



**UNIVERSITEIT  
GENT**

# INLEIDENDE BEGRIPPEN

Statistiek in de criminologie, hoofdstuk 2 / prof. dr. Christophe Vandeviver

# INLEIDING

# INLEIDING

- Doel
  - Inzicht verwerven in begrippenkader van kwantitatief criminologisch onderzoek.
- Onderzoekspopulaties, onderzoekseenheden (*cases*), en steekproeven (*samples*)
  - Steeds toevalssteekproeven (*aselect*, *willekeurig*, *at random*).
  - Belang van afbakening en representativiteit.
  - Bv. eerstejaarsstudenten, straatsegmenten.
- Variabelen of kenmerken
  - Enkel kenmerken die verschillen (spreiding, variabiliteit) tussen onderzoekseenheden in overweging nemen.
  - Verschillen tussen onderzoekseenheden voorspellen (onzekerheid modelleren, reduceren en verklaren).
  - Afhankelijke (te verklaren, *y*) en onafhankelijke (verklarende, *x*) variabelen.
  - Bv. Seksueel slachtofferschap, spreiding van criminaliteit, e.a.
- Nieuwsgierigheid en onderzoeks vragen
  - Bv. Wie wordt seksueel slachtoffer en in welke mate? (>< Wie wordt *geen* seksueel slachtoffer en waarom *niet*?)
  - Bv. Waar wordt criminaliteit gepleegd en waarom? (>< Waar wordt *geen* criminaliteit gepleegd en waarom *niet*?)

