Exam 1 Practice Exam 1: Long List –solutions, 18.05, Spring 2022

This is a big list of practice problems for Exam 1. It includes all the problems in other sets of practice problems and many more!

1 Counting and Probability

Problem 1. *A full house in poker is a hand where three cards share one rank and two cards share another rank. How many ways are there to get a full-house? What is the probability of getting a full-house?*

Solution: We build a full-house in stages and count the number of ways to make each stage:

Stage 1. Choose the rank of the pair: $\binom{13}{1}$.

Stage 2. Choose the pair from that rank, i.e. pick 2 of 4 cards: $\binom{4}{2}$.

Stage 3. Choose the rank of the triple (from the remaining 12 ranks): $\binom{12}{1}$.

Stage 4. Choose the triple from that rank: $\binom{4}{3}$.

Number of ways to get a full-house: $\binom{13}{1}\binom{4}{2}\binom{12}{1}\binom{4}{3}$

Number of ways to pick any 5 cards out of 52: $\binom{52}{5}$

Probability of a full house: $\dfrac{\binom{13}{1}\binom{4}{2}\binom{12}{1}\binom{4}{3}}{\binom{52}{5}} \approx 0.00144$

Problem 2. *There are 3 arrangements of the word DAD, namely DAD, ADD, and DDA. How many arrangements are there of the word PROBABILITY?*

Solution: Sort the letters: A BB II L O P R T Y. There are 11 letters in all. We build arrangements by starting with 11 'slots' and placing the letters in these slots, e.g

$$\underline{A}\ \underline{B}\ \underline{I}\ \underline{B}\ \underline{I}\ \underline{L}\ \underline{O}\ \underline{P}\ \underline{R}\ \underline{T}\ \underline{Y}$$

Create an arrangement in stages and count the number of possibilities at each stage:

Stage 1: Choose one of the 11 slots to put the A: $\binom{11}{1}$

Stage 2: Choose two of the remaining 10 slots to put the B's: $\binom{10}{2}$

Stage 3: Choose two of the remaining 8 slots to put the I's: $\binom{8}{2}$

Stage 4: Choose one of the remaining 6 slots to put the L: $\binom{6}{1}$

Stage 5: Choose one of the remaining 5 slots to put the O: $\binom{5}{1}$

Stage 6: Choose one of the remaining 4 slots to put the P: $\binom{4}{1}$

Stage 7: Choose one of the remaining 3 slots to put the R: $\binom{3}{1}$

Stage 8: Choose one of the remaining 2 slots to put the T: $\binom{2}{1}$

Stage 9: Use the last slot for the Y: $\binom{1}{1}$

Number of arrangements:

$$\binom{11}{1}\binom{10}{2}\binom{8}{2}\binom{6}{1}\binom{5}{1}\binom{4}{1}\binom{3}{1}\binom{2}{1}\binom{1}{1} = 11 \cdot \frac{10 \cdot 9}{2} \cdot \frac{8 \cdot 7}{2} \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 9979200$$

Note: choosing 11 out of 1 is so simple we could have immediately written 11 instead of belaboring the issue by writing $\binom{11}{1}$. We wrote it this way to show one systematic way to think about problems like this.

Problem 3. ***(a)*** *How many ways can you arrange the letters in the word STATISTICS? (e.g. SSSTTTIIAC counts a one arrangement.)*

(b) *If all arrangements are equally likely, what is the probabilitiy the two 'i's are next to each other.*

Solution: **(a)** Create an arrangement in stages and count the number of possibilities at each stage:

Stage 1: Choose three of the 10 slots to put the S's: $\binom{10}{3}$

Stage 2: Choose three of the remaining 7 slots to put the T's: $\binom{7}{3}$

Stage 3: Choose two of the remaining 4 slots to put the I's: $\binom{4}{2}$

Stage 4: Choose one of the remaining 2 slots to put the A: $\binom{2}{1}$

Stage 5: Use the last slot for the C: $\binom{1}{1}$

Number of arrangements:

$$\binom{10}{3}\binom{7}{3}\binom{4}{2}\binom{2}{1}\binom{1}{1} = 50400.$$

(b) The are $\binom{10}{2} = 45$ equally likely ways to place the two I's.

There are 9 ways to place them next to each other, i.e. in slots 1 and 2, slots 2 and 3, ..., slots 9 and 10.

So the probability the I's are adjacent is $9/45 = 0.2$.

Problem 4. *In a ballroom dancing class the students are divided into group A and group B. There are six people in group A and seven in group B. If four As and four Bs are chosen and paired off, how many pairings are possible?*

Solution: Build the pairings in stages and count the ways to build each stage:
Stage 1: Choose the 4 from group A: $\binom{6}{4}$.
Stage 2: Choose the 4 from group B: $\binom{7}{4}$
We need to be careful because we don't want to build the same 4 couples in multiple ways.
Line up the 4 A's A_1, A_2, A_3, A_4
Stage 3: Choose a partner from the 4 Bs for A_1: 4.
Stage 4: Choose a partner from the remaining 3 Bs for A_2: 3
Stage 5: Choose a partner from the remaining 2 Bs for A_3: 2
Stage 6: Pair the last B with A_4: 1

Number of possible pairings: $\binom{6}{4}\binom{7}{4}4!$.

Note: we could have done stages 3-6 in one go as: Stages 3-6: Arrange the 4 Bs opposite the 4 As: 4! ways.

Problem 5. *Suppose you pick two cards from a deck of 52 playing cards. What is the probability that they are both queens?*

Solution: Using choices (order doesn't matter):
Number of ways to pick 2 queens: $\binom{4}{2}$. Number of ways to pick 2 cards: $\binom{52}{2}$.

All choices of 2 cards are equally likely. So, probability of 2 queens $= \dfrac{\binom{4}{2}}{\binom{52}{2}}$

Using permutations (order matters):
Number of ways to pick the first queen: 4. No. of ways to pick the second queen: 3.
Number of ways to pick the first card: 52. No. of ways to pick the second card: 51.
All arrangements of 2 cards are equally likely. So, probability of 2 queens: $\frac{4\cdot 3}{52\cdot 51}$.

Problem 6. *Suppose that there are ten students in a classroom. What is the probability that no two of them have a birthday in the same month?*

Solution: We assume each month is equally likely to be a student's birthday month.
Number of ways ten students can have birthdays in 10 different months:

$$12 \cdot 11 \cdot 10 \ldots \cdot 3 \;=\; \frac{12!}{2!}$$

Number of ways 10 students can have birthday months: 12^{10}.

Probability no two share a birthday month: $\dfrac{12!}{2!\,12^{10}} \;=\; 0.00387$.

Problem 7. *20 politicians are having a tea party, 6 Democrats and 14 Republicans. To prepare, they need to choose:*

3 people to set the table, 2 people to boil the water, 6 people to make the scones.
Each person can only do 1 task. (Note that this doesn't add up to 20. The rest of the people don't help.)

(a) *In how many different ways can they choose which people perform these tasks?*

(b) *Suppose that the Democrats all hate tea. If they only give tea to 10 of the 20 people, what is the probability that they only give tea to Republicans?*

(c) *If they only give tea to 10 of the 20 people, what is the probability that they give tea to 9 Republicans and 1 Democrat?*

Solution: (a) There are $\binom{20}{3}$ ways to choose the 3 people to set the table, then $\binom{17}{2}$ ways to choose the 2 people to boil water, and $\binom{15}{6}$ ways to choose the people to make scones. So the total number of ways to choose people for these tasks is

$$\boxed{\binom{20}{3}\binom{17}{2}\binom{15}{6} = \frac{20!}{3!\,17!}\cdot\frac{17!}{2!\,15!}\cdot\frac{15!}{6!\,9!} = \frac{20!}{3!\,2!\,6!\,9!} = 775975200.}$$

(b) The number of ways to choose 10 of the 20 people is $\binom{20}{10}$ The number of ways to choose 10 people from the 14 Republicans is $\binom{14}{10}$. So the probability that you only choose 10 Republicans is

$$\frac{\binom{14}{10}}{\binom{20}{10}} = \frac{\frac{14!}{10!\,4!}}{\frac{20!}{10!\,10!}} \approx 0.00542$$

Alternatively, you could choose the 10 people in sequence and say that there is a 14/20 probability that the first person is a Republican, then a 13/19 probability that the second one is, a 12/18 probability that third one is, etc. This gives a probability of

$$\frac{14}{20}\cdot\frac{13}{19}\cdot\frac{12}{18}\cdot\frac{11}{17}\cdot\frac{10}{16}\cdot\frac{9}{15}\cdot\frac{8}{14}\cdot\frac{7}{13}\cdot\frac{6}{12}\cdot\frac{5}{11}.$$

(You can check that this is the same as the other answer given above.)

(c) You can choose 1 Democrat in $\binom{6}{1} = 6$ ways, and you can choose 9 Republicans in $\binom{14}{9}$ ways, so the probability equals

$$\frac{6\cdot\binom{14}{9}}{\binom{20}{10}} = \frac{6\cdot\frac{14!}{9!\,5!}}{\frac{20!}{10!\,10!}} = \frac{6\cdot 14!\,10!\,10!}{9!\,5!\,20!}.$$

Problem 8. *Let A and B be two events. Suppose the probability that neither A or B occurs is 2/3. What is the probability that one or both occur?*

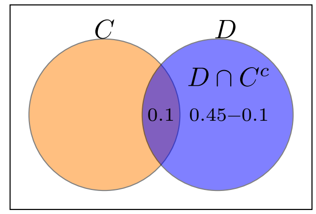
Solution: We are given $P(A^c \cap B^c) = 2/3$ and asked to find $P(A \cup B)$.

$A^c \cap B^c = (A \cup B)^c \;\Rightarrow\; P(A \cup B) = 1 - P(A^c \cap B^c) = \boxed{1/3.}$

Problem 9. *Let C and D be two events with $P(C) = 0.25$, $P(D) = 0.45$, and $P(C \cap D) = 0.1$. What is $P(C^c \cap D)$?*

Solution: D is the disjoint union of $D \cap C$ and $D \cap C^c$.

So, $P(D \cap C) + P(D \cap C^c) = P(D)$
$\Rightarrow\ P(D \cap C^c) = P(D) - P(D \cap C) = 0.45 - 0.1 = \boxed{0.35.}$
(We never use $P(C) = 0.25$.)



Problem 10. *You roll a four-sided die 3 times. For this problem we'll use the sample space with 64 equally likely outcomes.*

(a) *Write down this sample space in set notation.*

(b) *List all the outcomes in each of the following events.*

(i) A = 'Exactly 2 of the 3 rolls are fours'

(ii) B = 'At least 2 of the 3 rolls are fours'

(iii) C = 'Exactly 1 of the second and third rolls is a 4'

(iv) $A \cap C$

Solution: (a) Writing all 64 possibilities is too tedius. Here's a more compact representation

$$\{(i, j, k) \,|\, i, j, k \text{ are integers from 1 to 4}\}$$

(b) (i) Here we'll just list all 9 possibilities
{(4,4,1), (4,4,2), (4,4,3), (4,1,4), (4,2,4), (4,3,4), (1,4,4), (2,4,4), (3,4,4)}

(ii) This is the same as (i) with the addition of (4,4,4).
{ (4,4,1), (4,4,2), (4,4,3), (4,1,4), (4,2,4), (4,3,4), (1,4,4), (2,4,4), (3,4,4), (4,4,4)}

(iii) This is list is a little longer. If we're systematic about it we can still just write it out.

{(1,4,1), (2,4,1), (3,4,1), (4,4,1),
(1,4,2), (2,4,2), (3,4,2), (4,4,2),
(1,4,3), (2,4,3), (3,4,3), (4,4,3),
(1,1,4), (2,1,4), (3,1,4), (4,1,4),
(1,2,4), (2,2,4), (3,2,4), (4,2,4),
(1,3,4), (2,3,4), (3,3,4), (4,3,4)}

(iv) {(4,4,1), (4,4,2), (4,4,3), (4,1,4), (4,2,4), (4,3,4)}

Problem 11. *Suppose we have 8 teams labeled T_1, ..., T_8. Suppose they are ordered by placing their names in a hat and drawing the names out one at a time.*

(a) *How many ways can it happen that all the odd numbered teams are in the odd numbered slots and all the even numbered teams are in the even numbered slots?*

Solution: Slots 1, 3, 5, 7 are filled by T_1, T_3, T_5, T_7 in any order: 4! ways.

Slots 2, 4, 6, 8 are filled by T_2, T_4, T_6, T_8 in any order: 4! ways.

