

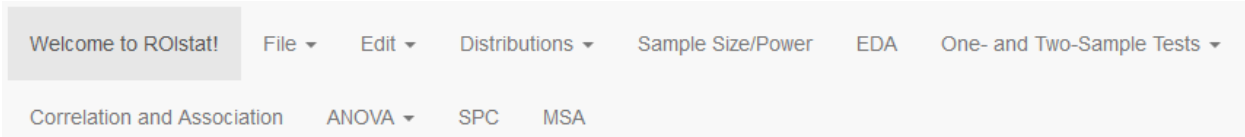
Using ROIstat

Once ROIstat is installed and running, here is how to use it for the various lessons in the book.

The Basics

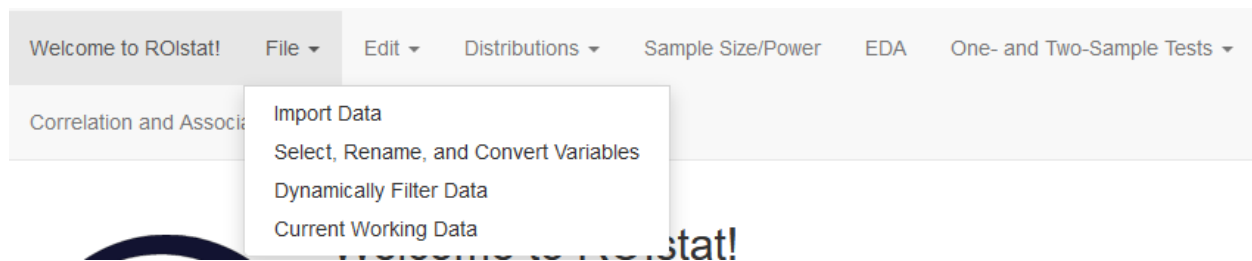
ROIstat is an app that is set up to have a familiar user interface.

Along the top you have a familiar menu set up:



File

Clicking on **File** will bring up these options:



Import Data

ROIstat can import data from:

- **The R Environment** – this is useful if you are working in both R and ROIstat, perhaps using R for data manipulation and ROIstat for the analysis.
- **Local File** – this is just a file that is on your computer somewhere. There are a large number of formats that can be imported (anything rio can import, ROIstat can), so you can use most plain data files as well as save files from a number of analysis programs. You may need to select “All Files” rather than “All Supported Types” when you get to the File Upload page to see them.
- **Paste Data In** – If you have copied data from a spreadsheet or word processor, you can paste it directly into a window and ROIstat will do its best to parse it into a data file.
- **Google Sheets** – If you have data stored in Google Sheets, you can import it directly into ROIstat.

The Local File and Paste Data In options are pre-selected, since they are the most common, but you can activate the other two just by clicking the check box next to their names:

Import Data

Select where the data file is located then click "Open Import Window" below:

- ☐ R Environment
- ☒ Local File
- ☒ Paste Data In
- ☐ Google Sheets

Open import window

Your original file is only accessed here. Nothing is changed in the original file.

Once you click the **Open Import Window** button above, you will see your active options for importing the data. Select your option and the import window will change to accommodate your choice.

For the **R Environment** it will have a drop-down list of the data frames that are in the selected environment. All you have to do is select the one you want:

Import data to be used in application

Import View Update

How to import data?

- ☒ Environment
- ☐ External file
- ☐ Copy / Paste
- ☐ Google sheets

Select a data.frame:

List of data.frame...

Select an environment in which to search:

Global Environment

No data selected! Use a data.frame from your environment or from the environment of a package.

Import data

For **External files** you will see the familiar Browse button interface that will allow you to navigate to the folder where the data are stored:

Import data to be used in application

Import

View

Update

</> Environment

External file

Copy / Paste

Googlesheets

Upload a file:

Browse...

No file selected

No file selected: You can import .csv, .txt, .xls, .xlsx, .rds, .fst, .sas7bdat, .sav files

Import data

To **Copy/Paste**, just paste what you have in the box:

Import data to be used in application

Import

View

Update

</> Environment

External file

Copy / Paste

Googlesheets

Paste data here:

Add a label to data

Nothing pasted yet! Please copy and paste some data in the dialog box above.

Import data

And for **Google Sheets**, just enter the link to the file:

Import data to be used in application

Import

View

Update

How to import data?

</> Environment

External file

Copy / Paste

Googlesheets

Enter a shareable link to a GoogleSheet:

Nothing pasted yet!

 Please paste a valid GoogleSheet link in the dialog box above.

Import data

You can also click on the **View** tab to see the data you have selected before importing it. The **Update** tab allows you to change the names or classes, see if there is missing data, and choose different date and time formats or decimal separators to handle whatever the specifics of your file might be. You can also deselect variables to be imported. Click **Apply Changes** to use the new settings.

Import data to be used in application

Import

View

Update

Data has 25 observations and 4 variables.

<input checked="" type="checkbox"/>	Name	Class	New class	Missing
<input checked="" type="checkbox"/>	Production Line	integer	integer	
<input checked="" type="checkbox"/>	Production Line	integer	integer	
<input checked="" type="checkbox"/>	Production Line	integer	integer	
<input checked="" type="checkbox"/>	Production Line	integer	integer	

Date format:

%Y-%m-%d

Date to use as origin to convert date/datetime:

1970-01-01

Decimal separator:

0.00 .

1

Select, rename and convert variables in table above, then apply changes by clicking button below.

Apply changes

Import data

Clicking on **Import Data** will bring a copy of the data into ROIstat. Note that nothing you do in ROIstat changes the original file.

Select, Rename, and Convert Variables

This provides the same function as the **Update** tab did in **Data Import**, namely you can change the name, class and formats of the data you have imported. This page also gives you a snapshot of the data you have loaded as well as the data after the changes you have made are executed. If you want to reduce the number of columns of data, deselected the check box next to the name of the data. It doesn't delete it, but it will not show up when selecting data in the app. It is a quick way to declutter if you have a big data set. You again have to click **Apply Changes** for them to take effect.

Select, rename and convert variables

Update & select variables

Data has 25 observations and 4 variables.

<input checked="" type="checkbox"/>	Name	Class	New class	Missing v...	Comp
<input checked="" type="checkbox"/>	Productio	integer	integer ▾	0	
<input checked="" type="checkbox"/>	Productio	integer	integer ▾	0	
<input checked="" type="checkbox"/>	Productio	integer	integer ▾	0	
<input checked="" type="checkbox"/>	Productio	integer	integer ▾	0	

① Select, rename and convert variables in table above, then apply changes by clicking button below.

➔ Apply changes

Original data:

	Production Line 1	Production Line 2	Production Line 3	Production L
1	1182	1254	874	
2	716	956	864	
3	904	707	1472	
4	821	1472	1120	
5	917	794	1145	
6	675	1039	1191	
7	887	915	1086	
8	1124	690	1114	
9	1185	444	1502	
10	1151	853	846	
11	892	333	1066	
12	1236	1013	1173	
13	956	731	1304	
14	1211	877	1291	
15	1280	1399	825	
16	968	891	713	
17	1493	894	1278	
18	104	1018	602	
19	798	354	1102	
20	1172	1036	768	
21	1299	971	775	
22	656	870	1224	
23	1066	683	882	
24	817	1529	1455	
25	896	565	1506	

Dynamically Filter Data

This allows you to perform some simple filtering on your data sets. If the column is marked as integer, numeric or date, it will give you a slider with a lower and upper marker. This will allow you to set lower and upper filters for each column. If there is missing data, it will give you the option to delete each row with missing data.

Current Working Data

This shows you the data as selected and filtered by the previous steps.

Edit Settings

Here you can change the default color palette for ROIstat. The default “R4” color palette is usually adequate for most uses and accommodates most color-blind users. Drop down the list to choose other palettes including Okabe-Ito for color vision deficiencies. A demo of the colors is shown for each selection in a graph below.

Using ROIstats in *Data-Driven Leadership*

From now on, I’ll show you how to use ROIstats to replicate the activities in Data-Driven Leadership. I’ll just go through in order. Anything in **bold** indicates something that you will see or select in the app.

Chapter 3 – Understanding and Presenting Data

Making a Run Chart

Once you load **PRODUCTION.CSV**, select the **SPC** menu item. You will see this:

Control Charts

Variables

Attributes

Limits Calculations

1. Set Up Chart

Mean ☐ Individuals ☒


Subgroups are:
☒ A row across multiple columns
☐ Defined by a column

Select Data Columns

Nothing selected

☐ Plot as run chart (no limits)?

2. Select Chart Types and Limits



This allows you to set up run charts and statistical process control (SPC) charts. We are only looking to set up a run chart to start with.

Step 1 is to set up the chart. Since the data is a list of individual observations, select **Individuals**. Next you need to specify which column you want to chart. In the book it was Production Line 2, so choose that in the drop down labeled **Select Data Column**. You can leave the **Select Set Column** on None.

Now select the **Plot as run chart (no limits)?** checkbox. After a moment it will create the chart. You can hover your mouse over a point to see information about it.

You can change where the centerline is by selecting either the mean or the median on the slider switch. You can adjust the base font size to make the numbers and labels larger or smaller. You can select a variety of formats to download the graph.


Making a Control Chart

Choosing the right chart and the right calculation for limits for control charts is a whole topic by itself, but here is how to make the ones in the book.

If you have just made the run chart as above, unselect the checkbox. Step 2 – Select Chart Types and Limits will appear. Otherwise, make sure that **Individuals** is selected and you have chosen Production Line 2 as the data column.

Click on the dropdown icon to get a list of the various ways you can make control charts. Once you do that, a chart will be plotted. By default, it will have the following selections, which are correct for this example.

2. Select Chart Types and Limits

 Click to select charts and limits

Location

X ▼

Location Centerline

Mean ▼

Span of MR

2 ▼

X Limit Calculation

Average Moving Range of X ▼

Standard Errors

3 ▼

Dispersion

Moving Range ▼

Dispersion Centerline

Mean ▼


Moving Range Limit Calculations

Average MR ▼

Click on the dropdown button again to hide the limits selection box.

By changing the data column, you can plot Production Line 1 with its special cause.

There are a variety of customizations you can make to the chart here:

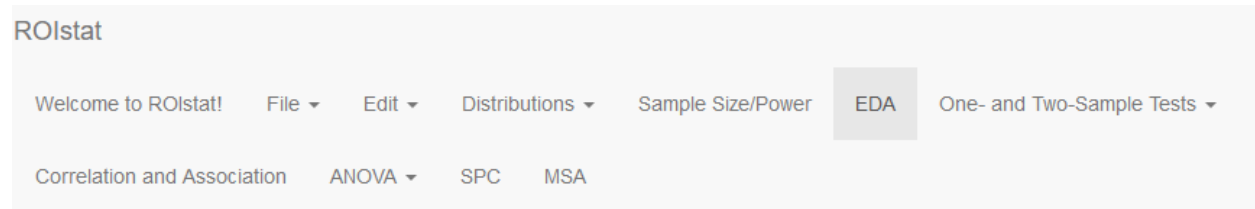
 ☐ Show Analysis ☐ Show plot data **Decimals**

4 ▼

The gear dropdown menu allows you to customize which pattern detection rules the control chart uses and which features are used on the graph. The **Show Analysis** slider performs a number of analyses on the chart. The **Show Plot Data** slider creates a table with the data in case you want to copy and paste that and create a custom chart of your own. The Decimals box controls the number of decimals displayed throughout.

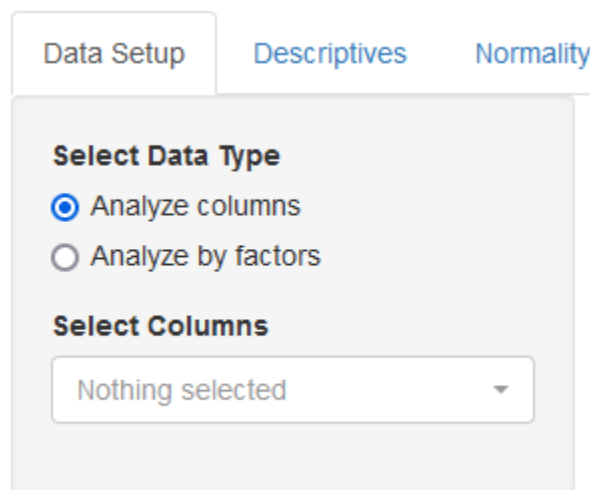
Getting Statistics

A quick way to get statistics from your loaded data is to click on the **EDA** menu item.



EDA stands for “Exploratory Data Analysis” and it is often the first step to get to know your data set.

When you first enter EDA, you will have to tell the app how your data is structured.



In the case of **PRODUCTION.CSV**, each column is a production line, so we want to analyze by columns. Clicking on the **Select Columns** dropdown will show all the columns. You can click one or more of these to explore. As you add data you will see a sample of the data appear to the right. This is helpful to see if you have the data you expected selected.

If you then click on the Descriptives tab, you will get a variety of common descriptive statistics.

[Data Setup](#)[Descriptives](#)[Normality Tests](#)[Boxplots](#)[Histograms](#)[Quantiles](#)[Conf. Int.](#)[Nat. Tol.](#)

Descriptive Statistics

Decimals

5

Select Statistics
Mean, Standard Deviation, Var ▼

Search:

dv.name	n	mean	sd	min	max	range
1 Production.Line.1	25	976.24	281.69876	104	1493	1389
2 Production.Line.2	25	891.52	308.22153	333	1529	1196
3 Production.Line.3	25	1087.12	261.2075	602	1506	904
4 Production.Line.4	25	950.52	135.96756	630	1216	586

Showing 1 to 4 of 4 entries

The **Decimals** box controls the number of decimals displayed. If you wish to add or remove to the statistics displayed, click on the **Select Statistics** dropdown list. You will find a comprehensive list of statistics grouped by type. If you select a large number of statistics to display, the table will extend to the right, and you may have to scroll the window to see them all.

Making a Histogram

Histograms are also created in the EDA menu item. Once you have set up your data, click on the **Histograms** tab.

Histograms



If you have more than one group you are graphing, the app will wrap them as above. If you want to compare the different groups of data on the same axis, make sure that **Display on same axis?** is checked. If your groups are at different scales, unchecking this will graph each group on whatever scale makes sense for it.

You can select which type of plot by clicking on the radio buttons under **Plot Style**.

You can add a normal curve or specifications by checking those boxes.

You can modify the histograms by changing the bin width or the number of bins as well as the center point of a bin.

You can change the size of the histogram by using the slider on the top to change width and the side to change height.

You can customize the title and x-axis label as well as increase the size of the font in the gear dropdown menu.

The slider switch **Show frequency distribution?** Will create a frequency distribution table based on your selections under chart modifications.

If you select the **Kernal Density** plot, different options will appear. In addition to being able to add the **Normal Curve** and **Draw Specs**, you can **Extend the Density** beyond the observed values and **Add a Rug Plot**.

