Contents

1 Probability

1.1 Probability

1. Fill in the details of the proof of Theorem 1.8. Also, prove the monotone decreasing case.

   - For the readers' convenience we restate the Continuity of Probabilities theorem. If $A_n \to A$ then $P(A_n) \to P(A)$ as $n \to \infty$. Here $A_n \to A_n$ means that either A_n is monotone increasing ($A_n \subseteq A_{n+1}$) and we define $A = \bigcup_{n=1}^{\infty} A_n$, or, A_n is monotone decreasing ($A_n \supseteq A_{n+1}$) and we define $A = \bigcap_{n=1}^{\infty} A_n$.
   - We fill in the details now. First of all we want to show that $B_i \cap B_j = \emptyset$ for all $i \neq j$. Suppose without loss of generality that $i < j$ and note that $B_i \subseteq A_i$ then
   $$B_j = A_j \backslash \bigcup_{k=1}^{j-1} A_k$$
   and $B_i \subseteq A_i \subseteq \bigcup_{k=1}^{j-1}$ as such $B_i \cap B_j = \emptyset$.
   - To see that $A_n = \bigcup_{i=1}^{n} B_i$ let $x \in A_n$ then there exists a minimal $k = k(x)$ such that $x \in A_k$, i.e., for all $k' < k : x \notin A_{k'}$. Then $x \notin \bigcup_{i=1}^{k-1} A_i$ and therefore $x \in B_k$. Because x are arbitrary it folows that $A_n \subseteq \bigcup_{i=1}^{n} B_i$. On the other hand
   $$\bigcup_{i=1}^{n} \underbrace{B_i}_{\subseteq A_i} \subseteq \bigcup_{i=1}^{n} A_i = A_n,$$
   where we have used that A_n is monotone increasing. The property that
   $$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$
   holds is identical to the finite case, an element is part of the (countably) infinite union if there exists some minimal i such that ...
   - For the monotone decreasing case we instead want to define $B_n := A_n \backslash \bigcup_{i>n} A_i$.

2. Prove the statements in equation (1.1).

   - This can immediately be seen by noting that $A \cup B = (A \backslash B) \cup (A \cap B) \cup (B \backslash A)$ is a disjoint union.

3. Let Ω be a sample space and let $A_1, A_2, \ldots,$ be events. Define $B_n = \bigcup_{i=n}^{\infty} A_i$ and $C_n = \bigcap_{i=n}^{\infty} A_i$

   (a) Show that (B_i) is monotone decreasing and that (C_i) is monotone increasing.

(b) Show that $\omega \in \bigcap_{n=1}^{\infty} B_n$ if and only if ω belongs to an infinite number of events $A_1, A_2, \ldots$.

- Let $\omega \in \bigcap_{n=1}^{\infty} B_n$ be such that ω does not belong to an infinite number of events $A_1, A_2, \ldots$, i.e., there exists $N \in \mathbb{N}$ such that for all $k > N$ it follows that $\omega \notin A_k$. But then $\omega \notin B_k$ for all $k > N$ and as such does not lie in the intersection over all the B_k, which is a contradiction.
- Suppose ω lies in infinitely many $A_1, A_2, \ldots$. Then there exists a sequence $(A_{n_i})_{i \in \mathbb{N}}$ such that $n_i < n_j$ for all $i < j$ and such that $\omega \in A_{n_i}$ for all $i \in \mathbb{N}$. In that case

$$\omega \in B_{n_i}$$

for all $i \in \mathbb{N}$, in particular $\omega \in \cap_{i=\mathbb{N}} B_{n_i}$. Because B_n is monotone decreasing the statement follows.

(c) Show that $\omega \in \bigcup_{n=1}^{\infty} C_n$ if and only if ω belongs to all the events $A_1, A_2, \ldots$ except possibly a finite number of those events.

