EE 302 Division 1.
Homework 5 Solutions.

Problem 1. A fair four-sided die (with faces labeled 0, 1, 2, 3) is thrown once to determine how many times a fair coin is to be flipped: if N is the number that results from throwing the die, we flip the coin N times. Let K be the total number of heads resulting from the coin flips. Determine and sketch each of the following probability mass functions for all values of their arguments:

(a) $p_N(n)$.

Solution. Since this is a fair die, N is equally likely to be any one of the four numbers:

$$p_N(n) = \begin{cases} 1/4 & n = 0, 1, 2, 3 \\ 0 & \text{otherwise.} \end{cases}$$

(b) $p_{K|N}(k|2)$.

Solution. The distribution of the number K of heads in 2 flips of a fair coin is binomial with parameters $n = 2$, $p = 1/2$:

$$p_{K|N}(k|2) = \begin{cases} \begin{pmatrix} 2 \\ k \end{pmatrix} \left(\frac{1}{2}\right)^2 & k = 0, 1, 2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 1/4 & k = 0, 2 \\ 1/2 & k = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Indeed, the probability that both flips come up heads is 1/4, the probability that neither of the two flips comes up heads is 1/4, and the probability to get one tail and one head is 1/2.

(c) $p_{N|K}(n|2)$.

Solution. For this part, and the next two parts, drawing a table of the joint PMF of N and K is very helpful. Keep in mind that the joint PMF is the most complete description of two discrete random variables: anything about these random variables can be obtained from the joint PMF. For this problem, the joint PMF can be obtained as follows:

$$\begin{aligned} p_{N,K}(n,k) &= \mathbf{P}(N = n, K = k) = \mathbf{P}(K = k|N = n)\mathbf{P}(N = n) \\ &= p_{K|N}(k|n)p_N(n) = \frac{1}{4}p_{K|N}(k|n), \text{ for } n = 0, 1, 2, 3, \end{aligned} \tag{1}$$

where we used the fact that the four outcomes of throwing the die are equally likely. To determine the conditional probability distributions $p_{K|N}(k|n)$ for each value of n, note that:

- If $N = 0$, then we do not flip the coin at all, and K must be zero. Therefore, $p_{K|N}(0|0) = 1$ and $p_{K|N}(k|0) = 0$ for $k \neq 0$.
- If $N = 1$, then we flip once, and have equal chances to get one head ($K = 1$) or no heads ($K = 0$): $p_{K|N}(0|1) = p_{K|N}(1|1) = 1/2$.
- The case $N = 2$ was considered in Part (b).

- Finally, if we flip the coin three times, the probability to get three heads is the same as the probability to get no heads, and is equal to 1/8; and the probability to get one head is the same as the probability to get two heads, and is equal to 3/8: $p_{K|N}(0|3) = p_{K|N}(3|3) = 1/8$, $p_{K|N}(1|3) = p_{K|N}(2|3) = 3/8$.

An alternative way to determine these conditional probabilities is to note–just like we did in Part (b)–that the distribution of the number K of heads in n tosses of a fair coin is binomial with parameters n and $p = 1/2$:

$$p_{K|N}(k|n) = \begin{cases} \begin{pmatrix} n \\ k \end{pmatrix} \left(\frac{1}{2}\right)^n & 0 \le k \le n \\ 0 & \text{otherwise.} \end{cases}$$

To get the joint PMF of N and K, Eq. (1) tells us to multiply the conditional probabilities by 1/4. We therefore have the following joint PMF:

n \ k	0	1	2	3
0	1/4	0	0	0
1	1/8	1/8	0	0
2	1/16	1/8	1/16	0
3	1/32	3/32	3/32	1/32

To get the PMF $p_{N|K}(n|2)$, we take the probabilities from the $k = 2$ column of the table, and normalize them by the probability of the event $K = 2$. The probability of the event $K = 2$ is the sum of all probabilities in that column: $\mathbf{P}(K = 2) = 1/16 + 3/32 = 5/32$. Therefore,

$$p_{N|K}(n|2) = \frac{p_{N,K}(n,2)}{p_K(2)} = \frac{p_{N,K}(n,2)}{5/32} = \begin{cases} 2/5 & n = 2, \\ 3/5 & n = 3, \\ 0 & \text{otherwise.} \end{cases}$$

(d) $p_K(k)$.

