Chapter 3 Part 2

STAT 5700: Probability

Prof. Katie Fitzgerald, PhD

Fall 2025

Contents

3.4 Binomial Distribution

Bernoulli Experiments

A **Bernoulli experiment** is a random experiment that has two mutually exclusive and exhaustive outcomes, often referred to as "success" and "failure."

Lots of real world occurrences fit this binary structure (e.g. heads/tails, life/death, disease/no disease, graduate/not graduate). Note, sometimes choosing which of the two outcomes to call a "success" is arbitrary.

A sequence of **Bernoulli trials** occurs when a Bernoulli experiment is performed several *independent* times, and the probability of success, p, remains the same from trial to trial.

Example 1: flipping a fair coin 5 times corresponds to 5 Bernoulli trials with probability of success $p = 0.5$.
Example 2: Suppose the probability that a seed will germinate is 0.8, and we consider germination to be a "success." If we plant 10 seeds and can assume the germination of one seed is independent of the germination of another seed, this corresponds to 10 Bernoulli trials with $p = 0.8$.

Bernoulli Distribution

Let the random variable Y be the outcome of a Bernoulli experiment, where $Y = 1$ denotes a "success" and $Y = 0$ denotes a "failure." The probability distribution of Y is given by

$$p(y) = p^y(1-p)^{1-y}, \quad y = 0, 1,$$

and we say that Y follows the **Bernoulli distribution**, or that Y is a **Bernoulli** random variable.

Sometimes we will denote the probability of failure as $q = (1 - p)$.

Exercise: Find the mean and variance for the Bernoulli distribution. That is, find $E(Y)$ and $V(Y)$ when Y is a Bernoulli random variable.

Exercise:
Joel Embiid, star player drafted as the #3 overall pick by the Philadelphia 76ers, made 86% of his free throws in his rookie season. Suppose Embiid is fouled while attempting a 3 point shot with 1 second left in the game and the Sixers down by 2 points. What is the probability the Sixers win the game? Tie? Lose? Develop a random variable (and probability distribution) to answer these questions.

Binomial distribution

In a sequence of Bernoulli trials, we are often interested in the total number of successes but not the actual order of their occurrences.

If we let the random variable Y equal the number of observed successes in n Bernoulli trials, then the possible values of Y are $0, 1, 2, ..., n$.

If y successes occur, where $y = 0, 1, 2, ..., n$, then $n - y$ failures occur. The number of ways of selecting y positions for the y successes in the n trials is

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

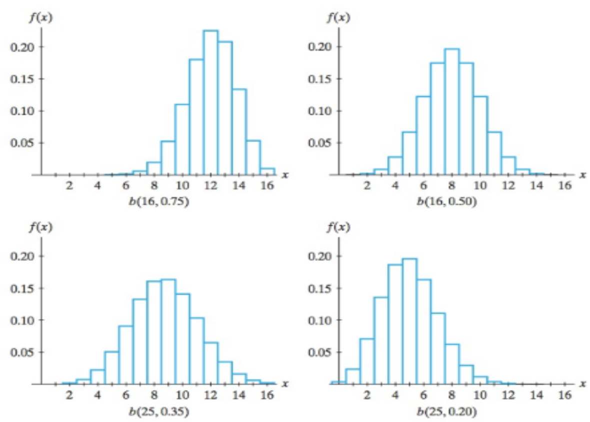
Since the trials are independent and since the probabilities of success and failure on each trial are, respectively, p and $q = 1 - p$, the probability of each of these ways is $p^y(1-p)^{n-y}$.

Thus, $p(y)$, the probability distribution of Y, is the sum of the probabilities of the $\binom{n}{y}$ mutually exclusive events; that is,

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

If Y has probability distribution $p(y) = \binom{n}{y} p^y (1-p)^{n-y}$, we say Y follows the **binomial distribution** and the probabilities $p(y)$ are called the **binomial probabilities**.
A binomial experiment satisfies the following properties:

1. A Bernoulli (success–failure) experiment is performed n times, where n is a (non-random) constant.
2. The trials are independent.
3. The probability of success on each trial is a constant p; the probability of failure is $q = 1 - p$.
4. The random variable Y equals the number of successes in the n trials.



Sometimes we write $Y \sim b(n, p)$, which is read as "Y is distributed as a binomial random variable with parameters n and p". That is, if we say $Y \sim b(23, 0.14)$, we mean Y is the number of successes in $n = 23$ Bernoulli trials that each have probability of success $p = 0.14$.

