40.001: Probability

Week 1: Counting principles, combinatorics (Chapter 1)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

1. Introduction
2. Counting principles
   - Binomial theorem

Course description

Read the *Course Description* on eDimension very carefully.

It contains information on office hours, assessment format, textbook and weekly topics.

The first homework set will be made available in Week 2.

Probability

- Probability models uncertainty and guides decision making under uncertainty.
- It is extremely important in any field that has quantitative data (science, engineering, finance, ...).
- It is a prerequisite for many ESD classes. It will also improve your logical thinking and problem solving skills.
- To do well:
  - Attend all classes, practice lots, read the textbook. There is no easy way.
  - You will need to take your own notes, and summarize them.
  - Do not use the 'solution manual' found online.
  - Do not forget freshmore maths, in particular, integration, proofs.

Why do we need so much maths?

Because humans are naturally illogical and irrational, have little intuition when estimating chances, and cannot cope with very large or very small numbers.

But in science, engineering, finance, etc, all decision making has to be rational.

Example 1

Many people would find the following statistic counter-intuitive:

Pick a random person on earth; what is the chance that this person lives in Singapore?

Another counter-intuitive example

Example 2

Susan is 30 years old, single, outspoken, and bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?

1. Susan is a bank teller.
2. Susan is a bank teller and is active in the feminist movement.

An example from combinatorics

Example 3

1. How many ways are there of holding up 3 fingers (using both hands)?

(A) 30 (B) 60 (C) 120 (D) 240

2. How many ways are there of holding up 7 fingers (using both hands)?

(A) 30 (B) 60 (C) 120 (D) 240

A demonstration

Example 4

In this room, what is the probability that at least 2 people share the same birthday?

Click here and enter your birthday.

A ridiculous comparison

Example 5

Which of the following will take longer?

1. Deal out 1 deck of cards every minute, until all possible decks are exhausted.

2. Brush off a layer of atoms in a 1cm $\times$ 1cm square every second, until a cube of diamond 1 light-year in each dimension disappears.

Solution

Situation 1 will take $52!/60/24/365 \approx 1.5 \times 10^{62}$ years.

Situation 2: 1 light-year $\approx 9.46 \times 10^{15}$ m; diamond bond length $\approx 154 \times 10^{-12}$ m, so it will take $\approx 1.7 \times 10^{54}$ years.

Overview of this course

This week: combinatorics (counting cleverly). Some of you have seen this already, but we will cover it with a higher level of rigour.

Next week: compute simple probabilities, by counting how many ways an event can occur, divided by the total number of outcomes.

Subsequent weeks: more complicated (such as conditional) probabilities; expected values.

End goal: understand the weak law of large numbers and the central limit theorem.

WLLN and CLT

The WLLN very roughly says that in the long run, the empirical average of some random events (e. g. repeating the same experiment) converges to the theoretical expected value.

Example: if you toss a fair coin repeatedly, then the proportion of H's will eventually get closer and closer to $1/2$, even if (say) the tosses started with 10 T's.

The CLT very roughly says that actually, the empirical average of these random events converges to a *bell curve* centred around the expected value.

Analogy: each random event is like the weather (short term and unpredictable); but if we consider the weather over the long term (and take some averages), then we get the climate, which is quite predictable.

Outline

Addition and multiplication principles

Basic enumeration is built up from two principles: addition and multiplication. The **addition principle** applies when you pick one object from multiple, disjoint categories:

Addition principle example

In how many ways can you type a letter or a digit on a keyboard?

Answer: $26 + 26 + 10 = 62$.

The **multiplication principle** applies when you perform an activity in multiple steps:

Multiplication principle example

You have 3 cars to choose from, and each can be fitted with 2 varieties of tyres and comes in 5 colours. How many choices for different cars do you have?

Answer: $3 \times 2 \times 5 = 30$.

Multiplication principle

Multiplication principle

If an activity is performed in k steps, where step 1 can be done in n_1 ways, step 2 in n_2 ways, ..., then the number of different possible activities is $n_1 \times n_2 \times \cdots \times n_k$.

Exercises

1. How many possible number plates of the format *3 letters - 4 numbers - 1 letter* can there be?

2. How many 8-character passwords can there be, if each character is either a letter or a digit? What if no repeats are allowed?

3. Suppose you can choose from t_1 T-shirts, t_2 shirts, t_3 singlets, b_1 pairs of pants, b_2 pairs of shorts, s_1 pairs of running shoes, s_2 pairs of sandals, or no shoes. What is the total number of ways to get dressed?

Permutations

Example

In how many ways can we arrange the letters of the word 'SINGAPORE'?

Answer: SINGAPORE has 9 letters with *no repetitions*.
$9 \times 8 \times 7 \times \cdots \times 2 \times 1 = 9! = 362880.$

Permutations

The number of *ordered* arrangements (permutations) of n *distinct* objects is given by $n!$ (n *factorial*).

More generally, the number of ordered arrangements of k objects from n distinct objects is given by

$$n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!} =: {}^{n}P_k.$$

Permutations – exercises

Exercise

A probability class has 6 boys and 4 girls. Students are ranked according to performance (no ties are allowed).

1. How many different rankings are possible?
2. If boys and girls are only ranked among themselves, how many rankings are possible?

10!, 4!6!

Exercise

A six-person committee composed of A, B, C, D, E, F is to select a chairperson, secretary, and treasurer.

1. In how many ways can this be done?
2. What if either A or B must be the chairperson?
3. What if F must hold an office?

6!/3!, 2x5x4, 3x5x4

Permutations with repetitions

Example

How many ways are there to rearrange the letters of 'KAYA'?

If the two A's were different, e. g. A_1 and A_2, then there would be $4!$ ways.
But they are the same, so we are counting everything $2!$ times.
So the answer is $4!/2! = 12$.

Permutations with repetitions

The number of permutations of n objects, of which n_1 are alike, n_2 are alike, ..., n_r are alike, is

$$\frac{n!}{n_1!\,n_2!\cdots n_r!}.$$

Exercise: how many signals, consisting of 9 flags in a line, can be made from 4 identical blue flags, 3 identical red flags, and 2 identical yellow flags?

Combinations

In some situations the order does not matter, for example, poker hands, or choosing 2 pillars to form a Capstone team.

Combination

The number of ways to *choose* k objects from n objects, when the order does not matter, is given by

$$\binom{n}{k} = {}^{n}C_k := \frac{{}^{n}P_k}{k!} = \frac{n!}{k!(n-k)!}.$$

Some useful identities:

$$\binom{n}{0} = 1, \quad \binom{n}{1} = n, \quad \binom{n}{2} = \frac{n(n-1)}{2}, \quad \binom{n}{k} = \binom{n}{n-k}.$$

Combinations – exercises

Exercise

1. In how many ways can 8 people in a race finish 1st, 2nd, and 3rd?

2. How many ways can 3 people from the 8 be disqualified?

Challenge

In a set of 8 antennas, 3 are defective. The antennas are otherwise indistinguishable. How many ways are there to place them in a row, such that no two defectives are consecutive?

6C3

Permutations with repetitions, another view

Example

If 12 people are to be divided into 3 committees of sizes 3, 4 and 5, how many divisions are possible?

Solution: $\binom{12}{3} \times \binom{12-3}{4} = \frac{12!}{3!\,4!\,5!}$.

In general, if you have n distinct objects to be divided into r distinct piles of sizes $n_1, n_2, \dots, n_r$ (where $n_1 + \cdots + n_r = n$), and the order within each pile does not matter, then the number of ways is given by

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n_r}{n_r} = \frac{n!}{n_1!\,n_2!\cdots n_r!}.$$

Note that this is the same as the formula for permutation with repetitions.

Binomial theorem

Motivating example:

$$(a+b)^2 = (a+b)(a+b) = a^2 + ab + ba + b^2 = 1a^2 + 2ab + 1b^2.$$

In general, in the expansion of $(a+b)^n$, the coefficient of a^k is the number of ways to select k of the a's from the n brackets. Moreover, the other $(n-k)$ choices must be b's.

Binomial theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k}.$$

Therefore, $\binom{n}{k}$ are also called *binomial coefficients*. They form *Pascal's triangle*.

Binomial theorem – exercises

1. Expand $(3a+2)^4$.
2. Simplify $\displaystyle\sum_{k=0}^{n}\binom{n}{k}$.
3. Find at least two ways to show that

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

40.001: Probability

Week 1: Counting principles, combinatorics (Chapter 1)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Passwords

Q1. Which password below is most secure?

A. Why!Asks2Me

B. !sec&ures!

C. sutdstu1new

D. cx3y24z8T7

The intended answer to this security quiz question is probably D, since it looks most random.

Is it really?

Passwords

Assume that passwords of type D are random characters up to length 10.

Assume that passwords of type A are 3 random English words separated by 2 random characters.

Say there are 72 possible characters (including punctuation etc).

For passwords of type D, there are

$$\sum_{i=1}^{10} 72^i \approx 3.8 \times 10^{18}$$

possibilities.

