

통계기법 점검표 작성시 타당성 평가방법

이상복¹, 김말숙²

요약

점검표 설계상의 타당성 평가이론에 근거한 점검표 설계 단계별 타당성 평가와 점검표 사용목적에 따른 타당성 평가방법을 제안하여 한국의류학회지를 대상으로 평가이론에 입각하여 점검표의 각 단계별 타당성 평가기준을 설정하였다.

주제어: 점검표, 타당성 평가, 통계적 방법, 통계상담

1. 서론

오늘날을 정보화 사회라 일컫는만큼 컴퓨터 과학의 발달로 세계는 점차 복잡하고 다양해지며, 여기에 다량의 정보를 신속하고 정확하게 처리하여 분석하기 위한 통계정보 이용에 있어서 연구자들이 많은 통계적 오류를 범하게 된다. 따라서, 통계적 방법을 활용하는 연구자가 통계학에 대한 이해의 부족에서 범할 수 있는 통계적 오류를 줄이고, 통계적 방법 활용의 타당성 평가를 위한 점검표에 대한 연구가 Schor와 Karten(1966), Ford와 Tortora(1978), White(1979), 윤기중 등(1987), Tabachnick와 Fidell(1989), 최종후(1989, 1995) 등에 의해 활발히 이루어졌다. 그러나, 통계적 방법을 활용하는 연구자가 통계상담, 통계적 방법을 활용한 논문에 대한 타당성 평가, 통계패키지의 이용 등에 있어서 많은 어려움이 있다.

본 논문에서는 웹상에서 점검표를 이용하여 통계상담을 할 수 있는 기초자료가 되고, 또한 연구자의 연구방법과 점검표 사용목적에 따른 연구 논문의 통계적 타당성 평가방법을 제안한다. 이를위한 웹상의 양방향 원격 통계상담이 가능한 점검표의 설계는 그림1과 같이 여섯 단계로 이루어진다. 1단계에서는 연구방법에 따라 조사와 실험으로 분리하여 각기 다른 점검표를 사용한다. 2단계에는 1단계의 조사나 실험으로 분리된 후 점검표의 사용목적이 통계상담 필요 유무인지, 아니면 연구 논문의 타당성 평가인지에 따라 두 가지로 분리된다. 3단계에서는 연구자의 통계학에 대한 지식 정도를 파악하기 위한 것으로 다섯 개의 질문 항목으로 구성된다. 4단계는 연구설계를 점검하는 단계로 조사설계와 실험설계로

¹712-702 경북 경산시 하양읍 금락1리 330번지 대구효성가톨릭대학교 정보통계학과 부교수

²712-702 경북 경산시 하양읍 금락1리 330번지 대구효성가톨릭대학교 전산교육원

분리하여 점검하게 되며, 5단계에서 점검표 사용 목적이 통계적 타당성 평가라면 통계기법의 타당성을 평가한다. 6단계에서는 결론 단계로써 점검표 사용 목적에 따라 달라진다. 점검표 사용 목적이 통계상담 필요 유무를 알기 위한 것이라면 2단계의 통계학의 지식수준과 3단계의 연구설계의 타당성 평가를 기초로 하여 연구자의 통계상담의 필요 유무와 평가 결과에 대한 해설을 첨부함으로써 오류를 검토 수정하여 통계적 오류를 줄일 수 있게 한다.

