Week 1: Bayes’ theorem and independence

Recap

- Probability space
  - Sample space *S* - set of outcomes
  - Events - subsets of sample space (technical condition)
  - Probability P - function from events to the closed interval [0,1]
  - Axioms
    - $P(S) = 1$
    - $E_1, E_2, \dots$. disjoint: $P(E_1 \cup E_2 \cup \dots) = P(E_1) + P(E_2) + \dots$
  - Properties
    - $P(\Phi)=0$, $P(E^c) = 1 - P(E)$, $P(E) = P(E \cap F) + P(E \setminus F)$, $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
- Conditional probability - given event B has occurred
  - $P(A \mid B) = P(A \cap B) / P(B)$ or $P(A \cap B) = P(B)\, P(A \mid B)$
- Next set of lectures
  - More with conditional probability - Law of total probability and Bayes' theorem
  - Independence and repeated trials of an experiment

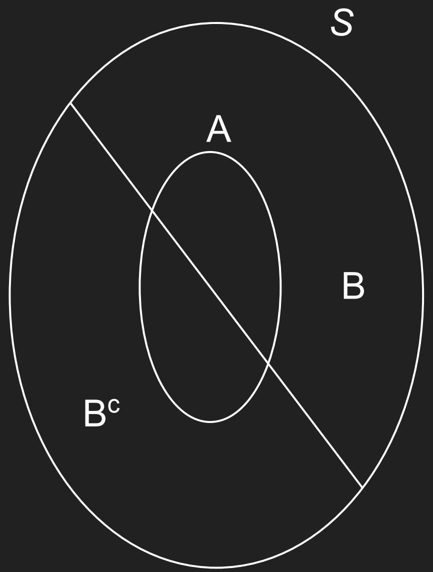
Law of total probability

Law of total probability

$$\boxed{P(A) = P(A \cap B) + P(A \cap B^c) = P(A \mid B)\, P(B) + P(A \mid B^c)\, P(B^c)}$$

Proof

- A: disjoint union of $A \cap B$ and $A \cap B^c$
- By Axiom 2, $P(A) = P(A \cap B) + P(A \cap B^c)$
- Using conditional probability on each term above, we get the result



Example: Two urns with coloured marbles

Pick an urn at random and then pick a marble at random from the chosen urn.

$P(\text{red} \mid \text{urn } 1) = 7/13$, $P(\text{red} \mid \text{urn } 2) = 5/13$

$P(\text{blue} \mid \text{urn } 1) = 6/13$, $P(\text{blue} \mid \text{urn } 2) = 8/13$

$$
\begin{aligned}
P(\text{red}) &= P(\text{red} \mid \text{urn } 1)\, P(\text{urn } 1) + P(\text{red} \mid \text{urn } 2)\, P(\text{urn } 2) \\
&= (7/13)(1/2) + (5/13)(1/2) = 6/13
\end{aligned}
$$

Example: Economic model

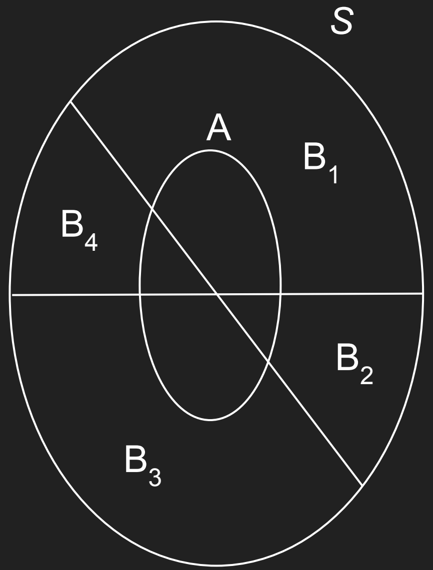
An economic model predicts that if interest rates rise, then there is a 60% chance that unemployment will increase, but that if interest rates do not rise, then there is only a 30% chance that unemployment will increase. If the economist believes there is a 40% chance that interest rates will rise, what should she calculate is the probability that unemployment will increase?

More generally….

$B_1, B_2, B_3, \ldots.$: Partition of S

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots = P(A \mid B_1)\, P(B_1) + P(A \mid B_2)\, P(B_2) + \ldots$$

Same proof as before



Example: Tricks with coins

A man possesses 5 coins - 2 are double-headed, 1 is double-tailed, 2 are normal. He picks a coin at random and tosses it. What is the probability that he sees a head?

Bayes’ theorem

Example: Two urns with coloured marbles

Pick an urn at random and then pick a marble at random from the chosen urn.

P(urn 1) = P(urn 2) = 1/2

P(red | urn 1) = 7/13, P(red | urn 2) = 5/13

P(blue | urn 1) = 6/13, P(blue | urn 2) = 8/13

What about P(urn 1 | red) or P(urn 1 | blue)?

