STAT 110: Week 12

University of Otago

Outline

- Compare different study designs
  - Focus on understanding relationship between variables
- Experiments
  - Introduce randomized control trial
- Observational data
  - Confounding
- Correlation and causation
- Causal inference

Data: Whickham smoking study

- A twenty year study was conducted on 1314 women in Whickham, England[1]
  - Variable 1: Information on mortality (alive/dead)
  - Variable 2: Smoking status at baseline (smoker: yes/no)
- Represent data in a contingency table

```
##         Smoking_status
## Outcome   No  Yes  Sum
##   Alive  502  443  945
##   Dead   230  139  369
##   Sum    732  582 1314
```

- We could compare survival probability in the two smoking categories
  - We know how to do this: `prop.test`

[1] There were also many other variables collected

Whickham smoking study

```r
prop.test(x = c(502, 443), n = c(732, 582))

##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(502, 443) out of c(732, 582)
## X-squared = 9, df = 1, p-value = 0.003
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.1252 -0.0256
## sample estimates:
## prop 1 prop 2
##  0.686  0.761
```

Whickham smoking study

- $\hat{p}_1$: estimate of 20-year survival probability for non-smokers is 0.686
- $\hat{p}_2$: estimate of 20-year survival probability for smokers is 0.761
- The 95% confidence interval for $p_1 - p_2$ is (-0.125, -0.026)
  - We are 95% confident that the survival probability of non-smokers is between 0.026 and 0.125 lower than that of smokers
- What is going on?!?
  - Let's ponder that while we look into different ways of collecting data to understand relationships between variables

Experiments and observational data

- Focus on studies that explore relationships between variables
  - Whickham study: between mortality and smoking status
- Look at two classes of study design:
  - Experiment
    - Researchers observe outcome after assigning treatments
    - Treatment: variable that could potential change an outcome
  - Observational study
    - Researchers observe outcome without manipulating any variables

Experiment: Example

- Does studying while listening to classical music improve test performance?
  - Let's suppose we have 50 participants
- Assign 'treatments'
  - Assign 25 to study to classical music during a 30 minute study period
  - Assign other 25 to study in silence during 30 minute study period
- Both groups took a standardized test immediately after the study period
- Compare the scores of the two groups
  - We have the tools to analyze this data!
  - Two independent groups

Experiments: randomization

- Randomization is an important principle when designing an experiment
  - Researchers randomly allocate treatments to experimental units, e.g.
    - allocate fertilizer A or B (treatments) to plots of land (units)
    - allocate new drug or existing drug (treatments) to participants (units)
    - allocate stressful task / neutral task (treatments) to participants (units)
- Idea: avoid systematic differences between the treatment groups
- Example: randomly allocate music / silent study to participants
  - The distribution of other variables should be approximately the same in the two groups
    - The only difference between the two groups is the treatment
    - Example: Distribution of intelligence should be approximately same in both treatments
- The use of randomization allows us to make causal interpretations
  - Example: evidence that an increase in test score is caused by studying to classical music

Experiment: Example 2

- Does a new drug reduce deaths in heart attack patients?
- Use randomization to assign 'treatment'
  - One group received the new drug
  - Other group received no drug treatment
- Compare the number of deaths over some time period in the two groups
- This experiment raises a number of other considerations

Experiment: control group

- The control group is an important part of the experimental
  - It helps determine a baseline (to compare against)
- Example 2: Put yourself into the shoes of someone in the study
  - In drug group: you receive a brand new drug that you hope will help
  - In no drug group: downcast, knowing you missed out on an improved chance of survival
- Often studies will introduce a placebo (fake treatment)
  - Example 2: Participants in the 'no drug' group receive a sugar pill
  - There can be real improvements in those receiving placebos: placebo effect
- Blind study: participants do not know if they are receiving treatment or placebo
- Double blind study: the doctor does not know if the participant is receiving treatment or placebo

RCTs: issues

- The experiments outlined above are called randomized control trials (RCTs)
  - They are the gold standard for understanding relationship between variables
- They also have challenges, including
  - Cost and time: RCTs are expensive and time-consuming to design, implement and monitor
  - Generalizability: the sampling frame may not match the population of interest
    - e.g. if evaluating a possible treatment for a particular disease, an RCT will generally not include the oldest and sickest individuals
  - Ethical considerations, including:
    - Is may not be ethical to assign ‘treatments’, e.g. smoking
    - Having participants continue in placebo group once treatment determined to be effective
    - Informed consent

Experiments

- There is much more to experimental design than RCTs
- There are approaches for reducing variability: blocking
  - Blocks are groups of similar experimental units
    - Example 2: we might have 'low-risk' block and 'high-risk' block
  - Blocking helps us isolate the effect of treatment by controlling for block variability
- There are experimental designs for more complex situations
  - e.g. factorial designs (multiple treatments), cross-over designs (each unit receives multiple treatments), etc
- Details of these extensions are not important (for this course)
  - Good to know that many extensions exist
- Study design explored in STAT 311

Observational studies

- Observational studies: researchers observe participants without intervention
  - Common in many fields: ecology, earth science, epidemiology, social science, genetics, economics, psychology, ...
- Observational study designs include
  - Cross-sectional study: collect data at a single point in time
  - Cohort study
    - Follow groups (cohorts) of participants and observe the occurrence of outcomes
  - Case-control study
    - Participants grouped on the basis of their outcome status
    - Look back in time for potential factors that might have contributed to the outcome

Observational data

- Observational data: researchers observe participants without intervention
  - There is no randomization in the variables that could influence the outcome
  - There may be important variables that can distort the story if omitted
- The Whickham smoking study is an example of observational data
  - Whickham study: Could there be an important omitted variable?