Histograms

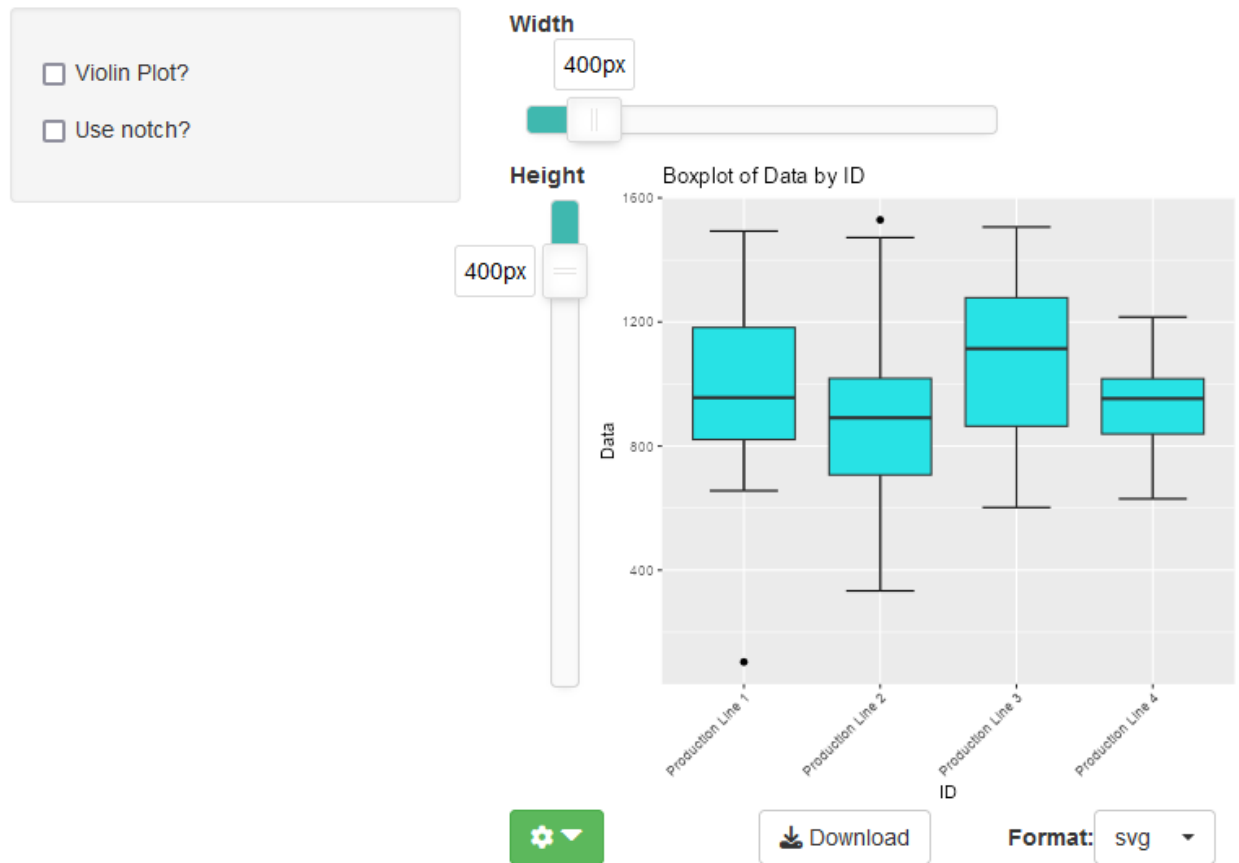


You can select the format and download the graphic.

Making a Boxplot

Boxplots are created in the EDA menu item. Once you have set up your data, click on the **Boxplot** tab.

Boxplots

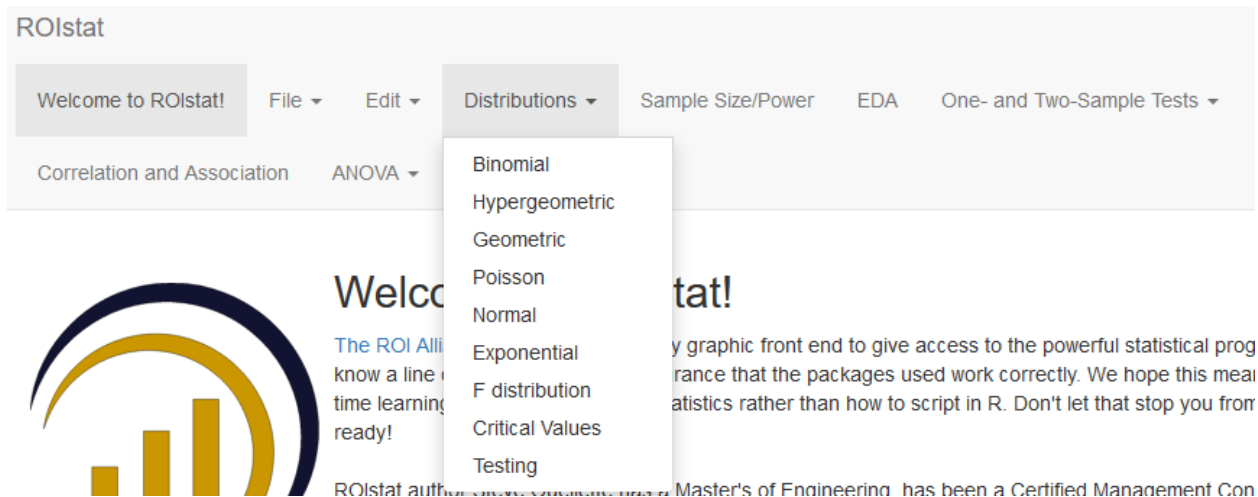


You can select a violin plot or show the 95% confidence interval of the median as a notch by clicking on the check boxes. You can change the width and height by using the sliders.

The gear dropdown will allow you to customize the chart title, x- and y-axis labels and choose a larger font. You can select the format and download the graphic.

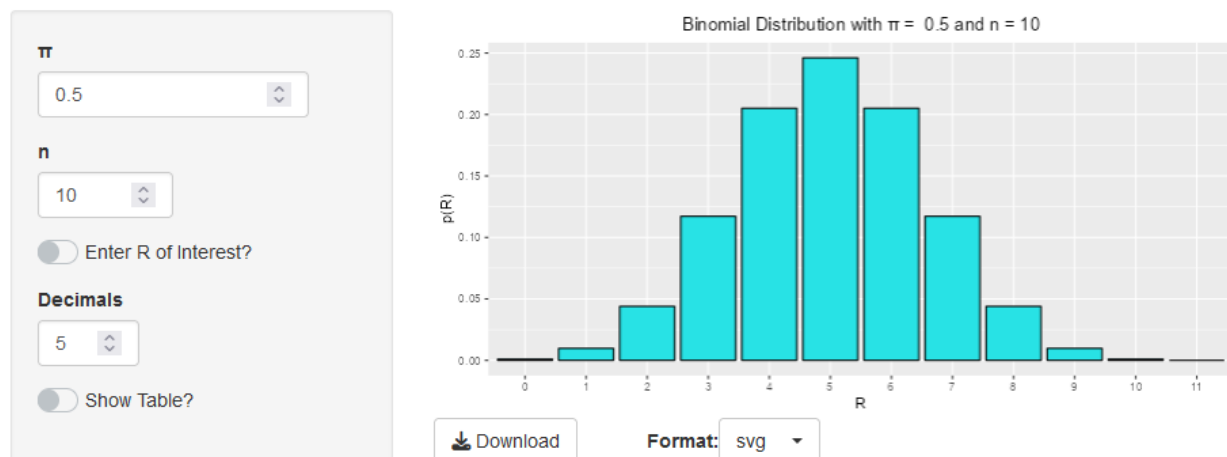
[Chapter 4 – Probability Distributions](#)

You can access various probability distributions and tests on the **Distributions** menu item.



Binomial Distribution

Selecting the binomial distribution shows you this:



You can change the population average π by typing in the box labeled π and the sample size by typing in the box labeled n . The example in the book had a $\pi = 0.45$ and an $n = 10$.

If you are interested in a specific point, choose **Enter R of Interest?** and you can select a particular count. You can choose to show the graph with that R and less or just less than the R you chose. The graph updates with the selection and adds the probabilities of R, R and below, and R and above.

π

n

☒ Enter R of Interest?

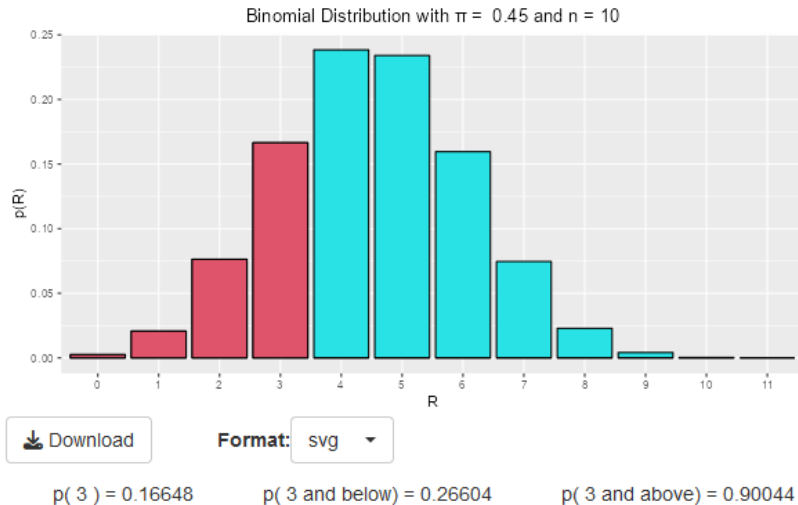
One or Two Tails?
☒ One-Tail
☐ Two-Tails

R

Include R?
☒ \leq
☐ $<$

Decimals

☒ Show Table?



Show entries Search:

	x	p.at.x	eq.and.above	eq.and.below
	0	0.00253	1	0.00253
	1	0.02072	0.99747	0.02326
	2	0.0763	0.97674	0.09956
	3	0.16648	0.90044	0.26604
	4	0.23837	0.73396	0.5044
	5	0.23403	0.4956	0.73844
	6	0.15957	0.26156	0.89801
	7	0.0746	0.10199	0.97261
	8	0.02289	0.02739	0.9955
	9	0.00416	0.0045	0.99966
	10	0.00034	0.00034	1

Showing 1 to 11 of 11 entries Previous Next

Decimals controls the number of decimals displayed here. Clicking **Show Table?** will print out a table with all the probabilities for each R, R and below, and R and above. You can expand the number of entries on this table to the length you need. You can save the graphic in a variety of formats by selecting the format and clicking the Download button.

Poisson Distribution

The Poisson distribution has similar options as the binomial. You create one by selecting **Poisson** on the **Distributions** dropdown menu item.

You can change λ . In the book λ was 4.2. You can Enter X of interest? by clicking on the slider switch. In the book we were interested in $X = 2$. We can display the table by clicking on **Show Table?** You can also control the number of **Decimals**.

λ

4.2

☒ Enter X of Interest?

One or Two Tails?

☒ One-Tail

☐ Two-Tails

X

2

Include X?

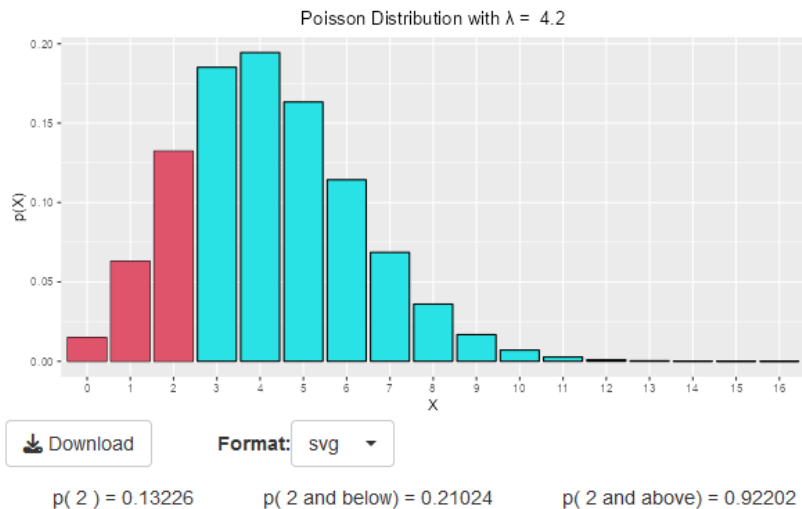
☒ \leq

☐ $<$

Decimals

5

☒ Show Table?



Show 25 entries

Search:

	x	p.at.x	eq.and.above	eq.and.below
	0	0.015	1	0.015
	1	0.06298	0.985	0.07798
	2	0.13226	0.92202	0.21024
	3	0.18517	0.78976	0.3954
	4	0.19442	0.6046	0.58983
	5	0.16332	0.41017	0.75314
	6	0.11432	0.24686	0.86746
	7	0.06859	0.13254	0.93606
	8	0.03601	0.06394	0.97207
	9	0.01681	0.02793	0.98887
	10	0.00706	0.01113	0.99593
	11	0.00269	0.00407	0.99863
	12	0.00094	0.00137	0.99957
	13	0.0003	0.00043	0.99987
	14	0.00009	0.00013	0.99997
	15	0.00003	0.00003	0.99999
	16	0.00001	0.00001	1

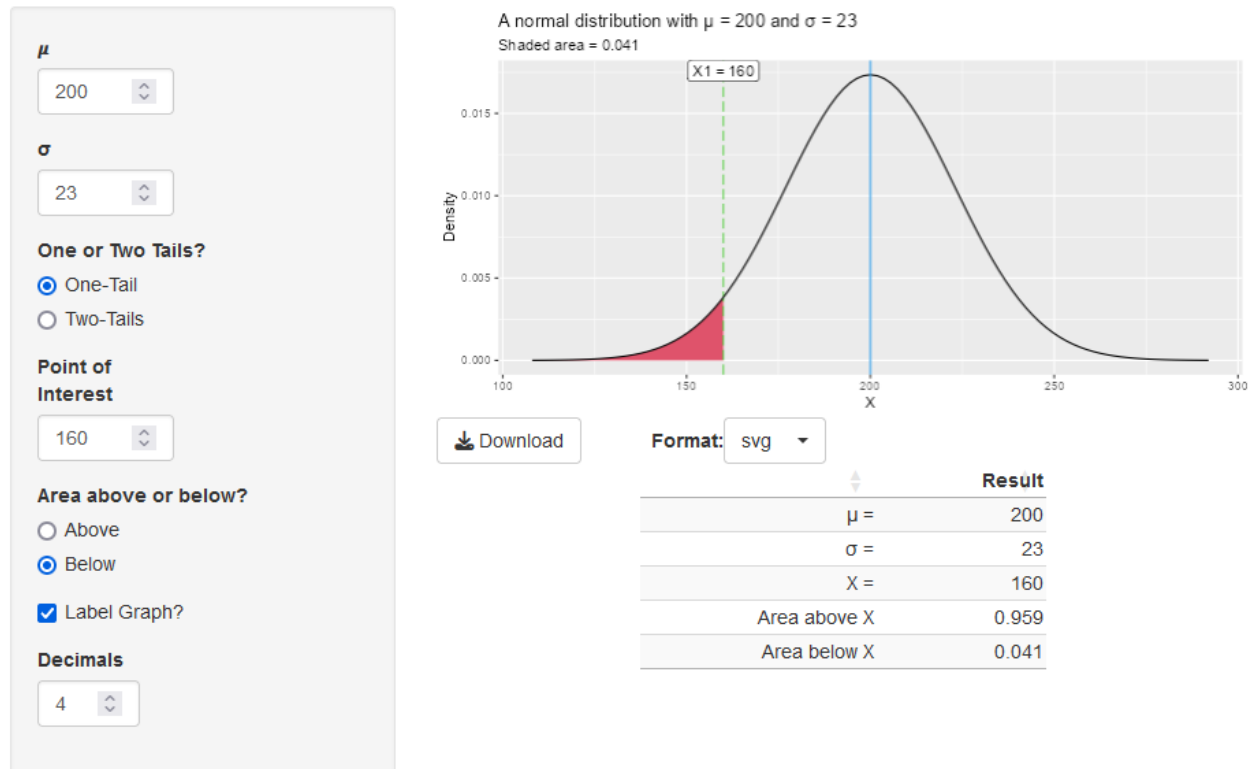
Showing 1 to 17 of 17 entries

Previous 1 Next

Normal Distribution

You can make the normal distribution by selecting **Normal** from the **Distributions** dropdown.