Passwords

For passwords of type A, there are about 2×10^5 words in English (not even counting capitalization and simple substitutions), so there are at least

$$(2 \times 10^5)^3 \times 72^2 \approx 4.1 \times 10^{19}$$

possibilities. Therefore A is actually better.

Refer to this XKCD comic.

Outline

A summary

Given n distinct objects, we have studied:

- The number of *ordered* arrangements of k of them, *repeats allowed*: n^k, multiplication principle.
- The number of *ordered* arrangements of k of them, *no repeats* allowed: ${}^nP_k$, permutation.
- The number of *unordered* arrangements of k of them, *no repeats* allowed: ${}^nC_k$, combination.

There is one more case.

Multichoose

Question: how many ways to buy 6 books, chosen from 3 titles?

This is an example of: given n distinct objects, find the number of *unordered* arrangements of k of them, *repeats* allowed.

Here $n = 3$, $k = 6$.

To solve this problem, we use a trick: represent the k objects as k dots, and separate them into n categories using $(n-1)$ bars:

$$\bullet \mid \bullet \ \bullet \ \bullet \mid \bullet \ \bullet$$

The number of ways this can be done is

$$\binom{k+n-1}{n-1},$$

since there are $(k+n-1)$ items in total and we choose $(n-1)$ of them to be bars.

Variations on what we have done

Exercise – multichoose

What is the answer to the books question on the last slide?

8C2=28

Exercise – forming blocks

1. How many ways are there to sit A, B, C, D, E in a row, if A and B must be next to each other?
2. How many ways are there to sit A, B, C, D, E in a row, if A must be to the left of B?

4!2!, 5!/2!

Exercise – circular symmetry

How many ways are there to sit n people around a circle?

(n-1)!

More variations on what we have done

Exercise – avoid double counting

How many integers between 1 and 61 are divisible by 3 or 5?

20+12-4

Exercise – consider the complement

How many n-character passwords are there that contains at least one digit? (Each character must be either a letter or a digit.)

62^n-52^n

40.001: Probability

Week 2: Sample space and events, axioms of probability (Chapter 2)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

Sample space, events

In probability, we often deal with experiments whose outcomes are random.

The set of all possible outcomes of an experiment is called the sample space, which we can denote by S.

We are often only interested in a *subset* of the sample space, which is known as an event; we can denote it by E.

'E is a subset of S' is written as $E \subseteq S$. Before we proceed, let us recall some set theory.

Set theory, revisited

A *set* is a well-defined collection of distinct objects. We call these objects 'elements' of the set.

Examples:

- All the days in a non-leap year form a set, Y.
- $\mathbb{Z}$ is the set of all integers; $\mathbb{R}$ is the set of all real numbers.
- $S_1 = \{\text{Heads}, \text{Tails}\}$ is the sample space of the experiment of flipping a coin.
- $S_2 = \{1, 2, 3, 4, 5, 6\}$ is the sample space of the experiment of rolling a dice.
- $E_2 = \{2, 4, 6\}$ is a subset of S_2; it corresponds to the event that the dice roll results in an even number.

Set notations

The symbol $\in$ means 'is an *element* of'; the symbol $\subseteq$ means 'is a *subset* of'.

For example,

$$3 \in S_2, \quad 7 \notin S_2, \quad E_2 \subseteq S_2, \quad S_1 \nsubseteq S_2.$$

Also,

$$E_2 \notin S_2, \qquad 6 \nsubseteq E_2.$$

The number of elements in a set S (the *cardinality* of S) is denoted by $|S|$. For example, $|Y| = 365$.

The *empty set* is denoted by $\varnothing$.

Set operations

Let U be called a universal set if all sets under consideration are subsets of U (e. g. U could be the sample space of an experiment).

Then the **complement** of a set A is the set of all elements in U but not in A. It is denoted by A^c. Example: $E_2^c = \{1, 3, 5\}$.

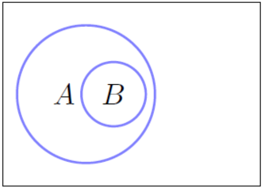
The **union** of sets A and B, denoted by $A \cup B$, is the set whose elements are either in A *or* B (or both).

The **intersection** of sets A and B, denoted by $A \cap B$, is the set whose elements are in both A *and* B.

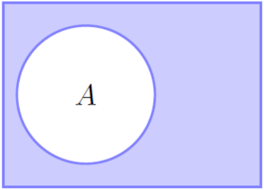
These notions can be represented on Venn diagrams.

Venn diagrams

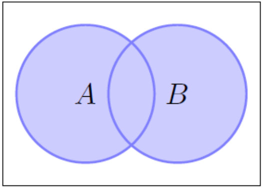
$B \subseteq A$



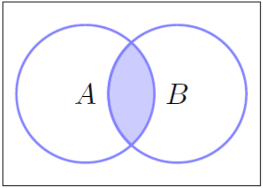
A^c



$A \cup B$



$A \cap B$



Set exercises

Note: the textbook uses a different notation for intersection.

Exercises

1. What is $\varnothing^c$?
2. What is $A \cap A^c$?
3. Use Venn diagrams to show that

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

4. Use Venn diagrams to prove *de Morgan's laws*,

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

Outline

Interpretations of probability

There are many interpretations of what a probability actually represents, and experts do not completely agree with each other.

Two major interpretations are:

- *Frequentist*: the probability of E is the limiting frequency that E occurs when we repeat an experiment under the same conditions.

  Problem: some experiments cannot be repeated. For example, how would you answer $\mathbb{P}$(you will pass this course), or $\mathbb{P}$(Shakespeare wrote *Hamlet*)?

- *Bayesian*: probability measures one's degree of belief, and the belief can be updated when new evidence arises.

  Problem: belief is subjective. For example, how would you find $\mathbb{P}$(the sun will rise tomorrow)?

Axioms of probability

To overcome these discrepancies, modern probability is developed from three *axioms* (self-evident assumptions), and is independent of any interpretation.

Consider an experiment with sample space S. For each event E, there exists a number $\mathbb{P}(E)$ (the *probability* of E), and it satisfies:

The Axioms of probability

1. $$0 \leq \mathbb{P}(E) \leq 1.$$

2. $$\mathbb{P}(S) = 1.$$

3. For any (infinite) sequence of *mutually exclusive* events $E_1, E_2, \ldots$ (that is, $E_i \cap E_j = \varnothing$ for all pairs of distinct i, j),

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

Some consequences of the Axioms

The Axioms are really a minimal set of assumptions. For instance, we can show using the Axioms that $\mathbb{P}(\varnothing) = 0$:

Take $E_1 = S$, $E_2 = E_3 = \cdots = \varnothing$. Then Axiom 3 gives

$$\mathbb{P}(S) = \mathbb{P}(S) + \sum_{i=2}^{\infty} \mathbb{P}(\varnothing),$$

which leads to the desired result.

A consequence which follows is that, for any *finite* sequence of mutually exclusive events $E_1, E_2, \ldots, E_n$,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mathbb{P}(E_i).$$

Some simple calculations

Example 1

1. Consider tossing a *fair* coin. $S = \{\mathsf{H}, \mathsf{T}\}$. Fairness means $\mathbb{P}(\{\mathsf{H}\}) = \mathbb{P}(\{\mathsf{T}\})$.

Using the finite version of Axiom 3, we get

$$\mathbb{P}(S) = \mathbb{P}(\{\mathsf{H}\} \cup \{\mathsf{T}\}) = \mathbb{P}(\{\mathsf{H}\}) + \mathbb{P}(\{\mathsf{T}\}).$$

Using Axiom 2, this equates to 1.
Therefore $\mathbb{P}(\{\mathsf{H}\}) = \mathbb{P}(\{\mathsf{T}\}) = \frac{1}{2}$.

Example 2

Similarly, with an *unloaded* dice,

$$\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = \frac{1}{6}.$$

Hence, $\mathbb{P}(\{2, 4, 6\}) = \frac{1}{2}$ using Axiom 3.

Exercises

Prove the following results using the Axioms

1. If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.
2. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.
3. For any events A and B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

The last result is an example of the *inclusion-exclusion principle*. The corresponding formula for 3 events is

$$\begin{aligned} \mathbb{P}(A \cup B \cup C) = {} & \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ & - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) - \mathbb{P}(C \cap A) \\ & + \mathbb{P}(A \cap B \cap C). \end{aligned}$$

(The pattern continues when more events are involved.)

Reduction to enumeration

A useful consequence of the Axioms is that, for a finite sample space S where all the *basic events* are equally likely,

$$\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{\#\text{ ways } E \text{ can happen}}{\#\text{ possible outcomes}}.$$

Exercises

1. In a lottery, you choose a set of 6 numbers *without replacement* from 45 numbers. What is the probability that it will be the same set of 6 numbers drawn by a machine?

1/(45 C 6)

2. Two unloaded dice are rolled. Find the probability that their sum is 8.

5/36

3. What is the probability of drawing either a diamond or an ace from a deck of 52 cards?

4/13

More exercises

4. A box has 5 black and 3 white balls. Two balls are randomly drawn one after the other, without replacement. What is the probability that both balls drawn are black? What about at least 1 white?

