

Práce do kurzu SOC414 – Úkol č.2

Model 0, df=0

. logit pref_lok [w=freq]
(frequency weights assumed)

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

Tabulka č.1: Shrnutí logitového modelu 0

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

Model 1

. logit pref_lok rasa reg_puv souc_lok [w=freq]
(frequency weights assumed)

Iteration 0: log likelihood = -5569.8597
Iteration 1: log likelihood = -4086.4936
Iteration 2: log likelihood = -4027.6249
Iteration 3: log likelihood = -4026.6055
Iteration 4: log likelihood = -4026.6051

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

Tabulka č.2: Shrnutí logitového modelu 1

pref_lok	Coef.	Std. Err.	z	P> z	[95% Conf.	Interval]
rasa	7616412	.0619217	12.30	0.000	.6402768	.8830055
reg_puv	2.613202	.0623686	41.90	0.000	2.490961	2.735442
souc_lok	1.548657	.0627151	24.69	0.000	1.425738	1.671576
cons	-2.75682	.0785685	-35.09	0.000	-2.910812	-2.602829

. fitstat

Measures of Fit for logit of pref_lok
McFadden's Adj R2: 0.276
BIC': -3059.534

Model 2, se vzájemnými interakcemi proměnných rasa a region původu

. gen new1 = rasa* reg_puv
. logit pref_lok rasa reg_puv souc_lok new1 [w=freq]
(frequency weights assumed)

Iteration 0: log likelihood = -5569.8597
 Iteration 1: log likelihood = -4086.4629
 Iteration 2: log likelihood = -4027.6445
 Iteration 3: log likelihood = -4026.5826
 Iteration 4: log likelihood = -4026.582

Logit estimates
 Number of obs = 8036
 LR chi2(4) = 3086.56
 Prob > chi2 = 0.0000
 Log likelihood = -4026.582 Pseudo R2 = 0.2771

Tabulka č.3: Shrnutí logitového modelu 2

pref_lok	Coef.	Std. Err.	z	P> z 	[95% Conf.]	Interval
rasa	.7499211	.0824443	9.10	0.000	.5883332	.9115089
reg_puv	2.601823	.0817194	31.84	0.000	2.441656	2.76199
souc_lok	1.548204	.0627763	24.66	0.000	1.425165	1.671244
new1	.026637	.1239074	0.21	0.830	-.2162169	.269491
cons	-2.748653	.0872098	-31.52	0.000	-2.919581	-2.577725

. fitstat, saving (mod2)
 Measures of Fit for logit of pref_lok
 McFadden's Adj R2: 0.276
 BIC': -3050.589

Model 2, se vzájemnými interakcemi proměnných rasa a současná lokace

. gen new2 = rasa* souc_lok

. logit pref_lok rasa reg_puv souc_lok new2 [w=freq]
 (frequency weights assumed)

Iteration 0: log likelihood = -5569.8597
 Iteration 1: log likelihood = -4082.4322
 Iteration 2: log likelihood = -4023.7739
 Iteration 3: log likelihood = -4022.8103
 Iteration 4: log likelihood = -4022.8099

Logit estimates
 Number of obs = 8036
 LR chi2(4) = 3094.10
 Prob > chi2 = 0.0000
 Log likelihood = -4022.8099 Pseudo R2 = 0.2778

Tabulka č.4. Shrnutí logitového modelu 3

pref_lok	Coef.	Std. Err.	z	P> z 	[95% Conf.]	Interval
rasa	1.010216	.1096133	9.22	0.000	.7953779	1.225054
reg_puv	2.625313	.0624892	42.01	0.000	2.502836	2.74779
souc_lok	1.715656	.0877449	19.55	0.000	1.54368	1.887633
new2	-.3396039	.1232897	-2.75	0.006	-.5812473	-.0979605
cons	-2.888961	.0925044	-31.23	0.000	-3.070266	-2.707656