In the book we had a distribution with a $\mu = 200$ and $\sigma = 23$ and were interested in the probability below 160.



The graph can display the area **Above** or **Below** the point of interest. If the label for the point of interest is in the way, you can unselect the **Label Graph?** check box. **Decimals** controls the number of decimals reported. You can save the graphic in a variety of formats.

If you want to know the probability **Inside** or **Outside** of two points of interest, select **Two-Tails** and enter the two points of interest:

μ

σ

One or Two Tails?
☐ One-Tail
☒ Two-Tails

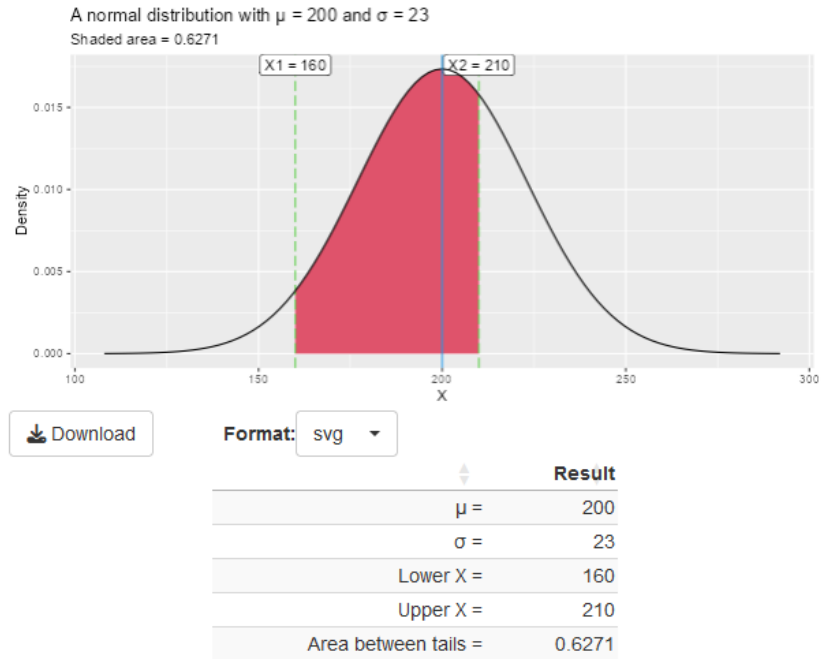
Lower Tail

Upper Tail

Area between or outside of points?
☒ Inside
☐ Outside

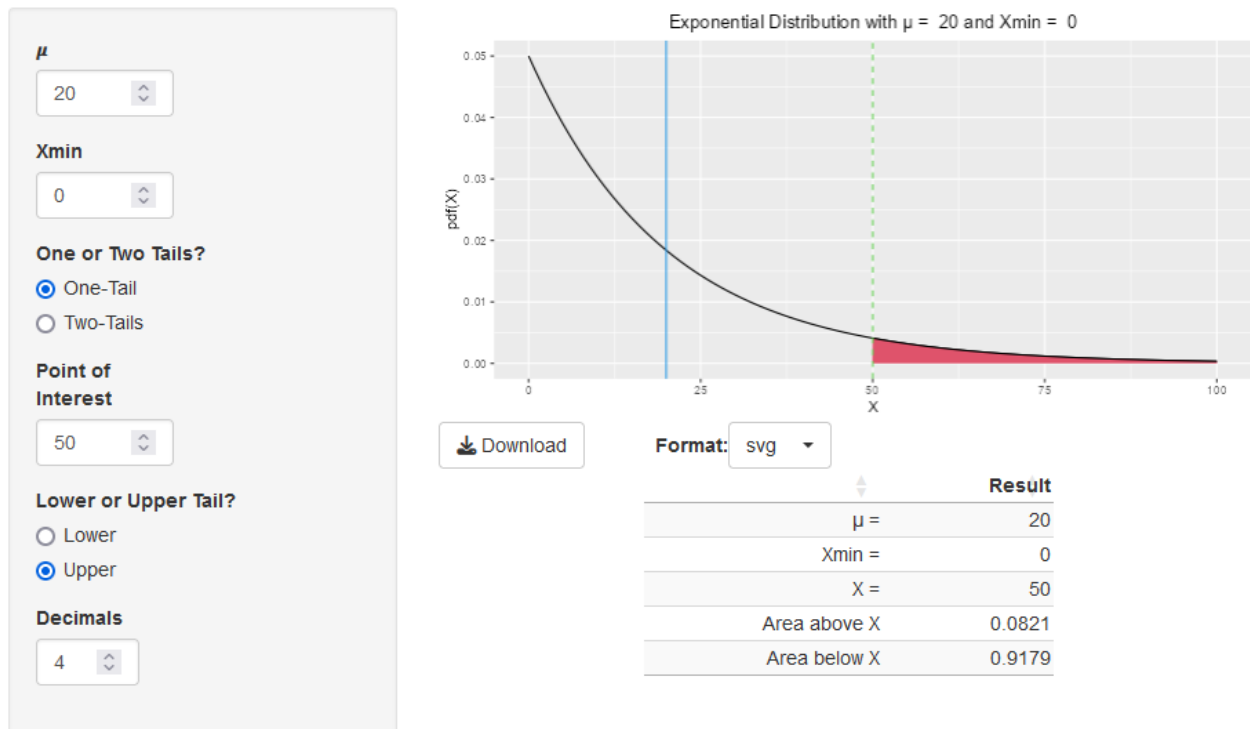
☒ Label Graph?

Decimals



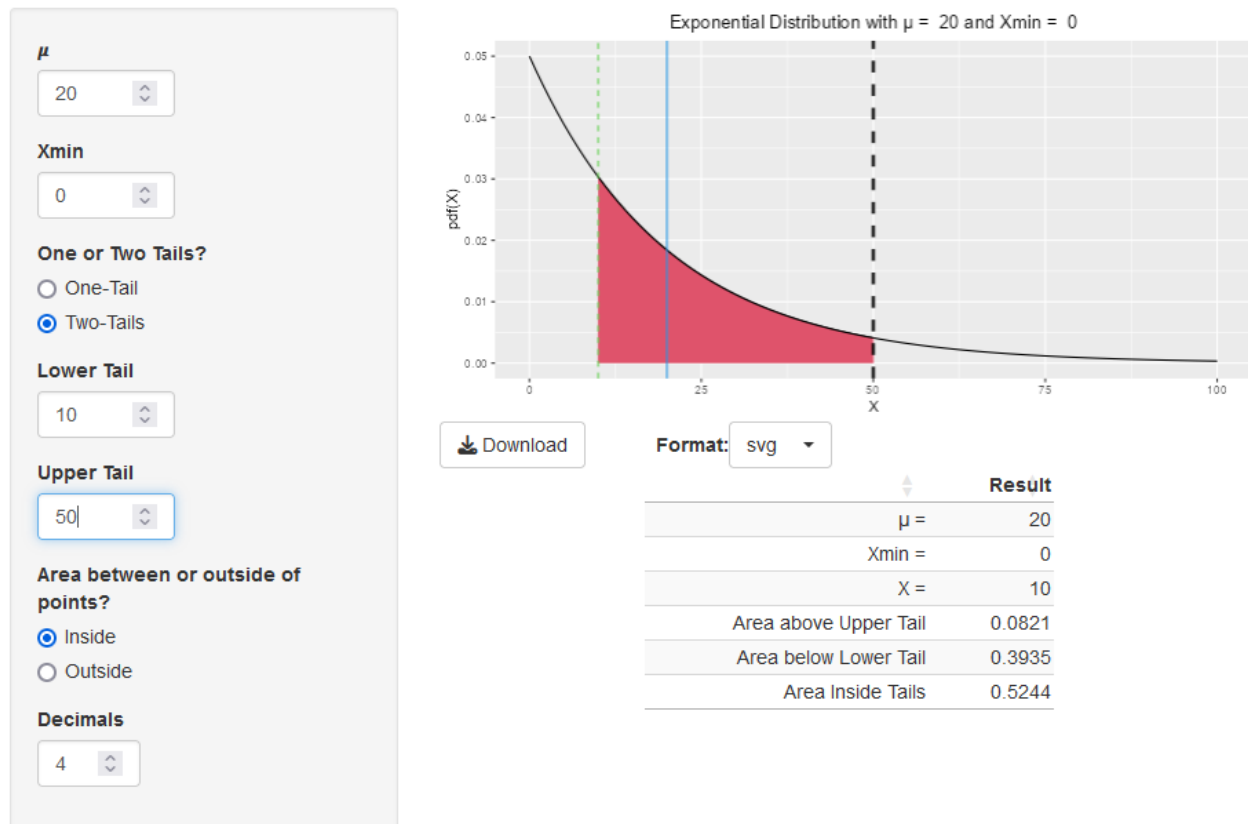
Exponential Distribution

To create an exponential distribution, select **Exponential** from the **Distribution** dropdown menu item. In the book we were interested in the probability for the **Point of Interest** = 50 for an exponential distribution with a $\mu = 20$ and an $X_{min} = 0$.



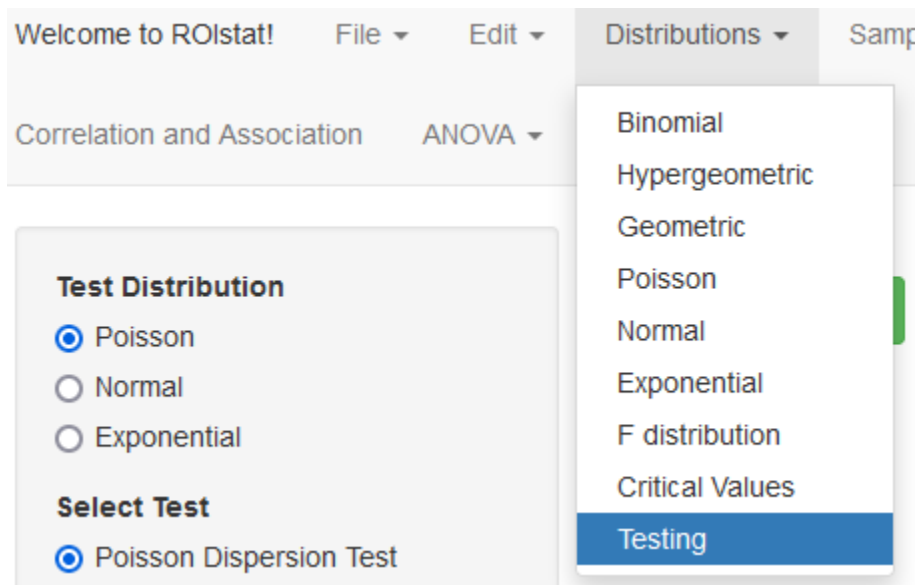
The graph can display the area for the **Lower** or **Upper** tails. You can control the number of **Decimals**.

If you want to know the probability **Inside** or **Outside** of two points of interest, select **Two-Tails** and enter the two points of interest:



Chapter 5 – Distribution Testing

Distribution testing requires that you have data loaded to test. It is located under the **Distribution** dropdown menu.



Testing for Poisson Distributions

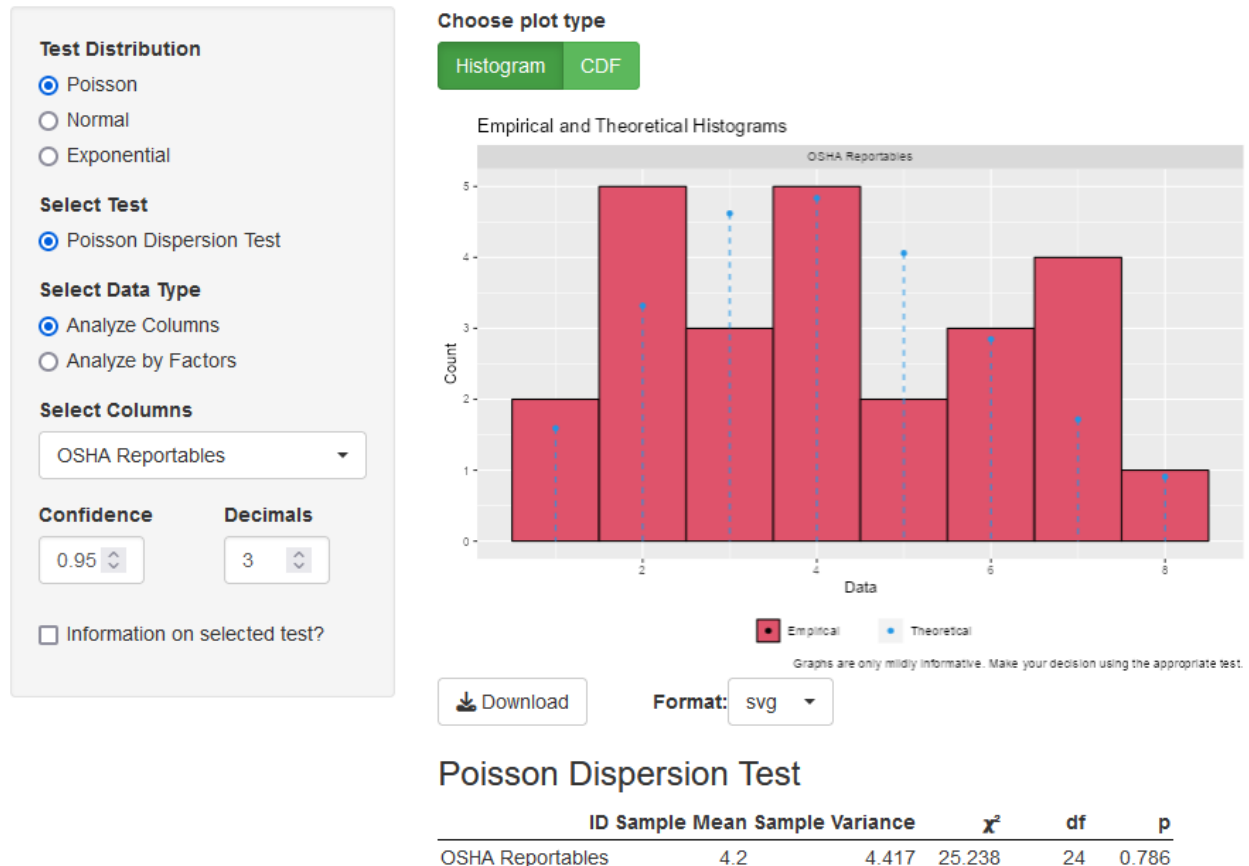
Here is the data for the Poisson testing example:

OSHA
Reportables

6
8
1
7
3
4
4
7
6
7
4
5
5
2
7
2
2
3
1
3
4
6
2
2
4

You can copy and paste this into ROlstat.