(5*4)/(8*7), 9/14

5. Find the probability of obtaining a *four of a kind* in a hand of 5 cards drawn from a standard deck (that is, the cards have denominations XXXXY).

13*48/(52 C 5)

6. In a class of 40 students, what is the probability that at least one student's phone number ends in 88? (Assume that the last 4 digits of phone numbers are random.)

1-(99/100)^40

7. Three couples sit randomly in a row. Find the probability that at least one couple sits together.

Hint: inclusion-exclusion.

Solution to exercise 7

Let A, B and C respectively denote the events that couple 1, couple 2 and couple 3 sit together.

We want to find $\mathbb{P}(A \cup B \cup C)$. Using the formula on the bottom of slide 14:

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) &= \frac{2 \times 5!}{6!} + \frac{2 \times 5!}{6!} + \frac{2 \times 5!}{6!} \\ &\quad - \frac{2^2 \times 4!}{6!} - \frac{2^2 \times 4!}{6!} - \frac{2^2 \times 4!}{6!} \\ &\quad + \frac{2^3 \times 3!}{6!} \\ &= \frac{2}{3}.\end{aligned}$$

40.001: Probability

Week 2: Sample space and events, axioms of probability (Chapter 2)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

1. Harder problems
   - Hypergeometric
   - Birthday problem
   - Derangement

Hypergeometric distribution

Hypergeometric distribution

A committee of k people is chosen randomly from n men and m women. What is the probability that it ends up with r men?

There are $\binom{n+m}{k}$ ways to choose a committee without restrictions.

For the committee to have r men, it must contain $k-r$ women. The number of ways to choose such a committee is $\binom{n}{r}\binom{m}{k-r}$.

Therefore,

$$\mathbb{P}(r \text{ men on a committee of } k) = \frac{\binom{n}{r}\binom{m}{k-r}}{\binom{n+m}{k}}.$$

Exercise

An exam has 2 sections, each with 10 questions. If a student randomly chooses 12 questions to answer, what is the probability that 6 questions come from each section?

Birthday problem

What is the probability that among n people, at least 2 people share the same birthday?

We first calculate the probability of no two people share the same birthday. Ignore leap years, and assume that a person is equally likely to be born on any day of the year.

The 1st person can be born on any of the 365 days, the 2nd person can be born on any of the remaining 364 days, etc. So this probability is given by

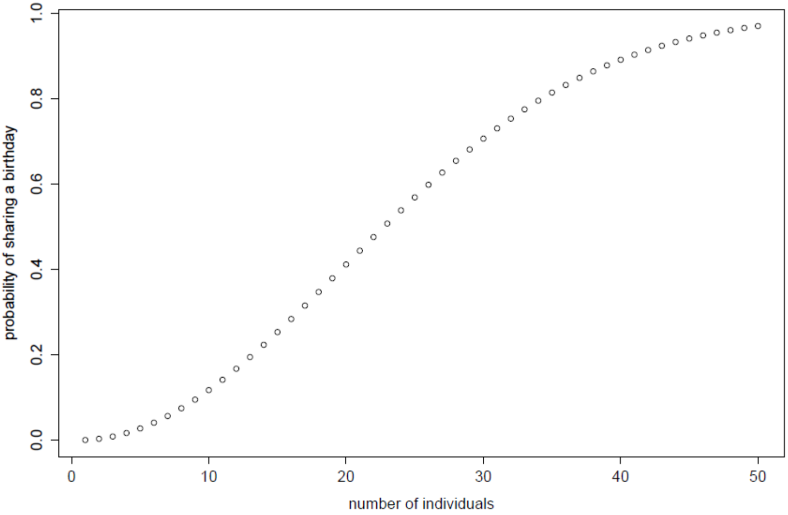
$$\frac{365 \times 364 \times \cdots \times (366-n)}{365 \times 365 \times \cdots \times \quad 365}.$$

Therefore the answer to the birthday problem is

$\mathbb{P}$(at least 2 people among n share the same birthday)

$$= 1 - \frac{365!}{(365-n)!\,365^n}.$$

Birthday problem



Also, check out this article.

Birthday problem – generalization

Suppose now that there are d days in a year. The probability $P(d,n)$ that at least 2 people share the same birthday is

$$P(d,n) = 1 - \left(1 - \frac{1}{d}\right)\left(1 - \frac{2}{d}\right)\cdots\left(1 - \frac{n-1}{d}\right).$$

Using the Taylor approximation $e^{-x} \approx 1 - x$ when $|x|$ is small, we get

$$\begin{aligned} P(d,n) &\approx 1 - e^{-1/d}\, e^{-2/d} \cdots e^{-(n-1)/d} \\ &= 1 - e^{-n(n-1)/(2d)}. \end{aligned}$$

For this probability to be $\frac{1}{2}$, $n \approx \sqrt{\ln(4)d}$.

Example

How likely is it for lottery numbers (6 numbers chosen from 1–45) to repeat? What about arrangements of a deck of cards?

Derangement

A *derangement* is a permutation in which no object appears in its original position.

For example: in how many ways can 3 hats be returned to 3 men so that no man gets his own hat?

The number of derangements of n objects, D_n, is

$$D_n = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right).$$

This formula can be proved using the general inclusion-exclusion principle (let E_i be the event that the ith man gets his own hat). However we will compute D_n using a different approach.

A recurrence relation

Denote the men by $A_1, A_2, \ldots, A_n$.

There are $(n-1)$ ways for A_1 to pick a hat. Say he picks A_i's hat. Now there are two cases:

If A_i picks A_1's hat, then there are D_{n-2} ways for the remaining men to pick their hats.

If A_i does not pick A_1's hat, then we may pretend that A_1's hat is actually A_i's (so he is not allowed to pick it); hence there are D_{n-1} ways for the remaining men to pick their hats.

Therefore,

$$D_n = (n-1)\big(D_{n-1} + D_{n-2}\big).$$

Together with the initial conditions $D_1 = 0$, $D_2 = 1$, this formula allows us to compute D_n for any n.

It can be shown that the two formulas given for D_n agree.

Exercises

1. If 5 hats are randomly returned to 5 men, what is the probability that at least 2 men receive their own hats?

1 - D5/5! - 5 D4/5! = 31/120

2. Two decks of cards are dealt side-by-side. What is the probability that no two cards match?

roughly 1/e

40.001: Probability

Week 3: Conditional probability (Chapter 3)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

Motivation

Problem (diamonds)

There are 3 drawers: one has 2 diamonds, another has 2 rubies, and the last has 1 diamond and 1 ruby. You randomly pick a drawer, reach in and take out a diamond. What is the probability that the other gem in the drawer is a diamond?

Problem ('Monty Hall')

You face 3 doors, 2 of which conceal a goat and the other conceals a car. You choose a door at random with the hope of winning the prize behind it, and the host at random opens a door that you didn't choose to reveal a goat. Should you switch your choice?

You can play the game here.

Conditional probability – introduction

Roll 2 fair dice. Let B be the event that the sum of the two numbers is 10.

$$\mathbb{P}(B) = \mathbb{P}(\{(4,6),(5,5),(6,4)\}) = \frac{|B|}{|S|} = \frac{3}{36}.$$

Let A be the event that the first dice shows a 4.

$$\mathbb{P}(A) = \mathbb{P}(\{(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\}) = \frac{|A|}{|S|} = \frac{6}{36}.$$

What is the probability that the first dice shows a 4, *given* that the sum is 10?

$$\mathbb{P}(A|B) = \frac{|\{(4,6)\}|}{|\{(4,6),(5,5),(6,4)\}|} = \frac{1}{3}.$$

Note that this also equals

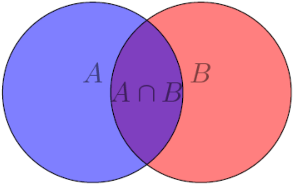
$$\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/36}{3/36}.$$

This is an example of *conditional probability*.

Conditional probability – intuition

The probability of A *given* B is written as $\mathbb{P}(A|B)$.

Given that B has happened, we can think of the sample spaced as being restricted to B.



In the case where all the basic events are equally likely,

$$\mathbb{P}(A|B) = \frac{|A \cap B|}{|B|} = \frac{|A \cap B|}{|S|}\frac{|S|}{|B|} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

In general, we use the above formula as a definition of conditional probability.

Conditional probability – definition

Definition (conditional probability)

If $\mathbb{P}(B) > 0$, then

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

An immediate consequence is the formula

$$\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B) = \mathbb{P}(A)\mathbb{P}(B|A).$$

Conditional probability – example

There are 100 applicants for a teaching position.

	PhD	no PhD
≥ 2 yrs teaching	21	15
< 2 yrs teaching	28	36

One of them is selected at random.

1. Given that an applicant has at least 2 years of teaching experience, what is the probability that s/he has a PhD?

2. Given that an applicant has a PhD, what is the probability that s/he has at least 2 years of teaching experience?

Conditional probability – exercises

Exercise 1

A box has 5 black and 3 white balls. Two balls are randomly drawn one after the other, without replacement. What is the probability that both balls drawn are black?