Measures of Fit for logit of pref_lok
 McFadden's Adj R2: 0.277
 BIC': -3058.133

Model 4, se vzájemnými interakcemi proměnných region původu a současná lokace

. gen new3 = reg_puv* souc_lok

. logit pref_lok rasa reg_puv souc_lok new3 [w=freq]
(frequency weights assumed)

Iteration 0: log likelihood = -5569.8597

Iteration 1: log likelihood = -4068.1458

Iteration 2: log likelihood = -4015.7105

Iteration 3: log likelihood = -4014.8474

Iteration 4: log likelihood = -4014.847

Logit estimates	Number of obs = 8036
	LR chi2(4) = 3110.03
	Prob > chi2 = 0.0000
Log likelihood = -4014.847	Pseudo R2 = 0.2792

Tabulka č.5: Shrnutí logitového modelu 4

pref_lok	Coef.	Std. Err.	z	P> z 	[95% Conf.	Interval]
rasa	.7426203	.0622675	11.93	0.000	.6205782	.8646624
reg_puv	2.180223	.1065738	20.46	0.000	1.971342	2.389103
souc_lok	1.219707	.0891185	13.69	0.000	1.045038	1.394376
new3	.5957729	.1216427	4.90	0.000	.3573576	.8341882
cons	-2.490725	.0921376	-27.03	0.000	-2.671312	-2.310139

. fitstat, saving (mod4)

Measures of Fit for logit of pref_lok

McFadden's Adj R2: 0.278
BIC': -3074.059

Interpretace dat na základě nejvhodnějšího modelu

Na základě údaje BIC' jsme vyhodnotili jako nejvhodnější model č.4. Následně jsme vypočítali koeficienty jednotlivých proměnných (viz tabulka č.6).

. listcoef, percent
logit (N=8036): Percentage Change in Odds

Odds of: 1 vs 0

Tabulka č.6: Koeficienty nezávislých proměnných logitového modelu 4

pref_lok	b	z	P> z 	%	%StdX	SDofX
rasa	0.74262	11.926	0.000	110.1	44.8	0.4988
reg_puv	2.18022	20.457	0.000	784.8	197.5	0.5000
souc_lok	1.21971	13.686	0.000	238.6	75.6	0.4616
new3	0.59577	4.898	0.000	81.4	33.5	0.4854

Všechny nezávislé proměnné vyšly jako statisticky signifikantní.

Bílý voják má oproti černochovi 45% šanci, že bude preferovat umístění v jižní jednotce.
Vojáci z jihu mají oproti severanům 198% šanci, že budou preferovat jižní jednotku.

Vojáci v současnosti umístění v jižní jednotce mají ve srovnání s vojáky na severu 76% šanci preference jižní jednotky.

Predikce pravděpodobnosti preferencí u jednotlivých typů vojáků (danou funkci demonstrujeme pouze u vybraných tří typů vojáků)

. prvalue, x(rasa=1 reg_puv=0)

logit: Predictions for pref_lok

Pr(y=1|x): 0.2710 95% ci: (0.2470,0.2965)
Pr(y=0|x): 0.7290 95% ci: (0.7035,0.7530)

rasa reg_puv souc_lok new3
x= 1 0 .5 .25

Bílý voják původem ze severu má 73% šanci, že bude preferovat umístění jednotky na severu.

. prvalue, x(rasa=0 reg_puv=1)

logit: Predictions for pref_lok

Pr(y=1|x): 0.6102 95% ci: (0.5830,0.6368)
Pr(y=0|x): 0.3898 95% ci: (0.3632,0.4170)

rasa reg_puv souc_lok new3
x= 0 1 .5 .25

Černý voják z jihu má 61% šanci preferovat umístění v jižní jednotce.

. prvalue, x(rasa=0 reg_puv=1 souc_lok=0)

logit: Predictions for pref_lok

Pr(y=1|x): 0.4597 95% ci: (0.4309,0.4888)
Pr(y=0|x): 0.5403 95% ci: (0.5112,0.5691)

rasa reg_puv souc_lok new3
x= 0 1 0 .25

Černý voják, jižan, umístěný v severní jednotce, má 54% šanci preferovat severní jednotku.