# 3. Random Variable and Univariate Distributions

[3. Random Variable and Univariate Distributions](#header-n0)  
 [3.1 Random Variables](#header-n3)  
 [3.2 Cumulative Distribution Function](#header-n42)  
 [3.3 Discrete Random Variable](#header-n81)  
 [3.4 Continuous Random Variables](#header-n129)  
 [3.5 Functions of Random Variable](#header-n205)  
 [3.5.1 Discrete Case](#header-n207)  
 [3.5.2 Continuous Case](#header-n254)  
 [Method 1: The CDF approach](#header-n255)  
 [Method 2: The transformation approach](#header-n371)  
 [3.6 Mathematical Expectations](#header-n382)  
 [3.7 Moments](#header-n401)  
 [3.7.1 Mean](#header-n410)  
 [3.7.2 Variance](#header-n436)  
 [3.7.3 Skewness](#header-n477)  
 [3.7.4 Kurtosis](#header-n482)  
 [3.8 Quantile](#header-n494)  
 [3.9 Moment Generating Function](#header-n536)  
 [3.9.1 Discrete Distributions](#header-n561)  
 [3.9.2 Continuous Distributions](#header-n576)  
 [3.9.3 Uniqueness of MGF](#header-n590)  
 [3.10 Characteristic Function](#header-n653)

## 3.1 Random Variables

**Definition (Random Variable)**: A random variable is a -measurable mapping from the sample space to such that to each there exists a corresponding unique real number (i.e. point function).

**Remarks**: For each outcome is a real number. The collection of all possible values that the random variable can take, also called the range of constitutes a new sample space, denoted as .

**Example**: Toss a coin three times.

The associated -algebra is Define as "number of heads in $ s " $, then

For any set where is a -field associated with we can define a probability function in terms of the original probability function as

* so-called the induced probability function, is indeed a probability function.
* Require additional condition to ensure that

**Definition (Measurable Function)**: A function is -measurable if for every real number , the set

* This is a regulation condition to guarantee that exists.
* Equivalently, is -measurable if for every Borel set .
* Let and be -measurable functions and .
  + Then and are -measurable;
  + is -measurable if for all
* Let be a sequence of -measurable functions. Then
  + and are -measurable;
  + if exists then is -measurable.
* Let be a -measurable function and be a Borel-measurable real valued function, then is also -measurable.

## 3.2 Cumulative Distribution Function

**Definition (Cumulative Distribution Function)**: The cumulative distribution function of a random variable is defined as

**Example**: Toss a coin three times. Define as "number of heads in Then,

and therefore,

**Example**: The experiment consists of shooting once at a circular target of radius Assume that it is certain that the target will be hit, and the probability of hitting a particular section is the ratio of the area of to the area of . Define random variable as the distance between the hitting point and the center.

* Clearly, So for and for
* For

Properties of CDF:

1. is non-decreasing, i.e. for any
2. is right-continuous, i.e. for all

**Remark**:

* If a function satisfies properties (1),(2) and (3), then there is a random variable such that
* In some textbooks, the CDF is defined as . Then is left-continuous under this definition.

Properties of CDF:

* for
* Suppose and are two CDF's, then for is also a

**Definition (Identical Distribution)**: Two random variables and are identically distributed if for every Borel set ,

**Theorem**: Two random variables and are identically distributed if and only if for all

## 3.3 Discrete Random Variable

**Definition (Discrete Random Variable)**: If a random variable can only take a countable number of values, then is called a discrete random variable (DRV).

**Definition (Probability Mass Function)**: The probability mass function of a DRV is defined as

**Definition (Support of DRV)**: The collection of the points at which a DRV has a positive probability is called the support of , denoted as

Properties of PMF:

* for all

**Example (Discrete Uniform Distribution)**: A DRV follows uniform distribution if its PMF

The CDF of uniform distributed DRV is

**Example (Bernoulli Distribution)**: A DRV follows a Bernoulli(p) distribution if its PMF

**Example (Binomial Distribution)**: A DRV follows a binomial distribution if its PMF

Remark: Toss a coin times independently. Each time the head has probability The total number of heads follows distribution.

**Example (Poisson Distribution)**: A DRV follows a Poisson distribution if its PMF

**Remarks**:

* Support of Poisson distribution is an infinite countable set.
* In a Poisson process with intensity the total number of occurrences over follows a Poisson distribution.
* **Poisson distribution can be used to describe the number of jumps in financial markets in a certain period**.