Do this using the second formula on slide 6.

Exercise 2

A man draws 2 cards randomly without replacement from 4 kings and 4 aces.

1. If he has at least 1 ace, what is the probability that he has 2 aces?

2. If he has the ace of hearts, what is the probability that he has 2 aces?

Some guidelines to problem solving

1. Work out what the question is asking.
2. Define the events explicitly.
3. Choose an appropriate formula to use.
4. Work out each probability in the formula separately.

Conditional probability – solutions

Exercise 1

Let A = first ball is black, B = second ball is black.

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A) = \frac{5}{8} \times \frac{4}{7} = \frac{5}{14}.$$

Note how this is consistent with, though slightly different from, using the multiplication principle.

Exercise 2

1. $\mathbb{P}(2 \text{ aces} \,|\, \text{at least 1 ace}) = \mathbb{P}(2 \text{ aces})/\mathbb{P}(\text{at least 1 ace})$.

   $\mathbb{P}(2 \text{ aces}) = (4 \times 3)/(8 \times 7) = 3/14$.

   $\mathbb{P}(\text{at least 1 ace}) = 1 - \mathbb{P}(2 \text{ kings}) = 1 - 3/14 = 11/14$.

Therefore the answer is $3/11$.

2. A similar calculation gives $3/7$. Note that you can check your answers here by listing the sample space.

Law of total probability

If events $B_1, B_2, \ldots, B_n$ are mutually exclusive and *exhaustive* (that is, $\bigcup_{i=1}^n B_i = S$), then for any event A, we have

Law of total probability

$$\mathbb{P}(A) = \mathbb{P}(B_1)\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\mathbb{P}(A|B_2) + \cdots + \mathbb{P}(B_n)\mathbb{P}(A|B_n).$$

Example

An insurance company classifies accident proneness as high (0.1 of the population, $\mathbb{P}(\text{accident}|\text{highly prone}) = 0.3$), medium ($0.3$ of the population, $\mathbb{P} = 0.1$) and low (everyone else, $\mathbb{P} = 0.02$).

An accident happened; what is the probability that it happened to someone highly prone?

$\mathbb{P}(\text{accident}) = 0.1 \times 0.3 + 0.3 \times 0.1 + 0.6 \times 0.02 = 0.072.$

So the answer is $(0.1 \times 0.3)/0.072 = 5/12$.

Exercise

1. In a box of n chocolates, k of them are poisoned. Three chocolates are drawn at random, *with* replacement. What is the probability that the first one is poisoned? What about the second one? The third one?

2. Now repeat the problem but *without* replacement.

k/n in all cases

(The observation here holds true in general.)

Diamonds problem

$$\begin{aligned}
\mathbb{P}(\text{2nd is diamond} \mid \text{1st is diamond}) &= \frac{\mathbb{P}(\text{both are diamonds})}{\mathbb{P}(\text{1st is diamond})} \\
&= \frac{1/3}{1/3 \times 1 + 1/3 \times 0 + 1/3 \times 1/2} \\
&= \frac{2}{3}.
\end{aligned}$$

Some things to note:

- Intuition: each of the 3 diamonds has equal chance of being chosen.
- Just because there are two seemingly symmetric outcomes doesn't mean that the probability of each is $1/2$.
- The problem can become easier to understand once you try to write a program to simulate it.

Monty Hall problem

Intuition: the host is actually giving you extra information. By not switching (i. e. ignoring this information), your probability of winning remains the same as the probability that your initial guess is right, which is $1/3$.

Solution using conditional probability: label the doors; say you picked door 1 and the host opened door 2. Then the probability of winning by switching is

$$\mathbb{P}(\text{car behind 3}|\text{host opened 2}) = \frac{\mathbb{P}(\text{car behind } 3 \cap \text{host opened 2})}{\mathbb{P}(\text{host opened 2})}$$

$$= \frac{1/3}{\mathbb{P}(\text{car behind } 1 \cap \text{host opened 2}) + \mathbb{P}(\text{car behind } 3 \cap \text{host opened 2})}$$

$$= \frac{1/3}{1/3 \times 1/2 + 1/3}$$

$$= \frac{2}{3}.$$

40.001: Probability

Week 3: Bayes' rule, independence (Chapter 3)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Reminder

Homework 1 is due on Thursday, at 6pm.

Show working; provide enough details so that someone who's taking the class (but not excelling at it) can follow your reasoning.

Outline

Bayes' rule

Bayes' rule (also known as Bayes' theorem) allows us to work out $\mathbb{P}(A|B)$ from $\mathbb{P}(B|A)$.

Bayes' rule

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

where $\mathbb{P}(B)$ can be computed from the law of total probability:

$$\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c).$$

Example

Suppose a breathalyzer displays false positive in 5% of the cases in which the driver is sober, but never fails to detect a truly drunk person. 1 in 200 drivers is driving drunk. If a breathalyzer shows a positive result, what is the probability the driver really is drunk?

9%

Exercises

1. A rare disease affects 0.5% of the population. A test for the disease is 98% accurate. If a person gets tested positive, what is the probability that s/he actually has the disease?

2. In a crime investigation, it is 70% certain that suspect A is guilty. It is known that the criminal is left-handed, a trait found in 10% of the population. If suspect A also turns out to be left-handed, then how certain now can we be of his guilt?

3. Suspect X is convicted purely due to a DNA match. The probability of a match by chance is 10^{-6}. However, there are 5×10^5 eligible candidates in the area who could be guilty. What is the probability that X is guilty?

Solutions

1. Let $P =$ tested positive, and $D =$ has the disease.

$$\mathbb{P}(D|P) = \frac{\mathbb{P}(P|D)\mathbb{P}(D)}{\mathbb{P}(P|D)\mathbb{P}(D) + \mathbb{P}(P|D^c)\mathbb{P}(D^c)}$$

$$= \frac{0.98 \times 0.005}{0.98 \times 0.005 + 0.02 \times 0.995} \approx 20\%.$$

2. Let $G = A$ is guilty, and $L = A$ is left-handed. We wish to find $\mathbb{P}(G|L)$, which can be thought of as an 'update' of the previous knowledge $\mathbb{P}(G)$, given the new information of left-handedness.

$$\mathbb{P}(G|L) = \frac{\mathbb{P}(L|G)P(G)}{\mathbb{P}(L|G)\mathbb{P}(G) + \mathbb{P}(L|G^c)\mathbb{P}(G^c)}$$

$$= \frac{1 \times 0.7}{1 \times 0.7 + 0.1 \times 0.3} \approx 0.96.$$

Solutions

3. Let $G = X$ is guilty, and $M = X$ has a DNA match.

$$\mathbb{P}(G|M) = \frac{\mathbb{P}(M|G)P(G)}{\mathbb{P}(M|G)\mathbb{P}(G) + \mathbb{P}(M|G^c)\mathbb{P}(G^c)}$$

$$= \frac{1 \times (5 \times 10^5)^{-1}}{1 \times (5 \times 10^5)^{-1} + 10^{-6} \times (1 - (5 \times 10^5)^{-1})} \approx \frac{2}{3}.$$

Intuition: on average, you'd expect another '0.5 person' in the area to have a match...

Outline

1. Bayes
2. Independence
   - Distributions

Independence

Definition: if for two events A and B, $\mathbb{P}(A|B) = \mathbb{P}(A)$, then we say that A and B are independent.

A and B are *independent* if *any* of the following holds:

- $\mathbb{P}(A|B) = \mathbb{P}(A)$
- $\mathbb{P}(B|A) = \mathbb{P}(B)$
- $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

These three conditions are equivalent.

- For independent events A and B, the knowledge of A does not alter the chances of B, and vice versa.
- Coin tosses, dice rolls, and drawing balls with replacement are independent events.
- Events that are 'independent' in real life can be treated as independent in probability, however the reverse is not true.

Independence, continued

Example

Roll 2 fair dice. Let:

- $E =$ the sum of the two dice is 7,
- $F =$ the first dice shows a 4,
- $G =$ the second dice shows a 3.

Not surprisingly, F and G are independent.

More surprisingly, E and F are independent, also E and G are independent!

However, E and $F \cap G$ are not independent.

Motivated by situation like the above, we define *three* events A, B, C to be independent if they are pairwise independent, *and* $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

Independence, exercises

Exercises

1. If A and B are independent, show that A and B^c are also independent.

2. Let $S =$ a person is a smoker, and $C =$ a person has lung cancer. Suppose we know that

$$\mathbb{P}(C|S) > \mathbb{P}(C|S^c).$$

What relationship can you deduce for $\mathbb{P}(S)$, $\mathbb{P}(C)$ and $\mathbb{P}(S \cap C)$?
Are S and C independent?
Which is larger, $\mathbb{P}(S|C)$ or $\mathbb{P}(S|C^c)$?

Distributions

Consider repeatedly tossing a weighted coin, where $\mathbb{P}(\mathsf{T}) = p$.

Geometric distribution

What is the probability that the first tail appears on the nth toss?

