

Test di normalità

Federico Plazzi

24 Ottobre 2016

Test di normalità

Test di Shapiro e Wilk

Test di normalità

Test di Shapiro e Wilk

- ▶ **Ipotesi nulla (H_0): i nostri valori hanno una distribuzione normale.**

Test di normalità

Test di Shapiro e Wilk

- ▶ **Ipotesi nulla (H_0): i nostri valori hanno una distribuzione normale.**
- ▶ Mettiamo in ordine crescente le osservazioni;

Test di normalità

Test di Shapiro e Wilk

- ▶ **Ipotesi nulla (H_0): i nostri valori hanno una distribuzione normale.**
- ▶ Mettiamo in ordine crescente le osservazioni;
- ▶ per n osservazioni, poniamo $m = n/2$ se n è pari e $m = (n - 1)/2$ se n è dispari;

Test di normalità

Test di Shapiro e Wilk

- ▶ **Ipotesi nulla (H_0): i nostri valori hanno una distribuzione normale.**
- ▶ Mettiamo in ordine crescente le osservazioni;
- ▶ per n osservazioni, poniamo $m = n/2$ se n è pari e $m = (n - 1)/2$ se n è dispari;
- ▶ calcoliamo la statistica del test di Shapiro e Wilk, W :

$$W = \frac{(\sum_{i=1}^m k_i (X_{n+1-i} - X_i))^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (1)$$

$k \setminus n$	2	3	4	5	6	7	8	9	10	
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5868	0.5739	
2	-	0.0000	0.1677	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291	
3	-	-	-	0.0000	0.0875	0.1401	0.1743	0.1976	0.2141	
4	-	-	-	-	-	0.0000	0.0561	0.0947	0.1224	
5	-	-	-	-	-	-	-	0.0000	0.0399	
$k \setminus n$	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1506	0.1707	0.1802	0.1876	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	-	-	0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8	-	-	-	-	0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9	-	-	-	-	-	-	-	0.0163	0.0303	0.0422
10	-	-	-	-	-	-	-	-	0.0000	0.0140
$k \setminus n$	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043	0.3018	0.2992	0.2968	0.2944
3	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533	0.2522	0.2510	0.2499	0.2487
4	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151	0.2152	0.2151	0.2150	0.2148
5	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836	0.1840	0.1857	0.1864	0.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	0.1092	0.1150	0.1201	0.1245	0.1263	0.1316	0.1346	0.1372	0.1395	0.1415
8	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089	0.1128	0.1162	0.1192	0.1219
9	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876	0.0923	0.0965	0.1002	0.1036
10	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672	0.0728	0.0778	0.0822	0.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	-	-	0.0000	0.0107	0.0200	0.0284	0.0358	0.0424	0.0483	0.0537
13	-	-	-	-	0.0000	0.0094	0.0178	0.0253	0.0320	0.0381
14	-	-	-	-	-	-	0.0000	0.0084	0.0159	0.0227
15	-	-	-	-	-	-	-	-	0.0000	0.0076

Test di normalità

Test di Shapiro e Wilk

- ▶ **Ipotesi nulla (H_0): i nostri valori hanno una distribuzione normale.**
- ▶ Mettiamo in ordine crescente le osservazioni;
- ▶ per n osservazioni, poniamo $m = n/2$ se n è pari e $m = (n - 1)/2$ se n è dispari;
- ▶ calcoliamo la statistica del test di Shapiro e Wilk, W :

$$W = \frac{(\sum_{i=1}^m k_i (X_{n+1-i} - X_i))^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (1)$$

- ▶ più W è vicino a 1, più i nostri dati si distribuiscono in modo normale.

Test di normalità

Test di Shapiro e Wilk

- ▶ Per sapere *quanto* il nostro W è vicino ad 1, cerchiamo sulle tavole con quale probabilità si ottiene lo stesso valore *per caso* quando l'ipotesi nulla è *vera*, cioè quando la distribuzione *reale* dei nostri dati è effettivamente normale.

