CHAPTER 5

Discrete random variables

5.1. Definition, properties, expectation, moments

As before, suppose S is a sample space.

Definition 5.1 (Random variable)

A *random variable* is a real-valued function on S. Random variables are usually denoted by $X, Y, Z, \ldots$. A *discrete random variable* is one that can take on only countably many values.

Example 5.1. If one rolls a die, let X denote the outcome, i.e. taking values $1, 2, 3, 4, 5, 6$.

Example 5.2. If one rolls a die, let Y be 1 if an odd number is showing, and 0 if an even number is showing.

Example 5.3. If one tosses 10 coins, let X be the number of heads showing.

Example 5.4. In n trials, let X be the number of successes.

Definition (PMF or density of a random variable)

For a discrete random variable X, we define the *probability mass function (PMF)* or the *density of* X by

$$p_X(x) := \mathbb{P}(X = x),$$

where $\mathbb{P}(X = x)$ is a standard abbreviation for

$$\mathbb{P}(X = x) = \mathbb{P}\left(X^{-1}(x)\right).$$

Note that the pre-image $X^{-1}(x)$ is the event $\{\omega \in S : X(\omega) = x\}$.

Suppose X is a discrete random variable taking on values $\{x_i\}_{i \in \mathbb{N}}$, then

$$\sum_{i \in \mathbb{N}} p_X(x_i) = \mathbb{P}(S) = 1.$$

Let X be the number showing if we roll a die. The *expected number* to show up on a roll of a die should be $1 \cdot \mathbb{P}(X = 1) + 2 \cdot \mathbb{P}(X = 2) + \cdots + 6 \cdot \mathbb{P}(X = 6) = 3.5$. More generally, we define

Definition 5.2 (Expectation of a discrete random variable)

For a discrete random variable X we define the *expected value* or *expectation* or *mean* of X as

$$\mathbb{E}X := \sum_{\{x : p_X(x) > 0\}} x p_X(x)$$

provided this sum converges absolutely. In this case we say that the expectation of X is well-defined.

We need absolute convergence of the sum so that the expectation does not depend on the order in which we take the sum to define it. We know from calculus that we need to be careful about the sums of conditionally convergent series, though in most of the examples we deal with this will not be a problem. Note that $p_X(x)$ is nonnegative for all x, but x itself can be negative or positive, so in general the terms in the sum might have different signs.

Example 5.5. If we toss a coin and X is 1 if we have heads and 0 if we have tails, what is the expectation of X?

Solution:

$$p_X(x) = \begin{cases} \frac{1}{2}, & x = 1 \\ \frac{1}{2}, & x = 0 \\ 0, & \text{all other values of } x. \end{cases}$$

Hence $\mathbb{E}X = (1)(\frac{1}{2}) + (0)(\frac{1}{2}) = \frac{1}{2}$.

Example 5.6. Suppose $X = 0$ with probability $\frac{1}{2}$, 1 with probability $\frac{1}{4}$, 2 with probability $\frac{1}{8}$, and more generally n with probability $1/2^{n+1}$. This is an example where X can take infinitely many values (although still countably many values). What is the expectation of X?

Solution: Here $p_X(n) = 1/2^{n+1}$ if n is a nonnegative integer and 0 otherwise. So

$$\mathbb{E}X = (0)\tfrac{1}{2} + (1)\tfrac{1}{4} + (2)\tfrac{1}{8} + (3)\tfrac{1}{16} + \cdots .$$

This turns out to sum to 1. To see this, recall the formula for a geometric series

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}.$$

If we differentiate this, we get

$$1 + 2x + 3x^2 + \cdots = \frac{1}{(1-x)^2}.$$

We have

$$\begin{aligned}\mathbb{E}X &= 1(\tfrac{1}{4}) + 2(\tfrac{1}{8}) + 3(\tfrac{1}{16}) + \cdots \\ &= \tfrac{1}{4}\Big[1 + 2(\tfrac{1}{2}) + 3(\tfrac{1}{4}) + \cdots\Big] \\ &= \tfrac{1}{4}\frac{1}{(1-\frac{1}{2})^2} = 1.\end{aligned}$$

Example 5.7. Suppose we roll a fair die. If 1 or 2 is showing, let $X = 3$; if a 3 or 4 is showing, let $X = 4$, and if a 5 or 6 is showing, let $X = 10$. What is $\mathbb{E}X$?

Solution: We have $\mathbb{P}(X = 3) = \mathbb{P}(X = 4) = \mathbb{P}(X = 10) = \frac{1}{3}$, so

$$\mathbb{E}X = \sum x\mathbb{P}(X = x) = (3)(\tfrac{1}{3}) + (4)(\tfrac{1}{3}) + (10)(\tfrac{1}{3}) = \tfrac{17}{3}.$$

Example 5.8. Consider a discrete random variable taking only positive integers as values with $\mathbb{P}(X = n) = \frac{1}{n(n+1)}$. What is the expectation $\mathbb{E}X$?

Solution: First observe that this is indeed a probability since we can use telescoping partial sums to show that

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$$

Then

$$\mathbb{E}X = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(X = n) = \sum_{n=1}^{\infty} n \cdot \frac{1}{n(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n+1} = +\infty,$$

so the expectation of X is infinite.

If we list all possible values of a discrete random variable X as $\{x_i\}_{i\in\mathbb{N}}$, then we can write

$$\mathbb{E}X = \sum_{\{x:p_X(x)>0\}} xp_X(x) = \sum_{i=1}^{\infty} x_i p_X(x_i).$$

We would like to show that the expectation is linear, that is, $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$.

We start by showing that we can write the expectation of a discrete random variable in a slightly different form. Note that in our definition of the expectation we first list all possible values of X and weights with probability that X attains these values. That is, we look at the range of X. Below we instead look at the domain of X and list all possible outcomes.

