Stat 230: Probability

(Sections 002 and 004)

Instructor: Adam Kolkiewicz

Department of Statistics and
and Actuarial Science

- Go over syllabus/office hours/other organizational issues
- What is probability? (Chapter 1)
- Probability basics (Chapter 2)

- Syllabus
- Tutorials
  - Will be used for:
    - 3 Tutorial Assignments (practice problems)
    - 3 Tutorial Tests
    - workshops and midterms reviews
  - Please pay attention to the tutorial schedule!
  - No tutorial today

- My office: M3 4009

  My office hours:

  Monday: 3:30 pm - 4:30 pm
  Friday: 11:30 am - 12:30 pm

  Or by appointment (arranged via email)

- Advice for doing well in this course:
  - Do the problems in the Course Notes!
  - If the problem has a solution, do not look into it before attempting to solve it on your own.
  - Take all of the tutorial assignments and tests.
  - Do not stay behind the material currently presented in class, as we cover quite a lot in this course.

- About the slides used in class:
  - They are based on the Course Notes, but in some places they provide more explanation.
  - They do not replace the Notes.
  - Be ready to write some additional information in class.

Introduction to Probability

- Probability Theory: A subdiscipline of mathematics concerned with describing and modeling uncertain experiments.

- Statistics: The study of the collection and analysis of data.

$$\text{Statistics} \iff \text{Data} \iff \text{Uncertainty} \iff \text{Probability}$$

- Clicker Question

- Typically we cannot eliminate uncertainty, but we can describe it and quantify it using the theory of probability.

- Although probability arose out of the study of gambling games, currently it has applications in a large number of disciplines, including

  - life-sciences
  - pharmacology
  - physics (e.g., quantum physics)
  - finance, insurance, economics[1]
  - business decisions
  - computer science (machine learning, AI, “big data”)
  - engineering
  - meteorology

---

[1] A great reading about the role of risk in our society is “Against the Gods” by P.L. Bernstein.

- Some of the many results you get if you search internet for “what are the chances”:
    - What are the chances that climate change is “natural”?
    - What are the chances of getting asylum?
    - What are the chances of getting alzheimer?
    - What are the chances of Trump being reelected?
    - What are the chances the price of the Apple stock will increase by more than 10% over the next year?
    - What are the chances of you coming into being?

- In most of the above situations, an accurate description of all the sources of uncertainty is impossible.
- In this course we will illustrate the theory of probability using much simpler situations:
    - Roll a six-sided red die and a six-sided green die. Find the probability that the total number of pips showing on the top faces is 5.
    - Toss a coin twice (there are two outcomes possible: head/tail). Find the probability of getting one head.
    - Find the probability a bridge hand (13 cards picked at random from a standard deck without replacement[2]) that has at least 1 ace.

[2] A standard deck has 13 cards in each of four suits, hearts, diamonds, clubs and spades, for a total of 52 cards. There are four aces in the deck, one of each suit.

Ways in which probabilities are defined:

1. The **classical** definition: The probability of some event is
$$\frac{\text{number of ways the event can occur}}{\text{the total number of possible outcomes}},$$
provided all outcomes are equally likely.
For example, the probability of rolling a 2 with a six sided die is 1/6.

2. The **relative frequency** definition: The probability of an event is the (limiting) proportion (or fraction) of times the event occurs in a very long series of repetitions of an experiment.
For example, the probability of rolling a 2 is 1/6 since if you roll the die many times, about 1/6th of the time the outcome will be a 2.

3. The **subjective** probability definition: The probability of an event is a measure of how sure the person making the statement is that the event will happen.

   For example, after considering all available data, an economist may say that the chances of recession in Canada next year are 10%.

   Subjective probabilities can be included into a more formal decision-making procedure in the context of “Bayesian Statistics”.

Each of the three approaches can be criticized[3].

Figure: Persi Diaconis, Probability of a coin flip landing Heads (standard US quarter) is approximately 50.4% under “typical” coin flip conditions.

[3] Thanks to my colleague Prof. Greg Rice for the picture and some other graphs used in the slides.

Mathematical Probability Models

- The difficulties in defining "probability" can be overcome by treating it as a mathematical system based on a set of axioms.