Solution: $4! \cdot 4! = 576$.

(b) *What is the probability of this happening?*

Solution: There are 8! ways to fill the 8 slots in any way.

Since each outcome is equally likely the probabilitiy is $\dfrac{4!\cdot 4!}{8!} = \dfrac{576}{40320} = 0.143 = 1.43\%$.

2 Conditional Probability and Bayes' Theorem

Problem 12. *More cards! Suppose you want to divide a 52 card deck into four hands with 13 cards each. What is the probability that each hand has a king?*

Solution: Let H_i be the event that the i^{th} hand has one king. We have the conditional probabilities

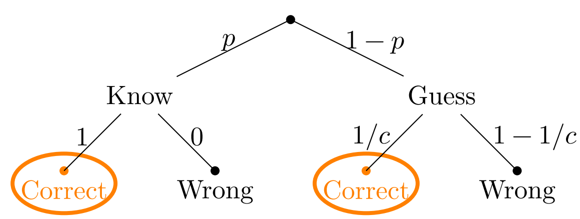
$$P(H_1) = \frac{\binom{4}{1}\binom{48}{12}}{\binom{52}{13}}; \quad P(H_2|H_1) = \frac{\binom{3}{1}\binom{36}{12}}{\binom{39}{13}}; \quad P(H_3|H_1\cap H_2) = \frac{\binom{2}{1}\binom{24}{12}}{\binom{26}{13}}$$

$P(H_4|H_1\cap H_2\cap H_3) = 1$

$$\begin{aligned} P(H_1\cap H_2\cap H_3\cap H_4) &= P(H_4|H_1\cap H_2\cap H_3)\,P(H_3|H_1\cap H_2)\,P(H_2|H_1)\,P(H_1) \\ &= \frac{\binom{2}{1}\binom{24}{12}\binom{3}{1}\binom{36}{12}\binom{4}{1}\binom{48}{12}}{\binom{26}{13}\binom{39}{13}\binom{52}{13}}. \end{aligned}$$

Problem 13. *Suppose you are taking a multiple-choice test with c choices for each question. In answering a question on this test, the probability that you know the answer is p. If you don't know the answer, you choose one at random. What is the probability that you knew the answer to a question, given that you answered it correctly?*

Solution: The following tree shows the setting



Let C be the event that you answer the question correctly. Let K be the event that you actually know the answer. The left circled node shows $P(K\cap C) = p$. Both circled nodes together show $P(C) = p + (1-p)/c$. So,

$$P(K|C) \;=\; \frac{P(K\cap C)}{P(C)} \;=\; \frac{p}{p+(1-p)/c}$$

Or we could use the algebraic form of Bayes' theorem and the law of total probability: Let G stand for the event that you're guessing. Then we have,
$P(C|K) = 1$, $P(K) = p$, $P(C) = P(C|K)P(K) + P(C|G)P(G) = p + (1-p)/c$. So,

$$P(K|C) = \frac{P(C|K)P(K)}{P(C)} = \frac{p}{p+(1-p)/c}$$

Problem 14. *Corrupted by their power, the judges running the popular game show* America's Next Top Mathematician *have been taking bribes from many of the contestants. Each episode, a given contestant is either allowed to stay on the show or is kicked off.*

If the contestant has been bribing the judges they will be allowed to stay with probability 1. If the contestant has not been bribing the judges, they will be allowed to stay with probability 1/3.

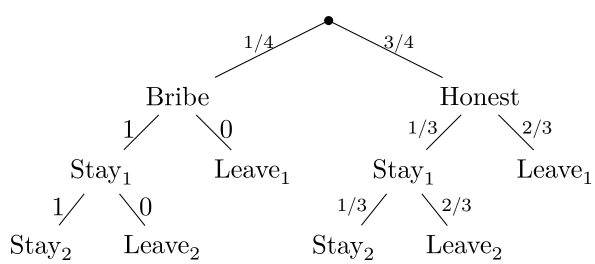
Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds, i.e., if a contestant bribes them in the first round, they bribe them in the second round too (and vice versa).

(a) *If you pick a random contestant who was allowed to stay during the first episode, what is the probability that they were bribing the judges?*

(b) *If you pick a random contestant, what is the probability that they are allowed to stay during both of the first two episodes?*

(c) *If you pick random contestant who was allowed to stay during the first episode, what is the probability that they get kicked off during the second episode?*

Solution: The following tree shows the setting. Stay_1 means the contestant was allowed to stay during the first episode and stay_2 means the they were allowed to stay during the second.



Let's name the relevant events:
B = the contestant is bribing the judges
H = the contestant is honest (not bribing the judges)
S_1 = the contestant was allowed to stay during the first episode
S_2 = the contestant was allowed to stay during the second episode
L_1 = the contestant was asked to leave during the first episode
L_2 = the contestant was asked to leave during the second episode

(a) We first compute $P(S_1)$ using the law of total probability.

$$P(S_1) = P(S_1|B)P(B) + P(S_1|H)P(H) = 1 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{2}.$$

We therefore have (by Bayes' rule) $P(B|S_1) = P(S_1|B)\dfrac{P(B)}{P(S_1)} = 1 \cdot \dfrac{1/4}{1/2} = \dfrac{1}{2}$.

(b) Using the tree we have the total probability of S_2 is

$$P(S_2) = \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{3}$$

(c) We want to compute $P(L_2|S_1) = \dfrac{P(L_2 \cap S_1)}{P(S_1)}$.

From the calculation we did in part (a), $P(S_1) = 1/2$. For the numerator, we have (see the tree)

$$P(L_2 \cap S_1) = P(L_2 \cap S_1|B)P(B) + P(L_2 \cap S_1|H)P(H) = 0 \cdot \frac{1}{4} + \frac{2}{9} \cdot \frac{3}{4} = \frac{1}{6}$$

Therefore $\boxed{P(L_2|S_1) = \dfrac{1/6}{1/2} = \dfrac{1}{3}.}$

Problem 15. *Consider the Monty Hall problem. Let's label the door with the car behind it a and the other two doors b and c. In the game the contestant chooses a door and then Monty chooses a door, so we can label each outcome as 'contestant followed by Monty', e.g ab means the contestant chose a and Monty chose b.*

(a) *Make a 3×3 probability table showing probabilities for all possible outcomes.*

(b) *Make a probability tree showing all possible outcomes.*

(c) *Suppose the contestant's strategy is to switch. List all the outcomes in the event 'the contestant wins a car'. What is the probability the contestant wins?*

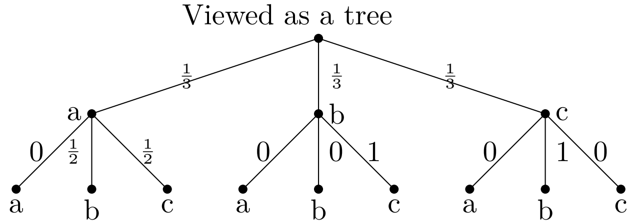
(d) *Redo part (c) with the strategy of not switching.*

Solution: **(a)** and **(b)** In the tree the first row is the contestant's choice and the second row is the host's (Monty's) choice.

Viewed as a table

		Contestant		
		a	b	c
	a	0	0	0
Host	b	1/6	0	1/3
	c	1/6	1/3	0

Viewed as a tree



(b) With this strategy the contestant wins with {bc, cb}. The probability of winning is $P(bc) + P(cb) = 2/3$. (Both the tree and the table show this.)

(c) $\{ab, ac\}$, probability $= 1/3$.

Problem 16. *Two dice are rolled.*

A = 'sum of two dice equals 3'
B = 'sum of two dice equals 7'
C = 'at least one of the dice shows a 1'

(a) *What is $P(A|C)$?*

(b) *What is $P(B|C)$?*

(c) *Are A and C independent? What about B and C?*

Solution: Sample space =

$$\Omega = \{(1,1),\ (1,2),\ (1,3),\ \ldots,\ (6,6)\} = \{(i,j)\,|\,i,j = 1,\ 2,\ 3,\ 4,\ 5,\ 6\}.$$

(Each outcome is equally likely, with probability 1/36.)

$A = \{(1,2),\ (2,1)\}$,

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

$C = \{(1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ (1,6),\ (2,1),\ (3,1),\ (4,1),\ (5,1),\ (6,1)\}$

(a) $P(A|C) = \dfrac{P(A \cap C)}{P(C)} = \dfrac{2/36}{11/36} = \dfrac{2}{11}$..

(b) $P(B|C) = \dfrac{P(B \cap C)}{P(C)} = \dfrac{2/36}{11/36} = \dfrac{2}{11}$..

(c) $P(A) = 2/36 \neq P(A|C)$, so they are not independent. Similarly, $P(B) = 6/36 \neq P(B|C)$, so they are not independent.

Problem 17. *There is a screening test for prostate cancer that looks at the level of PSA (prostate-specific antigen) in the blood. There are a number of reasons besides prostate cancer that a man can have elevated PSA levels. In addition, many types of prostate cancer develop so slowly that that they are never a problem. Unfortunately there is currently no test to distinguish the different types and using the test is controversial because it is hard to quantify the accuracy rates and the harm done by false positives.*

For this problem we'll call a positive test a true positive if it catches a dangerous type of prostate cancer. We'll assume the following numbers:

Rate of prostate cancer among men over 50 = 0.0005
True positive rate for the test = 0.9
False positive rate for the test = 0.01

Let T be the event a man has a positive test and let D be the event a man has a dangerous type of the disease. Find $P(D|T)$ and $P(D|T^c)$.

Solution: You should write this out in a tree! (For example, see the solution to the next problem.)

We compute all the pieces needed to apply Bayes' rule. We're given
$P(T|D) = 0.9 \ \Rightarrow\ P(T^c|D) = 0.1, \quad P(T|D^c) = 0.01 \ \Rightarrow\ P(T^c|D^c) = 0.99.$
$P(D) = 0.0005 \ \Rightarrow\ P(D^c) = 1 - P(D) = 0.9995.$

We use the law of total probability to compute $P(T)$:

$$P(T) = P(T|D)\,P(D) + P(T|D^c)\,P(D^c) = 0.9 \cdot 0.0005 + 0.01 \cdot 0.9995 = 0.010445$$

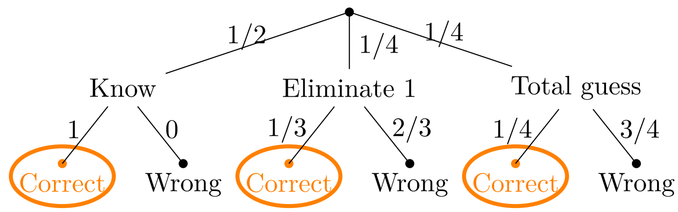
Now we can use Bayes' rule to answer the questions:

$$P(D|T) = \frac{P(T|D)\,P(D)}{P(T)} = \frac{0.9 \times 0.0005}{0.010445} = 0.043$$
$$P(D|T^c) = \frac{P(T^c|D)\,P(D)}{P(T^c)} = \frac{0.1 \times 0.0005}{0.989555} = 5.0 \times 10^{-5}$$

Problem 18. *A multiple choice exam has 4 choices for each question. A student has studied enough so that the probability they will know the answer to a question is 0.5, the probability that they will be able to eliminate one choice is 0.25, otherwise all 4 choices seem* *equally plausible. If they know the answer they will get the question right. If not they have to guess from the 3 or 4 choices.*

As the teacher you want the test to measure what the student knows. If the student answers a question correctly what's the probability they knew the answer?

Solution: We show the probabilities in a tree:



For a given problem let C be the event the student gets the problem correct and K the event the student knows the answer.

The question asks for $P(K|C)$.

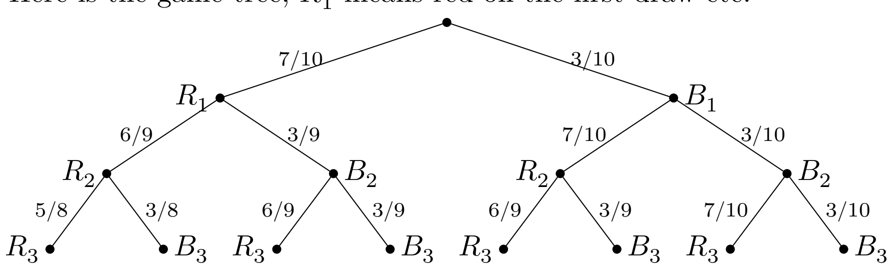
We'll compute this using Bayes' rule:

$$P(K|C) = \frac{P(C|K)\,P(K)}{P(C)} = \frac{1 \cdot 1/2}{1/2 + 1/12 + 1/16} = \frac{24}{31} \approx 0.774 = 77.4\%$$

Problem 19. *Suppose you have an urn containing 7 red and 3 blue balls. You draw three balls at random. On each draw, if the ball is red you set it aside and if the ball is blue you put it back in the urn. What is the probability that the third draw is blue?*

(If you get a blue ball it counts as a draw even though you put it back in the urn.)