Example: Flu test

1% of people in a city have Swine Flu.

Flu test: 95% of people with Swine Flu test positive
2% of people without the disease will test positive

A person is randomly chosen from the city and tests positive. What is the probability that the person actually has Swine Flu?

$P(\text{Swine Flu}) = 0.01$

$P(\text{positive} \mid \text{Swine Flu}) = 0.95$ and $P(\text{positive} \mid \text{Swine Flu}^c) = 0.02$

What is $P(\text{Swine Flu} \mid \text{positive})$?

Problem

Two events: A and B

P(A) and P(B) are known.

P(A | B) and P(A | B^c) are either known or easy to find.

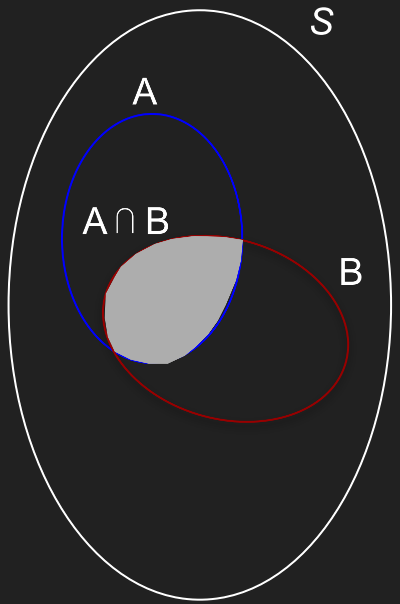
Question: Find P(B | A) and P(B^c | A).

Bayes' theorem

A, B: events with $P(A) > 0$, $P(B) > 0$

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

$$P(B \mid A) = P(B)\,P(A \mid B) / P(A)$$



Bayes' theorem along with law of total probability can be used to solve the problems described in the previous slides.

Example: Two urns with coloured marbles

Pick an urn at random and then pick a marble at random from the chosen urn.

Event B: urn 1 is chosen, and Event A: red marble is picked

$P(B) = 1/2$, $P(B^c) = 1/2$

$P(A \mid B) = 7/13$, $P(A \mid B^c) = 5/13$. By law of total probability, $P(A) = 6/13$

$P(\text{urn } 1 \mid \text{red}) = P(B \mid A) = P(B)\, P(A \mid B) / P(A) = (1/2)(7/13) / (6/13) = 7/12$

Example: Flu test

1% of people in a city have Swine Flu.

Flu test: 95% of people with Swine Flu test positive
2% of people without the disease will test positive

A person is randomly chosen from the city and tests positive. What is the probability that the person actually has Swine Flu?

B: Person has Swine Flu, A: Person tests positive

$P(B) = 0.01$, $P(B^c) = 0.99$

$P(A \mid B) = 0.95$ and $P(A \mid B^c) = 0.02 \Rightarrow P(A) = 0.95 \times 0.01 + 0.02 \times 0.99 = 0.0293$

Bayes' theorem: $P(B \mid A) = 0.01 \times 0.95 / 0.0293 = 0.3242\ldots = 32.42\%$ (surprise!)

Example: Multiple choice question (MCQ)

A student attempting an MCQ with 4 choices (of which one is correct) knows the correct answer with probability 3/4. If she does not know, she guesses a random choice. Given that a question was answered correctly, what is the conditional probability that she knew the answer?

Example: Roll a die and toss coins

You first roll a fair die, then toss as many fair coins as the number that showed on the die. Given that 5 heads are obtained, what is the probability that the die showed 5?

Independence

Independence of two events

DEFINITION (**Independence**) Two events A and B are independent if

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

- Motivation
  - If $P(B) > 0$, $P(A \mid B) = P(A)$
  - Probability of A in original space is equal to the conditional probability of A given B in the conditional space
  - Probability of A is unaffected by occurrence of B, i.e. A and B are independent
- Rules for multiplication and addition
  - When Event 1 and Event 2 are independent, P(Event 1 and Event 2) = P(Event 1) P(Event 2)
    - Multiply probabilities to find the probability of “and” of independent events
  - Contrast with disjoint or mutually exclusive events A and B: $P(A \cup B) = P(A) + P(B)$
    - Add probabilities to find probability of “or” of disjoint events

Example: Toss a coin thrice

Uniform{ HHH, HHT, HTH, HTT, THH, THT, TTH, TTT }

A: First toss is heads, B: Second toss is heads

A = { HHH, HHT, HTH, HTT }, B = { HHH, HHT, THH, THT }

$A \cap B$ = { HHH, HHT }

$P(A \cap B) = 1/4 = 1/2 \times 1/2 = P(A)\ P(B)$

A and B are independent

What about A and B^c? What about 1st toss if heads and 3rd toss is heads?

