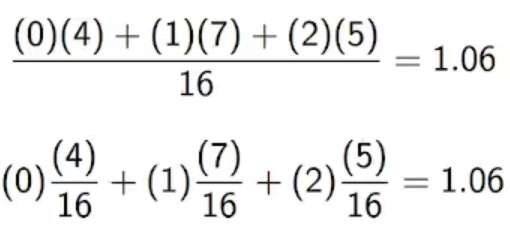
Mean of Random Variable

If 2 coins are tossed 16 times and X is the # of heads that occur per toss, then the values of X are 0, 1, 2.

Suppose that experiment yields no heads, 1 head, and 2 heads a total of 4, 7, and 5 times respectively.

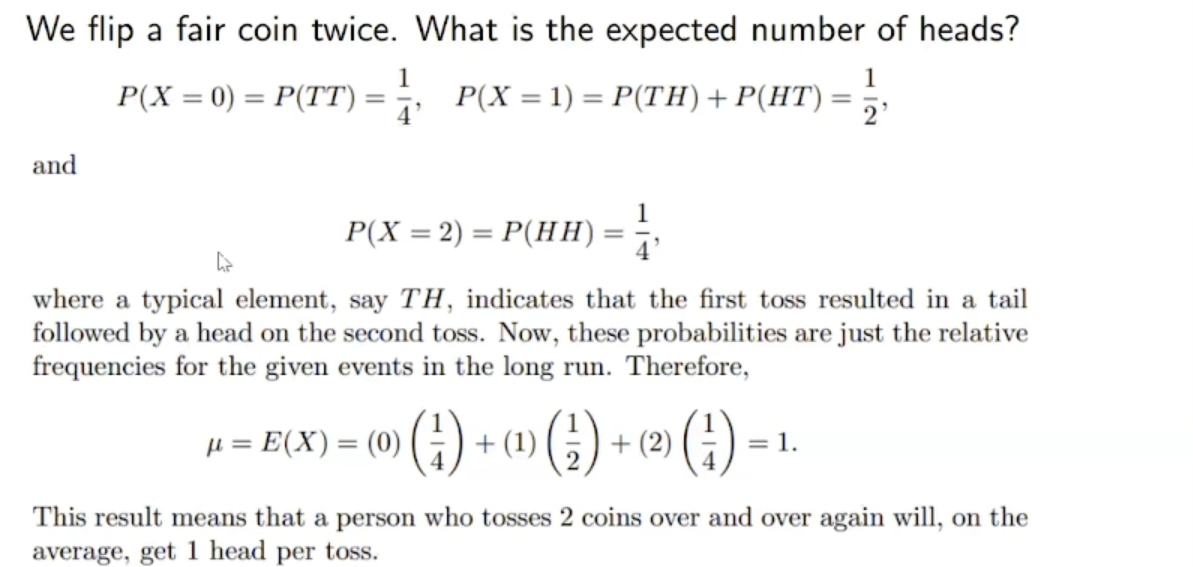
The average number of heads per toss of the 2 coins is then:



Average of head per toss is 1.06

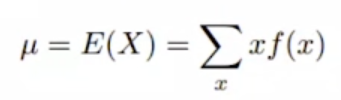
Mathematical Expectation

Let us start with an example:

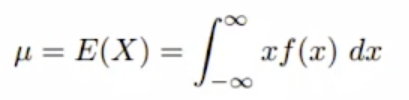
* We flip a fair coin twice. What is the expected number of heads?
* 
* 1 is expected value. It is kind of average of distribution of values that we expect to see # of heads.

