

통계적 추론에서 유의수준, 유의확률과
모수기지의 실무적 해석에 의한 적용방안
-Identification of Implementation Strategy by
Practical Interpretations of Significance
Level, Significance Probability, and Known
Parameters in Statistical Inferences-

최 성 운*

Abstract

The research presents a guideline for quality practitioners to provide a full comprehension of differences in theoretical and practical interpretations of assumed sampling errors of and significance probability of calculated p-value. Besides, the study recommends the use of statistical inferences methods with known parameters to identify the improvement effects. In practice, the quality practitioners obtain the known parameters through systematic quality Database (DB) activities.

Keywords: Understanding, Theoretical and Practical Interpretations, Sampling Errors, Significance Probability, Known Parameters, Quality DB Activities

* 가천대학교 산업공학과 교수

1. 서론

장기간의 공정개선을 추구하는 관리도(Control Chart)와 더불어 치공구개선 등 단기간의 일시적인 개선효과를 파악하기 위해 가설검정(Hypothesis Testing)의 품질통계기법들이 사용된다, 특히 가설검정은 개선후(New)와 개선전(Old)의 비교, 시제품 테스트와 설계 목표치와의 비교, 시험분석 측정값과 KS규격과의 비교 등 개선효과, 시제품 테스트의 확인, KS규격의 확인 등이 활용된다. 만약 가설검정[4]에서 대립가설이 유의적으로 판정되는 경우 새로운 개선변화에 대한 크기를 알아보기 위해 구간추정(Interval Estimation)[1]을 실시한다.

가설검정에서 가설설정시 개선전의 모수 값은 첨자에 Old의 약자인 μ_o 로, 새로운 개선후의 모수값 μ 와 구별하여 사용한다. 귀무가설 $H_o : \mu = \mu_o$ 에서 개선전의 μ_o 를 오른쪽에 위치하는 것은 $x = 3$ 의 수학적 표현에서와 같이 모르는 변수는 오른쪽에, 아는 숫자를 오른쪽에 위치하는 것이 관례화되었기 때문이다. 여기서 μ_o 의 개선 전 모수 값은 \bar{x} 의 통계량 값이 오랫동안 안정되게 계산되어 사용했을 경우 가정할 수 있는 값으로 실제 현장에서는 검사성적서와 작업일지의 DB(Data Base)에 의한 품질통계기법의 적용이 선행되어야 한다. 정규분포에서 μ, σ 두 모수를 모두 알아야 사용할 수 있거나 가설검정에서 개선전의 μ_o, σ_o 를 알고 있다는 의미의 실무적 적용과 해석이 없이, 임의로 \bar{x} 와 s의 통계량 값을 샘플의 크기($n \geq 30$)만을 고려하여 대체하여 적용하고 있다.

특히 대기업과 같이 품질 DB 인프라가 잘 구축되어 개선전의 통계량 값이 축적되어 있는 경우 데이터 행위(Data Behavior)에 대한 기술적, 실무적 판단을 고려하여 개선전의 모수 값의 사용여부를 결정할 수 있다. 만약 안정된 개선전의 모수값과 새로운 개선후의 모수값을 비교하는 1개 모집단의 가설검정을 사용하지 않고 개선전, 후 모두 샘플링에 의한 불안정한 통계값으로 비교하는 2개 모집단의 가설검정으로 적용[2]했을 경우 개선효과는 샘플링오차 α 와 더불어 신뢰할 수 없는 판정결과로 이어진다.

또한 가설검정에서 샘플링 오차 α, β 를 모두 고려하는 경우 샘플의 크기 n이 커지므로 이 중 중요하다고 생각되는 오차 중 하나를 선택하는 방법이 실무에서 많이 활용되고 있다. 그러나 가설 검정에서 β 보다 α 를 유의수준으로 선정한 이유가 귀무가설과 대립가설의 뒤 바뀌는 오류관점에서 단순히 설명하거나, 실제 예를 들어 설명하는 경우 상황에 따른 중요도가 달라져 품질 실무자가 샘플링 오차를 이해하고 적용하는 데 혼란을 빚고 있다.

따라서 본 연구에서는 개선효과를 파악하기 위해 사용되는 가설검정과 구간추정에서의 샘플링 오차의 이론적, 실무적 의미와 신뢰수준, P-Value[3,5]의 적용방안과 모수기지(Known Parameter)의 가설검정의 유용성 및 적용방안을 모집단 수의 관점에서 유형화하여 제시한다. 본 연구의 기여도는 품질 실무자에게 샘플링 오차의 실무적 해석, 검사성적서와 작업일지의 축적된 활동에 따른 모수 기지의 가설검정과 구간 추정의 적용방안을 제시하는 데 있다.

