



Fundamental Concepts in Medical Statistics



Håkon K. Gjessing

Professor/Principal Investigator

⇒ Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo
– a Norwegian Research Council Centre of Excellence

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



Kathmandu
Thursday, 1 October 2026



<https://gjessing.page/fundamental-concepts-in-medical-statistics-kathmandu-1-7-october-2026>



Fundamental concepts in medical statistics, Kathmandu 1-7 October 2026

+ Intro to R & R Studio pre-course 30 September 2026

COURSE RESOURCE PAGE

Organizers and Supporters

Course coordinator

Arun K Sharma, Department of Pediatrics

Organizers

Department of Pediatrics, MMC, Institute of Medicine and Research Department, Institute of Medicine

In collaboration with

- Centre for Fertility and Health, Norwegian Institute of Public Health (<https://www.fhi.no/en/more/research-centres/Centre-for-fertility-and-health>)
- Bergen Centre for Ethics and Priority Setting in Health, University of Bergen, Norway (<https://www.uib.no/bceps>), Director Ingvild Fossgard Sandøy

- 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 previous courses)
- but it should all be very relevant

Schedule: Thursday, 1 October

FOCUS OF THE DAY: Overview, and fundamental concepts			
Timing	Planned activity	Moderator	
09:00- 10:00	Welcome and startup • Welcome/introduction • Quick overview of course and practical info • Quick overview of basic concepts (Participants are expected to have basic understanding of/experience with basic statistical concepts)	Håkon K. Gjessing	
10:00- 12:00	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.	Håkon K. Gjessing	
12:00- 13:00	Lunch		
13:00- 14:00	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi	
14:00- 15:00	Sample size and statistical power / statistical precision • What it means	Håkon K. Gjessing	

		• How to restrict the probability of Type I/Type II errors • How to compute sample size needed for sufficient precision of effect estimate • How to compute sample size for study planning, simple situations	
15:00- 15:15	Tea break		
15:15- 16:30	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi	

Schedule: Friday, 2 October

FOCUS OF THE DAY: Generalized linear models (GLMs)			
	Timing	Planned activity	Moderator
	09:00- 11:00	Linear regression • Standard linear regression • Multiple linear regression, confounding, interpretation • Basic regression diagnostics	Håkon K. Gjessing
	11:00- 12:00	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi
	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	Håkon K. Gjessing
	15:00- 15:15	Tea break	
	15:15- 16:30	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi

Schedule: Monday, 5 October

FOCUS OF THE DAY: Mixed bag with focus on RCTs, other experimental designs			
	Timing	Planned activity	Moderator
	09:00- 10:00	More on Binomial Regression • Link functions • How to choose suitable effect measures, such as RD, RR, OR • What are the implications of these choices?	Håkon K. Gjessing

	10:00- 11:00	Interaction / effect measure modification • Basic concepts • Scale dependence • Examples How to estimate, variable coding, and dummy variables • Testing / estimating differences in effects across subgroups	Håkon K. Gjessing
	11:00- 12:00	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi
	12:00- 13:00	Lunch	
	13:00- 14:00	Some fundamental ideas behind RCTs • 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	Håkon K. Gjessing
	14:00- 14:15	Tea break	
	14:15- 15:00	A "Modeling Case study" using binomial, Poisson, and time-to-event in COVID-19 modeling	Håkon K. Gjessing
	15:15- 15:45	Weighted analyses • Frequency weights • Sampling weights • Importance weights	Håkon K. Gjessing
	15:45- 16:30	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi

Schedule: Tuesday, 6 October

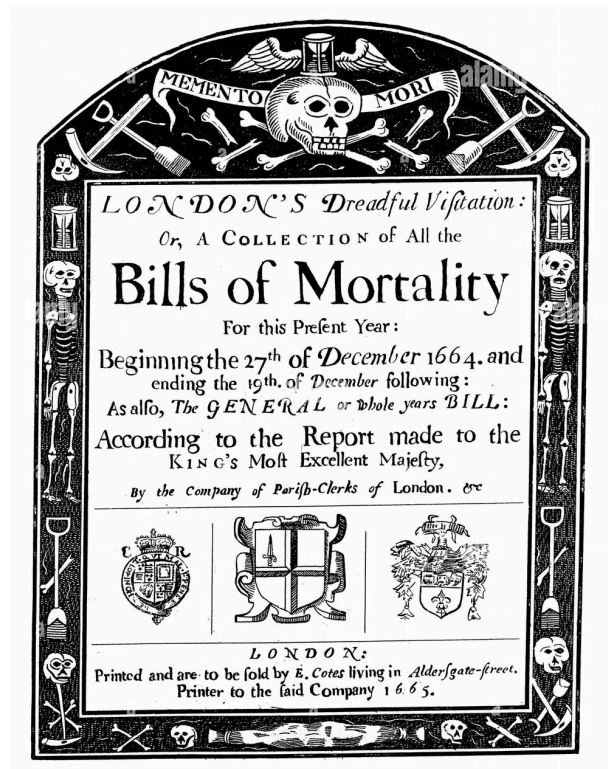
FOCUS OF THE DAY: Count data, standardization		
Timing	Planned activity	Moderator
09:00- 11:00	Poisson regression • How and when to use • How to handle varying exposure time, use of offset • Examples from Covid-19 pandemic analyses	Håkon K. Gjessing
11:00- 12:00	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi
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	Håkon K. Gjessing
15:00- 15:15	Tea break	
15:15- 16:30	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi

FOCUS OF THE DAY: Survival and event history analysis

Timing	Planned activity	Moderator
09:00- 11:00	Survival Analysis / Event History Analysis • Introduction to survival/event history analysis • Kaplan-Meier • Cox regression	Håkon K. Gjessing
11:00- 12:00	Exercises	Dagrun Slettebø Daltveit, Ayushi and Devashish Tripathi
12:00- 13:00	Lunch	
13:00- 14:00	Additional topics in survival analysis • Alternative models, such as additive hazards model	Håkon K. Gjessing
14:00- 15:15	Summary discussion and evaluation	All
15:15- 15:30	Tea	
15:30- 16:00	Valedictory session	All

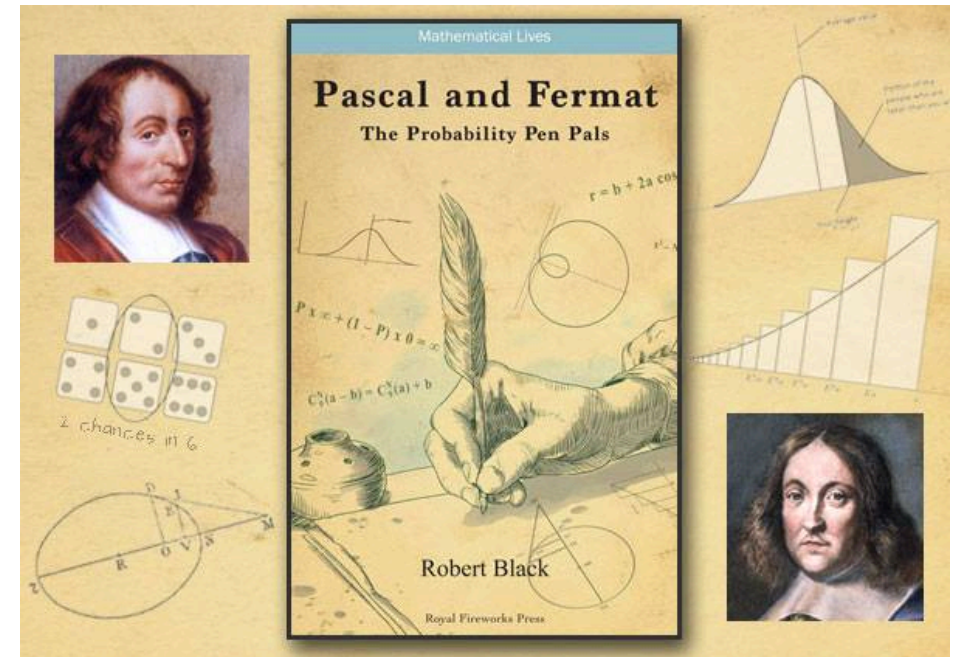
Statistics (1665)

