

# Informatie tentamen Applied Statistics

Ook: instroomtoets Master Data Driven Business

Het tentamen Applied Statistics zal bestaan uit 40 vierkeuzevragen. Je krijgt hiervoor één uur de tijd. Voor toelating tot de Master Data Driven Business moet je minimaal 75% van de vragen, dus minimaal 30 vragen, goed beantwoorden.

## Stof

De stof voor het tentamen bestaat uit paragrafen uit het boek *Discovering Statistics using IBM SPSS Statistics* (Field, 2018) en de acht online statistiekmodules in Numworx (zie <https://www.uu.nl/onderwijs/open-en-gepersonaliseerd-statistiekonderwijs/lesmateriaal>).

Uit het boek zijn de volgende hoofdstukken/paragrafen tentamenstof:

- Hoofdstuk 1: helemaal
- Hoofdstuk 2: helemaal
- Hoofdstuk 3: paragrafen 3.1, 3.2 (t/m 3.2.2) en 3.7 (t/m 3.7.2)
- Hoofdstuk 4: alles behalve paragrafen 4.9 en 4.10
- Hoofdstuk 5: paragrafen 5.1 t/m 5.4 en 5.6 t/m 5.8 (t/m 5.8.1)
- Hoofdstuk 6: paragrafen 6.1 t/m 6.5
- Hoofdstuk 8: paragrafen 8.1 t/m 8.4, 8.7 en 8.8
- Hoofdstuk 9: paragrafen 9.1, 9.2, 9.7 en 9.8
- Hoofdstuk 10: alles behalve paragrafen 10.8.4, 10.8.5, 10.9.5 en 10.9.6
- Hoofdstuk 11: paragraaf 11.3.1 en 11.4.1
- Hoofdstuk 19: paragrafen 19.2 t/m 19.3.2, 19.7 en 19.8 (behalve 19.8.3)

## Vragen en leerdoelen per onderwerp

Onderwerp	Niveau en beoordelingscriterium			Totaal	Totaal %
	Reproductievragen (Criterium AS-1)	Begripsvragen (Criterium AS-2)	Toepassingsvragen (Criterium AS-3)		
Onderzoeksontwerp	1	2	1	4	10
Dataverzameling	2	2	2	6	15
Data-analyse: beschrijvende statistiek en visualisaties	2	4	2	8	20
Data-analyse: betrouwbaarheidsintervallen en statistische modellen	2	2	4	8	20
Data-analyse: hypothesetoetsen	2	2	4	8	20
Interpretatie en conclusies	1	3	2	6	15
<b>Totaal</b>	<b>10</b>	<b>15</b>	<b>15</b>	<b>40</b>	
<b>Totaal %</b>	<b>25</b>	<b>37,5</b>	<b>37,5</b>		<b>100</b>

De tabel hierboven geeft aan welke onderwerpen worden getoetst in het tentamen en hoeveel vragen er, naar schatting, voor elk onderwerp in het tentamen zullen zitten. Voor elk onderwerp zullen vragen van drie typen voorkomen:

- **Reproductievragen:** vragen naar het herkennen van elementen uit de kwantitatieve onderzoekscyclus

- **Begripsvragen:** vragen waarin je, in concrete organisatiecontexten, elementen uit de kwantitatieve onderzoekscyclus moet benoemen
- **Toepassingsvragen:** vragen waarin je, in concrete organisatiecontexten, elementen uit de kwantitatieve onderzoekscyclus moet uitvoeren.

In de tabel is uitgesplitst hoeveel reproductievragen, begripsvragen en toepassingsvragen er per onderwerp ongeveer in het tentamen voor zullen komen. Hieronder volgt een overzicht van de precieze stof, onderwerpen en leerdoelen voor elk van de zes onderwerpen in de tabel hierboven.

## Onderzoeksontwerp

Aantal vragen in tentamen (schatting): 1 reproductievraag, 2 begripsvragen, 1 toepassingsvraag

Stof boek: 1.1 t/m 1.5, 1.7, 2.1 t/m 2.6

Stof Numworx: module 1 en 5 (deel)

### Onderwerpen

- Onderzoeksproces, theorieën, hypotheses (1.1 t/m 1.5)
- Onderzoeksmethoden (1.7)
  - Experimenteel of correlatieel
  - Het beschrijven van (systematische en niet-systematische) variatie
  - Nut en effect van randomisatie
  - Opzet met afhankelijke of onafhankelijke groepen
- Belang van statistiek (2.1 t/m 2.6 en Numworx 1 en 5)
  - Idee van statistische modellen
  - Populatie & steekproef
  - Representatieve en aselechte steekproeven

