

Normal Probability Plots for Normality

Jea-Young Lee¹⁾ Seong-Won Rhee²⁾

Abstract

The goodness of fit statistics of normality plots are obtained using the Receiver Operating Characteristic (ROC) method. This work is intended to compare with Shapiro-Wilk W statistic. We will use and discuss an accuracy of the test and the best cut-off value which minimizes the sum of the type I and II error probabilities.

1. 서 론

진단법(또는 결정 방법)의 정확도 및 두 진단법의 우수성에 대한 검정 방법으로 널리 사용되는 통계적 기법인 ROC 분석이 Metz (1978)에 의해 임상의학 부문에 소개되었으며, Hanley와 McNeil (1982, 1983), Altman (1992), Schoonjans 등 (1995) 그리고 Lee와 Rhee (1998) 등에 의하여 계속 연구되었다.

한편, 통계 분석의 현실적인 측면에서 정규성 가정은 매우 중요한 역할을 한다. 이런 이유로 정규성 가정의 적합도 검정에 관한 연구는 통계학 연구에 있어서 기본적인 부분이면서 또한 최근까지 연구가 계속되는 중요한 주제이다. 대표적인 연구로, 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 Kim (1997)이 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 플롯과 Q-Q 플롯 등이 있다. 이들 연구는 컴퓨터의 비약적인 발달로 최근 들어 더욱 활발해 저서, Jackson 등 (1989), Endrenyi와 Patel (1991), Lee와 Rhee (1997) 그리고 Lee, Woo 그리고 Choi (1998)에 의해 연구되었다.

본 논문에서는 먼저 대표적인 정규 그래프인 P-P 플롯과 Q-Q 플롯을 이용하여 정규성 가정에 대한 적합도 검정의 방법에 대하여 논의하고 ROC 분석 기법에 대하여 소개한다(2절). 다음으로 논의된 이들 두 적합도 검정과 정규성 검정에 있어서 가장 널리 사용되는 Shapiro-Wilk의 적합도 검정과 ROC 분석을 통하여 그들간의 효율성과 정확도를 검정하고자 한다(3절).

1) (712-749) 경북 경산시 대동 214-1, 영남대학교, 통계학과, 부교수

2) (706-010) 대구광역시 수성구 범어동 223-5 동일빌딩 609호, (주)리서치넷, 시스템개발실장

2. 정규 그래프의 정규성 적합도 검정과 ROC 분석

통계량을 이용하는 검정과 달리 그래프를 이용하는 검정에서는 명확한 검정 규칙(기각역)이 잘 알려져 있지 않다. 그래서 그래프 사이의 효율성의 비교는 물론이고 그래프와 다른 검정 통계량 사이의 효율성의 비교에 대한 논의가 제기되고 있는 실정이다. 이에 대한 시도로 Lee와 Rhee (1998)는 정규성 그래프의 변동 통계량에 관한 연구를 하였다. 본 논문에서는 그래프에 대한 일반적인 검정 규칙을 적용하려고 한다.

잘 알려진 바와 같이 정규 그래프에서의 정규성 검정의 척도는 정규분포에서 기대되는 직선에 주어진 자료의 적합도로 생각할 수 있다. 이 적합도는 관측자가 시각적으로 파악해야하므로 관측자 개인의 주관적인 판단에 크게 의존할 수밖에 없다. 하지만 이 적합도의 정도는 기대 직선과 자료 사이의 편차들의 크기에 크게 의존하고 있다. 그래서 이 적합도에 대한 기준으로 이 편차들의 표본 분산(sample variance)을 제시하고 그에 대한 효율성을 몇 가지 분포에서의 모의실험을 통하여 알아보기로 한다.

2.1 Q-Q 플롯에서의 정규성 적합도 검정

평균이 0(zero)이고 표준편차가 1인 정규분포를 따르는 확률표본을 X_1, \dots, X_n 라고 하고 그것들의 순서통계량을 $X_{1:n}, \dots, X_{n:n}$ 이라 하고, 표준정규분포의 누적분포함수(CDF)의 역함수를 $\Phi^{-1}(x)$ 라고 하자. 그러면 Q-Q 플롯은 좌표평면 위에 $(\Phi^{-1}(\frac{i-c}{n-2c+1}), x_{i:n})$ 좌표로 구성된다. 이때, y 좌표의 기대값을 구하면 다음과 같다.

