Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 2, pp. 429 ~ 436

정규성 그래프의 검정력 비교

이제영[1] · 이성원[2]

요약

정규분포에 관한 검정에 있어서 P-P 플롯과 Q-Q 플롯에서 정규분포에서의 기대 직선의 잔차들을 이용한 두 가지 통계량을 제시하고, 이 통계량들의 경험적 분위수(empirical quantile)를 구하였다. 그리고 이들 통계량의 검정력을 Shapiro-Wilk의 W 통계량과의 비교를 통하여 분석하였다.

주제어: 정규성 검정, Q-Q 플롯, P-P 플롯

1. 서론

정규성 검정에 관한 연구는 통계학 연구에 있어서 이론적으로나 실질적으로 매우 중요한 역할을 하고 있으며, 최근까지 연구되는 중요한 주제이다. 이와같은 정규성 검정은 크게 통계량을 이용한 적합도 검정과 그래프를 이용한 검정으로 연구 분야를 나눌 수 있다.

통계량을 이용한 적합도 검정으로는 Shapiro와 Wilk (1965)가 W 검정을 제안하였고, 이어 Shapiro와 Francia (1972)가 W' 검정을 제안하였다. 계속해서 많은 학자들이 여러 가지 통계량들을 제시하여 연구를 하였으며, 최근에는 Lin과 Mudholker (1980)가 표본평균과 표본분산의 독립성을 기초로 Z 통계량을 제안하여 정규성 검정을 시도하였고, Looney (1995)는 다중정규성(multivariate normality)에 대한 연구를 하였다. 또한, Fang과 Case (1996)는 표본평균과 표본표준편차의 상관을 이용한 정규성 검정을 연구하였고, Kim (1997)은 이변량 정규분포에 대한 연구를 하였다.

한편, 이와 같은 통계량을 이용한 정규성 검정 방법 외에도, 그래프에 의하여 시각적으로 표본의 정규성을 검정하는 방법이 통계학뿐만 아니라 의학부문에서도 계속 연구되어 왔다. 대표적인 그래프로는 정규확률 그래프를 위시하여 P-P, Q-Q 플롯을 들 수 있다. 이들 연구는 컴퓨터의 발달로 최근 들어 더욱 활발해 져서, Jackson *et al.* (1989), Endrenyi와

[1](712-749) 경상북도 경산시 대동 214-1, 영남대학교 통계학과, 부교수
[2](712-749) 경상북도 경산시 대동 214-1, 영남대학교 통계학과

Patel (1991), Lee와 Rhee (1997), 그리고 Lee *et al.* (1998) 등에 의해 연구되었다. 하지만, 이러한 정규성 그래프에 대한 적합도 검정은 잘 이루어 지지 않고 있었다.

본 논문에서는 정규성 검정에 대한 대표적인 그래프인 P-P 플롯과 Q-Q 플롯을 이용하여 적합도 검정을 제안하고(2절), 모의실험을 통하여 이 검정 통계량의 경험적 분위수를 제시하며, 아울러 이들 통계량과 정규성 검정에 있어서 가장 널리 사용되는 Shapiro-Wilk의 W 통계량에 대한 검정력을 비교하고자 한다(3절).

2. Q-Q 플롯과 P-P 플롯의 정규성 검정통계량

Q-Q 플롯이나 P-P 플롯에서 정규성을 검정하기 위해서는 정규분포에서 기대되는 직선으로부터의 편차를 연구자 개인의 경험에 의한 주관적인 기준에 의존하여 판단할 수 밖에 없다. 하지만 이 편차의 정도를 수치적으로 표현하여 정규성 검정의 척도로 삼으면 개인의 주관적인 요소를 배제시킬 수 있다.

