MATH 205: Statistical methods

September 20th, 2021

Lecture 5: Independence

Tentative schedule

Date	Theme/Topic	Labs	Assignments
Sep 1	Syllabus		
Sep 8	Chapter 1: Describing dataset	Section 2: Handling data	
Sep 13 - 15	Chapter 2: Looking at Relationships	Section 3: Univariate data	
Sep 20-22	Chapter 3: Basic Ideas in Probability	Section 4: Bivariate Data	Homework 1 (due 09/22)
Sep 27-29	Chapters 3-4	Section 4: Correlation	
Oct 4-6	Chapter 4: Random variables and expectations	Section 6: Random data	Homework 2 (due 10/06)
Oct 11-13	Chapter 5: Useful distributions	Section 7: The central limit theorem	
Oct 18-20	Chapter 6: Samples and populations	Section 9: Confidence interval estimation	Homework 3 (due 10/20)
Oct 25-27	Review and midterm exam		Midterm: Oct 27 (lecture), Oct 25-27 (labs)
Nov 1-3	Chapter 7: The significance of evidence	Section 10: Hypothesis testing	
Nov 8-10	Goodness of Fit	Section 12: Goodness of Fit	Homework 4 (due 11/10)
Nov 15-17	Linear Regression	Section 13: Linear regression	
Nov 22-24	Thanksgiving break		
Nov 29 - Dec 1	One-Way Analysis of Variance	Section 15: Analysis of variance	Homework 5 (due 12/01)
Dec 6-8	Selected topics + Review		
Exam week			

Last lecture

- Sample space and events
- Basic properties of probability
- Advanced properties of probability
- Compute probability
  - Computing event probabilities by counting outcomes
  - Computing probabilities by reasoning about sets

Sample space and events

1. An experiment: is any action, process, or phenomenon whose outcome is subject to uncertainty
2. An outcome: is a result of an experiment
   Each run of the experiment results in one outcome
3. A sample space: is the set of all possible outcomes of an experiment
4. An event: is a subset of the sample space.
   An event occurs when one of the outcomes that belong to it occurs

Different views of probability

- Frequentist: The probability of an outcome is the frequency of that outcome in a very large number of repeated experiments
- Bayesian: Probability is a quantification of a belief about how often an outcomes occurs

Bottom line: Probability is a function defined on the set of events of an experiment

Basic properties of probability

Useful Facts 3.1 (Basic Properties of the Probability Events)
We have

- The probability of every event is between zero and one; in equations

$$0 \le P(\mathcal{A}) \le 1$$

  for any event $\mathcal{A}$.
- Every experiment has an outcome; in equations,

$$P(\Omega) = 1.$$

- The probability of disjoint events is additive; writing this in equations requires some notation. Assume that we have a collection of events $\mathcal{A}_i$, indexed by i. We require that these have the property $\mathcal{A}_i \cap \mathcal{A}_j = \varnothing$ when $i \neq j$. This means that there is no outcome that appears in more than one $\mathcal{A}_i$. In turn, if we interpret probability as relative frequency, we must have that

$$P(\cup_i \mathcal{A}_i) = \sum_i P(\mathcal{A}_i)$$

Useful Facts 3.2 (Properties of the Probability of Events)

- $P(\mathcal{A}^c) = 1 - P(\mathcal{A})$
- $P(\varnothing) = 0$
- $P(\mathcal{A} - \mathcal{B}) = P(\mathcal{A}) - P(\mathcal{A} \cap \mathcal{B})$
- $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B})$

Others

- If $A \subset B$, then $P(A) \leq P(B)$.
- For any events A, B

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

Independence

Independence

- Some experimental results do not affect others
- Example: if I flip a coin twice, whether I get heads on the first flip has no effect on whether I get heads on the second flip
- We refer to events with this property as independent.

Independence

Definition

Two events A and B are independent if and only if

$$P(A \cap B) = P(A)P(B)$$

Dependent events: example

Toss a fair dice:

- A: the event that the die comes up with an odd number of spots
- B: the event that the number of spots is larger than 3.
- $P(A) = P(B) = 1/2$
- If we know that A has occurred, then we know the die shows either 1, 3, or 5 spots. One of these outcomes belongs to B, and two do not. $P(A \cap B) = 1/6$.
- This means that knowing that A has occurred tells you something about whether B has occurred.

$\rightarrow$ These events are interrelated.

Independence: example

Problem

A red die and a white die are rolled. Let event

$$A = \{4 \text{ on the red die}\}$$

and event

$$B = \{\text{sum of dice is odd}\}.$$

Show that A and B are independent.

Independence

Problem

Prove that if A and B are independent, then A and B^c are independent as well.

Problem

Prove that if A *and* B *are mutually exclusive events, and* $P(A) > 0, P(B) > 0$, *then they are dependent.*

Independence

Problem

If $P(A) = 0.5$, $P(B) = 0.2$ *and* $P(A \cup B) = 0.65$. *Are* A *and* B *independent?*

Problem

I search a DNA database with a sample. Each time I attempt to match this sample to an entry in the database, there is a probability of an accidental chance match of 10^{-4}*. Chance matches are independent. There are 20,000 people in the database. What is the probability I get at least one match, purely by chance?*