$$E[X_{i:n}] \cong \Phi^{-1}\left(\frac{i-c}{n-2c+1}\right), \quad 0 \leq c \leq 1.$$

c 값에 대해서는 여러 가지 연구를 통하여 0, 3/8, 1/2일 때가 널리 알려져 있다(Montgomery, 1996 and SAS Procedure Guide, 1988). 그러므로 표본이 정규분포를 따를 때, 그 표본의 관측값을 표준화(standardized)시켜 생각하면, 이 Q-Q 플롯에서의 기대되는 직선의 형태는 $y=x$ 이다. 이 경우, 정규성 검정의 적합도는 자료의 값과 그 값에 대한 기대값의 편차들의 표본 분산으로부터 검정될 수 있다. 그래프에서 기대되는 직선으로부터 편차를 구하여 그 편차의 표본 분산, S_{QQ}^2 를 아래와 같이 구할 수 있다(Lee와 Rhee, 1998).

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

이때 $L_n = \sum_{i=1}^n \left\{ \Phi^{-1}\left(\frac{i-c}{n-2c+1}\right) - x_{i:n} \right\}^2$ 은 DeWet과 Venter (1972)에 의해 성질이 밝혀진 표준화된 관측값에 대한 정규성의 검정 통계량이다. 이 S_{QQ}^2 통계량은 Q-Q 플롯에서 좌표들의 $y=x$ 직선에 대한 잔차의 표본분산이라고 할 수 있다. 이때 S_{QQ}^2 통계량의 값이 크면 우리는 정규성 가설에 대한 적합도 검정에서 가설을 기각할 것이다.

2.2 P-P 플롯에서의 정규성 적합도 검정

앞의 2.1에서와 같이 평균이 0(zero)이고 표준편차가 1인 정규분포를 따르는 확률표본을 X_1, \dots, X_n 라고 하고 그것들의 순서통계량을 $X_{1:n}, \dots, X_{n:n}$ 이라 하고, 표준정규분포의 누적분포 함수(CDF)를 $\Phi(x)$ 라고 하자. 그리고 표본분포함수를 $\Phi_n(x)$ 라고 두자. 그러면 P-P 플롯은 좌표 평면 위에 $(\Phi(X_{i:n}), \Phi_n(X_{i:n}))$ 좌표로 구성된다. 이때 $\Phi(X_{i:n})$ 가 $BETA(i, n-i+1)$ 분포를 따르므로, x 좌표의 기대값을 구하면 다음과 같다.

$$E[\Phi(X_{i:n})] = \frac{i}{n+1}.$$

따라서 표준화된(standardized) 표본이 정규분포를 따를 때, P-P 플롯에서의 기대되는 직선의 형태는 $y = \frac{n+1}{n}x$ 이고, 정규성 검정의 적합도는 자료의 값과 그 값에 대한 기대값의 편차들의 표본 분산으로 둔다. 즉, 그래프에서 이 기대 직선으로부터 편차를 구하여 그 편차의 표본 분산, S_{PP}^2 를 아래와 같이 구할 수 있다.

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

이렇게 구한 S_{PP}^2 통계량도 역시 P-P 플롯에서 좌표들이 기대 직선으로부터 많이 흩어져 있을수록 값이 커지는 표본분산이 된다. 이때, Q-Q 플롯에서와 유사하게 S_{PP}^2 통계량의 값이 크면 우리는 정규성 가설에 대한 적합도 검정에서 가설을 기각할 것이다.

여기에서 정의된 S_{QQ}^2 와 S_{PP}^2 통계량에 대한 효율성을 다음에 소개한 ROC 분석을 통하여 대표적인 정규성에 대한 적합도 검정통계량인 Shapiro-Wilk의 W 통계량과의 비교를 통하여 검정하고자 한다.

