

# 스크리닝 검사로부터 불량률 추정에 관한 연구<sup>1)</sup>

## - 알려지지 않은 민감도와 특이도의 경우 -

김 태 규  
한남대학교 정보통계학과

### Estimation of Defect Rate from the Screening Test

#### - the case of unknown sensitivity and specificity -

Tai-Kyoo Kim  
Department of Information Statistics, Han Nam University

Key Words : Screening Test, Sensitivity, Specificity

#### Abstract

Consider the problem of estimating the confidence interval of defect rate of a specified population through a screening test. Relationship between defect rate of the population and sensitivity and specificity of the screening test is discussed. More general statistical problem is introduced with the assumption of unknown imperfect specificity. An improved version of Rogan & Gladden(1978)'s estimator is proposed and discussed. Simulation studies are performed to compare the efficiency of the proposed estimator with current estimators.

## 1. 서 론

일반적으로 제품이나 시스템은 모집단으로부터 일부를 시료로 추출하여 불량여부를 검사한 후 불량률을 계산하게 된다. 하지만 제품의 불량을 검사하기가 경비나 시간이 많이 소요될 때는 간편하게 검사 절차를 간소화한 스크리닝 검사(screening test)를 통해서 불량 여부를 판단하는 경우도 있다.

스크리닝 검사는 주로 의학분야 등에서 질병의 조기진단이나 환자의 정밀검사 여부를 판단하는 수단으로 매우 광범위하게 사용되

고 있다. 그러나, 스크리닝 검사가 비용이나 시간 측면에서 매우 간편하게 사용될 수 있는 장점이 있지만 스크리닝 검사가 완벽한 검사가 아니기 때문에 이를 통해서 불량률을 추정할 때는 특히 주의를 기울여야한다. 스크리닝 검사의 정확도를 판단하는 기준으로 민감도(sensitivity)나 특이도(specificity)가 사용된다. 따라서 이들 측도가 완벽하지 않은 경우에 이를 통해서 불량률을 추정할 때에는 적절한 통계적 보정이 이루어져야 한다.

최근까지 이러한 문제는 완벽한 특이도를

가정하고 불량률을 추정하는 문제가 주종을 이루어 왔다(Parker(1996), Park, Won & Kim(1999) 등). 따라서 본 연구에서는 특이도가 완벽하지 않은 경우에 불량률을 추정하는 방법에 대하여 논하고자 하며, 이들 결과와 기존의 연구결과를 모의실험을 통해 비교해보고자 한다.

## 2. 스크리닝 검사로부터 불량률 추정 방법

불량의 속성을 가진 제품이 발생하는 사건을  $D$ , 스크리닝 검사에서 불량이라는 결과가 나타나는 사건을  $T$ 라 하면, 스크리닝 검사의 민감도는  $\pi_{se} = P(T|D)$ , 특이도는  $\pi_{sp} = P(\bar{T}|\bar{D})$ 로 표현할 수 있다. 관심있는 로트의 불량률을  $\pi_d = P(D)$ 라 할 때 이들의 관계는 다음과 같이 표현할 수 있다.

[표 1] 스크리닝 검사의 결과표

Program Results	$T$	$\bar{T}$	계
Defect Statue			
$D$	$\pi_{se}\pi_d$	$(1-\pi_{se})\pi_d$	$\pi_d$
$\bar{D}$	$(1-\pi_{sp})(1-\pi_d)$	$\pi_{sp}(1-\pi_d)$	$1-\pi_d$
계	$\pi_{sd}$	$1-\pi_{sd}$	1

일반적으로 스크리닝 검사의 정확도를 표현하는 민감도와 특이도는 스크리닝 검사가 형성될 때 추정된다.  $x_2$ 와  $x_3$ 를 각각 크기  $n_2$ 와  $n_3$ 의 표본에서 계산된 불량률의 수와

우량의 수라 할 때, 민감도는  $\hat{\pi}_{se} = x_2/n_2$ 로 추정되며, 특이도는  $\hat{\pi}_{sp} = x_3/n_3$ 로 각각 추정된다.