Due to independence, the answer is $(1-p)^{n-1}p$.

Binomial distribution

Toss the coin n times; what is the probability that exactly k tails are obtained?

Exactly k tails implies exactly $n-k$ heads. Using independence and considering all the ways to *arrange* the k tails, we get

$$\mathbb{P}(k \text{ tails}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Binomial distribution

The binomial distribution can be used to model situations where an experiment is repeated n times, each experiment has a probability of success p, and is independent of any other experiment. Each such experiment is known as a *Bernoulli* trial.

Exercises

1. In a family of 4 children, what is the probability that there are 2 boys and 2 girls?

2. If you can hit a target with probability 0.85, what is the probability that you can hit it at least 9 out of 10 times?

3. Airlines often overbook their flights. Suppose each passenger has a 3% chance of not showing up. What is the probability that everyone gets a seat in a 100-seat flight if 102 people were booked?

2: 0.5443; 3: about 81%

40.001: Probability

Week 4: More conditional probability (Chapter 3)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

Information

Homework 1 solutions and some extra notes for Week 3 are uploaded on eDimension.

Homework 2 is available today.

On **showing working**:

- Not showing working is a *bad habit*, and needs to be corrected now.
    - Would you only write down the answer in a project report or a business report, especially in a collaborative environment?
- You will lose marks for not showing working.

Show working

- If you make a rough draft when solving a problem, just transfer the draft to the final version of your homework.
  - Doing so also gives you an opportunity to check your work.
  - If your rough draft is too wordy, practise making it more concise.
  - You don't have to write everything in mathematical symbols; it's fine to explain your reasoning in normal sentences.
- If you don't make a rough draft but prefer to do everything in one line or in your head, *change this habit* now.
  - Breaking down problems into smaller steps is the only way to approach difficult problems; it also reduces your error rate.

Example

Example: HW1 Q6. A fair coin is tossed four times. Determine the probability that a sequence of three or more consecutive tails is obtained.

- (Solution 1) $\frac{3}{16}$ **0/1**
- (Solution 2) $\mathbb{P}(3 \text{ or more consecutive tails}) = \frac{3}{16}$ **0/1**
- (Solution 3) 3 ways to get 3 or more consecutive tails.
  $\mathbb{P}(3 \text{ or more consecutive tails}) = \frac{3}{2^4} = \frac{3}{16}$. **0.5/1**
- (Solution 4) 3 ways to get 3 or more consecutive tails: TTTH, HTTT, TTTT. 2^4 possible outcomes in total.
  $\mathbb{P}(3 \text{ or more consecutive tails}) = \frac{3}{2^4} = \frac{3}{16}$. **1/1**

Copying

On **copying homework**:

- We are keeping track of who is (probably) copying.
- Direct copying will result in a 0 for the assessment item, regardless of whether you copied, or allowed others to copy from you.
- Repeated offenders will be sent to a disciplinary committee. Likely penalties include: failure of subject; termination of scholarship; suspension; termination of candidature.

Outline

Conditional probability, revisited

It may be checked that $\mathbb{P}_C(A) := \mathbb{P}(A|B)$, when $\mathbb{P}(B) > 0$, is a probability function. Namely,

- $0 \leq \mathbb{P}_C(E) \leq 1$.
- $\mathbb{P}_C(S) = 1$.
- For mutually exclusive events E_i,

$$\mathbb{P}_C\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}_C(E_i).$$

Therefore, all the results proven for probabilities also apply to $\mathbb{P}_C$. For instance,

- $\mathbb{P}_C(E^c) = 1 - \mathbb{P}_C(E)$.
- $\mathbb{P}_C(E_1 \cup E_2) = \mathbb{P}_C(E_1) + \mathbb{P}_C(E_2) - \mathbb{P}_C(E_1 \cap E_2)$.

Conditional independence

E_1 and E_2 are said to be *conditionally independent* given B if

$$\mathbb{P}_C(E_1 \cap E_2) = \mathbb{P}_C(E_1)\,\mathbb{P}_C(E_2),$$

or equivalently,

$$\mathbb{P}\big((E_1 \cap E_2)|B\big) = \mathbb{P}(E_1|B)\,\mathbb{P}(E_2|B).$$

Example

An insurance company has two types of clients:

- Accident-prone (30%): $\mathbb{P}(\text{accident in a year}\,|A) = 0.4$,
- Not accident-prone (70%): $\mathbb{P}(\text{accident in a year}\,|A^c) = 0.2$.

Assume that for each client type, accidents in different years are independent.

What is the probability that a new client will have an accident in the 2nd year, given that she had an accident in the 1st year?

Solution

- $A =$ client is accident-prone,
- $A_1 =$ accident in 1st year,
- $A_2 =$ accident in 2nd year.

We want to find $\mathbb{P}(A_2|A_1)$.

Expand it out and apply the law of total probability:

$$\mathbb{P}(A_2|A_1) = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)}$$
$$= \frac{\mathbb{P}(A)\,\mathbb{P}(A_1 \cap A_2|A) + \mathbb{P}(A^c)\,\mathbb{P}(A_1 \cap A_2|A^c)}{\mathbb{P}(A)\,\mathbb{P}(A_1|A) + \mathbb{P}(A^c)\,\mathbb{P}(A_1|A^c)}.$$

Now make use of the assumption of conditional independence: $\mathbb{P}(A_1 \cap A_2|A) = \mathbb{P}(A_1|A)\,\mathbb{P}(A_2|A)$ etc. Therefore

$$\mathbb{P}(A_2|A_1) = \frac{0.3 \times 0.4 \times 0.4 + 0.7 \times 0.2 \times 0.2}{0.3 \times 0.4 + 0.7 \times 0.2} \approx 0.29.$$

Outline

Problem of points

Player A needs n points to win a game, while player B needs m points to win. The probability of A winning a point is p, that of B winning a point is $1-p$. Assume independence.

What is the probability that A wins the game?

There are many ways to solve this problem, and it is not obvious that the answers they give agree!

One way is to count the *arrangements* of $n-1$ points for A and i points for B, where $i \leq m-1$, followed by 1 more point for A.

Let the desired probability be $P_{n,m}$. Then

$$P_{n,m} = \sum_{i=0}^{m-1} \binom{n-1+i}{i} p^n (1-p)^i.$$

Problem of points

Another way is to consider the outcome of the first round, and use the law of *total probability*. Then

$$P_{n,m} = p\,P_{n-1,m} + (1-p)\,P_{n,m-1}.$$

Together with the boundary conditions $P_{i,0} = 0$ and $P_{0,j} = 1$ (where $i, j > 0$), this recurrence allows us to determine the value of $P_{n,m}$ for all $n, m > 0$.

A third way is to imagine that the game *continues* for $m+n-1$ rounds (which may go beyond the official end). A wins the game if and only if he wins at least n of these rounds.

Therefore

$$P_{n,m} = \sum_{k=n}^{m+n-1} \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}.$$

Exercise

A and B are two mutually exclusive outcomes of an experiment, with $\mathbb{P}(A) = a$ and $\mathbb{P}(B) = b$. If independent trials of the experiment are repeatedly conducted, what is the probability that A occurs before B?

Find two ways to do this problem.

Solution

Method 1: infinite *geometric* sum. Let $C = (A \cup B)^c$. Then A occurs before B if and only if the sequence of outcomes begins with A, or CA, or CCA, ...

(We don't actually care what happens afterwards, since A has already occurred before B.)

$$\begin{aligned}\mathbb{P}(E) := \mathbb{P}(A \text{ before } B) &= a + (1-a-b)a + (1-a-b)^2 a + \cdots \\ &= \frac{a}{1-(1-a-b)} = \frac{a}{a+b}.\end{aligned}$$

Method 2: consider the outcome of the first trial, and use the law of *total probability*.

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(A \text{ first})\,\mathbb{P}(E|A \text{ first}) + \mathbb{P}(B \text{ first})\,\mathbb{P}(E|B \text{ first}) \\ &\quad + \mathbb{P}(C \text{ first})\,\mathbb{P}(E|C \text{ first}) \\ &= a \times 1 + b \times 0 + (1-a-b) \times \mathbb{P}(E).\end{aligned}$$

We then solve the above equation for $\mathbb{P}(E)$!

Gambler's ruin

A coin has $\mathbb{P}(\mathsf{H}) = p$, $\mathbb{P}(\mathsf{T}) = 1 - p = q$. Every time it shows H, player A wins \$1 from player B; otherwise he loses \$1. They continue doing this until one wins all the money. If A starts with \$$k$ and B with \$$(N - k)$, what is the probability that A wins?

