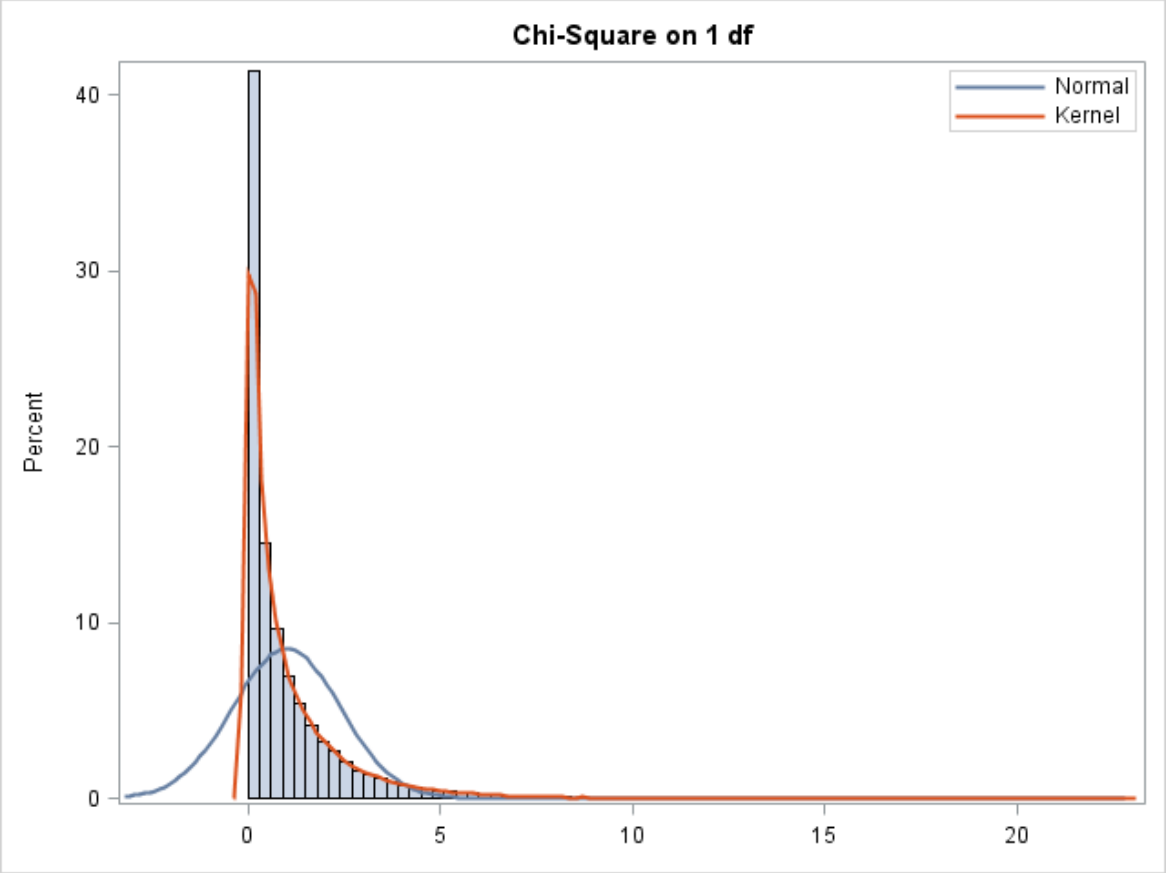
Simulating the Chi-Square Distribution: Annotated Output

Here is output from one run of **ProbChisq.sas**, which is found on my SAS Programs page. I used Proc SGPLOT to plot the empirical sampling distribution as histograms. I used the DENSITY function to overlay two line plots, one a smooth curve based on the actual scores in the sampling distribution (a kernel plot) and the other a smoothed curve showing what a normal distribution with the same mean and variance would look like.