만약  $n_1$ 개 제품의 스크리닝 검사에서  $x_1$ 개의 불량률이 얻어졌다면 불량률을

$\hat{\pi}_{sd} = x_1/n_1$ 로 계산할 수 있겠지만 이 추정량은 실제 불량률이 아니라 스크리닝 검사에서 얻은 불량률이라는 것이 올바른 표현이다. 왜냐하면 스크리닝 검사가 불완전하기 때문이다.

스크리닝 검사에서 로트의 불량률을 추정하려면 [표 1]에서 얻어낸 다음과 같은 관계를 이용해서 얻어낼 수 있다.

$$\pi_d = \frac{\pi_{sd} - (1 - \pi_{sp})}{\pi_{se} - (1 - \pi_{sp})} \quad (1)$$

즉, 불량률  $\pi_d$ 의 추정량은

$$\hat{\pi}_d = \frac{\hat{\pi}_{sd} - (1 - \hat{\pi}_{sp})}{\hat{\pi}_{se} - (1 - \hat{\pi}_{sp})} \quad (2)$$

이고 이 추정량의 점근적인 분산은 다음과 같다.

$$\begin{aligned} var(\hat{\pi}_d) = & \frac{\pi_{sd}(1-\pi_{sd})}{n_1(\pi_{se} + \pi_{sp} - 1)^2} + \frac{\pi_{se}(1-\pi_{se})(\pi_{sd} + \pi_{sp} - 1)^2}{n_2(\pi_{se} + \pi_{sp} - 1)^4} \\ & + \frac{\pi_{sp}(1-\pi_{sp})(\pi_{se} - \pi_{sd})^2}{n_3(\pi_{se} + \pi_{sp} - 1)^4} \quad (3) \end{aligned}$$

ROGAN & GLADEN(1979)은 정규근사를 활용해서  $\pi_d$ 의 신뢰구간을 추정하는 방법을 제안했다. 하지만 이 방법은 불량률  $\pi_d$ 가 낮을 때 좋지 않은 결과를 준다고 알려져 왔다.

본 연구논문은 식(2)의 불량률 추정량을 2차 Taylor 전개를 수행해서 보다 정확한 근사 방법을 제안하고자 한다. 실제 ROGAN & GLADEN의 방법이 1차 Taylor 전개에 의존하고 있기 때문에 2차 Taylor 전개에 의한 추정량은 약간 더 복잡하겠지만 정확성은 매우 향상되리라 생각된다.

$$\widehat{\pi}_d = \frac{\widehat{\pi}_{sd} - (1 - \widehat{\pi}_{sp})}{\widehat{\pi}_{se} - (1 - \widehat{\pi}_{sp})} \equiv f(\widehat{\pi}_{sd}, \widehat{\pi}_{se}, \widehat{\pi}_{sp}).$$

2차 Taylor 전개에 의하면 위 식은 다음과 같이 표현할 수 있다.

$$\begin{aligned} \widehat{\pi}_d &= \frac{\pi_{sd} - (1 - \pi_{sp})}{\pi_{se} - (1 - \pi_{sp})} + (\widehat{\pi}_{sd} - \pi_{sd}) \frac{\partial f}{\partial \pi_{sd}} + (\widehat{\pi}_{se} - \pi_{se}) \frac{\partial f}{\partial \pi_{se}} + (\widehat{\pi}_{sp} - \pi_{sp}) \frac{\partial f}{\partial \pi_{sp}} \\ &+ \frac{(\widehat{\pi}_{sd} - \pi_{sd})^2}{2!} \frac{\partial^2 f}{\partial \pi_{sd}^2} + \frac{(\widehat{\pi}_{se} - \pi_{se})^2}{2!} \frac{\partial^2 f}{\partial \pi_{se}^2} + \frac{(\widehat{\pi}_{sp} - \pi_{sp})^2}{2!} \frac{\partial^2 f}{\partial \pi_{sp}^2} \\ &+ (\widehat{\pi}_{sd} - \pi_{sd})(\widehat{\pi}_{se} - \pi_{se}) \frac{\partial^2 f}{\partial \pi_{sd} \partial \pi_{se}} + (\widehat{\pi}_{se} - \pi_{se})(\widehat{\pi}_{sp} - \pi_{sp}) \frac{\partial^2 f}{\partial \pi_{se} \partial \pi_{sp}} \\ &+ (\widehat{\pi}_{sd} - \pi_{sd})(\widehat{\pi}_{sp} - \pi_{sp}) \frac{\partial^2 f}{\partial \pi_{sd} \partial \pi_{sp}} \end{aligned}$$

