Axioms of Probability

Let S be a sample space (set of all possible outcomes),

* Axiom 1: P(E) >= 0 for all events E S
  + P(E) = 0 when event is empty set
* Axiom 2: P(S) = 1
* Axiom 3: Any countable sequence of disjoint sets (mutually exclusive events) E1, E2, . . . satisfies

Diagram

Description automatically generated with low confidence



For 2 events:

* 

You need to subtract the intersection if events are not mutually exclusive.

Probability Function and Random Variables

Let S = {s1, s2, …} be a sample space

* P : S 🡪 [0, 1]

such that and P() = 0

If sth changes, we call it variable. It varies.

Random variable also changes but it changes depend on random process. Event is random, outcomes are called random variables.

A random variable’s values depend on the outcomes of a random process:

* P(X = x)

X is a random variable, x is the values.

Probability of variable X, when it has a certain value.

Random Variable

A random variable is a function that associates a real number with each element in the sample space.

We shall use a capital letter, say X, to denote a random variable and its corresponding small letter, x in this case, for one of its values.

There are some electronical component. Pieces are defected (D) or non-defected (N).

In the electronic component testing illustration above, we notice that the random variable X assumes the value 2 for all elements in the subset

E = {DDN, DND, NDD}

of the sample space S. That is, each possible value of X represents an event that is a subset of the sample space for the given experiment.

If your interest is finding 2 defected ones and 1 non-defected one, then you can write down the Event set (E) like above. You don’t see any number up until now. We are not able to do any computation. These are random variables. These are things that change with a certain property.

Example:

* Table

  Description automatically generated
* ---
* Y is random variable which is defined by the # of red balls. Random variable is sth changing depending on the event or experiment you are working on.
* y is # of red balls. It is a value. It can be 2, 1, 1, or 0.
* 4 outcomes are possible (RR, RB, BR, BB)

Example:

* Graphical user interface, text

  Description automatically generated

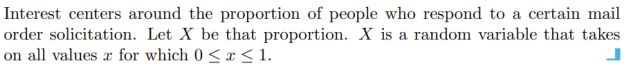
Discrete vs Continuous Sample Space

If a sample space contains a finite number of possibilities (Such as rolling a die, tossing a coin…) or an unending sequence with as many elements as there are whole numbers (They should be countable. Whole numbers means positive integers.), it is called a discrete sample space. Points isolated from other points. Sonlu sayıda da olabilir sonsuz sayıda da olabilir. Sonsuz sayıda olacaksa da hiçbir şeyin buçuğu küsüratı olmayacak şekilde hepsini tam bir şekilde integer harici bir şeye ihtiyaç duymadan sayabiliyor olmamız lazım.

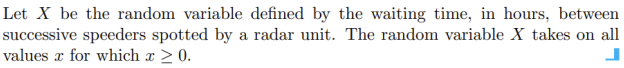
The outcomes of some statistical experiments may be neither finite nor countable. Such is the case, for example, when one conducts an investigation measuring the distances that a certain make of automobile will travel over a prescribed test course on 5 liters of gasoline. Assuming distance to be a variable measured to any degree of accuracy, then clearly we have an infinite number of possible distances in the sample space that cannot be equated to the number of whole numbers. Or, if one were to record the length of time for a chemical reaction to take place, once again the possible time intervals making up our sample space would be infinite in number and uncountable. We see now that all sample spaces need not be discrete.

If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a continuous sample space. Pouring a water to cup. Amount goes up continuously as you pour water. So amount of water there is continuous sample space. Getting exactly 1 liter of water is possible in theory but impossible in practice.

Example:

* 

Example:

* 
* ---
* Is sample space with time countable or not? Time is continuous.