**Theorem (Poisson Limit Theorem)**: Denote as the probability that the event occurs in a random experiment and it is related to the number of total experiments . is the number that occurs in experiments. If as then we have

**Example (Geometric Distribution)**: The geometric distribution is the probability distribution of the number of Bernoulli trials required to obtain the first success. The PMF of a geometric distributed DRV is

**Remarks**:

* The geometric distribution is the simplest of **the waiting time distributions**.
* The geometric distribution has the so-called "memoryless" property in the sense that for integers we have

## 3.4 Continuous Random Variables

**Definition (Continuous Random Variable)**: A random variable is called continuous (CRV) if its cumulative distribution function is continuous for all

**Definition (Absolute Continuity & Probability Density Function)**: The cumulative distribution function of a random variable is called **absolutely continuous** with respect to Lebesgue measure if there exists a Borel-measurable function such that

The function is called a probability density function (PDF) of .

* An absolutely continuous CDF is continuous; For some continuous CDF, absolute continuity may not hold.
* For those 's where is differentiable, .

**Theorem (Properties of PDF)**: A function is a PDF of a CRV if and only if

1. for almost everywhere and

Remark:

* For any nonnegative function with finite integral, i.e. is a , where is called the normalizing constant.
* For a given continuous random variable, CDF is unique but is not.
* The probability density functions of a continuous random variable can be different on a set of measure The value of **the PDF can be changed arbitrarily on a sequence of countable points without altering the distribution** of For example, represent the same distribution.

**Definition (Support of CRV)**: The support of a CRV with PDF is defined as

**Example (Continuous Uniform Distribution)**: A CRV follows a uniform distribution on if its PDF

**Example (Normal Distribution)**: A normally distributed random variable has the PDF

where and .

|  |  |
| --- | --- |
| PDF of Normal Distributions | CDF of Normal Distributions |
|  |  |

**Example (Log-normal Distribution)**: follows a log-normal distribution if its PDF

If log-normal then .

**Example (Exponential Distribution)**: CRV follows Exponential distribution if its PDF

where .

**Remark**:

* The PDF can be written in expression with
* Exponential distribution is popular in modeling **duration** between financial events or economic events because of its "memoryless" property:

**Example (Gamma Distribution)**: A CRV follows a Gamma distribution if its PDF

where and is the gamma function.

* Two parameters: shape and scale
* Gamma Exponential
* Gamma

**Example (Beta Distribution)**: A CRV follows a Beta distribution if its PDF

where and is the beta function.

* The support of beta distribution is [0,1]

|  |  |
| --- | --- |
| PDF of gamma distribution | PDF of beta distribution |
|  |  |

## 3.5 Functions of Random Variable

Suppose is a Borel-measurable function (i.e. the preimage of any Borel set under is a Borel set), then is also a random variable. What is the distribution function of the new random variable ?

### 3.5.1 Discrete Case

If is a discrete random variable with PMF then the PMF of can by obtained by using

where is the set of all possible values of in the sample space of such that .

Moreover, if function be strictly **monotonic**

for y in support.

**Example**: Suppose random variable has the distribution

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|  | 0.1 | 0.2 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 |

For function , the distribution is

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 4 | 9 | 16 |
|  | 0.1 | 0.1+0.2 | 0.1+0.1 | 0.2 | 0.2 |

**Example (Binomial transformation)**: a CRV has a binomial distribution and its PMF is

Consider the random variable ,

### 3.5.2 Continuous Case

#### Method 1: The CDF approach

The basic idea is first to find the CDF of and then its by differentiation,

1. Identify the possible values of (i.e. the support of ). For this purpose, it is useful to plot the function
2. Find
3. Differentiate the CDF with respect to :

If function be strictly **monotonic**, let

Monotone: one-to-one mapping

* Increasing
* Decreasing

Transformation (Increasing)

Transformation (decreasing)

**Theorem**: Let have cdf let and let and be defined as above.

1. If is an increasing function on for .
2. If is a decreasing function on and is a continuous random variable, for

**Example**: Suppose a CRV has a PDF

Find the PDF of the following :

1) with

If since Support the support of is given by So for

then . Therefore,

Similarly, if the support of is So for ,

and then . It follows that

2)

Observe that for in Support then if and if . For

By differentiation, Therefore, we have

3)

Since the Support we have Support Thus, for ,

By differentiation, . It follows that

**Example**: Random variable follows double exponential (or Laplace) distribution if

where Find the PDF of the following

1) has PDF

Since Support and therefore for For

By differentiation,

for Note follows exponential distribution.

