Random Variables

23 July 2024
Revision: 1325

Aziz Manva

azizmanva@gmail.com

[Type here]

TABLE OF CONTENTS

1. Random Variables

1.1 Descriptive Statistics

A. Frequency

1.1: Cumulative Frequency

Example 1.2

The table shows mark of students in an exam out of 30. Calculate cumulative frequency.

	f_i	Cumulative Frequency
< 10	4	4
$11 - 20$	6	10
$21 - 30$	10	20

B. Central Tendency

1.3: Mean

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum x_i}{n}$$

➤ You add all the data points, and then divide by the number of data points.

1.4: Mean

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Example 1.5

I buy four oranges costing three dollars each, six apples costing two dollars each, and ten mangoes costing a dollar each. Find the average cost of one piece of fruit.

	f_i	x_i	$f_i x_i$
Oranges	4	3	12
Apples	6	2	12
Mangoes	10	1	10
	$\sum f_i = 20$		$\sum f_i x_i = 34$

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{34}{20} = \frac{17}{10} = 1.7$$

C. Measures of Dispersion

Measures of dispersion measure how far are values from the mean. Two data sets that have the same mean can have values that are close, or values that are far away.

Example 1.6
$$X = \{0,100,200\}, \quad Y = \{99,100,101\}$$
A. Is the average of X and Y the same?
B. Are the datasets equally close to their mean?

Part A
$$\bar{x} = \bar{y} = 100$$

Part B
Y has values further away from the mean as compared to X.

1.7: Range
$$Range = Max - Min$$

Example 1.8
$$X = \{0,100,200\}, \quad Y = \{99,100,101\}$$

$$Range \ for \ X = 200 - 0 = 200$$
$$Range \ for \ Y = 101 - 99 = 2$$

The range for Y is much smaller than the range for X.

1.9: Deviation from the Mean
$$x_i - \bar{x}$$

Example 1.10
$$X = \{0,100,200\}, \quad Y = \{99,100,101\}$$

Deviation from the mean
$$for \ X = \{-100,0,100\}$$
$$for \ Y = \{-1,0,1\}$$

1.11: Mean Deviation from the Mean
Mean deviation from the mean is zero.

➢ Hence, this is not a useful measure of deviation.

1.12: Standard Deviation
The standard deviation is the square root of the average squared deviation from the mean.

$$Population \ Standard \ Deviation = \sigma_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

➢ Standard deviation measures how far the values are from the mean.
➢ If all values are the same, the standard deviation will be zero.
➢ If values are far from the mean, the standard deviation will be high.

➢

Example 1.13

Without calculating, state which dataset has higher standard deviation:
$$A = \{1001, 1002, 1003\}, \qquad B = \{1, 3, 5\}$$

In A, the average is 1002, and hence we write:
$$A = \{1001, 1002, 1003\} = \{\bar{x} - 1, \bar{x}, \bar{x} + 1\}$$

In B, the average is 3, and hence we write:
$$B = \{1, 3, 5\} = \{\bar{x} - 2, \bar{x}, \bar{x} + 2\}$$

Since B has values which are farther away from the mean than A has, B has greater standard deviation.

Example 1.14

Find the population standard deviation for:
$$4, 8, 9$$

Calculate the average first:
$$\bar{x} = \frac{4 + 8 + 9}{3} = \frac{21}{3} = 7$$

Calculate the squared deviations: given by the square of the difference from the average for each value:
$$(x_1 - \bar{x})^2 = (4 - 7)^2 = (-3)^2 = 9$$
$$(x_2 - \bar{x})^2 = (8 - 7)^2 = 1^2 = 1$$
$$(x_3 - \bar{x})^2 = (9 - 7)^2 = 2^2 = 4$$

Take the sum of the squared deviations, and divide by the number of values
$$s_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} = \sqrt{\frac{9 + 1 + 4}{3}} = \sqrt{\frac{14}{3}}$$

1.15: Standard Deviation

$$Sample\ Standard\ Deviation = s_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}}$$

Example 1.16

Find the sample standard deviation for:
$$4, 8, 9$$

$$\bar{x} = \frac{4 + 8 + 9}{3} = \frac{21}{3} = 7$$
$$(x_1 - \bar{x})^2 = (4 - 7)^2 = (-3)^2 = 9$$
$$(x_2 - \bar{x})^2 = (8 - 7)^2 = 1^2 = 1$$
$$(x_3 - \bar{x})^2 = (9 - 7)^2 = 2^2 = 4$$

$$s_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{9 + 1 + 4}{3 - 1}} = \sqrt{\frac{14}{2}} = \sqrt{7}$$

1.2 Discrete RVs

A. Summary
This will be useful after you have the studied the topic.

1.17: PMF/PDF

$$PMF = p(x) = P(X = x)$$
$$PDF = f(x) \Rightarrow P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)\, dx$$

The probability must be nonnegative
The probability must sum to 1.

1.18: CDF

CDF is integral of PDF

$$CDF = \int PDF \text{ or } PMF$$

CDF is more useful for PDF.

1.19: Moments

Mean

$$E[X] = \sum_x x \cdot p(x), \qquad E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Mean of Transformations
Linearity of expectation

$$E[X + Y] = E[X] + E[Y]$$

Multiple of a random variable

$$E[cX] = cE[X]$$
$$E[aX + bY] = aE[X] + bE[Y]$$

1.20: LOTUS
LOTUS (Law of The Unconscious Statistician)

$$E[g(x)] = \sum_x g(x) \cdot p(x), \qquad E[g(x)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx$$

Variance

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

Variance of Transformations

$$Var[cX] = c^2 Var[X]$$
$$Var[X + Y] = Var[X] + Var[Y] + 2Cov[X, Y]$$

1.21: Calculation of Probabilities

Calculation of probabilities
Back Calculation

1.22: MGF and PGF

The moment generating function of a random variable X, if it exists, is given by:
$$M_X(t) = E[e^{tX}]$$

Raw, central and standardized moments
Using MGF to calculate moments

1.23: Transformations

1.24: Order Statistics

B. Definition

1.25: Random Variable
A random variable is a variable or a process with a numerical outcome.

Example
If you toss a coin, the outcomes are $\{Heads, Tails\}$, which are non-numeric. We can assign $Heads = 0, Tails = 1$, and then say:
$$X = \begin{cases} 1, & if\ the\ toss\ is\ Tails \\ 0, & if\ the\ toss\ is\ Heads \end{cases}$$

1.26: Notation
  ➢ Random variables are usually assigned capital letters X, Y, Z
  ➢ The values that random variables can take are usually assigned small letters x, y, z.

The probability that the random variable X takes on the value x is written as:
$$\boldsymbol{P\{X = x\}}$$

1.27: Classification
Variables can be classified as
  ➢ Discrete
  ➢ Continuous
They can also be classified as
  ➢ Nominal
  ➢ Ordinal
  ➢ Interval
  ➢ Ratio

1.28: Probability Mass Function (PMF)
A probability mass function is a function that assigns probability to the outcomes of a sample space.

It is common, though not required to use $p(x)$ as the function for the probability that the random variable X

takes the value x:
$$P(X = x)$$

Example 1.29

I toss a weighted coin. The probability of heads is $\frac{1}{3}$. Write the probability mass function of the random variable X, which takes value 1, when the coin is heads, and value 0, when the coin is tails as a table.

$$P(X = 1) = \frac{1}{3}$$
$$P(X = 0) = \frac{2}{3}$$

x	1	0
$p(x)$	$\dfrac{1}{3}$	$\dfrac{2}{3}$

Example 1.30

I roll a six-sided die. And the probability of getting x on a particular roll is given by the random variable X, which has the following probability distribution function:
$$P(X = x) = \frac{x}{21}, \quad 1 \le x \le 6, x \in \mathbb{N}$$
Tabulate this probability distribution.

Solution

Die Roll	1	2	3	4	5	6	Total
Probability	$\dfrac{1}{21}$	$\dfrac{2}{21}$	$\dfrac{3}{21}$	$\dfrac{4}{21}$	$\dfrac{5}{21}$	$\dfrac{6}{21}$	1

Substitute $x = 1$ in $P(X = x) = \frac{x}{21}$ to get:
$$P(X = 1) = \frac{1}{21}$$
Substitute $x = 2$ in $P(X = x) = \frac{x}{21}$ to get:
$$P(X = 2) = \frac{2}{21}$$

Example 1.31

The probability distribution on the roll of a six-sided die is given below. State the probability distribution function for the table below.

Die Roll	1	2	3	4	5	6	Total
Probability	$\dfrac{1}{63}$	$\dfrac{2}{63}$	$\dfrac{4}{63}$	$\dfrac{8}{63}$	$\dfrac{16}{63}$	$\dfrac{32}{63}$	1

Solution

The denominator in each probability is constant.
The numerators are successive powers of 2:

$$1, 2, 4, \ldots, 32 \rightarrow 2^0, 2^1, 2^2, \ldots, 2^5 \rightarrow 2^{1-1}, 2^{2-1}, 2^{3-1}, \ldots, 2^{6-1}$$

And we

$$P(X = x) = \frac{2^{x-1}}{63}, \quad 1 \le x \le 6, x \in \mathbb{N}$$

C. Calculation of Probabilities

1.32: Calculation of Probabilities
➢ The probability mass function can be used to calculate probabilities.
➢ This is particularly useful if the PMF is given directly to you in the question.

Example 1.33
Tabulate the probability mass function of the number of heads when a fair coin tossed twice.
Use this PMF to calculate the probability that the number of heads is less than or equal to 1?

x	0	1	2
$P(X = x) = p(x)$	$\dfrac{1}{4}$	$\dfrac{1}{2}$	$\dfrac{1}{4}$

$$P(X \le 1) = P(X = 0) + P(X = 1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

D. Properties of a PMF

1.34: Range of Probability
A probability mass function is associated with the probability that a random variable X takes the value x. It is a specific case of probability.
Hence, a probability mass function must obey all known laws of probability. In particular:
➢ Probabilities are non-negative: $P(X = x) \ge 0$
➢ The sum of probabilities for all outcomes is 1: $\sum_x p(x) = 1$

Since the probabilities are nonnegative, and their sum is 1, we can conclude that:
$$P(X = x) \le 1$$

Example 1.35
I toss a fair coin twice, and count the number of heads X. Verify that the sum of probabilities is 1.

x	0	1	2
$P(X = x) = p(x)$	$\dfrac{1}{4}$	$\dfrac{1}{2}$	$\dfrac{1}{4}$

$$\sum_x p(x) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

Example 1.36
A particular very unfair coin is ten times more likely to come up heads than tails. (Heads and tails are the only possible outcomes).

State the probability mass distribution of this unfair coin as a table.

Method I
We also know that the sum of the probabilities must be 1:
$$P(H) + P(T) = 1$$

Substitute $P(H) = 10 \times P(T)$:
$$10P(T) + P(T) = 1$$
$$P(T) = \frac{1}{11}$$
$$P(H) = 10 \times \frac{1}{11} = \frac{10}{11}$$

Coin	Heads	Tails
Probability	$\dfrac{10}{11}$	$\dfrac{1}{11}$

Shortcut
Since $P(Heads)$ is ten times the probability of tails, we simply divide the probabilities in the ratio
$$10:1$$

Since $10 + 1 = 11$
$$P(Heads) = 10 \times \frac{1}{11}, \qquad P(Tails) = 1 \times \frac{1}{11}$$

Example 1.37
The weatherman has made the following predictions:
- ➢ The probability it will rain is three times the probability that it will be sunny.
- ➢ The probability it will be sunny is three the probability that it will be overcast.

(There are exactly three mutually exclusive possibilities for the weather: Rain, Sunny, Overcast)

State the probability distribution function:
- A. As a table
- B. As a formula

As a Table
$$P(Rain) = r, \qquad P(Sunny) = s, \qquad P(Overcast) = o$$

$$r = 3s$$
$$s = 3o$$

$$r + s + o = 1$$
$$9o + 3o + o = 1$$
$$o = \frac{1}{13} \Rightarrow s = 3o = \frac{3}{13} \Rightarrow r = 3s = \frac{9}{13}$$

Weather	Overcast	Sunny	Rain
	0	1	2

Probability	$\dfrac{1}{13}$	$\dfrac{3}{13}$	$\dfrac{9}{13}$

As a Formula

$$P(X = x) = \frac{3^x}{13}, \ \ 0 \le x \le 2, x \in \mathbb{N}$$

1.38: Normalizing Constant

If a function $f(x)$ is otherwise suitable to be made into a PMF, but the sum of the probabilities is not 1, we can multiply by a *normalizing constant* to make it unity.

Unity mean to make it one= 1

Example 1.39

Determine the normalizing constant k if the probability distribution on the roll of a six-sided die has the probability mass function:

$$P(X = x) = kx$$

$$p(1) + p(2) + \cdots p(6) = 1$$
$$k + 2k + \cdots + 6k = 1$$
$$21k = 1$$
$$k = \frac{1}{21}$$

Example 1.40: Geometric Series

Calculate the value of the normalizing constant k so that the function below is a valid PMF.

$$f(x) = k \left(\frac{1}{2}\right)^x, \qquad x = \mathbb{W}$$

Consider the probability distribution function given by

$$P(X = x) = k \left(\frac{1}{2}\right)^x, x = 0,1,2,3, ...$$

The sum of the probabilities must be 1:

$$k \left(\frac{1}{2}\right)^0 + k \left(\frac{1}{2}\right)^1 + k \left(\frac{1}{2}\right)^2 + \cdots = 1$$

Take k common:

$$k \left[1 + \left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \cdots\right] = 1$$

The sum of the geometric series with $a = 1, r = \frac{1}{2}$ is $\frac{a}{1-r} = \frac{1}{1-\frac{1}{2}} = 2$

$$2k = 1 \Rightarrow k = \frac{1}{2}$$

E. Expressions of a Random Variable

Example 1.41

x	0	1	2	3
$P(X = x)$	$\dfrac{1}{6}$	$\dfrac{1}{2}$	$\dfrac{1}{5}$	$\dfrac{2}{15}$

A. Find the probability distribution of $Y = X^2 - 2X$
B. Find
  I. $P(Y = 0)$
  II. $P(Y < 3)$

x	0	1	2	3
$y = x^2 - 2x$	0	-1	0	3
$P(Y = y)$	$\dfrac{1}{6}$	$\dfrac{1}{2}$	$\dfrac{1}{5}$	$\dfrac{2}{15}$

$$P(Y = 0) = \frac{1}{6} + \frac{1}{5}$$
$$P(Y < 3) = 1 - \frac{2}{15} = \frac{13}{15}$$

F. Cumulative Probabilities
The probability that a random variable X takes the value x is given by
$$P(X = x) = p(x)$$

These are individual probabilities, or the probability that X takes a specific value.

1.42: Cumulative Probabilities
We can also calculate the probability that the random variable X takes a value less than or equal to x.
$$F(x) = P(X \leq x) = \sum_{X \leq x} P(X)$$

Example 1.43
I toss a fair coin twice, and count the number of heads X. Write the cumulative probability distribution.

x	0	1	2
$P(X = x) = p(x)$	$\dfrac{1}{4}$	$\dfrac{1}{2}$	$\dfrac{1}{4}$
$F(x)$	$\dfrac{1}{4}$	$\dfrac{3}{4}$	1

1.44: Values of the Cumulative Distribution Function
The probability that a random variable takes a value less than $-\infty$ is zero, and hence
$$F(-\infty) = 0$$

The probability that a random variable takes a value less than ∞ is one, and hence
$$F(\infty) = 1$$

1.45: Non decreasing property

Cumulative probabilities are non-decreasing.

This is because the CDF is the sum of probabilities.
And probabilities are non-negative.

G. Symmetric Distribution

1.46: Symmetric Distribution
A symmetric distribution will have a vertical line of symmetry.

Example 1.47
I toss a fair coin thrice, and count the number of heads X. Is the probability distribution symmetric?

x	0	1	2	3
$P(X = x) = p(x)$	$\dfrac{1}{8}$	$\dfrac{3}{8}$	$\dfrac{3}{8}$	$\dfrac{1}{8}$

$$Yes$$

1.48: Symmetric Distribution
In a symmetric distribution:
$$Mean = Median = Mode$$

H. Using Probability Rules

Example 1.49
I have an urn filled with blue, red and green balls. I draw a ball from the urn and replace it. I expect to get two red balls for every three blue balls that I draw. Also, for every three red balls that I draw I expect to get two green balls.
I draw a ball from the urn and replace it. Then I repeat the process. In all, I draw two balls.

