

正态性检验

正态性检验用于检查某观测值是否符合正态分布。检验正态分布的方法是将样本数据的直方图与标准正态曲线进行比较，或者将样本数据的标准化后的分位数与正态分布的标准分位数比较，简称 Q-Q 图。

在 Q-Q 图中，样本数据和正态数据的相关性可以反映数据是否符合正态分布，对于正态数据，在 Q-Q 图中的散点必须近似于一条直线，这样表明高度正相关。此时也很容易观察到异常值。

该模块用于单变量正态性检验，检验方法有 Anderson - Darling, Cramér - von-Mise, Lilliefors (Kolmogorov-Smirnov), Pearson's 和 Shapiro - Francia 检验。Cramér - von-Mises 检验可用于小样本 ($n \leq 25$)。对于样本量 ≥ 200 可以采用 Anderson - Darling 检验。

使用本模块可以同时输入多个变量，每个变量分别统计。

例，打开练习项目 DEMO（或下载练习数据：

www.empowerstats.com/empowerStats/exdata/demo1.xls），输入界面：

正态性检验

标题: 正态性检验

选择分析对象: 所有数据记录

选择变量

变量

Age, years

Height, m

FEV1

FVC

刷新 保存 查看结果

输出结果：

Age, years : 正态性检验

方法	统计量	p 值
Anderson-Darling normality test	11.4258	<0.000001
Cramer-von Mises normality test	2.0218	<0.000001
Lilliefors (Kolmogorov-Smirnov) normality test	0.1276	<0.000001
Pearson chi-square normality test	215.5904	<0.000001
Shapiro-Francia normality test	0.9563	<0.000001

Age, years : 频数分布

分组区间下限	分组区间上限	分组区间中值	组内频数	百分数
15	20	17.5	70	8.4337
20	25	22.5	83	10.0000
25	30	27.5	125	15.0602
30	35	32.5	168	20.2410
35	40	37.5	65	7.8313
40	45	42.5	84	10.1205
45	50	47.5	63	7.5904
50	55	52.5	56	6.7470
55	60	57.5	45	5.4217
60	65	62.5	38	4.5783
65	70	67.5	19	2.2892
70	75	72.5	11	1.3253
75	80	77.5	3	0.3614

Height, m : 正态性检验

方法	统计量	p 值
Anderson-Darling normality test	1.6088	0.000389
Cramer-von Mises normality test	0.2680	0.000785
Lilliefors (Kolmogorov-Smirnov) normality test	0.0509	0.000047
Pearson chi-square normality test	213.2232	<0.000001
Shapiro-Francia normality test	0.9947	0.008458

Height, m : 频数分布

分组区间下限	分组区间上限	分组区间中值	组内频数	百分数
1.35	1.4	1.375	5	0.6305
1.4	1.45	1.425	21	2.6482
1.45	1.5	1.475	101	12.7364
1.5	1.55	1.525	169	21.3115
1.55	1.6	1.575	184	23.2030
1.6	1.65	1.625	150	18.9155
1.65	1.7	1.675	110	13.8714
1.7	1.75	1.725	45	5.6747
1.75	1.8	1.775	7	0.8827
1.8	1.85	1.825	1	0.1261

FEV1 : 正态性检验

方法	统计量	p 值
Anderson-Darling normality test	1.4031	0.001246
Cramer-von Mises normality test	0.1772	0.010500
Lilliefors (Kolmogorov-Smirnov) normality test	0.0428	0.002779
Pearson chi-square normality test	34.9429	0.112888
Shapiro-Francia normality test	0.9919	0.000696

FEV1 : 频数分布

分组区间下限	分组区间上限	分组区间中值	组内频数	百分数
0	0.5	0.25	3	0.4076
0.5	1	0.75	20	2.7174
1	1.5	1.25	18	2.4457
1.5	2	1.75	30	4.0761
2	2.5	2.25	53	7.2011
2.5	3	2.75	115	15.6250
3	3.5	3.25	137	18.6141
3.5	4	3.75	126	17.1196
4	4.5	4.25	120	16.3043
4.5	5	4.75	61	8.2880
5	5.5	5.25	36	4.8913
5.5	6	5.75	11	1.4946
6	6.5	6.25	3	0.4076
6.5	7	6.75	1	0.1359
7	7.5	7.25	1	0.1359
7.5	8	7.75	1	0.1359

FVC : 正态性检验

方法	统计量	p 值
Anderson-Darling normality test	1.1779	0.004464
Cramer-von Mises normality test	0.1966	0.005931
Lilliefors (Kolmogorov-Smirnov) normality test	0.0370	0.018218
Pearson chi-square normality test	32.6576	0.172265
Shapiro-Francia normality test	0.9943	0.008131

FVC : 频数分布

分组区间下限	分组区间上限	分组区间中值	组内频数	百分数
1	1.5	1.25	1	0.1359
1.5	2	1.75	3	0.4076
2	2.5	2.25	19	2.5815
2.5	3	2.75	41	5.5707
3	3.5	3.25	77	10.4620
3.5	4	3.75	110	14.9457
4	4.5	4.25	130	17.6630
4.5	5	4.75	113	15.3533
5	5.5	5.25	95	12.9076
5.5	6	5.75	69	9.3750
6	6.5	6.25	38	5.1630
6.5	7	6.75	24	3.2609
7	7.5	7.25	9	1.2228
7.5	8	7.75	6	0.8152

8	8.5	8.25	0	0.0000
8.5	9	8.75	1	0.1359

基本统计

统计指标	N	均数	标准差	最小值	5%位数	10%位数	25%位数	50%位数	75%位数	90%位数	95%位数	最大值
Age, years	830	37.8849	13.9423	15.6	18.6	21.4	27.6	34	47.9	58.41	63.9	77
Height, m	793	1.5856	0.0773	1.37	1.47	1.49	1.53	1.58	1.64	1.69	1.714	1.81
FEV1	736	3.4469	1.1135	0.42	1.4375	2.05	2.78	3.455	4.2	4.81	5.2225	7.57
FVC	736	4.5171	1.1579	1.34	2.7525	3.105	3.6975	4.42	5.2225	6.07	6.5725	8.81