Solution: Here is the game tree, R_1 means red on the first draw etc.



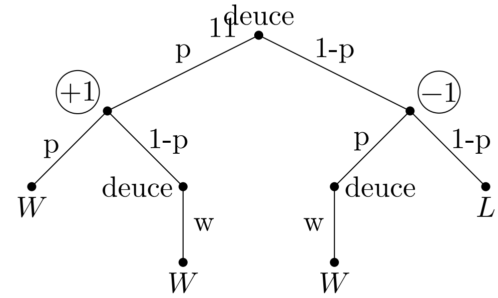
Summing the probability to all the B_3 nodes we get

$$P(B_3) = \frac{7}{10}\cdot\frac{6}{9}\cdot\frac{3}{8} + \frac{7}{10}\cdot\frac{3}{9}\cdot\frac{3}{9} + \frac{3}{10}\cdot\frac{7}{10}\cdot\frac{3}{9} + \frac{3}{10}\cdot\frac{3}{10}\cdot\frac{3}{10} = 0.350.$$

Problem 20. *Some games, like tennis or ping pong, reach a state called* deuce. *This means that the score is tied and a player wins the game when they get* two *points ahead of the other player. Suppose the probability that you win a point is p and this is true independently for all points. If the game is at deuce what is the probability you win the game?*

This is a tricky problem, but amusing if you like puzzles.

Solution: Let W be the event you win the game from deuce and L



the event you lose. For convenience, define $w = P(W)$.
The figure shows the complete game tree through 2 points. In the third level we just abreviate by indicating the probability of winning from deuce.
The nodes marked +1 and -1, indicate whether you won or lost the first point.
Summing all the paths to W we get

$$w = P(W) = p^2 + p(1-p)w + (1-p)pw = p^2 + 2p(1-p)w \;\Rightarrow\; \boxed{w = \frac{p^2}{1-2p(1-p)}}.$$

Problem 21. (Bayes formula)
A student takes a multiple-choice exam. Suppose for each question they either know the answer or gamble and choose an option at random. Further suppose that if they knows the answer, the probability of a correct answer is 1, and if they gamble, this probability is 1/4. To pass, students need to answer at least 60% of the questions correctly. The student has "studied for a minimal pass," i.e., with probability 0.6 they know the answer to a question. For a single question, given that they answers it correctly, what is the probability that they actually knew *the answer?*

For a given problem let C be the event the student gets the problem correct and K the event the student knows the answer.

The question asks for $P(K|C)$.

We'll compute this using Bayes' rule: $P(K|C) = \dfrac{P(C|K)\,P(K)}{P(C)}$.

We're given: $\quad P(C|K) = 1, \quad P(K) = 0.6.$

Law of total prob.:

$$P(C) = P(C|K)\,P(K) \;+\; P(C|K^c)\,P(K^c) = 1 \cdot 0.6 \;+\; 0.25 \cdot 0.4 = 0.7.$$

Therefore $P(K|C) = \dfrac{0.6}{0.7} = 0.857 = 85.7\%.$

3 Independence

Problem 22. *Suppose that $P(A) = 0.4, P(B) = 0.3$ and $P((A \cup B)^C) = 0.42$. Are A and B independent?*

Solution: We have $P(A \cup B) = 1 - 0.42 = 0.58$ and we know because of the inclusion-exclusion principle that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Thus,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.4 + 0.3 - 0.58 = 0.12 = (0.4)(0.3) = P(A)P(B).$$

So, A and B are independent.

Problem 23. *Suppose now that events A, B and C are mutually independent with*

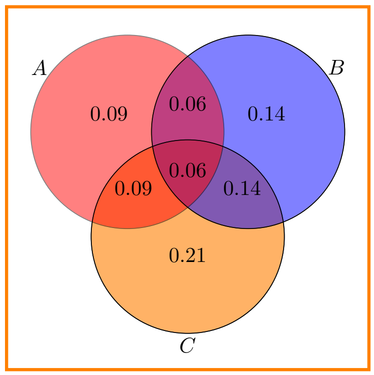
$$P(A) = 0.3, \quad P(B) = 0.4, \quad P(C) = 0.5.$$

Compute the following: (Hint: Use a Venn diagram)
(i) $P(A \cap B \cap C^c)$ (ii) $P(A \cap B^c \cap C)$ (iii) $P(A^c \cap B \cap C)$

Solution: By the mutual independence we have

$$P(A \cap B \cap C) = P(A)P(B)P(C) = 0.06 \qquad P(A \cap B) = P(A)P(B) = 0.12$$
$$P(A \cap C) = P(A)P(C) = 0.15 \qquad P(B \cap C) = P(B)P(C) = 0.2$$

We show this in the following Venn diagram



Note that, for instance, $P(A \cap B)$ is split into two pieces. One of the pieces is $P(A \cap B \cap C)$ which we know and the other we compute as $P(A \cap B) - P(A \cap B \cap C) = 0.12 - 0.06 = 0.06$. The other intersections are similar.

We can read off the asked for probabilities from the diagram.
(i) $P(A \cap B \cap C^c) = 0.06$
(ii) $P(A \cap B^c \cap C) = 0.09$
(iii) $P(A^c \cap B \cap C) = 0.14$.

Problem 24. *You roll a twenty-sided die. Determine whether the following pairs of events are independent.*

(a) *'You roll an even number' and 'You roll a number less than or equal to 10'.*

(b) *'You roll an even number' and 'You roll a prime number'.*

Solution: E = even numbered = $\{2, 4, 6, 8, 10, 12, 14, 16, 18, 20\}$.

L = roll ≤ 10 = $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

B = roll is prime = $\{2, 3, 5, 7, 11, 13, 17, 19\}$ (We use B because P is not a good choice.)

(a) $P(E) = 10/20$, $P(E|L) = 5/10$. These are the same, so the events are independent.

(b) $P(E) = 10/20$. $P(E|B) = 1/8$. These are not the same so the events are not independent.

Problem 25. *Suppose A and B are events with $0 < P(A) < 1$ and $0 < P(B) < 1$.*

***(a)** If A and B are disjoint can they be independent?*

***(b)** If A and B are independent can they be disjoint?*

***(c)** If $A \subset B$ can they be independent?*

Solution: The answer to all three parts is 'No'. Each of these answers relies on the fact that the probabilities of A and B are strictly between 0 and 1.

To show A and B are not independent we need to show either $P(A \cap B) \neq P(A) \cdot P(B)$ or $P(A|B) \neq P(A)$.

(a) No, they cannot be independent: $A \cap B = \emptyset \Rightarrow P(A \cap B) = 0 \neq P(A) \cdot P(B)$.

(b) No, they cannot be disjoint: same reason as in part (a).

(c) No, they cannot be independent: $A \subset B \Rightarrow A \cap B = A$
$\Rightarrow P(A \cap B) = P(A) > P(A) \cdot P(B)$. The last inequality follows because $P(B) < 1$.

4 Expectation and Variance

Problem 26. *Directly from the definitions of expected value and variance, compute $E[X]$ and $Var(X)$ when X has probability mass function given by the following table:*

X	*-2*	*-1*	*0*	*1*	*2*
pmf	*1/15*	*2/15*	*3/15*	*4/15*	*5/15*

Solution: We compute

$$E[X] = -2 \cdot \frac{1}{15} + -1 \cdot \frac{2}{15} + 0 \cdot \frac{3}{15} + 1 \cdot \frac{4}{15} + 2 \cdot \frac{5}{15} = \frac{2}{3}.$$

Thus

$$\begin{aligned}\text{Var}(X) &= E[(X - \frac{2}{3})^2] \\ &= \left(-2 - \frac{2}{3}\right)^2 \cdot \frac{1}{15} + \left(-1 - \frac{2}{3}\right)^2 \cdot \frac{2}{15} + \left(0 - \frac{2}{3}\right)^2 \cdot \frac{3}{15} + \left(1 - \frac{2}{3}\right)^2 \cdot \frac{4}{15} + \left(2 - \frac{2}{3}\right)^2 \cdot \frac{5}{15} \\ &= \frac{14}{9}.\end{aligned}$$

Problem 27. *Suppose that X takes values between 0 and 1 and has probability density function $2x$. Compute $\text{Var}(X)$ and $\text{Var}(X^2)$.*

Solution: We will make use of the formula $\text{Var}(Y) = E[Y^2] - E[Y]^2$. First we compute

$$E[X] = \int_0^1 x \cdot 2x dx = \frac{2}{3}$$

$$E[X^2] = \int_0^1 x^2 \cdot 2x dx = \frac{1}{2}$$

$$E[X^4] = \int_0^1 x^4 \cdot 2x dx = \frac{1}{3}.$$

Thus,

$$\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

and

$$\text{Var}(X^2) = E[X^4] - (E[X^2])^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Problem 28. *The random variable X takes values -1, 0, 1 with probabilities 1/8, 2/8, 5/8 respectively.*

(a) *Compute $E[X]$.*

(b) *Give the pmf of $Y = X^2$ and use it to compute $E[Y]$.*

(c) *Instead, compute $E[X^2]$ directly from an extended table.*

(d) *Compute $\text{Var}(X)$.*

(a) Solution: We have

X values:	-1	0	1
prob:	1/8	2/8	5/8
X^2	1	0	1

So, $\boxed{E[X] = -1/8 + 5/8 = 1/2.}$

(b) Solution:

Y values:	0	1
prob:	2/8	6/8

$\Rightarrow$ $\boxed{E[Y] = 6/8 = 3/4.}$

(c) Solution: The change of variables formula just says to use the bottom row of the table in part (a): $E[X^2] = 1 \cdot (1/8) + 0 \cdot (2/8) + 1 \cdot (5/8) = 3/4$ (same as part (b)).

(d) Solution: $\boxed{\text{Var}(X) = E[X^2] - E[X]^2 = 3/4 - 1/4 = 1/2.}$

Problem 29. *Suppose X is a random variable with $E[X] = 5$ and $\text{Var}(X) = 2$. What is $E[X^2]$?*

Solution: Use $\text{Var}(X) = E[X^2] - E[X]^2 \;\Rightarrow\; 2 = E[X^2] - 25 \;\Rightarrow\; \boxed{E[X^2] = 27.}$

Problem 30. *Compute the expectation and variance of a Bernoulli(p) random variable.*

Solution: Make a table:

X:	0	1
prob:	(1-p)	p
X^2	0	1.

From the table, $\boxed{E[X] = 0 \cdot (1-p) + 1 \cdot p = p.}$

Since X and X^2 have the same table $E[X^2] = E[X] = p$.

Therefore, $\boxed{\text{Var}(X) = p - p^2 = p(1-p).}$

Problem 31. *Suppose 100 people all toss a hat into a box and then proceed to randomly pick out of a hat. What is the expected number of people to get their own hat back.*

Hint: express the number of people who get their own hat as a sum of random variables whose expected value is easy to compute.

Solution: Let X be the number of people who get their own hat.

Following the hint: let X_j represent whether person j gets their own hat. That is, $X_j = 1$ if person j gets their hat and 0 if not.

We have, $X = \sum_{j=1}^{100} X_j$, so $E[X] = \sum_{j=1}^{100} E[X_j]$.

Since person j is equally likely to get any hat, we have $P(X_j = 1) = 1/100$. Thus, $X_j \sim \text{Bernoulli}(1/100) \;\Rightarrow\; E[X_j] = 1/100 \;\Rightarrow\; \boxed{E[X] = 1.}$

Problem 32. *Suppose I play a gambling game with even odds. So, I can wager b dollars and I either win or lose b dollars with probability $p = 0.5$.*

I employ the following strategy to try to guarantee that I win some money.

I bet \$1; if I lose, I double my bet to \$2, if I lose I double my bet again. I continue until I win. Eventually I'm sure to win a bet and net \$1 (run through the first few rounds and you'll see why this is the net).

If this really worked casinos would be out of business. Our goal in this problem is to understand the flaw in the strategy.

(a) *Let X be the amount of money bet on the last game (the one I win). X takes values 1, 2, 4, 8, ... Determine the probability mass function for X. That is, find $p(2^k)$, where k is in {0, 1, 2, ...}.*

Solution: It is easy to see that (e.g. look at the probability tree) $P(2^k) = \dfrac{1}{2^{k+1}}$.