Back to the Whickham smoking study

- Another variable collected in the study is age (at baseline)
  - Let's look at two age groups: 18 – 64 and 65+

- Age 18 – 64

```
##          Smoking_status
## Outcome    No  Yes  Sum
##    Alive  474  437  911
##    Dead    65   95  160
##    Sum    539  532 1071
```

- $\hat{p}_1$: 0.879
- $\hat{p}_2$: 0.821

- Age 65+

```
##          Smoking_status
## Outcome    No Yes Sum
##    Alive   28   6  34
##    Dead   165  44 209
##    Sum    193  50 243
```

- $\hat{p}_1$: 0.145
- $\hat{p}_2$: 0.12

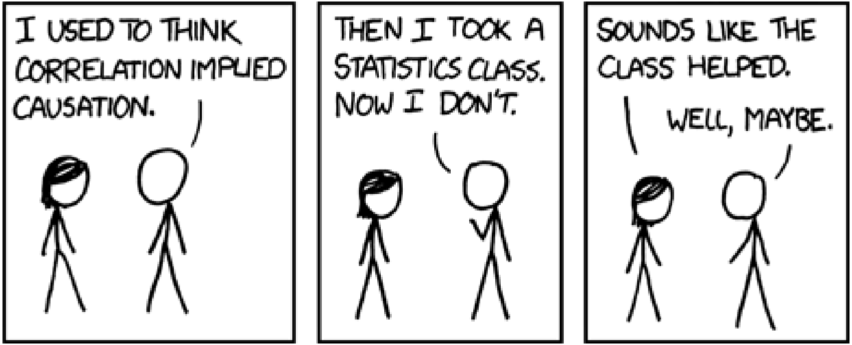
Whickham smoking study

- In each age group
  - Estimated 20-year survival probability is higher for non-smokers
- The estimated 20-year survival probability is (much) lower for those over 65
  - Not that surprising
- The proportion of smokers differs considerable between the two age classes
  - Proportion of smokers (young): $\frac{532}{1071} = 0.497$
  - Proportion of smokers (old): $\frac{50}{243} = 0.206$
- There are fewer smokers among 'old' than 'young' (as a proportion)
  - One possible explanation: many of those who may have been recruited as 'old' smokers died before the study began

Confounding

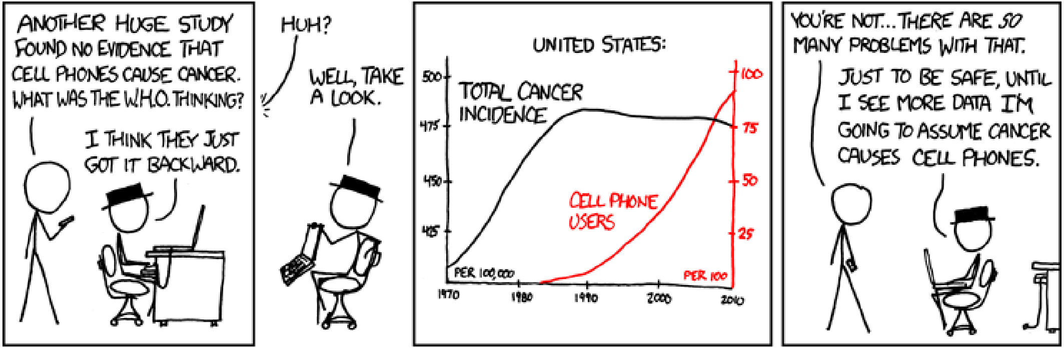
- Whickham study: Age is an example of a confounding variable
- Confounding variable: influences the predictor variable and the outcome variable
  - Spurious relationship: two variables are associated but not causally related
- Whickham study: Positive association between smoking status and survival
  - Highly unlikely to be a causal relationship
- We have been careful not to make causal interpretations
  - Difference in two means, difference in two proportions, linear regression
- Association/correlation: Comparing two (sub)populations
- Causation: a change in x causes a change in y

Correlation and causation[2]



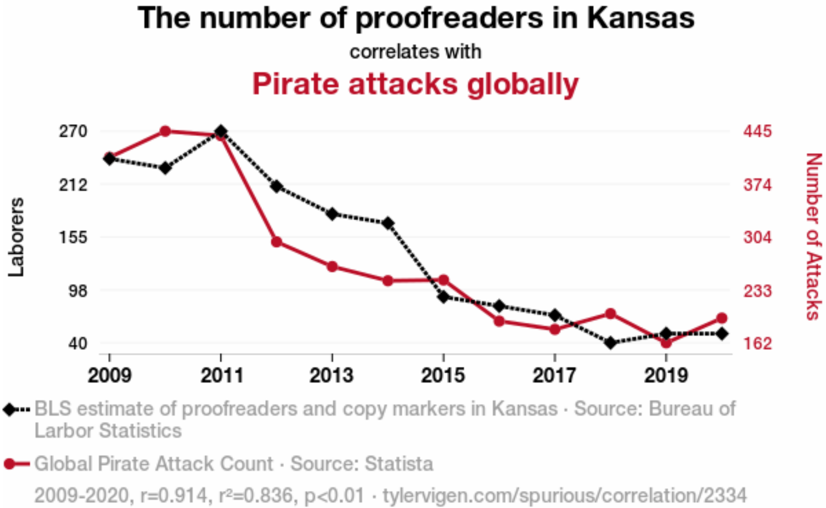
[2] https://xkcd.com/552

Correlation and causation[3]