Solution. The probability of $K = k$ is obtained by summing the probabilities in the corresponding column:

$$p_K(k) = \begin{cases} 15/32 & k = 0 \\ 11/32 & k = 1 \\ 5/32 & k = 2 \\ 1/32 & k = 3 \\ 0 & \text{otherwise.} \end{cases}$$

(e) Also determine the conditional PMF for random variable N, given that the experimental value k of random variable K is an odd number.

Solution. The event "K is odd" corresponds to the columns $k = 1$ and $k = 3$. Note that when K is odd, the only possible values for N are 1, 2 and 3. The probability of the event "$N = 1$ and K is odd" is 1/8; the probability of the event "$N = 2$ and K is odd" is 1/8; and the probability of the event "$N = 3$ and K is odd" is also 3/32+1/32=1/8. Thus, conditioned on K being odd, the events $N = 1$, $N = 2$, and $N = 3$ are equally likely:

$$p_{N|K}(n|K \text{ is odd}) = \frac{\mathbf{P}(N = n, K \text{ is odd})}{\mathbf{P}(K \text{ is odd})} = \frac{1/8}{\mathbf{P}(K \text{ is odd})}, \text{ for } n = 1, 2, 3. \quad (2)$$

Since the conditional PMF must sum up to one, these probabilities have to be 1/3 each (another way to show this is to simply notice that, according to our result in Part (d), the probability of the event "K is odd" is $11/32+1/32 = 3/8$):

$$p_{N|K}(n|K \text{ is odd}) = \begin{cases} 1/3 & n = 1, 2, 3 \\ 0 & \text{otherwise.} \end{cases}$$

Problem 2. Let X and Y be independent random variables. Random variable X has a discrete uniform distribution over the set $\{1, 2, 3\}$, and Y has a discrete uniform distribution over the set $\{1, 3\}$. Let $V = X + Y$, and $W = X - Y$.

(a) Are V and W independent? Explain without calculations.

Solution. No. Independence of V and W would imply that:

$$p_{W|V}(w|v) = p_W(w),$$

which means the distribution of W cannot depend on the value of V. But in fact, unconditionally, W can have several different experimental values, whereas conditioned on $V = 6$, both X and Y must be 3 and so $W = 0$ with probability 1. Thus $p_{W|V}(w|6) \neq p_W(w)$, and therefore V and W are dependent.

(b) Find and plot $p_V(v)$. Also, determine $E[V]$ and $\text{var}(V)$.

(c) Find and show in a diagram $p_{V,W}(v, w)$.

Solution. To visualize this situation, it is helpful to look at the joint PMF of X and Y. Since both X and Y are uniform, and since they are independent, each of the six possible pairs of their experimental values has probability 1/6. In addition to these probabilities, let us tabulate the corresponding experimental values v and w of random variables V and W:

$y \setminus x$	1	2	3
1	1/6 $v = 2$ $w = 0$	1/6 $v = 3$ $w = 1$	1/6 $v = 4$ $w = 2$
3	1/6 $v = 4$ $w = -2$	1/6 $v = 5$ $w = -1$	1/6 $v = 6$ $w = 0$

From this table, it is easy to determine the joint probability mass function of V and W:

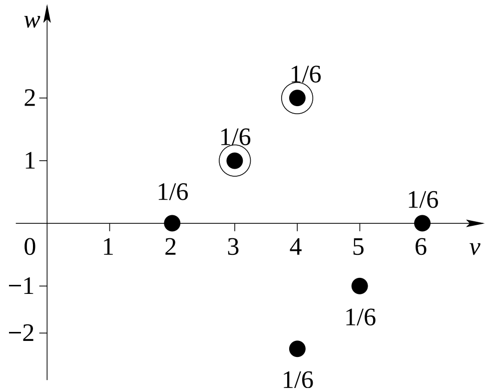


Figure 1: A sketch of the joint probability mass function of V and W: there are six equally likely pairs of values, each has probability $1/6$. The two circled points comprise the event $W > 0$ considered in Part (d).

