Neyman-Pearson 검정과 Fisher 검정에 의한 비모수 통계의 고찰
- Review of Nonparametric Statistics by Neyman-Pearson Test and Fisher Test -

최 성 운*
Sungwoon Choi*

Abstract

This paper reviews nonparametric statistics by Neyman-Pearson test and Fisher test. Nonparametric statistics deal with the small sample with distribution-free assumption in multi-product and small-volume production. Two tests for various nonparametric statistic methods such as sign test, Wilcoxon test, Mann-Whitney test, Kruskal-Wallis test, Mood test, Friedman test and run test are also presented with the steps for testing hypotheses and test of significance.
Keywords : Nonparametric Statistics, Small Sample, Neyman-Pearson Test, Fisher Test, Testing Hypotheses, Test of Significance

1. 서 론

수요자의 다양한 요구에 따라 품질 시스템도 소품종 대량생산에서 다품종 소량생산으로 급격히 변화되고 있다. 소품종 대량생산에서는 많은 데이터를 쉽게 구할 수 있어 분포를 가정하는 모수 통계(Parametric Statistics)를 주로 사용하나 다품종 소량생산에서는 소수의 데이터만을 구할 수 밖에 없어 순위, 위치, 순서에 의하고 모수 또는 분포를 가정하지 않는(Distribution-free) 비모수 통계(Nonparametric Statistics)를 활용한다.

Neyman-Pearson 검정(Testing Hypotheses)은 연구가설에 대하여 의사결정(Decision Making)을 내리기 위하여 미리 유의수준 α값을 설정하는 방법이다. 이 방법은 i)가설설정 ii)검정관측 통계량계산 iii)유의수준α 기각치에 의한 판정의 3단계로 구성된다.

* 경원대학교 산업공학과

대부분의 비모수 통계 관련 교재는 이 방법을 사용하며 유의수준 α의 기각치는 사람이 손으로 계산하기 어려우므로 컴퓨터가 수치적분에 의해 구해 놓은 표를 표준화 변수를 이용하여 구한다. 그러나 품질통계에서 α=5%, 1% 설정할 경우 설정근거의 객관성에 문제의 소지가 있다. [2,3]

Fisher 검정(Test of Significance)은 경험적 증거의 강약만을 평가하기 위해 미리 유의수준 α를 설정하지 않고 귀무가설을 기각하기 위한 강한 반증 즉 아주 작은 α를 P-Value 값으로 구하고 이용자의 연구상황에 맞춰 해석을 융통성 있게 하는 것이다.

이 방법은 i)가설설정 ii)P-Value 계산 iii)유의성 판정의 3단계로 구성된다.

MINITAB, SPSS, SAS 등의 통계 패키지에서는 이 방법을 사용하며 관측통계량이 갖는 P-Value는 사람이 손으로 구할 수 없으나 컴퓨터는 쉽게 구할 수 있기 때문에 별도의 표준화 수치표가 없어도 이용할 수 있는 장점이 있다. [2,3]

이렇듯 비모수 통계의 대부분의 교재는 사람이 손으로 계산할 수 없는 기각치 값을 표준화 수치표로 이용하는[4,5] Neyman-Pearson 검정을 소개하고 있고 MINITAB등의 통계 패키지 관련 교재는 컴퓨터가 쉽게 계산할 수 있는 P-Value를 활용하는[1] Fisher 검정을 제시하고 있다.

따라서 본 연구에서는 다품종 소량생산에서 소수의 데이터를 이용하는 비모수 통계를 대상으로 Neyman-Pearson 검정과 Fisher 검정에 의한 검정단계를 비교, 제시하고자 한다.

2. Neyman-Pearson 검정에 의한 비모수 통계[1]

Neyman-Pearson 검정은 i)가설설정 ii)검정관측통계량계산 iii)유의수준α 기각치에 의한 판정 등 3단계로 구성된다.

2.1 1개 모집단 중앙값 검정

2.1.1 부호 검정(Sign Test)

모수검정에서 1개 모집단 모평균 Z_0검정에 해당한다.

1) 가설설정

H_0: $M = M_0$

H_1: i) $M \neq M_0$

ii) $M > M_0$

iii) $M < M_0$

여기서 M:개선후 중앙값

M_0:개선전 중앙값

2) 검정관측 통계량 계산

$$Z_0 = \frac{n^+ - 0.5n}{(0.25n)^{1/2}} \qquad (1)$$

여기서 n : 데이터 수

n^+:M_0보다 큰 데이터의 수

3) 유의수준 α기각치를 표준 정규분포표에서 구해서 Z_0가 기각치보다 클 경우 H_0 기각(유의적)

2.1.2 Wilcoxon 검정(Signed Rank Test)

모수검정에서 1개모집단 모평균 t_0검정에 해당한다.