2. 가설검정과 구간추정의 적용

2.1 유의수준과 신뢰수준

가설검정(Hypothesis Test)은 개선효과, 시제품 테스트 결과 확인, 핸드북에 대한 실험분석치의 비교 등 품질개선 활동에서 가장 많이 사용되는 통계적 기법이다. 가설검정에서는 제1종오차(Type I Error), 유의수준(Significance Level) α 와 제 2종 오차(Type II Error)인 β 를 동시에 작게 하려면(또는 신뢰수준($1-\alpha$)와 검출력($1-\beta$)를 동시에 크게 하려면) 샘플의 크기 n 이 커지게 된다. 따라서 적은 n 을 요구하는 실무 검정을 수행하려면 α, β 둘 중 하나를 중요도에 의해 포기해야 한다. 이 경우 α, β 의 의미해석을 각 기법의 용도에 맞게 기술하면 <표 1>과 같다.

<표 1> 품질개선 기법에서 $\alpha, \beta, (1-\alpha), (1-\beta)$ 의 의미

	가설검정	샘플링검사	관리도
α	$H_0 \rightarrow H_1$	양품 \rightarrow 불량품 (생산자 위험)	우연원인 \rightarrow 이상원인
β	$H_1 \rightarrow H_0$	불량품 \rightarrow 양품 (소비자 위험)	이상원인 \rightarrow 우연원인
$(1-\alpha)$	$H_0 \rightarrow H_0$	양품 \rightarrow 양품	우연원인 \rightarrow 우연원인
$(1-\beta)$	$H_1 \rightarrow H_1$	불량 \rightarrow 불량	이상원인 \rightarrow 이상원인

<표 1>에서 관리도인 경우 $\alpha = 0.27\%$ 로 작게, 가설검정에서 $\alpha = 5\%, 1\%$ 로 관리도보다 크게 설정한다. 관리도(Control Chart)와 같이 Lot또는 Batch에 의한 Subgroup의 수가 25개이상(Subgroup Size도 $\bar{x}-R$ 관리도인 경우 4-5개)으로 전체 데이터가 100개 이상으로 장기간의 공정개선의 원인을 Monitoring하는 경우 생산자를 보호하는 취지로 α 를 적게 설계하였다. 이는 수요가 공급을 초과하는 경우 α 로 인한 생산지연에 의한 손실이 β 로 인한 소비자클레임 손실보다 큰 경우에 해당한다. 가설검정에서는 관리도보다 α 를 크게 설정하는 이유는 관리도에서 하나의 Subgroup에 대해 실시하는 것이 검정으로 관리도보다 샘플의 크기 n 이 작게 요구되기 때문이다. 따라서 품질실무자는 공정의 장기간의 변화와 근본적인 원인에 대한 개선을 요구할 시 관리도를, 치공구개선 같은 약간의 개선 효과를 평가할 시 가설검정방법을 사용할 것을 권고한다. 그러면 실무에서 사용되는 검정에서는 샘플링오차의 상대적 중요도에 대한 이해를 위해 <표 2>와 같은 예를 들어 α, β 를 설명한다.

<표 2> 실무 검정에서의 α , β 의 의미

용도	개선효과	질병	화재	죄	성적
사용자	품질실무자	의사	소방관	판사	선생님
α	변화無→변화有	병無 →병有	화재無→화재有	죄無 →죄有	정답→틀린답
β	변화有→변화無	병有 →병無	화재有→화재無	죄有 →죄無	틀린답→정답

<표 2>에서 개선효과, 질병, 화재의 예인 경우는 β 가 더 중요하며 죄와 성적인 예에서는 α 가 더 중요한 것으로 사람들에게 인식된다. 따라서 가설검정의 α , β 에 대한 의미는 사례에 따라 사람들에게 다르게 인식되므로 신중하게 제시되어야 한다. 가설검정에서는 관리도와 다르게 n 이 좀 더 큰 관점에서 α 가 상대적으로 크게 설계되었다는 점과 <표 2>의 개선효과 예에서 α 보다 β 가 중요하므로 β 를 작게 즉 α 를 크게 설계했다는 해석이 가능하다. 이에 대한 해석이 명확하지 않은 경우 α , β 를 동시에 고려하여 n 을 결정한 후 가설검정을 실시할 수 있다. 가설검정에서 남에게 실수가 많은 검정결과를 줄 경우 $\alpha=5\%$ 로, 실수가 적은 검정결과를 받고 싶은 경우 $\alpha=1\%$ 의 유의수준을 사용한다.