Let $P_k := \mathbb{P}(A \text{ wins starting with } \$k)$; conditioned on the outcome of the first toss, we get

$$P_k = p\,P_{k+1} + q\,P_{k-1}.$$

Hence, for $1 \leq k \leq N - 1$,

$$\begin{aligned} p\,P_k + q\,P_k &= p\,P_{k+1} + q\,P_{k-1}, \\ P_{k+1} - P_k &= \frac{q}{p}(P_k - P_{k-1}). \end{aligned}$$

It follows after some work that

$$P_2 - P_1 = \frac{q}{p}\,P_1,\ P_3 - P_2 = \left(\frac{q}{p}\right)^2 P_1, \ldots, P_k - P_{k-1} = \left(\frac{q}{p}\right)^{k-1} P_1.$$

Gambler's ruin

Add them up:

$$P_k - P_1 = P_1 \left[\frac{q}{p} + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{k-1} \right].$$

When $p \neq 1/2$, this geometric series simplifies to

$$P_k = \frac{1 - (q/p)^k}{1 - q/p} P_1.$$

Since $P_N = 1$, by setting $k = N$ above, we can solve for P_1.
Finally, we arrive at the answer

$$P_k = \frac{1 - (q/p)^k}{1 - (q/p)^N}.$$

Exercises

1. Find the solution for a fair coin ($p = 1/2$).
2. What is the probability that the game goes on forever?

40.001: Probability

Week 4: Random variables (Chapter 4)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

What is a random variable?

A random variable is a variable whose possible values are *numerical* representations of a random experiment.

Examples:

- The sum of two fair dice rolls.
- The number of T's in n tosses of a coin.
- The number of coin tosses until a H shows up.
- The number of emails you receive in a day.
- The time until a light-bulb expires.

More formally, a random variable is a function from the sample space to the real numbers.

Note: a good understanding of random variables is crucial for many ESD courses.

Random variable – example

We can assign *probabilities* to the possible values of a random variable.

Example: roll two fair dice, and define the random variable X to be the sum of the two numbers obtained.

Then X can take the values $2, 3, \ldots, 12$ with positive probabilities.

n (values of X)	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = n)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Convention: usually a random variable is denoted by an upper case letter, and the values it can take are denoted by a lower case letter. Confusingly, often the same letter (with different cases) is used.

Random variable – exercises

1. Toss a fair coin 3 times. Let the random variable X denote the number of tails obtained. Find the possible values that X can take, and their probabilities.

2. A weighted coin with $\mathbb{P}(\mathsf{H}) = p$ is tossed repeatedly until a H shows up. Let the random variable Y denote the total number of coin tosses. Find $\mathbb{P}(Y = n)$.

Probability mass function

When the *range* of a random variable X is *discrete*, we define the **probability mass function** (pmf) of X by

$$p(a) := \mathbb{P}(X = a).$$

If $a_1, a_2, \ldots, a_n$ (n may be ∞) are the values that X can take, then $\displaystyle\sum_{i=1}^{n} p(a_i) = 1$.

The **cumulative distribution function** (cdf) is defined as

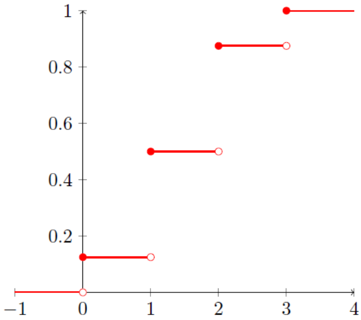
$$F(x) = \sum_{a \leq x} p(a).$$

- F is a non-decreasing function of x, where $x \in \mathbb{R}$.
- $0 \leq F(x) \leq 1$.
- $\mathbb{P}(a < X \leq b) = F(b) - F(a)$.

Example

The cdf for exercise 1 on slide 5 is given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1/8 & \text{if } 0 \le x < 1, \\ 1/2 & \text{if } 1 \le x < 2, \\ 7/8 & \text{if } 2 \le x < 3, \\ 1 & \text{if } x \ge 3. \end{cases}$$



Exercises

3. From a box containing 4 black balls and 2 white balls, 3 balls are drawn without replacement. Find the probability mass function for the number of white balls. Check that the sum is 1.

1/5, 3/5, 1/5

4. If the probability mass function of a random variable is given by $p(i) = c\,3^i/i!$, where $i = 0, 1, 2, 3, \ldots$, what is the value of c?

exp(-3)

5. Let W be the *maximum* of 3 rolls of a fair dice. Find the pmf of W.

Hint: find the cdf first.

Solution

5. For $i = 1, 2, \ldots, 6$, $\mathbb{P}(W \leq i) = \dfrac{i^3}{6^3}$.

The above actually gives the cdf.

Note that it also holds for $i = 0$.

The pmf is then given by

$$p(i) = \mathbb{P}(W \leq i) - \mathbb{P}(W \leq i - 1) = \frac{i^3}{6^3} - \frac{(i-1)^3}{6^3}.$$

40.001: Probability

Week 5: Expected value (Chapter 4)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

Homework

There is a **makeup class** for Section 2, on Friday 13 Oct, 2 – 4pm, in LT3. (This is due to Deepavali on Wednesday 18 Oct.)

- Homework 2 is due at 6pm on Thursday.
- For each homework, 2 questions are marked fully (out of 3 points), the rest are marked based on work shown and correctness of answer (out of 1 point).
- It is a good habit to check your own answers. In many cases, this can be done by doing the question in two different ways.

Exam

- Midterm exam: Friday 3 Nov, 2:30 – 4:30pm, in **CC13**, **CC14** and **Capstone 9 & 10**.

- You are only allowed a two-sided, handwritten A4 resource sheet. Calculators, phones, abacuses etc are **not** allowed.

  (You can compute ${}^nP_k$ and ${}^nC_k$ using the formulas. Values of the exponential function will be supplied.)

- There will be a problem session before the midterm, on 1 Nov, 4:30 – 6:30pm.

Outline

Intuition

The expected value of a random variable is its *average*.

(Toy example: work out the average salary in a company.)

Consider the following special case: events $E_1, E_2, \ldots, E_n$ partition a finite sample space S; let X be a random variable that takes the value x_i on E_i.

If all the basic events are equally likely, then the average of X is given by

$$\frac{x_1|E_1| + x_2|E_2| + \cdots + x_n|E_n|}{|S|} = \sum_{i=1}^{n} x_i \, \mathbb{P}(X = x_i).$$

In general, we use the above sum as the definition of expected value.

Expected value

Definition

The *expected value*, or *expectation*, of a discrete random variable X which takes values $x_1, x_2, \ldots, x_n$ (n may be ∞), is

$$\mathbb{E}(X) := \sum_{i=1}^{n} x_i \, p(x_i).$$

Note: sometimes the expected value may be ∞.

Exercises

1. Roll a fair dice and let X represent the number obtained. What is the expected value of X?
2. Roll 2 fair dice; calculate the expected value of the sum.

Expected value – examples

A lottery game costs \$2 to enter, and consists of picking 6 numbers from $1-45$ without replacement. If the numbers you pick correspond to what the machine picks (order does not matter), then you win \$10 million. Find the expected value of your winnings.

-\$0.77

One interpretation of the above example is that, if you keep playing the game many times, then your average winnings per game would approach the expected value.

In many situations, often a goal is to maximize the expected value (of your winnings, investments, etc), but not always.

E. g. consider this scenario: you are given 2 options: either take away \$950, or play to win \$1 million with winning probability $1/1000$ (you get nothing if you don't win). It would be sensible to pick the 1st option.

Another example

In sports betting, probabilities are artificially inflated so that they add up to *more* than 1.

Consider this example: 3 horses enter a race, and the true *odds* and probabilities of each horse winning are:

A: 1-1 $(\mathbb{P} = 0.500)$
B: 2-1 $(\mathbb{P} = 0.333)$
C: 5-1 $(\mathbb{P} = 0.167)$

The bookmaker can artificially increase everything by 20%, so the new odds and probabilities are:

A: 4-6 $(\mathbb{P} = 0.6)$
B: 6-4 $(\mathbb{P} = 0.4)$
C: 4-1 $(\mathbb{P} = 0.2)$

Suppose now he accepts some bets totaling \$1200, and assume they are in the correct proportions. Whichever horse wins, he will be expected to pay out \$1000, thus making a \$200 profit.

Expected value of a function of an RV

Example 1

Toss a fair coin 3 times; let X denote the number of tails obtained. Find $\mathbb{E}(X)$, $\mathbb{E}(X)^2$, and $\mathbb{E}(X^2)$.

3/2, 9/4, 3

It is not hard to generalize from the above example and deduce:

Theorem

For a function g,

$$\mathbb{E}(g(X)) = \sum_{i=1}^{n} g(x_i)\, p(x_i).$$

- $\mathbb{E}(X^n)$ is called the nth moment of X.
- For constants a and b,

$$\mathbb{E}(a\,X + b) = a\,\mathbb{E}(X) + b.$$

Expected value of a sum of RV's

Example 2

Toss a fair coin 3 times; let X_1 denote the number of tails obtained. Now repeat this experiment, and let X_2 denote the number of tails in the 2nd experiment.

What is the random variable $X_1 + X_2$?

n (values of $X_1 + X_2$)	0	1	2	3	4	5	6
$\mathbb{P}(X_1 + X_2 = n)$	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

Note that in this example, even though X_1 and X_2 have the same pmf, $X_1 + X_2$ and $2\,X_1$ are very different!