Proposition 5.1

If X is a random variable on a *finite* sample space S, then

$$\mathbb{E}X = \sum_{\omega\in S} X(\omega)\,\mathbb{P}(\{\omega\}).$$

Proof. For each $i \in \mathbb{N}$ we denote by S_i the event $\{\omega \in S : X(\omega) = x_i\}$. Then $\{S_i\}_{i\in\mathbb{N}}$ is a partition of the space S into disjoint sets. Note that since S is finite, then each set S_i is finite too, moreover, we only have a finite number of sets S_i which are non-empty.

$$\begin{aligned}
\mathbb{E}X &= \sum_{i=1}^{\infty} x_i p(x_i) = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i) = \sum_{i=1}^{\infty} x_i \left(\sum_{\omega\in S_i} \mathbb{P}(\{\omega\}) \right) \\
&= \sum_{i=1}^{\infty} \left(\sum_{\omega\in S_i} x_i \mathbb{P}(\{\omega\}) \right) = \sum_{i=1}^{\infty} \sum_{\omega\in S_i} X(\omega)\mathbb{P}(\{\omega\}) \\
&= \sum_{\omega\in S} X(\omega)\mathbb{P}(\{\omega\}),
\end{aligned}$$

where we used properties of sets $\{S_i\}_{i=1}^{\infty}$. □

Proposition 5.1 is true even if S is *countable* as long as $\mathbb{E}X$ is well-defined. First, observe that if S is countable then the random variable X is necessarily discrete. Where do we need to use the assumption that all sums converge absolutely? Note that the identity

$$\sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i) = \sum_{\omega\in S} X(\omega)\mathbb{P}(\{\omega\})$$

is a re-arrangement in the first sum, which we can do as long as the sums (series) converge absolutely. Note that if either the number of values of X or the sample space S is finite, we can use this argument.

Proposition 5.1 can be used to prove linearity of the expectation.

Theorem 5.1 (Linearity of expectation)

If X and Y are discrete random variables defined on the same sample space S and $a \in \mathbb{R}$, then

(i) $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$,

(ii) $\mathbb{E}[aX] = a\mathbb{E}X$,

as long as all expectations are well-defined.

Proof. Consider a random variable $Z := X + Y$ which is a discrete random variable on the sample space S. We use $\mathbb{P}(X = x, Y = y)$ to denote the probability of the event

$$\{\omega \in S : X(\omega) = x\} \cap \{\omega \in S : Y(\omega) = y\}.$$

Denote by $\{x_i\}_{i\in\mathbb{N}}$ the values that X is taking, and by $\{y_j\}_{j\in\mathbb{N}}$ the values that Y is taking. Denote by $\{z_k\}_{k\in\mathbb{N}}$ the values that Z is taking. Since we assume that all random variables have well-defined expectations, we can interchange the order of summations freely. Then by the *law of total probability* (Proposition 4.4) twice we have

$$
\begin{aligned}
\mathbb{E}Z &= \sum_{k=1}^{\infty} z_k \mathbb{P}(Z = z_k) = \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} z_k \mathbb{P}(Z = z_k, X = x_i) \right) \\
&= \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} z_k \mathbb{P}(X = x_i, Y = z_k - x_i) \right) \\
&= \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} z_k \mathbb{P}(X = x_i, Y = z_k - x_i, Y = y_j).
\end{aligned}
$$

Now $\mathbb{P}(X = x_i, Y = z_k - x_i, Y = y_j)$ will be 0, unless $z_k - x_i = y_j$. For each pair (i, j), this will be non-zero for only one value k, since the z_k are all different. Therefore, for each i and j

$$
\begin{aligned}
&\sum_{k=1}^{\infty} z_k \mathbb{P}(X = x_i, Y = z_k - x_i, Y = y_j) \\
&\qquad = \sum_{k=1}^{\infty} (x_i + y_j) \mathbb{P}(X = x_i, Y = z_k - x_i, Y = y_j) \\
&\qquad = (x_i + y_j) \mathbb{P}(X = x_i, Y = y_j).
\end{aligned}
$$

Substituting this to the above sum we see that

$$
\begin{aligned}
\mathbb{E}Z &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (x_i + y_j) \mathbb{P}(X = x_i, Y = y_j) \\
&\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i \mathbb{P}(X = x_i, Y = y_j) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} y_j \mathbb{P}(X = x_i, Y = y_j) \\
&= \sum_{i=1}^{\infty} x_i \left(\sum_{j=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j) \right) + \sum_{j=1}^{\infty} y_j \left(\sum_{i=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j) \right) \\
&= \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i) + \sum_{j=1}^{\infty} y_j \mathbb{P}(Y = y_j) = \mathbb{E}X + \mathbb{E}Y,
\end{aligned}
$$

where we used the *law of total probability* (Proposition 4.4) again.

Note that if we have a countable sample space all these sums converge absolutely and so we can justify writing this similarly to Proposition 5.1 as

$$
\begin{aligned}
\mathbb{E}[X + Y] &= \sum_{\omega \in S} (X(\omega) + Y(\omega)) \mathbb{P}(\omega) \\
&= \sum_{\omega \in S} (X(\omega) \mathbb{P}(\omega) + Y(\omega) \mathbb{P}(\omega)) \\
&= \sum_{\omega \in S} X(\omega) \mathbb{P}(\omega) + \sum_{\omega \in S} Y(\omega) \mathbb{P}(\omega) \\
&= \mathbb{E}X + \mathbb{E}Y.
\end{aligned}
$$

For $a \in \mathbb{R}$ we have

$$\mathbb{E}[aX] = \sum_{\omega \in S} (aX(\omega))\, \mathbb{P}(\omega) = a \sum_{\omega \in S} X(\omega)\mathbb{P}(\omega) = a\mathbb{E}X$$

since these sums converge absolutely as long as $\mathbb{E}X$ is well-defined. $\square$

Using induction on the number of random variables linearity holds for a collection of random variables $X_1, X_2, \dots, X_n$.

Corollary

If $X_1, X_2, \dots, X_n$ are random variables, then

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}X_1 + \mathbb{E}X_2 + \cdots + \mathbb{E}X_n.$$

Example 5.9. Suppose we roll a die and let X be the value that is showing. We want to find the expectation $\mathbb{E}X^2$.

