

KPA rating 데이터 보정을 통한 성숙도 설문서(MQ) 선정 모델의 신뢰성 평가

김우송, 이은서, 이경환
(kws0198^o, eslee, kwlee)@object.cau.ac.kr

Reliability Test of Maturity Questionnaire Selection Model Through KPA Rating Data Calibration

Woo-Song Kim^o, Eun-Seo Lee, Kyung-Whan Lee

School of Computer Science and Engineering, Chung-Ang University

요 약

소프트웨어 공학이 소프트웨어 시스템에 관한 방법론, 기술 및 툴 등의 유지보수와 개발에 중점을 두어 왔는데 최근에는 프로세스 개선과 프로세스 능력수준의 향상에 초점을 두는 방향으로 발전하면서 CMM 및 SPICE 활동이 증가하고 있다. 이와 같은 심사기법은 대규모의 회사에서 주로 시행되고 있어서 중소기업의 조직을 위한 간략한 심사기법의 도입이 요구되는 있는 상황이다. 본 논문에서 제시하는 심사기법은 CMM 심사를 위한 KPA 설문서의 rating 방법을 응용한 것으로서 SPICE 심사를 받은 국내 기업 중 일부 회사를 대상으로 하였다. 이 방법론에 대한 신뢰성의 평가는 아직 미흡한 상황이다. 이를 위해서 통계학적 접근방법을 도입하였는데 사용된 통계 기법은 상관계수를 통한 가설검정이다. 그 결과 성숙도 설문서(MQ) 선정모델의 적합성을 통계적 기법을 통해서 검증하였다.

$$\rho = \frac{\sigma_{xy}}{\sigma_x \times \sigma_y}$$

$$\sigma_{xy} = E(xy) - E(x) \times E(y)$$

σ_x, σ_y 는 두 변수 x, y 의 표준편차
 $E(x), E(y)$ 는 두 변수 x, y 의 평균

1. 서 론

소프트웨어 공학이 소프트웨어 시스템에 관한 방법론, 기술 및 툴 등의 유지보수와 개발에 중점을 두었다면 최근에는 프로세스 개선과 프로세스 능력수준의 향상에 초점을 두는 방향으로 발전하고 있다. 프로세스 개선에 대한 수요가 증가하면서 여러 조직들이 SPICE(Software Process Improvement Capability Determination) 및 CMM(Capability Maturity Model) 심사를 통한 프로세스 개선활동을 수행하고 있다. 하지만 이와 같은 심사는 시간과 비용이 많이 소모되기 때문에 주로 대규모 조직에서 수행되고 있고 중소기업의 조직에서는 SPICE 및 CMM 활동이 미흡하다. 또한 대규모 조직에서도 경영상의 이유로 이와 같은 심사를 받기가 곤란할 수 있다. 그래서 시간과 비용을 줄이면서 심사를 받을 수 있는 간소화된 심사기법의 필요성이 대두되고 있다. 여기에서 소개하는 심사기법은 CMM에서 사용하고 있는 KPA(Key Process Area) rating 방법을 응용한 방법이다. CMM의 KPA 방식은 프로세스 영역을 18개의 핵심프로세스 영역(KPA)으로 나누어서 각각의 영역에 대한 rating을 합산해서 산출해 내는 방식이다 [1]. 한편 본 논문에서 제시한 방법은 MQ(Maturity Questionnaire)의 KPA rating인데 CMM의 KPA 방법에서 사용하는 핵심 프로세스 영역에서 프로세스 개선을 위해 중요하다고 여겨지는 항목들을 pareto rule에 의해 선정하는 방식이다. 하지만 아직까지는 이 방식에 대한 체계화된 평가가 이루어지지 않고 있는 상황이다. 그래서 본 논문에서는 KPA 항목설정의 적합성을 통해서 MQ 선정모델의 신뢰성을 상관계수를 사용해서 통계적으로 검증해 보았다.

2. 관련연구

2.1 통계적 가설검증을 위한 용어 정의

상관계수(ρ)란 두 변수간에 존재하는 상호연관성을 수치적으로 표현한 것으로서 $-1 < \rho < 1$ 사이의 값을 취하며 값이 0일때는 두 변수간의 상관관계가 없다는 것을 나타낸다. 계산식은 다음과 같다.

t 분포란 어떤 샘플들의 모집단의 분포가 표준정규분포를 따를 때 이 샘플들이 이루는 분포이다 [2].