- Let $\omega \in \bigcup_{n=1}^{\infty} B_n$ then there exists $n \geq 1$ such that $\omega \in B_n$. This means that $\omega \in \bigcap_{i=n}^{\infty}$, i.e., ω belongs to all A_i where $i \geq n$.
- Now suppose ω belongs to all events $A_1, A_2, \ldots$ except for a finite number of events. In that case there exist $N \in \mathbb{N}$ such that $\forall k \geq N : \omega \in A_k$. This means that

$$\omega \in \bigcap_{i \geq k} A_i,$$

i.e., $\omega \in C_k$. In particular this shows

$$\omega \in C_k \subseteq \bigcup_{r \geq 1} C_r.$$

4. Let $\{A_i : i \in I\}$ be a collection of events where I is an arbitrary index set. Show that

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c, \quad \left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c$$

holds.

Hint: First prove this for $I = \{1, 2, \ldots, n\}$.

- I guess they want us to show it using induction in the hint, I'll just do it directly.

$$\begin{aligned}
\bigcap_i A_i^c &= \{x : \forall i : x \notin A_i\} \\
&= \{x : \neg (\exists i : x \in A_i)\} \\
&= \{x : \exists i : x \in A_i\}^c = \left(\bigcup_{i \in I} A_i \right)^c
\end{aligned}$$

The other direction is analogous.

5. Suppose we toss a fair coin until we get exactly two heads. Describe the sample space S. What is the probability that exactly k tosses are required?

6. Let $\Omega = \{0, 1, \dots\}$. Prove that there does not exist a uniform distribution on Ω (i.e., if $P(A) = P(B)$ whenever $|A| = |B|$, then P cannot satisfy the axioms of probability.)

   - Note that $P(\{0\}) \neq 0$ because otherwise $P = 0$, which would mean P is not a probability function. But then it follows for all $n \in \mathbb{N}$:

   $$P(\{0, \dots, n\}) = P\left(\bigcup_{k=0}^{n} \{k\}\right) = \sum_{k=0}^{n} P(\{k\}) = (n+1)P(\{0\}) \to \infty$$

   for $n \to \infty$, which is a contradiction to $P(\Omega) = +1$.

7. Let $A_1, A_2, \dots$ be events. Show that

   $$\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n).$$

   Hint: Define $B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$. Then show that the B_n are disjoint and that

   $$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n.$$

   - This was basically already shown in the first exercise, the only thing we mention here is that we only have to consider the case where the series on the right hand side converges, because the left hand side is always ≤ 1.

8. Suppose that $P(A_i) = 1$ for each i. Prove that

   $$P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1.$$

   - Note that $P(A_i^c) = 1 - P(A_i) = 0$ and that

   $$P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1 - P\left(\bigcup_{i=1}^{\infty} A_i^c\right).$$

   The latter can be calculated as follows:

   $$\begin{aligned} 0 &\leq P\left(\bigcup_{i=1}^{\infty} A_i^c\right) \\ &\leq \sum_{i=1}^{\infty} P(A_i^c) = 0 \end{aligned}$$

where we have used that probability measures are subadditive. As such $P\left(\bigcup_{i=1}^{\infty} A_i^c\right) = 0$ and therefore

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1.$$

9. For fixed B such that $P(B) > 0$, show that $P(\cdot|B)$ satisfies the axioms of probability.

Axiom 1: Let A be arbitrary then $P(A|B) = \frac{P(A,B)}{P(B)} \geq P(A,B) \geq 0$.

Axiom 2: $P(\Omega|B) = \frac{P(\Omega,B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

Axiom 3:

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) &= \frac{1}{P(B)} P\left(\bigcup_{i=1}^{\infty} A_i,\right) \\ &= \frac{1}{P(B)} \sum_{i=1}^{\infty} P(A_i, B) \\ &= \sum_{i=1}^{\infty} P(A_i|B). \end{aligned}$$

10. You have probably heard it before. Now you can solve it rigorously. It is called the "Monty Hall Problem." A prize is placed at random behind one of three doors. You pick a door. To be concrete, let's suppose you always pick door 1. Now Monty Hall chooses one of the other two doors, opens it and shows you that it is empty. He then gives you the opportunity to keep your door or switch to the other unopened door.Should you stay or switch? Intuition suggests it doesn't matter. The correct answer is that you should switch. Prove it. It will help to specify he sample space and the relevant events carefully. Thus write $\Omega = \{(\omega_1, \omega_2) : \omega_i \in \{1, 2, 3\}\}$ where ω_1 is where the prize is and ω_2 is the door Monty opens.