John Graunt, Bills of Mortality

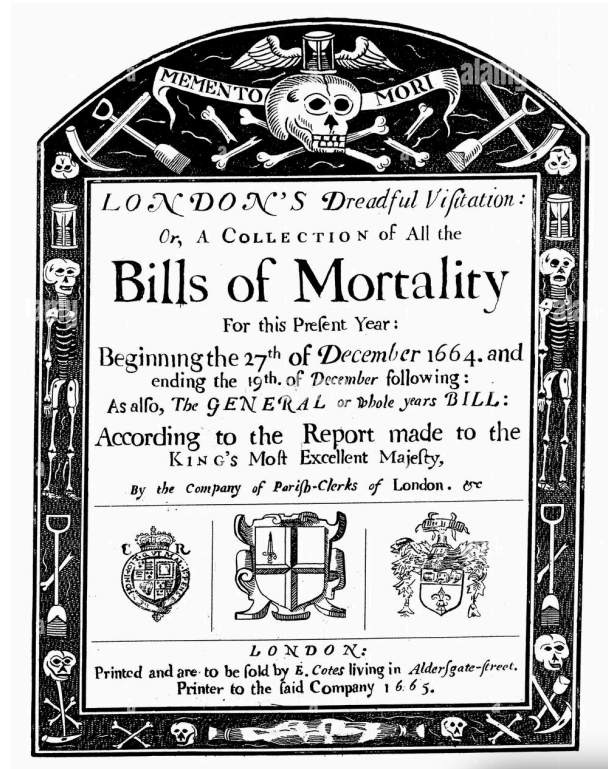


Probability Theory (≈ 1654)

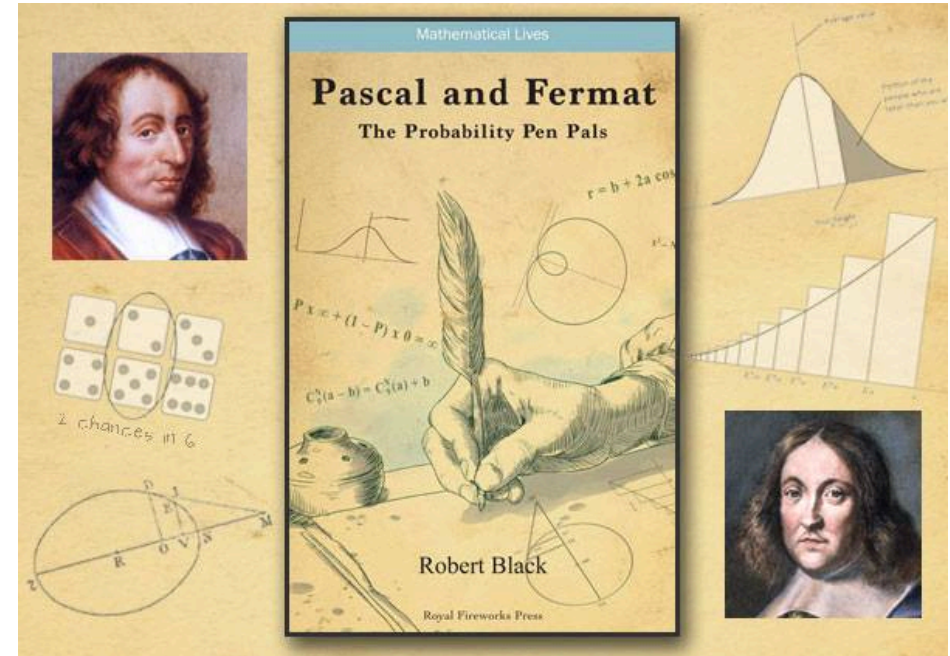
Pascal + Fermat, How to fairly end a game



Statistics (1665)



Probability Theory (≈ 1654)



Love Child: Mathematical Statistics

What's in a name?

Statistics

- Same root as State, Status, and Station
- The term **Statistical Science** has been coined as more suitable
- “Standard” statistical modelling is sometimes known as **Frequentist Statistics**
- Another approach (with many similarities and some differences) is known as **Bayesian Statistics**
- **Mathematical Statistics** is heavy on mathematical theory, sometimes with applications in mind...
- **Data Science** is a term stolen by people with a 2 week course in Python programming and linear regression (but now claimed by some statistics departments)
- **Statistical Learning** combines machine learning techniques with statistical modeling
- **Machine Learning** (and AIs) are quite heavy on the computational side.

Why Medical Statistics?

New England Journal of Medicine

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

The eleven most important developments in medicine over the past millennium

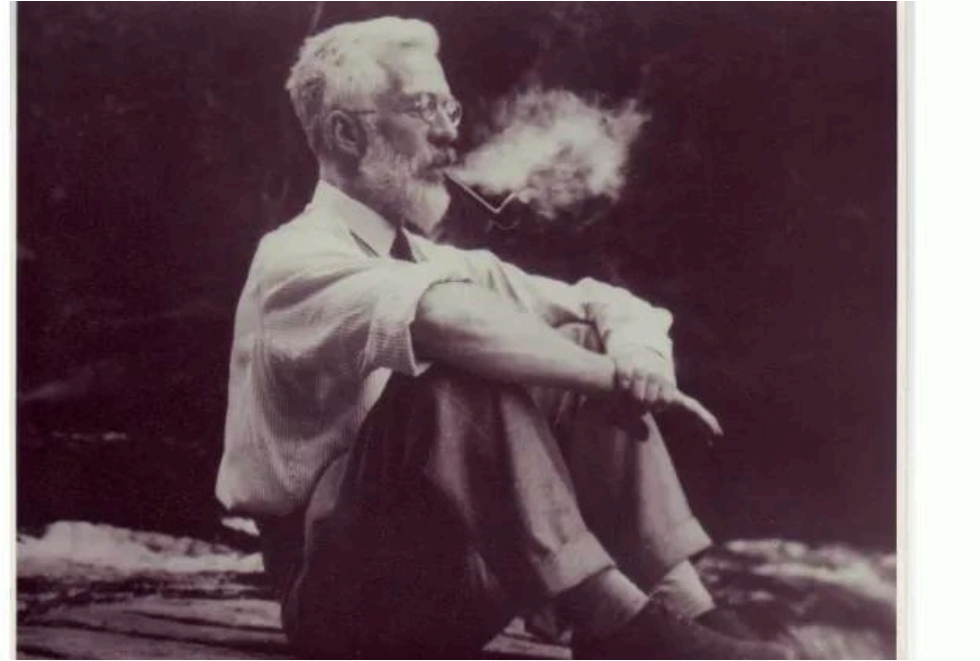
- Elucidation of human anatomy and physiology
- Discovery of cells and their substructures
- Elucidation of the chemistry of life
- *Application of statistics to medicine*
- Development of anesthesia
- Discovery of the relation of microbes to disease
- Elucidation of inheritance and genetics
- Knowledge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

Before randomization

Apparent treatment effect = True treatment effect
+ Effect of observed confounders
+ Effect of unobserved confounders
+ Random noise

After randomization

Apparent treatment effect = True treatment effect
+ **Noise from** observed confounders
+ **Noise from** unobserved confounders
+ Random noise



Ronald Fisher

A digression into

Statistical Software

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/>
 - BUT: With AIs, writing sensible R code has become much easier
- Stata
 - Stata commands, manuals, and documentation are in general of **AMAZING quality!**
- SPSS (Statistical Pseudo Science System)
- SAS
- Python
- Julia



wizard



=



muggle

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

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)

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!”