신뢰구간은 신뢰폭(信賴幅)이라고도 한다. 구간추정(區間推定)에서는 미리 주어진 하나의 수 α 에 대하여 $Pr(\theta_1 < \theta < \theta_2) = 1 - \alpha$ 가 성립되는 θ_1 과 θ_2 를 표본 x_1, x_2, \dots, x_n 으로부터 구한다. 이때 $1 - \alpha$ 를 신뢰도 또는 신뢰계수라고 하며, 구간 (θ_1, θ_2) 을 신뢰도 $1 - \alpha$ 의 신뢰구간이라고 한다. [2]

귀무가설은 설정한 가설이 진실할 확률이 극히 적어 처음부터 버릴 것이 예상되는 가설이다.

대립가설은 귀무가설에 반대되는 가설이다.

2.2 SPICE와 CMM의 KPA

2.2.1 SPICE

미국의 CMM과 더불어 소프트웨어 프로세스의 능력을 평가하는 ISO 표준으로서 총 40개 범주의 프로세스 영역을 평가한다. [3]

2.2.2 CMM의 KPA

CMM에서 사용하고 있는 평가기법으로서 전체 프로세스 영역을 18개 세부영역으로 나누어서 프로세스 능력을 평가한다. CMM에서는 프로세스 능력(Capability)을 성숙도(maturity) 수준을 기준으로 5단계로 구분한다. 핵심 프로세스 영역은 조직이 소프트웨어를 개선하기 위해 어떤 부분을 집중하여 수행하는지 파악할 수 있다. 본 연구에서 채용한 성숙도 설문서는 SEI가 1998년에 개발한 Maturity Questionnaire로서 CMM/KPA의 Key Practice Questionnaire에 대응하여 만든 설문서(MQ, Maturity Questionnaire)이다. [3] [6].

2.3 MQ 선정모델

2.3.1 MQ(Maturity Questionnaire)

CMM의 KPA rating 설문서를 응용한 것으로서 본 논문에서 통계 검증을 위해 사용된 설문서이다. <표1>은 MQ 모델링에 사용된 기초표본(1)의 MQ rating 데이터이다.

<표 1> MQ rating 데이터

구분	100	90	80	70	60	50	40	30	20	10	합	
1. 고객만족도 향상	100	75	50	50	100	75	75	50	50	50	7143	2047
2. 고객만족도 향상	100	75	50	50	100	75	75	50	50	50	7500	1630
3. 고객만족도 향상	100	75	50	50	100	75	75	50	50	50	7143	1537
4. 고객만족도 향상	100	75	50	50	100	75	75	50	50	50	7143	2087
5. 고객만족도 향상	100	75	50	50	100	75	75	50	50	50	7500	1830
6. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
7. 고객만족도 향상	100	75	50	50	100	75	75	50	50	50	7500	1800
8. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
9. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1631
10. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
11. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
12. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
13. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
14. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
15. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
16. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
17. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
18. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
19. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
20. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
21. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
22. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
23. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
24. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
25. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
26. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
27. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
28. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
29. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750
30. 고객만족도 향상	75	50	50	50	100	75	75	50	50	50	4750	1750

3. 가설검증을 위한 MQ의 KPA rating 데이터 수집

가설검증을 위해서 SPICE 심사와 CMM/KPA에 대한 SEI 성숙도 설문서 rating에 대한 데이터를 수집했다. 심사는 심사의 목적과 일치하는 문서화된 절차에 따라서 수행되었다. 그리고 심사 프로세스는 계획, 데이터 수집, 데이터 확인, 프로세스 rating, 보고와 같은 행위들을 최소한도로 포함하였다. 심사의 신뢰성은 ISO Guide 62와 TR 15504, part 6에 의해 교육 및 인증되었으며, 성숙도 설문서에 대해 rating한 책임있는 관계자를 확인하였다. SPICIE(ISO TR 15504, ver.3.3, 1998)로 해당 업체를 심사한 자격있는 심사원이 마찬가지로 SEI 성숙도 설문서를 rating했다. 심사원들이 심사한 프로젝트는 다음과 같이 구성되어 있다[4][5].