이 식의 점근적인 평균과 분산은 다음과 같다.

$$\begin{aligned} E(\widehat{\pi}_d) &= \pi_d + \frac{\pi_d}{(\pi_{sp} + \pi_{sp} - 1)^2} \frac{\pi_{se}(1 - \pi_{se})}{n_2} \\ &+ \frac{(1 - \pi_d)}{(\pi_{sp} + \pi_{sp} - 1)^2} \frac{\pi_{sp}(1 - \pi_{sp})}{n_3} \end{aligned} \quad (4)$$

$$\begin{aligned} Var(\widehat{\pi}_d) &= \frac{2}{(\pi_{se} + \pi_{sp} - 1)^4} \left( \frac{\pi_{se}(1 - \pi_{se})\pi_d^2}{n_2} + \frac{\pi_{sp}(1 - \pi_{sp})(1 - \pi_d)^2}{n_3} \right) \\ &+ \frac{\pi_{sd}(1 - \pi_{sd})}{4n_1(\pi_{se} + \pi_{sp} - 1)^4} \left( \frac{\pi_{se}(1 - \pi_{se})}{n_2} + \frac{\pi_{sp}(1 - \pi_{sp})}{n_3} \right) \\ &+ \frac{\pi_{se}(1 - \pi_{se})\pi_{sp}(1 - \pi_{sp})(\pi_{se} + \pi_{sp} - 12(\pi_{sd} + \pi_{se} - 1))^2}{4n_2n_3(\pi_{se} + \pi_{sp} - 1)^5} \end{aligned} \quad (5)$$

이들 평균과 분산의 추정식은 각각의 표본에서 얻어진 비율을 대체하여 얻어질 수 있다. 이러한 형식의 추정량의 점근적인 분포는 Gastwirth(1987)의 정리 A1에 의해 정규분포를 따름을 알 수 있다. 따라서 이를 이용하면  $(1 - \alpha) \times 100\%$  신뢰구간을 얻을 수 있다.

Fieller(1954)는 여러 개의 추정량의 비에 관한 일반적인 신뢰구간을 추정하는 방법을 제안하였다. Fieller의 방법을 이용해서 표준 정규분포를 활용하면  $\pi_d$ 의 점근적인 신뢰구간을 계산하면 다음과 같다.

$$\frac{(\widehat{\pi}_d - (1 - \widehat{\pi}_{se})) - \widehat{\pi}_d(\widehat{\pi}_{se} - (1 - \widehat{\pi}_{sp}))}{\sqrt{\frac{\widehat{\pi}_{sd}(1 - \widehat{\pi}_{sd})}{n_1} + \frac{\widehat{\pi}_{se}(1 - \widehat{\pi}_{se})}{n_2} + \widehat{\pi}_d^2 \left( \frac{\widehat{\pi}_{se}(1 - \widehat{\pi}_{se})}{n_2} + \frac{\widehat{\pi}_{sp}(1 - \widehat{\pi}_{sp})}{n_3} \right)}} \quad (6)$$

식 (6)의  $(1 - \alpha) \times 100\%$  신뢰하한과 신뢰상한은 다음과 같이 계산될 수 있다.

