KU LEUVEN | Universiteit Antwerpen ASOE | Antwerp School of Education | UITWISKELING

Bomen, kruistabellen, verzamelingen en stochasten: Kansrekening in de 2^de^ en 3^de^ graad

Johan Deprez en Gilberte Verbeeck
johan.deprez@uitwiskeling.be en gilberte.verbeeck@uitwiskeling.be

slides: zie syllabus (met kadertjes) en digitaal ('zonder' kadertjes)

Inleiding

Wie zijn wij?

Kennismaking: wie zijn jullie?

- D – D/A – A?
- 2de graad – 3de graad?
- Ervaring met kansbomen?
- Ervaring met kruistabellen?

- Afspraak: als je voorbeeld kent → zwijgen bij groepswerk en observeren

UITWISKELING

35/3 Teltechnieken

33/2 Eén probleem drie representatie

34/2 Probleem-oplossend denken

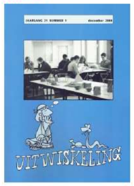
36/4 Besmettings-ziekten

41/1 Bomen, kruistabellen…

zie digitale syllabus

40/3 Spiraalaanpak: niet meer voor kansen

18/4 Bayes in kansbomen



21/1 Wiskunde en biologie

De aanleiding: nieuwe minimumdoelen

- De leerlingen lossen telproblemen op met behulp van boomdiagrammen en venndiagrammen; met kenniselementen: somregel, productregel, complementregel (2de graad D 06.17 en D/A 06.15);
- **De leerlingen bepalen kansen met behulp van kruistabellen, boomdiagrammen en de wet van Laplace; met kenniselement: verband tussen relatieve frequentie en kans (3de graad D 06.13);**
- **De leerlingen bepalen kansen met behulp van boomdiagrammen en met de wet van Laplace; met kenniselement: verband tussen relatieve frequentie en kans (3de graad D/A 06.06).**

Kansbomen en kruistabellen

2 voorbeelden: 4 (Z-R) groepen – maak boomdiagram

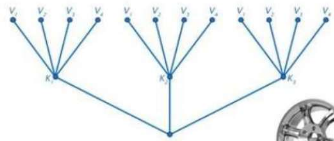
- **Voorbeeld 1 (start groep Z):**
  - 6 kaarten uit een kaartspel: Harten 2, H5, H8, HDame, Schoppen Aas en SDame
  - Trek 2 kaarten zonder ze terug te leggen
- **Voorbeeld 2 (start groep R):**

In een bedrijf kiezen alle werknemers één van de drie mobiliteitsformules:
1 een elektrische fiets;
2 een abonnement op het openbaar vervoer;
3 een abonnement op autodelen.
Van de werknemers die dichtbij wonen, dat is op minder dan 10 km van het bedrijf, kiezen er 72 voor de eerste formule, 16 voor de tweede en 12 voor de derde.
Bij de werknemers die veraf wonen, kiezen er 20 voor de eerste formule, 24 voor de tweede en 16 voor de derde.

Boomdiagram: terminologie

- **Terminologie:**
  - Boomdiagram of Telboom of Kansboom
  - Benoem in elke boom:
    1. Wortel
    2. Knopen
    3. Takken
    4. Bladeren
    5. Generatie
  - Voeg steeds de legende toe: knopen, bladeren

Bij de aankoop van een nieuwe wagen kun je, binnen het basispakket, kiezen uit drie verschillende kleuren en vier verschillende soorten velgen. **Schets een boomdiagram en bepaal hoeveel combinaties er mogelijk zijn.**



Legende

Soort velg

Kleur

Boomdiagram: terminologie

- **Voorbeeld boomdiagram:**

In een bedrijf kiezen alle werknemers één van de drie mobiliteitsformules:
1 een elektrische fiets;
2 een abonnement op het openbaar vervoer;
3 een abonnement op autodelen.
Van de werknemers die dichtbij wonen, dat is op minder dan 10 km van het bedrijf, kiezen er 72 voor de eerste formule, 16 voor de tweede en 12 voor de derde.
Bij de werknemers die veraf wonen, kiezen er 20 voor de eerste formule, 24 voor de tweede en 16 voor de derde.

- **Terminologie:**
  - Benoem in de boom:
    1. Wortel 2. Knopen 3. Takken 4. Bladeren 5. Generatie
  - Legende: knooppunten (woonplaats/formule) en bladeren (combinatie)
  - Geen telboom
  - Geen productregel

Telboom: terminologie

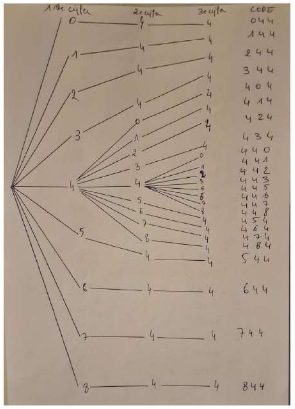
- **Voorbeeld telboom:**
  - 6 kaarten uit een kaartspel: H2, H5, H8, HD, SA en SD
  - Trek 2 kaarten zonder ze terug te leggen
  - Aantal mogelijkheden?
- **Terminologie:**
  - Benoem in de boom:
    1. Wortel 2. Knopen 3. Takken 4. Bladeren 5. Generatie
  - Legende: knooppunten (kaart 1/2) en bladeren (getrokken kaarten)
  - Telboom (vb. van boomdiagram)
  - Niet vereenvoudigd of wel – verloren informatie?
  - Productregel

