FUNDAMENTAL CONCEPTS IN MEDICAL STATISTICS

... with applications to randomized controlled trials

Håkon K. Gjessing

Professor/Principal Investigator

Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo

Department of Global Public Health and Primary Care, University of Bergen

Makerere

7-13 June 2023

CENTRE FOR FERTILITY AND HEALTH

COURSE HOME PAGES

https://gjessing.super.site + https://muele.mak.ac.ug

Fundamental concepts in medical statistics with applications to randomized controlled trials, Kampala 7-13 June 2023

COURSE RESOURCE PAGE

Organizers

Makerere University
College of Health Sciences
SCHOOL OF PUBLIC HEALTH

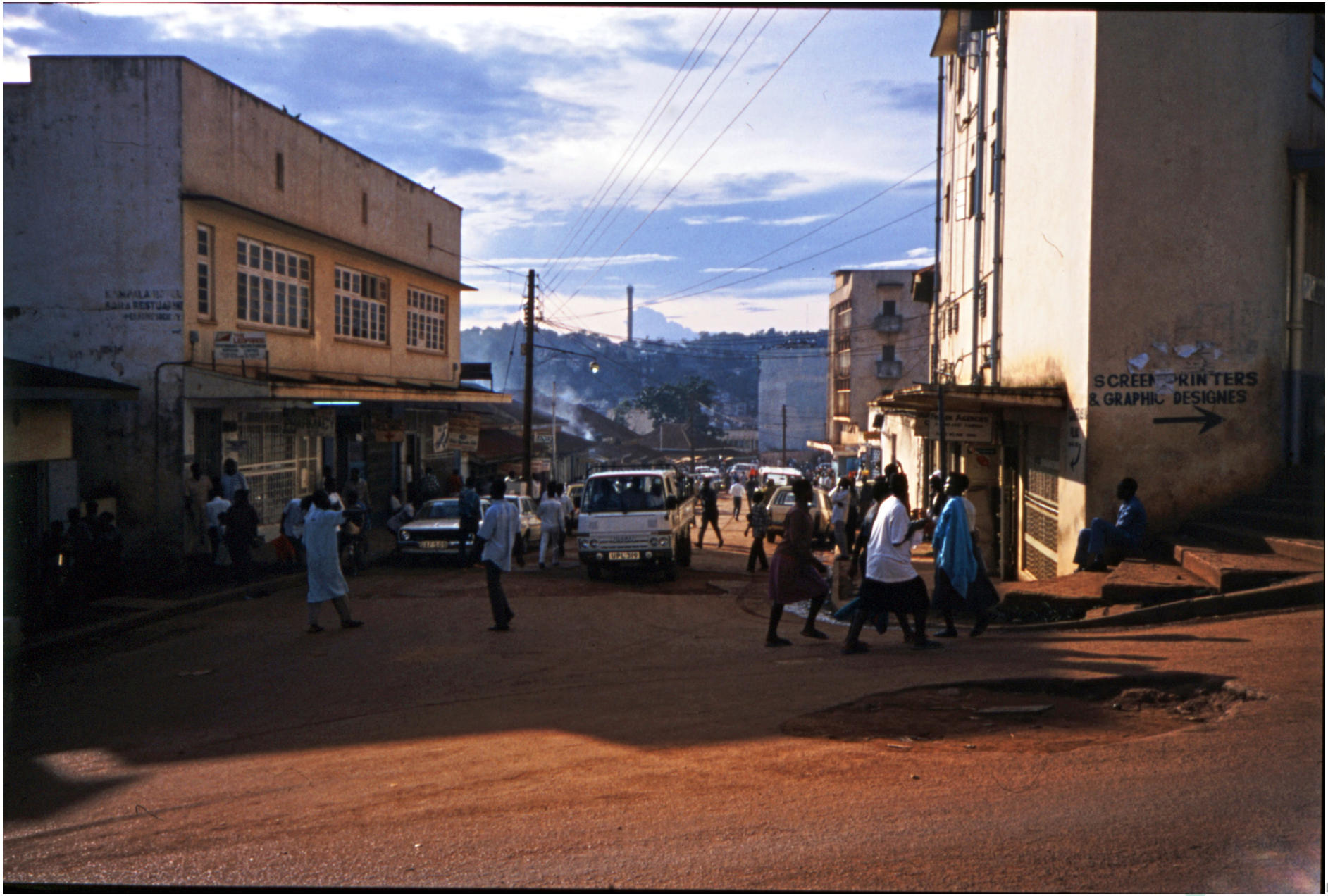
Victoria Nankabirwa
Lecturer

CIS MAC

Halvor Sommerfelt
Professor

CENTRE FOR FERTILITY AND HEALTH

KAMPALA, 1993



University Main Building, 1993

University Guest House, 1993

At the museum, 1993

About this course...

- More on principles & ideas rather than recipes
- The exercises will support with hands on practical examples
- Selection of topics mostly based on personal interests (and what I had available)
- .... but it should all be very relevant

CENTRE FOR FERTILITY AND HEALTH

SCHEDULE: WEDNESDAY, 7 JUNE

Wednesday 7

FOCUS OF THE DAY: Overview, and fundamental concepts

09:00 - 10:00 Welcome and startup

- Welcome/introduction round
- Quick overview over course + practical info
- Quick overview over some basic stuff (but note that participants are expected to have a basic understanding of/experience with basic statistical concepts)

10:00 - 12:00 Some fundamentals

- Central limit theorem (most importantly, why it's useful)
- General background on Confidence Intervals
- Testing, p-values, one-sided vs. two-sided
- Confidence intervals vs. testing
- Transformation of data, the log-transform. General discussion of measurement scales

12:00 - 13:00 LUNCH

13:00 - 14:00 Exercises

14:00 - 15:00 Sample size and statistical power / statistical precision

- What it means
- How to restrict the probability of Type I/Type II errors, i.e. achieve sufficient statistical power at a given test significance level
- ... or sometimes better, compute sample size needed for sufficient precision of your effect estimate
- How to compute sample size for study planning, simple situations

15:00 - 16:00 Exercises

CENTRE FOR FERTILITY AND HEALTH

SCHEDULE: THURSDAY, 8 JUNE

- **Thursday 8**

  FOCUS OF THE DAY: Generalized linear models (GLMs)

  - **09:00 - 11:00 Linear regression**
    - Standard linear regression
    - Multiple linear regression, confounding, interpretation
    - Basic regression diagnostics
  - **11:00 - 12:00 Exercises**
  - **12:00 - 13:00 LUNCH**
  - **13:00 - 15:00 Binomial regression**
    - Binomial regression (including logistic regression)
      - Different link functions: logit, log, and identity
      - Corresponding interpretations: logistic, log, and additive
      - Odds Ratio (OR), Relative Risk (RR), and Risk Difference (RD)
      - Basic regression diagnostics
    - Poisson regression (Introduced here as a GLM, but will for the most part be covered on Monday, in the context of event history analyses.)
  - **15:00 - 16:00 Exercises**

CENTRE FOR FERTILITY AND HEALTH

SCHEDULE: FRIDAY, 9 JUNE

- **Friday 9**

  FOCUS OF THE DAY: RCTs, other experimental designs

  - **09:00 - 10:00 Fundamentals**
    - The causal reasoning behind randomization
    - Adjustment for baseline variables? Other adaptations.
    - Regression to the mean, need for trial comparator arm, etc.
    - Dropouts, intention to treat
  - **10:00 - 11:00 Multiple testing, multiple estimation**
    - Problems with multiple testing
    - "Winner's curse" in multiple estimation
    - p-hacking, multiple subgroups, etc.
    - Some extreme examples from genetics
  - **11:00 - 12:00 Exercises**
  - **12:00 - 13:00 LUNCH**
  - **13:00 - 13:30 Weighted analyses**
    - Various types of weighting
      - Sampling
      - Frequency
      - Importance (although not really important here)
  - **13:30 - 15:00 Interaction / effect measure modification**
    - Basic concepts
    - Scale dependence
    - Examples
    - How to estimate, variable coding, such as Stata's approach, and dummy variables
    - Testing / estimating differences in effects across subgroups
  - **15:00 - 16:00 Exercises**

CENTRE FOR FERTILITY AND HEALTH

Schedule: Saturday, 10 June

Saturday 10

FOCUS OF THE DAY: Self study/group work

CENTRE FOR FERTILITY AND HEALTH

Schedule: Monday, 12 June

- **Monday 12**

  FOCUS OF THE DAY: Count data, standardization

  - **09:00 - 11:00 Poisson regression**
    - How and when to use
    - How to handle varying exposure time, use of offset
    - Examples from Covid-19 pandemic analyses
  - **11:00 - 12:00 exercises**
  - **12:00 - 13:00 LUNCH**
  - **13:00 - 15:00 Standardization methods**
    - Standardized averages for continuous data
    - Standardized risks and rates
    - Standardized Mortality Ratios (SMRs)
    - Model-based standardization, examples of how to present results
  - **15:00 - 16:00 Exercises**

CENTRE FOR FERTILITY AND HEALTH

SCHEDULE: TUESDAY, 13 JUNE

Tuesday 13

FOCUS OF THE DAY: Survival and event history analysis. Recap of central topics from the course.