(b) *Compute $E[X]$.*

Solution: $E[X] = \sum_{k=0}^{\infty} 2^k \dfrac{1}{2^{k+1}} = \sum \dfrac{1}{2} = \infty$. Technically, $E[X]$ is undefined in this case.

(c) *Use your answer in part (b) to explain why the stategy is a bad one.*

Solution: Technically, $E[X]$ is undefined in this case. But the value of ∞ tells us what is wrong with the scheme. Since the average last bet is infinite, I need to have an infinite amount of money in reserve.

This problem and solution is often referred to as the **St. Petersburg paradox**

Problem 33. *Suppose you roll a fair 6-sided die 100 times (independently), and you get \$3 every time you roll a 6.*

Let X_1 be the number of dollars you win on rolls 1 through 25.
Let X_2 be the number of dollars you win on rolls 26 through 50.
Let X_3 be the number of dollars you win on rolls 51 through 75.
Let X_4 be the number of dollars you win on rolls 76 throught 100.

Let $X = X_1 + X_2 + X_3 + X_4$ be the total number of dollars you win over all 100 rolls.

***(a)** What is the probability mass function of X?*

(b) *What is the expectation and variance of X?*

(c) *Let $Y = 4X_1$. (So instead of rolling 100 times, you just roll 25 times and multiply your winnings by 4.)*

(i) What are the expectation and variance of Y?

(ii) How do the expectation and variance of Y compare to those of X? (That is, are they bigger, smaller, or equal?) Explain (briefly) why this makes sense.

Solution: (a) There are a number of ways to present this.

Let T be the total number of times you roll a 6 in the 100 rolls. We know $T \sim \text{Binomial}(100, 1/6)$. Since you win \$3 every time you roll a 6, we have $X = 3T$. So, we can write

$$P(X = 3k) = \binom{100}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{100-k}, \quad \text{for } k = 0,\ 1,\ 2,\ \dots,\ 100.$$

Alternatively we could write

$$P(X = x) = \binom{100}{x/3}\left(\frac{1}{6}\right)^{x/3}\left(\frac{5}{6}\right)^{100-x/3}, \quad \text{for } x = 0,\ 3,\ 6,\ \dots,\ 300.$$

(b) $E[X] = E[3T] = 3E[T] = 3 \cdot 100 \cdot \frac{1}{6} = 50$,

$\text{Var}(X) = \text{Var}(3T) = 9\text{Var}(T) = 9 \cdot 100 \cdot \frac{1}{6} \cdot \frac{5}{6} = 125$.

(c) (i) Let T_1 be the total number of times you roll a 6 in the first 25 rolls. So, $X_1 = 3T_1$ and $Y = 12T_1$.

Now, $T_1 \sim \text{Binomial}(25, 1/6)$, so

$$E[Y] = 12E[T_1] = 12 \cdot 25 \cdot 16 = 50.$$

and

$$\text{Var}(Y) = 144\text{Var}(T_1) = 144 \cdot 25 \cdot \frac{1}{6} \cdot \frac{5}{6} = 500.$$

(ii) The expectations are the same by linearity because X and Y are the both

$3 \times 100 \times$ a Bernoulli(1/6) random variable.

For the variance, $\text{Var}(X) = 4\text{Var}(X_1)$ because X is the sum of 4 *independent* variables all identical to X_1. However $\text{Var}(Y) = \text{Var}(4X_1) = 16\text{Var}(X_1)$. So, the variance of Y is 4 times that of X. This should make some intuitive sense because X is built out of more independent trials than X_1.

Another way of thinking about it is that the difference between Y and its expectation is four times the difference between X_1 and its expectation. However, the difference between X and its expectation is the sum of such a difference for X_1, X_2, X_3, and X_4. It's probably the case that some of these deviations are positive and some are negative, so the absolute value of this difference for the sum is probably less than four times the absolute value of this difference for one of the variables, i.e. the deviations are likely to cancel to some extent.

5 Probability Mass Functions, Probability Density Functions and Cumulative Distribution Functions

Problem 34. *Suppose that $X \sim Bin(n, 0.5)$. Find the probability mass function of $Y = 2X$.*

Solution: For $y = 0, 2, 4, \ldots, 2n$,

$$P(Y = y) = P(X = \frac{y}{2}) = \binom{n}{y/2}\left(\frac{1}{2}\right)^n.$$

Problem 35. *(a) Suppose that X is uniform on $[0,1]$. Compute the pdf and cdf of X.*

(b) If $Y = 2X + 5$, compute the pdf and cdf of Y.

(a) Solution: We have $f_X(x) = 1$ for $0 \le x \le 1$. The cdf of X is

$$F_X(x) = \int_0^x f_X(t)dt = \int_0^x 1dt = x.$$

(b) Solution: Since X is between 0 and 1 we have Y is between 5 and 7. Now for $5 \le y \le 7$, we have

$$F_Y(y) = P(Y \le y) = P(2X + 5 \le y) = P(X \le \frac{y-5}{2}) = F_X(\frac{y-5}{2}) = \frac{y-5}{2}.$$

Differentiating $P(Y \le y)$ with respect to y, we get the probability density function of Y, for $5 \le y \le 7$,

$$f_Y(y) = \frac{1}{2}.$$

Problem 36. *(a) Suppose that X has probability density function $f_X(x) = \lambda e^{-\lambda x}$ for $x \ge 0$. Compute the cdf, $F_X(x)$.*

(b) If $Y = X^2$, compute the pdf and cdf of Y.

(a) Solution: We have cdf of X,

$$F_X(x) = \int_0^x \lambda \mathrm{e}^{-\lambda x} dx = 1 - \mathrm{e}^{-\lambda x}.$$

Now for $y \ge 0$, we have

(b) Solution:

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(X \le \sqrt{y}) = 1 - \mathrm{e}^{-\lambda\sqrt{y}}.$$

Differentiating $F_Y(y)$ with respect to y, we have

$$f_Y(y) = \frac{\lambda}{2} y^{-\frac{1}{2}} \mathrm{e}^{-\lambda\sqrt{y}}.$$

Problem 37. *Suppose that X is a random variable that takes on values 0, 2 and 3 with probabilities 0.3, 0.1, 0.6 respectively. Let $Y = 3(X-1)^2$.*

(a) What is the expectation of X?

(b) What is the variance of X?

(c) What is the expection of Y?

(d) *Let $F_Y(t)$ be the cumulative density function of Y. What is $F_Y(7)$?*

(a) Solution: We first make the probability tables

X	0	2	3
prob.	0.3	0.1	0.6
Y	3	3	12

So, $E[X] = 0 \cdot 0.3 + 2 \cdot 0.1 + 3 \cdot 0.6 = 2$

(b) Solution: $E[X^2] = 0 \cdot 0.3 + 4 \cdot 0.1 + 9 \cdot 0.6 = 5.8 \;\Rightarrow\; \mathrm{Var}(X) = E[X^2] - E[X]^2 = 5.8 - 4 = 1.8$.

(c) Solution: $E[Y] = 3 \cdot 0.3 + 3 \cdot 0.1 + 12 \cdot 6 = 8.4$.

(d) Solution: From the table we see that $F_Y(7) = P(Y \le 7) = 0.4$.

Problem 38. *Let T be the waiting time for customers in a queue. Suppose that T is exponential with pdf $f(t) = 2\mathrm{e}^{-2t}$ on $[0, \infty)$.*

Find the pdf of the rate at which customers are served $R = 1/T$.

Solution: The CDF for T is

$$F_T(t) = P(T \le t) = \int_0^t 2\mathrm{e}^{-2u}\, du = -\mathrm{e}^{-2u}\Big|_0^t = 1 - \mathrm{e}^{-2t}.$$

Next, we find the CDF of R. R takes values in $(0, \infty)$.

For $0 < r$,

$$F_R(r) = P(R \le r) = P(1/T < r) = P(T > 1/r) = 1 - F_T(1/r) = \mathrm{e}^{-2/r}.$$

We differentiate to get $f_R(r) = \dfrac{d}{dr}\left(\mathrm{e}^{-2/r}\right) = \dfrac{2}{r^2}\mathrm{e}^{-2/r}$.

Problem 39. *A continuous random variable X has PDF $f(x) = x + ax^2$ on [0,1]*

Find a, the CDF and $P(0.5 < X < 1)$.

Solution: First we find the value of a:

$$\int_0^1 f(x)\, dx = 1 = \int_0^1 x + ax^2\, dx = \frac{1}{2} + \frac{a}{3} \;\Rightarrow\; a = 3/2.$$

The CDF is $F_X(x) = P(X \le x)$. We break this into cases:

(i) $b < 0$, so $F_X(b) = 0$.

(ii) $0 \le b \le 1$, so $F_X(b) = \int_0^b x + \frac{3}{2}x^2\, dx = \frac{b^2}{2} + \frac{b^3}{2}$.

(iii) $1 < x$, so $F_X(b) = 1$.

Using F_X we get

$$P(0.5 < X < 1) = F_X(1) - F_X(0.5) = 1 - \left(\frac{0.5^2 + 0.5^3}{2}\right) = \frac{13}{16}.$$

Problem 40. (PMF of a sum)
Suppose X and Y are independent and $X \sim$ Bernoulli$(1/2)$ and $Y \sim$ Bernoulli$(1/3)$. Determine the pmf of $X + Y$

Solution: First we'll give the joint probability table:

$Y \backslash X$	0	1	
0	1/3	1/3	2/3
1	1/6	1/6	1/3
	1/2	1/2	1

We'll use the joint probabilities to build the probability table for the sum.

$X+Y$	0	1	2
(X,Y)	(0,0)	(0,1), (1,0)	(1,1)
prob.	1/3	$1/6+1/3$	1/6
prob.	1/3	1/2	1/6

Problem 41. *Let X be a discrete random variable with pmf p given by:*

x	-2	-1	0	1	2
$p(x)$	1/15	2/15	3/15	4/15	5/15

(a) *Let $Y = X^2$. Find the pmf of Y.*

(b) *Find the value the cdf of X at -3/2, 3/4, 7/8, 1, 1.5, 5.*

(c) *Find the value the cdf of Y at -3/2, 3/4, 7/8, 1, 1.5, 5.*

Solution: (a) Note: $Y = 1$ when $X = 1$ or $X = -1$, so

$$P(Y = 1) = P(X = 1) + P(X = -1).$$

Values y of Y	0	1	4
pmf $p_Y(y)$	3/15	6/15	6/15

(b) and **(c)** To distinguish the distribution functions we'll write F_X and F_Y.

Using the tables in part (a) and the definition $F_X(a) = P(X \le a)$ etc. we get

a	-3/2	3/4	7/8	1	1.5	5
$F_X(a)$	1/15	6/15	6/15	10/15	10/15	1
$F_Y(a)$	0	3/15	3/15	9/15	9/15	1