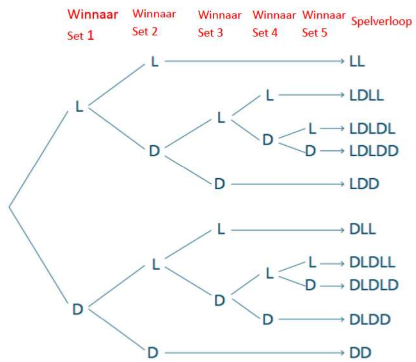
Didactische tips

- **Voorbeeld kaarten is geen goed voorbeeld**
  - Niet generiek, dus mag zeker niet het enige voorbeeld zijn
  - Niet vereenvoudigde telboom – aversie voor wiskundig getalenteerde leerlingen
  - Niet altijd leefwereld leerlingen, noch realistisch
- **Hoe probleemoplossend denken aanleren?**
  - Start van een probleem (zoals hier)
  - Inwerken in voorbeeld (stok kaarten voorradig – leerlingen laten trekken, vraagjes bij mobiliteitsprobleem)
  - Differentiëren

Generierieke voorbeelden van telbomen

Legende!





2 voorbeelden: 4 groepen – maak kansboom bij probleem

- **Voorbeeld 1:**
  - 6 kaarten uit een kaartspel: Harten 2, H5, H8, HDame, Schoppen Aas en SDame

Trek achtereenvolgens 2 kaarten zonder ze terug te leggen en let enkel op de kleur

- **Voorbeeld 2:**

In een bedrijf kiezen alle werknemers één van de drie mobiliteitsformules:
1 een elektrische fiets;
2 een abonnement op het openbaar vervoer;
3 een abonnement op autodelen.
Van de werknemers die dichtbij wonen, dat is op minder dan 10 km van het bedrijf, kiezen er 72 voor de eerste formule, 16 voor de tweede en 12 voor de derde.
Bij de werknemers die veraf wonen, kiezen er 20 voor de eerste formule, 24 voor de tweede en 16 voor de derde.

Doe alle werknemers in een pot en trek er willekeurig 1 werknemer uit

Kansen berekenen: los de vragen bij jouw vb op

- **Vb 1: Trek 2 kaarten zonder terugleggen uit 6 (4H+2S) en let enkel op kleur**
  - Wat is de kans
    1. om als 1ste kaart een harten kaart te trekken?
    2. om als 2de kaart een schoppen te trekken als de eerste kaart een harten was?
    3. om eerst een schoppen kaart en dan een harten kaart te trekken
    4. om als 2de kaart een harten te trekken (dus zonder dat je weet wat de 1ste kaart is)?
    5. op 1 schoppen en 1 harten, ongeacht de volgorde waarin ze getrokken worden?
    6. dat je een schoppen kaart als eerste trok als je weet dat de tweede kaart die je trok een harten was?
- **Vb 2: Trek 1 werknemer uit een pot van alle werknemers**
  - Wat is de kans dat
    1. de getrokken werknemer iemand is die voor formule 2 gekozen heeft?
    2. de getrokken werknemer iemand is die dichtbij woont en voor formule 2 gekozen heeft?
    3. een werknemer die voor formule 2 gekozen heeft dichtbij woont?
    4. een werknemer veraf woont of voor formule 1 koos?

Van telboom naar kansboom bij voorbeeld 1

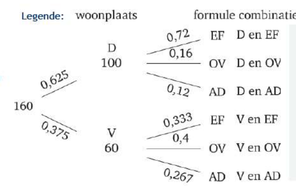
- **Voorbeeld kansboom:**
  - Trek 2 kaarten zonder terugleggen uit 6, let enkel op kleur
  - Wat is de kans
    1. om als 1ste kaart een harten kaart te trekken?
    2. om als 2de kaart een schoppen te trekken als de eerste kaart een harten was?
    3. om eerst een schoppen kaart en dan een harten kaart te trekken?
    4. om als 2de kaart een harten te trekken (dus zonder dat je weet wat de 1ste kaart is)?
    5. op 1 schoppen en 1 harten, ongeacht de volgorde waarin ze getrokken worden?
    6. dat je een schoppen kaart als eerste trok als je weet dat de tweede kaart die je trok een harten was?
- **Terminologie:**
  - Kansboom (i.p.v. boomdiagram) vertrekt vanuit vereenvoudigde telboom. Legende!
  - Bij de takken schrijven we kansen. De bladeren geven nu kansen = productregel
  - Vraag 4 en 5 = somregel

Van boomdiagram naar kansboom bij voorbeeld 2

- **Voorbeeld kansboom:**
  - Werknemers wonen dichtbij of veraf en kiezen voor één van de drie mobiliteitsformules
  - Wat is de kans dat
    1. de getrokken werknemer iemand is die voor formule 2 (abon Openbaar Vervoer) gekozen heeft?
    2. de getrokken werknemer iemand is die dichtbij woont en voor formule 2 gekozen heeft?
    3. een werknemer die voor formule 2 gekozen heeft dichtbij woont?
    4. een werknemer veraf woont of voor formule 1 (Elektrische Fiets) koos?
- **Kansboom opstellen in stappen:**
  1. Gegevens voorstellen in een boomdiagram
  2. Aantal werknemers dat dichtbij en veraf woont tellen en hieruit totaal aantal bepalen.