09:00 - 11:00 Survival Analysis / Event History Analysis

- Introduction to survival/event history analysis
- Kaplan-Meier
- Cox regression

11:00 - 12:00 Exercises

12:00 - 13:00 LUNCH

13:00 - 14:00 Additional topics in survival analysis

- Alternative models, such as additive hazards model
- Competing risk

14:00 - 15:00 Summary, discussion, evaluation

CENTRE FOR FERTILITY AND HEALTH

Why Medical Statistics?

CENTRE FOR
FERTILITY AND HEALTH

Statistics in the Medical Field

New England Journal of Medicine

Editorial, Jan. 6, 2000, p. 42-49

The eleven most important developments in medicine over the past millennium

- Elucidation of human anatomy and physiology
- Discovery of cells and their substructures
- Elucidation of the chemistry of life
- *Application of statistics to medicine*
- Development of anesthesia
- Discovery of the relation of microbes to disease
- Elucidation of inheritance and genetics
- Knowledge of the immune system
- Development of body imaging
- Discovery of antimicrobial agents
- Development of molecular pharmacotherapy

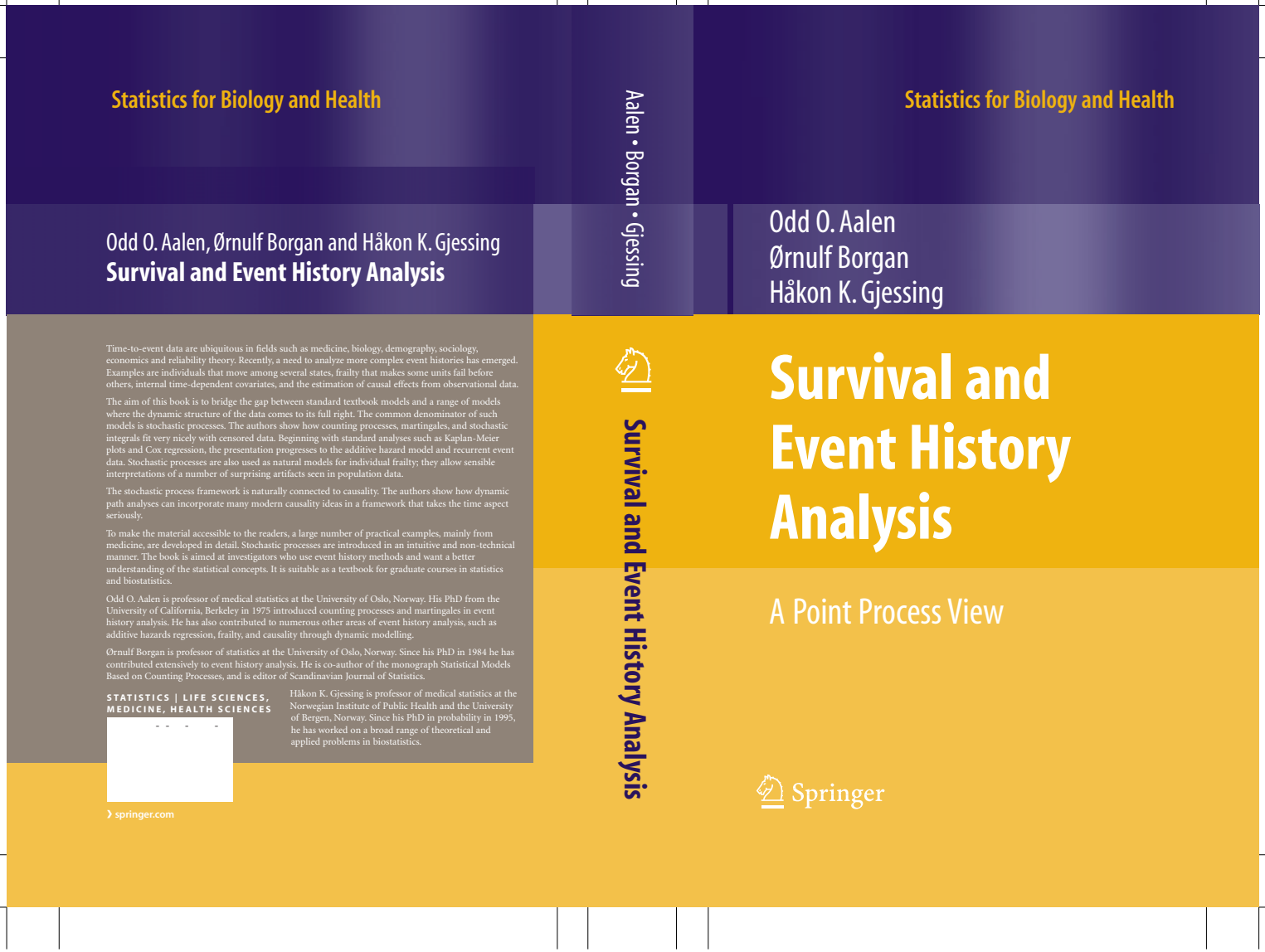
CENTRE FOR FERTILITY AND HEALTH

Application of statistics to medicine

"Sir David Cox's 1972 paper on proportional-hazards regression ignited the fields of survival analysis and semiparametric inference (using partial specification of the probability distribution of the outcomes under investigation). Rapid improvements in computer support were essential to the growing role of empirical investigation and statistical inference."

CENTRE FOR FERTILITY AND HEALTH

TOTALLY INDECENT SELF-PROMOTION



CENTRE FOR FERTILITY AND HEALTH

A digression into

Statistical Software

CENTRE FOR FERTILITY AND HEALTH

RELEVANT SOFTWARE

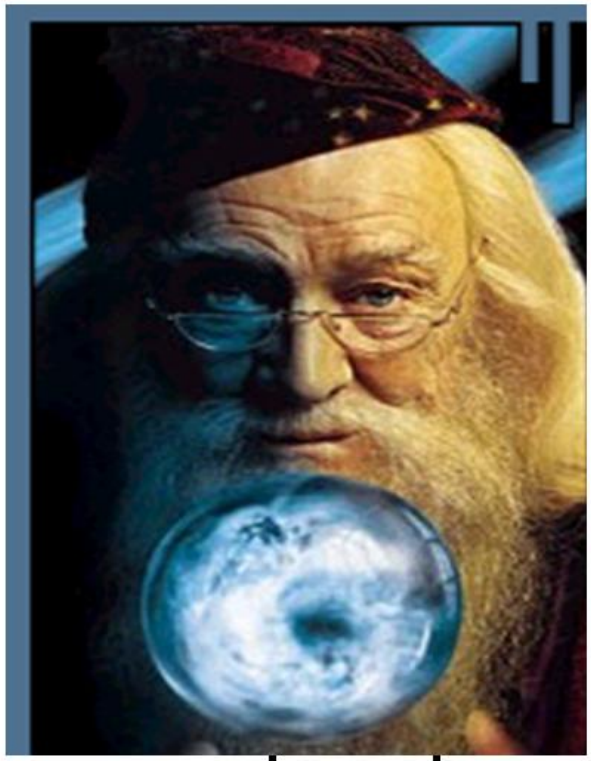
- R
  - Many users find RStudio helpful. Free, open-source version
    `https://posit.co/products/open-source/rstudio/`
  - Some simple GUIs are also available
    `https://r4stats.com/2022/06/20/updated-comparison-of-r-guis/`
- Stata
  - Stata commands, manuals, and documentation are in general of

    — AMAZING quality! —
- SPSS (Statistical Pseudo Science System)
- SAS
- Python
- Julia

CENTRE FOR FERTILITY AND HEALTH

R USERS

=



wizard

CENTRE FOR FERTILITY AND HEALTH

SPSS USERS

SPSS =

muggle

CENTRE FOR FERTILITY AND HEALTH

What type of "Restaurant" is your Favorite Software?

