

## Hoofdstuk 1

1. Wat is het verschil tussen populatie en sample?
2. Wat is het verschil tussen simple random sampling en systematic sampling?
3. Wat is het verschil tussen een parameter en een statistiek?
4. Wat zijn non sampling errors?
5. Wat is het verschil tussen descriptive statistics en inferential statistics?
6. Numerieke variabelen kunnen twee vormen aannemen, welke?
7. Wat is het verschil tussen kwalitatieve en kwantitatieve gegevens?
8. Welke meetniveaus worden voor kwalitatieve gegevens gebruikt?
9. Welke meetniveaus worden voor kwantitatieve gegevens gebruikt?
10. Wat zijn tabellen en grafieken die voornamelijk worden gebruikt voor nominaal meetniveau?
11. Wat zijn grafieken die voornamelijk worden gebruikt om numerieke variabelen te beschrijven?

## Hoofdstuk 2

1. Wat zijn numerieke maten en op welke vragen geven ze antwoord?
2. Welke maten worden gebruikt bij categorische gegevens?
3. Welke maten worden gebruikt bij numerieke gegevens?
4. Hoe weet je naar welke kant de scheefheid van de verdeling is?
5. Hoe weet je of de verdeling symmetrisch is?
6. Wat is de vijf-nummer samenvatting?
7. Hoe bereken je de interkwartiel afstand?
8. Welke grafiek geeft de vorm van de verdeling in termen van de vijf-nummer samenvatting?
9. Hoe bereken je de variantie?
10. Wat is het verschil tussen Chebyshev's theorem en de z-score?
11. Wat hebben covariantie en correlatiecoëfficiënt met elkaar gemeen?

## Hoofdstuk 3

1. Wat is zijn basis uitkomsten en wat hebben deze te maken met de steekproefruimte
2. Wanneer zijn de events collectively exhaustive?
3. Wat is het verschil tussen permutaties en combinaties?
4. Wat is het verschil tussen klassieke waarschijnlijkheid en subjectieve waarschijnlijkheid?
5. Wat zijn de 5 kansregels?
6. Wat zijn mogelijke gevolgen indien verschillende metingen worden gedaan over een zelfde steekproefgroep?

## Hoofdstuk 4

1. Wat zijn de vereiste eigenschappen van kansverdeling discrete random variabelen?
2. Wat is de verwachte waarde (expected value) van een discrete random variabele  $X$ ?
3. Hoe berekenen we de standaardafwijking?
4. Wat zijn de eigenschappen van een binominale kansverdeling?
5. Wat zijn de mogelijke uitkomsten van een binomiale verdeling (Bernoulli)?
6. Wat zijn de eigenschappen van gezamenlijke kansen?
7. Hoe weet je of de gezamenlijke verdeelde random variabelen  $X$  en  $Y$  onafhankelijk zijn?
8. Wat is een portfolio analyse?

## Hoofdstuk 5

1. Wat is de kans dat een continue random variabele in een bepaalde reeks valt?

Published on *WorldSupporter* ([www.worldsupporter.org](http://www.worldsupporter.org))

2. Wat zijn de eigenschappen van kansdichtheid functie?
3. Hoe ziet de kansdichtheidsfunctie eruit bij een normale verdeling?
4. Wat is de notatie van de normale verdeling?
5. Wat is de formule voor een normaal verdeelde random variabele en Z?

## Hoofdstuk 6

1. Waarom wordt er vaak gebruik gemaakt van een steekproef?
2. Wat zegt de central limit theorem?
3. Wat is de acceptatie interval?
4. Wat is de steekproef proportie?

## Hoofdstuk 7

1. Wat is het verschil tussen een schatter en een schatting?
2. Wat is de bias van een onpartijdige schatter?
3. Wat is de meest efficiënte schatter?
4. Hoe wordt de betrouwbaarheidsinterval van een schatting bepaald?
5. Hoe worden de grenzen van een betrouwbaarheidsinterval aangegeven?
6. Op welke manier kan de margin of error verminderd worden?
7. Wanneer wordt de student's t verdeling gebruikt? $n$

## Hoofdstuk 8

1. Wanneer zijn samples afhankelijk?
2. Waar staat  $d$  voor?
3. Wat betekent het als de uitkomst tussen twee drugs  $\mu_x - \mu_y$  positief is?
4. Welke variabele gebruik je wanneer de populatievariantie onbekend is?
5. Hoe worden deze geobserveerde sample variaties in de interval formule gebruikt indien als gelijk beschouwd?
6. Wanneer mag er gebruik worden gemaakt van het verschil tussen twee populatie proporties?
7. Hoe luidt de interval formule van het verschil tussen twee populatie proporties?