### Leerdoelen

- Je weet wat het verschil is tussen een theorie en een hypothese, en welke rol theorieën en hypotheses in onderzoek spelen
- Je weet wat het verschil is tussen een probleem en een onderzoeksvraag.
- Je weet hoe correlatieel en experimenteel onderzoek van elkaar verschillen
- Je weet wat causaliteit inhoudt en wat schijnverbanden zijn
- Je weet wat systematische en niet-systematische variatie is
- Je weet wat randomisatie is, waarom het belangrijk is en welk effect het heeft op onderzoeksresultaten.
- Je kent het verschil tussen onafhankelijke en afhankelijke groepen in een onderzoeksontwerp
- Je weet waarvoor we statistische modellen maken en wat statistische modellen globaal beschrijven
- Je weet wat het beschrijven van een populatie op basis van een steekproef inhoudt en welke rol toeval en standaardfout hierbij spelen

## Dataverzameling

Aantal vragen in tentamen (schatting): 2 reproductievragen, 2 begripsvragen, 2 toepassingsvragen

Stof boek: 1.6, 1.8.6, 2.7

Stof Numworx: module 3 en 4

### Onderwerpen

- Variabelen (1.6)

- Onafhankelijk en afhankelijk
- Meetniveau
- Meetfouten, validiteit en betrouwbaarheid
- Verdelingen (1.8.6, Numworx 3)
- Normale verdeling (1.8.6, Numworx 4)
- Verdeling van het steekproefgemiddelde en standaardfout (2.7)

#### *Leerdoelen*

- Je weet wat een variabele is en welke rol variabelen spelen in onderzoek
- Je weet wat afhankelijke en onafhankelijke variabelen zijn
- Je kunt aangeven of variabelen van nominaal, ordinaal, interval of ratio meetniveau zijn
- Je kent mogelijke oorzaken van meetfouten
- Je weet wat validiteit is, wat betrouwbaarheid is en hoe deze met elkaar samenhangen.
- Je weet wat een verdeling is en welke rol verdelingen spelen in onderzoek
- Je weet wat de normale verdeling is en welke rol de normale verdeling speelt in onderzoek
- Je weet wat de verdeling van het steekproefgemiddelde is
- Je weet wat een standaardfout inhoudt, hoe je de standaardfout bepaalt en hoe je deze interpreteert

#### Data-analyse: beschrijvende statistiek en visualisaties

Aantal vragen in tentamen (schatting): 2 reproductievragen, 4 begripsvragen, 2 toepassingsvragen

Stof boek: 1.8, 1.9, 5.1 t/m 5.4, 5.6 t/m 5.8.1

Stof Numworx: module 2

#### *Onderwerpen*

- Centrummaten en spreidingsmaten (1.8 en Numworx 2)
  - Gemiddelde, mediaan, modus
  - Standaardafwijking, bereik, interkwartielafstand
- Visualisaties (5.1 t/m 5.4 en 5.6 t/m 5.8.1, Numworx 2)
  - Histogram
  - Staafdiagram
  - Boxplot
  - Spreidingsdiagram
  - (Kruis)tabellen

#### *Leerdoelen*

- Je weet hoe je de verschillende centrummaten (gemiddelde, mediaan, modus) kunt bepalen
- Je weet hoe je de verschillende spreidingsmaten (standaardafwijking, bereik, interkwartielafstand) kunt bepalen
- Je weet wat beschrijvende statistiek is, hoe dit verschilt van inferentiële/toetsende statistiek en bij welk onderzoeksdoel je welke vorm inzet.
- Je weet welke centrum- en spreidingsmaten je voor variabelen van verschillende meetniveaus kunt gebruiken
- Je weet welke invloed uitschieters op de verschillende centrum- en spreidingsmaten kunnen hebben
- Je weet hoe je verschillende visualisaties van data ((kruis)tabel, histogram, staafdiagram, boxplot en spreidingsdiagram) kunt aflezen
- Je weet welke visuele representaties je voor variabelen van verschillende meetniveaus kunt gebruiken

- Je weet hoe je in verschillende visualisaties centrum- en spreidingsmaten terug kunt zien

### Data-analyse: betrouwbaarheidsintervallen en statistische modellen

Aantal vragen in tentamen (schatting): 2 reproductievragen, 2 begripsvragen, 4 toepassingsvragen

Stof boek: 2.1 t/m 2.8, 8.1 t/m 8.4, 8.7 en 8.8, 9.1, 9.2, 9.7 en 9.8, 11.3.1 en 11.4.1

Stof Numworx: module 5 (deel) en 6

#### Onderwerpen

- Betrouwbaarheidsintervallen (2.8):
  - Berekenen
  - Interpretieren
  - Voor zowel populatiegemiddelde als andere parameters uit populatie
- Correlatie (8.1 t/m 8.4, 8.7 en 8.8)
  - Berekenen/bepalen
  - Interpretieren
  - Significantie beoordelen en interpreteren
  - Verschil Pearson/Spearman/Kendal
- Regressie (9.1, 9.2, 9.7 en 9.8):
  - Model bepalen en uitleggen
  - Regressiecoëfficiënten aflezen en interpreteren
  - Significantie model en regressiecoëfficiënten beoordelen en interpreteren
  - Voorspellingen doen met behulp van model
- Moderatoren en mediators (11.3.1 en 11.4.1)