2.3 ROC 분석

의학 부문에서의 질병에 대한 진단(diagnostic decision)이나 통계학 부문에서의 가설에 대한 검정은 서로 의미가 일치하는 바가 있다. 여기에서는 Lee와 Rhee (1998)에 의해 국내에 소개된 ROC 분석에 대하여 몇 가지 기본적인 개념을 다시 알아보기로 한다. 옳은 진단이나 검정을 내릴 비율을 정확도(accuracy)라고 정의하고, 특히 귀무가설이 참일 때 귀무가설을 채택할 비율을 특이도(specificity, Spec.)라고 하며 대립가설이 참일 때 대립가설을 채택할 비율을 민감도(sensitivity, Sens.)라고 한다. 이것은 다음과 같이 나타낼 수 있다.

$$\text{Specificity} = \frac{H_0 \text{를 옳게 채택할 경우의 수}}{\text{실제 } H_0 \text{가 참인 경우의 수}} \times 100,$$

$$\text{Sensitivity} = \frac{H_1 \text{를 옳게 채택할 경우의 수}}{\text{실제 } H_1 \text{가 참인 경우의 수}} \times 100,$$

표 1. Shapiro-Wilk W의 ROC 분석 결과

Sample Size	분포	Shapiro-Wilk W				
		ROC Area	Best Cut-off			
			Criteria	Sens.	Spec.	Accuracy
10	균일	0.618	0.9269	57.6	60.8	59.20
	지수	0.830	0.8902	66.5	84.5	75.50
	코쉬	0.840	0.8726	66.7	90.5	78.60
20	균일	0.780	0.9449	73.2	70.6	71.90
	지수	0.967	0.9194	90.9	89.9	90.40
	코쉬	0.954	0.9010	85.4	96.0	90.70
50	균일	0.977	0.9543	94.0	91.0	92.50
	지수	1.000	0.9291	99.5	99.5	99.50
	코쉬	0.998	0.9308	98.5	99.4	98.95

$$\text{Accuracy} = \frac{\text{옳은 진단을 한 경우의 수}}{\text{전체 경우의 수}} \times 100$$

$$= \text{Sensitivity} \times (\text{전체에서 } H_1 \text{이 참인 비율})$$

$$+ \text{Specificity} \times (\text{전체에서 } H_0 \text{이 참인 비율}).$$

이 Spec.과 Sens.은 통계적으로 $(1 - \text{제1종 오류}) \times 100\%$ 와 $(1 - \text{제2종 오류}) \times 100\%$ 의 의미와 유사하다. 특히, Sens.는 바로 검정력(power)을 의미하고 있다. 우리는 여기에서 진단의 정확도를 측정하고 가장 바람직한 진단의 기준을 파악하기 위하여 ROC 분석을 하게된다. 결국 진단의 기준(threshold)이나 검정 통계량의 기각역을 변화시켜가면서 좌표평면 위에 (1-Spec., Sens.) 좌표를 표시하는 그래프를 ROC 곡선이라고 부른다. 이 곡선과 x 축 사이의 면적은 정확도에 대한 척도로 사용된다. 즉, θ 를 아래와 같은 확률로 정의한다면, 이 θ 는 이 주어진 통계량이 $H_0: normal$ 에서와 $H_1: abnormal$ 에서 어느 정도 옳게 구별될 수 있는가에 대한 척도이다.

$$\theta \text{ (or } 1 - \theta) = \Pr[H_0 \text{에서의 통계량 값} < H_1 \text{에서의 통계량 값}],$$

이때 우리는 ROC 분석을 통하여 일정한 유의수준에서의 검정력 등을 구할 수도 있으며, 또한 진단의 정확도를 가장 높일 수 있는 최적 기각치의 값도 구할 수 있다. 이 최적 기각치 값은 정확도의 측면에서 발생하는 모든 오류(제1종 오류 + 제2종 오류)를 가장 최소로 하는 기각역으로 기존의 일정한 제1종 오류의 수준에서 제2종 오류의 수준을 줄여나가는 방법보다 많은 정보를 제공하여 주므로 Altman (1992)에 의하여 중요성이 강조되었다. 예를 들어 어떤 검정에 있어서 오류가 발생할 때 지불하는 비용이 제1종 오류와 제2종 오류에 관계없이 동일하다면 일정한 유의수준 하에서 정한 기각역보다 최적 기각치 하에서 정한 기각역에서 적은 비용이 발생하므로 최적 기각치를 이용하는 것이 더 바람직하다고 할 수 있다(Metz, 1978). 그리고 나아가서 두 진단법이나 검정법에 대한 비교에서도 ROC 면적(area)을 이용하여 정확도의 측면에서 우수성을 밝힐 수 있다.