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!”

- But how about **SAS**, then?

Yes, how about SAS, then?



R: Wonderfully clear & pedagogical help pages

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

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.

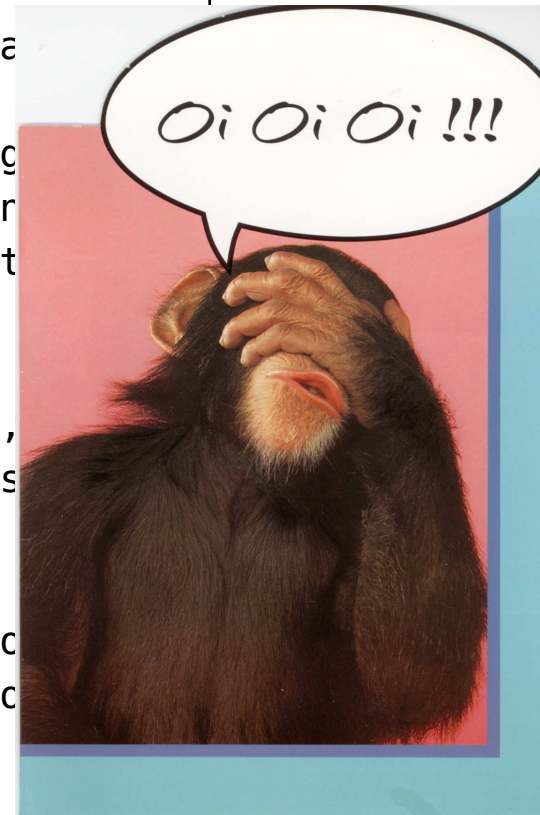
R: Wonderfully clear & pedagogical help pages

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

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

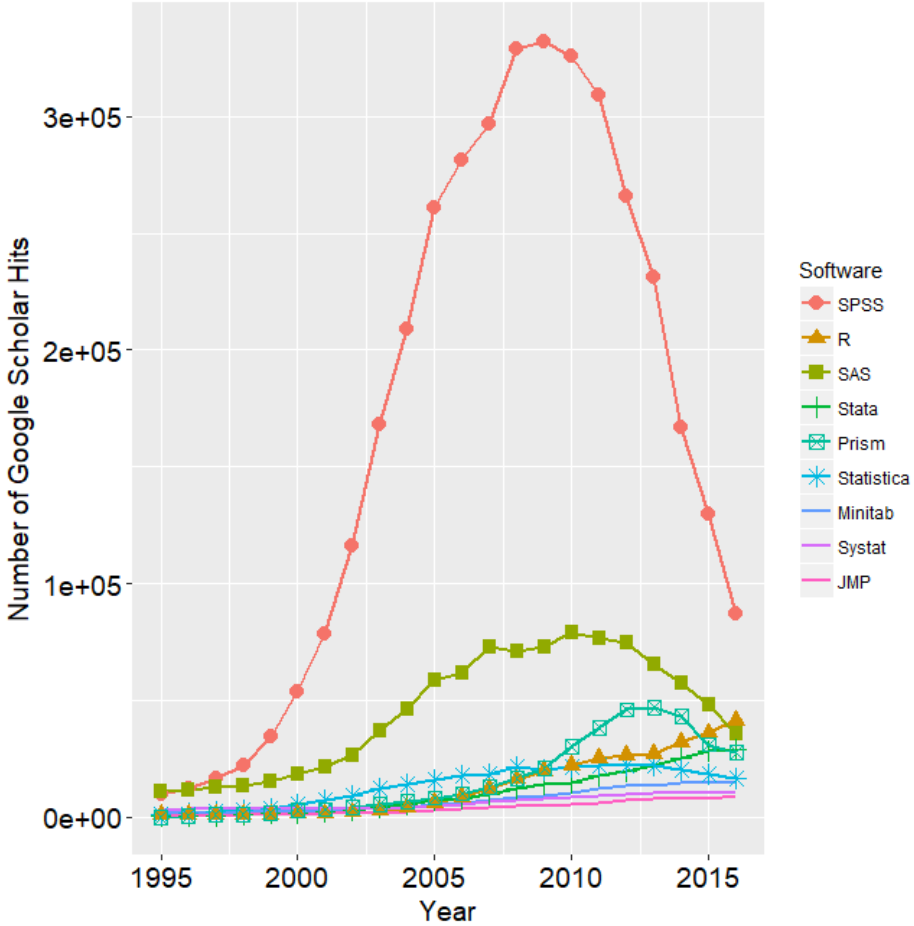
If the calls to FUN return vectors of different lengths, returns a list of length `prod(dim(X)[MARGIN])` with dimensions to MARGIN if this has length greater than one.

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.