**PROBABILITY DISTRIBUTIONS**

Discrete Probability Distributions

A discrete random variable assumes each of its values with a certain probability. In the case of tossing a coin three times, the random variable X, representing the number of heads, assumes the value 2 with probability 3/8, since 3 of the 8 equally likely sample points result in two heads and one tail. If one assumes equal weights for the simple events in Example 3.2, the probability that no employee gets back the right helmet, that is, the probability that M assumes the value 0, is 1/3. The possible values m of M and their probabilities are



Note that the values of m exhaust all possible cases and hence the probabilities add to 1.

Frequently, it is convenient to represent all the probabilities of a random variable X by a formula. Such a formula would necessarily be a function of the numerical values that we shall denote by f(x), g(x), r(x), and so forth. Therefore, we write f(x) = P(X = x); that is, f(3) = P(X = 3). The set of ordered pairs (x, f(x)) is called the probability function, probability mass function, or probability distribution of the discrete random variable X.

Random variableı nasıl tanımlarsan ona göre alacağı valuelar değişecek. Random variable, herhangi olan bir olayla onu sayı olarak hesaplayabileceğin bir şekle getirmek demek.

The set of ordered pairs (x, f(x)) is a probability function, probability mass function, or probability distribution of the discrete random variable X if, for each possible outcome x,

1. f(x) >= 0
2. P(X = x) = f(x)
   1. Prob. of random variable being a certain value x is f(x)

Discrete Example:

Text

Description automatically generated

Denominators are sample space.

***Cumulative Distribution Function:***

* The cumulative distribution function F(x) of a discrete random variable X with probability distribution f(x) is
  + A picture containing text, watch

    Description automatically generated
* For the random variable M, the number of correct matches in Example 3.2 (we didn’t go through it, just look piecewise, staircase function), we have
  + A picture containing text

    Description automatically generated
* The cumulative distribution function of M is
  + A picture containing text

    Description automatically generated
  + When you calculate 1 <= m < 3, you first do m < 3 then you do m < 1 and subtract it.

Remark: One should pay particular notice to the fact that the cumulative distribution function is a monotone nondecreasing function defined not only for the values assumed by the given random variable but for all real numbers. Nondecreasing çünkü her seferinde pozitif bir değer ekleye ekleye gidiyorsun.

Example – discrete (Example 3.9 is skipped but values f(0), f(1), f(2), f(3), and f(4) that calculated there is given):

* Text

  Description automatically generated
* Chart

  Description automatically generated
* x < 0, not defined, empty set, probability is 0
* Chart, line chart, box and whisker chart

  Description automatically generated

Continuous Probability Distributions

A continuous random variable has a probability of 0 of assuming exactly any of its values. Consequently, its probability distribution cannot be given in tabular form.

Probability of getting a point among infinitely many points is 0. Aralık olarak düşünüp işlem yaparız, nokta olarak düşünmeyiz.

We shall concern ourselves with computing probabilities for various intervals of continuous random variables such as P(a < X < b), P(W ≥ c), and so forth. Note that when X is continuous,

P(a < X ≤ b) = P(a < X < b) + P(X = b) = P(a < X < b).

That is, it does not matter whether we include an endpoint of the interval or not.

This is not true, though, when X is discrete.

We cannot make a table. We can either draw a plot or write a formula.

Although the probability distribution of a continuous random variable cannot be presented in tabular form, it can be stated as a formula. Such a formula would necessarily be a function of the numerical values of the continuous random variable X and as such will be represented by the functional notation f(x). In dealing with continuous variables, f(x) is usually called the probability density function (mass function is for discrete case, density function is for continuous case), or simply the density function, of X. Since X is defined over a continuous sample space, it is possible for f(x) to have a finite number of discontinuities. However, most density functions that have practical applications in the analysis of statistical data are continuous and their graphs may take any of several forms, some of which are shown in Figure 3.4. Because areas will be used to represent probabilities and probabilities are positive numerical values, the density function must lie entirely above the x axis.