Problem 42. *Suppose that the cdf of X is given by:*

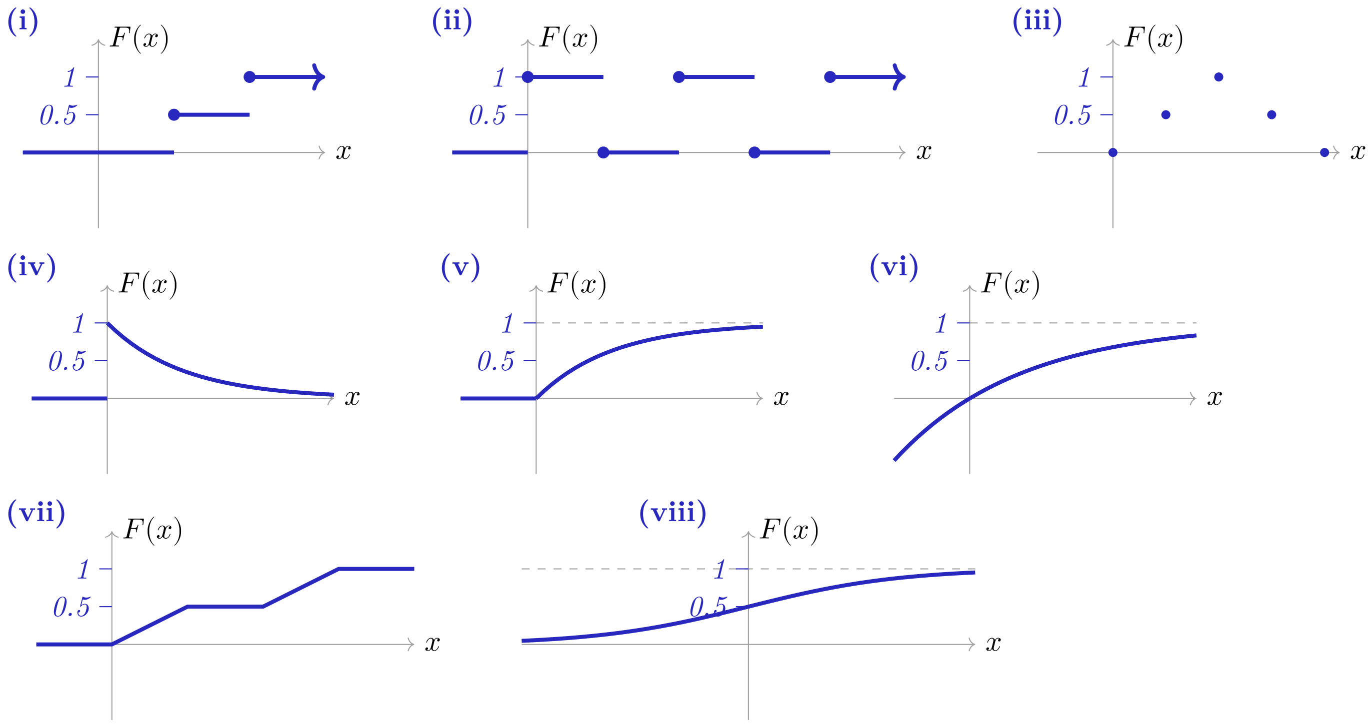
$$F(a) = \begin{cases} 0 & \text{for } a < 0 \\ \frac{1}{5} & \text{for } 0 \le a < 2 \\ \frac{2}{5} & \text{for } 2 \le a < 4 \\ 1 & \text{for } a \ge 4. \end{cases}$$

Determine the pmf of X.

Solution: The jumps in the distribution function are at 0, 2, 4. The value of $p(a)$ at a jump is the height of the jump:

a	0	2	4
$p(a)$	1/5	1/5	3/5

Problem 43. *For each of the following say whether it can be the graph of a cdf. If it can be, say whether the variable is discrete or continuous.*



Solution: (i) yes, discrete, (ii) no, (iii) no, (iv) no, (v) yes, continuous (vi) no (vii) yes, continuous, (viii) yes, continuous.

Problem 44. *Suppose X has range [0,1] and has cdf*

$$F(x) = x^2 \quad \text{for } 0 \le x \le 1.$$

Compute $P(\frac{1}{2} < X < \frac{3}{4})$.

Solution: $P(1/2 \le X \le 3/4) = F(3/4) - F(1/2) = (3/4)^2 - (1/2)^2 = \boxed{5/16}$.

Problem 45. *Let X be a random variable with range $[0,1]$ and cdf*

$$F(X) = 2x^2 - x^4 \quad \text{for } 0 \le x \le 1.$$

(a) *Compute $P(\frac{1}{4} \le X \le \frac{3}{4})$.*

(b) *What is the pdf of X?*

Solution: (a) $P(1/4 \le X \le 3/4) = F(3/4) - F(1/4) = \boxed{11/16 = 0.6875.}$

(b) $f(x) = F'(x) = 4x - 4x^3$ in [0,1].

6 Distributions with Names

Problem 46. Exponential Distribution

Suppose that buses arrive are scheduled to arrive at a bus stop at noon but are always X minutes late, where X is an exponential random variable with probability density function $f_X(x) = \lambda e^{-\lambda x}$. Suppose that you arrive at the bus stop precisely at noon.

(a) *Compute the probability that you have to wait for more than five minutes for the bus to arrive.*

Solution: We compute

$$P(X \geq 5) = 1 - P(X < 5) = 1 - \int_0^5 \lambda e^{-\lambda x} dx = 1 - (1 - e^{-5\lambda}) = e^{-5\lambda}.$$

(b) *Suppose that you have already waiting for 10 minutes. Compute the probability that you have to wait an additional five minutes or more.*

Solution: We want $P(X \geq 15 | X \geq 10)$. First observe that $P(X \geq 15, X \geq 10) = P(X \geq 15)$. From similar computations in (a), we know

$$P(X \geq 15) = e^{-15\lambda} \qquad\qquad P(X \geq 10) = e^{-10\lambda}.$$

From the definition of conditional probability,

$$P(X \geq 15 | X \geq 10) = \frac{P(X \geq 15, X \geq 10)}{P(X \geq 10)} = \frac{P(X \geq 15)}{P(X \geq 10)} = e^{-5\lambda}$$

Note: This is an illustration of the memorylessness property of the exponential distribution.

Problem 47. Normal Distribution: *Throughout these problems, let ϕ and Φ be the pdf and cdf, respectively, of the standard normal distribution Suppose Z is a standard normal random variable and let $X = 3Z + 1$.*

(a) *Express $P(X \leq x)$ in terms of Φ*

Solution: We have

$$F_X(x) = P(X \leq x) = P(3Z + 1 \leq x) = P(Z \leq \frac{x-1}{3}) = \Phi\left(\frac{x-1}{3}\right).$$

(b) *Differentiate the expression from (a) with respect to x to get the pdf of X, $f(x)$. Remember that $\Phi'(z) = \phi(z)$ and don't forget the chain rule*

Solution: Differentiating with respect to x, we have

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{dx}} F_X(x) = \frac{1}{3}\phi\left(\frac{x-1}{3}\right).$$

Since $\phi(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}$, we conclude

$$f_X(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2\cdot 3^2}},$$

which is the probability density function of the $N(1, 9)$ distribution. **Note:** The arguments in (a) and (b) give a proof that $3Z+1$ is a normal random variable with mean 1 and variance 9. See Problem Set 3, Question 5.

(c) *Find $P(-1 \leq X \leq 1)$*

Solution: We have

$$P(-1 \le X \le 1) = P\left(-\frac{2}{3} \le Z \le 0\right) = \Phi(0) - \Phi\left(-\frac{2}{3}\right) \approx 0.2475$$

(d) *Recall that the probability that Z is within one standard deviation of its mean is approximately 68%. What is the probability that X is within one standard deviation of its mean?*

Solution: Since $E[X] = 1$, $\text{Var}(X) = 9$, we want $P(-2 \le X \le 4)$. We have

$$P(-2 \le X \le 4) = P(-3 \le 3Z \le 3) = P(-1 \le Z \le 1) \approx 0.68.$$

Problem 48. Transforming Normal Distributions
Suppose $Z \sim N(0,1)$ and $Y = \mathrm{e}^Z$.

(a) *Find the cdf $F_Y(a)$ and pdf $f_Y(y)$ for Y. (For the CDF, the best you can do is write it in terms of Φ the standard normal cdf.)*

Solution: Note, Y follows what is called a *log-normal distribution.*

$F_Y(a) = P(Y \le a) = P(e^Z \le a) = P(Z \le \ln(a)) = \boxed{\Phi(\ln(a)).}$

Differentiating using the chain rule:

$$f_y(a) = \frac{d}{da}F_Y(a) = \frac{d}{da}\Phi(\ln(a)) = \frac{1}{a}\phi(\ln(a)) = \boxed{\frac{1}{\sqrt{2\pi}\,a}\mathrm{e}^{-(\ln(a))^2/2}.}$$

(b) *We don't have a formula for $\Phi(z)$ so we don't have a formula for quantiles. So we have to write quantiles* in terms *of Φ^{-1}.*

(i) Write the 0.33 quantile of Z in terms of Φ^{-1}

(ii) Write the 0.9 quantile of Y in terms of Φ^{-1}.

(iii) Find the median of Y.

Solution: (i) The 0.33 quantile for Z is the value $q_{0.33}$ such that $P(Z \le q_{0.33}) = 0.33$. That is, we want

$$\Phi(q_{0.33}) = 0.33 \;\Leftrightarrow\; \boxed{q_{0.33} = \Phi^{-1}(0.33)}.$$

(ii) We want to find $q_{0.9}$ where

$$F_Y(q_{0.9}) = 0.9 \;\Leftrightarrow\; \Phi(\ln(q_{0.9})) = 0.9 \;\Leftrightarrow\; \boxed{q_{0.9} = \mathrm{e}^{\Phi^{-1}(0.9)}}.$$

(iii) As in (ii) $q_{0.5} = \mathrm{e}^{\Phi^{-1}(0.5)} = \mathrm{e}^0 = \boxed{1}$.

Problem 49. (Random variables derived from normal random variables)
Let X_1, X_2, ...X_n be i.i.d. $N(0,1)$ random variables.

Let $Y_n = X_1^2 + ... + X_n^2$.

(a) *Use the formula $\text{Var}(X_j) = E[X_j^2] - E[X_j]^2$ to show $E[X_j^2] = 1$.*

Solution: $\text{Var}(X_j) = 1 = E[X_j^2] - E[X_j]^2 = E[X_j^2]$. QED

(b) *Set up an integral in x for computing $E[X_j^4]$.*

For 3 extra credit points, use integration by parts show $E[X_j^4] = 3$.

(If you don't do this, you can still use this result in part c.)

Solution: $E[X_j^4] = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} x^4 \mathrm{e}^{-x^2/2}\,dx.$

(Extra credit) By parts: let $u = x^3$, $v' = x\mathrm{e}^{-x^2/2} \;\Rightarrow\; u' = 3x^2$, $v = -\mathrm{e}^{-x^2/2}$

$E[X_j^4] = \dfrac{1}{\sqrt{2\pi}} \left[x^3 \mathrm{e}^{-x^2/2} \Big|_{infty}^{\infty} + \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} 3x^2 \mathrm{e}^{-x^2/2}\,dx \right]$

The first term is 0 and the second term is the formula for $3E[X_j^2] = 3$ (by part (a)). Thus, $E[X_j^4] = 3$.

(c) *Deduce from parts (a) and (b) that* $\mathrm{Var}(X_j^2) = 2$.

Solution: $\mathrm{Var}(X_j^2) = E[X_j^4] - E[X_j^2]^2 = 3 - 1 = 2$. QED

(d) *Use the Central Limit Theorem to approximate $P(Y_{100} > 110)$.*

Solution: $E[Y_{100}] = E[100X_j^2] = 100. \quad \mathrm{Var}(Y_{100}) = 100\mathrm{Var}(X_j) = 200.$

The CLT says Y_{100} is approximately normal. Standardizing gives

$P(Y_{100} > 110) = P\left(\dfrac{Y_{100} - 100}{\sqrt{200}} > \dfrac{10}{\sqrt{200}} \right) \approx P(Z > 1/\sqrt{2}) = \boxed{0.24}.$

This last value was computed using R: `1 - pnorm(1/sqrt(2),0,1)`.

Problem 50. More Transforming Normal Distributions

(a) *Suppose Z is a standard normal random variable and let $Y = aZ + b$, where $a > 0$ and b are constants.*

Show $Y \sim N(b, a^2)$ (remember our notation for normal distributions uses mean and variance).

Solution: Let $\phi(z)$ and $\Phi(z)$ be the PDF and CDF of Z.
$F_Y(y) = P(Y \le y) = P(aZ + b \le y) = P(Z \le (y-b)/a) = \Phi((y-b)/a)$.
Differentiating:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \Phi((y-b)/a) = \frac{1}{a}\phi((y-b)/a) = \frac{1}{\sqrt{2\pi}\,a}\,\mathrm{e}^{-(y-b)^2/2a^2}.$$

Since this is the density for $\mathrm{N}(b, a^2)$ we have shown $Y \sim \mathrm{N}(b, a^2)$.

(b) *Suppose $Y \sim N(\mu, \sigma^2)$. Show $\dfrac{Y - \mu}{\sigma}$ follows a standard normal distribution.*

Solution: By part (a), $Y \sim \mathrm{N}(\mu, \sigma^2) \;\Rightarrow\; Y = \sigma Z + \mu$. But, this implies $(Y - \mu)/\sigma = Z \sim \mathrm{N}(0,1)$. QED

Problem 51. (Sums of normal random variables)
Let X, Y be independent random variables where $X \sim N(2,5)$ and $Y \sim N(5,9)$ (we use the notation $N(\mu, \sigma^2)$). Let $W = 3X - 2Y + 1$.

(a) *Compute $E[W]$ and* $\mathrm{Var}(W)$.