Table 6. Percentage points of the W test* for $n = 3(1)50$

n	Level								
	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
4	.687	.707	.748	.792	.935	.987	.992	.996	.997
5	.686	.715	.762	.806	.927	.979	.986	.991	.993
6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
7	.730	.760	.803	.838	.928	.972	.979	.985	.988
8	.749	.778	.818	.851	.932	.972	.978	.984	.987
9	.764	.791	.829	.859	.935	.972	.978	.984	.986
10	.781	.806	.842	.869	.938	.972	.978	.983	.986
11	0.792	0.817	0.850	0.876	0.940	0.973	0.979	0.984	0.986
12	.805	.828	.859	.883	.943	.973	.979	.984	.986
13	.814	.837	.866	.889	.945	.974	.979	.984	.986
14	.825	.846	.874	.895	.947	.975	.980	.984	.986
15	.835	.855	.881	.901	.950	.975	.980	.984	.987
16	0.844	0.863	0.887	0.906	0.952	0.976	0.981	0.985	0.987
17	.851	.869	.892	.910	.954	.977	.981	.985	.987
18	.858	.874	.897	.914	.956	.978	.982	.986	.988
19	.863	.879	.901	.917	.957	.978	.982	.986	.988
20	.868	.884	.905	.920	.959	.979	.983	.986	.988
21	0.873	0.888	0.908	0.923	0.960	0.980	0.983	0.987	0.989
22	.878	.892	.911	.926	.961	.980	.984	.987	.989
23	.881	.895	.914	.928	.962	.981	.984	.987	.989
24	.884	.898	.916	.930	.963	.981	.984	.987	.989
25	.888	.901	.918	.931	.964	.981	.985	.988	.989
26	0.891	0.904	0.920	0.933	0.965	0.982	0.985	0.988	0.989
27	.894	.906	.923	.935	.965	.982	.985	.988	.990
28	.896	.908	.924	.936	.966	.982	.985	.988	.990
29	.898	.910	.926	.937	.966	.982	.985	.988	.990
30	.900	.912	.927	.939	.967	.983	.985	.988	.990
31	0.902	0.914	0.929	0.940	0.967	0.983	0.986	0.988	0.990
32	.904	.915	.930	.941	.968	.983	.986	.988	.990
33	.906	.917	.931	.942	.968	.983	.986	.989	.990
34	.908	.919	.933	.943	.969	.983	.986	.989	.990
35	.910	.920	.934	.944	.969	.984	.986	.989	.990
36	0.912	0.922	0.935	0.945	0.970	0.984	0.986	0.989	0.990
37	.914	.924	.936	.946	.970	.984	.987	.989	.990
38	.916	.925	.938	.947	.971	.984	.987	.989	.990
39	.917	.927	.939	.948	.971	.984	.987	.989	.991
40	.919	.928	.940	.949	.972	.985	.987	.989	.991
41	0.920	0.929	0.941	0.950	0.972	0.985	0.987	0.989	0.991
42	.922	.930	.942	.951	.972	.985	.987	.989	.991
43	.923	.932	.943	.951	.973	.985	.987	.990	.991
44	.924	.933	.944	.952	.973	.985	.987	.990	.991
45	.926	.934	.945	.953	.973	.985	.988	.990	.991
46	0.927	0.935	0.945	0.953	0.974	0.985	0.988	0.990	0.991
47	.928	.936	.946	.954	.974	.985	.988	.990	.991
48	.929	.937	.947	.954	.974	.985	.988	.990	.991
49	.929	.937	.947	.955	.974	.985	.988	.990	.991
50	.930	.938	.947	.955	.974	.985	.988	.990	.991

* Based on fitted Johnson (1949) S_B approximation, see Shapiro & Wilk (1965a) for details.

Test di normalità

Q-Q plot

- ▶ Un altro modo per verificare la normalità di una distribuzione, anche se più descrittivo, è il Q-Q plot (*plot quantile-quantile*).

Test di normalità

Q-Q plot

- ▶ Un altro modo per verificare la normalità di una distribuzione, anche se più descrittivo, è il Q-Q plot (*plot quantile-quantile*).
- ▶ In un Q-Q plot, i quantili osservati (reali) vengono confrontati con i quantili attesi nel caso la distribuzione fosse normale.

Test di normalità

Q-Q plot

- ▶ Un altro modo per verificare la normalità di una distribuzione, anche se più descrittivo, è il Q-Q plot (*plot quantile-quantile*).
- ▶ In un Q-Q plot, i quantili osservati (reali) vengono confrontati con i quantili attesi nel caso la distribuzione fosse normale.
- ▶ Se i punti si dispongono lungo una retta, la distribuzione approssima bene la normale.

Test di normalità

Q-Q plot

- ▶ Un altro modo per verificare la normalità di una distribuzione, anche se più descrittivo, è il Q-Q plot (*plot quantile-quantile*).
- ▶ In un Q-Q plot, i quantili osservati (reali) vengono confrontati con i quantili attesi nel caso la distribuzione fosse normale.
- ▶ Se i punti si dispongono lungo una retta, la distribuzione approssima bene la normale.

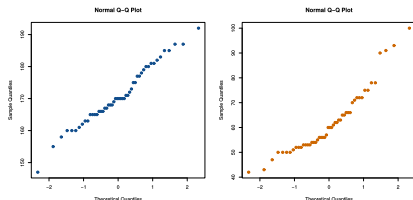


Figura: Q-Q plot di una distribuzione che approssima bene quella normale (a sinistra, $p = 0,6284$) e di una che se ne allontana significativamente (a destra, $p = 0,001179$)

Asimmetria e curtosi

Coefficiente di asimmetria

$$g_1 = \frac{\sum(x_i - \bar{X})^3}{\sigma^3} \quad (2)$$

Coefficiente di curtosi

$$g_2 = \frac{\sum(x_i - \bar{X})^4}{\sigma^4} \quad (3)$$