확률표본 $X_1,\ldots,X_n$의 순서통계량을 $X_{1:n},\ldots,X_{n:n}$이라 하고, 이들 표준화된 확률변수 X_i가 정규분포를 따를 때, 이 확률변수의 누적분포함수(CDF)를 $\Phi(x)$라고 하자. 그러면 Q-Q 플롯은 좌표평면 위에 $\left(\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right), x_{i:n}\right)$ 좌표로 구성된다. 이때, y 좌표의 기대값은

$$E[X_{i:n}] \cong \Phi^{-1}\left(\frac{i-c}{n-2c+1}\right)$$

이고, 이때 c 값에 대해서는 여러 가지 연구를 통하여 0, $\frac{3}{8}$, $\frac{1}{2}$ 일 때가 바람직하다고 알려져 있다(Montgomery, 1996). 그러므로 표본이 정규분포를 따를 때, 그 표본의 관측값을 표준화(standardized)시켜 생각하면, 이 Q-Q 플롯에서의 기대되는 직선의 형태는 $y = x$이고, 우리는 이 직선으로부터의 흩어진 정도를 정규성 검정의 척도로 사용한다. 이 주관적인 흩어진 정도를 계산하기 위하여, 통계학에서 흔히 사용되는 변동에 대한 유용한 통계량인 표본의 분산을 이용하였다. 즉, 기대되는 직선으로부터 잔차를 구하여 그 잔차의 표본분산(sample variance), S^2_{QQ}를 아래와 같이 구할 수 있다.

$$\begin{aligned}
S^2_{QQ} &= \frac{1}{n-1}\sum_{i=1}^{n}\left\{\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n} - \overline{(\Phi^{-1}(\cdot)-x)}\right\}^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left\{\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n} - \frac{1}{n}\sum_{i=1}^{n}\left(\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n}\right)\right\}^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left\{\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n} - \frac{1}{n}\sum_{i=1}^{n}\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) + \frac{1}{n}\sum_{i=1}^{n}x_{i:n}\right\}^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left\{\Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n}\right\}^2 \\
&= \frac{1}{n-1}L_n
\end{aligned}$$

이때 $L_n = \sum_{i=1}^{n} \left\{ \Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n} \right\}^2$ 이다. 이 L_n 통계량은 DeWet과 Venter (1972)에 의해 성질이 밝혀진 표준화된 관측값에 대한 정규성의 검정 통계량이며, 이때 이들은 $c = 0$를 사용하였다. 이 S^2_{QQ} 통계량은 Q-Q 플롯에서 좌표들의 정규분포에서의 기대 직선 $y = x$에 대한 잔차들의 표본분산이라고 할 수 있다. 본 논문에서는 c의 값을 $\frac{3}{8}$으로 하여 모의실험을 실시하였다.

같은 방법으로 P-P 플롯에서도 기대되는 직선의 잔차들의 표본분산, S^2_{PP}을 다음과 같이 구할 수 있다.

$$S^2_{PP} = \frac{1}{n-1}\sum_{i=1}^{n}\left\{\left(\frac{i}{n} - \frac{n+1}{n}\Phi(x_{i:n})\right) - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{i}{n} - \frac{n+1}{n}\Phi(x_{i:n})\right)\right\}^2$$

이렇게 구한 S^2_{PP} 통계량도 역시 P-P 플롯에서 좌표들이 기대되는 직선으로부터 많이 흩어져 있을수록 값이 커지는 통계량이 된다.

그리고 널리 알려져 있는 Shapiro-Wilk의 W 통계량은 다음과 같이 표현된다(Shapiro와 Wilk, 1965). 만약 X_i가 표준정규분포를 따른다면, Shapiro-Wilk의 W 통계량은

$$W = \frac{(\sum_{i=1}^{n} a_i x_{i:n})^2}{\sum_{i=1}^{n} {x_{i:n}}^2}$$

이고, 여기서 $\mathbf{a}' = (a_1, ..., a_n) = \frac{\mathbf{m}'V^{-1}}{(\mathbf{m}'V^{-1}V^{-1}\mathbf{m}')^{\frac{1}{2}}}$, $\mathbf{m}' = (E(X_{1:n}), ..., E(X_{n:n}))$, $V = \{v_{ij}\} = COV(X_{i:n}, X_{j:n})$이다.

여기에서 정의된 S^2_{QQ}과 S^2_{PP} 통계량들의 분위수를 모의실험을 통하여 구해보고, 또한 이 분위수를 이용한 검정의 검정력을 Shapiro-Wilk의 W 통계량과의 비교를 통하여 이들 통계량의 효율성을 알아보고자 한다.