- STATA
  - Michelin restaurant
  - Excellent, detailed menu and top quality
  - Can meet the cooks or attend cooking classes

CENTRE FOR FERTILITY AND HEALTH

What type of "Restaurant" is your Favorite Software?

- STATA
  - Michelin restaurant
  - Excellent, detailed menu and top quality
  - Can meet the cooks or attend cooking classes
- R
  - YOU ARE THE COOK...
  - All the tools & ingredients are there for you to use
  - The sky is the limit...
  - (...or it all ends in a total mess)

CENTRE FOR FERTILITY AND HEALTH

What type of "Restaurant" is your Favorite Software?

- STATA
  - Michelin restaurant
  - Excellent, detailed menu and top quality
  - Can meet the cooks or attend cooking classes
- R
  - YOU ARE THE COOK...
  - All the tools & ingredients are there for you to use
  - The sky is the limit...
  - (...or it all ends in a total mess)
- SPSS
  - McDonald's

CENTRE FOR FERTILITY AND HEALTH

What type of "Restaurant" is your Favorite Software?

- STATA
  - Michelin restaurant
  - Excellent, detailed menu and top quality
  - Can meet the cooks or attend cooking classes
- R
  - YOU ARE THE COOK...
  - All the tools & ingredients are there for you to use
  - The sky is the limit...
  - (...or it all ends in a total mess)
- SPSS
  - McDonald's
  - Motto: "More than just Burgers & Fries!"

CENTRE FOR FERTILITY AND HEALTH

What type of "Restaurant" is your Favorite Software?

- STATA
  - Michelin restaurant
  - Excellent, detailed menu and top quality
  - Can meet the cooks or attend cooking classes
- R
  - YOU ARE THE COOK...
  - All the tools & ingredients are there for you to use
  - The sky is the limit...
  - (...or it all ends in a total mess)
- SPSS
  - McDonald's
  - Motto: "More than just Burgers & Fries!"
- How about SAS, then?

CENTRE FOR FERTILITY AND HEALTH

YES, HOW ABOUT SAS, THEN?

R: Wonderfully clear & pedagogical help pages

Excerpt from an R help page for the frequently used *apply* function

apply

If each call to FUN returns a vector of length n, then apply returns an array of dimension c(n, dim(X)[MARGIN]) if n > 1. If n equals 1, apply returns a vector if MARGIN has length 1 and an array of dimension dim(X)[MARGIN] otherwise. If n is 0, the result has length 0 but not necessarily the ‘‘correct’’ dimension.

If the calls to FUN return vectors of different lengths, apply returns a list of length prod(dim(X)[MARGIN]) with dim set to MARGIN if this has length greater than one.

In all cases the result is coerced by as.vector to one of the basic vector types before the dimensions are set, so that (for example) factor results will be coerced to a character array.

CENTRE FOR FERTILITY AND HEALTH

R: WONDERFULLY CLEAR & PEDAGOGICAL HELP PAGES

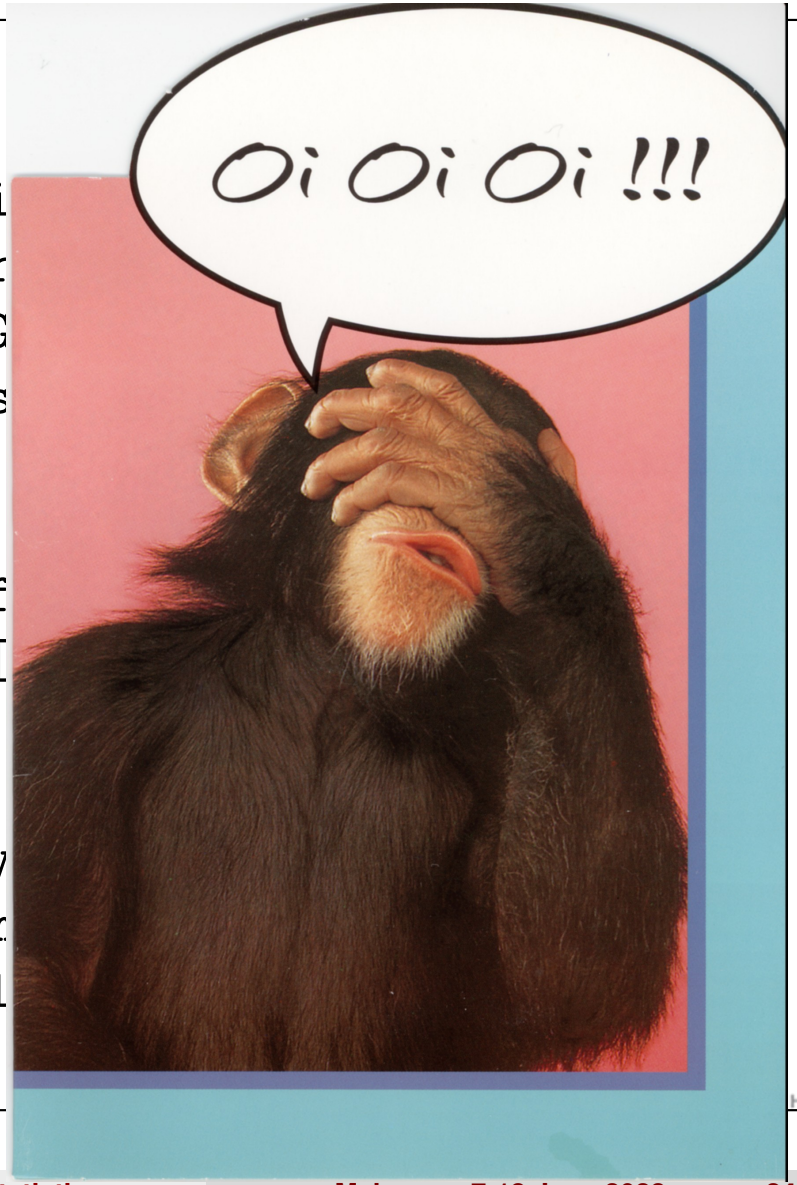
Excerpt from an R help page for the frequently used *apply* function