Solution: Let $Y = X^2$, so that $\mathbb{P}(Y = 1) = \frac{1}{6}$, $\mathbb{P}(Y = 4) = \frac{1}{6}$ etc. and

$$\mathbb{E}X^2 = \mathbb{E}Y = (1)\tfrac{1}{6} + (4)\tfrac{1}{6} + \cdots + (36)\tfrac{1}{6}.$$

We can also write this as

$$\mathbb{E}X^2 = (1^2)\tfrac{1}{6} + (2^2)\tfrac{1}{6} + \cdots + (6^2)\tfrac{1}{6},$$

which suggests that a formula for $\mathbb{E}X^2$ is $\sum_x x^2 \mathbb{P}(X = x)$. This turns out to be correct.

The only possibility where things could go wrong is if more than one value of X leads to the same value of X^2. For example, suppose $\mathbb{P}(X = -2) = \frac{1}{8}, \mathbb{P}(X = -1) = \frac{1}{4}, \mathbb{P}(X = 1) = \frac{3}{8}, \mathbb{P}(X = 2) = \frac{1}{4}$. Then if $Y = X^2$, $\mathbb{P}(Y = 1) = \frac{5}{8}$ and $\mathbb{P}(Y = 4) = \frac{3}{8}$. Then

$$\mathbb{E}X^2 = (1)\tfrac{5}{8} + (4)\tfrac{3}{8} = (-1)^2\tfrac{1}{4} + (1)^2\tfrac{3}{8} + (-2)^2\tfrac{1}{8} + (2)^2\tfrac{1}{4}.$$

But even in this case $\mathbb{E}X^2 = \sum_x x^2 \mathbb{P}(X = x)$.

Theorem 5.2

For a discrete random variable X taking values $\{x_i\}_{i=1}^{\infty}$ and a real-valued function g defined on this set, we have

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

Proof. Let $Y := g(X)$, then

$$\begin{aligned} \mathbb{E}Y &= \sum_y y\mathbb{P}(Y = y) = \sum_y y \sum_{\{x: g(x) = y\}} \mathbb{P}(X = x) \\ &= \sum_x g(x)\mathbb{P}(X = x). \end{aligned}$$

$\square$

Example 5.10. As before we see that $\mathbb{E}X^2 = \sum x^2 p_X(x)$. Also if $g(x) \equiv c$ is a constant function, then we see that the expectation of a constant is this constant

$$\mathbb{E}g(X) = \sum_{i=1}^{\infty} cp(x_i) = c \sum_{i=1}^{\infty} p(x_i) = c \cdot 1 = c.$$

Definition (Moments)

$\mathbb{E}X^n$ is called the *nth moment* of a random variable X. If $M := \mathbb{E}X$ is well defined, then

$$\operatorname{Var}(X) = \mathbb{E}(X - M)^2$$

is called the *variance* of X. The square root of $\operatorname{Var}(X)$ is called the *standard deviation* of X

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

By Theorem 5.2 we know that the nth moment can be calculated by

$$\mathbb{E}X^n = \sum_{x : p_X(x) > 0} x^n p_X(x).$$

The variance measures how much spread there is about the expected value.

Example 5.11. We toss a fair coin and let $X = 1$ if we get heads, $X = -1$ if we get tails. Then $\mathbb{E}X = 0$, so $X - \mathbb{E}X = X$, and then $\operatorname{Var} X = \mathbb{E}X^2 = (1)^2 \frac{1}{2} + (-1)^2 \frac{1}{2} = 1$.

Example 5.12. We roll a die and let X be the value that shows. We have previously calculated $\mathbb{E}X = \frac{7}{2}$. So $X - \mathbb{E}X$ equals

$$-\frac{5}{2}, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{5}{2},$$

each with probability $\frac{1}{6}$. So

$$\operatorname{Var} X = (-\tfrac{5}{2})^2 \tfrac{1}{6} + (-\tfrac{3}{2})^2 \tfrac{1}{6} + (-\tfrac{1}{2})^2 \tfrac{1}{6} + (\tfrac{1}{2})^2 \tfrac{1}{6} + (\tfrac{3}{2})^2 \tfrac{1}{6} + (\tfrac{5}{2})^2 \tfrac{1}{6} = \tfrac{35}{12}.$$

Using the fact that the expectation of a constant is the constant we get an alternate expression for the variance.

Proposition 5.2 (Variance)

Suppose X is a random variable with finite first and second moments. Then

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

PROOF. Denote $M := \mathbb{E}X$, then

$$\begin{aligned}\operatorname{Var} X &= \mathbb{E}X^2 - 2\mathbb{E}(XM) + \mathbb{E}(M^2) \\ &= \mathbb{E}X^2 - 2M^2 + M^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2.\end{aligned}$$

$\square$

5.2. Further examples and applications

5.2.1. Discrete random variables. Recall that we defined a discrete random variable in Definition 5.1 as the one taking countably many values. A *random variable* is a function $X : S \longrightarrow \mathbb{R}$, and we can think of it as a numerical value that is random. When we perform an experiment, many times we are interested in some quantity (a function) related to the outcome, instead of the outcome itself. That means we want to attach a numerical value to each outcome. Below are more examples of such variables.

Example 5.13. Toss a coin and define

$$X = \begin{cases} 1 & \text{if outcome is heads (H)} \\ 0 & \text{if outcome is tails (T).} \end{cases}$$

As a random variable, $X(H) = 1$ and $X(T) = 0$. Note that we can perform computations on real numbers but directly not on the sample space $S = \{H, T\}$. This shows the need to covert outcomes to numerical values.

Example 5.14. Let X be the amount of liability (damages) a driver causes in a year. In this case, X can be any dollar amount. Thus X can attain any value in $[0, \infty)$.

Example 5.15. Toss a coin 3 times. Let X be the number of heads that appear, so that X can take the values $0, 1, 2, 3$. What are the associated probabilities to each value?

