Samenvatting Statistiek

Distribution mean standard deviation

Y in population *μ* *σ*

Y in sample  S

Y meta study *μ*= *μ* = *σ/*

**Statische schattingen** zijn processen van het interpeteren van dedata door (bijvoorbeeld) een steekproef te nemen.

Het doel van een **statische schatting** kunnen zijn:

* Een schatting van een bepaalde eigenschap van de populatie
* De precieze van deze schatting bepalen

***Standard error of the mean***:

Hoe zeker ben ik er van dat mijn gemiddelde overeenkomt met de daadwerkelijke meetwaardes.

SE= sd */*

Als x een normale verdeling heeft dan:

 - *μ = Z*

*σ/*

*mean = μ*

*standaarddeviatie = SD*

*standard normal verdeling = z*

*standard error of the mean = SE*

Bepalen of 2 populaties verschillend zijn of niet. **Distribution of “difference in means”** values

De hypothese dat *μ1* en  *μ2* niet gelijk zijn heet de ***alternatieve hypothese***

Ha : *μ1 != μ2*

De hypothese dat *μ1 en μ2* wel gelijk zijn heet de ***null hypothese***

Ho : *μ1 = μ2*

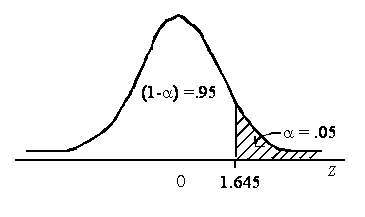
Je gaat uit van de **null hypothese** maar test de alternatieve hypothese.

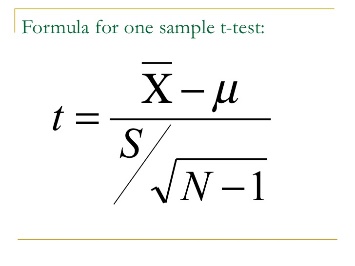
De t-test is een standaard methode om te kiezen tussen de 2 hypotheses.

***Alternative hypothese:*** Ha : *μ1 != μ2*

***Directional alternative:*** Ha : *μ1 < μ2*

***Another directional alternative:*** Ha : *μ1 > μ2*

*[](https://www.google.nl/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwjT_La45LPeAhWDEVAKHYieA7gQjRx6BAgBEAU&url=https://www.kean.edu/~fosborne/bstat/07amean.html&psig=AOvVaw0m_bx2ccFUcwhNloiIqEtB&ust=1541181942330212)One-tailed t.test*

[](https://www.google.nl/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwjzlbGE5rPeAhVQaFAKHVulAwAQjRx6BAgBEAU&url=https://www.slideshare.net/physlee/12-14621964&psig=AOvVaw0cCC9XbCya97HWR3A5n0Lv&ust=1541182403155735)

μ = steekproefgemiddelde

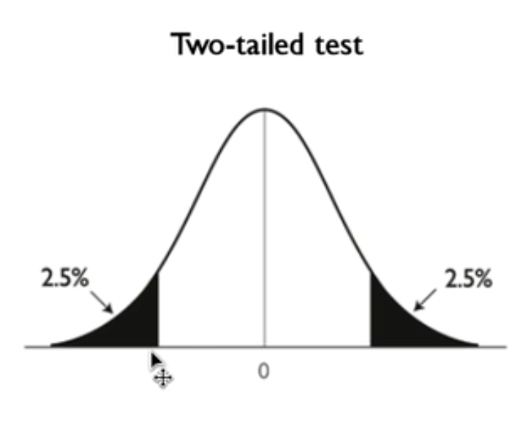
σ = steekproefstandaarddeviatie

Ho : *μ1 = μ2*

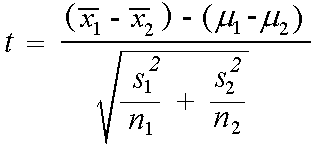
Ha : *μ1 < μ2*

Ha : *μ1 > μ2*

alternative = “less” or “greater”