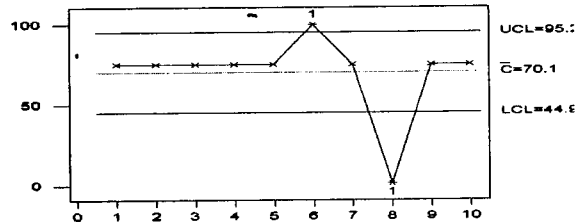
- 심사한 표본들의 수: 4개의 기초표본
- 프로젝트 형태: 디지털 시스템(제품)의 내장형 소프트웨어
- 프로젝트 기간: 3-4개월
- 개발 인력 수: 해당 사업부에 근무하고 있는 100-400
- 개발자의 경험: 3년 이상의 경력자가 70%

4. MQ 선정모델의 신뢰성 평가

4.1 심사원의 KPA 심사 rating에 관한 평가

Data collection한 4개의 기초표본을 대상으로 rating을 평가하였다.

1) 관리도 기법을 이용 - MQ 각 항목에 관한 심사원의 rating에 대하여 $\mu \pm 3\sigma$ 를 관리한계선으로 잡고 이 범위를 넘어가는 rating은 제외시켰다. <그림2>에서는 10개의 샘플링 데이터 중 관리한계선을 넘어가는 데이터의 예가 제시된다. 그림에 제시된 빨간선은 관리한계선으로서 이 선을 넘어가는 데이터는 삭제하는데 6번 데이터와 8번 데이터는 관리한계선을 넘어갔으므로 상관계수를 계산할 때 삭제한다.



<그림2> A 샘플의 rating 데이터에 대한 관리도

2) 관리한계선에 근접한 데이터를 삭제하는 보정기법 1차로 관리한계선을 벗어나는 rating이 2개가 될 경우 상한과 하한의 편차가 심한 2개를 선택하여 제외한다. 즉 관리한계선에 근접한 심사원의 rating 데이터를 1차 보정하고 다시 그 기준을 적용해서 벗어나는 rating 데이터는 삭제하는 방법이다.

4.2 KPA 일반항목과 MQ 항목간의 연관성 평가

본 논문에서 사용한 MQ는 Pareto rule을 이용하여 작성되었기 때문에 CMM의 KPA 세부항목들과 약간의 차이가 있다. 그래서 상관계수를 사용하여 MQ 항목선정의 적합성 여부에 대한 평가를 통계적으로 검증해보았다.

4.2.1 상관계수로 기초표본(1)의 MQ에 대한 신뢰성 평가

A 샘플의 rating 데이터에 관한 가설의 설정 및 검증은 다음과 같이 진행하였다.

2.3.2 Pareto rule

상위 20%가 전체를 결정짓는다는 법칙으로서 80-20 법칙이라고도 한다. 이탈리아 경제학자 파레토가 사회현상 속에서 발견한 원리로 고객의 20%가 매출의 80%를 차지한다는 것이다. 즉 핵심적인 소수가 다수보다 더 많은 일을 하고 있다는 것으로서 본 논문에서는 MQ 항목으로 선정된 상위 항목들이 전체항목의 대표로 사용되는 것을 보여주기 위해 채택되었다.

2.3.3 MQ 모델

(1) MQ 선택

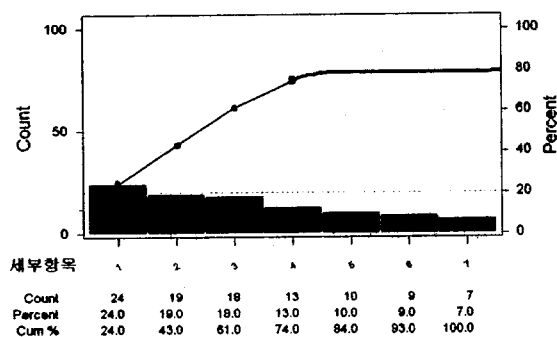
CMM/KPA의 Key Practice Questionnaire에 대응하여 SEI가 1998년에 만든 Maturity 설문서로서 CMM level을 심사할 수 있는 설문항목이다.