    - Note that $\{(1,1),(2,2),(3,3)\}$ are invalid because he'll never open the door with the price. The staying strategy wins on $\{(1,2),(1,3)\}$, the switching strategy wins on $\{(2,3),(3,2)\}$. Note that this means, if $\omega_1 \neq 1$, then we are guaranteed to win. As such we win whenever $\omega_1 \in \{2,3\}$, as such the winning probability is 2/3 for the switching strategy.
    - This is a pretty subtle problem, the intuition is that him opening a door does not grant new information. If $\omega_1 = 1$ then his reveal is arbitrary and we lose on switching. If $\omega_1 = 2, 3$ then he's forced to reveal the nonempty door, as such either the remaining door is the price or $\omega_1 = 1$. As such when we switch we are guaranteed to win whenever $\omega_1 = 2, 3$.

11. Suppose that A and B are independent events. Show that A^c and B^c are independent events.

   - $$\begin{aligned} P(A^c, B^c) &= P((A \cup B)^c) \\ &= 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A, B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= (1 - P(A))(1 - P(B)) \\ &= P(A^c)P(B^C). \end{aligned}$$

12. There are three cards. The first is green on both sides, the second is red on both sides and the third is green on one side and red on the other. We choose a card at random and we see one side (also chosen at random) . If the side we see is green, what is the probability that the other side is also green? Many people intuitively answer $\frac{1}{2}$. Show that the correct answer is $\frac{2}{3}$.

13. Suppose that a fair coin is tossed repeatedly until both a head and tail have appeared at least once.

   (a) Describe the sample space Ω.

   (b) What is the probability that three tosses will be required?

14. Show that if $P(A) = 0$ or $p(A) = 1$ then A is independent of every other event. Show that if A is independent of itself then $P(A)$ is either 0 or 1.

   - $P(A) = P(A, A) = P(A)P(A) = P(A)^2 \iff P(A)(P(A) - 1)$.

15. The probability that a child has blue eyes is $1/4$. assume independence between children. Consider a family with 3 children.

   (a) If it is known that at least one child has blue eyes, what ist he probability that at least two children have blue eyes?

      - Assuming that having blue eyes is independent from the other children then the chance that none of the siblings have blue eyes is $\left(\frac{3}{4}\right)^2 = \frac{9}{16}$. As such the probability is $7/16$.

   (b) If it is known that the youngest child has blue eyes, what is the probability that at least two children have blue eyes?

      - While this doesn't seem intuitive this actually changes the probabilities. To see this note that in the first problem the arrangements of having at least 2 blue eyed children were BBB, BBG, BGB, GBB where the children are ordered by age. But now we disregard the case GBB as such the probability becomes smaller.
      - The probability is then $\left(\frac{1}{4}\right)^2 + \frac{1}{4}\frac{3}{4} + \frac{3}{4}\frac{1}{4}$

16. Prove Lemma 1.14.

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

   - $$\begin{aligned} P(AB) &= P(A,B)\frac{P(B)}{P(B)} \\ &= P(A|B)P(B) \\ &= P(B|A)\frac{P(A)}{P(B)}P(B) \\ &= P(B|A)P(A). \end{aligned}$$

17. Show that
$$P(A,B,C) = P(A|BC)P(B|C)P(C).$$

18. Suppose k events form a partition of the sample space Ω, i.e., they are disjoint and
$$\bigcup_{i=1}^{k} A_i = \Omega.$$
Assume that $P(B) > O$. Prove that if $P(A_1|B) < P(A_1)$ then $P(A_i|B) > P(A_i)$ for some $i = 2, \ldots, k$.