# BEGRIPPENKADER

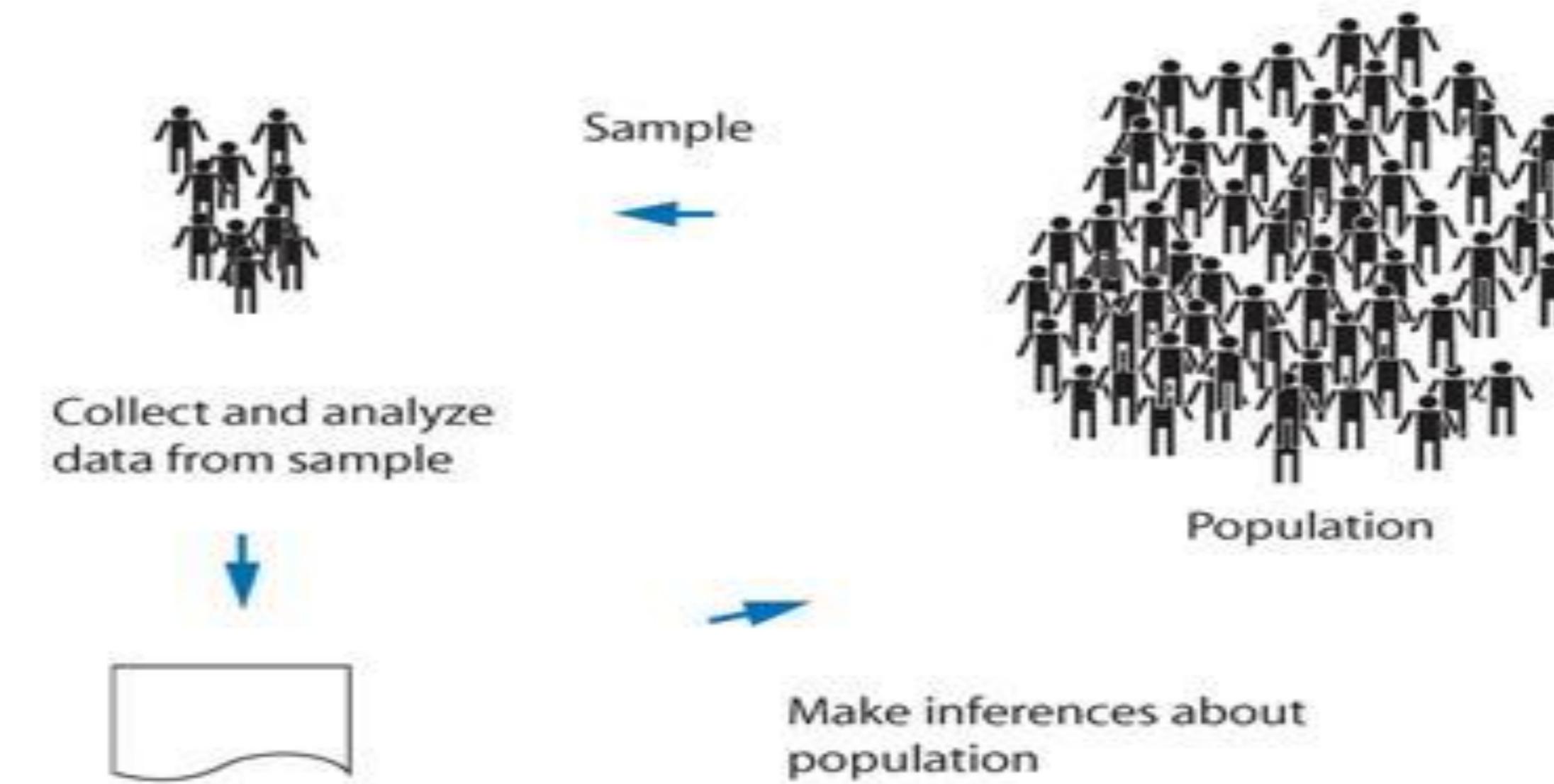
---

# STATISTIEK

# BESCHRIJVENDE & INFERENTIËLE STATISTIEK

- Beschrijvende statistiek
  - Kwantitatieve samenvatting (beschrijving) van de kenmerken van een steekproef.
  - Klemtoon bij datareductie & focus op steekproef.
  - Bv. maten van centraliteit en spreiding (gemiddelde, variantie).
- Inductieve of inferentiële statistiek
  - Kwantitatieve uitspraken over populatie a.d.h.v. steekproef.
  - Hypothesetoetsing: nulhypothese  $H_0$  & alternatieve hypothese  $H_a$ .
  - Klemtoon bij veralgemening & focus op populatie.
  - Bv. regressie-analyse.

# BESCHRIJVENDE & INFERENTIELLE STATISTIEK



# TYPES ONDERZOEKSVRAGEN

- Beschrijvende onderzoeks vragen
  - Kwantitatieve samenvatting (beschrijving) van het fenomeen.
  - Bv. Wat is de genderratio (m/v-verhouding) bij seksueel-slachtofferschap?
- Exploratieve onderzoeks vragen
  - Kwantitatieve verkenning van het fenomeen.
  - Bv. Is er een verband tussen het onderschrijven van traditionele visies op gender en slachtofferschap van seksueel geweld?
- Verklarende onderzoeks vragen
  - Kwantitatieve verklaring van de geobserveerde verschillen (variabiliteit) in een uitkomstvariabele a.d.h.v. kenmerken die eigen zijn aan een theoretisch paradigma.
  - Bv. In hoeverre verklaart de woongeschiedenis van een inbreker zijn doelwitkeuze?

# UNI-, BI- EN MULTIVARIATE STATISTIEK

- Univariaat
  - Eén variabele.
  - Maten van centraliteit & spreiding.
  - Bv. Gemiddelde, standaardafwijking.
- Bivariaat
  - Twee variabelen.
  - Samenhang (covariatie, correlatie).
  - Bv. Odds-ratio, covariantie, Pearson correlatie.
- Multivariaat
  - Meer dan twee variabelen.
  - Regressie.
  - Bv. OLS regressie, binomiale logistische regressive.

# DISCRETE & CONTINUE VARIABELEN

- Discrete variabelen
  - Verzameling van de natuurlijke getallen  $\mathbb{N}$ .
  - Aftelbaar oneindige uitkomstenverzameling: in principe steeds +1 mogelijk.
  - Bv. Aantal kinderen, aantal delicten.
- Continue variabelen
  - Verzameling van de reële getallen  $\mathbb{R}$ .
  - Overaftelbare oneindige uitkomstenverzameling: in principe steeds preciezer mogelijk (schaal  $\sim$  cm  $>$  mm  $>$  nm).
  - Bv. Tijd, afstand.

# VARIABILITEIT

# VARIABILITEIT EN VERSCHIL

- Statistiek draait om het begrijpen (beschrijven, verklaren, en beperken) van verschillen (variabiliteit, onzekerheid).
  - Bv. Waarom verschillen eerstejaarsstudenten in de mate waarin ze seksueel slachtoffer worden?
  - Bv. Waarom verschillen straten in de mate waarin er criminaliteit voorkomt?
- Onderzoekseenheden waarover men uitspraken wil doen, moeten dus onderling verschillen (variabiliteit) op de bestudeerde kenmerken (variabelen).
  - Bv. Mate waarin eerstejaarsstudenten seksueel slachtoffer worden verschilt tussen eerstejaarsstudenten (veelen geen, sommigen een beetje, weinigen zeer veel) (zgn. ‘hot’ victims).
  - Bv. Niet alle straten in een stad kennen evenveel criminaliteit (veel geen, sommige een beetje, weinig zeer veel) (zgn. hot spots of crime).
- Verschil impliceert ‘onzekerheid’  $\Delta$ .
  - Uiteindelijk is het doel om die ‘onzekerheid’  $\Delta$  te modeleren en reduceren.
  - Indien géén verschil, dan vervalt een variabele tot een constante  $c$  en is er zekerheid (want steeds dezelfde waarde).