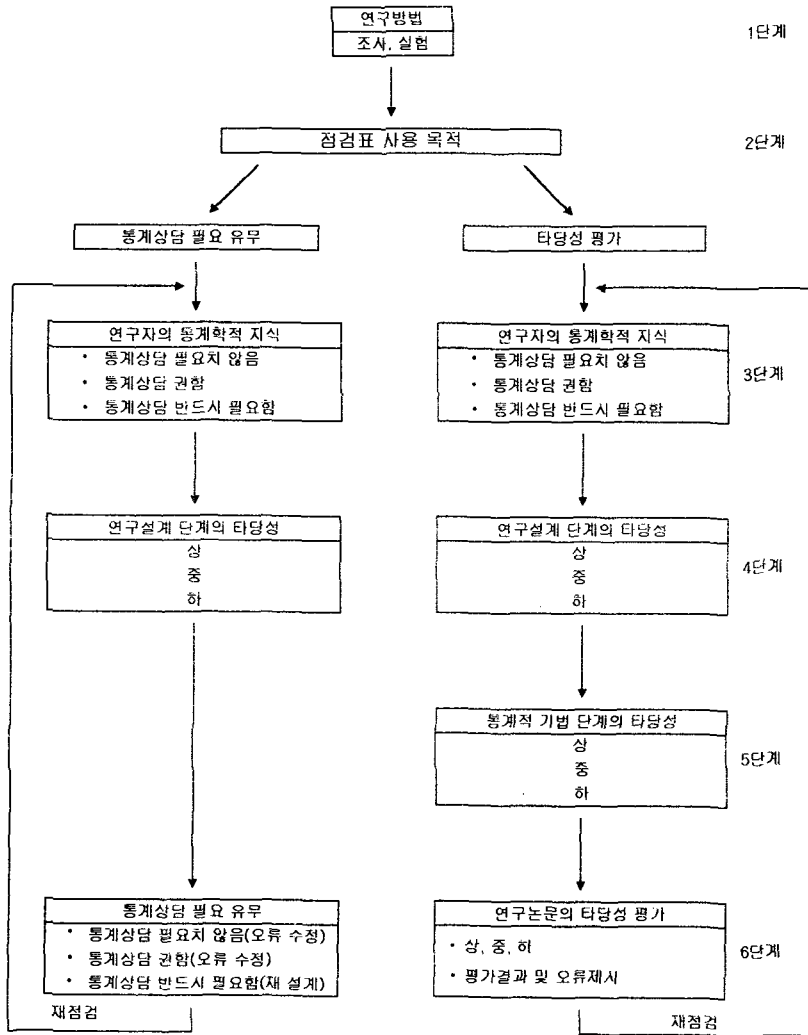


그림 1 원격 통계상담을 위한 점검표의 설계 또한 통계적 타당성 평가를 위한 목적이라면 3단계의 통계학적 지식 정도, 4단계 연구설계,

표 1: 연구설계(4단계)

조사설계	정확하게 수행됨	잘못수행됨	해당사항없음	실험설계	정확하게 수행됨	잘못수행됨	해당사항없음
1. 연구문제 2. 연구가설 3. 모집단의 정의 4. 표본추출방법 5. 표본의 크기 6. 표본의 대표성 7. 측정의 타당도 8. 측정의 신뢰도 9. 예비조사 실시 10. 수집된 자료의 검토				1. 연구문제 2. 연구가설 3. 실험계획방안 4. 실험의 통제 5. 피험자의 표집과 배치 6. 실험처치 7. 반복화 8. 확률화 9. 블럭화 10. 검사 또는 측정기구 11. 예비실험 12. 수집된 자료의 검토			

5단계 통계적 방법의 타당성 평가를 종합하여 연구 논문의 통계적 타당성 평가점수와 평가 결과를 제시하여 오류를 수정하게 한다. 본 논문에서 제안된 점검표는 표1과 같이 연구설계(조사설계와 실험설계)와 표2의 통계적 기법으로 구성되어 있다.

점검표의 타당성 평가이론은 2절에서 제안되며, 3절에서는 제시된 평가이론에 입각하여 점검표의 각 단계별 평가방법을 제시하여 조사연구와 실험연구를 겸한 한국의류학회지에 게재된 논문을 대상으로 타당성 평가기준을 설정한다. 4절에서는 점검표 사용목적에 따른 타당성 평가방법을 설정한다.

2 타당성 평가이론

평가(evaluation)란 사물 또는 그 속성에 대한 가치판단을 의미하며, 어떤 사물 또는 그 속성에 대한 가치판단을 내리자면 반드시 판단의 준거(criteria) 또는 표준(standard)이 있어야 한다. 그러나 평가에의 접근을 대변해 주는 여러 평가 모형(model)마다 평가는 다르게 정의하기 때문에 어떤 접근과 모형을 선택하느냐에 따라 평가의 개념은 다를 수밖에 없다.

