Week 1 – Course overview and Probability basics

Juan Pablo Lewinger

8/28/2025

PM 618 - Theory of Statistics for the Health Sciences

- Instructor: Juan Pablo Lewinger
  - Office hours by appointment
  - lewinger@usc.edu

Introduction to probability and statistics for the health sciences

- First half: Probability
    - e.g. sample space, random variables, expectation and variance, random vectors, key parametric families of distributions, central limit theorem, law of large numbers
- Second Half: Statistics theory
    - E.g. Sampling distributions. Estimation, efficiency. Maximum likelihood. Asymptotic normality. Hypothesis Testing. Confidence Intervals.

Required background

- Basic data analysis methods including linear regression.
- Working knowledge of univariate and multivariate calculus: multivariable derivation and integration
- Basic matrix algebra: matrix multiplication, inverse of a matrix
- A level of mathematical maturity
- Familiarity with a high-level programming language such as R or Python is highly desirable.

Class format

Flipped model:

- Students watch videos and read materials beforehand in preparation for Thursday's in-person meetings
- Readings from Textbooks: 1) Introduction to probability, statistics, and random processes 2) Foundations of Statistics with R
- Easy quizz on the videos and readings at the beginning of each class
- In class review/explanation of key concepts (1 ~ 1.5 hours)
- In class work on problems in small groups + discussion (~ 2 - 2.5 hours)

Evaluation

- Midterm exam 25%
- Final exam 30%
- In-class quizzes 5%
- Assignments/Problem sets submission 10%
- In-class work 30%

What's expected from you?

- Watch videos and read material before class
- Come to class
- Participate in class: ask questions, work on problems, present problems solutions to rest of class
- Work on problem sets individually or in groups but submit your own solutions
- If you are stuck with a problem ask instructor
- Use of ChatGPG and other chatbots:
  - *Ok* for getting programming help with R
  - *Not ok* for solving homework problems involving math derivations

Texbooks

"Introduction to probability, statistics, and random processes, available at https://www.probabilitycourse.com, Kappa Research LLC, 2014.

Foundations of Statistics with R Darrin Speegle and Bryan Clair, available at https://mathstat.slu.edu/~speegle/_book_spring_2020/

Probability basics

- Random phenomena are all around us: outcome of a coin flip, weather patterns, stock market, getting disease
- Random means non-deterministic, i.e. phenomena where there is an associated uncertainty about the outcome; we cannot predict perfectly (next day weather, roll of a die, price of Apple stock)
- Inability to predict perfectly could be due to lack of full information (e.g. flipping coin) or intrinsic (quantum mechanics)
- Error in measurements: anything we measure has uncertainty associated with it
- Randomness does not imply complete lack of pattern (e.g. a multiple coin toss will have about half heads and half tails), meteorologists can assess (and quantify!) whether the chance of rain is high or low

Probability basics

- Probability is the branch of mathematics that allows us to model randomness and studies properties of random phenomena
- Probability has application in computer science, machine learning, artificial intelligence, finance, **statistics**
- Probability allows as to model observations/data arising from phenomena/processes that are or can be modeled as random
- Statistics can be thought of as providing solutions to the inverse problem of probability: given observations/data infer properties about the underlying random phenomenon/process that generated the data
- Probability can be defined formally (at different levels of mathematical rigor) or can be treated more intuitively

Sample space

Definition: the set of all possible outcomes of an experiment of interest

Examples:

- Single coin tossing: $\Omega = \{H, T\}$
- Month of birth of a randomly chosen person:
  $\Omega = \{Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec\}$
- Whether a Youtube video will be clicked when presented to a potential viewer
  $\Omega = \{YES, NO\}$

Event space

Event Space: all possible events (collection of outcomes) we will consider.

For discrete sample spaces event space is typically all possible subsets of Ω

- Example: single coin toss
  - Sample space: $\Omega = \{H, T\}$
  - Event space: $\mathcal{F} = \{\emptyset, \{T\}, \{H\}, \{H, T\}\}$
- Example: birth month of a randomly chosen person
  - $\Omega = \{Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec\}$
  - $\mathcal{F} = \{\emptyset, \{Jan\}, \dots, \{Dec\}, \{Jan, Feb\}, \dots, \{Jan, \dots, Dec\}\}$
- Q: how many elements in $\mathcal{F}$?

