Chapter 3
Discrete Random Variables and Probability Distributions

Part 3: Some Common Discrete Random Variable Distributions

Section 3.4 Discrete Uniform Distribution
Section 3.5 Bernoulli trials and Binomial Distribution

Others sections will cover more of the common discrete distributions:
Geometric, Negative Binomial, Hypergeometric, Poisson

Common Discrete Random Variable Distributions

A random variable (r.v.) following any of the distributions below is limited to only discrete values.

- Discrete Uniform
- Bernoulli
- Binomial
- Geometric
- Negative Binomial
- Hypergeometric
- Poisson

Some of these distributions have mass (i.e. positive probability) at only a finite number of values, such as $\{1, 2, 3\}$ or $\{-2, -1, 0, 1, 2\}$.

Some of these discrete r.v. distributions have mass at a countably infinite number of values, like $\{0, 1, 2, 3, \ldots\}$

Discrete Uniform Distribution

Definition (Discrete Uniform Distribution)

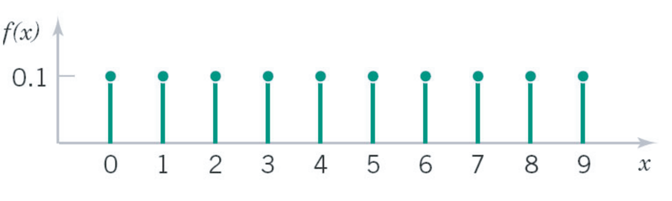
A random variable X has a discrete uniform distribution if each of the n values in its range, say $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$f(x_i) = \frac{1}{n}$$

where $f(x)$ represents the probability mass function (PMF).

Example (Discrete Uniform Distribution)

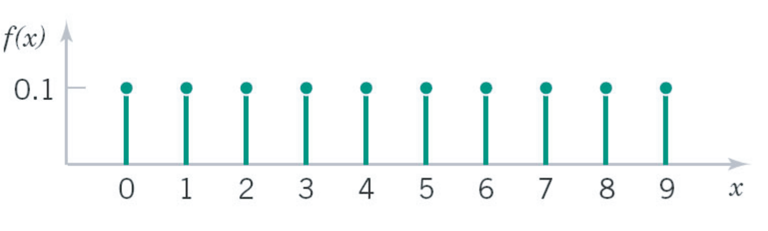
One example for $n = 10$ on consecutive integers from 0 to 9:



Discrete Uniform Distribution

Example (Discrete Uniform Distribution, cont.)

Let X represent a random variable taking on the possible values of $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, and each possible value has equal probability. This is a discrete uniform distribution and the probability for each of the 10 possible value is $P(X = x_i) = f(x_i) = \frac{1}{10} = 0.10$



Discrete Uniform Distribution - Mean and Variance

Definition (Mean and Variance for Discrete Uniform Distribution)

Suppose X is a discrete uniform random variable on the consecutive integers $a, a+1, a+2, \ldots, b$ for $a \leq b$.

The mean of X is

$$\mu = E(X) = \frac{b+a}{2}$$

The variance of X is

$$\sigma^2 = \frac{(b-a+1)^2-1}{12}$$

NOTE: If you compute the mean and variance by their definitions (i.e. using the possible x-values from a to b, $f(x_i) = \frac{1}{n}$, $E(X) = \sum x \cdot f(x)$,etc.), you will derive the above formulas. But for the special distributions, you just need to know how to use the above formulas to get the mean and variance, not derive it yourself.

Discrete Uniform Distribution - Mean and Variance

Example (Discrete Uniform Distribution)

What is the mean and variance of the random variable X described on the previous page? i.e. X is distributed uniform discrete on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

ANS:

Binomial Distribution

- Suppose a trial has only two outcomes, denoted by S for success and F for failure with $P(S) = p$ and $P(F) = 1 - p$.

  For example, a coin toss where a Head is a success S and a Tail is a failure F.

- Such a trial is called a Bernoulli trial.
- If we perform a random experiment by repeating n independent Bernoulli trials, then the random variable X representing the number of successes in the n trials has a binomial distribution.
- The possible values for binomial random variable X depends on the number of Bernoulli trials independently repeated, and is $\{0, 1, 2, \ldots, n\}$.
- Prior to the experiment, the number of successes to occur is unknown, but you could have as few as 0 successes, or as many as n successes.

Binomial Distribution

Example (Binomial Distribution with $p(\text{success}) = 0.4$)

Suppose 40% of a very large population of registered voters favor candidate Obama. A random sample of $n = 5$ voters will be selected, and X, the number favoring Obama out of 5, is to be observed.

What is the probability of getting no one who favors Obama,
i.e. what is $P(X = 0)$?

ASSUMPTION FOR THE ABOVE CALCULATION:
We assume here that this is a very large population. Thus, we assume that drawing one person at random without replacement who favors Obama <u>will not</u> substantially change the probability in subsequent draws. This means we assume we have <u>5 independent trials</u>, with the same probability of choosing an Obama supporter each time. And each trial is labeled as S or F.

Example (Binomial Distribution with $p(\text{success}) = 0.4$, cont.)

We'll consider 'picking someone who favors Obama' a success and X is the number of successes. (The terms *success* and *failure* are just labels).

$$\begin{aligned} p = P(success) &= 0.40 \\ 1 - p = P(failure) &= 0.60 \end{aligned}$$

Either 'Yes'(S) or 'No' (F) on each of 5 draws.

X (the number out of 5 favoring Obama) follows a *binomial distribution*.

What is the probability of getting 0 persons who favors Obama?

$$\begin{aligned} P(X = 0) &= \underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)} \qquad \{\text{independence between events}\} \\ &= (0.6)^5 \\ &= 0.07776 \end{aligned}$$

Binomial Distribution

Example (Binomial Distribution with $p(\text{success}) = 0.4$, cont.)

What is the probability of getting 1 person who favors Obama?
(5 configurations of Y and N...)

$$\begin{aligned} P(X=1) = {} & \underset{\text{Y}}{(0.4)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)} \\ & + \underset{\text{No}}{(0.6)}\underset{\text{Y}}{(0.4)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)} \\ & \quad \ldots \\ & + \underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{No}}{(0.6)}\underset{\text{Y}}{(0.4)} \\ = {} & \binom{5}{1}(0.4)^1(0.6)^4 = 0.25920 \end{aligned}$$

Example (Binomial Distribution with $p(\text{success}) = 0.4$, cont.)

What is the probability of getting 2 person who favors Obama?
$P(X = 2) = ?$

How many configurations of 2 Yes' and 3 No's can we have?
We have 5 'slots' to fill.

probability

$\binom{5}{2} = \frac{5!}{2!\,3!} = 10$

Y Y N N N	(0.4)(0.4)(0.6)(0.6)(0.6)
Y N Y N N	(0.4)(0.6)(0.4)(0.6)(0.6)
⋮	⋮
N N N Y Y	(0.6)(0.6)(0.6)(0.4)(0.4)

10 configurations

$$P(X = 2) = \binom{5}{2} (0.4)^2 (0.6)^3 = 10 \cdot (0.4)^2 (0.6^3) = 0.34560$$

Binomial Distribution

Example (Binomial Distribution with $p(\text{success}) = 0.4$, cont.)

We'll finish out the probability distribution for X...

$$P(X = 3) = \binom{5}{3}(0.4)^3(0.6)^2 = 10 \cdot (0.4)^3(0.6)^2 = 0.23040$$

$$P(X = 4) = \binom{5}{4}(0.4)^4(0.6)^1 = 5 \cdot (0.4)^4(0.6)^1 = 0.07680$$

$$P(X = 5) = \binom{5}{5}(0.4)^5(0.6)^0 = 1 \cdot (0.4)^5 = 0.01024$$

Note: $\sum_{i=0}^{5} P(X = i) = 1$ as this is a legitimate discrete probability distribution for $X \in \{0, 1, 2, 3, 4, 5\}$.

Binomial Distribution

Definition (Binomial Distribution)

A random experiment consists of n Bernoulli trials such that

1. The trials are independent
2. Each trial results in only two possible outcomes labeled as 'success' and 'failure' (dichotomous)
3. The probability of a success in each trial denoteD as p, remains constant

The random variable X that equals the number of trials that result in a success is a binomial random variable with parameters p and n and $0 < p < 1$ and $n = 1, 2, \ldots$.

The probability mass function (PMF) of X is

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

for $x = 0, 1, 2, \ldots, n$

Binomial Distribution

Example (Sampling water)

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant.

Let $X =$ the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with $p = 0.10$ and $n = 18$.

1. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

ANS:

Binomial Distribution

Example (Sampling water,cont.)

2. Find the probability that $3 \leq X \leq 5$.

ANS:

3. Find the probability that $X \geq 2$.

ANS:

Binomial Distribution

Definition (Mean and Variance for Binomial Distribution)

If X is a binomial random variable with parameters p and n, then

the mean of X is

$$\mu = E(X) = np$$

the variance of X is

$$\sigma^2 = V(X) = np(1-p)$$

NOTATION:
If X follows a binomial distribution with parameters p and n, we sometimes just write

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

Binomial Distribution

Example (Sampling water,cont.)

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Let $X =$ the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with $p = 0.10$ and $n = 18$.

Compute the expected value and variance of X with $X \sim Bin(18, 0.10)$.
ANS:

$$E(X) =$$

$$V(X) =$$

If X follows a binomial distribution, then X is a discrete random variable.

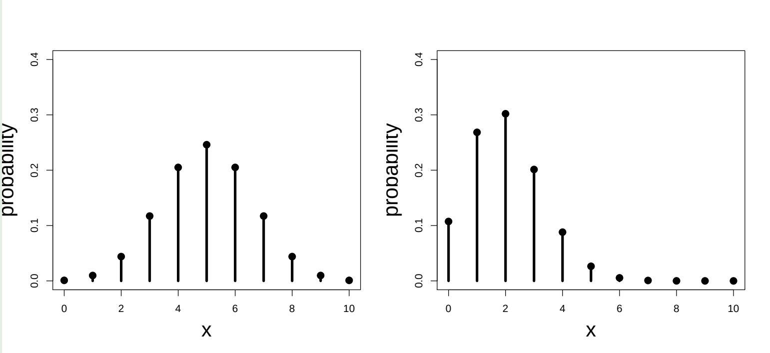
Binomial Distribution

What does the binomial distribution look like? Well, it depends on the parameters p and n. Here we see a few different binomial distributions...

Example (Binomial Distributions PMF)

$X \sim Bin(10, 0.5)$
equal chance of success/failure

$X \sim Bin(10, 0.2)$
small chance of success



Binomial Distribution

What does the binomial distribution look like? Well, it depends on the parameters p and n. Here we see a few different binomial distributions...

Example (Binomial Distributions PMF)

$X \sim Bin(10, 0.8)$
large chance of success

$X \sim Bin(10, 0.9)$
even larger chance of success