Solution:

$$\begin{aligned}
\mathbb{P}(X = 0) &= \mathbb{P}((T, T, T)) = \frac{1}{2^3} = \frac{1}{8}, \\
\mathbb{P}(X = 1) &= \mathbb{P}((T, T, H), (T, H, T), (H, T, T)) = \frac{3}{8}, \\
\mathbb{P}(X = 2) &= \mathbb{P}((T, H, H), (H, H, T), (H, T, H)) = \frac{3}{8}, \\
\mathbb{P}(X = 3) &= \mathbb{P}((H, H, H)) = \frac{1}{8}.
\end{aligned}$$

Example 5.16. Toss a coin n times. Let X be the number of heads that occur. This random variable can take the values $0, 1, 2, \ldots, n$. From the binomial formula we see that

$$\mathbb{P}(X = k) = \frac{1}{2^n} \binom{n}{k}.$$

Example 5.17. Suppose we toss a fair coin, and we let X be 1 if we have H and X be 0 if we have T. The probability mass function of this random variable is

$$p_X(x) = \begin{cases} \frac{1}{2} & x = 0 \\ \frac{1}{2} & x = 1, \\ 0 & \text{otherwise} . \end{cases}$$

Often the probability mass function (PMF) will already be given and we can then use it to compute probabilities.

Example 5.18. The PMF of a random variable X taking values in $\mathbb{N} \cup \{0\}$ is given by

$$p_X(i) = e^{-\lambda} \frac{\lambda^i}{i!}, i = 0, 1, 2, \ldots,$$

where λ is a positive real number.

(a) Find $\mathbb{P}(X = 0)$.
Solution: by definition of the PMF we have

$$\mathbb{P}(X = 0) = p_X(0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda}.$$

(b) Find $\mathbb{P}(X > 2)$.
Solution: note that

$$\begin{aligned} \mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leqslant 2) \\ &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2) \\ &= 1 - p_X(0) - p_X(1) - p_X(2) \\ &= 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}. \end{aligned}$$

5.2.2. Expectation. We defined the expectation in Definition 5.2 in the case when X is a discrete random variable X taking values $\{x_i\}_{i \in \mathbb{N}}$. Then for a random variable X with PMF $p_X(x)$ the expectation is given by

$$\mathbb{E}[X] = \sum_{x : p_(x) > 0} x p_X(x) = \sum_{i=1}^{\infty} x_i p_X(x_i).$$

Example 5.19. Suppose again that we have a coin, and let $X(H) = 0$ and $X(T) = 1$. What is $\mathbb{E}X$ if the coin is not necessarily fair?

$$\mathbb{E}X = 0 \cdot p_X(0) + 1 \cdot p_X(1) = \mathbb{P}(T).$$

Example 5.20. Let X be the outcome when we roll a fair die. What is $\mathbb{E}X$?

$$\mathbb{E}X = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2} = 3.5.$$

Note that in the last example X can never be 3.5. This means that the expectation may not be a value attained by X. It serves the purpose of giving an average value for X.

Example 5.21. Let X be the number of insurance claims a person makes in a year. Assume that X can take the values $0, 1, 2, 3\dots$ with $\mathbb{P}(X=0) = \frac{2}{3}$, $\mathbb{P}(X=1) = \frac{2}{9}, \dots, \mathbb{P}(X=n) = \frac{2}{3^{n+1}}$. Find the expected number of claims this person makes in a year.

Solution: Note that X has infinite but countable number of values, hence it is a discrete random variable. We have that $p_X(i) = \frac{2}{3^{i+1}}$. We compute using the definition of expectation,

$$\begin{aligned}
\mathbb{E}X &= 0 \cdot p_X(0) + 1 \cdot p_X(1) + 2 \cdot p_X(2) + \cdots \\
&= 0 \cdot \frac{2}{3} + 1\frac{2}{3^2} + 2\frac{2}{3^3} + 3\frac{2}{3^4} + \cdots \\
&= \frac{2}{3^2}\left(1 + 2\frac{1}{3} + 3\frac{1}{3^2} + 4\frac{1}{3^3} + \cdots\right) \\
&= \frac{2}{9}\left(1 + 2x + 3x^2 + \cdots\right), \text{ where } x = \frac{1}{3} \\
&= \frac{2}{9}\frac{1}{(1-x)^2} = \frac{2}{9\left(1-\frac{1}{3}\right)^2} = \frac{2}{2^2} = \frac{1}{2}.
\end{aligned}$$

Example 5.22. Let $S = \{1, 2, 3, 4, 5, 6\}$ and assume that $X(1) = X(2) = 1$, $X(3) = X(4) = 3$, and $X(5) = X(6) = 5$.

(1) Using the initial definition, the random variable X takes the values $1, 3, 5$ and $p_X(1) = p_X(3) = p_X(5) = \frac{1}{3}$. Then

$$\mathbb{E}X = 1 \cdot \frac{1}{3} + 3\frac{1}{3} + 5\frac{1}{3} = \frac{9}{3} = 3.$$

(2) Using the equivalent definition, we list all of $S = \{1, 2, 3, 4, 5, 6\}$ and then

$$\mathbb{E}X = X(1)\mathbb{P}(\{1\}) + \cdots + X(6) \cdot \mathbb{P}(\{6\}) = 1\frac{1}{6} + 1\frac{1}{6} + 3\frac{1}{6} + 3\frac{1}{6} + 5\frac{1}{6} + 1\frac{1}{6} = 3.$$

5.2.3. The cumulative distribution function (CDF).

We implicitly used this characterization of a random variable, and now we define it.

Definition 5.3 (Cumulative distribution function)

Let X be a random variable. The *cumulative distribution function* (CDF) or the *distribution function* of X is defined as

$$F_X(x) := \mathbb{P}(X \leqslant x),$$

for any $x \in \mathbb{R}$.

Note that if X is discrete and p_X is its PMF , then

$$F(x_0) = \sum_{x \leqslant x_0} p_X(x).$$

Example 5.23. Suppose that X has the following PMF

$$\begin{aligned}
p_X(0) &= \mathbb{P}(X = 0) = \frac{1}{8} \\
p_X(1) &= \mathbb{P}(X = 1) = \frac{3}{8} \\
p_X(2) &= \mathbb{P}(X = 2) = \frac{3}{8} \\
p_X(3) &= \mathbb{P}(X = 3) = \frac{1}{8}.
\end{aligned}$$

Find the CDF for X and plot the graph of the CDF.

Solution: summing up the probabilities up to the value of x we get the following

$$F_X(x) = \begin{cases} 0 & -\infty < x < 0, \\ \frac{1}{8} & 0 \leqslant x < 1, \\ \frac{4}{8} & 1 \leqslant x < 2, \\ \frac{7}{8} & 2 \leqslant x < 3, \\ 1 & 3 \leqslant x < \infty. \end{cases}$$

This is a step function shown in the Figure below on page 71.

