

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD.

CHE384, From Data to Decisions: Measurement, Uncertainty, Analysis, and Modeling

## Lecture 16

### Shapiro-Wilk Test for Normality

Chris A. Mack  
Adjunct Associate Professor

<http://www.lithoguru.com/scientist/statistics/>

© Chris Mack, 2016 Data to Decisions 1

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD.

## Shapiro-Wilk Test

- **Null Hypothesis:** The sample (of size  $n$ ) comes from a normal distribution
- Test statistic:
 
$$W = \frac{(\sum_{i=1}^n a_i x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \begin{array}{l} x_i = i^{th} \text{ smallest value of } x \\ a_i = \text{Shapiro - Wilk constant} \end{array}$$
- Reject null hypothesis if  $W < W_\alpha$  (p-value  $< \alpha$ )

S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)", *Biometrika*, 52(3-4), 591-611 (1965).

© Chris Mack, 2016 Data to Decisions 2

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD.

## Shapiro-Wilk Coefficients Table

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
a1	0.7071	0.7071	0.4872	0.4444	0.4431	0.4233	0.4092	0.3988	0.3919	0.3861	0.3815	0.3779	0.3750	0.3727	0.3708	0.3692	0.3678	0.3665	0.3653	0.3642	0.3632	0.3622	0.3612	0.3602	0.3592	0.3582
a2		0.1677	0.2413	0.2806	0.3031	0.3184	0.3284	0.3351	0.3393	0.3421	0.3438	0.3447	0.3450	0.3450	0.3447	0.3441	0.3434	0.3425	0.3414	0.3401	0.3387	0.3372	0.3356	0.3339	0.3321	0.3302
a3			0.0875	0.1403	0.1748	0.1979	0.2161	0.2290	0.2367	0.2421	0.2450	0.2465	0.2468	0.2460	0.2449	0.2435	0.2418	0.2398	0.2375	0.2349	0.2321	0.2291	0.2259	0.2225	0.2189	0.2151
a4				0.0596	0.0967	0.1224	0.1429	0.1588	0.1707	0.1780	0.1821	0.1839	0.1843	0.1835	0.1815	0.1792	0.1757	0.1719	0.1678	0.1634	0.1587	0.1537	0.1484	0.1428	0.1369	0.1307
a5					0.0399	0.0695	0.0922	0.1099	0.1232	0.1321	0.1367	0.1382	0.1377	0.1353	0.1319	0.1274	0.1228	0.1180	0.1130	0.1078	0.1023	0.0965	0.0904	0.0840	0.0773	0.0704
a6						0.0293	0.0539	0.0722	0.0842	0.0917	0.0950	0.0953	0.0937	0.0903	0.0859	0.0813	0.0765	0.0715	0.0663	0.0609	0.0553	0.0495	0.0435	0.0373	0.0309	0.0243
a7							0.0200	0.0399	0.0542	0.0627	0.0669	0.0679	0.0663	0.0629	0.0585	0.0538	0.0489	0.0438	0.0385	0.0330	0.0273	0.0214	0.0153	0.0091	0.0028	0.0000

For more tables, see <http://www.real-statistics.com/statistics-tables/shapiro-wilk-table/>

© Chris Mack, 2016 Data to Decisions 3

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD.

## Shapiro-Wilk Critical Values

[www.real-statistics.com/statistics-tables/shapiro-wilk-table/](http://www.real-statistics.com/statistics-tables/shapiro-wilk-table/)