Example: Throw a die

A: even, B: odd

$A = \{ 2, 4, 6\}$, $B = \{ 1, 3, 5 \}$

$A \cap B = \Phi$

$P(A \cap B) = 0 \neq P(A)\, P(B)$

Point to remember:

- Disjoint events are never independent
- Why? If A and B are disjoint and B occurs, A definitely did not occur
    - Occurrence of B definitely impacts conditional probability of A!
- For events to be independent, they should have a non-empty intersection

Example: Throw a die

A: even number, B: multiple of 3

$A = \{ 2, 4, 6\}, B = \{ 3, 6 \}$

$A \cap B = \{ 6 \}$

$P(A \cap B) = 1/6 = 1/2 \times 1/3 = P(A) P(B)$

A and B are independent

Typical points of confusion:

- A and B have an intersection. How can they be independent?
- There is only one throw of the die. How can two events be independent when there is only one throw?

Example: Card from a pack

A: Card is an spade, B: Card is a king

$A \cap B$: Card is spade-king

$P(A \cap B) = 1/52 = 1/4 \times 1/13 = P(A)\, P(B)$

A and B are independent

Typical points of confusion:

- Isn't there a spade-king card? How can card being a spade be independent of card being a king when there is a spade king card?
- There is only one card drawn. How can two events be defined and be independent when there is only one card?

Mutual independence of three events

DEFINITION (**Mutual Independence**) Events A, B, C are mutually independent if

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

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

- Two constraints
  - A and B are independent, A and C are independent, B and C are independent
  - Additional constraint on $A \cap B \cap C$ - this is important
- Example: Toss a coin twice
  - $A = \{ HH, TT \}$, $B = \{ HH, HT \}$, $C = \{ HH, TH \}$
  - A and B, A and C, B and C: independent pairwise
  - However, $P(A \cap B \cap C) = 1/4 \neq P(A)\,P(B)\,P(C)$
  - Pairwise independent but not mutually independent

Mutual independence of multiple events

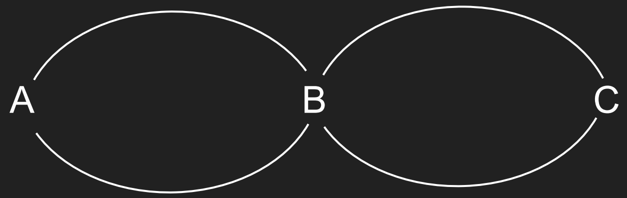
DEFINITION (**Mutual Independence**) Events $A_1, A_2, \ldots, A_n$ are mutually independent if, for all i1, i2, …, ik,

$$P(A_{i1} \cap A_{i2} \cap \ldots \cap A_{ik}) = P(A_{i1})\, P(A_{i2}) \ldots P(A_{ik})$$

- Almost 2^n constraints
  - Prob{Intersection of any subset of events} = Product of the Prob{events}
- Interesting result: A and B are independent => A and B^c are independent
  - $P(A \cap B^c) = P(A \setminus B) = P(A) - P(A \cap B) = P(A)(1 - P(B)) = P(A)\, P(B^c)$
  - Intuitive: B does not affect A => B^c does not affect A
  - Two events are independent => Complement of one event and other event are independent
- Using the above twice: A and B are independent => A^c and B^c are independent
- Extension: n events are mutually independent => any subset with or without complementing are independent as well

Example: Roads

Two roads each connect A and B and B and C. Each of the four roads gets blocked with probability p independent of all other roads. What is the probability that there is an open route from A to B given that there is no open route from A to C?



Repeated independent trials - Bernoulli, Binomial and Geometric Distribution

Bernoulli trials

Example: Incidence of a disease

- Suppose you want to find out how many people in a city suffer from a disease.
- Trial: Select a random person from a population and test for a disease.
- Repeat trial n times. Assume each trial is independent.

Observation: number of times the test was positive

- Questions: What is the distribution of the above number?
- More interesting question: What fraction of people have the disease?

A single Bernoulli trial

Setting: Occurrence of Event A in a sample space is considered “success”. Non-occurrence of A is considered “failure”. Let $p = P(A)$.

Bernoulli trial: Sample space is {success, failure} with $P(\text{success}) = p$

or

{0, 1} with $P(1) = p$, $P(0) = 1 - p$

This distribution is denoted Bernoulli(p).

In previous example, a Bernoulli trial is a framework for one test.

Repeated Bernoulli trials

Setting: Repeat a Bernoulli trial multiple times independently.

Perform n independent Bernoulli(p) trials.