표 2. Q-Q 플롯 검정통계량의 ROC 분석 결과

Sample Size	분포	Q-Q 플롯 검정통계량				
		ROC Area	Best Cut-off			
			Criteria	Sens.	Spec.	Accuracy
10	균일	0.561	0.0560	64.4	46.2	55.30
	지수	0.813	0.1008	67.0	81.5	74.25
	코쉬	0.860	0.1290	69.5	91.1	80.30
20	균일	0.666	0.0370	78.7	47.6	63.15
	지수	0.954	0.0744	88.8	88.7	88.75
	코쉬	0.969	0.0904	90.4	93.7	92.05
50	균일	0.899	0.0288	88.1	78.4	83.25
	지수	0.999	0.0656	99.1	98.9	99.00
	코쉬	0.999	0.0805	98.8	99.7	99.25

3. W 통계량과 Q-Q, P-P 플롯의 ROC 분석

이미 정규성 검정의 척도로 잘 알려진 Shapiro-Wilk의 W 통계량과 앞에서 제시한 두 정규 그래프들을 몇 가지 표본의 크기($n=10, 20, 50$)에 대하여 정규분포와 균일(uniform)분포, 지수(exponential)분포, 코쉬(Cauchy)분포의 표본을 모의실험을 통하여 각각 2000표본을 생성하여 ROC 분석 기법으로 최적 기각치 측면에서 효율성을 살펴보고, 또한 ROC 면적에서 정확도에 대한 우수성을 검정하도록 한다. 이때 정규분포에 대한 비교를 이 세 가지 분포에 대하여 한정시킨 것은 균일분포와 지수분포 그리고 코쉬분포는 손쉽게 구할 수 있고, 또한 그 분포의 형태가 정규분포에 비하여 꼬리부분이 짧거나 길거나 치우쳐진 형태를 취하고 있어서 다양한 형태에 대한 비교가 가능하기 때문이다.

먼저, 각 검정법에 대한 ROC 분석을 실시하여 보기로 한다. 표 1에서는 Shapiro-Wilk의 W 통계량에 대한 ROC 분석을 통하여 요약한 결과를 나타낸 것이다. 표 1에서 표본의 크기가 10이고 비교하는 분포가 균일분포일 때 전체 통계량 자료 4000개 중에서, 임의의 정규분포에서의 W 값이 임의의 균일분포에서의 W 값보다 클 확률($\hat{\theta}$)이 세 번째 열의 ROC 면적 항목에 0.618로 나타나 있다. 또한 이때 기각역을 0.9269로 잡게 되면 정규분포에서의 발생한 표본을 정규분포라고, 균일분포에서 발생한 표본을 정규분포가 아니라고 판단할 정확도가 59.20으로 최대가 됨을 의미하고 있다. 이것은 제 1종 오류와 제 2종 오류의 합이 이 기각역에서 최소가 됨을 의미하는 것이다. 이런 방법으로 표 1을 분석하면, 균일분포에서 정확도가 다른 분포에 비하여 떨어짐을 알 수 있다. 표 2와 표 3은 Q-Q 플롯과 P-P 플롯에 대한 ROC 분석 결과를 나타낸 것이다. 표 1과 같은 방법으로 분석하면, 역시 균일분포에서 정확도가 다른 분포에 비하여 떨어짐을 알 수 있다.

다음으로는 각 검정법 사이의 정확도를 비교하여 보기로 한다. 먼저, 균일분포에서 정확도를 알아보면, W 가 59.20(표본크기 10), 71.90(20), 92.50(50)이고 Q-Q 플롯은 이보다 낮은 55.30(10), 63.15(20), 83.25(50)로 나타났으며, P-P 플롯에서는 61.10(10)은 W 보다 높았으나 나머지는 69.25(20), 83.70(50)으로 낮았다. 지수분포에서는, W 가 75.50(10), 90.40(20), 99.50(50)이고 Q-Q 플롯은 이보다 낮은 74.25(10), 88.75(20), 99.00(50)으로 나타났으며, P-P 플롯에서는 70.35(10),

표 3. P-P 플롯 검정통계량의 ROC 분석 결과

Sample Size	분포	P-P 플롯 검정통계량				
		ROC Area	Best Cut-off			
			Criteria	Sens.	Spec.	Accuracy
10	균일	0.640	0.0060	66.0	56.2	61.10
	지수	0.776	0.0078	69.2	71.5	70.35
	코쉬	0.819	0.0117	59.8	91.2	75.50
20	균일	0.753	0.0034	69.8	68.7	69.25
	지수	0.922	0.0047	82.1	86.2	84.15
	코쉬	0.949	0.0061	85.5	94.7	90.10
50	균일	0.911	0.0017	82.5	84.9	83.70
	지수	0.996	0.0028	96.7	97.3	97.00
	코쉬	0.998	0.0038	98.9	99.8	99.35

84.15(20), 97.00(50)으로 낮았다. 한편, 코쉬분포에서는, Q-Q 플롯이 80.30(10), 92.05(20), 99.25(50)로 가장 높았고 W 는 이보다 낮은 78.60(10), 90.70(20), 98.95(50)로 나타났으며, P-P 플롯에서는 75.50(10), 90.10(20), 99.35(50)로 낮았다.

