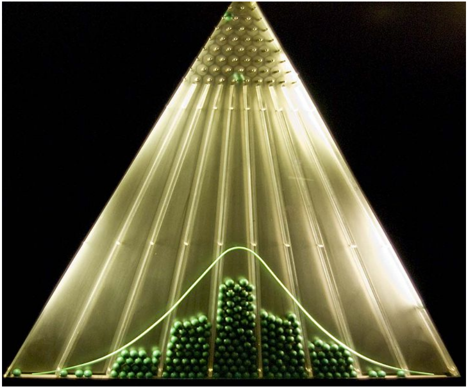
Ch. 3 – Discrete Random Variables



Random Variables

Given a set of outcomes Ω, a **random variable** is a number that depends on the outcome. A random variable is **discrete** if its possible values can be listed in a sequence $x_1, x_2, \ldots$

Random Variables

Given a set of outcomes Ω, a **random variable** is a number that depends on the outcome. A random variable is **discrete** if its possible values can be listed in a sequence $x_1, x_2, \ldots$

Example: Suppose we toss a fair coin 3 times. The set of outcomes is

$$\Omega = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$$

Let X be the number of heads. Then X is a random variable:

$$\begin{aligned} TTT &: X = 0 & HTT &: X = 1 \\ TTH &: X = 1 & HTH &: X = 2 \\ THT &: X = 1 & HHT &: X = 2 \\ THH &: X = 2 & HHH &: X = 3 \end{aligned}$$

The possible values of X are 0, 1, 2, and 3.

Probability Mass Function

In the previous example, we can calculate the probability of the number of heads X being each of the values 0, 1, 2, and 3:

$$P(X = 0) = P(\{TTT\}) = 1/8$$
$$P(X = 1) = P(\{HTT, THT, TTH\}) = 3/8$$
$$P(X = 2) = P(\{HHT, HTH, THH\}) = 3/8$$
$$P(X = 3) = P(\{HHH\}) = 1/8$$

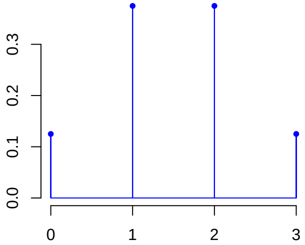
Probability Mass Function

In the previous example, we can calculate the probability of the number of heads X being each of the values 0, 1, 2, and 3:

$$
\begin{aligned}
P(X = 0) &= P(\{TTT\}) = 1/8 \\
P(X = 1) &= P(\{HTT, THT, TTH\}) = 3/8 \\
P(X = 2) &= P(\{HHT, HTH, THH\}) = 3/8 \\
P(X = 3) &= P(\{HHH\}) = 1/8
\end{aligned}
$$

The **probability mass function** (pmf), $f(x) = P(X = x)$, describes the probability of each possible value:

x	0	1	2	3
f(x)	1/8	3/8	3/8	1/8



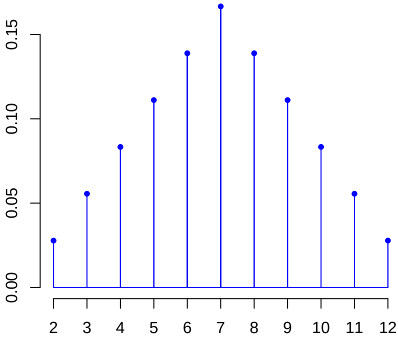
Sum of Two Random Variables

Suppose we roll two six-sided dice, and let X and Y be the results. Their sum is a random variable $X + Y$ with values 2, 3, . . . , 12. What is the probability mass function of $X + Y$?

Sum of Two Random Variables

Suppose we roll two six-sided dice, and let X and Y be the results. Their sum is a random variable $X + Y$ with values 2, 3, . . . , 12. What is the probability mass function of $X + Y$?

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12



x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Expected Value

Definition

Given a discrete random variable X with probability mass function $f(x)$, the **expected value** of X (or **mean** of X) is

$$E(X) = \sum x \cdot f(x)$$

where in the sum x ranges over all possible values of X. For $E(X)$ we will sometimes write μ_X or just μ.

Expected Value

Definition

Given a discrete random variable X with probability mass function $f(x)$, the **expected value** of X (or **mean** of X) is

$$E(X) = \sum x \cdot f(x)$$

where in the sum x ranges over all possible values of X. For $E(X)$ we will sometimes write μ_X or just μ.

Example: Recall that if we toss a fair coin three times, the number of heads X has probability mass function

x	0	1	2	3
f(x)	1/8	3/8	3/8	1/8

The expected value of the number of heads is then

$$E(X) = \sum x \cdot f(x) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1.5$$

Expected Value

If X is a discrete random variable X and $h(x)$ is a function, then $h(X)$ is also a discrete random variable, and

$$E[h(X)] = \sum h(x) \cdot f(x)$$

where $f(x)$ is the probability mass function of X.

Expected Value

If X is a discrete random variable X and $h(x)$ is a function, then $h(X)$ is also a discrete random variable, and

$$E[h(X)] = \sum h(x) \cdot f(x)$$

where $f(x)$ is the probability mass function of X.

Example: Again let X be the number of heads when tossing a fair coin 3 times. What is the expected value of X^2?

Expected Value

If X is a discrete random variable X and $h(x)$ is a function, then $h(X)$ is also a discrete random variable, and

$$E[h(X)] = \sum h(x) \cdot f(x)$$

where $f(x)$ is the probability mass function of X.

Example: Again let X be the number of heads when tossing a fair coin 3 times. What is the expected value of X^2?

Recall the pmf of X:

x	0	1	2	3
f(x)	1/8	3/8	3/8	1/8

Therefore, applying the formula above with $h(x) = x^2$,

$$\begin{aligned} E(X^2) &= \sum x^2 \cdot f(x) \\ &= 0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = \frac{24}{8} = 3 \end{aligned}$$

Properties of Expected Value

Let X and Y be random variables, and let c be a constant. Then

1. $E(c) = c$
2. $E(cX) = cE(X)$
3. $E(X + Y) = E(X) + E(Y)$

Example

Someone offers to let you play a game where you pay him \$10, toss a fair coin 3 times, and then he pays you back $3X + 5$ dollars, where X is the number of times the coin comes up heads.

If you play, what is the expected amount of money that you will you be paid? On average, would this game work in your favor?

Example

Someone offers to let you play a game where you pay him \$10, toss a fair coin 3 times, and then he pays you back $3X + 5$ dollars, where X is the number of times the coin comes up heads.

If you play, what is the expected amount of money that you will you be paid? On average, would this game work in your favor?

Recall that if you toss a coin 3 times, the expected number of times the coin comes up heads is $E(X) = 1.5$.

Example

Someone offers to let you play a game where you pay him \$10, toss a fair coin 3 times, and then he pays you back $3X + 5$ dollars, where X is the number of times the coin comes up heads.

If you play, what is the expected amount of money that you will you be paid? On average, would this game work in your favor?

Recall that if you toss a coin 3 times, the expected number of times the coin comes up heads is $E(X) = 1.5$. Therefore, using properties of expected value,

$$
\begin{aligned}
E(3X+5) &= E(3X) + E(5) \\
&= 3E(X) + 5 \\
&= 3(1.5) + 5 = 9.5
\end{aligned}
$$

Example

Someone offers to let you play a game where you pay him \$10, toss a fair coin 3 times, and then he pays you back $3X + 5$ dollars, where X is the number of times the coin comes up heads.

If you play, what is the expected amount of money that you will you be paid? On average, would this game work in your favor?

Recall that if you toss a coin 3 times, the expected number of times the coin comes up heads is $E(X) = 1.5$. Therefore, using properties of expected value,

$$\begin{aligned} E(3X+5) &= E(3X) + E(5) \\ &= 3E(X) + 5 \\ &= 3(1.5) + 5 = 9.5 \end{aligned}$$

So the expected payback is \$9.50. Since you have to pay \$10, this game would not work in your favor on average.

Variance and Standard Deviation

Definition

Given a discrete random variable X with probability mass function $f(x)$ and mean $\mu = E(X)$, the **variance** of X is

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

The variance of X is sometimes written as σ_X^2 or just σ^2. The **standard deviation** of X is the square root of the variance:

$$\sigma = \sqrt{V(X)}$$

Variance and Standard Deviation

Definition

Given a discrete random variable X with probability mass function $f(x)$ and mean $\mu = E(X)$, the **variance** of X is

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

The variance of X is sometimes written as σ_X^2 or just σ^2. The **standard deviation** of X is the square root of the variance:

$$\sigma = \sqrt{V(X)}$$

Example: Again let X be the number of heads when tossing a fair coin 3 times. We know that $\mu = E(X) = 1.5$. What is $V(X)$?

Variance and Standard Deviation

Definition

Given a discrete random variable X with probability mass function $f(x)$ and mean $\mu = E(X)$, the **variance** of X is

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

The variance of X is sometimes written as σ_X^2 or just σ^2. The **standard deviation** of X is the square root of the variance:

$$\sigma = \sqrt{V(X)}$$

Example: Again let X be the number of heads when tossing a fair coin 3 times. We know that $\mu = E(X) = 1.5$. What is $V(X)$?

$$\begin{aligned} V(X) = E[(X - \mu)^2] &= \sum (x - \mu)^2 \cdot f(x) \\ &= (0 - 1.5)^2 \cdot \frac{1}{8} + (1 - 1.5)^2 \cdot \frac{3}{8} + (2 - 1.5)^2 \cdot \frac{3}{8} + (3 - 1.5)^2 \cdot \frac{1}{8} \\ &= 0.75 \end{aligned}$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof: $\quad V(X) = E[(X - \mu)^2]$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:
$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \end{aligned}$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:

$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \end{aligned}$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:

$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:

$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Example: Again let X be the number of heads when tossing a fair coin 3 times. We know that $\mu = E(X) = 1.5$.

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:
$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Example: Again let X be the number of heads when tossing a fair coin 3 times. We know that $\mu = E(X) = 1.5$. Then

$$V(X) = E(X^2) - \mu^2$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:
$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Example: Again let X be the number of heads when tossing a fair coin 3 times. We know that $\mu = E(X) = 1.5$. Then

$$\begin{aligned} V(X) &= E(X^2) - \mu^2 \\ &= \left(0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8}\right) - 1.5^2 \end{aligned}$$

Shortcut Formula for Variance

Let X be a discrete random variable with mean $E(X) = \mu$.

$$V(X) = E(X^2) - \mu^2$$

Proof:
$$\begin{aligned} V(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Example: Again let X be the number of heads when tossing a fair coin 3 times. We know that $\mu = E(X) = 1.5$. Then

$$\begin{aligned} V(X) &= E(X^2) - \mu^2 \\ &= \left(0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8}\right) - 1.5^2 \\ &= 0.75 \end{aligned}$$

Bernoulli Random Variable

One of the simplest kinds of random variables is one which takes only two possible values: 0 and 1. We call X a **Bernoulli random variable** with parameter p if

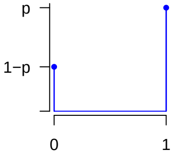
$$P(X = 1) = p$$
$$P(X = 0) = 1 - p$$

Bernoulli Random Variable

One of the simplest kinds of random variables is one which takes only two possible values: 0 and 1. We call X a **Bernoulli random variable** with parameter p if

$$P(X = 1) = p$$
$$P(X = 0) = 1 - p$$

x	0	1
$f(x)$	1-p	p

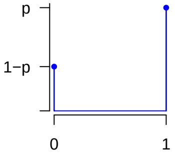


Bernoulli Random Variable

One of the simplest kinds of random variables is one which takes only two possible values: 0 and 1. We call X a **Bernoulli random variable** with parameter p if

$$P(X = 1) = p$$
$$P(X = 0) = 1 - p$$

x	0	1
$f(x)$	1-p	p



For example, if we toss a fair coin, then the outcome is a Bernoulli random variable with parameter $p = 1/2$, with 0 representing tails, and 1 representing heads.

