

## Tratamentul prin analiză de regresie a variabilelor de control în studiile non-experimentale\*

Dan Ispas<sup>1</sup>

### Abstract

This short presentation focuses on the procedure for statistical treatment of control variables in nonexperimental studies, through multiple regression analysis. Indications are illustrated in a step-by-step manner, based on an actual study. Some recommendations are made, based on research literature, regarding the optimal deployment of this technique for the statistical control of variables.

**Keywords:** regression analysis, control variables

### Résumé

Cette brève présentation explique comment les variables de contrôle peuvent être traitées dans les études expérimentales, en utilisant une analyse de régression multiple. Les indications sont illustrées étape par étape, avec l'aide d'une étude concrète. Des recommandations basées sur la littérature en domaine, sont formulées à propos de la modalité dans ce contrôle statistique peut être utilisé de façon optimale.

**Mots clés:** analyse de régression, variables de contrôle

### Rezumat

Această scurtă prezentare discută modalitatea în care se pot trata variabilele de control în studiile experimentale, prin utilizarea analizei de regresie multiplă. Indicațiile sunt ilustrate pas cu pas pe baza unui studiu concret. Sunt făcute recomandări bazate pe literatura de specialitate, cu privire la modalitatea în care acest control statistic poate fi utilizat în mod optim.

**Cuvinte-cheie:** analiză de regresie, variabile de control

Variabilele de control sunt variabile incluse în studii pentru a putea elimina explicațiile alternative ale rezultatelor. În literatura de specialitate (Becker, 2005) sunt prezentate două metode primare pentru tratamentul variabilele de control: manipularea experimentală sau controlul statistic. Să presupunem că genul participanților este o variabilă pe care vrem să o controlăm în studiul nostru. De exemplu, folosind manipularea experimentală putem controla natura eșantionului astfel încât să avem în studiu doar femei sau doar bărbați. Folosind controlul statis-

tic, putem include în analize variabila categoricală gen (1=bărbați, 2=femei). Cum majoritatea studiilor în psihologia industrial-organizațională sunt studii non-experimentale, în acest tutorial vom discuta controlul statistic, ca metodă de tratament al variabilelor de control.

\* Acest articol este un tutorial bazat pe sursele citate în referințe.

<sup>1</sup> Department of Psychology, Illinois State University, Normal, Illinois, USA.  
Adresa pentru corespondență: dispas@gmail.com.

### Tratamentul prin analiză de regresie a variabilelor de control în studiile non-experimentale

#### Controlul statistic folosind regresia multiplă

În general, scopul oricărui studiu este examinarea relației dintre variabila predictor (X) și variabila criteriu (Y). Acest lucru se face, în general, prin modele multivariate sau prin modele factoriale. Modelele multivariate se bazează pe corelație și pe diverse aplicații ale acesteia, pentru punerea în evidență a relațiilor dintre variabile, pe când în modelele factoriale obiectivul cercetării este comparația dintre grupuri. Ne vom concentra în acest tutorial asupra modelelor multivariate.

În studiile multivariate non-experimentale, care reprezintă cea mai mare parte a studiilor din domeniul psihologiei muncii și organizațională, nu este posibil să manipulăm variabila X. Însă este posibil ca o altă variabilă (C) să fie responsabilă pentru relația dintre X și Y și este de datoria noastră să verificăm dacă acesta este cazul. Astfel, când examinăm relația dintre X și Y este important să controlăm statistic variabila C. Cea mai folosită strategie pentru control statistic este folosirea regresiei ierarhice. Variabila control C este în introdusă în pasul 1, iar variabila predictor X este introdusă în pasul 2. Schimbarea în  $R^2$  între pasul 1 și pasul 2 (corelația semi-parțială ridicată la pătrat) estimează varianța unică în criteriul Y explicată de predictorul X.

O altă strategie constă în includerea atât a variabilei X, cât și a variabilei C, în aceeași ecuație de regresie. Ambele strategii ne oferă informații echivalente: semnificația statistică a corelației semi-parțiale ridicată la pătrat este identică cu semnificația statistică a coeficientului de regresie (Pedhazur, 1997 apud Breaugh, 2008).

#### Ilustrație empirică

O metodă des folosită este includerea ca variabilă de control a unor trăsături de personalitate atunci când examinăm relația dintre atitu-

dini și comportamente la locul de muncă. De exemplu, Ispas și Ilie (2008) au examinat relația dintre satisfacția la locul de muncă și comportamentele contraproductive (CWB). Un factor important în determinarea satisfacției la locul de muncă este personalitatea, mai exact, *core self-evaluations* (CSE, Judge et al., 2003). Fundamentul teoretic și empiric al relației dintre CSE și satisfacția la locul de muncă este prezentat în Bono și Judge (2003).

Pentru aceste motive, am inclus ca variabilă de control CSE în ecuația de regresie ce include drept predictor satisfacția la locul de muncă și drept criteriu CWB. Am folosit un eșantion de 196 de angajați. Satisfacția la locul de muncă a fost măsurată cu scala de 3 itemi (alfa = .88) validată de Camman et al. (1979). CWB au fost măsurate folosind scala de 27 de itemi (alfa = .91) din Fox și Spector (1999). CSE au fost măsurate folosind scala de 12 itemi (alfa = .91) din Judge et al. (2003).

Pentru analiză, am folosit regresia ierarhică. În pasul 1, am introdus variabila de control, CSE. În pasul 2, am introdus variabila predictor - satisfacția la locul de muncă. Variabila dependentă (criteriu) a fost CWB. În SPSS 18, analiza se efectuează selectând Analyze à Regression à Linear, apoi introducem variabila (variabilele) de control în Block 1 și variabila predictor în Block 2. De asemenea, este necesară selectarea opțiunii „R square change” din submeniul „Statistics”.

