One-sample t-toets

Table of Contents

<div class="navbar-header">  
 <button type="button" class="navbar-toggle collapsed" data-toggle="collapse" data-target="#navbar">  
 <span class="icon-bar"></span>  
 <span class="icon-bar"></span>  
 <span class="icon-bar"></span>  
 </button>  
 <a class="navbar-brand" href="Index.html">Statistisch Handboek Studiedata</a>  
</div>  
<div id="navbar" class="navbar-collapse collapse">  
 <ul class="nav navbar-nav navbar-start">  
 <li>

Toetsmatrix (R)

Toetsmatrix (Python)

Begrippenlijst

Over

Verantwoording

Feedback

Nieuwsbrief

</ul>  
 <ul class="nav navbar-nav navbar-right">  
  
 </ul>  
</div><!--/.nav-collapse -->

Disclaimer: Het peer review proces voor deze toets is nog niet afgerond; daarom is deze pagina nog in concept.

# Toepassing

Gebruik de *one sample t-toets* om het gemiddelde van de steekproef te vergelijken met een bekend gemiddelde of norm van de totale populatie.[[1]](#footnote-21)

# Onderwijscasus

De opleidingsdirecteur van de opleiding Werktuigbouwkunde wil weten of het gemiddelde eindexamencijfer voor de exacte vakken (Wiskunde, Natuurkunde en Scheikunde; WNS) van vwo studenten anders is dan het landelijk gemiddelde (6,8).[[2]](#footnote-24) Met deze gegevens probeert zij een inschatting te maken van het niveau van de studenten en kan zij bepalen of het curriculum van de inleidende vakken genoeg aansluit bij eerstejaars studenten met een vwo vooropleiding.

Dit onderzoek vertaalt zich in de volgende combinatie van hypothesen, waarbij de nulhypothese zo geformuleerd is dat er geen effect of verschil is en de alternatieve hypothese zo geformuleerd is dat er wel een effect of verschil is.

*H0*: Het gemiddelde eindexamencijfer voor de exacte vakken van vwo studenten die beginnen aan de Bachelor Werktuigbouwkunde is gelijk aan het landelijk gemiddelde: µ = 6,8.

*HA*: Het gemiddelde eindexamencijfer voor de exacte vakken van vwo studenten die beginnen aan de Bachelor Werktuigbouwkunde is anders dan het landelijk gemiddelde: µ ≠ 6,8.

# Assumpties

Om een valide resultaat te bereiken moeten, voordat de toets kan worden uitgevoerd, de data aan een aantal voorwaarden voldoen.

## Normaliteit

De *t-toets* gaat ervan uit dat de data van de steekproef normaal verdeeld zijn. Ga er bij meer dan 100 observaties vanuit dat de t-toets robuust genoeg is om uit te voeren zonder dat de data een normale verdeling volgen.[[3]](#footnote-29)

Controleer de assumptie van normaliteit met de volgende stappen:  
1. Controleer de data visueel met een histogram, een boxplot of een Q-Q plot.  
2. Toets of de data normaal verdeeld zijn met de *Kolmogorov-Smirnov test* of bij een kleinere steekproef (n < 50) met de *Shapiro-Wilk test*.[[4]](#footnote-31), [[5]](#footnote-33)

Als blijkt dat de data niet normaal verdeeld zijn, transformeer dan de data eventueel en bepaal daarna of deze wel normaal verdeeld zijn.[[6]](#footnote-35)

Als er geen sprake is van normaliteit, gebruik de [Wilcoxon signed rank toets](07-Wilcoxon-signed-rank-toets-R.html).[[7]](#footnote-37), [[8]](#footnote-38)

# Uitvoering

Er is een dataset ingeladen met de gemiddelde eindexamencijfers van WNS van eerstejaars Werktuigbouwkunde: Gemiddeld\_cijfer\_WNS.

## De data bekijken

Gebruik head() en tail() om de structuur van de data te bekijken.

## Eerste 5 observaties  
head(Gemiddeld\_cijfer\_WNS)  
## [1] 7.400000 7.366667 7.366667 7.666667 5.933333 6.500000  
  
## Laatste 5 observaties  
tail(Gemiddeld\_cijfer\_WNS)  
## [1] 6.333333 6.133333 8.066667 5.166667 7.700000 6.366667

Inspecteer de data met length(), mean()en sd() om meer inzicht te krijgen in de data.

## Gemiddelde en standaarddeviatie  
length(Gemiddeld\_cijfer\_WNS)  
## [1] 124  
mean(Gemiddeld\_cijfer\_WNS)  
## [1] 7.128763  
sd(Gemiddeld\_cijfer\_WNS)  
## [1] 0.7850378

* Gemiddeld cijfer WNS (standaardafwijking): 7,13 (0,79). *n* = 124.