    A. Tabulate the probability distribution of drawing a ball from the urn.
    B. If you draw two balls, find the probability of getting:
        I.   Two Blue Balls
      II.   A Red and a Green Ball
    III.   Balls of two different colours

Part A

Red	Blue	Green
2	3	
3		2
6	9	4

Normalize the probabilities:

$$6x + 9x + 4x = 1 \Rightarrow x = \frac{1}{19}$$

Outcome	Red	Blue	Green
Probability	$\dfrac{6}{19}$	$\dfrac{9}{19}$	$\dfrac{4}{19}$

Part B
Two Blue Balls

$$Two\ Blue\ Balls = P(B) \times P(B) = \frac{9}{19} \times \frac{9}{19} = \frac{81}{361}$$

Red and a Green Ball
The order is not specified, so it can be in any order

$$P(R) \times P(G) + P(G) \times P(R) = 2 \times \frac{6}{19} \times \frac{4}{19} = \frac{48}{361}$$

Balls of Two different Colours

$$P(Red + Blue) = 2 \times P(R) \times P(B) = 2 \times \frac{6}{19} \times \frac{9}{19} = \frac{108}{361}$$

$$P(Blue + Green) = 2 \times P(B) \times P(G) = 2 \times \frac{9}{19} \times \frac{4}{19} = \frac{72}{361}$$

Finally, the probability that we want is:
$$\frac{48}{361} + \frac{108}{361} + \frac{72}{361} = \frac{228}{361}$$

I. Reimann Zeta Properties

1.50: Reimann Zeta Function
The Reimann Zeta function is defined as

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n^s}, \qquad s \in \mathbb{C}$$

Where
$$\mathbb{C}\ is\ the\ set\ of\ complex\ numbers$$

➢ s is a complex number.
➢ When $s = 1$, we get the harmonic series, which diverges
➢ The summation converges when $Re(s) > 1$.

Example 1.51
The PMF of the zeta distribution, with real parameter $s > 1$, defined over the natural numbers k is
$$P(X = k) = \frac{1}{k^s} \cdot \frac{1}{\zeta(s)}$$

Verify that it is a valid PMF.
Assume for this question that $\zeta(s)$ is a real number.

Sum of the probabilities should be 1:

$$\sum_{k=1}^{\infty} \frac{1}{k^s} \cdot \frac{1}{\zeta(s)}$$

Since $\frac{1}{\zeta(s)}$ is independent of k, it can be "moved out" of the summation sign:

$$= \frac{1}{\zeta(s)} \sum_{k=1}^{\infty} \frac{1}{k^s}$$

Substitute $\sum_{k=1}^{\infty} \frac{1}{k^s} = \zeta(s)$

$$= \frac{1}{\zeta(s)} \cdot \zeta(s) = 1$$

Probabilities should be non-negative:

$$0 \le \frac{1}{k^s} \cdot \frac{1}{\zeta(s)} = P(X = k)$$

And since the sum of the probabilities is 1, and the probabilities are non-negative, each individual probability is less than 1.

Example 1.52

The *PMF* of the zeta distribution, with real parameter $s > 1$ is

$$P(X = k) = \frac{1}{k^s} \cdot \frac{1}{\zeta(s)}, \qquad k \in \mathbb{N}, \qquad \zeta(s) \in \mathbb{R}$$

Calculate the probability that a random variable X that follows the zeta distribution is a multiple of k, where k is some natural number.

The values that X can take are:

$$\{1,2,3, \dots \}$$

The values whose probability we want to calculate are:

$$\{k, 2k, 3k, \dots \}$$

Since we want multiples of k, we want the probability:

$$P(X = k) + P(X = 2k) + P(X = 3k) \dots$$

Use the PMF:

$$= \frac{1}{k^s} \cdot \frac{1}{\zeta(s)} + \frac{1}{2^s k^s} \cdot \frac{1}{\zeta(s)} + \cdots$$

Factor out $\frac{1}{k^s}$:

$$= \frac{1}{k^s} \left(\frac{1}{\zeta(s)} + \frac{1}{2^s} \cdot \frac{1}{\zeta(s)} + \cdots \right)$$

Note that $\frac{1}{\zeta(s)} + \frac{1}{2^s} \cdot \frac{1}{\zeta(s)} + \cdots$ is precisely the PMF of the zeta distribution and hence it sums to 1:

$$= \frac{1}{k^s} (1) = \frac{1}{k^s}$$

1.3 Mean and Variance

A. Definition

1.53: Mean and Variance

$$\text{Mean} = \sum_{i=1}^{n} x_i p(x) \qquad\qquad \text{Variance} = \sum_{i=1}^{n} (x_i - \bar{x})^2 p(x)$$

A discrete probability distribution is defined as a frequency distribution.
Hence, the mean and variance of a probability distribution can be calculated using the formula for frequency distributions:

$$\text{Mean} = \frac{1}{n} \sum_{i=1}^{n} x_i f_i \qquad\qquad \text{Variance} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 f_i$$

However, in a probability distribution, the frequencies are probabilities. So, we replace the frequencies with probabilities (using the probability mass function that we just defined).

$$\text{Mean} = \frac{1}{n} \sum_{i=1}^{n} x_i p(x) \qquad\qquad \text{Variance} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 p(x)$$

The sum of the probabilities must, by definition, be 1. Hence, the term $\frac{1}{n}$ vanishes.

$$\text{Mean} = \sum_{i=1}^{n} x_i p(x) \qquad\qquad \text{Variance} = \sum_{i=1}^{n} (x_i - \bar{x})^2 p(x)$$

Example 1.54
Find the pmf, mean and variance of X, if, on tossing a fair coin,
$$X = \begin{cases} 1, & \text{if the result is heads} \\ 0, & \text{if the result is tails} \end{cases}$$

Since the coin is fair, the results are equiprobable (equally possible):
$$p(0) = p(1) = \frac{1}{2}$$

We can tabulate the pmf:

X	0	1
$p(x)$	$\dfrac{1}{2}$	$\dfrac{1}{2}$

$$E(X) = \sum x p(x) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$\sum (x_i - \mu)^2 p_i = \left(0 - \frac{1}{2}\right)^2 \cdot \frac{1}{2} + \left(1 - \frac{1}{2}\right)^2 \cdot \frac{1}{2} = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

$$E[X^2] = \sum x^2 p(x) = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} = 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

Both formulae give us the same answer.

1.55: Expected Value
The mean of a random variable is the $probability - weighted$ average of the outcomes of the random variable.

$$Mean = E[X] = \mu = \sum_x x \cdot p(x)$$

➢ μ is the Greek letter mu.
➢ $E[X]$ is an operator (or function). It is not $E \cdot X$
➢ x is the value taken by the random variable.
➢ $p(x)$ is the probability that X takes the value x.
➢ The other names for mean are:
  ➢ Mathematical expectation
  ➢ Expected Value

Example 1.56: Expected Value

I roll a standard n-sided die. Let X be the number of dots on the face that comes up on top.
 A. Tabulate the probability distribution of X.
 B. Write a formula for $P(X = x)$
 C. What is the expected value of X?

Part A

x	1	2	.	.	.	n
$p(x)$	$\dfrac{1}{n}$	$\dfrac{1}{n}$	.	.	.	$\dfrac{1}{n}$

Part B

$$P(X = x) = \frac{1}{n}$$

Part C

To find the expected value:
$$1 \cdot p(1) + 2 \cdot p(2) + \cdots + n \cdot p(n)$$

Substitute the values from above:
$$= 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + \cdots + n \cdot \frac{1}{n}$$

Factor $\frac{1}{n}$:
$$= \frac{1}{n}(1 + 2 + \cdots + n)$$

Substitute $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$:
$$= \frac{1}{n}\left(\frac{n(n+1)}{2}\right) = \frac{n+1}{2}$$

1.57: Variance

Variance is the expected value of the sum of the squared deviations from the mean:
$$Variance = \sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 \cdot p(x) = E[X^2] - (E[X])^2$$

Consider a random variable X with mean μ.

The difference between the values $x_1, x_2, \ldots, x_n$ is given by:
$$x_i - \mu$$

Squaring the above gives:
$$(x_i - \mu)^2$$

The sum of these values is:
$$\sum_{i=1}^{n}(x_i - \mu)^2$$

And the above expression has expected value
$$\sum_{x}(x - \mu)^2 \cdot p(x)$$

Which is the variance.

1.58: Interpretation of Variance
Variance is a measure of dispersion. It quantifies how spread out a distribution is from its mean.

Example 1.59

Consider two students with marks in quizzes as follows:
$$Student\ 1: \frac{7}{10}, \frac{6.5}{10}, \frac{7.5}{10}$$
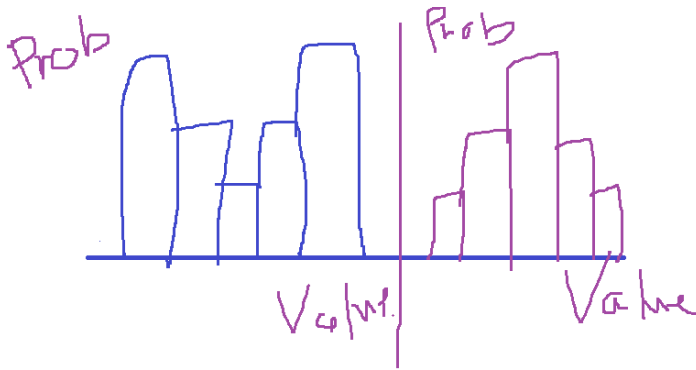$$Student\ 2: \frac{4}{10}, \frac{7}{10}, \frac{10}{10}$$

A. What is the average for Students 1 and 2.
B. You want to send Students 1 and 2 to perform in math competitions. You know that to win in Competition 1, you need $\frac{9}{10}$. To win in competition 2, you need $\frac{6}{10}$. You would like to maximize your chances of winning in both competitions. What strategy should you use?

$$Average\ for\ 1 = Average\ for\ 1 = \frac{7}{10}$$

For Competition 1, student 1 is not likely to succeed.
Hence, you should send student 2.

For Competition 2, student 2 may or not may not succeed. But student 1 is highly likely to succeed.
Hence, you should send student 1.

Example 1.60

➤ Both distributions have the same mean, but the distribution on the left is more spread out since there is a greater probability for extreme values.
➤ The distribution on the right is clustered around the mean.
➤ Both distributions are approximately symmetrical.

1.61: Standard Deviation
Standard deviation is the square root of the variance:
$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_x (x - \mu)^2 \cdot p(x)}$$

Example 1.62
I toss a fair coin twice, and count the number of heads X. Calculate the
    A. probability mass function for X.
    B. mean of X.
    C. variance of X.
    D. standard deviation of X.

X is the random variable that counts the number of heads in two tosses of a fair coin.

x	0	1	2	
$P(X = x) = p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	
$x \cdot p(x)$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\mu = 1$
$x - \mu$	-1	0	1	
$(x - \mu)^2$	1	0	1	
$(x - \mu)^2 \cdot p(x)$	$\frac{1}{4}$	0	$\frac{1}{4}$	$\sigma^2 = \frac{1}{2}$

$$Mean = \mu = \sum_x x \cdot p(x) = 0 + \frac{1}{2} + \frac{1}{2} = 1$$

$$Variance = \sigma^2 = \sum_x (x - \mu)^2 \cdot p(x) = \frac{1}{4} + 0 + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$$

$$Standard\ Deviation = \sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{2}}$$

(Calculator) Example 1.63
50 texts with probability 0.75, and 250 texts with probability 0.25.

		Plan 1	Plan 2
	$P(x)$	x_1	x_2
50 Texts	0.75	30	40
250 Texts	0.25	70	60

The expected cost is:
$$\mu_{Plan\ 1} = \sum x\, P(x) = 30\left(\frac{3}{4}\right) + 70\left(\frac{1}{4}\right) = \frac{90+70}{4} = \frac{160}{4} = 40$$
$$\mu_{Plan\ 2} = \sum xP(x) = 40\left(\frac{3}{4}\right) + 60\left(\frac{1}{4}\right) = \frac{120+60}{4} = \frac{180}{4} = 45$$
Plan 2 has higher expected value.

The variance and standard deviation are:
$$\sigma^2_{Plan\ 1} = \sum (x-\mu)^2 P(x) = (30-40)^2\left(\frac{3}{4}\right) + (70-40)^2\left(\frac{1}{4}\right) = 100\left(\frac{3}{4}\right) + 900\left(\frac{1}{4}\right) = \frac{300+900}{4} = 300$$
$$\sigma_{Plan\ 1} = \sqrt{\sigma^2_{Plan\ 1}} = \sqrt{300} = 17.32$$
$$\sigma^2_{Plan\ 2} = \sum (x-\mu)^2 P(x) = (40-45)^2\left(\frac{3}{4}\right) + (60-45)^2\left(\frac{1}{4}\right) = 25\left(\frac{3}{4}\right) + 225\left(\frac{1}{4}\right) = \frac{75+225}{4} = 75$$

$$\sigma_{Plan\ 2} = \sqrt{\sigma^2_{Plan\ 2}} = \sqrt{75} = 8.66$$

	Plan 1	Plan 2
μ	40	17.32
σ	45	8.66

B. Properties of Mean and Variance

1.64: Properties of Mean and Variance
If c is a constant and X is a random variable, then:
$$E[cX] = cE[X]$$
$$Var[cX] = c^2 Var[X]$$
$$SD[cX] = |c|SD[X]$$

Example 1.65
  A. X is a random variable with mean 5 and standard deviation 3. Determine the mean and standard deviation of $3X$.
  B. The number of monthly international phone calls that Anaya makes is a random variable with mean 12 and variance 2. The cost of each phone call is $5. Find the mean and the standard deviation of the monthly cost of the international phone calls that Anaya makes.

Part A
$$Mean[3X] = E[3X] = 3E[X] = 3(5) = 15$$
$$SD[3X] = 3SD[X] = 3(3) = 9$$

Part B
Let the random variable

$$X = No. of\ monthly\ international\ phone\ calls$$

We want to find:
$$E[5X] = 5E[X] = 5 \cdot 12 = 60$$
$$SD[5X] = 5SD[X] = 5\sqrt{2}$$

1.66: Constant Values
The expectation of a constant value is the constant value itself:
$$E[c] = c$$

The variance of a constant value is zero:
$$Var[c] = 0$$

➢ This are "trivial" properties that are useful when solving "non-trivial" questions.

1.67: Linearity of Expectation
The mean of the sum of two random variables is the sum of the respective means. In other words, you can just add the means:
$$E[X + Y] = E[X] + E[Y]$$

The variance of the sum of two random variables is the sum of the respective variance if the variables are independent:
$$Var[X + Y] = Var[X] + Var[Y]\ if\ X\ and\ Y\ are\ independent$$

➢ Note that the property does not apply to standard deviation, even if the variables are independent.
➢ To add standard deviations, first convert to variances, then add them, and then convert back to standard deviation.

1.68: Adding a Constant: Mean
The expected value of the sum of a constant a and a random variable X is the sum of the constant and the expected value of the random variable:
$$E[a + X] = a + E[X]$$

$$E[a + X]$$
Using linearity of expectation:
$$= E[a] + E[X]$$

Since the expected value of a constant is the constant itself:
$$a + E[X]$$

1.69: Adding a Constant: Mean
The variance of the sum of a constant a and a random variable X remains unchanged. It is the variance of the random variable only:
$$Var[a + X] = Var[X]$$

$$Var[a + X]$$
Since X is independent of a:
$$Var[a] + Var[X]$$

Since the variance of a constant is zero:
$$Var[X]$$

1.70: Linear Combinations of Random Variables
$$E[aX + bY] = aE[X] + bE[Y]$$

$$E[aX + bY]$$

Using linearity of expectation:
$$= E[aX] + E[bY]$$

Using the constant multiple rule:
$$= aE[X] + bE[Y]$$

1.71: Linear Combinations of Random Variables
If X and Y are independent:
$$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$$

$$Var[aX + bY]$$

Using linearity of expectation:
$$= Var[aX] + Var[bY]$$

Using the constant multiple rule:
$$a^2 Var[X] + b^2 Var[Y]$$

Example 1.72
If X and Y are independent, expand:
$$Var[aX - bY]$$
Use a change of variable.
$$Var[aX + (-b)Y]$$

$$= a^2 Var[X] + (-b)^2 Var[Y] = a^2 Var[X] + b^2 Var[Y]$$

Example 1.73
The runs scored by a cricketer in an innings of a test match are a random variable X with mean 50 and standard deviation 10. The runs in two different innings are independent of each other.
   A. Determine an expression for the random variable Y that tracks the average score in a test match with two innings.
   B. Show that $E[Y] = E[X]$
   C. Show that $SD[Y] < SD[X]$

Part A
Let X_1 and X_2 be random variables that follow the probability distribution X. We write
$$X_1, X_2 \sim X(\mu = 50, \sigma = 10)$$

Then the random variable Y is defined as:
$$Y = \frac{X_1 + X_2}{2} = \frac{1}{2}X_1 + \frac{1}{2}X_2$$

Part B
We want to calculate
$$E[Y]$$

Substitute the definition from Part A:
$$= E\left[\frac{1}{2}X_1 + \frac{1}{2}X_2\right]$$

Using linearity of expectation
$$\frac{1}{2}E[X_1] + \frac{1}{2}E[X_2]$$

Since X_1 and X_2 both follow the distribution X, substitute $X_1 = X, X_2 = X$:
$$= \frac{1}{2}E[X] + \frac{1}{2}E[X] = E[X]$$

Part C
$$Var[Y] = Var\left[\frac{X_1 + X_2}{2}\right] = \frac{Var[X_1] + Var[X_2]}{4} = \frac{Var[X] + Var[X]}{4} = \frac{1}{2}Var[X]$$
$$Var[Y] = \frac{1}{2}Var[X]$$
$$SD[Y] = \sqrt{Var[Y]} = \sqrt{\frac{1}{2}Var[X]} = \frac{1}{\sqrt{2}}SD[Y]$$

Example 1.74
A baker notes that pastry sales on a weekday have a mean of 20, and a standard deviation of 3. Pastry sales on a weekend have a mean of 40, and a standard deviation of 7. Pastry sales on any given day are independent of factors other weekday/weekend. Estimate the mean and standard deviation of the pastry sales over a week.