$$\begin{aligned} \pi_d^L(L) &= \frac{1}{(\pi_{se} + \pi_{sp} - 1)^2 - z_{\alpha/2}^2 \left( \frac{\pi_{se}(1 - \pi_{se})}{n_2} + \frac{\pi_{sp}(1 - \pi_{sp})}{n_3} \right)} \\ &\times \left( (\pi_{sd} + \pi_{se} - 1)(\pi_{se} + \pi_{sp} - 1) - z_{\alpha/2}^2 \sqrt{(\pi_{se} + \pi_{sp} - 1)^2 A - z_{\alpha/2}^2 B} \right) \end{aligned} \quad (7)$$

$$\begin{aligned} \pi_d^U(U) &= \frac{1}{(\pi_{se} + \pi_{sp} - 1)^2 - z_{\alpha/2}^2 \left( \frac{\pi_{se}(1 - \pi_{se})}{n_2} + \frac{\pi_{sp}(1 - \pi_{sp})}{n_3} \right)} \\ &\times \left( (\pi_{sd} + \pi_{se} - 1)(\pi_{se} + \pi_{sp} - 1) + z_{\alpha/2}^2 \sqrt{(\pi_{se} + \pi_{sp} - 1)^2 A - z_{\alpha/2}^2 B} \right) \end{aligned} \quad (8)$$

여기서

$$A = \left( \frac{\pi_{sd}(1-\pi_{sd})}{n_1} + \frac{\pi_{se}(1-\pi_{se})}{n_2} + \frac{2\pi_{sp}(1-\pi_{sp})}{n_3} \right)$$

$$B = \left( \frac{\pi_{sd}(1-\pi_{sd})}{n_1} + \frac{\pi_{sp}(1-\pi_{sp})}{n_3} \right)$$

이다. 이 방법은 ROGAN & GLADEN(1979)의 방법을 약간 변형한 식으로 일반적인 신뢰구간을 계산할 때 참고적인 정보를 줄 수 있을 것으로 생각된다.

### 3. 모의실험

민감도와 특이도가 알려져 있지 않은 경우 위에서 언급한 세 가지 방법을 비교하기 위하여 모의실험을 수행하였다. 시료의 크기가 컸을 때는 위의 세 가지 방법들이 모두 주어진 신뢰수준을 유지할 것으로 판단되어 비교적 적은 표본을 이용해서 모의실험을 수행하였다. 즉, 스크리닝 검사 시료인

$n_1 = 100$  이고 민감도와 특이도를 계산하는

$n_2 = n_3 = 500$  을 가정하였다. 실제 불량률

$\pi_d$ 가 0.01, 0.05, 0.1인 경우에 민감도  $\pi_{se}$ 와

특이도  $\pi_{sp}$ 가 각각 0.8, 0.9, 0.95를 가정하고

95% 신뢰구간을 계산한 후 실제 주어진 불량률  $\pi_d$ 를 포함하는 여부를 계산하였다. 각 방법의 효율을 비교하기 위해 이런 실험을 각각 2,000번 반복하여 포함 확률과 평균 신뢰구간 길이를 계산한 후 이를 비교하였다.

여기에서, 실제 불량률이 0.01인 [표 2]를 참고하면 Fieller 방법은 90% 수준을 유지하고 있음을 볼 수 있고, Rogan & Gladden 방법이 95% 수준을, 2차 Taylor 전개에 의

한 방법은 95%를 상회하면서 약간하는 보수적인 결과를 보여주고 있다.

실제 불량률이 0.05인 [표 3]을 참고하면 Fieller 방법은 90% 수준을 믿고 있음을 볼 수 있고, Rogan & Gladden 방법이 민감도가 높아질 때 95% 수준을 맞추지 못하는 경향이 있고, 2차 Taylor 전개에 의한 방법은 95%를 유지하고 있음을 볼 수 있다.