(2) Pareto rule에 의한 MQ 모델의 Tailoring

KPA 세부항목에서 MQ 항목을 선정한 기준은 Pareto rule인데 세부항목 중에서 누적 백분율이 80%가 되는 4개의 MQ항목이 선정되었다. <표 2>와 <그림 1>에서는 세부항목 중에서 상위 4개(1,2,3,4)의 항목들이 MQ 항목으로 선정되는 예와 이에 대한 Pareto 차트를 보여준다. 최종적으로 Pareto rule을 적용하여 테일러링된 MQ모델이 <표 1>의 MQ rating 데이터이다.

<표 2> Pareto rule에 의한 MQ 항목 설정의 예

구분	세부항목에 대한 중요도(%)											
	1	2	3	4	5	6	7	8	9	10		
1	25	25	20	30	24	25	20	20	30	25	20	24
2	20	15	20	15	20	20	24	20	15	20	20	19
3	20	15	15	15	20	10	20	23	15	20	25	18
4	15	15	15	10	13	15	10	12	13	10	15	13
5	12	10	10	10	11	10	8	9	10	10	10	10
6	3	12	10	15	5	9	8	10	11	7	9	9
7	5	8	10	5	7	11	10	6	6	8	11	7
8	100	100	100	100	100	100	100	100	100	100	100	100



<그림 1> Rating할 MQ 선정을 위한 Pareto 차트

1) 가설의 설정 및 검증

1. KPA 일반항목의 속성과 MQ 항목의 속성간에 관계없다는 가설을 세운다
2. 가설검증에 사용된 KPA 일반항목과 MQ 항목은 이변량 정규분포에 따른다는 가정에 기초를 둔다.
3. 가설검증 절차
 1. 유의수준 α 와 검정통계량 $t = \sqrt{S^2}$ 에 기초를 둔다.
신뢰구간은 99.7%수준으로 정한다.
 2. 귀무가설 $H_0: \rho = 0$ 의 검정(ρ 는 상관계수)
 3. H_0 를 양측대립가설 $H_1: \rho \neq 0$ 에 대해 검정하기 위해 $t_1 < -t_{\alpha/2}$ 또는 $t_1 > t_{\alpha/2}$ 이면 H_0 를 기각한다. 여기서 $t_1 > t_{\alpha/2}$ 는 $P(t_1 > t_{\alpha/2}) = \alpha/2$ 인 기각치이며 확률변수 t_1 는 자유도 $\nu = (n-2)$ 인 t분포를 따른다

2) 기초표본(1)에 대한 가설 검증

4. 기초표본(1)의 KPA 일반항목과 MQ 항목간에 상관관계가 없다는 가설을 검증
 1. 표본상관계수 $r = 298/319 = 0.9335$
 2. $S^2 = \{(1-0.9335^2)/(9-2)\}^{1/2} = 5.8234$
 3. $\rho = 0$ 일때 t통계량 $t_{1,2} = r/\sqrt{S^2}$ 은 자유도 $(n-2)$ 인 t분포를 따르기 때문에 $r/\sqrt{S^2} = 5.8234$, 자유도 = 7과 같이 된다.
 4. 기각역이 $P(t_{1,2} > t_{\alpha/2}) = 0.0015$ 의 t값은 ± 3.79 가 된다.
 5. t통계량이 기각치를 초과하였으므로 귀무가설을 기각한다. 즉 KPA 일반항목과 MQ 항목간의 연관성은 신뢰도 99.7% 수준에서 연관있다고 결론을 내린다

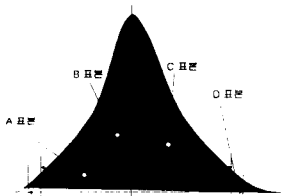
기초표본(1)에 대한 가설 검증의 결과 신뢰도 99.7% 수준에서 KPA 일반항목과 MQ 항목간에 관계가 없다는 귀무가설을 기각하고 두 개의 항목간에 관계가 있다는 대립가설을 채택함으로써 MQ 선정모델의 적합성이 검증되었다.

4.2.2 데이터 보정 전 4개의 기초표본 모두에 대한 가설 검증

전체 기초표본에 대한 상관계수와 검정량 t값, 그리고 이에 대한 신뢰구간이 <표 3>와 <그림3>에 제시된다.

