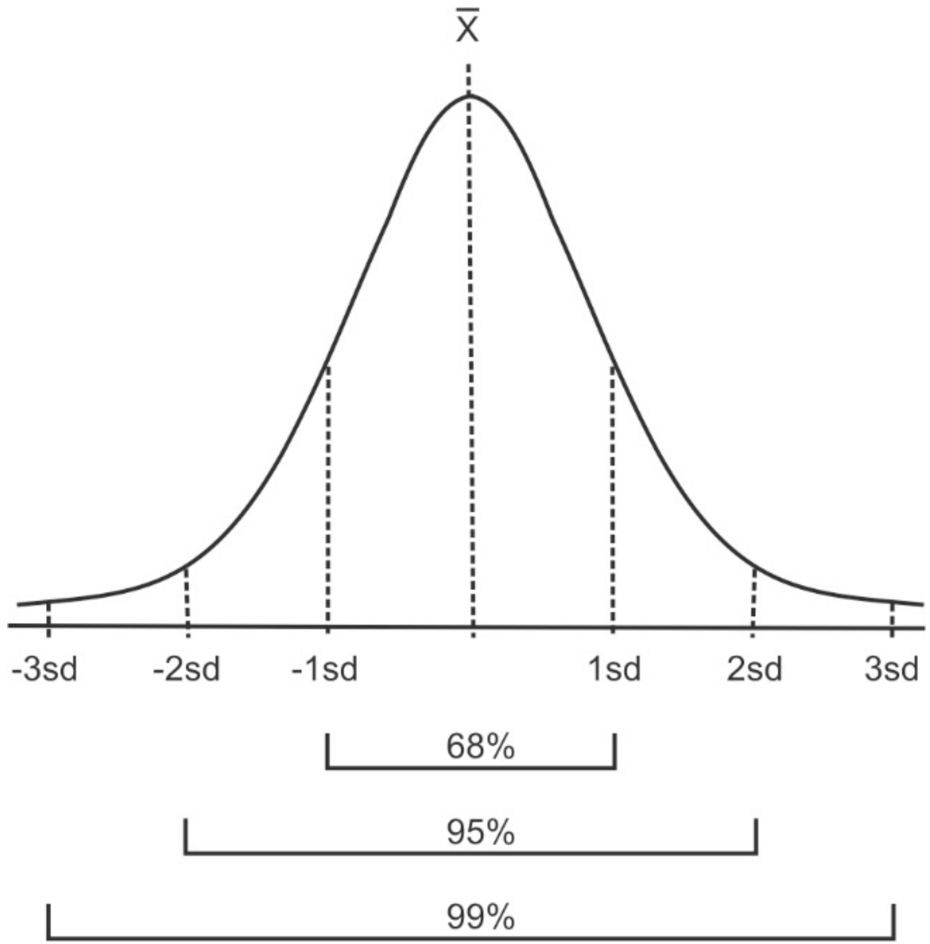
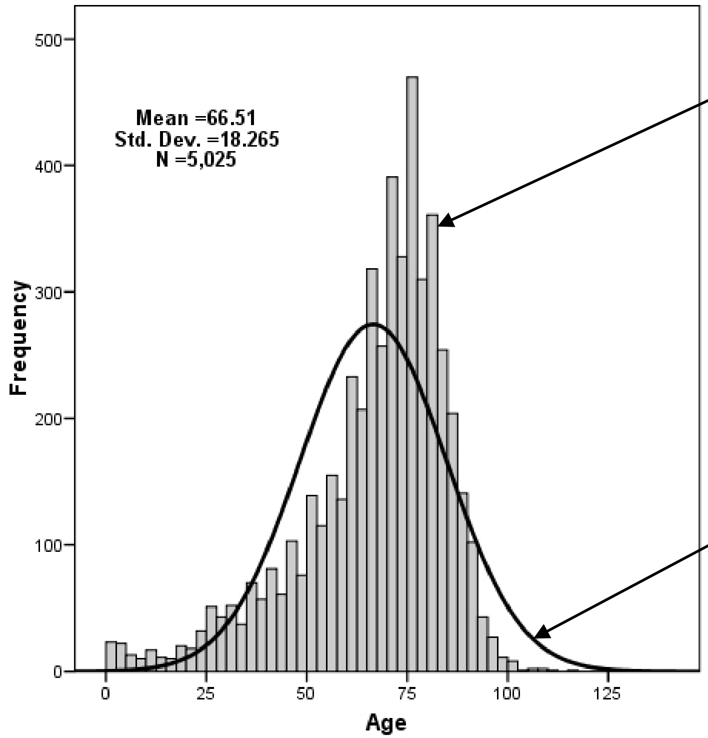
Testing for Normality

For each mean and standard deviation combination a theoretical normal distribution can be determined. This distribution is based on the proportions shown below.



This theoretical normal distribution can then be compared to the actual distribution of the data.



The ***actual*** data distribution that has a mean of 66.51 and a standard deviation of 18.265.

Theoretical normal distribution ***calculated*** from a mean of 66.51 and a standard deviation of 18.265.

Are the actual data statistically different than the computed normal curve?