# VARIABILITEIT IMPLICEERT ONZEKERHEID

- Verschillen (variabiliteit) op een kenmerk (variabele) impliceren onzekerheid  $\Delta$ .
  - Bv. We kunnen niet met zekerheid zeggen welke eerstejaarsstudent wel of niet slachtoffer zal worden.
  - Bv. We kunnen niet met zekerheid zeggen in welke straat er delicten zullen voorkomen.
- Onzekerheid  $\Delta$  op verdeling van het te verklaren kenmerk (afhankelijke variabele  $y$ ) tussen de onderzoekseenheden trachten we te modelleren (bv. regressie) en te verklaren (begrijpen) om ultiem die onzekerheid te beperken.
  - Bv. Waarom wordt de ene eerstejaarsstudent wél seksueel slachtoffer en de andere niet?
  - Bv. Waarom komt criminaliteit voor in de ene straat en in de andere niet?

# GEBREK AAN VARIABILITEIT IS EEN PROBLEEM

- Als een kenmerk niet varieert (verschilt) tussen onderzoekseenheden, vervalt dat kenmerk tot een constante  $c$  zonder onzekerheid  $\Delta$  ( $\neq$  variabele!).
  - Bv. Indien elke eerstejaarstudent in dezelfde mate seksueel slachtoffer wordt (of niet), kan je geen kenmerken van eerstejaarsstudenten gebruiken om de mate van seksueel slachtofferschap te verklaren.
  - Bv. Indien criminaliteit in dezelfde mate voorkomt over alle straten heen (of niet), kan je geen kenmerken van die straten hanteren om criminaliteit te verklaren.
- Ultiem valt er dan niets te begrijpen of te verklaren.
  - Er is geen onzekerheidsprobleem meer om op te lossen.
- Zonder variabiliteit, geen kwantitatief onderzoek want elke oplossing is perfect gedefinieerd (deterministisch) en er kan dus geen onzekerheid  $\Delta$  onderzocht of verklaard ( $\sim$  gemodelleerd, beperkt) worden.

# WAAROM IS VARIABILITEIT ZO BELANGRIJK?

- Ultiem trachten we verschillen op een kenmerk tussen onderzoekseenheden te verklaren.
- Daarvoor zoeken we naar andere kenmerken (zgn. onafhankelijke variabelen  $x$ ) waarvan de verschillen samenhangen (dezelfde mate van variabiliteit of verschil vertonen, zgn. co-variabiliteit) met de waargenomen verschillen  $\Delta$  op het te verklaren kenmerk  $y$  (zgn. afhankelijke variabele  $y$ ).
  - Bv. Waarom wordt de ene eerstejaarsstudent wél slachtoffer van criminaliteit en de andere niet? Misschien speelt geslacht een rol! Maar ... als alle onderzoekseenheden mannelijke eerstejaarsstudenten zijn, kan je niet nagaan of geslacht een rol speelt (er zijn geen verschillen op geslacht).
  - Bv. Waarom komt criminaliteit voor in de ene straat en in de andere niet? Misschien speelt de aanwezigheid van een café een rol! Maar .. als alle onderzoekseenheden straten zonder cafés zijn, kan je niet nagaan of de aanwezigheid van een café een rol speelt (er zijn geen verschillen op de aanwezigheid van een café).
- Dit impliceert dat een constant kenmerk  $c$  geen verklaring kan zijn voor waargenomen verschillen  $\Delta$  op het te verklaren kenmerk  $y$ .
  - M.a.w. een constante  $c$  voegt geen informatie toe die toelaat de mate van onzekerheid  $\Delta$  te beperken.
- Wanneer is er voldoende spreiding?
  - Elk kenmerk (variabele) dient min. 2 verschillende waarden te hebben (bv. man, vrouw; bv. café, geen café), én
  - Per waarde min. 1 onderzoekseenheid.