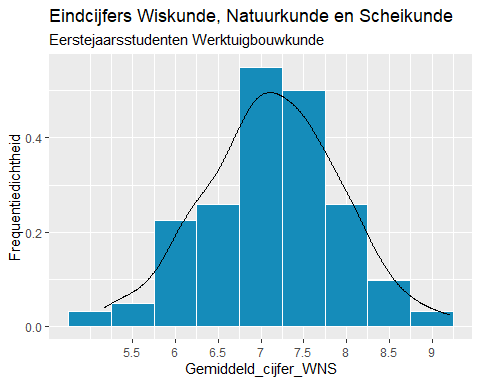
## Visuele inspectie van normaliteit

Geef de verdeling van de data visueel weer met een histogram, Q-Q plot en boxplot.

### Histogram

Focus bij het analyseren van een histogram op de symmetrie van de verdeling, de hoeveelheid toppen (modaliteit) en mogelijke uitbijters. Een normale verdeling is symmetrisch, heeft één top en geen uitbijters.[[9]](#footnote-44), [[10]](#footnote-46)

## Histogram met ggplot2  
library(ggplot2)  
  
ggplot(data.frame(Gemiddeld\_cijfer\_WNS),   
 aes(x = Gemiddeld\_cijfer\_WNS)) +  
 geom\_histogram(aes(y = ..density..),  
 binwidth = 0.5,   
 color = "white",   
 fill = "#158CBA") +  
 geom\_density(alpha = .2, adjust = 1) +  
 ylab("Frequentiedichtheid") +  
 xlab("Gemiddeld\_cijfer\_WNS") +  
 scale\_x\_continuous(labels = as.character(seq(5.5, 9, 0.5)),   
 breaks = seq(5.5, 9, 0.5)) +  
 labs(title = "Eindcijfers Wiskunde, Natuurkunde en Scheikunde",   
 subtitle = "Eerstejaarsstudenten Werktuigbouwkunde")



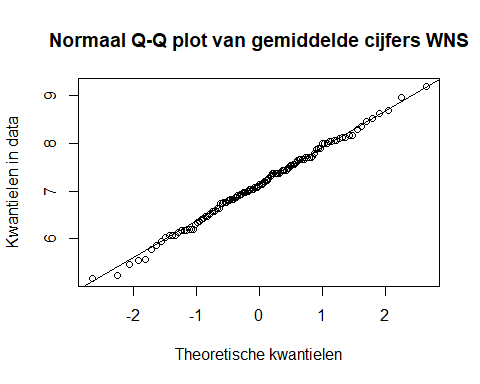
De histogram lijkt symmetrisch, heeft één top en geen outliers. De data is dus bij benadering normaal verdeeld.

### Q-Q plot

Gebruik qqnorm() en qqline() met pch = 1om een Q-Q plot te maken, met als datapunten kleine cirkels.

Als over het algemeen de meeste datapunten op de lijn liggen, kan aangenomen worden dat de data normaal verdeeld zijn.

## Q-Q plot  
qqnorm(Gemiddeld\_cijfer\_WNS,   
 pch = 1,  
 main = "Normaal Q-Q plot van gemiddelde cijfers WNS",  
 ylab = "Kwantielen in data",  
 xlab = "Theoretische kwantielen")  
qqline(Gemiddeld\_cijfer\_WNS)

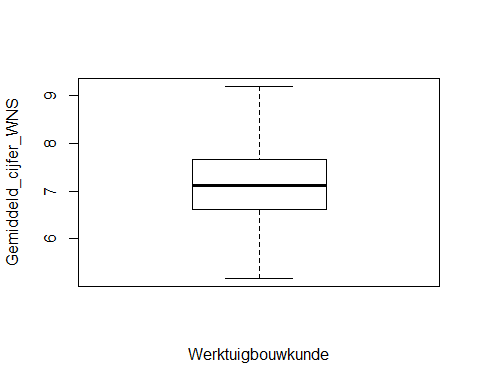


In deze casus liggen de meeste punten op de lijn. Bij de uiteinden liggen de punten dichtbij de lijn. Deze Q-Q plot duidt dus op een goede benadering van de normaalverdeling.

### Boxplot

De box geeft de middelste 50% van de tentamencijfers weer. De zwarte lijn binnen de box is de mediaan. In de staarten of snorreharen zitten de eerste 25% en de laatste 25%.. Cirkels visualiseren mogelijke uitbijters.[[11]](#footnote-51) Hoe meer de boxen overlappen, hoe waarschijnlijker er geen significant verschil is tussen de groepen.

