How do I perform a Kolmogorov-Smirnov Test in SAS?

Authored by

stats writer

July 1, 2024

The Kolmogorov-Smirnov Test is a statistical test used to compare two probability distributions. In SAS, the test can be performed by using the PROC UNIVARIATE procedure. This procedure allows the user to specify the two datasets to be compared and the desired level of significance. The output of the test includes a D statistic and corresponding p-value, which can be used to determine if the two distributions are significantly different. The Kolmogorov-Smirnov Test in SAS is a useful tool for analyzing data and making informed decisions based on the comparison of probability distributions.

Perform a Kolmogorov-Smirnov Test in SAS

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

Example: Kolmogorov-Smirnov Test in SAS

First, let's create a dataset in SAS with a sample size of n = 20:

```
/*create dataset*/
data my_data;
input Values;
```

```
datalines;
5.57
8.32
8.35
8.74
8.75
9.38
9.91
9.96
10.36
10.65
10.77
10.97
11.15
11.18
11.47
11.64
11.88
12.24
13.02
13.19
;
run;
```

Next, we'll use proc univariate to perform a Kolmogorov-Smirnov test to determine if the sample is normally distributed:

```
/*perform Kolmogorov-Smirnov test*/
proc univariatedata=my_data;
histogram Values / normal(mu=est sigma=est);
run;
```

At the bottom of the output we can see the test statistic and corresponding p-value of the Kolmogorov-Smirnov test:

The UNIVARIATE Procedure
Fitted Normal Distribution for Values

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	10.375
Std Dev	Sigma	1.826721

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.10983186	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.04020411	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq	0.29089867	Pr > A-Sq	>0.250

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	5.57000	6.12541
5.0	6.94500	7.37031
10.0	8.33500	8.03396
25.0	9.06500	9.14290
50.0	10.71000	10.37500
75.0	11.55500	11.60710
90.0	12.63000	12.71604
95.0	13.10500	13.37969
99.0	13.19000	14.62459

The test statistic is 0.1098 and the corresponding p-value is >0.150.

Recall that a Kolmogorov-Smirnov test uses the following null and alternative hypotheses:

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

Since the p-value from the test is not less than .05, we

fail to reject the null hypothesis.

This means we can assume that the dataset is normally distributed.

Additional Resources

The following tutorials explain how to perform a Kolmogorov-Smirnov test in other statistical software: