Statistical Methods - MATH2715 - University of Leeds

https://leedsforlife.leeds.ac.uk/Broadening/Module/MATH2715

Objectives:

- manipulate univariate and bivariate probability distributions, including moments and transformations;
- use univariate moment generating functions to derive the classic limit theorems of probabilty;
- understand the principles of statistical modelling, from data collection to model assessment and refinement;
- deal with robustness problems in statistical estimation;
- carry out elementary Bayesian statistical modelling.

Assessment

Courswork: 20 % of total assessment

- 15 % R programming assignments;
- 5 % In-course assessment.

Exam: 80 % of total assesment

Standard exam (closed essays), duration 2 hr 0 mins :

- part A: Answer all questions in Section A (20 Marks in total);
- part B: Answer all questions in Section B (60 Marks in total).

Teaching

10 Credits

- 22 lectures (2 per week);
- 10 Workshops (1 per week starting 2nd week);
- 5 handouts/assignments ($\sim$ every other week).

Teaching

10 Credits

- 22 lectures (2 per week);
- 10 Workshops (1 per week starting 2nd week);
- 5 handouts/assignments ($\sim$ every other week).

Resources

- text books
  - Mathematical Statistics and Data Analysis (by John A. Rice);
  - Introduction to Statistics - Online Edition -D.M.Lane et al.
- introductory R document by Stuart Barber: http://www1.maths.leeds.ac.uk/statistics/R/Rintro.pdf;
- Free R course on DataCamp: https://www.datacamp.com/courses/free-introduction-to-r.

Can you define the word "Statistics"?

Can you define the word "Statistics"?

Did you say any of those?

- involves mathematics;
- relies upon calculations of numbers;
- relies heavily on how the samples are chosen;
- relies on how the statistics are interpreted.

Can you define the word "Statistics"?

Did you say any of those?

- involves mathematics;
- relies upon calculations of numbers;
- relies heavily on how the samples are chosen;
- relies on how the statistics are interpreted.

Misleading situations

- the numbers may be right, but the interpretation may be wrong...

Examples: can you identify a major flaw?

Mr Whippy Ice Cream Van

- Mr Whippy renewed his ice-cream van in June with a brand new advert on it.
- Mr Whippy noticed that in July and August he had a 35% increase in ice-creams sales;
- Mr Whippy, very proud, concluded that the new advertisement was effective!

Examples: can you identify a major flaw?

Examples: can you identify a major flaw?

Misleading situation

Major flaw:

- ice cream consumption generally increases in summer regardless of advertisements!
- so called "history effect": outcomes are the result of *another variable* (in this case time).

Examples: can you identify a major flaw?

Examples: can you identify a major flaw?

Churches and Crimes

Think of the following observations:

- The more churches in a city, the more crime there is;
- hence, churches lead to crime.

Examples: can you identify a major flaw?

Examples: can you identify a major flaw?

Misleading situation

Major flaw:

- Both increased number of churches and increased crime rates can be explained by larger populations;
- in bigger cities, there are both more churches and more crime;
- recognize that a third variable can cause both situations!

Same question as before: What are statistics ?

Same question as before: What are statistics ?

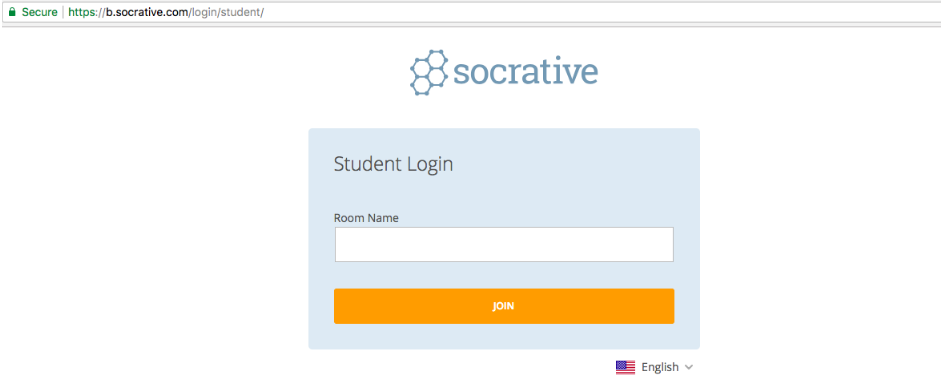
not only facts and figures!

Range of techniques and procedures for:

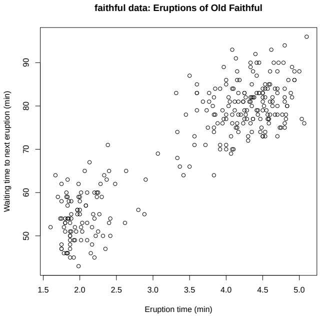
- analyzing
- interpreting
- displaying
- making decisions based on data

Join Socrative, Sudent login, ROOM NAME: CUTILLO01

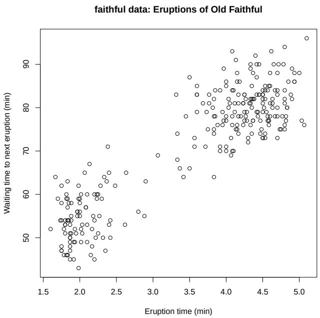
- `https://www.socrative.com/`
- Socrative Student on the App Store - iTunes - Apple
- Socrative Student - Apps on Google Play



Examples: Old Faithful Geyser Data



Examples: Old Faithful Geyser Data



Old Faithful geyser in Yellowstone National Park, Wyoming, USA

Waiting time between eruptions and their duration.

- Can you identify any trend in data?
- Statistics is about extracting meaningful conclusions from noisy data!
- We will eventually address this question.

How did we get the previous picture?

We used the R code:

```
> data(faithful)
> attach(faithful)
> f.tit < −"faithfuldata : Eruptions of Old Faithful"
> plot(faithful[, -3], main = f.tit,
  +     xlab = "Eruption time (min)",
  +     ylab = "Waiting time to next eruption (min)")
```

Algorithm 1: Eruption Time VS Waiting Time Plot

- We need to have a good grasp of what we mean by *trend* or *noise* in the data;
- we lay the groundwork with some probability;
- we need to get more formal!

What are Probabilities?

Probabilities are numbers assigned to events. They must satisfy the following properties:

- $P(\Omega) = 1$.
- $P(A) \geq 0 \ \forall A \subset \Omega$.
- If A_1 and A_2 are disjoint, then $P(A_1 \cup A_2) = P(A_1) + P(A_2)$. It follows that :
- $P(A^c) = 1 - P(A)$.
- $P(\emptyset) = 0$.
- $P(A) \leq P(B)$ whenever $A \subset B$.

Examples in a finite sample space

Flipping two pennies

- If we flip two pennies, what will Ω be?
- Given the previous rules, can you compute the $P(\textit{one head and one tail})$?
- The answer is $\frac{1}{2}$. Let's work it out together.

Rolling two dice

- If we roll two dice, what will Ω be?
- Given the previous rules, can you compute the $P(\textit{sum is } 4)$?
- The answer is $\frac{3}{36}$. Let's work it out together.

The Bayes′ Rule

Conditional Probabilities

The conditional probability of A given that B is known to have occurred is $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Law of Total Probability

Let $B_1, B_2, \ldots, B_n$ be a disjoint collection of sets each having positive probability whose union is all of Ω. Then $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$.

Bayes′ Rule.

If, in addition to the assumptions above, $P(A) > 0$, then

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} \tag{1}$$

Independence

Intuitively

We would say A and B are independent if, knowing A had happened gave no information about whether B had happened and viceversa that is: $P(A|B) = P(A)$ and $P(B|A) = P(B)$. Given that:

$$P(A \cap B) = \frac{P(A|B)}{P(B)} \tag{2}$$

then:

$$P(A \cap B) = P(A)P(B). \tag{3}$$

Definition

If, in addition to the assumptions above, $P(A) > 0$, then

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} \tag{4}$$

Bayes' Rule: Example

Case study

Facts:

- Probability that a driver is drunk is $P(B) = 1\%$.
- Probability that an Alcohol Test will give a positive result (indicating a high alcohol concentration) (event A), for a driver who is drunk, is $P(A|B) = 80\%$.
- A sober driver will test positive 10% of the times. That is, $P(A|B^c) = 10\%$.

Question

If a driver tests positive for Alcohol test, what is the probability that he/she actually is drunk? That is, what is $P(B|A)$?

Most estimated the answer to be 75%.

Bayes′ Rule: Example

Case study

- Prob. that a driver is drunk is $P(B) = 1\%$.
- Prob. that an Alcohol Test will give a positive result (event A), for a drunk driver, is $P(A|B) = 80\%$.
- A sober driver will test positive 10% of the times: $P(A|B^c) = 10\%$.

Question

If a driver tests positive for alcohol test, what is the probability that he/she is actually drunk? That is, what is $P(B|A)$?

Apply Bayes′ Rule

Let's compute:

$$P(B|A) = \frac{(0.8)(0.01)}{(0.8)(0.01) + (0.1)(0.99)} = \frac{8}{8 + 99} \sim 7.5\%. \tag{5}$$