   - Suppose this statement is wrong, i.e. $P(A_1|B) < P(A_1)$ and for all $i = 2, \ldots, k$ we have $P(A_i|B) \leq P(A_i)$. It then follows that
$$\begin{aligned} \sum_{i=1}^{k} P(A_i|B) &= \underbrace{P(A_1|B)}_{<P(A_1)} + \sum_{i=2}^{k} \underbrace{P(A_i|B)}_{\leq P(A_i)} \\ &< P(A_1) + \sum_{i=1}^{k} P(A_i) = 1. \end{aligned}$$
Multiplying $\sum_{i=1}^{n} P(A_i|B) < 1$ by $P(B) > 0$ yields
$$\sum_{i=1}^{n} P(A_i|B)P(B) < P(B).$$
This is a contradiction because $\sum_{i=1}^{n} P(A_i|B)P(B) = \sum_{i=1}^{n} P(A_i, B) = P\left(\bigcup_{i=1}^{n} A_i, B\right) = P(\Omega, B) = P(B)$.

19. Suppose that 30% of computer owners use a Macintosh, 50% use Windows, and 20% use Linux. Suppose that 65% of the Mac users have succumbed to a computer virus, 82% of the Windows users get the virus, and 50% of the Linux users get the virus. We select a person at random and learn that her system was infected with the virus. What is the probability that she is a Windows user?

- Denote the operating systems by M, W, L and having the virus by V. Using this notation we obtain $P(M) = 0.3$, $P(W) = 0.5$ and $P(L) = 0.2$. Furthermore $P(V|M) = 0.65$, $P(V|W) = 0.82$ and $P(V|L) = 0.5$ We can use the law of total probability to calculate the probability of a random person having a virus as being

$$\begin{aligned} P(V) &= P(V, M) + P(V, W) + P(V, L) \\ &= P(V|M)P(M) + P(V|W)P(W) + P(V|L)P(L) \\ &= 0.65 \cdot 0.3 + 0.82 \cdot 0.5 + 0.5 \cdot 0.2 \\ &= 0.705 = 70.5\%. \end{aligned}$$

  Now we can use Bayes' Theorem to calculate

$$P(W|V) = P(V|W)\frac{P(W)}{P(V)} = 0.82\frac{0.5}{0.705} = 0.5816 = 58.16\%$$

20. A box contains 5 coins and each has a different probability of showing heads. Let P_1, P_2, P_3, P_4, P_5 denote the probability of heads on each coin. Suppose that

$$p_1 = 0, \quad p_2 = 1/4, \quad p_3 = 1/2, \quad p_4 = 3/4, \quad p_5 = 1.$$

    Let H denote "heads is obtained" and let C_i denote the event that coin i is selected.

    (a) Select a coin at random and toss it. Suppose a head is obtained. What is the posterior probability that coin i was selected ($i = 1, ..., 5$)? In other words, find $P(C_i|H)$ for $i = 1, ..., 5$.

    - Note that $P(C_i) = \frac{1}{5}$ and $P(H|C_i) = p_i$. The total probability of gettings heads is

$$\begin{aligned} P(H) &= \sum_{i=1}^{5} P(H|C_i)P(C_i) \\ &= \frac{1}{5}\sum_{i=1}^{5} P(H|C_i) \\ &= \frac{1}{5}\sum_{i=1}^{5} p_i \\ &= \frac{1}{5}(0 + 1/4 + 1/2 + 3/4 + 1) \\ &= 0.5 = 50\%. \end{aligned}$$

We can use Bayes' theorem to calculate

$$\begin{aligned}P(C_i|H) &= P(H|C_i)\frac{P(C_i)}{P(H)}\\ &= P(H|C_i)\frac{\frac{1}{5}}{0.5}\\ &= P(H|C_i)\frac{2}{5}.\end{aligned}$$

Using this we obtain

i	1	2	3	4	5
$P(C_i\|H)$	0	1/10	1/5	3/10	2/5

(b) Toss the coin again. What is the probability of another head? In other words find $P(H_2|H_1)$ where H_j = "heads on toss j."

- Suppose we have chosen the *ith* coin then the probability of getting heads is $P(H|C_i)$. The probabilities of the individual coins have been calculated before and we just have to take the weighted sum over those.

```
import numpy as np

mC = np.array([0, 1/10, 1/5, 3/10, 2/5])
mP = np.array([0, 0.25, 0.5, 0.75, 1])
np.dot(mC, mP)

> 0.75
```

(c) Find $P(C_i|B_4)$ where B_4 = "first head is obtained on toss 4."