Mean and Variance of Bernoulli Random Variable

Let X be a Bernoulli random variable with parameter p. This means that the probability mass function (pmf) of X is

x	0	1
$f(x)$	1-p	p

What is the expected value and variance of X?

Mean and Variance of Bernoulli Random Variable

Let X be a Bernoulli random variable with parameter p. This means that the probability mass function (pmf) of X is

x	0	1
$f(x)$	1-p	p

What is the expected value and variance of X?

$$\begin{aligned}\mu = E(X) &= \sum x \cdot f(x) \\ &= 0 \cdot (1-p) + 1 \cdot p = p\end{aligned}$$

Mean and Variance of Bernoulli Random Variable

Let X be a Bernoulli random variable with parameter p. This means that the probability mass function (pmf) of X is

x	0	1
$f(x)$	1-p	p

What is the expected value and variance of X?

$$\begin{aligned}\mu = E(X) &= \sum x \cdot f(x) \\ &= 0 \cdot (1-p) + 1 \cdot p = p\end{aligned}$$

$$\begin{aligned}V(X) &= E((X-\mu)^2) = E((X-p)^2) \\ &= \sum (x-p)^2 \cdot f(x) \\ &= (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p \\ &= p^2(1-p) + p(1-p)^2 \\ &= p(1-p)(p+(1-p)) = p(1-p)\end{aligned}$$

Independent Random Variables

Recall that two events A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

Independent Random Variables

Recall that two events A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

Definition

We say that discrete random variables X and Y are **independent** if for any possible values a and b of X and Y respectively,

$$P(X = a \cap Y = b) = P(X = a)P(Y = b)$$

Independent Random Variables

Recall that two events A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

Definition

We say that discrete random variables X and Y are **independent** if for any possible values a and b of X and Y respectively,

$$P(X = a \cap Y = b) = P(X = a)P(Y = b)$$

This definition generalizes to several random variables: We say that $X_1, X_2, X_3, \ldots$ are **independent** if for all k and all possible values $a_1, a_2, a_3, \ldots, a_k$ of $X_1, X_2, X_3, \ldots, X_k$ respectively,

$$P(X_1 = a_1 \cap \cdots \cap X_k = a_k) = P(X_1 = a_1) \cdots P(X_k = a_k)$$

Problem

Suppose that when we spin a coin, it only comes up heads with probability .4. If we spin the coin 3 times, find the probability mass function for the number of times we get heads.

Problem

Suppose that when we spin a coin, it only comes up heads with probability .4. If we spin the coin 3 times, find the probability mass function for the number of times we get heads.

We may represent the outcomes of the three spins as independent Bernoulli random variables Y_1, Y_2, Y_3 each with parameter .4, where $Y_i = 1$ if the ith spin is heads and $Y_i = 0$ if the ith spin is tails. Then the number of heads is $X = Y_1 + Y_2 + Y_3$,

Problem

Suppose that when we spin a coin, it only comes up heads with probability .4. If we spin the coin 3 times, find the probability mass function for the number of times we get heads.

We may represent the outcomes of the three spins as independent Bernoulli random variables Y_1, Y_2, Y_3 each with parameter .4, where $Y_i = 1$ if the ith spin is heads and $Y_i = 0$ if the ith spin is tails. Then the number of heads is $X = Y_1 + Y_2 + Y_3$, and

$$\begin{aligned}
P(X = 0) &= P(\{TTT\}) \\
&= P(Y_1 = 0 \cap Y_2 = 0 \cap Y_3 = 0) \\
&= P(Y_1 = 0)P(Y_2 = 0)P(Y_3 = 0) \\
&= (.6)(.6)(.6) = .216
\end{aligned}$$

There are three ways to get heads exactly once:

$$\begin{aligned} P(X=1) &= P(\{HTT, THT, TTH\}) \\ &= P(\{HTT\}) + P(\{THT\}) + P(\{TTH\}) \end{aligned}$$

There are three ways to get heads exactly once:

$$\begin{aligned} P(X=1) &= P(\{HTT, THT, TTH\}) \\ &= P(\{HTT\}) + P(\{THT\}) + P(\{TTH\}) \end{aligned}$$

We calculate,

$$\begin{aligned} P(\{HTT\}) &= P(Y_1 = 1 \cap Y_2 = 0 \cap Y_3 = 0) \\ &= P(Y_1 = 1)P(Y_2 = 0)P(Y_3 = 0) \\ &= (.4)(.6)(.6) = .144 \end{aligned}$$

There are three ways to get heads exactly once:

$$\begin{aligned} P(X=1) &= P(\{HTT, THT, TTH\}) \\ &= P(\{HTT\}) + P(\{THT\}) + P(\{TTH\}) \end{aligned}$$

We calculate,

$$\begin{aligned} P(\{HTT\}) &= P(Y_1 = 1 \cap Y_2 = 0 \cap Y_3 = 0) \\ &= P(Y_1 = 1)P(Y_2 = 0)P(Y_3 = 0) \\ &= (.4)(.6)(.6) = .144 \end{aligned}$$

Similarly $P(\{THT\}) = P(\{TTH\}) = .144$. Therefore,

$$\begin{aligned} P(X=1) &= P(\{HTT\}) + P(\{THT\}) + P(\{TTH\}) \\ &= 3P(\{HTT\}) = 3(.144) = .432 \end{aligned}$$

By the same kind of reasoning,

$$\begin{aligned} P(X=2) &= P(\{HHT, HTH, THH\}) \\ &= 3P(\{HHT\}) \\ &= 3P(Y_1=1)P(Y_2=1)P(Y_3=0) \\ &= 3(.4)(.4)(.6) = .288 \end{aligned}$$

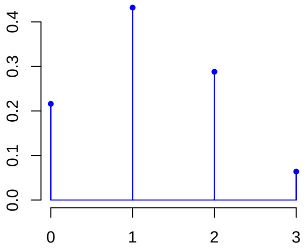
$$\begin{aligned} P(X=3) &= P(\{HHH\}) \\ &= P(Y_1=1)P(Y_2=1)P(Y_3=1) \\ &= (.4)(.4)(.4) = .064 \end{aligned}$$

Putting this together, we found

$$P(X = 0) = .216$$
$$P(X = 1) = .432$$
$$P(X = 2) = .288$$
$$P(X = 3) = .064$$

Therefore, the probability mass function for X is

x	0	1	2	3
f(x)	.216	.432	.288	.064



Binomial Distributions

Let $Y_1, Y_2, \ldots, Y_n$ be independent Bernoulli random variables with parameter p. Then their sum

$$X = Y_1 + Y_2 + \cdots + Y_n$$

is a **binomial random variable** with parameters n and p. We write $X \sim \text{Bin}(n, p)$

Binomial Distributions

Let $Y_1, Y_2, \ldots, Y_n$ be independent Bernoulli random variables with parameter p. Then their sum

$$X = Y_1 + Y_2 + \cdots + Y_n$$

is a **binomial random variable** with parameters n and p. We write $X \sim \text{Bin}(n, p)$

In other words, given a sequence of n independent trials, each with probability p of success, the binomial random variable X counts the number of successes. The possible values of X are 0, 1, ..., n.

Binomial Distributions

Let $Y_1, Y_2, \ldots, Y_n$ be independent Bernoulli random variables with parameter p. Then their sum

$$X = Y_1 + Y_2 + \cdots + Y_n$$

is a **binomial random variable** with parameters n and p. We write $X \sim \text{Bin}(n, p)$

In other words, given a sequence of n independent trials, each with probability p of success, the binomial random variable X counts the number of successes. The possible values of X are 0, 1, ..., n.

For example, if we toss a coin 3 times, then the number of heads X is a binomial random variable with parameters $n = 3$ and $p = .5$.

If we toss a fair coin 4 times, what is the probability that we get exactly 2 heads?

Binomial Random Variables

If we toss a fair coin 4 times, what is the probability that we get exactly 2 heads?

We may list the 16 equally likely outcomes:

TTTT	TTTH	TTHT	TTHH
THTT	THTH	THHT	THHH
HTTT	HTTH	HTHT	HTHH
HHTT	HHTH	HHHT	HHHH

If we toss a fair coin 4 times, what is the probability that we get exactly 2 heads?

We may list the 16 equally likely outcomes:

TTTT	TTTH	TTHT	**TTHH**
THTT	**THTH**	**THHT**	THHH
HTTT	**HTTH**	**HTHT**	HTHH
HHTT	HHTH	HHHT	HHHH

Out of the 16 outcomes, 6 involve getting exactly 2 heads. Therefore,

$$P(X = 2) = 6/16 = 3/8 = .375$$

Binomial Random Variables

If we toss a fair coin 4 times, what is the probability that we get exactly 2 heads?

We may list the 16 equally likely outcomes:

TTTT	TTTH	TTHT	TTHH
THTT	THTH	THHT	THHH
HTTT	HTTH	HTHT	HTHH
HHTT	HHTH	HHHT	HHHH

Out of the 16 outcomes, 6 involve getting exactly 2 heads. Therefore,

$$P(X = 2) = 6/16 = 3/8 = .375$$

However, it would be nice to be able to solve this without having to list all the outcomes.

Counting

How many ways are there to rearrange the letters $ABCD$?

Counting

How many ways are there to rearrange the letters *ABCD*?

ABCD	ABDC	ACBD	ACDB	ADBC	ADCB
BACD	BADC	BCAD	BCDA	BDAC	BDCA
CABD	CADB	CBAD	CBDA	CDAB	CDBA
DABC	DACB	DBAC	DBCA	DCAB	DCBA

Counting

How many ways are there to rearrange the letters *ABCD*?

ABCD	ABDC	ACBD	ACDB	ADBC	ADCB
BACD	BADC	BCAD	BCDA	BDAC	BDCA
CABD	CADB	CBAD	CBDA	CDAB	CDBA
DABC	DACB	DBAC	DBCA	DCAB	DCBA

There are

$$
\begin{aligned}
& 4 \text{ choices for which letter to put in the first position} \\
\times\ & 3 \text{ choices for which letter to put in the second position} \\
\times\ & 2 \text{ choices for which letter to put in the third position} \\
\times\ & 1 \text{ choice for which letter to put in the fourth position} \\
=\ & 4 \cdot 3 \cdot 2 \cdot 1 = 24
\end{aligned}
$$

Counting

How many ways are there to rearrange the letters *ABCD*?