Exercise:
A manufacturing process has historically produced defective items every 1 out of 20. Five objects are selected independently from the production line.

a. What is the probability none of the selected items are defective?

b. What is the probability of 1 or fewer defective?

c. What is the probability more than 2 are defective?

Binomial cdf

Like in the previous example, cumulative probabilities are often of interest when working with the Binomial distribution. Recall that cumulative probabilities are given by the cumulative distribution function

$$F(y) = P(Y \leq y), \qquad -\infty < y < \infty.$$

The cdf of the Binomial distribution is given by

$$F(Y) = P(Y \leq y) = \sum_{y=0}^{y} \binom{n}{y} p^y (1-p)^{n-y}$$

Binomial expansion

The binomial expansion is given by

$$(a+b)^n = \sum_{y=0}^{n} \binom{n}{y} a^y b^{n-y}$$

Exercise:
Use the binomial expansion to show that the probability distribution of the Binomial distribution is a valid probability distribution.

3.5 Geometric Distribution

Recall: the Binomial distribution allows us to model the number of successes in n independent trials. However, sometimes we are interested in how long we have to wait until the first success happens. You have seen examples of this already, but here we will formally define the probability distribution as well as its mean and variance.

Definition 3.8 A random variable Y is said to have a *geometric probability distribution* if and only if

$$p(y) = (1-p)^{y-1} p, \qquad y = 1, 2, ... \text{and } 0 \leq p \leq 1$$

Here, Y is the number of the trial on which the first success occurs, and we could write $Y \sim geom(p)$. Just as with the Binomial distribution, we are assuming independent trials and that the probability of success is p on any individual trial.

Theorem 3.8: If $Y \sim geom(p)$, then

$$\mu = E(Y) = \frac{1}{p}$$

$$\sigma^2 = V(X) = \frac{1-p}{p^2}$$

PROOF of E(Y) (derivation of V(Y) is left as Exercise 3.85):

Example: Suppose that the probability of engine malfunction during any one-hour period is p = 0.02. Find the probability that a given engine will survive two hours.

If Y is the number of one-hour intervals until the first malfunction, find the mean and standard deviation of Y.

3.6 Negative Binomial Distribution

What if we are interested in knowing the number of the trail on which the 2nd, 3rd, or 4th success occurs?

Definition 3.9: A random variable Y is said to have a **negative binomial probability distribution** if and only if

$$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$$

We would write $Y \sim negbinom(p, r)$, where Y is the number of the trial on which the rth success occurs, and each trial is independent with probability of success p.

Theorem 3.9: If Y is a random variable with a negative binomial distribution,

$$\mu = E(Y) = \frac{r}{p}$$

$$\sigma^2 = V(X) = \frac{r(1-p)}{p^2}$$

Example: A geological study indicates that an exploratory oil well drilled in a particular region should strike oil with probability 0.2. Find the probability that the third oil strike comes on the 5th well drilled.

Example: Suppose that a basketball coach requires players to make 10 free throws at the end of practice before they are able to leave. For a player that has a 72% free throw percentage, how many free throws will she have to shoot on average? Define an appropriate random variable, it's probability distribution, and find the mean and standard deviation.

3.7 Hypergeometric Distribution

Suppose there is a population of size N in which exactly r elements have a particular feature, which will be considered a "success." The hypergeometric distribution allows us to model the probability of y successes in n draws from the population.

A random variable Y is said to have a *hypergeometric probability distribution* if and only if

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}},$$

where y is an integer $0, 1, 2, \dots n$ subject to the restrictions $y \leq r$ and $n - y \leq N - r$.

Theorem 3.10: If Y is a random variable with a hypergeometric distribution,

$$\mu = E(Y) = \frac{nr}{N}$$

$$\sigma^2 = V(Y) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Example: From a group of 20 PhD engineers, 10 are randomly selected for employment. What is the probability that the 10 selected include all the 5 best engineers in the group of 20? What is the average number of top 5 candidates that will be selected using this strategy?

3.8 Poisson Distribution

The count of the number of occurrences of an event in a continuous interval is an **approximate Poisson process**, with parameter $\lambda > 0$ if:

1. The number of occurances in nonoverlapping intervals are independent
2. The probability of exactly one occurance in a sufficiently short subinterval of length h is approximately λh.
3. The probability of two or more occurances in a sufficiently short subinterval is essentially 0.