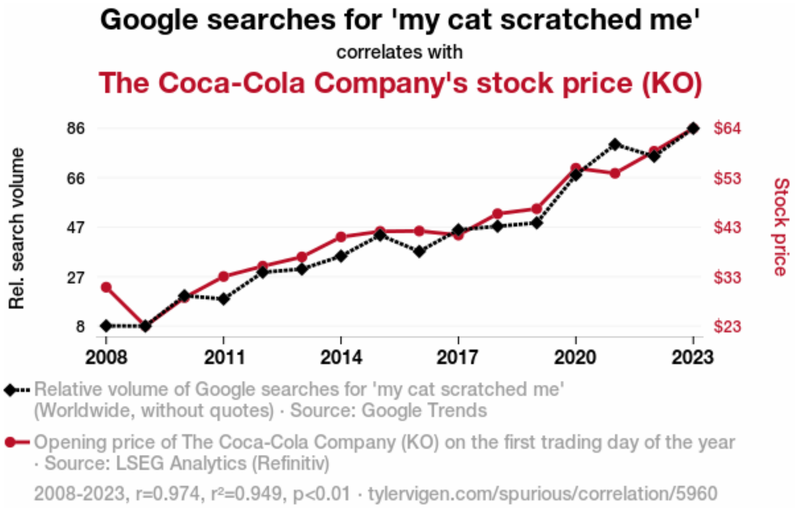


[3] https://xkcd.com/925

Correlation and causation

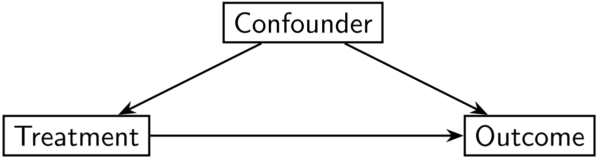


Correlation and causation

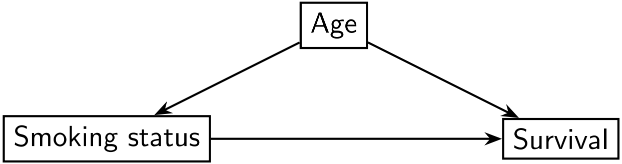


Causal diagram

- Confounding variables can be represented on a causal diagram



- For the Whickham smoking study



Causal inference from observational data

- Observational data is everywhere
  - Ecology, medical records, genetics, education data, economics
- Are there ways to try and infer causation from observational data?
- Causal inference (from observational data) is a large and active research area
- We must make assumptions about the causal relationship (what causes what)
  - Can be specified using a causal diagram
- Translate assumptions about causal relationship to an estimate of the causal relationship between a particular variable and the outcome
- More details about causal inference: see STAT 311

Summary

- Compared experiments and observational study designs
- Experiments (RCTs)
  - Randomization
  - Blinding
  - Causal interpretation of effects
- Observational data
  - Confounding
  - Correlation and causation
  - Causal inference
- Whickham example: correlation is not causation with observational data

Outline

- We are looking big picture
- How do we use statistics in 'the real world'
- Discuss the replication crisis
  - Crisis largely caused by poor statistical and scientific practice
  - Focused in psychology, but relevant in all disciplines
- Explore how it relates to what we have been taught
- Controversial topic
  - Everyone has an opinion
  - Try and provide a balanced view

Replication crisis

- BBC: Most scientists 'can't replicate studies by their peers'
- Northwestern: 'An Existential Crisis' for Science
  - "the replication crisis refers to a pattern of scientists being unable to obtain the same results previous investigators found"
- Nature: 1,500 scientists lift the lid on reproducibility
  - "More than 70% of researchers have tried and failed to reproduce another scientist's experiments" (based on a survey of 1576 scientists)
- Science: Estimating the reproducibility of psychological science
  - Authors replicated 100 studies published in three psychology journals
  - Found 39% of effects were replicated

p-values

- The use of p-values and statistical testing took a lot of heat
  - "The p-value plays into the human need for certainty and has led to the reproducibility crisis in may fields"
  - Scientific American
    - "The concept of statistical significance ... has emerged as an obvious part of the problem"
    - "The current culture of statistical significance testing, interpretation, and reporting has to go"

p-values

- Others defended p-values
- Claim the problem is not with p-values themselves, but in how they are used in modern science
- Scientific American (same article as above)
  - "p-values themselves are not necessarily the problem. They are a useful tool when considered in context."
  - "Statistical significance is supposed to be like a right swipe on Tinder. It indicates just a certain level of interest. But unfortunately, that's not what statistical significance has become. People say, 'I've got 0.05, I'm good.' The science stops."
  - "a p-value shouldn't be a gatekeeper … Let's take a more holistic and nuanced and evaluative view."

A closer look at the problem?

- The problem is often called p-hacking
  - When we collect or select data or statistical analyses until non-significant results become significant
- Motivated by a desire for p-value $< \alpha$
  - Publication bias: it can be difficult to publish null (non-signficant) results
    - No evidence of an effect
- There are many things we (as researchers) can do to make a significant results more likely
  - Destroy the validity of p-values (and confidence intervals) in the process

A fishing expedition?

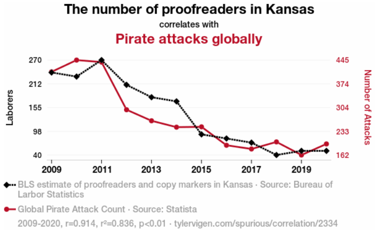
- We might look at many variables until we find combinations that are significantly related
  - There could be multiple predictor and/or outcome variables
    - e.g. GWAS (genome-wide association study)
    - Compares the DNA of individuals with a specific trait to those without
    - It is common to have more than one million possible predictors
- Often we explore variables 'formally'
  - We fit multiple models, calculating CI or p-value until we find significance
- It can also be 'informal'
  - Determine which variables are 'of interest' while plotting and exploring the data

Multiple comparisons

- This is another example of multiple comparisons
  - Recall: ANOVA
- Example: suppose we have outcome variable y and we make up (or otherwise procure) k completely unrelated variables
  - Fit a regression model between y and each variable in turn
  - If we have k variables with $\alpha = 0.05$, the probability of having at least one significant predictor is
    - $k = 5$: probability = 0.23
    - $k = 10$: probability = 0.4
    - $k = 50$: probability = 0.92
    - $k = 100$: probability = 0.994
    - Probability found using complement

Recall: Spurious relationships

- We saw examples of spurious relationships previously, e.g.



- The website that is from has a section 'Why this works'
  - "I have 25,153 variables in my database. I compare all these variables against each other to find ones that randomly match up. That's 632,673,409 correlation calculations! This is called 'data dredging'. Instead of starting with a hypothesis and testing it, I instead abused the data to see what correlations shake out. It's a dangerous way to go about analysis"
- If we compare enough variables, we will find a significant correlation

Another xkcd cartoon

- We can’t fit this one on the slide!

Multiple Comparisons

- To some degree multiple comparisons can be accounted for
  - The Bonferroni correction is a popular, general approach (e.g. in GWAS)
- If we conduct m tests, use significance level $\alpha^* = \alpha/m$
- e.g. if we perform $m = 1\,000\,000$ tests with $\alpha = 0.05$
  - Significance level is $\alpha^* = \frac{0.05}{1000000} = 0.00000005$
- The Bonferroni adjustment is simple to specify and use
  - It is a general approach
- It ensures the family-wise error rate is less than α
  - It is conservative
- It is difficult to account for 'informal' tests
  - How do you quantify the decisions made by eye?

Other types of p-hacking

- There are many other choices that we can make when exploring data that can effect the results that can be just as problematic
- Many of these choices are a part of good model building practice
  - Discuss more in STAT 210
- Easily abused if hunting for a significant result
- Best seen with an example (next slide)

Example

- Example below is taken from here
  - We conduct a study testing whether symmetrical faces are more attractive than asymmetrical ones
    - Find no overall difference in attractiveness
  - So, we test whether the effect differed as a function of the gender of the participant and the gender of the face
    - Find that men found symmetry attractive for faces of both genders, whereas women found symmetry attractive in women's faces but asymmetry attractive in men's faces
    - But, not significant
  - We examine the data more closely. We notice that some faces were rated as maximally attractive by almost everyone
    - Delete those faces from the analysis because they might obscure a real effect
  - Other participants were older than the rest and their ratings don't seem to fit the pattern
    - Delete those faces from the analysis because they might obscure a real effect
  - Now the difference between men and women becomes statistically significant
    - 'Eureka!' we cry

HARKing

- Suppose we then took that significant result and presented it without an explanation of how we got there
  - Example of HARKing (Hypothesizing after the results are known)
- We explore the data and multiple models (formally or informally) and then present the result as if it was the hypothesis we had in mind all along
- The problem:
  - p-values and confidence intervals lose their validity

Preregistration

- One approach for improving transparancy is to use preregistration
- Simple concept: create a permanent record of our study plans before we look at (or even collect) the data
- It is possible to preregister any (and every) detail of the study
  - Data-collection plans, analysis code, competing hypotheses, etc
- This does not eliminate exploration
  - It can make sense to divert from the plan
  - It is then clear which hypotheses were confirmatory (specified in advance) and what aspects were exploratory and driven by the data
- Open Science: preregistration

ASA statement on p-values

- We have seen many mentions about the ASA statement on p-values
- There are six principles
  - Principle 1: P-values can indicate how incompatible the data are with a specified statistical model
    - This is consistent with how we have discussed p-values: they measure incompatibility between the data and the model given by the null hypothesis
  - Principle 2: P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
    - These are common misconceptions

ASA statement on p-values

- Principle 3: Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
  - "The widespread use of 'statistical significance' (generally interpreted as " < 0.05) as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process."
  - "Researchers should bring many contextual factors into play to derive scientific inferences, including the design of a study, the quality of the measurements, the external evidence for the phenomenon under study, and the validity of assumptions that underlie the data analysis."
- Principle 4: Proper inference requires full reporting and transparency
  - Exactly what we have been describing above
  - "Conducting multiple analyses of the data and reporting only those with certain p-values (typically those passing a significance threshold) renders the reported p-values essentially uninterpretable"

ASA statement on p-values

- Principle 5: A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
  - “Smaller p-values do not necessarily imply the presence of larger or more important effects, and larger p-values do not imply a lack of importance or even lack of effect.”
  - “Any effect, no matter how tiny, can produce a small p-value if the sample size or measurement precision is high enough”
- Principle 6: By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis
  - “Researchers should recognize that a p-value without context or other evidence provides limited information.”
  - “A p-value near 0.05 taken by itself offers only weak evidence against the null hypothesis”

Concluding remarks

- The ASA statement on p-values concludes with:
  - "Good statistical practice, as an essential component of good scientific practice, emphasizes principles of good study design and conduct, a variety of numerical and graphical summaries of data, understanding of the phenomenon under study, interpretation of results in context, complete reporting and proper logical and quantitative understanding of what data summaries mean. No single index should substitute for scientific reasoning."
- We like certainty. We like black and white
- The problem is that, with data, we can't be certain. Things are grey

Outline

- Look at approaches for estimation
  - Maximum likelihood estimation
  - Bayesian inference

What is statistics?

