

Statistiek voor de financiële beroepen

Statistiek voor de financiële beroepen

Sven Janssen

Opmaak: Sander Pinkse Boekproductie, Amsterdam

ISBN 978 90 3723 458 9

© 2016 Uitgeverij Edu'Actief b.v.

Gehele of gedeeltelijke overneming of reproductie van de inhoud van deze uitgave, op welke wijze dan ook, zonder voorafgaande schriftelijke toestemming van de auteursrechthebbende is verboden, behoudens de beperkingen bij de wet gesteld. Het verbod betreft ook gehele of gedeeltelijke bewerking. De uitgever is met uitsluiting van ieder ander gerechtigd de door derden verschuldigde vergoedingen voor kopiëren, als bedoeld in artikel 17 Auteurswet 1912 en in het KB van 20 juni 1974 (Stb. 351, 1974) ex artikel 16b Auteurswet 1912, te innen en/of daartoe in en buiten rechte op te treden.

Correspondentie inzake overneming of reproductie richten aan:

Edu'Actief, Postbus 1056, 7940 KB Meppel, info@edu-actief.nl, 0522-235 235.

Inhoud

Voorwoord 7

HOOFDSTUK 1

Inleiding in de statistiek

- 1.1 Inleiding 9
- 1.2 Populatie en steekproef 10
- 1.3 Gegevens verzamelen 11
- 1.4 Gegevens verwerken 12
- Samenvatting 16
- Theorievragen 17
- Opgaven 19

HOOFDSTUK 2

Frequentieverdelingen

- 2.1 Frequenties 27
- 2.2 Absolute frequenties 27
- 2.3 Relatieve frequenties 28
- 2.4 Cumulatieve frequenties 29
- Samenvatting 31
- Theorievragen 32
- Opgaven 33

HOOFDSTUK 3

Diagrammen

- 3.1 Inleiding 39
- 3.2 Lijndiagram 39
- 3.3 Kolomdiagram 40
- 3.4 Cirkeldiagram 41
- Samenvatting 43
- Opgaven 44

HOOFDSTUK 4

Gegevensanalyse

- 4.1 Inleiding 49
- 4.2 Het rekenkundig gemiddelde 50
- 4.3 De modus 52
- 4.4 De mediaan 54
 - Samenvatting 59
 - Theorievragen 60
 - Opgaven 62

HOOFDSTUK 5

Indexcijfers

- 5.1 Inleiding 77
- 5.2 Enkelvoudige indexcijfers 78
- 5.3 Samengestelde indexcijfers 80
- 5.4 Consumentenprijsindexcijfer 83
 - Samenvatting 87
 - Theorievragen 89
 - Opgaven 91

Register 106

Voorwoord

In ons dagelijks leven worden we voortdurend geconfronteerd met statistieken. In principe zullen niet veel mensen die hun MBO-diploma behaald hebben, statistisch werk gaan verrichten. Echter, juist omdat het dagelijkse leven doordrenkt is met statistiek, is basale kennis van dit vakgebied essentieel. Het is als het ware basiskennis.

Wat precies deze basiskennis zou moeten zijn, staat beschreven in het eerder verschenen kwalificatiedossier en de hieruit afgeleide toetstermen. Dit boek is dan ook geschreven aan de hand van deze toetstermen. Voorheen werd dit onderdeel verwerkt in het boek ‘Rekenvaardigheid voor financiële toepassingen’. Inmiddels maakt het onderwerp ‘Statistiek’ geen onderdeel meer uit van de SPL-toetsmatrijzen. Ten onrechte, naar mijn idee. Vooral met het oog op de vervolgstudie van MBO-ers mag Statistiek niet ontbreken in de opleiding Financiële Beroepen.

Ik denk daarnaast dat het belangrijk is dat studenten niet alleen een aantal trucjes uit het hoofd leren, maar vooral ook inzicht krijgen in de statistiek. Op deze manier kun je hetgeen in dit boek behandeld wordt, ook vertalen naar de dagelijkse praktijk. Ik heb mijn best gedaan om een vak dat moeilijk gevonden wordt, zo uit te leggen dat het eenvoudiger wordt.