Outcome: 0 or 1 (Trial 1), 0 or 1 (Trial 2), 0 or 1 (Trial 3), … … , 0 or 1 (Trial n)

Sample space: 2^n outcomes

Eg: n = 3, *S* = { 000, 001, 010, 011, 100, 101, 110, 111 }

n = 4, *S* = { 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111 }

Probabilities: Use independence.
Let n = 3.
P(000) = P(trial 1 is 0 and trial 2 is 0 and trial 3 is 0) = (1-p) x (1-p) x (1-p) = $(1-p)^3$
P(101) = P(trial 1 is 1 and trial 2 is 0 and trial 3 is 1) = p x (1-p) x p = $p^2(1-p)$

Example: Toss a fair coin 5 times

Toss a fair coin 5 times. H is 0 and T is 1. Sample space has 32 equally likely outcomes.

For example, P(HHHHH) = 1/32, P(HTHTH) = 1/32, P(THTHT) = 1/32 etc.

P(0 tails) = P(HHHHH) = 1/32

P(1 tail) = P({THHHH, HTHHH, HHTHH, HHHTH, HHHHT}) = 5/32

P(2 tails) = P({TTHHH, THTHH, THHTH, THHHT, HTTHH,
HTHTH,HTHHT, HHTTH, HHTHT, HHHTT}) = 10/32

P(3 tails) = P({TTTHH, TTHTH, TTHHT, THTTH, THTHT,
THHTT, HTTTH, HTTHT, HTHTT, HHTTT}) = 10/32

P(4 tails) = P({TTTTH, TTTHT, TTHTT, THTTT, HTTTT}) = 5/32

P(5 tails) = P(TTTTT) = 1/32

Example: Toss a coin 5 times (using uniform distribution)

Toss a fair coin 5 times. H is 0 and T is 1. Sample space has 32 outcomes.

P(0 tails) = $5C_0$ favourable out of 32 = 1/32 = P(5 heads)
P(1 tail) = $5C_1$ favourable out of 32 = 5/32 = P(4 heads)
P(2 tails) = $5C_2$ favourable out of 32 = 10/32 = P(3 heads)
P(3 tails) = $5C_3$ favourable out of 32 = 10/32 = P(2 heads)
P(4 tails) = $5C_4$ favourable out of 32 = 5/32 = P(1 head)
P(5 tails) = $5C_5$ favourable out of 32 = 1/32 = P(0 heads)

P(at least 4 tails) = P(4 tails or 5 tails) = 5/32 + 1/32 = 6/32
P(at most 3 heads) = P(0 or 1 or 2 or 3 heads) = 1/32+5/32+10/32+10/32 = 26/32

Computing probabilities for Bernoulli trials

Perform n independent Bernoulli(p) trials.

Ex: $n = 10$
$P(0110001011) = (1-p) \times p \times p \times (1-p) \times (1-p) \times (1-p) \times p \times (1-p) \times p \times p = p^5 (1-p)^5$

In general,

$$P(b_1 b_2 \ldots b_n) = p^w (1-p)^{n-w}, \text{ where } w = \text{no of 1s in } b_1 b_2 \ldots b_n$$

- The actual sequence of 0s and 1s does not matter. Only the number of successes matters in the probability computation.

Ex: $P(b_1 b_2 \ldots b_{200} \text{ with 36 1s}) = p^{36} (1-p)^{200-36} = p^{36} (1-p)^{164}$

Example: Toss a biased coin 5 times

H is 0 and T is 1. $P(H) = 1/3$, $P(T) = 2/3$.

$P(HHHHH) = (1/3)^5$, $P(TTTTT) = (2/3)^5$

$P(1 \text{ tail}) = P(\{THHHH, HTHHH, HHTHH, HHHTH, HHHHT\}) = 5(2/3)(1/3)^4$

$P(2 \text{ tails}) = P(\{TTHHH, THTHH, THHTH, THHHT, HTTHH,$
$HTHTH, HTHHT, HHTTH, HHTHT, HHHTT\}) = 5(2/3)^2(1/3)^3$

$P(3 \text{ tails}) = P(\{TTTHH, TTHTH, TTHHT, THTTH, THTHT,$
$THHTT, HTTTH, HTTHT, HTHTT, HHTTT\}) = 5(2/3)^3(1/3)^2$

$P(4 \text{ tails}) = P(\{TTTTH, TTTHT, TTHTT, THTTT, HTTTT\}) = 5(2/3)^4(1/3)$

Revisit Example: Incidence of a disease

- Suppose a fraction of the people in a city have a disease.

- Trial: Select a random person from a population and test for a disease.
- We will assume that the probability that a person tests positive is ‘p’.
  - Can ‘p’ be assumed to be the fraction having the disease? When is this a good assumption?

- Repeat trial n times. Assume each trial is independent.

Observation: number of times the test was positive

- Questions: What is the distribution of the above number in terms of n and p?

Binomial distribution

Binomial(n, p): Binomial distribution

Perform n independent Bernoulli(p) trials.