- First note that $P(C_i) = P(C_i, H_1) + P(C_i, H_1^c)$ as such

$$P(C_i, H_1^c) = P(C_i) - P(C_i, H_1) = P(C_i) - P(C_i|H_1)P(H_1) = \frac{1}{5} - P(C_i|H_1)\frac{1}{2}$$

as such

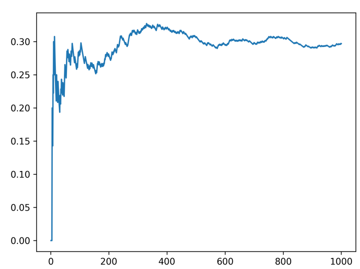
i	1	2	3	4	5
$P(C_i\|H_i^c)$	0.2	0.15	0.1	0.05	0

$P(H_1, H_2) = P(H_2|H_1)P(H_1) = \frac{3}{4}$

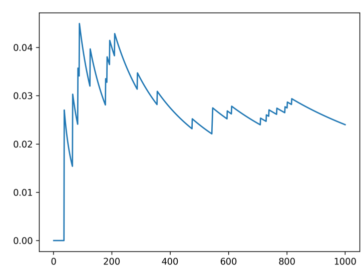
21. (Computer Experiment.) Suppose a coin has probability p of falling heads up. If we flip the coin many times, we would expect the proportion of heads to be near p. We will make this formal later. Take $p = 0.3$ and $n = 1000$ and simulate n coin flips. Plot the proportion of heads as a function of n. Repeat for $p = 0.03$.

- ```
  import numpy as np
  import matplotlib.pyplot as plt
  import math
  ```
  ```

```python
def Ex1_21(p = 0.3, n = 1000, draw = True, save = False):
    rand = np.random.random(size = 1000)
    _range = np.arange(1, n + 1)

    result = np.cumsum(rand < p) / _range

    fig, ax = plt.subplots()

    ax.plot(_range, result)

    if draw:
        plt.show()

    if save:
        # p = 0.213, n = 1000 -> 21pct_1000
        str_form = f"{math.floor(p * 100)}pct_{n}"
        fig.savefig(f"Ex1_21-{str_form}.png")

Ex1_21(p = 0.3, draw = True, save = True)
Ex1_21(p = 0.03, draw = True, save = True)
```



(a) 30% Probability



(b) 3% Probability

22. (Computer Experiment.) Suppose we flip a coin n times and let P denote the probability of heads. Let X be the number of heads. We call X a binomial random variable, which is discussed in the next chapter. Intuition suggests that X will be close to n p. To see if this is true, we can repeat this experiment many times and average the X values. Carry out a simulation and compare the average of the X's to n p . Try this for $p = 0.3$ and $n = 10$, $n = 100$, and $n = 1000$.

- ```python
  import numpy as np

  def Ex1_22(p = 0.3, n = 1000, num_iter = 100):
      sum = 0
      for i in range(num_iter):
          rand = np.random.random(size = n)
  ```
  ```

```
        sum += np.sum(rand < p)

    mean = sum / num_iter
    expected = n * p
    delta = abs(mean - expected) / (expected)
    print(f"mean = {mean}, delta = {delta * 100:.2f}%")

for n in [10, 100, 1000]:
    Ex1_22(n = n)

> mean = 2.93, delta = 2.33%
> mean = 30.2, delta = 0.67%
> mean = 297.83, delta = 0.72%
```