- How do we collect data to ensure
  - Representative of the population
  - We can explore the scientific questions (hypothesis) of interest
- Describe a statistical model for the data
  - Describes the variability of the data
  - Estimate the parameters
    - Quantify the uncertainty about the parameters
  - Interpret these estimates
    - In the context of (scientific) application
  - Predict new observations
  - Visualise data and model output

Estimation in context

- Let's focus on how we estimate parameters
  - Recall: estimate parameters with statistics
- In many of the cases we've seen, we've relied on these as being 'obvious'
  - Estimate population mean with sample mean
  - Estimate population variance with sample variance
  - Estimate p with sample proportion
- Others were more complicated
  - Used least squares to estimate β_0 and β_1
- There are many complex situations with no obvious estimators
  - e.g. a model with a parameter related to the skew (shape) of the data
- Can we find general strategies to estimate parameters?

Estimation

- We will look at two such estimation approaches (there are several!)
  - Maximum likelihood estimation
    - Extensively used in applied statistical work
    - The estimators we have used this semester are maximum likelihood estimators
  - Bayesian statistics
    - A (very) different approach
    - Use a different definition of probability
    - Seen increasing use in last 30 years
- Looking at a basic understanding of both approaches
  - We will not sweat the details
  - Useful to see: if you continue doing research you will likely come across these approaches

Example: Palmer penguins

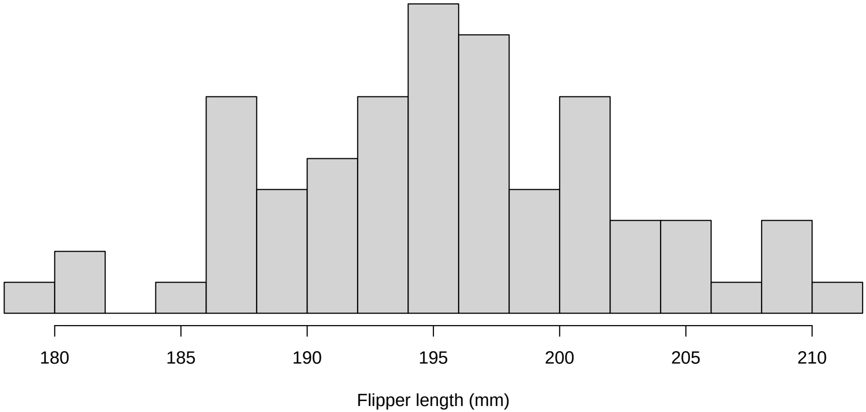
- Flipper and bill length of chinstrap penguins on Palmer archipelago
  - We will focus on flipper lengths

```
penguin = read.csv('penguin.csv')
```

```
head(penguin)
```

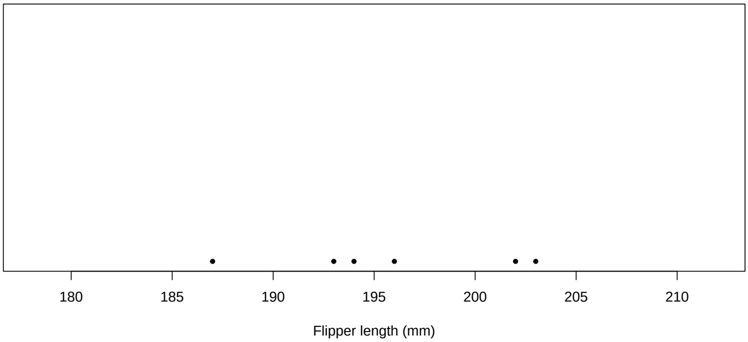
```
##   bill_length_mm flipper_length_mm
## 1           46.5               192
## 2           50.0               196
## 3           51.3               193
## 4           45.4               188
## 5           52.7               197
## 6           45.2               198
```

Palmer penguins: flipper length



Palmer penguins: flipper length

- To help us understand we reduce the sample down to 6 (randomly chosen) observations

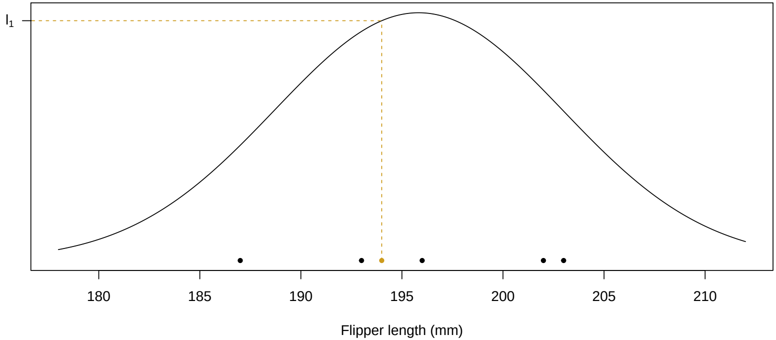


Maximum likelihood estimation

- Idea: we specify a likelihood function
  - For a given value of the parameters μ and σ
  - The function gives us a numerical value for how 'likely' the parameters are given the data we have observed
- Maximum likelihood
  - Find the value of the parameters that are most likely
- The likelihood function is given by the probability density function (pdf) of the model
  - Penguins: normal pdf
- Look at a graphical representation