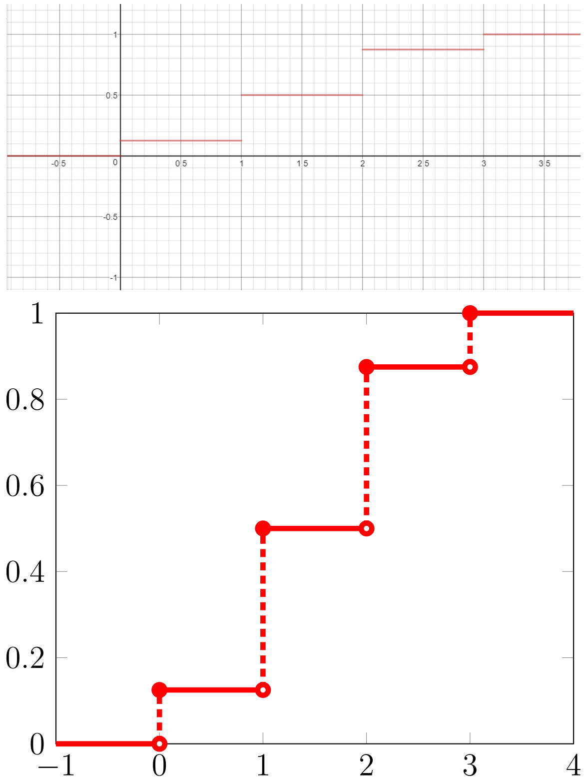
Proposition 5.3 (Properties of cumulative distribution functions(CDF))

1. F is nondecreasing , that is, if $x < y$, then $F(x) \leqslant F(y)$.
2. $\lim_{x\to\infty} F(x) = 1$.
3. $\lim_{x\to-\infty} F(x) = 0$.
4. F is right continuous, that is, $\lim\limits_{u \downarrow x} F_X(u) = F_X(x)$, where $u \downarrow x$ means that u approaches x from above (from the right).

Example 5.24. Let X have distribution

$$F_X(x) = \begin{cases} 0 & x < 0, \\ \frac{x}{2} & 0 \leqslant x < 1, \\ \frac{2}{3} & 1 \leqslant x < 2, \\ \frac{11}{12} & 2 \leqslant x < 3, \\ 1 & 3 \leqslant x. \end{cases}$$

(a) Compute $\mathbb{P}(X < 3)$.
Solution: We have that $\mathbb{P}(X < 3) = \lim\limits_{n\to\infty} \mathbb{P}\left(X \leqslant 3 - \frac{1}{n}\right) = \lim\limits_{n\to\infty} F_X\left(3 - \frac{1}{n}\right) = \frac{11}{12}$.



Above are two Graphs for $F_X(x)$ in Example 5.23.

(b) Compute $\mathbb{P}(X = 1)$.
Solution: We have that

$$\mathbb{P}(X = 1) = \mathbb{P}(X \leqslant 1) - \mathbb{P}\left(X < 1\right) = F_X(1) - \lim_{x \to 1} \frac{x}{2} = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$$

(c) Compute $\mathbb{P}(2 < X \leqslant 4)$.
Solution: We have that

$$\mathbb{P}(2 < X \leqslant 4) = F_X(4) - F_X(2) = \frac{1}{12}.$$

5.2.4. Expectation of a function of a random variable. Given a random variable X we would like to compute the expected value of expressions such as X^2, e^X or $\sin X$. How can we do this?

Example 5.25. Let X be a random variable whose PMF is given by

$$\begin{aligned} \mathbb{P}\left(X = -1\right) &= 0.2, \\ \mathbb{P}\left(X = 0\right) &= 0.5, \\ \mathbb{P}\left(X = 1\right) &= 0.3. \end{aligned}$$

Let $Y = X^2$, find $\mathbb{E}[Y]$.

Solution: Note that Y takes the values 0^2, $(-1)^2$ and 1^2, which reduce to 0 or 1. Also notice that $p_Y(1) = 0.2 + 0.3 = 0.5$ and $p_Y(0) = 0.5$. Thus, $\mathbb{E}[Y] = 0 \cdot 0.5 + 1 \cdot 0.5 = 0.5$.

Note that $\mathbb{E}X^2 = 0.5$. While $(\mathbb{E}X)^2 = 0.01$ since $\mathbb{E}X = 0.3 - 0.2 = 0.1$. Thus in general

$$\mathbb{E}X^2 \neq (\mathbb{E}X)^2 .$$

In general, there is a formula for $g(X)$ where g is function that uses the fact that $g(X)$ will be $g(x)$ for some x such that $X = x$. We recall Theorem 5.2. If X is a discrete distribution that takes the values $x_i, i \geq 1$ with probability $p_X(x_i)$, respectively, then for any real valued function g we have that

$$\mathbb{E}\left[g\left(X\right)\right] = \sum_{i=1}^{\infty} g\left(x_i\right) p_X(x_i).$$

Note that

$$\mathbb{E}X^2 = \sum_{i=1}^{\infty} x_i^2 p_X(x_i)$$

will be useful.

Example 5.26. Let us revisit the previous example. Let X denote a random variable such that

$$\begin{aligned}\mathbb{P}\left(X = -1\right) &= 0.2\\ \mathbb{P}\left(X = 0\right) &= 0.5\\ \mathbb{P}\left(X = 1\right) &= 0.3.\end{aligned}$$

Let $Y = X^2$. Find $\mathbb{E}Y$.

Solution: We have that

$$\mathbb{E}X^2 = \sum_{i=1}^{\infty} x_i^2 p_X(x_i) = (-1)^2(0.2) + 0^2(0.5) + 1^2(0.3) = 0.5$$

5.2.5. Variance. The variance of a random variable is a measure of how spread out the values of X are. The expectation of a random variable is quantity that help us differentiate between random variables, but it does not tell us how spread out its values are. For example, consider

$$\begin{aligned} X &= 0 \text{ with probability } 1\\ Y &= \begin{cases} -1 & p = \frac{1}{2}\\ 1 & p = \frac{1}{2}\end{cases}\\ Z &= \begin{cases} -100 & p = \frac{1}{2}\\ 100 & p = \frac{1}{2}\end{cases}.\end{aligned}$$

What are the expected values? The are $0, 0$ and 0. But there is much greater spread in Z than Y and Y than X. Thus expectation is not enough to detect spread, or variation.