Outcome: number of successes, which we will denote B(n,p) or B in short

Sample space: { 0, 1, 2, …, n }

Ex: n = 3

$P(B = 0) = P(\text{trials result in } 000) = (1 - p)^3$

$P(B = 1) = P(\text{trials result in } 001 \text{ or } 010 \text{ or } 100) = p(1 - p)^2+p(1 - p)^2+p(1 - p)^2 = 3p(1-p)^2$

$P(B = 2) = P(\text{trials result in } 011 \text{ or } 101 \text{ or } 110) = p^2(1 - p)+p^2(1 - p)+p^2(1 - p) = 3p^2(1-p)$

$P(B = 3) = P(\text{trials result in } 011) = p^3$

Binomial(5, p)

n = 5. B = no of successes in n Bernoulli(p) trials

P(B = 0) = P(trials result in 00000) = $(1 - p)^5$

P(B = 1) = P(trials result in 00001 or 00010 or 00100 or 01000 or 10000)
= $p(1 - p)^4 + p(1 - p)^4 + p(1 - p)^4 + p(1 - p)^4 + p(1 - p)^4 = 5p(1-p)^2$

P(B = 2) = P(trials result in 00011 or 00101 or 01001 or 10001 or 00110
01010 or 10010 or 01100 or 10100 or 11000)
= $p^2(1 - p)^3 + p^2(1 - p)^3 + p^2(1 - p)^3 + p^2(1 - p)^3 + p^2(1 - p)^3 + p^2(1 - p)^3 +p^2(1 - p)^3 +$
$p^2(1 - p)^3 + p^2(1 - p)^3 + p^2(1 - p)^3 = 10p^2(1-p)^3$

P(B = 3) = P(trials result in 3 1s)=(no of favourable outcomes)$p^3(1-p)^2=5C_3p^3(1-p)^2=10p^3(1-p)^2$

P(B = 4) = P(trials result in 4 1s)=(no of favourable outcomes)$p^4(1-p) = 5C_4p^4(1-p) = 5p^4(1-p)$

P(B = 5) = P(trials result in 5 1s) = p^5

Binomial(n,p): small n

Probability of number of successes

	B = 0	B = 1	B = 2	B = 3	B = 4	B = 5	B = 6
n = 1	$1 - p$	p					
n = 2	$(1 - p)^2$	$2p(1 - p)$	p^2				
n = 3	$(1 - p)^3$	$3p(1 - p)^2$	$3p^2(1 - p)$	p^3			
n = 4	$(1 - p)^4$	$4p(1 - p)^3$	$6p^2(1 - p)^2$	$4p^3(1 - p)$	p^4		
n = 5	$(1 - p)^5$	$5p(1 - p)^4$	$10p^2(1 - p)^3$	$10p^3(1 - p)^2$	$5p^4(1 - p)$	p^5	
n = 6	$(1 - p)^6$	$6p(1 - p)^5$	$15p^2(1 - p)^4$	$20p^3(1 - p)^3$	$15p^4(1 - p)^2$	$6p^5(1 - p)$	p^6

Binomial(n, p): Expression for probability

In general, what is $P(B(n,p) = k)$? ($k = 0, 1, 2, \ldots, n$)

$P(B(n,p) = k) = P(\text{trial results in } b_1 b_2 \ldots . b_n \text{ with exactly } k \text{ 1s})$

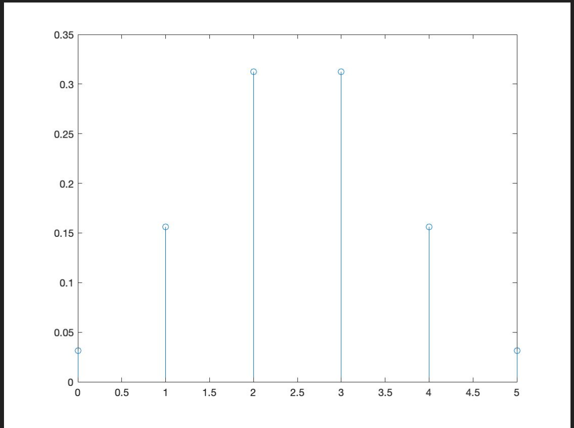
$= (\text{no of } b_1 b_2 \ldots . b_n \text{ with exactly } k \text{ 1s})\ p^k (1 - p)^{n - k}$

