STAT 305: Chapter 5

Part IV

Amin Shirazi

Course page:
ashirazist.github.io/stat305_s2020.github.io

Chapter 5.4: Joint Distributions and Independence

Working with Multiple Random Variables

Joint Distributions

We often need to consider two random variables together - for instance, we may consider

- the length and weight of a squirrel,
- the loudness and clarity of a speaker,
- the blood concentration of Protein A, B, and C and so on.

This means that we need a way to describe the probability of two variables *jointly*. We call the way the probability is simultaneously assigned the "joint distribution".

Joint distribution of discrete random variables

For several discrete random variable, the device typically used to specify probabilities is a *joint probability function*. The two-variable version of this is defined.

> A **joint probability function (joint pmf)** for discrete random variables and is a nonnegative function , giving the probability that (simultaneously) takes the values and takes the values . That is,

Properties of a valid joint pmf:

- for all
-

Joint distribution of discrete random variables

So we have probability functions for , probability functions for and now a probability function for and together - that's a lot of s floating around though! In order to be clear which function we refer to when we refer to , we also add some subscripts

Suppose and are two discrete random variables.

- we may need to identify the *joint probability function* using ,
- we may need to identify the probability function of by itself (aka the *marginal probability function* for) using ,
- we may need to identify the probability function of by itself (aka the *marginal probability function* for) using

Joint pmf

For the discrete case, it is useful to give in a **table**.

Two bolt torques, cont'd

Recall the example of measure the bolt torques on the face plates of a heavy equipment component to the nearest integer. With

Joint pmf

the joint probability function,   , is

y\x	11	12	13	14	15	16	17	18	19	20
20	0	0	0	0	0	0	0	2/34	2/34	1/34
19	0	0	0	0	0	0	2/34	0	0	0
18	0	0	1/34	1/34	0	0	1/34	1/34	1/34	0
17	0	0	0	0	2/34	1/34	1/34	2/34	0	0
16	0	0	0	1/34	2/34	2/34	0	0	2/34	0
15	1/34	1/34	0	0	3/34	0	0	0	0	0
14	0	0	0	0	1/34	0	0	2/34	0	0
13	0	0	0	0	1/34	0	0	0	0	0

Calculate:

-
-

Joint Distributions

Discrete RVs

By summing up certain values of , probabilities associated with and with patterns of interest can be obtained.

Consider:

y\x	11	12	13	14	15	16	17	18	19	20
20										
19										
18										
17										
16										
15										
14										
13										

Discrete RVs

y\x	11	12	13	14	15	16	17	18	19	20
20										
19										
18										
17										
16										
15										
14										
13										

Discrete RVs

y\x	11	12	13	14	15	16	17	18	19	20
20										
19										
18										
17										
16										
15										
14										
13										

Marginal Distribution

Joint Distributions

Discrete RVs

Marginal distributions

In a bivariate problem, one can add down columns in the (two-way) table of to get values for the probability function of , and across rows in the same table to get values for the probability distribution of , .

The individual probability functions for discrete random variables and with joint probability function are called **marginal probability functions**. They are obtained by summing values over all possible values of the other variable.

Connecting Joint and Marginal Distributions

Discrete RVs

Use: Joint to Marginal for Discrete RVs

Let and be discrete random variables with joint probability function Then the marginal probability function for can be found by:

and the marginal probability function for can be found by:

`

Example: [Torques, cont'd]

Find the marginal probability functions for and from the following joint pmf.

y\x	11	12	13	14	15	16	17	18	19	20
20	0	0	0	0	0	0	0	2/34	2/34	1/34
19	0	0	0	0	0	0	2/34	0	0	0
18	0	0	1/34	1/34	0	0	1/34	1/34	1/34	0
17	0	0	0	0	2/34	1/34	1/34	2/34	0	0
16	0	0	0	1/34	2/34	2/34	0	0	2/34	0
15	1/34	1/34	0	0	3/34	0	0	0	0	0
14	0	0	0	0	1/34	0	0	2/34	0	0
13	0	0	0	0	1/34	0	0	0	0	0

Joint Distributions

Discrete RVs

Getting marginal probability functions from joint probability functions begs the question whether the process can be reversed.

> **Can we find joint probability functions from marginal probability functions?**

Conditional Distribution

Joint Distributions

Discrete RVs

Conditional Distribution

Conditional Distribution of Discrete Random Variables

When working with several random variables, it is often useful to think about what is expected of one of the variables, given the values assumed by all others.