## Hoofdstuk 9

1. Wat is het verschil tussen de nul-hypothese en de alternatieve hypothese?
2. Wanneer is een hypothese eenzijdig en wanneer tweezijdig?
3. Wat zijn de twee mogelijke fouten bij het toetsen van hypothesen?
4. Wat gebeurt er als een uitkomst in de rejection region valt?
5. Wanneer wordt op basis van de p-waarde de nul-hypothese verworpen?
6. Hoe bepaal je de sterkte van een toets?

## Hoofdstuk 10

1. Hoe formuleren we de hypothesen wanneer je de verschillen tussen twee normale populatiegemiddelden tweezijdig test?
2. Wat is het verschil in  $t$  of  $z$  bij een eenzijdige en een tweezijdige toets?
3. Met welke verdeling testen we de gelijkheid van de variaties tussen twee normaal verdeelde populaties?
4. Hoe formuleren we de hypothesen wanneer je de gelijkheid van de variaties tussen twee normaal verdeelde populaties test?
5. Op welke aanname zijn de toetsen gebaseerd?

## Hoofdstuk 11

1. Waar staan  $b_1$  en  $b_0$  voor in een least squares regressielijn?
2. Welke onderdelen zitten er in de analyse van variantie?
3. Wat zegt de determinatiecoëfficiënt over de regressie?
4. Van welke twee hoeveelheden is de variantie van de helling coëfficiënt afhankelijk?
5. Wanneer kunnen regressiemodellen worden gebruikt bij het voorspellen van de afhankelijke variabele?
6. Wat voor een gevolg heeft een wijdere interval?
7. Welke drie hoeveelheden hebben invloed op de voorspelling en betrouwbaarheid intervallen en op welke manier?
8. Wat is het verschil tussen diversifiable risico en nondiversifiable risico?
9. Noem 3 voorbeelden van diversifiable risico.
10. Hoe kan het diversifiable risico gecontroleerd worden?
11. Wat geeft de beta coëfficiënt aan?
12. Wat betekent het als de beta coëfficiënt 1 is?
13. Hoe weet je of het rendement van een bedrijf heftiger reageert op de markt?
14. Uit welke onderdelen bestaat de required return op een investering?
15. Wat zijn extreme punten?
16. Wat zijn outlier punten?

## Hoofdstuk 14

1. Waarvoor worden goodness-of-fit toetsen gebruikt?
2. Welke variabele wordt er gebruikt bij een goodness-of-fit toets?
3. Welke kans beschrijft  $H_0$  bij een goodness-of-fit toets?
4. Hoe bereken je de vrijheidsgraden van een chi-kwadraat random variabele wanneer de populatie parameters geschat zijn?
5. Wat wordt er getoetst met de Jarque-Bera toets?
6. Uit welke twee onderdelen bestaat een Jarque-Bera toets?
7. Hoe formuleren we de nul-hypothese bij een chi-kwadraat random variabele voor contingency tabel?
8. Wat wordt er berekend met een sign toets?
9. Hoe formuleren we de nul-hypothese van een sign toets?
10. Wat is een nadeel van de sign toets?
11. Welke methode kan gebruikt worden om de grootte van het verschil te betrekken in de toets?
12. Wanneer benadert de Mann-Whitney U toets de normale verdeling?
13. Wat is de spearman rank correlatiecoëfficiënt?
14. Wat is de nul-hypothese van een runs toets?

# Antwoorden

## Hoofdstuk 1

1. De populatie is de complete set van items waar de onderzoeker in geïnteresseerd is. Een sample is het deel van de populatie waarover een steekproef wordt gedaan.
2. Bij simple random sampling worden objecten random gekozen. Bij systematic sampling wordt elk j-ste item uit de populatie geselecteerd.
3. Een parameter beschrijft een specifiek kenmerk van de populatie. Een statistiek beschrijft een specifiek kenmerk van een sample.
4. De populatie waarvan de steekproef is genomen is niet de relevante populatie. Deelnemers kunnen onnauwkeurige of geen eerlijke antwoorden geven. Deelnemers geven geen antwoorden.
5. Descriptive is het samenvatten en presenteren van data op een informatie manier. Inferential is het gebruiken van de gegevens om voorspellingen en schattingen te maken.
6. Discrete variabelen en continue variabelen.
7. Bij kwalitatieve gegevens is er geen meetbare betekenis tussen het verschil in aantallen en bij kwantitatieve gegevens wel.
8. Nominaal en ordinaal
9. Interval en ratio
10. Staafdiagram, kruis tabel, cirkeldiagram, pareto diagram
11. Frequentieverdeling, histogram, ogive, stamdiagram, scatterdiagram