23. (Computer Experiment.) Here we will get some experience simulating conditional probabilities. Consider tossing a fair die. Let $A = \{2, 4, 6\}$ and $B = \{1, 2, 3, 4\}$. Then, $P(A) = 1/2$, $P(B) = 2/3$ and $P(AB) = 1/3$. Since $P(AB) = P(A)P(B)$, the events A and B are independent. Simulate draws from the sample space and verify that $\hat{P}(AB) = \hat{P}(A)\hat{P}(B)$ where $\hat{P}(A)$ is the proportion of times A occurred in the simulation and similarly for $\hat{P}(AB)$ and $\hat{P}(B)$. Now find two events A and B that are not independent. Compute $\hat{P}(A)$,$\hat{P}(B)$ and $\hat{P}(AB)$. Compare the calculated values to their theoretical values. Report your results and interpret.

- ```
  import numpy as np

  labels = ['P(A)', 'P(B)', 'P(A, B)', 'P(A)P(B)',
            'P(A, B) - P(A)P(B)']
  pct_str = lambda x : f"{x * 100:.2f}%"

  def Ex1_23_helper(probabilities):
      for label, prob in zip(labels, probabilities):
          print(f"\t{label}: {pct_str(prob)}")

  def Ex1_23(A, B, n = 10000):
      cap = A.intersection(B)

      # rand does [) so I need to increment high to get 6
      rand = np.random.randint(low = 1, high = 6 + 1, size = n)
      prob_A = sum(np.isin(rand, list(A))) / n
      prob_B = sum(np.isin(rand, list(B))) / n
      prob_cap = sum(np.isin(rand, list(cap))) / n

      prob_prod = prob_A * prob_B

      probabilities = [prob_A, prob_B, prob_cap, prob_prod,
                       prob_cap - prob_prod]
  ```
  ```

```
    print("Experimental Values")
    Ex1_23_helper(probabilities)

A, B = {2, 4, 6}, {1, 2, 3, 4}
print(f"A = {A}, B = {B}, AB = {A.intersection(B)}")
print("Theoretical Values")
Ex1_23_helper([1/2, 2/3, 1/3, 1/3, 0])
Ex1_23(A, B)
print()

A, B = {1, 2, 3}, {3, 4, 5, 6}
print(f"A = {A}, B = {B}, AB = {A.intersection(B)}")
print("Theoretical Values")
Ex1_23_helper([1/2, 2/3, 1/6, 1/3, -1/6])
Ex1_23(A, B)

> A = {2, 4, 6}, B = {1, 2, 3, 4}, AB = {2, 4}
> Theoretical Values
>         P(A): 50.00%
>         P(B): 66.67%
>         P(A, B): 33.33%
>         P(A)P(B): 33.33%
>         P(A, B) - P(A)P(B): 0.00%
> Experimental Values
>         P(A): 50.18%
>         P(B): 67.42%
>         P(A, B): 33.70%
>         P(A)P(B): 33.83%
>         P(A, B) - P(A)P(B): -0.13%
>
> A = {1, 2, 3}, B = {3, 4, 5, 6}, AB = {3}
> Theoretical Values
>         P(A): 50.00%
>         P(B): 66.67%
>         P(A, B): 16.67%
>         P(A)P(B): 33.33%
>         P(A, B) - P(A)P(B): -16.67%
> Experimental Values
>         P(A): 50.20%
>         P(B): 66.30%
>         P(A, B): 16.50%
>         P(A)P(B): 33.28%
>         P(A, B) - P(A)P(B): -16.78%
```

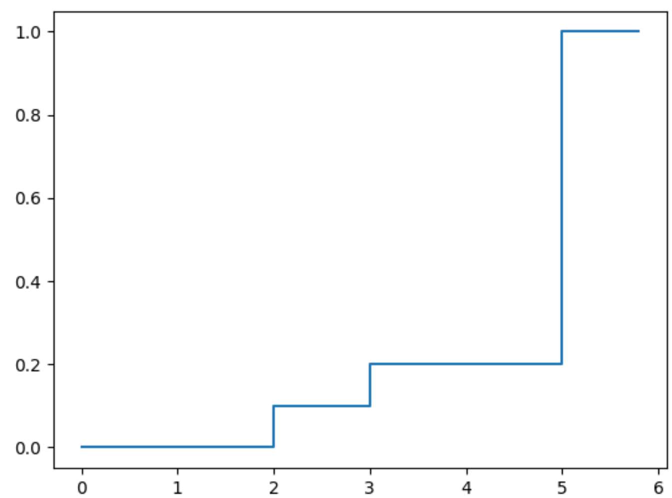
1.2 Random Variables

1. Show that
$$P(X = x) = F(x^+) - F(x^-).$$

   - Note that (∞, x) and $\{x\}$ are disjoint as such
$$P(X \leq x) = P(X < x) + P(X = x).$$
     This allows us to calculate:
$$\begin{aligned} P(X = x) &= P(X \leq x) - P(X < x) \\ &= F(x) - P(X < x) \\ &= F(x) - \lim_{\varepsilon \searrow 0} P(X \leq x - \varepsilon) \\ &= F(x) - \lim_{y \searrow x} P(X \leq y) \\ &= F(x) - \lim_{y \searrow x} F(y) \\ &= F(x^+) - F(x^-), \end{aligned}$$
     where we have used that $F(x) = F^+(x)$ by definition of CDFs.

2. Let X be such that $P(X = 2) = P(X = 3) = 1/10$ and $P(X = 5) = 8/10$. Plot the CDF F. Use F to find $P(2 < X \leq 4.8)$ and $P(2 \leq X \leq 4.8)$.

   -
```
import numpy as np
import matplotlib.pyplot as plt

