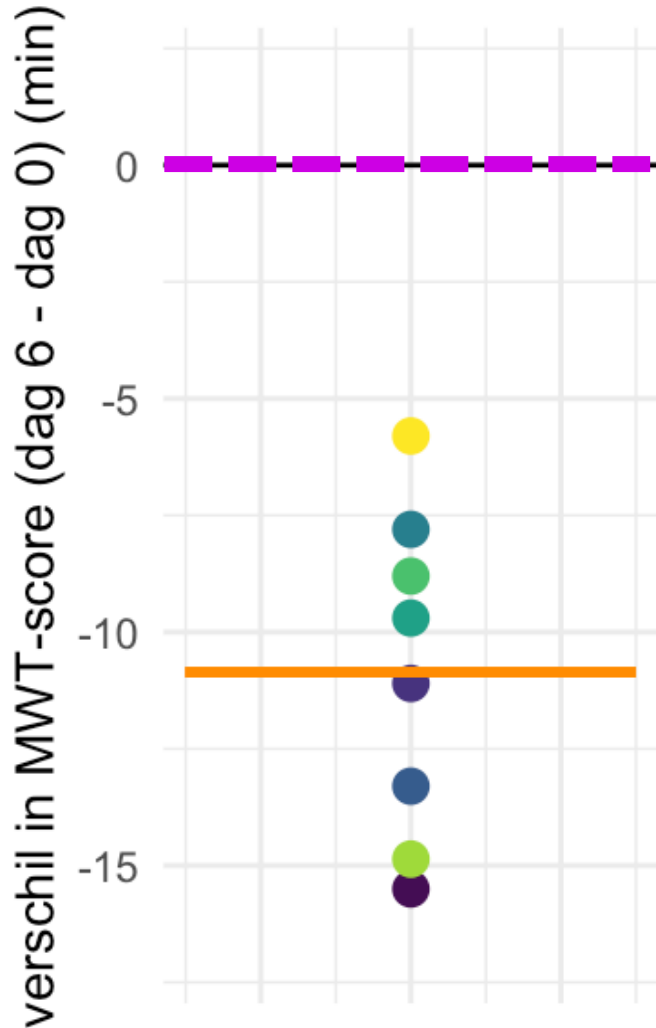
- Zijn de deelnemers op dag 6 gemiddeld slaperiger?
- Grote verschil: in beide steekproeven zijn dezelfde proefpersonen gebruikt.
- De twee steekproeven zijn dus niet onafhankelijk!
- Normale Model niet op beide steekproeven van toepassing.

Gepaarde gegevens



- De data bestaan uit paren van datapunten.
- Niet geïnteresseerd in onderlinge verschillen, alleen in *verschillen* tussen MWT-score op dag 6 en dag 0 *van dezelfde* persoon.
- Vraag is:
Zijn die *verschillen* systematisch negatief?

t -toets voor gepaarde steekproeven = t -toets voor één steekproef, voor verschillen.



Het verschil tussen dag 6 en dag 0: variabele D .

Stap 1: Neem het Normale Model aan voor variabele D .

Stap 2: $H_0: \mu_D = 0$; $H_A: \mu_D \neq 0$.

Stap 3: Toetsingsgrootte: $t = \frac{\bar{D}}{SE_{\bar{D}}}$, $SE_{\bar{D}} = s_D / \sqrt{n}$.

Verdeling: t -verdeling, $df = n - 1$.

Stap 4: Uitrekenen: $t = -8,86$.

Stap 5: P -waarde:

```
n <- 8
(P <- 2 * pt(-abs(t), df = n - 1))
```

Stap 6: Conclusie.

```
[1] 4.723134e-05
```

Gemakkelijker: gebruik `t-test()`



```
dag0 <- c(20, 13, 17, 14, 20, 17, 20, 14)
dag6 <- c(4.5, 1.9, 3.7, 6.2, 10.3, 8.2, 5.14, 8.2)

t.test(dag0, dag6, paired = TRUE)
```

Paired t-test

data: dag0 and dag6

$t = 8.86$, $df = 7$, $p\text{-value} = 4.723e-05$

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:

7.959764 13.755236

sample estimates:

mean difference

10.8575

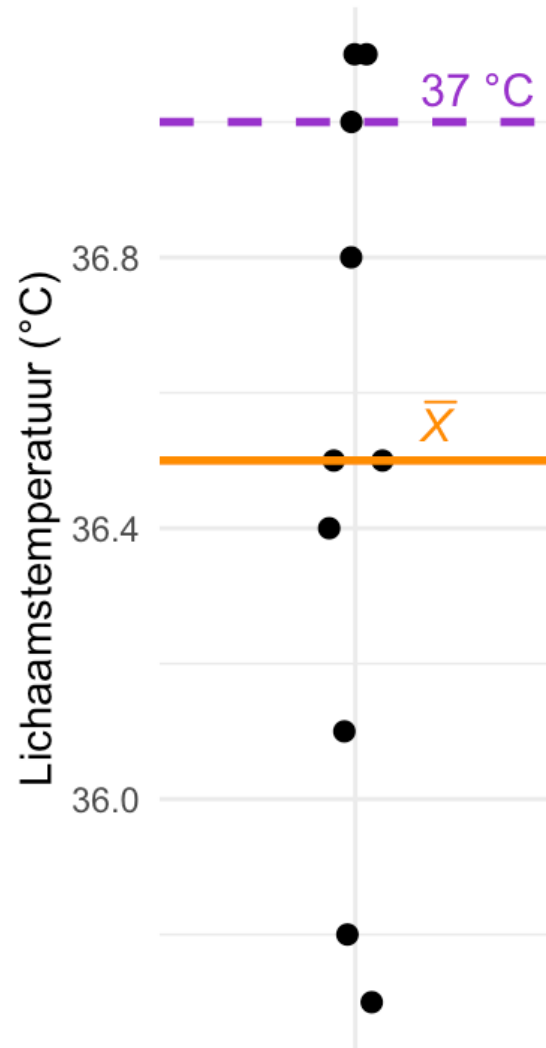
Wanneer t-toets voor gepaarde steekproeven?

- Er zijn twee steekproeven, condities, behandelingen.
- Bij ieder datapunt in steekproef 1 hoort een datapunt in steekproef 2.
- Dat levert gepaarde gegevens op.
- We zijn niet geïnteresseerd in de onderlinge verschillen tussen de paren, maar in de verschillen *binnen* de paren.
- Die verschillen geven aan wat het effect is van de conditie of behandeling.

Inhoud

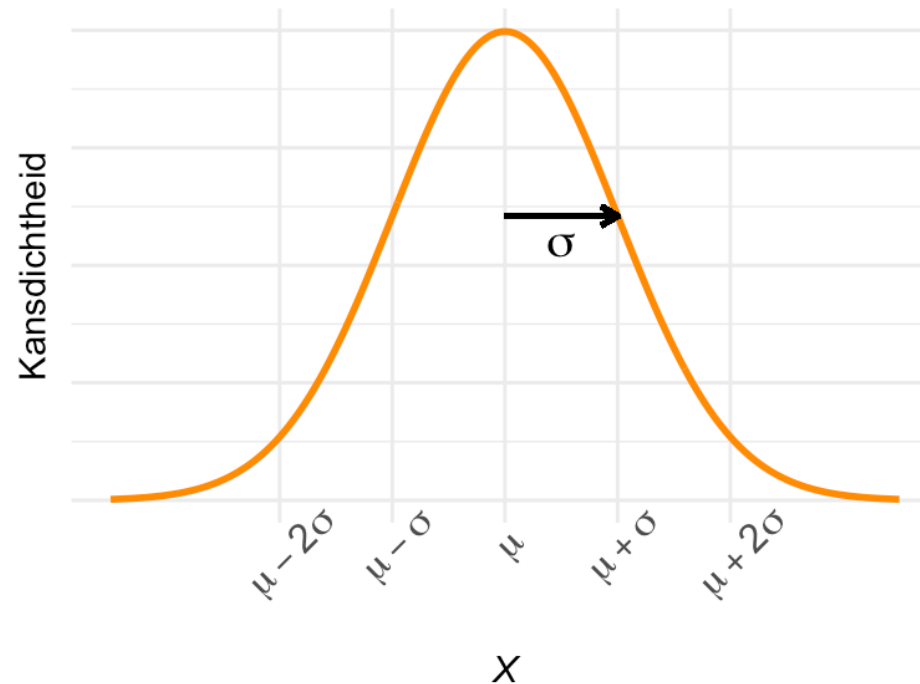
1. De t -toets voor één steekproef
2. De t -toets voor twee onafhankelijke steekproeven
3. De t -toets voor gepaarde steekproeven
4. De statistische modellen van de t -toetsen