However, $\mathbb{E}(X_1 + X_2) = \mathbb{E}(2\,X_1) = 3$.

Expected value of a sum of RV's

Theorem

For random variables X_i, not necessarily independent,

$$\mathbb{E}\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} \mathbb{E}(X_i).$$

Proof: it suffices to prove this for 2 random variables. A different way to write $\mathbb{E}(X)$ is

$$\mathbb{E}(X) = \sum_{s_i \in S} X(s_i)\, p(s_i),$$

where s_i is an outcome in the sample space S, and X is a function. (Compare with slide 6.) Then

$$\begin{aligned}\mathbb{E}(X_1 + X_2) &= \sum_{s_i \in S} \big(X_1(s_i) + X_2(s_i)\big)\, p(s_i)\\ &= \sum_{s_i \in S} X_1(s_i)\, p(s_i) + \sum_{s_i \in S} X_2(s_i)\, p(s_i) = \mathbb{E}(X_1) + \mathbb{E}(X_2).\end{aligned}$$

Problems

1. Re-do Exercise 2 and the $\mathbb{E}(X)$ part of Example 1 using the previous theorem. Generalize.
2. Simplify $\mathbb{E}(X - \mathbb{E}(X))$ and $\mathbb{E}\big((X - \mathbb{E}(X))^2\big)$.
3. (Harder) A experiment involves tossing a weighted coin ($\mathbb{P}(\mathsf{H}) = p$) until a H appears or n tosses are completed. Let Z be the number of coin tosses; find $\mathbb{E}(Z)$.

Solution for 3.

For $i = 1, 2, \ldots, n-1$,

$$\mathbb{P}(Z = i) = (1-p)^{i-1}p.$$

Also, $\mathbb{P}(Z = n) = (1-p)^{n-1}$.

Therefore,

$$\mathbb{E}(Z) = \left[\sum_{i=1}^{n-1} i(1-p)^{i-1}p\right] + n(1-p)^{n-1}.$$

How can we deal with sums of the form

$$\sum_{i=0}^{n-1} i\, r^{i-1}?$$

Trick: we know that

$$\sum_{i=0}^{n-1} r^i = \frac{1-r^n}{1-r},$$

and we can differentiate both sides with respect to r!

Solution, continued

The result is:

$$\sum_{i=0}^{n-1} i\, r^{i-1} = \frac{1-r^n}{(1-r)^2} - \frac{n\, r^{n-1}}{1-r}.$$

Back to finding $\mathbb{E}(Z)$. Let $r = 1-p$ in the above formula, multiply both sides by p, then add the extra term $n(1-p)^{n-1}$.

We finally find that

$$\mathbb{E}(Z) = \frac{1-(1-p)^n}{p}.$$

40.001: Probability

Week 5: Variance, discrete random variables (Chapter 4)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

1. Variance
2. Discrete random variables
   - Bernoulli
   - Binomial
   - Hypergeometric
   - Geometric

Variance

Consider these random variables:

- X_1, which takes the values ± 1 with probability $1/2$ each,
- X_2, which takes the values ± 100 with probability $1/2$ each.

They both have expectation 0, but are clearly very different: X_2 is much more spread out.

For a random variable X, let $\mu = \mathbb{E}(X)$. A measure of the *spread* of X around μ is the variance, defined as

$$\mathrm{Var}(X) := \mathbb{E}\big((X-\mu)^2\big).$$

An equivalent formula is

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Variance – properties

- From the first formula, we see that the variance is *never negative*.
- Using the first formula, we can show that, for constants a, b,

$$\mathsf{Var}(a\,X + b) = a^2\,\mathsf{Var}(X).$$

- The positive square root of Var is the *standard deviation*:

$$\mathsf{SD}(X) = \sqrt{\mathsf{Var}(X)}.$$

Exercises

1. Compute $\mathsf{Var}(X_1)$ and $\mathsf{Var}(X_2)$ for the X_i defined on the previous slide.
2. Find $\mathsf{Var}(X)$ if X represents the outcome of a fair dice roll.

2: 35/12

Outline

1. Variance

2. Discrete random variables
   - Bernoulli
   - Binomial
   - Hypergeometric
   - Geometric

Bernoulli random variable

X is called a *Bernoulli* random variable if it can only take two values, 1 (success) or 0 (failure), where

$$\mathbb{P}(X = 1) = p, \qquad \mathbb{P}(X = 0) = 1 - p.$$

We denote this by $X \sim \text{Bernoulli}(p)$.

It is easy to check that for a Bernoulli random variable X,

$$\mathbb{E}(X) = p, \qquad \text{Var}(X) = p\,(1 - p).$$

Binomial random variable

Suppose X is the number of successes in n independent Bernoulli trials, each with probability of success p. Then X is also the *sum* of n independent Bernoulli random variables, each with parameter p.

X is called a *binomial* random variable, and we denote this by $X \sim \text{binomial}(n, p)$.

We have already seen that

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

The expectation and variance are given by

$$\mathbb{E}(X) = n\,p, \qquad \text{Var}(X) = n\,p\,(1-p).$$

Binomial – $\mathbb{E}$

To prove that $\mathbb{E}(X) = n\,p$, we can use the observation that a binomial random variable is the sum of Bernoulli random variables.

Alternatively, we can use the *differentiation* trick. Start with the binomial theorem,

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n.$$

Differentiate both sides with respect to x, then multiply by x:

$$\sum_{k=0}^{n} k \binom{n}{k} x^k y^{n-k} = n\,x\,(x+y)^{n-1}.$$

Now, letting $x = p$, $y = 1 - p$ gives $\mathbb{E}(X)$. Performing the trick again, we can find $\mathbb{E}(X^2)$ and hence $\mathsf{Var}(X)$.

Binomial – Var

Another way to find Var(X) is via the representation $X = X_1 + X_2 + \cdots + X_n$, where each $X_i \sim$ Bernoulli(p).

Then

$$\begin{aligned}\mathbb{E}(X^2) &= \mathbb{E}\big(X_1^2 + X_2^2 + \cdots + X_n^2 + n(n-1) \text{ terms of the form } X_iX_j\big) \\ &= n\,\mathbb{E}(X_i^2) + n(n-1)\,\mathbb{E}(X_iX_j), \quad \text{where } i \neq j.\end{aligned}$$

This leads to $\mathbb{E}(X^2) = np + n(n-1)p^2$, therefore

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = np + n(n-1)p^2 - (np)^2 = n\,p\,(1-p).$$

Binomial – shape

As k goes from 0 to n, the pmf of a binomial random variable first increases, then decreases, with a peak around $\mathbb{E}(X)$.

(This can be proven by analyzing $p(k)/p(k-1)$.)

A visual demonstration can be found here.

Exercise

In a game that costs \$1 to play, three fair dice are rolled. A person wins $\$k$ if she bets on a number and the number appears k times. Find the expectation and variance of her winnings.

Y ˜ binomial(3, 1/6)
E(Y-1) = -1/2 , Var(Y-1) = 5/12

Hypergeometric random variable

A committee of k people is chosen randomly from n men and m women. Let X be the number of men. Then (as we have seen) X is a *hypergeometric* random variable, with

$$\mathbb{P}(X = r) = \frac{\binom{n}{r}\binom{m}{k-r}}{\binom{n+m}{k}}.$$

By writing X as an appropriate sum of random variables, it can be shown that $\mathbb{E}(X) = \frac{k\,n}{n+m}$.

Alternatively, $\mathbb{E}(X)$ may be computed by manipulating binomial coefficients:

$$\mathbb{E}(X) = \sum_r \frac{r\binom{n}{r}\binom{m}{k-r}}{\binom{n+m}{k}} = \sum_r \frac{n\binom{n-1}{r-1}\binom{m}{k-r}}{\binom{n+m}{k}}$$

$$= \sum_r \frac{n\binom{n+m-1}{k-1}}{\binom{n+m}{k}} \frac{\binom{n-1}{r-1}\binom{m}{k-r}}{\binom{n+m-1}{k-1}} = \frac{n\binom{n+m-1}{k-1}}{\binom{n+m}{k}} \cdot 1 = \frac{k\,n}{n+m}.$$

Hypergeometric – shape

More work is required to show that

$$\mathrm{Var}(X) = k\frac{n}{n+m}\left(1 - \frac{n}{n+m}\right)\frac{n+m-k}{n+m-1}.$$

If n and m are *large* in relation to k, then selection without replacement can be approximated by selection with replacement.

Then the pmf of the hypergeometric random variable can be approximated by a binomial$\left(k, \frac{n}{n+m}\right)$ random variable.

Indeed, with $p = \frac{n}{n+m}$, we see that $\mathrm{Var}(X) \approx k\, p\, (1-p)$.

Geometric random variable

For independent Bernoulli trials with probability of success p, let X be the number of trials required for the first success to occur.