Let
$$WD = No.\,of\,pastries\,sold\,on\,a\,weekday$$
$$WE = No.\,of\,pastries\,sold\,on\,a\,weekday$$

$$E[5WD + 2WE] = 5E[WD] + 2E[WE] = 5 \cdot 20 + 2 \cdot 40 = 100 + 80 = 180$$

$$Var[5WD + 2WE] = 25Var[WD] + 4Var[WE] = 25 \cdot 3^2 + 4 \cdot 7^2 = 225 + 196 = 421$$
$$SD[5WD + 2WE] = \sqrt{Var[5WD + 2WE]} = \sqrt{421} = 20.52$$

Example 1.75
A follows a normal distribution with mean 10% and standard deviation 0.03. B follows a normal distribution with mean 8% and standard deviation 0.04. Determine the mean and standard deviation of the difference $A - B$. A and B are independent.

$$A \sim N(\mu = 10\%, \sigma^2 = 0.03^2)$$
$$B \sim N(\mu = 8\%, \sigma^2 = 0.04^2)$$

The mean is:
$$E[A - B] = E[A] - E[B] = 10\% - 8\% = 2\%$$

The variance is:

$$Var[A - B] = Var[A] + Var[B] = 0.05^2$$
$$SD[A - B] = 0.05$$

C. Advanced Topics

1.76: Formula for Mean

For a random variable X with non-negative support:
$$E(X) = \int_0^\infty \left(1 - F(x)\right) \, dx$$

Since $1 - F_X(x) = P(X \geq x) = \int_x^\infty f_X(t) \, \mathrm{d}t$,

$$\boxed{\int_0^\infty \left(1 - F_X(x)\right) \, \mathrm{d}x} = \int_0^\infty P(X \geq x) \, \mathrm{d}x = \int_0^\infty \int_x^\infty f_X(t) \, \mathrm{d}t\mathrm{d}x$$

Then change the order of integration:

$$= \int_0^\infty \int_0^t f_X(t) \, \mathrm{d}x\mathrm{d}t = \int_0^\infty [xf_X(t)]_0^t \, \mathrm{d}t = \int_0^\infty tf_X(t) \, \mathrm{d}t$$

Recognizing that t is a dummy variable, or taking the simple substitution $t = x$ and $\mathrm{d}t = \mathrm{d}x$,

$$= \int_0^\infty xf_X(x) \, \mathrm{d}x = \boxed{E(X)}$$

1.4 Mean and Variance-II

A. Geometric Distribution

Example 1.77: Expectation
Arithmetic Geometric Series
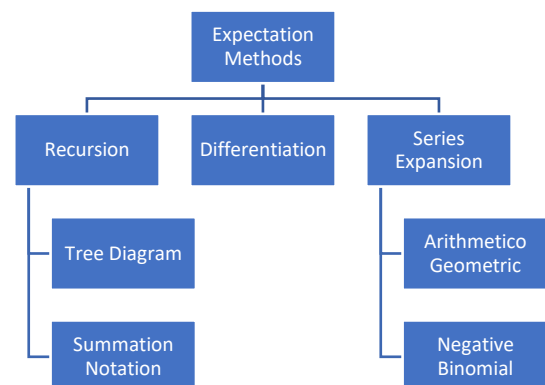Negative Binomial Series

Substitute the PMF of the geometric random variable in the definition of expectation

$$E(X) = \sum_{n=1}^\infty np(n) = \sum_{n=1}^\infty nq^{n-1}p$$

Add and subtract 1 from n:

$$= \sum_{n=1}^\infty (n - 1 + 1)q^{n-1}p$$

[Expectation Methods]
[Recursion] [Differentiation] [Series Expansion]
[Tree Diagram]
[Summation Notation]
[Arithmetico Geometric]
[Negative Binomial]

Breakup the summation into maroon and purple terms

$$= \sum_{n=1}^{\infty} (n-1)q^{n-1}p + \sum_{n=1}^{\infty} 1 \cdot q^{n-1}p = qE(X) + 1$$

The purple term is the sum of probabilities of a geometric random variable and is, hence, 1:

$$\sum_{n=1}^{\infty} 1 \cdot q^{n-1}p = \sum_{n=1}^{\infty} q^{n-1}p = 1$$

$$= \sum_{m=0}^{\infty} mq^m p = q \sum_{m=0}^{\infty} mq^{m-1}p = qE(X)$$

> Substitute $m = n - 1$
> Take q out of the summation
> The expression after the summation sign is the same expression as E(X)
Add back the purple and maroon terms

Solve this as an equation in E(X)

$$E(X) = qE(X) + 1$$
$$E(X)(1 - q) = 1$$
$$E(X) = \frac{1}{1-q} = \frac{1}{p}$$

Step II – Solve the equation

$pE(X) = 1$

$E(X) = \dfrac{1}{p}$

Collect like terms on the left-hand side
Take $E(X)$ common
Substitute $p = 1 - q$
Solve for $E(X)$

Example 1.78: Expectation of X²

Step I – Create an equation

$$\sum_{n=1}^{\infty} n^2 p(n) = \sum_{n=1}^{\infty} n^2 q^{n-1} p$$

Substitute the pmf of the geometric random variable in the definition of expectation of a function
Add and subtract 1 from n

$$= \sum_{n=1}^{\infty} (n-1+1)^2 q^{n-1} p$$

$$= \sum_{n=1}^{\infty} [(n-1)^2 + 2(n-1) + 1] q^{n-1} p$$

Expand $(n-1+1)^2$ using the formula for $(a+b)^2$ and letting $a = (n-1)^2$

$$= \sum_{n=1}^{\infty} (n-1)^2 q^{n-1} p + \sum_{n=1}^{\infty} 2(n-1) q^{n-1} p + \sum_{n=1}^{\infty} q^{n-1} p$$

Expand $(n-1+1)^2$ using the formula for $(a+b)^2$ and letting $a = (n-1)^2$

Expectation of X - Step II – Solve the equation

$E(X) = qE(X) + 1$	Solve this as an equation in E(X)
$E(X) - qE(X) = 1$	Collect like terms on the left-hand side
$E(X)(1-q) = 1$	Take $E(X)$ common
$pE(X) = 1$	Substitute $p = 1 - q$
$E(X) = \dfrac{1}{p}$	Solve for $E(X)$

B. Negative Binomial Distribution

Example 1.79: Expectation

$E(Y) = E\left(\sum_{i=1}^{r} X_i\right)$	Y is the sum of X independent, identical geometrically distributed random variables.
$= \sum_{i=1}^{r} E(X_i)$	Sum of expectations is the expectation of sums
$= \sum_{i=1}^{r} \dfrac{1}{p}$	Substitute the expectation of each X
$= \dfrac{r}{p}$	Carry out the summation

Example 1.80: Variance

Variance of Y	
$$\text{Var}(Y) = \text{Var}\left(\sum_{i=1}^{r} X_i\right)$$	Y is the sum of X independent, identical geometrically distributed random variables.
$$= \sum_{i=1}^{r} \text{Var}(X_i)$$	The variance of sums is the sum of the variances (We can only do this since the variables are independent)
$$= \sum_{i=1}^{r} \text{Var}\left(\frac{1-p}{p^2}\right)$$	Substitute the Variance of X
$$= \frac{r(1-p)}{p^2}$$	Carry out the summation

C. Poisson Distribution

1.81: Expectation

$$\text{Mean} = \text{Variance} = \lambda$$

The mean and variance of a binomial distribution are:
$$E(X) = np$$
$$\text{Var}(X) = np(1-p)$$

Substituting $\lambda = np$ in the above:
$$E(X) = \lambda$$
$$\text{Var}(Y) = \lambda(1-p)$$

Since p is small, $(1-p)$ is close to 1. Hence:
$$\text{Var}(Y) = \lambda$$

1.82: Expectation

Calculation of Expectation

Use the definition of expectation, and substitute the PDF:
$$\sum_{i=0}^{\infty} ip(i) = \sum_{i=0}^{\infty} ie^{-\lambda}\frac{\lambda^i}{i!}$$

Make the following three changes:
➢ The zeroth term is zero, and hence it vanishes.
➢ Move a single power of λ out of the summation sign.
➢ Cancel i and make use of $i! = i(i-1)!$

$$\lambda \sum_{i=1}^{\infty} e^{-\lambda}\frac{\lambda^{i-1}}{(i-1)!}$$

Substitute $j = i-1$,

$$\lambda \sum_{j=0}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!} = \lambda \sum_{i=0}^{\infty} p(j) = \lambda$$

Step	$E(X)$	Justification	$E(X^2)$

1	$\displaystyle\sum_{i=0}^{\infty} ip(i)$	Substitute the definition of expectation	$\displaystyle\sum_{i=0}^{\infty} i^2 p(i)$
2	$\displaystyle\sum_{i=0}^{\infty} ie^{-\lambda}\frac{\lambda^i}{i!}$	Substitute the definition of $p(i)$	$\displaystyle\sum_{i=0}^{\infty} i^2 e^{-\lambda}\frac{\lambda^i}{i!}$
3	$\displaystyle\sum_{i=1}^{\infty} ie^{-\lambda}\frac{\lambda^i}{i!}$	The zeroth term is zero, and hence it vanishes	$\displaystyle\sum_{i=1}^{\infty} i^2 e^{-\lambda}\frac{\lambda^i}{i!}$
4	$\displaystyle\lambda\sum_{i=1}^{\infty} ie^{-\lambda}\frac{\lambda^{i-1}}{i!}$	Move a single power of λ out of the summation sign. The power inside reduces	$\displaystyle\lambda\sum_{i=1}^{\infty} i^2 e^{-\lambda}\frac{\lambda^{i-1}}{i!}$
5	$\displaystyle\lambda\sum_{i=1}^{\infty} e^{-\lambda}\frac{\lambda^{i-1}}{(i-1)!}$	Cancel i and make use of $i! = i(i-1)!$	$\displaystyle\lambda\sum_{i=1}^{\infty} ie^{-\lambda}\frac{\lambda^{i-1}}{(i-1)!}$
6	$\displaystyle\lambda\sum_{j=0}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}$	Substitute $j = i - 1$	$\displaystyle\lambda\sum_{j=1}^{\infty}(j+1)e^{-\lambda}\frac{\lambda^j}{j!}$
7	$\displaystyle\lambda\sum_{i=0}^{\infty} p(j)$	Definition of $p(j)$	No Steps here. Read the next cell below
8	λ	Sum of probabilities is 1	
9		Writing the sum as two sums $$\sum_{j=1}^{\infty} je^{-\lambda}\frac{\lambda^j}{j!} = \lambda$$	$\displaystyle\lambda\left[\sum_{j=1}^{\infty} je^{-\lambda}\frac{\lambda^j}{j!} + \sum_{j=1}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}\right]$
10		$$\sum_{j=1}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!} = 1$$	$\lambda[\lambda + 1]$

D. Hypergeometric

Example 1.83: Expectation
Determine the expectation of the hypergeometric probability distribution using the definition of the PMF.

Let $X \sim HGeom(w, b, n)$ where

$$w = white\ balls$$
$$b = black\ balls$$
$$n = no.\ of\ balls\ drawn\ without\ replacement$$

Then:

$$E(X) = \sum_{k=0}^{n} k \cdot P(X = k)$$

Substitute the PMF of the hypergeometric probability distribution:

$$= \sum_{k=0}^{n} k\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$$

Substitute $k \cdot \binom{w}{k} = k \cdot \frac{w}{k}\binom{w-1}{k-1} = w\binom{n-1}{k-1}$. Drop the $k = 0$ from the index of summation since that term is

zero, and hence the summation is unchanged:

$$= \sum_{k=1}^{n} \frac{w \binom{w-1}{k-1}\binom{b}{n-k}}{\frac{w+b}{n}\binom{w+b-1}{n-1}}$$

Move $\frac{w}{\frac{w+b}{n}} = \frac{wn}{w+b}$ out of the summation sign (valid since the expressions are independent of k).

$$= \frac{wn}{w+b} \sum_{k=1}^{n} \frac{\binom{w-1}{k-1}\binom{b}{n-k}}{\binom{w+b-1}{n-1}}$$

Use a change of variable. Substitute $j = k - 1 \Rightarrow k = j + 1 \Rightarrow n - k = n - (j+1) = (n-1) - j$

$$= \frac{wn}{w+b} \sum_{j=0}^{n-1} \frac{\binom{w-1}{j}\binom{b}{(n-1)-j}}{\binom{w+b-1}{n-1}}$$

Note that $\sum_{j=0}^{n-1} \frac{\binom{w-1}{j}\binom{b}{(n-1)-j}}{\binom{w+b-1}{n-1}}$ follows $HGeom(w-1, b, n-1)$ and hence has sum 1. Substitute 1:

$$= \frac{wn}{w+b}(1) = \frac{wn}{w+b}$$

> ## Example 1.84
> Mean and Variance of hypergeometric distribution

1.5 Bernoulli Distribution

A. Bernoulli
The Bernoulli is more important as a building block then because of its direct usage.

1.85: Bernoulli Distribution

Define a trial, or an experiment (called *Bernoulli* experiment) to have *exactly* two outcomes
- ➢ success - with probability p
- ➢ failure - with probability q
    - ✓ Since the sum of probabilities must equal 1, therefore $p + q = 1$, and hence $q = 1 - p$

The meaning of this *mathematical* success and failure does not have to be the same as real life success and failure.
- ➢ Success can be defined as the death of a patient in a hospital, or an airplane engine failing in midair.

Definition
X is called a Bernoulli random variable, if:
- ➢ X = 1, when the experiment is successful
- ➢ X = 0, when the experiment is a failure

1.86: Bernoulli Distribution

Parameter, shape and notation

Parameter

A single parameter, p, describes the Bernoulli distribution completely.

Shape

If p = 0.5, then the Bernoulli distribution is symmetrical.

As p varies from 0.5, the distribution becomes increasingly skewed.

Notation

If X is a random variable that follows a Bernoulli distribution, we write

$$X \sim B(1, p)$$

We read the notation as "X follows a Binomial distribution with 1 trial and probability of success p".

The B represents a binomial distribution (discussed next), of which the Bernoulli distribution is a special case.

Properties

Probability Mass Function

The probability mass function of $X \sim B(1, p)$ is given by

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

Example 1.87: Expectation - Using the formula

Expectation and Variance		
Expectation of X	**Expectation of X²**	**Variance of X**
$E(X)$ $= \sum xp(x)$ $= 0 \cdot (1-p) + 1 \cdot p$ $= 0 + p$ $= p$	$E(X^2)$ $= \sum x^2 p(x)$ $= 0 \cdot (1-p) + 1^2 \cdot p$ $= 0 + p$ $= p$	$Var(X)$ $= E(X^2) - [E(X)]^2$ $= p - p^2$ $= p(1-p)$
Alternate Calculation of Variance		
$E[(X-\mu)^2] = \sum (x-\mu)p(x)$		Substituting the definition of expectation
$= (0 - p)^2 (1 - p) + (1 - p)^2 (p)$ $= p^2 (1 - p) + (1 - p)^2 (p)$		Substituting the values of the probability mass function
$= p(1 - p)(p + 1 - p)$ $= p(1-p)$		Taking $p(1-p)$ common

B. Bernoulli Distribution

1.88: Bernoulli Distribution

A random variable X follows a Bernoulli distribution if
- ➢ it has exactly two outcomes: success, and failure.
- ➢ Probability of success $= p$

- ➢ Traditionally, success is traditionally assigned a value of 1, and failure is assigned a value of 0.
- ➢ Success does not have to mean success in the common-sense understanding. For example, the death of a patient in a hospital can be termed a "success".

