

조절회귀분석의 통계적 검정력에 대한 연구

한 인 수*

〈 목 차 〉

- | | |
|-------------------------------------|------------------------|
| I. 서론 | IV. 통계적 검정력의 향상을 위한 제안 |
| II. 조절효과 검증방법의 하나로서의
조절회귀분석(MMR) | V. 결 론 |
| III. MMR법과 통계적 검정력의 문제 | |

I. 서론

최근에 조직이나 노사관계의 연구에서 조절변수를 상정한 연구모형이 늘고 있고 또 조절효과를 검증하는 연구들이 증가하고 있다. 그것은 사회현상이 기존에 상정한 것처럼 독립변수와 종속변수로만은 설명하기가 어려울 정도로 복잡해지고 있기 때문일 것이다. 특히 경영연구에서는 단일의 최선의 보편적 원리보다는 '상황에 좌우된다'(It depends on)는 입장이 지지를 획득하고 있는데, 이는 결국 제 3의 변수를 가정하고 있다는 것을 의미하는 것이고 따라서 이를 검증해야 할 필요도 늘고 있는 것이다.

이러한 조절변수들의 조절효과를 통계적으로 검증하는데는 여러 가지 방법이 쓰이고 있지만 그 중에서도 가장 널리 쓰이는 것은 단연 조절회귀분석(moderated multiple regression: MMR)의 방법이다. 이는 통계적 검정력이나 또는 이용자의 편리면에서 다른 조절회귀방법들 보다 압도적으로 많이 활용되고 있다.

그러나 이 방법의 사용이 점증함에도 불구하고 그 유효성에 대해서는 회의도 늘고 있다. 그것은 연구자들이 매우 튼튼한 이론적 기초를 바탕으로 조절변수를 상정했음에도 불구하고 좀처럼 조절효과가 실증적으로 입증되지 못하는 실망적인 연구결과들이 너무 많이 발견되고 있기 때문이다. 이에 일각에서는 조절효과가 검증되지 못하는 이유가 기

* 충남대학교 경영학과 교수

본적으로 조절효과가 존재하지 않기 때문이 아니라 우리들이 널리 쓰고 있는 조절회귀분석의 방법 자체에 또는 사용과정에서 파워(power) 즉 통계적 검정력의 문제에 기인하는 것인 아닌가하는 의문이 제기되었고 또 이를 뒷받침하는 실증연구결과들이 많이 제시되게 되었다.

따라서 본고는 최근 사용이 빈번히 늘고 있는 조절회귀분석의 통계적 검정력의 문제를 살펴보고 동시에 이 통계적 검정력을 높일 수 있는 방법을 제시해보고자 한다. 그럼으로써 조직과 노사관계의 많은 연구들에서 명백히 조절효과가 있음에도 불구하고 우리가 사용하는 방법론의 문제 때문에 이를 구명하지 못하는 잘못을 범하지 않는데 일조를 하고자 한다.

II. 조절효과 검증방법의 하나로서의 조절회귀분석(MMR)

조절변수의 검증방법으로는 하위집단별 상관계수 비교법(subgroup-based correlation coefficients: SCC) 등을 비롯해서 여러 가지 방법이 있지만 가장 널리 쓰이는 것은 조절회귀분석(moderated multiple regression: MMR)이다. 이는 한 변수 X 가 제3변수 Y 를 예측하는데 있어 조절변수 Z 와 통계적으로 의미있는 상호작용을 하는지를 통계적으로 검증하는 것이다. 보다 구체적으로는 Y 와 X 간의 관계가 Z 에 의해서 조절된다는 다음과 모델을 설정하고 모델에서 $\beta_3=0$ 이라는 귀무가설을 기각하면 조절효과가 존재하는 것으로 추론하는 방법이다.

구체적으로는 우선 종속변수 Y 에 대해 X 와 Z 가 예측변수 역할을 하는 방정식(1)을 설정한다.

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + e \quad (1)$$

다음에 회귀식에 예측변수의 곱셈항($X*Z$)을 넣은 다음의 방정식을 설정한다.

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X*Z + \epsilon \quad (2)$$

조절효과의 통계적 유의성 검증을 위해서는 방정식(1)과 방정식(2)의 결정계수를 비교한다. 즉 R_1^2 과 R_2^2 와의 차이 즉 ΔR^2 의 유의성 여부에 의해 조절효과의 여부를 판단하게 된다. 이 때 사용되는 F통계량은 다음의 공식에 의해 계산된다.

$$F = \{[(R_2^2 - R_1^2) / (k_2 - k_1)] / [(1 - R_2^2) / (N - k_2 - 1)]\} \quad (3)$$

이러한 MMR법은 기존의 통계 패키지들, 즉 BMDP, SAS, SPSS들을 이용한 계층적 회귀분석(hierarchical multiple regression)을 통해 쉽게 시행될 수 있다. 우선 Y 를 예측하는데 있어 X 와 Z 를 투입하여 회귀분석을 실시한다(방정식 1). 다음 단계로서 $X*Z$ 항을 투입하여 회귀분석을 실시한다(방정식 2). 모든 통계 패키지들은 두 모델 모두의 R^2 를 제시해 주며 두 R^2 의 차이에 기초한 F통계량도 제시해 주는데 이 수치의 유의성은 $X*Z$ 상호작용의 존재 즉 조절효과의 존재를 표시해 주는 것이다.

조절효과를 분석한 실제 예를 제시하면 다음과 같다. <표 1>은 종업원의 조직몰입이 조직시민행동에 영향을 미침에 있어 상사에 대한 종업원의 신뢰가 조절변수 역할을 하는지 여부를 조절회귀분석을 이용해 분석한 결과이다(이규범, 1997).