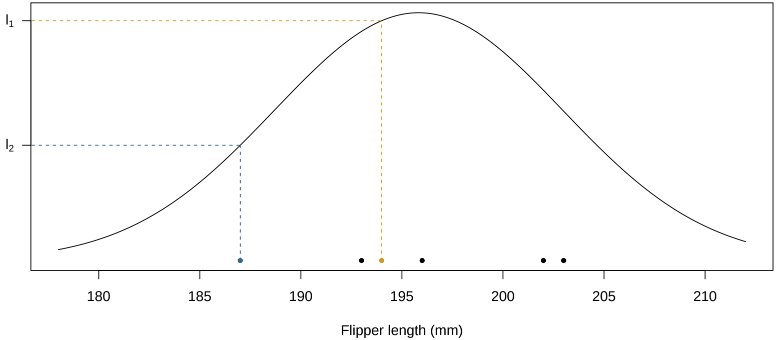
Maximum likelihood: normal model

- Likelihood when $\mu = 195.824$ and $\sigma = 7.132$
  - Find the likelihood of first observation (in gold)
  - Given by the value on the y-axis (denoted l_1)



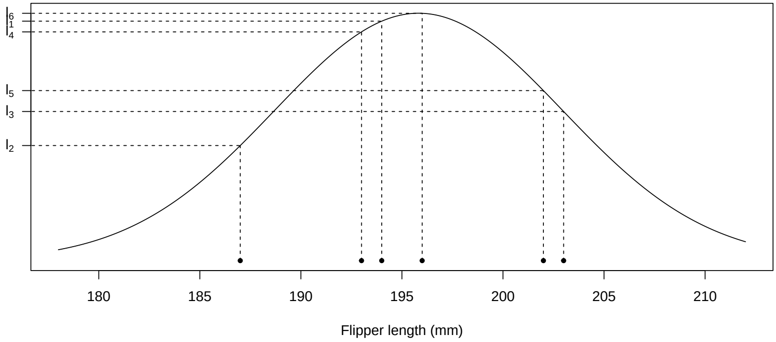
Maximum likelihood: normal model

- Likelihood when $\mu = 195.824$ and $\sigma = 7.132$
  - Find the likelihood of second observation (in blue)
  - Given by the value on the y-axis (denoted l_2)



Maximum likelihood: normal model

- Likelihood when $\mu = 195.824$ and $\sigma = 7.132$
  - Find the likelihood for all observations



Maximum likelihood: normal model

- We want the joint (or combined) likelihood
  - Multiply together the likelihood for each observation: $l_1 \times l_2 \times \ldots \times l_6$
  - Usually just called the likelihood
- We find the value of μ and σ so that the joint likelihood is as large as possible
  - Hence the name, maximum likelihood
- For many models we can find the maximum likelihood estimator mathematically
  - Normal model: $\hat{\mu} = \bar{y}$
  - Linear regression with normal errors: least squares and maximum likelihood estimators are the same
  - Binomial model: $\hat{p} = \frac{x}{n}$
- We've been using maximum likelihood without realising it!

Maximum likelihood in practice

- We have demonstrated this for a normal model
- Same process can be used for any statistical model
  - General approach for estimating a model
- Maximum likelihood estimation is explored more in STAT 270 and 370
  - How do we find maximum likelihood estimators mathematically?
  - What is the sampling distribution?
  - What is the standard error?
  - Are maximum likelihood estimators ‘good’ estimators?

Bayesian inference

- In the past 40 years, Bayesian inference has surged in popularity
  - Increasingly used in application areas

International Journal of Psychology

IUPsyS

Empirical Article | Open Access

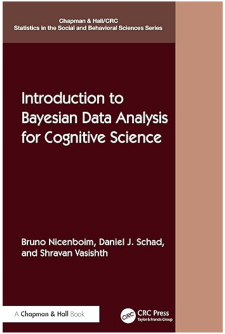
Navigating the Bayes maze: The psychologist's guide to Bayesian statistics, a hands-on tutorial with R code

Udi Alter, Miranda A. Too, Robert A. Cribbie

First published: 19 December 2024 | https://doi.org/10.1002/ijop.13271

Bayesian inference

- In the past 40 years, Bayesian inference has surged in popularity
  - Increasingly used in application areas



Introduction to Bayesian Data Analysis for Cognitive Science (Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences) 1st Edition

by Bruno Nicenboim (Author), Daniel J. Schad (Author), Shravan Vasishth (Author)

See all formats and editions

Savings Pre-order Price Guarantee. Terms

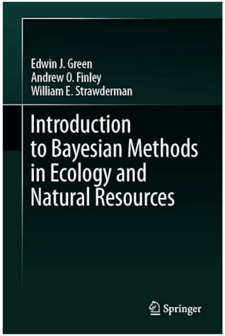
This book introduces Bayesian data analysis and Bayesian cognitive modeling to students and researchers in cognitive science (e.g., linguistics, psycholinguistics, psychology, computer science), with a particular focus on modeling data from planned experiments. The book relies on the probabilistic programming language Stan and the R package brms, which is a front-end to Stan. The book only assumes that the reader is familiar with the statistical programming language R, and has basic high school exposure to pre-calculus mathematics; some of the important mathematical constructs needed for the book are introduced in the first chapter.

