

ANALÝZA KATEGORIZOVANÝCH DAT V SOCIOLOGII, ÚKOL Č. 2
Petr Fučík

VÝSLEDKY Z SPSS:

Model s konstantou:

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-,016	,022	,542	1	,462	,984

Model se všemi proměnnými:

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 1 ORIGIN	2,613	,062	1755,550	1	,000	13,643
RACE	,762	,062	151,291	1	,000	2,142
CAMP	1,549	,063	609,770	1	,000	4,705
Constant	-2,757	,079	1231,175	1	,000	,063

a. Variable(s) entered on step 1: ORIGIN, RACE, CAMP.

pozn: použitá metoda: Enter

interpretace Exp(B) (odds ratios): nejsilnějším prediktorem preference umístění v severních či jižních vojenských táborech je geografický původ vojáků. Seveřané tedy logicky preferují severní tábory, běloši taktéž severní tábory a ti, kteří byli umístěni v severních táborech preferují silněji sever.

Signifikace testu dobré shody:

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	24,978	5	,000

Classification Table^a

		Predicted		
		preference tabora		Percentage Correct
Observed	sever	jih		
Step 1 preference tabora	sever	3475	576	85,8
	jih	1228	2757	69,2
Overall Percentage				77,6

a. The cut value is ,500

Model s interakcemi:

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	ORIGIN by RACE	,105	,131	,647	1	,421	1,111
	CAMP by ORIGIN	,601	,145	17,202	1	,000	1,825
	CAMP by RACE	-,018	,145	,015	1	,903	,983
	RACE	,710	,146	23,760	1	,000	2,033
	ORIGIN	2,132	,145	215,012	1	,000	8,429
	CAMP	1,224	,145	71,293	1	,000	3,401
	Constant	-2,464	,142	301,343	1	,000	,085

a. Variable(s) entered on step 1: ORIGIN * RACE , CAMP * ORIGIN , CAMP * RACE , RACE, ORIGIN, CAMP.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	,505	5	,992

Podle výsledků modelování vč. Interakcí je zřejmé, že změna modelu po jejich zařazení není signifikantní. Pouze v případě interakce umístění v táboře a geografického původu jde o signifikantní změnu modelu. Dle výsledků testu dobré shody je však nutné model odmítnout pro signifikantní odchylky mezi předikovanými a napozorovanými daty.

VÝSLEDKY ZE STATY:

MODEL KONSTANTY

Iteration 0: log likelihood = -5569.8597

Logit estimates	Number of obs =	8036
	LR chi2(0) =	0.00
	Prob > chi2 =	.
Log likelihood = -5569.8597	Pseudo R2 =	0.0000

pref	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
-----+-----					
_cons	-.0164265	.0223113	-0.74	0.462	-.0601558 .0273029

MODEL SE VŠEMI PROMĚNNÝMI

Logit estimates
Log likelihood = -4026.6051
Number of obs = 8036
LR chi2(3) = 3086.51
Prob > chi2 = 0.0000
Pseudo R2 = 0.2771

pref	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
camp	4.705148	.2950837	24.69	0.000	4.160927	5.320549
origin	13.64266	.8508743	41.90	0.000	12.07288	15.41656
race	2.141788	.1326232	12.30	0.000	1.897006	2.418157

BIC pro model s konstantou: -64168,018
BIC pro plný model: -3059,534

PREDIKCE HODNOT ZÁVISLE PROMĚNNÉ:

(provedeny na základě modelu se všemi proměnnými, pouze pro některé kombinace hodnot nezávisle proměnných)

logit: Predictions for pref

Pr(y=1|x): 0.0597 95% ci: (0.0516,0.0690)
Pr(y=0|x): 0.9403 95% ci: (0.9310,0.9484)

camp origin race
x= 0 0 0

logit: Predictions for pref

Pr(y=1|x): 0.8972 95% ci: (0.8844,0.9088)
Pr(y=0|x): 0.1028 95% ci: (0.0912,0.1156)

camp origin race
x= 1 1 1

černoch, pocházející z jihu, umístěný v jižním táboře bude s pravděpodobností 0,95 preferovat umístění v jižním táboře.

Seveřan, běloch, umístěný v severním táboře bude preferovat sever s pravděpodobností 0,90.

Běloch umístěný na severu, původem z jihu preferuje s 61% pravděpodobností jih.

Atd.