Van boomdiagram naar kansboom bij voorbeeld 2

- **Voorbeeld kansboom:**
  - Werknemers wonen dichtbij of veraf en kiezen voor één van de drie mobiliteitsformules
  - Wat is de kans dat
    1. de getrokken werknemer iemand is die voor formule 2 (abon Openbaar Vervoer) gekozen heeft?
    2. de getrokken werknemer iemand is die dichtbij woont en voor formule 2 gekozen heeft?
    3. een werknemer die voor formule 2 gekozen heeft dichtbij woont?
    4. een werknemer veraf woont of voor formule 1 (Electrische Fiets) koos?
- **Kansboom opstellen in stappen:**
  1. Gegevens voorstellen in een boom
  2. Aantal werknemers dat dichtbij en ve en hieruit totaal aantal bepalen.
  3. Kansen bepalen en bij takken en blad

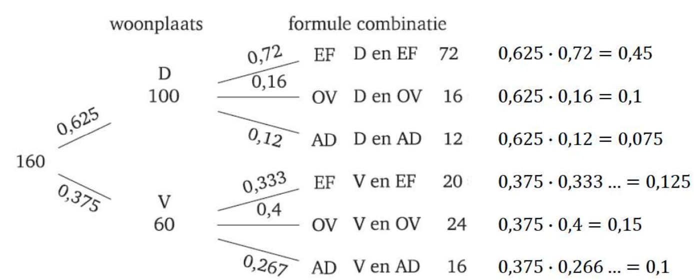


Van boomdiagram naar kansboom bij voorbeeld 2

- **Voorbeeld kansboom:**
  - Werknemers wonen dichtbij of veraf en kiezen voor één van de drie mobiliteitsformules
  - Wat is de kans dat
    1. de getrokken werknemer iemand is die voor formule 2 (abon Openbaar Vervoer) gekozen heeft?
    2. de getrokken werknemer iemand is die dichtbij woont en voor formule 2 gekozen heeft?
    3. een werknemer die voor formule 2 gekozen heeft dichtbij woont?
    4. een werknemer veraf woont of voor formule 1 (Elektrische Fiets) koos?
- **Oplossingen**

Uitgebreide kansboom bij voorbeeld 2

- **Terminologie:**
  - Kansboom met aantallen: uitgebreide kansboom



Kansen berekenen vanuit boomdiagram voorbeeld 2

- **Voorbeeld boomdiagram:**
  - Werknemers wonen dichtbij of veraf en kiezen voor één van de drie mobiliteitsformules
  - Wat is de kans dat
    1. de getrokken werknemer iemand is die voor formule 2 (abon Openbaar Vervoer) gekozen heeft?
    2. de getrokken werknemer iemand is die dichtbij woont en voor formule 2 gekozen heeft?
    3. een werknemer die voor formule 2 gekozen heeft dichtbij woont?
    4. een werknemer veraf woont of voor formule 1 (Electrische Fiets) koos?
- **oplossingen**

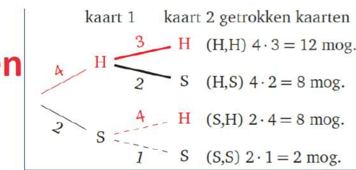
Productregel

Productregel

- **één naam, maar in feite verschillende regels**
  - enerzijds: productregel voor aantallen in vereenvoudigde telbomen
  - anderzijds: productregel voor kansen in kansbomen
- **we illustreren dit aan de hand van het voorbeeld met de kaarten**

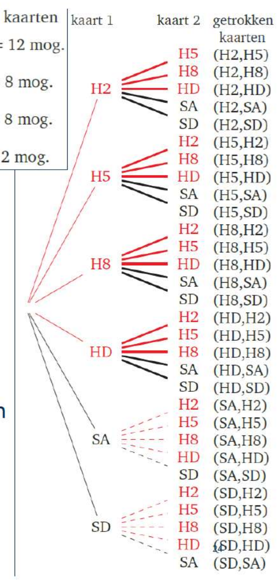
  2 kaarten trekken zonder teruglegging uit boek van 6 kaarten: H2, H5, H8, HD, SA, SD

Productregel voor aantallen



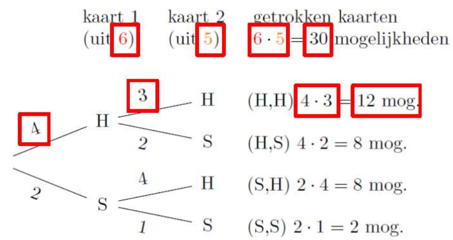
- **van telboom …**

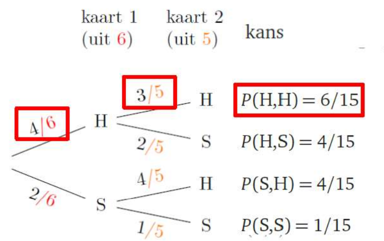
  # mogelijkheden om (H,H) te trekken:
  - 4 dunne rode lijnen die …
  - … elk vertakken in 3 dikke rode lijnen
  - totaal: 4 · 3 mogelijkheden
- **… naar vereenvoudigde telboom**
  - 4 dunne rode lijnen worden één dunne rode lijn met ‘4’ erbij
  - 4 pakketjes van 3 dikke rode lijnen worden één dikke rode lijn met een ‘3’ erbij
  - **productregel met aantallen**: 4 · 3 mogelijkheden