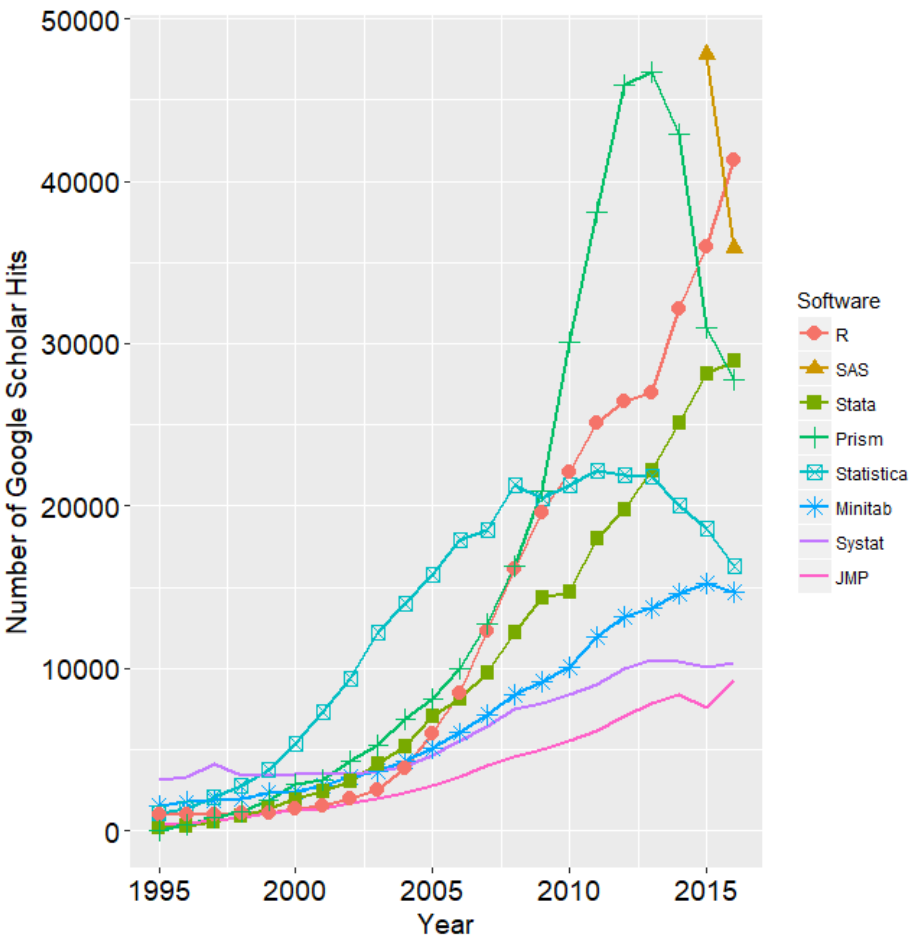
Google Scholar hits, by software

<http://r4stats.com/popularity>



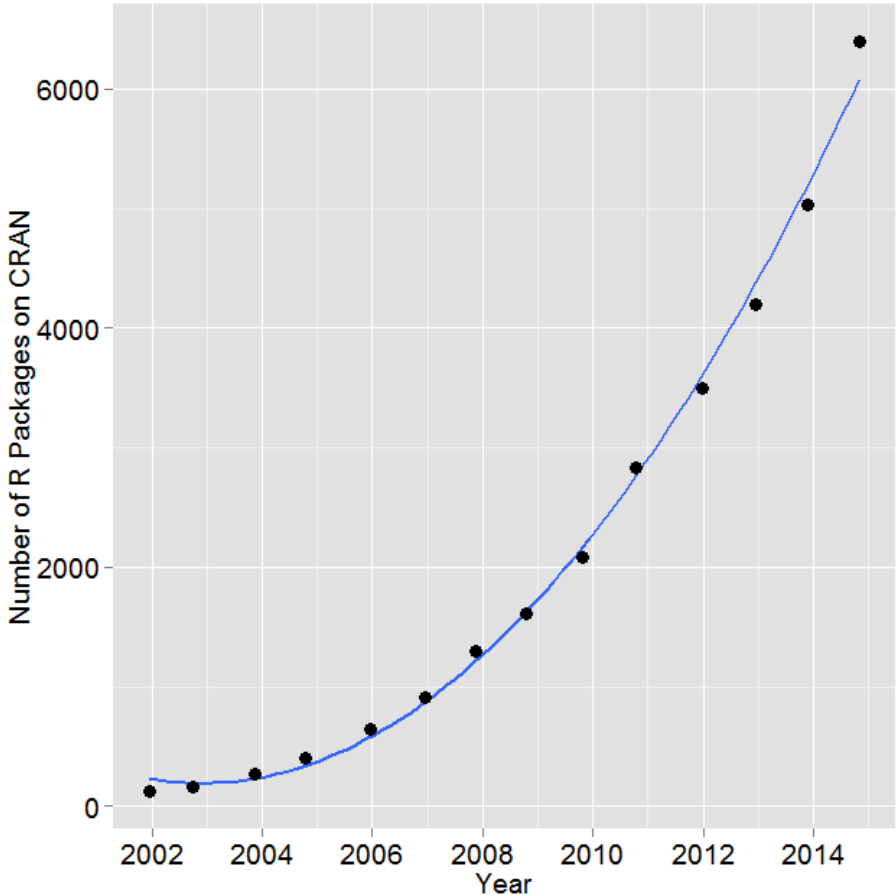
Google Scholar hits, by software (minus SPSS, SAS)