Het Normale Model bij de drie toetsen

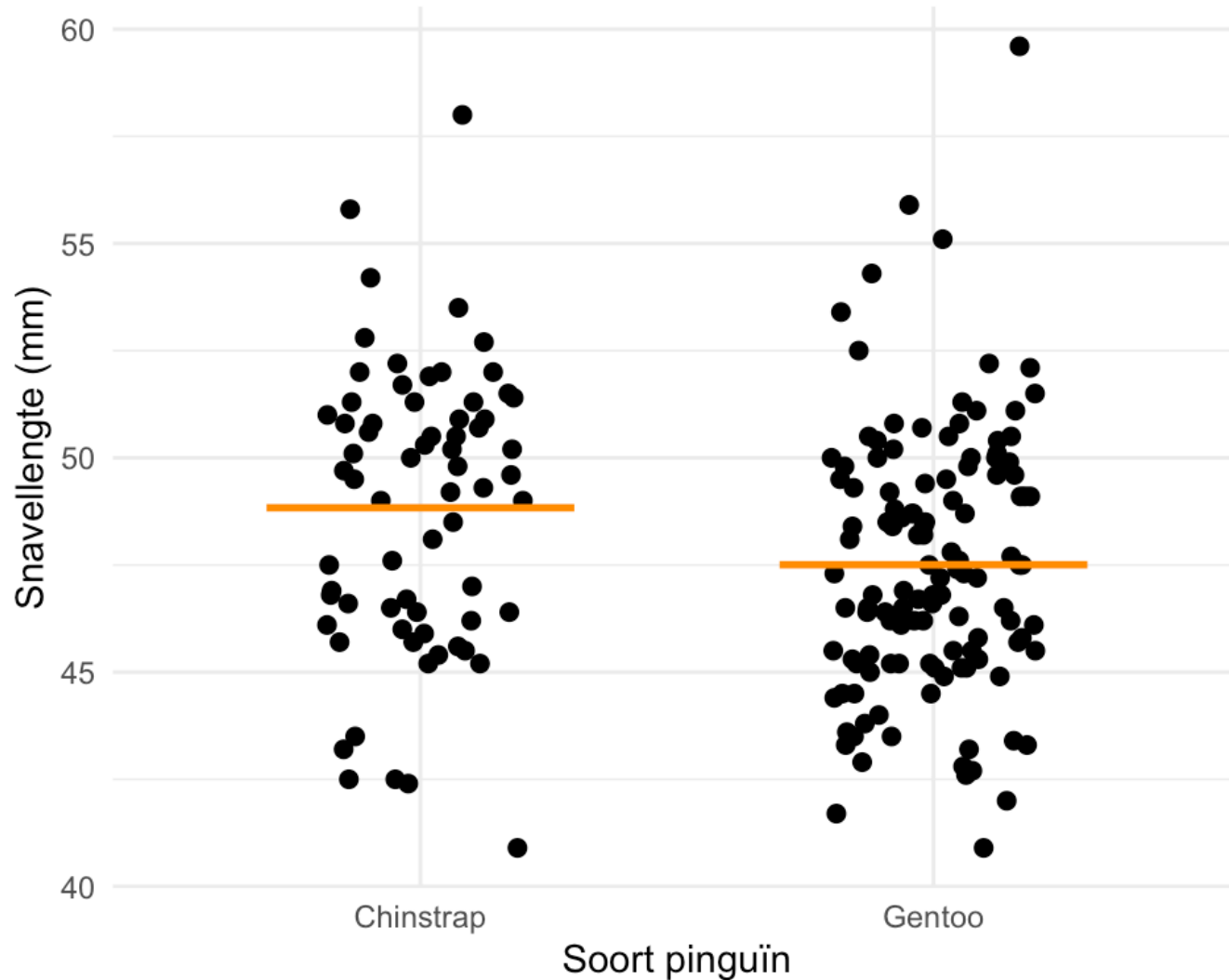


Één steekproef:

De variabele X is normaal verdeeld in de populatie.