apply

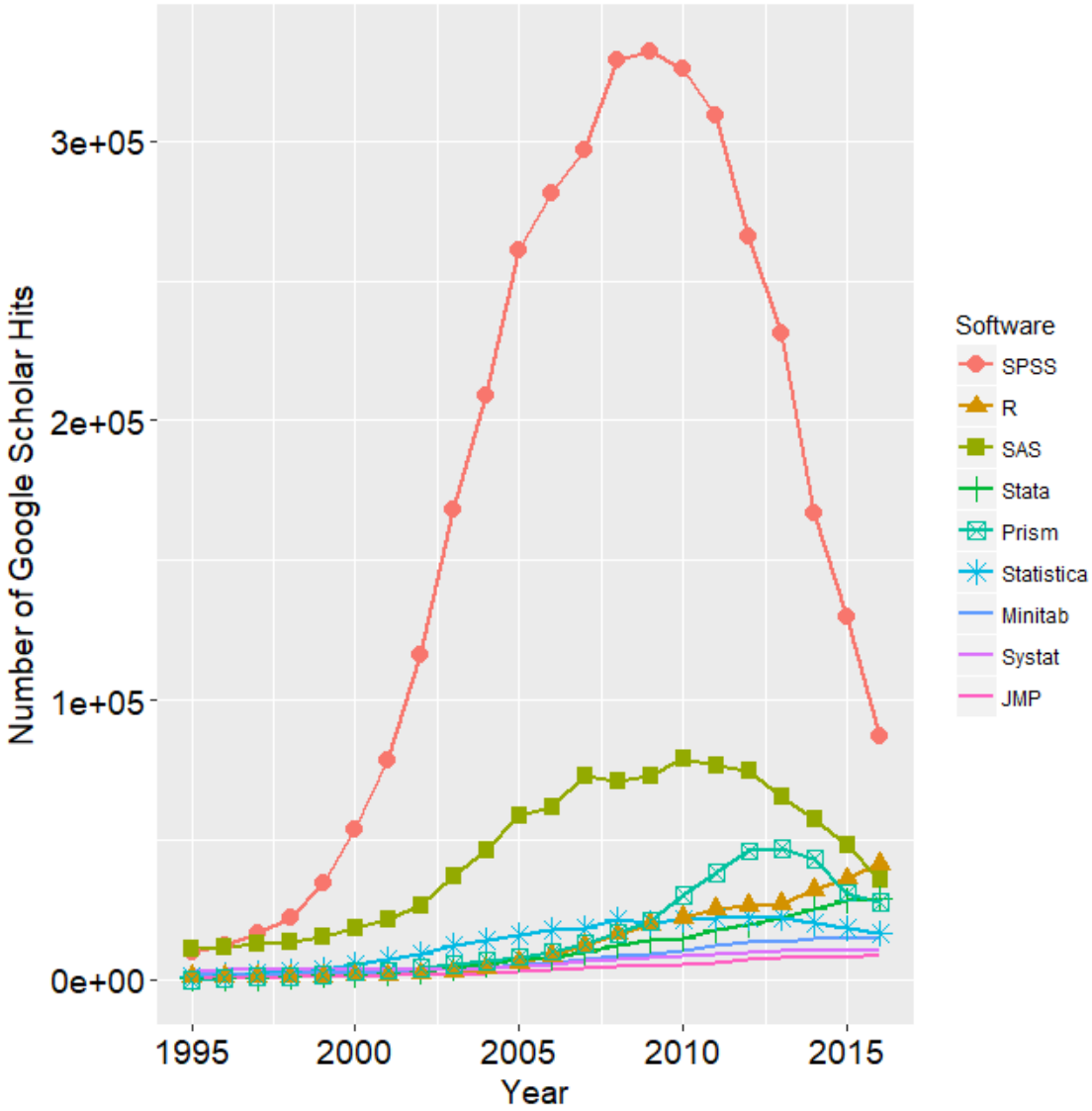
```
If each call to FUN returns a vector
returns an array of dimension c(n, di
If n equals 1, apply returns a vector
and an array of dimension dim(X)[MARG
the result has length 0 but not neces
dimension.

If the calls to FUN return vectors of
returns a list of length prod(dim(X)[
to MARGIN if this has length greater

In all cases the result is coerced by
the basic vector types before the dim
that (for example) factor results wil
character array.
```

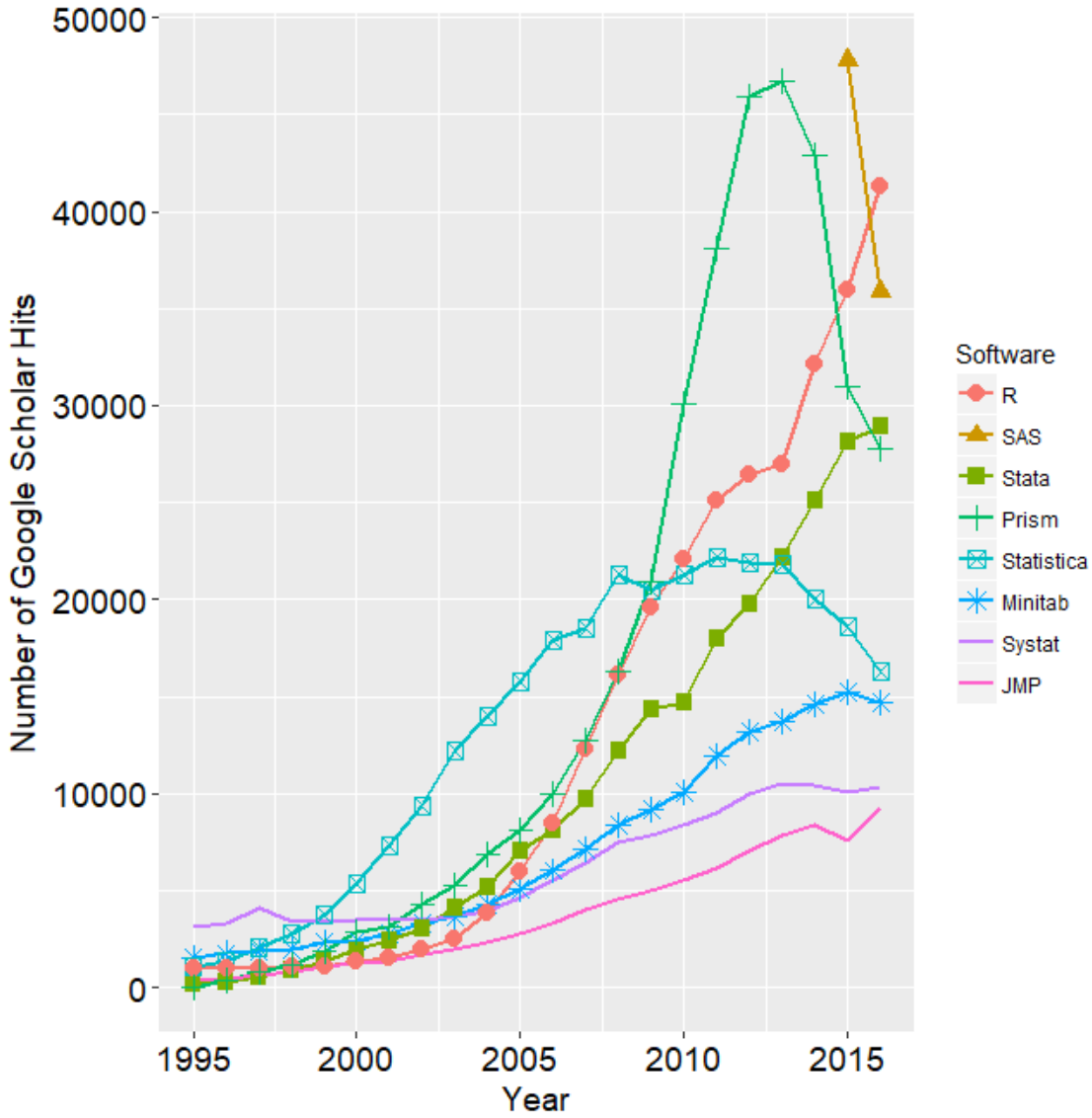


Google Scholar hits, by software (http://r4stats.com/popularity)