3. 검정통계량의 분위수와 검정력 비교

앞 절에서 정의한 S^2_{QQ}과 S^2_{PP} 통계량들의 경험적(empirical) 분위수를 구하기 위하여 다음과 같은 절차를 반복 수행하였다. 먼저 주어진 표본 크기에 대하여 SAS 프로그램을 이용하여 정규분포로부터 표본을 2000개씩 생성하고 각각의 표본에 대한 검정통계량들을 구하였다. 이렇게 구한 통계량들을 순서대로 나열하여 상위 $1-\alpha$ 분위수를 구하였다. 이 과정을 표본 크기를 변화시켜 가면서 계속 수행하여 표 3.1과 표 3.2를 구하였다.

다음으로 이들 분위수를 이용하여 두 그래프의 검정력을 Shapiro-Wilk의 W 통계량과의 비교를 통하여 알아보았다. 표 3.3은 유의수준 5%에서 표본 크기 10, 20, 50에 대한 모의실험 결과를 나타내고 있다. 각 표본 크기에 대하여 균일(uniform) 분포, 코쉬(Cauchy) 분포, 그리고 지수(exponential) 분포에서 각각 2000개씩의 표본을 생성하여, Shapiro-Wilk의 W와 Q-Q 플롯과 P-P 플롯의 검정통계량들을 구하여 검정력을 구하였다.

Shapiro-Wilk의 W 통계량과 비교하면서 결과를 살펴보면, 먼저 Q-Q 플롯은 균일 분포에 대하여서는 검정력이 다소 떨어지고 있으나 코쉬 분포에 대하여는 조금 앞서고 있고 지

표 1: Q-Q 플롯의 정규성 검정 통계량의 경험적(empirical) 분위수

SIZE	MEAN	P50	P75	P90	P95	P99
10	0.071254	0.059169	0.090602	0.12908	0.15909	0.22433
11	0.066747	0.056733	0.083838	0.11773	0.14582	0.21684
12	0.062125	0.052164	0.076093	0.10928	0.13777	0.21171
13	0.058675	0.050825	0.073531	0.10133	0.13086	0.18363
14	0.057949	0.049058	0.071316	0.10521	0.12966	0.18044
15	0.054500	0.045511	0.067304	0.09653	0.11977	0.18632
16	0.051513	0.043126	0.064592	0.09220	0.11564	0.15770
17	0.051329	0.043309	0.063121	0.09243	0.11023	0.16999
18	0.048340	0.041060	0.061046	0.08485	0.09996	0.14233
19	0.046973	0.039442	0.057378	0.08286	0.10461	0.15643
20	0.044101	0.036927	0.055167	0.07693	0.09310	0.14866
25	0.038743	0.032989	0.047859	0.06757	0.08572	0.12280
30	0.033735	0.028775	0.042110	0.05792	0.07053	0.11242
35	0.029614	0.025563	0.036306	0.05074	0.06258	0.08679
40	0.026520	0.022888	0.032251	0.04519	0.05409	0.08335
45	0.024639	0.021386	0.030422	0.04189	0.05094	0.06996
50	0.022883	0.019590	0.027960	0.03822	0.04807	0.07248
55	0.020783	0.017846	0.025467	0.03515	0.04309	0.06558
60	0.019079	0.016762	0.023366	0.03223	0.04001	0.05422
65	0.018021	0.015828	0.022108	0.03071	0.03668	0.05393
70	0.017551	0.015140	0.021954	0.02944	0.03523	0.05075
75	0.016119	0.014228	0.019671	0.02671	0.03252	0.04601
80	0.015169	0.013237	0.018527	0.02553	0.03084	0.04257
85	0.014311	0.012450	0.017401	0.02408	0.02912	0.04083
90	0.013638	0.012066	0.016836	0.02235	0.02729	0.03596
95	0.013125	0.011807	0.016123	0.02158	0.02564	0.03416
100	0.012474	0.011016	0.015277	0.02074	0.02436	0.03627
200	0.007040	0.006154	0.008679	0.01139	0.01415	0.02021
300	0.004980	0.004418	0.006084	0.00809	0.00987	0.01348