<표 1> 조절회귀분석에 의한 조절효과 검증의 예시

모델	사용변수	R^2	ΔR^2	ΔF 및 유의성	Beta	t값 및 유의성
방정식(5)	조직몰입	.086	-	18.989**	.292	4.358**
방정식(6)	조직몰입	.168	.083	20.154**	.194	2.864**
	상사에 대한 신뢰				.304	4.489**
방정식(7)	조직몰입	.175	.006	1.500	.615	1.756
	상사에 대한 신뢰				.598	2.399*
	조직몰입×상사에 대한 신뢰				-.591	-1.225

결과변수: 조직시민행동, 예측변수: 조직몰입, 조절변수: 상사에 대한 신뢰

** : $p < .01$ * : $p < .05$

분석결과 우선 방정식 (5)에서 조직몰입은 조직시민행동을 예측해 주는 예측변수로 기 능함을 알 수 있다. 그리고 상사에 대한 신뢰를 또 하나의 예측변수로 도입한 방정식 (6)

의 분석결과 상사에 대한 신뢰 역시 종업원의 조직시민행동을 설명해 주는 또 하나의 예측변수임을 볼 수 있다. 마지막으로 방정식 (7)의 분석결과 이 두 예측변수의 곱으로 생성된 새로운 변수는 유의한 예측변수로 나타나지 않음을 볼 수 있다. 단계적 회귀분석을 통해 도출된 R^2 의 증분을 검토해 보면 .006으로 유의성이 없기 때문에 조직몰입×상사에 대한 신뢰는 회귀식에 영향을 미치지 못한다는 결론을 내릴 수 있다. 즉, 상사에 대한 신뢰는 조절변수로 작용한다기보다는 또 하나의 예측변수로 기능한다고 볼 수 있는 것이다.

이러한 조절효과 검증방법으로서의 MMR법은 조절효과 검증을 위해 가장 널리 사용되는 방법으로 자리 잡고 있다. 조절회귀분석이 다른 종류의 조절효과 검증방법들, 특히 널리 사용되는 하위집단 상관계수 비교법(SCC법) 보다 선호되는 이유는 다음과 같은 이유들 때문이다.

우선 첫째로 하위집단 비교법은 조절변수를 둘 또는 세 개의 하위집단으로 나누는데 이렇게 구분변수의 측정보다는 조절회귀분석에서처럼 연속변수로의 측정쪽이 훨씬 많은 정보를 활용할 수 있다는 점이다. 즉, 2점 척도보다는 5점 내지 7점 척도의 측정치가 훨씬 유용한 정보를 기초로 분석이 가능하다는 것이다. 이러한 입장에서 등간척도 이상으로 측정하여 MMR법을 사용해야하고 또 사용할 수 있음에도 불구하고 SCC법을 사용하는 경향에 대해서는 당연히 비판이 제기되고 있다(Stone-Romero & Anderson, 1994). 즉 조절변수를 등간척도로 측정해놓고 SCC법에 의한 조절효과의 검증을 위해 조절변수별로 2 혹은 3, 또는 그 이상의 집단으로 분할하는 것은 잘못된 것이라는 것이다(Erwin & McClelland, 2001). 예컨대 조절효과의 검증을 위해서 평균이나 중위수를 기준으로 두 집단으로 구분하는 경우가 많았다. 이러한 방법을 채택하는 이유는 보다 조절효과를 분명히 드러내리라는 추측에 근거한 것이다.

그러나 이러한 방법은 결코 조절효과를 포착하는데 유효한 방법이 못된다. 연속변수로 측정을 한 조절효과의 검증에서도 낮은 파워(검정력)가 문제가 되는데, 하물며 중위수 분할방법에 의하여 표본을 쪼개면 효과적인 샘플의 크기가 반으로 줄어들어 더욱 파워를 약화시키게 만드는 결과를 초래하는 것이다. 뿐만 아니라 자료들을 몇 개의 하위집단으로 분류하는 방식은 있지도 않은 상호작용효과를 거짓으로 만들어 낼 수 있는 가능성도 있다(Maxwell & Delany, 1993). 따라서 조절변수를 기초로 집단을 분할하는 분석방식은 1종 오류와 2종 오류를 범할 가능성을 증대시키므로 바람직하지 못하다는 것이다.

또 하나 조절회귀분석이 여타의 조절효과 분석기법들 특히 하위집단별로 상관계수의 동질성을 비교하는 방법보다 더 선호되는 가장 큰 이유는 조절회귀분석의 결과가 다른 방법이 주지 못하는 유용한 정보를 제공해주고 있기 때문이다(Aguinis & Stone-Romero, 1997). 즉, MMR법은 각 하위집단들에 대해 기울기의 차이와 관련된 정보를 제공해 준다. 그리고 이 정보는 각 집단별로 차별적인 예측을 하는데 매우 귀중한 역할을 하게 되는 것이다.

이상과 같은 이유로 MMR법은 조절효과 분석의 가장 강력한 기법으로 널리 사용되게 되었다. 이는 조절변수가 연속변수(예: 나이등)나 구분변수(성별)에 불문하고 적용되고 있다. 구분변수의 경우에는 물론 더미변수를 이용한 분석이 채용되게 된다. 한 연구 [Cortina, 1993]에 따르면 *Journal of Applied Psychology* 1991년부터 1992년간에 수록된 논문들 중에서 조절효과의 발견을 위해 이 MMR법이 123번이나 사용될 정도로 보편화되고 있다.

III. MMR법과 통계적 검정력의 문제

조직연구에서 조절효과의 분석기법으로 MMR법의 사용이 증가하는 과정에서 이를 사용하는 학자들간에 상당한 어려움이 생기게 되었다. 그것은 좀처럼 조절효과가 발견되지 않았기 때문이다. 학자들이 튼튼한 이론적 기초위에서 그들의 모델을 수립하였음에도 불구하고 조절효과의 입증은 어려웠다.