The distribution to model the process is the **Poisson distribution**:

Let Y denote the number of occurrences in a "unit" length (the "unit" of interest), then $Y \sim$ Poisson(λ) with probability distribution:

$$p(y) = \frac{\lambda^y e^{-\lambda}}{y!}, \qquad y = 0, 1, 2, \ldots$$

In a Poisson distribution, the parameter λ can be interpreted as the expected rate of occurrences. It turns out that

$\lambda = E(Y) = V(Y)$

Example:
An old backup system was a computer tape, and flaws occurred on these tapes. In a particular situation, flaws (bad records) on a used computer tape occurred on the average of one flaw per 1200 feet. If one assumes a Poisson distribution, what is the distribution of Y, the number of flaws in a 4800-foot roll? What is the probability of 0 flaws on the 4800-foot roll?

Maclaurin series expansion (from Calculus)

The Maclaurin series expansion of a function $g(x)$ is given by:

$$g(x) = \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} x^n = g(0) + g'(0)x + \frac{g''(0)}{2!}x^2 + \frac{g'''(0)}{3!}x^3 + \dots$$

Therefore, a Maclaurin series expansion of

$$g(x) = e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Exercise: Use the Maclaurin series expansion to show that the probability distribution of the Poisson distribution is a valid probability distribution.

Exercise
Assume that a policyholder is four times more likely to file exactly two claims as to file exactly three claims. Assume also that the number Y of claims of this policyholder is Poisson. Determine the expectation $E(Y^2)$.

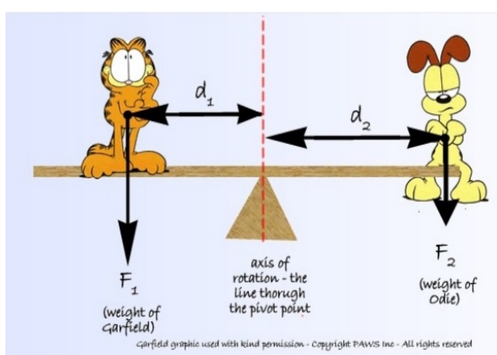
3.9 Moments and Moment Generating Functions

Special Expectations: Moments

Definition: The kth **moment** of a random variable X is the expected value of X^k and is denoted by $E(X^k)$, for each integer k. That is,

$$E(X^k) = \sum_{x \in \mathbb{S}} x^k p(x)$$

The term “moment” comes from physics: if the quantities $p(x)$ are point masses acting perpendicularly to the $x-$axis at distances x from the origin, $E(X^1)$ would be the $x-$coordinate of the center of gravity, and $E(X^2)$ would be the moment of inertia.



First Moment = Mean

Note that the first moment where $k = 1$, we have

$$\begin{aligned} E(Y^1) &= \sum_{y \in \mathbb{S}} y^1 p(y) \\ &= E(Y) = \sum_{y \in \mathbb{S}} y p(y) \\ &= \mu \end{aligned}$$

Therefore, we usually refer to the first moment as μ, the mean of Y.

Example:
Suppose Y is a random variable with support $\{1, 2, 3\}$ and probability distribution is given by $p(1) = 0.5$, $p(2) = 0.2$, $p(3) = 0.3$. Find the mean and show that the negative distances from the mean balance the positive.

Special Expectations: Central Moments

Definition The kth **central moment** of a random variable Y is the expected value of $(Y - \mu)^k$ and is denoted by $E[(Y - \mu)^k]$, for each integer k. That is,

$$E[(Y - \mu)^k] = \sum_{y \in \mathbb{S}} (y - \mu)^k p(y)$$

Recall that $\mu = E(Y)$ is the mean of Y, so the central moments are sometimes refered to as **moments about the mean**.

Exercise:
What is $E(Y - \mu)$?

Moment-generating functions

Definition 3.14 Let Y be a discrete random variable with probability distribution $p(y)$ and support $\mathbb{S}$. If there is a positive number h such that

$$E(e^{tY}) = \sum_{y \in \mathbb{S}} e^{ty} p(y)$$

exists and is finite for $-h < t < h$, then the function defined by $m(t) = E(e^{tY})$ is called the **moment-generating function** of Y. This function is often abbreviated as mgf.

$m(t) = E(e^{tY})$ is called the moment-generating function, because by taking derivatives of $m(t)$ at $t = 0$ can generate expressions for all the moments of a random variable Y!