ABCD	ABDC	ACBD	ACDB	ADBC	ADCB
BACD	BADC	BCAD	BCDA	BDAC	BDCA
CABD	CADB	CBAD	CBDA	CDAB	CDBA
DABC	DACB	DBAC	DBCA	DCAB	DCBA

There are

$$\begin{aligned} & 4 \text{ choices for which letter to put in the first position} \\ \times\ & 3 \text{ choices for which letter to put in the second position} \\ \times\ & 2 \text{ choices for which letter to put in the third position} \\ \times\ & 1 \text{ choice for which letter to put in the fourth position} \\ =\ & 4 \cdot 3 \cdot 2 \cdot 1 = 24 \end{aligned}$$

In general, the number of ways to rearrange n distinct symbols is

$$n! = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1$$

Counting

How many ways are there to rearrange the letters *BANANA*?

Counting

How many ways are there to rearrange the letters *BANANA*?

Simply counting them all, we find there are 60:

BANANA	BANAAN	BAANAN	ABANAN	ABANNA	BAANNA
BANNAA	ABNANA	BNAANA	BNANAA	NBANAA	NBAANA
NABANA	ANBANA	ANBAAN	NABAAN	NBAAAN	BNAAAN
ABNAAN	BAAANN	ABAANN	ANABAN	NAABAN	AANBAN
AABNAN	AABANN	AABNNA	ABNNAA	AANBNA	ANABNA
ANBNAA	NABNAA	NAABNA	BNNAAA	NBNAAA	ANNBAA
NANBAA	NNABAA	NNBAAA	NANABA	NAANBA	ANANBA
ANNABA	NNAABA	AANNBA	ANAABN	AANABN	AANANB
ANAANB	ANANAB	ANNAAB	AANNAB	NANAAB	NAANAB
NNAAAB	NAAANB	NAAABN	AAANBN	AAANNB	AAABNN

Counting

How many ways are there to rearrange the letters *BANANA*?

How many ways are there to rearrange the letters *BANANA*?

If the 6 letters were all distinct, there would be $6! = 720$ ways of rearranging them.

How many ways are there to rearrange the letters *BANANA*?

If the 6 letters were all distinct, there would be $6! = 720$ ways of rearranging them.

However, given any particular rearrangement, there are $3! = 6$ ways of rearranging the A's among themselves and $2! = 2$ ways of rearranging the N's among themselves, with no effect.

How many ways are there to rearrange the letters *BANANA*?

If the 6 letters were all distinct, there would be $6! = 720$ ways of rearranging them.

However, given any particular rearrangement, there are $3! = 6$ ways of rearranging the A's among themselves and $2! = 2$ ways of rearranging the N's among themselves, with no effect.

Therefore we are overcounting by a factor of $3! \cdot 2!$, and the correct number of rearrangements is

$$\frac{6!}{3! \cdot 2!} = \frac{720}{12} = 60$$

How many ways are there to rearrange the letters $HHTT$?

Counting

How many ways are there to rearrange the letters $HHTT$?

Reasoning as before, the are $\frac{4!}{2! \cdot 2!} = 6$ ways to rearrange them.

Counting

How many ways are there to rearrange the letters $HHTT$?

Reasoning as before, the are $\frac{4!}{2! \cdot 2!} = 6$ ways to rearrange them.

If we toss a coin 6 times, what is the probability that it comes up heads exactly 3 times?

Counting

How many ways are there to rearrange the letters $HHTT$?

Reasoning as before, the are $\frac{4!}{2! \cdot 2!} = 6$ ways to rearrange them.

If we toss a coin 6 times, what is the probability that it comes up heads exactly 3 times?

Out of the $2^6 = 64$ equally likely outcomes, the number of ways of getting exactly 3 heads is the number of rearrangements of $HHHTTT$, which is $\frac{6!}{3! \cdot 3!} = 20$. So the probability of getting exactly 3 heads is

$$P(X = 3) = \frac{20}{64} = \frac{5}{16} = .3125$$

Binomial Coefficients

By the same reasoning, the number of ways to rearrange the word $HH\cdots HTT\cdots T$, where there are n symbols, k of which are H's and $n-k$ of which are $T's$, is $\frac{n!}{k!(n-k)!}$.

Binomial Coefficients

By the same reasoning, the number of ways to rearrange the word $HH \cdots HTT \cdots T$, where there are n symbols, k of which are H's and $n - k$ of which are $T's$, is $\frac{n!}{k!(n-k)!}$.

Choosing a rearrangement of such a word is the same as choosing k out of the n positions to contain an H.

Binomial Coefficients

By the same reasoning, the number of ways to rearrange the word $HH \cdots HTT \cdots T$, where there are n symbols, k of which are H's and $n-k$ of which are $T's$, is $\frac{n!}{k!(n-k)!}$.

Choosing a rearrangement of such a word is the same as choosing k out of the n positions to contain an H. Therefore,

The number of ways to choose k objects out of n objects is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

These numbers are called **binomial coefficients**.

Binomial Coefficients

By the same reasoning, the number of ways to rearrange the word $HH\cdots HTT\cdots T$, where there are n symbols, k of which are H's and $n-k$ of which are $T's$, is $\frac{n!}{k!(n-k)!}$.

Choosing a rearrangement of such a word is the same as choosing k out of the n positions to contain an H. Therefore,

The number of ways to choose k objects out of n objects is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

These numbers are called **binomial coefficients**.

For example, if a batch contains 30 widgets, the number of ways of choosing 2 for inspection is

$$\binom{30}{2} = \frac{30!}{2!28!} = \frac{30 \cdot 29}{2} = 435$$

Binomial Random Variables

A binomial random variable $X \sim \text{Bin}(n, p)$ has pmf

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Proof: Write X as the sum of n independent Bernoulli random variables $Y_1, \ldots, Y_n$ each with parameter p:

$$X = Y_1 + Y_2 + \cdots + Y_n$$

Reasoning as in the previous slides, we then calculate

$$\begin{aligned} P(X = x) &= \binom{n}{x} P(Y_1 = 1 \cap \cdots \cap Y_x = 1 \cap Y_{x+1} = 0 \cap \cdots \cap Y_n = 0) \\ &= \binom{n}{x} P(Y_1 = 1) \cdots P(Y_x = 1) P(Y_{x+1} = 0) \cdots P(Y_n = 0) \\ &= \binom{n}{x} p^x (1-p)^{n-x} \end{aligned}$$

Example

Suppose your friend claims he can make a basketball free throw shot 60% of the time. You ask him to demonstrate, and he only makes 2 out of 7 shots. If your friend's claim is correct, what is the probability that he would make 2 or fewer out of 7 shots?

Example

Suppose your friend claims he can make a basketball free throw shot 60% of the time. You ask him to demonstrate, and he only makes 2 out of 7 shots. If your friend's claim is correct, what is the probability that he would make 2 or fewer out of 7 shots?

If the friend's claim is correct, then the number of shots X is binomial, $X \sim \mathrm{Bin}(7, .6)$.

Example

Suppose your friend claims he can make a basketball free throw shot 60% of the time. You ask him to demonstrate, and he only makes 2 out of 7 shots. If your friend's claim is correct, what is the probability that he would make 2 or fewer out of 7 shots?

If the friend's claim is correct, then the number of shots X is binomial, $X \sim \text{Bin}(7, .6)$. Then

$$\begin{aligned}
P(X \le 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\
&= \binom{7}{0}(.6)^0(.4)^7 + \binom{7}{1}(.6)^1(.4)^6 + \binom{7}{2}(.6)^2(.4)^5 \\
&= 1(.6)^0(.4)^7 + 7(.6)^1(.4)^6 + 21(.6)^2(.4)^5 \\
&= .0016384 + .0172032 + .0774144 \\
&= .096256
\end{aligned}$$

Example

Suppose your friend claims he can make a basketball free throw shot 60% of the time. You ask him to demonstrate, and he only makes 2 out of 7 shots. If your friend's claim is correct, what is the probability that he would make 2 or fewer out of 7 shots?

If the friend's claim is correct, then the number of shots X is binomial, $X \sim \text{Bin}(7, .6)$. Then

$$\begin{aligned}
P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\
&= \binom{7}{0}(.6)^0(.4)^7 + \binom{7}{1}(.6)^1(.4)^6 + \binom{7}{2}(.6)^2(.4)^5 \\
&= 1(.6)^0(.4)^7 + 7(.6)^1(.4)^6 + 21(.6)^2(.4)^5 \\
&= .0016384 + .0172032 + .0774144 \\
&= .096256
\end{aligned}$$

So assuming your friend is as good as he claims, there is about a 10% chance that he would do this poorly.

Cumulative Distribution Function (CDF)

Given a discrete random variable X, recall that the probability mass function $f(x)$ is defined by

$$f(x) = P(X = x)$$

Cumulative Distribution Function (CDF)

Given a discrete random variable X, recall that the probability mass function $f(x)$ is defined by

$$f(x) = P(X = x)$$

The **cumulative distribution function** (cdf), written $F(x)$, is defined by

$$F(x) = P(X \leq x)$$

Cumulative Distribution Function (CDF)

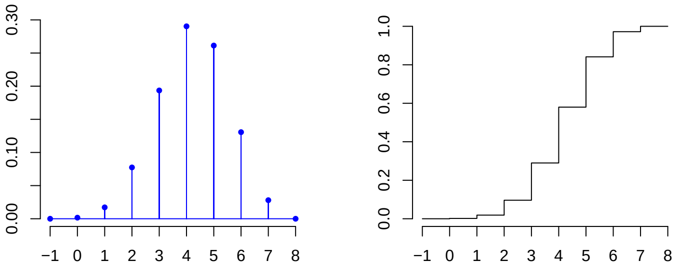
Given a discrete random variable X, recall that the probability mass function $f(x)$ is defined by

$$f(x) = P(X = x)$$

The **cumulative distribution function** (cdf), written $F(x)$, is defined by

$$F(x) = P(X \leq x)$$

For example, given $X \sim \text{Bin}(7, .6)$, $f(x)$ and $F(x)$ are as follows:



Properties of Variance

Let X and Y be random variables, and let c be a constant. Then

1. $V(c) = 0$
2. $V(cX) = c^2 V(X)$
3. $V(X + c) = V(X)$
4. If X and Y are independent, $V(X + Y) = V(X) + V(Y)$.

Properties of Variance

Let X and Y be random variables, and let c be a constant. Then

1. $V(c) = 0$
2. $V(cX) = c^2 V(X)$
3. $V(X + c) = V(X)$
4. If X and Y are independent, $V(X + Y) = V(X) + V(Y)$.

Example: Let X be the number of heads when tossing a fair coin three times, so $X = Y_1 + Y_2 + Y_3$ where Y_1, Y_2, and Y_3 are independent Bernoulli random variables with parameter $p = .5$. By our formula for the variance of a Bernoulli,

$$V(Y_i) = p(1 - p) = .5(1 - .5) = .25$$

Properties of Variance

Let X and Y be random variables, and let c be a constant. Then

1. $V(c) = 0$
2. $V(cX) = c^2 V(X)$
3. $V(X + c) = V(X)$
4. If X and Y are independent, $V(X + Y) = V(X) + V(Y)$.

Example: Let X be the number of heads when tossing a fair coin three times, so $X = Y_1 + Y_2 + Y_3$ where Y_1, Y_2, and Y_3 are independent Bernoulli random variables with parameter $p = .5$. By our formula for the variance of a Bernoulli,

$$V(Y_i) = p(1 - p) = .5(1 - .5) = .25$$

Therefore,

$$\begin{aligned} V(X) &= V(Y_1 + Y_2 + Y_3) \\ &= V(Y_1) + V(Y_2) + V(Y_3) \\ &= .25 + .25 + .25 = .75 \end{aligned}$$

Mean and Variance of Binomial Random Variables

A binomial random variable $X \sim \text{Bin}(n, p)$ has mean and variance $E(X) = np$ and $V(X) = np(1-p)$.