Definition: Expected Value

Let X be a random variable with probability distribution f(x). The mean, or expected value, of X is:



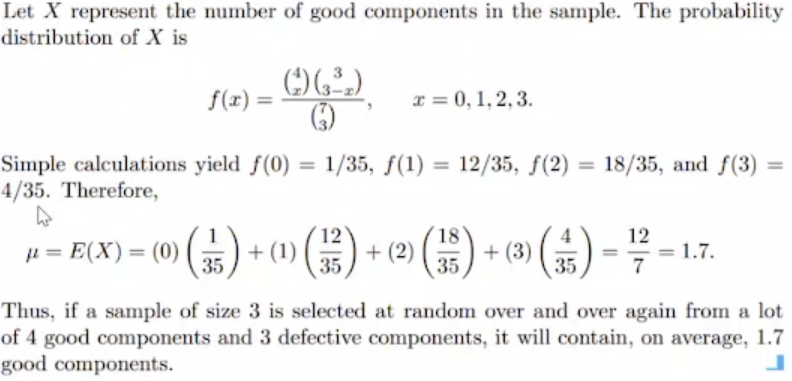
if X is discrete, and



if x is continuous.

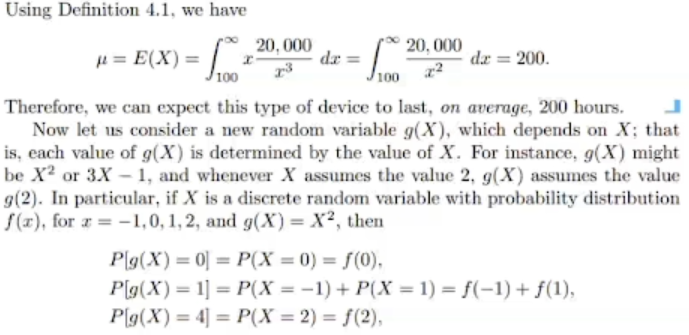
mean is special kind of weighted (possible values probability function can get) average.

Example (Discrete Case):

* A lot containing 7 components is sampled by a quality inspector;
  + the lot contains 4 good components and 3 defective components.
* A sample of 3 is taken by the inspector.
* Find the expected value of the number of good components in this sample.
* 
* f(0) + f(1) + f(2) + f(3) = 1
* Little more than half of the components is good so expected value of getting good out of 3 of them is little more than half (1.5).

Example (Continuous Case):

* Let X be the random variable that denotes the life in hours of a certain electronic device.
* The probability density function is
  + A picture containing graphical user interface

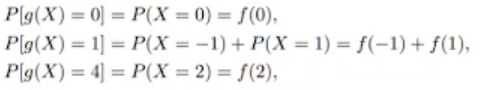
    Description automatically generated
  + After 100 hours, probability of your machine to broken gets higher.
* Find the expected life of this type of device.
* 

Expected value is based on X. For each sum, we find probability of X equals that value. Instead of X, we can define probability of variable by some function of X.

Now let us consider a new random variable g(X), which depends on X; that is, each value of g(X) is determined by the value of X.

For instance, g(X) might be X2 or 3X - 1, and whenever X assumes the value 2, g(X) assumes the value g(2).

In particular, if X is a discrete random variable with probability distribution f(x), for x = -1, 0, 1, 2, and g(X) = X2, then



and so the probability distribution of g(X) may be written

Graphical user interface, text, application, chat or text message

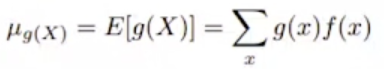
Description automatically generated

By the definition of the expected value of a random variable, we obtain

Text

Description automatically generated

Theorem 4.1:

* Let X be a random variable with probability distribution f(x). The expected value of the random variable g(X) is
  + 
* if X is discrete, and
  + A picture containing text, watch, gauge

    Description automatically generated
* if X is continuous.

g(X) 🡪 Function of random variable.

Instead of defining expected value based on just 1 random variable, we can define based on a function.