#### Leerdoelen

- Je weet wat een betrouwbaarheidsinterval inhoudt
- Je weet wat de betrouwbaarheid van een betrouwbaarheidsinterval zegt en hoe deze betrouwbaarheid samenhangt met de nauwkeurigheid of precisie van het interval
- Je kunt een interpretatie geven van een betrouwbaarheidsinterval, zowel voor het populatiegemiddelde als voor andere parameters uit de populatie
- Je weet wat een correlatiecoëfficiënt inhoudt
- Je kunt zowel de richting als de sterkte van een correlatiecoëfficiënt interpreteren
- Je kunt de significantie van een correlatiecoëfficiënt bepalen en interpreteren
- Je kent het verschil tussen drie veelgebruikte correlatiecoëfficiënten: Pearsons  $r$ , Spearman's  $\rho$  en Kendal's  $\tau$ .
- Je kunt SPSS-output over correlatiecoëfficiënten lezen en interpreteren
- Je kunt de algemene formule van het lineaire regressiemodel (met één voorspellende variabele) uitleggen
- Je kunt regressiecoëfficiënten voor lineaire regressiemodellen met één voorspellende variabele aflezen en interpreteren
- Je kunt de significantie van een lineair regressiemodel en regressiecoëfficiënten beoordelen en interpreteren
- Je kunt de fit en het percentage verklaarde variantie van een lineair regressiemodel interpreteren
- Je kunt voorspellingen doen met behulp van een lineair regressiemodel met één voorspellende variabele
- Je kunt SPSS-output over regressiemodellen lezen en interpreteren
- Je weet wat moderatoren en mediators zijn

## Data-analyse: hypothesetoetsen

Aantal vragen in tentamen (schatting): 2 reproductievragen, 2 begripsvragen, 4 toepassingsvragen

Stof boek: 2.9, 2.10, Hoofdstuk 10, 19.2 t/m 19.3.2, 19.7, 19.8

Stof Numworx: module 7 en 8

### Onderwerpen

- Verschillende typen toetsen:
  - T-toets voor onafhankelijke groepen
  - T-toets voor afhankelijke groepen
  - Chi-kwadraattoets
  - Keuze voor meest geschikte toets
- Hypotheses opstellen en interpreteren
- Toets uitvoeren en resultaten interpreteren
- Een- en tweezijdig toetsen
- Significantieniveau en p-waarde
- Type-I en Type-II fouten en power

### Leerdoelen

- Je kunt de logische redeneerlijn van de procedure van hypothesetoetsen uitleggen
- Je weet wat een nulhypothese en alternatieve hypothese inhouden en hoe deze worden ingezet bij statistische toetsen
- Je weet wat statistische significantie inhoudt en hoe je het al dan niet ontbreken van statistische significantie bij een analyse kunt interpreteren
- Je weet wat een p-waarde is en wat het significantieniveau is
- Je kent het verschil tussen afhankelijke en onafhankelijke t-toetsen en kunt bij een gegeven onderzoeksopzet de meest geschikte toets kiezen
- Je kunt SPSS-output voor onafhankelijke en afhankelijke t-toetsen en chi-kwadraattoetsen lezen en interpreteren
- Je kunt de resultaten van onafhankelijke en afhankelijke t-toetsen en van chi-kwadraattoetsen interpreteren in de context van de variabelen
- Je weet wat eenzijdige en tweezijdige toetsing inhouden
- Je kunt op basis van meetniveaus van de onafhankelijke en afhankelijke variabelen vaststellen welke statistische toets geschikt is.
- Je weet wat Type-I en Type-II fouten zijn, hoe je de kansen hierop kunt beïnvloeden en hoe de kansen hierop samenhangen met de power van de toets.

## Interpretatie en conclusies

Aantal vragen in tentamen (schatting): 1 reproductievraag, 3 begripsvragen, 2 toepassingsvragen

Stof boek: 3.1, 3.2 (t/m 3.2.2) en 3.7 (t/m 3.7.2), 6.1 t/m 6.5, 10.8.6, 10.9.7,

### Onderwerpen

- Effectgrootte
- Bias
- Interpretatie

### Leerdoelen

- Je weet wat effectgroottes zijn en waarvoor deze worden gebruikt

- Je kunt de effectgroottes bij t-toetsen interpreteren
- Je weet wat bias in statistische analyses is
- Je kunt mogelijke oorzaken van bias benoemen (met name uitschieters en afwijkende verdelingen van data)
- Je kunt de resultaten van statistische analyses, inclusief effectgroottes en kanttekeningen rondom eventuele bias, interpreteren in de context van het probleem.