Mean and Variance of Binomial Random Variables

A binomial random variable $X \sim \text{Bin}(n, p)$ has mean and variance $E(X) = np$ and $V(X) = np(1-p)$.

Proof: Write X as $X = Y_1 + Y_2 + \cdots + Y_n$ where $Y_1, \ldots, Y_n$ are independent Bernoulli random variables with parameter p. Then

$$E(Y_i) = p, \quad V(Y_i) = p(1-p)$$

Mean and Variance of Binomial Random Variables

A binomial random variable $X \sim \text{Bin}(n, p)$ has mean and variance $E(X) = np$ and $V(X) = np(1-p)$.

Proof: Write X as $X = Y_1 + Y_2 + \cdots + Y_n$ where $Y_1, \ldots, Y_n$ are independent Bernoulli random variables with parameter p. Then

$$E(Y_i) = p, \quad V(Y_i) = p(1-p)$$

Therefore,

$$\begin{aligned} E(X) &= E(Y_1 + \cdots + Y_n) \\ &= E(Y_1) + \cdots + E(Y_n) \\ &= p + \cdots + p = np \end{aligned}$$

$$\begin{aligned} V(X) &= V(Y_1 + \cdots + Y_n) \\ &= V(Y_1) + \cdots + V(Y_n) \\ &= p(1-p) + \cdots + p(1-p) = np(1-p) \end{aligned}$$

Problem

A manufacturer produces widgets which work with probability .4. Suppose we test widgets one at a time until we find one that works. Let X be the number of bad widgets we try before we find one that works. What is the probability that $X = 2$?

Problem

A manufacturer produces widgets which work with probability .4. Suppose we test widgets one at a time until we find one that works. Let X be the number of bad widgets we try before we find one that works. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying $Y_1 = 0, Y_2 = 0, Y_3 = 1$. Therefore,

$$\begin{aligned} P(X = 2) &= P(Y_1 = 0, Y_2 = 0, Y_3 = 1) \\ &= P(Y_1 = 0)P(Y_2 = 0)P(Y_3 = 1) \\ &= (.6)(.6)(.4) = .144 \end{aligned}$$

Geometric Random Variable

Suppose we have a sequence of independent trials each with probability p of success. Let X be the number of failures before the first success. The possible values of X are 0, 1, 2, ... We say that X is a **geometric random variable** with parameter p.

The pmf of a geometric random variable with parameter p is $f(x) = p(1-p)^x$.

Geometric Random Variable

Suppose we have a sequence of independent trials each with probability p of success. Let X be the number of failures before the first success. The possible values of X are 0, 1, 2, . . . We say that X is a **geometric random variable** with parameter p.

The pmf of a geometric random variable with parameter p is $f(x) = p(1-p)^x$.

Proof: We may model the trials as a sequence of independent Bernoulli random variables $Y_1, Y_2, \ldots$ each with parameter p. Then we can calculate the pmf of X:

$$
\begin{aligned}
P(X = x) &= P(Y_1 = 0, Y_2 = 0, \ldots, Y_x = 0, Y_{x+1} = 1) \\
&= P(Y_1 = 0)P(Y_2 = 0) \cdots P(Y_x = 0)P(Y_{x+1} = 1) \\
&= (1-p)^x p
\end{aligned}
$$

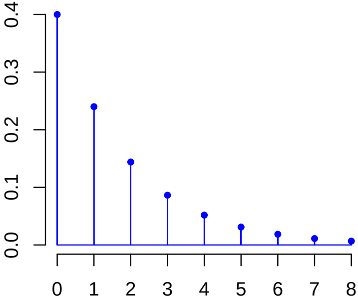
Example

A manufacturer produces widgets which work with probability .4. Suppose we test widgets one at a time until we find one that works. Let X be the number of bad widgets we try before we find one that works. What is the pmf of X?

Example

A manufacturer produces widgets which work with probability .4. Suppose we test widgets one at a time until we find one that works. Let X be the number of bad widgets we try before we find one that works. What is the pmf of X?

The pmf of a geometric random variable is $f(x) = p(1-p)^x$. In this case, $p = .4$, so this becomes $f(x) = .4(.6)^x$.



Geometric Sums

Let $r \neq 1$ be a real number. Then

$$\sum_{j=0}^{n} r^j = 1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

Geometric Sums

Let $r \neq 1$ be a real number. Then

$$\sum_{j=0}^{n} r^j = 1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

Example: $$1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^5 = \frac{1 - (1/2)^6}{1 - 1/2}$$

$$= \frac{\frac{63}{64}}{1/2} = \frac{63}{32}$$

Geometric Sums

Let $r \neq 1$ be a real number. Then

$$\sum_{j=0}^{n} r^j = 1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

Example: $$1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^5 = \frac{1 - (1/2)^6}{1 - 1/2} = \frac{\frac{63}{64}}{1/2} = \frac{63}{32}$$

Proof:

$$\begin{aligned}(1 - r)(1 + r + r^2 + \cdots + r^n) &= 1 + r + r^2 + \cdots + r^n \\ &\quad - (r + r^2 + \cdots + r^n + r^{n+1}) \\ &= 1 - r^{n+1}\end{aligned}$$

Now divide by $1 - r$.

Infinite Geometric Sums

Let r be a real number with $|r| < 1$. Then

$$\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$$

Infinite Geometric Sums

Let r be a real number with $|r| < 1$. Then

$$\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$$

Proof:

$$\sum_{j=0}^{\infty} r^j = 1 + r + r^2 + \cdots = \lim_{n\to\infty} \sum_{j=0}^{n} r^j$$

$$= \lim_{n\to\infty} \frac{1-r^{n+1}}{1-r} = \frac{1}{1-r}$$

Infinite Geometric Sums

Let r be a real number with $|r| < 1$. Then

$$\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$$

Proof:

$$\sum_{j=0}^{\infty} r^j = 1 + r + r^2 + \cdots = \lim_{n\to\infty} \sum_{j=0}^{n} r^j$$

$$= \lim_{n\to\infty} \frac{1-r^{n+1}}{1-r} = \frac{1}{1-r}$$

Example:

$$\sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$$

$$= \frac{1}{1-\frac{1}{2}} = 2$$

Sum of PMF of Geometric Random Variable

Sum of PMF of Geometric Random Variable

If $f(x)$ is the pmf of a geometric random variable, we can confirm that the values of $f(x)$ add up to 1 (as they must for any pmf), using the formula $\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$ from the previous slide:

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} p(1-p)^x$$

Sum of PMF of Geometric Random Variable

If $f(x)$ is the pmf of a geometric random variable, we can confirm that the values of $f(x)$ add up to 1 (as they must for any pmf), using the formula $\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$ from the previous slide:

$$\begin{aligned}\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} p(1-p)^x \\ &= p\sum_{x=0}^{\infty}(1-p)^x\end{aligned}$$

Sum of PMF of Geometric Random Variable

If $f(x)$ is the pmf of a geometric random variable, we can confirm that the values of $f(x)$ add up to 1 (as they must for any pmf), using the formula $\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$ from the previous slide:

$$\begin{aligned}\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} p(1-p)^x \\ &= p\sum_{x=0}^{\infty}(1-p)^x \\ &= p \cdot \frac{1}{1-(1-p)}\end{aligned}$$

Sum of PMF of Geometric Random Variable

If $f(x)$ is the pmf of a geometric random variable, we can confirm that the values of $f(x)$ add up to 1 (as they must for any pmf), using the formula $\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}$ from the previous slide:

$$\begin{aligned}\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} p(1-p)^x \\ &= p\sum_{x=0}^{\infty}(1-p)^x \\ &= p \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1\end{aligned}$$

CDF of Geometric Distribution

The cdf of a geometric random variable X with parameter p is $F(x) = 1 - (1-p)^{x+1}$.

CDF of Geometric Distribution

The cdf of a geometric random variable X with parameter p is $F(x) = 1 - (1-p)^{x+1}$.

Proof: For any integer $t \geq 0$, we calculate

$$F(t) = P(X \leq t) = \sum_{x=0}^{t} P(X = x)$$

CDF of Geometric Distribution

The cdf of a geometric random variable X with parameter p is $F(x) = 1 - (1-p)^{x+1}$.

Proof: For any integer $t \geq 0$, we calculate

$$\begin{aligned} F(t) = P(X \leq t) &= \sum_{x=0}^{t} P(X = x) \\ &= \sum_{x=0}^{t} p(1-p)^x \end{aligned}$$

CDF of Geometric Distribution

The cdf of a geometric random variable X with parameter p is $F(x) = 1 - (1-p)^{x+1}$.

Proof: For any integer $t \geq 0$, we calculate

$$\begin{aligned} F(t) = P(X \leq t) &= \sum_{x=0}^{t} P(X = x) \\ &= \sum_{x=0}^{t} p(1-p)^x \\ &= p \sum_{x=0}^{t} (1-p)^x \end{aligned}$$

CDF of Geometric Distribution

The cdf of a geometric random variable X with parameter p is $F(x) = 1 - (1-p)^{x+1}$.

Proof: For any integer $t \geq 0$, we calculate

$$\begin{aligned}
F(t) = P(X \leq t) &= \sum_{x=0}^{t} P(X = x) \\
&= \sum_{x=0}^{t} p(1-p)^x \\
&= p\sum_{x=0}^{t}(1-p)^x \\
&= p \cdot \frac{1-(1-p)^{t+1}}{1-(1-p)}
\end{aligned}$$

CDF of Geometric Distribution

The cdf of a geometric random variable X with parameter p is $F(x) = 1 - (1-p)^{x+1}$.