## Boxplot  
boxplot(Gemiddeld\_cijfer\_WNS, xlab = "Werktuigbouwkunde", ylab = "Gemiddeld\_cijfer\_WNS")



De boxplot geeft de spreiding van het gemiddelde eindexamencijfer voor de exacte vakken weer van de studenten Werktuigbouwkunde. De box en staarten zien er symmetrisch uit, wat een indicatie is van een normale verdeling.[[12]](#footnote-53)

## Toetsen van normaliteit

Om te controleren of de data normaal verdeeld zijn, kan de normaliteit getoetst worden. Twee veelgebruikte toetsen zijn: de *Kolmogorov-Smirnov test* en de *Shapiro-Wilk test*.

### Kolmogorov-Smirnov

De *Kolmogorov-Smirnov test* toetst het verschil tussen twee verdelingen. Standaard toetst deze test het verschil tussen een normale verdeling en de verdeling van de steekproef.De Lilliefors correctie is vereist als het gemiddelde en de standaardafwijking niet van tevoren bekend of bepaald zijn, wat meestal het geval is bij een steekproef. Als de p < 0,05 is, is de verdeling van de data statistisch significant verschillend van de normale verdeling.

## Kolmogorov-Smirnov test  
library(nortest)  
lillie.test(Gemiddeld\_cijfer\_WNS)

##   
## Lilliefors (Kolmogorov-Smirnov) normality test  
##   
## data: Gemiddeld\_cijfer\_WNS  
## D = 0.041104, p-value = 0.8745

De p-waarde is 0,87, dus er is geen statistisch significant verschil gevonden tussen de verdeling van de steekproef en de normale verdeling. De *one sample-t-toets* kan uitgevoerd worden.

### Shapiro-Wilk Test

De *Shapiro-Wilk test* is een soortgelijke test als de *Kolmogorov-Smirnov test* en vooral geschikt bij kleine steekproeven (n < 50). Als de p < 0,05 is, is de verdeling van de data significant verschillend van de normale verdeling. Er is een subset van Gemiddeld\_cijfer\_WNS ingeladen: Gemiddeld\_cijfer\_WNS\_n30. De subset bevat 30 studenten. Voor een relatief kleine steekproef als deze is de *Shapiro-Wilk Test* geschikt.

## Shapiro-Wilk test  
shapiro.test(Gemiddeld\_cijfer\_WNS\_n30)

##   
## Shapiro-Wilk normality test  
##   
## data: Gemiddeld\_cijfer\_WNS\_n30  
## W = 0.9289, p-value = 0.04591

De p-waarde is 0,05, dus er is geen statistisch significant verschil gevonden tussen de verdeling van de steekproef en de normale verdeling. De *one sample-t-toets* kan uitgevoerd worden.

## One sample t-toets

Gebruik t.test() om een t-toets uit te voeren. Gebruik het argument mu = 6.8 om het gemiddelde te specificeren waarmee wordt vergeleken en specifieer welke alternatieve hypothese er getoetst wordt. De verwachting is dat de studenten hoger scoren, maar omdat het relevant is om te weten of de studenten ook lager scoren dan het landelijk gemiddelde, is er voor gekozen om tweezijdig te toetsen. Gebruik hiervoor alternative = "two.sided". Gebruik de hele dataset Gemiddeld\_cijfer\_WNS met *n* = 124.

## T-test  
t.test(Gemiddeld\_cijfer\_WNS, mu = 6.8, alternative = "two.sided")

##   
## One Sample t-test  
##   
## data: Gemiddeld\_cijfer\_WNS  
## t = 4.6634, df = 123, p-value = 7.97e-06  
## alternative hypothesis: true mean is not equal to 6.8  
## 95 percent confidence interval:  
## 6.989216 7.268311  
## sample estimates:  
## mean of x   
## 7.128763

* Vrijheidsgraden, *df* = *n* -1 = 124-1 = 123
* *t* 123 = 4,66, *p* < 0,0001
* p-waarde < 0,05, dus de H0 wordt verworpen [[13]](#footnote-58)
* 95%-betrouwbaarheidsinterval: bij het herhalen van het experiment met verschillende steekproeven van de populatie zal 95% van de betrouwbaarheidsintervallen de daadwerkelijke parameter bevatten, het gemiddeld eindexamencijfer exacte vakken. In deze casus is het interval tussen 6,99 en 7,27. Aangezien 6.8 niet in dit interval zit, verschilt het gemiddelde significant van 6.8.
* Het gemiddelde van de steekproef is 7,13

# Rapportage

De *one sample t-toets* is uitgevoerd om te toetsen of het gemiddelde eindexamencijfer voor de exacte vakken van vwo studenten die Werktuigbouwkunde zijn gaan studeren anders is dan het landelijk gemiddelde. Het gemiddelde van de steekproef (*M* = 7,13, *SD* = 0,79) is significant verschillend van het landelijk gemiddelde van 6,8, *t* 123 = 4,66, *p* < 0,0001. De resultaten ondersteunen de conclusie dat het gemiddelde eindexamencijfer voor de exacte vakken van studenten Werktuigbouwkunde met een vwo vooropleiding hoger ligt dan het landelijk gemiddelde.