n	0.01	0.02	0.05	0.10	0.20	0.25	0.50	0.75	0.90	0.95	0.975	0.99
1	0.7071	0.7071	0.4872	0.4444	0.4431	0.4233	0.4092	0.3988	0.3919	0.3861	0.3815	0.3779
2	0.6897	0.6897	0.4798	0.4370	0.4357	0.4159	0.4018	0.3914	0.3845	0.3787	0.3741	0.3705
3	0.6713	0.6713	0.4699	0.4271	0.4258	0.4060	0.3919	0.3815	0.3746	0.3688	0.3642	0.3606
4	0.6529	0.6529	0.4685	0.4257	0.4244	0.4046	0.3905	0.3801	0.3732	0.3674	0.3628	0.3592
5	0.6345	0.6345	0.4671	0.4243	0.4230	0.4032	0.3891	0.3787	0.3718	0.3660	0.3614	0.3578
6	0.6161	0.6161	0.4657	0.4229	0.4216	0.4018	0.3877	0.3773	0.3704	0.3646	0.3600	0.3564
7	0.5977	0.5977	0.4643	0.4215	0.4202	0.4004	0.3863	0.3759	0.3690	0.3632	0.3586	0.3550
8	0.5793	0.5793	0.4629	0.4201	0.4188	0.3990	0.3849	0.3745	0.3676	0.3618	0.3572	0.3536
9	0.5609	0.5609	0.4615	0.4187	0.4174	0.3976	0.3835	0.3731	0.3662	0.3604	0.3558	0.3522
10	0.5425	0.5425	0.4601	0.4173	0.4160	0.3962	0.3821	0.3717	0.3648	0.3590	0.3544	0.3508
11	0.5241	0.5241	0.4587	0.4159	0.4146	0.3948	0.3807	0.3703	0.3634	0.3576	0.3530	0.3494
12	0.5057	0.5057	0.4573	0.4145	0.4132	0.3934	0.3793	0.3689	0.3620	0.3562	0.3516	0.3480
13	0.4873	0.4873	0.4559	0.4131	0.4118	0.3920	0.3779	0.3675	0.3606	0.3548	0.3502	0.3466
14	0.4689	0.4689	0.4545	0.4117	0.4104	0.3906	0.3765	0.3661	0.3592	0.3534	0.3488	0.3452
15	0.4505	0.4505	0.4531	0.4103	0.4090	0.3892	0.3751	0.3647	0.3578	0.3520	0.3474	0.3438
16	0.4321	0.4321	0.4517	0.4089	0.4076	0.3878	0.3737	0.3633	0.3564	0.3506	0.3460	0.3424
17	0.4137	0.4137	0.4503	0.4081	0.4068	0.3870	0.3729	0.3625	0.3556	0.3498	0.3452	0.3416
18	0.3953	0.3953	0.4489	0.4071	0.4058	0.3860	0.3719	0.3615	0.3546	0.3488	0.3442	0.3406
19	0.3769	0.3769	0.4475	0.4063	0.4050	0.3852	0.3711	0.3607	0.3538	0.3480	0.3434	0.3398
20	0.3585	0.3585	0.4461	0.4053	0.4040	0.3842	0.3701	0.3597	0.3528	0.3470	0.3424	0.3388
21	0.3401	0.3401	0.4447	0.4045	0.4032	0.3834	0.3693	0.3589	0.3520	0.3462	0.3416	0.3380
22	0.3217	0.3217	0.4433	0.4037	0.4024	0.3826	0.3685	0.3581	0.3512	0.3454	0.3408	0.3372
23	0.3033	0.3033	0.4419	0.4029	0.4016	0.3818	0.3677	0.3573	0.3504	0.3446	0.3400	0.3364
24	0.2849	0.2849	0.4405	0.4021	0.4008	0.3810	0.3669	0.3565	0.3496	0.3438	0.3392	0.3356
25	0.2665	0.2665	0.4391	0.4013	0.4000	0.3802	0.3661	0.3557	0.3488	0.3430	0.3384	0.3348
26	0.2481	0.2481	0.4377	0.4005	0.3992	0.3794	0.3653	0.3549	0.3480	0.3422	0.3376	0.3340
27	0.2297	0.2297	0.4363	0.3997	0.3984	0.3786	0.3645	0.3541	0.3472	0.3414	0.3368	0.3332
28	0.2113	0.2113	0.4349	0.3989	0.3976	0.3780	0.3639	0.3535	0.3466	0.3408	0.3362	0.3326
29	0.1929	0.1929	0.4335	0.3981	0.3968	0.3782	0.3641	0.3537	0.3468	0.3410	0.3364	0.3328
30	0.1745	0.1745	0.4321	0.3973	0.3960	0.3774	0.3633	0.3529	0.3460	0.3402	0.3356	0.3320
31	0.1561	0.1561	0.4307	0.3965	0.3952	0.3766	0.3625	0.3521	0.3452	0.3394	0.3348	0.3312
32	0.1377	0.1377	0.4293	0.3957	0.3944	0.3758	0.3617	0.3513	0.3444	0.3386	0.3340	0.3304
33	0.1193	0.1193	0.4279	0.3949	0.3936	0.3752	0.3611	0.3507	0.3438	0.3380	0.3334	0.3298
34	0.1009	0.1009	0.4265	0.3941	0.3928	0.3744	0.3603	0.3499	0.3430	0.3372	0.3326	0.3290
35	0.0825	0.0825	0.4251	0.3933	0.3920	0.3736	0.3595	0.3491	0.3422	0.3364	0.3318	0.3282
36	0.0641	0.0641	0.4237	0.3925	0.3912	0.3728	0.3587	0.3483	0.3414	0.3356	0.3310	0.3274
37	0.0457	0.0457	0.4223	0.3917	0.3904	0.3720	0.3579	0.3475	0.3406	0.3348	0.3302	0.3266
38	0.0273	0.0273	0.4209	0.3909	0.3896	0.3712	0.3571	0.3467	0.3398	0.3340	0.3294	0.3258
39	0.0089	0.0089	0.4195	0.3901	0.3888	0.3704	0.3563	0.3459	0.3390	0.3332	0.3286	0.3250
40	0.0000	0.0000	0.4181	0.3893	0.3880	0.3696	0.3555	0.3451	0.3382	0.3324	0.3278	0.3242