Proof: For any integer $t \geq 0$, we calculate

$$\begin{aligned} F(t) = P(X \leq t) &= \sum_{x=0}^{t} P(X = x) \\ &= \sum_{x=0}^{t} p(1-p)^x \\ &= p\sum_{x=0}^{t}(1-p)^x \\ &= p \cdot \frac{1-(1-p)^{t+1}}{1-(1-p)} \\ &= 1-(1-p)^{t+1} \end{aligned}$$

CDF of Geometric Distribution

Alternate proof: The event $X \leq x$ means that there are no more than x failures before the first success, which is the same as saying that a success occurs among the first $x+1$ trials:

CDF of Geometric Distribution

Alternate proof: The event $X \leq x$ means that there are no more than x failures before the first success, which is the same as saying that a success occurs among the first $x+1$ trials:

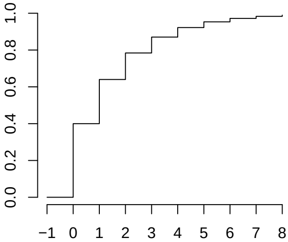
$$\begin{aligned} P(X \leq x) &= P(Y_1 = 1 \cup Y_2 = 1 \cup \cdots \cup Y_{x+1} = 1) \\ &= 1 - P(Y_1 = 0 \cap Y_2 = 0 \cap \cdots \cap Y_{x+1} = 0) \\ &= 1 - P(Y_1 = 0)P(Y_2 = 0) \cdots P(Y_{x+1} = 0) \\ &= 1 - (1-p)^{x+1} \end{aligned}$$

CDF of Geometric Distribution

Alternate proof: The event $X \leq x$ means that there are no more than x failures before the first success, which is the same as saying that a success occurs among the first $x+1$ trials:

$$\begin{aligned} P(X \leq x) &= P(Y_1 = 1 \cup Y_2 = 1 \cup \cdots \cup Y_{x+1} = 1) \\ &= 1 - P(Y_1 = 0 \cap Y_2 = 0 \cap \cdots \cap Y_{x+1} = 0) \\ &= 1 - P(Y_1 = 0)P(Y_2 = 0) \cdots P(Y_{x+1} = 0) \\ &= 1 - (1-p)^{x+1} \end{aligned}$$

Example: Here is a graph of the cdf for $p = .4$:



Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$E(X) = \sum x \cdot f(x) = \sum_{x=0}^{\infty} xp(1-p)^x$$

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$E(X) = \sum x \cdot f(x) = \sum_{x=0}^{\infty} xp(1-p)^x$$

$$= p(1-p) \sum_{x=0}^{\infty} x(1-p)^{x-1}$$

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$
\begin{aligned}
E(X) = \sum x \cdot f(x) &= \sum_{x=0}^{\infty} xp(1-p)^x \\
&= p(1-p)\sum_{x=0}^{\infty} x(1-p)^{x-1} \\
&= -p(1-p)\sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^x
\end{aligned}
$$

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$E(X) = \sum x \cdot f(x) = \sum_{x=0}^{\infty} xp(1-p)^x$$

$$= p(1-p) \sum_{x=0}^{\infty} x(1-p)^{x-1}$$

$$= -p(1-p) \sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^x$$

$$= -p(1-p) \frac{d}{dp} \sum_{x=0}^{\infty} (1-p)^x$$

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$
\begin{aligned}
E(X) = \sum x \cdot f(x) &= \sum_{x=0}^{\infty} xp(1-p)^x \\
&= p(1-p)\sum_{x=0}^{\infty} x(1-p)^{x-1} \\
&= -p(1-p)\sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\frac{1}{1-(1-p)}
\end{aligned}
$$

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$\begin{aligned}
E(X) = \sum x \cdot f(x) &= \sum_{x=0}^{\infty} xp(1-p)^x \\
&= p(1-p) \sum_{x=0}^{\infty} x(1-p)^{x-1} \\
&= -p(1-p) \sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^x \\
&= -p(1-p) \frac{d}{dp} \sum_{x=0}^{\infty} (1-p)^x \\
&= -p(1-p) \frac{d}{dp} \frac{1}{1-(1-p)} \\
&= -p(1-p) \frac{d}{dp} \frac{1}{p}
\end{aligned}$$

Mean of Geometric Distribution

We can use calculus to find the mean of a geometric random variable X:

$$
\begin{aligned}
E(X) = \sum x \cdot f(x) &= \sum_{x=0}^{\infty} xp(1-p)^x \\
&= p(1-p)\sum_{x=0}^{\infty} x(1-p)^{x-1} \\
&= -p(1-p)\sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\frac{1}{1-(1-p)} \\
&= -p(1-p)\frac{d}{dp}\frac{1}{p} = -p(1-p)\cdot\frac{-1}{p^2} = \frac{1}{p} - 1
\end{aligned}
$$

Example

If we roll 5 six-sided dice at once, and all the dice turn up the same number, this is called a "Yahtzee". The probability of getting a Yahtzee is $1/6^4 = 1/1296$. If we keep rolling 5 dice until we get a Yahtzee, what is the expected value of the number of times that we must try?

Example

If we roll 5 six-sided dice at once, and all the dice turn up the same number, this is called a "Yahtzee". The probability of getting a Yahtzee is $1/6^4 = 1/1296$. If we keep rolling 5 dice until we get a Yahtzee, what is the expected value of the number of times that we must try?

Solution: The number of failures X until we succeed is a geometric random variable with parameter $p = 1/1296$. We calculate

$$E(X) = \frac{1}{p} - 1 = \frac{1}{1/1296} - 1 = 1296 - 1 = 1295$$

Example

If we roll 5 six-sided dice at once, and all the dice turn up the same number, this is called a "Yahtzee". The probability of getting a Yahtzee is $1/6^4 = 1/1296$. If we keep rolling 5 dice until we get a Yahtzee, what is the expected value of the number of times that we must try?

Solution: The number of failures X until we succeed is a geometric random variable with parameter $p = 1/1296$. We calculate

$$E(X) = \frac{1}{p} - 1 = \frac{1}{1/1296} - 1 = 1296 - 1 = 1295$$

Now, the number of tries until we succeed is $X + 1$, which has expected value

$$E(X + 1) = E(X) + 1 = 1295 + 1 = 1296$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$E(X^2) = \sum x^2 \cdot f(x) = \sum_{x=0}^{\infty} x^2 p(1-p)^x$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$E(X^2) = \sum x^2 \cdot f(x) = \sum_{x=0}^{\infty} x^2 p(1-p)^x$$

$$= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$
\begin{aligned}
E(X^2) &= \sum x^2 \cdot f(x) = \sum_{x=0}^{\infty} x^2 p(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x \\
&= -p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x
\end{aligned}
$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$
\begin{aligned}
E(X^2) = \sum x^2 \cdot f(x) &= \sum_{x=0}^{\infty} x^2 p(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x \\
&= -p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\frac{1}{p}
\end{aligned}
$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$
\begin{aligned}
E(X^2) = \sum x^2 \cdot f(x) &= \sum_{x=0}^{\infty} x^2 p(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x \\
&= -p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\frac{1}{p} = p(1-p)\frac{d}{dp}(1-p)\left(\frac{-1}{p^2}\right)
\end{aligned}
$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$
\begin{aligned}
E(X^2) = \sum x^2 \cdot f(x) &= \sum_{x=0}^{\infty} x^2 p(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x \\
&= -p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\frac{1}{p} = p(1-p)\frac{d}{dp}(1-p)\left(\frac{-1}{p^2}\right) \\
&= p(1-p)\frac{d}{dp}\left(\frac{-1}{p^2}+\frac{1}{p}\right)
\end{aligned}
$$

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$
\begin{aligned}
E(X^2) = \sum x^2 \cdot f(x) &= \sum_{x=0}^{\infty} x^2 p(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x \\
&= -p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\frac{1}{p} = p(1-p)\frac{d}{dp}(1-p)\left(\frac{-1}{p^2}\right) \\
&= p(1-p)\frac{d}{dp}\left(\frac{-1}{p^2}+\frac{1}{p}\right) = 1 - \frac{3}{p} + \frac{2}{p^2}
\end{aligned}
$$

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

Variance of Geometric Distribution

By a similar technique, we can find the variance:

$$
\begin{aligned}
E(X^2) &= \sum x^2 \cdot f(x) = \sum_{x=0}^{\infty} x^2 p(1-p)^x \\
&= -p(1-p)\frac{d}{dp}\sum_{x=0}^{\infty} x(1-p)^x \\
&= -p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\sum_{x=0}^{\infty}(1-p)^x \\
&= p(1-p)\frac{d}{dp}(1-p)\frac{d}{dp}\frac{1}{p} = p(1-p)\frac{d}{dp}(1-p)\left(\frac{-1}{p^2}\right) \\
&= p(1-p)\frac{d}{dp}\left(\frac{-1}{p^2}+\frac{1}{p}\right) = 1-\frac{3}{p}+\frac{2}{p^2} \\
V(X) &= E(X^2) - [E(X)]^2 \\
&= 1-\frac{3}{p}+\frac{2}{p^2}-\left(\frac{1}{p}-1\right)^2 = \frac{1}{p^2}-\frac{1}{p}
\end{aligned}
$$

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial.

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial.

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial. In other words, $Y_7 = 1$ and $Y_1 + \cdots + Y_6 = 4$.

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial. In other words, $Y_7 = 1$ and $Y_1 + \cdots + Y_6 = 4$. Now, $Y_1 + \cdots + Y_6$ has a binomial distribution with parameters $n = 6$ and $p = .8$.

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial. In other words, $Y_7 = 1$ and $Y_1 + \cdots + Y_6 = 4$. Now, $Y_1 + \cdots + Y_6$ has a binomial distribution with parameters $n = 6$ and $p = .8$. Therefore

$$P(X = 7) = P(Y_1 + \cdots + Y_6 = 4 \cap Y_7 = 1)$$

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial. In other words, $Y_7 = 1$ and $Y_1 + \cdots + Y_6 = 4$. Now, $Y_1 + \cdots + Y_6$ has a binomial distribution with parameters $n = 6$ and $p = .8$. Therefore

$$\begin{aligned} P(X = 7) &= P(Y_1 + \cdots + Y_6 = 4 \cap Y_7 = 1) \\ &= P(Y_1 + \cdots + Y_6 = 4)P(Y_7 = 1) \end{aligned}$$

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial. In other words, $Y_7 = 1$ and $Y_1 + \cdots + Y_6 = 4$. Now, $Y_1 + \cdots + Y_6$ has a binomial distribution with parameters $n = 6$ and $p = .8$. Therefore

$$\begin{aligned} P(X = 7) &= P(Y_1 + \cdots + Y_6 = 4 \cap Y_7 = 1) \\ &= P(Y_1 + \cdots + Y_6 = 4)P(Y_7 = 1) \\ &= \binom{6}{4}(.8)^4(.2)^2 \cdot .8 \end{aligned}$$

Problem

A manufacturer produces widgets which work with probability .8. Suppose we test widgets one at a time. Let X be the number of bad widgets we test until we find 5 that work. What is the probability that $X = 2$?

Solution: Let Y_i be the Bernoulli random variable representing whether the ith widget works. Saying $X = 2$ is the same as saying that we have 2 failures before our 5th success, meaning our 5th success occurs on the 7th trial. This means the 7th trial is a success, and there are 4 successes up through the 6th trial. In other words, $Y_7 = 1$ and $Y_1 + \cdots + Y_6 = 4$. Now, $Y_1 + \cdots + Y_6$ has a binomial distribution with parameters $n = 6$ and $p = .8$. Therefore

$$\begin{aligned} P(X = 7) &= P(Y_1 + \cdots + Y_6 = 4 \cap Y_7 = 1) \\ &= P(Y_1 + \cdots + Y_6 = 4)P(Y_7 = 1) \\ &= \binom{6}{4}(.8)^4(.2)^2 \cdot .8 \approx .197 \end{aligned}$$

Negative Binomial Distribution

Recall that a binomial random variable X counts the number of successes out of a fixed number n of independent trials. In contrast, a **negative binomial** random variable X counts the number of failures before we achieve a fixed number r of successes. The possible values of X are $0, 1, 2, \ldots$.

Negative Binomial Distribution

Recall that a binomial random variable X counts the number of successes out of a fixed number n of independent trials. In contrast, a **negative binomial** random variable X counts the number of failures before we achieve a fixed number r of successes. The possible values of X are $0, 1, 2, \ldots$.

In the last slide we saw an example of a negative binomial random variable with $r = 5$ and $p = .8$. In the same way we can write down the pmf of a general negative binomial random variable:

$$P(X = x) = \binom{x + r - 1}{r - 1} p^r (1 - p)^x$$

Negative Binomial Distribution

Recall that a binomial random variable X counts the number of successes out of a fixed number n of independent trials. In contrast, a **negative binomial** random variable X counts the number of failures before we achieve a fixed number r of successes. The possible values of X are $0, 1, 2, \ldots$.