Solution: $E[W] = 3E[X] - 2E[Y] + 1 = 6 - 10 + 1 = \boxed{-3}$

$\text{Var}(W) = 9\text{Var}(X) + 4\text{Var}(Y) = 45 + 36 = \boxed{81}$

(b) *It is known that the sum of independent normal distributions is normal. Compute $P(W \leq 6)$.*

Solution: Since the sum of independent normal is normal part (a) shows: $W \sim N(-3, 81)$.
Let $Z \sim N(0,1)$. We *standardize* W: $P(W \leq 6) = P\left(\frac{W+3}{9} \leq \frac{9}{9}\right) = P(Z \leq 1) \approx \boxed{0.84.}$

Problem 52. *Let $X \sim U(a,b)$. Compute $E[X]$ and $\text{Var}(X)$.*

Solution: Method 1

$U(a,b)$ has density $f(x) = \dfrac{1}{b-a}$ on $[a,b]$. So,

$$E[X] = \int_a^b x f(x)\,dx = \frac{1}{b-a}\int_a^b x\,dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)} = \boxed{\frac{a+b}{2}.}$$

$$E[X^2] = \int_a^b x^2 f(x)\,dx = \frac{1}{b-a}\int_a^b x^2\,dx = \left.\frac{x^3}{3(b-a)}\right|_a^b = \frac{b^3-a^3}{3(b-a)}.$$

Finding $\text{Var}(X)$ now requires a little algebra,

$$\begin{aligned}\text{Var}(X) = E[X^2] - E[X]^2 &= \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4} \\ &= \frac{4(b^3-a^3) - 3(b-a)(b+a)^2}{12(b-a)} = \frac{b^3 - 3ab^2 + 3a^2b - a^3}{12(b-a)} = \frac{(b-a)^3}{12(b-a)} = \boxed{\frac{(b-a)^2}{12}.}\end{aligned}$$

Method 2

There is an easier way to find $E[X]$ and $\text{Var}(X)$.

Let $U \sim \text{U}(a,b)$. Then the calculations above show $E[U] = 1/2$ and $(E[U^2] = 1/3 \Rightarrow \text{Var}(U) = 1/3 - 1/4 = 1/12$.

Now, we know $X = (b-a)U + a$, so $E[X] = (b-a)E[U] + a = (b-a)/2 + a = (b+a)/2$ and $\text{Var}(X) = (b-a)^2\text{Var}(U) = (b-a)^2/12$.

Problem 53. *In $n+m$ independent Bernoulli(p) trials, let S_n be the number of successes in the first n trials and T_m the number of successes in the last m trials.*

(a) *What is the distribution of S_n? Why?*

Solution: $S_n \sim \text{Binomial}(n,\, p)$, since it is the number of successes in n independent Bernoulli trials.

(b) *What is the distribution of T_m? Why?*

Solution: $T_m \sim \text{Binomial}(m,\, p)$, since it is the number of successes in m independent Bernoulli trials.

(c) *What is the distribution of $S_n + T_m$? Why?*

Solution: $S_n + T_m \sim \text{Binomial}(n+m,\, p)$, since it is the number of successes in $n+m$ independent Bernoulli trials.

(d) *Are S_n and T_m independent? Why?*

Solution: Yes, S_n and T_m are independent. We haven't given a formal definition of independent random variables yet. But, we know it means that knowing S_n gives no information about T_m. This is clear since the first n trials are independent of the last m.

Problem 54. *Compute the median for the exponential distribution with parameter λ.*

Solution: The density for this distribution is $f(x) = \lambda\, \mathrm{e}^{-\lambda x}$. We know (or can compute) that the distribution function is $F(a) = 1 - \mathrm{e}^{-\lambda a}$. The median is the value of a such that $F(a) = 0.5$. Thus, $1 - \mathrm{e}^{-\lambda a} = 0.5 \;\Rightarrow\; 0.5 = \mathrm{e}^{-\lambda a} \;\Rightarrow\; \log(0.5) = -\lambda a \;\Rightarrow\; \boxed{a = \log(2)/\lambda.}$

Problem 55. *Pareto and the 80-20 rule.*

Pareto was an economist who used the Pareto distribution to model the wealth in a society. For a fixed baseline m, the Pareto density with parameter α is

$$f(x) = \frac{\alpha\, m^\alpha}{x^{\alpha+1}} \qquad \textit{for } x \geq m.$$

Assume X is a random variable that follows such a distribution.

(a) *Compute $P(X > a)$ (you may assume $a \geq m$).*

Solution: $P(X > a) \;=\; \displaystyle\int_a^\infty \frac{\alpha\, m^\alpha}{x^{\alpha+1}} \;=\; -\frac{m^\alpha}{x^\alpha}\Big|_a^\infty \;=\; \boxed{\frac{m^\alpha}{a^\alpha}.}$

(b) *Pareto's principle is often paraphrased as the 80-20 rule. That is, 80% of the wealth is owned by 20% of the people. The rule is only exact for a Pareto distribution with $\alpha = \log(5)/\log(4) = 1.16$.*

Suppose $\alpha = m = 1$. Compute the 0.80 quantile for the Pareto distribution.

In general, many phenomena follow the power law described by $f(x)$. You can look up 'Pareto principle' in Wikipedia to read more about this.

Solution: We want the value $q_{0.8}$ where $P(X \leq q_{0.8}) = 0.8$.

This is equivalent to $P(X > q_{0.8}) = 0.2$. Using part (a) and the given values of m and α we have $\dfrac{1}{q_{0.8}} = 0.2 \;\Rightarrow\; \boxed{q_{0.8} = 5.}$

7 Joint Probability, Covariance, Correlation

Problem 56. (**Another Arithmetic Puzzle**)
Let X and Y be two independent Bernoulli(0.5) random variables. Define S and T by:

$$S = X + Y \quad \textit{and} \quad T = X - Y.$$

(a) *Find the joint and marginal pmf's for S and T.*

(b) *Are S and T independent.*

Solution: **(a)** $S = X + Y$ takes values 0, 1, 2 and $T = X - Y$ takes values -1, 0, 1.

First we make two tables: the joint probability table for X and Y and a table given the values (S, T) corresponding to values of (X, Y), e.g. $(X, Y) = (1, 1)$ corresponds to $(S, T) = (2, 0)$.

$X \backslash Y$	0	1
0	1/4	1/4
1	1/4	1/4

Joint probabilities of X and Y

$X \backslash Y$	0	1
0	0,0	1,-1
1	1,1	2,0

Values of (S, T) corresponding to X and Y

We can use the two tables above to write the joint probability table for S and T. The marginal probabilities are given in the table.

$S \backslash T$	-1	0	1	
0	0	1/4	0	1/4
1	1/4	0	1/4	1/2
2	0	1/4	0	1/4
	1/4	1/2	1/4	1

Joint and marginal probabilities of S and T

(b) No probabilities in the table are the product of the corresponding marginal probabilities. (This is easiest to see for the 0 entries.) So, S and T are not independent

Problem 57. *Data is taken on the height and shoe size of a sample of MIT students. Height is coded by 3 values: 1 (short), 2 (average), 3 (tall) and shoe size is coded by 3 values 1 (small), 2 (average), 3 (large). The joint counts are given in the following table.*

Shoe \ Height	1	2	3
1	234	225	84
2	180	453	161
3	39	192	157

Let X be the coded shoe size and Y the height of a random person in the sample.

***(a)** Find the joint and marginal pmf of X and Y.*

***(b)** Are X and Y independent?*

Solution: (a) The joint distribution is found by dividing each entry in the data table by the total number of people in the sample. Adding up all the entries we get 1725. So the joint probability table with marginals is

$Y \backslash X$	1	2	3	
1	$\frac{234}{1725}$	$\frac{225}{1725}$	$\frac{84}{1725}$	$\frac{543}{1725}$
2	$\frac{180}{1725}$	$\frac{453}{1725}$	$\frac{161}{1725}$	$\frac{794}{1725}$
3	$\frac{39}{1725}$	$\frac{192}{1725}$	$\frac{157}{1725}$	$\frac{388}{1725}$
	$\frac{453}{1725}$	$\frac{839}{1725}$	$\frac{433}{1725}$	1

The marginal distribution of X is at the right and of Y is at the bottom.

(b) X and Y are dependent because, for example,

$$P(X = 1 \text{ and } Y = 1) = \frac{234}{1725} \approx 0.136$$

is not equal to

$$P(X = 1)P(Y = 1) = \frac{453}{1725} \cdot \frac{543}{1725} \approx 0.083.$$

Problem 58. *Let X and Y be two continuous random variables with joint pdf*

$$f(x, y) = cx^2y(1 + y) \quad \textit{for } 0 \leq x \leq 3 \textit{ and } 0 \leq y \leq 3,$$

and $f(x, y) = 0$ otherwise.

(a) *Find the value of c.*

(b) *Find the probability $P(1 \leq X \leq 2,\, 0 \leq Y \leq 1)$.*

(c) *Determine the joint cdf, $F(a, b)$, of X and Y for a and b between 0 and 3.*

(d) *Find marginal cdf $F_X(a)$ for a between 0 and 3.*

(e) *Find the marginal pdf $f_X(x)$ directly from $f(x, y)$ and check that it is the derivative of $F_X(x)$.*

(f) *Are X and Y independent?*

Solution: **(a)** Total probability must be 1, so

$$1 = \int_0^3 \int_0^3 f(x, y)\, dy\, dx = \int_0^3 \int_0^3 c(x^2y + x^2y^2)\, dy\, dx = c \cdot \frac{243}{2},$$

(Here we skipped showing the arithmetic of the integration) Therefore, $c = \dfrac{2}{243}$.

(b)

$$\begin{aligned} P(1 \leq X \leq 2,\, 0 \leq Y \leq 1) &= \int_1^2 \int_0^1 f(x, y)\, dy\, dx \\ &= \int_1^2 \int_0^1 c(x^2y + x^2y^2)\, dy\, dx \\ &= c \cdot \frac{35}{18} \\ &= \frac{70}{4374} \approx 0.016 \end{aligned}$$

(c) For $0 \leq a \leq 3$ and $0 \leq b \leq 3$. we have

$$F(a, b) = \int_0^a \int_0^b f(x, y) dy\, dx = c\left(\frac{a^3b^2}{6} + \frac{a^3b^3}{9}\right)$$

(d) Since $y = 3$ is the maximum value for Y, we have

$$F_X(a) = F(a, 3) = c\left(\frac{9a^3}{6} + 3a^3\right) = \frac{9}{2}c\, a^3 = \frac{a^3}{27}$$

(e) For $0 \leq x \leq 3$, we have, by integrating over the entire range for y,

$$f_X(x) = \int_0^3 f(x, y)\, dy = cx^2\left(\frac{3^2}{2} + \frac{3^3}{3}\right) = c\frac{27}{2}x^2 = \frac{1}{9}x^2.$$

This is consistent with (c) because $\frac{d}{dx}(x^3/27) = x^2/9$.

(f) Since $f(x, y)$ separates into a product as a function of x times a function of y we know X and Y are independent.

Problem 59. *Let X and Y be two random variables and let r, s, t, and u be real numbers.*
(a) *Show that* $\text{Cov}(X + s, Y + u) = \text{Cov}(X, Y)$.

(b) *Show that* $\text{Cov}(rX, tY) = rt\text{Cov}(X, Y)$.

(c) *Show that* $\text{Cov}(rX + s, tY + u) = rt\text{Cov}(X, Y)$.

Solution: **(a)** First note by linearity of expectation we have $E[X + s] = E[X] + s$, thus $X + s - E[X + s] = X - E[X]$.
Likewise $Y + u - E[Y + u] = Y - E[Y]$.
Now using the definition of covariance we get

$$\begin{aligned}\text{Cov}(X + s, Y + u) &= E[(X + s - E[X + s]) \cdot (Y + u - E[Y + u])] \\ &= E[(X - E[X]) \cdot (Y - E[Y])] \\ &= \text{Cov}(X, Y).\end{aligned}$$

(b) This is very similar to part (a).

We know $E[rX] = rE[X]$, so $rX - E[rX] = r(X - E[X])$. Likewise $tY - E[tY] = s(Y - E[Y])$. Once again using the definition of covariance we get

$$\begin{aligned}\text{Cov}(rX, tY) &= E[(rX - E[rX])(tY - E[tY])] \\ &= E[rt(X - E[X])(Y - E[Y])] \\ &\qquad \text{(Now we use linearity of expectation to pull out the factor of } rt) \\ &= rtE[(X - E[X](Y - E[Y]))] \\ &= rt\text{Cov}(X, Y)\end{aligned}$$

(c) This is more of the same. We give the argument with far fewer algebraic details

$$\begin{aligned}\text{Cov}(rX + s, tY + u) &= \text{Cov}(rX, tY) \text{ (by part (a))} \\ &= rt\text{Cov}(X, Y) \text{ (by part (b))}\end{aligned}$$

Problem 60. *Derive the formula for the covariance:* $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$.