실제연구에서는 일반적으로 조절효과가 없다는 것이 귀무가설이 되고 이를 기각함으로써 조절효과를 입증하게 되는 과정을 밟게 된다. 그런데 조절회귀분석의 과정에서 귀무가설을 기각하지 못하고 귀무가설이 그대로 채택되는 경우가 많았던 것이다. 즉 조절효과가 없다는 검증들이 이루어지게 된 것이다.

이러한 연구결과는 이론적으로 조절효과가 존재하지 않는다면 별 문제가 없는 것이었지만, 이론적 기초로 볼 때 학자들에게는 수궁하기 힘들었기 때문에 그들은 이 기법이 지닌 통계적 검정력, 즉 파워의 문제에 관심을 갖게 만들었다. 즉, 이 기법이 근본적으로 또는 그 사용과정에서 잘못된 귀무가설(이 경우 조절효과가 없다는 것)을 받아들이는 경향(type II error)이 큰 것은 아닌가, 다시 말해 잘못된 귀무가설을 기각하는 힘, 즉

파워(power)가 약한 것은 아닌가 하는 문제의식을 불러 일으키게 된 것이다.

그래서 Cronbach(1987)가 MMR법의 상용과 관련된 통계적 검정력(파워) 연구의 필요성을 제기한 이래 MMR법의 사용과 관련된 파워의 문제에 대해서도 많은 연구가 이루어졌다. 구체적으로는 조절효과가 있음에도 불구하고 없다는 결론을 나오게 만드는 경향, 즉 조절효과가 없다는 귀무가설을 기각하지 못하고 대신에 조절효과를 가정한 연구자의 연구모형을 잘못 기각하게 만드는 확률, 즉 약한 파워를 만들어내는 제반 조건과 이를 해결하는 방안들에 대해 연구가 이루어지게 된 것이다.

1. MMR의 파워에 영향을 미치는 제반요인

MMR의 통계적 검정력에 부정적인 영향을 미치는 요인 내지 가공사실(artifacts)들, 즉 어떤 조건하에서 조절효과가 있음에도 잘 포착이 안될 수 있는가에 관해 다음과 같은 것들이 구명되고 있다.

가. 총 표본 크기

표본의 크기는 어떤 통계적 추론에서나 통계적 검정력에 긍정적인 영향을 미친다(Cohen, 1988). MMR의 경우에도 이는 통계적 검정력에 영향을 미치는 요인들 중 단일요인으로는 가장 중요한 요인일 것이다. MMR의 통계적 검정력에 미치는 효과에 대한 최근의 연구들(Alexander & DeShon, 1994; Stone-Romero & Anderson, 1994)에 의하면 표본크기가 120일때는 조절효과가 작은 경우 포착이 안되고 120이하에서는 중간크기나 큰 조절효과까지 포착이 안될 수도 있다는 것이 밝혀지고 있다.

나. 조절변수에 기초한 하위집단간의 불균등한 표본 크기

조절변수를 기초해 하위집단을 나누었을 때 하위집단간의 표본크기가 다르다는 것도 통계적 검정력에 영향을 미친다. 이는 조절변수가 인종이나 성별처럼 구분변수(Z)의 경우에 두드러지는데, 이 경우 실제연구에서는 지배 집단($Z=1, m_1$)과 소수집단($Z=2, n_2$)간에 표본크기가 상당히 차이가 날 수 있고 이는 통계적 검정력을 약화시킨다. 인사관리연구에서 성별, 인종별 조절변수의 효과가 잘 안 나타나는 데는 이러한 약한 검정력의 문제가 숨겨져 있는 수가 있다.

통계적으로 하위집단 규모가 서로 다른 상황($n_1 \neq n_2$)에서 한 하위집단의 크기가 n_1 으로 고정되어 있으면 통계적 검정력은 다른 하위집단의 표본크기가 얼마든 그 하위집단의 두배, 즉 $2(n_1)$ smd 이용한 검증의 검정력(파워)을 초과할 수 없다는 것이다.

덧붙여 MMR에서는 조절효과는 결국 기준변수와 곱셈항($X*Z$)의 준 편상관관계(semi-partial correlation)라고 볼 수 있는데 이 상관계수의 최대치는 하위집단의 비율이 동일할 때($p=.50$) 나타나며 p 가 .50을 이탈할수록 상관관계의 최대치는 낮아지고 그럼으로써 조절변수를 포착할 통계적 검정력은 동시에 떨어지게 된다.

다. 예측변수의 범위제한

실제의 현장연구에서는 예측변수의 범위제한(range restriction)이 불가피한 경우가 많은데, 이도 조절효과 검증의 통계적 검정력을 떨어뜨리는 요인들중 하나가 된다. 범위제한이란 모집단의 구성원들이 표본에 뽑힐 확률이 동일하지 않음으로써 발생하는 무작위표본이나 편기된 표본의 예이다. 예컨대 경영연구, 특히 인사관리의 연구에서 선발절차가 범위제한의 주요한 요인 중의 하나이다. 즉, 영어시험 점수가 인사선발의 기준이 되는 경우 영어시험점수가 일정기준을 넘어서는 경우(x)만 채용이 된다. 그런데 영어점수라는 예측변수(X)와 직무성적이라는 기준변수(Y)간의 상관을 구한다면 예측변수는 x 이상이라는 범위의 제한을 받게 되고, 그 결과 분산이 작아져 실제보다 작은 상관계수가 구해지게 될 가능성이 높아지게 된다.