실제 불량률이 0.1인 [표 4]를 참고하면 Fieller 방법은 여전히 90% 수준에서 움직이고 있고, Rogan & Gladden 방법이 민감도가 높아질 때 95% 수준을 맞추지 못하고 있으며, 2차 Taylor 전개에 의한 방법은 비교적 95% 수준을 잘 유지하고 있음을 볼 수 있다. 따라서 전체적으로 2차 Taylor 전개에 의한 방법이 95% 신뢰수준을 잘 유지하는 효과적인 방법이라는 결론을 얻을 수 있었다.

### 4. 결 론

일반적으로 고가제품에 불량률 추정에 시간이 많이 걸리는 경우, 직접 제품의 불량률을 추정하기 보다는 적당한 스크리닝 과정에 의해서 불량률을 추정하는 경우가 많아 지고 있다. 하지만 스크리닝 검사에 의해서 불량률을 추정하는 경우 스크리닝 검사의 불완벽성 때문에 불량률의 추정이 과소 혹은 과대 추정될 수 있다. 기존에 스크리닝 검사에서 불량률을 추정하는 Rogan & Gladden 방법이나 Fieller 방법은 간단하여 이를 사용하는 사람들에게 선호될 수 있지만 불량률에 편의를 제공한다. 본 연구에서 제안하는 새로운 방법은 Rogan & Gladden 방법을 확장한 방법으로 추정량의 통계적 최적성을 그대로 가

[표 2] 불량률 1%인 경우 평균신뢰구간길이와 포함확률

$\pi_d$	$\pi_{se}$	$\pi_{sp}$	$\pi_{sd}$	방법	평균신뢰 구간길이	포함확률
0.01	0.8	0.8	0.206	Fieller	0.1364	0.9085
				Rogan & Gladden	0.1553	0.9520
				Proposed	0.2012	0.9755
		0.9	0.107	Fieller	0.0961	0.9070
				Rogan & Gladden	0.1062	0.9615
				Proposed	0.1286	0.9725
		0.95	0.0575	Fieller	0.0722	0.8915
				Rogan & Gladden	0.0778	0.9545
				Proposed	0.0910	0.9745
	0.9	0.8	0.207	Fieller	0.1200	0.9080
				Rogan & Gladden	0.1346	0.9555
				Proposed	0.1644	0.9760
		0.9	0.108	Fieller	0.0850	0.9045
				Rogan & Gladden	0.0934	0.9545
				Proposed	0.1089	0.9785
		0.95	0.0585	Fieller	0.0649	0.9000
				Rogan & Gladden	0.0695	0.9580
				Proposed	0.0788	0.9745
	0.95	0.8	0.2075	Fieller	0.1113	0.8930
				Rogan & Gladden	0.1258	0.9475
				Proposed	0.1506	0.9730
		0.9	0.1085	Fieller	0.0803	0.8955
				Rogan & Gladden	0.0884	0.9590
				Proposed	0.1015	0.9740
0.95		0.059	Fieller	0.0627	0.8885	
			Rogan & Gladden	0.0669	0.9525	
			Proposed	0.0748	0.9690	

[표 3] 불량률이 5%인 경우 평균신뢰구간길이와 포함확률

$\pi_d$	$\pi_{se}$	$\pi_{sp}$	$\pi_{sd}$	방법	평균신뢰 구간길이	포함확률
0.05	0.8	0.8	0.23	Fieller	0.1770	0.8955
				Rogan & Gladden	0.1839	0.9415
				Proposed	0.2288	0.9615
		0.9	0.135	Fieller	0.1359	0.8960
				Rogan & Gladden	0.1380	0.9495
				Proposed	0.1604	0.9605
		0.95	0.0875	Fieller	0.1123	0.9000
				Rogan & Gladden	0.1127	0.9300
				Proposed	0.1257	0.9615
	0.9	0.8	0.235	Fieller	0.1572	0.9040
				Rogan & Gladden	0.1618	0.9495
				Proposed	0.1919	0.9695
		0.9	0.14	Fieller	0.1231	0.8875
				Rogan & Gladden	0.1247	0.9405
				Proposed	0.1404	0.9570
		0.95	0.0925	Fieller	0.1062	0.8885
				Rogan & Gladden	0.1063	0.9145
				Proposed	0.1157	0.9490
	0.95	0.8	0.2375	Fieller	0.1466	0.8875
				Rogan & Gladden	0.1515	0.9355
				Proposed	0.1764	0.9580
		0.9	0.1425	Fieller	0.1191	0.9030
				Rogan & Gladden	0.1201	0.9305
				Proposed	0.1334	0.9540
0.95		0.095	Fieller	0.1015	0.9010	
			Rogan & Gladden	0.1015	0.9025	
			Proposed	0.1094	0.9545	