Select the Poisson radio button under Test Distribution. You will need to tell the app how the data is configured. In this case, the data is a column, so we Analyze by Columns. Select Columns to get the OSHA data:



Confidence lets you choose the confidence of both the test and any confidence intervals that are reported. **Decimals** controls the number of decimals reported.

You can get a pop-up window for more **Information on the selected test**.

There are two types of plots that you can select. For the Poisson these are a **Histogram** displaying the data and the theoretical distribution and the Cumulative Density Function, or **CDF**.

The results of the test are displayed in a table.

You can download the graphic in a number of formats.

Testing for Normal Distributions

There are a couple of ways to test for normality.

If you are already exploring the data in EDA, you can get them quickly by clicking on the **Normality Test** tab. The example data was [OVEN.CSV](#).

Normality Tests

Confidence
0.95

Decimals
5

☒ Auto Select Normality Test?

Search:

dv.name	n	missing	mean	var	sd	g3.skewness	g3test.p	g4.kurtosis	g4test.p
1 Temperature	25	0	200.59937	506.25402	22.50009	0.63616	0.16072	1.41781	0.13863

Showing 1 to 1 of 1 entries

You can control the **Confidence** of the test and the number of **Decimals** reported. Here you can have the app automatically select the best normality tests depending on the sample size. If you wish to manually control what tests are reported, turn off the **Auto Select Normality Test?** slider. You will then see buttons for the three tests available here.

Normality Tests

Confidence
0.95

Decimals
5

☐ Auto Select Normality Test?

Select Tests

Anderson-Darling (n < 25)

Shapiro-Wilk (n < 25)

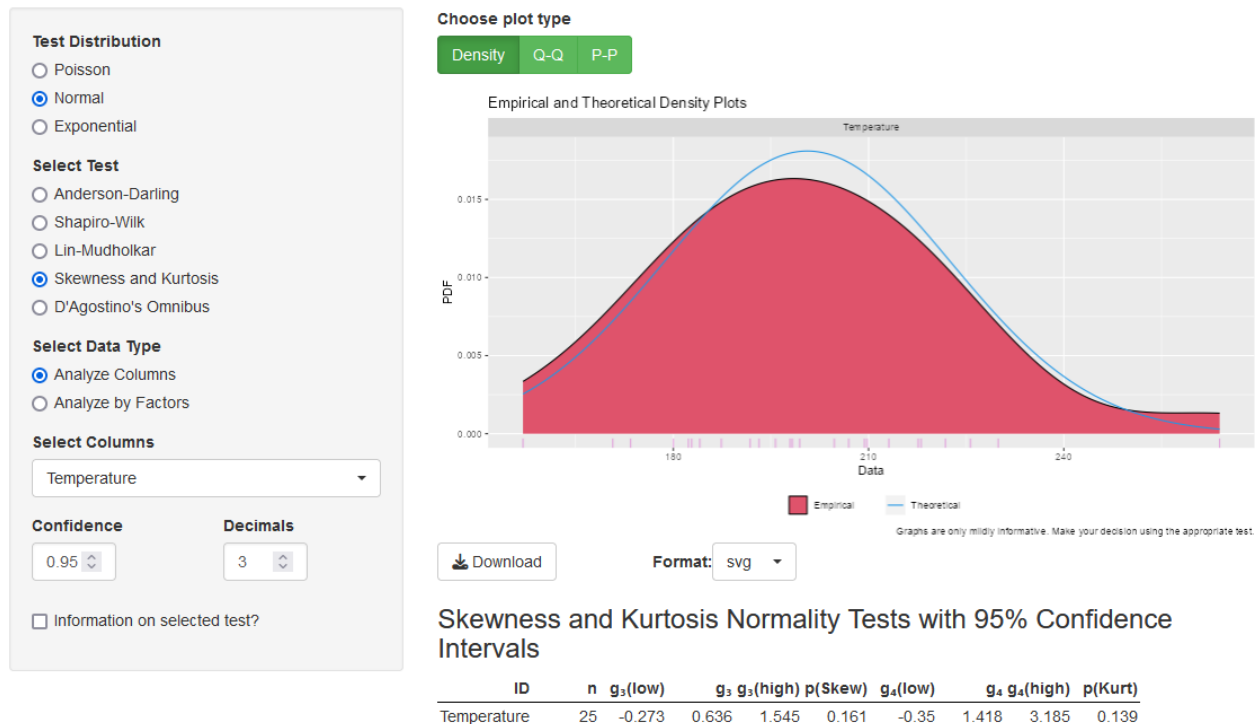
Skewness and Kurtosis (n > 25)

Search:

dv.name	n	missing	mean	var	sd	adtest.AA	adtest.p
1 Temperature	25	0	200.59937	506.25402	22.50009	0.21424	0.85049

Showing 1 to 1 of 1 entries

If you want more options for normality testing, click on **Distributions** and then **Testing**, then select **Normal** under **Test Distribution**. You will see a list of more tests, and each test will give you more information than the ones available under **EDA**. Select how the data is configured and whatever data sets you want to test. The oven temperature data is in a column.



Confidence lets you choose the confidence of both the test and any confidence intervals that are reported. **Decimals** controls the number of decimals reported.

You can get a pop-up window for more **Information on the selected test**.

There are three types of plots that you can select. For the normal these are a **Density** plot displaying the data and the theoretical distribution, the quantile-quantile (**Q-Q**) and probability-probability (**P-P**) plots, showing the data as points and the theoretical distribution as a line. If the data is perfectly normal, you would expect to see all the dots on the line for these two plots.

The results of the test are displayed in a table.

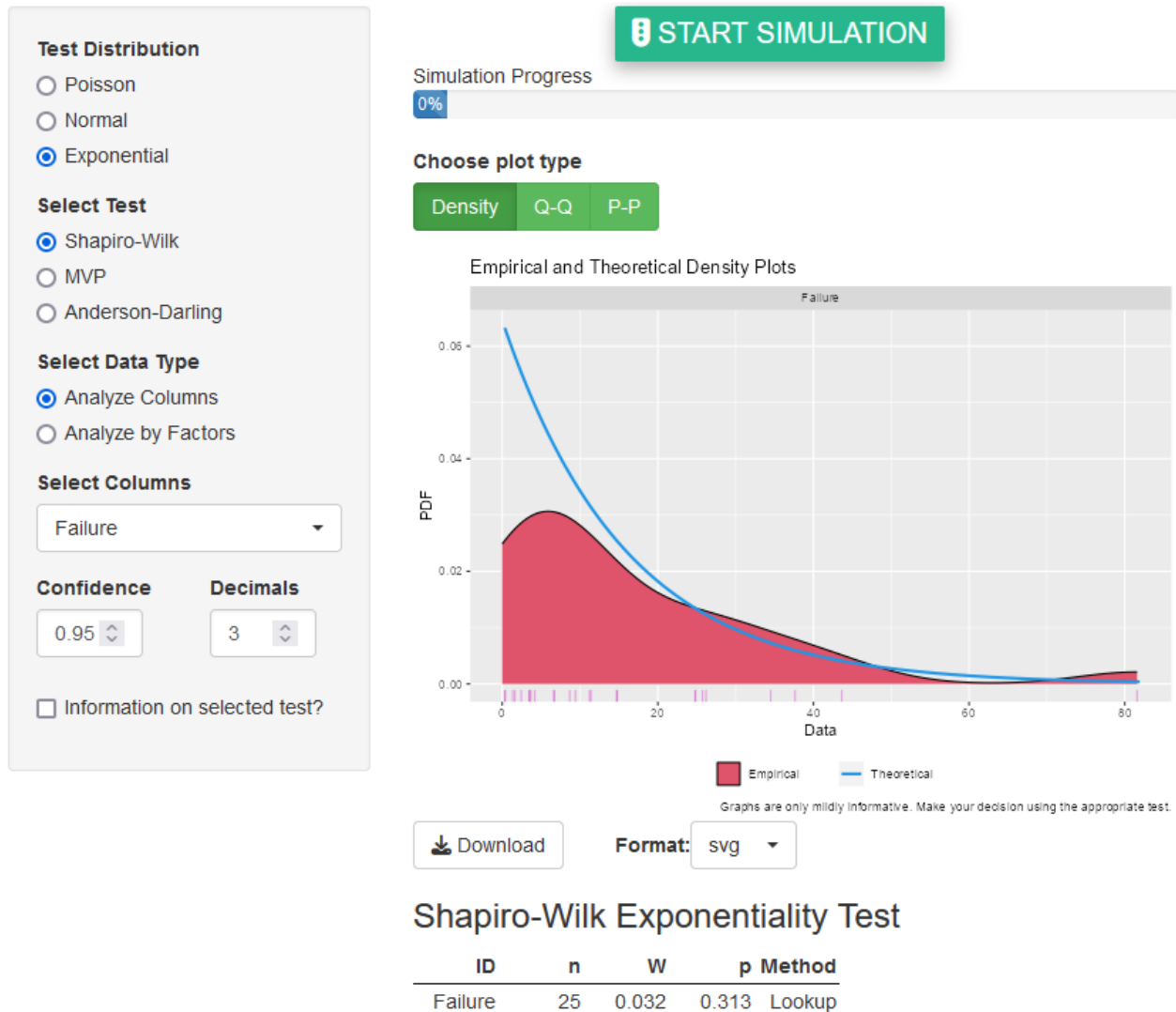
You can download the graphic in a number of formats.

Testing for Exponential Distributions

The exponential tests are located under the **Distributions** dropdown in **Testing**. The example data was [TTF1.csv](#).

Select the **Test Distribution** to be **Exponential**. You will see the available tests under **Select Test**.

The **Shapiro-Wilk** and **MVP** tests are Monte Carlo tests, so they will both have a **Start Simulation** button when you have the data set up and are ready to run the tests. The Time to Failure data is in a column.



Confidence lets you choose the confidence of both the test and any confidence intervals that are reported. **Decimals** controls the number of decimals reported.

You can get a pop-up window for more **Information on the selected test**.

There are three types of plots that you can select. For the exponential these are a **Density** plot displaying the data and the theoretical distribution, the quantile-quantile (**Q-Q**) and probability-probability (**P-P**) plots, showing the data as points and the theoretical distribution as a line. If the data is perfectly exponential, you would expect to see all the dots on the line for these two plots.

The results of the test are displayed in a table.

You can download the graphic in a number of formats.

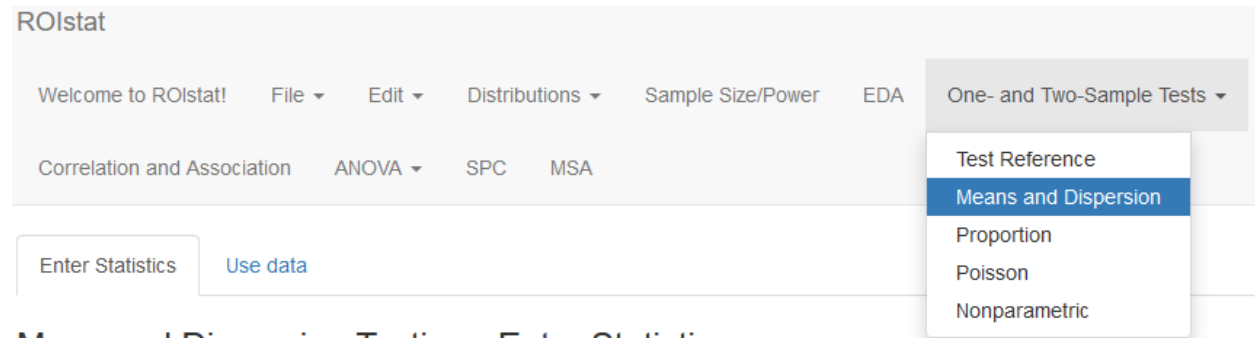
Chapter 6 – Estimation

Central Limit Theorem- Getting a z-Score and Probability for the RSD

The Central Limit Theorem is built into hypothesis testing, so you can get quick results for some of the discussion by using one-sample tests to “trick” the app into doing the work.

In the example we know that the data has an average of 9.237, a standard deviation of 6.552 and a sample size of 30. We are testing to see if the sample could have come from a population with $\mu_0 = 10$.

Select **One- and Two-Sample Tests** from the menu dropdown and then **Means and Dispersion**.



You will **Enter Statistics**, so select that tab.

All we need from the app is the z-score and probability of the RSD so we can choose **One Sample**. Since we used a normal distribution, select **Is σ definitively known?** as **Yes, use z**.

Since we are asking if the sample is lower than the population, select

Alternative Hypothesis for Means

μ of sample is less than μ_0 ▼

Enter the sample statistics and the population average.

Mean and Dispersion Testing - Enter Statistics

One or Two Sample?

☒ One Sample

☐ Two Sample

Is σ definitively known?

☐ No, use t

☒ Yes, use z

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

μ of sample is less than μ_0

Enter Statistics and Parameters

\bar{X} : 9.237 μ_0 : 10

σ : 6.552

n : 30

Results

One-Sample Z Test For Means

$\bar{X} = 9.237$ $\mu_0 = 10$

$\sigma = 6.552$

$n = 30$

95 % confidence interval for μ : 6.8924 to 11.5816

Test for μ of sample is less than μ_0 : **$z = -0.6378$ $p = 0.2618$**

Power to reject the null if the observed difference was real = 15.7 %

You can find the z-score and the p-value at the highlight, with a little rounding difference.

