**Samenvatting beschrijvende statistiek**: formules, definities,werkwijzen en eigenschappen

VERZAMELINGEN

**Kardinaalgetal** = het aantal elementen van een verzameling

**Machtsverzameling** = de verzameling van alle mogelijke deelverzamelingen

: doorsnede  
: unie  
: verschil  
: complement  
: kardinaalgetal  
: deelverzameling  
: machtsverzameling  
: productverzameling

**Eigenschappen van verzamelingen:**  
A⋂B = B⋂A  
A⋂(B⋂C) = (A⋂B) ⋂C  
A⋂∅ = ∅  
A⋂(B⋃C) = (A⋂B) ⋃(A⋂C)  
(A⋂B)c = Ac⋃BcA⋃B = B⋃A  
A⋃(B⋃C) = (A⋃B) ⋃C  
A⋃∅ = A  
A⋃(B⋂C) = (A⋃B) ⋂(A⋃C)  
(Ac)c = A  
(A⋃B)c = Ac⋂Bc

**Partitie** = opsplitsing van een verzameling in een stel niet-lege en niet-overlappende deelverzamelingen

**Productverzameling** = verzameling van alle geordende koppels

**Eigenschappen van het sommatieteken:**

ℕ: de natuurlijke getallen (0,1,2,3, …)  
ℤ: de gehele getallen (…, -2, -1, 0, 1, 2, …)  
ℚ: breuken  
ℝ: de reële getallen  
∅: lege verzameling  
𝒰: universele verzameling

BESCHRIJVENDE STATISTIEK

**Kwalitatieve variabele** = het bereik bestaat uit een aantal waarden (categorieën) waarover geen verdere claims worden gemaakt

**Kwantitatieve variabele** = het bereik bestaat uit numerieke waarden waarvoor ordening, optellen en aftrekken zinvol is

p(xj) =

**Lijndiagram, frequentiefunctie:**  
Y-as = frequentie  
X-as = X  
Lijntjes tekenen

**Staafdiagram, frequentiefunctie:**  
Y-as = frequentie  
X-as = X  
Staven tekenen (met spaties, WANT geen histogram)

**Histogram, proportiefunctie:**  
Y-as = proportie  
X-as = X  
staven tekenen (geen spatie, WANT geen staafdiagram)

**Polygoonvoorstelling:**  
🡪 middelpunten van de bovenste lijnstukken van de histogramstaven met elkaar verbinden

**Cumulatieve frequentie** = = het aantal observaties met waarde kleiner of gelijk aan Xi

|  |  |  |
| --- | --- | --- |
| naam | notatie | definitie |
| percentielen | Pc1 … Pc99 | x.01 x.02 x.03 … x.99 |
| decielen | D1 … D9 | x.1 x.2 x.3 … x.9 |
| kwartielen | Q1 … Q3 | x.25 x.50 x.75 |

**Cumulatieve proportie** =

**Cumulatieve proportiefunctie:**  
Y-as = F  
X-as = X  
Lijnen tekenen (gekleurde bol, lege bol)  
Laten doorlopen tot het oneindige

**r-de kwantiel zoeken:**  
GEVAL 1: geen xi geobserveerd met F(xi) = r  
🡪 kiezen voor het eerstvolgende getal 🡪 dat is dan je r-de kwantiel  
GEVAL 2: er is een xi geobserveerd met F(xi) = r  
🡪 het gemiddelde van dat getal en het eerstvolgende getal

**Klassegrenzen** = xi en xh  
**Klassemiddelpunt** =   
**Histogram met klassen:**  
Y-as = F  
X-as = X  
X is telkens het klassemiddelpunt van die klasse

**Gegroepeerde frequentietabel:**  
🡪 stam- en loofdiagram

**Modus** = elke waarde waarvoor freq(x) maximaal is

1 modus = unimodaal  
2 modi = bimodaal

**Mediaan** = (Mex): Pc50 = D5 = Q2

**Mediaan zoeken:**  
🡪orden de observaties naar grootte-orde en hernummer ze  
🡪als n oneven is:   
🡪als n even is: =

**Eigenschappen van het gemiddelde:**  
   
Als c dan   
Regel van Steiner:

**Eigenschap van mediaan:**  
Als c dan

**Bereik**: verzamelingen van alle functiewaarden

**Bereik**  
max(x) – min(x)

**Interkwartielbereik**  
Q3 – Q1

**Variantie**: de gemiddelde kwadratische afstand van de observaties t.o.v. hun gemiddelde

**Variantie**

**Standaarddeviatie**: een maat voor de spreiding van een variabele of van een verdeling

**Standaarddeviatie**

**Berekening van :**  
METHODE 1: berekenen per observatie en maat direct gebruik van de bovenstaande definities  
METHODE 2: maak gebruik van de volgende gelijkheid:  
   
 Bereken daartoe per observatie enz…  
METHODE 3: Vertrek van de frequentie- of proportietabel en maak gebruik van de formules  
 OF   
 OF

**Eigenschappen van variantie:**  
Eigenschap 1:  
 = het gemiddelde kwadratische verschil tussen alle paren observaties  
   
Eigenschap 2: ongelijkheid van Tchebychev  
 als dan   
 equivalent met als dan

**Boxplot:**Zie boek statistiek p.44-45

**Z-transformatie**

**Invloed van transformaties op frequentiefuncties:**  
als y = f(x)  
dan  
  
 f(X)=Y

**Invloed van transformaties op centrale tendensmaten:**  
Als y = f(x) met f(x) = ax + b  
dan  
  
Bijzonder geval:

**Invloed van transformaties op spreidingsmaten:**  
Stel y = f(x) met f(x) = ax + b  
dan  
  
Bijzonder geval:  
  
Gevolg:   
 Zx als a 0  
ZaX + b=  
 -Zx als a 0

BESCHRIJVENDE STATISTIEK MET 2 VARIABELEN

**Bivariate frequentiefuncties 🡪 bivariate frequentietabel**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Aantal anderen | | | | | | | | | | | | | | | | |  |
| Plaats | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | FreqX |
| Thuis | 4 | 5 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 15 |
| Werk | 1 | 3 | 2 | 1 | 1 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 |
| Buitenshuis | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 4 |
| FreqY | 7 | 8 | 4 | 5 | 1 | 3 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 30 |

**Bivariate proportiefuncties 🡪 bivariate proportietabel**  
Hetzelfde als bovenstaand voorbeeld maar dan met proporties

**Lijndiagrammen**  
Per plaats een andere lijndiagram

**Rug-aan-rug stam-en-loofdiagram**  
Hetzelfde als een stam-en-loofdiagram maar dan rug aan rug, 2 variabelen

**Rijconditionele proportie**  
=

**Kolomconditionele proportie**  
=

**Scatterdiagram**  
Zie boek p.61-62

**Conditionele gemiddelde**

**Conditionele variantie**

**Covariantie**

Zie verder boek p.67

**Eigenschappen covariantie**

**Correlatie**

**Eigenschappen correlatie**

**Algemene optimale voorspelling**

**Gekwadrateerde standaardfout van estimatie**

**Optimale lineaire voorspelling**  
b0 is de regressieconstante  
b1 is het regressiegewicht van X

**Somvariabelen**

**Oefeningen werkwijze:**  
-Eerst somvariabelen eigenschappen  
-Daarna lineaire transformaties

BESCHRIJVENDE STATISTIEK MET MEER DAN 2 VARIABELEN