# Combinatieleer:

**P = Wat je kiest  
N = Totaal mogelijkheden**

## Permutatie:

Gebruikt voor:  
Een permutatie van n elementen is een geordend n – tal van verschillende elementen gekozen uit een gegeven verzameling van n elementen.

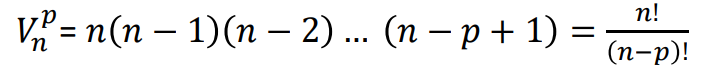
**A picture containing text

Description automatically generated**

Voorbeeld: 9! = 9\*8\*7\*6\*5\*4\*3\*2\*1 = 362 880  
Voorbeeld: 6! = 6\*5\*4\*3\*2\*1 = 720

## Variaties:

Gebruikt voor:  
Een variatie van p elementen uit n elementen is een geordend p – tal van p verschillende elementen gekozen uit de gegeven verzameling van n elementen.



Voorbeeld:

## Herhalingsvariaties

Gebruikt voor:  
Een herhalingsvariatie van p elementen uit n elementen is een geordend p – tal van elementen gekozen uit een gegeven verzameling van n elementen; waarbij hetzelfde element meermaals gekozen mag worden.

Text

Description automatically generated

Voorbeeld: Var 30 uit 4 = 4^30 = 1 152 921 505

Voorbeeld: var 2 uit 6 = 6^2 = 36

Volgorde van belang?

Herhalingsvariatie

Combinatie

Herhaling mogelijk?

Variatie (p n)

Permutatie (p = n)

# Kansrekening:

*A = deelverzameling (vb: dobbelsteen even getal)*

*U = Universum*

*Complement* ***van U = de lege verzameling***

## Definitie: voorwardelijk kans

De voorwaardelijke kans van een gebeurtenis A, gegeven dat een gebeurtenis B zich voorgedaan heeft, wordt genoteerd als P(A B) en wordt berekend als volgt:

Text

Description automatically generated

## Eigenschap: Productregel voor onafhankelijke gebeurtenissen

Twee gebeurtenissen A en B zijn onafhankelijk ⇔ P(A ∩ B) = P(A) \* P(B)

## Eigenschap: Wet van de totale kans

Als U = A1 ∪ A2 ∪Λ ∪ An met Ai ∩ Aj = φ ∀ i, j (i ≠ j)

Dan geldt voor elke gebeurtenis B:

Text

Description automatically generated

## Eigenschap: De regel van Bayes

Als U = A1 ∪ A2 ∪Λ ∪ An met Ai ∩ Aj = φ ∀ i, j (i ≠ j)

Dan geldt voor elke gebeurtenis B met P(B) > 0:

A picture containing text

Description automatically generated

Indien *A* en *B* twee deelverzamelingen zijn van het universum *U*, die corresponderen met het bezitten van de eigenschappen *E1* resp. *E2*, dan bevat de **unie** van *A* en *B* alle uitkomsten die tot ***A* en/of *B*** behoren.

Notatie: *A* ∪ *B*

Indien *A* en *B* twee deelverzamelingen zijn van het universum *U*, die corresponderen met het bezitten van de eigenschappen *E1* resp. *E2*, dan bevat de verzameling *A* min *B* alle uitkomsten die tot ***A* maar niet tot *B*** behoren. Notatie: *A* \ *B  = A* ∩ *Bc*

**6 behoort tot a en 2 en 4 NIET**

Twee **elkaar uitsluitende (of disjuncte)** gebeurtenissen zijn gebeurtenissen die geen gemeenschappelijke uitkomsten hebben. Ze kunnen dus niet samen optreden.