To get the marginal PMF for V, we need to sum $p_{V,W}$ over w, for each v, i.e., take vertical sums in our picture of the joint PMF of V and W:

$$p_V(v) = \begin{cases} \frac{1}{6}, & v = 2, 3, 5, 6, \\ \frac{1}{3}, & v = 4, \\ 0, & \text{otherwise.} \end{cases}$$

The PMF of V is depicted in figure 2. Since it is symmetric about $v = 4$, we have: $E[V] = 4$.

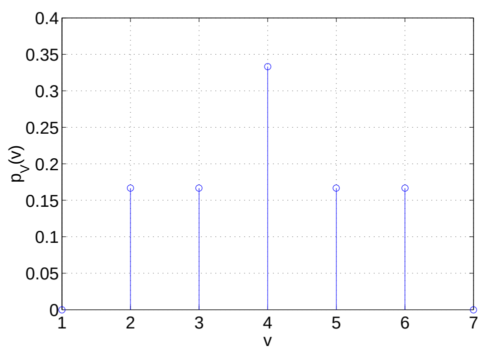


Figure 2: $p_V(v)$.

Also,

$$\begin{aligned} \text{var}(V) &= E[(V - E[V])^2] = E[(V-4)^2] = \sum_v (v-4)^2 p_V(v) \\ &= \frac{1}{6} \cdot [(2-4)^2 + (3-4)^2 + (5-4)^2 + (6-4)^2] + \frac{1}{3} \cdot (4-4)^2 = \frac{1}{6}(4+1+1+4) \\ &= \frac{5}{3}. \end{aligned}$$

(d) Find $E[V|W > 0]$.

Solution. The event $W > 0$ corresponds to the two circled points in the picture of the joint PMF of V and W. Conditioned on this event, there are two equally likely values that V can assume: 3 and 4. The conditional expectation is therefore 3.5.

(e) Find the conditional variance of W given the event $V = 4$.

Solution. Looking at our sketch of $p_{V,W}(v, w)$ again, we see that, conditioned on $V = 4$, there are two equally likely values for W, 2 and -2:

$$p_{W|V}(w|4) = \begin{cases} \frac{1}{2} & w = -2, 2 \\ 0 & \text{otherwise.} \end{cases}$$

This conditional PMF is symmetric about 0 and therefore $E[W|V = 4] = 0$, and

$$\text{var}[W|V = 4] = E[W^2|V = 4] = 2^2 \cdot \frac{1}{2} + (-2)^2 \cdot \frac{1}{2} = 4.$$

(f) Find and plot the conditional PMF $p_{X|V}(x|v)$, for all values.

Solution. Looking back at the table of $p_{X,Y}$ and corresponding values of V and W that we constructed in Parts (b), (c), we see that, given $V = 2, 3, 5, 6$, X has to be 1,2,2,3, respectively:

$$p_{X|V}(x|2) = \begin{cases} 1 & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x|3) = \begin{cases} 1 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x|5) = \begin{cases} 1 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x|6) = \begin{cases} 1 & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

When $V = 4$, X is equally likely to be 1 or 3, and therefore we get:

$$p_{X|V}(x|4) = \begin{cases} 1/2 & x = 1 \\ 1/2 & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

These five conditional probability mass functions are depicted in Figure 3.

Problem 3. Joe and Helen each know that the a priori probability that her mother will be home on any given night is 0.6. However, Helen can determine her mother's plans for the night at 6pm, and then, at 6:15pm, she has only one chance each evening to shout one of two code words across the river to Joe. He will visit her with probability 1 if he thinks Helen's message means "Ma will be away," and he will stay home with probability 1 if he thinks the message means "Ma will be home."