Productregel voor kansen

- **van vereenvoudigde telboom …**
  - productregel met aantallen
  - bv. aantal mogelijkheden voor (H,H) is 4 · 3
- **… naar kansboom**
  - kans op (H,H) =
  - $= \frac{12}{30} = \boxed{} = \boxed{}$
  - $=$ product van de kansen op de takken
  - productregel voor kansen





(Kruis)tabellen

2 voorbeelden: maak tabel met aantal mogelijkheden

Je mag tip vragen

- **Voorbeeld 2**

  In een bedrijf kiezen alle werknemers één van de drie mobiliteitsformules:
  1 een elektrische fiets;
  2 een abonnement op het openbaar vervoer;
  3 een abonnement op autodelen.
  Van de werknemers die dichtbij wonen, dat is op minder dan 10 km van het bedrijf, kiezen er 72 voor de eerste formule, 16 voor de tweede en 12 voor de derde.
  Bij de werknemers die veraf wonen, kiezen er 20 voor de eerste formule, 24 voor de tweede en 16 voor de derde.

- **Voorbeeld 1**
  - 6 kaarten uit een kaartspel: Harten 2, H5, H8, HDame, Schoppen Aas en SDame
  - Trek 2 kaarten zonder ze terug te leggen en let enkel op kleur.

tip

Kruistabellen: oplossing voorbeelden

- **Vb 2: mobiliteitsformules**
- **Vb 1: 2 kaarten zonder terugleggen uit 4 H en 2 S, let enkel op kleur**

Van kruistabellen met frequenties naar kansen

- **Voorbeeld 2:**
- **Voorbeeld 1:**

Kansen berekenen bij kruistabellen

- **Vb 2: Trek 1 werknemer uit een pot van alle werknemers**
  - Wat is de kans dat
    1. de getrokken werknemer iemand is die voor formule 2 gekozen heeft? $0{,}25$
    2. de getrokken werknemer iemand is die dichtbij woont en voor formule 2 gekozen heeft? $0{,}1$
    3. een werknemer die voor formule 2 gekozen heeft dichtbij woont? $\frac{0{,}1}{0{,}25} = 0{,}4$
    4. een werknemer veraf woont of voor formule 1 koos? $0{,}375 + 0{,}45 = 0{,}825$

Kansen berekenen bij kruistabellen

- **Vb 1: Trek 2 kaarten zonder terugleggen uit 6 (4H+2S) en let enkel op kleur**
  - Wat is de kans
    1. om als 1^{ste} kaart een harten kaart te trekken? P(H,X)= $\frac{2}{3}$
    2. om als 2^{de} kaart een schoppen te trekken als de eerste kaart een harten was? $\frac{\frac{4}{15}}{\frac{2}{3}} = \frac{2}{5}$
    3. om eerst een schoppen kaart en dan een harten kaart te trekken? P(S,H)= $\frac{4}{15}$
    4. om als 2^{de} kaart een harten te trekken (dus zonder dat je weet wat de 1^{ste} kaart is)? P(X,H)= $\frac{2}{3}$
    5. op 1 schoppen en 1 harten, ongeacht de volgorde waarin ze getrokken worden? P(S,H)+P(H,S)= $\frac{4}{15} + \frac{4}{15} = \frac{8}{15}$
    6. dat je een schoppen kaart als eerste trok als je weet dat de tweede kaart die je trok een harten was? $\frac{\frac{4}{15}}{\frac{2}{3}} = \frac{2}{5}$

Kruistabellen en kansbomen in de maatschappij: Covid

- Carl ging eind februari 2020 met vijf vrienden skiën in Ischgl (Oostenrijk), een ski-oord dat intussen bekend is als broeihaard van destijds de coronabesmettingen. Van zijn vijf vrienden kwamen er vier besmet terug. In België laat Carl zich testen op corona. Het testresultaat is negatief. Mag Carl besluiten dat hij niet besmet is?
- Werktekst individueel of in duo  PAUZE TOT 15u20
- Enkele begrippen
  - **Sensitiviteit** drukt uit hoeveel procent van de besmette mensen ook positief test.
  - **Specificiteit** drukt uit hoeveel procent van de niet-besmette mensen ook negatief test.
  - **Vals** negatief of positief
  - **Prevalentie** is het percentage besmette personen in een populatie.

Werken met absolute aantallen

- **Sensitiviteit: 71%**
- **Specificiteit: 98%** **schattingen**
- **Prevalentie: 3%**
- **Er zijn geen aantallen gegeven. We kiezen om te redeneren voor 100 000 Belgen**
- Maak een boomdiagram en een kruistabel die de informatie over al dan niet besmet zijn en al dan niet positief testen weergeeft. Kies of je eerst het boomdiagram of eerst de kruistabel maakt. Vergeet de legende niet!
  - Bij het boomdiagram: Is jouw boomdiagram een telboom, een kansboom of een uitgebreide kansboom? Kan je de knooppunten van volgorde veranderen?
  - Bij de kruistabel: Kan je je kruistabel op verschillende manieren maken?