<표 3> 보정전 데이터에 대한 가설검정 결과

일반항목과 MQ 항목	상관계수	검정량 t	신뢰구간
기초표본(1)	0.9335	5.8234	99.9%
기초표본(2)	0.9167	5.6239	99.9%
기초표본(3)	0.9317	7.7127	99.9%
기초표본(4)	0.8319	4.2465	99.7%



<그림 3>상관계수에 대한 신뢰구간

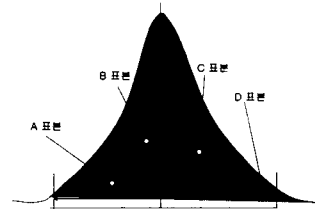
이 결과는 데이터에 대한 보정이 없을 때 신뢰구간 99.7% 수준에서 KPA rating 데이터의 일반항목과 MQ 항목간의 속성이 서로 연관있다는 것을 통계 검증으로 보여 주었다. <그림 3>에서는 D표본을 제외한 3개 표본이 신뢰구간 99.9%안에 포함되어 있는 것을 볼 수 있다.

4.2.3 데이터 보정 후 4개 기초표본에 대한 가설검정

데이터 보정전의 수치는 일반항목과 일반항목에 따른 세부항목간의 연관관계가 상당하다는 것을 보여주지만 절대적인 수치에는 조금 미흡하다. 그래서 본 논문에서 제시한 보정방법을 사용해서 다시 MQ rating 데이터에 대한 상관계수를 구하였다.<표 4>와 <그림 4>에서는 그 결과가 제시되어 있다.

<표 4> 보정후 데이터에 대한 가설검정 결과

일반항목과 MQ 항목	상관계수	검정량 t	신뢰구간
기초표본(1)	0.9412	6.4427	99.9%
기초표본(2)	0.9283	6.1235	99.9%
기초표본(3)	0.9385	8.2324	99.9%
기초표본(4)	0.8917	5.5725	99.9%



<그림 4> 상관계수에 대한 신뢰구간

5. 보정하지 않은 데이터에 대한 가설 검증과 보정한 데이터에 대한 가설 검증의 결과비교

보정전 데이터에 대한 상관계수의 신뢰구간을 구했을 경우에는 3개 표본만이 신뢰구간 99.9%에 포함되고 나머지 1개표본은 신뢰구간이 99.7%였던 반면 본 논문에서 제시한 보정기법을 사용해서 상관계수를 구한 결과 신뢰구간 99.9% 수준에서 4개표본 모두 일반항목과 MQ 항목간의 연관관계가 있다고 통계적으로 검증할 수 있었다. 이를 통해서 MQ 선정모델이 신뢰구간 99.9% 수준에서 적합하다는 결론을 이끌어 낼 수 있었다.

6. 결론 및 향후 연구방향

본 논문에서는 MQ의 KPA rating 데이터에 대한 신뢰성을 상관계수를 통해 평가하고 그 결과로 MQ 선정모델의 신뢰성을 통계적으로 검증하였고 KPA rating 데이터에 대한 보정방법을 제시하여 기존의 99.7% 신뢰구간을 99.9% 구간으로 증가시켰다. 하지만 실제 raw 데이터에서의 rating값이 SPICE수준과는 약간의 차이가 발생하였다. SPICE 심사가 정확히 수행된다는 점에서 KPA rating에 대한 보정이 필요할 것이다. 현재 MQ 항목에 대한 보정모델이 연구되고 있다. 향후 연구 과제로는 MQ 선정모델의 신뢰성을 이용해서 KPA 일반항목의 보정모델에 관해서 연구하고자 한다.

참고문헌

- [1] www.comjournal.co.kr/trend/Tech/tech06116.asp
- [2] 6 시그마 경영 수행기법, 건국대학교 출판부, 2000.
- [3] Kyung Whan, Lee, "Modeling for High Depending Computing", Keynote Speech, Proceeding of the 5th Korea Conference on Software Engineering, Feb. 20-22, 2003.
- [4] KSPICE, "SPICE Assessments in Korea", 2002.
- [5] KASPA, "SPICE 심사사례", www.kaspa.org, 2002.
- [6] Boehm, Software Cost Estimation, COCOMO II, PH, 2000, PP 34-40