2) has PDF

Again, Support so for For

By differentiation,

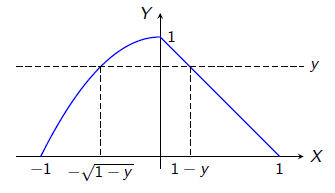
for . And for

A Weibull distribution is given by

follows Weibull distribution with

**Example**: Suppose has PDF and

Find the PDF of .



Note that Support so for

By differentiation,

for and for or .

**Example (Probability Integral Transform)**: Suppose has a continuous distribution which is strictly monotonically increasing. Find the PDF of .

The support of is the unit interval Because is continuous and strictly monotonically increasing, its inverse function exists and is also strictly increasing. Thus, for ,

If follows that the PDF

The application of probability integral transformation

* Let and is the CDF of some continuous random variable . If is strictly increasing, then has distribution (Why?)
* It can be used to simulate random samples of any CRV : inverse transform sampling.
  + Generate a realization from uniform distribution
  + Solve for from the equation .
  + The value is a realization of with the specified distribution .
* When is a DRV, the probability integral transform is no longer uniformly distributed. (How to generate random numbers from a discrete probability distribution via uniform distribution?)
* The result that provides a basis for goodness-of-fit tests of distributional models. To **check whether a probability model is correctly specified**, one can first compute the probability integral transform and then check if follows the distribution using a sample where .
* This is the basic idea behind the popular **Kolmogorov-Smirnov test** for a hypothesized distribution model.

#### Method 2: The transformation approach

**Theorem (Univariate Transformation)**: Let be a CRV with PDF and let function be strictly **monotonic** and differentiable over the support of . Then the PDF of the random variable is

for any in the support of where is the unique number in the support of such that .

**Theorem**: Suppose are disjoint regions and . Suppose for all and for each is strictly monotonic and differentiable on region Then the PDF of is given by

for all in the support of .

**Example (Square transformation)**: Suppose is a continuous random variable. For the CDF of is . Because is continuous, we can drop the equality from the left endpoint and obtain

The PDF of can now be obtained from the CDF by differentiation:

## 3.6 Mathematical Expectations

**Definition (Expectation)**: Suppose is a random variable with PMF or PDF . Then the expectation of a measurable function is defined as

where is the support of .

**Remarks**:

* can be considered as the **weighted average** of
* If we say does not exist.
* If is a constant, then
* The expectation is a linear operator, namely,
* is also a random variable. Let the PMF or PDF of Y be then we can also compute by

## 3.7 Moments

**Definition -th moment**: The -th **moment** of a random variable is defined as

Similarly, the -th **central moment** of a random variable is defined as

Relationship between uncentered moments and centered moments:

and

### 3.7.1 Mean

**Definition (Mean)**: The mean of a random variable is defined as

where is the support of .

**Remarks**:

* The mean is also the expectation of (i.e. ), or the first moment of .
* is a measure of central tendency for the distribution of .
* The mean exists if and only if for or for

**Example (Binomial Mean)**: lf has a binomial distribution, its pmf is given by

where is a positive integer, , and for every fixed pair and the pmf sums to 1. Calculate the .

Since the term vanishes. Let and Subbing and into the last sum (and using the fact that the limits and correspond to

**Example (Cauchy distribution)**: A random variable follows Cauchy distribution if its PDF

where is the location parameter and is the scale parameter.

Suppose follows Cauchy (0,1) distribution, then

so **mean of Cauchy distribution does not exist**.

**Theorem**: Suppose exists. Then

### 3.7.2 Variance

**Definition (Variance & Standard Deviation)**: The variance of random variable is defined as

where is the support of . The quantity is called the standard deviation of .

**Remarks**:

* is a measure of the degree of spread of a distribution around its mean.
* In economics, it is interpreted as a measure of **uncertainty** or risk. It is often called a measure of **volatility** of .
* If then with probability 1 and there is no variation in . This is so-called **degenerate distribution**.