There are several methods of assessing whether data are normally distributed or not. They fall into two broad categories: *graphical* and *statistical*. The some common techniques are:

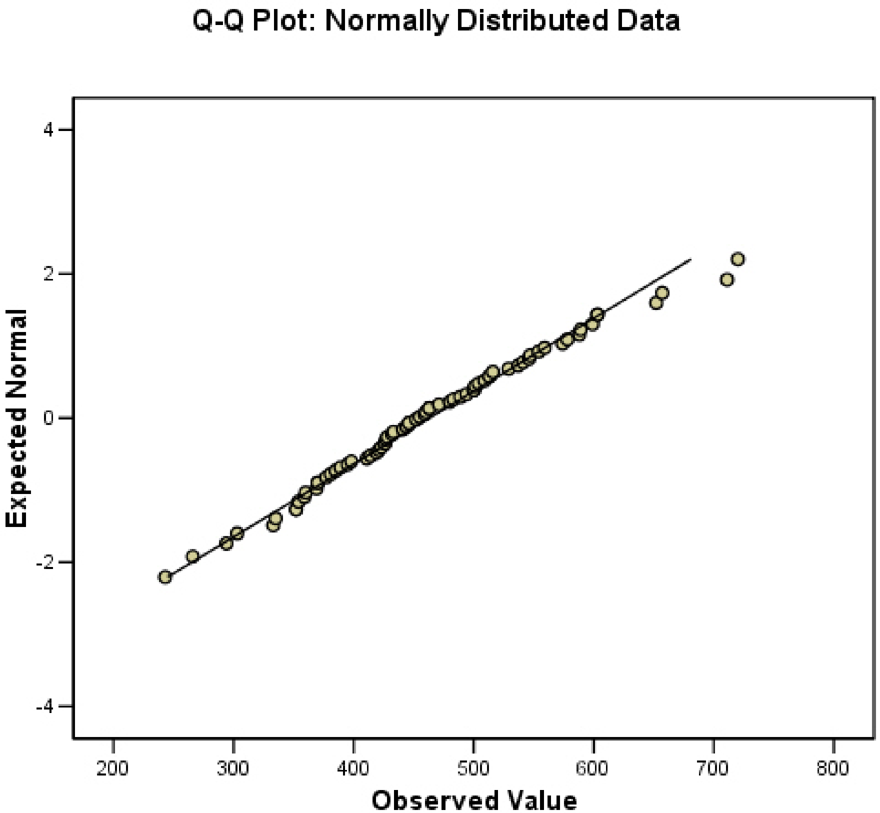
Graphical

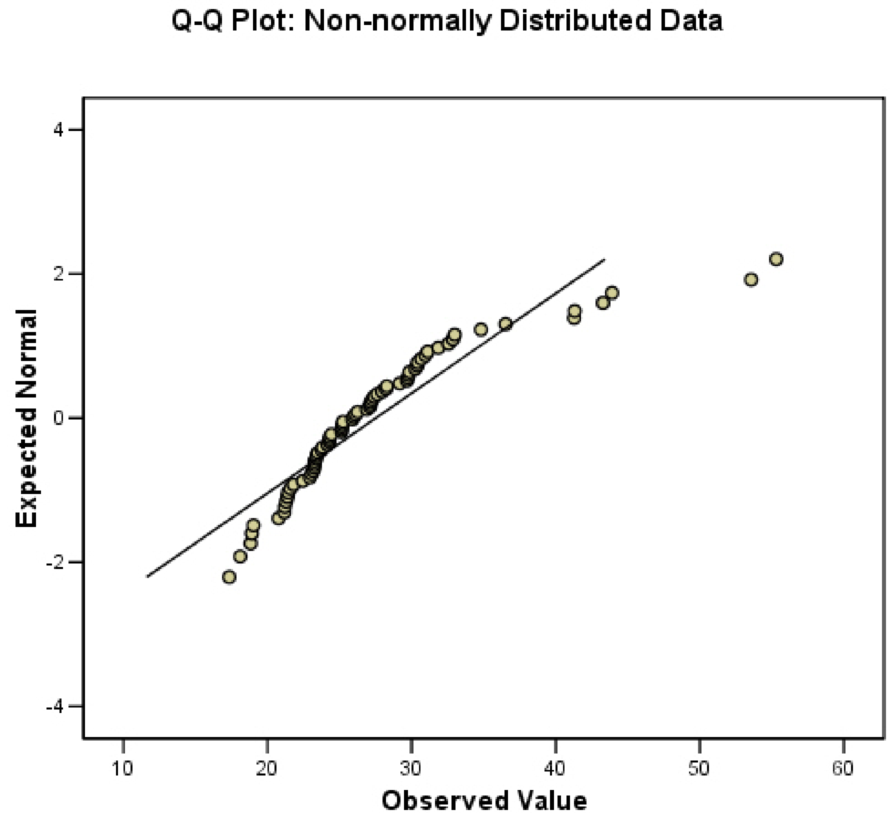
- Q-Q probability plots
- Cumulative frequency (P-P) plots

Statistical

- W/S test
- Jarque-Bera test
- Shapiro-Wilks test
- Kolmogorov-Smirnov test
- D'Agostino test