평가모형에는 목표달성모형, 의사결정모형, 판단중심모형, 자연주의모형, 법정재판모형 등이 있다. 판단중심모형(judgment oriented evaluation model)은 평가란 본질적으로 전문가에 의해 전문적인 지식과 기술을 바탕으로 하여 평가하고자 하는 대상 또는 내용의 가

표 2: 통계적 기법(5단계)

통계적 기법	정확하게 수행됨	잘못수행 됨	해당사항 없음
1. 통계적 가설 2. 통계적 기법 3. 사용된 통계적 기법의 적절성 4. 특수한 분석법을 사용한 경우의 그 분석법에 대한 설명 5. 연구목적과 통계적 결론의 일치 여부 6. 결론을 정당화하기 위한 모든 통계량의 제시			

치를 체계적으로 판단하는 활동이라고 보는 입장으로 전문가의 지식과 기술 그리고 자질과 경험을 중시하고 판단을 내리기 위해 어떤 표준을 사용할 것인가에 많은 관심을 기울인다. 평가를 어떻게 정의하든 가치판단이 개입 되기 때문에 판단중심모형으로 분류되지 않을 평가모형이 없다고 할 정도로 판단중심 평가모형은 많다. 판단중심모형에는 Flexner의 평가인정모형, Stake의 종합실상모형, Scriven의 탈목표모형, Eisner의 감정 비평모형 등이 있다. 평가인정모형(accreditation evaluation model)은 발명가가 출원한 발명품을 특허청에서 심사, 인정하거나 공업진흥청에서 제품이 규격과 성능을 심사하여 KS마크를 부여하듯이 전문가들이 어떤 제품이나 기술, 조직체의 구성과 시설, 프로그램 등이 미리 정해 놓은 표준 혹은 요건을 충족시키고 있는가를 검토하여 인정여부를 결정하는 것을 말한다. 그러므로 평가인정모형에서 가장 중요한 부분은 표준의 설정과 그것을 기준으로 하는 판단이다. 스테이크(Stake, 1967)는 그의 논문, 교육평가 의 실상에서 교사들이 자신의 수업방식을 정확히 이해하고 수업이론의 발전에 기여하려고 하면, 전체적 평가국면을 정확하게 심리 검토해야 한다고 주장하면서 제안한 교육평가의 틀을 종합실상모형(countenance model)이라 한다.

평가준거를 기준으로 분류한 평가유형은 두 가지가 있으며, 어떠한 성과를 평가하는 준거를 주어진 목표(표준)로 하느냐 또는 평가를 실시해서 받은 점수의 평균으로 하느냐에 따라 평가는 준거기준 평가와 규준기준평가로 나누어진다. 규준기준평가와 준거기준평가는 사용하는 준거를 중심으로 구분 한 평가유형으로 어느 평가에 입각하여 평가하느냐에 따라 평가의 내용선정과 평가 절차에 뚜렷한 차이가 있다. 규준기준평가에서의 규준은 어떤 평가에서 받은 점수의 평균점을 기준으로 개개인의 성취도가 평균이상 어느 정도 또는 평균이하 어느 정도 인지를 평가할 때의 평균점을 평가의 규준이라 하고 사전에 할당해 놓은 평점 배당비율을 기준으로 하여 성취도를 서로 비교, 평가하는 방법이다. 준거기준평가라고 할 때의 준거는 주어진 목표를 의미하며 목표를 평가의 준거로 삼아 목표의 도달여부

와 그 정도를 확인, 점검하는 평가이다. 준거기준평가는 획득한 점수를 정규분포에 입각하여 평균이나 미리 정해 놓은 비율을 근거로 해서 상대적으로 비교하지 않고 오직 주어진 목표를 기준으로 알아야 할 것을 알고 있느냐(목표도달), 모르고 있느냐(목표미달)를 판정한다고 해서 절대평가라고 부르기도 한다.

본 논문에서의 타당성 평가는 판단중심모형에서 Flexner의 평가인정모형에 입각하여 연구설계 단계와 통계적 기법 단계에서 갖추어야 할 목표(표준)를 표1의 점검항목으로 설정한 후 그 항목을 평가의 준거로 삼아 목표도달 여부를 결정하는 준거기준평가 방식을 취한다. 연구자와 연구 논문의 타당성 평가기준의 설정은 준거기준평가의 타당성 평가 평균 점수와 표준편차를 이용한 규준기준평가 방식을 취한다. 따라서, 점검표 설계상의 각 단계별 타당성 평가점수를 이용하여 연구자와 연구 논문의 타당성 정도를 상, 중, 하로 평가한다.

3. 단계별 평가방법

연구자의 통계학적 지식 정도에 대한 평가기준의 설정은 모 대학교 대학원생을 대상으로 설문조사 한 결과를 기초로 하였다. 점검표의 각 단계에 대한 평가기준 설정을 위하여 국내 과학학술지중에서 자연과학 분야이면서 동시에 사회과학 분야의 성격을 가지며, 조사연구와 실험연구를 적절히 겸한 한국의류학회지를 선정하였다. 타당성 평가는 1997년 한국의류학회지에 게재된 논문중 조사연구방법을 사용한 12편(33편중 36%)의 논문과 실험연구방법을 사용한 12편(21편중 57%)의 논문을 임의 선정하여 제안된 점검표를 이용하여 실시하였다. 그 타당성 평가 결과를 기초로 하여 타당성의 평가기준을 설정하였다.