**Theorem**

**Remark**: is called the second central moment and is called the second moment.

**Theorem**: If then

**Example (Portfolio selection)**: Assume that the investor likes higher return but lower risk. That is, his utility function is such that (the more expected return, the better and (the smaller risk, the better . An example of is

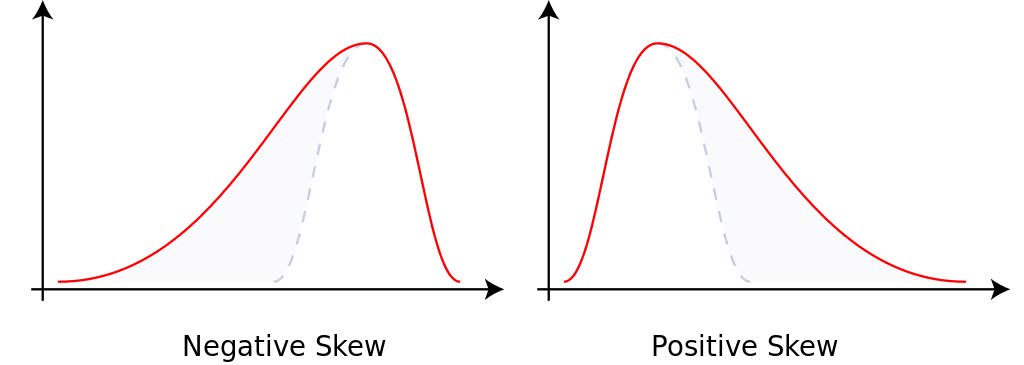
Assume that the investor has totally I dollars to be split between stock and saving What is the portfolio that maximizes the utility function?

* The rate of return on stocks is a random variable with mean and variance .
* The rate of return on saving is constant which can be considered as a random variable with mean and variance Usually .
* The return of a portfolio is then
* It is maximized when
* If (i.e. the investor does not care risk), is maximized at .
* If then is maximized at

### 3.7.3 Skewness

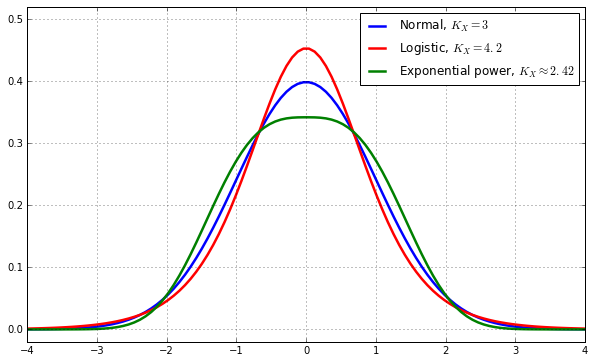
**Definition (Skewness)**: The third central moment is a measure of "skewness" (or asymmetry) of the distribution of . Skewness is defined as

The skewness has been used to measure financial crashes. Negative (or positive) skewness indicates a higher (or lower) probability of experiencing large losses than large gains.



### 3.7.4 Kurtosis

**Definition (Kurtosis)**: The fourth central moment is a measure of how heavy the tail of a distribution is. Kurtosis is defined as



Remarks:

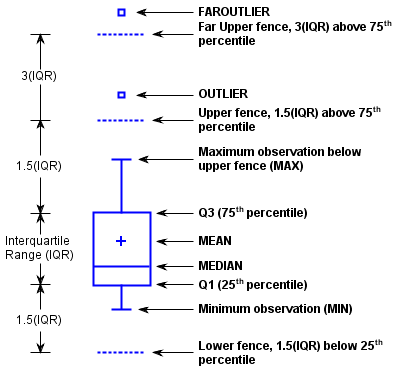
* Kurtosis of normal distribution is 3, regardless of the values of parameters and .
* The excess kurtosis of a random variable is defined as .
* A distribution with positive excess kurtosis has a more acute peak around the mean and fatter tails. A distribution with negative excess kurtosis has a lower, wider peak and thinner tails.

## 3.8 Quantile

**Definition (-quantile)**: Suppose random variable has a Let then the -quantile of the distribution is defined as which satisfies

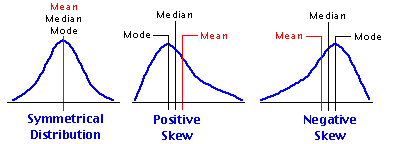
**Remarks**:

* When is continuous and strictly increasing,
* In case has flat regions or is discontinuous, we can define the -quantile as
* For -quantile we have
* 0.5-quantile is called the median. -quantile and -quantile are called lower quartile and upper quartile.
* 0-quantile is the minimum value; 1-quantile is the maximum value.



**Difference between mean and median**:

* Median is the cutoff point that divides the population in half.
* Mean can be misleading when used to measure the location of highly skewed data. In contrast, **median is a more robust** measure of the central tendency of a distribution in the sense that it is not much affected by a few outliers.
* Median is the optimal solution for minimizing the mean absolute error, that is,
* while mean is the optimal solution for minimizing the mean squared error, that is,
* For symmetric distribution, e.g. normal distribution, mean and median are the same. For skewed distributions, mean and median are different.