그리고 ROC 곡선으로 각 검정법들에 대한 비교를 할 수도 있는데, 특히 표본 크기 20인 경우에 세 가지 검정에 대한 ROC 곡선을 그림 1에 나타내었다. 그림 1을 통하여 세 가지 검정법의 정확도에 대한 비교를 시각적으로 쉽게 알 수 있다. (a)는 균일분포에 대한 그래프인데, W (ROC Area=0.780), P-P 플롯(ROC Area=0.753), Q-Q 플롯(ROC Area=0.666) 순으로 나타났고, (b) 지수분포에서는 W (ROC Area=0.967), Q-Q 플롯(ROC Area=0.954), P-P 플롯(ROC Area=0.922) 순으로 나타났고, (c) 코쉬분포에서는 Q-Q 플롯(ROC Area=0.969), W (ROC Area=0.954), P-P 플롯(ROC Area=0.949) 순으로 나타났다. 이때 ROC 면적은 ROC 곡선과 가로축 사이의 면적이 된다. Hanley와 McNeil (1983)은 두 곡선의 면적 사이의 차를 검정하는 기법을 제시하였다. 그 기법으로 각 검정법들을 두 개씩 쌍으로 하여 면적 사이의 차가 있는지를 검정한 결과, 거의 모든 경우에서 통계적으로 유의한 차이가 있었고, 단지 표본의 크기가 50이고 지수분포일 때 W 검정과 Q-Q 플롯 사이의 검정에서 유의확률이 0.093으로 나타났으며, 역시 표본의 크기는 50이고 이번에는 코쉬분포일 때 W 검정과 P-P 플롯 그리고 P-P 플롯과 Q-Q 플롯 사이의 검정에서 유의확률이 각각 0.313과 0.126으로 나타나서 통계적으로 유의한 차이가 없음을 알 수 있었다.

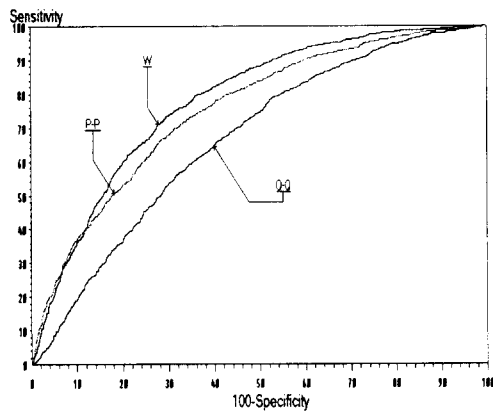
4. 결 론

그래프를 통한 정규성 검정은 정규성 검정뿐만 아니라 부수적으로 다양한 정보에 제공하여 주는 장점 때문에 널리 사용되고 있다. 그러나 정규성 검정이 어떤 검정 법칙이나 확률을 갖고 이루어지지 않아서, 정확한 결론을 내리려면 다시 Shapiro-Wilk의 W 검정과 같은 통계량을 이용한 검정을 실시해야 하는 단점이 있다. 이를 개선하기 위하여 본 논문에서는 그래프, Q-Q 플롯과 P-P 플롯에서 바로 적합도 검정을 실시하는 방법을 제안하였다. 그리고 이들 그래프에 대한 적합도 검정과 Shapiro-Wilk의 W 검정을 ROC 기법을 이용하여 비교하였다. 비교 결과, 소표본(표본