가설검정에서 유의적일 경우 즉, 개선효과가 있을 경우(H_0 기각) 새로운 모수에 대한 변화의 크기를 통계량의 구간값으로 파악하는 방법이 구간추정(Interval Estimation)이다. 구간추정에서는 공식의 절약적인 표현을 위해 가설검정의 유의수준 α 와 검정통계량으로 신뢰수준($1-\alpha$)와 구간추정식을 유도한다. 즉, 정규분포의 양측검정의 바깥면적인 α 는 신뢰수준인 안쪽면적의 $(1-\alpha)$ 의 Z 통계량의 위치가 같다는 원리를 적용한다. 신뢰수준 $(1-\alpha)$ 는 불완전한 넓은 구간의 통계량으로 모수에 대한 구간추정결과를 줄 경우 99%로, 좀 더 정밀한 구간추정의 결과를 받고 싶은 경우 95%로 받는다.

<표 1>에서 가설검정과 관리도와 다르게 샘플링검사에서는 α 와 β 를 용도에 따라서 동시에 사용하는 경우(규준형검사)가 있는데 통상 $\alpha=5\%$, $\beta=10\%$ 로 생산자 위험을 소비자 위험보다 적게 설계해 주고 있는데, 이는 과거 2차 세계대전 당시 미국에서 군수물자를 납품하는 민간 생산 업체에 대한 인센티브 차원에서 효율적인 검수를 수행했던 근대적 방법으로 완벽한 고객만족의 품질을 목표로 하는 기업에서는 적용할 수 없는 낡은 기법이다.

2.2 P-Value의 의미와 산출방법

유의수준 α 는 정규분포의 양쪽꼬리에 대한 면적으로 이를 손으로 계산할 수 없어 면적에 영향을 주는 여러 변수를 하나의 표준화 변수로 변수변환하는 것이 Z , t , χ^2 , F 검정 통계량이다. 이 표준화 변수로 정규분포의 면적을 컴퓨터가 적분 계산해 놓은 면적을 찾는 것이 통계책의 부록의 손에 의한 방법으로 MINITAB에 의한 품질개선을 실시하는 실무자에게는 무용지물의 없어져야 할 방법이다. 과거 면적에 대한 표준화 변수와 검정통계량의 값으로 유의성을 판정하는 대신, 주어진 α 의 면적에 대해 검정통계량에 의해 계산되어진 P-Value를 비교하여 유의성을 판정한다. P-Value에 대한

해석은 두가지로 가능하다. 첫째로 P-Value는 검정통계량이 갖는 귀무가설 H_0 분포의 꼬리면적으로 유의수준 $\alpha = 5\%, 1\%$ 보다 작은 경우 샘플링 오차 범위내에 존재하므로 H_0 가설을 기각한다는 해석이다. 두 번째 해석은 검정통계량 값이 $\alpha = 5\%, 1\%$ 의 기각치(Critical Value)보다 큰 경우 P-Value가 α 보다 작아지며 H_0 가설을 기각할 수 있다. MINITAB이 검정에서는 검정통계량과 P-Value 모두를 제공해 주나 품질 실무자는 P-Value와 2.1쪽의 α 의 사용방법을 지침으로 H_0 가설의 기각여부를 판정할 수 있다. Z 우측검정에서 $Z=2.10$ 인 경우 MINITAB의 정규분포에서 평균 0, 표준편차 1, 입력상수 2.10을 대입하면 좌측면적=0.982136이 나오므로 P-Value는 우측면적 P-Value= $1-0.982136 \approx 0.018$ 로 산출된다. 좌측검정은 Z의 음수값을 양수값으로 바꾸면 우측검정과 같이 계산된다. Z 양측검정에서 $Z=2.70$ 인 경우 MINITAB의 정규분포에서 평균 0, 표준편차 1, 입력상수 2.70을 대입하면 좌측면적=0.996533이 나오므로 양측면적 P-Value= $(1-0.996533) \times 2 \approx 0.007$ 로 산출된다.

2.3 그래프에 의한 가설검정과 구간추정

수치 데이터를 일정한 모양의 그림으로 표현한 것이 분포이고 1.0점 만점으로 상대도수에 의해 비교하기 위한 척도가 확률로 확률분포는 효과적이지 못하지만 효율적인 측면에서 시각화의 장점으로 인해 통계적 기법에서는 기본 가정의 전제조건이 된다. 이와같이 기하학적, 도형적 그림은 가시관리의 관점에서 이해가 쉬어 MINITAB에서도 가설검정과 구간추정을 위해 Box-Plot, 신뢰구간 그래프 등을 제공한다.