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

***Probability Density Function:***

Definition 3.6: The function f(x) is a probability density function (pdf) for the continuous random variable X, defined over the set of real numbers, if

1. f(x) >= 0, for all x R
2. = 1
3. P(a < X < b) =

Example:

* Text, letter

  Description automatically generated
* First we need to make sure that function takes just positive values.
* Why we start with – infinity to + infinity and we ended up with -1 and 2, because it is 0 elsewhere. Sum of all 0s is 0.
* Even though this 1 is included, it doesn’t matter because equality doesn’t change anything because property of continuous case 🡪 probability at one certain point is 0

***Cumulative Distribution Function:***

The cumulative distribution function F(x) of a continuous random variable X with density function f(x) is



As an immediate consequence of this definition, one can write the 2 results



if the derivative exists.

Example:

* Text, letter

  Description automatically generated
* It is just defined between -1 and 2. So we concern about this area. All area up until -1 is 0.
* For x more than 2, we will have probability of sample space which is 1.

**REVISION & EXAMPLES**

Random variables can be defined on 2 different sample space:

* discrete 🡪 points
* continuous 🡪 intervals

Discrete

We deal with points here!

We can count them. Even there are infinitely many of them, we can still count them just like natural numbers.

Distribution of X (Collection of all possibilities):

f(x) = P(X = x) ------------> probability mass function (pmf)

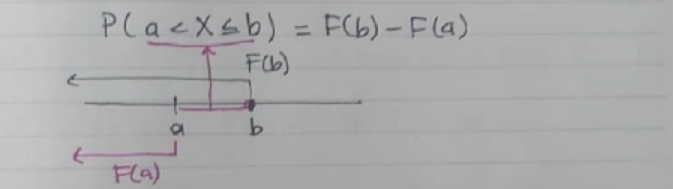
input is value of random variable, it means we are looking for prob. of that random variable at   
 that certain value

*Cumulative Distribution Function:*

F(x) = P(X <= x) =

F(x) is non decreasing function of x with:

F(x) = 0 and F(x) = 1



Continuous

We deal with intervals here!

P(X = a) = 0 for any a.

**P(a < X <= b)** = P(a < X < b) + P(X = b) = **P(a < X < b)**

P(a < X < b) = -----> probability density function

A picture containing text, document, linedrawing

Description automatically generated

Example:

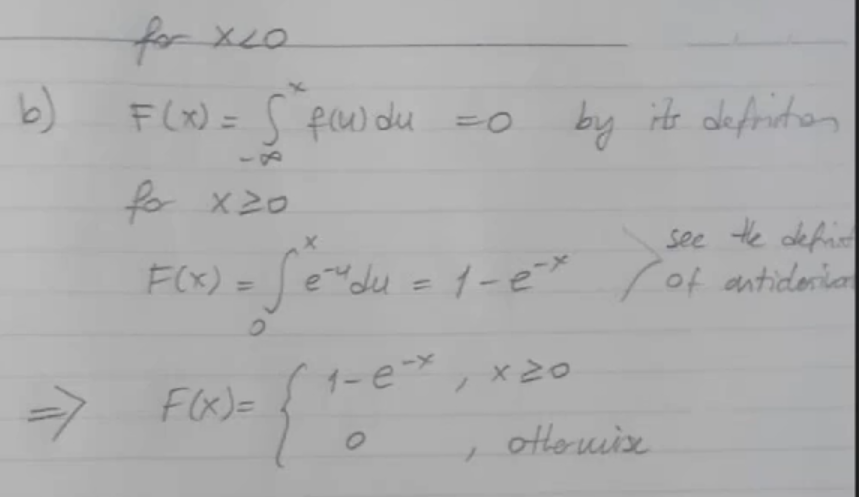
Text, letter

Description automatically generated

Example:

A piece of paper with writing

Description automatically generated with medium confidence



see the definition of antiderivative

Text, letter

Description automatically generated

You can use either difference between two cumulative distribution functions OR directly integration of the pdf.