

## **4 Domein STATISTIEK - versie 1.2**

(Op initiatief van USolv-IT werd deze boomstructuur mede in overleg met het Universitair Centrum voor Statistiek (K.U.Leuven) afgesproken.)

### **4.1 Beschrijvende statistiek**

#### **4.1.1 Type data**

**Discrete - continue variabele**

**Nominaal - ordinaal - interval - ratio**

#### **4.1.2 Frequentietabellen**

**Eén dimensionaal**

**Kruistabellen**

1. Gezamenlijke frequentie
2. Marginale frequentie
3. Voorwaardelijke frequentie
4. Associatiematen

#### **4.1.3 Visualisatie**

**Staafdiagram**

**Taartdiagrammen**

**Stamdiagrammen (stem-and-leaf plot)**

**Histogram**

1. Op basis van gegroepeerde gegevens
2. Op basis van ruwe gegevens

**Empirische verdelingsfunctie**

**Boxplot**

**Puntenwolk**

**Bubble plot**

## **Tijdsreeks**

### **4.1.4 Numerische maten**

#### **Centrummaten**

1. Gemiddelde
2. Mediaan
3. Modus

#### **Spreidingsmaten**

1. Range
2. Variantie
3. Standaarddeviatie
4. Interkwartielafstand
5. Variatiecoëfficiënt
6. Gemiddelde absolute afwijking

#### **Andere numerische maten**

1. Kwantielen
2. Gewogen gemiddelde
3. Meetkundig gemiddelde
4. Scheefheid
5. Covariantie
6. Correlatie

### **4.1.5 Standaardiseren**

### **4.1.6 Aanpassen van een lijn aan gegevens**

1. Lineaire functie
2. Logaritmische functie
3. Exponentiële functie
4. Machtsfunctie
5. Andere functies

#### **4.1.7 Economische toepassingen**

1. Lorenzkromme
2. Gini-coëfficiënt

### **4.2 Kansrekenen**

#### **4.2.1 Kans**

1. Het begrip kans
2. Combinatoriek en telproblemen (in kansvraagstukken)

#### **4.2.2 Kansruimte**

#### **4.2.3 Voorwaardelijke kans**

1. Definitie en eigenschappen
2. Onafhankelijkheid
3. Stelling van Bayes

#### **4.2.4 Kansverdelingen univariaat**

##### **Discrete kansverdelingen**

1. Kansdichtheid (kansfunctie)
2. Cumulatieve verdelingsfunctie
3. Verwachtingswaarde en variantie
  - Verwachtingswaarde
  - Variantie en standaardafwijking
4. Momentgenererende functie

##### **Specifieke discrete kansverdelingen**

1. Uniforme verdeling
2. Bernoulli verdeling
3. Binomiale verdeling
4. Poisson verdeling
5. Geometrische verdeling
6. Hypergeometrische verdeling
7. Negatief binomiale verdeling

### **Continue kansverdelingen**

1. Kansdichtheid (dichtheidsfunctie)
2. Cumulatieve verdelingsfunctie
3. Verwachtingswaarde en variantie
  - Verwachtingswaarde
  - Variantie en standaardafwijking
4. Momentgenererende functie

### **Specifieke continue kansverdelingen**

1. Uniforme verdeling
2. Normale verdeling
3. t-verdeling
4. Exponentiële verdeling
5. Gamma verdeling
6. Chi-kwadraatverdeling
7. F-verdeling
8. Logistische verdeling
9. Beta-verdeling
10. Pareto-verdeling

### **Transformaties**

#### **4.2.5 Centrale limietstelling**

#### **Centrale limietstelling**

#### **Normale benadering voor binomiale kansen**

#### **4.2.6 Kansverdelingen (multivariaat)**

#### **Discrete kansverdelingen**

1. Gezamenlijke kansdichtheid
2. Gezamenlijke kansverdeling
3. Som, verschil, product, transformaties
4. Covariantie, correlatie

## **Continue kansverdelingen**

1. Gezamenlijke kansdichtheid
2. Gezamenlijke kansverdeling
3. Som, verschil, product, transformaties
4. Covariantie, correlatie
5. De bivariaat normale verdeling

## **4.3 Inferentie**

### **4.3.1 Gegevens verwerven: steekproeftrekkingen, experimenten**

### **4.3.2 Inferentie voor gemiddelde**

#### **Betrouwbaarheidsinterval**

1. Bij gekende standaardafwijking
2. Bij ongekende standaardafwijking

#### **Hypothesetoetsen**

1. Parametrisch, bij gekende standaardafwijking
  - Hypothesetest en assumpties (normaliteit)
  - Power
  - Steekproefgrootte
2. Parametrisch, bij ongekende standaardafwijking
  - Hypothesetest en assumpties (normaliteit)
  - Power
3. Niet-parametrisch
  - Hypothesetest
  - Power
  - Steekproefgrootte

### **4.3.3 Inferentie van 2 gemiddeldes**

#### **Betrouwbaarheidsinterval**

1. Onafhankelijke steekproeven
2. Gepaard

## Hypothesetoetsen

### 1. Parametrisch

- Onafhankelijke steekproeven - gelijke varianties
  - Hypothesetest en assumpties
  - Power
  - Steekproefgrootte
- Onafhankelijke steekproeven - ongelijke varianties
  - Hypothesetest en assumpties
  - Power
  - Steekproefgrootte
- Gepaard
  - Hypothesetest en assumpties
  - Power
  - Steekproefgrootte

### 2. Niet-parametrisch

- Niet gepaard
  - Hypothesetest en assumpties
  - Power
  - Steekproefgrootte
- Gepaard
  - Hypothesetest en assumpties
  - Power
  - Steekproefgrootte

