**Multiple Random Variables and Random Vectors**

Table of Contents

[Multivariate Join Probability Mass Function 4](#_Toc63693514)

[Marginal PMF 4](#_Toc63693515)

[Degree of Freedom 6](#_Toc63693516)

[Multivariate Joint Cumulative Distributive Functions 8](#_Toc63693517)

[Marginal CDF 8](#_Toc63693518)

[Well-Known Multivariate Discrete Random Variables 9](#_Toc63693519)

[Multivariate Hypergeometric Distributions 9](#_Toc63693520)

[Multinomial Distributions 10](#_Toc63693521)

[Multivariate Continuous Random Variables 12](#_Toc63693522)

[Multivariate Joint Cumulating Distributive Functions 12](#_Toc63693523)

[Multivariate Joint Probability Density Functions 12](#_Toc63693524)

[Random Vectors 15](#_Toc63693525)

[Distribution Functions 16](#_Toc63693526)

[Joint Distribution Functions 17](#_Toc63693527)

[Expected Value Vectors 18](#_Toc63693528)

[Expected Value Matrices 18](#_Toc63693529)

[Covariance 18](#_Toc63693530)

[Correlation 21](#_Toc63693531)

[Covariance and Correlation of Two Random Vectors 24](#_Toc63693532)

Consider a scenario where a single random experiment has multiple random variables associated with it. We will only be considering the cases where all the random variables are from discrete sets or all of them are from continuous sets. All of the random variables must be from the same random experiment.

Say we have a set of discrete random variables, , , , , each of which are associated with a set of values, , , , respectively. For a single instance of the experiment, the random variables have the values , , , respectively.

## Multivariate Join Probability Mass Function

In this scenario, the probability model we will be using is the multivariate joint PMF, MJPMF.

Essentially, this gives us the probability of getting a single point in an -dimensional space.

Of course, to be a PMF at all, a few conditions have to be met:

* ,

### Marginal PMF

For MJPMF’s of random variables, we can define marginal PMFs for random variables, where . If , the marginal PMF will be a marginal joint PMF.

Say . Thus, the MJPMF is .

Of course, all other combinations are possible, such as . We just need to take the sum of the PMFs for the other remaining random variables.

Example

Say an access point (AP) forwards data from many different devices to a particular router. Say the AP forwards data packets, of which

* are lost due to channel errors
* are dropped
* are delivered successfully

From the packets, we observe random packets.

We can have three random variables here,

From here, we want to fine the MJPMF and the marginal PMF or marginal joint PMF.

There are two conditions here that are implied that we must abide by:

* , ,

This is just going to be the number of ways we can pick lost packets, dropped packets and delivered packets divided by the number of ways packets can be randomly observed from packets.

Notice the last condition. Essentially, since we are only picking packets, the values of and are limited to . Additionally, since the maximum value of is , is limited to .

### Degree of Freedom

Say we want to fine the marginal PMF for .

Consider that . Then, cannot be more than , since the total number of packets is limited to . If , then has to be . It is said that the degree of freedom is in this case. This means that, even though we have three random variables, we cannot independently choose values for all three.

Since we have defined a particular value for , the next question is to ask for the value of . One we have a value for as well, we can no longer choose the value of . Even if we had chosen and , the value of would have to be . It is fixed. Thus, we could only independently choose values for two of the random variables.

Because the value of is fixed once we have picked a value for and one for , we just need to loop over all the values of .

has been brought outside since it is not related to .

Thus, the second part of the equation, , essentially tells us the number of ways to pick dropped and/or delivered packets, we do not care which. However, that is just . Thus,

Similarly,

For a marginal joint PMF of two values, we again cannot choose any values for the third random variable. Thus, we do not even need to sum anything.

## Multivariate Joint Cumulative Distributive Functions

The MJCDF of random variables is given by

### Marginal CDF

## Well-Known Multivariate Discrete Random Variables

### Multivariate Hypergeometric Distributions

With multivariate hypergeometric distributions, we deal with items of types are mixed together. is the number of items of the -th type, where is between and . Thus,

From here, we will pick items randomly and without replacement.

Finally, we need to find . This is given by

Looking at this, it should be obvious that this is what we were working with in the previous lecture.

### Multinomial Distributions

Multinomial distributions are the same as multivariate hypergeometric distributions, except that we replace items picked. Thus, the probability of picking an item of the -th type, , remains constant at and .

This can directly be related to binomial random variables.

Here, the term is called the multinomial coefficient, similar to the binomial coefficient.

Say and we are picking items. Thus,

Of course, this is the binomial distribution exactly.