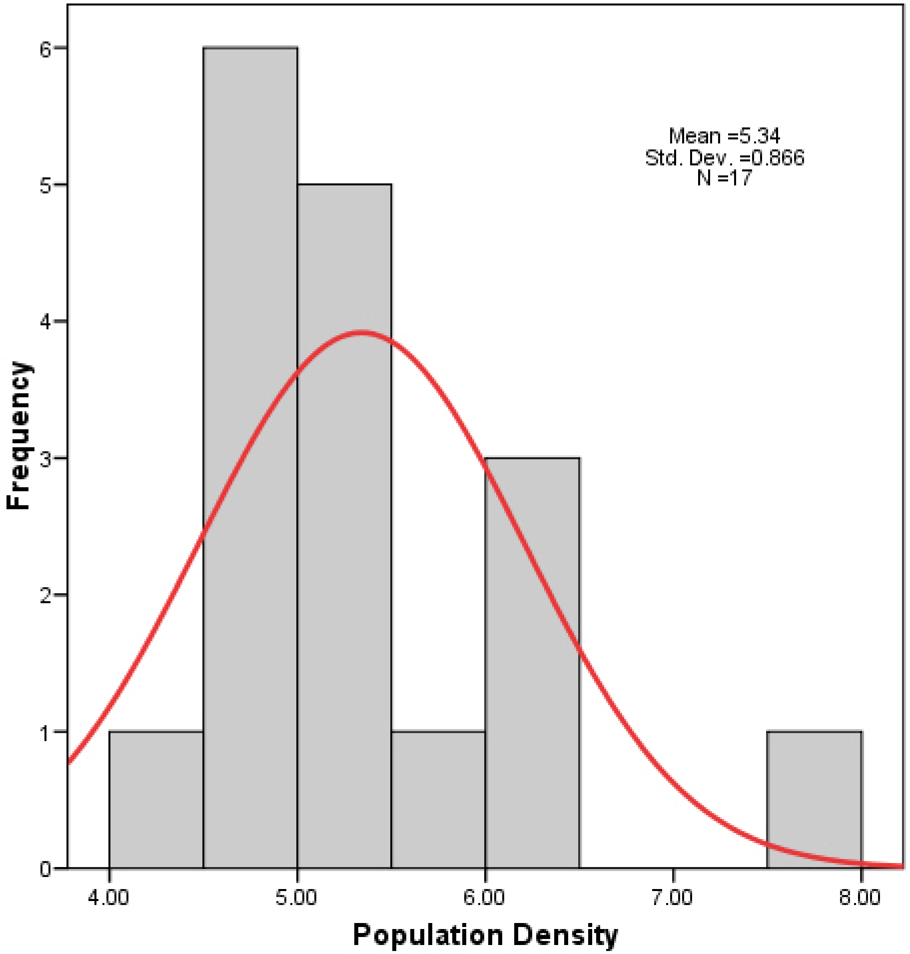
Q-Q plots display the observed values against normally distributed data (represented by the line).

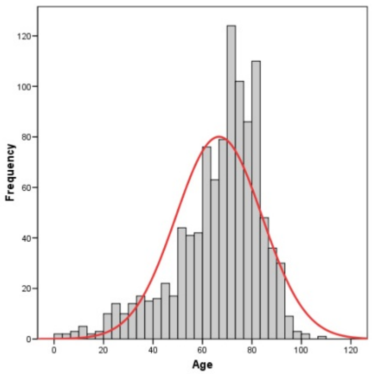


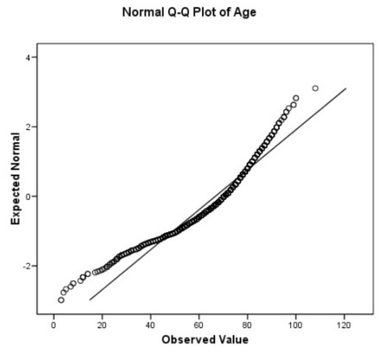


Normally distributed data fall along the line.

Graphical methods are typically not very useful when the sample size is small. This is a histogram of the last example. These data do not ‘look’ normal, but they are not statistically different than normal.



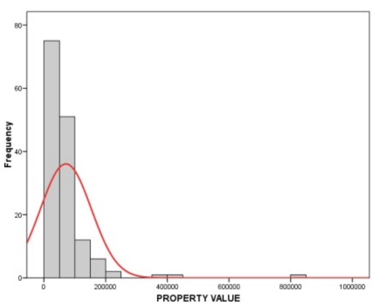


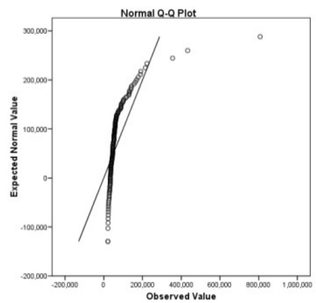


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Age	.110	1048	.000	.931	1048	.000

a. Lilliefors Significance Correction

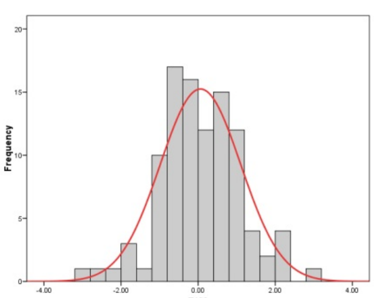


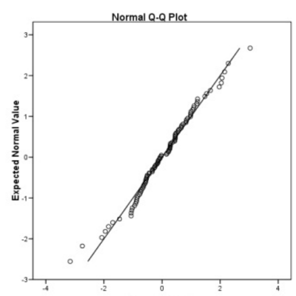


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
TOTAL_VALU	.283	149	.000	.463	149	.000

a. Lilliefors Significance Correction





Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Z100	.071	100	.200*	.985	100	.333

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Statistical tests for normality are more precise since actual probabilities are calculated.

Tests for normality calculate the probability that the sample was drawn from a normal population.

The hypotheses used are:

H_o: The sample data are not significantly different than a normal population.

H_a: The sample data are significantly different than a normal population.

Typically, we are interested in finding a difference between groups. When we are, we 'look' for small probabilities.

- If the probability of finding an event is rare (less than 5%) and we actually find it, that is of interest.

When testing normality, we are not 'looking' for a difference.

- In effect, we want our data set to be NO DIFFERENT than normal. We want to accept the null hypothesis.

So when testing for normality:

- Probabilities > 0.05 mean the data are normal.

- Probabilities < 0.05 mean the data are NOT normal.

Remember that LARGE probabilities denote normally distributed data. Below are examples taken from SPSS.

Normally Distributed Data

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Asthma Cases	.069	72	.200*	.988	72	.721

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Non-Normally Distributed Data

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Average PM10	.142	72	.001	.841	72	.000

a. Lilliefors Significance Correction

Normally Distributed Data

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Asthma Cases	.069	72	.200*	.988	72	.721

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

In SPSS output above the probabilities are greater than 0.05 (the typical alpha level), so we accept H_o… these data are not different from normal.

Non-Normally Distributed Data

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Average PM10	.142	72	.001	.841	72	.000

a. Lilliefors Significance Correction

In the SPSS output above the probabilities are less than 0.05 (the typical alpha level), so we reject H_o… these data are significantly different from normal.