In the last slide we saw an example of a negative binomial random variable with $r = 5$ and $p = .8$. In the same way we can write down the pmf of a general negative binomial random variable:

$$P(X = x) = \binom{x + r - 1}{r - 1} p^r (1 - p)^x$$

A geometric random variable is simply a negative binomial random variable with $r = 1$.

Negative Binomial as Sum of Geometric

A negative binomial random variable X counts the number of failures it takes to get r successes, given that the trials are independent and each have probability p of success.

Negative Binomial as Sum of Geometric

A negative binomial random variable X counts the number of failures it takes to get r successes, given that the trials are independent and each have probability p of success.

The number of failures before the first success is a geometric random variable Y_1 with parameter p.

Negative Binomial as Sum of Geometric

A negative binomial random variable X counts the number of failures it takes to get r successes, given that the trials are independent and each have probability p of success.

The number of failures before the first success is a geometric random variable Y_1 with parameter p.

After the first success, the number of additional failures until the next success is an independent geometric random variable Y_2.

Negative Binomial as Sum of Geometric

A negative binomial random variable X counts the number of failures it takes to get r successes, given that the trials are independent and each have probability p of success.

The number of failures before the first success is a geometric random variable Y_1 with parameter p.

After the first success, the number of additional failures until the next success is an independent geometric random variable Y_2.

After k successes, the number of additional failures until the next success is an independent geometric random variable Y_{k+1}.

Negative Binomial as Sum of Geometric

A negative binomial random variable X counts the number of failures it takes to get r successes, given that the trials are independent and each have probability p of success.

The number of failures before the first success is a geometric random variable Y_1 with parameter p.

After the first success, the number of additional failures until the next success is an independent geometric random variable Y_2.

After k successes, the number of additional failures until the next success is an independent geometric random variable Y_{k+1}.

Therefore, we can express a negative binomial random variable X as a sum of r independent geometric random variables:

$$X = Y_1 + \cdots + Y_r$$

Mean and Variance of Negative Binomial

Expressing a negative binomial random variable X as a sum of independent geometric random variables allows us to compute its mean and variance.

$$X = Y_1 + \cdots + Y_r$$

Mean and Variance of Negative Binomial

Expressing a negative binomial random variable X as a sum of independent geometric random variables allows us to compute its mean and variance.

$$X = Y_1 + \cdots + Y_r$$

We know the mean and variance of geometric random variables:

$$E(Y_i) = \frac{1}{p} - 1$$
$$V(Y_i) = \frac{1}{p^2} - \frac{1}{p}$$

Mean and Variance of Negative Binomial

Expressing a negative binomial random variable X as a sum of independent geometric random variables allows us to compute its mean and variance.

$$X = Y_1 + \cdots + Y_r$$

We know the mean and variance of geometric random variables:

$$E(Y_i) = \frac{1}{p} - 1$$
$$V(Y_i) = \frac{1}{p^2} - \frac{1}{p}$$

Therefore the mean and variance of $X = Y_1 + \cdots + Y_n$ is

$$E(X) = E(Y_1) + \cdots + E(Y_r) = r\left(\frac{1}{p} - 1\right)$$
$$V(X) = V(Y_1) + \cdots + V(Y_r) = r\left(\frac{1}{p^2} - \frac{1}{p}\right)$$

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

- There are $\binom{12}{4} = 495$ ways to choose 4 balls from the 12 balls in the bag. Each of these 495 outcomes is equally likely.

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

- There are $\binom{12}{4} = 495$ ways to choose 4 balls from the 12 balls in the bag. Each of these 495 outcomes is equally likely.
- Drawing exactly 2 red balls means that the remaining 2 balls drawn are green.

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

- There are $\binom{12}{4} = 495$ ways to choose 4 balls from the 12 balls in the bag. Each of these 495 outcomes is equally likely.
- Drawing exactly 2 red balls means that the remaining 2 balls drawn are green.
- The number of ways to choose 2 of 5 red balls is $\binom{5}{2} = 10$.

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

- There are $\binom{12}{4} = 495$ ways to choose 4 balls from the 12 balls in the bag. Each of these 495 outcomes is equally likely.
- Drawing exactly 2 red balls means that the remaining 2 balls drawn are green.
- The number of ways to choose 2 of 5 red balls is $\binom{5}{2} = 10$.
- The number of ways to choose 2 of 7 green balls is $\binom{7}{2} = 21$.

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

- There are $\binom{12}{4} = 495$ ways to choose 4 balls from the 12 balls in the bag. Each of these 495 outcomes is equally likely.
- Drawing exactly 2 red balls means that the remaining 2 balls drawn are green.
- The number of ways to choose 2 of 5 red balls is $\binom{5}{2} = 10$.
- The number of ways to choose 2 of 7 green balls is $\binom{7}{2} = 21$.
- So the total number of ways to choose 2 red balls and 2 green balls is $10 \cdot 21 = 210$.

Another Counting Problem

Suppose a bag contains 5 red balls and 7 green balls. If we draw 4 balls at random, what is the probability that exactly 2 are red?

Solution:

- There are $\binom{12}{4} = 495$ ways to choose 4 balls from the 12 balls in the bag. Each of these 495 outcomes is equally likely.
- Drawing exactly 2 red balls means that the remaining 2 balls drawn are green.
- The number of ways to choose 2 of 5 red balls is $\binom{5}{2} = 10$.
- The number of ways to choose 2 of 7 green balls is $\binom{7}{2} = 21$.
- So the total number of ways to choose 2 red balls and 2 green balls is $10 \cdot 21 = 210$.
- The probability that this occurs is therefore $\frac{210}{495} = 42/99$.

Hypergeometric Distribution

In general if we select n individuals at random from a population of size N, where M individuals are of type A, and $N - M$ are of type B, then the number X of selected individuals of type A is a **hypergeometric** random variable:

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Hypergeometric Distribution

In general if we select n individuals at random from a population of size N, where M individuals are of type A, and $N - M$ are of type B, then the number X of selected individuals of type A is a **hypergeometric** random variable:

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Example: Suppose that out of a batch of 10 widgets, 7 are defective. If we randomly select 3 of the 10 widgets for inspection, what is the probability that we will find exactly 1 defective?

Hypergeometric Distribution

In general if we select n individuals at random from a population of size N, where M individuals are of type A, and $N - M$ are of type B, then the number X of selected individuals of type A is a **hypergeometric** random variable:

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Example: Suppose that out of a batch of 10 widgets, 7 are defective. If we randomly select 3 of the 10 widgets for inspection, what is the probability that we will find exactly 1 defective?

Here X is hypergeometric with $n = 3$, $N = 10$, $M = 7$, so

$$P(X = 1) = \frac{\binom{7}{1}\binom{3}{2}}{\binom{10}{3}} = \frac{7 \cdot 3}{120} = 7/40$$

Example – Animal Tagging

Researchers catch and tag 5 animals of a species thought to be near extinction in a certain region. After the animals have mixed back into the population, 10 animals from the population are randomly selected. Let X be the number of tagged animals out of these 10. If there are actually 25 animals of this type in the region, what is the probability that (a) $X = 2$? (b) $X \leq 2$?

Example – Animal Tagging

Researchers catch and tag 5 animals of a species thought to be near extinction in a certain region. After the animals have mixed back into the population, 10 animals from the population are randomly selected. Let X be the number of tagged animals out of these 10. If there are actually 25 animals of this type in the region, what is the probability that (a) $X = 2$? (b) $X \leq 2$?

$$P(X = 2) = \frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} \approx .385$$

Example – Animal Tagging

Researchers catch and tag 5 animals of a species thought to be near extinction in a certain region. After the animals have mixed back into the population, 10 animals from the population are randomly selected. Let X be the number of tagged animals out of these 10. If there are actually 25 animals of this type in the region, what is the probability that (a) $X = 2$? (b) $X \leq 2$?

$$P(X = 2) = \frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} \approx .385$$

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{\binom{5}{0}\binom{20}{10}}{\binom{25}{10}} + \frac{\binom{5}{1}\binom{20}{9}}{\binom{25}{10}} + \frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} \approx .699 \end{aligned}$$

Sampling Without Replacement

Suppose that out of a batch of 20 widgets, 5 are defective. If we randomly draw two widgets from the 20, we may represent the outcome using two Bernoulli random variables, Y_1 and Y_2, where $Y_1 = 1$ if the first widget drawn is defective, and $Y_2 = 1$ if the second widget drawn is defective.

Suppose that out of a batch of 20 widgets, 5 are defective. If we randomly draw two widgets from the 20, we may represent the outcome using two Bernoulli random variables, Y_1 and Y_2, where $Y_1 = 1$ if the first widget drawn is defective, and $Y_2 = 1$ if the second widget drawn is defective.

However, Y_1 and Y_2 are dependent, because if the first widget drawn is defective, this reduces the probability that the second widget drawn will be defective:

$$P(Y_1 = 1) = P(Y_2 = 1) = 5/20 = .25$$
$$P(Y_2 = 1 \mid Y_1 = 1) = 4/19 \approx .211$$

Sampling Without Replacement

Suppose that out of a batch of 20 widgets, 5 are defective. If we randomly draw two widgets from the 20, we may represent the outcome using two Bernoulli random variables, Y_1 and Y_2, where $Y_1 = 1$ if the first widget drawn is defective, and $Y_2 = 1$ if the second widget drawn is defective.

However, Y_1 and Y_2 are dependent, because if the first widget drawn is defective, this reduces the probability that the second widget drawn will be defective:

$$P(Y_1 = 1) = P(Y_2 = 1) = 5/20 = .25$$
$$P(Y_2 = 1 \mid Y_1 = 1) = 4/19 \approx .211$$

This situation is called **sampling without replacement** because once a widget is drawn, it is removed from the batch and may not be drawn again. The total number of defective widgets drawn in this way, $X = Y_1 + Y_2$, is a *hypergeometric* random variable.

Sampling With Replacement

Again suppose that out of a batch of 20 widgets, 5 are defective. Now draw two widgets at random, but this time after drawing the first widget, return it to the batch before randomly choosing the second widget. Thus there is a chance that the same widget will be chosen twice. This is called **sampling with replacement**.

Sampling With Replacement

Again suppose that out of a batch of 20 widgets, 5 are defective. Now draw two widgets at random, but this time after drawing the first widget, return it to the batch before randomly choosing the second widget. Thus there is a chance that the same widget will be chosen twice. This is called **sampling with replacement**.

In this case, the two Bernoulli random variables Y_1 and Y_2 are independent, and the total number of defective widgets drawn $X = Y_1 + Y_2$ is a *binomial* random variable with $n = 2$ and $p = 5/20 = 1/4$.

Sampling With Replacement

Again suppose that out of a batch of 20 widgets, 5 are defective. Now draw two widgets at random, but this time after drawing the first widget, return it to the batch before randomly choosing the second widget. Thus there is a chance that the same widget will be chosen twice. This is called **sampling with replacement**.

In this case, the two Bernoulli random variables Y_1 and Y_2 are independent, and the total number of defective widgets drawn $X = Y_1 + Y_2$ is a *binomial* random variable with $n = 2$ and $p = 5/20 = 1/4$.