최근의 연구들에서는 범위의 제한을 안 받는 $X*Z$ 점수보다 제한을 받는 $X*Z$ 점수의 분산이 작고 그 결과 파워에 부정적인 영향을 미친다는 연구결과가 발표되고 있다. 예컨대 총표본 크기가 300이고 범위제한이 없는 경우 중간크기의 조절효과를 포착할 파워는 0.8정도 되었다. 반면 모집단의 상위 80%만을 표본으로한 경우 파워는 추천수준 0.8을 크게 하회하는 0.51로 줄어들었다. 즉 약간의 범위제한에도 파워는 크게 손상을 입을 수 있다는 것이다.

라. 오차분산의 이질성

예측변수의 범위제한과 함께 변수의 분포와 관련된 문제로 오차분산의 이질성(error variance heterogeneity)의 문제가 있다. 이는 조절변수가 이분변수 등의 구분변수일 경우의 검증에서 문제가 된다. 검증의 기초가 되는 하위집단간 오차분산 동질성의 가정

이 위반이 되고 있는 것이다. 모집단에 조절효과가 존재할 때 조절변수에 기초한 하위집단내의 X 와 Y 간의 상관관계는 다를 수 밖에 없고, 따라서 오차항은 필연적으로 달라지게 된다.

이러한 오차분산의 이질성은 파워를 상당히 줄이고 있는 것으로 나타나고 있다 [Alexander & DeShon, 1984].

마. 변수들의 측정오차

연구에 있어 실제로 완벽하게 신뢰성 있는 측정치를 사용하지 못한다는 문제, 즉 측정오차(measurement error)가 존재한다는 것도 파워를 제약하는 문제 중의 하나이다. 상관관계를 구하는 경우 변수들 측정치의 비신뢰성으로 인해 상관계수가 실제보다 희석(attenuated)되는 것처럼 MMR의 경우에도 예측변수 측정치의 비신뢰성이 존재한다면 관찰된 회귀계수는 희석이 되는 효과를 낼 것이다. 다시 말해 예측변수가 완벽한 신뢰성을 유지하지 못한다면 표본에 기초한 곱셈항의 회귀계수(b_3)는 모집단의 계수(β_3)를 과소반영하게 되는 것이다.

X 와 Z 간의 상관관계가 0일 때 곱셈항의 신뢰성은 X 의 신뢰성 $X*Z$ 의 신뢰성이 된다. 예컨대 X 의 신뢰성과 Z 의 신뢰성이 모두 0.7이면 곱셈항의 신뢰성은 0.49가 된다. X 와 Z 간의 상관관계가 증가하면 곱셈항의 신뢰성은 다소 증가하겠지만 측정오차의 존재는 β 의 수치를 줄이는 바이어스를 낳게 하고 결과적으로 검증의 파워를 약화시키는 것이다. 표본크기가 증대되면 파워가 다소 증가하기는 하지만 근본적인 문제해결은 못된다.

최근의 한 연구(Aguinis & Stone-Romero, 1997)는 MMR을 사용하는 연구자들을 힘빠지게 만드는데 그들에 의하면 경영이나 사회과학의 연구에서 비교적 적당하다고 생각되는 신뢰성의 수준, 즉 0.8(Nunally, 1978)에 있어서조차도 조절효과를 포착할 파워는 추천된 수준 0.8(Cohen, 1988)보다 훨씬 작다는 것이다. 예컨대 측정오차가 전혀 없는 완벽한 신뢰성의 측정치를 가지는 경우 표본크기가 300일 경우 조절효과 포착의 파워는 0.85를 넘는다. 그런데 측정치의 신뢰성이 0.8의 수준이라면 파워는 0.45로 떨어지게 된다.

바. 예측변수간 상호관련성

MMR분석을 실시하는데 투입되는 예측변수간의 상호관련성, 즉 다중공선성(multi-

collinearity)의 문제도 파워에 부정적인 영향을 주는 요인으로 지적된다. 흔히 MMR분석에서는 예측변수들(X 와 Z)과 이 두 변수들의 곱셈항($X*Z$)이 사용된다. 따라서 X 와 Z 간에는 상호관련성이 없다하더라도 X 와 $X*Z$ 그리고 Z 와 $X*Z$ 간에는 높은 상호관련성이 존재할 수밖에 없다. 그리고 이러한 예측변수간 상호관련성, 즉 다중공선성은 회귀계수를 불안정하게 만들고 오차를 크게 만듦으로써 파워를 감소시키게 한다(Morris et al., 1986).

그러나 한편으로는 MMR의 파워에 미치는 다중공선성의 문제가 그렇게 심각한 것은 아니라는 주장도 제기되고 있다. 다중공선성의 문제를 깊이 다룬 Morris의 연구에서는 예측변수가 상당히 많은 상황을 상정하고 있는데 실제로 조절효과를 분석하는 많은 상황은 대개 예측변수가 조절변수를 포함하여 두 개인 경우가 많다. 이러한 전형적인 상황에서는 다중공선성은 파워에 그다지 심각한 문제가 안 된다는 것이다(Cronbach, 1987).

그러나 이러한 주장자들도 어쨌든 다중공선성이 큰 경우에는 문제가 되므로 이를 해소할 수 있는 방안을 사용할 것을 추천하고 있다.

사. 척도의 조악성

기준변수의 '척도의 조악성'(scale coarseness)이라는 것도 파워에 영향을 미칠 수 있다. 기준변수의 조악성이란 예측변수에 비해 기준변수가 충분한 척도점을 갖지 못하는 현상을 말한다. 척도점의 부족은 정보의 손실을 가져오고 이는 조절효과와 포착을 방해하는 결과를 낳게 한다.