Through this book, the reader will be able to develop a practical ability to apply Bayesian modeling within their own field. The book begins with an informal introduction to foundational topics such as probability theory, and univariate and bi-/multivariate discrete and continuous random variables. Then, the application of Bayes' rule for statistical inference is introduced with several simple analytical examples that require no computing software; the main insight here is that the posterior distribution of a parameter is a compromise between the prior and the likelihood functions. The book then gradually builds up the regression framework using the brms package in R,

https://bruno.nicenboim.me/bayescogsci/

Bayesian inference

- In the past 40 years, Bayesian inference has surged in popularity
  - Increasingly used in application areas



Read sample

Introduction to Bayesian Methods in Ecology and Natural Resources 1st ed. 2020 Edition, Kindle Edition

by Edwin J. Green (Author), Andrew O. Finley (Author), William E. Strawderman (Author) | Format: Kindle Edition

5.0 ★★★★★ (1)

See all formats and editions

This book presents modern Bayesian analysis in a format that is accessible to researchers in the fields of ecology, wildlife biology, and natural resource management. Bayesian analysis has undergone a remarkable transformation since the early 1990s. Widespread adoption of Markov chain Monte Carlo techniques has made the Bayesian paradigm the viable alternative to classical statistical procedures for scientific inference. The Bayesian approach has a number of desirable qualities, three chief ones being: i) the mathematical procedure is always the same, allowing the analyst to concentrate on the scientific aspects of the problem; ii) historical information is readily used, when appropriate; and iii) hierarchical models are readily accommodated.

This monograph contains numerous worked examples and the requisite computer programs. The latter are easily modified to meet new situations. A primer on probability distributions is also included because these form the basis of Bayesian inference.

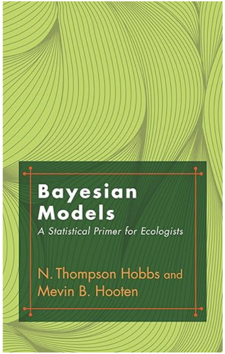
Researchers and graduate students in Ecology and Natural Resource Management will find this book a valuable reference.

Read more

ISBN-13	Edition	Publisher	Publication date	Language
978-3030607500	1st ed. 2020	Springer	November 26, 2020	English

Bayesian inference

- In the past 40 years, Bayesian inference has surged in popularity
  - Increasingly used in application areas



Bayesian Models: A Statistical Primer for Ecologists

by N. Thompson Hobbs (Author), Mevin B. Hooten (Author)

4.7 ★★★★☆ (40)   4.3 on Goodreads 24 ratings   See all formats and editions

Bayesian modeling has become an indispensable tool for ecological research because it is uniquely suited to deal with complexity in a statistically coherent way. This textbook provides a comprehensive and accessible introduction to the latest Bayesian methods—in language ecologists can understand. Unlike other books on the subject, this one emphasizes the principles behind the computations, giving ecologists a big-picture understanding of how to implement this powerful statistical approach.

Bayesian Models is an essential primer for non-statisticians. It begins with a definition of probability and develops a step-by-step sequence of connected ideas, including basic distribution theory, network diagrams, hierarchical models, Markov chain Monte Carlo, and inference from single and multiple models. This unique book places less emphasis on computer coding, favoring instead a concise presentation of the mathematical statistics needed to understand how and why Bayesian analysis works. It also explains how to write out properly formulated hierarchical Bayesian models and use them in computing, research papers, and proposals.

Read more

Report an issue with this product or seller

ISBN-10	ISBN-13	Publisher	Publication date	Language
0691159289	978-0691159287	Princeton University Press	August 4, 2015	English

Bayesian inference: Applied probability

- Bayesian statistics return to earlier discussions about probability
- Recall: Mentioned there are several interpretations of probability
  - We previously relied on a frequentist definition
- Bayesian statistics: interprets probability as a measure of belief in the occurrence of an event
  - Often called subjective or personal probability
- Example: The All Blacks played France at Lancaster Park on 26 June 1994
  - What is the probability the All Blacks won?

Bayesian inference: Applied probability

- Bayesian statistics return to earlier discussions about probability
- Recall: Mentioned there are several interpretations of probability
  - We previously relied on a frequentist definition
- Bayesian statistics: interprets probability as a measure of belief in the occurrence of an event
  - Often called subjective or personal probability
- Example: The All Blacks played France at Lancaster Park on 26 June 1994
  - What is the probability the All Blacks won?
  - Does it make sense to use probability here?

Bayesian inference: Applied probability

- Bayesian statistics return to earlier discussions about probability
- Recall: Mentioned there are several interpretations of probability
  - We previously relied on a frequentist definition
- Bayesian statistics: interprets probability as a measure of belief in the occurrence of an event
  - Often called subjective or personal probability
- Example: The All Blacks played France at Lancaster Park on 26 June 1994
  - What is the probability the All Blacks won?
  - Does it make sense to use probability here?
  - The event has happened (it is fixed and not random)
    - Probability makes no sense under frequentist interpretation
  - It seems reasonable to use probability
    - Using probability as measure of belief (in the All Blacks winning)