[표 4] 불량률이 10%인 경우 평균신뢰구간길이와 포함확률

$\pi_d$	$\pi_{se}$	$\pi_{sp}$	$\pi_{sd}$	방법	평균신뢰 구간길이	포함확률
1.0	0.8	0.8	0.26	Fieller	0.2190	0.9100
				Rogan & Gladden	0.2176	0.9425
				Proposed	0.2642	0.9755
		0.9	0.17	Fieller	0.1745	0.9175
				Rogan & Gladden	0.1728	0.9145
				Proposed	0.1972	0.9675
		0.95	0.125	Fieller	0.1480	0.9045
				Rogan & Gladden	0.1467	0.9045
				Proposed	0.1614	0.9385
	0.9	0.8	0.27	Fieller	0.1948	0.9010
				Rogan & Gladden	0.1932	0.9215
				Proposed	0.2251	0.9545
		0.9	0.18	Fieller	0.1579	0.9070
				Rogan & Gladden	0.1563	0.9140
				Proposed	0.1737	0.9385
		0.95	0.135	Fieller	0.1357	0.9175
				Rogan & Gladden	0.1347	0.9160
				Proposed	0.1449	0.9470
	0.95	0.8	0.275	Fieller	0.1881	0.9065
				Rogan & Gladden	0.1864	0.9135
				Proposed	0.2132	0.9610
		0.9	0.185	Fieller	0.1535	0.9085
				Rogan & Gladden	0.1519	0.9135
				Proposed	0.1668	0.9455
0.95		0.14	Fieller	0.1300	0.9065	
			Rogan & Gladden	0.1290	0.9000	
			Proposed	0.1386	0.9440	

지고 있으면서 추정량의 편의를 줄여주는 방법으로 생각될 수 있고, 이는 모의실험에서 확인되었다. 이 방법은 스크리닝 검사에서 불량률이나 질병률과 같은 비율을 추정하는 문제에 있어서 기존의 방법에 대한 좋은 대

안으로 활용될 수 있을 것으로 생각된다.

시간적, 비용적인 제약때문에 스크리닝 검사의 활용이 계속 늘어날 것으로 판단되지만, 본 연구에서 제시된 방법을 포함하여 기존에 제시된 방법들이 모두 대표본 근사를

이론적 배경으로 하고 있다. 따라서 앞으로 소표본에서의 스크리닝검사에 의한 불량률 추정에 관한 연구는 계속 되어야 할 것으로 생각된다.

### 참고문헌

- [1] Kim, T. (2001), Confidence Interval for Defect Rate from Screening Test, 한남대학교논문집, 제31집, 165-172.
- [2] Fieller (1954). Some problem in interval estimation. Journal of Royal Statistical Society. Series B 16, 175-185.
- [3] Gastwirth, J.L. (1987). The statistical precision of medical screening process: Application to polygraph and AIDS antibodies test data. Statistical Science 2, 213-238.
- [4] Park, S., Won, T. and Kim, T. (1999). Confidence interval for defect rate from screening test. Proceeding of Autumn Conference of The Society of Korean Quality Management. 67-72.
- [5] Parker, R.A. (1996). An exact confidence interval for the ratio of two related proportions. Biometrical Journal 38, 373-384.
- [6] Rogan, W.J. and Gladen, B. (1978). Estimating prevalence form the results of a screening test. American Journal of Epidemiology 107, 71-76.