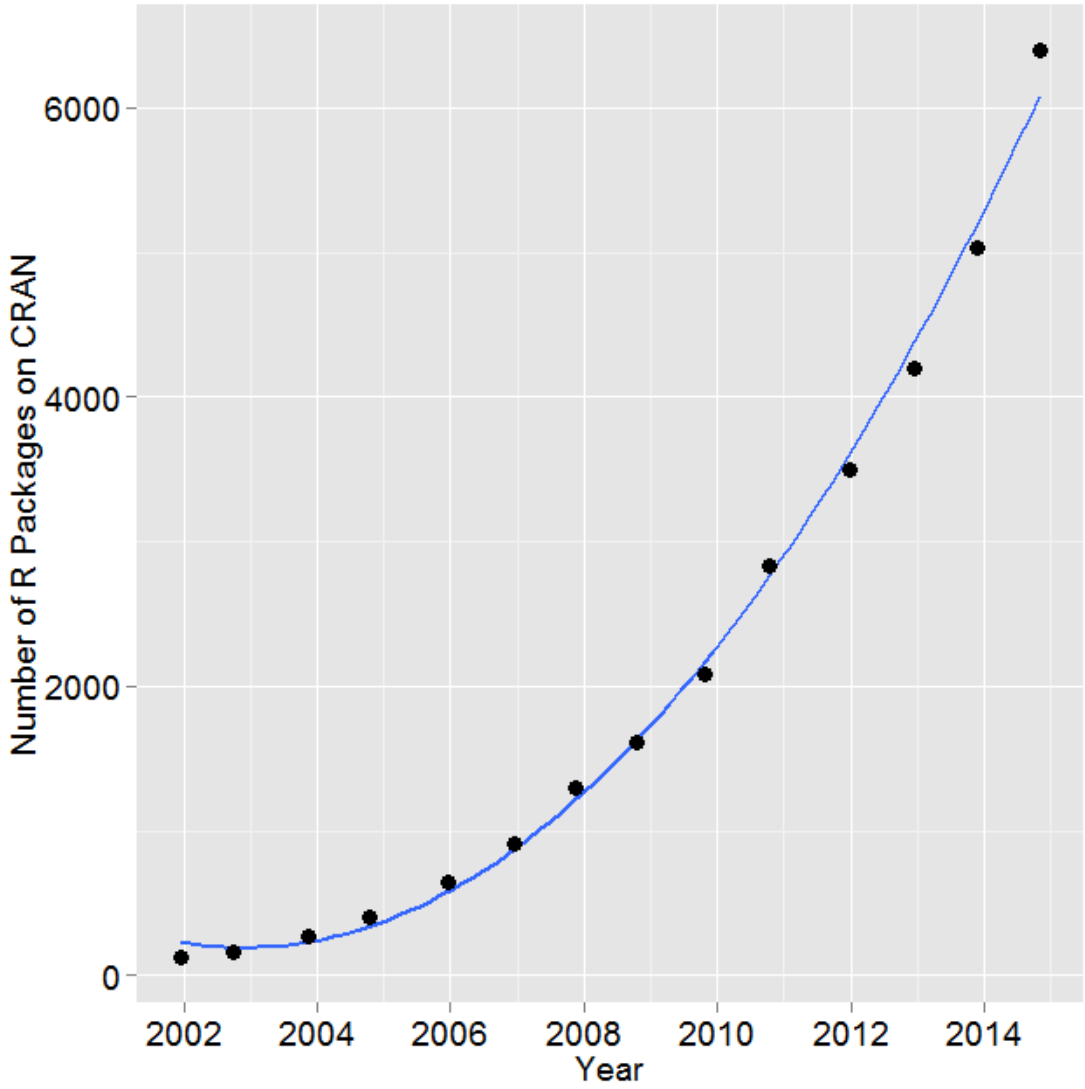
CENTRE FOR FERTILITY AND HEALTH

Google Scholar hits, by software (minus SPSS, SAS) (http://r4stats.com/popularity)



CENTRE FOR FERTILITY AND HEALTH

Number of R packages on CRAN (http://r4stats.com/popularity)



CENTRE FOR FERTILITY AND HEALTH

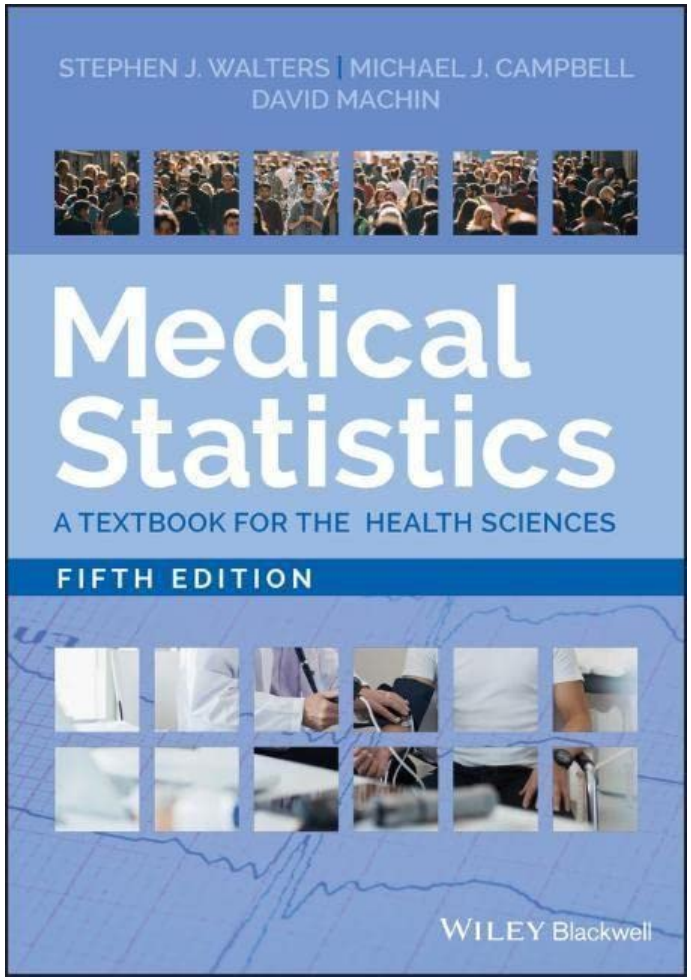
Other Software

- Python
  - (Closer to) a compiled language
  - Inherently faster, strong on very large data files
  - Suited for larger software development projects
  - Two weeks' course in Python $\longrightarrow$ "Data Scientist"(??)

- Julia
  - A "modernized" R
  - Just-In-Time (JIT) compiled
  - Faster and better structured than R
  - Can use R, Python and other libraries
  - Will it catch on?

CENTRE FOR FERTILITY AND HEALTH

TEXTBOOK



- Pretty decent
- Wide (but shallow) coverage of important topics
- Good as a reference text