Solution: Using linearity of expectation, we have

$$\begin{aligned}\text{Cov}(X, Y) &= E\left[(X - E[X])(Y - E[Y])\right] \\ &= E\left[XY - E[X]Y - E[Y]X + E[X]E[Y]\right] \\ &= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y].\end{aligned}$$

Problem 61. (**Arithmetic Puzzle**)

The joint and marginal pmf's of X and Y are partly given in the following table.

$X \backslash Y$	1	2	3	
1	1/6	0	...	1/3
2	...	1/4	...	1/3
3	...	...	1/4	...
	1/6	1/3	...	1

(a) *Complete the table.*

(b) *Are X and Y independent?*

Solution: **(a)** The marginal probabilities have to add up to 1, so the two missing marginal probabilities can be computed: $P(X = 3) = 1/3$, $P(Y = 3) = 1/2$. Now each row and column has to add up to its respective margin. For example, $1/6 + 0 + P(X = 1, Y = 3) = 1/3$, so $P(X = 1, Y = 3) = 1/6$. Here is the completed table.

$X \backslash Y$	1	2	3	
1	1/6	0	1/6	1/3
2	0	1/4	1/12	1/3
3	0	1/12	1/4	1/3
	1/6	1/3	1/2	1

(b) No, X and Y are not independent.

For example, $P(X = 2, Y = 1) = 0 \neq P(X = 2) \cdot P(Y = 1)$.

Problem 62. (**Simple Joint Probability**)
Let X and Y each have range {1,2,3,4}. The following formula gives their joint pmf

$$P(X = i, Y = j) = \frac{i + j}{80}$$

Compute each of the following:

(a) $P(X = Y)$.

(b) $P(XY = 6)$.

(c) $P(1 \leq X \leq 2,\ 2 < Y \leq 4)$.

Solution: First we'll make the table for the joint pmf. Then we'll be able to answer the questions by summing up entries in the table.

$X \backslash Y$	1	2	3	4
1	2/80	3/80	4/80	5/80
2	3/80	4/80	5/80	6/80
3	4/80	5/80	6/80	7/80
4	5/80	6/80	7/80	8/80

(a) $P(X = Y) = p(1, 1) + p(2, 2) + p(3, 3) + p(4, 4) = \boxed{20/80 = 1/4.}$

(b) $P(XY = 6) = p(2, 3) + p(3, 2) = \boxed{10/80 = 1/8.}$

(c) $P(1 \le X \le 2, 2 < Y \le 4)$ = sum of 4 orange probabilities in the upper right corner of the table = $\boxed{20/80 = 1/4.}$

Problem 63. *Toss a fair coin 3 times. Let X = the number of heads on the first toss, Y the total number of heads on the last two tosses, and F the number of heads on the first two tosses.*

(a) *Give the joint probability table for X and Y. Compute* $\text{Cov}(X, Y)$.

Solution: (a) X and Y are independent, so the table is computed from the product of the known marginal probabilities. Since they are independent, $\text{Cov}(X, Y) = 0$.

$Y \backslash X$	0	1	P_Y
0	1/8	1/8	1/4
1	1/4	1/4	1/2
2	1/8	1/8	1/4
P_X	1/2	1/2	1

(b) *Give the joint probability table for X and F. Compute* $\text{Cov}(X, F)$.

Solution: (b) The sample space is Ω = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

$P(X = 0, F = 0) = P(\{TTH, TTT\}) = 1/4.$
$P(X = 0, F = 1) = P(\{THH, THT\}) = 1/4.$
$P(X = 0, F = 2) = 0.$
$P(X = 1, F = 0) = 0.$
$P(X = 1, F = 1) = P(\{HTH, HTT\}) = 1/4.$
$P(X = 1, F = 2) = P(\{HHH, HHT\}) = 1/4.$

$F \backslash X$	0	1	P_F
0	1/4	0	1/4
1	1/4	1/4	1/2
2	0	1/4	1/4
P_X	1/2	1/2	1

$\text{Cov}(X, F) = E[XF] - E[X]E[F]$.
$E[X] = 1/2, \quad E[F] = 1, \; E[XF] = \sum x_i y_j p(x_i, y_j) = 3/4.$
$\Rightarrow \; \text{Cov}(X, F) = 3/4 - 1/2 = \boxed{1/4.}$

Problem 64. Covariance and Independence
Let X be a random variable that takes values -2, -1, 0, 1, 2; each with probability 1/5. Let $Y = X^2$.

(a) *Fill out the following table giving the joint frequency function for X and Y. Be sure to include the marginal probabilities.*

X / Y	-2	-1	0	1	2	total
0						
1						
4						
total						

Solution:

Y \ X	-2	-1	0	1	2	
0	0	0	1/5	0	0	1/5
1	0	1/5	0	1/5	0	2/5
4	1/5	0	0	0	1/5	2/5
	1/5	1/5	1/5	1/5	1/5	1

Each column has only one nonzero value. For example, when $X = -2$ then $Y = 4$, so in the $X = -2$ column, only $P(X = -2, Y = 4)$ is not 0.

(b) *Find $E[X]$ and $E[Y]$.*

Solution: Using the marginal distributions: $E[X] = \frac{1}{5}(-2 - 1 + 0 + 1 + 2) = 0.$

$E[Y] = 0 \cdot \dfrac{1}{5} + 1 \cdot \dfrac{2}{5} + 4 \cdot \dfrac{2}{5} = 2.$

(c) *Show X and Y are not independent.*

Solution: We show the probabilities don't multiply:

$P(X = -2,\ Y = 0) = 0 \neq P(X = -2) \cdot P(Y = 0) = 1/25.$

Since these are not equal X and Y are not independent. (It is obvious that X^2 is not independent of X.)

(d) *Show $\text{Cov}(X, Y) = 0$.*

This is an example of uncorrelated but non-independent random variables. The reason this can happen is that correlation only measures the linear dependence between the two variables. In this case, X and Y are not at all linearly related.

Solution: Using the table from part (a) and the means computed in part (d) we get:

$$\begin{aligned}\text{Cov}(X, Y) &= E[XY] - E[X]E[Y] \\ &= \frac{1}{5}(-2)(4) + \frac{1}{5}(-1)(1) + \frac{1}{5}(0)(0) + \frac{1}{5}(1)(1) + \frac{1}{5}(2)(4) \\ &= 0.\end{aligned}$$

Problem 65. Continuous Joint Distributions

Suppose X and Y are continuous random variables with joint density function $f(x, y) = x+y$ on the unit square $[0, 1] \times [0, 1]$.

(a) *Let $F(x, y)$ be the joint CDF. Compute $F(1, 1)$. Compute $F(x, y)$.*

Solution: $F(a, b) = P(X \le a, Y \le b) = \displaystyle\int_0^a \int_0^b (x + y)\, dy\, dx.$

Inner integral: $\quad xy + \left.\dfrac{y^2}{2}\right|_0^b = xb + \dfrac{b^2}{2}.$

Outer integral: $\quad \left.\dfrac{x^2}{2}b + \dfrac{b^2}{2}x\right|_0^a = \dfrac{a^2 b + ab^2}{2}.$

So $\boxed{F(x, y) = \dfrac{x^2 y + xy^2}{2}}$ and $\boxed{F(1, 1) = 1.}$

(b) *Compute the marginal densities for X and Y.*

Solution: $f_X(x) = \int_0^1 f(x,y)\,dy = \int_0^1 (x+y)\,dy = xy + \left.\frac{y^2}{2}\right|_0^1 = \boxed{x + \frac{1}{2}.}$

By symmetry, $\boxed{f_Y(y) = y + 1/2.}$

(c) *Are X and Y independent?*

Solution: To see if they are independent we check if the joint density is the product of the marginal densities.

$$f(x,y) = x+y,\ f_X(x)\cdot f_Y(y) = (x+1/2)(y+1/2).$$

Since these are not equal, $\boxed{X \text{ and } Y \text{ are not independent.}}$

(d) *Compute $E[X]$, $E[Y]$, $E[X^2+Y^2]$, $\mathrm{Cov}(X,Y)$.*

Solution: $E[X] = \int_0^1\int_0^1 x(x+y)\,dy\,dx = \int_0^1 \left[\left. x^2y + x\frac{y^2}{2}\right|_0^1\right] dx = \int_0^1 x^2 + \frac{x}{2}\,dx = \boxed{\frac{7}{12}.}$

(Or, using (b), $E[X] = \int_0^1 x f_X(x)\,dx = \int_0^1 x(x+1/2)\,dx = 7/12.$)

By symmetry $\boxed{E[Y] = 7/12.}$

$E[X^2+Y^2] = \int_0^1\int_0^1 (x^2+y^2)(x+y)\,dy\,dx = \boxed{\frac{5}{6}.}$

$E[XY] = \int_0^1\int_0^1 xy(x+y)\,dy\,dx = \boxed{\frac{1}{3}.}$

$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{3} - \frac{49}{144} = \boxed{-\frac{1}{144}.}$

Problem 66. Correlation

Flip a coin 3 times. Use a joint pmf table to compute the covariance and correlation between the number of heads on the first 2 and the number of heads on the last 2 flips.

Solution: Let X = the number of heads on the first 2 flips and Y the number in the last 2. Considering all 8 possibe tosses: HHH, HHT etc we get the following joint pmf for X and Y

Y/X	0	1	2	
0	1/8	1/8	0	1/4
1	1/8	1/4	1/8	1/2
2	0	1/8	1/8	1/4
	1/4	1/2	1/4	1

Using the table we find

$$E[XY] = \frac{1}{4} + 2\frac{1}{8} + 2\frac{1}{8} + 4\frac{1}{8} = \frac{5}{4}.$$

We know $E[X] = 1 = E[Y]$ so

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = \frac{5}{4} - 1 = \frac{1}{4}.$$

Since X is the sum of 2 independent Bernoulli(0.5) we have $\sigma_X = \sqrt{2/4}$

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{1/4}{(2)/4} = \frac{1}{2}.$$

Problem 67. Correlation
Flip a coin 5 times. Use properties of covariance to compute the covariance and correlation between the number of heads on the first 3 and last 3 flips.

Solution: As usual let X_i = the number of heads on the i^{th} flip, i.e. 0 or 1.

Let $X = X_1 + X_2 + X_3$ the sum of the first 3 flips and $Y = X_3 + X_4 + X_5$ the sum of the last 3. Using the algebraic properties of covariance we have

$$\begin{aligned}
\text{Cov}(X, Y) &= \text{Cov}(X_1 + X_2 + X_3, X_3 + X_4 + X_5) \\
&= \text{Cov}(X_1, X_3) + \text{Cov}(X_1, X_4) + \text{Cov}(X_1, X_5) \\
&+ \text{Cov}(X_2, X_3) + \text{Cov}(X_2, X_4) + \text{Cov}(X_2, X_5) \\
&+ \text{Cov}(X_3, X_3) + \text{Cov}(X_3, X_4) + \text{Cov}(X_3, X_5)
\end{aligned}$$

Because the X_i are independent the only non-zero term in the above sum is $\text{Cov}(X_3 X_3) = \text{Var}(X_3) = \dfrac{1}{4}$ Therefore, $\text{Cov}(X, Y) = \frac{1}{4}$.

We get the correlation by dividing by the standard deviations. Since X is the sum of 3 independent Bernoulli(0.5) we have $\sigma_X = \sqrt{3/4}$

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{1/4}{(3)/4} = \frac{1}{3}.$$

8 Law of Large Numbers, Central Limit Theorem

Problem 68. (Table of normal probabilities)
Use the table of standard normal probabilities to compute the following. (Z is the standard normal.)

(a) *(i)* $P(Z \leq 1.5)$ *(ii)* $P(-1.5 < Z < 1.5)$ $P(Z > -0.75)$.

(b) *Suppose* $X \sim N(2, (0.5)^2)$. *Find (i)* $P(X \leq 2)$ *(ii)* $P(1 < X \leq 1.75)$.

Solution: (a) (i) 0.9332 (ii) 0.9332 - 0.0668 = 0.8664

(iii) By symmetry $= P(Z < 0.75) = 0.7734$. (Or we could have used $1 - P(Z > -0.75)$.)

(b) (i) Since 2 is the mean of the normal distribution, $P(X \leq 2) = 0.5$.