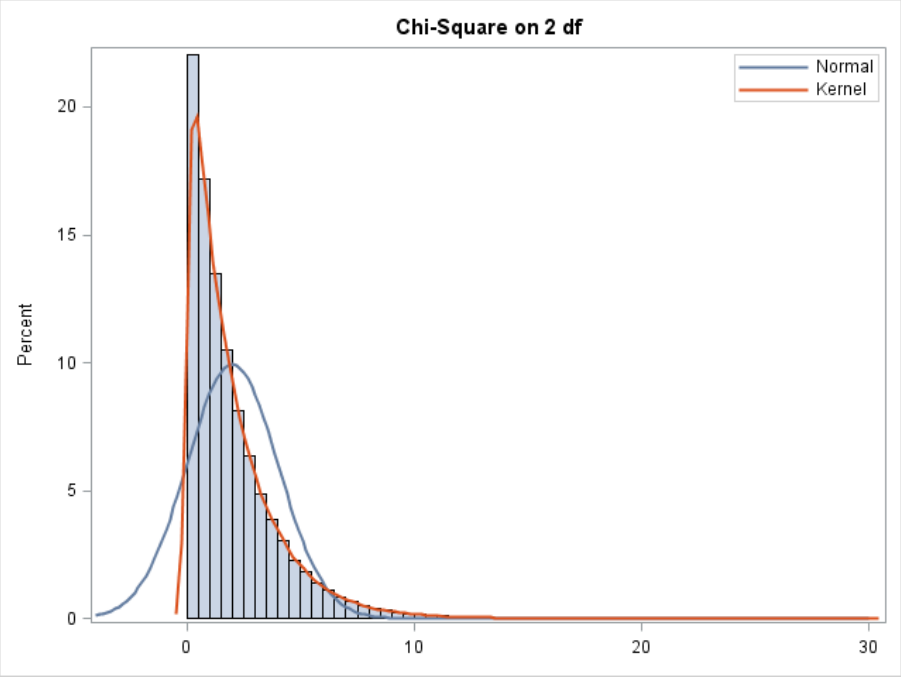
The Chi-Square on one degree of freedom certainly does not look like a normal distribution!

Chi-Square on 1 df

The MEANS Procedure

Analysis Variable : zsq		
Mean	Variance	Skewness
0.9973671	1.9746282	2.8502811

A chi-square distribution always has a mean equal to its degrees of freedom and a variance equal to twice its degrees of freedom. The values obtained here a very close to the expected 1 and 2.

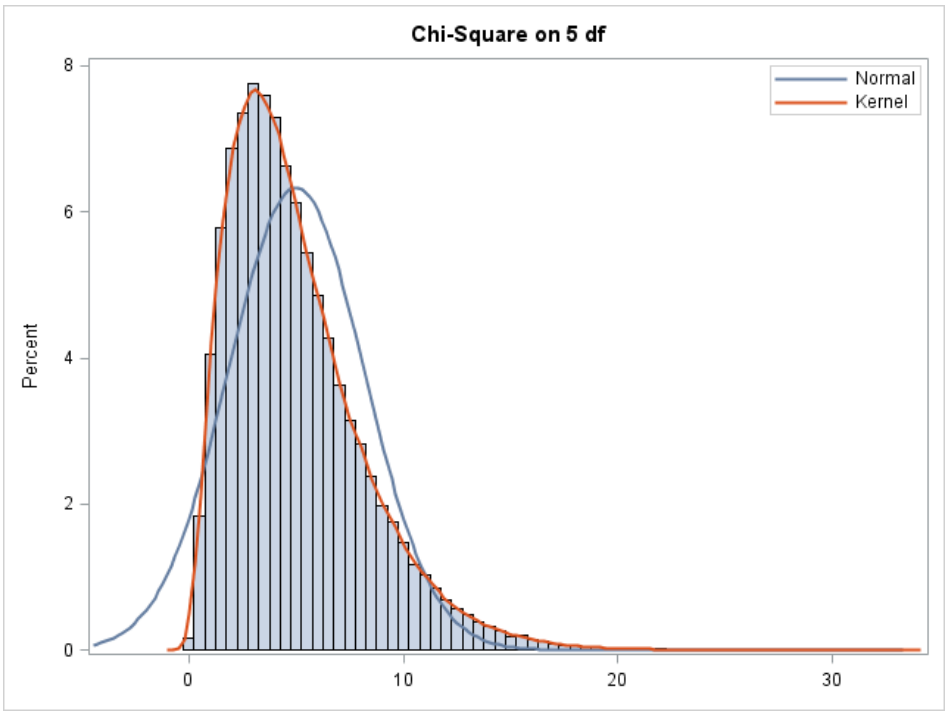


Chi-Square on 2 df

The MEANS Procedure

Analysis Variable : chisq2		
Mean	Variance	Skewness
2.0023596	4.0212408	2.0488650

Notice that increasing the degrees of freedom from one to two has produced a drop in the skewness, but the distribution is still distinctly not normal. Again, the obtained values of mean and variance are very close to what is expected.