Important: As the sample size ***increases***, normality parameters becomes ***MORE*** restrictive and it becomes harder to declare that the data are normally distributed. So for very large data sets, normality testing becomes less important.

Three Simple Tests for Normality

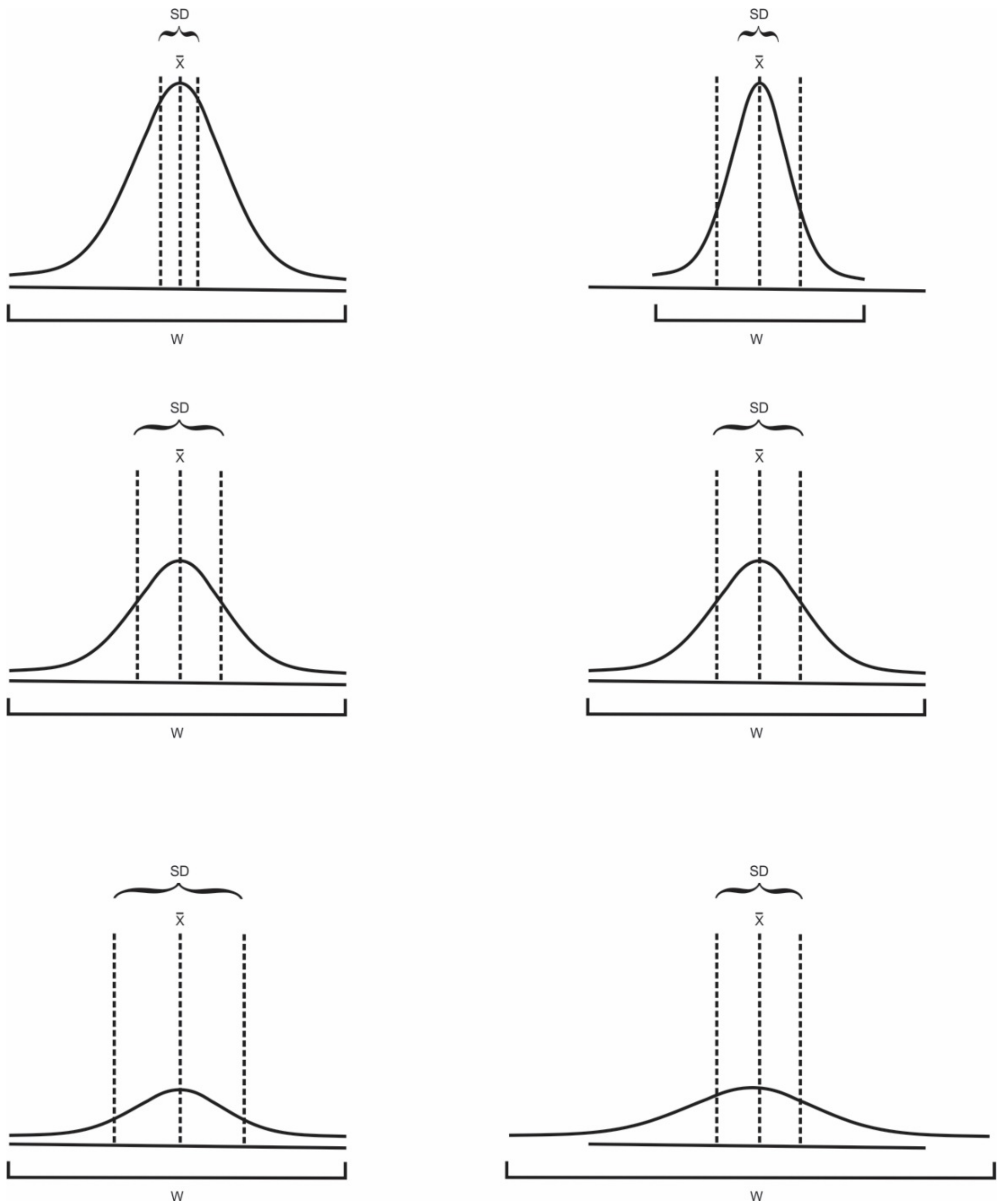
W/S Test for Normality

- A fairly simple test that requires only the sample standard deviation and the data range.
- Should not be confused with the Shapiro-Wilk test.
- Based on the q statistic, which is the 'studentized' (meaning t distribution) range, or the range expressed in standard deviation units. Tests kurtosis.

$$q = \frac{w}{s}$$

where q is the test statistic, w is the range of the data and s is the standard deviation.

Range constant,
SD changes



Range changes,
SD constant

Village	Population Density
Aranza	4.13
Corupo	4.53
San Lorenzo	4.69
Cheranatzicurin	4.76
Nahuatzen	4.77
Pomacuaran	4.96
Sevina	4.97
Arantepacua	5.00
Cocucho	5.04
Charapan	5.10
Comachuen	5.25
Pichataro	5.36
Quinceo	5.94
Nurio	6.06
Turicuaro	6.19
Urapicho	6.30
Capacuaro	7.73

Standard deviation (s) = 0.866
Range (w) = 3.6
n = 17

$$q = \frac{w}{s}$$

$$q = \frac{3.6}{0.866} = 4.16$$

$$q_{Critical\ Range} = 3.06\ to\ 4.31$$

The W/S test uses a critical range. IF the calculated value falls WITHIN the range, then accept H_o. IF the calculated value falls outside the range then reject H_o.

Since 3.06 < *q=4.16* < 4.31, then we accept H_o.