1) 가설설정

H_0: $M = M_0$

H_1: i) $M \neq M_0$

ii) $M > M_0$

iii) $M < M_0$

2) 검정관측 통계량 계산

$$Z_0 = \frac{WI - \frac{n(n+1)}{4}}{\left(\frac{n(n+1)(2n+1)}{24}\right)^{1/2}} \qquad (2)$$

여기서 WI=M_0를 초과하는 Walsh 평균의 수

+0.5(M_0와 같은 Walsh 평균의 수)

Walsh평균= $\frac{x_i + x_j}{2}$ $i = 1, 2, ..., n$ $j = i, i+1, ..., n$

3) 유의수준 α기각치를 표준 정규분포표에서 구해서 Z_0가 기각치보다 클 경우 H_0 기각(유의적)

2.2 2개모집단 중앙값 검정 : Mann-Whitney 검정(Wilcoxon Rank Sum Test)

모수검정에서 두 모분산이 거의 같다고 생각되는 경우 두 모평균차 t_0검정에 해당한다.

1) 가설설정

H_0: $M_1 = M_2$
H_1: ⅰ)$M_1 \neq M_2$
ⅱ)$M_1 > M_2$
ⅲ)$M_1 < M_2$

2) 검정관측통계량 계산

$$Z_0 = \frac{MH - \frac{n_1(n_1+n_2-1)}{2}}{\left(\frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)}\sum_{i=1}^{n_1+n_2} - \frac{n_1 n_2 (n_1+n_2+1)^2}{4(n_1+n_2-1)}\right)^{1/2}} \quad (3)$$

여기서 MH : 첫 수준의 순위합
n_1 : 첫 수준의 반복수
n_2 : 두 번째 수준의 반복수
r_i : 두 수준 전체의 순위

3) 유의수준 α기각치를 표준정규분포표에서 구해서 Z_0가 기각치보다 클 경우 H_0기각(유의적)

2.3 세 개 이상 모집단의 중앙값 검정

2.3.1 Kruskal-Wallis 검정

모수검정에서 반복수가 다른 1원배치법 분산분석에 해당한다.

1) 가설설정

H_0: $M_1 = M_2 = \cdots = M_l$
H_1: $M_1 \neq M_2 \neq \cdots \neq M_l$

2) 검정관측통계량 계산

$$KW = \frac{12}{\sum_{i=1}^{l} m_i (\sum_{i=1}^{l} m_i + 1)} \sum_{i=1}^{l} \frac{1}{m_i} \left[r_{i.} - \frac{m_i (\sum_{i=1}^{l} m_i + 1)}{2} \right]^2 \quad (4)$$

여기서 m_i : 각 수준의 반복수

l : 수준수

$r_{i.}$: 전체 수준의 순위 r_{ij}에 의한 각 수준의 합

3) 유의수준 α기각치를 $\chi^2(l-1:\alpha)$표에서 구하고 $KW > \chi^2(l-1:\alpha)$이면 H_0기각 (유의적)

2.3.2 Mood 검정

모수검정에서 반복수가 다른 1원배치법 분산분석에 해당한다.

1) 가설설정

H_0: $M_1 = M_2 = \cdots = M_l$

H_1: $M_1 \neq M_2 \neq \cdots \neq M_l$

2) 검정관측통계량 계산

		A_1 A_2 $\cdots$ A_l	$O_{.j}$
M보다 작은 갯수	1	O_{ij}	
M보다 큰 갯수	2		
$O_{i.}$			O

M : 전 수준에서의 중앙값

O_{ij} : $i = 1, 2, \cdots, l \quad j = 1, 2$

$$E_{ij} = \frac{O_{i.} O_{.j}}{O}$$

$$\chi_0^2 = \sum_{i=1}^{l} \sum_{j=1}^{2} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (5)$$

3) 유의수준 α기각치를 $\chi^2(l-1:\alpha)$표에서 구하고 $\chi_0^2 > \chi^2(l-1:\alpha)$이면 H_0기각(유의적)

2.3.3 Friedman 검정

모수검정에서 난괴법(RBD : Randomized Block Design)의 분산분석에 해당한다.

1) 가설설정

H_0: $M_1 = M_2 = \cdots = M_l$
H_1: $M_1 \neq M_2 \neq \cdots \neq M_l$

2) 검정관측통계량 계산

$$FR = \frac{12}{lm(l+1)} \left[\sum_{i=1}^{l} r_{i.} - \frac{m(l+1)}{2} \right]^2 \qquad (6)$$

여기서 l : 수준수
m : Block 수
$r_{i.}$: 각 Block 내에서 순위 r_{ij}의 합

3) 유의수준 α기각치를 $\chi^2(l-1:\alpha)$표에서 구하고 $FR > \chi^2(l-1:\alpha)$이면 H_0기각(유의적)

2.4 런 검정

관리도에서 중심선의 한쪽에 연속해서 나타나는 수를 런(Run)이라 하며 이상원인(Assignable Cause)으로 판정할 경우 사용된다.