<http://r4stats.com/popularity>

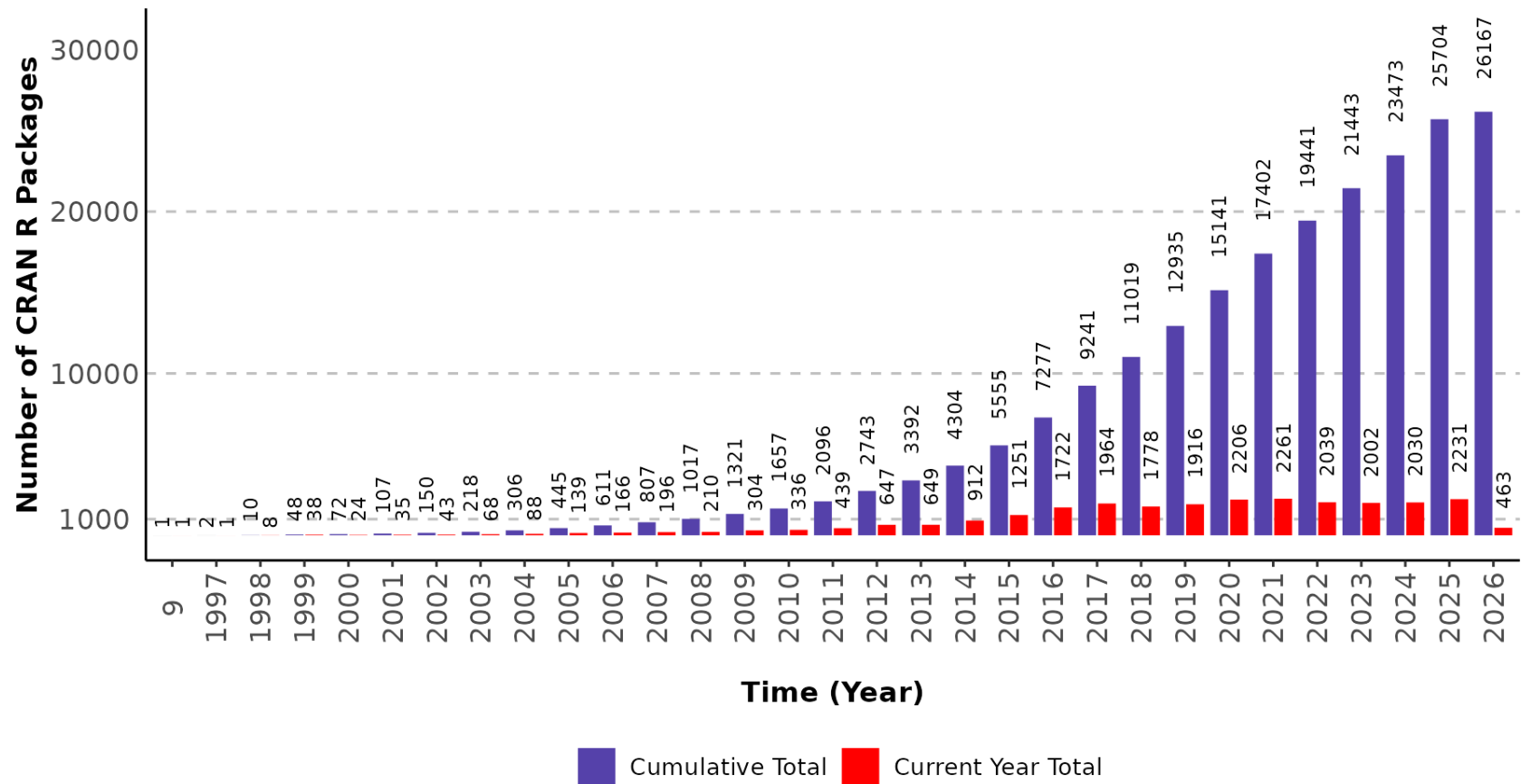


Number of R packages on CRAN, cumulative

<http://r4stats.com/popularity>



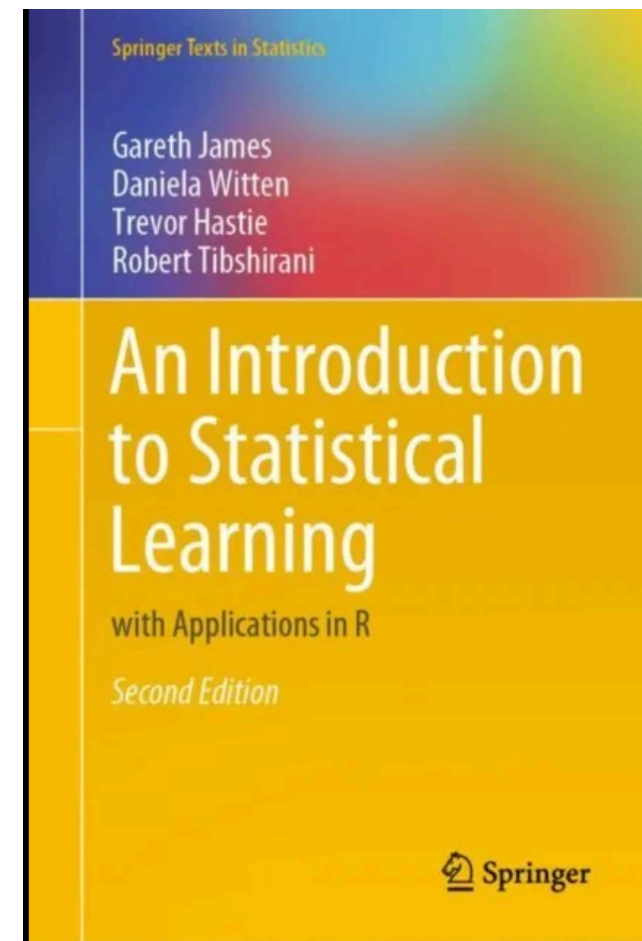
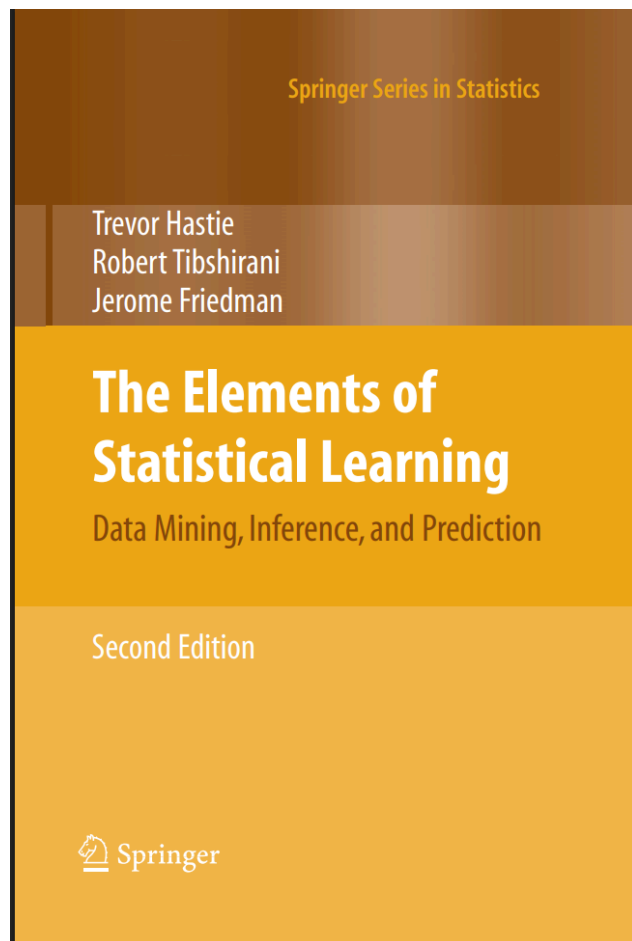
Number of R packages (on CRAN), cumulative

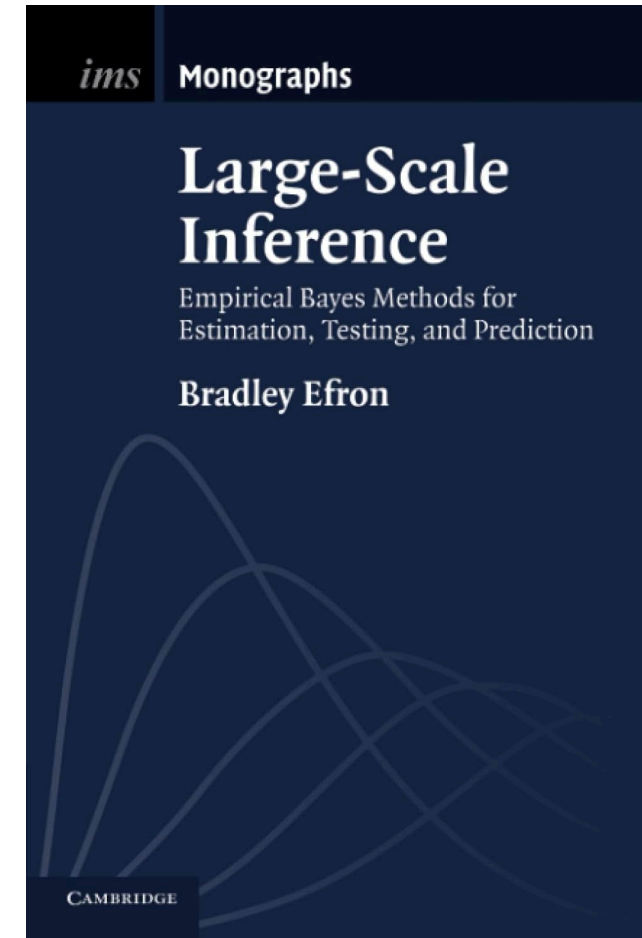
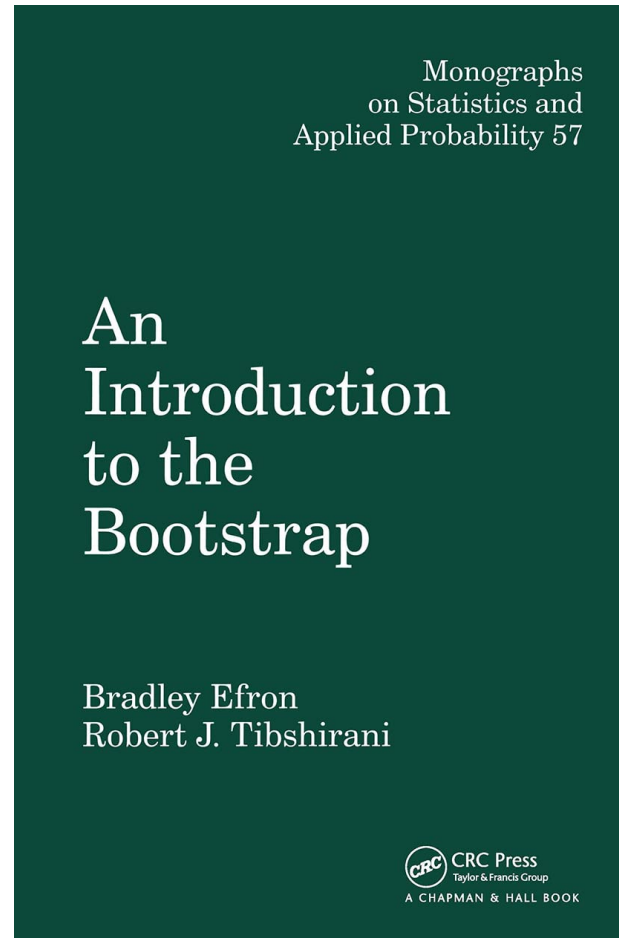


Data source: www.RPKG.net, 2026-03-17

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 → “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?