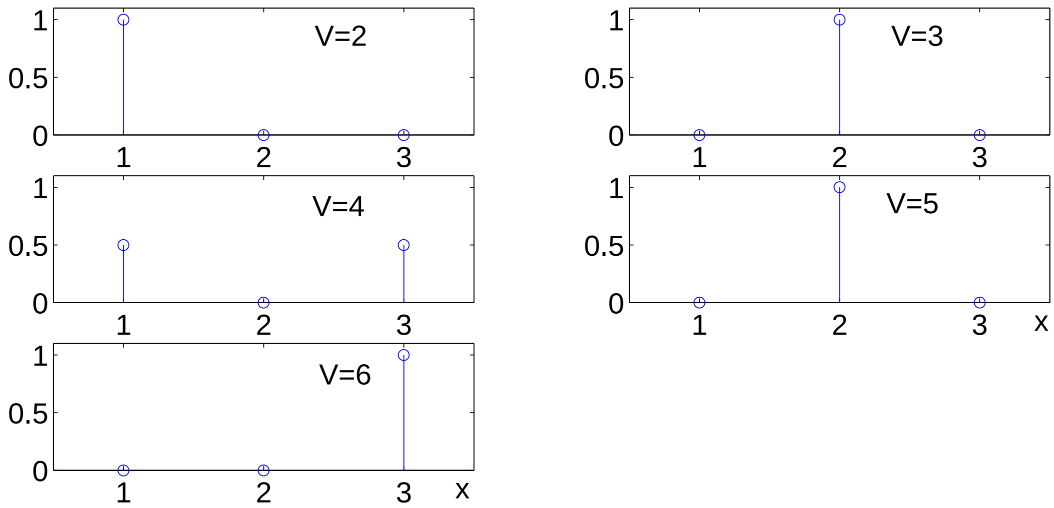


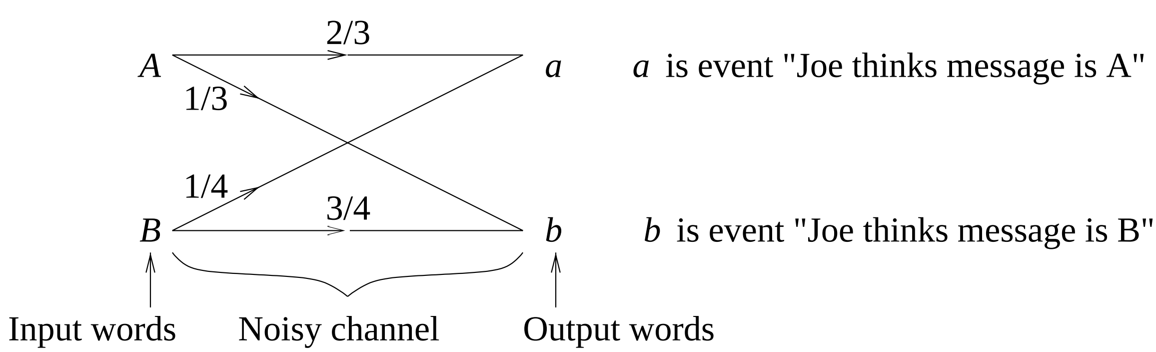
Figure 3: Conditional PMF $p_{X|V}(x|v)$ in Problem 2(f), as a function of x, for each of the five possible experimental values of V.

But Helen has a meek voice, and the river is channeled for heavy barge traffic. Thus she is faced with the problem of coding for a noisy channel. She has decided to use a code containing only the code words A and B.

The channel is described by:

$$\mathbf{P}(a|A) = \frac{2}{3},$$
$$\mathbf{P}(a|B) = \frac{1}{4},$$
$$\mathbf{P}(b|A) = \frac{1}{3},$$
$$\mathbf{P}(b|B) = \frac{3}{4},$$

and these events are defined in the following sketch:



(a) In order to minimize the probability of error between transmitted and received messages, should

Helen and Joe agree to use code I or code II?

Code I: $A =$ Ma away, $B =$ Ma home.
Code II: $A =$ Ma home, $B =$ Ma away.