No of $b_1 b_2 \ldots . b_n$ with exactly k 1s $= n\,C_k = n! / (k!\,(n-k)!)$

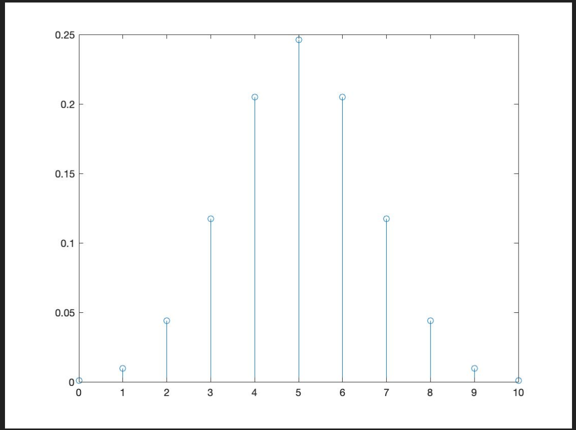
$$P(B(n,p) = k) = n\,C_k\, p^k (1 - p)^{n - k}$$

Binomial(n, 0.5)

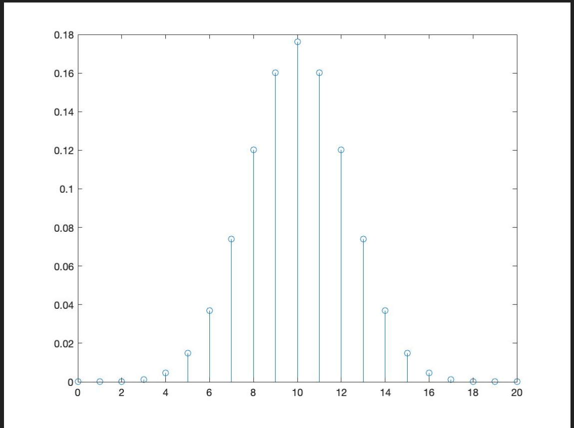
n = 5



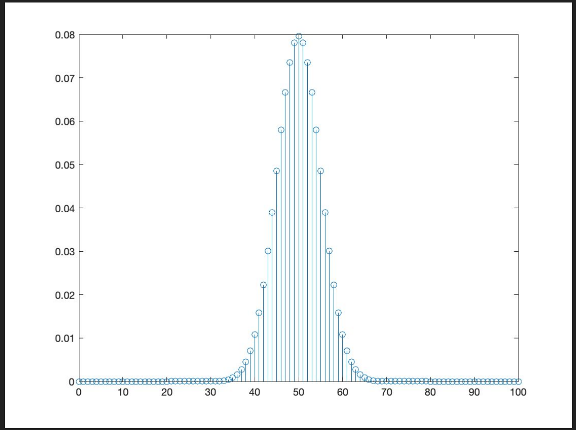
n = 10



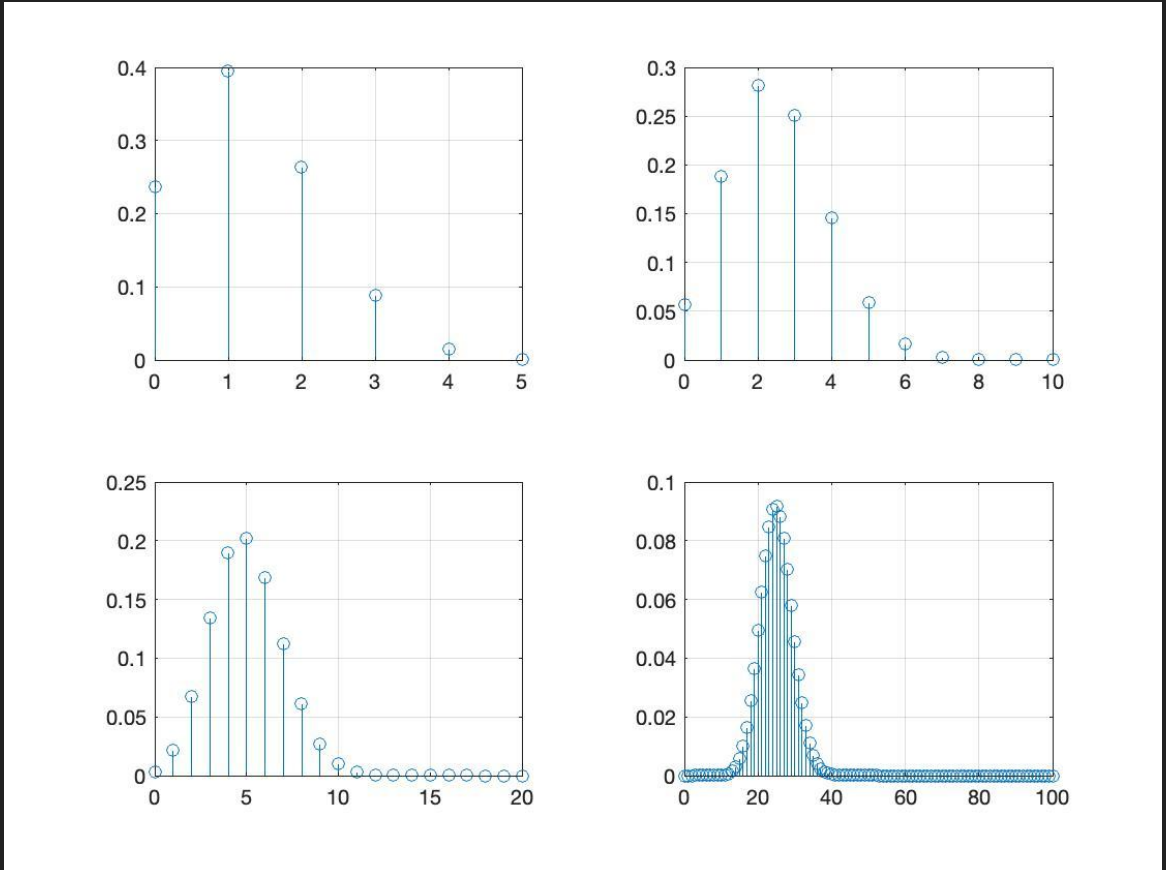
n = 20



n = 100



p = 0.25, n = 5, 10, 20, 100



Some observations

- Starts at $(1-p)^n$ -> increases and reaches a peak -> falls to p^n

- Where is the peak? Near np
  - Exactly, floor(p(n+1)), i.e. largest integer <= p(n+1)

- P(B = 0 or B = 1 or B = 2 or …. or B = n) = 1

  P(B = 0) + P(B = 1) + P(B = 2) + …. + P(B = n) = 1

  $(1-p)^n + nC_1\, p(1-p)^{n-1} + nC_2\, p^2(1-p)^{n-2} + \ldots + p^n = 1$ (important identity)

  Check the above identity for n = 1, 2, 3, 4….

Example: Incidence of a disease

Each person has a disease with probability 0.1 independently. Out of 100 random persons tested for the disease, what is the probability that 20 persons test positive? Assume that the disease can be tested accurately with no false positives.

Think about this one: Such questions will come later!

Each person has a disease with probability p independently. Out of 100 random persons tested for the disease, suppose that 20 persons test positive. What is p? Assume that the disease can be tested accurately with no false positives.

Example: Coin toss

Suppose a fair coin is tossed 10 times.

(a) What is the probability that the number of heads is a multiple of 3?