Werken met absolute aantallen: Tip bij kruistabel

Werken met absolute aantallen: kruistabel

- **Sensitiviteit: 71%**
- **Specificiteit: 98%**
- **Prevalentie: 3%**
- **100 000 Belgen**

Werken met absolute aantallen: boomdiagram

- **Sensitiviteit: 71%**
- **Specificiteit: 98%**
- **Prevalentie: 3%**
- **100 000 Belgen**

Vals positieven en negatieven bij een kruistabel

- **Terminologie: vals positief en vals negatief**

1. **Omcirkel/fluoresceer het aantal vals positieven en vals negatieven**
2. **Geef het aantal personen van de groep waar Carl toe behoort.**

Vals positieven en negatieven bij een boomdiagram

1. **Bepaal welke opeenvolging van takken je bij het aantal vals positieven en vals negatieven brengt.**
2. **Hoe vind je het aantal personen van de groep waar Carl toe behoort?**

Mag Carl besluiten dat hij niet besmet is?

- **Carl test negatief. Mag hij dan besluiten dat hij niet besmet is?**

1. **Welke kans moet je bepalen om Carl te informeren?**
   Wat is de kans dat Carl niet besmet is, als je weet dat hij een negatief testresultaat had?
2. **Bepaal deze kans. Kies of je de boom of de kruistabel gebruikt.**
   Uit kruistabel en uitgebreide kansboom: $\frac{95\,060}{95\,930} = 0{,}991$

Carla test positief mag zij besluiten dat ze besmet is?

1. **Welke kans moet je bepalen om Carla te informeren?**
   Wat is de kans dat Carla besmet is, als je weet dat zij een positief testresultaat had?
2. **Bepaal deze kans. Kies of je de boom of de kruistabel gebruikt.**
   Uit kruistabel en uitgebreide kansboom: $\frac{2130}{4070} = 0{,}523$

Conclusies over vals positieven en vals negatieven

- **99,1% heeft de kans om niet besmet te zijn bij een negatieve test. Er is dus een klein percentage 0,9% dat bij een negatieve test toch besmet is. Deze personen hebben een vals gevoel van veiligheid.**
- **52,3% heeft de kans om besmet te zijn bij een positieve test. Dit is vrij laag. Ongeveer 48% (bijna de helft) van de positief geteste personen zal dus vals positief zijn en dus niet- besmet zijn ondanks de positieve test. Zij maken zich onnodig ongerust.**

Prevalentie van 3% naar 80%

- Carl behoort tot de risicogroep. Op basis van een groot onderzoek bij alle teruggekeerden van Ischl weet men dat de prevalentie in deze risicogroep 80% bedraagt. Onderzoek de volgende vragen:
  1. Indien zijn testresultaat negatief is, wat is dan de kans dat hij niet besmet is?
  2. Wat was zijn kans op besmetting, indien hij positief getest zou hebben?
- Antwoord via kruistabel:

Prevalentie van 3% naar 80%

- **Antwoord via kansboom:**

Conclusie: gericht testen

- **Prevalentie 3%:** **Het heeft geen zin iedereen te testen**
  - 0,9% zal bij een negatieve test toch besmet zijn. Deze personen hebben een vals gevoel van veiligheid.
  - Ongeveer 48% (bijna de helft) van de positief geteste personen zal niet- besmet zijn. Zij maken zich onnodig ongerust.
- **Prevalentie 80%:**
  - Ongeveer 46% zal bij een negatieve test toch besmet zijn. Deze personen hebben een vals gevoel van veiligheid.
  - Ongeveer 1% van de positief geteste personen zal niet- besmet zijn. Zij maken zich onnodig ongerust.

Verdachte gevallen testen: ziektesymptomen, uit risico gebied komen, contact met besmette mensen

Uitgebreide kansboom

Voorbeeld

Voorbeeld

Borstkanker is de meest voorkomende vorm van kanker bij vrouwen. Omdat deze aandoening beter bestreden kan worden naarmate de kanker vroeger wordt ontdekt, wordt aangeraden dat vrouwen regelmatig een mammografie laten uitvoeren.

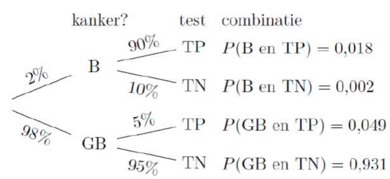
Zo'n test is echter niet onfeilbaar. We nemen hieronder aan dat de mammografie slechts bij 90 % van de vrouwen met borstkanker een 'positief' resultaat oplevert, dat wil zeggen dat er kanker wordt gedetecteerd[5]. Bij vrouwen zonder borstkanker is de kans op een verkeerdelijk positief resultaat 5 %. Verder nemen we aan dat ongeveer 2 % van de vrouwen tussen 40 en 50 borstkanker ontwikkelt[6].