Confidence Intervals

Most tests will give you confidence intervals, so we can again use this to our advantage.

Binomial Confidence Intervals

The example with customer satisfaction is a binomial distribution, so we can use the binomial proportiono test to get confidence intervals. This is under **One- and Two-Sample Tests, Proportion**.

One- and Two-Sample Tests ▾

- Test Reference
- Means and Dispersion
- Proportion**
- Poisson
- Nonparametric

The first scenario is a point estimate of 90% based on 10 interviews. Since we just want the confidence interval, the null population parameter doesn't matter. Just enter the statistic and sample size, along with the desired confidence level, to get the confidence interval.

Proportion Testing - Enter Statistics

One or Two Sample?

☒ One Sample

☐ Two Sample

Confidence

0.95 ▾

Decimals

4 ▾

Alternative hypothesis for proportions

π of sample is not equal to π_0 ▾

Enter Statistics and Parameters

p_1 0.9 ▾ π_0 : 0.5 ▾

n_1 10 ▾

Results

One-Sample Proportion Test (Exact)

$p = 0.9$ $\pi_0 = 0.5$

$n = 10$

95 % confidence interval for π : 0.555 to 0.9975

Exact test for π of sample is not equal to π_0 : $p = 0.0215$ *

Power to reject the null if the observed difference was real = 73.61 %

The second scenario is also 90% but based on a sample size of 1,000:

Proportion Testing - Enter Statistics

One or Two Sample?

☒ One Sample

☐ Two Sample

Confidence

0.95

Decimals

4

Alternative hypothesis for proportions

π of sample is not equal to π_0

Enter Statistics and Parameters

p_1 0.9 π_0 : 0.5

n_1 1000

Results

One-Sample Proportion Test (Exact)

$p = 0.9$ $\pi_0 = 0.5$

$n = 1000$

95 % confidence interval for π : 0.8797 to 0.9179

Exact test for π of sample is not equal to π_0 : $p = 0 *$

Power to reject the null if the observed difference was real = 100 %

Normal Confidence Intervals

We can use the same trick as above to get confidence intervals for a normal distribution. In the example, we are looking for the confidence interval for a normal distribution with an average of 42, a standard deviation of 7 from a sample size of 100. In order to get the normal distribution confidence intervals, you need to select **Yes, use z** for **Is σ definitively known?** Since we only want the confidence interval, it doesn't matter what you put in for the null parameter.

Mean and Dispersion Testing - Enter Statistics

One or Two Sample?

☒ One Sample

☐ Two Sample

Is σ definitively known?

☐ No, use t

☒ Yes, use z

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

μ of sample is not equal to μ_0

Enter Statistics and Parameters

\bar{X} : 42 μ_0 : 10

σ : 7

n : 100

Results

One-Sample Z Test For Means

$\bar{X} = 42$ $\mu_0 = 10$

$\sigma = 7$

$n = 100$

95 % confidence interval for μ : 40.628 to 43.372

Test for μ of sample is not equal to μ_0 : $z = 45.7143$ $p = 0^*$

Power to reject the null if the observed difference was real = 100 %

Chapter 8 – Sample Size

Calculating Power

In the ZapCareer example to calculate power, the effect size is 2, $\alpha = 0.05$, $\sigma = 4.6816$, and $n = 30$. Select Sample Size/Power from the menu.

ROIstat

Welcome to ROIstat! File Edit Distributions **Sample Size/Power** EDA One- and Two-Sample Tests

Correlation and Association ANOVA SPC MSA

This is a problem about the averages, so select **Calculate the sample size or power for: Means**. The example uses the normal distribution, so **Select the Test: One-sample Mean z**. Select the **Power** checkbox and enter the statistics.

Calculate the sample size or power for:

- ☒ Means
- ☐ Variances
- ☐ Proportions (binomial)
- ☐ Rates (Poisson)
- ☐ ANOVA
- ☐ Correlations

Select the Test

- ☒ One-sample Mean z
- ☐ Two-sample Mean z Independent
- ☐ One-sample Mean t Independent
- ☐ Two-sample Mean t equal variance Independent
- ☐ Two-sample Mean t unequal variance Independent
- ☐ Two-sample Mean t Dependent

Alternative is:

Equal to the null ▼

☒ Power

α : 0.05

n : 30

σ : 4.6816

Δ : 2

Power Calculations - One-Sample Mean: z test, σ known Two-Tail

$$\alpha = 0.05$$

$$\beta = 0.352$$

$$n = 30$$

$$\sigma = 4.6816$$

$$\Delta = 2$$

$$\text{Power} = 0.648$$

Calculating Sample Size

To calculate the sample size, we set it up the same way as above, uncheck the **Power** checkbox, and enter the required $\beta = 0.1$.

Calculate the sample size or power for:

☒ Means

☐ Variances

☐ Proportions (binomial)

☐ Rates (Poisson)

☐ ANOVA

☐ Correlations

Select the Test

☒ One-sample Mean z

☐ Two-sample Mean z Independent

☐ One-sample Mean t Independent

☐ Two-sample Mean t equal variance Independent

☐ Two-sample Mean t unequal variance Independent

☐ Two-sample Mean t Dependent

Alternative is:

Equal to the null ▼

☐ Power

α : 0.05 ▼

β : 0.1 ▼

σ : 4.6816 ▼

Δ : 2 ▼

Sample Size Calculations - One-Sample Mean: z test, σ known Two-Tail

$\alpha = 0.05$ $\beta = 0.1$

$\sigma = 4.6816$

$\Delta = 2$

$n_{calc} = 57.5738$

$n = 58$ Power = 0.9

Chapter 10 – One-Sample Hypothesis Tests

One-Sample z-Test for Known σ

The data in the example is [TTF ZAPCAREER.TXT](#). We have already reviewed how to generate a run chart and kernel density plot above.

Once you have the data loaded, go to the One- and Two-Sample Tests dropdown menu and select Means and Dispersion:

ROIstat

Welcome to ROIstat! File ▼ Edit ▼ Distributions ▼ Sample Size/Power EDA One- and Two-Sample Tests ▼

Correlation and Association ANOVA ▼ SPC MSA

Enter Statistics Use data

Test Reference

Means and Dispersion

Proportion

Poisson

Nonparametric

Mean and Dispersion Testing - Enter Statistics

Tell the app you are using data by clicking on the **Use data** tab.

Now you need to tell the app that the **Data is in Columns** under **How is the data configured?**

Select **Time to Fill Position (ZapCareer)** under **Analyze which column(s)?**

Since you only selected one column, ROIstat knows you are performing a one-sample test. We are assuming that the standard deviation is known, so select **Yes, use z** under **Analyze which column(s)?**

The statistics are automatically calculated for the sample, though you can override the calculated standard deviation. Since we believe we know that the standard deviation is 4.6816, enter that for σ_0 .

We are testing against a population average of 8.8428, so enter that as μ_0 . Select an $\alpha = 0.05$ by choosing a **Confidence** of 0.95. Select the number of desired **Decimal** places. It is a non-directional test, so make sure the **Alternative Hypothesis for Means** is μ of sample is not equal to μ_0 .

Mean and Dispersion Testing - Data

How is the data configured?

☒ Data in Columns

☐ Use Reference Column

Analyze which column(s)?

☐ Sample

☒ Time to Fill Position (ZapCareer)

Is σ definitively known?

☐ No, use t

☒ Yes, use z

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

μ of sample is not equal to μ_0

$\bar{X} = 6.4007$

$\sigma_0 : 4.6816$

$n = 58$

$\mu_0 : 3$

Results

One-Sample Z Test For Means

$\bar{X} = 6.4007$ $\mu_0 = 3$

$\sigma = 4.6816$

$n = 58$

95 % confidence interval for $\mu : 5.1959$ to 7.6055

Test for μ of sample is not equal to $\mu_0 : z = 5.5321$ $p = 0^*$

Power to reject the null if the observed difference was real = 99.98 %

One-Sample t -test for Unknown σ

Using the same data, just select **No, use t** under **Is σ definitively known?** You may have to re-enter the null hypothesis mean.

Mean and Dispersion Testing - Data

How is the data configured?

☒ Data in Columns

☐ Use Reference Column

Analyze which column(s)?

☐ Sample

☒ Time to Fill Position (ZapCareer)

Is σ definitively known?

☒ No, use t

☐ Yes, use z

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

μ of sample is not equal to μ_0

$\bar{X} = 6.4007$

$s = 3.8124$

$n = 58$

Results

One-Sample t Test For Means

$\bar{X} = 6.4007$ $\mu_0 = 8.8428$

$s = 3.8124$

$n = 58$

95 % confidence interval for μ : 5.3983 to 7.4031

σ : 3.2231 to 4.6676

Test for μ of sample is not equal to μ_0 : $t = -4.8784$ $df = 57$ $p = 0 *$

Power to reject the null if the observed difference was real = 99.77 %

☐ One-sample test for variance?

μ_0 : 8.8428

One-Sample χ^2 Test for Variance

This is still [TTF ZAPCAREER.TXT](#), so it is set up as above, but you click the **One-sample test for variance?** checkbox. You will need to enter 4.6816 for the null σ_0 . It is a non-directional test, so select the **Alternative Hypothesis for Means** is σ of sample is not equal to σ_0 . Note that you will not see the checkbox if you have the z-test selected, since you would only use the z-test if you already know the true variance.

Mean and Dispersion Testing - Data

How is the data configured?

☒ Data in Columns

☐ Use Reference Column

Analyze which column(s)?

☐ Sample

☒ Time to Fill Position (ZapCareer)

Is σ definitively known?

☒ No, use t

☐ Yes, use z

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

σ of sample is not equal to σ_0

$s = 3.8124$

$n = 58$

Results

One-Sample Chi-Square Test For Variance

$s = 3.8125$ $\sigma_0 = 4.6816$

$n = 58$

95 % confidence interval for σ : 3.2231 to 4.6676

Test for σ of sample is not equal to σ_0 : $\chi^2 = 37.8002$ $df = 57$ $p = 0.0468 *$

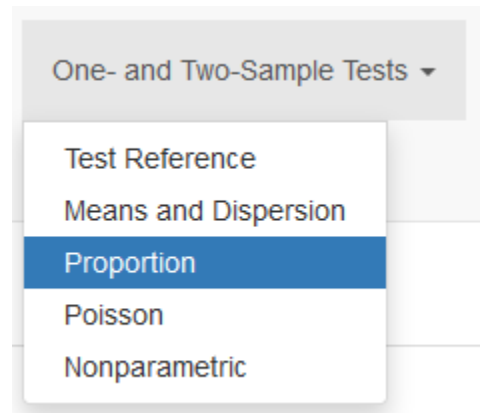
Power to reject the null if the observed difference was real = 53.76 %

☒ One-sample test for variance?

σ_0 : 4.6816

One-Sample Proportion Test

Choose **Proportion** on the **One- and Two-Sample Tests** menu dropdown.



The example in the book just uses statistics, so select the **Enter Statistics** tab. This is a **One Sample** test. Enter $p = 0.0667$, $n = 30$, $\pi_0 = 0.5$ with an $\alpha = 0.05$ indicated by **Confidence** = 0.95. This is a non-directional test, so make sure that the **Alternative hypothesis for proportions** is **π of samples is not equal to π_0** .

Proportion Testing - Enter Statistics

One or Two Sample?

☒ One Sample
☐ Two Sample

Confidence

0.95

Decimals

4

Alternative hypothesis for proportions

π of sample is not equal to π_0

Enter Statistics and Parameters

p_1 0.0667 π_0 : 0.5

n_1 30

Results

One-Sample Proportion Test (Exact)

$p = 0.0667$ $\pi_0 = 0.5$

$n = 30$

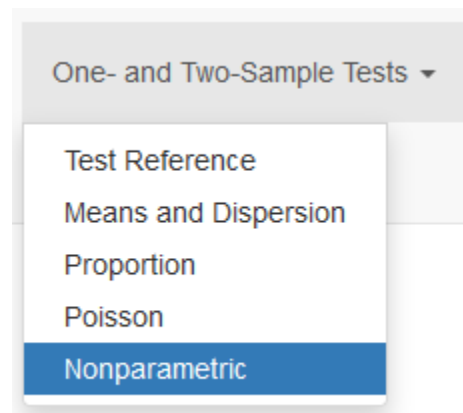
95 % confidence interval for π : 0.0082 to 0.2207

Exact test for π of sample is not equal to π_0 : $p = 0^*$

Power to reject the null if the observed difference was real = 100 %

One-Sample Sign Test for Location

This test is a nonparametric test, which just means that there are few assumptions. You can find it under **One- and Two-Sample Tests** then select **Nonparametric Tests**.



The example in the book uses counts, so make sure the **Enter Statistics** tab is selected and then you can just enter the counts and the null hypothesis directly into the boxes. Choose the **Sign Test for Location** under **Select Test**. The problem calls for an $\alpha = 0.1$, so enter a **Confidence** of 0.9. You can get **More information about the selected test** by checking the box. This is a non-directional test, so make sure the **Alternative Hypothesis** is **M of sample is not equal to M_0** .

Nonparametric Tests

Enter Statistics

Use Data

Confidence
0.9

One Sample or Two?
One-Sample

Select Test
☒ Sign Test for Location
☐ Wilcoxon Test for Location

Decimals
4

☐ More information about selected test

Alternative Hypothesis
M of sample is not equal to M_0

n_{above} : 59

n_{equal} : 20

n_{below} : 21

Results
Sign Test for Location
Method: One-Sample Proportion Test (Exact)
 $n_{above} = 59$
 $n_{equal} = 20$
 $n_{below} = 21$
 $n_{observed} = 100$
 $p_{below} = 0.2625$
 $n_{included} = 80$
Test for proportion below = 0.5
90 % confidence interval for π_{below} : 0.1831 to 0.3556
Test for M of sample is not equal to M_0 : $p_{below} = 0.2625$ $p = 0 *$
Power to reject the null if the observed difference was real = 0.9975

One-Sample Exact Poisson Test

Choose **Poisson** under the **One- and Two-Sample Tests** menu dropdown.

One- and Two-Sample Tests

Test Reference

Means and Dispersion

Proportion

Poisson

Nonparametric

This is a **One Sample** test. You can select **Confidence** and **Decimals**. The example in the book used counts, so select the **Enter Statistics** tab and enter the counts and null hypothesis into the boxes.

Poisson Rate Testing - Enter Statistics

One or Two Sample?

☒ One Sample
☐ Two Sample

Confidence

0.95

Decimals

4

Alternative hypothesis for rates

λ of sample is not equal to λ_0

Enter Statistics and Parameters

c_1 : 30 λ_0 : 8.6

n_1 : 6

Results

Exact Poisson test

$c = 30$ $\lambda_0 = 8.6$
 $n = 6$

95 % confidence interval for λ : 3.3735 to 7.1378

Exact test for λ of sample is not equal to λ_0 : $p = 0.0017$ *

Chapter 11 – Correlation, Regression, and Association

Calculating r and One-Sample Test for Correlation, $\rho = 0$

The example uses the data file [FACESPACEADS.CSV](#). After that is loaded, click on the **Correlation and Association** menu item.