표 2: P-P 플롯의 정규성 검정 통계량의 경험적(empirical) 분위수

SIZE	MEAN	P50	P75	P90	P95	P99
10	0.0065095	0.0055636	0.0081452	0.011198	0.013463	0.017824
11	0.0059077	0.0051677	0.0073327	0.010076	0.011793	0.016961
12	0.0052100	0.0045089	0.0063962	0.009105	0.011022	0.015050
13	0.0047436	0.0041017	0.0058663	0.008237	0.010101	0.013428
14	0.0045099	0.0039464	0.0057611	0.007748	0.009156	0.012397
15	0.0041889	0.0036423	0.0052701	0.007321	0.008734	0.011917
16	0.0037590	0.0032868	0.0046619	0.006323	0.007695	0.010285
17	0.0036756	0.0031813	0.0045584	0.006292	0.007616	0.010507
18	0.0033340	0.0028741	0.0042072	0.005785	0.006854	0.009797
19	0.0032050	0.0028010	0.0040801	0.005499	0.006492	0.009323
20	0.0029694	0.0025929	0.0037054	0.005146	0.006009	0.008738
25	0.0024077	0.0020541	0.0029595	0.004197	0.005183	0.006920
30	0.0019304	0.0016598	0.0024032	0.003301	0.004014	0.005431
35	0.0016398	0.0014162	0.0020385	0.002846	0.003403	0.004624
40	0.0014541	0.0012889	0.0018059	0.002447	0.002875	0.004174
45	0.0013057	0.0011068	0.0016088	0.002290	0.002771	0.003831
50	0.0011607	0.0010093	0.0014435	0.001989	0.002437	0.003447
55	0.0010452	0.0008977	0.0013039	0.001802	0.002238	0.003008
60	0.0009458	0.0008144	0.0011817	0.001611	0.001920	0.002760
65	0.0008832	0.0007608	0.0010969	0.001505	0.001851	0.002588
70	0.0008324	0.0007248	0.0010504	0.001423	0.001674	0.002479
75	0.0007606	0.0006547	0.0009526	0.001315	0.001572	0.002251
80	0.0007090	0.0006080	0.0008690	0.001241	0.001506	0.002124
85	0.0006723	0.0005831	0.0008375	0.001137	0.001382	0.002003
90	0.0006249	0.0005441	0.0007849	0.001070	0.001277	0.001793
95	0.0005983	0.0005213	0.0007361	0.000999	0.001220	0.001749
100	0.0005564	0.0004798	0.0006963	0.000965	0.001137	0.001648
200	0.0002758	0.0002385	0.0003427	0.000482	0.000565	0.000763
300	0.0001896	0.0001635	0.0002309	0.000329	0.000402	0.000580

표 3: 5% 유의수준에서 모의실험을 통한 세 통계량의 검정력 비교

		W	Q-Q plot	P-P plot
n=10	uniform	7.60	4.60	12.15
	Cauchy	59.55	62.25	54.65
	exponential	43.55	40.25	32.95
n=20	uniform	20.40	10.00	25.70
	Cauchy	86.00	89.60	85.60
	exponential	83.05	79.95	68.70
n=50	uniform	88.75	43.00	56.65
	Cauchy	99.40	99.65	99.55
	exponential	99.99	99.75	98.20

수분포에 대하여 조금 떨어지고 있다. 다음으로 P-P 플롯은 Q-Q 플롯에서와는 달리 균일분포에서는 소표본(N=10, 20)에서 검정력이 다른 검정보다 높게 나타났지만 지수 분포에서는 다름 검정에 비하여 낮게 나타났고 코쉬 분포에서는 조금 낮게 나타났다. 전반적으로, 소표본($N \leq 20$)의 경우에서는 두 그래프가 Shapiro-Wilk의 W와 비교해서 검정력 면에서 별로 뒤지지 않음을 알 수 있었다.

