Introduction to Probability
Lecture 3: Expectation properties, variance, discrete distributions
Mateja Jamnik, Thomas Sauerwald

University of Cambridge, Department of Computer Science and Technology
email: {mateja.jamnik, thomas.sauerwald}@cl.cam.ac.uk
Outline

Properties of expectation

Variance

Bernoulli discrete random variable

Binomial discrete random variable
Properties of expectation: linearity

Linearity of expectation

Expectations preserve linearity: if $a$ and $b$ are constants, then

$$E\left[aX + b\right] = aE\left[X\right] + b$$

Proof:

Let the event be a roll of a 6-sided die, $X$ its random variable, and $Y$ another random variable where $Y = 3X + 1$. What are the expected values $E\left[X\right]$ and $E\left[Y\right]$?

Answer

We know from last time that $E\left[X\right] = 3.5$. Thus $E\left[Y\right] = 3 \cdot 3.5 + 1 = 11.5$. 

Properties of expectation: linearity

**Linearity of expectation**

Expectations preserve linearity: if $a$ and $b$ are constants, then

$$E[aX + b] = aE[X] + b$$

**Proof:**

Let the event be a roll of a 6-sided die, $X$ its random variable, and $Y$ another random variable where $Y = 3X + 1$. What are the expected values $E[X]$ and $E[Y]$?

Answer: 

---

**Example**

Let the event be a roll of a 6-sided die, $X$ its random variable, and $Y$ another random variable where $Y = 3X + 1$. What are the expected values $E[X]$ and $E[Y]$?
Properties of expectation: additivity

Additivity of expectation

Expectation of a sum is equal to the sum of expectations: if $X$ and $Y$ are any random variables on the same sample space then

$$E[X + Y] = E[X] + E[Y]$$
Properties of expectation: additivity

Additivity of expectation

Expectation of a sum is equal to the sum of expectations: if \( X \) and \( Y \) are any random variables on the same sample space then

\[
E[X + Y] = E[X] + E[Y]
\]

Example

Let the events be rolls of 2 dice, and \( X \) the random variable for the roll of die 1, and \( Y \) for the roll of die 2. What is the expected value of the sum of the rolls of the two dice?

Answer  

\[
E[X + Y] = E[X] + E[Y] = 3.5 + 3.5 = 7
\]
Law of the unconscious statistician (LOTUS)

Let $X$ be a random variable, and $Y$ another random variable that is a function of $X$, so $Y = g(X)$. Let $p(x)$ be a PMF of $X$. Then

$$E[Y] = E[g(X)] = \sum_{x:p(x)>0} g(x)p(x)$$

Note how now we no longer need to know PMF of $Y$. 

LOTUS is also known as the expected value of a function of a random variable. Note that the properties of expectation let you avoid defining difficult PMFs.
Properties of expectation: LOTUS

Law of the unconscious statistician (LOTUS)

Let $X$ be a random variable, and $Y$ another random variable that is a function of $X$, so $Y = g(X)$. Let $p(x)$ be a PMF of $X$. Then

$$
E[Y] = E[g(X)] = \sum_{x:p(x)>0} g(x)p(x)
$$

Note how now we no longer need to know PMF of $Y$.

- LOTUS is also known as expected value of a function of a random variable.
Law of the unconscious statistician (LOTUS)

Let $X$ be a random variable, and $Y$ another random variable that is a function of $X$, so $Y = g(X)$. Let $p(x)$ be a PMF of $X$. Then

$$
E[Y] = E[g(X)] = \sum_{x: p(x) > 0} g(x)p(x)
$$

Note how now we no longer need to know PMF of $Y$.

- LOTUS is also known as **expected value of a function of a random variable**.
- Note that the properties of expectation let you avoid defining difficult PMFs.
Law of the unconscious statistician (LOTUS)

Let $X$ be a random variable, and $Y$ another random variable that is a function of $X$, so $Y = g(X)$. Let $p(x)$ be a PMF of $X$. Then

$$E[Y] = E[g(X)] = \sum_{x:p(x)>0} g(x)p(x)$$

Note how now we no longer need to know PMF of $Y$.

- LOTUS is also known as **expected value of a function of a random variable**.

- Note that the properties of expectation let you avoid defining difficult PMFs.

- Let $X$ be a discrete RV, then:
  - $E[X^2]$ is know as the **second moment of $X$**.
  - $E[X^n]$ is know as the $n^{th}$ moment of $X$.  