ROIstat

Welcome to ROIstat! File ▼ Edit ▼ Distributions ▼ Sample Size/Power EDA One- and Two-Sample Tests ▼

Correlation and Association ANOVA ▼ SPC MSA

Select the **Use Data** tab. This is a **One-Sample** test. Choose **Pearson r** under **Select Test**. You can adjust the **Confidence** and **Decimals** and get **More information about the selected test**.

Since different measures of association have different requirements for how you set up your data, there is also a checkbox to learn about the **Data Structure Requirements** for the selected test.

You then select which column you want for the **X Data** and **Y Data**. The Pearson r will be calculated and displayed along with the One-Sample Test for Correlation. In the example, we are testing to see if there is

any correlation at all, so make sure that $\rho_0 = 0$. It is a non-directional test, so make sure the **Alternative Hypothesis** is ρ of sample is not equal to ρ_0 .

Correlation and Association

ANOVA ▾SPCMSA

Enter Statistics

Use Data

Scatterplot

Correlation and Association

Confidence

0.95 ▾

One Sample or Two?

One-Sample ▾

Select Test

☒ Pearson r

☐ Spearman Rank r_s

☐ Kendall's τ

☐ ϕ / Cramér's V

☐ Biserial

☐ Point Biserial

☐ Yule's Q

☐ Goodman and Kruskal's γ

☐ Tetrachoric Correlation

☐ Cohen's κ

☐ Kendall's Concordance

☐ J-Index of Predictive Efficiency

Decimals

4 ▾

☐ More information about selected test

☐ Data Structure Requirements

X Data

Ad Spend ▾

Y Data

Revenue ▾

Alternative Hypothesis

ρ of sample is not equal to ρ_0 ▾

ρ_0 :

0 ▾

Results

One-Sample Test for Pearson Product Moment Correlation

$r = 0.5697$

$\rho_0 = 0$

$r^2 = 0.3246$

$n = 23$

95 % confidence interval for ρ : 0.2059 to 0.7952

Test for ρ of sample is not equal to ρ_0 : $t = 3.1769$ $df = 21$ $p = 0.0045$ *

Power to reject the null if the observed difference was real = 82.49 %

Note that the coefficient of determination, r^2 , is also displayed in the output.

One-Sample Approximate Test for Correlation, $\rho = X$

This test only requires that you change the null parameter from 0 to whatever null you are testing. In the example, the $\rho_0 = 0.8$. Again, it is a non-directional test, so make sure the **Alternative Hypothesis** is ρ of sample is not equal to ρ_0 .

Correlation and Association

Confidence
0.95

One Sample or Two?
One-Sample

Select Test
☒ Pearson r
☐ Spearman Rank r_s
☐ Kendall's τ
☐ ϕ / Cramér's V
☐ Biserial
☐ Point Biserial
☐ Yule's Q
☐ Goodman and Kruskal's γ
☐ Tetrachoric Correlation
☐ Cohen's κ
☐ Kendall's Concordance
☐ J-Index of Predictive Efficiency

Decimals
4

☐ More information about selected test
☐ Data Structure Requirements

X Data
Ad Spend

Y Data
Revenue

Alternative Hypothesis
 ρ of sample is not equal to ρ_0

ρ_0 : 0.8

Results
One-Sample Test for Pearson Product Moment Correlation
 $r = 0.5697$ $\rho_0 = 0.8$
 $r^2 = 0.3246$
 $n = 23$
95 % confidence interval for ρ : 0.2059 to 0.7952
Test for ρ of sample is not equal to ρ_0 : $z = -2.0191$ $p = 0.0435$ *
Power to reject the null if the observed difference was real = 52.36 %

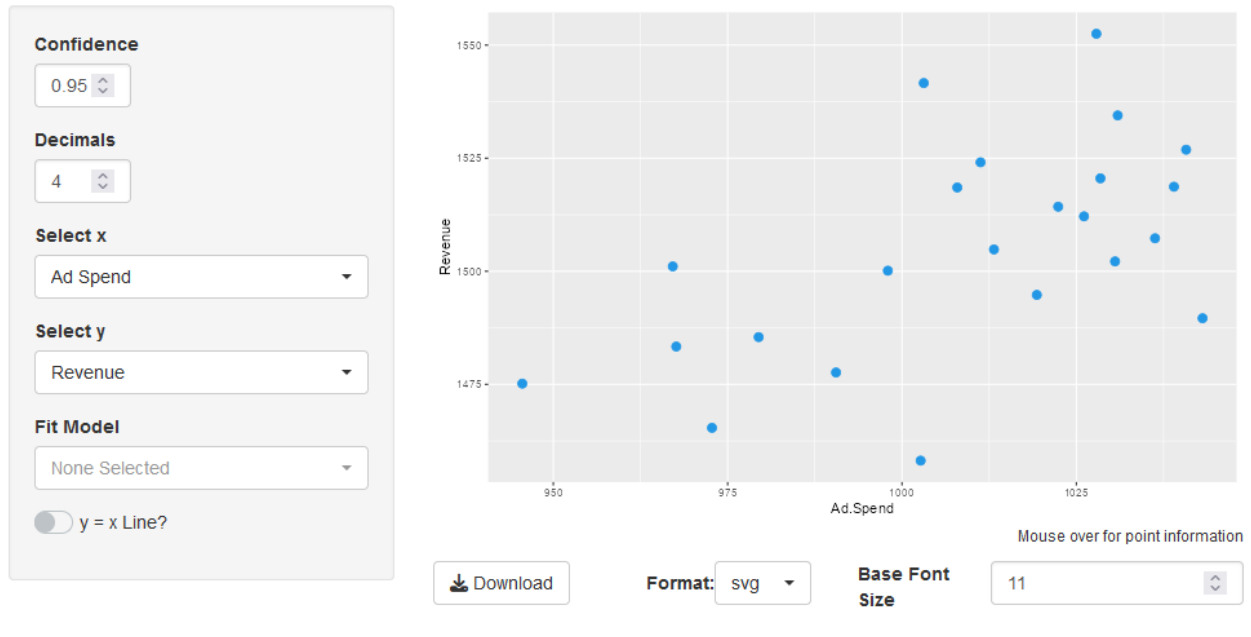
Linear Regression

Making a Scatterplot

The example in the book again uses the data file [FACESPACEADS.csv](#). Once the data is loaded select **Correlation and Association** from the menu, then the **Scatterplot** tab. You need to **Select X** and **Select Y**. For regression, it matters which one you select for each, so make sure you select **Ad Spend** for x and **Revenue** for y.

[Enter Statistics](#)[Use Data](#)[Scatterplot](#)

Scatterplot

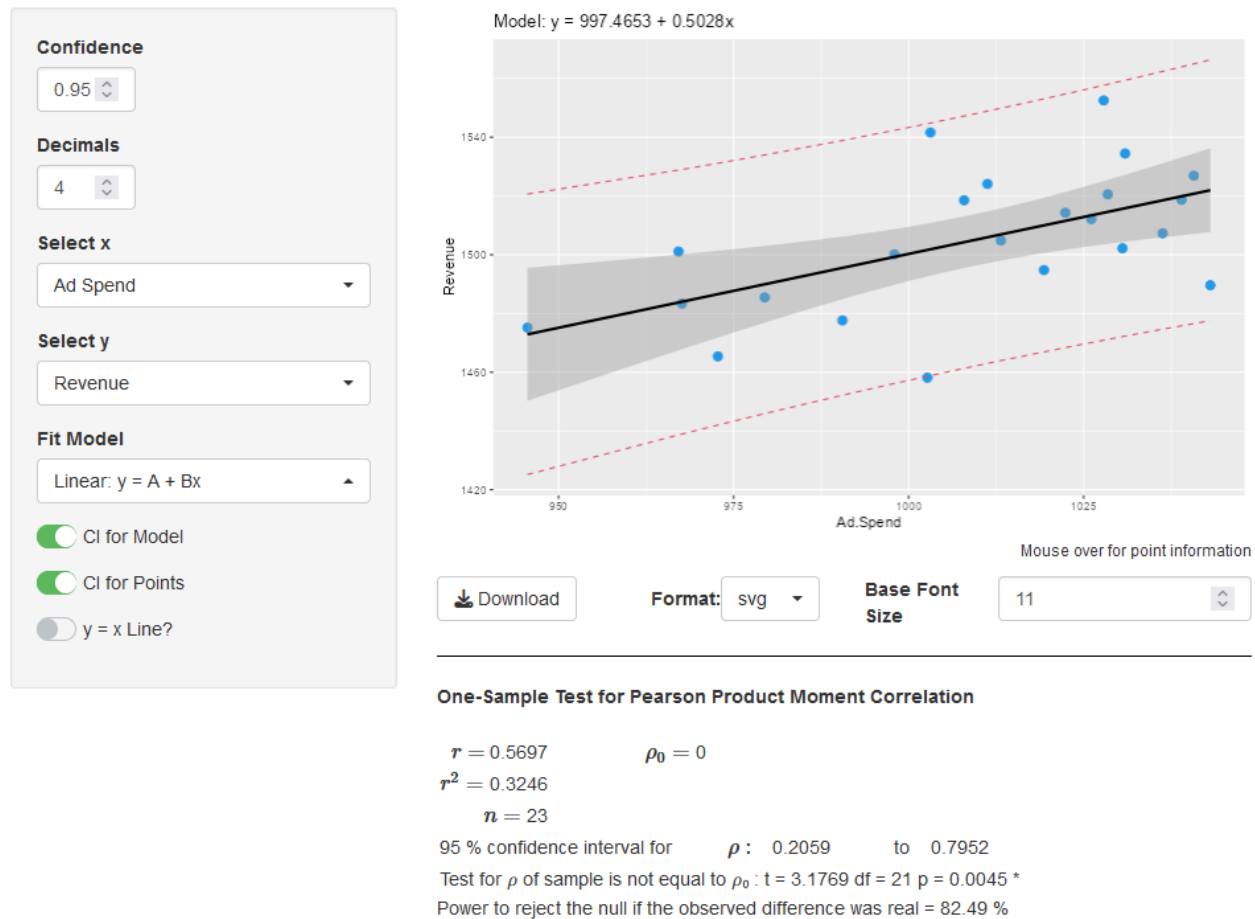


You can select **Confidence** and **Decimals**. You can adjust the **Base Font Size** and **Download** the graph in a variety of **Formats**. You can add a **y = x Line** by clicking on the slider. You can get information for each point by hovering the mouse pointer over it.

Adding a Regression Line

Once the scatterplot is set up, you can add a regression line for a variety of **Fit Models**. The example in the book is a simple linear model. Selecting a model opens up new options where you can turn **CI for Model** and **CI for Points** on and off. Selecting a model also creates a One-Sample Pearson r test for $\rho_0 = 0$ for that model.

Scatterplot



Correlation of Ordinal Data – Spearmans r_s

The example in the book uses [SWEAT.CSV](#). Once that is loaded, select the **Correlation and Association** menu item. We want to **Use Data**, so select that tab.

This is a **One-Sample** test. Choose the **Spearman Rank r_s** test under **Select Test**. Then select the two columns **Scoville** and **Sweat**. You can select the **Confidence** and **Decimals**, and get **More information about the selected test** and **Data Structure Requirements**. It is a non-directional test, so make sure the **Alternative Hypothesis** is ρ_s is not equal to 0.

Correlation and Association

Confidence
0.95

One Sample or Two?
One-Sample

Select Test
☐ Pearson r
☒ Spearman Rank r_s
☐ Kendall's τ
☐ ϕ / Cramér's V
☐ Biserial
☐ Point Biserial
☐ Yule's Q
☐ Goodman and Kruskal's γ
☐ Tetrachoric Correlation
☐ Cohen's κ
☐ Kendall's Concordance
☐ J-Index of Predictive Efficiency

Decimals
4

☐ More information about selected test
☐ Data Structure Requirements

X₁ Data
Scoville

X₂ Data
Sweat

Alternative Hypothesis
 ρ_s is not equal to 0

Results
One-Sample Spearman Rank Correlation Coefficient Test
 $r_s = 0.8707$ $\rho_s = 0$
 $r_s^2 = 0.7581$
 $n = 20$
95 % confidence interval for ρ_s : 0.6514 to 0.9557
Test for ρ_s is not equal to 0 : t = 7.5103 df = 18 p = 0 *

Association of Nominal Data- Cramér's V

This test is also located in the **Correlation and Association** menu item. The example in the book uses counts, so select the **Enter Statistics** tab.

Choose the **ϕ / Cramers V** radio button under **Select Test**. This will change the data input area to show four boxes. This is where you enter the counts from the contingency table. It is a non-directional test, so make sure the **Alternative Hypothesis** is **ϕ of sample is not equal to 0**.

You can select the **Confidence** or get **More information about the selected test**.

Correlation and Association

Confidence

0.95

One Sample or Two?

One-Sample

Select Test

☐ Pearson r
☐ Spearman Rank r_s
☐ Kendall's τ
☒ ϕ / Cramers V

Decimals

4

☐ More information about selected test

Alternative Hypothesis

ϕ is not equal to 0

n_{11} :

12

n_{21} :

18

n_{12} :

47

n_{22} :

46

Results

Cramer's V and Fisher Exact Test

$n_{11} = 12$ $n_{21} = 18$
 $n_{12} = 47$ $n_{22} = 46$
 $V = 0.0906$ $\chi^2 = 1.0092$
Test for ϕ is not equal to 0 : p = 0.4014

To analyze a contingency table that is more than 2 x 2, you will need to use data, as in the exercise. The data can be set up in short form, where each combination is listed along with a count, or long form, where each line is one count.

To set up the short form, select **Frequency** for **How is your data configured?** Then choose whichever column you want for the **Row Data** and **Column Data**. The **Weight** is the count.

How is your data configured?

☐ Independent ☒ Frequency

Row Data

Line

Column Data

Scrap Reason

Weight

Count

If you use the long form, select Independent for **How is your data configured?** Then choose whichever column you want for the **Row Data** and **Column Data**.

How is your data configured?

☒ Independent ☐ Frequency

Row Data

Line

Column Data

Scrap Reason

ROIstat handles both numeric and text descriptions for your contingency tables.

Chapter 12 – Two-Sample Hypothesis Tests

Independent Tests

Two-Sample Independent Test for the Average and Dispersion

The example in the book uses the data file [ZAP_V_ACTUALLY.CSV](#). Once that is loaded, click on the **One- and Two-Sample Tests** menu dropdown, and then **Means and Dispersion**. We will Use Data, so select that tab.

Each vendor is in a column, so select **Data in Columns** for **How is the data configured?** Then click the checkbox next to each column under **Analyze which column(s)?** The top column selected is Group 1 and the lowest column selected is Group 2.