[](https://www.google.nl/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwjDyuyk57PeAhUPKVAKHYezBOoQjRx6BAgBEAU&url=https://www.sophia.org/one-tailed-and-two-tailed-tests-top&psig=AOvVaw2yHc7ayrFgZNWXcJdjSkz3&ust=1541182713443113)

*Two-tailed t.test*

[](https://www.google.nl/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwj5lsvU5rPeAhURbVAKHVIjBAMQjRx6BAgBEAU&url=http://www.stat.yale.edu/Courses/1997-98/101/meancomp.htm&psig=AOvVaw2oVhixFixP_ybmNaOAedAa&ust=1541182452416754)

Ho : *μ1 = μ2*

Ha : *μ1 != μ2*

Alternative = “two.sided”

Pvalue > ἀ

Ho : niet verwerpen

Pvalue < ἀ

Ho : verwerpen

Ha : accepteren

Paired = TRUE als

* Proefpersonen voor en na het zelfde zijn
* Als je wil testen of m1 dezelfde resultaten geeft als m2

*Type 1 fout 🡪 Ho ten onrechten verwerpen, Ha accepteren*

*Type 2 fout 🡪 we vergeten Ho te verwerpen en nemen ten onrechte aan. H0 accepteren*

Waarom een gepaard design:

* Mindere bias
* Betere precisie

*One way anova*

Er is slechts 1 variabele (waarin wij geïnteresseerd zijn) die in de verschillende groepen varieert, die wij onderzoeken. Varieren e

* Normale verdeling van de gegevens
* De individuen zijn onafhankelijk van elkaar
* De groepen zijn onafhankelijk van elkaar
* SD zijn enigszins gelijk

**Is er een effect tussen (meer dan) 2 groepen?**

Betrouwbaarheidsinterval van 95%

Y +- T \* (SE)

T= 2,57 🡪 0,0025

Y = mean

SE = (sd */)*

Welke 2 gegevens heb je nog nodig om de toetsingsgrootheid te kunnen berekenen?

* Gemiddelde en SD

Welke 2 gegevens, die je berekent uit de resultaten bij deze monsters, heb je nog nodig om de toetsingsgrootheid te kunnen berekenen?

* Randomization en replication

**Voorbeeld tentamen statistiek thema 5 voor bi2**

*Het tentamen is individueel op papier met behulp van de computer en duurt 90 minuten.*

*Cesuur: 55 punten is voldoende*

***Neem voor de onbetrouwbaarheid bij alle opdrachten 5% .***

**Opdracht 1 ( 15 punten )**

Om te bepalen of hamburgers (bereid) in een snackbar meer dan 10 gram verzadigd vet per 100 gram bevatten neem je 5 monsters.

**a.** Welke toets kun je hier het beste gebruiken?

*De one sample t-test. (of de gewone t-toets ) 1-zijdig toetsen*

**b.** Welke 2 gegevens, die je berekent uit de resultaten bij deze monsters, heb je nog nodig om de toetsingsgrootheid te kunnen berekenen?

*Je hebt nog nodig het* ***steekproefgemiddelde*** *en de* ***steekproefstandaarddeviatie.***

**Opdracht 2 ( 20 punten )**

In een studie naar de relatie tussen dieet en serum-cholesterolgehalte, werden 16 vrouwen aselect in vier groepen verdeeld, die elk een ander dieet kregen.

Dit leverde de volgende cholesterolwaarden:

**Dieet 1 5,5 5,4 6,1 5,9**

**Dieet 2 5,4 5,8 5,3 6,8**

**Dieet 3 6,1 6,3 6,7 7,1**

**Dieet 4 5,9 5,8 6,1 6,5**

Toets of de cholesterolwaarden in de 4 groepen gemiddeld hetzelfde zijn.

a. Formuleer de nulhypothese en de alternatieve hypothese in woorden.

***Nulhypothese:*** *de cholesterolwaarden in de 4 groepen zijn gemiddeld hetzelfde*

***Alternatieve hypothese:*** *de cholesterolwaarden in de 4 groepen zijn gemiddeld niet hetzelfde*

b. Welke toets kun je hier het beste gebruiken?