Solution. By the total probability theorem, the probability of error is:

$$\begin{aligned}\mathbf{P}(\text{error}) &= \mathbf{P}(\text{error}|A)\mathbf{P}(A) + \mathbf{P}(\text{error}|B)\mathbf{P}(B) \\ &= \mathbf{P}(b|A)\mathbf{P}(A) + \mathbf{P}(a|B)\mathbf{P}(B) \\ &= \frac{1}{3}\mathbf{P}(A) + \frac{1}{4}\mathbf{P}(B) = \frac{1}{3}\mathbf{P}(A) + \frac{1}{4}(1 - \mathbf{P}(A)) \\ &= \frac{1}{4} + \frac{1}{12}\mathbf{P}(A).\end{aligned}$$

To minimize the probability of error, they need to choose the coding scheme for which the value of $\mathbf{P}(A)$ is the smallest, i.e. Code I for which $\mathbf{P}(A) = \mathbf{P}(\text{Ma away}) = 0.4$ (because for Code II $\mathbf{P}(A) = \mathbf{P}(\text{Ma home}) = 0.6$).

(b) Helen and Joe put the following cash values (in dollars) on all possible outcomes of a day:

Ma home and Joe comes	-30
Ma home and Joe does not come	0
Ma away and Joe comes	+30
Ma away and Joe does not come	-5

Joe and Helen make all their plans with the objective of maximizing the expected value of each day of their continuing romance. Which of the above codes will maximize the expected cash value per day of this romance?

Solution. Suppose random variable V is their value per day. It follows from the definition of expectation that:

$$\begin{aligned}E[V] &= -30\mathbf{P}(V = -30) + 0\mathbf{P}(V = 0) + 30\mathbf{P}(V = 30) - 5\mathbf{P}(V = -5) \quad (3) \\ &= -30\mathbf{P}(\text{Ma home and Joe comes}) \\ &\quad +30\mathbf{P}(\text{Ma away and Joe comes}) \\ &\quad -5\mathbf{P}(\text{Ma away and Joe doesn't come})\end{aligned}$$

We therefore need to work out the probabilities of these three events. For the first one, from the definition of conditional probability, we have:

$$\begin{aligned}\mathbf{P}(\text{Ma home and Joe comes}) &= \mathbf{P}(\text{Joe comes}|\text{Ma home})\mathbf{P}(\text{Ma home}) \\ &= 0.6\mathbf{P}(\text{Joe comes}|\text{Ma home})\end{aligned}$$

Supposing they use Code I, we have:

$$\begin{aligned}\mathbf{P}(\text{Joe comes}|\text{Ma home}) &= \mathbf{P}(a|B) = 1/4 \\ \mathbf{P}(\text{Ma home and Joe comes}) &= 0.6 \cdot \frac{1}{4} = 0.15\end{aligned}$$

Similarly, we can get the remaining probabilities for Code I:

$$\begin{aligned}\mathbf{P}(\text{Ma away and Joe comes}) &= 0.4\mathbf{P}(a|A) = 0.8/3 \\ \mathbf{P}(\text{Ma away and Joe doesn't come}) &= 0.4\mathbf{P}(b|A) = 0.4/3\end{aligned}$$

We now substitute these probabilities back into Eq. 3 to get:

$$E[V|CodeI] = -30 \cdot 0.15 + 30 \cdot \frac{0.8}{3} - 5 \cdot \frac{0.4}{3} 2\frac{5}{6} \approx 2.83.$$

If we assume they use Code II and go through the above process again, the expected value will be:

$$E[V|CodeII] = -30 \cdot 0.2 + 30 \cdot 0.3 - 5 \cdot 0.1 = 2.5.$$

We can see that Code I is also better in the sense of maximizing the daily expected value of their romance.

(c) Clara is not quite so attractive as Helen, but at least she lives on the same side of the river. What would be the lower limit of Clara's expected value per day which would make Joe decide to give up Helen?

Solution. The maximum expected daily value is $2\frac{5}{6}$ if he chooses to date Helen. If the expected value of Clara is greater than that, he will change his mind (if his objective is to maximize the expected value). So Clara's expected value per day should be at least $2\frac{5}{6}$ to compete with Helen.

(d) What would be the maximum rate which Joe would pay the phone company for a noiseless wire to Helen's house which he could use once per day at 6:15pm?