By selecting two columns, ROlstat knows you are doing a two-sample test. Make sure that the **Independent** test is selected under **Independent or Dependent Test?** and that the **Alternative Hypothesis for Means** is μ_1 is not equal to μ_2 .

We will only use the t-test with real data, so under **Is σ definitively known?** Choose **No, use t**. We recommend that you **Select** the **Welch** test.

You can set your **Confidence** and **Decimals**.

Mean and Dispersion Testing - Data

How is the data configured?
☒ Data in Columns
☐ Use Reference Column

Analyze which column(s)?
☒ ZapCareer
☒ Actually

Is σ definitively known?
☒ No, use t
☐ Yes, use z

Select:
☒ Welch (recommended)
☐ Fisher (unknown but equal variance)
☐ Choose based on variance test (not recommended)

Confidence

Decimals

Alternative Hypothesis for Means

Independent or Dependent Test?

$\bar{X}_1 = 6.6346$
 $s_1 = 4.7807$
 $n_1 = 25$

$\bar{X}_2 = 4.088$
 $s_2 = 2.47$
 $n_2 = 25$

Results
Two-Sample t Test For Means (Unequal Variances)

$\bar{X}_1 = 6.6346$
 $s_1 = 4.7807$
 $n_1 = 25$

$\bar{X}_2 = 4.088$
 $s_2 = 2.47$
 $n_2 = 25$

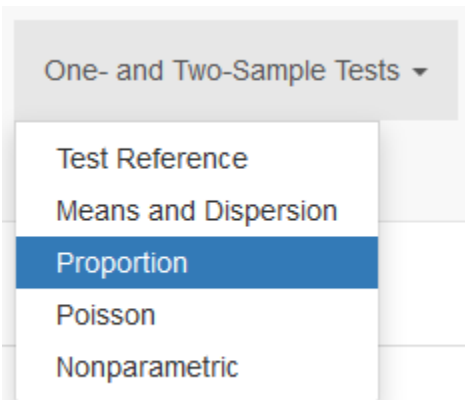
95 % confidence interval for
 μ_1 : 4.6612 to 8.6079
 μ_2 : 3.0685 to 5.1076
 σ_1 : 3.7329 to 6.6507
 σ_2 : 1.9286 to 3.4361
 $\mu_1 - \mu_2$: 0.3638 to 4.7293

Tests for $\sigma_1^2 = \sigma_2^2$:
F-test F = 3.7462 df = 24 / 24 p = 0.002 *
Levene's Test (ADA) t = 3.2112 df = 36.4458 p = 0.0028 *
ADM $n-1$ t = 2.7918 df = 32.4811 p = 0.0087 *
Test for μ_1 is not equal to μ_2 : t = 2.3662 df = 35.9606 p = 0.0235 *

The output gives you the results from a number of different dispersion tests.

Two-Sample Independent Proportion Test

Click on **Proportion** under the **One- and Two-Sample Tests** menu dropdown.



The problem in the book enters the proportions, so make sure the **Enter Statistics** tab is selected. This is a **Two Sample** test. You can adjust the **Confidence** and **Decimals**.

Proportion Testing - Enter Statistics

One or Two Sample?

☐ One Sample

☒ Two Sample

Confidence

0.95

Decimals

4

Alternative hypothesis for proportions

π_1 is not equal to π_2

Enter Statistics and Parameters

p_1 0.07 p_2 0.01

n_1 100 n_2 100

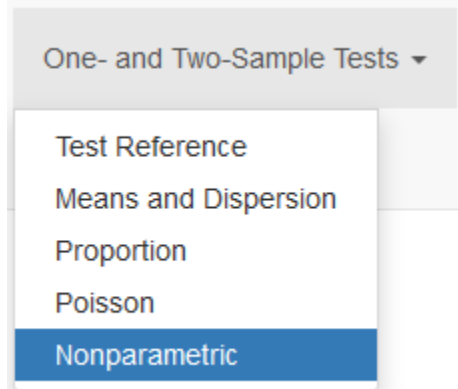
Results

Two-Sample Proportion Test - Fisher Exact Test

$p_1 = 0.07$	$p_2 = 0.01$		
$n_1 = 100$	$n_2 = 100$		
95 % confidence interval for π_1 :		0.0286	to 0.1389
		π_2 :	3e-04 to 0.0545
Exact test for π_1 is not equal to π_2 : p = 0.0649			

Wilcoxon-Mann-Whitney Test

The example uses the data set **CHILIHEAD.DAT**. Once the data is loaded, select the **Nonparametric** tests under the **One- and Two-Sample Tests** menu dropdown.



Since we have data, select the Use Data tab. In this example, the data is set up to have one column with the product name and one column with the rating, so select **Use Reference Column** under **How is the data configured?**

You then need to **Select the Factor** column and **Select the Data** column. Select **Group 1** as **Gut Bomb** and **Group 2** as **The Reaper** for the analysis. Under **Select Test** click on the **Wilcoxon-Mann-Whitney U**. You can adjust the **Confidence** and **Decimals** as needed. This calculation can take a while if there is a large data set, so to get the analysis started click on the **Start Calculation** button.

Nonparametric Tests

[Enter Statistics](#) [Use Data](#)

Confidence

0.95

How is the data configured?

☐ Data in Columns

☒ Use Reference Column

Select Factor

Recipe

Select Data

Rating

Group 1

Gut Bomb

Group 2

The Reaper

Independent Only

Independent

Select Test

☒ Wilcoxon-Mann-Whitney U

☐ Two-Sample Median

Decimals

4

☐ More information about selected test

Alternative Hypothesis

M₁ is not equal to M₂

$\widetilde{X}_1 = 4$

$n_1 = 15$

$\widetilde{X}_2 = 4$

$n_2 = 15$

Results

Method: Mann-Whitney U Test

$S_1 = 281.5$ $S_2 = 183.5$

$n_1 = 15$ $n_2 = 15$

95 % confidence selected

Test for M₁ is not equal to M₂ : U = 63.5 p = 0.0408 *

This calculation can take a long time with large sample sizes

START CALCULATION

Two-Sample Independent Test for Rates (Poisson)

This example uses the data file [BURNT_CHIPS.CSV](#). Once that is loaded, select **Poisson** under the **One- and Two-Sample Tests** menu dropdown.

One- and Two-Sample Tests ▾

Test Reference
 Means and Dispersion
 Proportion
Poisson
 Nonparametric

Since we have a data set, select the **Use Data** tab. The **Data is in Columns**, so select that for **How is the data configured?** Select the **old** and **new** columns by clicking on the checkboxes next to them. By clicking on two columns, ROIstat knows you are doing a two-sample test. You can adjust the **Confidence** and **Decimals**. This is a non-directional test, so make sure the **Alternative hypothesis for rates** is λ_1 is not equal to λ_2 .

Poisson Rate Testing - Use Data

How is the data configured?
☒ Data in Columns
☐ Use Reference Column
Analyze which column(s)?
☐ V1
☒ old
☒ new
Confidence

Decimals

Alternative hypothesis for rates

Statistics from Data
 $\lambda_1 = 9.945$
 $n_1 = 200$

$\lambda_2 = 2.215$
 $n_2 = 200$

Results
Two-Sample Poisson Test
 $c_1 = 1989$
 $n_1 = 200$

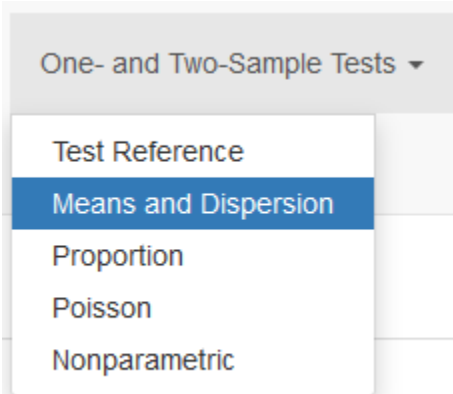
$c_2 = 443$
 $n_2 = 200$

95 % confidence interval for λ_1 : 9.5127 to 10.3919
 λ_2 : 2.0135 to 2.4312
 Exact test for λ_1 is not equal to λ_2 : p = 0 *

Dependent Tests

Repeated Measures t-Test for Averages

The example uses [IND_DEP_DATA_LONG.CSV](#), so once that is loaded select **Means and Dispersion** under the **One- and Two-Sample Tests** menu dropdown.



The data **Uses** a **Reference Column**, so select that under **How is the data configured?** Select the **Factor** as **Treat** and **Select the Data** as **Measure**. Then select the two **Groups** as **Before** and **After**. This is a **Dependent** test, so select that under **Independent or Dependent Test?** This example is a repeated measures *t*-test, so it is **Dependency by Nature (\bar{D})** so make sure that is selected. It is a non-directional hypothesis so select **Δ of sample is not equal to Δ** under **Alternative Hypothesis for Means**. As usual, you can select the **Confidence** and **Decimals**.

Mean and Dispersion Testing - Data

How is the data configured?

☐ Data in Columns

☒ Use Reference Column

Select Factor

Treat

Select Data

Measure

Group 1

Before

Group 2

After

Dependency by

☒ Nature (\bar{D})

☐ Design (Mean Difference)

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

Δ of sample is not equal to Δ

$\bar{D} = -0.9655$

$s_D = 0.4973$

$n = 20$

Independent or Dependent Test?

Dependent

$\Delta :$ 0

Results

Dependent Samples t Test for Means (D-bar method)

$\bar{D} = -0.9655$ $\Delta = 0$

$s_D = 0.4973$

$n = 20$

95 % confidence interval for $\Delta :$ -1.1982 to -0.7327

$\sigma_D :$ 0.3782 to 0.7263

If normal use Two Dependent Sample t Test For Variance

Normal Dependent Test for $\sigma_1^2 = \sigma_2^2 :$ $t = 0.793$ $df = 18$ $p = 0.4381$

If non-normal use Two-Sample Dependent t Test for Equality of ADM_{n-1} (D-Bar Method)

Non-normal Dependent Test for $ADM_{n-1} :$ $t = 1.4025$ $df = 18$ $p = 0.1778$

\bar{D} Test for μ_1 is not equal to $\mu_2 :$ $t = -8.6824$ $df = 19$ $p = 0 *$

Matched Pairs t-test for Averages and Dispersion

The example uses [ZAP_v_ACTUALLY_PAURED.csv](#), so once that is loaded select **Means and Dispersion** under the **One- and Two-Sample Tests** menu dropdown.

One- and Two-Sample Tests

- Test Reference
- Means and Dispersion**
- Proportion
- Poisson
- Nonparametric

The **Data is in Columns**, so select that under **How is the data configured?** Select the two columns **TTFP ZapCareer** and **TTFP Actually** under **Analyze which column(s)?** This is a **Dependent** test, so select that under **Independent or Dependent Test?** This example is a paired t-test, so it is **Dependency by Design (Mean Difference)** so make sure that is selected. It is a non-directional hypothesis so select **μ_1 is not equal to μ_2** under **Alternative Hypothesis for Means**. As usual, you can select the **Confidence** and **Decimals**.

Mean and Dispersion Testing - Data

How is the data configured?

☒ Data in Columns

☐ Use Reference Column

Analyze which column(s)?

☐ Difficulty score

☒ TTFP ZapCareer

☒ TTFP Actually

☐ Difference

Dependency by

☐ Nature (\bar{D})

☒ Design (Mean Difference)

Confidence

0.95

Decimals

4

Alternative Hypothesis for Means

μ_1 is not equal to μ_2

$\bar{X}_1 = 5.9356$
 $s_1 = 3.0867$
 $n = 20$

$\bar{X}_2 = 4.3932$
 $s_2 = 2.9899$
 $r_{12} = 0.9493$

Results

Dependent Samples t Test For Means - Difference of Means Method (Equal Variances)

$\bar{X}_1 = 5.9356$ $\bar{X}_2 = 4.3932$
 $s_1 = 3.0867$ $s_2 = 2.9899$
 $n = 20$ $r_{12} = 0.9493$

95 % confidence interval for

μ_1 :	4.491	to	7.3802
μ_2 :	2.9939	to	5.7925
σ_1 :	2.3474	to	4.5083
σ_2 :	2.2738	to	4.3669
ρ_{12} :	0.8738	to	0.9801
$\mu_1 - \mu_2$:	1.0873	to	1.9976

Correlation Test $\rho_{12} = 0$: t = 12.8056 df = 18 p = 0 *

If normal use Two Dependent Sample t Test For Variance

Normal Dependent Test for $\sigma_1^2 = \sigma_2^2$: t = 0.4298 df = 18 p = 0.6724

If non-normal use Two-Sample Dependent t Test for Equality of ADMn-1 (Mean-Difference Method)

Non-normal Dependent Test for ADM_{n-1} : t = 0.3344 df = 17 p = 0.7422

Paired Test for μ_1 is not equal to μ_2 : t = 7.0924 df = 19 p = 0 *

The dependent t-test for variance and dispersion are included in the output, as well as the test for correlation.

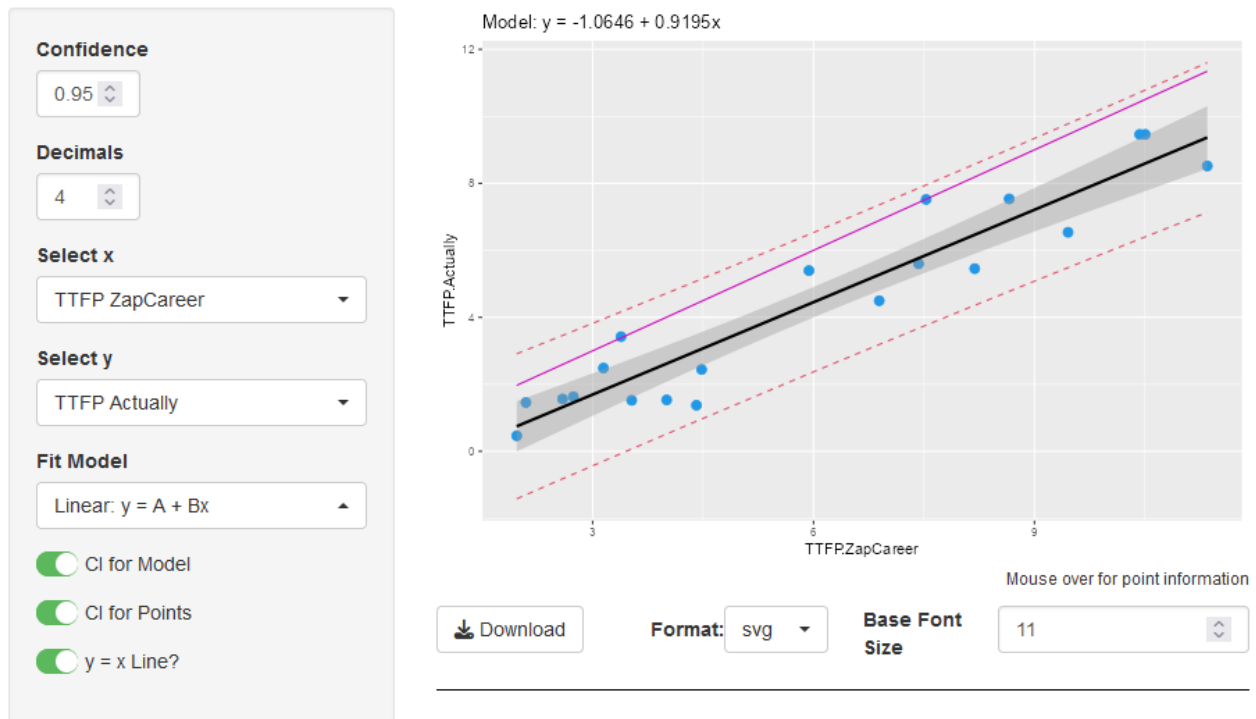
To make the scatter plot with the iso-line, select Correlation and Association from the menu, then the Scatterplot tab.