Example 1.89

I have an urn with 3 blue balls and some red balls. The total number of balls is 8. I draw a ball from the urn, and assign the random variable X as 1 if the ball is blue, and 0 otherwise.
- A. Determine the distribution that X follows.
- B. Write it in a probability distribution.
- C. Calculate the mean and the variance of X.

$$X \sim Bernouli$$
$$Success = P(Blue) = \frac{3}{8}$$
$$Failure = P(Red) = \frac{5}{8}$$

	Failure $x = 0$	Success $x = 1$	
$P(X = x)$	$\frac{5}{8}$	$\frac{3}{8}$	
$xp(x)$	0	$\frac{3}{8}$	$\mu = \frac{3}{8}$
$(x - \mu)$	$-\frac{3}{8}$	$\frac{5}{8}$	

| $(x - \mu)^2$ | $\dfrac{9}{64}$ | $\dfrac{25}{64}$ | |
| $(x - \mu)^2 p(x)$ | $\dfrac{45}{512}$ | $\dfrac{75}{512}$ | $\sigma^2 = \dfrac{120}{512} = \dfrac{15}{64}$ |

Example 1.90

X follows a Bernoulli distribution with probability of success p.
   A. Determine the probability of failure, and write the probability distribution of X as a table.
   B. Determine the mean, variance, and standard deviation of X.

Part A
Let probability of failure be q.
$$p + q = 1 \Rightarrow q = 1 - p$$

x	0	1
$P(X = x)$	$1 - p$	p
$x - \mu$	$-p$	$1 - p$

Part B
Mean
$$\sum_x xP(x) = (0)(1 - p) + (1)(p) = p$$

Variance
$$Variance = \sigma^2 = \sum_x (x - \mu)^2 P(x)$$
$$= (-p)^2(1 - p) + (1 - p)^2(p)$$
$$= p^2 - p^3 + (1 - 2p + p^2)(p)$$
$$= p^2 - p^3 + p - 2p^2 + p^3$$
$$= p^2 + p - 2p^2$$
$$= -p^2 + p$$
$$= p(1 - p)$$

Standard Deviation
$$Standard\ Deviation = \sigma = \sqrt{\sigma^2} = \sqrt{p(1 - p)}$$

1.91: Bernoulli Distribution: Summary

If X follows a Bernoulli distribution with parameter p, we write
$$X \sim Bernoulli(p)$$

	$Mean = \mu$	$Variance = \sigma^2$	$SD = \sigma$
X	p	$p(1 - p)$	$\sqrt{p(1 - p)}$

Example 1.92

In an urn with 8 balls, exactly 3 are blue. I draw a ball from the urn, and assign the random variable X as 1 if the ball is blue, and 0 otherwise.
   A. Determine the distribution that X follows.
   B. Write it in a probability distribution.
   C. Calculate the mean and the variance of X.

$$p = \frac{3}{8} \Rightarrow 1 - \frac{5}{8}$$
$$\mu = p = \frac{3}{8}$$

$$\sigma^2 = p(1-p) = \frac{3}{8}\left(\frac{5}{8}\right) = \frac{15}{64}$$

Example 1.93
A student attempts to solve a statistics question. His probability of success is known to be 0.4. If we consider a success as 1, and a failure as zero, what is the standard deviation of the outcome?

$$p = 0.4 \Rightarrow SD = \sqrt{0.4 \cdot 0.6} = \sqrt{0.24} = \sqrt{\frac{24}{100}} = \frac{2\sqrt{6}}{10}$$

1.94: Constant Property: Mean and Variance
$$E[aX] = aE[X]$$
$$Var[aX] = a^2 Var[X]$$
$$SD[aX] = |a|SD[X]$$

Example 1.95
A student attempts a question, with probability of success 0.4. Successful answers get 10 marks, and 0 marks on failure. Calculate the expected value and the standard deviation of his marks.

Let X be a random variable that has the outcome for the statistics question.
$$X \sim Bernoulli(0.4), \quad Success \Rightarrow X = 1, \quad Failure \Rightarrow X = 0$$

$$E[10X] = 10E[X] = 10 \times 0.4 = 4$$
$$SD[10X] = 10SD[X] = 10 \cdot \frac{2\sqrt{6}}{10} = 2\sqrt{6}$$

1.6 Binomial Distribution

A. Definition

- Repeat a Bernoulli trial n times (e.g. toss a coin n times)
    - ✓ The outcome for each trial is a Bernoulli random variable
- A random variable represents each outcome:
    - ✓ X_1 – The first outcome
    - ✓ X_2 – The second outcome
    - ✓ X_n – The nth outcome
- If $X_1, X_2 \ldots X_n$ are
    - ✓ n independent, distributed identically (*iid*) Bernoulli random variables
    - ✓ with common probability of success p for each random variable,
    - ✓ then the sum of the variables (say, Y) written as
$$Y = \sum_{k=1}^{n} X_k$$
    follows a binomial distribution with parameters n and p.
- Independent, identically distributed (written in short form as *iid*) is an important, high-frequency term
    - ✓ Independence – the trials are not co-related
    - ✓ Identically distributed – the probability of success does not change from trial to trial

- Given $X_1, X_2 \ldots X_n$ *iid* Bernoulli random variables with parameter p

$$Y = \sum_{k=1}^{n} X_k \sim B(n, p)$$

That is,

Y follows a Binomial distribution with parameters n and p

➢ Exactly two parameters completely describe a binomial random variable
  ✓ n - the number of trials
  ✓ p – the success of probability, which is assumed to be the same in each trial

1.96: Binomial Distribution
A random variable X follows a binomial distribution if it consists of n independent, identically distributed Bernoulli distributions.
  ➢ The random variable X takes the value of the number of successes in the n trials.

Example 1.97
I have a bowl with 10 jellybeans of which 4 are red, and 6 are green. I draw 3 jellybeans without replacement. Consider each drawing of a jellybean as a trial. Are the trials independent?

Trials are not independent since the probabilities will change after one jellybean is drawn.

1.98: Binomial Probabilities
We can also define a binomial random variable without reference to the Bernoulli.

Binomial random variable requires n independent, identically distributed trials.
Each trial has exactly two outcomes ($success$) and $failure$ with:
$$P(success) = p, \qquad P(failure) = 1 - p = q$$

Trials can refer to any events that occurs repeatedly. For example
  ➢ Tossing a coin
  ➢ Drawing a ball from an urn with replacement
  ➢ Raining on a particular day
  ➢ Being able to solve a problem

In the case of Binomial probabilities, it is assumed that:
  ➢ Trials are independent: Success or failure in one trial does not affect success or failure in other trials.
  ➢ Trials are identical: Each trial has same probability of success or failure.

B. Mean and Variance

Example 1.99
Determine the mean and the variance of a Binomial distribution with n trials and probability of success p.

Let
$$X \sim Binomial(n, p)$$
$$X = X_1 + X_1 + X_2 + \cdots + X_n$$

Mean
Note that X consists of n independent, identically distributed Bernoulli distributions:
$$E[X] = E[X_1 + X_2 + \cdots + X_n]$$

Using linearity of expectation:
$$= E[X_1] + E[X_2] + \cdots + E[X_n]$$

But note that each Bernoulli distribution has *success parameter = mean = p*:
$$= \underbrace{p + p + \cdots + p}_{n \ times} = np$$

Variance

Note that X consists of n independent, identically distributed Bernoulli distributions:
$$Var[X] = Var[X_1 + X_2 + \cdots + X_n]$$

Using the linearity of variance for independent random variables:
$$= Var[X_1] + Var[X_2] + \cdots + Var[X_n]$$

But note that each Bernoulli distribution has *variance = p(1 − p)*:
$$= \underbrace{p(1-p) + p(1-p) + \cdots + p(1-p)}_{n \ times} = np(1-p)$$

Standard Deviation

$$Standard \ Deviation = \sigma = \sqrt{\sigma^2} = \sqrt{np(1-p)}$$

Example 1.100
Expectation and Variance – Bogus Proof

Expectation $= E(Y) = E\left(\sum_{k=1}^{n} X_k\right) = E(nX) = nE(X) = np$

Variance $= Var(Y) = Var\left(\sum_{k=1}^{n} X_k\right) = Var(nX) = n^2 Var(X) = n^2 p(1-p)$

Finding the error

$E(Y) = E\left(\sum_{k=1}^{n} X_k\right) = \boldsymbol{E(nX)} = nE(X) = np$

$Var(Y) = Var\left(\sum_{k=1}^{n} X_k\right) = \boldsymbol{Var(nX)} = n^2 Var(X) = n^2 p(1-p)$

The X's are identically distributed, but not identical! Furthermore, they are independent.
Hence, they can be added together, but they are not the same value.

Note: Be careful when manipulating random variables, and expectations. Make sure each step is due to a valid property.

Example 1.101: Expectation
Using linearity of expectations

Expectation of Y $= \boldsymbol{E(Y)}$

$= E\left(\sum_{k=1}^{n} X_k\right)$ 　　　　　　　　　Substituting the definition

$$= \sum_{k=1}^{n} E(X_k) \qquad\qquad E(X+Y) = E(X) + E(Y)$$

$$= \sum_{k=1}^{n} p \qquad\qquad \text{Substituting the expectation of each } X_k$$

$$= np \qquad\qquad \text{Carrying out the summation}$$

Example 1.102: Expectation
Determine the expectation of the binomial probability distribution using the definition of the PMF.

Let $X \sim Bin(n, p)$. Then:

$$E(X) = \sum_{k=0}^{n} k \cdot P(X = k)$$

Substitute the PMF of the binomial probability distribution:

$$= \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k q^{n-k}$$

Substitute $k \cdot \binom{n}{k} = k \cdot \dfrac{n}{k} \binom{n-1}{k-1} = n \binom{n-1}{k-1}$:

$$= \sum_{k=0}^{n} n \binom{n-1}{k-1} p^k q^{n-k}$$

Move np out of the summation sign (valid since it is independent of k). Drop the $k = 0$ from the index of summation since that term is zero, and hence the summation is unchanged:

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k}$$

Use a change of variable. Substitute $j = k - 1 \Rightarrow k = j + 1 \Rightarrow n - k = n - (j+1) = (n-1) - j$

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{(n-1)-j}$$

Note that $\sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{(n-1)-j}$ follows $Bin(n-1, p)$ and hence has sum 1. Substitute 1:

$$= np(1) = np$$

Example 1.103: Variance

Variance of Y

$$\text{Var}(Y) = \text{Var}\left(\sum_{k=1}^{n} X_k \right) = \sum_{k=1}^{n} \text{Var}(X_k) = \sum_{k=1}^{n} p(1-p) = np(1-p)$$

Interchanging summation and variance (third step) requires the X's to be independent.
Not meeting the condition of independence would change the variance since there would be $n(n-1)$ covariance

terms.

1.104: Binomial Distribution: Summary

If X follows a Binomial distribution with n trials, and probability of success p,
$$X \sim Binomial(n, p)$$

	$Mean = \mu$	$Variance = \sigma^2$	$SD = \sigma$
X	np	$np(1-p)$	$\sqrt{np(1-p)}$

Example 1.105

Ralph tosses a weighted coin that lands on heads with probability 0.3. If tosses are independent of each other, calculate the expected value and the variance of the number of heads in 5 tosses.

Let the number of heads in 5 tosses be X.
$$X \sim Binomial(n = 5, p = 0.3)$$

$$E[X] = np = 5(0.3) = 1.5$$
$$Var[X] = np(1-p) = 1.5 \times 0.7 = 1.05$$

Example 1.106

Every day, I visit the gym with probability 0.1. The probability on each day is independent of the probability on other days. Calculate the expected value and the standard deviation of the number of days I will visit the gym in a week.

Let the number of days I visit the gym in a week be X.
$$X \sim Binomial(n = 7, p = 0.3)$$

$$E[X] = np = 7(0.1) = 0.7$$
$$SD[X] = \sqrt{np(1-p)} = \sqrt{7(0.1)(0.9)} = \sqrt{0.63}$$

C. Probability Mass Function

1.107: PMF on Binomial

The probability of x successes for a binomial distribution with n trials and probability of success p is:
$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$
Where

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

Use s to denote a success and f to denote a failure:
$$\underbrace{sss \dots s}_{x\ successes} \ , \underbrace{fff \dots f}_{n-x} \ failures$$

If you have n trials and you want x successes, and you use s the probability of getting exactly x successes followed by $n - x$ failures is:
$$p^x (1-p)^{n-x}$$

However, order does not matter, and hence, we need multiply by the number of ways that x successes can be

arranged among n trials:
$$\binom{n}{x}$$

Hence, the final probability is:
$$\binom{n}{x} p^x (1-p)^{n-x}$$

Example 1.108
I toss a coin four times that has $P(Heads) = \frac{3}{4}$. Derive the probability mass function of the number of heads.

Since I toss a coin four times, we can make each coin toss a trial. Then:
$$No. = n = 4$$

Let X be the number of successes in the four trials:
$$No.\ of\ Successes = No.\ of\ Heads = r \in \{0,1,2,3,4\}$$

Define a head to be a success. Then:
$$P(Success) = p = \frac{3}{4}$$

And since heads and tails are complementary, the probability of tails
$$= P(Failure) = 1 - p = \frac{1}{4}$$

Substitute $n = 4, p = \frac{3}{4}, 1 - p = \frac{1}{4}$ in $\binom{n}{r} p^r (1-p)^{n-r}$:
$$P(X = r) = \binom{4}{r}\left(\frac{3}{4}\right)^r \left(\frac{1}{4}\right)^{n-r}$$

D. Evaluating Probabilities

Example 1.109
With a fair coin, what is the probability of getting
    A.   two heads in seven tosses
    B.   three heads in five tosses
    C.   two heads in eight tosses
    D.   three heads in ten tosses

Let X be the number of heads in seven tosses.
$$X \sim Binomial\left(n, p = \frac{1}{2}\right)$$

Part A
The pdf for the above binomial is:
$$P(X = x) = \binom{7}{x}\left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{n-x} = \binom{7}{x}\left(\frac{1}{2}\right)^n = \binom{7}{x}\left(\frac{1}{2}\right)^7 = \binom{7}{x} \cdot \frac{1}{128}$$
$$P(X = 2) = \binom{7}{x} \cdot \frac{1}{128} = \frac{7!}{2!\,5!}\left(\frac{1}{128}\right) = 21\left(\frac{1}{128}\right) = \frac{21}{128}$$

Example 1.110

I visit the gym every day. At the gym, I am equally likely to run on the treadmill or do weight training when I visit the gym (which I do every day). What is the probability that I run on the treadmill exactly four days in the second week of February?

Example 1.111
A chess player won the last six out of ten games in his skill bracket. If this is representative of his future performance, and wins are independent of each other, calculate the probability of exactly 2 wins in his next five games.