Solution. If they use a perfect telephone line, Joe would come whenever Ma is away and he would never come if Ma is home. It follows that:

$$\begin{aligned} \mathbf{P}(\text{Ma home and Joe comes}) &= 0 \\ \mathbf{P}(\text{Ma away and Joe comes}) &= \mathbf{P}(\text{Ma away}) = 0.4 \\ \mathbf{P}(\text{Ma away and Joe doesn't come}) &= 0 \end{aligned}$$

We now substitute these probabilities back into Eq. (3) to get:

$$E[V|\text{Perfect Channel}] = 30 \cdot 0.4 = 12.$$

If the daily rate is less than the increase in their expected daily value, he would be willing to pay the phone company. So the answer is:

$$12 - 2\frac{5}{6} = 9\frac{1}{6} \approx 9.17.$$

(e) How much is it worth to Joe and Helen to double her mother's probability of being away from home? Would this be a better or worse investment than spending the same amount of money for a telephone line (to be used once a day at 6:15pm) with the following properties:

$$\begin{aligned} \mathbf{P}(a|A) = \mathbf{P}(b|B) &= 0.9, \\ \mathbf{P}(b|A) = \mathbf{P}(a|B) &= 0.1? \end{aligned}$$

Solution. First, let us see how much the daily value will be if her mother's probability of being away is doubled. Because we do not know which code is better in this case, we still need to calculate the expected values for both codes. For Code I, similarly to what we did in Part (b), we get:

$$E[V|\text{Code I}] = -30 \cdot 0.05 + 30 \cdot \frac{1.6}{3} - 5 \cdot \frac{0.8}{3} = 13\frac{1}{6}.$$

If we assume they use Code II and go through the above process again, the expected value will be:

$$E[V|\text{Code II}] = -30 \cdot \frac{0.2}{3} + 30 \cdot 0.6 - 5 \cdot 0.2 = 15$$

If they choose Code II, their expected daily value will be increased to 15. Again, they will be willing to pay any amount which is smaller than the increment in the expected value, which is $15 - 2\frac{5}{6} = 12\frac{1}{6}$.

Note that we do not need to calculate the expected value with the imperfect telephone line because it will surely be less than the expected value for the noiseless channel which we calculated above, and which is $12 < 12\frac{1}{6}$. So, it is better to invest in doubling Ma's probability of being away from home.

Problem 4. Evaluate the following summation without too much work:

$$\sum_{n=0}^{N} \binom{N}{n} n^2 A^n (1-A)^{N-n},$$

where $0 < A < 1$ and $N > 2$.

Solution. Consider a binomial random variable K with the following PMF:

$$p_K(n) = \binom{N}{n} A^n (1-A)^{N-n},$$

where N can be thought of as the total number of independent Bernoulli trials and A the probability of success in one trial. We then see that the summation we need to evaluate is equal to the expected value of K^2, i.e.:

$$E[K^2] = \sum_{n=0}^{N} n^2 \binom{N}{n} A^n (1-A)^{N-n}.$$

On the other hand, we have:

$$\begin{aligned} E[K^2] &= \text{var}[K] + E[K]^2 \\ &= NA(1-A) + (NA)^2 = N^2A^2 - NA^2 + NA. \end{aligned}$$

Problem 5. The joint PMF of discrete random variables X and Y is given by:

$$p_{X,Y}(x,y) = \begin{cases} Cx^2\sqrt{y}, & \text{for } x = -5, -4, \ldots, 4, 5 \text{ and } y = 0, 1, \ldots, 10. \\ 0, & \text{otherwise.} \end{cases}$$

Here, C is some constant. What is $E[XY^3]$? (**Hint**: This question admits a short answer/explanation. Do not spend time doing calculations.)

Solution. We first note that $p_{X,Y}(x,y) = p_{X,Y}(-x,y)$, since $Cx^2\sqrt{y} = C(-x)^2\sqrt{y}$. By definition, $E[XY^3] = \sum_{x,y} xy^3 p_{X,Y}(x,y)$. Consider the contribution $xy^3 p_{X,Y}(x,y)$ of each pair (x,y) to the overall sum. When $x = 0$, the term $xy^3 p_{X,Y}(x,y)$ is equal to zero and thus contributes nothing to the overall sum. For each pair (x,y) with $x \neq 0$, there is another pair $(-x,y)$ which contributes $-xy^3 p_{X,Y}(-x,y) = -xy^3 p_{X,Y}(x,y)$. Thus, the contribution of each pair (x,y) is canceled out by the contribution of $(-x,y)$. Therefore, $E[XY^3] = 0$.