Critical Values of q for the W/S Normality Test
Taken from Kanji, 1994 Table 14

Columns *a* denote the lower boundaries or the left-sided critical values.
Columns *b* denote the upper boundaries or the right-sided critical values.

n	Level of significance α											
	0.000		0.005		0.01		0.025		0.05		0.10	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
3	1.732	2.000	1.735	2.000	1.737	2.000	1.745	2.000	1.758	1.999	1.782	1.997
4	1.732	2.449	1.82	2.447	1.87	2.445	1.93	2.439	1.98	2.429	2.04	2.409
5	1.826	2.828	1.98	2.813	2.02	2.803	2.09	2.782	2.15	2.753	2.22	2.712
6	1.826	3.162	2.11	3.115	2.15	3.095	2.22	3.056	2.28	3.012	2.37	2.949
7	1.871	3.464	2.22	3.369	2.26	3.338	2.33	3.282	2.40	3.222	2.49	3.143
8	1.871	3.742	2.31	3.585	2.35	3.543	2.43	3.471	2.50	3.399	2.59	3.308
9	1.897	4.000	2.39	3.772	2.44	3.720	2.51	3.634	2.59	3.552	2.68	3.449
10	1.897	4.243	2.46	3.935	2.51	3.875	2.59	3.777	2.67	3.685	2.76	3.57
11	1.915	4.472	2.53	4.079	2.58	4.012	2.66	3.903	2.74	3.80	2.84	3.68
12	1.915	4.690	2.59	4.208	2.64	4.134	2.72	4.02	2.80	3.91	2.90	3.78
13	1.927	4.899	2.64	4.325	2.70	4.244	2.78	4.12	2.86	4.00	2.96	3.87
14	1.927	5.099	2.70	4.431	2.75	4.34	2.83	4.21	2.92	4.09	3.02	3.95
15	1.936	5.292	2.74	4.53	2.80	4.44	2.88	4.29	2.97	4.17	3.07	4.02
16	1.936	5.477	2.79	4.62	2.84	4.52	2.93	4.37	3.01	4.24	3.12	4.09
17	1.944	5.657	2.83	4.70	2.88	4.60	2.97	4.44	3.06	4.31	3.17	4.15
18	1.944	5.831	2.87	4.78	2.92	4.67	3.01	4.51	3.10	4.37	3.21	4.21
19	1.949	6.000	2.90	4.85	2.96	4.74	3.05	4.56	3.14	4.43	3.25	4.27
20	1.949	6.164	2.94	4.91	2.99	4.80	3.09	4.63	3.18	4.49	3.29	4.32
25	1.961	6.93	3.09	5.19	3.15	5.06	3.24	4.87	3.34	4.71	3.45	4.53
30	1.966	7.62	3.21	5.40	3.27	5.26	3.37	5.06	3.47	4.89	3.59	4.70
35	1.972	8.25	3.32	5.57	3.38	5.42	3.48	5.21	3.58	5.04	3.70	4.84
40	1.975	8.83	3.41	5.71	3.47	5.56	3.57	5.34	3.67	5.16	3.79	4.96
45	1.978	9.38	3.49	5.83	3.55	5.67	3.66	5.45	3.75	5.26	3.88	5.06
50	1.980	9.90	3.56	5.93	3.62	5.77	3.73	5.54	3.83	5.35	3.95	5.14
55	1.982	10.39	3.62	6.02	3.69	5.86	3.80	5.63	3.90	5.43	4.02	5.22
60	1.983	10.86	3.68	6.10	3.75	5.94	3.86	5.70	3.96	5.51	4.08	5.29
65	1.985	11.31	3.74	6.17	3.80	6.01	3.91	5.77	4.01	5.57	4.14	5.35
70	1.986	11.75	3.79	6.24	3.85	6.07	3.96	5.83	4.06	5.63	4.19	5.41
75	1.987	12.17	3.83	6.30	3.90	6.13	4.01	5.88	4.11	5.68	4.24	5.46
80	1.987	12.57	3.88	6.35	3.94	6.18	4.05	5.93	4.16	5.73	4.28	5.51
85	1.988	12.96	3.92	6.40	3.99	6.23	4.09	5.98	4.20	5.78	4.33	5.56
90	1.989	13.34	3.96	6.45	4.02	6.27	4.13	6.03	4.24	5.82	4.36	5.60
95	1.990	13.71	3.99	6.49	4.06	6.32	4.17	6.07	4.27	5.86	4.40	5.64
100	1.990	14.07	4.03	6.53	4.10	6.36	4.21	6.11	4.31	5.90	4.44	5.68
150	1.993	17.26	4.32	6.82	4.38	6.64	4.48	6.39	4.59	6.18	4.72	5.96
200	1.995	19.95	4.53	7.01	4.59	6.84	4.68	6.60	4.78	6.39	4.90	6.15
500	1.998	31.59	5.06	7.60	5.13	7.42	5.25	7.15	5.47	6.94	5.49	6.72
1000	1.999	44.70	5.50	7.99	5.57	7.80	5.68	7.54	5.79	7.33	5.92	7.11

Source: Sachs, 1972

Since we have a critical range, it is difficult to determine a probability range for our results. Therefore we simply state our alpha level.

The sample data set is not significantly different than normal ($W/S_{4.16}$, $p > 0.05$).

Jarque–Bera Test

A goodness-of-fit test of whether sample data have the skewness and kurtosis matching a normal distribution.

$$k_3 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{ns^3} \qquad k_4 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^4}{ns^4} - 3$$

$$JB = n\left(\frac{(k_3)^2}{6} + \frac{(k_4)^2}{24}\right)$$

Where ***x*** is each observation, ***n*** is the sample size, ***s*** is the standard deviation, $\boldsymbol{k}_3$ is skewness, and $\boldsymbol{k_4}$ is kurtosis.