Let the number of wins be the random variable X.
$$p = \frac{6}{10} = 0.6$$
Substitute $n = 5, p = 0.6$ in $\binom{n}{x} p^x (1-p)^{n-x}$
$$\binom{5}{2}(0.6)^2(0.4)^3 = \frac{5!}{2!\,3!} \cdot 0.36 \cdot 0.064 = 0.2304$$

E. Evaluating Probabilities: Addition Rule

1.112: Addition Rule
For mutually exclusive events A and B:
$$P(A \cup B) = P(A) + P(B)$$

(Calculator) Example 1.113
A chess player won the last six out of ten games in his skill bracket. If this is representative of his future performance, and wins are independent of each other, calculate the probability of at least eight wins in his next nine games.

$$P(X \geq 8) = P(X = 8) + P(X = 9) = \binom{9}{8}(0.6)^8(0.4)^1 + \binom{9}{9}(0.6)^9(0.4)^0 \approx 0.0705$$

Example 1.114
An electronic system contains three cooling components that operate independently. The probability of each component's failure is 0.05. The system will overheat if and only if at least two components fail. Calculate the probability that the system will overheat. (SoA/P)

Let the number of components that fail be X.
$$X \sim Binomial(3,0.95)$$

The system will fail if and only if at least two components fail. Hence, we need to consider two possibilities:
Case I: Two Components Fail
Substitute $n = 3, p = 0.95, x = 2$ in $\binom{n}{x} p^x (1-p)^{n-x}$
$$P(X = 2) = \binom{3}{2}\underbrace{(0.95)^1}_{1\ Success}\underbrace{(1-0.95)^{3-2}}_{2\ Failures} = 3 \times \left(\frac{19}{20}\right) \times \left(\frac{1}{20}\right)^2 = \frac{57}{8000}$$
Case II: Three Components Fail
Substitute $n = 3, p = 0.95, x = 3$ in $\binom{n}{x} p^x (1-p)^{n-x}$

$$P(X = 3) = \binom{3}{3} \underbrace{(0.95)^0}_{0\ Success} \underbrace{(0.05)^3}_{3\ Failures} = 1 \times \left(\frac{1}{20}\right)^3 = \frac{1}{8000}$$

$$\frac{57}{8000} + \frac{1}{8000} = \frac{58}{8000} = \frac{29}{4000}$$

(Calculator) Example 1.115

A company prices its hurricane insurance using the following assumptions. (i) In any calendar year, there can be at most one hurricane. (ii) In any calendar year, the probability of a hurricane is 0.05. (iii) The numbers of hurricanes in different calendar years are mutually independent. Using the company's assumptions, calculate the probability that there are fewer than 3 hurricanes in a 20-year period. (SoA/P)

$$n = 20, p = 0.05, 1 - p = 0.95$$

$$\underbrace{\binom{20}{0}(0.05)^0(0.95)^{20}}_{P(X=0)} + \underbrace{\binom{20}{1}(0.05)^1(0.95)^{19}}_{P(X=1)} + \underbrace{\binom{20}{2}(0.05)^2(0.95)^{18}}_{P(X=2)}$$

Simplify the binomial coefficients:

$$= (0.95)^{20} + 20(0.05)^1(0.95)^{19} + 190(0.05)^2(0.95)^{18}$$

Calculate:

$$0.9245$$

F. Evaluating Probabilities: Complementary Probability

1.116: Complementary Probability Rule: Revision

Since mutually exclusive and collectively exhaustive probabilities add up to 1, the sum of the probability of an event and its complement is also 1:
$$P(A) + P(A') = 1$$

1.117: Complementary Probability for Discrete Distributions
$$P(X \geq k) = 1 - P(X \leq k)$$

We can expand the above to write:
$$1 - [P(X = 0) + P(X = 1) + \cdots + P(X = k - 1)]$$

(Calculator) Example 1.118

A chess player won the last six out of ten games in his skill bracket. If this is representative of his future performance, and wins are independent of each other, calculate the probability of at least one win in one next seven games.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{7}{0}(0.6)^0(0.4)^7 \approx 0.9983$$

Example 1.119

A study is being conducted in which the health of two independent groups of ten policyholders is being monitored over a one-year period of time. Individual participants in the study drop out before the end of the study with probability 0.2 (independently of the other participants). Calculate the probability that at least nine participants complete the study in one of the two groups, but not in both groups? (SoA/P)

The number of participants who continue follows a binomial distribution with:
$$n = 10, \qquad p = 0.8, \qquad 1 - p = 0.27$$

In a single group,

$$P(X \geq 9) = \binom{10}{9}(0.8)^9(0.2)^1 + (0.8)^{10} = 0.376$$
$$P(X < 9) = 1 - 0.376 = 0.624$$

Across both groups, we need an XOR probability. That is, exactly one of the groups should meet the condition above, given by:

$$\underbrace{(0.376)}_{\substack{Group\ I \\ X \geq 9}}\underbrace{(0.624)}_{\substack{Group\ II \\ X < 9}} + \underbrace{(0.376)}_{\substack{Group\ II \\ X \geq 9}}\underbrace{(0.624)}_{\substack{Group\ I \\ X < 9}} = 0.469$$

1.120: Special Cases

$P(X = 0) = (1-p)^{n-k}$	(Multiplication Rule for independent events
$P(X = 1) = np(1-p)^{n-1}$	
$P(X = n - 1) = np^{n-1}(1-p)$	
$P(X = n) = p^n$	

Applications:
$P(X \leq k)$
$= 1 - P(X > k)$ Complement Rule
$= P(X = 0) + P(X = 1) + \ldots + P(X = k)$ Sum of Probabilities Rule

G. Converting to Binomial

1.121: Converting to Binomial

Example 1.122
I have an urn that has 33 red balls, 33 green balls, and 33 blue balls. I draw a ball (with replacement) from the urn nine times. What is the probability that I get exactly nine red balls? eight or more green balls? two or more blue balls? (Keep numbers with large exponents in exponent form for the answers to this question).

Nine Red Balls
This is not binomial because it has more than two outcomes. But, we can classify the events as:
$$P(Red\ Ball) = \frac{33}{99} = \frac{1}{3} \Rightarrow P(Not\ Red) = 1 - \frac{1}{3} = \frac{2}{3}$$

Now, the probability of getting nine red balls:
$$\binom{9}{9}\left(\frac{1}{3}\right)^9\left(\frac{2}{3}\right)^0 = \frac{1}{3^9}$$

Eight or More Green Balls
As above, we get:
$$P(Green\ Ball) = \frac{33}{99} = \frac{1}{3} \Rightarrow P(Not\ Green) = 1 - \frac{1}{3} = \frac{2}{3}$$
Let the number of green balls drawn be G.
$$P(G = 8) = \binom{9}{8}\left(\frac{1}{3}\right)^8\left(\frac{2}{3}\right)^1 = 9 \times \frac{1}{3^8} \times \frac{2}{3} = \frac{18}{3^9}$$
$$P(G = 9) = \frac{1}{3^9}$$
$$P(G \geq 8) = P(G = 8) + P(G = 9) = \frac{18}{3^9} + \frac{1}{3^9} = \frac{19}{3^9}$$

Two or More Blue Balls

We need to find:
$$P(B=2) + P(B=3) + \cdots + P(B=9)$$
One option is to stoically calculate the individual probabilities in the above, which is eight different probabilities.
$$P(B=0) = \binom{9}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^9 = \frac{512}{3^9}$$
$$P(B=1) = \binom{9}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^8 = 9 \times \frac{1}{3} \times \frac{2^8}{3^8} = \frac{256 \times 9}{3^9}$$

Then the expression we need to find is:
$$1 - P(B=0) - P(B=1) = 1 - \frac{256 \times 2}{3^9} - \frac{256 \times 9}{3^9} = \frac{3^9 - 256 \times 11}{3^9}$$

Example 1.123: Ratio of Probabilities

Phillip flips an unfair coin eight times. This coin is twice as likely to come up heads as tails. How many times as likely is Phillip to get exactly three heads than exactly two heads? (AOPS, Alcumus, Binomial Theorem, Counting and Probability)

$$\binom{8}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^5 : \binom{8}{2}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^6$$

$$\left(\frac{8 \times 7 \times 6}{6}\right)\left(\frac{2}{3}\right) : \left(\frac{8 \times 7}{2}\right)\left(\frac{1}{3}\right) = 2 : \frac{1}{2} = 4 : 1 \Rightarrow 4\ Times$$

(Calculator) Example 1.124: Back Calculations

A company establishes a fund of 120 from which it wants to pay an amount, C, to any of its 20 employees who achieve a high-performance level during the coming year. Each employee has a 2% chance of achieving a high-performance level during the coming year. The events of different employees achieving a high-performance level during the coming year are mutually independent. Calculate the maximum value of C for which the probability is less than 1% that the fund will be inadequate to cover all payments for high performance. (SoA/P)

The probability that the fund will be inadequate must be
$$< 1\%$$

By complementary probability, the probability that the fund is adequate must be
$$> 99\%$$

Consider this a binomial distribution with each employee getting a high performance level as a trial.
$$n = no.\ of\ trials = 20$$
$$p = probability\ of\ sucess = 2\% = 0.02$$

$$P(X=0) = \binom{20}{0}(0.02)^0(0.98)^{20} = 0.668$$
$$P(X=1) = \binom{20}{1}(0.02)^1(0.98)^{19} = 0.272$$
$$P(X=2) = \binom{20}{2}(0.02)^2(0.98)^{19} = 0.053$$
Note that:

$$P(X = 0) + P(X = 1) = 0.668 + 0.272 = 0.94 < 0.99$$
$$P(X = 0) + P(X = 1) + P(X = 2) = 0.668 + 0.272 + 0.053 = 0.993 > 0.99$$

Hence, if we consider upto two employees, we cover 99% of the cases.
Hence, the maximum value of C is:
$$\frac{120}{2} = 60$$

H. Geometrical Counting

Example 1.125[1]
(AMC 10B 2004/23)

Problem

Each face of a cube is painted either red or blue, each with probability 1/2. The color of each face is determined independently. What is the probability that the painted cube can be placed on a horizontal surface so that the four vertical faces are all the same color?

(A) $\frac{1}{4}$ (B) $\frac{5}{16}$ (C) $\frac{3}{8}$ (D) $\frac{7}{16}$ (E) $\frac{1}{2}$

Let X be the number of pairs of opposite faces that have the same color. X can take the values:
$$X \in \{0,1,2,3\}$$

Opposite faces have the same color with probability
$$\frac{1}{2}$$

Case I: $X = 0$ Or $X = 1$
If $X = 0$ OR $X = 1$, then four vertical faces will not be the same color.

Case II: $X = 3$
3 opposite pairs have the same color. This does not mean that all faces have the same color. Rather the possible cases are:
$$RRR, BBB, RRB, BBR$$
However, in all of these cases, we get four vertical faces that are the same color.

$$P(X = 3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8} = \frac{2}{16}$$
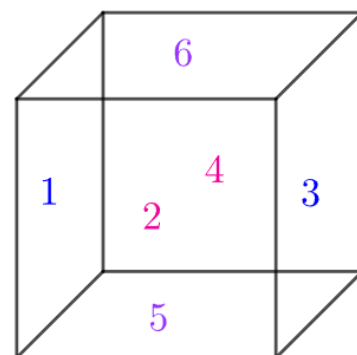
Case III: $X = 2$
If 2 opposite pairs have the same color, then the possible cases are:
$$RR, BB, RB, BR$$

Hence, the successful probability (if this case holds) is
$$\frac{2}{4} = \frac{1}{2}$$

---

[1] A simpler solution for this can be found in the Note on Probability.

Using $n = 3, r = 2, p = q = \frac{1}{2}$, the probability that:

$$P(X = 2) = \binom{n}{r} p^r q^{n-r} = \binom{3}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^1 = \frac{3}{8}$$

And the probability of being successful via this case is:

$$\frac{3}{8} \cdot \frac{1}{2} = \frac{3}{16}$$

The final probability is:

$$\frac{3}{16} + \frac{2}{16} = \frac{5}{16}$$

1.7 Uniform Distribution

A. Uniform Distribution

All outcomes have equal probability when
  ➢ Tossing a fair coin
  ➢ Throwing a fair die
  ➢ Checking the suit when drawing from a standard deck of cards
  ➢ Drawing a card from a standard deck of cards
An equiprobable sample space results in a uniform distribution
X is called a standard uniform(discrete) random variable, if
  ➢ X takes integer values from 1 to n
  ➢ all values of X have the same probability

Hence, X has the probability mass function

$$p(x) = P(X = x) = \frac{1}{n} \qquad\qquad x = 1, 2, 3, \dots n$$

1.126: Standard Uniform Distribution
The standard uniform distribution takes integer values from 1 to n, each with equal probability
$$\frac{1}{n}$$

It is the distribution encountered in common probability questions such as
  ➢ Tossing a fair coin
  ➢ Rolling a fair die
  ➢ Picking a card from a well shuffled pack of cards

Example 1.127
Find the mean of the standard uniform distribution.

The expectation of X is

$$E(X) = \sum_{x=1}^{n} xp(x) = \sum_{x=1}^{n} x\frac{1}{n} = \frac{1}{n}\sum_{x=1}^{n} x$$

Use the formula for the sum of the first n natural numbers and simplify:

$$= \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$$

Example 1.128
Calculate the variance of the standard uniform distribution

Expectation of X^2

$$= \sum_{x=1}^{n} x^2 p(x) = \sum_{x=1}^{n} x^2 \frac{1}{n} = \frac{1}{n} \sum_{x=1}^{n} x^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$
$$= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$$
$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$
$$= \frac{2(n+1)(2n+1) - 3(n+1)^2}{12}$$
$$= \frac{(n+1)[2(2n+1) - 3(n+1)]}{12}$$
$$= \frac{(n+1)[4n+2-3n-3]}{12}$$
$$= \frac{(n+1)[n-1]}{12}$$
$$\frac{n^2-1}{12}$$

1.129: Uniform Distribution
The uniform distribution takes on all valid values with equal probability.

> ➢ has equal probability of all values
> ➢ has a minimum value of a, and a maximum value of b
> ➢ has a range of
>> ✓ $b - a + 1$

Definition
Y is called a uniform(discrete) random variable, if
> ➢ Y takes integer values from a to b
> ➢ all values of Y have the same probability

Hence, Y has the probability mass function

$$p(y) = P(Y = y) = \frac{1}{b - a + 1} \qquad\qquad y = a, a+1, a+2, \dots b$$

Example 1.130
In a game, the player spins a wheel divided into three equal sectors (red, blue and green). If the spinner lands on a
> ➢ red region, he wins 5 points
> ➢ blue region, he wins 6 points
> ➢ green region, he wins 7 points

If the random variable X represents the number of points won, find, for X,
probability mass function
expectation
variance

Using geometric probability, the probability mass function is:
$$p(x) = P(X = x) = \frac{1}{3} \qquad x = 5, 6, 7$$
Tabulating the probability mass function:

X	5	6	7
$p(x)$	$\dfrac{1}{3}$	$\dfrac{1}{3}$	$\dfrac{1}{3}$

Using the definition
From the table:
$$E(X) = \sum_{x=1}^{n} xp(x) = 5 \cdot \left(\frac{1}{3}\right) + 6 \cdot \left(\frac{1}{3}\right) + 7 \cdot \left(\frac{1}{3}\right) = \frac{5+6+7}{3} = 6$$

Since all the probabilities are equal the formula for expectation reduces to the arithmetic average of the possible outcomes.

Using the properties of symmetry
We want to use the example to learn how to apply the properties we have learnt.
In a symmetric distribution:
$$\text{Mean} = \text{median} = \text{mode}$$
Therefore:
$$\text{Mean} = \text{middlemost value} = 6$$

Note that this distribution does not have a mode.

Using the formula
From the table:
$$E(X^2) = \sum_{x=1}^{n} x^2 p(x) = 5^2 \cdot \left(\frac{1}{3}\right) + 6^2 \cdot \left(\frac{1}{3}\right) + 7^2 \cdot \left(\frac{1}{3}\right) = \frac{5^2 + 6^2 + 7^2}{3} = \frac{25 + 36 + 49}{3} = \frac{110}{3} = 36.67$$

Variance $= E(X^2) - [E(X)]^2 = 36.67 - 6^2 = 36.67 - 36 = 0.67$

Example 1.131: Expectation - Using symmetry

In this case, using this method is the easiest. In a symmetric distribution:
$$\text{Mean} = \text{median} = \text{mode}$$
Therefore:
$$\text{Mean} = \text{median} = \frac{a+b}{2}$$

Note that this distribution does not have a mode.

Example 1.132: Expectation - Using the formula

$$E(X) = \sum_{x=a}^{b} xp(x) = \sum_{x=a}^{b} x \frac{1}{b-a+1}$$

$$= \frac{1}{b-a+1} \sum_{x=a}^{b} x$$

$$= \frac{1}{b-a+1} \left(\sum_{x=1}^{b} x - \sum_{x=1}^{a-1} x \right)$$

$$= \frac{1}{b-a+1} \left(\frac{b(b+1)}{2} - \frac{(a-1)a}{2} \right)$$

$$= \frac{1}{b-a+1} \left(\frac{b^2+b-a^2+a}{2} \right)$$

$$= \frac{1}{b-a+1} \left(\frac{b^2-a^2+b+a}{2} \right)$$

$$= \frac{1}{b-a+1} \left(\frac{(b+a)(b-a)+(b+a)}{2} \right)$$

$$= \frac{1}{b-a+1} \left(\frac{(b+a)(b-a+1)}{2} \right)$$

$$= \frac{a+b}{2}$$

Since all the probabilities are equal the formula for expectation reduces to the arithmetic average of all the values which is the same as the average (midpoint) of the maximum and minimum values

1.8 Other Distributions

A. Geometric Distribution
Like the binomial, the geometric random variable also builds on Bernoulli trials
  - ➢ The binomial random variable represents the *number of successes* in *n trials*.
  - ➢ The geometric random variable represents the *number of trials* till the *first success*.
If X represents
  - ➢ the number of
      - ✓ independent
      - ✓ identical (probability of success for each trial *p*, with $0 < p < 1$)
      - ✓ Bernoulli trials
  - ➢ required before the first success is achieved
then X is a geometric random variable

Example 1.133
Introduction to Geometric Random Variables: MIT 6.012

1.134: PMF of Geometric Distribution

$p(n)$
$= P\{X = n\}$
$=$(Probability of $n-1$ failures followed by one success)

$$= (\text{Probability of } n-1 \text{ failures})(\text{Probability of success})$$
$$= (1-p)^{n-1}p$$
$$= q^{n-1}p$$

1.135: Geometric Distribution

If you have independent, identical Bernoulli trials probability of success p, then the random variable Y that counts the number of trials till the first success has a geometric distribution.
$$Y \sim Geometric(p)$$

1.136: PDF of Geometric Distribution

If you have independent, identical Bernoulli trials, then the probability that the first success is obtained on the kth trial is:
$$P(Y=k) = (1-p)^{k-1}p$$

There are k trials. Probability of success is p, probability of failure is $1-p$.