# MEETNIVEAUS

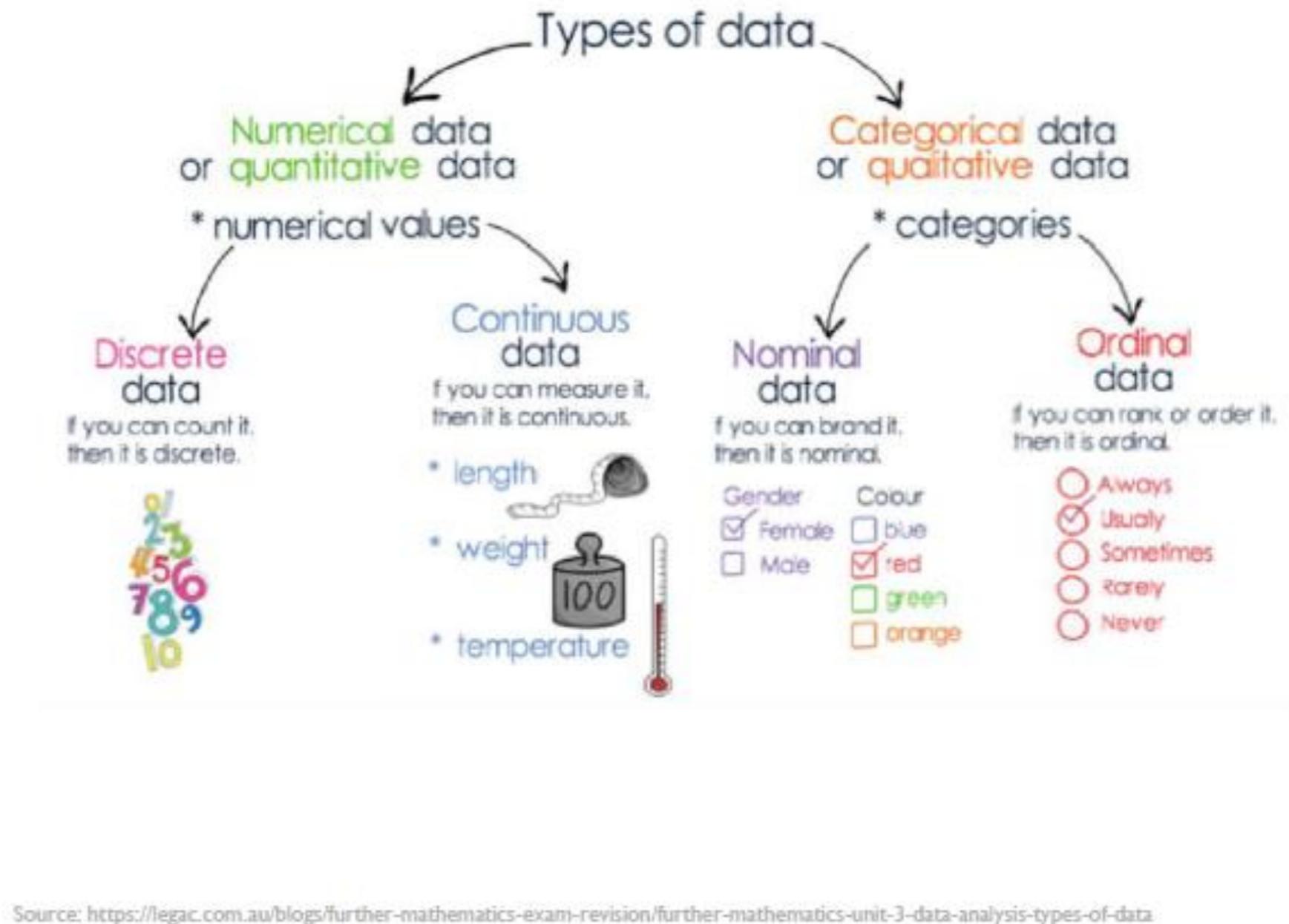
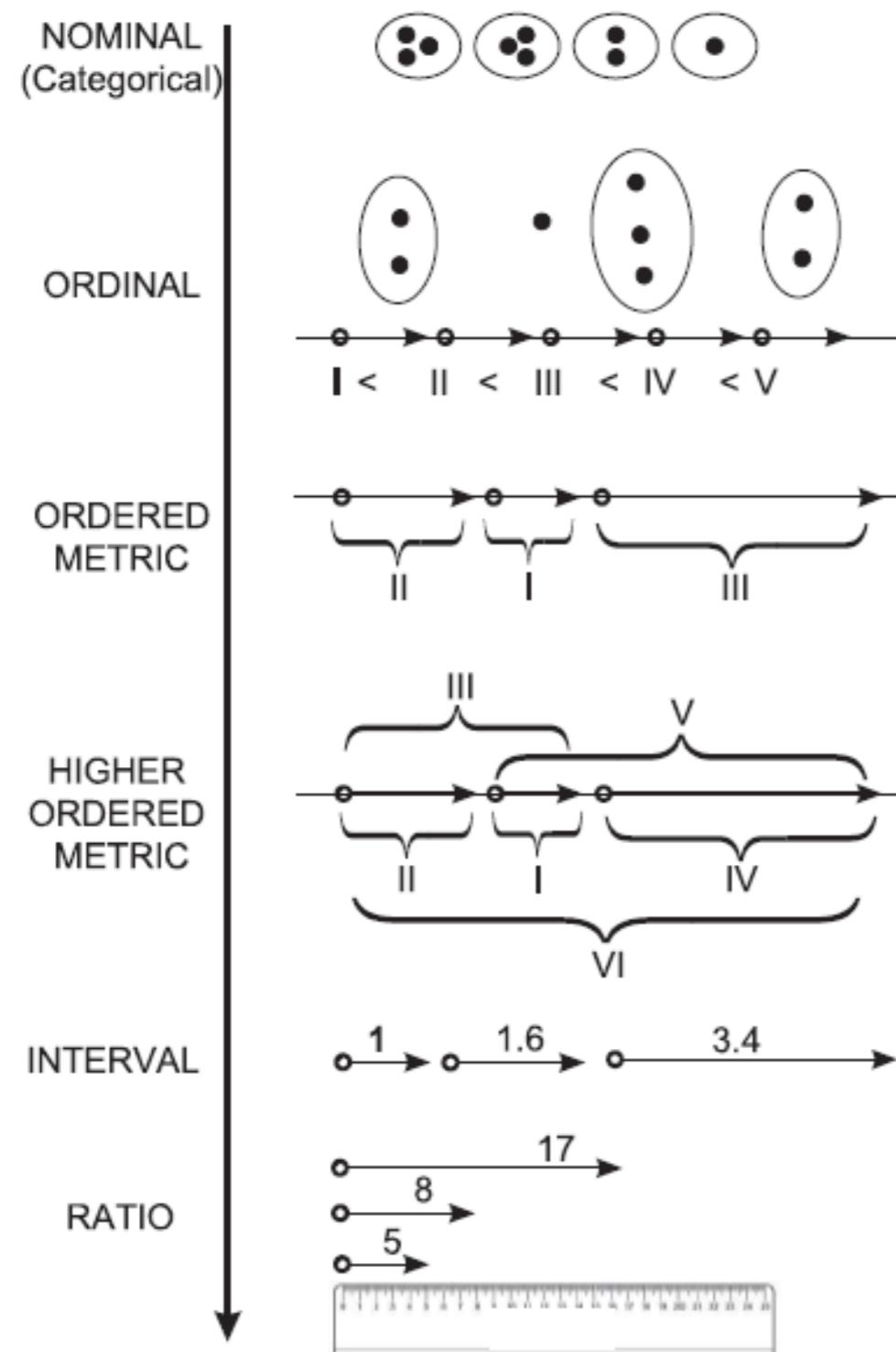
# CATEGORISCHE MEETNIVEAUS