In this case, the size of the batch (20 widgets) is irrelevant. If it had been a batch of 2000 widgets with 500 defective, it would not change the distribution of X. All that matters in this case is the *proportion* of defective widgets.

Relationship between Binomial and Hypergeometric

If the size of the batch is very large (say, 10000 widgets) and only a few widgets are drawn, then it makes little difference whether we sample with or without replacement, because it is very unlikely that any widget would be chosen more than once anyway. In this case, the hypergeometric and binomial distributions are practically identical.

Relationship between Binomial and Hypergeometric

If the size of the batch is very large (say, 10000 widgets) and only a few widgets are drawn, then it makes little difference whether we sample with or without replacement, because it is very unlikely that any widget would be chosen more than once anyway. In this case, the hypergeometric and binomial distributions are practically identical.

In mathematical terms, the pmf of a hypergeometric random variable approaches the pmf of a binomial random variable, in the limit as we increase the population size N while keeping the same proportion $p = M/N$.

Binomial as Limit of Hypergeometric

Given a hypergeometric random variable X with $M/N = p$ and n held constant while $N \to \infty$,

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Binomial as Limit of Hypergeometric

Given a hypergeometric random variable X with $M/N = p$ and n held constant while $N \to \infty$,

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \frac{\dfrac{M(M-1)\cdots(M-x+1)}{x!} \cdot \dfrac{(N-M)\cdots(N-M-n+x+1)}{(n-x)!}}{\dfrac{N(N-1)\cdots(N-n+1)}{n!}}$$

Binomial as Limit of Hypergeometric

Given a hypergeometric random variable X with $M/N = p$ and n held constant while $N \to \infty$,

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \frac{\dfrac{M(M-1)\cdots(M-x+1)}{x!} \cdot \dfrac{(N-M)\cdots(N-M-n+x+1)}{(n-x)!}}{\dfrac{N(N-1)\cdots(N-n+1)}{n!}}$$

$$= \frac{n!}{x!(n-x)!} \frac{M(M-1)\cdots(M-x+1)}{N(N-1)\cdots(N-x+1)} \cdot \frac{(N-M)\cdots(N-M-n+x+1)}{(N-x)\cdots(N-n+1)}$$

Binomial as Limit of Hypergeometric

Given a hypergeometric random variable X with $M/N = p$ and n held constant while $N \to \infty$,

$$
\begin{aligned}
P(X = x) &= \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \\
&= \frac{\dfrac{M(M-1)\cdots(M-x+1)}{x!} \cdot \dfrac{(N-M)\cdots(N-M-n+x+1)}{(n-x)!}}{\dfrac{N(N-1)\cdots(N-n+1)}{n!}} \\
&= \frac{n!}{x!(n-x)!} \frac{M(M-1)\cdots(M-x+1)}{N(N-1)\cdots(N-x+1)} \cdot \frac{(N-M)\cdots(N-M-n+x+1)}{(N-x)\cdots(N-n+1)} \\
&= \binom{n}{x} \frac{p(p-\frac{1}{N})\cdots(p-\frac{x-1}{N})}{1(1-\frac{1}{N})\cdots(1-\frac{x-1}{N})} \cdot \frac{(1-p)\cdots(1-p-\frac{n-x-1}{N})}{(1-\frac{x}{N})\cdots(1-\frac{n-1}{N})}
\end{aligned}
$$

Binomial as Limit of Hypergeometric

Given a hypergeometric random variable X with $M/N = p$ and n held constant while $N \to \infty$,

$$
\begin{aligned}
P(X = x) &= \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \\
&= \frac{\dfrac{M(M-1)\cdots(M-x+1)}{x!}\cdot\dfrac{(N-M)\cdots(N-M-n+x+1)}{(n-x)!}}{\dfrac{N(N-1)\cdots(N-n+1)}{n!}} \\
&= \frac{n!}{x!(n-x)!}\,\frac{M(M-1)\cdots(M-x+1)}{N(N-1)\cdots(N-x+1)}\cdot\frac{(N-M)\cdots(N-M-n+x+1)}{(N-x)\cdots(N-n+1)} \\
&= \binom{n}{x}\frac{p(p-\frac{1}{N})\cdots(p-\frac{x-1}{N})}{1(1-\frac{1}{N})\cdots(1-\frac{x-1}{N})}\cdot\frac{(1-p)\cdots(1-p-\frac{n-x-1}{N})}{(1-\frac{x}{N})\cdots(1-\frac{n-1}{N})} \\
&\to \binom{n}{x}p^x(1-p)^{n-x}
\end{aligned}
$$

Binomial as Limit of Hypergeometric

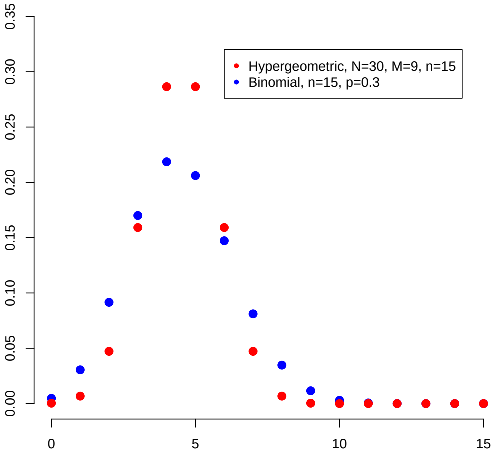
Given a hypergeometric random variable X with $M/N = p$ and n held constant while $N \to \infty$,

$$
\begin{aligned}
P(X = x) &= \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \\
&= \frac{\dfrac{M(M-1)\cdots(M-x+1)}{x!} \cdot \dfrac{(N-M)\cdots(N-M-n+x+1)}{(n-x)!}}{\dfrac{N(N-1)\cdots(N-n+1)}{n!}} \\
&= \frac{n!}{x!(n-x)!} \frac{M(M-1)\cdots(M-x+1)}{N(N-1)\cdots(N-x+1)} \cdot \frac{(N-M)\cdots(N-M-n+x+1)}{(N-x)\cdots(N-n+1)} \\
&= \binom{n}{x} \frac{p(p-\frac{1}{N})\cdots(p-\frac{x-1}{N})}{1(1-\frac{1}{N})\cdots(1-\frac{x-1}{N})} \cdot \frac{(1-p)\cdots(1-p-\frac{n-x-1}{N})}{(1-\frac{x}{N})\cdots(1-\frac{n-1}{N})} \\
&\to \binom{n}{x} p^x (1-p)^{n-x}
\end{aligned}
$$

This agrees with the pmf of a binomial, $\text{Bin}(n, p)$.

Binomial as Limit of Hypergeometric

If we draw 15 widgets from a population with 30% defective, the number of defective units is a hypergeometric random variable, but if we increase the population size while keeping the same proportion defective, it approaches binomial.



Mean of Hypergeometric Distribution

Given a population of size N, of which M are type A, if we sample n random individuals without replacement, then the number of type A individuals selected is a hypergeometric random variable X.

Mean of Hypergeometric Distribution

Given a population of size N, of which M are type A, if we sample n random individuals without replacement, then the number of type A individuals selected is a hypergeometric random variable X.

We may express X as the sum of Bernoulli random variables, $X = Y_1 + \cdots + Y_n$, where $Y_i = 1$ if the ith individual sampled is a success. Each draw has probability M/N of being a success, so $E(Y_i) = M/N$. Therefore,

$$\begin{aligned} E(X) &= E(Y_1 + \cdots + Y_n) \\ &= E(Y_1) + \cdots + E(Y_n) \\ &= \frac{M}{N} + \cdots \frac{M}{N} = n \cdot \frac{M}{N} \end{aligned}$$

Mean of Hypergeometric Distribution

Given a population of size N, of which M are type A, if we sample n random individuals without replacement, then the number of type A individuals selected is a hypergeometric random variable X.

We may express X as the sum of Bernoulli random variables, $X = Y_1 + \cdots + Y_n$, where $Y_i = 1$ if the ith individual sampled is a success. Each draw has probability M/N of being a success, so $E(Y_i) = M/N$. Therefore,

$$\begin{aligned} E(X) &= E(Y_1 + \cdots + Y_n) \\ &= E(Y_1) + \cdots + E(Y_n) \\ &= \frac{M}{N} + \cdots \frac{M}{N} = n \cdot \frac{M}{N} \end{aligned}$$

If we set $p = M/N$, the probability of success for each draw, then we may write $E(X) = np$; this is the same mean as a binomial, $\mathrm{Bin}(n, p)$.

Variance of Hypergeometric Distribution

Although a hypergeometric random variable X is the sum of Bernoulli random variables, $X = Y_1 + \cdots + Y_n$, the random variables $Y_1, \ldots, Y_n$ are dependent. Therefore we *cannot* find the variance of X by simply summing the variances of $Y_1, \ldots, Y_n$.

Variance of Hypergeometric Distribution

Although a hypergeometric random variable X is the sum of Bernoulli random variables, $X = Y_1 + \cdots + Y_n$, the random variables $Y_1, \ldots, Y_n$ are dependent. Therefore we *cannot* find the variance of X by simply summing the variances of $Y_1, \ldots, Y_n$.

Later we will show that the variance of X is

$$V(X) = \frac{N-n}{N-1} \cdot np(1-p)$$

Variance of Hypergeometric Distribution

Although a hypergeometric random variable X is the sum of Bernoulli random variables, $X = Y_1 + \cdots + Y_n$, the random variables $Y_1, \ldots, Y_n$ are dependent. Therefore we *cannot* find the variance of X by simply summing the variances of $Y_1, \ldots, Y_n$.

Later we will show that the variance of X is

$$V(X) = \frac{N-n}{N-1} \cdot np(1-p)$$

We call $\frac{N-n}{N-1}$ the *finite population correction factor*. When the population size N is large compared to the sample size n, the correction factor is approximately 1 and the variance is approximately $np(1-p)$, the variance of a binomial, $\text{Bin}(n, p)$.

Poisson Process

Consider a process where events occur at random times, such as

- The arrival times of customers at a store
- Clicks of a Geiger counter exposed to a radioactive material
- Webpage requests on an internet server
- Incoming calls to a customer service center

Poisson Process

Consider a process where events occur at random times, such as

- The arrival times of customers at a store
- Clicks of a Geiger counter exposed to a radioactive material
- Webpage requests on an internet server
- Incoming calls to a customer service center

Such a process is called a **Poisson process** if the following assumptions hold:

1. The mean number of events which occur in a time interval of length t is λt, where λ is a constant, called the **rate** of the Poisson process.
2. Events occur only one at a time.
3. The number of events which occur in a time interval is independent of the number and timing of past events.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.
- Let X_k be the number of events which occur in the subinterval I_k, so $X = X_1 + X_2 + \cdots + X_n$.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.
- Let X_k be the number of events which occur in the subinterval I_k, so $X = X_1 + X_2 + \cdots + X_n$.
- By assumption (1), $E(X_k) = \lambda \cdot \frac{t}{n} = \frac{\mu}{n}$.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.
- Let X_k be the number of events which occur in the subinterval I_k, so $X = X_1 + X_2 + \cdots + X_n$.
- By assumption (1), $E(X_k) = \lambda \cdot \frac{t}{n} = \frac{\mu}{n}$.
- By assumption (2), if n is large, then the probability of more than one event occuring in any given interval I_k is very small.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.
- Let X_k be the number of events which occur in the subinterval I_k, so $X = X_1 + X_2 + \cdots + X_n$.
- By assumption (1), $E(X_k) = \lambda \cdot \frac{t}{n} = \frac{\mu}{n}$.
- By assumption (2), if n is large, then the probability of more than one event occuring in any given interval I_k is very small.
- So we may approximate X_k as a Bernoulli random variable with parameter $p = \frac{\mu}{n}$.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.
- Let X_k be the number of events which occur in the subinterval I_k, so $X = X_1 + X_2 + \cdots + X_n$.
- By assumption (1), $E(X_k) = \lambda \cdot \frac{t}{n} = \frac{\mu}{n}$.
- By assumption (2), if n is large, then the probability of more than one event occuring in any given interval I_k is very small.
- So we may approximate X_k as a Bernoulli random variable with parameter $p = \frac{\mu}{n}$.
- By assumption (3), the random variables $X_1, \ldots, X_n$ are independent. Thus X is approximately binomial, $\text{Bin}(n, \frac{\mu}{n})$.