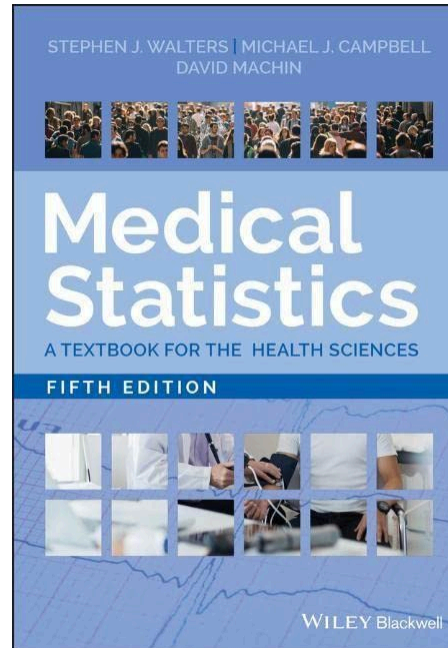
Machine Learning / AI

- Typically computationally highly demanding
- Typically very data hungry
- Have been developed for decades, but now expands along with the recent huge increases in available data and computing power
- Often appears like a “Black box”
- Typically focused purely on *prediction*, and less on structured models
- ... but a lot of recent work tries to provide more structure, transparency and control

In contrast to, for instance, **Causal Modeling**:

- Highly structured modeling approach to provide (if possible) a causal interpretation of results

Recommended reading ahead of course



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

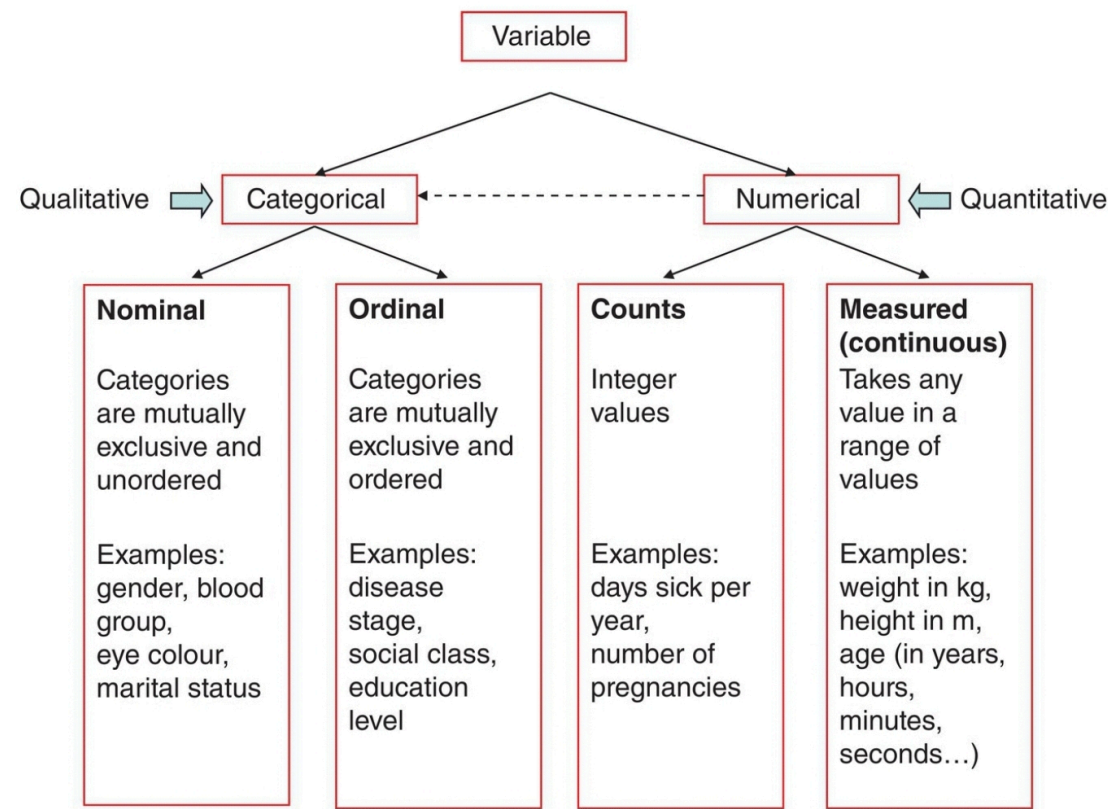
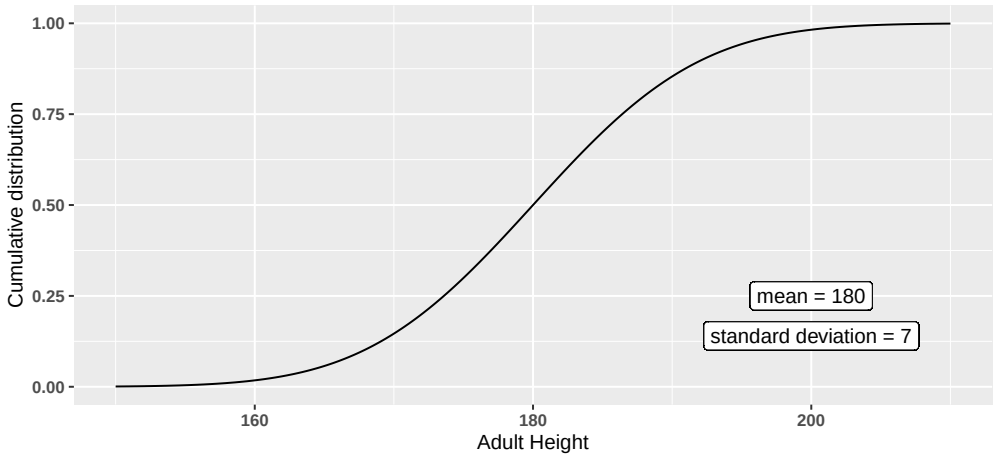
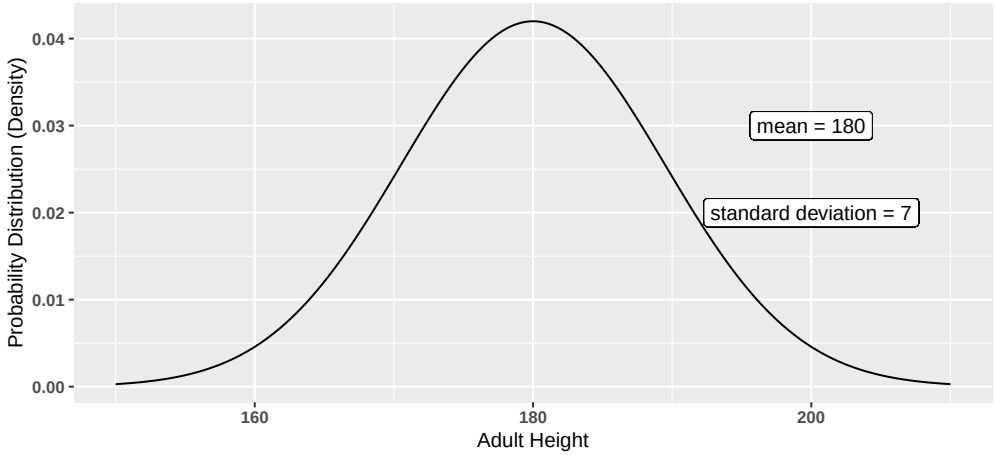