3.1 연구자의 통계학적 지식 단계(3단계)

(1) 문항별 배점

연구자의 통계학에 대한 지식 정도를 파악하여 통계상담 필요 유무를 결정하기 위한 것으로 조사항목은 표3과 같다. 이것은 Efron이 제시한 통계학 지식의 기본적 3대 방향인 수리적 측면, 데이터 분석, computing 측면을 골고루 측정하기 위한 것이다. 조사항목은 통계학의 지식 수준, 통계패키지 수강, 연구 논문의 통계적 방법 적용 경험, 통계패키지 사용정도, 결과에 대한 이해 정도로 다섯 개 항목으로 구성하였다. 각 항목당 점수는 20점씩 부여되어 총 100점 만점으로 계산된다. 문항 1, 2, 3의 질문에 대한 응답형태는 5점 척도를 사용하며, 문항 답지별 부분점수는 각 척도의 순서에 따라 0, 5, 10, 15, 20점의 가중치를 부여한다. 그러나, 문항 4와 5는 사용 가능한 통계적 기법을 10가지 제시한 후 모두 선택하게 하는 다답형 문항으로 답지별 2점의 가중치를 주게 된다.

(2) 조사결과의 점수 분포

연구자의 통계학에 대한 지식 정도를 알아보기 위하여 모 대학교 대학원을 자연계 열과 사회계열로 층화한 후 랜덤하게 학과를 선정하여 해당학과 전체 대학원생을 대상으로 면

표 3: 통계학적 지식 점수 도수분포표

조사항목	배점	통계영역
1. 통계학 과목의 수강 정도		수리
수강 경험 없음	0	데이터분석
1강좌 수강	5	
2강좌 수강	10	
3강좌 수강	15	
4강좌 이상 수강	20	
2. 통계패키지 강좌 수강		수리
전혀 없음	0	Computing
1회 수강	5	
2회 수강	10	
3회 수강	15	
4회 이상 수강	20	
3. 연구논문의 통계적 방법 적용 경험		수리
전혀 없음	0	데이터분석
1 2편	5	
3편	10	
4편	15	
5편 이상	20	
4. 통계패키지 사용 정도		Computing
사용경험 없음		
사용할 수 있는 통계기법의 수는?()	각 2점	
기초통계() t검정() 2검정()	(20점)	
분산분석() 회귀분석() 인자분석()		
다변량 분산분석() 군집분석() 판별분석()		
비모수 검정()		
5. 통계패키지에 의해 출력된 Output에 대한 이해		수리
이해할 수 없음	0	데이터분석
Output를 이해할 수 있는 통계기법의 수는?()	각 2점	
기초통계() t검정() 2검정()	(20점)	
분산분석() 회귀분석() 인자분석()		
다변량 분산분석() 군집분석() 판별분석()		
비모수 검정()		

표 4: 계열별 t검정 결과

계열	인원	평균	표준편차	T	Prob> T
사회계열	28	15	16	-2.87	.0059
자연계열	26	25	22		
계열전체	54	18	20		
등분산성 검토: H0: Variances are equal F=1.89 Prob>F=0.1099					

표 5: 통계학적 지식 단계의 평가기준

통계상담유무	통계상담 필요치 않음		통계상담 권함		통계상담필요함
	통계상담 필요치 않음		통계상담 권함		
평가기준	58점	100점 ($\bar{X} + 2\sigma$)	38점	57점 ($\bar{X} + 1\sigma$)	38점미만

접조사를 실시하였다. 70부를 배부한 결과 54부가 회수되어 77%의 회수율을 보였다. 연구자의 통계학적 지식에 관한 계열별 t검정 결과는 표4와 같다. 계열별 t검정 결과는 표4에서와 같이 등분산성을 만족한다(F=1.89 Prob>F=0.1099). "계열별 평균 차가 없다"는 귀무가설이 유의수준 5%에서 기각되어 계열별 차이가 있음을 알 수 있다(t값=-2.87, Prob> |T|=0.0059).