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

Poisson Distribution

Recall from calculus that

$$\lim_{n \to \infty} \left(1 + \frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, $\text{Bin}(n, \frac{\mu}{n})$, so

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, Bin$(n, \frac{\mu}{n})$, so

$$P(X = x) \approx \binom{n}{x}(\mu/n)^x(1-\mu/n)^{n-x}$$

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, $\text{Bin}(n, \frac{\mu}{n})$, so

$$\begin{aligned} P(X=x) &\approx \binom{n}{x}(\mu/n)^x(1-\mu/n)^{n-x} \\ &= \frac{n(n-1)\cdots(n-x+1)}{x!}(\mu/n)^x(1-\mu/n)^{n-x} \end{aligned}$$

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, $\text{Bin}(n, \frac{\mu}{n})$, so

$$\begin{aligned}
P(X=x) &\approx \binom{n}{x}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{x!}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{nn\cdots n}(1-\mu/n)^n(1-\mu/n)^{-x}\cdot\frac{\mu^x}{x!}
\end{aligned}$$

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, $\text{Bin}(n, \frac{\mu}{n})$, so

$$\begin{aligned}
P(X=x) &\approx \binom{n}{x}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{x!}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{nn\cdots n}(1-\mu/n)^n(1-\mu/n)^{-x}\cdot\frac{\mu^x}{x!} \\
&\to 1\cdot e^{-\mu}\cdot 1\cdot\frac{\mu^x}{x!} \qquad\qquad (\text{as } n\to\infty)
\end{aligned}$$

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, $\text{Bin}(n, \frac{\mu}{n})$, so

$$\begin{aligned}
P(X=x) &\approx \binom{n}{x}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{x!}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{nn\cdots n}(1-\mu/n)^n(1-\mu/n)^{-x}\cdot\frac{\mu^x}{x!} \\
&\to 1\cdot e^{-\mu}\cdot 1\cdot\frac{\mu^x}{x!} \qquad\qquad (\text{as } n\to\infty) \\
&= \frac{e^{-\mu}\mu^x}{x!}
\end{aligned}$$

Poisson Distribution

Given a Poisson process with rate λ, the number X of events which occur in a time interval of length t is a **Poisson** random variable with mean $\mu = \lambda t$. The possible values of X are $0, 1, 2, \dots$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Poisson Distribution

Given a Poisson process with rate λ, the number X of events which occur in a time interval of length t is a **Poisson** random variable with mean $\mu = \lambda t$. The possible values of X are $0, 1, 2, \ldots$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Example: Suppose that at a small store, customers arrive at an average rate of 6 per hour. What is the probability that during a given hour only 3 or fewer customers will arrive?

Poisson Distribution

Given a Poisson process with rate λ, the number X of events which occur in a time interval of length t is a **Poisson** random variable with mean $\mu = \lambda t$. The possible values of X are $0, 1, 2, \ldots$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Example: Suppose that at a small store, customers arrive at an average rate of 6 per hour. What is the probability that during a given hour only 3 or fewer customers will arrive?

Assuming a Poisson process,

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} + \frac{e^{-6}6^3}{3!} \\ &= e^{-6}(1 + 6 + 18 + 36) = 61e^{-6} \approx .151 \end{aligned}$$

Example

Suppose that at random times a system suffers breakdowns requiring immediate repairs. If the system breaks down at a rate of once per year, what is the probability that the system will break down 3 or more times in one year?

Example

Suppose that at random times a system suffers breakdowns requiring immediate repairs. If the system breaks down at a rate of once per year, what is the probability that the system will break down 3 or more times in one year?

Solution: The given information suggests that the number X of breakdowns in a year is a Poisson random variable with mean 1.

Example

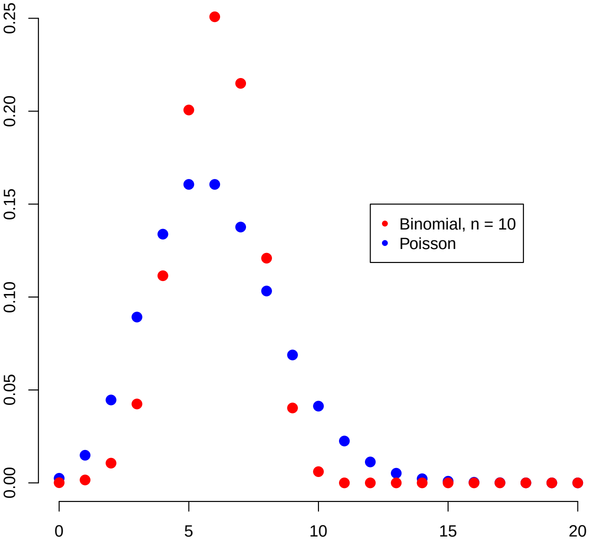
Suppose that at random times a system suffers breakdowns requiring immediate repairs. If the system breaks down at a rate of once per year, what is the probability that the system will break down 3 or more times in one year?

Solution: The given information suggests that the number X of breakdowns in a year is a Poisson random variable with mean 1.

$$\begin{aligned}
P(X \geq 3) &= 1 - P(X \leq 2) \\
&= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\
&= 1 - \frac{e^{-1}1^0}{0!} - \frac{e^{-1}1^1}{1!} - \frac{e^{-1}1^2}{2!} \\
&= 1 - e^{-1}\left(1 + 1 + \frac{1}{2}\right) \\
&= 1 - \frac{5e^{-1}}{2} \approx .080
\end{aligned}$$

Poisson as Limit of Binomial Distribution

Compare a $\text{Bin}(n, 6/n)$ with a Poisson with $\mu = 6$:



Poisson Distribution

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Poisson Distribution

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore, the pmf $f(x)$ of a Poisson random variable satisfies

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!}$$

Poisson Distribution

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore, the pmf $f(x)$ of a Poisson random variable satisfies

$$\begin{aligned}\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\ &= e^{-\mu}\sum_{x=0}^{\infty} \frac{\mu^x}{x!}\end{aligned}$$

Poisson Distribution

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore, the pmf $f(x)$ of a Poisson random variable satisfies

$$\begin{aligned}\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\ &= e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \\ &= e^{-\mu} e^{\mu}\end{aligned}$$

Poisson Distribution

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore, the pmf $f(x)$ of a Poisson random variable satisfies

$$\begin{aligned}
\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \\
&= e^{-\mu} e^{\mu} \\
&= 1
\end{aligned}$$

Poisson Distribution

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore, the pmf $f(x)$ of a Poisson random variable satisfies

$$\begin{aligned}
\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \\
&= e^{-\mu} e^{\mu} \\
&= 1
\end{aligned}$$

which shows that $f(x)$ is a valid pmf.

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$E(X) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!}$$

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$E(X) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!}$$
$$= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^x}{(x-1)!}$$

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!} \\ &= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^x}{(x-1)!} \\ &= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \end{aligned}$$

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^{x}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x}}{(x-1)!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x}}{x!}
\end{aligned}
$$

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= \mu
\end{aligned}
$$

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= \mu
\end{aligned}$$

This is what we expected based on the definition.

Variance of Poisson Distribution

We can also calculate the variance:

Variance of Poisson Distribution

We can also calculate the variance:

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!}$$

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!}
\end{aligned}
$$

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!}
\end{aligned}
$$

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^{x+1}}{x!} + \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!}
\end{aligned}
$$

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^{x+1}}{x!} + \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{(x-1)!} + \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!}
\end{aligned}
$$

Variance of Poisson Distribution

We can also calculate the variance:

$$\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^{x+1}}{x!} + \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{(x-1)!} + \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x}}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+2}}{x!} + \mu
\end{aligned}$$

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^{x+1}}{x!} + \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{(x-1)!} + \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+2}}{x!} + \mu = \mu^2 + \mu
\end{aligned}
$$

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^{x+1}}{x!} + \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{(x-1)!} + \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+2}}{x!} + \mu = \mu^2 + \mu
\end{aligned}
$$

So $V(X) = E(X^2) - [E(X)]^2 = \mu^2 + \mu - \mu^2 = \mu.$

Summary

The **probability mass function (pmf)** of a discrete random variable X is

$$f(x) = P(X = x)$$

The **cumulative distribution function (cdf)** is

$$F(x) = P(X \leq x)$$

A **Bernoulli** random variable X with parameter p takes the value 1 with probability p and the value 0 with probability $1 - p$.

The **binomial coefficient** $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is the number of ways of choosing a subset of k objects from a set of n objects.

Summary

The **expected value** or **mean** of a discrete random variable X is $E(X) = \sum x \cdot f(x)$.

1. $E(c) = c$
2. $E(cX) = cE(X)$
3. $E(X + Y) = E(X) + E(Y)$

The **variance** is $V(X) = E[(X - \mu)^2] = E(X^2) - \mu^2$ where $\mu = E(X)$. The **standard deviation** is $\sigma = \sqrt{V(X)}$.

Let X and Y be random variables, and let c be a constant. Then

1. $V(c) = 0$
2. $V(cX) = c^2 V(X)$
3. $V(X + c) = V(X)$
4. If X and Y are independent, $V(X + Y) = V(X) + V(Y)$.

Summary

Distribution	Random variable X	pmf	Mean	Variance
Binomial	Number of successes out of n independent trials	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n$	np	$np(1-p)$
Geometric	Number of failures until first success	$p(1-p)^x$ $x = 0, 1, 2, \dots$	$\frac{1}{p} - 1$	$\dfrac{1}{p^2} - \dfrac{1}{p}$
Negative Binomial	Number of failures until r successes	$\binom{x+r-1}{r-1} p^r (1-p)^x$ $x = 0, 1, 2, \dots$	$r(\frac{1}{p} - 1)$	$r\left(\dfrac{1}{p^2} - \dfrac{1}{p}\right)$
Hyper-geometric	Number of type A in a random sample of size n from a population of size N containing M of type A	$\dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ $x = 0, \dots, n$	np where $p = \frac{M}{N}$	$\frac{N-n}{N-1} \cdot np(1-p)$
Poisson	Number of events occuring over an interval of time, where events occur at random times	$\dfrac{e^{-\mu}\mu^x}{x!}$ $x = 0, 1, 2, \dots$	μ	μ