Figure 2.1 Broad classification of the different types of data with examples.

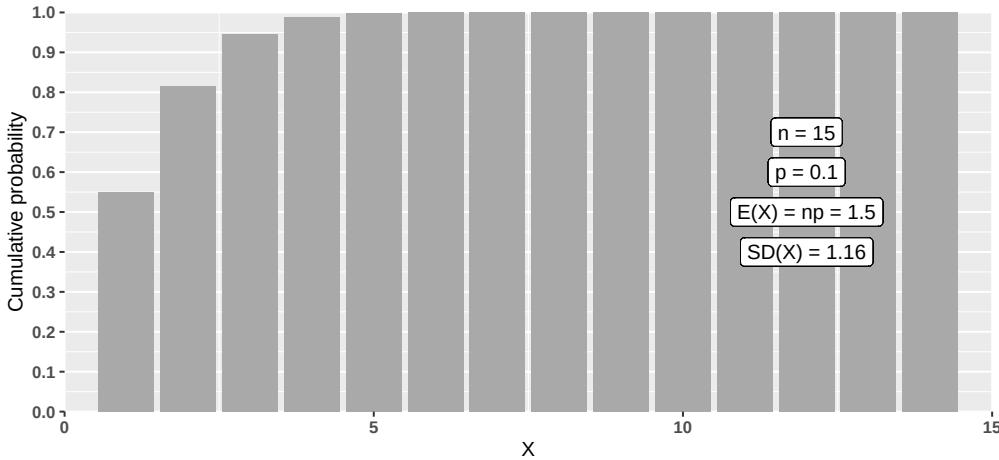
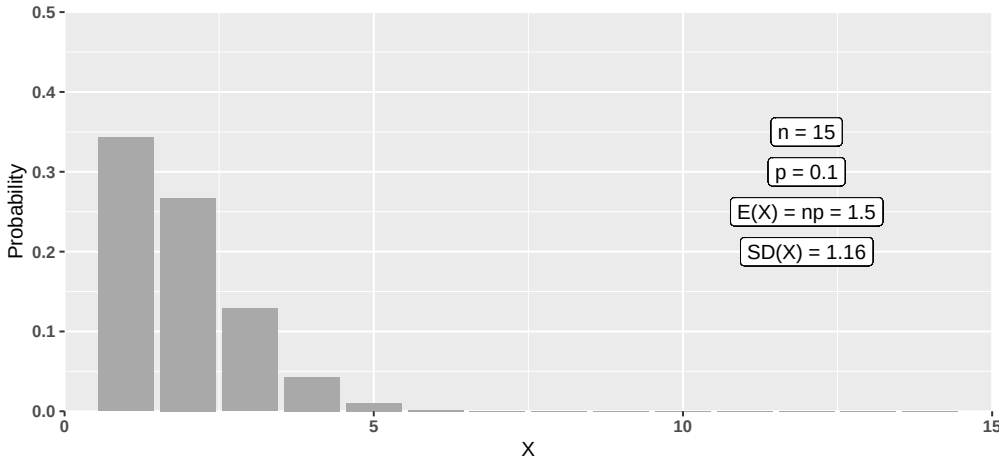
A Random variable X

- 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.
 - OK for discrete distributions, e.g. a Poisson distribution
 - Not useful for a continuous distribution since $P(X = 15) = 0$ in all continuous distributions!
- For continuous distributions, intervals make more sense: $P(X > 15)$, $P(X \leq 15)$ etc.
- E.g. birth weight: $P(BW < 2500)$ is the risk of Low Birth Weight
- Conditional probability:
 $P(BW < 2500 \mid GA = 258)$ is the probability of being Low Birth Weight conditional on Gestational Age at birth being 258 days.

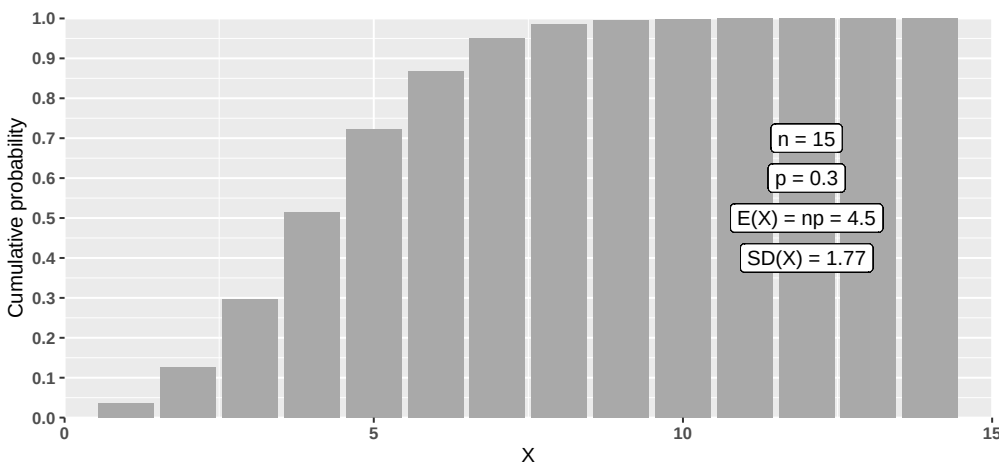
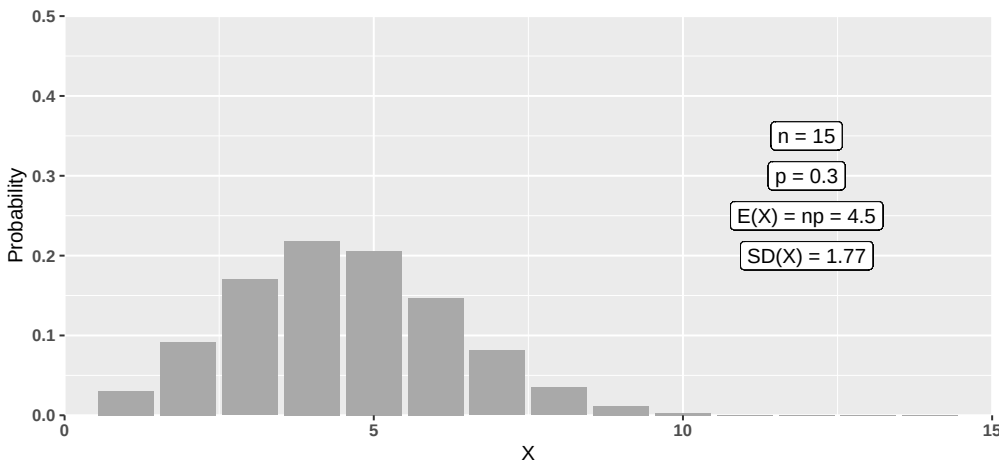
Normal distribution