$\bar{x} = 5.34$

$s = 0.87$

Village	Population Density	Mean Deviates	Mean Deviates^3	Mean Deviates^4
Aranza	4.13	-1.21	-1.771561	2.14358881
Corupo	4.53	-0.81	-0.531441	0.43046721
San Lorenzo	4.69	-0.65	-0.274625	0.17850625
Cheranatzicurin	4.76	-0.58	-0.195112	0.11316496
Nahuatzen	4.77	-0.57	-0.185193	0.10556001
Pomacuaran	4.96	-0.38	-0.054872	0.02085136
Sevina	4.97	-0.37	-0.050653	0.01874161
Arantepacua	5.00	-0.34	-0.039304	0.01336336
Cocucho	5.04	-0.30	-0.027000	0.00810000
Charapan	5.10	-0.24	-0.013824	0.00331776
Comachuen	5.25	-0.09	-0.000729	0.00006561
Pichataro	5.36	0.02	0.000008	0.00000016
Quinceo	5.94	0.60	0.216000	0.12960000
Nurio	6.06	0.72	0.373248	0.26873856
Turicuaro	6.19	0.85	0.614125	0.52200625
Urapicho	6.30	0.96	0.884736	0.84934656
Capacuaro	7.73	2.39	13.651919	32.62808641
			12.595722	37.433505

$$k_3 = \frac{12.6}{(17)(0.87^3)} = 1.13 \qquad k_4 = \frac{37.43}{(17)(0.87^4)} - 3 = 0.843$$

$$JB = 17\left(\frac{(1.13)^2}{6} + \frac{(0.843)^2}{24}\right) = 17\left(\frac{1.2769}{6} + \frac{0.711}{24}\right)$$

$$JB = 17(0.2128 + 0.0296)$$

$$JB = 4.12$$

The Jarque-Bera statistic can be compared to the χ^2 distribution (table) with 2 degrees of freedom (df or ***v***) to determine the critical value at an alpha level of 0.05.

Critical Values of the Chi-Squared (χ^2) Distribution

Taken from Rholf and Sokal Table 14

$\nu \backslash \alpha$	.995	.975	.9	.5	.1	.05	.025	.01	.005	.001	α / ν
1	0.000	0.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	10.828	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	13.816	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	16.266	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	18.467	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	20.515	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	22.458	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	24.322	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	26.124	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	27.877	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	29.588	10

The critical χ^2 value is 5.991. Our calculated Jarque-Bera statistic is 4.12 which falls between 0.5 and 0.1, which is greater than the critical value.

Therefore we accept H_o that there is no difference between our distribution and a normal distribution (Jarque-Bera $\chi^2_{4.12}$, 0.5 > p > 0.1).

D'Agostino Test

- A very powerful test for departures from normality.

- Based on the D statistic, which gives an upper and lower critical value.

$$D = \frac{T}{\sqrt{n^3 SS}} \qquad where \qquad T = \sum\left(i - \frac{n+1}{2}\right)X_i$$

where D is the test statistic, SS is the sum of squares of the data and n is the sample size, and *i* is the order or rank of observation *X*. The df for this test is n (sample size).

- First the data are ordered from smallest to largest or largest to smallest.

Village	Population Density	i	Mean Deviates^2
Aranza	$\boxed{4.13}$	1	1.46410
Corupo	4.53	2	0.65610
San Lorenzo	4.69	3	0.42250
Cheranatzicurin	4.76	4	0.33640
Nahuatzen	4.77	5	0.32490
Pomacuaran	4.96	6	0.14440
Sevina	4.97	7	0.13690
Arantepacua	5.00	8	0.11560
Cocucho	5.04	9	0.09000
Charapan	5.10	10	0.05760
Comachuen	5.25	11	0.00810
Pichataro	5.36	12	0.00040
Quinceo	5.94	13	0.36000
Nurio	6.06	14	0.51840
Turicuaro	6.19	15	0.72250
Urapicho	6.30	16	0.92160
Capacuaro	7.73	17	5.71210
		Mean = $\boxed{5.34}$	SS = $\boxed{11.9916}$

$$(4.13-5.34)^2 = -1.21^2 = 1.46410$$

$$\frac{n+1}{2} = \frac{17+1}{2} = 9$$

$$T = \sum (i-9)X_1$$

$$T = (1-9)4.13 + (2-9)4.53 + (3-9)4.69 + \ldots(17-9)7.73$$

$$T = 63.23$$

$$D = \frac{63.23}{\sqrt{(17^3)(11.9916)}} = 0.26050$$

$$D_{Critical} = 0.2587, 0.2860$$

Df = n = 17

If the calculated value falls within the critical range, accept H_o.

Since 0.2587 < *D = 0.26050* < 0.2860 accept H_o.

The sample data set is not significantly different than normal ($D_{0.26050}$, $p > 0.05$).

Use the next lower *n* on the table if your sample size is NOT listed.

Critical Values for the D'Agostino D Normality Test
Taken from Zar, 1981 Table B.22