Stel nu dat een willekeurige vrouw tussen 40 en 50 een positief testresultaat krijgt na zo'n mammografie. Wat is de kans dat ze borstkanker heeft?

uit: Delta Nova 5/6 Kansrekenen 3/4

Voorbeeld: twee oplossingen

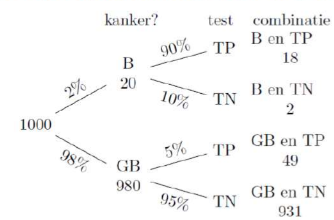
met een 'gewone' kansboom



$$P(B\ als\ TP) = \frac{P(B\ en\ TP)}{P(B\ en\ TP)+P(GB\ en\ TP)}$$

$$= \frac{0{,}018}{0{,}018+0{,}049} = 0{,}268\ldots \approx 0{,}27$$

een kansboom met een extra'tje

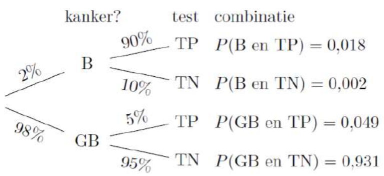


'gedachtenexperiment':

- **(fictieve) groep van 1000 vrouwen**
- **doe alsof zij zich exact gedragen zoals de kansen aangeven**

Voorbeeld: twee oplossingen

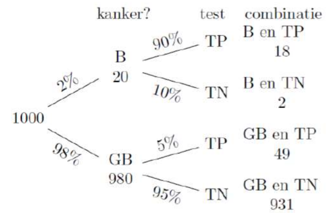
met een 'gewone' kansboom



$$P(B\ als\ TP) = \frac{P(B\ en\ TP)}{P(B\ en\ TP)+P(GB\ en\ TP)}$$

$$= \frac{0{,}018}{0{,}018+0{,}049} = 0{,}268\ldots \approx 0{,}27$$

een kansboom met een extra'tje



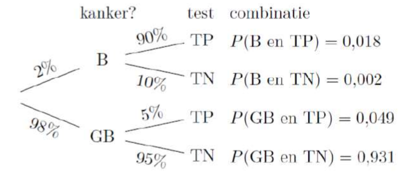
$P(B\ als\ TP)$?

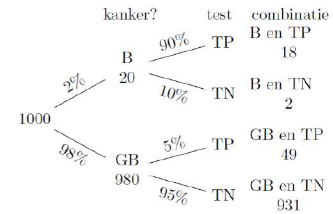
Twee oplossingen: voorbeeld

Vrij grote kans op vals positief…

Wordt de kans op vals positief groter of juist kleiner als de prevalentie van borstkanker …

- **… lager is (bv. 0,5% i.p.v. 2%)?**
- **… hoger is (bv. 8% i.p.v. 2%)?**





Uitgebreide kansboom

- **(onze) benaming: uitgebreide kansboom**
- **redeneren met aantallen (natuurlijke getallen) …**
- **… is toegankelijker dan …**
- **redeneren met kansen (breuken, percentages, decimale getallen)**
- **Dit wordt ondersteund door vakdidactisch onderzoek!**
- **Heb je gemerkt dat we dit eigenlijk al gedaan hebben (zie bv. slide 20)**

Uitgebreide kansboom

- **Geen verplichte leerinhoud, maar erg nuttig om dit in de klas te gebruiken en als tip mee te geven met leerlingen!**
- **disclaimer**
  - het is een gedachtenexperiment met 'ideale' uitkomsten;
  - het blijft belangrijk dat leerlingen weten dat er bij een echt kansexperiment variabiliteit optreedt in de uitkomsten!

De stap naar verzamelingen en stochastische veranderlijken

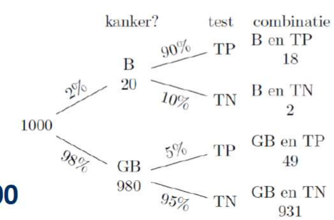
De stap naar verzamelingen en stochastische veranderlijken

- **kansrekening met verzamelingen**
  - <u>niet</u> verplicht (of wel?: zie MD ‘kansen … met behulp van de regel van Laplace’)
  - … maar (zeker voor een deel) interessant voor iedereen
  - … en is bekend bij alle (?) leraren
- **kansrekening met stochastische veranderlijken**
  - enkel verplicht voor specifieke eindtermen gevorderde wiskunde en/of statistiek
- **tijdens deze nascholing**
  - hoe hangt het allemaal aan elkaar?
  - vooral verzamelingen, slechts heel beperkt over stochastische veranderlijken
  - voor meer: zie digitale syllabus (artikel uit Uitwiskeling)

Van kansboom naar verzamelingen

- **borstkanker-voorbeeld**
- **kansexperiment: we hebben een groep van 1000 vrouwen en trekken er willekeurig één uit**
- **tabel met alle vrouwen en hun ‘kenmerken’**
- **de <u>uitkomstenverzameling</u> is**

$$U = \{v001, v002, \dots, v1000\}$$

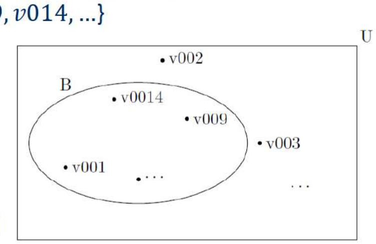


ID	kanker?	test?
v001	ja	positief
v002	neen	negatief
v003	neen	negatief
v004	neen	negatief
v005	neen	negatief
v006	neen	negatief
v007	neen	negatief
v008	neen	negatief
v009	ja	positief
v010	neen	negatief

Van kansboom naar verzamelingen

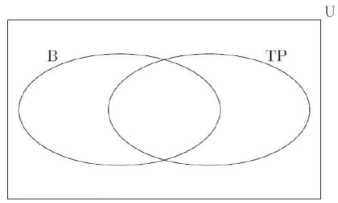
- **uitkomstenverzameling** $U = \{v001, v002, \dots, \}$
- **<u>gebeurtenis</u> = een verzameling van uitkomsten, meestal uitkomsten die aan een bepaalde eigenschap voldoen, bv.**