> For discrete random variables and with joint probability function , the **conditional probability function of given** is *a function of*
>
> and the **conditional probability function of given** is *a function of*

Joint Distributions

Discrete RVs

Conditional Distribution

Example: [Torque, cont'd]

y\x	11	12	13	14	15	16	17	18	19	20
20	0	0	0	0	0	0	0	2/34	2/34	1/34
19	0	0	0	0	0	0	2/34	0	0	0
18	0	0	1/34	1/34	0	0	1/34	1/34	1/34	0
17	0	0	0	0	2/34	1/34	1/34	2/34	0	0
16	0	0	0	1/34	2/34	2/34	0	0	2/34	0
15	1/34	1/34	0	0	3/34	0	0	0	0	0
14	0	0	0	0	1/34	0	0	2/34	0	0
13	0	0	0	0	1/34	0	0	0	0	0

Find the following probabilities:

-

Joint Distributions

Discrete RVs

Conditional Distribution

Example: [Torque, cont'd]

-
-
-

Independence

Let's start with an example. Look at the following joint probability distribution and the associated marginal probabilities.

$y \backslash x$	1	2	3	$f_Y(y)$
3	0.08	0.08	0.04	0.20
2	0.16	0.16	0.08	0.40
1	0.16	0.16	0.08	0.40
$f_X(x)$	0.40	0.40	0.20	1.00

What do you notice?

Joint Distributions

Discrete RVs

Conditional Distribution

Independence

Discrete random variables and are **independent** if their joint distribution function is the product of their respective marginal probability functions. This is,

> independence means that
>
> If this does not hold, then and are **dependent**

If and are not only independent but also have the same marginal distribution, then they are **independent and identically distributed (iid)**.

Chapter 5.5: Functions of Random Variables

Results and Theorems

Functions of Random Variables

A random variable can be thought of as a function whose input is an outcome and whose output is a real number. When we take a function of the value the random variable takes, the resulting value still depends on the outcome of a random experiment - in other words: functions of random variables are random variables.

This means that a function of a random variable will have probabilities attached to the value it takes, based on the value taken by the random variable. It also means functions of random variables will have:

> - probability functions (if discrete) or probability density functions (if continuous)
> - cumulative probability functions (if discrete) or cumulative density functions (if continuous)
> - expected values and variances …

Functions of RVs

Linear Combinations

Linear combinations

For engineering purposes, it often suffices to know the mean and variance for a function of several random variables, (as opposed to knowing the whole distribution of). When is **linear**, there are explicit functions.

Proposition: If are **independent** random variables and are constants, then consider

> U is itself a random variable as it is a linear combination of n *independent* random variables

Linear combinations [cont'd]

U, as a random variable has mean

and variance

Functions of RVs

Linear Combinations

Example:

Say we have two independent random variables and with , and . Find the mean and variance for

- U = 3 + 2X - 3Y

- V = -4X + 3Y

Functions of RVs

Linear Combinations

Example:

Say and

and and are independent.

Calculate the mean and variance of .

Functions of RVs

Linear Combinations

Sample Mean

A particularly important use of functions of random variables concerns iid random variables where each − for . Then we can define the random variable as follows

— — —

> Note that is a random variable

We can then find the mean and variance of this random variable.

as they relate to the population parameters and .

For **independent** variables with common mean and variance ,

What is the point?

It does not matter if we are working with discrete or continuous random variables, as long as we have an independent and identically distributed (iid) sample of size with the same mean and the same variance , the random variable has

and

> The point is that the variance of a sample mean of size is the population variance devided by the sample size which makes it smaller
>
> i.e. as the sample size increases, the variability of the sample mean decreases.

Functions of RVs

Linear Combinations

Sample Mean

Example:[Seed lengths]

One botanist measured the length of seeds from the same plant. The seed lengths measurements are . Suppose it is known that the seed lengths are iid with mean mm and variance mm.

Calculate the mean and variance of the average of seed measurements.

Central Limit Theorem

The Most Important Result in Statistics

Central limit theorem

One of the most frequently used statistics in engineering applications is the sample mean. We can relate the mean and variance of the probability distribution of the sample mean to those of a single observation when an iid model is appropriate.

In the case of the sample mean, if the sample size is large enough, we can also approximate the *shape* of the *probability distribution function* of the sample mean!