#### 4.3.4 Inferentie voor 1 variantie, standaardafwijking

##### Betrouwbaarheidsinterval

##### Hypothesetoets

#### 4.3.5 Inferentie voor 2 varianties, standaardafwijkingen

##### Betrouwbaarheidsinterval

##### Hypothesetoets

#### 4.3.6 Inferentie voor 1 proportie

##### Betrouwbaarheidsinterval

### **Hypothesetoetsen**

1. Hypothesetest
2. Power
3. Steekproefgrootte

#### **4.3.7 Inferentie voor een verschil van proporties**

##### **Betrouwbaarheidsinterval**

### **Hypothesetoetsen**

1. Hypothesetest
2. Power
3. Steekproefgrootte

#### **4.3.8 Inferentie voor één correlatie**

#### **4.3.9 Inferentie voor kruistabellen**

##### **Odds en odds ratio**

##### **Chi-kwadraat toets**

##### **Fisher exact test**

#### **4.3.10 Inferentie voor verdelingen**

##### **Normale verdeling**

1. Chi kwadraat aanpassingstoets
2. Kolmogorov-Smirnov en Shapiro Wilk test
3. QQ-plot
4. Transformatie van verdeling

##### **Andere verdelingen**

1. Chi kwadraat aanpassingstoets
2. Kolmogorov-Smirnov test
3. QQ-plot
4. Transformatie van verdeling

## **4.4 Lineaire regressie**

### **4.4.1 Enkelvoudige lineaire regressie**

#### **Hypothesetoetsen**

1. F-toets

#### **Schatting van de parameters**

#### **Goodness of fit**

1.  $\mathbb{R}^2$

#### **Betrouwbaarheidsintervallen en hypothesetesten voor parameters**

#### **Voorspellingsintervallen**

#### **Diagnostiek van de residu's, nagaan van onderstellingen**

1. Normale QQ plot
2. Residu plot
3. Cook's distance (D)
4. Leverage points (h)
5. Influential points, Dfbetas
6. Outlier
7. Test seriele correlatie (Durbin Watson)
8. Test constante variantie (Levene)

### **4.4.2 Meervoudige lineaire regressie**

#### **Hypothesetoetsen**

1. F-toets

#### **Schatting van de parameters**

#### **Goodness of fit**

1.  $\mathbb{R}^2$
2. adjusted  $\mathbb{R}^2$

## Betrouwbaarheidsintervallen en hypothesetesten voor parameters

### Voorspellingsintervallen

#### Diagnostiek van de residu's, nagaan van onderstellingen

1. Normale QQ plot
2. Partial Residual plot
3. Outliers
4. Leverage points (h)
5. Cook's distance (D)
6. Influential points, Dfbetas

#### Multicollineariteit

1. Correlatie
2. VIF

#### Selectie van variabelen

1. Stepwise linear regression
2. Mallows  $C_p$
3. R-kwadraat

#### Interactie tussen variabelen

#### 4.4.3 Polynomiale regressie

#### Hypothesetoetsen

1. F-toets

#### Schatting van de parameters

#### Goodness of fit

1.  $R^2$

## Betrouwbaarheidsintervallen en hypothesetesten voor parameters

### Voorspellingsintervallen

### **Diagnostiek van de residu's, nagaan van onderstellingen**

1. Normale QQ plot
2. Partial Residual plot
3. Outliers
4. Leverage points (h)
5. Cook's distance (D)
6. Influential points, Dfbetas

### **Multicollineariteit**

### **Interactie tussen variabelen**

## **4.5 Variantie analyse**

### **4.5.1 1-factor variantie analyse gelijke sample sizes**

#### **anova tabel**

1. F-toets
2. hypothesetesten

#### **comparisons**

1. pairwise comparisons
2. contrasts
3. multiple comparisons
  - Tukey
  - Scheffe
  - Bonferroni
  - Dunnett

### **Diagnostiek van de residu's**

1. Residual plot
2. Normale QQ plot
3. Test constante variantie (Levene)

### **Power van de test**

## **Gevolgen van afwijkingen van de assumpties**

### **4.5.2 2-factor variantie analyse gelijke sample sizes**

#### **anova tabel**

1. F-toets
2. hypothesetesten hoofdeffecten
3. hypothesetesten interactie

#### **comparisons op significante interactie**

1. pairwise comparisons
2. contrasts
3. multiple comparisons
  - Tukey
  - Scheffe
  - Bonferroni
  - Dunnet

#### **comparisons op hoofdeffecten**

1. pairwise comparisons
2. contrasts
3. multiple comparisons
  - Tukey
  - Scheffe
  - Bonferoni
  - Dunnet

#### **Diagnostiek van de residu's**

1. Residual plot
2. Normale QQ plot
3. Test constante variantie (Levene)

#### **Power**

### **4.5.3 2-factor variantie analyse ongelijke sample size**

#### **anova tabel**

1. F-toets
2. Hypothesetesten hoofdeffecten
3. Hypothesetesten interactie

#### **comparisons op significante interactie**

1. pairwise comparisons
2. contrasts
3. multiple comparisons
  - Bonferroni

#### **comparisons op hoofdeffecten**

1. pairwise comparisons
2. contrasts
3. multiple comparisons
  - Bonferroni

#### **Diagnostiek van de residu's**

1. Residual plot
2. Normale QQ plot
3. Test constante variantie (aangepaste Levene)

## **4.6 Logistieke regressie**

### **4.6.1 Binaire respons variabele**

#### **Schatting van de parameters**

#### **Goodness of fit**

1. Partial Deviance Test

#### **Diagnostiek van de residu's**

## **4.6.2 Multinomiale respons variabele**

### **Schatting van de parameters**

#### **Goodness of fit**

1. Deviance test
2. Partial deviance test

#### **Diagnostiek van de residu's**