---

Intro to Probability

Properties of expectation

5
Let $X$ be a discrete random variable that ranges over the values $\{-1, 0, 1\}$, and respective probabilities $P[ X = -1 ] = 0.2$, $P[ X = 0 ] = 0.5$ and $P[ X = 1 ] = 0.3$. Let another random variable $Y = X^2$ (second moment). What is $E[ Y ]$?

Note that $Y = g(X) = X^2$ and $E[ Y ] = E[ g(X) ] = \sum_{x: p(x) > 0} g(x) p(x)$, thus
Outline

Properties of expectation

Variance

Bernoulli discrete random variable

Binomial discrete random variable
Spread in the distribution

Expectation is a useful statistic, but it does not give a detailed view of the PMF. Consider these 3 distributions (PMFs).
Spread in the distribution

Expectation is a useful statistic, but it does not give a detailed view of the PMF. Consider these 3 distributions (PMFs).

- Expectation is the same for all distributions: $E[X] = 3$.
- First has the greatest spread, the third has the least spread.
- But the "spread" or "dispersion" of $X$ in the distribution is very different!
Spread in the distribution

Expectation is a useful statistic, but it does not give a detailed view of the PMF. Consider these 3 distributions (PMFs).

- Expectation is the same for all distributions: $E[X] = 3$.
- First has the greatest spread, the third has the least spread.
- But the "spread" or "dispersion" of $X$ in the distribution is very different!
- **Variance**, $V[X]$ defines a formal quantification of "spread".
- Several ways to quantify: it uses average square distance from the mean.

Intro to Probability

Variance
Definition of variance

Variance

The variance of a discrete random variable \( X \) with expected value (mean) \( \mu \) is:

\[
V[ X ] = E[ (X - \mu)^2 ]
\]

When computing the variance, we often use a different form of the same equation:

\[
V[ X ] = E[ X^2 ] - (E[ X ])^2
\]

Proof:

Note:

- \( V[ X ] \geq 0 \)
- AKA: Second central moment, or square of the standard deviation
Example with a die roll

Let $X$ be the value on one roll of a 6-sided fair die. Recall that $E[ X ] = \frac{7}{2} = 3.5$. What is $V[ X ]$?

Answer

Using $V[ X ] = E[ X^2 ] - (E[ X ])^2$:

Using $V[ X ] = E[ (X - \mu)^2 ] = E[ (X - E[ X ])^2 ]$:
Example of spread

Example

Let $X$, $Y$ and $Z$ be discrete random variables with the range $X : \{10\}$ and probability 1, and $Y : \{11, 9\}$ and $Z : \{110, -90\}$ with equal probabilities $\frac{1}{2}$. Compute expectation and variance for $X$, $Y$ and $Z$.

Answer

- $E[X] = \sum x \cdot p(x) = 10 \cdot 1 = 10$
  - $E[X^2] = \sum x^2 \cdot p(x) = 100 \cdot \frac{1}{2} = 50$
  - $V[X] = E[X^2] - 20 \cdot 10 + 100 = 50 - 200 + 100 = 0$

- $E[Y] = \sum y \cdot p(y) = (11 \cdot \frac{1}{2}) + (9 \cdot \frac{1}{2}) = 10$
  - $E[Y^2] = \sum y^2 \cdot p(y) = 121 + 9 \cdot \frac{1}{2} = 122.5$
  - $V[Y] = E[Y^2] - 20 \cdot 10 + 100 = 122.5 - 200 + 100 = 22.5$

- $E[Z] = \sum z \cdot p(z) = (110 \cdot \frac{1}{2}) + (-90 \cdot \frac{1}{2}) = 10$
  - $E[Z^2] = \sum z^2 \cdot p(z) = 12100 + 81 \cdot \frac{1}{2} = 12140.5$
  - $V[Z] = E[Z^2] - 20 \cdot 10 + 100 = 12140.5 - 200 + 100 = 12040.5$
Let $X$, $Y$ and $Z$ be discrete random variables with the range $X : \{10\}$ and probability 1, and $Y : \{11, 9\}$ and $Z : \{110, -90\}$ with equal probabilities $\frac{1}{2}$. Compute expectation and variance for $X$, $Y$ and $Z$.

a) $\mathbb{E}[X] = \sum_x x p(x) = 10 \cdot 1 = 10$

$\mathbb{V}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[(X - 10)^2\right]$

$= (X - 10)^2 p(x) = 0^2 \cdot 1 = 0$

\begin{itemize}
\item \textbf{Answer} \\
\end{itemize}
Example of spread

Let $X$, $Y$ and $Z$ be discrete random variables with the range $X : \{10\}$ and probability 1, and $Y : \{11, 9\}$ and $Z : \{110, -90\}$ with equal probabilities $\frac{1}{2}$. Compute expectation and variance for $X$, $Y$ and $Z$.