CENTRE FOR FERTILITY AND HEALTH

DATA TYPES

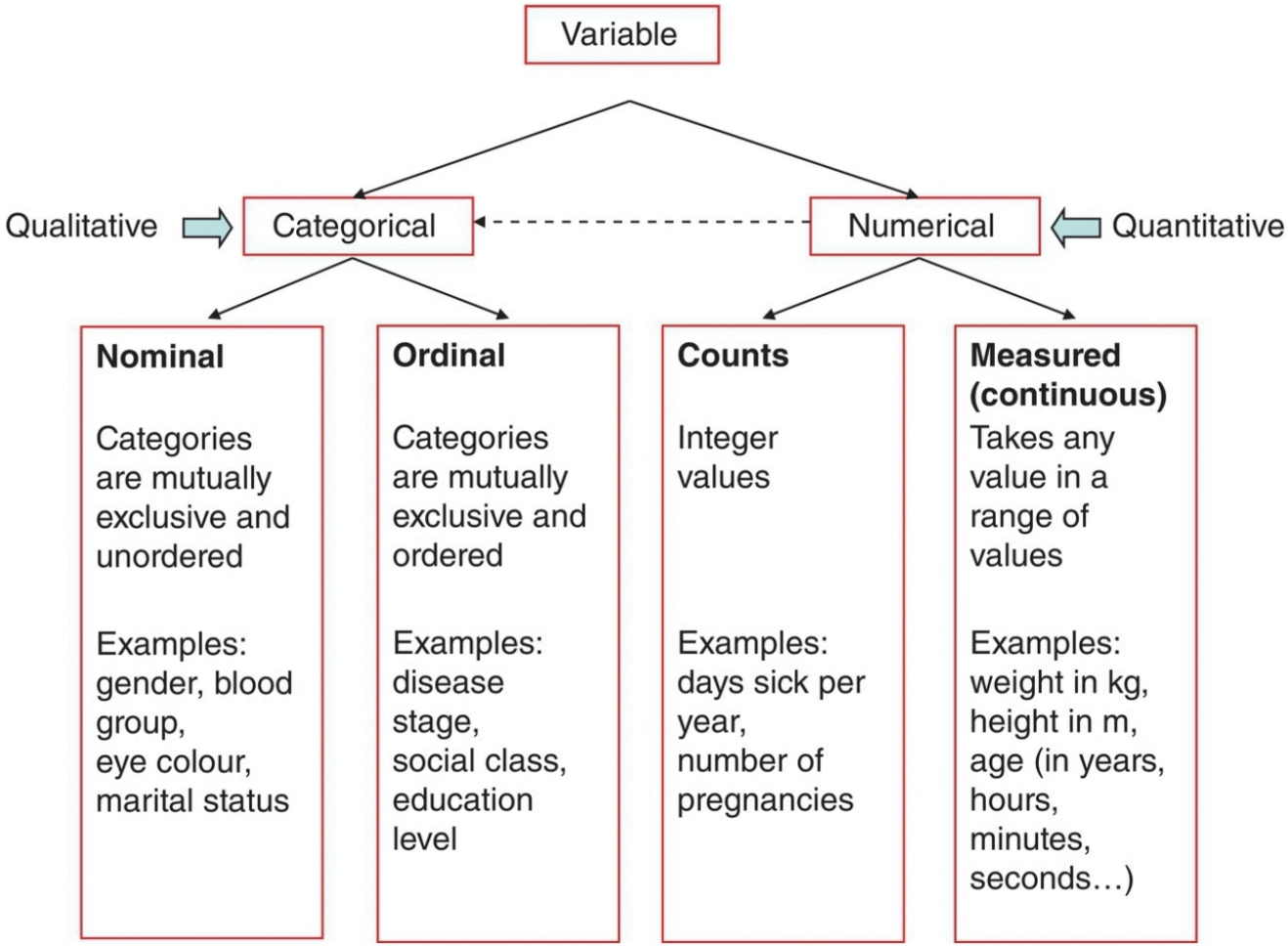


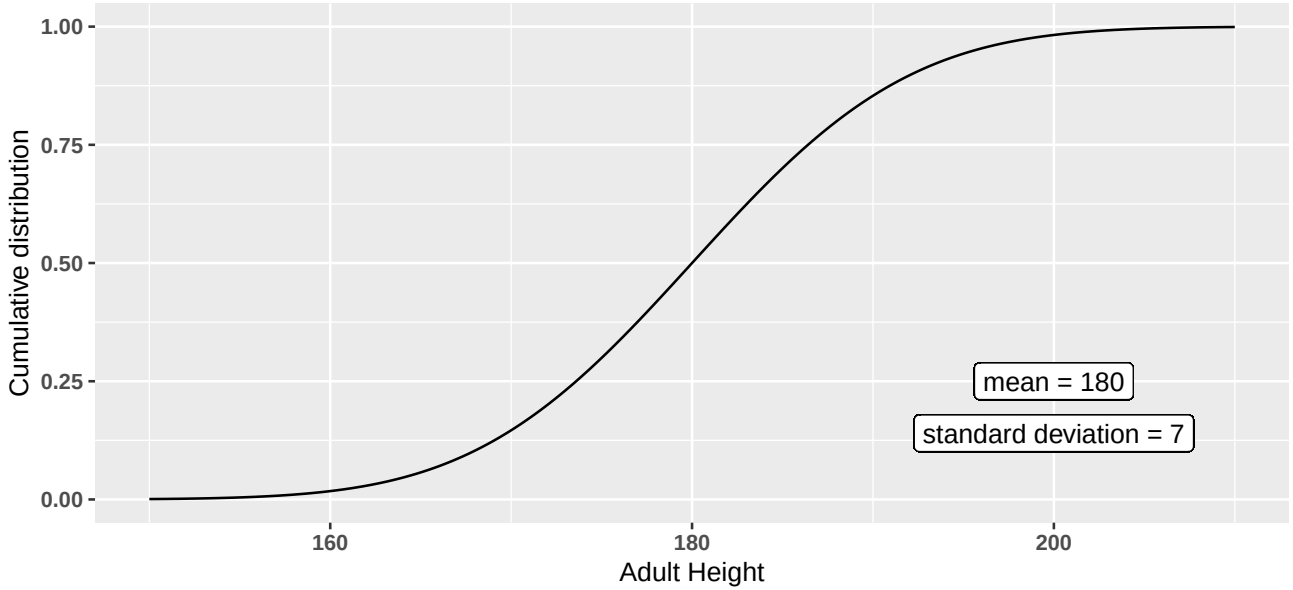
Figure 2.1 Broad classification of the different types of data with examples.

CENTRE FOR FERTILITY AND HEALTH

A **Random variable** X

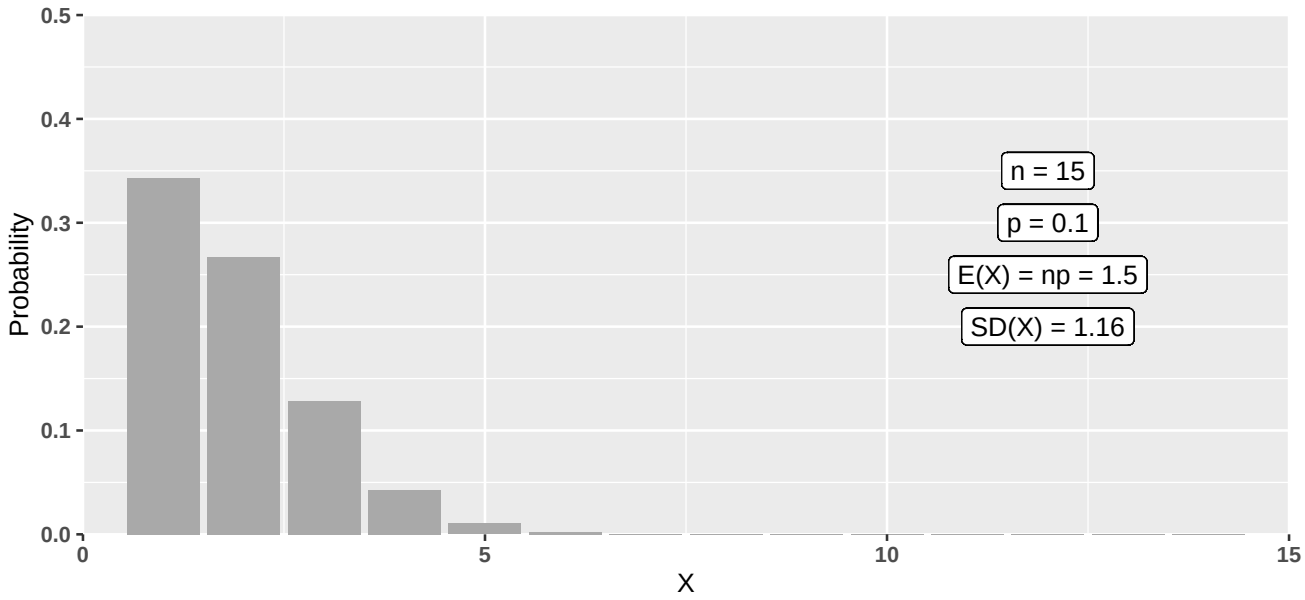
- X is *discrete* or *continuous* (or both)
- Typical examples are Binomial and Poisson (discrete) and Normal (continuous)
- $P(X = 15)$ is the probability that X is equal to 15, say.
- This is OK for e.g. a Poisson distribution, but not useful for a continuous distribution ($P(X = 15) = 0$ in all continuous distributions!)
- For continuous distributions, intervals make more sense: $P(X > 15)$, $P(X \leqslant 15)$ etc.
- E.g. birth weight: $P(BW < 2500)$ is the risk of Low Birth Weight
- Conditional probability:
  E.g. $P(BW < 2500|GA = 258)$, i.e. the probability of being Low Birth Weight conditional on Gestational Age at birth being 258 days.