Deze pagina maakt onderdeel uit van het Statistisch Handboek Studiedata, ontwikkeld binnen de zone Veilig en betrouwbaar benutten van studiedata van het Versnellingsplan. R code is uitgevoerd met R versie 3.6.3; Python code is uitgevoerd in Python 3.7. © 2020 Versnellingsplan - Statistisch Handboek Studiedata - Licentie Laatst gewijzigd op:05-08-2020

1. Van Geloven, N. (25 mei 2016). *T-toets* [Wiki Statistiek Academisch Medisch Centrum](https://wikistatistiek.amc.nl/index.php/T-toets#one_sample_t-toets). [↑](#footnote-ref-21)
2. Centraal Instituut voor Toetsontwikkeling (2018). *Examenverslag 2018*. <https://www.cito.nl/onderwijs/voortgezet-onderwijs/centrale-examens-voortgezet-onderwijs/tools-en-informatie-voor-docenten/examenverslagen/examenverslag-2018> [↑](#footnote-ref-24)
3. Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. Annu Rev Public Health, 23, 151-69. doi: 10.1146/annurev.publheath.23.100901.140546 <http://rctdesign.org/techreports/arphnonnormality.pdf> [↑](#footnote-ref-29)
4. Laerd statistics (2018). [Testing for Normality using SPSS Statistics](https://statistics.laerd.com/spss-tutorials/testing-for-normality-using-spss-statistics.php). [↑](#footnote-ref-31)
5. Normaliteit. (14 juli 2014). [UvA Wiki Methodologiewinkel](https://wiki.uva.nl/methodologiewinkel/index.php/Normaliteit). [↑](#footnote-ref-33)
6. Bereken de logaritme van de variabele en de inverse (1 gedeeld door de variabele), en bekijk of deze normaal verdeeld zijn. Als dat zo is, mag je die gebruiken voor de analyse. [↑](#footnote-ref-35)
7. De [Wilcoxon signed rank toets](07-Wilcoxon-signed-rank-toets-R.html) maakt een rangschikking van de data. Hierdoor is de test verdelingsvrij en is normaliteit geen assumptie. Ook zijn uitbijters minder van invloed op het eindresultaat. Toch wordt er voor deze test minder vaak gekozen, doordat bij het maken van een rankschikking de data informatie verliest. Als de data wel normaal verdeeld zijn heeft de [Wilcoxon signed rank toets](07-Wilcoxon-signed-rank-toets-R.html) minder onderscheidend vermogen dan wanneer de *one sample t-toets* uitgevoerd zou worden. [↑](#footnote-ref-37)
8. Prabhakaran, S. (2016-2017). *Statistical Tests*. <http://r-statistics.co/Statistical-Tests-in-R.html> [↑](#footnote-ref-38)
9. Outliers (13 augustus 2016). [UvA Wiki Methodologiewinkel](https://wiki.uva.nl/methodologiewinkel/index.php/Outliers). [↑](#footnote-ref-44)
10. Uitbijters kunnen bepalend zijn voor de uitkomst van toetsen. Bekijk of de uitbijters valide uitbijters zijn en niet een meetfout of op een andere manier incorrect verkregen data. Het weghalen van uitbijters kan de uitkomst ook vertekenen, daarom is het belangrijk om verwijderde uitbijters te melden in een rapport. [↑](#footnote-ref-46)
11. Outliers (13 augustus 2016). [UvA Wiki Methodologiewinkel](https://wiki.uva.nl/methodologiewinkel/index.php/Outliers). [↑](#footnote-ref-51)
12. Uitbijters kunnen bepalend zijn voor de uitkomst van toetsen. Bekijk of de uitbijters valide uitbijters zijn en niet een meetfout of op een andere manier incorrect verkregen data. Het weghalen van uitbijters kan de uitkomst ook vertekenen, daarom is het belangrijk om verwijderde uitbijters te melden in een rapport. [↑](#footnote-ref-53)
13. In dit voorbeeld wordt uitgegaan van een waarschijnlijkheid van 95% c.q. een p-waardegrens van 0,05. De grens is naar eigen inzicht aan te passen; houd hierbij rekening met type I en type II fouten. [↑](#footnote-ref-58)