  The mathematical approach presented in this course uses the following components of a probability model:

  (i) **sample space:** this is the set of all possible outcomes of the random experiment

  (ii) **set of all subsets of the sample space:** these are the "events" to which we are going to assign probabilities

  (iii) **mechanism for assigning probabilities to events:** these are numbers from $[0, 1]$.

- In this part we are looking at the definitions of:
  - sample space
  - event, simple event, compound event
  - probability distribution
  - the term “odds”
- We will also learn about:
  - a method of assigning probabilities to events
  - uniform probability model

Example: Suppose we roll a six sided die. We may ask questions like:

- what is the probability that the number is 6?
- what is the probability that the number is less than 6?
- what is the probability that the number is even?

For these questions, the events of interest are:

$$
\begin{aligned}
A &= \text{the rolled number is 6} \\
B &= \text{the rolled number is less than 6} \\
C &= \text{the rolled number is even}
\end{aligned}
$$

Example. Consider drawing a 5 card hand at random from a standard 52 card deck of playing cards (13 kinds: A,2,3,4,...,10,J,Q,K, in 4 suits: $\clubsuit, \diamondsuit, \heartsuit, \spadesuit$).

(i) What is the probability that the hand contains at least 3 K's?

(ii) What is the probability that the hand contains 1 or fewer A's?

The events of interest:

A = the hand contains at least 3 K's

B = the hand contains 1 or fewer A's.

Definition. A **sample space** S is a *set* of distinct outcomes of an experiment with the property that in a single trial of the experiment one and only one of these outcomes occurs[4].

Example.

Experiment	S
Flip a Coin	$\{Head, Tail\}$
Number of Heads when flipping a coin	$\{0, 1\}$
Count the number when rolling a die	$\{1, 2, \ldots, 6\}$
Number of car accidents on 401	$\{0, 1, 2, 3, 4, ...\}$
The lifetime of a policyholder	$\{t \in \mathbb{R}, \quad t > 0\}$

[4] In this course typically we will be able to identify all possible outcomes of an experiments. In practice sometimes this may be difficult.

- Often there will be more than one way of defining a sample space.

  For example: in the experiment where we flip a coin 3 times, we can use:

  Sample space 1:

  $$S := \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

  Sample space 2:

  $$S := \{0 \text{ Heads}, 1 \text{ Head}, 2 \text{ Heads}, 3 \text{ Heads}\}.$$

  The second sample space is simpler than the first one, which encodes exactly each outcome, but it may not be suitable to answer some questions.

Definition. A sample space S is said to be **discrete** if it is finite, or "countably infinite". Otherwise a sample space is said to be **non-discrete**.

S	Discrete/Non-Discrete
$\{Heads, Tails\}$	discrete
$\{1, 2, 3, 4, ...\}$	discrete
$\{t, \quad t > 0\}$	non-discrete

Remark: A set is "countable" if its elements can be placed in a one-to-one correspondence with a subset of the natural numbers $\{0, 1, 2, \ldots\}$. Examples:

- $\{1, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \ldots\}$
- $\{1, -1, 2, -2, 3, -3, \ldots\}$
- the set of rational numbers

Sets that are not countable: intervals $(a, b) \subseteq \mathbb{R}$.

Definition. An **event** is a subset[5] of a sample space S.

Thus: A is an event iff[6] $A \subseteq S$.

> We say the event A **has occurred** iff the outcome of an experiment belongs to A.

Example:

Event	Subset
The coin flip is "Heads"	$\{Heads\}$
The rolled number is even	$\{2, 4, 6\}$
The policyholder dies before the age of 40	$(0, 40)$

[5] If S is not discrete then we consider only the so-called measurable subsets, but we are going to ignore this fact in this course.

[6] if and only if

Example: Suppose we roll a six sided die. In this experiment

$$S = \{1, 2, 3, 4, 5, 6\}.$$

For the questions

- what is the probability that the number is 6?
- what is the probability that the number is less than 6?
- what is the probability that the number is even?

the events of interest are:

$$\begin{aligned} A &= \text{the rolled number is } 6 = \{6\} \\ B &= \text{the rolled number is less than } 6 = \{1, 2, 3, 4, 5\} \\ C &= \text{the rolled number is even} = \{2, 4, 6\}. \end{aligned}$$

Note that the event A may occur in only one way, while events B and C may occur in more than one way.