**Example (Value at Risk)**: The value at risk (VaR) at level of a portfolio over a certain time horizon is defined as

where is the return on the portfolio over the holding period , and is the information available at time . which is the negative conditional quantile of portfolio return at level is the threshold that actual loss will exceed with probability .

## 3.9 Moment Generating Function

**Definition (Moment Generating Function)**: The moment generating function(MGF) of a random variable is defined as

Remarks:

* may not exist for some If the expectation does not exist for any small neighborhood of 0, then we say that MGF does not exist for the distribution of .
* The existence of the MGF implies the existence of an infinite set of moments.

**Theorem**: If exists for in some neighborhood of then

1. for ,
2. the MGF of is

for all in a small neighborhood of 0.

**Proof**:

Assuming that we can differentiate under the integral sign, we have

Thus, .

Proceeding in an analogous manner, we can establish that

### 3.9.1 Discrete Distributions

**Example (Binomial distribution)**: For binomial distribution with PMF

**Proof**:

Using the binomial formula hence, letting and we have

**Example (Poisson distribution)**: For Poisson distribution with PMF

**Proof**:

**Example (Uniform distribution)**: For continuous uniform distribution on its PDF

Then

### 3.9.2 Continuous Distributions

**Example (Normal distribution)**: For normal distribution with PDF

**Proof**:

Define which implies hence,

**Example (Gamma mgf)**

Gamma pdf :

Mgf:

**Example (Log-normal distribution)**: Suppose follows a log-normal distribution. Let then and therefore

The MGF does not exists for .

### 3.9.3 Uniqueness of MGF

Recall that random variables and are identically distributed if two CDF's and are the same, i.e.

* If and are identically distributed, then for any ,
* Identity of the distributions of and does not imply . Example

**Example**: Suppose a fair coin is tossed times, and let be the number of heads obtained, be the number of tails obtained. Then but and possibly .

**Theorem (Uniqueness of MGF)**: Suppose two random variables and with MGF's and existing in a neighborhood of .

1. Then and have the same and for all in , if and only if for all .
2. If and have bounded support, then for all if and only if for all integers

**Remarks**:

* If the MGF exists in a neighborhood of it uniquely characterizes a distribution function.
* Given some MGF suppose we can find some that corresponds to Then must be the only distribution that generates
* Uniqueness of MGF can be used to prove CLT.

**Example**: A DRV has Then its PMF

**Example**: Suppose Recall the normal distribution has

Let and then So

**Example**: If a DRV has for where What is the probability distribution of

Note that

where since

is the PMF of .

Existence of the MGF

* The existence of the MGF implies the existence of
* The existence of all moments is not equivalent to the existence of moment generating function in a neighborhood of

**Example (Log-normal distribution)**: Suppose follows a log-normal distribution. Then

but the MGF

for .

**The set of moments does not uniquely characterize a distribution function.**

**Example**: Consider two distributions with support :

for for

**Theorem**: Let and be two both of which have bounded support. Then for all if and only if for all integers

**Theorem (Convergence of MGF)**: Suppose is a sequence of random variables, each with and Furthermore, suppose that

for all in a neighborhood of where is a MGF of a random variable with Then

for all continuous points of .

**Example (Poisson approximation)**: The MGF of the binomial distribution is

For every when and we have

which is the MGF of the Poisson distribution with parameter .

## 3.10 Characteristic Function

**Definition (Characteristic Function)**: The characteristic function of a random variable with is defined as

where and .

For continuous random variable, the **characteristic function is the** **Fourier transform of the PDF**, so the PDF can be recovered from the characteristic function by

Properties of characteristic function:

* For any probability distribution, the characteristic function always exists and is bounded, i.e. for any ,
* is continuous over
* If the MGF exists for in some neighborhood of then for all

**Theorem**: Suppose the -th moment of exists. Then is differentiable up to order and

**Theorem (Uniqueness of Characteristic Function)**: Suppose two random variables and have characteristic functions and respectively. Then and are identically distributed if and only if for all

**Remark**: It is important to check all on the entire real line. But for MGF's, it is only necessary to check in a neighborhood of

**Theorem (Convergence of Characteristic Function)**: Let be a sequence of random variables with CDF's and characteristic functions . Let be a random variable with CDF and characteristic function . Let .

1. If for all continuous points of then for every
2. Further, if for all then for all continuous points of .

**Leibnitz's Rule**: if and are differentiable respect to then