# Data representatie

## Python

Zie bestanden op laptop

## Frequentietabellen

Calendar

Description automatically generated with medium confidence

**Klassengrenzen**: zijn de kleinste en grootste grens van een klasse, in die zin dat de onderste grens in die klasse wel en de bovenste grens niet kan bereikt worden. Zo bevat de klasse [15.89; 17.89[ alle getallen die groter of gelijk zijn aan 15.89 en strikt kleiner dan 17.89

**Klassenbreedte**: is het verschil tussen de grootste en kleinste klassengrens van een klasse.

• **Klassenmidden**: is de helft van de som van de grootste en kleinste klassengrens van een klasse. Zo is het klassenmidden van de klasse [15.89 ; 17.89[ gelijk aan 16.89.

• **Klassenmidden**: is de helft van de som van de grootste en kleinste klassengrens van een klasse. Zo is het klassenmidden van de klasse [15.89 ; 17.89[ gelijk aan 16.89.

## Kengetallen

Een kengetal is in de beschrijvende statistiek een getal dat een eigenschap van een gegevensverzameling (dataset) samenvat. Kengetallen zijn vergelijkbaar met parameters in de mathematische statistiek. Met weinige kengetallen kan een overzicht van de verdeling van een grootheid over de data verkregen worden.

Table

Description automatically generated

A picture containing graphical user interface

Description automatically generated

## Boxplots

Chart, box and whisker chart

Description automatically generated

kwartielen:

minimum, 1e kwartiel (eerste kwart), de helft, de helft + 1e kwartiel, maximum

## Soort gegevens

Categorische gegevens

* Nominaal: verschillende waarden waar geen verschil zit tussen beide
* Ordinaal: duidelijke rangorde

Numerieke gegevens

* Discrete gegevens: gehele getallen
* Continue gegevens: bijna elke numerieke waarde

## Verstrengelende factoren

Voorbeeld: Een ijsverkoper aan zee beweert dat er een causaal verband is tussen het aantal mensen dat in zee zwemt en het aantal ijsjes dat hij verkoopt. Dit is een foutieve conclusie. Het aantal mensen dat in zee zwemt en het aantal ijsjes dat aan zee verkocht wordt, wordt nl beïnvloed door een andere variabele: temperatuur. Omdat beiden afhangen van temperatuur, lijkt er een verband te zijn tussen beide. Temperatuur noemen we een verstrengelende factor.

## MAR vs MCAR

= Missing value

MCAR = Missing Completely at Random: onafhankelijk van de waargenomen en niet-geobserveerde gegevens

MAR = Missing at Random: er is een systematische relatie tussen de neiging tot ontbrekende waarden en de waargenomen gegevens, maar niet de ontbrekende gegevens.

## Outlier

= Waarneming die ongewoon klein / groot is tov andere waarnemingen

Oorzaak: onjuiste waarde & juiste waarde (zeldzaam)

Detecteren: Boxplot, z-scores

## Populatie & Steekproef

Populatie = gehele groep waarover sprake is.

Steekproef = specifieke groep die je hebt onderzoekt

## Select vs aselect

Selecte steekproef: Populatie = steekproef

Aselecte steekproef: te onderzoeken elementen willekeurig uit populatie

## Symmetrisch

= Mediaan in het midden

## Variantie

Variantie = Maat van spreiding

Table

Description automatically generated

# Machine Learning

## Supervised vs Unsupervised

A picture containing table

Description automatically generated

Supervised Machine Learning

= gelabelde datasets om algoritmen te trainen die data classificeren of uitkomsten nauwkeurig voorspellen

* Classificatie
* Regressie

Unsupervised Machine Learning

= gebruikt machine learning-algoritmen om ongelabelde datasets te analyseren en te clusteren

* Clustering

## Klassieke aanpak & ML aanpak

Diagram

Description automatically generatedDiagram

Description automatically generated

## Stappen ML proces

A screenshot of a computer

Description automatically generated

## Video

https://hogeschoolpxl-my.sharepoint.com/personal/20005733\_pxl\_be/\_layouts/15/onedrive.aspx?id=%2Fpersonal%2F20005733\_pxl\_be%2FDocuments%2FGezichtherkkening%2Emp4&parent=%2Fpersonal%2F20005733\_pxl\_be%2FDocuments&ga=1