n	$\alpha =$ 0.20	0.10	0.05	0.02	0.01
10	0.2632, 0.2835	0.2573, 0.2843	0.2513, 0.2849	0.2436, 0.2855	0.2379, 0.2857
12	0.2653, 0.2841	0.2598, 0.2849	0.2544, 0.2854	0.2473, 0.2859	0.2420, 0.2862
14	0.2669, 0.2846	0.2618, 0.2853	0.2568, 0.2858	0.2503, 0.2862	0.2455, 0.2865
16	0.2681, 0.2848	0.2634, 0.2855	$\boxed{0.2587, 0.2860}$	0.2527, 0.2865	0.2482, 0.2867
18	0.2690, 0.2850	0.2646, 0.2855	0.2603, 0.2862	0.2547, 0.2866	0.2505, 0.2868
20	0.2699, 0.2852	0.2657, 0.2857	0.2617, 0.2863	0.2564, 0.2867	0.2525, 0.2869
22	0.2705, 0.2853	0.2670, 0.2859	0.2629, 0.2864	0.2579, 0.2869	0.2542, 0.2870
24	0.2711, 0.2853	0.2675, 0.2860	0.2638, 0.2865	0.2591, 0.2870	0.2557, 0.2871
26	0.2717, 0.2854	0.2682, 0.2861	0.2647, 0.2866	0.2603, 0.2870	0.2570, 0.2872
28	0.2721, 0.2854	0.2688, 0.2861	0.2655, 0.2866	0.2612, 0.2870	0.2581, 0.2873
30	0.2725, 0.2854	0.2693, 0.2861	0.2662, 0.2866	0.2622, 0.2871	0.2592, 0.2872
32	0.2729, 0.2854	0.2698, 0.2862	0.2668, 0.2867	0.2630, 0.2871	0.2600, 0.2873
34	0.2732, 0.2854	0.2703, 0.2862	0.2674, 0.2867	0.2636, 0.2871	0.2609, 0.2873
36	0.2735, 0.2854	0.2707, 0.2862	0.2679, 0.2867	0.2643, 0.2871	0.2617, 0.2873
38	0.2738, 0.2854	0.2710, 0.2862	0.2683, 0.2867	0.2649, 0.2871	0.2623, 0.2873
40	0.2740, 0.2854	0.2714, 0.2862	0.2688, 0.2867	0.2655, 0.2871	0.2630, 0.2874
42	0.2743, 0.2854	0.2717, 0.2861	0.2691, 0.2867	0.2659, 0.2871	0.2636, 0.2874
44	0.2745, 0.2854	0.2720, 0.2861	0.2695, 0.2867	0.2664, 0.2871	0.2641, 0.2874
46	0.2747, 0.2854	0.2722, 0.2861	0.2698, 0.2866	0.2668, 0.2871	0.2646, 0.2874
48	0.2749, 0.2854	0.2725, 0.2861	0.2702, 0.2866	0.2672, 0.2871	0.2651, 0.2874
50	0.2751, 0.2853	0.2727, 0.2861	0.2705, 0.2866	0.2676, 0.2871	0.2655, 0.2874
60	0.2757, 0.2852	0.2737, 0.2860	0.2717, 0.2865	0.2692, 0.2870	0.2673, 0.2873
70	0.2763, 0.2851	0.2744, 0.2859	0.2726, 0.2864	0.2708, 0.2869	0.2687, 0.2872
80	0.2768, 0.2850	0.2750, 0.2857	0.2734, 0.2863	0.2713, 0.2868	0.2698, 0.2871
90	0.2771, 0.2849	0.2755, 0.2856	0.2740, 0.2862	0.2721, 0.2866	0.2707, 0.2870
100	0.2774, 0.2849	0.2759, 0.2855	0.2745, 0.2860	0.2727, 0.2865	0.2714, 0.2869
120	0.2779, 0.2847	0.2765, 0.2853	0.2752, 0.2858	0.2737, 0.2863	0.2725, 0.2866
140	0.2782, 0.2846	0.2770, 0.2852	0.2758, 0.2856	0.2744, 0.2862	0.2734, 0.2865
160	0.2785, 0.2845	0.2774, 0.2851	0.2763, 0.2855	0.2750, 0.2860	0.2741, 0.2863
180	0.2787, 0.2844	0.2777, 0.2850	0.2767, 0.2854	0.2755, 0.2859	0.2746, 0.2862
200	0.2789, 0.2843	0.2779, 0.2848	0.2770, 0.2853	0.2759, 0.2857	0.2751, 0.2860
250	0.2793, 0.2841	0.2784, 0.2846	0.2776, 0.2850	0.2767, 0.2855	0.2760, 0.2858
300	0.2796, 0.2840	0.2788, 0.2844	0.2781, 0.2848	0.2772, 0.2853	0.2766, 0.2855
350	0.2798, 0.2839	0.2791, 0.2843	0.2784, 0.2847	0.2776, 0.2851	0.2771, 0.2853
400	0.2799, 0.2838	0.2793, 0.2842	0.2787, 0.2845	0.2780, 0.2849	0.2775, 0.2852
450	0.2801, 0.2837	0.2795, 0.2841	0.2789, 0.2844	0.2782, 0.2848	0.2778, 0.2851
500	0.2802, 0.2836	0.2796, 0.2840	0.2791, 0.2843	0.2785, 0.2847	0.2780, 0.2849
600	0.2804, 0.2835	0.2799, 0.2839	0.2794, 0.2842	0.2788, 0.2845	0.2784, 0.2847
700	0.2805, 0.2834	0.2800, 0.2838	0.2796, 0.2840	0.2791, 0.2844	0.2787, 0.2846
800	0.2806, 0.2833	0.2802, 0.2837	0.2798, 0.2839	0.2793, 0.2842	0.2790, 0.2844
900	0.2807, 0.2833	0.2803, 0.2836	0.2799, 0.2838	0.2795, 0.2841	0.2792, 0.2843
1000	0.2808, 0.2832	0.2804, 0.2835	0.2800, 0.2838	0.2796, 0.2840	0.2793, 0.2842
1250	0.2809, 0.2831	0.2806, 0.2834	0.2803, 0.2836	0.2799, 0.2839	0.2797, 0.2840
1500	0.2810, 0.2830	0.2807, 0.2833	0.2805, 0.2835	0.2801, 0.2837	0.2799, 0.2839
1750	0.2811, 0.2830	0.2808, 0.2832	0.2806, 0.2834	0.2803, 0.2836	0.2801, 0.2838
2000	0.2812, 0.2829	0.2809, 0.2831	0.2807, 0.2833	0.2804, 0.2835	0.2802, 0.2837

For each significance level, α is given a pair of critical values. If the calculated D is $\leq$ the first member of the pair, or $\geq$ the second, then, the null hypothesis of population normality is rejected.

Different normality tests produce vastly different probabilities. This is due to where in the distribution (central, tails) or what moment (skewness, kurtosis) they are examining.