Event operations

We require that union of events and intersection of events are also events:

$A, B \in \mathcal{F} \implies A \cup B \text{ and } A \cap B \in \mathcal{F}$

$\mathcal{F}$ is closed under unions and intersections

E.g. $A = \text{First semester} = \{Jan, Feb, Mar, Apr, May, Jun\}$

$B = \{May, Jun, July\}$

$A \text{ or } B = A \cup B = \{Jan, Feb, Mar, Apr, May, Jun, Jul\}$

$A \text{ and } B = A \cap B = \{May, Jun\}$

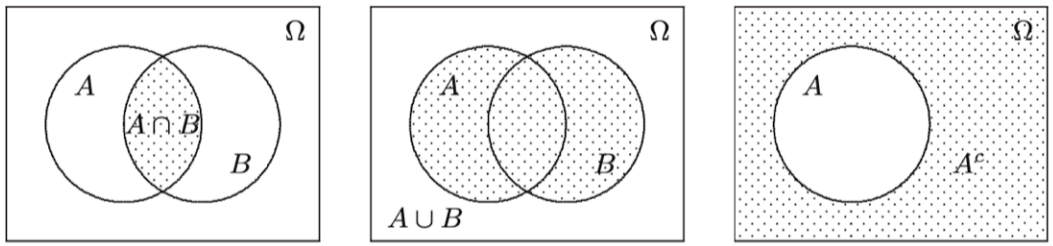
Complements

$$\text{not } A = A^c = \{Jul, Aug, Sep, Oct, Nov, Dec\}$$

$$A \setminus B = A - B = \{Jan, Feb, Mar\}$$

$$A^c = \Omega \setminus A = \Omega - A$$

$\mathcal{F}$ is closed under union, intersection, and set difference



Intersection $A \cap B$ Union $A \cup B$ Complement A^c

Disjoint events

A and $B = \emptyset$ i.e., no elements in common

E.g. $A =$ First semester $= \{Jan, Feb, Mar, Apr, May, Jun\}$

$B =$ Second semester $= \{Jul, Aug, Sep, Oct, Nov, Dec\}$

DeMorgan's Laws

. $(A \cup B)^c = A^c \cap B^c$

. $(A \cap B)^c = A^c \cup B^c$

Example: Let J be the event "John is guilty" and M the event "Mary is guilty."

$(J \cap M)^c =$ Not true that both John and Mary are guilty

$J^c \cup M^c =$ Either John or Mary are not guilty

$(J \cup M)^c =$ Not true that either John or Mary (or both) are guilty

$J^c \cap M^c =$ Neither John nor Mary are guilty

- (prove DeMorgan's Laws to practice with set operations)

Probability functions

A probability function is a 'set function' that assigns a real number to each event in $\mathcal{F}$:

$P : \mathcal{F} \rightarrow \mathbb{R}$ such that:

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$ (i.e. additive on disjoint sets)

The probability reflects the chances an event occurs, 0 being impossible and 1 being certain

Probability functions

Example: Fair coin

$P(\{H\}) = P(\{T\}) = \frac{1}{2}$ (we will simplify notation as $P(H) = P(T)$)

$P(\emptyset) = 0$

$P(\{H, T\}) = P(\Omega) = 1$

Probability function

Example: birth month of a randomly chosen person

$$P(Jan) = P(Feb) \ldots \ldots P(Dec) = \frac{1}{12}$$

Or perhaps a more reasonable assignment of probabilities would be proportional to the number of days in each month:

$$P(Jan) = 31/365, P(Feb) = 28/365, \ldots$$

(This shows that it is us, the users who assign probabilities; probabilities are not 'laws of nature')

Probability function

Any probability in a discrete sample space can be constructed like in the previous examples:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

$$p_1 + p_2 + \dots + p_n = 1, \qquad p_1 \geq 0,\ p_2 \geq 0,\ \dots\ p_n \geq 0$$

$$P(A) = \sum_{w_i \in A} p_i, \quad \forall A \subset \Omega$$

is a probability function.

Exercise: show this

Probability function

Property 3 holds for any number of disjoint events:

$P(A \cup B \cup C) = P(A) + P(B) + P(C), A \cap B = \emptyset, A \cap C = \emptyset, B \cap C = \emptyset$

E.g. the sets $A = \{\text{First Trimester}\} = \{Jan, Feb, Mar\}$

$B = \{\text{Second Trimester}\} = \{Apr, May, Jun\}$

$C = \quad = \{Nov\}$

Are pairwise disjoint

$P(A \cup B \cup C) = P(\{Jan, Feb, Mar, Apr, May, Jun, Nov\}) = 7/12$

$P(A) = P(B) = 3/12, P(C) = 1/12$

Probability function

In general:

$$P(A_1 \cup \ldots \cup A_n) = P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