Remember that a Bernoulli experiment was one in which there were only two possible outcomes. A binomial distribution is a type of Bernoulli experiment. If instead we have multiple possible outcomes, one of the possible distributions is the multinomial distribution. Such experiments are sometimes called multinoulli or categorical distributions, since we have multiple categories.

We can compare Bernoulli experiments to tossing a coin and counting how many heads and how many tails we get. Multinoulli experiments would then be like rolling a die and counting how many outcomes of each category we get.

## Multivariate Continuous Random Variables

With multivariate continuous random variables, we have continuous random variables defined from the same sample space, from to .

### Multivariate Joint Cumulating Distributive Functions

MJCDFs are defined as

### Multivariate Joint Probability Density Functions

MJPDFs can be defined with the help of MJCDFs. The -th order partial derivative of the MJCDF for a multivariate continuous random variable is its MJPDF.

MJPDFs have a few conditions they need to abide by:

#### Marginal PDFs

The marginal PDF of a single random variable or the marginal joint PDF of two random variables or the marginal multivariate joint PDF of random variables, where , can be found.

Example

Say we have three random variables, , and , for which the MJCDF is given as

From this, we need to find the MJPDF and the marginal PDF.

Now say we want to find the marginal PDF for and .

If we wanted to find the marginal PDF of , we could do this from the original PDF, or from the marginal PDF of and .

We can even find the marginal CDFs by integrating the marginal PDFs.

## Random Vectors

Say we have a set of random variables which are

* either all discrete or all continuous
* defined on a single sample space
* jointly distributed

In this scenario, the random variable can be represented by a vector. This vector is called a random vector.

Random vectors are denoted using bold, capital letters. If bold letters are inconvenient, for example when writing on physical paper, double lines may be used.

Random Variable

Value of a random variable

or Random Vector

or Value of random vector

Random vectors are column vectors by default, unless otherwise mentioned. However, since writing the vectors as columns would take a huge amount of space, they are also frequently written in their transposed form.

### Distribution Functions

Just by looking at the CDF of a random vector, we can tell that we can represent a multivariate random variable in a much more concise manner, which is exactly why it is used. Additionally, we can use matrix algebra with random vectors.

Similarly,

Example

Since there are three different values in , there must also be three different random variables involved here. Thus,

Thus,

If we want to find the CDF,

### Joint Distribution Functions

If we have a pair of random vectors, and , and we know that they are either both distribute or continuous, the distribution functions are represented as:

Another way of looking at this is as though is a random vector such that , and we are just finding the distribution functions for .

### Expected Value Vectors

If is a random vector such that , then the individual expectations of each of the vectors in can be expressed as a vector. This vector is called the expected value vector.

The expected value vector can also be expressed as .

### Expected Value Matrices

Similar to random vectors, we can also have random matrices, which are matrices of random variables. We will not discuss those in depth here. For random matrices, we can have corresponding expected value matrices.

### Covariance

We have previously seen covariances when discussing joint random variables. For the joint random variables and , the covariance was defined as

This was also alternatively written as

In this second form, the value was given a special name, the correlation, . This special name was given since the correlation will crop up in many different places. Thus,

From this, it would make sense that the covariance of a random vector is just the covariance of the random variables that make up the vector. However, by definition, covariance refers to just two random variables. As such, the covariance of a random vector is the covariance of all possible pairs of random variables in the vector.

If a vector has random variables, each of those random variables can make pairs with random variables, including with themselves. Thus, there are pairs.

There are two points to notice about the covariance of a random vector. Firstly, notice the covariance of pairs with the same random variables.

Thus, the covariance of pairs with the same random variables is just the variance of that random variable. This means every element of the matrix along the diagonal is going to be a variance.

Secondly, notice that the matrix for the covariance of the random vector is symmetric, since .

From these two points, we can make another observation. A symmetric matrix is the product of a column vector and its transpose. Thus, if we consider the following column vector and its transpose,

This product can also be written as . The result is an matrix. The expectation of this matrix is .

Thus,

Similar to normal covariances, we also have alternative formulae for this.

Here, is the correlation of . It is also a matrix.

### Correlation

We know that .

Notice that all the values along the diagonal are just the squares of the random variables.

Thus, the correlation matrix is also a symmetric matrix. The values along the diagonal are the expected values of the squares of the component random variables, and all other values are the expected values of all possible pairs. Note that each of these is a correlation as well, i.e. .

Example

Let and for . We need to find , and .

To be able to find the expectation, we first need to find the PDFs of the individual random variables.

### Covariance and Correlation of Two Random Vectors

Here, we can also define a derived random vector from the random vector as

where is a matrix, is a vector and is a vector.

Further,

For example, we found values for , and in the previous example. Thus, provided a specific and , we can easily calculate all of these values.