The probability that the first $k-1$ trials will result in failure:
$$(1-p)^{k-1}$$

The probability that the kth trial results in success
$$= p$$

The probability of k trials, with the first $k-1$ resulting in failure, and the kth resulting in success is given by:
$$P(Y=k) = (1-p)^{k-1}p$$

1.137: Mean of Geometric Distribution

$$E[Y] = \frac{1}{p}$$

1.138: Variance of Geometric Distribution

$$Var[G] = \frac{1-p}{p^2}$$

Example 1.139

Anshuman is calling customers to get leads for a business. The probability of getting a lead on any individual call is 10%. Calls are independent of each other. What is the
  A. expected number of calls till Anshuman gets his first lead?
  B. standard deviation of the number of calls till Anshuman gets his first lead?
  C. probability that Anshuman get his first lead in the fourth call?

Let the number of calls that makes till his first lead be
$$Y \sim Geometric(p = 0.1)$$

Part A

$$E[Y] = \frac{1}{p} = \frac{1}{0.1} = 10$$

Part B

$$SD[Y] = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1-0.1}{0.01}} = \sqrt{\frac{0.9}{0.01}} = \sqrt{90} = 3\sqrt{10}$$

Part C

$$P(Y = 5) = (1 - 0.1)^3(0.1) = 0.9^3(0.1) = 0.729(0.1) = 0.0729$$

B. Poisson Distribution

Poisson is the approximation to the binomial distribution when
➢ n is large (number of events is large)
➢ p is small (probability of success for each individual event is small)

Define a parameter λ:
$$\lambda = np$$

Since:
$$n > 0$$
$$p > 0$$

therefore:
$$\lambda > 0$$

Why do the conditions for n and p apply? Consider what is the meaning of zero, or negative n and p.
The Poisson distribution can also be derived independently of the binomial.

1.140: Suitability

It is important to check suitability of a distribution before applying it to situation. Applying the wrong model will result in incorrect calculation of probabilities.

	n = Number of	P = Probability of
Typos in a book	Characters	A single typo
People in a community who survive to age 110	People in a community	The person being greater than 110
Wrong telephone numbers dialed in a day	Telephone numbers dialed in a day	A wrong telephone number being dialed
Packages of dog food sold in a particular store each day	Number of customers who walk into the store in a day	Packages of dog food sold in a day
Customers entering a post office on a given day to send a parcel to a specific small town	Customers entering the post office	Sending in a parcel to spec. small town
Vacancies occurring during a year in a government cadre	Total Vacancies | people in govt. cadre	Vacancy in a govt. cadre | retirement of an individual in a govt. cadre
α-particles discharged in a fixed period of time from some radioactive material	Atoms in radioactive material	α-particle being emitted from an individual atom

1.141: PMF

If X has the probability mass function below, then it follows a Poisson distribution:

$$p(i) = P\{X = i\} = \frac{\lambda^i}{i!\, e^\lambda} \qquad i = 0, 1, 2$$

Sum of probabilities for X is 1:

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^\lambda = 1$$

Use the sum of a *Maclaurin series* [a Taylor series expansion at $f(0)$] in the second-last step:

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} \ldots = \frac{1}{1} + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6} \ldots = e^\lambda$$

Takeaway:

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^\lambda$$

1.142: Poisson Probability Mass Function

In a Poisson probability distribution, the probability that are exactly x occurrences is given by:

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

Example 1.143

If 5% of the electric bulbs manufactured by a company are defective, what is the probability that in a sample of 100 bulbs, none is defective? (JMET 2011/76)

$$P(X = 0) = \frac{\lambda^0 e^{-5}}{0!} = e^{-5}$$

C. Negative Binomial Distribution

1.144: Background

The negative binomial distribution generalizes the geometric distribution
  ➢ The geometric random variable represents the *number of trials* till the *first success*.
  ➢ The negative binomial random variable represents the *number of trials* till r successes.

1.145: Definition

If Y represents the number of
  ➢ independent
  ➢ identical (probability of success for each trial = p, with $0 < p < 1$)
  ➢ Bernoulli trials
  required before the r successes are achieved
then Y is a negative binomial random variable

1.146: Probability Mass Function

Probability of r– 1 successes in n trials			
Number of failures	Number of ways to get successes	Probability of each success $= p$	Probability of each failure $= 1 - p$
= No. of trials – No. of successes $= (n-1) - (r-1)$ $= n - 1 - r + 1$ $= n - r$	**A** Order is not important Choose $(r-1)$ successes out of $(n-1)$ trials $\binom{n-1}{r-1}$	**B** Probability of r– 1 successes is p^{r-1}	**C** Probability of n – r failures $(1-p)^{n-r}$
	Multiply **A**, **B** and **C** to get the probability of r– 1 successes in n trials $\binom{n-1}{r-1} p^{r-1}(1-p)^{n-r}$		
Probability of rth success on the nth trial			
Since it is independent of the earlier trials, it is simply p			

$p(n) = P\{X = n\}$ =(Probability of r sucesses in n trials)

Break the above expression into the two probabilities calculated in the table:
= (Probability of $r - 1$ successes in n $- 1$ trials) $\times$ (Probability of rth sucess in n^{th} trial)
$$= \binom{n-1}{r-1} p^{r-1}(1-p)^{n-r} \cdot p$$

1.147: Properties

Y is the sum of X independent, identical geometrically distributed random variables.
The first X will give the number of trials required for the first success
The second X will give the number of trials required for the second success
.
.
.
The rth X will give the number of trials required for the rth success

D. Hypergeometric Distribution

1.148: Binomial as Limit of Hypergeometric

1.149: PMF of Hypergeometric Distribution

Given w white balls and b black balls in an urn, the probability of drawing k white balls when sampling n balls from the urn without replacement is:
$$P(X = k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$$

Where

> X is a random variable that follows a hypergeometric distribution.

E. Summary of Distributions

Bernoulli Distribution		
μ	σ	$P(X = k)$
p	$\sqrt{p(1-p)}$	$P(X = 1) = p$ $P(X = 0) = 1 - p$

- ➢ Takes value 1 with prob. of success p and value 0 with probability $1 - p$
- ➢ $\hat{p} = \dfrac{No.of\ successes}{No.of\ trials}$

Binomial	np	$np(1 - p)$	$\dfrac{n!}{k!\,(n-k)!} p^k (1 - p)^{n-k}$	If $i = 2$, the multinomial simplifies to the **Binomial**, which describes the probability of having exactly k successes in n iid Bernoulli trials. *Normal approximation to binomial:* ➢ Requires $n\hat{p} > 10$ AND $n(1 - \hat{p}) > 10$ ➢ Has same μ and σ as original binomial Breaks down on small intervals: To improve, decrease lower cutoff by 0.5, and increase upper cutoff by 0.5
Geometric	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$(1-p)^{n-1}p$	Describes the waiting time until a success for *iid* Bernoulli RVs. Cannot be approximated using the Normal distribution due to its skewness
Negative Binomial			$\dfrac{(n-1)!}{(k-1)!\,(n-k)!} p^k (1 - p)^{n-k}$	
Poisson	λ	$\sqrt{\lambda}$	$\dfrac{\lambda^k e^{-\lambda}}{k!}$	A random variable may follow a Poisson distribution if the event being considered is rare, the population is large, and the events occur independently of each other.
Hypergeometric				

1.9 Normal Distribution

A. Parametrization
The pdf of the Normal Distribution is mathematically complicated.

1.150: Probability Density Function
X is a normal random variable with parameters μ and σ² if its probability density function is given by

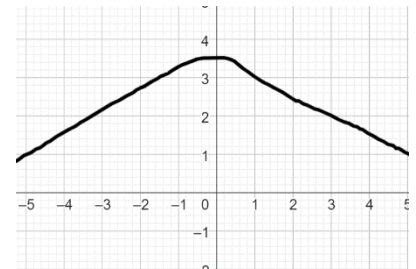$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

➢ The normal distribution, also known as the bell curve, is one of the most important distributions.
➢ The probability density function is *not used* in most (or all) of what follows.

1.151: Normal Distribution
A random variable X with mean μ and variance σ^2 that follows a Normal Distribution is written

$$\underset{\substack{Random \\ Variable}}{X} \sim \underset{\substack{N=Normal \\ Distribution}}{N(\mu,\sigma^2)}$$

➢ The normal distribution is completely described (parametrized) by its mean and variance.

Example 1.152
X is a random variable that follows a Normal Distribution with mean 3 and variance 4. Y is a random variable that follows a Normal Distribution with mean 3 and standard deviation 2 and mean 3. Compare X and Y.

$$X \sim N(3,4)$$
$$Y \sim N(3,2^2) = Y \sim N(3,4)$$

Since the above are equal, they represent the same normal distribution.

1.153: Range of Values for X
A random variable X that follows a normal distribution can take any value on the real number line. That is
$$X \ can \ be \ (-\infty, \infty)$$

➢ However, note that the major proportion of values lies close to the mean.

Example 1.154
X is a random variable that follows a Normal Distribution with mean 3 and variance 4. The probability that X takes a value greater than 2023 is:
   A. Zero
   B. Non-zero
   C. Cannot be determined
   D. Negative

$$Non-zero$$

1.155: Probability for a specific value
Since the normal distribution is continuous, the probability for any specific value is always zero.

The reason for this is technical and related to the way probability is defined for a continuous random variable.
$$P(a \le X \le b) := \int_a^b f(x) \, dx, \qquad f(x) \ is \ pdf \ of \ the \ random \ variable \ X$$

For a single value, $a = b$, and hence, we get:

$$P(a \leq X \leq a) = \int_{a}^{a} f(x)\, dx = F(a) - F(a) = 0$$

Example 1.156

Mark the correct option

X is a random variable that follows a Normal Distribution with mean 3 and variance 4. The probability that X takes the value 2023 is:
   A. Zero
   B. Non-zero
   C. Cannot be determined
   D. Negative

Zero

B. Symmetric, Skewness and Kurtosis

1.157: Symmetric Distribution

The normal distribution is symmetric.
   ➢ If you reflect it about its mean, the distribution remains unchanged.
   ➢ In a symmetric distribution $mean = median = mode$.

1.158: Mode of the Normal Distribution

The normal distribution is unimodal. That it has a single peak, which is its mode.

1.159: Skewness

Skewness refers to a graph which is non symmetric. Since the normal distribution is perfectly symmetric, it has
$$Skewness = 0$$

1.160: Kurtosis

Kurtosis is the measure of the tailedness of a distribution (which is how often outliers occur).
A normal distribution has
$$Kurtosis = 3$$

1.161: Excess Kurtosis

The normal distribution is used a benchmark, and hence kurtosis for other distributions is measured using the benchmark as:
$$Excess\ Kurtosis = Kurtosis - 3$$

1.162: Types of Distributions based on Kurtosis

Mesokurtic: Kurtosis=3
Platykurtic: Thin Tails = Kurtosis<3
Leptokurtic: Fat Tails = Kurtosis>3

1.163: Bell Shaped

The normal distribution has a typical bell shape. Changing the mean and the variance changes the shape, but still maintains the "bell shape".

   ➢ Increasing the mean moves the curve to the right. Decreasing the mean moves the curve to the left.

➢ Increasing the variance makes the graph less "clustered" around the mean. Decreasing the variance makes the graph more "clustered" around the mean.

1.164: Area under the curve
The sum of the areas under the curve of the normal probability distribution is 1.

Example 1.165
A. X is a random variable that follows a Normal Distribution with mean μ and variance σ^2. Determine the probability that X is less than μ.
B. X is a random variable that follows a Normal Distribution with mean μ and variance σ^2. Determine the probability that X is less than or equal to μ.

Part A
Since the distribution is symmetric, the areas to the right and the left of the mean are equal, and hence the probabilities are also equal:
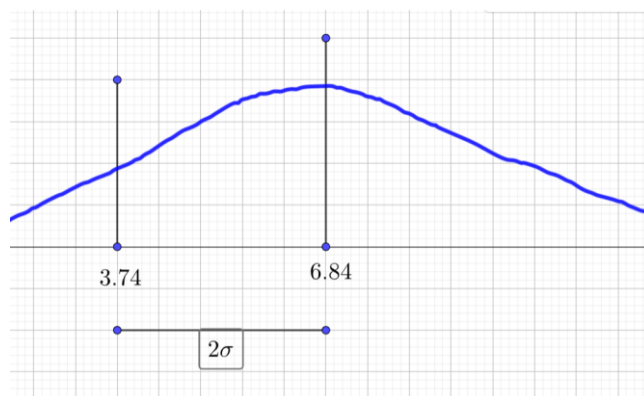$$P(X < \mu) = \frac{1}{2}$$

Part B
Since the probability of X taking a particular value in a continuous probability distribution is zero, the probability for $P(X \leq \mu)$ is the same as $P(X < \mu)$:
$$P(X \leq \mu) = P(X < \mu) = \frac{1}{2}$$

1.166: $68 - 95 - 99.7$ Rule

Example 1.167

Example 1.168

The probability that we want is:
$$P(\mu - \sigma < X < \mu + 2\sigma)$$

We can split the probability into two parts, and calculate each probability separately:
$$P(\mu - \sigma < X < \mu) + P(\mu < X < \mu + 2\sigma)$$

By symmetry,
$$P(\mu - \sigma < X < \mu) = \frac{68}{2} = 34\%$$
$$P(\mu < X < \mu + 2\sigma) = \frac{95}{2} = 47.5\%$$

And hence, the final answer is:
$$= 34 + 47.5 = 81.5\%$$

C. Standard Normal Distribution

Since the normal distribution is mathematically complicated to work with, we introduce the standard normal distribution. This distribution is a transformation of the normal distribution

1.169: Standard Normal Distribution

A random variable Z that follows a normal distribution with mean 0 and variance 1 is said to follow a standard normal distribution.

$$\underbrace{X}_{\substack{Random \\ Variable}} \sim \underbrace{Z(0,1)}_{\substack{Z=Normal \\ Distribution}}$$

> ➤ The standard normal distribution is a special case of the normal distribution

We convert a normal distribution with mean μ and standard deviation σ
$$X \sim N(\mu, \sigma)$$
$$(X - \mu) \sim N(0, \sigma)$$
$$\frac{(X - \mu)}{\sigma} \sim N(0, 1)$$

1.170: Area to the left

Areas to the left are mentioned in the Z table.

Example 1.171

Find the probability that $Z < 1.24$

Look up the row that says 1.2 for the first two digits, and the column that says "0.04" for the third digit

STANDARD NORMAL DISTRIBUTION: Table Values Represent A

Z	.00	.01	.02	.03	.04	.05
0.0	.50000	.50399	.50798	.51197	.51595	.51994
0.1	.53983	.54380	.54776	.55172	.55567	.55962
0.2	.57926	.58317	.58706	.59095	.59483	.59871
0.3	.61791	.62172	.62552	.62930	.63307	.63683
0.4	.65542	.65910	.66276	.66640	.67003	.67364
0.5	.69146	.69497	.69847	.70194	.70540	.70884
0.6	.72575	.72907	.73237	.73565	.73891	.74215
0.7	.75804	.76115	.76424	.76730	.77035	.77337
0.8	.78814	.79103	.79389	.79673	.79955	.80234
0.9	.81594	.81859	.82121	.82381	.82639	.82894
1.0	.84134	.84375	.84614	.84849	.85083	.85314
1.1	.86433	.86650	.86864	.87076	.87286	.87493
1.2	.88493	.88686	.88877	.89065	.89251	.89435
1.3	.90320	.90490	.90658	.90824	.90988	.91149
1.4	.91924	.92073	.92220	.92364	.92507	.92647
1.5	.93319	.93448	.93574	.93699	.93822	.93943
1.6	.94520	.94630	.94738	.94845	.94950	.95053

$$P(Z < 1.24) = 0.89251$$

Example 1.172
Find the probability that $Z < -0.75$

$$P(Z < -0.75) = 0.22663$$

1.173: Area to the right
To calculate areas to the right using the Z table, we use complementary probability.
$$P(A') = 1 - P(A)$$

Example 1.174
Find the probability that $Z > 0.84$

$$P(Z > 0.84)$$

Using complementary probability:
$$= 1 - P(Z < 0.84)$$

Using the $Z - Table$:
$$= 1 - 0.79955 = 0.20045$$

1.175: Area over an interval
To find area over an interval, subtract area to the left of the smaller value from the area to the left of the larger value.
$$P(a < Z < b) = P(Z < b) - P(Z < a)$$

Example 1.176
Find the probability that $-0.29 < Z < 1.37$

$$P(-0.29 < Z < 1.37)$$

Split into two different probabilities:
$$= P(Z < 1.37) - P(Z < -0.29)$$
$$= 0.91466 - 0.38591$$
$$= 0.52875$$

D. Transform to Z

1.177: Transformation from Normal to Standard Normal Distribution
If $X \sim N(\mu, \sigma^2)$, it can be converted into $Z \sim N(0,1)$ using the transformation:
$$Z = \frac{X - \mu}{\sigma}$$

➤ Subtracting μ is a horizontal shift
➤ Dividing by σ is a vertical scale.