Provided the events are pairwise disjoint:

$$A_k \cup A_l = \emptyset, \ \forall k, l \in \{1 \ldots n\}, \ k \neq l$$

Derived properties

If $A, B \in \mathcal{F}$

- $P(\emptyset) = 0$
- $0 \leq P(A) \leq 1$
- $P(A^c) = 1 - P(A)$
- $P(A - B) = P(A) - P(A \cap B)$
- $B \subset A \implies P(A - B) = P(A) - P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(Good practice exercise to show these)

Repeated experiments/Product of sample spaces

E.g. Flip a coin twice

$\Omega = \{(H,H),(H,T),(T,H),(T,T)\} = \{H,T\} \times \{H,T\} = \{H,T\}^2$

E.g. Flip a coin n times:

$\Omega = \{H,T\}^n$ all n-tuples with elements in H,T:

E.g. Flip a coin and then pick a month at random

$\Omega_1 = \{H,T\},\ \Omega_2 = \{Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec\}$

$\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1,\omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$

- Q: How many elements in Ω?

Repeated experiments/Product of sample spaces

If we have a probability function P_1 defined in Ω_1 and a probability P_2 defined in Ω_2 we can naturally defined a probability P in $\Omega^1 \times \Omega^2$ as:

$$P(\{\omega^1, \omega^2\}) = P_1(\omega^1)P_2(\omega^2)$$

E.g.

$P(\{H, Jul\}) = P(H) \times P(Jul) = \frac{1}{2} \times \frac{1}{12}$

(This is how we model independence, which will cover next week)

Uniform probability spaces

In many applications it makes sense to assign the same probability to all elements of a finite sample space

E.g. two coin flip:

$$\Omega = \{(H,H),(H,T),(T,H),(T,T)\} = \{H,T\} \times \{H,T\}$$

$$P(H,H) = P(H,T) = P(T,H) = P(T,T) = \frac{1}{4}$$

In general, a uniform probability space with $|\Omega| = n$, has $P(\omega) = \frac{1}{n} \ \forall\, \omega \in \Omega$

And $P(A) = \frac{|A|}{|\Omega|}$, the probability of an event is the number of elements in the event divided by the total number of elements in the sample space

E.g. Pick a single card from a well shuffled standard 52-card deck:

$$P(Ace) = \frac{4}{52}$$

$$P(\text{diamond suit}) = P(\diamondsuit) = \frac{12}{52} = \frac{1}{4}$$

Multiplicative counting principle

Many problems in probability theory require that we count the number of ways that a particular event can occur. This kind of counting falls under the area of mathematics called combinatorics.

The Multiplicative counting principle (MP).

Suppose that we perform r experiments such that the k^{th} experiment has n_k possible outcomes, for $k = 1, 2, \ldots, r$. Then there are a total of $n_1 \times n_2 \times n_3 \times n_r$ possible outcomes for the sequence of r experiments.

Example 1: Need to choose a password for an online account. Password must consist of two lowercase letters (a to z) followed by one capital letter (A to Z) followed by four digits (0,1, ,9).

Example 2: How many subsets does a set with n elements have?

Permutations

How many five-card hands are possible from a standard fifty-two card deck? (if order matters)

$52 \times 51 \times 50 \times 49 \times 48 = \frac{52!}{47!} = 311,875,200$ by the MP

In general, a k-permutation of n distinct objects is a way to arrange k objects out of the n in a row (order matters)

The number of k-permutations $p(n,k)$ is given by $p(n,k) = \frac{n!}{(n-k)!}$

n-permutations are often refer to as just permutations. There are $\frac{n!}{(n-n)!} = n!$ permutations

Combinations

How many five-card hands are possible from a standard fifty-two card deck? (if order **does not** matter)

$52 \times 51 \times 50 \times 49 \times 48 = \frac{52!}{47!}$ ordered arrangements

Each arrangement was counted $5!$ times so the number of unordered arrangements is $\frac{52!}{47!5!} = 2,598,960$

In general, a k-combination of n distinct objects is a way to arrange k objects out of the n when order does not matter

The number of k-combinations $c(n,k)$ is given by $c(n,k) = \frac{n!}{(n-k)!k!} = \binom{n}{k}$

$p(n,k) = c(n,k) \times k!$

Card problem

Suppose we deal a 5-card hand from a regular 52-card deck. Which is larger, $P(\text{One king})$ or $P(\text{Two hearts})$?