def Ex2_2(draw = True, save = False):
    x = np.arange(start=0, stop = 6, step = 0.2)
    F = (x >= 2) * 0.1 + (x >= 3) * 0.1 + (x >= 5) * 0.8
    #plt.step(x, F, where='post')

    if save:
        plt.savefig("Ex2_2.png")
    if draw:
        plt.show()

    right = np.where(abs(x - 4.8) < 0.1)[0]
    left = np.where(abs(x - 2) < 0.1)[0]
    print("P(2 <  X <= 4.8) = ", (F[right] - F[left])[0])
    print("P(2 <= X <= 4.8) = ", (F[right] - F[left - 1])[0])

Ex2_2(draw = False, save = False)

> P(2 <  X <= 4.8) =  0.1
> P(2 <= X <= 4.8) =  0.2
```



3. Prove Lemma 2.15.

4. Let X have probability density function

$$f(x) = \begin{cases} \frac{1}{4} & \text{if } 0 < x < 1 \\ \frac{3}{8} & \text{if } 3 < x < 5 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Find the cumulative distribution function of X.

   (b) Let $Y = 1/X$. Find the probability density function $f_Y(y)$ for Y. Hint. Consider three cases: $\frac{1}{5} \leq y \leq \frac{1}{3}$, $\frac{1}{3} \leq y \leq 1$, and $y \geq 1$.

5. Let X and Y be discrete random variables. Show that X and Y are independent if and only if $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ for all x and y.

6.

7.

8.

9. Let $X \sim \text{Exp}(\beta)$. Find $F(x)$ and $F^{-1}(q)$.

10. Let X and Y be independent. Show that $g(X)$ is independent of $h(Y)$ where g and h are functions.

11.

12. Prove Theorem 2.33

13. Let $X \sim \mathcal{N}(0,1)$ and let $Y = e^X$.

    (a) Find the PDF for Y. Plot it.

    (b) (Computer Experiment) Generate a vector $x = (x_1, \ldots, x_{10000})$ consisting of 10000 random standard Normals. Let $y = (y_1, \ldots, y_{10000})$ where $y_i = e^{x_i}$. Draw a histogram of y and compare it to the PDF you found in part (a).

14. Let (X,Y) be uniformly distributed on the unit disk $\{(x,y) : x^2+y^2 \leq 1\}$. Let $R = \sqrt{X^2+Y^2}$. Find the CFD and PDF of R.

15.

16.

17.

18.

19. Prove formula (2.12)

20.

21.

1.3 Expectation

1.

2.

3.

4.

5.

6. Prove Theorem 3.6 for discrete random variables

7.

8. Prove Theorem 3.17

9. (Computer Experiment) Let $X_1, X_2, \ldots, X_n$ be $\mathcal{N}(0,1)$ random variables and let

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n}.$$