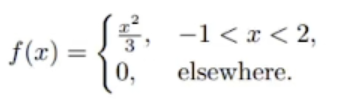
Example (Discrete Case):

* Suppose that the # of cars X that pass through a car wash between 4:00 pm and 5:00 pm on any sunny Friday has the following probability distribution:
* Text

  Description automatically generated with low confidence
* Let g(X) = 2X – 1 represent the amount of money, in dollars, paid to the attendant by the manager.
* Find the attendant’s expected earnings for this particular time.
* Text, letter

  Description automatically generated

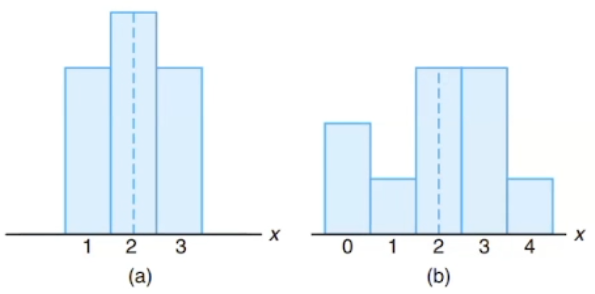
Example (Continuous Case):

* Let X be a random variable with density function
  + 
* Find the expected value of g(X) = 4X + 3
* A picture containing text, watch, clock, device

  Description automatically generated

Variance and Covariance

Figure 4.1, we have the histograms of 2 discrete probability distributions that have the same mean, =2, but differ considerably in variability, or the dispersion of their observations about the mean.



*Figure 4.1: Distributions with equal means and unequal dispersions.*

*Discrete distribution functions’ histograms*

With each distribution, mean is 2.

How the other values of the distribution function, not just the mean, vary 🡪 variance

Figuring out general behaviour of data is hard. Just knowing the mean is not enough. We will add more tools to come up with more knowledge of the variables and how they behave around mean.

First step is defining the probability distribution function.

We also need to find the mean.

Then we define standard deviation which is based on the variation of the values around the mean.

Variance

The most important measure of variability of a random variable X is obtained by applying Theorem 4.1 with g(X) = (X – )2.

X – 🡪 gives you the difference between the random variable and mean. So that can tell us how random variable values are far or close to mean (). Square makes negative things positive so even if the expected value is negative, we are still interested in distance. We don’t care if they are right or left of the mean.

The quantity is referred to as the variance of the random variable X or the variance of the probability distribution of X and is denoted by Var(X) or the symbol 2, or simply by when it is clear to which random variable we refer.

Definition:

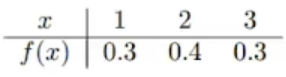
* Let X be a random variable with probability distribution f(x) and mean . The variance of X is:
* A picture containing text, watch

  Description automatically generated
* The positive square root of the variance, , is called the standard deviation of X.

In summary:

* The deviation of an observation from its mean:
  + x -
* Variance: A function of the expected value of X
  + g(X) = (X - )2
* Variance = = E[(X - )2] 🡪 E is expected value
* Standard Deviation: Positive square root of the variance
  + How much you are deviating from the mean, average of random values.

Example (Discrete Case):

* Let the random variable X represent the number of automobiles that are used for official business purposes on any given workday.
* The probability distribution for company A (Figure 4.1 – a) is:
  + 
* and that for company B (Figure 4.1 – b) is:
  + A picture containing diagram

    Description automatically generated
* Show that the variance of the probability distribution for company B is greater than that for company A.
* Text, letter

  Description automatically generated

Random variables are around the mean with the max distance of 0.6



# of automobiles that we expect in a given day to be used is 2 for company A

* When you see 0.6 and 1.6, you can say “Means are same but for company A random variable values are around the mean, for B values are scattered around.”.

Theorem 4.2:

* The variance of a random variable X is
  + = E(X2) - 2

Proof:

* Text, letter

  Description automatically generated

Example (Discrete Case):

* Let the random variable X represent the # of defective parts for a machine when 3 parts are sampled from a production line and tested.
* The following is the probability distribution of X.
* A picture containing graphical user interface

  Description automatically generated
* Using Theorem 4.2, calculate .
* Text

  Description automatically generated with medium confidence

Example (Continuous Case):

* The weekly demand for a drinking-water product, in thousands of liters, from a local chain of efficiency stores is a continuous random variable X having the probability density
* A picture containing text, watch, clock, gauge

  Description automatically generated
* Find the mean and variance of X
* Text

  Description automatically generated
* E(X2) = 🡪 don’t take square of f(x) !!!
* Expected value is not probability, don’t expect it to be less than 1 or sth. It depends on events we are interested in.