(3) 연구자의 통계학적 지식 단계의 평가기준

연구자의 통계학적 지식 정도에 대한 설문조사 결과를 토대로 통계상담 필요 유무의 평가기준은 평균 (\bar{X})과 표준편차 (σ)를 이용하여 표5와 같이 38점 ($\bar{X} + 1\sigma$) 으로 한다. 이와 같은 기준은 설문조사 결과에서 계열전체 평균이 18이고 표준편차는 20이므로 38점을 평가기준으로 하며, 38점 이상은 15%이므로 점검표 사용 목적이 통계적 오류를 줄이기 위한 것임에 비추어 볼 때 오류를 범할 위험을 적게 줄 수 있는 기준으로 적합하다고 생각되며, 이 기준은 앞으로 통계상담 DB가 구축되면 점차 수정해 나아가야 할 것이다. 38점 미만은 연구설계에서부터 모든 통계적 절차를 통계전문가에게 의뢰하게 하고, 38점 이상 즉 통계상담이 필요없는 연구자라도 58점 미만은 통계상담을 권한다. 58점 ($\bar{X} + 2\sigma$) 이상은 통계상담을 할 필요 없이 그 다음 단계의 통계적인 절차를 따르게 하지만, 이러한 연구자라도 통계전문가의 조언을 얻게 하여 통계적 방법의 이용에 따른 통계적 오류를 최대한 줄일 수 있도록 하는 것이 바람직하다.

3.2 연구설계 단계(4단계)

(1) 평가방법

점검항목의 배점 부여방법에는 점검항목의 통계학적 기여도에 따라 가중치를 준다든지

표 6: 연구설계 단계의 타당성 평가기준

평가기준		연구설계의 타당성	평가결과
조사설계	실험설계		
80점이상	96점이상	상 ($\bar{x} + 2\sigma$)	<ul style="list-style-type: none"> • 현 단계에서 연구논문이 적합 • 통계상담 필요치 않음 • 오류 재검토 요함
52점 이상 79점 이하	74점 이상 95점 이하	중 ($\bar{X} + 1\sigma$)	<ul style="list-style-type: none"> • 현 단계에서 연구논문이 적합 • 통계상담을 권함 • 오류 재검토 요함
52점 미만	74점 미만	하	<ul style="list-style-type: none"> • 현 단계에서 연구논문이 부적합 • 통계상담 반드시 필요함 • 의뢰하여 재작업

혹은 모든 항목에 대하여 동일한 배점을 주는 방법 등이 있다. 그러나, 이러한 방법들중 어느 방법을 사용하더라도 표본의 크기가 커질 경우 모든 결과가 같게 되므로, 각 항목당 배점을 동일하게 주게 된다. 연구설계 단계와 통계적 기법 단계에서 각 점검항목에 대한 목표달성 여부를 조사하여, 통계적 타당성 평가점수 계산은 교육평가 이론에서 문항의 어려운 정도를 나타내는 문항 곤란도를 추정하는 식 (1)을 이용한다.

$$P = \frac{R}{N} \times 100 \quad (1)$$

P = 문항 곤란도 지수, R = 문항에 대한 정답을 한 학생수, N = 총 사례수

따라서, 타당성 평가점수 계산을 위한 각 점검항목의 답지별 배점은 '정확하게 수행됨'은 1점, '잘못 수행됨'은 0점으로 처리하여, 전체 점검 항목수에 대한 '정확하게 수행됨' 항목수에서 '해당사항 없음' 항목 수를 뺀 값의 비를 백분율로 표시한다. 계산방법은 식 (2)과 같다.

$$\text{평가점수} = \frac{\text{'정확하게 수행됨' 항목수} - \text{'해당사항 없음' 항목수}}{\text{전체 점검항목수}} \times 100 \quad (2)$$

만약, 평가점수가 80이라면 점검하는 각 단계에서 점검 항목중 80% 정도는 목표를 달성했다고 볼 수 있다.

(2) 연구설계 단계의 타당성 평가기준

연구설계 단계의 타당성 평가의 기준은 표6과 같다. 이와 같은 평가기준의 설정은 1997년 한국의류학회지에 게재된 조사연구방법 논문 12편을 대상으로 조사한 결과를 기초로 하여 평균과 표준편차를 이용한다.

표 7: 통계적 기법 단계의 타당성 평가기준

평가기준	통계적기법의 타당성	평가결과
77점이상	상 ($\bar{x} + 2\sigma$)	<ul style="list-style-type: none"> • 현 단계에서 연구논문이 적합 • 통계상담 필요치 않음 • 오류 재검토 요함
41점 이상 76점 이하	중 ($\bar{X} + 1\sigma$)	<ul style="list-style-type: none"> • 현 단계에서 연구논문이 적합 • 통계상담을 권함 • 오류 재검토 요함
41점 미만	하	<ul style="list-style-type: none"> • 현 단계에서 연구논문이 부적합 • 통계상담 반드시 필요함 • 의뢰하여 재작업

3.3 통계적 기법 단계(5단계)

(1) 평가방법

통계적 기법 단계에서 점검항목의 답지별 배점 부여방법은 연구설계에서의 평가점수 계산과 동일하다. 즉, 점검항목의 답지별 배점은 '정확하게 수행됨'은 1점, '잘못 수행됨'은 0점으로 처리하여 전체 점검항목수에 대한 '정확하게 수행됨' 항목수에서 '해당사항 없음' 항목수를 뺀 값의 비를 백분율로 표시한다. 계산방법은 식 (2)과 동일하다.

(2) 통계적 기법 단계의 타당성 평가기준

통계적 기법 단계에 대한 타당성 평가는 연구방법에 따라 점수차는 나지만 같은 통계적 기법을 사용하므로 연구방법에 따라 평가방법을 분리하지 않고 하나의 평가방법을 이용하여 타당성 정도를 결정한다. 통계적 기법에 대한 타당성 평가 기준은 평균과 표준편차를 이용하여 표7과 같이 41점 \bar{X} 으로 한다.

4. 점검표 사용목적에 따른 평가방법(6단계)

점검표에서의 평가방법은 점검표의 사용목적에 따라 두 가지로 나누어 설정한다. 첫째, 통계적 기법 적용 단계에서의 통계적 오류를 줄이기 위하여 연구자의 통계학에 대한 지식 정도와 연구설계의 목표달성 정도에 따라 통계상담 필요 유무를 제시하기 위한 평가기준을 설정한다. 둘째, 통계적 타당성 평가를 목적으로 한 경우 연구자의 통계학에 대한 지식, 연구설계 단계 그리고 통계적 기법 적용의 타당성을 근거로 타당성 평가기준을 설정하여야 한다.

표 8: 통계상담 필요 유무 평가기준

연구방법	통계상담 필요치 않음		통계상담 필요함 (재 설계)
	통계상담 필요치 않음 (오류수정)	통계상담 권함 (오류수정)	
조사연구	69점이상	45점-68점	45점미만
실험연구	77점이상	45점-76점	45점미만

4.1 통계상담 필요 유무

(1) 평가방법

연구자의 통계학적 지식 정도와 연구설계 단계의 점검항목 달성 정도에 따라 그 다음 단계인 통계적 기법 적용에서의 통계상담 필요 유무를 제시한다. 통계상담 필요 유무의 평가기준은 식 (3)과 같이 연구자의 통계학적 지식 점수와 연구설계의 타당성 평가점수를 합한 평균값으로 한다.

$$\text{평가점수} = \frac{\text{통계학적 지식 점수} + \text{연구설계의 타당성 점수}}{2} \quad (3)$$

따라서 통계상담 필요 유무의 결정은 식 (3)에 의해 계산된 점수를 이용하여 연구방법에 따른 평가기준을 제시한다.

(2) 통계상담 필요 유무 평가기준

통계상담 필요 유무의 평가기준은 표8과 같다. 이와 같은 평가기준은 연구자의 통계학적 지식 단계의 평가기준인 표5와 표6의 조사설계 단계의 평가기준 점수를 식 (3)으로 계산하여 설정한 것이다. 조사연구에서 통계상담 필요 유무의 평가기준은 45점으로 한다. 45점 미만은 통계전문가에게 의뢰하여 연구설계에서 발견된 오류를 검토하여 재 설계하도록 하고, 45점 이상 즉 통계상담이 필요 없는 연구자라도 69점 미만은 오류를 수정 하고 통계상담을 권하여 다음 단계인 통계적 방법 적용에서의 통계적 오류를 최대한 줄일 수 있게 한다. 69점 이상은 통계상담이 필요 없지만 발견된 몇 가지 통계적 오류를 검토 수정하도록 한다. 또한 각 단계에서 연구자의 통계학적 지식 정도의 점수, 연구설계의 타당도 점수 그리고, 통계상담 여부를 결정하는 평균 평가점수를 제시하며, 발견된 오류와 종합적인 평가 결과의 해설도 제시한다.