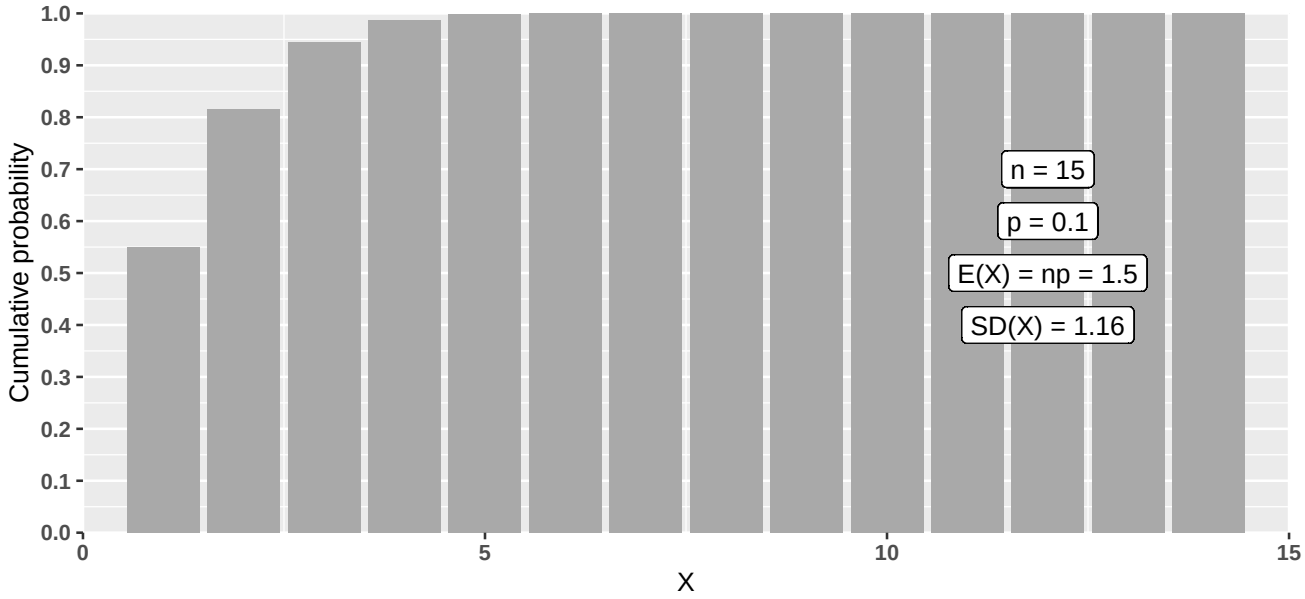
NORMAL DISTRIBUTION



CENTRE FOR FERTILITY AND HEALTH

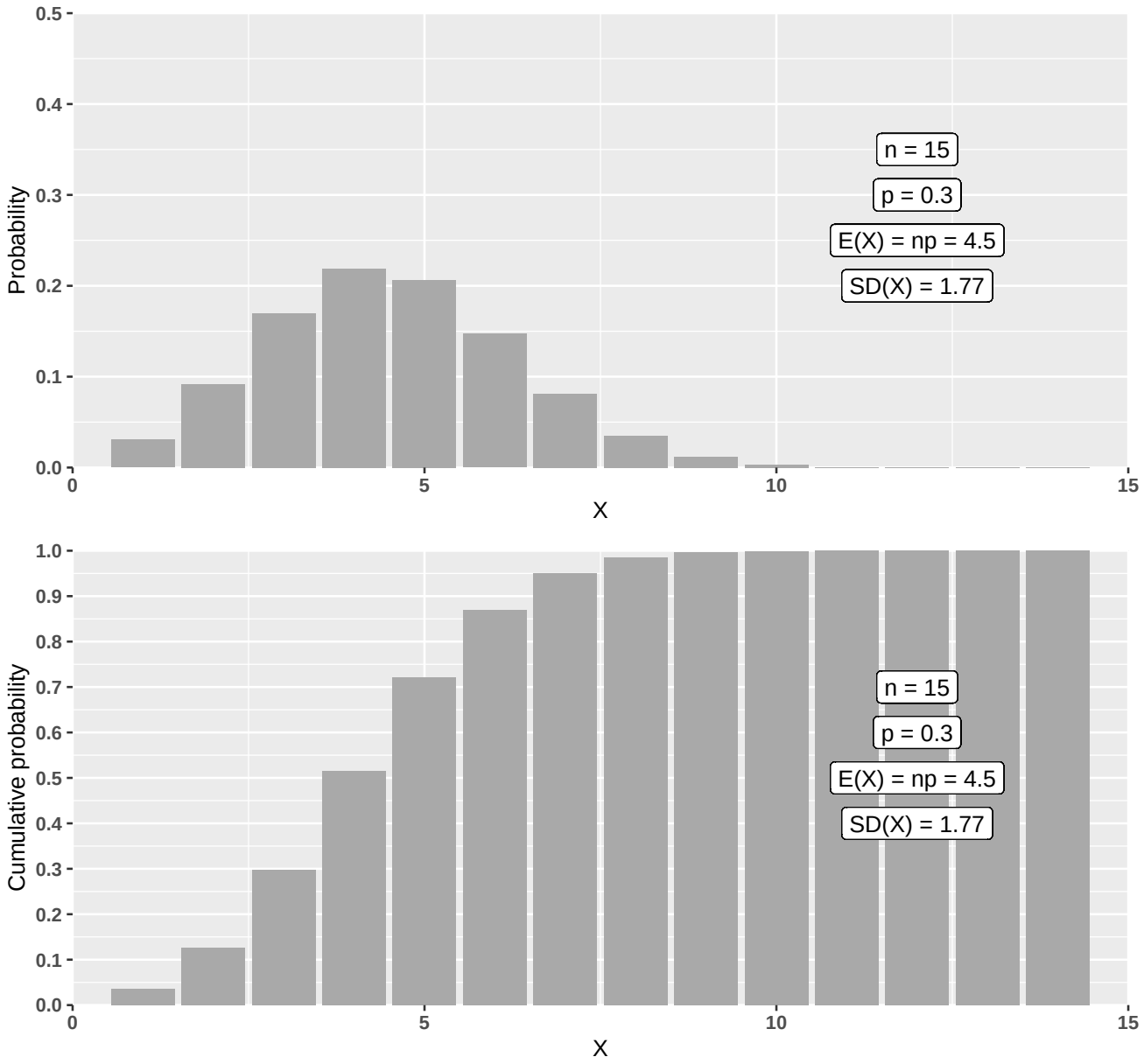
BINOMIAL DISTRIBUTION





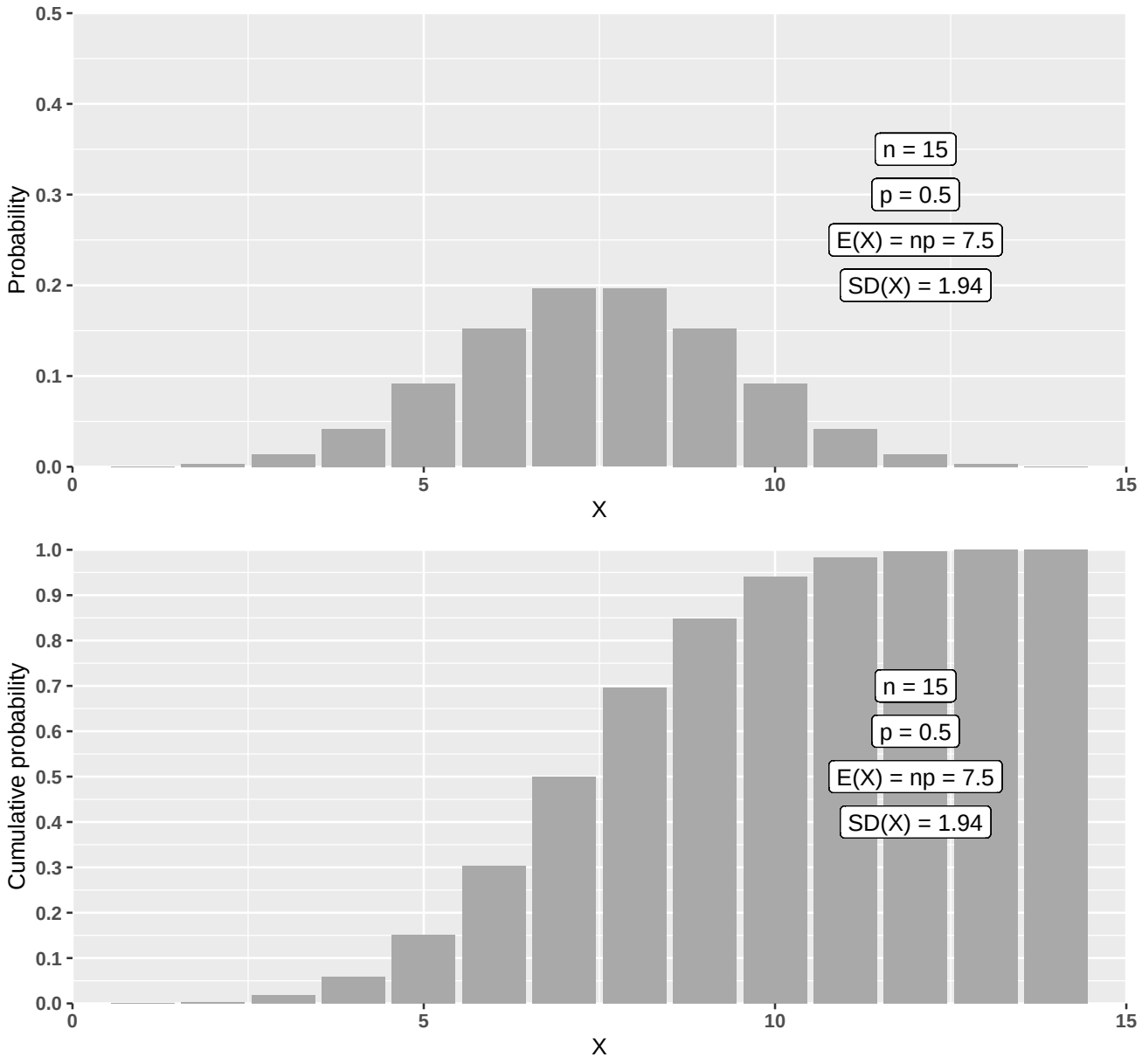
CENTRE FOR FERTILITY AND HEALTH

BINOMIAL DISTRIBUTION



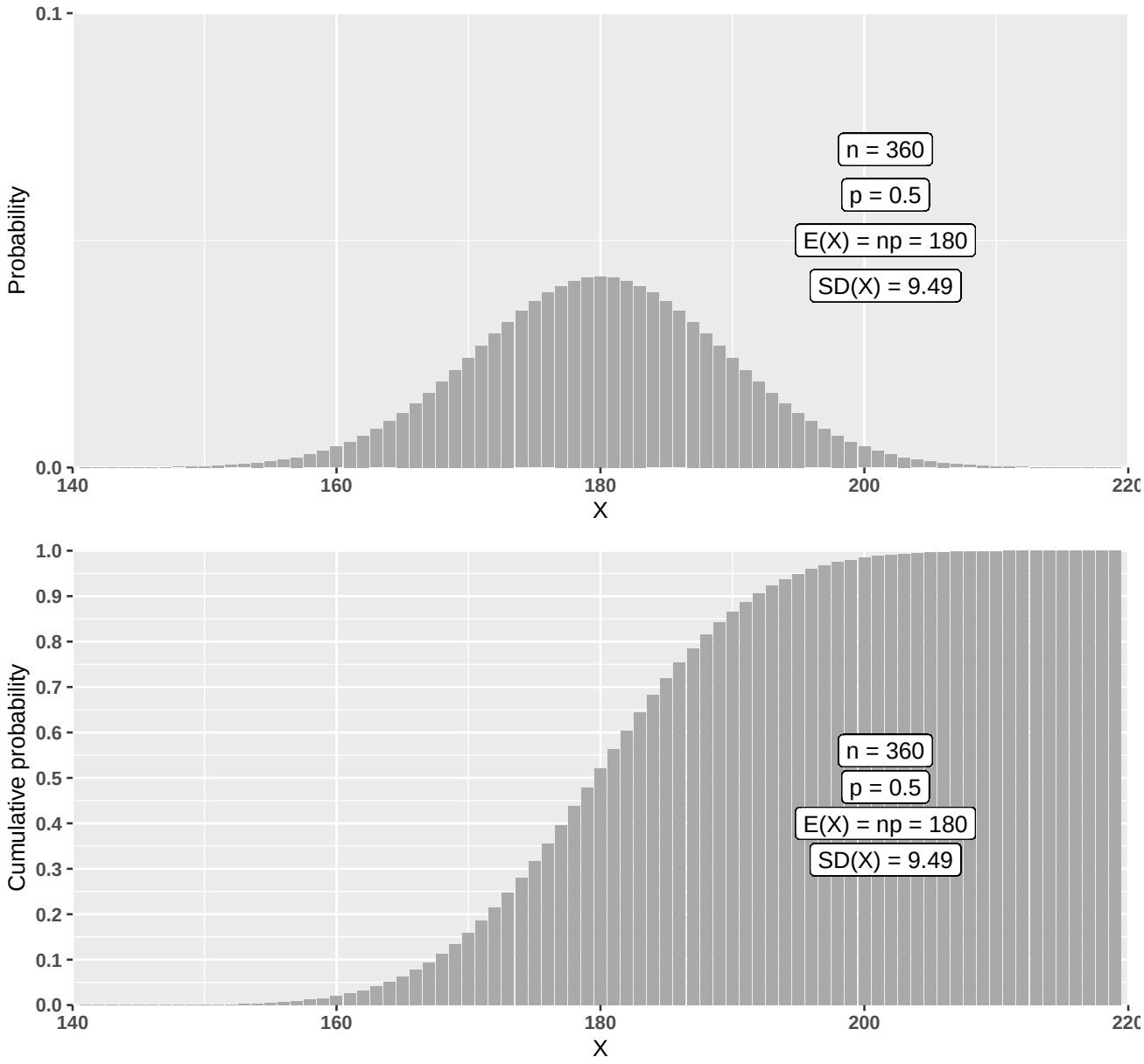
CENTRE FOR FERTILITY AND HEALTH

BINOMIAL DISTRIBUTION



CENTRE FOR FERTILITY AND HEALTH

BINOMIAL DISTRIBUTION



CENTRE FOR FERTILITY AND HEALTH

EXPECTED VALUE, STANDARD DEVIATION

- Expectation EX is the *mean* of X
  You often see
  - μ (mju) in a population
  - $\overline{X}$ in a sample
  - $\widehat{\mu} = \overline{X}$ since $\widehat{\mu}$ means *an estimate* of μ

$$\overline{X} = \frac{1}{n}\Sigma_{i=1}^{n} X_i = \frac{1}{n}\left(X_1 + X_2 + \cdots + X_n\right)$$

- $SD(X)$ is the *Standard deviation* of X
  You often see
  - σ (sigma) in a population
  - s or $SD(X)$ in a sample
  - $\widehat{\sigma} = s = SD(X)$ since $\widehat{\sigma}$ means *an estimate* of σ

$$s^2 = \frac{1}{n-1}\Sigma_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\left\{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \cdots + (X_n - \overline{X})\right\}^2$$

Student's t-tests

Traditionally,

- One-sample t-test
- Two independent samples t-test
- Two paired samples t-test (take difference and use one-sample)

- With small sample sizes, say, less than 30, a t-distribution (with degrees of freedom) is used instead of a Normal distribution

---