예컨대 예측변수 X 와 조절변수 Z 가 리커트형 7점 척도에 의해 측정되었다고 하자. 그러면 이의 곱셈항 $X*Z$ 은 49점 척도로 점수가 분산되게 된다. 이에 반해 기준변수 Y 를 상대적으로 '조악'하다고 할 수 있는 7점 척도로 측정하였다면 Y 와 $X*Z$ 간의 관련성에 대한 정보는 손상이 되며 조절효과가 과소평가되고 따라서 파워는 떨어지게 된다.

아. 제 요인들의 상호작용 효과

이상에서 설명한 요인들은 각기 독립적으로도 MMR분석의 파워에 영향을 미치기도 하지만 몇 개의 요인들의 상호작용효과도 큰 것으로 보고되고 있다. 이들 요인들에 주의를 기울이지 않는다면 상당한 크기의 조절효과도 포착이 안될 가능성이 있는 것이다. 한 연구(Aguinis & Stone-Romero, 1997)에 따르면 예측변수의 범위제한, 전체 표본크기,

조절변수로 구분된 하위집단간 표본크기의 차이 등의 요인들이 개별적으로 또는 상호작용하여 파워에 상당히 부정적으로 작용하는 것이 밝혀지고 있다.

IV. MMR분석에 있어 발생하는 파워감소이슈의 해결방안

앞서 설명한 요인들은 독립적으로 또는 상호작용해서 MMR분석에 있어 파워를 감소시킬 가능성이 있는데 심한 경우는 0.5까지 이를 경우도 있다. 이러한 현상은 당연히 나타나야 할 조절효과를 포착하지 못하게 함으로써 연구결과를 호도할 가능성이 매우 높다. 따라서 파워에 미치는 부정적 요인들의 연구와 함께 이를 개선할 수 있는 방안에도 많은 노력이 경주되고 있다. 아직 만족스러울만한 방안이 제시되지는 못했지만 다음과 같은 해결방안들이 모색되고 있다.

가. 사전 연구설계단계에서의 철저한 대비

연구설계를 조심스럽게 또 철저하게 하는 것은 모든 연구에서 아무리 강조해도 지나칠 문제가 아니지만 파워문제와 관련해서도 이는 매우 중요하다. 데이터를 수집하기 전에 범위제한의 문제나 측정문제(구성개념의 신뢰성있고 타당성있는 측정)를 충분히 고려한 연구설계를 할 필요가 있다.

사전연구단계에서의 대비책으로 예를 들면 다음과 같은 것들도 추천되고 있다. 첫째는 앞서 척도의 조악성이라는 문제의 해결로서 리커트의 5점 척도나 7점 척도처럼 눈금이 있는 척도대신 척도선상에 연구자가 체크를 하고 이의 길이를 실제로 재는 방법이 추천되고 있다[Russel & Bobko, 1992]. 그런데 이는 실제 시행에 시간이 많이 걸리고 또 잘못 잴 가능성이 크다는 단점이 있다. 다행히도 최근에는 이를 컴퓨터 프로그램으로 해결하는 방법이 제시되고 있다[Aguinis, Bommer & Pierce, 1996]. 즉, 응답자가 컴퓨터 프로그램 상의 척도에 자신의 응답을 표시하면 프로그램이 자동적으로 이를 수치화해주는 프로그램이다.

또 하나의 방안은 표본추출에 신경을 쓰는 것인데 극단적인 수치만을 보이는 표본을 뽑는 것이다[McClelland & Judd, 1993]. 이러한 방법은 변수의 분산을 높여주어 그 결과 분석의 파워를 올리는 효과를 갖는다. 그러나 여기서 얻어진 결론은 전체 모집단이

아니라 극단적인 점수를 갖는 표본들에만 일반화될 수 있다는 비판을 모면하기 힘들다.

나. 파워의 추정

앞서 설명한 것처럼 연구설계의 단계에서부터 파워를 높이려는 노력이 이루어질 수도 있으나 현실에서는 연구자가 상황에 대한 통제력이 별로 없는 경우가 많다. 즉, 연구대상이 한정된 상황에서 연구를 한다든지 신뢰성이 높은 측정치를 구하기가 어려운 경우가 있다. 이러한 상황에서는 연구자는 불가피하게 현재 가능한 대상에 의존할 수 밖에 없다. 이 때 연구자는 현재 자기의 특정상황(표본수, 범위제한, 측정오차 등)에서 MMR분석의 파워를 추정해볼 필요가 있다.

다행히도 이러한 추정을 용이하게 해주는 방안이 강구되고 있다. 예컨대 예측변수와 조절변수가 모두 연속척도일 경우 $\alpha=0.05$ 수준에서 우리가 용인할 수 있는 0.8의 파워를 낼 수 있는 샘플의 크기를 연구자에게 알려주는 표가 만들어져 활용되고 있다. 이러한 표를 이용하여 자료수집의 사전에 필요한 파워를 얻기 위한 충분한 표본의 크기를 결정할 수 있다. 또 예측변수가 연속척도이고 조절변수가 이분척도인 경우 (1) 총 표본크기 (2) 이분변수별 표본크기 (3) 하위집단별 예측변수와 기준변수간 상관계수 등의 수치를 근거해서 $\alpha=0.05$ 수준에서 파워를 측정해주는 컴퓨터 프로그램이 개발되어 있는데, 이는 연구자들에게 도움을 줄 수 있는 프로그램으로 평가되었다[Aguinis, Pierce, & Stone-Romero, 1994].

최근에 Aguinis[Aguinis, Boik, & Pierce, 2001] 등은 다시 이 프로그램을 개정하고 있다. 그것은 기존 프로그램이 다음의 다섯 가지 문제점을 지니고 있었기 때문이다.