When building a probability model it will be important to distinguish between the two types of events.

Definition. Let A be an event in a discrete sample space S.

If the event contains only one point, e.g. $A = \{a_1\}$, we call it a **simple event**.

If A is made up of two or more simple events, e.g. $A = \{a_1, a_2\}$, then A is called a **compound event.**

Example: when we roll a six sided die with

$$\begin{aligned} A &= \{6\} \\ B &= \{1,2,3,4,5\} \\ C &= \{2,4,6\}. \end{aligned}$$

the event A is simple while B and C are compound events.

Clicker Questions.

Assigning Probabilities

- Once we have defined a discrete sample space S, we need to assign a probability to each subset A of S.

  When doing this we have to follow the **Axioms of Probability**, which we will discuss in Chapter 4.

- One way of assigning probabilities so that the Axioms are satisfied is:

  (S1) based on the description of an experiment, assign probabilities to the simple events in S

  (S2) using the probabilities of the simple events, assign probabilities to the remaining subsets of S.

About the Course
Introduction to Probability (CH1)
Mathematical Probability Models (CH2)
Assigning Probabilities
Uniform Probability Model
The term "odds"

Definition 3 Let $S = \{a_1, a_2, a_3, \ldots\}$ be a discrete sample space. Assign numbers (probabilities) $P(a_i) : i = 1, 2, 3, \ldots$ to the $a_i\,'s$ such that the following two conditions hold:

(1) $0 \leq P(a_i) \leq 1$

(2) $\sum_{\text{all } i} P(a_i) = 1$.

Then the set of probabilities $\{P(a_i); i = 1, 2, \ldots\}$ is called a **probability distribution** on S.

Definition 4 Once a probability distribution is known, the probability $P(A)$ of an event $A \subseteq S$ is defined as

$$P(A) = \sum_{a \in A} P(a).$$

Thus, it is the sum of the probabilities for all the simple events that make up A.

Uniform Probability Model

- Under an additional assumption, the above method of assigning probabilities gives us the *classical definition of probability*.

Definition. We say that a sample space $S = \{a_1, a_2, .., a_N\}$ has **equally probable** (or **equally likely**) simple events if the probability of every individual outcome is the same.

- Note that for equally likely simple events we must have

$$P(a_i) = \frac{1}{N} \text{ for any } a_i \in S, \tag{1}$$

and for $A \subseteq S$

$$P(A) = \frac{|A|}{N},$$

where $|A| =$ the number of elements in A.

- The distribution (1) is referred to as a (*discrete*) *uniform distribution* (or *uniform probability model*).

Example. Roll a fair four-sided die twice. What is the probability that the second number is larger than the first one?

Example. Consider drawing a card at random[7] from a standard 52 card deck of playing cards (13 kinds: A,2,3,4,...,10,J,Q,K, in 4 suits: ♣, ♢, ♡, ♠).

What is the probability that the card is a diamond?

For the sample space

$$S = \{2\diamondsuit, 3\diamondsuit, \ldots, Q\diamondsuit, K\diamondsuit, A\diamondsuit, 2\heartsuit, 3\heartsuit, \ldots\},$$

the answer is: 13/52=1/4.

For the sample space

$$S = \{spade, heart, diamond, club\}$$

the answer is: 1/4.

[7] At "random" means that all the outcomes are equally likely.

The term “odds”

- The term “odds” is sometimes used in describing probabilities.

Definition. The odds in **favour** of an event A is the probability with which the event occurs divided by the probability it does not occur:

$$\frac{P(A)}{1 - P(A)},$$

assuming that $P(A) > 0$.

The odds **against** the event is the reciprocal of this.

- Note that there is $1 \leftrightarrow 1$ relation between the probability of an event and the odds in favour of the event.

  For example: if the probability that the selected card is diamond is 1/4, then the odds in favour of diamonds are

  $$\frac{1/4}{3/4} = 1/3,$$

  that is 1:3. The odds against diamonds are 3:1.