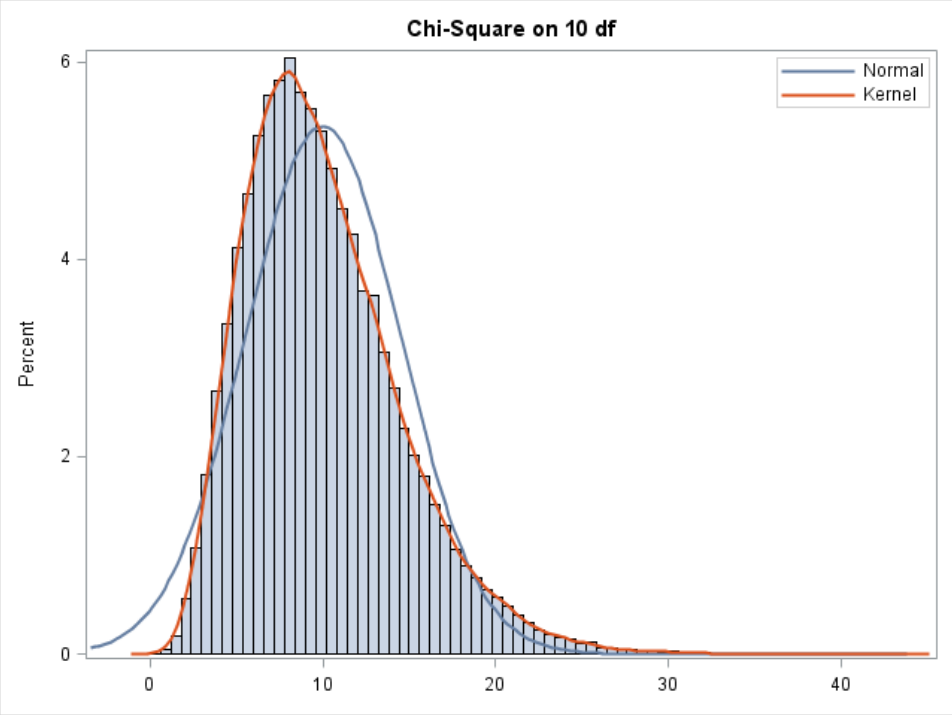


Chi-Square on 5 df

The MEANS Procedure

Analysis Variable : chisq5		
Mean	Variance	Skewness
4.9958670	9.9174648	1.2294911

Another drop in skewness, as the chi-square distribution starts looking more and more like a normal distribution as the degrees of freedom increase.

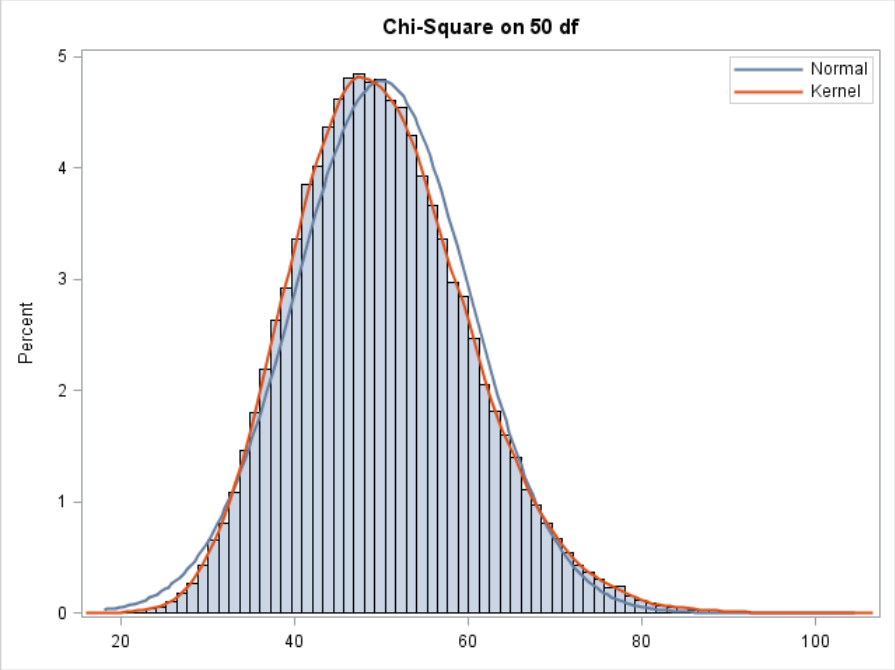


Chi-Square on 10 df

The MEANS Procedure

Analysis Variable : chisq10		
Mean	Variance	Skewness
9.9916046	20.0141543	0.8903758

And now I am going to jump to 50 degrees of freedom, with another reduction in skewness.



Chi-Square on 50 df

The MEANS Procedure

Analysis Variable : chisq50		
Mean	Variance	Skewness
49.9987584	99.9863955	0.3963213

On 50 degrees of freedom, the chi-square is looking even more like a normal distribution.

Karl L. Wuensch, October, 2016.