첫째는 기존 프로그램이 측정된 변수들의 측정오차를 고려하지 않고 있다는 점이다. 범주척도인 Z는 측정오차가 없다는 가정이 들어맞을지 모르나 X나 Y의 경우에도 그러하길 기대하기는 어려운 것이다

둘째, 이 프로그램은 몬테칼로(Monte Carlo)법에 기초하고 있는데 여기서는 MMR의 파워에 영향을 미치는 요인들에 대해 매우 제한된 값만이 포함되어 있다는 것이다. 예컨대 Aguinis와 Stone-Romero[Aguinis & Stone-Romero, 1997]가 만든 프로그램은 샘플 사이즈를 60과 300의 경우의 시뮬레이션 결과에 기초하여 만들어진 것이기 때문에 샘플 사이즈가 여기서 멀어지는 경우에는 정확한 파워의 추정이 어려워진다는 문제가 있다. 게다가 그들의 시뮬레이션은 X와 Y간의 부의 관계를 포함하고 있지 못하다. 그러

나 현실에는 두 관계가 부의 관계를 가지는 경우의 조절회귀분석이 필요한 경우도 있다. 예컨대 '낙관주의'라는 예측변수가 "수술로 부터의 회복시간"이라는 기준변수에 미치는 부정적 관계에서 성별변수의 조절효과를 검증하는 경우가 있을 수 있다. 이러한 경우에는 앞서의 컴퓨터 프로그램은 파워에 대한 부정확한 추정치를 제공해줄 가능성이 높아진다.

셋째, 기존의 프로그램은 X 변수를 일정시점에서 자르는 것을 고려할 수 있게 되어 있다. 그런데 이러한 절단은 X 변수가 정규분포를 이룰 때는 파워를 줄이는 한 요인으로 작용할 수 있다. 그러나 이러한 예 말고도 샘플의 분산이 모집단의 그것보다 작아지게 만드는 절단효과가 발생할 가능성이 있다. 예컨대 학교의 학점과 입사후 실적과의 관련성을 알아보는 경우 최고점수의 사람들은 입사제안을 거절함으로써 실제 분석 샘플집단의 분산이 모집단의 분산보다 작아지는 경우가 있다. 이 경우를 기존의 프로그램에서는 고려치 못하고 있다.

넷째, 기존의 프로그램은 조절변수가 범주변수인 경우 성별과 같이 이분변수의 경우만 파워를 계산할 수 있다. 그러나 실제로 보면 민족(한국인, 일본인, 중국인)처럼 범주가 셋 이상인 경우도 많은 데 이 경우는 파워의 계산이 힘들다는 문제를 가지고 있었다.

다섯째, 또 다른 기존 프로그램의 문제는 범주 변수인 조절변수가 둘 이상의 경우 파워의 계산이 불가능하다. 예컨대 범주변수로서 성별과 인종이 같이 들어가는 경우 성별 X 변수와의 상호작용, 인종별 X 변수와의 상호작용, 인종과 성별간의 상호작용 등에 대한 파워의 계산이 필요할 수가 있는데 이 경우 기존 프로그램으로는 파워를 계산할 수가 없다.

그러나 이러한 표나 컴퓨터 프로그램이 아직 충분한 도구가 되지는 못하고 있다. 우선 파워를 약화시키는 것으로 알려진 요인들 중 고려가 안 되는 요인들(범위제한이나 측정 오차 등)이 있을 뿐만 아니라 앞서 언급된 것과 같은 요인들의 독립적인 영향 뿐만 아니라 상호작용효과도 검토가 되어야하는 바, 이러한 점에서는 개선의 여지가 크다고 할 수 있다.

다. 평균집중화

앞서 MMR분석시 파워를 약화시키는 요인의 하나로 예측변수간 다중공선성의 문제를 언급한 바 있다. 이의 해결방법의 하나로써 평균집중화(mean centering)의 방법이 제시되고 있다(Aiken & West, 1991). 이는 간단한 자료변형으로 조절효과의 분석전에

각 관찰치로부터 표본평균을 차감함으로써 집중화시키는 방법이다. 이렇게 되면 모든 점수는 편차점수로 표시되고 그 평균은 0이 될 것이다.

이 방법의 주장자들에 따르면 다중공선성의 문제는 오차를 증가시킴으로써 파워가 약화된다는 문제를 야기하였는 바 이러한 평균집중화는 표준오차를 감소시키는 효과를 가져옴으로써 추정치의 정확성과 안정성을 증가시킬 수 있다는 것이다.

그러나 이러한 자료의 집중화 효과에 대해 회의적인 주장도 적지 않다. 실제 조절효과 검증이 이러한 자료의 변형에 의해 별로 영향을 받지 않는다는 것이다[Cronbach, 1987]. 또 영향이 있다하더라도 원점수를 사용한 것과 집중화된 변형된 자료를 사용할 때 조절효과 검증에 별 차이가 없음에 비추어 노력한 것만큼의 실효가 없다는 회의적인 연구결과도 나오고 있다[Kromery & Foster-Johnson, 1998].

라. 표본크기의 증대

연구의 진행과정에서 낮은 검정력이 발견되는 경우 표본크기를 증대시키는 것이 파워를 높이는 방안의 하나이다. 조절회귀분석에 있어 통계적 추론의 검정력은 표본의 수가 증가할수록 커지는 것이 사실이다[Cohen, 1988]. 그러나 현실적으로는 이런 방안을 채택하기가 어려운 경우가 많다.