Roermond, mei 2016
Sven Janssen

Inleiding in de statistiek

1.1 Inleiding

Het woord statistiek zal misschien voor velen een negatieve associatie hebben. Jullie vinden het misschien moeilijk, saai, wetenschappelijk, et cetera. Desalniettemin is de statistiek of zijn statistieken bijzonder belangrijk. Wij worden namelijk elke dag geconfronteerd met statistieken. Statistieken zijn niets meer en niets minder dan schematisch weergegeven resultaten. Alle soorten cijfers of resultaten die je in je leven tegenkomt zijn een soort statistiek. We kunnen namelijk statistiek maken van bijna alles. Denk hierbij aan het aantal verkeersslachtoffers in een jaar, de economische groei in een bepaalde periode, het aantal verkochte schoenen in een bepaalde periode of bijvoorbeeld de stand in de eredivisie. Kortom, van alles waar gegevens van zijn, kunnen we statistieken maken.

Daarnaast vormt statistiek een vakgebied dat zich bezighoudt met het verzamelen, ordenen, verwerken, samenvatten en analyseren van die gegevens. Om de stand in de eredivisie vast te stellen, moeten we eerst gegevens verzamelen over de resultaten van de gespeelde wedstrijden. Vervolgens moeten we per team deze gegevens ordenen om ze te kunnen verwerken. Bij het verwerken gaan we het aantal punten van elk team tellen. Deze verwerking kun je terugvinden op teletekstpagina 818 van de publieke omroep. Pas als dit alles gebeurd is, kun je de gegevens analyseren. Dit analyseren gebeurt tot in den treuren bij alle voetbalpraatprogramma's.

Een belangrijke organisatie die zich met statistiek bezighoudt, is het CBS. CBS staat voor Centraal Bureau voor de Statistiek. Deze organisatie houdt van bijzonder veel aspecten statistieken bij zoals economische groei, werkloosheid, werkgelegenheid, bevolkingsgroei, sterfte, doodsoorzaken, geboortes, gemiddeld opleidingsniveau, aantal overvallen, migratie, et cetera. Wanneer je betrouwbare gegevens nodig hebt, kun je naar de website van het CBS gaan om die gegevens op te zoeken (www.cbs.nl).

1.2 Populatie en steekproef

Wanneer je ergens een statistiek van wilt maken, zul je duidelijk moeten maken waar het over gaat. Hiervoor is het van belang dat je aangeeft wat de populatie is die je onderzoekt. Een populatie is een verzameling personen of objecten. Een object kan dus alles zijn, behalve een persoon. Computers, tafels, printers, televisies, telefoons, bakstenen, lampen, glazen, kastanjes zijn dus voorbeelden van objecten.

Het bepalen van de populatie is overigens lang niet altijd even eenvoudig. De populatie ‘woningen’ is bijvoorbeeld erg onduidelijk. Wanneer is een gebouw namelijk een woning? En kunnen binnen één gebouw meerdere woningen onderscheiden worden? Als je (statistisch) onderzoek gaat doen, zal je dus eerst duidelijk moeten weten wat je precies gaat onderzoeken en wat de populatie is.

Bij veel onderzoeken onderzoek je niet de hele populatie, maar slechts een deel van die populatie. Dit noemen we een steekproef. Een goede steekproef is aselect. Dit wil zeggen dat de mensen of objecten die in de steekproef zitten, willekeurig gekozen zijn. Iedereen of alles binnen de populatie die je onderzoekt, moet een even grote kans hebben om onderzocht te worden. Pas als de kans voor alles of iedereen binnen een populatie even groot is om in de steekproef te worden opgenomen, is de steekproef representatief.

Daarnaast moet een steekproef voldoende groot zijn. Hoe groot een steekproef moet zijn, kan wiskundig bepaald worden. Dit hoeven jullie niet te kunnen. Wel moeten jullie kunnen bepalen of de steekproef voldoende groot zou kunnen zijn. Je moet je realiseren dat wanneer een populatie bestaat uit 100 miljoen mensen en er een steekproef genomen wordt onder 100 mensen, dit niet tot elkaar in verhouding staat. Hoe groter de groep personen of objecten in de steekproef, hoe eerlijker en representatiever de resultaten zullen zijn. Met eerlijk bedoelen we hier niet ‘oprecht’. Eerlijk in deze context geeft aan dat de resultaten uit steekproef overeen komen met de resultaten die je zou krijgen als je de hele populatie zou meten.

Dit klinkt allemaal misschien heel eenvoudig, maar het kan ook betrekkelijk ingewikkeld worden. Het volgende voorbeeld maakt dit duidelijk. Rondom de verkiezingen worden er vaak peilingen gehouden. Deze peilingen geven aan welke partijen kunnen rekenen op veel stemmen en welke partijen niet. Bij zo'n peilingonderzoek vormen alle stemgerechtigde Nederlanders de populatie. Natuurlijk wordt lang niet deze hele populatie gevraagd waar ze op zou stemmen. Er wordt een steekproef genomen waarbij slechts een deel van de stemge-

rechtigden gevraagd wordt waar ze op zou stemmen. De vraag is vervolgens hoe je die steekproef gaat uitvoeren. Als je namelijk op de markt mensen gaat vragen waar ze op zouden stemmen, kom je andere types mensen tegen dan wanneer je bij een supermarkt mensen gaat ondervragen. Daarbij komt nog dat er veel verschil is in soorten supermarkten. Ook is er veel verschil in stemgedrag tussen mensen die in steden wonen en mensen die in dorpen wonen. En daarnaast zie je dat mensen in bepaalde regio's anders stemmen dan mensen die in andere regio's wonen. In het noorden van het land wordt anders gestemd dan in het zuiden. Dit zijn nog maar een paar voorbeelden van verschillen die er kunnen ontstaan. Het doen van een volledig representatieve steekproef is dus bijzonder moeilijk.

Wanneer de steekproef eenmaal afgerond is, wordt de steekproef 'opgeblazen'. De steekproef wordt dan verrekend naar de gehele populatie. Als een populatie bestaat uit 10.000 mensen en de steekproef bestaat uit 100 mensen, dan zeggen de resultaten iets over de gehele populatie. Om terug te gaan naar de peilingen bij de verkiezingen: in een klein land met 10.000 kiesgerechtigden worden verkiezingen gehouden. Er zijn twee partijen: partij A en partij B. Bij de peiling worden 100 mensen ondervraagd. Van deze 100 mensen zeggen 52 mensen op partij A te stemmen en 48 mensen op partij B. Op basis hiervan wordt dan geconcludeerd dat 5200 mensen op partij A zullen stemmen en 4800 mensen op partij B. Let wel: deze conclusie mag je dus alleen trekken als de steekproef representatief is.

1.3 Gegevens verzamelen

Wanneer je statistisch onderzoek gaat doen, zul je gegevens moeten verzamelen. Soms zijn de gegevens die je nodig hebt al bekend. Dit kan gebeurd zijn door onderzoeken van bijvoorbeeld het CBS. Wanneer gegevens nog niet bekend zijn, maken onderzoekers vaak gebruik van een enquête. Het ontwikkelen van een goede enquête is niet bijzonder eenvoudig.

- De enquête moet niet te lang zijn
- De vragen moeten eenduidig zijn
- De vragen moeten relevant zijn
- De gegeven antwoorden moeten eenduidig zijn

Toelichting:

1. De enquête moet niet te lang zijn.

Als een enquête te lang is, zullen er minder mensen reageren. Mensen zul-

len het te veel moeite vinden om mee te doen. Wanneer te weinig mensen reageren, is het resultaat dat uit de enquêtes komt niet meer representatief.

2. De vragen moeten eenduidig zijn.
Het moet voor iedereen duidelijk zijn wat er bedoeld wordt met een vraag. Een vraag mag niet op twee manieren geïnterpreteerd worden. Wanneer mensen een vraag namelijk op verschillende manieren geïnterpreteerd wordt, krijg je uiteenlopende antwoorden die niet met elkaar vergeleken kunnen worden. Dit zou betekenen dat de resultaten uit de enquête geen antwoord geven op datgene wat je wilt onderzoeken.
3. De vragen moeten relevant zijn.
Alle vragen in een enquête moeten toewerken naar hetgeen je wilt onderzoeken. Stel geen vragen die niet direct betrekking hebben op je onderzoek. Dit leidt mensen alleen maar af en zorgt voor onduidelijkheid.
4. De gegeven antwoorden moeten eenduidig zijn.
Niet alleen de vragen, maar ook de antwoorden moeten helder en duidelijk zijn. Wanneer mensen zelf antwoorden gaan formuleren, loop je het risico dat de antwoorden niet duidelijk zijn. Daarom worden er vaak meerkeuzevragen gesteld. Mensen hoeven dan alleen maar een kruisje of cirkeltje zetten om het antwoord dat volgens hen het meest juist is.

De antwoorden die uit zo'n onderzoek rollen zijn de gegevens die verwerkt moeten worden. Deze gegevens noemen we variabelen. Er zijn twee soorten variabelen:

- Kwantitatieve variabelen
Kwantiteit is een ander woord voor hoeveelheid. Bij kwantitatieve gegevens/variabelen gaat het dus om aantallen. De resultaten kun je dus uitdrukken in getallen.
- Kwalitatieve variabelen
Kwalitatieve variabelen kun je niet in getallen uitdrukken, maar in kenmerken. Denk hierbij aan de kleur van een auto, of een oordeel: goed, voldoende, onvoldoende.