Correlation and Association ANOVA SPC

Enter Statistics Use Data Scatterplot

Select **x** as **TTFP ZapCareer** and Select **y** as **TTFP Actually**. This will create a scatterplot. Add the line by selecting **Linear: $y = A + Bx$** under **Fit Model**. Include the model and point confidence intervals by clicking on **CI for Model** and **CI for Points**. Add the iso-line by clicking on **$y = x$ Line?** This analysis was done at $\alpha = 0.05$, so select a **Confidence** of 0.95 and adjust **Decimals** as you like.

Scatterplot



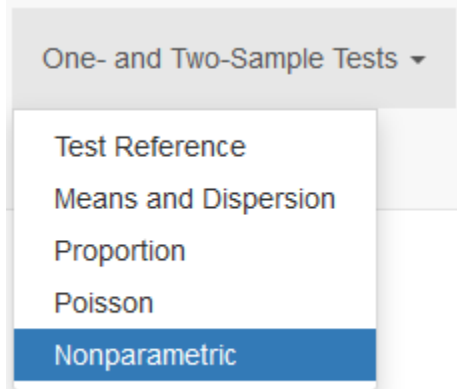
One-Sample Test for Pearson Product Moment Correlation

$r = 0.9493$ $\rho_0 = 0$
 $r^2 = 0.9011$
 $n = 20$
 95 % confidence interval for ρ : 0.8738 to 0.9801
 Test for ρ of sample is not equal to ρ_0 : $t = 12.8056$ $df = 18$ $p = 0^*$
 Power to reject the null if the observed difference was real = 100 %

Once you select a model, ROIstat will calculate the significance of the r for that model. You can change the **Base Font Size** and **Download** the graph in a variety of **Formats**.

McNemar's Test of Change

Select **Nonparametric** under the **One- and Two-Sample Tests** menu dropdown.



This test was at an $\alpha = 0.05$, so select a **Confidence** of 0.95 and choose the number of **Decimals**. We will be entering the counts directly, so select the **Enter Statistics** tab. This is a **Two-Sample Dependent** test, so select that under **One Sample or Two?** This will change the available test, so select **McNemar's Test of Change** under **Select Test**. This is a non-directional test, so select **Pass₁ Fail₂ is not equal to Fail₁ Pass₂** under the **Alternative Hypothesis**.

That will change the data entry to look like a contingency table. Enter the counts in the boxes.

Nonparametric Tests

Enter Statistics
Use Data

Confidence

0.95

One Sample or Two?

Two-Sample Dependent

Select Test

☐ Dependent Sign Test
☐ Dependent Wilcoxon Signed Ranks
☒ McNemar's Test of Change

Decimals

4

☐ More information about selected test

Alternative Hypothesis

Pass₁ Fail₂ is not equal to Fail₁ Pass₂

	Pass 2	Fail 2
Pass 1	186	0
Fail 1	12	2

Results

Method: McNemar's Test for Dependent Proportions (Exact)

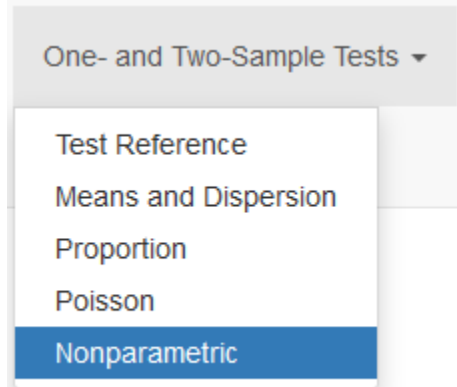
$P_1 F_2 = 0$ $p(b) = 0$
 $F_1 P_2 = 12$ $p(c) = 1$

95 % confidence

Test for Pass₁ Fail₂ is not equal to Fail₁ Pass₂ : p = 5e-04 *

Two-Sample Dependent Sign Test

This test is found by clicking **Nonparametric** under the **One- and Two-Sample Tests** menu dropdown.



You are given the counts in this problem, so select the **Enter Statistics** tab. The $\alpha = 0.05$ so enter a **Confidence** of 0.95. This is a **Two-Sample Dependent** test, so select that under **One Sample or Two?** then **Select the Dependent Sign Test**. You can adjust **Decimals** as needed. You can also get **More information about the selected test** by clicking on the checkbox.

This is a non-directional test, so make sure the **Alternative Hypothesis** is set to $x_n - y_n$ is not equal to 0.

Enter the counts into the boxes.

Nonparametric Tests

Enter Statistics

Use Data

Confidence
0.95

One Sample or Two?
Two-Sample Dependent

Select Test
☒ Dependent Sign Test
☐ Dependent Wilcoxon Signed Ranks
☐ McNemar's Test of Change

Decimals
4

☐ More information about selected test

Alternative Hypothesis
 $x_n - y_n$ is not equal to 0

n^+ : 8

n^- : 0

n^- : 2

Results
Two-Sample Dependent Sign Test for Location
Method: One-Sample Proportion Test (Exact)
 $n^+ = 8$
 $n^- = 0$
 $n^- = 2$
 $n_{observed} = 10$
 $n_{included} = 10$
Test for the probability that the difference in each pair is positive = 0.5
95 % confidence interval for π^+ : 0.4439 to 0.9748
Test for $x_n - y_n$ is not equal to 0 : $p^+ = 0.8$ $p = 0.1094$
,
Power to reject the null if the observed difference was real = 0.3758

Wilcoxon Signed-Rank Test

This example uses data, so once [CHILIHEAD_DEP.TXT](#) is loaded, click on **Nonparametric** under the **One- and Two-Sample Tests** menu dropdown.

One- and Two-Sample Tests

Test Reference

Means and Dispersion

Proportion

Poisson

Nonparametric

Select the **Use Data** tab. The **Data is in Columns** so select that under **How is the data configured?** Select the two columns **Current Package** and **New Package** under **Analyze which column(s)?** This is at an $\alpha = 0.05$ so select a **Confidence** of 0.95. You can adjust the Decimals shown.

This is a **Dependent** test, so select that under **Independent or Dependent Test?** The example is non-directional, so make sure the **Alternative Hypothesis** is set to $x_n - y_n$ is not equal to 0.

Nonparametric Tests

[Enter Statistics](#) [Use Data](#)

Confidence

How is the data configured?
☒ Data in Columns
☐ Use Reference Column

Analyze which column(s)?
☒ Current Package
☒ New Package
☐ Current-New

Independent or Dependent Test?

Select Test
☐ Dependent Sign Test
☒ Dependent Wilcoxon Signed Ranks
☐ McNemar's Test of Change

Decimals

☐ More information about selected test

Alternative Hypothesis

$n^+ = 2$
 $n^- = 3$
 $n^- = 25$

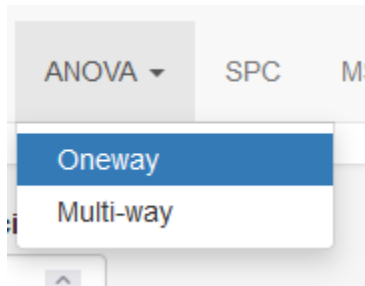
Results
Method: Dependent Wilcoxon Signed-Ranks Test
 $S^+ = 14$
 $S^- = 364$
 $n = 27$
95 % confidence
Test for $x_n - y_n$ is not equal to 0 : $W = 14$ $p = 0^*$

You can get **More information about the selected test** by clicking on the checkbox.

Chapter 13 – Analysis of Variance

One-Way Fixed Factor ANOVA – Means, Importance, and Dispersion Analysis

The example uses the data file [PACKAGE.CSV](#). After loading that, select **Oneway** under the **ANOVA** menu dropdown.



Make sure that **ANOVA** is selected first. This example uses an $\alpha = 0.05$, so set the **Confidence** to be 0.95. You can choose the number of **Decimals**. **Select the Factor** as **design** and **Select the Data** as **burst**. This example demonstrates the **Fixed – Fisher ANOVA**, so select that under **Factor and Test Type**. Select **Show Dispersion Tests?**

ANOVA

Post-hoc

Confidence

0.95

Decimals

4

Select Factor

design

Select Data

burst

Factor and Test Type

Fixed - Welch

Fixed - Fisher

Random

Kruskal-Wallis

☒ Show Dispersion Tests?

Oneway ANOVA

Fisher's One-way analysis of variance (assumes equal variances, robust if n is large)

Model : burst by design

Source	df	SS	MS	F	p
design	3	741.7917	247.2639	134.2609	0 *
Within	20	36.8333	1.8417		
Total	23	778.625			

Fixed Effect Importance:
 $\omega^2 = 94.3367\%$

Dispersion Analysis

If normally distributed within cells

Levene F(3 , 20) = 8.9177 p = 6e-04 *

If not normally distributed within cells

If n ≤ 10 *ADM* F(3 , 20) = 7.2126 p = 0.0018 *

If n > 10 *ADM*_{n-1} F(3 , 16) = 6.825 p = 0.0036 *

Importance is shown under the ANOVA source table.

One-Way Fixed Factor *Post-Hoc* Analysis and Graph

To continue the analysis above, select **Post-hoc**.

Select **Post-Hoc** as **Tukey (equal variances)** to get the Tukey's HSD test, or select **Games & Howell (unequal variances)** to get that test. The confidence for the test is set on the **ANOVA** tab.

ANOVA

Post-hoc

Decimals

4

☐ Add Lines?

Choose plot type

☐ Points Only

☐ Violin

☒ Boxplot

Select Post-Hoc

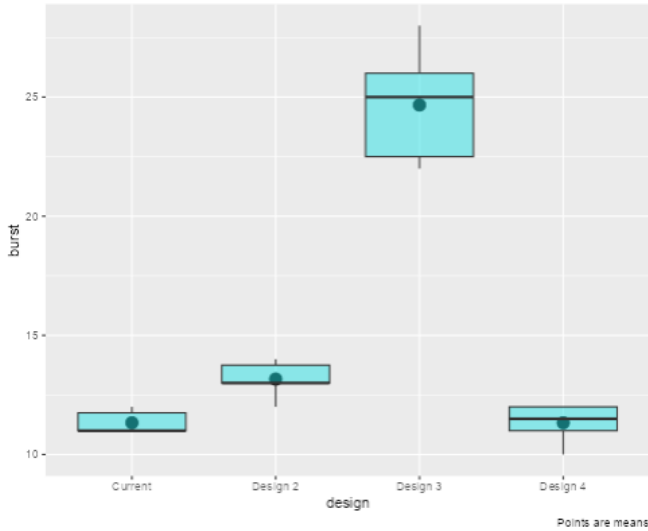
☒ Tukey (equal variances)

☐ Games & Howell(unequal variances)

☐ Post-hoc Details?

Oneway ANOVA Post-hoc

Confidence set on Oneway ANOVA tab



Download

Format: svg

Base Font
Size

11

Tukey's HSD

	Current	Design 4	Design 2	Design 3
Current				Reject
Design 4				Reject
Design 2				Reject
Design 3	Reject	Reject	Reject	

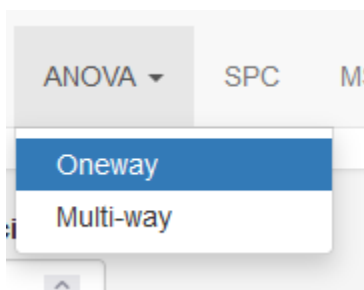


If you want the results of all the post-hoc tests, click the checkbox next to **Post-hoc Details?**

ROIstat automatically generates a graph for you. **You can Add Lines?** connecting the means or medians, you can **Choose a Plot Type** as **Points Only**, **Violin**, or **Boxplot**. You can change the **Base Font Size** and **Download** the graphic in a number of **Formats**.

One-Way Random Effects ANOVA – Variance Contribution, Importance, and Dispersion Analysis

The random effects ANOVA example uses the same data as the fixed effect example, so load the data file [PACKAGE.CSV](#) and select **Oneway** under the **ANOVA** menu dropdown.



Make sure that **ANOVA** is selected first. This example uses an $\alpha = 0.05$, so set the **Confidence** to be 0.95. You can choose the number of **Decimals**. **Select the Factor** as **design** and **Select the Data** as **burst**. This example demonstrates the **Random** effects ANOVA, so select that under **Factor and Test Type**. Optionally select **Show Dispersion Tests?**

ANOVA

Post-hoc

Confidence

0.95

Decimals

4

Select Factor

design

Select Data

burst

Factor and Test Type

Fixed - Welch

Fixed - Fisher

Random

Kruskal-Wallis

☒ Show Dispersion Tests?

Oneway ANOVA

Fisher's One-way analysis of variance (assumes equal variances, robust if n is large)

Model : burst by design

Source	df	SS	MS	F	p
design	3	741.7917	247.2639	134.2609	0 *
Within	20	36.8333	1.8417		
Total	23	778.625			

Random Effect Importance:

Treatment Variance = 40.9037 $\hat{\sigma}_{treat} = 6.3956$

Within Variance = 1.8417 $\hat{\sigma}_{within} = 1.3571$

Total Variance = 42.7454 $\hat{\sigma}_{total} = 6.538$

Intraclass Correlation = 95.6915 %

Dispersion Analysis

If normally distributed within cells

Levene $F(3, 20) = 8.9177$ $p = 6e-04 *$

If not normally distributed within cells

If $n \leq 10$ **ADM** $F(3, 20) = 7.2126$ $p = 0.0018 *$

If $n > 10$ **ADM_{n-1}** $F(3, 16) = 6.825$ $p = 0.0036 *$

The importance statistic ICC is reported along with the variance components under the ANOVA source table.

One-Way Random Effects Post-Hoc Analysis

To continue the analysis above, simply click on the Post-hoc tab to generate the random-effects graph and variance components. You can adjust the Decimals shown here.



You can change the **Base Font Size** and **Download** the graphic in a number of **Formats**.