**a)** $E[X] = \sum_x x p(x) = 10 \cdot 1 = 10$

\[
V[X] = E[(X - E[X])^2] = E[(X - 10)^2]
\]

\[
= (X - 10)^2 p(x) = 0^2 \cdot 1 = 0
\]

**b)** $E[Y] = (11)(0.5) + (9)(0.5) = 10$

\[
\]

\[
= (11 - 10)^2(0.5) + (9 - 10)^2(0.5) = 1
\]
Let $X$, $Y$ and $Z$ be discrete random variables with the range $X : \{10\}$ and probability 1, and $Y : \{11, 9\}$ and $Z : \{110, -90\}$ with equal probabilities $\frac{1}{2}$. Compute expectation and variance for $X$, $Y$ and $Z$.

a) $E[X] = \sum_x xp(x) = 10 \cdot 1 = 10$

\[
V[X] = E[(X - E[X])^2] = E[(X - 10)^2] = (X - 10)^2 p(x) = 0^2 \cdot 1 = 0
\]

b) $E[Y] = (11)(0.5) + (9)(0.5) = 10$

\[
V[Y] = E[(Y - E[Y])^2] = E[(Y - 10)^2] = (11 - 10)^2 (0.5) + (9 - 10)^2 (0.5) = 1
\]

c) $E[Z] = (110)(0.5) + (-90)(0.5) = 10$

\[
V[Z] = E[(Z - E[Z])^2] = E[(Z - 10)^2] = (110 - 10)^2 (0.5) + (-90 - 10)^2 (0.5) = 100^2 = 10000
\]
Standard deviation

- Standard deviation is a kind of average distance of a sample of the mean, i.e., a root mean square (RMS) average.
- Variance is the square of this average distance.
Standard deviation

- Standard deviation is a kind of average distance of a sample of the mean, i.e., a root mean square (RMS) average.
- Variance is the square of this average distance.

\[
\text{SD}[X] = \sqrt{V[X]}
\]

Note:
- \(E[X]\) and \(V[X]\) are real numbers, not RVs.
- \(V[X]\) is expressed in units of the values in the range of \(X^2\).
- \(SD[X]\) is expressed in units of the values in the range of \(X\).
- For the spread example above: \(SD[X] = 0, SD[Y] = 1, SD[Z] = 100.\)
Properties of variance

- Property 1: \( V[X] = E[X^2] - (E[X])^2 \)
Properties of variance

- **Property 1:** \( \text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \)
- **Property 2:** variance is **not** linear: \( \text{Var}(aX + b) = a^2 \text{Var}(X) \)
Properties of variance

- **Property 1:** \( \text{V}[X] = \text{E}[X^2] - (\text{E}[X])^2 \)

- **Property 2:** Variance is not linear: \( \text{V}[aX + b] = a^2 \text{V}[X] \)

Proof:

\[
\begin{align*}
\text{V}[aX + b] &= \text{E}[(aX + b)^2] - (\text{E}[aX + b])^2 \\
 &= \text{E}[a^2X^2 + 2abX + b^2] - (a\text{E}[X] + b)^2 \\
 &= a^2\text{E}[X^2] + 2ab\text{E}[X] + b^2 - (a^2(\text{E}[X])^2 + 2ab\text{E}[X] + b^2) \\
 &= a^2\text{E}[X^2] - (a^2(\text{E}[X])^2) = a^2(\text{E}[X^2] - (\text{E}[X])^2) \\
 &= a^2\text{V}[X]
\end{align*}
\]
Summary of expectation and variance for discrete RV

\[ E[X] = \sum_{x: p(x) > 0} x \cdot p(x) \]

\[ \text{Properties of Expectation} \]

\[ E[X + Y] = E[X] + E[Y] \]

\[ E[aX + b] = a \cdot E[X] + b \]

\[ E[g(X)] = \sum_x g(x) \cdot p_X(x) \]

\[ \text{Properties of Variance} \]

\[ V[X] = E[(X - \mu)^2] \]

\[ V[aX + b] = a^2 \cdot V[X] \]

\[ V[X] = E[X^2] - (E[X])^2 \]
Parametric/standard discrete random variables

- There is deluge of classic RV abstractions that show up in problems.
- They give rise to significant discrete distributions.
- If problem fits, use precalculated (parametric) PMF, expectation, variance and other properties by providing parameters of the problem.