Example 1.178
The mean weight of some apples is normally distributed with a mean of 12 ounces, and a standard deviation of 2.1 ounces. Calculate the Z value for an apple with a weight of 11.9 ounces.

$$X \sim N(\mu = 12, \sigma^2 = 2.1^2)$$

$$Z = \frac{X - \mu}{\sigma} = \frac{11.9 - 12}{2.1} = -\frac{0.1}{2.1} = -0.047$$

Example 1.179
X is a normal random variable with mean 304, and variance 64. Calculate the probability that X is greater than 290.

We want the probability that:
$$P(X > 290)$$

Convert to a standard normal random variable:
$$P\left(\frac{X - \mu}{\sigma} > \frac{290 - 304}{8}\right)$$
Simplify:
$$P(Z > -1.75)$$

Use complementary probability to rewrite:
$$1 - P(Z < -1.75)$$

Look up the probability in the Z table:
$$1 - 0.04006 = 0.95994$$

Example 1.180
Scores on the SAT follow a normal distribution with mean 505 and standard deviation 110. What percent of scores fall between 550 and 700.

We want the probability:

$$P(550 < X < 700)$$

Convert to Z:

$$P\left(\frac{550 - 505}{110} < \frac{X - \mu}{\sigma} < \frac{700 - 550}{110}\right)$$

Simplify:

$$P(0.41 < Z < 1.36)$$
$$= P(Z < 1.36) - P(Z < 0.41)$$
$$= 0.91309 - 0.65910$$
$$= 0.25399$$

E. Back Calculations

Example 1.181
Given $P(X < 5) = 0.05, \mu = 8$ find σ.

We know

$$P(X < 5) = 0.05$$

Convert to Z:

$$P\left(\frac{X - \mu}{\sigma} < \frac{5 - 8}{\sigma}\right) = 0.05$$
$$P\left(Z < \frac{-3}{\sigma}\right) = 0.05$$

Use the inverse normal to find the value of Z:

$$-\frac{3}{\sigma} = -1.64$$
$$\sigma = \frac{3}{1.64} = 1.83$$

F. Normal Random Variable

Notation
If X is a random variable that follows a normal distribution, we write
$$X \sim N(\mu, \sigma)$$

Summary Statistics
➢ Mean $= \mu$
➢ standard deviation $= \sigma$
➢ variance $= \sigma^2$

Background

The probability density function of a normal random variable is not mathematically tractable
➢ Closed-form solutions (a.k.a formulae) cannot be used
➢ Calculation of probabilities is done via tables
➢ Hence, we introduce the standard normal random variable

Substitution

If we make the substitution:

$$y = \frac{x - \mu}{\sigma}$$

then the probability density function simplifies to:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$$

Definition

Standard normal random variable is a normal distribution with
- ➤ mean $= \mu = 0$
- ➤ standard deviation $= \sigma = 1$
- ➤ variance $= \sigma^2 = 1$

We write

$$Z \sim N(0, 1)$$

Converting variables using change of scale and origin

Any normal random variable can be standardized using:

$$Z = \frac{X - \mu}{\sigma}$$

Calculating area under the curve

Convert the problem statement into a range for an applicable distribution. Then:
- ➤ Step I: Convert a Binomial / Poisson random variable into an equivalent normal random variable
  - ✓ Ensure the conditions for conversion are met
- ➤ Step II: Convert a normal random variable into a standard normal random variable
- ➤ Step III: Find the area under the curve for the standard normal random variable using the following cases
  - ✓ $Z < a$ - From a point till minus infinity
    - ▪ Same as $Z \leq a$ due to the properties of continuous distributions
  - ✓ $Z > a$ – From a point till positive infinity
    - ▪ Same as $Z \geq a$ due to the properties of continuous distributions
  - ✓ $a < Z < b$ – In a range

Skip steps which are not applicable.
For instance, if the problem discusses a normal random variable, start with Step II.

Example 1.182

The radius of a ball bearing is normally distributed with a mean of 5 mm, and a variance of 0.04 mm. Find the probability that the radius of a randomly selected ball bearing lies in the range 4.6 to 6 mm.

Let X be the random variable measuring the radius of the ball bearing.
Then:

$$P(4.6 < X < 6)$$

Convert X to Z using the transformation $Z = \frac{X - \mu}{\sigma}$

$$P\left(\frac{4.6 - 5}{0.2} < \frac{X - \mu}{\sigma} < \frac{6 - 5}{0.2}\right)$$

Simplify

$$P(-2 < Z < 5)$$

This can be broken up as
$$P(-2 < Z < 0) + P(0 < Z < 5)$$
Convert the first term to use positive numbers (using the symmetry of the normal distribution)
$$P(0 < Z < 2) + P(0 < Z < 5)$$
Convert the terms to use equivalent areas from infinity
$$[P(-\infty < Z < 2) - 0.5] + [P(-\infty < Z < 5) - 0.5]$$
Substitute the required probabilities
$$[0.9772 - 0.5] + [0.998 - 0.5]$$
Simplify
$$0.4772 + 0.4998$$

2. CONTINUOUS RVS
2.1 Probability Density Function

A. Definition
Rather than defining the continuous probability distribution, we will define the probability density function. This is the equivalent of a probability mass function for a discrete probability distribution.

A function f is called the probability density function of a random variable X if, for any interval [a, b]

The random variable X is said to follow a continuous probability distribution.

2.1: Continuous Random Variable
A continuous random variable can take any value on the real number line over its valid interval.

2.2: Probability Density Function (PDF)
The probability density function is the continuous version of a probability mass function.

For a discrete random variable, the probability is given by a
$$probability\ mass\ function$$

2.3: Probability as Area
For a continuous random variable, the probability is evaluated as the area. Hence, we must have:
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Since the sum of mutually exclusive and collectively exhaustive probabilities is 1. The discrete version of this property is:
$$\sum_a p(a) = 1$$

Example 2.4
A random variable X has the same probability for all values. It takes minimum value a and maximum b. Determine the probability density function for X.

Note: This is a uniform random variable.

Since the probability is the same, it is a constant probability:
$$f(x) = c, \quad for\ some\ constant\ c$$

Since the sum of the probabilities is 1:
$$\int_a^b c\,dx = 1$$

Integrate:
$$[cx]_a^b = 1$$

Evaluate:
$$c(b - a) = 1$$
$$c = \frac{1}{b - a}$$

$$PDF = f(x) = \begin{cases} \dfrac{1}{b - a}, & a < x < b \\ 0, & otherwise \end{cases}$$

2.5: Probability is nonnegative
Probabilities are nonnegative
$$f(x) \geq 0$$

If a function $f(x)$ is negative across some portion of its domain, it cannot be used as a probability function (unless it is modified).

B. Normalization

2.6: Normalizing Constant
If a function $f(x)$ is otherwise suitable to be made into a PDF, but the area under the curve is not 1, we can multiply by a *normalizing constant* to make it one.

Example 2.7
Calculate the normalizing constant to convert the function $f(x) = e^{-\lambda x}$, $x > 0$ into a PDF, and state the PDF. Also, the valid range for λ.

Note: This is the exponential random variable.

With the addition of the normalizing constant, the function becomes:
$$f(x) = ke^{-\lambda x}, \qquad x > 0$$

Since the sum of the probabilities must be 1:
$$\int_0^\infty ke^{-\lambda x}\, dx = 1$$

Integrate:
$$-k\left[\frac{e^{-\lambda x}}{\lambda}\right]_0^\infty = 1$$

The integral will converge when $\lambda > 0$ to:
$$\frac{k}{\lambda} = 1$$
$$k = \lambda$$

Since the integral only converges when $\lambda > 0$
$$Valid\ Range\ for\ \lambda: \lambda > 0$$

The PDF is:

$$PDF = f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$$

Example 2.8

$$f(x) = e^{-\lambda_1 x} + e^{-\lambda_2 x}, \qquad x, \lambda_1, \lambda_2 > 0$$

Calculate the normalizing constant to convert the function into a PDF, and state the PDF.

With the addition of the normalizing constant, the function becomes:

$$f(x) = ke^{-\lambda_1 x} + e^{-\lambda_2 x}, \qquad x > 0$$

Since the sum of the probabilities must be 1:

$$\int_0^\infty ke^{-\lambda_1 x}\, dx + \int_0^\infty ke^{-\lambda_2 x}\, dx = 1$$

Integrate:

$$-k\left[\frac{e^{-\lambda_1 x}}{\lambda}\right]_0^\infty - k\left[\frac{e^{-\lambda_2 x}}{\lambda}\right]_0^\infty = 1$$

This will evaluate to:

$$\frac{k}{\lambda_1} + \frac{k}{\lambda_2} = 1$$
$$\frac{k\lambda_2 + k\lambda_1}{\lambda_1 \lambda_2} = 1$$
$$k = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$

For $\lambda > 0$:

$$PDF = f(x) = \begin{cases} \dfrac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}\left(e^{-\lambda_1 x} + e^{-\lambda_2 x}\right), & x > 0 \\ 0, & otherwise \end{cases}$$

Example 2.9

Let the probability density function of a random variable X be:

$$f(x) = \begin{cases} x, & 0 \le x < \dfrac{1}{2} \\ c(2x - 1)^2, & \dfrac{1}{2} < x \le 1 \\ 0, otherwise \end{cases}$$

Then the value of c is equal to: (GATE 2016)

Since the area under the curve must be 1,

$$\int_0^{\frac{1}{2}} x\, dx + \int_{\frac{1}{2}}^1 c(2x - 1)^2\, dx = 1$$

C. Evaluation of Probabilities

2.10: Evaluating Probabilities

The probability for a continuous random variable with PDF $f(x)$ is given by:
$$P\{a \leq X \leq b\} = \int_a^b f(x)dx$$

Informally, the probability is given by the area under the curve of the probability density function.

Example 2.11
The probability function of a random variable X is given by:
$$f(x) = \begin{cases} \dfrac{1}{4}, & |x| \leq 1 \\ \dfrac{1}{4x^2}, & otherwise \end{cases}$$
Then $P\left(-\dfrac{1}{2} \leq x \leq 2\right)$ is (JAM-MS 2015)

By definition:
$$P\left(-\frac{1}{2} \leq x \leq 2\right) = \int_{-\frac{1}{2}}^{2} f(x)\, dx$$

Since the function is defined piece-wise, split the integral:
$$= \frac{1}{2}$$

Example 2.12
Let X be a continuous random variable with the probability density function:
$$f(x) = \begin{cases} \dfrac{x}{8}, & 0 < x < 2 \\ \dfrac{k}{8}, & 2 \leq x \leq 4 \\ \dfrac{6-x}{8}, & 4 < x < 6 \\ 0, & otherwise \end{cases}$$
Where k is a real constant. Then, $P(1 < X < 5)$ equals: (JAM-MS 2017)

Split the integral $\int_{-\infty}^{\infty} f(x)\, dx = 1$ piecewise:
$$\int_0^2 \frac{x}{8}\, dx + \int_2^4 \frac{k}{8}\, dx + \int_4^6 \frac{6-x}{8}\, dx = 1$$
$$\left[\frac{x^2}{16}\right]_0^2 + \frac{k}{8}[x]_0^2 + \int_4^6 \frac{6-x}{8}\, dx = 1$$
$$k = 2$$

$$P(1 < X < 5) = \int_1^2 \frac{x}{8}\, dx + \int_2^4 \frac{2}{8}\, dx + \int_4^5 \frac{6-x}{8}\, dx$$

$$\int_1^2 \frac{x}{8}\,dx = \frac{3}{16}$$

$$\int_2^4 \frac{2}{8}\,dx = \frac{1}{2} = \frac{8}{16}$$

$$\int_4^5 \frac{6-x}{8}\,dx = \frac{3}{16}$$

$$\frac{3}{16} + \frac{8}{2} + \frac{3}{16} = \frac{14}{16} = \frac{7}{8}$$

Example 2.13
Let X be a random variable with probability density function
$$f(x) = \begin{cases} 4x^k, & 0 < x < 1 \\ x - \dfrac{x^2}{2}, & 1 \le x < 2 \\ 0, & otherwise \end{cases}$$
Where k is a positive integer. Then $P\left(\frac{1}{2} < X < \frac{3}{2}\right)$ equals: (JAM-MS 2018)

Split the integral $\int_{-\infty}^{\infty} f(x)\,dx = 1$ piecewise:
$$\int_0^1 4x^k\,dx + \int_1^2 \left(x - \frac{x^2}{2}\right)dx = 1$$

Evaluate the integral:
$$4\left[\frac{x^{k+1}}{k+1}\right]_0^1 + \left[\frac{x^2}{2} - \frac{x^3}{6}\right]_1^2 = 1$$

Substitute $4\left[\frac{x^{k+1}}{k+1}\right]_0^1 = \frac{4}{k+1}$, $\left[\frac{x^2}{2} - \frac{x^3}{6}\right]_1^2 = \left(\frac{4}{2} - \frac{8}{6}\right) - \left(\frac{1}{2} - \frac{1}{6}\right) = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$:
$$\frac{4}{k+1} + \frac{1}{3} = 1 \Rightarrow \frac{4}{k+1} = \frac{2}{3} \Rightarrow 12 = 2k + 2 \Rightarrow k = 5$$

$$P\left(\frac{1}{2} < X < \frac{3}{2}\right) = \int_{\frac{1}{2}}^1 4x^5\,dx + \int_1^{\frac{3}{2}} \left(x - \frac{x^2}{2}\right)dx$$

$$= 4\left[\frac{x^6}{6}\right]_{\frac{1}{2}}^1 + \left[\frac{x^2}{2} - \frac{x^3}{6}\right]_1^{\frac{3}{2}}$$

$$= \frac{4}{6}\left(1 - \frac{1}{64}\right) + \left[\left(\frac{9}{8} - \frac{27}{48}\right) - \left(\frac{1}{2} - \frac{1}{6}\right)\right]$$

$$= \frac{85}{96}$$

Calculation of Probabilities – Discrete Probability Distribution			
Range of X	I	II	Probabilities in Column I and II are same

$p(x)$	$P\{X = x\}$		Substitute x into the probability mass function of the random variable
	$P\{X \leq x\}$	$P\{X < x + 1\}$	➤ Applicable when the random variable is restricted to the integers.
	$P\{X \geq x\}$	$P\{X > x - 1\}$	➤ Applicable (with modifications) when the random variable can take non-integer values.
$0, 1, 2, 3.....n$	$P\{X > 0\}$	$1 - P\{X = 0\}$	Based on the property of complement of a set
$0, 1, 2, 3.....n$	$P\{X < n\}$	$1 - P\{X = n\}$	

Calculation of Probabilities – Continuous Probability Distribution

$P\{X = x\}$	The probability of any single value is zero since		

$$\int_a^a f(x)dx = F(a) - F(a) = 0$$

$P\{X < x\}$	$P\{X \leq x\}$	$1 - P\{X \geq x\}$	$\int_{-\infty}^x f(x)dx = 1 - \int_x^\infty f(x)dx$
$P\{X > x\}$	$P\{X \geq x\}$	$1 - P\{X \leq x\}$	$\int_x^\infty f(x)dx = 1 - \int_{-\infty}^x f(x)dx$

2.2 Mean and Variance

A. Definition and Basics

2.14: Expectation of X (Mean)

The expectation of a continuous random variable X is:

$$Mean = E(X) = \int_{-\infty}^\infty x \cdot f(x)dx$$

Where

$$f(x) \text{ is the probability density function of } X$$

➤ $E(X)$ is notation for expectation of X, which is another name for the mean.
➤ The continuous version of the expectation replaces the summation (from the discrete version) with an integration to reflect the continuous nature of the variable.