Example 5.27. Calculate $\text{Var}(X)$ if X represents the outcome when a fair die is rolled.

Solution: recall that we showed Equation 5.2 to find the variance

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

Previously we calculated that $\mathbb{E}X = \frac{7}{2}$. Thus we only need to find the second moment

$$\mathbb{E}X^2 = 1^2\left(\frac{1}{6}\right) + \cdots + 6^2\frac{1}{6} = \frac{91}{6}.$$

Using our formula we have that

$$\operatorname{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

Another useful formula is the following.

Proposition 5.4

For any constants $a, b \in \mathbb{R}$ we have that $\operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X)$.

Proof. By Equation 5.2 and linearity of expectation

$$\begin{aligned}
\operatorname{Var}(aX+b) &= \mathbb{E}(aX+b)^2 - (\mathbb{E}(aX+b))^2 \\
&= \mathbb{E}\left(a^2X^2 + 2abX + b^2\right) - (a\mathbb{E}X + b)^2 \\
&= a^2\mathbb{E}X^2 + 2ab\mathbb{E}X + b^2 - a^2(\mathbb{E}X)^2 - 2ab\mathbb{E}X - b^2 \\
&= a^2\mathbb{E}X^2 - a^2(\mathbb{E}X)^2 = a^2\operatorname{Var}(X).
\end{aligned}$$

$\square$

5.3. Exercises

Exercise 5.1. Three balls are randomly chosen with replacement from an urn containing 5 blue, 4 red, and 2 yellow balls. Let X denote the number of red balls chosen.

(a) What are the possible values of X?
(b) What are the probabilities associated to each value?

Exercise 5.2. Two cards are chosen from a standard deck of 52 cards. Suppose that you win \$2 for each heart selected, and lose \$1 for each spade selected. Other suits (clubs or diamonds) bring neither win nor loss. Let X denote your winnings. Determine the probability mass function of X.

Exercise 5.3. A financial regulator from the FED will evaluate two banks this week. For each evaluation, the regulator will choose with equal probability between two different stress tests. Failing under test one costs a bank 10K fee, whereas failing test 2 costs 5K. The probability that the first bank fails any test is 0.4. Independently, the second bank will fail any test with 0.5 probability. Let X denote the total amount of fees the regulator can obtain after having evaluated both banks. Determine the cumulative distribution function of X.

Exercise 5.4. Five buses carry students from Hartford to campus. Each bus carries, respectively, 50, 55, 60, 65, and 70 students. One of these students and one bus driver are picked at random.

(a) What is the expected number of students sitting in the same bus that carries the randomly selected student?
(b) Let Y be the number of students in the same bus as the randomly selected driver. Is $\mathbb{E}[Y]$ larger than the expectation obtained in the previous question?

Exercise 5.5. Two balls are chosen randomly from an urn containing 8 white balls, 4 black, and 2 orange balls. Suppose that we win \$2 for each black ball selected and we lose \$1 for each white ball selected. Let X denote our winnings.

(a) What are the possible values of X?
(b) What are the probabilities associated to each value?

Exercise 5.6. A card is drawn at random from a standard deck of playing cards. If it is a heart, you win \$1. If it is a diamond, you have to pay \$2. If it is any other card, you win \$3. What is the expected value of your winnings?

Exercise 5.7. The game of roulette consists of a small ball and a wheel with 38 numbered pockets around the edge that includes the numbers $1-36, 0$ and 00. As the wheel is spun, the ball bounces around randomly until it settles down in one of the pockets.

(a) Suppose you bet \$1 on a single number and random variable X represents the (monetary) outcome (the money you win or lose). If the bet wins, the payoff is \$35 and you get your money back. If you lose the bet then you lose your \$1. What is the expected profit on a 1 dollar bet?

(b) Suppose you bet \$1 on the numbers $1 - 18$ and random variable X represents the (monetary) outcome (the money you win or lose). If the bet wins, the payoff is \$1 and you get your money back. If you lose the bet then you lose your \$1. What is the expected profit on a 1 dollar bet ?

Exercise 5.8. An insurance company finds that Mark has a 8% chance of getting into a car accident in the next year. If Mark has any kind of accident then the company guarantees to pay him \$10,000. The company has decided to charge Mark a \$200 premium for this one year insurance policy.

(a) Let X be the amount profit or loss from this insurance policy in the next year for the insurance company. Find $\mathbb{E}X$, the expected return for the Insurance company? Should the insurance company charge more or less on its premium?
(b) What amount should the insurance company charge Mark in order to guarantee an expected return of \$100?

Exercise 5.9. A random variable X has the following probability mass function: $p_X(0) = \frac{1}{3}$, $p_X(1) = \frac{1}{6}$, $p_X(2) = \frac{1}{4}$, $p_X(3) = \frac{1}{4}$. Find its expected value, variance, and standard deviation, and plot its CDF.

Exercise 5.10. Suppose X is a random variable such that $\mathbb{E}[X] = 50$ and $\operatorname{Var}(X) = 12$. Calculate the following quantities.

(a) $\mathbb{E}[X^2]$,
(b) $\mathbb{E}[3X + 2]$,
(c) $\mathbb{E}\left[(X + 2)^2\right]$,
(d) $\operatorname{Var}[-X]$,
(e) $\operatorname{SD}(2X)$.

Exercise 5.11. Does there exist a random variable X such that $\mathbb{E}[X] = 4$ and $\mathbb{E}[X^2] = 10$? Why or why not? (Hint: look at its variance)

Exercise 5.12. A box contains 25 peppers of which 5 are red and 20 green. Four peppers are randomly picked from the box. What is the expected number of red peppers in this sample of four?