1) 가설설정

H_0: 공정이 우연원인(정상, 관리상한)으로 이루어져 있다.

H_1: 공정이 이상원인(비 관리상태)으로 이루어져 있다.

2) 검정 관측통계량 계산

$$Z_0 = \frac{n_R - \dfrac{2n^-n^+ + 1}{n^- + n^+}}{\left(\dfrac{2n^-n^+(2n^+n^+ - n^-n^+)}{(n^- + n^+)^2(n^- + n^+ - 1)} \right)^{1/2}} \qquad (7)$$

여기서 n_R : 각 데이터에서 관리도의 중심값을 뺀 데이터의 부호로 판정된 런의 수
n^- : 중심값보다 작거나 같은 데이터의 수
n^+ : 중심값보다 큰 데이터의 수

3) 유의수준 α기각치를 표준 정규분포표에서 구해서 Z_0가 기각치보다 클 경우 H_0 기각(유의적)

3. Fisher 검정에 의한 비모수 통계[1]

Fisher 검정은 ⅰ)가설설정 ⅱ)P-Value 계산 ⅲ)연구자의 연구조건을 반영한 유의성 판정 등의 3단계로 구성된다. 따라서 Fisher 검정은 단계1의 가설설정은 Neyman-Pearson 검정과 동일하며 단계 3의 유의성 판정은 연구자의 판단에 의존하므로 단계 2의 P-Value 계산을 양쪽검정, 우측검정, 좌측검정 으로 나누어 소개한다. 우측검정일 경우 검정 통계량에 -0.5를 좌측검정일 경우 +0.5로 연속성 수정을 해주면 정확한 P-Value를 계산할 수 있으나 본 연구에서는 이를 생략하기로 한다.

3.1 1개 모집단 중앙값 검정

3.1.1 부호검정

1) 양쪽검정 P-Value $=2\times\Phi(\frac{n^{+}-0.5n}{(0.25n)^{1/2}})$

2) 우측검정 P-Value $=\Phi(\frac{n^{+}-0.5n}{(0.25n)^{1/2}})$

3) 좌측검정 P-Value $=1-\Phi(\frac{n^{+}-0.5n}{(0.25n)^{1/2}})$

여기서 Φ는 우측 면적으로 계산하는 누적 표준 정규분포값이다.

3.1.2 Wilcoxon 검정

1) 양쪽검정

$$\text{P-Value} = 2\times\Phi\left(\frac{WI-\frac{n(n+1)}{4}}{\left(\frac{n(n+1)(2n+1)}{24}\right)^{1/2}}\right)$$

2) 우측검정

$$\text{P-Value} = \Phi\left(\frac{WI-\frac{n(n+1)}{4}}{\left(\frac{n(n+1)(2n+1)}{24}\right)^{1/2}}\right)$$

3) 좌측검정

$$\text{P-Value} = 1-\Phi\left(\frac{WI-\frac{n(n+1)}{4}}{\left(\frac{n(n+1)(2n+1)}{24}\right)^{1/2}}\right)$$

여기서 Φ는 우측면적으로 계산하는 누적 표준 정규분포값이다.

3.2 2개모집단 중앙값 검정 : Mann-Whitney 검정

1) 양쪽검정 P-Value =

$$2\times\Phi\left(\frac{MH-\frac{n_1(n_1+n_2-1)}{2}}{\left(\frac{n_1n_2}{(n_1+n_2)(n_1+n_2-1)}\sum_{i=1}^{n_1+n_2}-\frac{n_1n_2(n_1+n_2+1)^2}{4(n_1+n_2-1)}\right)^{1/2}}\right)$$

2) 우측검정 P-Value =

$$\Phi\left(\frac{MH-\frac{n_1(n_1+n_2-1)}{2}}{\left(\frac{n_1n_2}{(n_1+n_2)(n_1+n_2-1)}\sum_{i=1}^{n_1+n_2}-\frac{n_1n_2(n_1+n_2+1)^2}{4(n_1+n_2-1)}\right)^{1/2}}\right)$$

3) 좌측검정 P-Value = 1- $\Phi\left(\frac{MH-\frac{n_1(n_1+n_2-1)}{2}}{\left(\frac{n_1n_2}{(n_1+n_2)(n_1+n_2-1)}\sum_{i=1}^{n_1+n_2}-\frac{n_1n_2(n_1+n_2+1)^2}{4(n_1+n_2-1)}\right)^{1/2}}\right)$

여기서 Φ는 우측면적으로 계산하는 누적 표준 정규분포값이다.