Then X is a *geometric* random variable, denoted by $X \sim \text{geometric}(p)$, with

$$\mathbb{P}(X = n) = (1-p)^{n-1}\, p.$$

Using the differentiation trick, we can show that

$$\mathbb{E}(X) = \frac{1}{p}, \qquad \mathrm{Var}(X) = \frac{1-p}{p^2}.$$

40.001: Probability

Week 6: Poisson random variable (Chapter 4)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Outline

Motivation

Suppose that a book has 500 pages, and the probability that a page is misprinted is 0.05.

The probability that the book has, say, 30 misprinted pages, is

$$\binom{500}{30} 0.05^{30}\, 0.95^{470}.$$

Can we find a way to easily estimate this number?

An approximation

Consider a binomial random variable X with *large* n, *small* p, and let $\lambda := np$.

$$\begin{aligned}\mathbb{P}(X=k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!}\,\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k}\,\frac{\lambda^k}{k!}\,\frac{(1-\lambda/n)^n}{(1-\lambda/n)^k}.\end{aligned}$$

As $n \to \infty$, the two expressions in blue both $\to 1$, while the expression in red $\to e^{-\lambda}$. Thus

$$\mathbb{P}(X=k) \approx e^{-\lambda}\,\frac{\lambda^k}{k!}.$$

Poisson random variable

A random variable X is said to satisfy a **Poisson** distribution with parameter $\lambda > 0$, if its pmf is given by

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

We denote this by $X \sim \text{Poisson}(\lambda)$.

The Poisson distribution is a good approximation to the binomial distribution (when n is large and p is small). As such, it can be used to model rare events, for example:

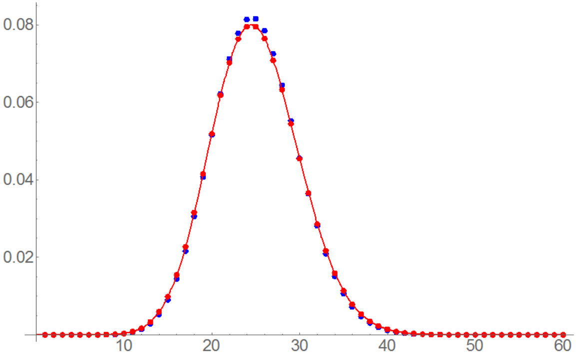
- The number of misprinted pages in a book.
- The number of arrivals in a grocery store.
- The number of phone calls you get today.
- The number of supernovas in a galaxy.

Poisson – example

We wanted to estimate the probability that a book has 30 misprinted pages.

Here $n = 500$, $p = 0.05$, so $\lambda = 25$.

$$\mathbb{P}(30 \text{ misprinted pages}) \approx e^{-25}\,\frac{25^{30}}{30!} \approx 0.045.$$



Poisson – exercises

1. If $X \sim \text{Poisson}(\lambda)$, then show that

$$\mathbb{E}(X) = \lambda = \text{Var}(X).$$

2. If the average number of typos on each slide is 1.5, find the probability that

   (a) There is at least one typo on one slide,

   (b) There is at least one typo, on two slides (assume independence.)

3. The probability that a base-pair in a strand of DNA undergoes mutation is 10^{-5}, independent of any other base-pairs. In a strand consisting of 10^6 base-pairs, what is the probability that 8 to 12 mutations occur?

Solutions

1. Use $\frac{k}{k!} = \frac{1}{(k-1)!}$, then *shift* the summation index.

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{\infty} k\, e^{-\lambda}\, \frac{\lambda^k}{k!} = \sum_{k=1}^{\infty} k\, e^{-\lambda}\, \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} e^{-\lambda}\, \frac{\lambda^{k-1}\lambda}{(k-1)!} \qquad \text{now let } j = k-1 \\
&= \sum_{j=0}^{\infty} \lambda\, e^{-\lambda}\, \frac{\lambda^j}{j!} = \lambda \cdot 1 = \lambda.
\end{aligned}$$

2. $\lambda = 1.5$. (a) $1 - e^{-1.5} \approx 0.78$. (b) $1 - e^{-1.5\times 2} \approx 0.95$.

3. This is tricky to compute with the binomial distribution, so we approximate with Poisson.
$\lambda = 10^6 \cdot 10^{-5} = 10$, so $X \sim \text{Poisson}(10)$, and

$$\mathbb{P}(8 \le X \le 12) \approx \sum_{k=8}^{12} e^{-10}\, \frac{10^k}{k!} \approx 0.57.$$

Poisson – harder example

Suppose on average, you can see one shooting star every 20 minutes. If you watch the sky for one hour each evening, on average how many days would it take till you see 6 or more shooting stars in an evening?

Let N = number of shooting stars seen in an evening, and X = number of days till you see 6 or more shooting stars in an evening.

Then $N \sim$ Poisson(3), and $\mathbb{P}(N \geq 6) \approx 0.0839 = p$.

Now X is a geometric random variable with probability p.
$\mathbb{E}(X) = \dfrac{1}{p} \approx 11.9$.

Birthday problem revisited

In the birthday problem, the probability that any two people share a birthday is $p = \dfrac{1}{365}$. The number of trials (comparisons) is $\dbinom{n}{2}$.

The trials are not independent, but it can be shown that the dependence is *weak* (e. g. they are pairwise independent), and the Poisson approximation can be used.

So $\lambda = \dbinom{n}{2} p = n(n-1)/730$, and

$$\mathbb{P}(\text{no 2 people have the same birthday}) \approx e^{-\lambda}\,\frac{\lambda^0}{0!} = \exp\left(\frac{-n(n-1)}{730}\right).$$

This agrees with the approximation we found in Week 2.

40.001: Probability

Week 6: Poisson process (Chapter 4)

ESD, SUTD

Term 4, 2017

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Information

- A seating plan for the midterm exam, and a reminder for the midterm survey, have been uploaded on eDimension.
- Please complete the **survey** between 30 Oct and 2 Nov.
- Homework 3 is out; due 6pm, Tuesday 31 Oct.
- Homework 2 solutions are available. Carefully read the last paragraphs of the **Q2 and Q3 solutions** and learn from them!
- To come: consolidated slides, extra notes.
- To come: past exam; review session on 1 Nov, 4:30–6:30pm, LT5.

Outline

Poisson process

The Poisson distribution also models the number of occurrences in a fixed time interval, when the *rate* per unit time interval is λ.

Notation: let $o(h)$ stand for any function $f(h)$ for which $\lim_{h \to 0} f(h)/h = 0$.

Let $N(t)$ be the number of events occurring in the time interval $[0, t]$. $N(t)$ is called a **Poisson process** with *rate* λ if:

1. The probability that exactly 1 event occurs in a given interval of length h is $\lambda h + o(h)$.
2. The probability that more than 1 event occurs in a given interval of length h is $o(h)$.
3. Independence: whatever occurs in one interval has no probabilistic effect on what occurs in other, non-overlapping intervals.

Poisson process – proof

We argue that the number of events occurring in an interval of length t is a *Poisson* random variable with parameter λt.

We first divide the interval $[0, t]$ into n equal sub-intervals. Then

$$\mathbb{P}(N(t) = k) = \mathbb{P}(A_n) + \mathbb{P}(B_n),$$

where A_n is the event that k of the sub-intervals contain exactly 1 event each, and the other $(n - k)$ sub-intervals contain 0 events.

By condition 2, it can be shown that $\lim_{n\to\infty} \mathbb{P}(B_n) = 0$.

By conditions 1 and 2, the probability that 0 events occur in a sub-interval of length t/n is $1 - \lambda t/n - o(t/n)$.

Poisson process – exercises

Then by condition 3,

$$\mathbb{P}(A_n) = \binom{n}{k}\left[\frac{\lambda t}{n} + o\left(\frac{t}{n}\right)\right]^k \left[1 - \frac{\lambda t}{n} - o\left(\frac{t}{n}\right)\right]^{n-k}.$$

As $n \to \infty$, we invoke the Poisson approximation to the binomial to deduce that

$$\mathbb{P}(N(t) = k) \to e^{-\lambda t}\frac{(\lambda t)^k}{k!}.$$

Therefore, Poisson random variables can be used to model the number of (rare) events during some fixed time span, such as the number of earthquakes, wars, machine failures, etc per year.

Exercise

Suppose earthquakes are a Poisson process with a rate of 2 per week.

(a) Find $\mathbb{P}$(at least 3 earthquakes occur during the next 2 weeks).

(b) Let T be the time in weeks until the next earthquake. Find $\mathbb{P}(T \leq t)$.

Outline

For fun

At a prom, attended by equal numbers of boys and girls, it is reported that on average, each boy danced with 4.6 girls, while each girl danced with 3.5 boys. (Each dance involved 1 boy and 1 girl.)

Are the reported figures possible?

A visualization of some of the concepts we have learned so far can be found here.

Connection with Statistics

The formulas we have learned for the expected value and variance apply to *random variables*.

You may have seen, and will see a different (though consistent) version of these formulas for *data values*, e. g.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2.$$

A card trick

See a card trick here.

- Secretly pick an integer k between 1 and 10.
- Count to the kth card.
- The value of this card determines the next card you count to; royal cards count as 5.
- Repeat until you cannot go on anymore; remember your final card.

A heuristic explanation is given in the link.