마. 통계적 보정치의 이용

범위제한이나 측정오차 때문에 발생하는 파워의 문제는 통계적 보정치를 사용하는 것도 하나의 방법일 수 있다[Ghiselli E. E., Campbell, J.P. & Zedeck, S. 1981]. 그러나 상관계수의 보정치를 이용하여 조절변수의 효과를 검증하는 방법이 과연 적절한 것인가에 대해 논란이 제기되고 있다. 보정치에 대한 표준오차를 모를 뿐더러 설혹 그것을 안다 하더라도 보정 후에는 그것이 증가할 것이므로 보정치를 사용한 실익이 없다는 것이다[Aguinis, 1995].

바. 구조방정식 모형의 이용

측정오차 때문에 발생하는 파워약화의 문제에는 구조방정식 모형(structural equation model)에 의한 상호작용 분석이 하나의 대안으로 제시되기도 한다. 구조방정식 모형에서는 복합지표를 이용하고 측정오차를 고려한 분석이 이루어지므로 만족할만한 상호작용

의 분석이 가능하다는 것이다(Jaccard & Wan, 1995). 문제는 LISREL과 같은 구조 방정식 모형분석의 컴퓨터 프로그램을 이용하기가 복잡하다는 것인데, 이는 LISREL 8 판에서 부터는 상당히 개선되고 있다. 그러나 그럼에도 불구하고 활용상의 어려움은 구조 방정식 모형에 의한 조절효과 검증에는 여전히 장애의 하나로 남아 있다.

사. 하위집단간 오차분산의 동질성을 요구하지 않는 검증법 사용

조절변수를 이분형 척도나 구분변수로 측정하여 MMR 분석을 할 때에는 F검증에 필요한 하위집단간 오차분산의 동질성이 유지되지 못한다는 문제점이 있었다. 이러한 문제의 해결방안으로는 χ^2 검정 같이 하위집단간 오차분산의 동질성을 요구하지 않는 검증법의 사용이 추천되고 있다. 그런데 χ^2 검증은 개념상 집단간 기울기의 동질성에 대한 검증이 아니라 상관계수의 동질성에 대한 검증이라는 문제점이 존재한다.

최근에는 이 문제에 대한 유효한 몇 개의 지침이 제시되고 있다.

첫째, 조절변수에 기초해 구분된 하위집단의 오차분산을 계산해 한 하위집단의 오차분산이 다른 하위집단의 그것의 1.5배가 넘지 않는다면 MMR에서 사용하는 F분석에 큰 부정적인 영향이 없다는 것이다(DeShon & Alexander, 1996). 즉, MMR분석을 해도 좋다는 것이다.

둘째, 한 집단의 오차분산이 다른 집단의 그것의 1.5배를 초과하면 이질성의 존재를 인정해 다른 통계분석을 실시하는 것이 좋다. 이 때 대안이 되는 것이 A분석(A statistics)과 J분석(J statistics)이다. A분석은 계산이 용이하다는 강점이 있고 J분석(DeShon & Alexander, 1996)은 작은 규모의 표본에 통계적 파워가 있다는 장점을 가지고 있다.

최근에 개발되어 비교적 생소한 이 두 검증방법을 기초통계량만 투입하면 쉽게 계산해주는 프로그램이 개발되어 웹상에 올라와 있어 용이하게 활용할 수 있다. 프로그램명은 ALTMMR(Alternative to MMR)이고 인터넷 주소는 <http://members.aol.com/IMSAP>이다.

아. 1종 오류의 증대

이론적으로 파워란 2종 오류를 범하지 않을 확율을 말하는 것으로써 이를 높이려면 1종 오류를 증대시키면 된다. 즉, 전통적인 기준보다 높은 예컨대 $\alpha=0.1$ 등의 기준을 적

용하는 것이다. 그러나 통상 사용하는 $\alpha = 0.05$ 보다 높은 기준은 연구자들이 사용하기를 꺼려할 뿐만 아니라 이 기준을 이용하여 작성된 연구결과를 심사하고 게재할 논문 편집자의 입장에서도 그리 환영받지는 못할 것이라는 점이 문제이다.

자. 회귀방식의 변경

조절효과가 잘 포착이 안 된다는 문제의 해결방법으로 회귀방식의 변경을 추천하고 있다. 즉 통상의 조절효과 검증에 있어서는 Y 를 예측하는데 X 와 Z 를 넣고 다음에 곱셈항 $X*Z$ 를 투입하는데 반해, 여기서 추천되는 것은 이를 역으로 하여 곱셈항 $X*Z$ 를 먼저 투입하는 것이다. 즉, 후방 회귀(backward regression)방식을 사용하는 것이다. 이 방법은 확실히 이 곱셈항이 설명하는 분산의 부분을 크게 할 것이고, 따라서 조절효과가 유의할 확율을 높여줄 것이다.

그러나 이러한 방식은 회귀분석의 기본원리를 위반하고 있기 때문에 사용해서는 안 된다는 주장이 제기되고 있다(Stone, 1988). 상호작용효과를 알아보려는 회귀분석에서는 곱셈항은 각 변수들보다 먼저 투입되어서는 안되며 또 곱셈항이 들어가는 회귀분석에서는 그것을 구성하는 변수들이 빠져서는 안된다는 것이다. 따라서 곱셈항을 먼저 투입하는 회귀분석방법은 조절효과를 포착하는데 유리한 방법일지 몰라도 사용해서는 안 되는 방법인 것이다.

V. 결 론

앞서 조절효과 검정을 위해 가장 널리 쓰이고 있는 MMR법의 사용시 발생할 수 있는 통계적 검정력의 문제들을 살펴보았다. 그리고 이를 해결하기 위한 방법들에 대해서도 제시하여 보았다.

이러한 논의들 과정에서 많은 조절효과의 검증과 관련되어 연구자들이 느꼈던 좌절이 사실은 방법상의 문제에 기인하였다는 사실을 알 수 있었다. 동시에 연구방법과 관련된 통계적 검정력의 문제에 조심한다면 조절효과의 포착이 가능할 수 있다는 희망적인 단서도 얻을 수 있었다.