## Hoodstuk 2

1. Het gemiddelde, de mediaan en de modus. Op vragen over de locatie van het centrum van een set gegevens.
2. Mediaan of modus
3. Gemiddelde of mediaan
4. Scheefheid is positief als de verdeling scheef naar rechts is, negatief als de verdeling scheef naar links is.
5. Dan is de scheefheid nul en is de verdeling klokvormig.
6. Deze verwijst naar de vijf beschrijvende maten: minimum, eerste kwartiel, mediaan, derde kwartiel en maximum.
7. Het derde kwartiel - het eerste kwartiel:  $Q3-Q1$ .
8. Box-and-whisker plot (boxplot)
9. De som van de gekwadraterde verschillen tussen elke waarneming en het gemiddelde, gedeeld door de steekproefomvang.
10. Chebyshey's theorem: Het aantal van de observaties in een steekproef die binnen  $k$  standaardafwijking van het gemiddelde afliggen. Z-score: een waarde die het aantal standaardafwijkingen aangeeft die een waarde van het gemiddelde af zit.
11. Beiden meten de richting van een lineaire relatie tussen twee variabelen. De correlatiecoëfficiënt geeft ook de sterkte van het verband tussen twee variabelen.

## Hoofdstuk 3

1. Basis uitkomsten zijn alle mogelijke uitkomsten van een random experiment. Alle basis uitkomsten bij elkaar noemen we de steekproefruimte. Het deel van de basis uitkomsten waar we in geïnteresseerd zijn noemen we een event.
2. Als de unie van een aantal events de hele steekproefomvang omvat.

Published on *WorldSupporter* ([www.worldsupporter.org](http://www.worldsupporter.org))

3. Bij permutaties wordt gelet op volgorde van elementen en bij combinaties niet.
4. Klassieke: hoe vaak een event zal plaatsvinden, er vanuit gaande dat alle uitkomsten in een steekproef gelijke kansen hebben om te gebeuren.  
Subjectieve waarschijnlijkheid: in hoeverre een individu gelooft in de kans dat een event zal gebeuren.
5. Complement regel, toevoegingsregel, voorwaardelijke kans, vermenigvuldigingsregel, statistische onafhankelijkheid.
6. Gezamenlijke kans, marginale kans, voorwaardelijke kans, overbetrokkenheid ratio

## Hoofdstuk 4

1.  $F(x_0)$  is tussen 0 en 1 voor elke  $x_0$ , de som van de individuele kansen is 1.
2. Het gemiddelde,  $\mu$ .
3. De wortel van de variantie.
4. Meerdere steekproeven waarbij steeds twee mutually exclusive uitkomsten mogelijk zijn. De kans op de uitkomst is hetzelfde voor elke proef. De kans op de uitkomst in een proef heeft geen invloed op de kans in andere proeven.
5. Succes en Falen
6.  $P(x,y)$  ligt tussen 0 en 1. De som van de gezamenlijke kansen  $P(x,y)$  van alle mogelijke paren moet 1 zijn.
7. Wanneer de covariantie 0 nul.
8. De lineaire combinatie van de gemiddelde waarden van de effecten in de portfolio.

## Hoofdstuk 5

1.  $P(a < X < b) = F(b) - F(a)$
2.  $F(x) > 0$  voor alle waarden van  $x$ . Het gebied onder de kansdichtheidsfunctie voor alle waarden van de random variabele is gelijk aan 1. De kans dat  $X$  ligt tussen  $a$  en  $b$  is het gebied onder de kansdichtheidsfunctie tussen deze punten. De cumulatieve verdelingsfunctie is het gebied onder de kansdichtheidsfunctie tot  $x_0$ .
3. Een symmetrische klokvormige kromme met het gemiddelde als centrum
4.  $X \sim N(\mu, \sigma^2)$
5.  $Z = \frac{X - \mu}{\sigma}$

## Hoofdstuk 6

1. Het is erg moeilijk om elk item in een populatie te meten en ook erg duur.
2. Als  $n$  groot is, de central limit theorem zegt dat de verdeling van  $Z$  de standaard normale verdeling nadert.
3. Een interval waarbinnen een steekproef gemiddelde een hoge kans van waarschijnlijkheid kent.
4. De proportie van de populatieleden die een karakteristiek bezitten waarin interesse is.

## Hoofdstuk 7

1. Een schatter is een random variabele die afhangt van de steekproef informatie. een schatting is een specifieke waarde van die random variabele.
2. Nul
3. De onpartijdige schatter met de kleinste variantie.
4.  $a < \theta < b$ , met  $100(1 - a)\%$

Published on *WorldSupporter* ([www.worldsupporter.org](http://www.worldsupporter.org))

5. Upper confidence limit en Lower confidence limit
6. Door de standaardafwijking naar beneden te brengen, de steekproefgrootte omhoog te brengen of een lager betrouwbaarheidslevel.
7. Als de populatie variantie onbekend is.