- We will cover the following RVs:
  1. Bernoulli
  2. Binomial
  3. Poisson
  4. Geometric
  5. Negative Binomial
  6. Hypergeometric
Outline

Properties of expectation

Variance

Bernoulli discrete random variable

Binomial discrete random variable
Bernoulli discrete random variable

A Bernoulli RV $X$ maps "success" of an experiment to 1 and "failure" to 0. It is AKA indicator RV, boolean RV. $X$ is "Bernoulli RV with parameter $p$", where $\mathbb{P}[\text{"success"}] = p$ and so PMF $p(1) = p$.

$$X \sim \text{Ber}(p)$$

Range: $\{0, 1\}$

PMF: $\mathbb{P}[X = 1] = p(1) = p$

$\mathbb{P}[X = 0] = p(0) = 1 - p$

Expectation: $\mathbb{E}[X] = p$

Variance: $\mathbb{V}[X] = p(1 - p)$

Examples: coin toss, random binary digit, if someone likes a film, the gender of a newborn baby, pas/fail of you taking an exam.
You watch a film on Netflix. At the end you click "like" with probability $p$. Define a RV representing this event.

**Answer**

$X$: 1 if "like"-d

$X \sim \text{Ber}(p)$

$P[X = 1] = p$, $P[X = 0] = 1 - p$
You watch a film on Netflix. At the end you click "like" with probability $p$. Define a RV representing this event.

Answer

Two fair 6-sided dice are rolled. Define a random variable $X$ for a successful roll of two 6's, and failure for anything else.

Answer
Outline

Properties of expectation

Variance

Bernoulli discrete random variable

Binomial discrete random variable
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: $\{0, 1, \ldots, n\}$

PMF: $k \in \{0, 1, \ldots, n\}$

$$P[X = k] = p(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Expectation: $E[X] = np$

Variance: $V[X] = np(1 - p)$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 

Intro to Probability
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: \{0, 1, \ldots, n\}
PMF: $k \in \{0, 1, \ldots, n\}$

$$P[X = k] = p(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Expectation: $E[X] = np$
Variance: $V[X] = np(1 - p)$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1.$
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: $\{0, 1, \ldots, n\}$

PMF: $k \in \{0, 1, \ldots, n\}$

$$P[X = k] = p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Expectation: $E[X] = np$

Variance: $V[X] = np(1-p)$

Examples: # heads in $n$ coin tosses, # of 1's in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

\[ X \sim Bin(n, p) \]

- **Range:** \{0, 1, \ldots, n\}
- **PMF:** $k \in \{0, 1, \ldots, n\}$

\[ P[X = k] = p(k) = \binom{n}{k} p^k (1 - p)^{n-k} \]

- **Expectation:** $E[X] = np$
- **Variance:** $V[X] = np(1 - p)$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 

---

**Intro to Probability**

Binomial discrete random variable

---
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: $\{0, 1, \ldots, n\}$

PMF: $k \in \{0, 1, \ldots, n\}$

$$P[X = k] = p(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Expectation: $E[X] = np$

Variance: $V[X] = np(1 - p)$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 
Binomial

A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: $\{0, 1, \ldots, n\}$

Probability that $X$ takes on the value $k$:

$$P[X = k] = p(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Expectation:
$$E[X] = np$$

Variance:
$$V[X] = np(1 - p)$$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: $\{0, 1, \ldots, n\}$

Probability that $X$ takes on the value $k$:

$$P[X = k] = p(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Expectation:

$$E[X] = np$$

Variance:

$$V[X] = np(1-p)$$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 

Intro to Probability
A Binomial RV $X$ represents the number of successes in $n$ successive independent trials of a Bernoulli experiment. $X \sim Bin(n, p)$ is a Binomial RV, where $p$ is the probability of success in a given trial:

$$X \sim Bin(n, p)$$

Range: \( \{0, 1, \ldots, n\} \)

PMF: \( k \in \{0, 1, \ldots, n\} \)

\[
P[X = k] = p(k) = \binom{n}{k} p^k (1-p)^{n-k}
\]

Expectation: $E[X] = np$

Variance: $V[X] = np(1-p)$

Examples: # heads in $n$ coin tosses, # of 1’s in randomly generated length $n$ bit string

Note: by Binomial theorem (revision), we can prove $\sum_{k=0}^{n} P[X = k] = 1$. 
Example

Let $X$ be the number of heads after a coin is tossed three times: $X \sim \text{Bin}(3, 0.5)$. What is the probability of each of the different values of $X$?