  B = verzameling van alle vrouwen met kanker = $\{v001, v009, v014, \dots\}$



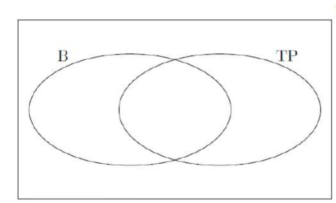
ID	kanker?	test?
v001	ja	positief
v002	neen	negatief
v003	neen	negatief
v004	neen	negatief
v005	neen	negatief
v006	neen	negatief
v007	neen	negatief
v008	neen	negatief
v009	ja	positief
v010	neen	negatief
v011	neen	positief

Van kansboom naar verzamelingen



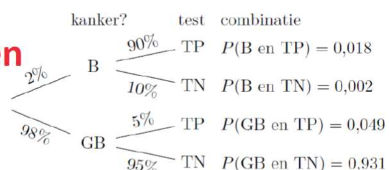
- andere gebeurtenissen: TP (test positief), GB (geen borstkanker), TN (test negatief)
- waar vind je GB, TN in de figuur?
- benoem de vier gebieden in de figuur
- 'samengestelde' gebeurtenissen kun je nu noteren via bewerkingen met verzamelingen

Van kansboom naar verzamelingen

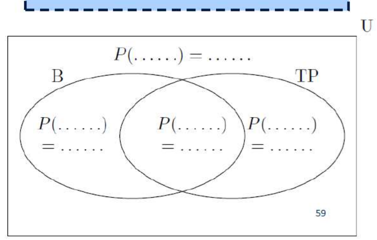


- er zijn nu twee manieren om samengestelde gebeurtenissen te benoemen
  - als logische uitspraak
    - bv. de vrouw die we getrokken hebben, heeft borstkanker en testte positief
    - bv. de vrouw … test niet positief
  - met verzamelingen,
    - bv. $B \cap TP$,
    - bv. $TN = TP^C$

Van kansboom naar verzamelingen



- <u>kans</u> van een uitkomst: hier heeft elke uitkomst kans 1/1000
- <u>kans</u> van een gebeurtenis = som van de kansen van alle elementen
- vul de vier kansen in!
  - zowel de getallen als de gebeurtenissen!
  - waar vind je de getallen in de kansboom?



Van kansboom naar verzamelingen

- we zijn hier heel voorzichtig tewerk gegaan
  - kansexperiment is groep van 1000 vrouwen waaruit we willekeurig 1 vrouw trekken
  - …
- wordt meestal niet zo in detail gedaan
  - ineens naar de vier gebeurtenissen met elk hun kans

De weg terug: van verzamelingen naar kruistabel

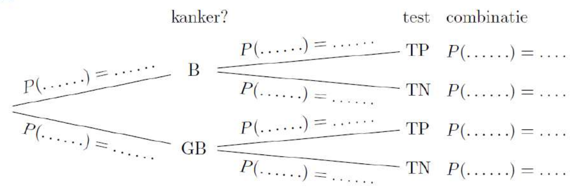
Vul in elke cel van de tabel de juiste kans in.

In vorige voorbeelden stonden de uitkomsten van het 'eerste' kansexperiment in de eerste kolom en nu staan ze in de eerste rij. Beide zijn OK!

test \ kanker?	B	GB	totaal
TP	$P(\ldots\ldots\ldots) = \cdots$	$P(\ldots\ldots\ldots) = \cdots$	$P(\ldots\ldots\ldots) = \cdots$
TN	$P(\ldots\ldots\ldots) = \cdots$	$P(\ldots\ldots\ldots) = \cdots$	$P(\ldots\ldots\ldots) = \cdots$
tot.	$P(\ldots\ldots\ldots) = \cdots$	$P(\ldots\ldots\ldots) = \cdots$	$P(\ldots\ldots\ldots) = \cdots$

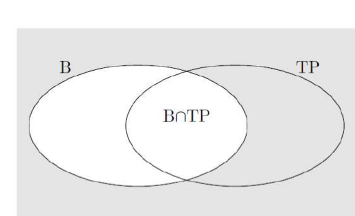
De weg terug: van kruistabel naar kansboom

- **vul zoveel mogelijk kansen in**
- **Twee vaststellingen**
  - de kansen op TP en op TN vind je niet als dusdanig in de kansboom, wel als je optelt
    - $P(TP) = P(TP \cap B) + P(TP \cap GB) = \cdots$
    - $P(TN) = \cdots$
  - de kansen op de tweede vertakking zijn voorwaardelijke kansen, bv. $P(TP|B)$



Voorwaardelijke kans

Voorwaardelijke kansen



bijvoorbeeld: $P(TP|B)$

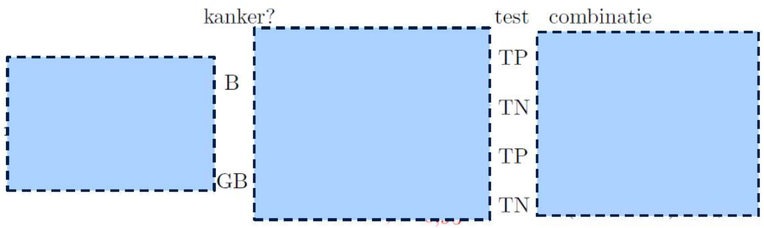
- **we beperken ons tot de vrouwen die borstkanker hebben**
- **de kans dat bij zo'n vrouw de test een positief resultaat oplevert:**

$$P(TP|B) = \frac{P(B \cap TP)}{P(B)}$$

- **we hebben voordien al informeel met voorwaardelijke kansen gewerkt**

De weg terug: van verzamelingen naar kansboom

- We vullen de kansboom nu verder aan!