Today often replaced by regression models

- Larger sample sizes are the norm
- Software chooses distribution automatically
- More flexible in terms of adjustments etc.
- Paired samples t-test replaced by mixed-effects models

CENTRE FOR FERTILITY AND HEALTH

Parametric versus non-parametric

The Good, The Bad, and The Ugly

The good(?), old days:

- Small sample size
- Test for Normal distribution
- If not very Normal, choose non-parametric (for instance the Wilcoxon-Mann-Whitney test)
- Else use t-test

Dilemma, model choice:

- Test for Normal distribution insensitive when n small
- When n is big, a normal distribution will often be rejected regardless
- ... BUT when n is big, the Central Limit Theorem helps you, so you usually don't need to assume a completely normal distribution anyway, and there is less need for the non-parametric

Dilemma, power:

- When n is small, parametric tests have little power...
- ... but non-parametric tests are even worse...
- And non-parametric tests usually only provide p-values

CENTRE FOR FERTILITY AND HEALTH

Conclusion:

Parametric versus non-parametric

The Good, The Bad, and The Ugly

Conclusion:

Get over it. Get more data.

CENTRE FOR FERTILITY AND HEALTH

Parametric versus non-parametric

The Good, The Bad, and The Ugly

Conclusion:

Get over it. Get more data.

(Having said that, used in the right places, non-parametric and semi-parametric methods can actually be great! Particularly when n is big!)

CENTRE FOR FERTILITY AND HEALTH