Answer

\begin{align*}
P[X = 0] &= \binom{3}{0} (0.5)^0 (1-0.5)^3 = 1/8 \\
P[X = 1] &= \binom{3}{1} (0.5)^1 (1-0.5)^2 = 3/8 \\
P[X = 2] &= \binom{3}{2} (0.5)^2 (1-0.5)^1 = 3/8 \\
P[X = 3] &= \binom{3}{3} (0.5)^3 (1-0.5)^0 = 1/8
\end{align*}
Binomial RV is sum of Bernoulli RVs

Let $X$ be a Bernoulli RV: $X \sim Ber(p)$. Let $Y$ be a Binomial RV: $Y \sim Bin(n, p)$. Binomial RV = sum of $n$ independent Bernoulli RVs:

$$Y = \sum_{i=1}^{n} X_i, \quad X_i \sim Ber(p)$$

$$E[Y] = E\left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i] = np$$

Note: $Ber(p) = Bin(1, p)$
An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

Answer
Another example

An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

Answer

- $X$: # of bad bottles in a case (20 bottles)
Another example

Example

An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

Answer

- $X$: # of bad bottles in a case (20 bottles)
- $P\left[\text{have to give money back}\right] = P\left[ X \geq 2 \right] = 1 - P\left[ X = 0 \right] - P\left[ X = 1 \right]$
Another example

An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

**Answer**

- $X$: # of bad bottles in a case (20 bottles)
- $P[\text{have to give money back}] = P[X \geq 2] = 1 - P[X = 0] - P[X = 1]$
- $X$ is a binomial RV with parameters $X \sim Bin(n = 20, p = 0.05)$. 

Recall, when $X \sim Bin(n, p)$ then $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$ thus 

$P[X \geq 2] = 1 - P[X = 0] - P[X = 1] = 1 - \binom{20}{0} 0.05^0 0.95^{20} - \binom{20}{1} 0.05^1 0.95^{19} = 0.26$. 

Example

Intro to Probability Binomial discrete random variable 23
An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

**Answer**

- $X$: # of bad bottles in a case (20 bottles)
- $\mathbb{P}[\text{have to give money back}] = \mathbb{P}[X \geq 2] = 1 - \mathbb{P}[X = 0] - \mathbb{P}[X = 1]
- $X$ is a binomial RV with parameters $X \sim Bin(n = 20, p = 0.05)$.
- Bernoulli trial: check if a bottle is bad
Another example

An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

\[ X: \# \text{ of bad bottles in a case (20 bottles)} \]
\[ \mathbb{P} [ \text{have to give money back} ] = \mathbb{P} [ X \geq 2 ] = 1 - \mathbb{P} [ X = 0 ] - \mathbb{P} [ X = 1 ] \]
\[ X \text{ is a binomial RV with parameters } X \sim Bin(n = 20, p = 0.05). \]
\[ \text{Bernoulli trial: check if a bottle is bad} \]
\[ \mathbb{P} [ \text{success} ] = \mathbb{P} [ \text{bottle is bad} ] = 0.05 \]
\[ \mathbb{P} [ \text{failure} ] = \mathbb{P} [ \text{bottle is good} ] = 0.95 \]
Another example

An off-licence sells cases of wine, each containing 20 bottles. The probability that a bottle is bad is 0.05. The off-licence gives a money-back guarantee that the case will contain no more than one bad bottle. What is the probability that the off-licence will have to give money back?

\[ X: \text{# of bad bottles in a case (20 bottles)} \]
\[ \mathbf{P}[\text{have to give money back}] = \mathbf{P}[X \geq 2] = 1 - \mathbf{P}[X = 0] - \mathbf{P}[X = 1] \]
\[ X \text{ is a binomial RV with parameters } X \sim \text{Bin}(n = 20, p = 0.05). \]
\[ \text{Bernoulli trial: check if a bottle is bad} \]
\[ \mathbf{P}[\text{success}] = \mathbf{P}[\text{bottle is bad}] = 0.05 \]
\[ \mathbf{P}[\text{failure}] = \mathbf{P}[\text{bottle is good}] = 0.95 \]
\[ \text{Recall, when } X \sim \text{Bin}(n, p) \text{ then } \mathbf{P}[X = k] = \binom{n}{k}p^k(1 - p)^{n-k} \text{ thus} \]
\[ \mathbf{P}[X \geq 2] = 1 - \mathbf{P}[X = 0] - \mathbf{P}[X = 1] \]
Visualising Binomial PMFs

\[ X \sim \text{Bin}(40, 0.3); \quad X \sim \text{Bin}(40, 0.5); \quad X \sim \text{Bin}(40, 0.7) \]