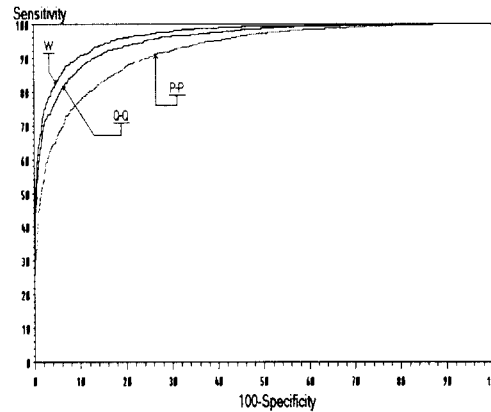
크기가 10 또는 20)에서는 각각 대립분포의 종류(균일, 지수, 코쉬분포)에 따라 서로 장단점이 있는 것으로 나타났으며 대표본(표본 크기가 50)에서는 그래프를 이용한 검정보다 Shapiro-Wilk의 W 검정이 대체적으로 다소 우수하나, 각각의 검정에서 정확도 자체가 매우 높아서(거의 100%) 큰 차이가 나타나지는 않았다.

이것을 통하여 정규성 검정을 위한 그래프를 적용할 경우에, 그래프를 이용한 적합도 검정을 실시하여도 정도 면에서 비교적 안정된 정규성 검정이 이루어진다고 추론을 할 수 있다.

a) Uniform Distribution



b) Exponential Distribution



c) Cauchy Distribution

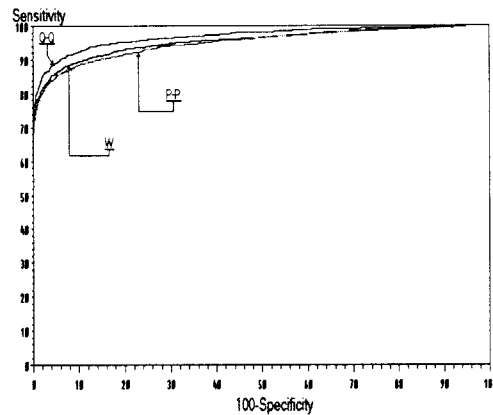


그림 1. 여러 가지 분포에 대한 ROC 곡선 비교 (표본의 크기 = 20)

참고문헌

- [1] Altman, D. G. (1992). *Practical statistics for medical research*, London : Chapman and Hall.
- [2] De Wet, T. and Venter, J. H. (1972). Asymptotic distributions of certain test criteria of normality, *South African Statistical Journal*, Vol. 6, 135-149.
- [3] Endrenyi, L. and Patel, M. (1991). A new, sensitive graphical method for detecting deviations from the normal distribution of drug responses: the NTV plot, *Br. J. clin. pharmac.*, Vol. 32, 159-166.
- [4] Hanley, J. A. and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, Vol. 143, 29-36.
- [5] Hanley, J. A. and McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases, *Radiology*, Vol. 148, 839-843.
- [6] 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, *Br. J. clin. pharmac.*, Vol. 28, 647-653.
- [7] Kim, N. (1997). 이변량 정규분포의 적합도 검정을 위한 통계량의 극한분포에 대한 연구, *한국통계학회논문집*, 4권 3호, 863-879.
- [8] Lee, J.-Y. and Rhee, S.-W. (1997). 특정분포에 따른 확률 Plot들의 정규성과 Bimodality 비교, *한국통계학회논문집*, 4권 1호, 243-254.
- [9] Lee, J.-Y. and Rhee, S.-W. (1998). Q-Q, P-P 플롯의 변동 통계량에 대한 ROC 분석, *한국통계학회논문집*, 5권 1호, 205-215.
- [10] 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.
- [11] Metz, C. E. (1978). Basic principles of ROC analysis, *Seminars in Nuclear Medicine*, Vol. 8, 283-298.
- [12] Montgomery, D. C. (1996). *Introduction to Statistical Quality Control*, Third Edition. John Wiley & Sons, Inc., 113-116.
- [13] SAS Procedures Guide (1988). Release 6.03 Edition, SAS Institute Inc., Cary, NC, USA. pp 293-299.
- [14] Schoonjans, F., Zalata, A., Depuydt, C. E., and Comhaire, F. H. (1995). MedCalc: a new computer program for medical statistics, *Computer Methods and Programs in Biomedicine*, Vol. 48, 257-262.
- [15] Shapiro, S. S. and Francia, R. S. (1972). An approximation analysis of variance test for normality, *J. Am. Stat. Assoc.*, Vol. 67, 215-216.
- [16] Shapiro, S. S. and Wilk, M. B. (1965). An analysis-of-variance test for normality (complete sample), *Biometrika*, Vol. 52, 591-611.