Example 2.15

Find the mean of the continuous random variable X with pdf

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

$$E[X] = \int_{-\infty}^\infty x \cdot f(x)dx = \int_0^1 x \cdot 2x \, dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$$

2.16: Expectation of X^2

The expectation of X^2 for a continuous random variable X is:

$$E(X^2) = \int_{-\infty}^\infty x^2 \cdot f(x)dx$$

2.17: Variance of X

The variance of a continuous random variable X is:
$$Var(X) = E[X^2] - (E[X])^2$$

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

$Var(X)$ can also be written $V(X)$

Example 2.18

Find the variance of the continuous random variable X with pdf
$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

We calculated above that the expectation for this random variable is:
$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_0^1 x \cdot 2x \, dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$$

And the expectation of X^2 is:
$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_0^1 x^2 \cdot 2x \, dx = \left[\frac{2}{4}x^4\right]_0^1 = \frac{1}{2}$$

Finally, the variance is:
$$Var(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9}{18} - \frac{8}{18} = \frac{1}{18}$$

B. Continuous Uniform Distribution

Example 2.19: Uniform

Find the mean and variance of a random variable X with PDF:
$$f(x) = \begin{cases} \dfrac{1}{b - a}, & a < x < b \\ 0, & otherwise \end{cases}$$

Calculate $E(X) = Mean$

Substitute the definition of $f(x)$ in $E(X) = \int_a^b x \cdot f(x)dx$
$$= \int_a^b x \frac{1}{b - a} dx = \frac{1}{b - a}\left[\frac{x^2}{2}\right]_a^b$$

Substitute the limits of integration and simplify:
$$E(X) = \frac{1}{b - a}\left(\frac{b^2 - a^2}{2}\right) = \frac{(b + a)(b - a)}{2(b - a)} = \underline{\frac{a + b}{2}}$$
Result I:$E(X)$

Calculate $E(X^2)$

Substitute the definition of $f(x)$ in $E(X^2) = \int_a^b x^2 \cdot f(x)dx$

$$= \int_a^b x^2 \frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

Substitute the limits of integration and simplify:

$$E(X^2) = \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right) = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3} = \underbrace{\frac{b^2+ab+a^2}{3}}_{\textit{Result II}:E(X^2)}$$

Calculate the Variance

Substitute Results I and II into the formula for variance:

$$= \underbrace{\frac{b^2+ab+a^2}{3}}_{E[X^2]} - \underbrace{\left(\frac{a+b}{2}\right)^2}_{(E[X])^2}$$

Expand the second term:

$$= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4}$$

Take common denominators:

$$= \frac{4(b^2+ab+a^2)-3(b^2+2ab+a^2)}{12}$$

Simplify:

$$\frac{b^2-2ab+a^2}{12}$$

Factor:

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

C. Exponential Distribution

Example 2.20: Exponential Distribution
Find the mean and variance of a random variable X with PDF
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$$

Calculate $E(X) = Mean$

Substitute the definition of $f(x)$ in $E(X) = \int_a^b x \cdot f(x)dx$

$$= \int_0^\infty x \cdot \lambda e^{-\lambda x}\, dx$$

Use integration by parts. Let:

$$u = x, dv = \lambda e^{-\lambda x}\, dx$$
$$du = dx, v = -e^{-\lambda x}$$

Substituting the above in $\int u\, dv = uv - \int v\, du$

$$= [1 \cdot -e^{-\lambda x}]_0^\infty - \int_0^\infty -e^{-\lambda x}\, dx$$

The first term evaluates to zero, and the minus signs cancel out in the second term:

$$= \int_0^\infty e^{-\lambda x}\, dx$$

Integrating gives:

$$= \left[-\frac{e^{-\lambda x}}{\lambda}\right]_0^\infty = \left(-\frac{e^{-\lambda \cdot \infty}}{\lambda}\right) - \left(-\frac{e^{-\lambda \cdot 0}}{\lambda}\right) = 0 + \frac{1}{\underbrace{\lambda}} = \underbrace{\frac{1}{\lambda}}_{\textcolor{purple}{Result\ I:E(X)}}$$

Calculate $E(X^2)$

Substitute the definition of $f(x)$ in $E(X^2) = \int_a^b x^2 \cdot f(x)dx$

$$= \int_0^\infty x^2 \cdot \lambda e^{-\lambda x}\, dx$$

Use integration by parts. Let:

$$u = x^2, dv = \lambda e^{-\lambda x}\, dx$$
$$du = 2x\, dx, v = -e^{-\lambda x}$$

Substituting the above in $\int u\, dv = uv - \int v\, du$

$$= [x^2 \cdot e^{-\lambda x}]_0^\infty - \int_0^\infty -e^{-\lambda x} \cdot 2x\, dx$$

The first term evaluates to zero by L'Hospital's Rule $\lim\limits_{x\to\infty} \frac{x^2}{e^{\lambda x}} = \lim\limits_{x\to\infty} \frac{2x}{e^{\lambda x}} = \lim\limits_{x\to\infty} \frac{2}{e^{\lambda x}} = 0$, and the minus signs cancel out in the second term:

$$= \int_0^\infty 2x \cdot e^{-\lambda x}\, dx = \frac{2}{\lambda} \int_0^\infty x \cdot \lambda e^{-\lambda x}\, dx$$

And note that the integral is now exactly $E[X]$:

$$= \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \underbrace{\frac{2}{\lambda^2}}_{\textcolor{purple}{Result\ II:E(X^2)}}$$

Calculate the Variance

Substitute Results I and II into the formula for variance:

$$= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{2-1}{\lambda^2} = \frac{1}{\lambda^2}$$

Example 2.21: Exponential Distribution
Calculate the mean and variance for a random variable X with PDF below by first calculating $E(X^n)$:
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$$

Step I: $E(X^n)$

$$E(X^n) = \int_0^\infty x^n \cdot \lambda e^{-\lambda x}\, dx$$

Use integration by parts. Let:
$$u = x^n, dv = \lambda e^{-\lambda x}\, dx$$
$$du = nx^{n-1}\, dx, v = -e^{-\lambda x}$$

Substituting the above in $\int u\, dv = uv - \int v\, du$
$$= \left[-x^n \lambda e^{-\lambda x}\right]_0^\infty - \int_0^\infty -e^{-\lambda x} n x^{n-1} dx$$

The first term evaluates to zero. The second term simplifies to:
$$\frac{n}{\lambda} \int \lambda e^{-\lambda x} x^{n-1}\, dx$$

Note the term in the integral is the expectation of X^{n-1}. Hence
$$E(X^n) = \frac{n}{\lambda} E(X^{n-1})$$

Step II: $E(X)$ and $E(X^2)$
$$Mean = E(X) = \frac{1}{\lambda} E(X^0) = \frac{1}{\lambda} E(1) = \frac{1}{\lambda} \cdot 1 = \frac{1}{\lambda}$$
$$E(X^2) = \frac{2}{\lambda} E(X^1) = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2}$$

Step III: Variance
Substituting the values from above:
$$Var(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{2-1}{\lambda^2} = \frac{1}{\lambda^2}$$

Example 2.22: Double Exponential Distribution
Let X is a random variable with density
$$f(x) = \frac{1}{4} e^{-\frac{|x|}{2}}, \qquad -\infty < x < \infty$$
Then
$$E(X)$$

$$E[X] = \int_{-\infty}^\infty x \cdot f(x) dx$$

But note that $f(x)$
$$f(-x) = f(x) \Rightarrow f(x) \text{ is even}$$

Hence:
$$\int_{-\infty}^\infty x \cdot f(x) dx$$

Example 2.23: Double Exponential Distribution
Let X is a random variable with density
$$f(x) = \frac{1}{4} e^{-\frac{|x|}{2}}, \qquad -\infty < x < \infty$$

Then
$$E(|X|)$$

2.3 Median
A. Definition
Recall that for a discrete data set, the median is the data point that occurs when the data are arranged in ascending (or descending order). For example:
$$\left\{1,3,\frac{1}{2},-4,7\right\} \Rightarrow \left\{-4,\frac{1}{2},1,3,7\right\} \Rightarrow Middle\ Value = 1$$

For a continuous random variable, the median is defined similarly. The median in the middle of the probability density function (with middle calculated in terms of probability, and not value of the random variable).

2.24: Median
The median m for a continuous random variable is the value of the random variable such that the area to its left is exactly half. Hence:
$$\int_{-\infty}^{m} f(x)\ dx = \frac{1}{2}$$

The median for a continuous random variable X is the value x such that
$$P(X < x) = \frac{1}{2}$$

Substituting $Median = m$ gives:
$$P(X < m) = \frac{1}{2}$$

Using the definition of probability gives the desired result:
$$\int_{-\infty}^{m} f(x)\ dx = \frac{1}{2}$$

2.25: Area to the left and right
$$P(X > m) = P(X < m) = \frac{1}{2}$$

Using complementary probability:
$$P(X > m) = 1 - P(X < m) = 1 - \frac{1}{2} = \frac{1}{2}$$

Example 2.26
Ranking
A continuous random variable median m. Rank the following in ascending order:
- A. $P(X < m)$
- B. $P(X \leq m)$
- C. $P(X > m)$
- D. $P(X \geq m)$

Aziz Manva (azizmanva@gmail.com)

$$A = B = C = D$$

B. Symmetry

2.27: Symmetry
A geometrical figure is symmetrical across a line if the two halves that the line divides it into are equal when "folded" over each other.

2.28: Symmetrical Distribution
A symmetrical distribution is a distribution where a horizontal line of symmetry exists.
For a symmetrical distribution:
$$Median = Value\ at\ Line\ of\ Symmetry$$

➤ This is because the line of symmetry divides the distribution into two parts of equal area.
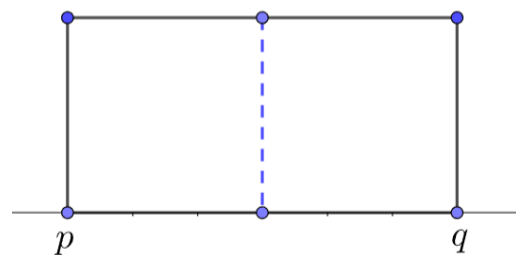➤ Symmetry is useful because you can do not need to integrate.

Example 2.29: Uniform Distribution
If a, b, p, q are real numbers, determine the median of the random variable X with the probability density function:
$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & p < x < q \\ 0, & otherwise \end{cases}$$

By symmetry:
$$Median = \frac{p+q}{2}$$

2.30: Even Function
A function is even if and only if it is symmetrical about the y axis. Algebraically,
$$f(x) = f(-x)$$

2.31: Algebraic Symmetry
A function $f(x)$ is symmetrical about the vertical line $x = a$ if
$$f(x + a) = f(-x + a)$$

An even function (which is symmetrical about $y = 0$) satisfies the functional equation:
$$f(x) = f(-x)$$

Translate the above function a units to the left, giving the desired result:
$$f(x) = f(-x)$$

Example 2.32
Let X be a random variable such that $E|X| < \infty$, and (IIT JAM-MS 2008)
$$P\left(X \geq \frac{1}{2} + x\right) = P\left(X \leq \frac{1}{2} - x\right)$$
Then:

A. $E(X) = \frac{1}{2}$ and $Median(X) = \frac{1}{2}$
B. $E(X) = \frac{1}{2}$ and $Median(X) > \frac{1}{2}$
C. $E(X) < \frac{1}{2}$ and $Median(X) = \frac{1}{2}$
D. $E(X) < \frac{1}{2}$ and $Median(X) > \frac{1}{2}$

The PDF is symmetrical about $x = \frac{1}{2}$ since:

$$P\left(X \geq \frac{1}{2} + x\right) = P\left(X \leq \frac{1}{2} - x\right)$$

For a symmetrical distribution symmetrical about $x = a$

$$Mean = Median = Mode = a$$

Therefore:

$$Mean = Median = \frac{1}{2}$$

C. Further Examples

Example 2.33: Exponential Distribution
The random variables X and Y have probability density functions:
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \lambda > 0 \\ 0, & otherwise \end{cases}, \qquad f_Y(y) = \begin{cases} me^{-my}, & y > 0 \\ 0, & otherwise \end{cases}$$

Given that m is the median of X, show that:
A. $m > 0$
B. λ is the median for Y.

Step I: Determine the median for X
$$\int_0^m \lambda e^{-\lambda x} \, dx = \left[-e^{-\lambda x}\right]_0^m = -e^{-\lambda m} - (-1) = 1 - e^{-\lambda m}$$

Equate this to $\frac{1}{2}$:

$$1 - e^{-\lambda m} = \frac{1}{2} \Rightarrow e^{-\lambda m} = \frac{1}{2}$$

Take the natural log both sides and solve for m:

$$-\lambda m = \ln\left(\frac{1}{2}\right) \Rightarrow m = \frac{-\ln 2}{-\lambda} = \frac{\ln 2}{\lambda}$$

Step II: Determine the median for Y
Note that Y has the same form as X. Hence:
$$Median_Y = \frac{\ln 2}{m} = \frac{\ln 2}{\frac{\ln 2}{\lambda}} = \ln 2 \times \frac{\lambda}{\ln 2} = \lambda$$

Example 2.34: Uniform Distribution
For $a, b, p, q \in \mathbb{R}$, the random variable X has the probability density function:

$$f_X(x) = \begin{cases} \dfrac{a}{b}, & p < x < q \\ 0, & otherwise \end{cases}$$

Determine the value of the median of in terms of a, b, q but not p.

Step I: Relate p and q:
The area under the curve for a PDF must sum to 1. Hence:

$$\int_p^q \frac{a}{b}\,dx = \frac{a}{b}[x]_p^q = \frac{a}{b}(q - p) = 1$$

Solve for p:

$$q - p = \frac{b}{a} \Rightarrow \underbrace{p = q - \frac{b}{a}}_{Equation\ I}$$

Step II: Find the Median

$$\int_p^m \frac{a}{b}\,dx = \frac{a}{b}[x]_p^m = \frac{a}{b}(m - p) = \frac{1}{2}$$

Hence:

$$m - p = \frac{b}{2a}$$

Solve for m and substitute $Equation\ I$:

$$m = p + \frac{b}{2a} = q - \frac{b}{a} + \frac{b}{2a} = q - \frac{2b}{2a} + \frac{b}{2a} = q - \frac{b}{2a}$$

Alternatively, the median of a constant (uniform) PDF is the average of its upper and lower limit:

$$Median = \frac{p + q}{2} = \frac{q - \dfrac{b}{a} + q}{2} = q - \frac{b}{2a}$$

2.4 Special Distributions
A. Uniform Distribution

2.35: Uniform Distribution
A random variable X with a uniform distribution, minimum value a, and maximum value b, has probability density function:

$$f(x) = \begin{cases} \dfrac{1}{b - a} & if\ a \le x \le b \\ 0\ otherwise \end{cases}$$

➢ The uniform distribution has equal probability for all equal width intervals

2.36: Cumulative Probability Density Function
A random variable X with a uniform distribution, minimum value a, and maximum value b, has cumulative probability density function:

$$f(x) = \begin{cases} \dfrac{x-a}{b-a} & if \ a \le x \le b \\ 0 \ otherwise \end{cases}$$

➤ The uniform distribution has equal probability for all equal width intervals

Example 2.37
Find $P(4 < X < 5)$ if X~U(4,10)

$$f(5) - f(4) = \frac{5}{10-4} - \frac{4}{10-4} = \frac{5}{6} - \frac{4}{6} = \frac{1}{6}$$

2.38: Cumulative Probability Density Function
A random variable X with a uniform distribution, minimum value a, and maximum value b, has cumulative probability density function:

$$f(x) = \begin{cases} \dfrac{x}{b-a} & if \ a \le x \le b \\ 0 \ otherwise \end{cases}$$

The uniform distribution has equal probability for all equal width intervals

Example 2.39

You are going to meet your friend. He will meet you anytime between 1 to 4 pm (with equal probability for any specific time).
What is the probability that
A. you meet him between 1 to 4 pm?
B. you meet him between 1 to 2 pm?
C. you meet him at 2 pm?a
Part A
1

Part B
1/3

Part C
0

B. Exponential Probability Distribution

2.40: PDF of the Exponential
A continuous random variable follows an exponential distribution when it has a probability density function
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

2.41: CDF of the Exponential
$$F(x) = 1 - e^{-\lambda x}$$

The CDF of X is given by:

$$F(x) = P(X < x) = \int f(x)\, dx = \int \lambda e^{-\lambda x}\, dx = -e^{-\lambda x} + C$$

To determine the value of C, note that

$$F(\infty) = -e^{-\infty} + C = 1$$
$$0 + C = 1$$
$$C = 1$$

Hence,

$$F(x) = 1 - e^{-\lambda x}$$

2.42: Cumulative Distribution Function

$F(a)$
Substitute the definition of the cumulative distribution function
$= P(X \le a)$
Substitute the definition of probability for a continuous random variable
$$= \int_0^a f(x)dx$$
Substitute the probability density function of an exponential random variable
$$= \int_0^a \lambda e^{-\lambda x}dx$$
Carrying out the integration:
$$= -e^{-\lambda x}\Big|_0^a$$
Substitute the limits of integration
$$= -e^{-\lambda a} - \left(-e^{-\lambda \cdot 0}\right)$$
$$= -e^{-\lambda a} - \left(-e^0\right)$$
$$= -e^{-\lambda a} - (-1)$$
$$= 1 - e^{-\lambda a}$$

2.43: Memory Less Property

We can derive the conditional distribution of a random variable, based on the occurrence of certain events.
➢ The expected lifetime of a light bulb that follows an exponential distribution is *s*.
➢ If the bulb has been used for *t* hours without failure,
➢ then what is the expected lifetime of the bulb
➢ Surprisingly, the incremental expected lifetime of the bulb (given that it has already been used for t hours without failure) is still *s*.
We can state this technically as
$P(X > s + t \mid X > t) = P(x > s)$

2.44: Connection to Poisson
If the number of occurrences of a particular event in a given time interval follows a Poisson distribution, then the time between events follows an exponential distribution

2.5 Further Topics