- nieuwe formulering voor productregel, bv.

Kansboom versus kruistabel versus venndiagram

Welke kansen vind je rechtstreeks in …?

	kansboom	kruistabel	venndiagram
$P(B), P(GB)$			
$P(TP), P(TN)$			
$P(B \cap TP)$, $P(B \cap TN)$, $P(GB \cap TP)$, $P(B \cap TN)$			
$P(TP \mid B)$, $P(TN \mid B)$, $P(TP \mid GB)$, $P(TN \mid GB)$			
$P(B \mid TP)$, $P(B \mid TN)$, $P(GB \mid TP)$, $P(GB \mid TN)$			

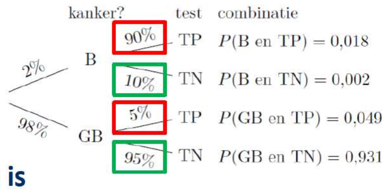
Kansboom versus kruistabel versus venndiagram

	kansboom	kruistabel	venndiagram
$P(B), P(GB)$			
$P(TP), P(TN)$			
$P(B \cap TP)$, $P(B \cap TN)$, $P(GB \cap TP)$, $P(B \cap TN)$			
$P(TP \mid B)$, $P(TN \mid B)$, $P(TP \mid GB)$, $P(TN \mid GB)$			
$P(B \mid TP)$, $P(B \mid TN)$, $P(GB \mid TP)$, $P(GB \mid TN)$			

- welke voorstelling is de beste?

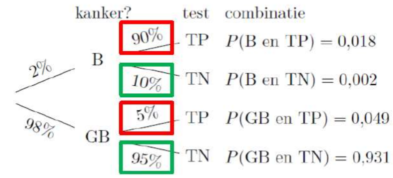
(On)afhankelijkheid

(On)afhankelijkheid in een kansboom



- **de kans op een positieve (en negatieve) test is <u>afhankelijk</u> van het al dan niet ziek zijn**
- **gebeurtenissen TP en TN zijn <u>afhankelijk</u> van de gebeurtenissen B en GB**
- **hoe zouden de kansen op de takken van de tweede generatie moeten zijn opdat TP en TN <u>on</u>afhankelijk zouden zijn van B en GB?**

(On)afhankelijke gebeurtenissen: definitie met verzamelingen



- **definitie**

  gebeurtenis A is onafhankelijk van gebeurtenis B asa $P(A|B) = P(A)$, en afhankelijk in het andere geval

- **hebben we <u>deze definitie</u> gebruikt bij het vaststellen van (on)afhankelijkheid bij de kansboom?**

(On)afhankelijke gebeurtenissen in een kruistabel

- **Eigenschap: A is onafhankelijk van B asa $P(A \cap B) = P(A) \cdot P(B)$**

$$\text{asa } P(A|B) = P(A) \text{ asa } \frac{P(A \cap B)}{P(B)} = P(A)$$

- **toegepast:**
  - is TP onafhankelijk van B?
  - neen! hoe zien we dit in de kruistabel?
  - is $P(B \cap TP) = P(B) \cdot P(TP)$?

	B	GB	totaal
TP	$P(B \cap TP)$	$P(GB \cap TP)$	$P(TP)$
TN	$P(B \cap TN)$	$P(GB \cap TN)$	$P(TN)$
tot.	$P(B)$	$P(GB)$	$P(U) = 1$

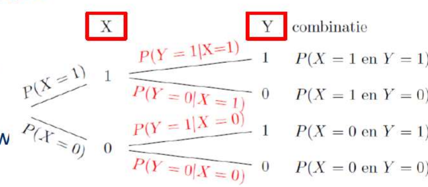
is | kans in een cel | = | marginale kans van de kolom | × | marginale kans van de rij | ?

De stap naar stochastische veranderlijken

test \ kanker?	B	GB	totaal
TP	$P(B \cap TP) = 0{,}018$	$P(GB \cap TP) = 0{,}049$	$P(TP) = 0{,}067$
TN	$P(B \cap TN) = 0{,}002$	$P(GB \cap TN) = 0{,}931$	$P(TN) = 0{,}933$
totaal	$P(B) = 0{,}02$	$P(GB) = 0{,}98$	$P(U) = 1$

Y \ X	1	0	
1	$P(X = 1 \text{ en } Y = 1)$	…	$P(Y = 1)$
0	…	…	…
	$P(X = 1)$	…	1

- **formaliseert legende uit kruistabel/ kansboom**
- **twee stochastische veranderlijken X en Y**
  - X drukt uit of de vrouw al dan niet borstkanker heeft: B: $X = 1$, GB: $X = 0$
  - Y drukt het resultaat van de test uit bij de vrouw TP: $Y = 1$, TN: $Y = 0$



Bedankt voor uw aandacht!
Vragen welkom!

Universiteit Antwerpen | ASOE | Antwerp School of Education

KU LEUVEN