5.4. Selected solutions

Solution to Exercise 5.1:

(a) X can take the values $0, 1, 2$ and 3.
(b) Since balls are withdrawn with replacement, we can think of *choosing red* as a success and apply Bernoulli trials with $p = \mathbb{P}(\text{red}) = \frac{4}{11}$. Then, for each $k = 0, 1, 2, 3$ we have

$$\mathbb{P}(X = k) = \binom{3}{k} \left(\frac{4}{11}\right)^k \cdot \left(\frac{7}{11}\right)^{3-k}.$$

Solution to Exercise 5.2: The random variable X can take the values $-2, -1, 0, 1, 2, 4$. Moreover,

$$\begin{aligned}
\mathbb{P}(X = -2) &= \mathbb{P}(2\spadesuit) = \frac{\binom{13}{2}}{\binom{52}{2}},\\
\mathbb{P}(X = -1) &= \mathbb{P}(1\spadesuit \text{ and } 1(\diamondsuit \text{ or } \clubsuit)) = \frac{13 \cdot 26}{\binom{52}{2}},\\
\mathbb{P}(X = 0) &= \mathbb{P}(2(\diamondsuit \text{ or } \clubsuit)) = \frac{\binom{26}{2}}{\binom{52}{2}},\\
\mathbb{P}(X = 1) &= \mathbb{P}(1\heartsuit \text{ and } 1\spadesuit) = \frac{13 \cdot 13}{\binom{52}{2}},\\
\mathbb{P}(X = 2) &= \mathbb{P}(1\heartsuit \text{ and } 1(\diamondsuit \text{ or } \clubsuit)) = \mathbb{P}(X = -1),\\
\mathbb{P}(X = 4) &= \mathbb{P}(2\heartsuit) = \mathbb{P}(X = -2).
\end{aligned}$$

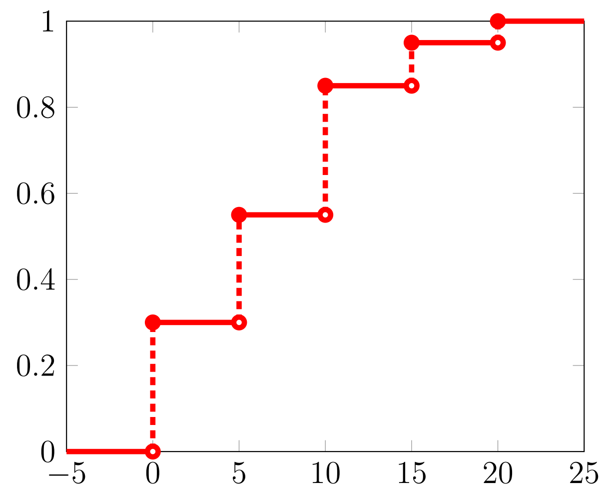
Thus the probability mass function is given by $p_X(x) = \mathbb{P}(X = x)$ for $x = -2, -1, 0, 1, 2, 4$ and $p_X(x) = 0$ otherwise.

Solution to Exercise 5.3: The random variable X can take the values $0, 5, 10, 15$ and 20 depending on which test was applied to each bank, and if the bank fails the evaluation or not. Denote by B_i the event that the *ith bank fails* and by T_i the event that *test i applied*. Then

$$\begin{aligned}
&\mathbb{P}(T_1) = \mathbb{P}(T_2) = 0.5, \mathbb{P}(B_1) = \mathbb{P}(B_1 \mid T_1) = \mathbb{P}(B_1 \mid T_2) = 0.4\\
&\mathbb{P}(B_2) = \mathbb{P}(B_2 \mid T_1) = \mathbb{P}(B_2 \mid T_2) = 0.5.
\end{aligned}$$

Since banks and tests are independent we have

$$\begin{aligned}
\mathbb{P}(X = 0) &= \mathbb{P}(B_1^c \cap B_2^c) = \mathbb{P}(B_1^c) \cdot \mathbb{P}(B_2^c) = 0.6 \cdot 0.5 = 0.3,\\
\mathbb{P}(X = 5) &= \mathbb{P}(B_1)\mathbb{P}(T_1)\mathbb{P}(B_2^c) + \mathbb{P}(B_1^c)\mathbb{P}(B_2)\mathbb{P}(T_2) = 0.25,\\
\mathbb{P}(X = 10) &= \mathbb{P}(B_1)\mathbb{P}(T_1)\mathbb{P}(B_2^c) + \mathbb{P}(B_1)\mathbb{P}(T_2)\mathbb{P}(B_2)\mathbb{P}(T_2) + \mathbb{P}(B_1^c)\mathbb{P}(B_2)\mathbb{P}(T_1) = 0.3\\
\mathbb{P}(X = 15) &= \mathbb{P}(B_1)\mathbb{P}(T_1)\mathbb{P}(B_2)\mathbb{P}(T_2) + \mathbb{P}(B_1)\mathbb{P}(T_2)\mathbb{P}(B_2)\mathbb{P}(T_1) = 0.1\\
\mathbb{P}(X = 20) &= \mathbb{P}(B_1)\mathbb{P}(T_1)\mathbb{P}(B_2)\mathbb{P}(T_1) = 0.05.
\end{aligned}$$



The graph of the probability distribution function for Exercise 5.3

The probability distribution function is given by

$$F_X(x) = \begin{cases} 0 & x < 0, \\ 0.3 & 0 \leqslant x < 5, \\ 0.55 & 5 \leqslant x < 10, \\ 0.85 & 10 \leqslant x < 15, \\ 0.95 & 15 \leqslant x < 20, \\ 1 & x \geqslant 20. \end{cases}$$

Solution to Exercise 5.4: Let X denote the number of students in the bus that carries the randomly selected student.

(a) In total there are 300 students, hence $\mathbb{P}(X = 50) = \frac{50}{300}$, $\mathbb{P}(X = 55) = \frac{55}{300}$, $\mathbb{P}(X = 60) = \frac{60}{300}$, $\mathbb{P}(X = 65) = \frac{65}{300}$ and $\mathbb{P}(X = 70) = \frac{70}{300}$. The expected value of X is thus

$$\mathbb{E}[X] = 50\frac{50}{300} + 55\frac{55}{300} + 60\frac{60}{300} + 65\frac{65}{300} + 70\frac{70}{300} \approx 60.8333.$$

(b) In this case, the probability of choosing a bus driver is $\frac{1}{5}$, so that

$$\mathbb{E}[Y] = \frac{1}{5}(50 + 55 + 60 + 65 + 70) = 60$$

which is slightly less than the previous one.