3.3 세 개 이상의 모집단 중앙값 검정

3.3.1 Kruskal-Wallis 검정

P-Value =

$$\chi^2\left(l-1 : \frac{12}{\sum_{i=1}^{l}m_i(\sum_{i=1}^{l}m_i+1)}\sum_{i=1}^{l}\frac{1}{m_i}\left[r_{i.}-\frac{m_i(\sum_{i=1}^{l}m_i+1)}{2}\right]^2\right)$$

여기서 $\chi^2\left(l-1 : \frac{12}{\sum_{i=1}^{l}m_i(\sum_{i=1}^{l}m_i+1)}\sum_{i=1}^{l}\frac{1}{m_i}\left[r_{i.}-\frac{m_i(\sum_{i=1}^{l}m_i+1)}{2}\right]^2\right)$은 우측면적을 계산하는 χ^2분포값이다.

3.3.2 Mood 검정

P-value = $\chi^2(l-1:\sum_{i=1}^{l}\sum_{j=1}^{2}\frac{(O_{ij}-E_{ij})^2}{E_{ij}})$

여기서 $\chi^2(l-1:\sum_{i=1}^{l}\sum_{j=1}^{2}\frac{(O_{ij}-E_{ij})^2}{E_{ij}})$은 우측면적을 계산하는 χ^2분포값이다.

3.3.3 Firedman 검정

P-Value = $\chi^2(l-1:\frac{12}{lm(l+1)}\left[\sum_{i=1}^{l}r_{i.}-\frac{m(l+1)}{2}\right]^2)$

여기서 $\chi^2(l-1:\frac{12}{lm(l+1)}\left[\sum_{i=1}^{l}r_{i.}-\frac{m(l+1)}{2}\right]^2)\left[\sum_{i=1}^{l}r_{i.}-\frac{m(l+1)}{2}\right]^2)$은 우측면적을 계산하는 χ^2분포값이다.

3.4 런 검정

1) 양쪽검정

P-Value = $2\times\Phi\left(\frac{n_R-\frac{2n^-n^++1}{n^-+n^+}}{\left(\frac{2n^-n^+(2n^+n^+-n^-n^+)}{(n^-+n^+)^2(n^-+n^+-1)}\right)^{1/2}}\right)$

2) 우측검정

P-value = $\Phi\left(\frac{n_R-\frac{2n^-n^++1}{n^-+n^+}}{\left(\frac{2n^-n^+(2n^+n^+-n^-n^+)}{(n^-+n^+)^2(n^-+n^+-1)}\right)^{1/2}}\right)$

3) 좌측검정

P-Value = $1-\Phi\left(\frac{n_R-\frac{2n^-n^++1}{n^-+n^+}}{\left(\frac{2n^-n^+(2n^+n^+-n^-n^+)}{(n^-+n^+)^2(n^-+n^+-1)}\right)^{1/2}}\right)$

여기서 Φ는 우측면적으로 계산하는 누적표준 정규분포값이다.

4. 결 론

본 연구에서는 다품종 소량생산과 같이 소표본을 대상으로 하는 비모수 통계를 대상으로 Neyman-Pearson 검정과 Fisher 검정의 3단계를 체계적으로 비교, 고찰하였다.

사람이 손으로 계산하는 경우 유의수준 α의 기각치 값을 수치표에 의해 구하는 Neyman-Pearson검정을 사용하는 것이 좋으나 유의수준 α의 설정에 대한 객관적인 의문의 여지가 있다. MINITAB의 통계패키지로 검정을 실시할 경우 검정관측통계량이 가지는 P-Value으로 유의성 판정을 하나 이용자의 주관적인 판단이 게재될 위험이 있다. 두 검정은 판정방법을 제외하고는 같은 원리의 통계적 방법을 사용하는 데도 불구하고 이용자는 검정의 단계에서 공식이 다르다는 이유로 큰 혼란을 겪을 수 있으므로 본 연구의 결과를 활용하면 효율적인 운영이 가능하다.

5. 참 고 문 헌

[1] 이승훈, Minitab을 이용한 공학통계 자료분석, 이레테크, 2006.
[2] 최성운, “샘플링 오차에 의한 품질통계 모형의 해석”, 대한안전경영과학회, In Press.
[3] 최종후 외, 학술논문과 통계적 기법, 자유아카데미, 1990.
[4] Conover W.J., Practical Nonparametric Statistics, 3rd Edition, Wiley, 1999.
[5] Wasserman L., All of Nonparametric Statistics, Springer, 2007.