평균추론은 두 그림의 중심 축 위치로 제품 스펙(Specification)의 기준치와 함께 정확도(편의, 치우침)의 평가를 내릴 수 없으며 두 그림의 산포의 폭넓이를 제품스펙의 허용차 또는 공차와 함께 정밀도의 평가를 시각적인 관점에서 효율적으로 확인할 수 있다.

3. 모수 기지의 검정, 추정의 유용성

<표 3>에서 3개 이상의 모집단을 최적의 제품기술이나 생산기술조건을 설정하는 품질실험설계에서 사용하는 방법이다. 식스시그마 프로젝트 수행시 DMAIC 5단계 프로세스 중 Analyze 단계에서 핵심인자 선정시 또는 품질분임조의 QC Story 15단계 중 현상파악, 결과분석에서 개선전과 개선후를 비교할 시 2개 모집단의 가설검정 방법을 대부분 사용한다. 1개 모집단을 적용할 수 있는 가설검정을 2개 모집단으로 적용할 경우 검정결과의 신뢰도에 큰 영향을 준다. 즉 1개 모집단은 개선전의 통계량 값인 s, \bar{x} 를 각각 σ_0, μ_0 로 장기축적하여 사용한다는 의미로 첨자 Old는 개선전의 값을 알고 있다(Known)는 의미이다. 물론 현장에서 검사성적서나 작업일지를 DB화하여 축적해서 안정되게 나오는 통계량 값을 모니터링하고 유지하는 어려움과 번거로움은 있으나 모수에 버금가는 개선전의 값을 기준으로, 개선 후 샘플이 크기 n 과 샘플링 오차 α 에 크게 영향을 받는 개선후의 통계량 값을 비교할 경우 실수는 오직 개선 후에서만 일

어날 수 있다. 그러나 개선전의 값을 축적하지 않아 2개 모집단의 가설검정을 적용할 경우, 두 모집단의 샘플의 크기 n_1, n_2 와 각각의 샘플링 오차 α 에 의한 불확실성이 커져 검정의 결과를 신뢰할 수 없게 될 수도 있다. 따라서 DB가 용이한 대기업의 경우 품질 개선시 <표 3>과 같은 1개 모집단의 가설검정을 적용해야 하며 폭의 비교시 χ^2 등분산검정과 축의 비교시 t 검정보다는 Z검정을 사용해야 한다,

<표 3> 모집단별 가설검정의 비교

	1개 모집단	2개 모집단	3개이상의 모집단
정밀도 검정 : 분포의 폭 비교 : 등분산검정	χ^2 검정	F검정	Bartlett 검정 Levene 검정
정확도검정 : 분포의 축 비교	1) σ Known Z 검정 2) σ Unknown t 검정	1) σ_1, σ_2 Known Z 차 검정 2) σ_1, σ_2 Unknown a) Pooled t차 검정 b) Unpooled t차 검정 3) Paired t 검정	ANOVA

4. 결 론

본 연구에서는 품질개선효과를 확인하기 위해 사용되는 가설검정과 구간추정의 올바른 적용방안을 제시하였다. 특히 가설검정에서 샘플링 오차의 중요도에 따른 유의수준의 의미와 신뢰수준의 사용방안을 제시하였으며 표준화 변수를 이용하는 기존의 검정 방법과의 차이를 P-Value의 의미와 함께 기술하였다. 또한 평균과 분산을 기하학적 의미에서의 축과 폭으로 제품스펙과의 비교에서의 정확도와 정밀도로 확장해석하여 모수 기지의 검정, 추정의 유용성과 함께 적용방안을 제시하였다.

5. 참 고 문 헌

[1] 최성운, “정확도 및 정밀도 관점에서의 통계적 품질 기법의 해석”, 대한안전경영과학회지, 9(1)(2007):119-131.
 [2] 최성운, “동질성 및 이질성 모집단의 가설검정과 구간추정의 비교”, 대한안전경영과학회 춘계학술대회 발표문집, (2009):365-370.
 [3] Lesik S., Applied Statistical Inference with MINITAB, Chapman and Hall/CRC, 2009.
 [4] Mendenhall W., Sincich T., Statistics for Engineering and the Sciences, 5 Edition, Prentice Hall, 2006.
 [5]. <http://www.minitab.com>