물론 여기에서 제시된 검정력 향상의 방안들이 완벽한 것은 못된다. 나름대로의 한계

들을 갖고 있다. 그러나 연구설계의 단계에서부터 검정력과 관련된 몇 가지 대처방안들을 강구하게 되면 조절회귀분석의 통계적 검정력은 크게 향상될 수 있다는 점에서 본고에서 제시된 몇 가지 해결안들은 연구자들에게 매우 유용하다고 할 수 있다. 뿐만 아니라 본 연구에서 제시된 바 있는 최근 개발된 웹 상의 몇가지 프로그램들도 검정력의 문제를 해결하는 쓸모 있는 도구로서의 역할을 할 수 있을 것이다.

참 고 문 헌

- 이규범, 『조직공정성과 조직시민행동의 관계에 관한 연구』, 충남대학교, 석사학위논문, 1997.
- Aguinis, H. (1995), Statistical power problems and moderated multiple regression in management research, *Journal of Management*, Vol. 21, no. 6, pp. 1141-1158 .
- Aguinis, H. (2002), Estimation of interaction effects in organization studies, *Organizational Research Methods*, Vol.5, pp.207-211.
- Aguinis, H., & Anderson L. E. (1994), Relative statistical power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects, *Journal of Applied Psychology*, 79, pp. 354-359.
- Aguinis, H., Boik R. J. & Pierce, C. A. (2001), A Generalized solution for approximating the power to detect effects of categorical moderator variables using multiple regression, *Organizational Research Methods*, Vol.4.
- Aguinis, H., Bommer, W. H., & Pierce, C. A. (1996), Improving the estimation of moderating effects by using computer-administered questionnaires, *Educational and Psychological Measurement*, Vol. 56, no. 4, pp. 1043-1047.
- Aguinis, H., Pierce, C. A. & Stone-Romero E., (1994), Estimating the power

- to detect dichotomous moderators with moderated multiple regression, *Educational and Psychological Measurement*, Vol. 54, pp. 690-692.
- Aguinis, H. & Stone-Romero E (1997), Methodological artifacts in moderated multiple regression and their effects on statistical power, *Journal of Applied Psychology*, Vol. 82, no. 1, pp. 192-206.
- Aiken, L & West, S. (1991), *Multiple Regression*, Newbury Park, CA: Sage.
- Alexander R. A. & DeShon R. D. (1994), Effects of error variance heterogeneity on the power of tests for regression slope differences, *Psychological Bulletin*, Vol. 115, no. 2, pp. 308-314.
- Bobko, P. & Russell, C.J. (1994), On theory, statistics, and the search for interactions in the organizations science, *Journal of Management*, Vol. 20, pp. 193-200.
- Cortina J. M. (1993), Interaction, nonlinearity, and multicollinearity: Implications for multiple regression, *Journal of Management*, Vol. 19, pp. 915-922.
- Cohen, J., & Cohen, P. (1983), *Applied multiple regression/correlational analysis for the behavioral science*, 2nd ed., Hillsdale, NJ: Erlbaum.
- Cronbach, L .J. (1987), Statistical tests for moderator variables: Flaws in analyses recently proposed, *Psychological Bulletin*, 102, pp. 414-417.
- DeShon, R. P., & Alexander, R. A. (1996), Alternative procedures for testing regression slope homogeneity when group error variances are unequal, *Psychological Methods*, Vol.1, pp.261-277.
- Erwin, J. R. & McClelland, G. H. (2001), Misleading heuristics and moderated multiple regression models, *Journal of Marketing Research*, Vol. 38, no. 1, pp.100-109.
- Ghiselli E. E., Campbell, J.P. & Zedeck, S. (1981), *Measurement theory for the behavioral science*, San Francisco, CA: Freeman.
- Jaccard J. J & Wan C. K. (1995), Measurement error in the analysis of interaction effects between continuous predictors using multiple

- regression multiple indicator and structural equation approaches. *Psychological Bulletin*, Vol. 117, No. 2, pp.348-356.
- Kromery, J. D. & Foster-Johnson, L. (1998), Mean centering in moderated multiple regression: Much ado about nothing, *Educational and Psychological Measurement*, Vol. 58, pp. 42-67.
- Maxwell S. E., & Delany, H. D. (1993). Bivariate median splits and spurious statistical significance, *Psychological Bulletin*, Vol. 113, no. 1, pp.181-190.
- McClelland G. H. & Judd C. M. (1993), Statistical difficulties of detecting interactions and moderator effects, *Psychological Bulletin*, Vol. 114, no. 2, pp. 376-390.
- Morris J. H., Sherman J. & Mansfield E. R. (1986), Failures to detect moderating effects with ordinary least squares moderated-regression: Some reasons and a remedy, *Psychological Bulletin*, 99, pp. 282-288.
- Stone, E. F. (1988), Moderator variables in research: A review and analysis of conceptual and methodological issues, pp. 191-229 in G. R .Ferris & K. M. Rowland(eds.), *Research in Personnel and human resources management*, Vol.6, Greenwich, CT: JAI Press.

Moderated Multiple Regression and Statistical Power Problems

In Soo Han*

ABSTRACT

Moderated multiple regression (MMR) has been increasingly used for detecting moderating effect in various management researches such as organizational behavior, human resources management, and industrial relations. But its wide use, recent researches has reported that this method has some flaws, that is, low statistical power problems. This article identifies several factors that reduce statistical power below acceptable levels in the process of using MMR approach for detecting moderating effect. Finally it proposes some solutions to increase statistical power in MMR use.

* Professor, Chung-nam National University.