Solution to Exercise 5.5(A): Note that $X = -2, -1, -0, 1, 2, 4$.

Solution to Exercise 5.5(B): below is the list of all probabilities.

$$\mathbb{P}(X=4)=\mathbb{P}(\{BB\})=\frac{\binom{4}{2}}{\binom{14}{2}}=\frac{6}{91},$$
$$\mathbb{P}(X=0)=\mathbb{P}(\{OO\})=\frac{\binom{2}{2}}{\binom{14}{2}}=\frac{1}{91}$$
$$\mathbb{P}(X=2)=\mathbb{P}(\{BO\})=\frac{\binom{4}{1}\binom{2}{1}}{\binom{14}{2}}=\frac{8}{91},$$
$$\mathbb{P}(X=-1)=\mathbb{P}(\{WO\})=\frac{\binom{8}{1}\binom{2}{1}}{\binom{14}{2}}=\frac{16}{91},$$
$$\mathbb{P}(X=1)=\mathbb{P}(\{BW\})=\frac{\binom{4}{1}\binom{8}{1}}{\binom{14}{2}}=\frac{32}{91},$$
$$\mathbb{P}(X=-2)=\mathbb{P}(\{WW\})=\frac{\binom{8}{2}}{\binom{14}{2}}=\frac{28}{91}$$

Solution to Exercise 5.6:

$$\mathbb{E}X=1\cdot\frac{1}{4}+(-2)\frac{1}{4}+3\cdot\frac{1}{2}=\frac{5}{4}$$

Solution to Exercise 5.7(A): The expected profit is $\mathbb{E}X=35\cdot\left(\frac{1}{38}\right)-1\cdot\frac{37}{38}=-\$0.0526$.

Solution to Exercise 5.7(B): If you will then your profit will be \$1. If you lose then you lose your \$1 bet. The expected profit is $\mathbb{E}X=1\cdot\left(\frac{18}{38}\right)-1\cdot\frac{20}{38}=-\$0.0526$.

Solution to Exercise 5.8(A): If Mark has no accident then the company makes a profit of 200 dollars. If Mark has an accident they have to pay him 10,000 dollars, but regardless they received 200 dollars from him as an yearly premium. We have

$$\mathbb{E}X=(200-10,000)\cdot(0.08)+200\cdot(0.92)=-600.$$

On average the company will lose \$600 dollars. Thus the company should charge more.

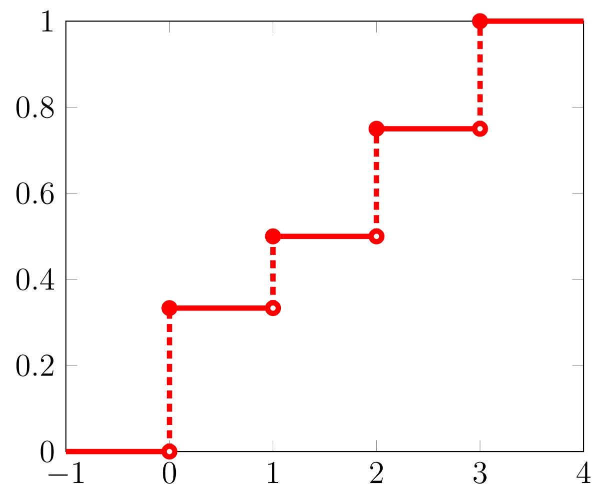
Solution to Exercise 5.8(B): Let P be the premium. Then in order to guarantee an expected return of 100 then

$$100=\mathbb{E}X=(P-10,000)\cdot(0.08)+P\cdot(0.92)$$

and solving for P we get $P=\$900$.

Solution to Exercise 5.9: we start with the expectation

$$\mathbb{E}X=0\cdot\frac{1}{3}+1\cdot\frac{1}{6}+2\cdot\frac{1}{4}+3\cdot\frac{1}{4}=\frac{34}{24}.$$



The plot of the CDF for Exercise 5.9

Now to find the variance we have

$$\begin{aligned}\operatorname{Var}(X) &= \mathbb{E}\left[X^2\right] - (\mathbb{E}X)^2 \\ &= 0^2 \cdot \frac{1}{3} - 1^2\frac{1}{6} + 2^2 \cdot \frac{1}{4} + 3^2 \cdot \frac{1}{4} - \left(\frac{34}{24}\right)^2 \\ &= \frac{82}{24} - \frac{34^2}{24^2} = \frac{812}{24^2}.\end{aligned}$$

Taking the square root gives us

$$\operatorname{SD}(X) = \frac{2\sqrt{203}}{24}.$$

Solution to Exercise 5.10(A): Since $\operatorname{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}X)^2 = 12$ then

$$\mathbb{E}\left[X^2\right] = \operatorname{Var}(X) + (\mathbb{E}X)^2 = 12 + 50^2 = 2512.$$

Solution to Exercise 5.10(B):

$$\mathbb{E}\left[3X+2\right] = 3\mathbb{E}\left[X\right] + \mathbb{E}\left[2\right] = 3 \cdot 50 + 2 = 152.$$

Solution to Exercise 5.10(C):

$$\mathbb{E}\left[(X+2)^2\right] = \mathbb{E}\left[X^2\right] + 4\mathbb{E}\left[X\right] + 4 = 2512 + 4 \cdot 50 + 4 = 2716.$$

Solution to Exercise 5.10(D):

$$\operatorname{Var}\left[-X\right] = (-1)^2 \operatorname{Var}(X) = 12$$

Solution to Exercise 5.10(E):

$$\operatorname{SD}(2X) = \sqrt{\operatorname{Var}(2X)} = \sqrt{2^2 \operatorname{Var}(X)} = \sqrt{48} = 2\sqrt{12}.$$

Solution to Exercise 5.11: Using the hint let's compute the variance of this random variable which would be $\operatorname{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}X)^2 = 10 - 4^2 = -6$. But we know a random variable cannot have a negative variance. Thus no such a random variable exists.