(ii) Standardizing,

$$P(1 < X \leq 1.75) = P\left(\frac{1-2}{0.5} < Z \leq \frac{1.75-2}{0.5}\right) = P(-2 < Z < -0.5) = 0.3085 - 0.0228 = \boxed{0.2857}.$$

Problem 69. *Suppose* $X_1, \ldots, X_{100}$ *are i.i.d. with mean 1/5 and variance 1/9. Use the central limit theorem to estimate* $P(\sum X_i < 30)$.

Solution: Standardize:

$$\begin{aligned} P\left(\sum_i X_i < 30\right) &= P\left(\frac{\sum X_i - \mu}{\sqrt{n}\,\sigma} < \frac{30 - n\mu}{\sqrt{n}\,\sigma}\right) \\ &\approx P\left(Z < \frac{30-20}{10/3}\right) \quad \text{(by the central limit theorem)} \\ &= P(Z < 3) \\ &= 0.9987 \quad \text{(from the table of normal probabilities)} \end{aligned}$$

Problem 70. All or None

You have \$100 and, never mind why, you must convert it to \$1000. Anything less is no good. Your only way to make money is to gamble for it. Your chance of winning one bet is p.

Here are two extreme strategies:

Maximum strategy: bet as much as you can each time. To be smart, if you have less than \$500 you bet it all. If you have more, you bet enough to get to \$1000.

Minimum strategy: bet \$1 each time.

If $p < 0.5$ (the odds are against you) which is the better strategy?

What about $p > 0.5$?

Solution: If $p < 0.5$ your expected winnings on any bet is negative, if $p = 0.5$ it is 0, and if $p > 0.5$ is is positive. By making a lot of bets the minimum strategy will 'win' you close to the expected average. So if $p \le 0.5$ you should use the maximum strategy and if $p > 0.5$ you should use the minumum strategy.

Problem 71. (Central Limit Theorem)

Let $X_1, X_2, \dots, X_{81}$ be i.i.d., each with expected value $\mu = E[X_i] = 5$, and variance $\sigma^2 = \mathrm{Var}(X_i) = 4$. Approximate $P(X_1 + X_2 + \cdots X_{81} > 369)$, using the central limit theorem.

Solution: Let $T = X_1 + X_2 + \ldots + X_{81}$. The central limit theorem says that

$$T \approx \mathrm{N}(81 * 5, 81 * 4) = \mathrm{N}(405, 18^2)$$

Standardizing we have

$$\begin{aligned} P(T > 369) &= P\left(\frac{T - 405}{18} > \frac{369 - 405}{18}\right) \\ &\approx P(Z > -2) \\ &\approx 0.975 \end{aligned}$$

The value of 0.975 comes from the rule-of-thumb that $P(|Z| < 2) \approx 0.95$. A more exact value (using R) is $P(Z > -2) \approx 0.9772$.

Problem 72. (Binomial $\approx$ normal)

Let $X \sim$ binomial(100,1/3).

An 'exact' computation in R gives $P(X \le 30) = 0.2765539$. Use the central limit theorem to give an approximation of $P(X \le 30)$

Solution: $X \sim$ binomial$(100, 1/3)$ means X is the sum of 100 i.i.d. Bernoulli$(1/3)$ random variables X_i.

We know $E[X_i] = 1/3$ and $\text{Var}(X_i) = (1/3)(2/3) = 2/9$. Therefore the central limit theorem says

$$X \approx \text{N}(100/3, 200/9)$$

Standardization then gives

$$P(X \leq 30) = P\left(\frac{X - 100/3}{\sqrt{200/9}} \leq \frac{30 - 100/3}{\sqrt{200/9}}\right) \approx P(Z \leq -0.7071068) \approx 0.239751$$

We used R to do these calculations The approximation agrees with the 'exact' number to 2 decimal places.

Problem 73. (**More Central Limit Theorem**)
The average IQ in a population is 100 with standard deviation 15 (by definition, IQ is normalized so this is the case). What is the probability that a randomly selected group of 100 people has an average IQ above 115?

Solution: Let X_j be the IQ of a randomly selected person. We are given $E[X_j] = 100$ and $\sigma_{X_j} = 15$.

Let $\overline{X}$ be the average of the IQ's of 100 randomly selected people. Then we know

$$E[\overline{X}] = 100 \quad \text{and} \quad \sigma_{\overline{X}} = 15/\sqrt{100} = 1.5.$$

The problem asks for $P(\overline{X} > 115)$. Standardizing we get $P(\overline{X} > 115) \approx P(Z > 10)$. $\boxed{\text{This is effectively 0.}}$

Problem 74. Hospitals (binomial, CLT, etc)

- *A certain town is served by two hospitals.*
- *Larger hospital: about 45 babies born each day.*
- *Smaller hospital about 15 babies born each day.*
- *For a period of 1 year, each hospital recorded the days on which more than 60% of the babies born were boys.*

(a) *Which hospital do you think recorded more such days?*

(i) The larger hospital.
(ii) The smaller hospital.
(iii) About the same (that is, within 5% of each other).

(b) *Let L_i (resp., S_i) be the Bernoulli random variable which takes the value 1 if more than 60% of the babies born in the larger (resp., smaller) hospital on the i^{th} day were boys. Determine the distribution of L_i and of S_i.*

(c) *Let L (resp., S) be the number of days on which more than 60% of the babies born in the larger (resp., smaller) hospital were boys. What type of distribution do L and S have? Compute the expected value and variance in each case.*

(d) *Via the CLT, approximate the 0.84 quantile of L (resp., S). Would you like to revise your answer to part (a)?*

(e) *What is the correlation of L and S? What is the joint pmf of L and S? Visualize the region corresponding to the event $L > S$. Express $P(L > S)$ as a double sum.*

(a) Solution: When this question was asked in a study, the number of undergraduates who chose each option was 21, 21, and 55, respectively. This shows a lack of intuition for the relevance of sample size on deviation from the true mean (i.e., variance).

(b) Solution: The random variable X_L, giving the number of boys born in the larger hospital on day i, is governed by a $\text{Bin}(45, 0.5)$ distribution. So L_i has a $\text{Ber}(p_L)$ distribution with

$$p_L = P(X > 27) = \sum_{k=28}^{45} \left(\begin{array}{c} 45 \\ k \end{array} \right) 0.5^{45} \approx 0.068$$

Similarly, the random variable X_S, giving the number of boys born in the smaller hospital on day i, is governed by a $\text{Bin}(15, 0.5)$ distribution. So S_i has a $\text{Ber}(p_S)$ distribution with

$$p_S = P(X_S > 9) = \sum_{k=10}^{15} \left(\begin{array}{c} 15 \\ k \end{array} \right) 0.5^{15} \approx 0.151$$

We see that p_S is indeed greater than p_L, consistent with (ii).

(c) Solution: Note that $L = \sum_{i=1}^{365} L_i$ and $S = \sum_{i=1}^{365} S_i$. So L has a $\text{Bin}(365, p_L)$ distribution and S has a $\text{Bin}(365, p_S)$ distribution. Thus

$$\begin{aligned} E[L] &= 365 p_L \approx 25 \\ E[S] &= 365 p_S \approx 55 \\ \text{Var}(L) &= 365 p_L (1 - p_L) \approx 23 \\ \text{Var}(S) &= 365 p_S (1 - p_S) \approx 47 \end{aligned}$$

(d) Solution: mean + sd in each case:

For L, $q_{0.84} \approx 25 + \sqrt{23}$.

For S, $q_{0.84} \approx 55 + \sqrt{47}$.

(e) Since L and S are independent, their joint distribution is determined by multiplying their individual distributions. Both L and S are binomial with $n = 365$ and p_L and p_S computed above. Thus

$$p_{l,s} P(L = i \text{ and } S = j) = p(i,j) = \binom{365}{i} p_L^i (1 - p_L)^{365-i} \binom{365}{j} p_S^j (1 - p_S)^{365-j}$$

Thus

$$P(L > S) = \sum_{i=0}^{364} \sum_{j=i+1}^{365} p(i,j) \approx 0.0000916$$

(We used R to do the computations.)

9 R Problems

R will not be on the exam. However, these problems will help you understand the concepts we've been studying.

Problem 75. R simulation
Consider X_1, X_2, ...all independent and with distribution $N(0,1)$. Let $\overline{X}_m$ be the average of X_1, ...X_n.

(a) *Give $E[\overline{X}_n]$ and $\sigma_{\overline{X}_n}$ exactly.*

Solution: $E[X_j] = 0 \;\Rightarrow\; E[\overline{X}_n] = 0$.

$$\text{Var}(X_j) = 1 \;\Rightarrow\; \text{Var}\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{1}{n} \;\Rightarrow\; \sigma_{\overline{X}_n} = \frac{1}{\sqrt{n}}.$$

(b) *Use a R simulation to estimate $E[\overline{X}_n]$ and $\text{Var}(\overline{X}_n)$ for $n = 1, 9, 100$. (You should use the `rnorm` function to simulate 1000 samples of each X_j.)*

Solution: Here's my R code:

```
x = rnorm(100*1000,0,1)
data = matrix(x, nrow=100, ncol=1000)
data1 = data[1,]
m1 = mean(data1)
v1 = var(data1)
data9 = colMeans(data[1:9,])
m9 = mean(data9)
v9 = var(data9)
data100 = colMeans(data)
m100 = mean(data100)
v100 = var(data100)
#display the results
print(m1)
print(v1)
print(m9)
print(v9)
print(m100)
print(v100)
```

Note if $x = [x_1, x_2, \ldots, x_n]$ then `var(x)` actually computes $\dfrac{\sum x_k}{n-1}$ instead of $\dfrac{\sum x_k}{n}$. There is a good reason for this which we will learn in the statistics part of the class. For now, it's enough to note that if $n = 1000$ the using n or $n-1$ won't make much difference.

Problem 76. R Exercise
Let X_1, X_2, X_3, X_4, X_5 be independent $U(0,1)$ random variables.

Let $X = X_1 + X_2 + X_3$ and $Y = X_3 + X_4 + X_5$.

Use the `runif()` function to simulate 1000 trials of each of these variables. Use these to estimate $\text{Cov}(X,Y)$.

Solution: `a = runif(5*1000,0,1)`

```
data = matrix(a,5,1000)
x = colSums(data[1:3,])
y = colSums(data[3:5,])
print(cov(x,y))
```

Extra Credit
Compute this covariance exactly.

Solution: Method 1 (Algebra)
First, if $i \neq j$ we know X_i and X_j are independent, so $\text{Cov}(X_i, X_j) = 0$.

$$\begin{aligned}
\text{Cov}(X,Y) &= \text{Cov}(X_1 + X_2 + X_3, X_3 + X_4 + X_5) \\
&= \text{Cov}(X_1, X_3) + \text{Cov}(X_1, X_4) + \text{Cov}(X_1, X_5) \\
&\qquad + \text{Cov}(X_2, X_3) + \text{Cov}(X_2, X_4) + \text{Cov}(X_2, X_5) \\
&\qquad + \text{Cov}(X_3, X_3) + \text{Cov}(X_3, X_4) + \text{Cov}(X_3, X_5) \\
&\quad \text{(most of these terms are 0)} \\
&= \text{Cov}(X_3, X_3) \\
&= \text{Var}(X_3) \\
&= \frac{1}{12} \qquad \text{(known variance of a uniform(0,1) distribution)}
\end{aligned}$$

Method 2 (Multivariable calculus)
In 5 dimensional space we have the joint distribution

$$f(x_1, x_2, x_3, x_4, x_5) = 1.$$

Computing directly

$$\begin{aligned}
E[X] = E[X_1 + X_2 + X_3] &= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \int_0^1 (x_1 + x_2 + x_3)\, dx_1\, dx_2\, dx_3, dx_4\, dx_5 \\
\text{first integral} &= \frac{1}{2} + x_2 + x_3 \\
\text{second integral} &= \frac{1}{2} + \frac{1}{2} + x_3 = 1 + x_3 \\
\text{third integral} &= \frac{3}{2} \\
\text{fourth integral} &= \frac{3}{2} \\
\text{fifth integral} &= \frac{3}{2}
\end{aligned}$$

So, $E[X] = 3/2$, likewise $E[Y] = 3/2$.

$$\begin{aligned}
E[XY] &= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \int_0^1 (x_1 + x_2 + x_3)(x_3 + x_4 + x_5)\, dx_1\, dx_2\, dx_3\, dx_4\, dx_5 \\
&= 7/3.
\end{aligned}$$

$\text{Cov}(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{12} = 0.08333.$