

How do I perform a Kolmogorov-Smirnov test in Excel?

Authored by
stats writer

July 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I perform a Kolmogorov-Smirnov test in Excel?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=165281>

The Kolmogorov-Smirnov test is a statistical method used to determine if a dataset follows a specific distribution. In order to perform this test in Excel, you will need to first organize your data into two separate columns, with one column containing the actual data and the other containing the expected values based on the distribution you want to test. Next, you will need to calculate the cumulative distribution function (CDF) for both columns. This can be done by using the "ECDF" function in Excel. Once the CDFs are calculated, you can use the "KSTEST" function to compare the two distributions and determine the p-value. If the p-value is greater than the specified alpha level, the data can be considered to follow the chosen distribution. However, if the p-value is less than the alpha level, the data does not fit the distribution. Using this method, you can easily perform a Kolmogorov-Smirnov test in Excel to analyze the distribution of your data.

Perform a Kolmogorov-Smirnov Test in Excel

The Kolmogorov-Smirnov test is used to determine whether or not a sample is .

This test is widely used because many statistical tests and procedures make the that the data is normally distributed.

The following step-by-step example shows how to perform a Kolmogorov-Smirnov test on a sample dataset in Excel.

Step 1: Enter the Data

First, let's enter the values for a dataset with a sample size of $n = 20$:

	A	B	C	D	E	F
1	Data					
2	5.57					
3	8.32					
4	8.35					
5	8.74					
6	8.75					
7	9.38					
8	9.91					
9	9.96					
10	10.36					
11	10.65					
12	10.77					
13	10.97					
14	11.15					
15	11.18					
16	11.47					
17	11.64					
18	11.88					
19	12.24					
20	13.02					
21	13.19					
22						
23						
24						
25						

Step 2: Calculate Actual vs. Expected Values from Normal Distribution

Next, we'll calculate the actual values vs. the expected values from the normal distribution:

	A	B	C	D	E	F	G	H	I	J
1	Data	Cumulative	Expected	(RANK-1)/N	NORM.S.INV	Actual	Difference		Mean	10.375
2	5.57	1	0.05	0	-1.64485363	0.00426	0.004264		Std. Dev.	1.82672
3	8.32	2	0.1	0.05	-1.28155157	0.1303	0.080302			
4	8.35	3	0.15	0.1	-1.03643339	0.13381	0.033814		Maximum	0.10983
5	8.74	4	0.2	0.15	-0.84162123	0.18538	0.035381			
6	8.75	5	0.25	0.2	-0.67448975	0.18685	0.013152			
7	9.38	6	0.3	0.25	-0.52440051	0.29298	0.042983			
8	9.91	7	0.35	0.3	-0.38532047	0.39953	0.099534			
9	9.96	8	0.4	0.35	-0.2533471	0.41014	0.060141			
10	10.36	9	0.45	0.4	-0.12566135	0.49672	0.096724			
11	10.65	10	0.5	0.45	0	0.55983	0.109832			
12	10.77	11	0.55	0.5	0.125661347	0.5856	0.085597			
13	10.97	12	0.6	0.55	0.253347103	0.62768	0.077682			
14	11.15	13	0.65	0.6	0.385320466	0.66431	0.064311			
15	11.18	14	0.7	0.65	0.524400513	0.67028	0.020278			
16	11.47	15	0.75	0.7	0.67448975	0.72556	0.025558			
17	11.64	16	0.8	0.75	0.841621234	0.75569	0.005688			
18	11.88	17	0.85	0.8	1.036433389	0.795	0.005004			
19	12.24	18	0.9	0.85	1.281551566	0.84636	0.003638			
20	13.02	19	0.95	0.9	1.644853627	0.92618	0.026184			
21	13.19	20	1	0.95		0.93834	0.011657			
22										
23										
24										
25										

Here is the formula we used in various cells:

B2: =ROW() - 1
 C2: =B2/COUNT(\$A\$2:\$A\$21)
 D2: =(B2-1)/COUNT(\$A\$2:\$A\$21)
 E2: =IF(C2<1, NORM.S.INV(C2),"")
 F2: =NORM.DIST(A2, \$J\$1, \$J\$2, TRUE)
 G2: =ABS(F2-D2)
 J1: =AVERAGE(A2:A21)
 J2: =STDEV.S(A2:A21)
 J4: =MAX(G2:G21)

Step 3: Interpret the Results

A Kolmogorov-Smirnov test uses the following null and

alternative hypotheses:

H0: The data is normally distributed. HA: The data is not normally distributed.

To determine if we should reject or fail to reject the null hypothesis we must refer to the Maximum value in the output, which turns out to be 0.10983.

This represents the maximum absolute difference between the actual values of our sample and the expected values from a normal distribution.

To determine if this maximum value is statistically significant, we must refer to a Kolmogorov-Smirnov Table of critical values and find the number equal to $n = 20$ and $\alpha = .05$.

Since our maximum value is not greater than this critical value, we fail to reject the null hypothesis.

This means we can assume that our sample data is normally distributed.

Additional Resources

The following tutorials explain how to perform other

common statistical tests in Excel:

ARABPSYCHOLOGY.COM