*ANOVA one way*

c. Voer de toets uit met behulp van Excel of R en vermeld in je uitwerking de berekende

p-waarde en de conclusie die je daaruit trekt.

*p-waarde is 0,109987203 Onvoldoende reden om Ho te verwerpen.*

**Opdracht 3 ( 15 punten )**

Het AT% in een populatie is normaal verdeeld.

We willen onderzoeken of het AT% in 2 populaties hetzelfde is.

We hebben de volgende hypothesen:

H0: gemiddeld AT% in populatie A is hetzelfde als in populatie B

H1: gemiddeld AT% in populatie A is niet hetzelfde als in populatie B

We doen 4 metingen in populatie A en vinden de volgende AT% : 32 ; 31; 31 en 30.

We doen 4 metingen in populatie B en vinden de volgende AT% : 32 ; 34; 34 en 33.

Geven deze uitkomsten je genoeg reden om de nulhypothese te verwerpen?

Motiveren m.b.v. een toets. Vermeld welke toets je gebruikt en relevante tussenuitkomsten

*Ja, we mogen de nulhypothese verwerpen.*

*Een 2-zijdige ” two sample t-test with equal variances” geeft als p-waarde 0,011695964, dit is kleiner dan 0,05, dus H0 verwerpen. We mogen deze test gebruiken want bij het uitvoeren van de F-test was de p-waarde 0,799800667, dus de aanname dat de 2 populaties dezelfde variantie hebben voor het AT% mag ik aanhouden.*

*Verder geeft de 2-zijdige ” two sample t-test with unequal variances” als p-waarde 0,012194187, dit is ook kleiner als 0,05!*

**Opdracht 4 ( 20 punten )**

8 mensen met hoge bloeddruk krijgen gedurende 4 weken een medicijn om de bloeddruk te verlagen. Toets of het medicijn werkt. Neem α = 0,01.

Onderdruk voor de behandeling onderdruk na 4 weken medicijngebruik

A 92 84

B 95 88

C 97 85

D 90 87

E 105 92

F 98 82

G 90 92

H 108 90

*μv = bloeddruk voor de behandeling en μn  = bloeddruk na 4 weken*

*H0: μn = μv*

*H1: μn < μv*

*Voer een 1-zijdige gepaarde t-toets uit, dan vind je p-waarde is 0,0028. Dus H0 verwerpen.*

*De behandeling verlaagt de bloeddruk.*

**Opdracht 5 ( 15 punten )**

Om te bepalen of twee analyse-methoden dezelfde ijzerconcentraties geven, onderzoek je 10 preparaten met beide methodes. De uitkomsten staan hieronder in milligram per 100 gram.

**preparaat 1 2 3 4 5 6 7 8 9 10**

**methode 1** 13 11 21 8 7 35 16 29 5 1

**methode 2** 11 12 20 9 9 32 14 28 6 11

a. Formuleer de nulhypothese en de alternatieve hypothese.

*Nulhypothese: de twee analyse-methoden geven dezelfde ijzerconcentraties*

*Alternatieve hypothese: de twee analyse-methoden geven niet dezelfde ijzerconcentraties*

b. Welke toets kun je hier het beste gebruiken?

*Paired t-test*

c. Voer de toets uit met behulp van Excel of R en vermeld in je uitwerking de berekende

p-waarde en de conclusie die je daaruit trekt.

*p- waarde is 0,362938239*

*Ho niet verwerpen*

**Opdracht 6 ( 10 punten )**

Bepaal een 95%-betrouwbaarheidsinterval voor de pH van tien liter vloeistof op grond van de volgende 6 pH-metingen: 9,20 9,18 9,05 9,11 9,10 9,08.

*De algemene vorm van een betrouwbaarheidsinterval is ± t·s/*

*Dus hier betrouwbaarheidsinterval is 9,120 ± 2,57 x (0,05831/wortel 6)*

*9,120 ± ( 2,57 x0,0238049)*

*dus 9,120 ± 0,0611787*