Real Statistics Using Excel

Everything you need to do real statistical analysis using Excel

Lilliefors Test for Normality

When the population mean and standard deviation are known we can use the one sample Kolmogorov-Smirnov test to test for normality, as described in Kolmogorov-Smirnov Test for Normality.

However, when the population mean and standard deviation are not known, but instead are estimated from the sample data, then the usual Kolmogorov-Smirnov test based on the critical values in the Kolmogorov-Smirnov Table yields results that are too conservative. Lilliefors created a related test that gives more accurate results in this case (see Lilliefors Test Table).

The Lilliefors Test uses the same calculations as the Kolmogorov-Smirnov Test, but the table of critical values in the Lilliefors Test Table is used instead of the Kolmogorov-Smirnov Table. Since the critical values in this table are smaller, the Lilliefors Test is less likely to show that data is normally distributed.

Example 1: Repeat Examples 1 and 2 of the Kolmogorov-Smirnov Test for Normality using the Lilliefors Test.

For Example 1 of Kolmogorov-Smirnov Test for Normality, using the Lilliefors Test Table, we have

$$D_{n,\alpha} = \frac{.895}{f(n)} \qquad f(n) = \frac{.83 + n}{\sqrt{n}} - .01$$

$$D_{1000,.05} = \frac{.895}{f(1000)} = \frac{.895}{31.64} = .0283 \qquad f(1000) = \frac{.83 + 1000}{\sqrt{1000}} - .01 = 31.64$$

Since $D_n = 0.0117 < 0.0283 = D_{n,\alpha}$, once again we conclude that the data is a good fit with the normal distribution. (Note that the critical value of .0283 is smaller than the critical value of .043 from the KS Test.)

For Example 2 of Kolmogorov-Smirnov Test for Normality, using the Lilliefors Test Table with n = 15 and α = .05, we find that D_n = 0.1875 < 0.2196 = $D_{n,\alpha}$, which confirms that the data is normally distributed.

Real Statistics Functions: The following functions are provided in the Real Statistics Resource Pack to automate the table lookup:

> **LCRIT**(*n, α, tails, interp*) = the critical value of the Lilliefors test for a sample of size *n*, for the given value of alpha (default .05) and *tails* = 1 (one tail) or 2 (two tails, default) based on the Lilliefors Test Table. If *interp* = TRUE (default) harmonic interpolation is used; otherwise linear interpolation is used.
>
> **LPROB**(*x, n, tails, iter, interp, txt*) = an approximate p-value for the Lilliefors test for the D_n value equal to *x* for a sample of size *n* and *tails* = 1 (one tail) or 2 (two tails, default) based on a linear interpolation (if *interp* = FALSE) or harmonic interpolation (if *interp* = TRUE, default) of the critical values in the Lilliefors Test Table, using *iter* number of iterations (default = 40).

Note that the values for α in the table in the Lilliefors Test Table range from .01 to .2 (for *tails* = 2) and .005 to .1 for *tails* = 1. When *txt* = FALSE (default), if the p-value is less than .01 (*tails* = 2) or .005 (*tails* = 1) then the p-value is given as 0 and if the p-value is greater than .2 (*tails* = 2) or .1 (*tails* = 1) then the p-value is given as 1. When *txt* = TRUE, then the output takes the form "< .01", "< .005", "> .2" or "> .1".

For Example 2 of Kolmogorov-Smirnov Test for Normality, $D_{n,\alpha}$ = LCRIT(15, .05, 2) = .2196 > .184 = D_n and p-value = LPROB(0.184, 15) = .182858 > .05 = α, and so once again we can't reject the null hypothesis that the data is normally distributed.

Real Statistics Support for KS Test

Click here for information about the Real Statistics functions that perform the Kolmogorov-Smirnov test both when the mean and standard deviation are specified and when they are estimated from the data. Both raw data and a data in the form of a frequency table are supported.

Lilliefors Distribution

Especially for values of α not found in the Lilliefors Test Table, we can use an approximation to the Lilliefors distribution. Click here for more information about this distribution, including some useful functions provided by the Real Statistics Resource Pack.

11 Responses to *Lilliefors Test for Normality*

Chris *says:*
August 10, 2019 at 6:43 am

Dear Charlie,

Thank you for your website, which is well written and particularly pedagogical.

I see a problem of principles in these tests of normality. In fact we don't test the hypothesis Ho with an accuracy of alpha but we test the hypothesis H1 (rejection) with this percentage.

For the KS test for example the higher the % (0.95 ; 0.99 ; 0.995 ; …) and the lower the chance not to conclude H1 and reject Ho, so the "easier" to conclude it would be a Gaussian! That makes no sense.

When the test passes with success, that does not mean we have 95 % (or more) it is a Gaussian. It means that we can't say with 95 % chance it is something different. But the probability it is really a normal dsitribution is not known.

So shouldn't we always take at least 50 % (meaning 50 % or *less*) if we want to conclude distribution is a Gaussian ? Indeed, to fairly conclude we have a "good" chance that it is a Gaussian, we should at least be allowed to say there is no 50 % chance it is something else…

Reply

Charles *says:*
August 10, 2019 at 9:47 am

Hello Chris,
This is the sort of issue we have with all statistical tests (at least the non-Bayesian tests). We don't know whether the data is really coming from a normal distribution whether the p-value is 50% or 2%. The value of 5% is arbitrary, but commonly used, compromise. Since rejection occurs for values less than alpha, the lower the alpha value the more

likely you are to declare the data as normally distributed. An alpha of 50% would increase the likelihood that you would declare the data as not normally distributed.
Charles

Reply

Mark G Filler *says:*
November 5, 2017 at 12:41 am

For LCRIT, I can't seem to get a value if n > 50. What am I doing wrong?

Reply

Charles *says:*
November 5, 2017 at 8:18 am

Mark,
I am not sure what you are doing wrong, but I just tried to use =LCRIT(60), and I got the value .114113. What version of Real Statistics are you using? You can find this out by entering the formula =VER()
Charles

Reply

Mark G Filler *says:*
November 5, 2017 at 8:46 pm

Charles

I am using 4.14 2010.

When I use Excel 2013 with the corresponding Real Statistics version, it works OK.

I don't like Excel 2013, so I guess this a cost of that attitude.

Reply

Mark G Filler *says:*
November 5, 2017 at 9:50 pm

Charles

Problem solved – I installed version 5.2 for Excel 2010 and LCRIT works for a sample size of 300.

Mark

Reply

Charles *says:*
November 6, 2017 at 8:28 am

Mark,
Good to hear.
Charles

David *says:*
August 7, 2017 at 9:28 pm

Hey Charles,

If I'm not mistaken, Dn from the Kolmogorov-Smirnov Test for Normality page should be Dn = 0.1875, not Dn = 0.184.

Thanks.

Reply

Charles *says:*
August 7, 2017 at 11:23 pm

David,
Yes you are correct. Thanks for catching this mistake. I really appreciate your helping in improving the Real statistics website.
Charles

Reply

Keith Wild *says:*
June 28, 2017 at 4:44 pm

Of the many tests regimes there are for tests for normality. Is there a list illustrating the order of preference for the test method according to the type of data you have?
I mean which test should I use for what type of data? It seems to be so easy to fudge a result as necessary according to the test method.

Reply

Charles *says:*

June 28, 2017 at 9:36 pm

Keith,
In general, I believe that the Shapiro-Wilk test is the best one to use. If you have a number of ties, then d'Agostino-Pearson is probably better.
Charles

Reply