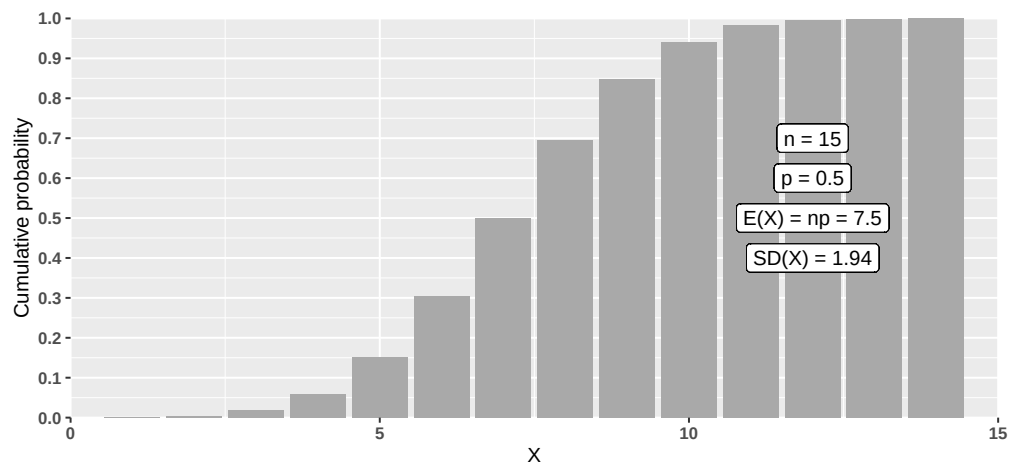
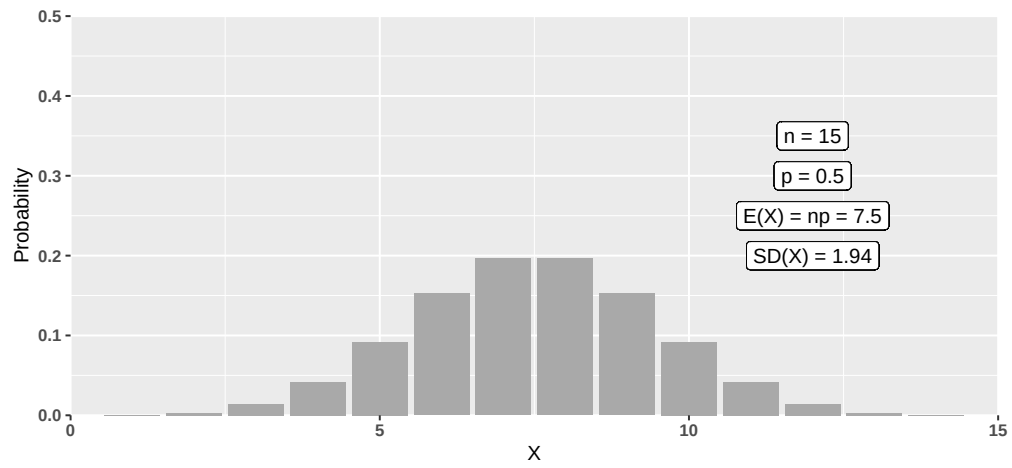
Binomial distribution



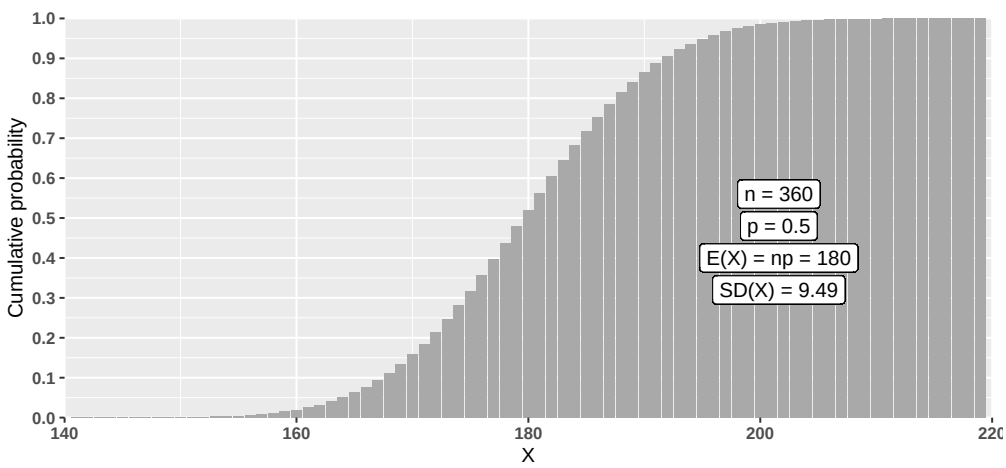
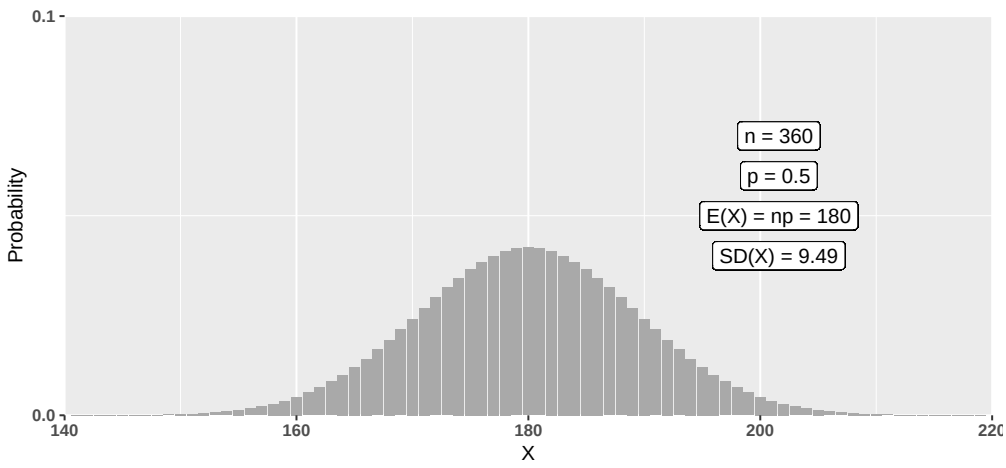
Binomial distribution



Binomial distribution



Binomial distribution



Expected value, standard deviation

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

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} (X_1 + X_2 + \dots + X_n)$$

Expected value, standard deviation

- $SD(X)$ is the *Standard deviation* of X

You often see

- σ (sigma) in a population
- s or $SD(X)$ in a sample
- $\hat{\sigma} = s = SD(X)$ since $\hat{\sigma}$ means *an estimate* of σ

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left\{ (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2 \right\}$$

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

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

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!**)

“Those who ignore Statistics
are condemned to reinvent it.”
— Bradley Efron