   Plot $\overline{X}_n$ versus n for $n = 1, \ldots, 10000$. Repeat for $X_1, X_2, \ldots, X_n \sim$ Cauchy. Explain why there is such a difference.

10.

11. (Computer Experiment: Simulating the Stock Market). Let $Y_1, Y_2, \ldots$ be independent random variables such that $P(Y_i = 1) = P(Y_i = -1) = \frac{1}{2}$. Let $X_n = \sum_{i=1}^n Y_i$. Think of $Y_i = 1$ as "the stock market increased by one dollar", $Y_i = -1$ as "the stock market decreased by one dollar", and X_n as the value of the stock on day n.

   (a) Find $E(X_n)$ and $Var(X_n)$.

   - By linearity $E(X_n) = \sum_{i=1}^n E(Y_i) = \sum_{i=1}^n (-1)\frac{1}{2} + (1)\frac{1}{2} = 0$. Note that $E(X_n)^2 = 0$, as such $Var(X_n) = E(X_n^2)$. To calculate this note that

   $$X_n^2 = \left(\sum_{i=1}^n Y_i\right)^2 = \sum_{i=1}^n Y_i^2 + 2\sum_{i<j} Y_i Y_j.$$

   Because $Y_i \in \{-1, 1\}$ it follows that $E(Y_i^2) = 1$ and as such

   $$E(X_n^2) = n+2\sum_{i<j} E(Y_i Y_j) = n+2\sum_{i<j} P(Y_i Y_j = +1) - P(Y_i Y_j = -1).$$

   Note that $Y_i Y_j = 1$ if and only if $(Y_i, Y_j) \in \{(1,1), (-1,-1)\}$ and $Y_i Y_j = -1$ if and only if $(Y_i, Y_j) \in \{(-1,1), (1,-1)\}$, as such both of those are equally liked. In total this means that

   $$V(X_n) = E(X_n^2) = n.$$

   (b) Simulate X_n and plot X_n versus n for $n = 1, 2, \ldots, 10000$. Repeat the whole simulation several times. Notice two things. First, it's easy to "see" patterns in the sequence even though it is random. Second, you will find that the four runs look very different even though they were generated the same way. How do the calculations in (a) explain the second observation?

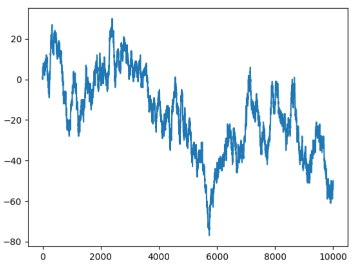
   - ```python
     import numpy as np
     import matplotlib.pyplot as plt

     def Ex3_11(n = 10000, draw = True, save = False, i = 0):
         # 2*0 - 1 = -1, 2 * 1 -1 = 1
         rand = 2 * np.random.randint(low = 0, high = 2, size = n) - 1

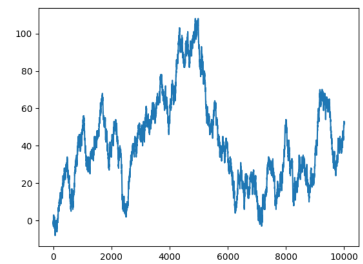
         fig, ax = plt.subplots()
         ax.plot(range(0, n), np.cumsum(rand))

         if draw:
             plt.show()

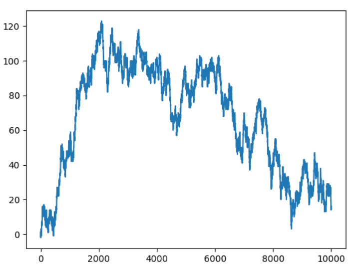
         if save:
             fig.savefig(f"Ex3_11-{i}.png")
     ```
     ```



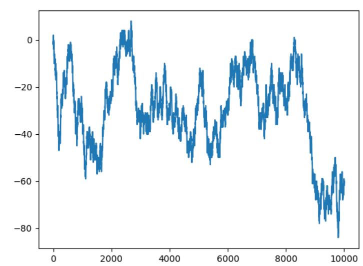
(a)



(b)



(c)



(d)

```
for i in range(4):
    Ex3_11(draw = True, save = True, i = i)
```

12.