Functions of RVs

Linear Combinations

Sample Mean

CLT

Central limit theorem

If are **independent** and **identically** distributed (iid) random variable (with mean and variance), then for large , the variable is approximately normally distributed. That is,

—

This is one of the **most important** results in statistics.

Functions of RVs

Linear Combinations

Sample Mean

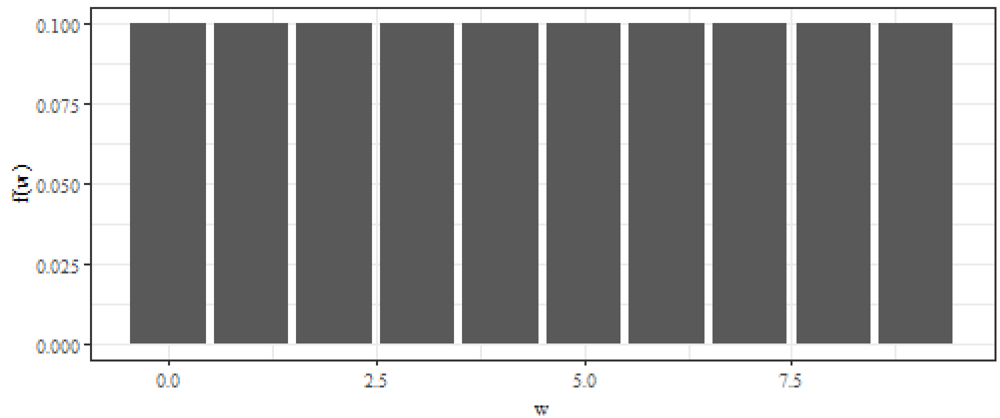
CLT

Example: [Tool serial numbers]

Consider selecting the last digit of randomly selected serial numbers of pneumatic tools. Let

A plausible model for the pair of random variables
is that they are independent, each with the marginal probability function

Example: [Tool serial numbers]



With and .

Using such a distribution, it is possible to see that

– has probability distribution

$\overline{w}$	$f(\overline{w})$	$\overline{w}$	$f(\overline{w})$	$\overline{w}$	$f(\overline{w})$	$\overline{w}$	$f(\overline{w})$	$\overline{w}$	$f(\overline{w})$
0.00	0.01	2.00	0.05	4.00	0.09	6.00	0.07	8	0.03
0.50	0.02	2.50	0.06	4.50	0.10	6.50	0.06	8.5	0.02
1.00	0.03	3.00	0.07	5.00	0.09	7.00	0.05	9	0.01
1.50	0.04	3.50	0.08	5.50	0.08	7.50	0.04		