## Hoofdstuk 8

1. Als de waarden in de ene sample beïnvloedt worden door de waarden in een andere sample.
2. Het gemiddelde van de verschillen.
3. Dat X effectiever is dan Y.
4. De geobserveerde sample variaties  $s_x^2$  en  $s_y^2$ .
5. Door de pooled sample variantie  $s_p^2$ .
6. Bij grote samples.
7.  $(\hat{p}_x - \hat{p}_y) \pm ME$

## Hoofdstuk 9

1. De nul-hypothese wordt als waarheid beschouwd. Wanneer de nul-hypothese wordt verworpen dan wordt de alternatieve hypothese als waarheid beschouwd.
2. Eenzijdig wanneer de alternatieve hypothese groter of kleiner dan. Tweezijdig wanneer alternatieve hypothese niet gelijk aan.
3. Type I error: het verwerpen van een kloppende nul-hypothese  
Type II error: een foute nul-hypothese niet verwerpen
4. Dan wordt de nul-hypothese verworpen
5. Wanneer er sprake is van significantie, de p-waarde is  $< 0.05$ .
6. Door  $1 - \beta$

## Hoofdstuk 10

1.  $H_0: \mu_x - \mu_y = 0$      $H_1: \mu_x - \mu_y \neq 0$
2. Bij een tweezijdige toets wordt de alfa gedeeld door 2.
3. Met de F verdeling.
4.  $H_0: \sigma_x^2 = \sigma_y^2$      $H_1: \sigma_x^2 \neq \sigma_y^2$
5. De aanname dat de onderliggende verdeling normaal is of dat de central limit theorie van toepassing is.

## Hoofdstuk 11

1.  $b_1$  staat voor de helling en  $b_0$  staat voor de intersectie met de y-as.
2. Sum of squares total, sum of squares error en sum of squares regression.
3. Een hogere waarde staat voor een betere regressie
4. De afstand van de punten op de regressielijn en de totale afwijking van de X waarden van het gemiddelde.
5. Als de toekomstige waarde van de onafhankelijke variabele gegeven is.
6. Hoe wijder de interval, hoe groter de onzekerheid rondom de puntvoorspelling.
7. Hoe groter  $s_e^2$ , hoe groter de voorspelling interval en de betrouwbaarheidsinterval.  
Hoe groter de steekproefomvang  $n$ , hoe smaller de voorspelling interval en de

Published on *WorldSupporter* ([www.worldsupporter.org](http://www.worldsupporter.org))

betrouwbaarheidsinterval.

Hoe groter de dispersion, hoe meer informatie, hoe preciezer de schatting.

8. Diversifiable risico is het risico verbonden aan specifieke bedrijven en industrieën.  
Non diversifiable risico is het risico verbonden aan de gehele economie.
9. Werk conflicten, nieuwe concurrentie, consumentenmarkt veranderingen.
10. Door een groter portfolio en door het opnemen van waarvan het rendement een negatieve correlatie kent.
11. Hoe afhankelijk het rendement van een bepaald bedrijf is van het gehele markt rendement.
12. Dan volgt het rendement van een bedrijf precies de markt.
13. Als de beta meer dan 1 is.
14. Risk-free rate, beta for investment en market return.
15. Punten die X waarden hebben die substantieel afwijken van de X waarden van andere punten.
16. Punten die substantieel afwijken in de Y richting van de voorspelde waarde.

## Hoofdstuk 14

1. Om de populatie van nominale data te beschrijven.
2. De chi-kwadraat random variabele,
3. De kans dat een waarneming in elke categorie valt.
4.  $(K-m-1)$ .
5. De normaliteit van een verdeling.
6. De scheefheid en kurtosis.
7.  $H_0$ : er bestaat geen verband tussen twee karakteristieken in de populatie.
8. De kans dat het verschil negatief of positief is.
9.  $H_0: P=0.5$
10. Dat het maar een beperkt deel van de informatie verwerkt, alleen het teken van verschil.
11. De Wilcoxon Signed rank toets.
12. Wanneer elk sample op z'n minst 10 observaties bevat.
13. Als  $x_i$  en  $y_i$  zijn gerangschikt in oplopende volgorde en de sample correlatie van deze rangen wordt berekend, de resulterende coëfficiënt wordt de Spearman rank correlatiecoëfficiënt genoemd
14.  $H_0: the series is random$