Bayesian inference: Applied probability

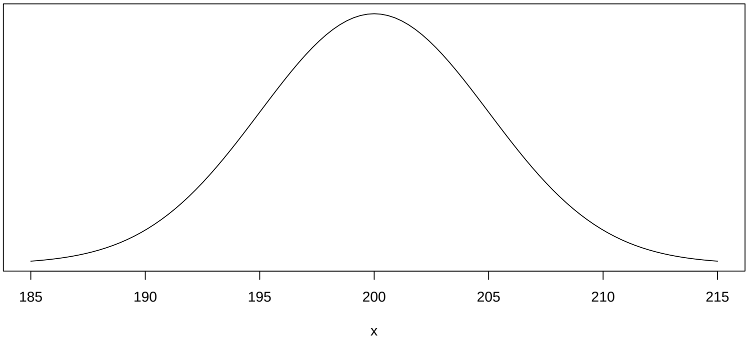
- Use probability to describe our 'belief' about the parameters
  - Use probability to describe uncertainty about the parameters
- There are two such probability distributions
  - Prior distribution: belief about the parameters before study conducted
  - Posterior distribution: belief about the parameters given data observed
- These are found using Bayes theorem (we saw this earlier!)
- The posterior distribution is what we use to get
  - Estimate (a point estimate)
  - Uncertainty (an interval estimate)

Bayesian inference: posterior distribution

- Posterior distribution found by combining (multiplying) likelihood and prior
  - Likelihood: same likelihood as above
  - Note: there is some additional mathematical complexity that we can ignore here
- We will look at the process graphically
- We will ignore many aspects
  - Mathematical details
  - In-depth discussion about the prior distribution
    - It has been a (historically) controversial aspect of Bayesian inference

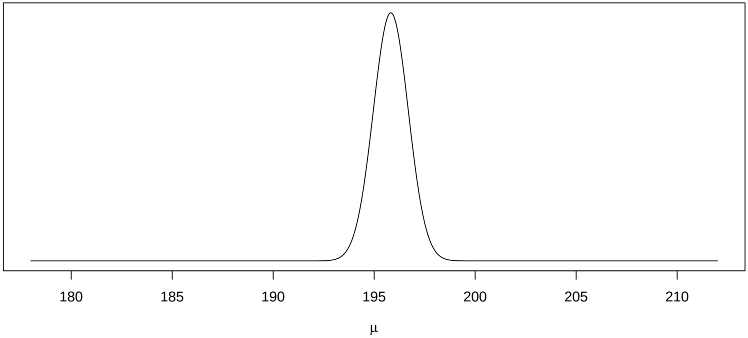
Prior

- The prior distribution describes belief about the parameters before study conducted
  - We may have a prior centered on mean flipper length of 200 mm



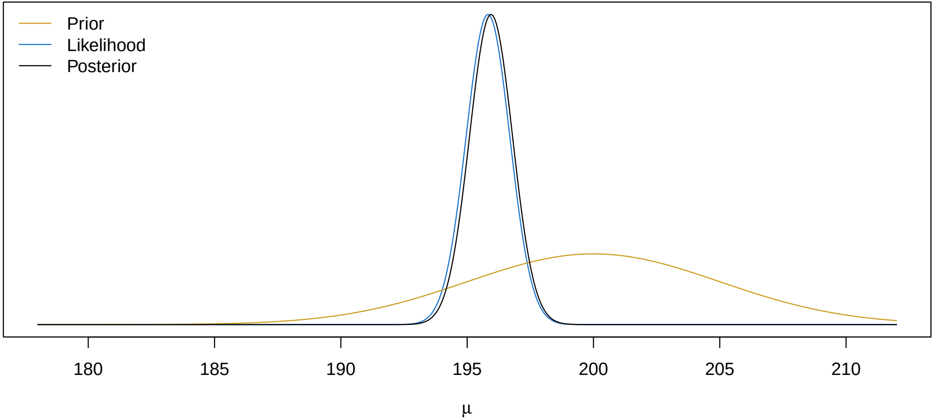
Likelihood

- The likelihood we describe above can be seen graphically
  - Likelihood for parameter μ (for a given value of σ)
    - The larger the value, the more likely the parameter value given the observed data



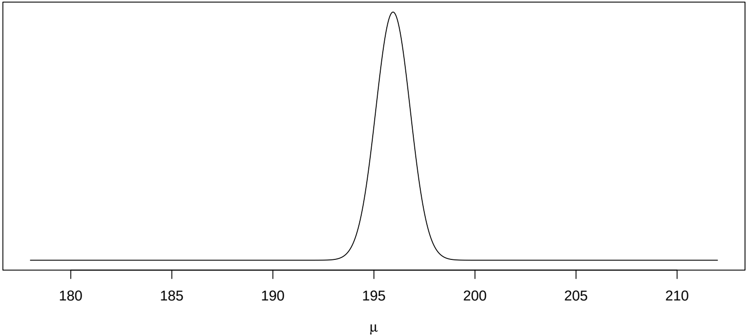
Posterior distribution

- The posterior distribution combines the likelihood and prior
  - Posterior $\propto$ Likelihood $\times$ Prior



Posterior distribution

- The posterior distribution describes our belief about the parameters given data observed
  - Probabilistic description of what value we think μ is
  - Can obtain a point and interval estimates
    - Summaries of the posterior distribution



Why has it become popular?

- For many complex (realistic) problems
  - We can fit Bayesian models, where ‘standard’ approaches are prohibitively difficult
  - There are software packages for fitting Bayesian models
- Revolutionized applied statistical modeling in the last 30 years
- Explore Bayesian modeling in STAT 371

Summary

- Introduced maximum likelihood and Bayesian modeling
  - Heard the terminology
  - Likely to come across one or both terms if continue into research involving data
- To delve deeper into these approaches
  - We need a better understanding of probability
  - We need some understanding of calculus
  - We explore the approaches in higher level courses (STAT 270, 370, 371)