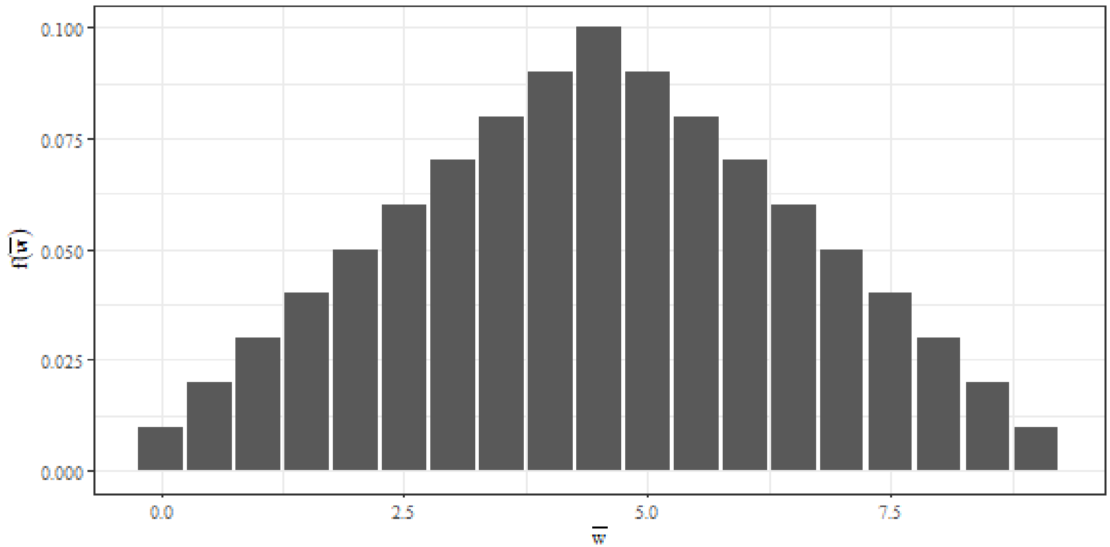
Functions of RVs

Linear Combinations

Sample Mean

CLT

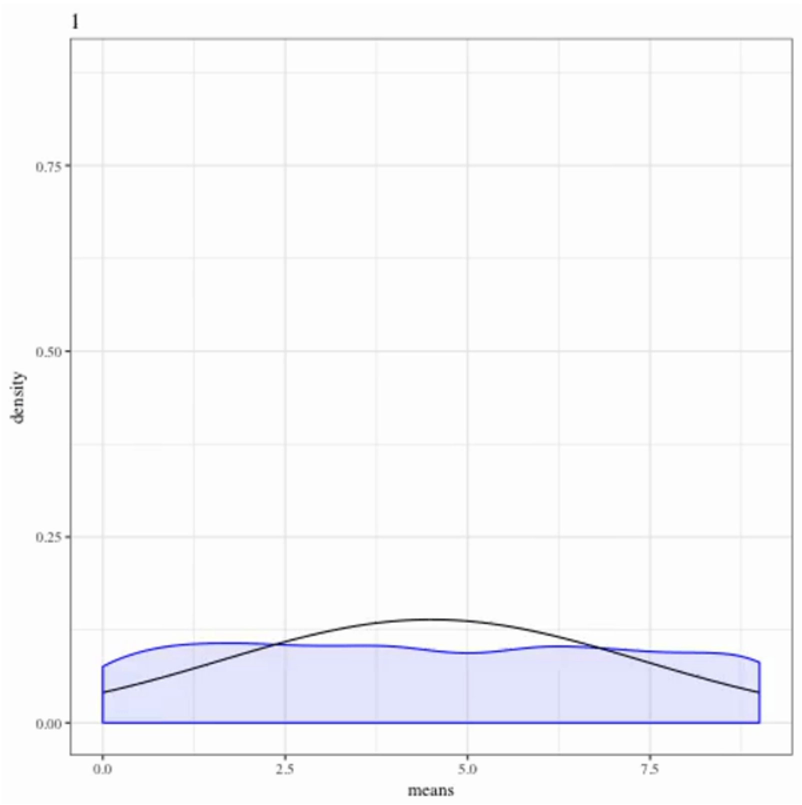
Example: [Tool serial numbers]



Comparing the two distributions, it is clear that even for a completely flat/uniform distribution of and a small sample size of , the probability distribution of looks more bell-shaped than the underlying distribution.

Now consider larger and larger sample sizes,
:

Watch how CLT works here



Functions of RVs

Linear Combinations

Sample Mean

CLT

Example: [Stamp sale time]

Imagine you are a stamp salesperson (on eBay). Consider the time required to complete a stamp sale as , and let

Each individual sale time should have an distribution. We want to consider approximating

.

Functions of RVs

Linear Combinations

Sample Mean

CLT

Example: [Cars]

Suppose a bunch of cars pass through certain stretch of road. Whenever a car comes, you look at your watch and record the time. Let be the time (in minutes) between when the car comes and the car comes for . Suppose you know the average time between cars is minute.

Find the probability that the average time gap between cars for the next 44 cars exceeds 1.05 minutes.

Example: [Baby food jars, cont'd]

The process of filling food containers appears to have an inherent standard deviation of measured fill weights on the order of . Suppose we want to calibrate the filling machine by setting an adjustment knob and filling a run of jars. Their sample mean net contents will serve as an indication of the process mean fill level corresponding to that knob setting.

You want to choose a sample size, , large enough that there is an chance the sample mean is within g of the actual process mean.

Example: [Printing mistakes]

Suppose the number of printing mistakes on a page follows some unknown distribution with a mean of and a variance of . Assume that number of printing mistakes on a printed page are iid.

- What is the approximate probability distribution of the average number of printing mistakes on 50 pages?

- Can you find the probability that the number of printing mistakes on a single page is less than 3.8?

Example: [Printing mistakes]

- Can you find the probability that the average number of printing mistakes on 10 pages is less than 3.8?

- Can you find the probability that the average number of printing mistakes on 50 pages is less than 3.8?