1.4 Gegevens verwerken

Het verwerken van gegevens doen we doorgaans in tabellen en grafieken. Voor zowel die tabellen als voor die grafieken zijn richtlijnen opgesteld. Het is name-

lijk de bedoeling dat iemand die totaal niet weet waar het onderzoek over gaat, toch kan begrijpen wat je met de tabel of de grafiek bedoelt.

De grafieken behandelen we in hoofdstuk 3. We hebben hiervoor gekozen omdat er verschillende soorten grafieken/diagrammen bestaan. Hier dient in een apart hoofdstuk bij te worden stilgestaan. Deze paragraaf gebruiken we als een inleiding van het maken van tabellen. In hoofdstuk 2 gaan we hier verder op in.

Een tabel bestaat uit kolommen en rijen/regels. Een plek waarop een kolom en een regel elkaar kruisen, noemen we een veld. Een veld is dus een ‘hokje’ waarin we een gegeven kunnen plaatsen. Om ervoor te zorgen dat een tabel voor een buitenstaander duidelijk is, is er een aantal richtlijnen bedacht:

- Titel. Iedere tabel hoort een titel te hebben die kort aangeeft waar de tabel over gaat.
- Kolomentitels; de kolommen zijn verticaal. Boven elke kolom dient kort te worden weergegeven welke informatie er uit die kolom te halen valt.
- Regeltitels. In de eerste (meest linkse) kolom, staan de titels van de regels. Dit zorgt ervoor dat ook duidelijk is waar elke afzonderlijke regel over gaat.
- Logische volgorde. De volgorde van de rijen en kolommen dient zo logisch mogelijk te zijn.
- Totalen. Het kan nuttig zijn om op de laatste regel en of laatste kolom de totalen van die regel of kolom te noteren.
- Bronvermelding. Als de gegevens uit de tabel van een externe bron zijn, dienen we dit te vermelden.
- Tekens. Binnen tabellen gebruiken we onder andere de volgende tekens:
 - * Bijzondere aanduiding. Een veld met een sterretje wordt vaak nog extra toegelicht.
 - X Gegeven is onbekend
 - Het getal is exact gelijk aan 0
 - o Het getal is na afronding 0, maar is zonder afronding niet gelijk aan 0

Hieronder volgen twee voorbeelden van tabellen.

Pieters BV is een onderneming met twee vestigingen: Roermond en Eindhoven. Om investeringen mogelijk te maken, heeft Pieters BV een lening bij de bank nodig. De bank eist echter onder andere de omzetcijfers over de afgelopen twee jaar op om te kunnen beoordelen of Pieters BV de lening krijgt. Pieters BV levert daarop onderstaand overzicht in.

Omzetten Pieters BV (* € 1000)				
Maanden	Omzetcijfers			
	Roermond		Eindhoven	
	2011	2012	2011	2012
Januari	35	34	39	38
Februari	39	36	41	42
Maart	40	40	44	44
April	43	45	51	50
Mei	49	50	54	56
Juni	52	54	59	59
Juli	52	56	61	60
Augustus	55	55	63	62
September	51	50	58	59
Oktober	47	46	52	54
November	40	42	44	47
December	38	39	44	43
Totaal	541	547	610	614

Figuur 1.1

Toelichting

Bovenstaande tabel heeft een titel: het gaat om de omzet van het bedrijf Pieters BV. Daarnaast staat in de titel dat de getallen euro's betreffen en dat elk getal met 1.000 vermenigvuldigd moet worden. Dit betekent dus dat wanneer iemand vraagt naar de omzet van januari 2011 van Roermond, het antwoord niet € 35 maar € 35.000 is. Voor Roermond, februari 2011 noteren we voluit dus € 39.000.

De kolommen hebben dusdanige titels gekregen dat een volstrekte buitenstaander kan begrijpen wat de tabel weergeeft. Zonder voorkennis kun je op basis van bovenstaande tabel van beide filialen over de afgelopen twee jaren de omzetten bekijken per maand, maar ook de totale omzet is af te lezen.

Een ander soort tabel volgt hieronder. Op basis van een uitdraai van de salarisadministratie van Van Eckeren BV is een tabel samengesteld van de inkomensverdeling van het personeel van de onderneming.

