

Modelling Statistically- Significant Skewness
in Samples from a Normal Distribution

Q. In sampling from a Normal distribution, when is Skew3 statistically significant?

$$\text{Skew3} = 3 * |\text{Mean-Median}| / \text{Std.Dev}$$

1. BACKGROUND:

Doane and Seward (2011) gave exact values for different sample sizes and levels of significance.

Since they used simulation, there was no analytic solution.

The goal of this paper is to fit their data between two simple analytic solutions.

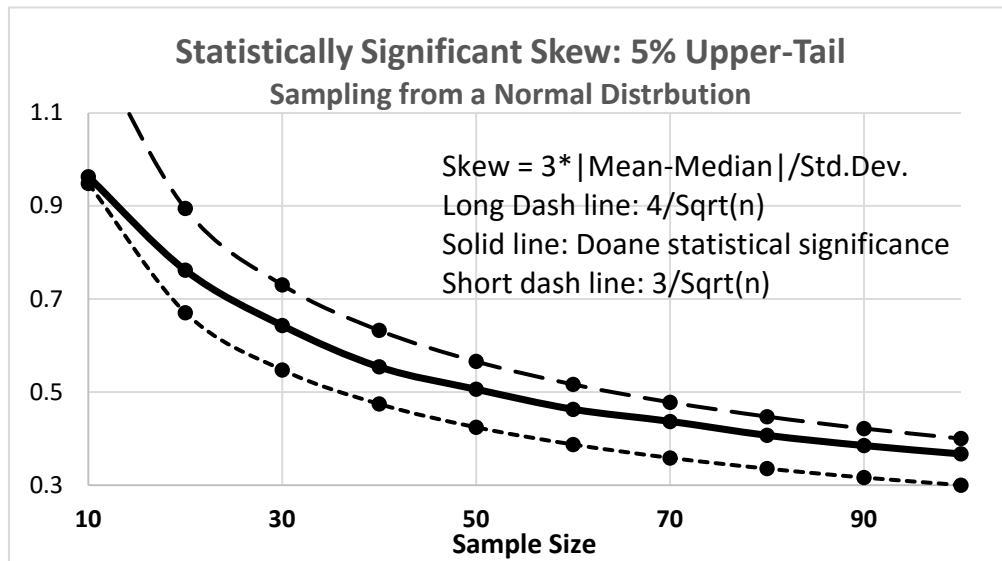
A	B	C	D	E	F	G	H	I
2. DATA		SIMPLE ANALYTIC FUNCTIONS: K / SQRT(n) for various K						
	"Exact"							
	2-tailed 90%					D13	=D\$12/SQRT(\$A13)	
	5% upper		----- Necessary -----	----- Mixed -----			----- Sufficient -----	
n	Skew3	3	3.04	3.05	3.66	3.67	4	
10	0.963	0.95	0.961	0.964	1.157	1.161	1.26	
20	0.762	0.67	0.680	0.682	0.818	0.821	0.89	
30	0.643	0.55	0.555	0.557	0.668	0.670	0.73	
40	0.554	0.47	0.481	0.482	0.579	0.580	0.63	
50	0.506	0.42	0.430	0.431	0.518	0.519	0.57	
60	0.463	0.39	0.392	0.394	0.473	0.474	0.52	
70	0.437	0.36	0.363	0.365	0.437	0.439	0.48	
80	0.407	0.34	0.340	0.341	0.409	0.410	0.45	
90	0.385	0.32	0.320	0.321	0.386	0.387	0.42	
100	0.367	0.30	0.304	0.305	0.366	0.367	0.40	

Doane and Seward (2011)

<http://www.amstat.org/publications/jse/v19n2/doane.pdf>

3.04/sqrt(n) is the highest necessary condition; 3.67/sqrt(n) is the lowest sufficient condition

3. GRAPH OF RESULTS



4. CONCLUSIONS;

Conclusions for $10 \leq n \leq 100$ when sampling from a Normal distribution:

Skew3 must be more than $3/\text{sqrt}(n)$ to be statistically significant -- a necessary condition

If Skew3 is more than $4/\text{sqrt}(n)$, then it is statistically significant -- a sufficient condition