Theorem

$$\frac{d^k}{dt^k} m(t)|_{t=0} = E(Y^k)$$

That is, the kth moment of Y is equal to the kth derivative of $m(t)$ evaluated at $t = 0$.

Example:
Let Y be a uniformly distributed random variable. Recall that the probability distribution of the uniform distribution is given by

$$p(y) = \frac{1}{m}, \qquad y = 1, 2, ..., m$$

Find an expression for the moment-generating function of the distribution. Then use the mgf to find the mean of Y.

Example:
If the moment-generating function of Y is $m(t) = \frac{2}{5}e^t + \frac{1}{5}e^{2t} + \frac{2}{5}e^{3t}$, find the mean, variance, and probability distribution of Y.

Moments of the Binomial Distribution

Exercise:

1. Find the mgf of the Binomial distribution.
2. Use the mgf to find the mean and the variance of the binomial distribution

Group Work

Problem 1

Some biology students were checking eye color in a large number of fruit flies. For the individual fly, suppose that the probability of white eyes is 1/4 and the probability of red eyes is 3/4, and that we may treat these observations as independent Bernoulli trials. What is the probability that at least four flies have to be checked for eye color to observe a white-eyed fly?

Problem 2

Suppose that Y is a random variable with a geometric distribution. Show that

a. $\sum_y p(y) = \sum_{y=1}^{\infty} q^{y-1}p = 1$
b. $\frac{p(y)}{p(y-1)} = q$, for $y = 2, 3, ...$ This ratio is less than 1, implying that the geometric probabilities are monotonically decreasing as a function of y. If Y has a geometric distribution, what value of Y is the most likely (has the highest probability)?

Problem 3

About 7 months into Donald Trump's 2nd term as president (August 2025), a Gallup poll found that a record low of 39% of adults approved of how the Supreme Court is handling its job.

a. Find the probability distribution for Y, the number of calls until the first person is found who *does* express approval of the U.S. Supreme Court.
b. On average, how many calls are needed until the 1st approval is found?
c. Find the probability distribution for Z, the number of calls until the 50th person is found who approves of the U.S. Supreme Court.
d. On average, how many calls are needed until the 50th approval is found?

Problem 4

The employees of a firm that manufactures insulation are being tested for indications of asbestos in their lungs. The firm is requested to send three employees who have positive indications of asbestos on to a medical center for further testing. If 40% of the employees have positive indications of asbestos in their lungs, find the probability that 10 employees must be tested in order to find three positives.

Problem 5

A jury of 6 persons was selected from a group of 20 potential jurors, of whom 8 were Black and 12 were White. The jury was supposedly randomly selected, but it contained only 1 Black member. Do you have any reason to doubt the randomness of the selection?

Problem 6

The number of typing errors made by a typist has a Poisson distribution with an average of four errors per page. If more than four errors appear on a given page, the typist must retype the whole page. What is the probability that a randomly selected page does not need to be retyped?

Problem 7

Let Y be a random variable with probability distribution $p(y) = \frac{y}{6}, \quad y = 1, 2, 3$

a) Find an expression for the moment generating function of Y. That is, write $E(e^{tY})$ as a sum.
b) Use the mgf to show that $E(Y) = 7/3$
c) Use the mgf to show that $E(Y^2) = 6$
d) Find $V(Y)$

Problem 8

Obtain an expression for the mgf of the Poisson distribution. Use the mgf to show that $E(Y) = V(Y) = \lambda$ for a Poisson random variable.

Problem 7

In a lab experiment involving inorganic syntheses of molecular precursors to organometallic ceramics, the final step of a five-step reaction involves the formation of a metal-metal bond. The probability of such a bond forming is $p = 0.2$. Let Y equal the number of successful reactions out of $n = 25$ such experiments.

a) What distribution is appropriate to model Y?
b) Write out the probability distribution of Y
c) Find the probability that $Y = 1$
d) Find the probability that Y is at least 1
e) Find the mean, variance, and standard deviation of Y

Problem 8

A random variable Y has a binomial distribution with mean 6 and variance 3.6. Find $P(Y = 4)$.

Problem 9

Return to the Joel Embiid example (he makes 87% of his free throws, he's fouled shooting a 3-pointer). Describe in detail how you could, in principle, perform a by hand simulation involving physical objects (e.g. coins, dice, spinners, etc.) to estimate the probability that the 76ers lose. Be sure to describe (1) what one repetition of the simulation entails, and (2) how you would use the results of many repetitions. Note: you do NOT need to compute any numerical values.