실험연구의 통계상담 필요 유무 평가기준은 연구자의 통계학적 지식 단계의 평가기준인 표 3과 조사설계 단계의 타당성 평가기준인 표 6의 점수를 식 (3)으로 계산하여 실험연구에서 통계상담 유무의 평가기준 점수를 45점으로 한다. 따라서, 45점 미만은 통계상담을 반드시 하게 하고, 45점에서 76점까지는 통계상담을 권하고, 77점 이상은 통계상담은 필요하지 않으나 발견된 오류를 수정하게 한다.

표 9: 연구논문의 종합적인 타당성 평가기준

평가기준		논문의 타당성	평가결과
조사설계	실험설계		
74점이상	81점이상	상	<ul style="list-style-type: none"> • 연구논문이 적합 • 오류 수정 요함
45점이상 73점이하	54점이상 80점이하	중	<ul style="list-style-type: none"> • 연구논문이 적합 • 통계상담을 권함 • 오류 수정 요함
45점미만	54점미만	하	<ul style="list-style-type: none"> • 연구논문이 부적합 • 통계상담을 의뢰하여 작업

4.2 연구논문의 종합적인 타당성 평가

(1) 평가방법

점검표 사용목적이 연구논문의 타당성 평가일 경우는 연구자의 통계학적 지식 정도, 연구설계 단계의 점검항목 달성정도 그리고, 통계적 기법 단계의 점검항목 달성 정도에 따라 연구 논문에 대한 종합적인 타당성 평가를 내린다. 연구 논문의 타당성 평가에서 연구자의 통계학적 지식 점수는 원 점수 그대로 활용하고, 연구설계와 통계적 방법에서의 타당성 평가점수의 비중은 원 점수의 2배로 하여 세 값을 합하여 평균값을 계산한다. 계산 방법은 식 (4)과 같다.

$$\text{점수} = \frac{\text{통계학적 지식정도} + \text{연구설계 점수} \times 2 + \text{통계기법 점수} \times 2}{2} \quad (4)$$

이와 같은 점수 계산은 점검표 사용목적이 연구논문의 종합적인 타당성 평가이기 때문에 이미 연구설계와 통계적 기법 적용이 끝난 경우에 있어서는 통계학적 지식 수준보다는 연구설계에서의 타당성 정도와 통계적 방법 적용에서의 타당 정도가 종합적인 타당성 평가에 더 큰 영향을 미치게 되므로 연구자의 통계학적 지식의 비중은 1로 하고 연구설계와 통계적 방법의 타당성 비중은 2로 하는 것이 합당하다. 연구논문의 종합적인 타당성 평가기준은 연구방법에 따라 다르게 설정한다.

(2) 연구논문의 종합적인 타당성 평가기준

연구논문의 종합적인 타당성 평가기준은 연구자의 통계학적 지식 단계의 평가 기준인 표5, 조사설계단계의 타당성 평가기준인 표6, 통계적 기법 단계에서의 타당성 평가기준인 표7를 기초로 하여 식 (4)에 의해 계산된 점수는 표9와 같다.

조사연구에 대한 통계적 타당성 평가기준을 45점으로 하고 45점 미만의 연구 논문을 부적합하다고 판정하고 통계전문가에게 의뢰한 후 조사설계와 통계적 기법 적용을 재 작업

하도록 한다. 45점 이상 연구 논문이 적합하다고 판정하지만, 74점 미만은 발견된 통계적 오류를 재검토하여 수정하게 하며 통계상담을 의뢰하여 통계전문가의 조언을 듣게 한다. 74점 이상은 발견된 몇 가지 통계적 오류를 검토 수정하도록 한다. 실험연구의 통계적 타당성 평가는 기준은 54점 미만의 연구 논문은 부적합 판정하고 통계전문가에게 의뢰한 후 실험설계와 통계적 기법 적용을 재 작업하도록 한다. 54점 이상의 연구논문은 적합하다고 판정하여, 54점에서 80점까지는 통계상담을 권하고, 81점 이상은 통계상담이 필요 없으므로 오류만 수정하게 한다.

5. 결론

본 논문은 연구방법과 점검표 사용목적에 적합한 점검표의 타당성 평가방법을 제안하였다. 제안된 타당성 평가 방법을 이용한 통계상담서버는 구현중에 있으며, 다음과 같은 몇 가지 문제점들이 앞으로 해결해야 할 과제로 남아 있다.