(b) What is the probability that the number of heads is even?

Example: Communicating bits

A bit (0 or 1) sent by Alice to Bob gets flipped with probability 0.1.
(a) If 5 bits are sent by Alice independently, what is the probability that at most 2 bits get flipped?
(b) If 10 bits are sent by Alice independently, what is the probability that at most 2 bits get flipped?

Geometric distribution

Example: Tossing coins till you get a head

Suppose you toss a fair coin repeatedly (independently) till you get a head. How many times will you toss the coin?

P(1) = P(heads in first toss) = 1/2

P(2) = P(tails in first toss and heads in second toss) = 1/2 x 1/2 = 1/4

P(3) = P(1T and 2T and 3H) = 1/2 x 1/2 x 1/2 = 1/8

…..

P(k) = P(1T and 2T and …. and (k-1)T and kH) = $(1/2)^k$

….

Example: Throw a die till you get 1

In Ludo, a player needs to repeatedly throw a die till she gets a 1. How many throws are needed?

$P(1) = 1/6$

$P(2) = P(\text{not 1 in first throw and 1 in second throw}) = 5/6 \times 1/6 = 5/36$

$P(3) = P(\text{not 1 and not 1 and 1}) = 5/6 \times 5/6 \times 1/6 = 25/216$

….

$P(k) = (5/6)^{k-1} (1/6) = 5^{k-1}/6^k$

….

Geometric(p): Geometric distribution

Perform independent Bernoulli(p) trials indefinitely.

Outcome: Number of trials needed for first success, which we denote G(p) or G

Sample space: { 1, 2, 3, 4, 5, 6, …. (goes on and on)}

$P(G = 1) = P(\text{first trial is success}) = p$

$P(G = 2) = P(\text{first trial is failure and second trial is failure}) = (1-p)\,p$