Inkomensverdeling Van Eckeren BV naar geslacht						
Brutoloon (€)	Geslacht			Geslacht, procentueel		Totaal
	Man	Vrouw	Totaal	Man	Vrouw	
€ 0 - < € 1000	6	12	18	2,7%	5,5%	8,2%
€ 1000 - < € 1500	13	12	25	5,9%	5,5%	11,4%
€ 1500 - < € 2000	28	25	53	12,7%	11,4%	24,1%
€ 2000 - < € 2500	35	37	72	15,9%	16,8%	32,7%
€ 2500 - < € 3000	21	18	39	9,5%	8,2%	17,7%
€ 3000 en hoger	9	4	13	4,1%	1,8%	5,9%
Totaal	112	108	220	50,9%	49,1%	100,0%

Figuur 1.2

Toelichting

Een verschil tussen deze tabel en de tabel uit figuur 1.1 is dat de klassenindeling anders is. In figuur 1.1 waren de klassen de maanden (januari, februari, et cetera). Nu zijn de klassen niet een enkel getal of gegeven, maar is er een klassenindeling gemaakt. Zo zijn er in totaal 18 mensen die een brutomaandloon hebben tussen de € 0 en € 1.000. Dit doen we om de tabel overzichtelijk te houden. Het is, zeker als er sprake is van een grote populatie, nagenoeg onmogelijk om een overzichtelijke tabel te maken van alle afzonderlijke inkomens. Je zou dan in bovenstaand voorbeeld van alle 220 personen afzonderlijk moeten aangeven hoe hoog zijn of haar brutomaandloon is. Om die reden wordt er in soortgelijke gevallen meestal voor gekozen om een groepen te maken.

Een term die bij bovenstaand verhaal belangrijk is, is klassenbreedte. De klassenbreedte is letterlijk wat het woord zegt: de breedte van de klasse. De eerste klasse is dus € 1.000 breed, de tweede klasse is € 500 breed, et cetera. De laatste klasse kent geen klassenbreedte. De laatste klasse is een zogenoemde open klasse.

Daarnaast zie je dat er in deze tabel ook nog kolommen zijn opgenomen voor de procentuele verhoudingen. Zo zijn de 6 mannen die tussen de € 0 en de € 1000 verdienen 2,7% van het totale personeel ($6 / 220 \times 100\% = 2,7\%$). Het is op deze manier dus ook logisch dat het veld helemaal rechtsonder het totaal op 100% staat.

Samenvatting

Statistiek is de wetenschap die zich bezighoudt met het verzamelen, ordenen, verwerken en analyseren van gegevens. Als je statistisch onderzoek gaat doen, is het van belang te weten welke populatie je gaat onderzoeken. Een populatie is een verzameling personen of objecten. Bij een statistisch onderzoek wordt vaak niet de hele populatie onderzocht. Er wordt dan een steekproef genomen. Een steekproef hoort aselekt te zijn. Letterlijk betekent dit niet select of ook wel willekeurig. Daarnaast moet het aantal mensen waaronder een steekproef gehouden wordt, in een redelijke verhouding staan tot de totale populatie. Als de steekproef voldoende groot is en aselekt is, is de steekproef representatief. Hoe groter de steekproef, hoe eerlijker de uitkomst van die steekproef zal zijn.

Om gegevens te verzamelen voor een statistisch onderzoek, worden er vaak enquêtes afgenomen. Bij het ontwikkelen van een enquête, is het belangrijk om op een aantal zaken te letten;

- De enquête moet niet te lang zijn
- De vragen moeten eenduidig zijn
- De vragen moeten relevant zijn
- De gegeven antwoorden moeten eenduidig zijn

Gegevens worden doorgaans verwerkt in tabellen en grafieken. Beide dienen zo te worden weergegeven dat een buitenstaander begrijpt wat er met de tabel of de grafiek bedoeld wordt. Om dit te realiseren, zijn er richtlijnen opgesteld om een tabel mee in te richten.

Theorievragen

1. Wat is de definitie van statistiek?

.....

.....

2. Ga naar de website van het CBS en zoek op hoeveel inwoners Nederland heeft, hoe groot de economische groei is en hoe groot de werkloosheid is.

.....

.....

3. Wat is een populatie?

.....

.....

4. Geef vier voorbeelden van mogelijke populaties die niet in het hoofdstuk genoemd worden.

.....

.....

.....

.....

5. Wat is een steekproef?

.....

.....