Problem 6. A random variable X is called a *shifted exponential* when its PDF has the following form:

$$f_X(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x-\alpha}{\theta}}, & x > \alpha \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the CDF of X.

Solution. Define $Y = X - \alpha$. The cumulative distribution function of this random variable is then:

$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(X - \alpha \leq y) = \mathbf{P}(X \leq y + \alpha) = F_X(y + \alpha). \tag{4}$$

Differentiating the leftmost part of this equation and the rightmost part, we get $f_Y(y)$ and $f_X(y+\alpha)$, respectively:

$$f_Y(y) = f_X(y + \alpha) = \begin{cases} \frac{1}{\theta} e^{-\frac{y}{\theta}}, & y > 0 \\ 0, & \text{otherwise,} \end{cases}$$

i.e. Y is an exponential random variable with mean θ. Using Eq. (4), as well as the exponential CDF derived in class, we have:

$$F_X(x) = F_Y(x - \alpha) = \begin{cases} 1 - e^{-\frac{x-\alpha}{\theta}}, & x > \alpha \\ 0, & \text{otherwise.} \end{cases}$$

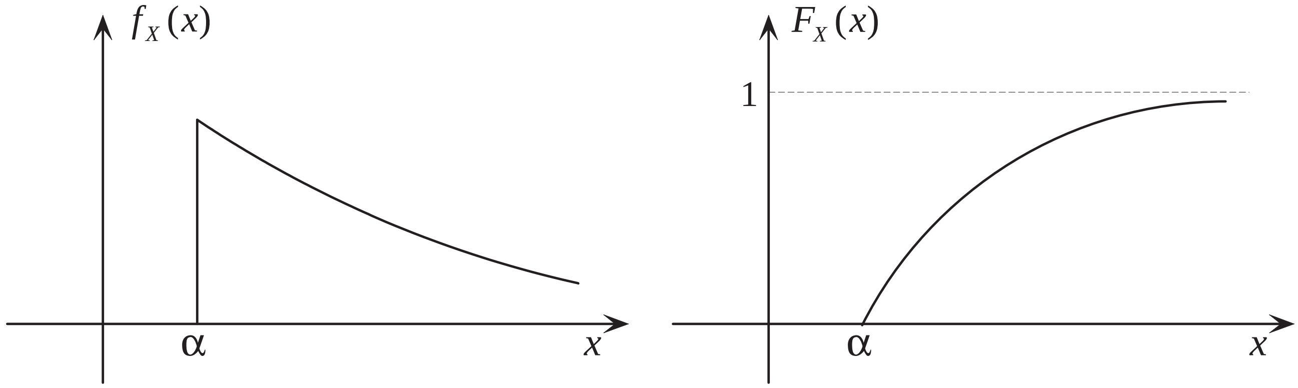
$F_X(x)$ is depicted in Figure 4(b).

(b) Calculate the mean and the variance of X.

Solution. It was derived in class that the mean and variance of the exponential random variable Y we defined in the last part are θ and θ^2 respectively. Since α is just a constant, we have:

$$\begin{aligned} E[X] &= E[Y + \alpha] = E[Y] + \alpha = \theta + \alpha \\ Var[X] &= Var[Y + \alpha] = Var[Y] = \theta^2 \end{aligned}$$

(c) Find the real number μ that satisfies: $F_X(\mu) = 1/2$. This number μ is called the *median* of the random variable.



(a) PDF of X (b) CDF of X

Figure 4: PDF and CDF of X

Solution. With the CDF we got in Part(a), we have:

$$
\begin{aligned}
1 - e^{-\frac{\mu-\alpha}{\theta}} &= \frac{1}{2} \\
\Rightarrow \frac{\mu-\alpha}{\theta} &= \ln 2 \\
\Rightarrow \mu &= \theta \ln 2 + \alpha.
\end{aligned}
$$