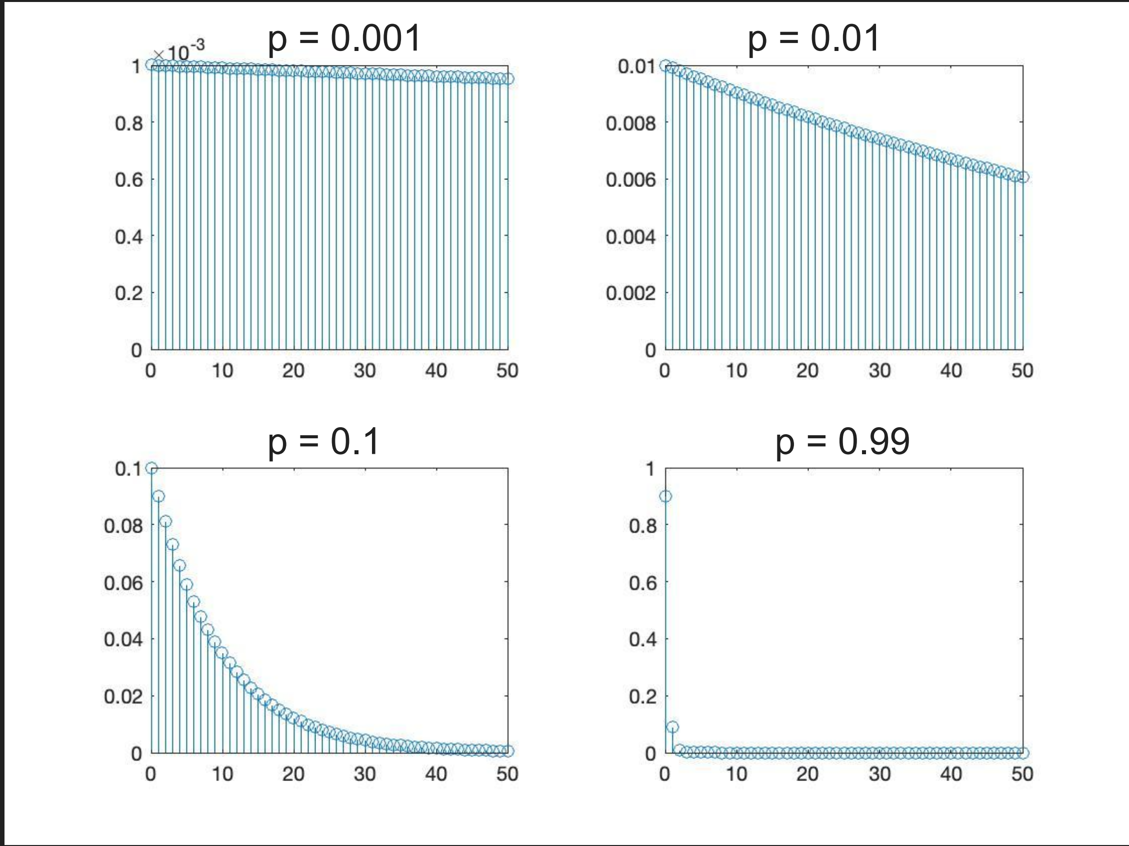
$P(G = 3) = P(\text{trial result: }001) = (1-p)^2p$

….

$P(G = k) = P(\text{first } k\text{-1 trials result in 0 and } k\text{-th trial is 1}) = (1-p)^{k-1}\,p$

….

Plot of geometric distribution (shown till k = 50)



Some observations

- Starts at p and keeps falling

- Keeps on decreasing but, if $p<1$, never goes all the way to zero

- $P(G \leq k) = P(G = 1 \text{ or } G = 2 \text{ or } \ldots\ldots \text{ or } G = k)$

  $P(G \leq k) = P(G = 1) + P(G = 2) + \ldots. + P(G = k)$

  $P(G \leq k) = p + (1 - p)p + (1 - p)^2 p + \ldots. + (1 - p)^{k-1} p$
  $= 1 - (1 - p)^k$ (important identity)

Example: Throw a die till you get 1

In Ludo, a player needs to repeatedly throw a die till she gets a 1. What is the probability that she needs lesser than 6 throws? What is the probability that she needs lesser than 11 throws? What is the probability that she needs lesser than 21 throws?

P(lesser than 6 throws) = P(G = 1 or …. or G = 5)

$= 1/6 + 5/6 \times 1/6 + (5/6)^2(1/6) + (5/6)^3(1/6) + (5/6)^4(1/6)$

$= 1 - (5/6)^5 = 0.5981$

P(lesser than 11 throws) = $1/6 + 5/6 \times 1/6 + (5/6)^2(1/6) + \ldots. + (5/6)^{10}(1/6)$

$= 1 - (5/6)^{11} = 0.8654$

P(lesser than 21 throws) = $1/6 + 5/6 \times 1/6 + (5/6)^2(1/6) + \ldots. + (5/6)^{20}(1/6)$

$= 1 - (5/6)^{21} = 0.9783$

Example: Basketball free throw contest

Player 1 is a 40% free throw shooter, while Player 2 is a 70% shooter. Each throw is independent of all previous throws. The two players alternate shooting with Player 1 starting till the first basket is scored.

(1) What is the probability that Player 1 wins before the 3rd round?
(2) What is the probability that Player 1 wins?