Normality Test	**Statistic**	**Calculated Value**	**Probability**	**Results**
W/S	q	4.16	> 0.05	Normal
Jarque-Bera	χ^2	4.15	$0.5 > p > 0.1$	Normal
D’Agostino	D	0.2605	> 0.05	Normal
Shapiro-Wilk	W	0.8827	0.035	Non-normal
Kolmogorov-Smirnov	D	0.2007	0.067	Normal

Normality tests using various random normal sample sizes:

Sample Size	K-S Prob
20	1.000
50	0.893
100	0.871
200	0.611
500	0.969
1000	0.904
2000	0.510
5000	0.106
10000	0.007
15000	0.005
20000	0.007
50000	0.002

Notice that as the sample size increases, the probabilities decrease. In other words, ***it gets harder*** to meet the normality assumption as the sample size increases since even small departures from normality are detected.

Which normality test should I use?

Kolmogorov-Smirnov:

- Not sensitive to problems in the tails.
- For data sets > 50.

Shapiro-Wilks:

- Doesn't work well if several values in the data set are the same.
- Works best for data sets with < 50, but can be used with larger data sets.

W/S:

- Simple, but effective.
- Not available in SPSS.

Jarque-Bera:

- Tests for skewness and kurtosis, very effective.
- Not available in SPSS.

D’Agostino:

- Powerful omnibus (skewness, kurtosis, centrality) test.
- Not available in SPSS.

When is non-normality a problem?

- Normality can be a problem when the sample size is small (< 50).
- Highly skewed data create problems.
- Highly leptokurtic data are problematic, but not as much as skewed data.
- Normality becomes a serious concern when there is “activity” in the tails of the data set.
    - Outliers are a problem.
    - “Clumps” of data in the tails are worse.

Testing for Outliers

Grubbs Test: $G_{Max} = \frac{x_n - \bar{x}}{s}$ *or* $G_{Min} = \frac{\bar{x} - x_n}{s}$

$df = n$

where x_n is the suspected outlier, $\bar{x}$ is the mean, and s is the standard deviation. G_{Max} is used when the suspect observation is greater than the mean and G_{Min} is used when it is less than the mean.

Ho: The suspected outlier is not different than the sample distribution.
Ha: The suspected outlier is different than the sample distribution.

Obs
15
7
6
6
5
5
5
4
4
3

$$G_{Max} = \frac{15 - 6}{3.37} = 2.671$$

The critical value for an n = 10 from Grubbs modified t table (G table) at an α = 0.05 is 2.18.

Since 2.671 > 2.18, reject Ho.

The suspected outlier is from a significantly different sample population (G_{Max}, 2.671, $p < 0.01$).

Critical Values of Grubb's Outlier (G) Test
Taken from Grubb 1969, Table 1

N	α=0.05	α=0.025	α=0.01
3	1.15	1.15	1.15
4	1.46	1.48	1.49
5	1.67	1.71	1.75
6	1.82	1.89	1.94
7	1.94	2.02	2.10
8	2.03	2.13	2.22
9	2.11	2.21	2.32
10	2.18	2.29	2.41
11	2.23	2.36	2.48

Testing for Outliers

Dixon Test: $Q = \frac{x_n - x_{n-1}}{x_n - x_1}$

$df = n$

where x_n is the suspected outlier, x_{n-1} is the next ranked observation, and x_1 is the last ranked observation.

Ho: The suspected outlier is not different than the sample distribution.
Ha: The suspected outlier is different than the sample distribution.

Obs
15
7
6
6
5
5
5
4
4
3

$$Q = \frac{15 - 7}{15 - 3} = 0.6667$$

The critical value for an n = 10 from Verma and Quiroz-Ruiz expanded Dixon table at an $\alpha = 0.05$ is 0.4122. Since 0.6667 > 0.4122, reject H_o.

The suspected outlier is from a significantly different sample population ($Q_{0.6667}$, $p < 0.005$).

Critical Values of Expanded Dixon Outlier Test

Taken from Verma and Quiroz-Ruiz, Table 2

n	CL	70%	80%	90%	**95%**	98%	**99%**	99.5%
	SL	30%	20%	10%	**5%**	2%	**1%**	0.5%
	α	0.30	0.20	0.10	**0.05**	0.02	**0.01**	0.005
3		0.6836	0.7808	0.8850	0.9411	0.9763	0.9881	0.9940
4		0.4704	0.5603	0.6789	0.7651	0.8457	0.8886	0.9201
5		0.3730	0.4508	0.5578	0.6423	0.7291	0.7819	0.8234
6		0.3173	0.3868	0.4840	0.5624	0.6458	0.6987	0.7437
7		0.2811	0.3444	0.4340	0.5077	0.5864	0.6371	0.6809
8		0.2550	0.3138	0.3979	0.4673	0.5432	0.5914	0.6336
9		0.2361	0.2915	0.3704	0.4363	0.5091	0.5554	0.5952
10		0.2208	0.2735	0.3492	$\boxed{0.4122}$	0.4813	0.5260	0.5658
11		0.2086	0.2586	0.3312	0.3922	0.4591	0.5028	0.5416
12		0.1983	0.2467	0.3170	0.3755	0.4405	0.4831	0.5208

These tests have several requirements:

1) The data are from a normal distribution
2) There are not multiple outliers (3+),
3) The data are sorted with the suspected outlier first.

If 2 observations are suspected as being outliers and both lie on the same side of the mean, this test can be performed again after removing the first outlier from the data set.

Caution must be used when removing outliers. Only remove outliers if you suspect the value was caused by an error of some sort, or if you have evidence that the value truly belongs to a different population.

If you have a small sample size, extreme caution should be used when removing any data.

Final Words Concerning Normality Testing:

1. Since it IS a test, state a null and alternate hypothesis.

2. If you perform a normality test, do not ignore the results.

3. If the data are not normal, use non-parametric tests.

4. If the data are normal, use parametric tests.

AND MOST IMPORTANTLY:

5. **If you have groups of data, you MUST test each group for normality.**