첫째, 학문 응용기술 분야별로 사용되는 통계적 방법들이 다양하므로 각 전문분야에서 요구되는 통계적 표준들을 설정하여 자체 전문 통계 점검표들을 연구 개발하여야 할 것이다.

둘째, 본 논문에서 제시된 연구자 평가기준은 한국의류학회지를 중심으로 설정하였다. 따라서, 통계상담 DB 구축을 통하여 타당성 평가방법과 평가기준의 설정을 보다 체계화하여 통계학적 지식 정도와 연구설계의 타당성 정도에 따라 연구자들을 평가하여 연구자의 수준에 적합한 구체적이고 일반화된 통계적 방법들의 제시가 요구된다.

참고문헌

1. Ford, B. L., and Tortora, R. D.(1978). The Consultant's Forum, *Biometrics*, 34, 299-304.
2. Fred N. Kerlinger (1973). *Foundations of Behavioral Research*, 2nd Ed. New York:Holt, Rinehart and Winston, 17-18.
3. Gorsuch, R. L.(1983). *Factor analysis*, 2nd Ed., Hillsdale, NJ:Lawrence Erlbaum.
4. Gosling, J. and Arnold, K.(1995). *The Java Programming Language*, Addison Wesley.
5. Kerlinger, 前掲書, 20-23.
6. Rossini, A. J.(1994). Teaching and Computing via Multimedia through the World Wide Web, *Statistical Computing and Graphical Newsletter*, 5(3).
7. Schor, S., and Karten, I.(1966). Statistical Evaluation of Medical Journal Manuscripts, *Journal of the American Medical Association*, 195(13), 1123-1128.

8. Tabachnick. B. G. and Fidell. L. S.(1989). *Using Multivariate Statistics*, 2nd Ed. New York: Harper and Row, Publishers, Inc.
9. Townsend, 前掲書, 45-46.
10. White, S. J.(1979). Statistical Errors in Papers in the British Journal of Psychiatry, *British Journal of Psychiatry*, 135, 336-342.
11. 김경동, 이은죽(1991). 사회조사연구방법, 박영사.
12. 변창진, 최진승, 문수백, 김진규, 권대훈(1997). 교육평가, 학지사.
13. 소영일(1994). 연구조사방법론, 박영사.
14. 이재창(1997). 국제화 정보화 시대의 통계학의 위상, 대한통계학회 대구경북지회 춘계학술논문발표회 초청강연.
15. 윤기중, 안윤기, 김병수(1987). 통계의 오용과 효율적 이용에 관한 연구, 산업과 경영, 24-2, 3-37.
16. 정인성의 2인(1996). 초고속정보통신망 시범 사업 관련 원격교육시범시스템의 교육적 활용방안 탐색, 정보과학지, 13(6), 23-42.
17. 차용근(1994). 사회과학연구방법, 세영사.
18. 최종후(1989). 국내 과학학술지에 나타난 통계적 기법 활용의 타당성 평가에 대한 연구, 고려대학교 박사학위 청구 논문.
19. 최종후, 이광진, 김현의(1995). 통계자료분석의 이해, 자유아카데미.
20. (1986, 1991, 1996, 1997). 한국의류학회, 한국의류학회지.

The method of evaluating the validity in making the checklist of statistical method

Sangbock Lee³ · Malsook Kim⁴

Abstract

This thesis is to propose the method of evaluating the validity in stages on the design of a checklist based on the estimate theory of the validity on the design of a checklist and the method of evaluating the validity accordy to the purpose for using checklist and establish the standards of estimating the validity in stages on the checklist owing to the estimate theory with the journal, Journal of the Korean Society of Clothing and Textiles .

We are embodying the statistical consulting server using the proposed estimate method of the validity and the followings are the questions to solve in future. First, there are many kinds of the statistical methods used in the fields of learing and applied technology. So we must establish the statistical standars required in each field and study and develop the special statistical checklists. Second, it is neded to systemize the estimate methods and standard on the validity by building the statistical consult DB, evaluate the researchers according to the degrees of statistical knowledge the validity of study design and propose the delicate and generalized statistical methods appropriate to the ability of researchers.

³Associate Professor, Department of Statistical information, Catholic University of Taegu-Hyosung, 712-702

⁴Computer education center, Catholic University of Taegu-Hyosung, 712-702