© Chris Mack, 2016 Data to Decisions 4

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD.

## Shapiro-Wilk in R

- It is best to run the Shapiro-Wilk test using a statistical software package
- In R,
  - Let  $X = (X_1, \dots, X_n)$  be the data vector
  - `shapiro.test(X)`
    - calculates  $W$  and the p-value
    - reject the null hypothesis of a normal distribution if the p-value is less than  $\alpha$ , your significance level

© Chris Mack, 2016 Data to Decisions 5

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD.

## Properties of the Test

- $0 < W \leq 1$ , defined for  $n \geq 3$
- $W$  is very similar to the correlation coefficient of a normal probability plot
- $W$  is independent of location and scale (add or multiply the data by a constant and  $W$  doesn't change)
- The Shapiro-Wilk test has the **best power** for a given significance compared to other popular tests
- This test is **sample-size biased**
  - A normal probability plot should be used to confirm any test results

© Chris Mack, 2016 Data to Decisions 6

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD

## Interpreting the Test

- Failure to reject the null hypothesis is **not proof** that the distribution is Normal
  - The **Default** is that the distribution is normal
  - It means that the data does not give enough evidence to reject this default assumption
  - Small sample sizes often don't have enough information to conclude much of anything
- Rejecting the null hypothesis does not tell you **how much** the distribution differs from normal
  - For a large data set, even a very small departure from normality will trigger a rejection

© Chris Mack, 2016 Data to Decisions 7

THE UNIVERSITY OF TEXAS  
AT AUSTIN

WHAT STARTS HERE CHANGES THE WORLD

## Lecture 16: What have we learned?

- Why do people like the Shapiro-Wilk test for normality?
- Be able to run the Shapiro-Wilk test in R
- What are the difficulties in interpreting the results of any hypothesis test for normality?

© Chris Mack, 2016 Data to Decisions 8