Statisticile descriptive și corelațiile dintre cele trei variabile sunt prezentate în tabelele 1 și 2. Din rezultatele regresiei ierarhice rezultă că, după ce controlăm pentru CSE, satisfacția la locul de muncă explică varianța incrementală în CWB:  $\Delta R^2 = .08$ ;  $F(1, 193) = 20.83$ ;  $p < .001$ . Aceste rezultate sunt vizualizate în Tabelul 3. Rezultă deci că după ce controlăm efectele CSE, satisfacția la locul de muncă explică varianță unică în predicția CWB.

**Tabelul 1.** Statistici descriptive

	Mean	Std. Deviation	N
CWB	39.37	11.841	196
CSE	40.19	9.811	196
SAT	12.52	4.086	196

**Tratamentul prin analiză de regresie a variabilelor de control în studiile non-experimentale**

**Tabelul 2.** Corelații

		cwbtot	csetot	sattot
Pearson Correlation	CWB	1.000	-.385	-.417
	CSE	-.385	1.000	.396
	SAT	-.417	.396	1.000
Sig. (1-tailed)	CWB	.	.000	.000
	CSE	.000	.	.000
	SAT	.000	.000	.
N	CWB	196	196	196
	CSE	196	196	196
	SAT	196	196	196

**Tabelul 3.** Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.385 <sup>a</sup>	.148	.144	10.957	.148	33.722	1	194	.000
2	.481 <sup>b</sup>	.231	.223	10.437	.083	20.829	1	193	.000

a. Predictors: (Constant), CSE.

b. Predictors: (Constant), CSE, SAT.

Dacă ne interesează doar corelațiile dintre variabile, putem analiza corelațiile parțiale dintre variabile. În exemplul de mai sus, corelația directă dintre satisfacție și CWB este  $-.417$ ,  $p < .001$ . După ce controlăm efectele CSE, corelația parțială dintre satisfacție și CWB este  $-.312$ ,  $p < .001$ . Deși redusă în magnitudine, corelația este substanțială și semnificativă statistic. Corelația parțială se selectează în SPSS astfel: Analyze à Correlate à Partial.

#### Recomandări pentru folosirea variabilelor de control

Data fiind ușurința cu care se poate realiza controlul statistic, unii cercetători pot fi tentați să abuzeze de această metodă, aplicând-o foarte des sau incluzând cât mai multe variabile de control în studiile lor. Această strategie nu este recomandată, datorită impactului negativ asupra puterii statistice.

Una din problemele care pot să apară în acest context și care face ca metoda descrisă să nu fie universal aplicabilă, este problema coliniarității dintre X și C, în contextul regresiei față de Y. Cu cât corelația dintre predictor și covarianța este mai mare, cu atât coeficienții ecuației de regresie sunt mai nesiguri, adică au erori standard mai mari. Acest fenomen conduce la reducerea puterii statistice a coeficienților beta. Fenomenul poate fi contracarat prin utilizarea unui eșantion mai mare. Coliniaritatea afectează și coeficienții de regresie parțială.

Mai jos prezentăm câteva recomandări extrase din literatura de specialitate (Becker, 2005; Breaugh, 2008; Spector & Brannick, 2010), care discută felul în care metoda descrisă aici poate fi aplicată corect și cu efect maxim.

1. Trebuie să se explice de ce C este variabilă de control și nu variabilă substanțială (C poate fi o altă variabilă predictor, variabilă mediatore sau variabilă moderatoare a relației dintre X și Y).

### Tratamentul prin analiză de regresie a variabilelor de control în studiile non-experimentale

2. Nu este necesar să se includă în analize variabile de control necorelate cu Y (decât dacă există motive pentru a suspecta că aceste variabile au efect supresor).

3. Trebuie să se includă doar variabile de control care au fundament teoretic.

4. Trebuie să se precizeze care este relația dintre variabila de control și variabilele principale din studiu. Există o relație cauzală sau doar corelațională?

5. Trebuie să se explice clar cum a fost măsurată fiecare variabilă de control.

6. Trebuie să se descrie clar metoda de analiză folosită (regresie, SEM).

7. Trebuie să se prezinte statistici descriptive pentru toate variabilele de control.

8. Trebuie să se prezinte informații privind validitatea și fidelitatea variabilelor de control.

9. Trebuie să se includă variabilele de control în matricea de corelații cu celelalte variabile.

10. Trebuie să se analizeze și să se includă în studiu rezultatele atât cu, cât și fără, variabilele de control.

11. Dacă este posibil, trebuie să se includă variabilele de control în ipotezele care vor fi testate. De exemplu: După ce controlăm variabila C, relația dintre X și Y este pozitivă/negativă.

### Referințe

- Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods, 8*, 274-289.
- Bono, J. E., & Judge, T. A. (2003). Core self-evaluations: A review of the trait and its role in job satisfaction and job performance. *European Journal of Personality, 17*, S5-S18.
- Breaugh, J. A. (2008). Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review, 18*, 282-293.
- Cammann, C., Fichman, M, Jenkins, D., & Klesh, J. (1979). *The Michigan Organizational Assessment Questionnaire*. Unpublished manuscript, University of Michigan, Ann Arbor.
- Fox, S., & Spector, P. E. (1999). A model of work frustration—aggression. *Journal of Organizational Behavior, 20*, 915-931.
- Ispas, D., & Ilie, A. (2008, April). *Negative emotions, core self-evaluations and counterproductive work behaviors*. Poster presented at the Annual Meeting of the Society for Industrial and Organizational Psychology.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale (CSES): Development of a measure. *Personnel Psychology, 56*, 303-331.
- Spector, P. A., & Brannick, M. T. (2011, in press). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*.