- Nominaal
  - Uitputtende en wederzijds uitsluitende categorieën.
  - Géén ordening mogelijk.
  - Bv. Geslacht, geboorteland.
- Ordinaal
  - Ordening is mogelijk o.b.v. de mate waarin een kenmerk voorkomt (meer of minder).
  - Sprake van een continuüm.
  - Bv. Opinievragen (helemaal niet akkoord—helemaal wel akkoord), opleidingsniveau (basisonderwijs, middelbaar, hoger, etc.).

# METRISCHE MEETNIVEAUS

- Interval
  - Vaste (gelijke) afstanden (met exacte betekenis) tussen categorieën.
  - Natuurlijk (absoluut) nulpunt ontbreekt ( $\neq 0$ ).
  - Bv. Temperatuur in Celsius, geboortejaar.
- Ratio
  - Natuurlijk (absoluut) nulpunt is 0.
  - Bv. Temperatuur in Kelvin, lengte, criminaliteit, slachtofferschap, inkomen.

# SAMENGEVAT



**FIGURE 4.4** Diagrammatic classification of scale types and their data. Adapted from Coxon, A.P.M., 1982. *The User's Guide to Multidimensional Scaling*. Heinemann Educational Books, London.

# Prof. Dr. Christophe Vandeviver

Onderzoeksprofessor Criminologie UGent

VAKGROEP CRIMINOLOGIE, STRAFRECHT &  
SOCIAAL RECHT

E [christophe.vandeviver@ugent.be](mailto:christophe.vandeviver@ugent.be)

T +32 9 264 67 38

-  Universiteit Gent
-  @ugent
-  Ghent University

[www.ugent.be](http://www.ugent.be)