결론적으로 대표본(N=50)에서는 Shapiro-Wilk의 W 검정이 다른 검정보다 우수하였으나, 소표본에서는 지수분포에 대하여서만 다른 검정보다 우수하고, 코쉬 분포에 대하여서는 Q-Q 플롯이, 균일 분포에 대하여서는 P-P 플롯이 더 우수한 것으로 나타났다. 특히 그래프를 이용한 두 검정은 통계량에 포함되지 않았던 패턴(pattern)까지도 분석할 수 있어서 효율적이라고 할 수 있을 것이다.

4. 결론

Q-Q 플롯과 P-P 플롯을 이용한 정규성 검정에서 S^2_{QQ}과 S^2_{PP} 통계량들을 사용한다면, 더욱 명확하게 결론을 지을 수 있고 또한 플롯을 통하여 패턴까지도 분석할 수 있어서 효율적이라고 할 수 있을 것이다. 특히 소표본에서 그 검정력이 뛰어난 바가 있어서 유용하다고 할 수 있다. 추가적으로 연구되어야 할 사항은 단순히 잔차들의 크기만으로 정규성 검정을 실시하는 것보다 그 잔차들의 패턴을 어떻게 검정통계량에 포함시킬 것인가에 대한 연구가 필요하다고 생각된다.

참고문헌

1. De Wet, T. and Venter, J. H. (1972). Asymptotic distributions of certain test criteria of normality, *South African Statistical Journal*, Vol. 6. 135-149.

2. Endrenyi, L. and Patel, M. (1991). A new, sensitive graphical method for detecting deviations from the normal distribution of drug responses: the NTV plot, *British Journal of clinical pharmacology*, Vol. 32. 159-166.

3. Fang, J. and Case, K. (1996). Using a correlation test for normality, *Journal of Quality Technology*, Vol. 28, 356-362.

4. Jackson, P. R., Tucker, G. T. and Woods, H. F. (1989). Testing for bimodality in frequency distributions of data suggesting polymorphisms of drug metabolism- histograms and probit plots, *British Journal of clinical pharmacology*, Vol. 28, 647-653.

5. Kim, N. (1997). 이변량 정규분포의 적합도 검정을 위한 통계량의 극한분포에 대한 연구, ⟨한국통계학회논문집⟩, 4권 3호. 863-879.

6. Lee, J.-Y. and Rhee, S.-W. (1997). 특정분포에 따른 확률 Plot들의 정규성과 Bimodality 비교, ⟨한국통계학회논문집⟩, 4권 1호, 243-254.

7. Lee, J.-Y., Woo, J. S., and Choi, D. W. (1998). Using a normal test variable (NTV) for clinical research, ⟨응용통계연구⟩, 11권 1호, 129-139.

8. Lin, C.-C. and Mudholkar, G. (1980). A simple test for normality to against asymmetric alternatives, *Biometrika*, Vol. 67, 455-461.

9. Looney, S. W. (1995). How to use tests for univariate normality to assess multivariate normality, *The American Statistician*, Vol. 49, 64-70.

10. Montgomery, D. C. (1996). *Introdution to Statistical Quality Control*, Third Edition. John Wiley & Sons, Inc., 113-116.

11. Shapiro, S. S. and Francia, R. S. (1972). An approximation analysis of variance test for normality, *Journal of American Statistical Association*, Vol. 67, 215-216.

12. Shapiro, S. S. and Wilk, M. B. (1965). An analysis-of-variance test for normality (complete sample), *Biometrika*, Vol. 52, 591-611.

Power Analysis for Normality Plots

Lee Jea-Young [3] · Rhee Seong-Won [4]

Abstract

We suggest test statistics for normality using Q-Q plot and P-P plot and obtain empirical quantitles of these statistics. Also the power compari son with Shapiro-Wilk's W is conducted by Monte Carlo study.

Key Words and Phrases: Normality test, Q-Q plot, P-P plot.

[3]Associate Professor, Department of Statistics, Yeungnam University, Kyongsan, Kyongbuk, 712-749, Korea
[4]Graduate, Department of Statistics, Yeungnam University, Kyongsan, Kyongbuk, 712-749, Korea