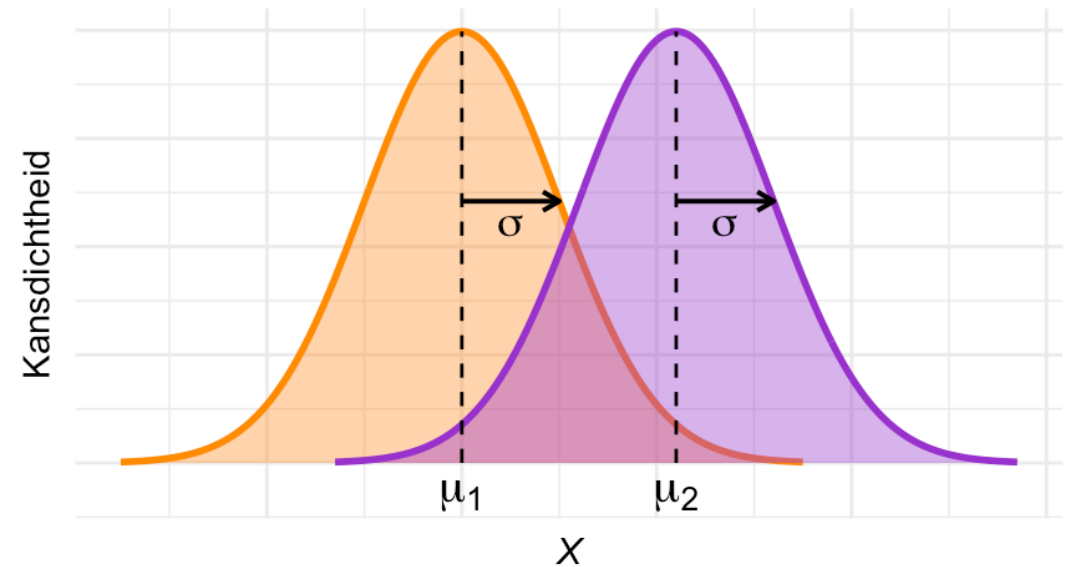


Het Normale Model bij de drie toetsen

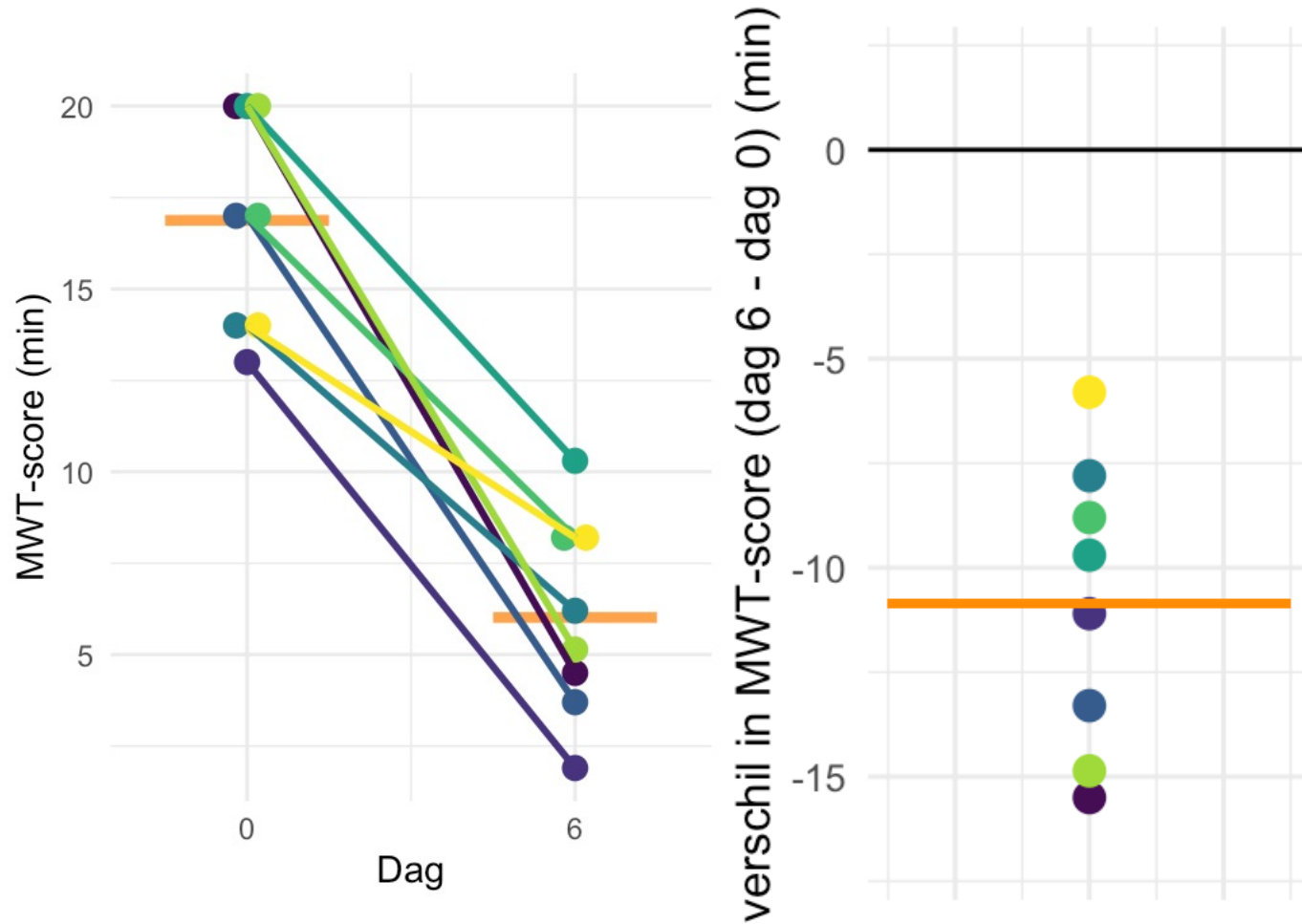


Twee onafhankelijke
steekproeven:

De variabele X is normaal
verdeeld in *beide* populaties.



Het Normale Model bij de drie toetsen



Twee gepaarde steekproeven:

De verschillen D zijn normaal verdeeld.