$$P(\text{One king}) = \frac{\binom{4}{1} \times \binom{48}{4}}{\binom{52}{5}}$$

$$P(\text{Two hearts}) = \frac{\binom{13}{2} \times \binom{39}{3}}{\binom{52}{5}}$$

```
4 * choose(48, 4) / choose(52, 5)
```

```
## [1] 0.2994736
```

```
choose(13, 2) * choose(39, 3) / choose(52, 5)
```

```
## [1] 0.2742797
```

De Mere's problem

- Dice game that played an important role in the historical development of probability.
- Chevalier de Méré had been betting that, in four rolls of a die, at least one six would turn up.
- He was winning consistently and, to get more people to play, he changed the game to bet that, in 24 rolls of two dice, a pair of sixes would turn up.
- De Méré lost with 24 and felt that 25 rolls were necessary to make the game favorable.
- **Was De Méré right?**

Simulating De Mere's problem

Single die roll in R:

```
sample(1:6, size = 1, replace = TRUE)
```

```
## [1] 2
```

Four rolls of one die

```
sample(1:6, size = 4, replace = TRUE)
```

```
## [1] 1 2 4 3
```

Simulating De Mere's problem

Checking if a six came up

```
any(sample(1:6, size = 4, replace = TRUE) == 6)
```

```
## [1] FALSE
```

Full simulation:

```
nreps = 1000
set.seed(2021)

results = numeric(0)
for (i in 1:nreps) results[i] = any(sample(1:6, size = 4, replace = TRUE) == 6)
mean(results)
```

```
## [1] 0.507
```

De Mere's problem

Questions:

1. Based on this simulation result, do you think the bet's favorable?
2. Derive/compute the actual probability (hint: use that all outcomes of the four rolls of a die are equally likely)
3. Simulate the second scheme (24 rolls of two dice). What can you say about the favorability of the bet?
4. Derive/compute the actual probability. How about for 25 rolls of two dice?

Birthday 'paradox'

How many people n do we need to have in a room to make it a favorable bet (probability of success greater than 1/2) that two people in the room will have the same birthday?

Assume all 365 b-days are equally likely.

1. Perform a simulation in R to answer this question (hint: use the base R function 'duplicate' to check that whether there are matching b-days)

2. Compute the probability by mathematical derivation and plot the probability as a function of n. (Hint: use the multiplication principle to count)

Infinite (countable) sample spaces

E.g. flip a coin until first heads appears:

What is the right probability space for this experiment?

$\Omega = \{1, 2, 3, \ldots\} = \mathbb{N}$

Sample space has to be infinite because no guarantee experiment will terminate in a finite number of steps!

If we assume that after k flips all k–tuples are equally likely what should the probability $P(k)$ be?

$k = 1 = \{H\} \implies P(1) = \frac{1}{2}$

$k = 2 = \{T, H\} \implies P(2) = \frac{1}{4}$

$\vdots$

$k = \{\underbrace{T, \ldots, T}_{\text{k-1 times}}, H\} \implies P(k) = \frac{1}{2^k}$

Does this result in a probability function?

Infinite (countable) sample spaces

For infinite sample spaces need to change additivity rule to countably additivity rule:

2′. $P(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$ provided thy are disjoint ($A_k \cup A_l = \emptyset \; \forall k, l \in \mathbb{N}, k \neq l$)

$\Omega = \{H, T\}^{\infty}$ all infinite sequences with elements in $\{H, T\}$

$$P(\Omega) = P(\{1, 2, 3 \ldots\}) = P(1) + P(2) + P(3) + \ldots = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1$$

(Used that for a geometric series: $1 + r + r^2 + \ldots = \sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$ if $0 \leq r < 1$)

Q: What's the probability that it'll take an even number of tosses until the first heads?

Finite and countable sets: a mathematical aside

A set is finite if its elements can be put in one-to-one correspondence with with $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$

E.g., the set of students in the classroom, the set of inhabitants in the world, the set of stars in the Milky Way

A set is countable if its elements can be put in one-to-one correspondence with with the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$

E.g. The set of natural numbers, the set of odd numbers ($n \longrightarrow 2n + 1$), the set of even numbers ($n \longrightarrow 2n$), the set of primes, the set of rational numbers ($\mathbb{Q}$)!!

Examples of infinite non-countable sets: $\mathbb{R}$, the set of irrational numbers $\mathbb{R} - \mathbb{Q}$, $\mathbb{R}^2$

Next week

Read IPS Ch 1: 1.0-1.3, 1.4-1.5

Remember there will be a quiz at the beginning of class!

Example quiz question:

A conditional probability is an update of the probability of an event after additional information about the event is known True or False?