

НИУ ВШЭ, ф-т социологии, 2-й курс
«Анализ социологических данных – 1»
Комиссия. 27 сентября 2012 г.

Часть I. В этой части экзамена вам нужно ответить на следующие вопросы.

1. Для сравнения средних значений в группах используется дисперсионный анализ. Почему нельзя сравнивать средние попарно с помощью нескольких t-тестов?
2. Что такое среднее арифметическое значение и медиана? Чем они отличаются друг от друга?
3. Что такое метод наименьших квадратов (МНК, OLS) и для чего он используется?
4. В случае применения обычной линейной регрессии, должна ли зависимая переменная быть нормально распределена?
5. Для решения каких социологических задач могут использоваться логарифмически-линейные модели? Приведите примеры.
6. Что такое гистограмма? Приведите пример – нарисуйте от руки любую гистограмму.
7. Чем отличаются порядковая и мультиномиальная логистическая регрессия?
8. При моделировании зависимости дохода от возраста в регрессионную модель часто добавляют возраст в квадрате. Зачем?
9. Объясните значение понятия «статистический контроль».
10. Почему эксперименты являются «золотым стандартом» причинно-следственного анализа в социальных науках?

Часть II. В этой части вы должны подробно проинтерпретировать представленные результаты модели мультиномиальной логистической регрессии.

Зависимая переменная: pomreg – тип программы, на которой учатся ученики в американских школах (общая – general, академическая – academic и профессиональная – vocational). Базовая категория – академическая программа.

Независимые переменные: ses – социально-экономический статус семьи (категориальная переменная, три категории, от 1 (низкий) до 3 (высокий)), write – умение ученика писать по-английски (оценка в баллах).

Подробно охарактеризуйте зависимость выбора типа программы от социально-экономического статуса учеников и их умения писать по-английски.

Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	254.986			
Final	206.756	48.230	6	.000

Pseudo R-Square

Cox and Snell	.214
Nagelkerke	.246
McFadden	.118

Parameter Estimates

type of program ^a		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B)	
								Lower Bound	Upper Bound
general	Intercept	1.689	1.227	1.896	1	.169			
	write	-.058	.021	7.320	1	.007	.944	.905	.984
	[ses=1]	1.163	.514	5.114	1	.024	3.199	1.168	8.764
	[ses=2]	.630	.465	1.833	1	.176	1.877	.754	4.669
	[ses=3]	0 ^b	.	.	0
vocation	Intercept	4.236	1.205	12.361	1	.000			
	write	-.114	.022	26.139	1	.000	.893	.855	.932
	[ses=1]	.983	.596	2.722	1	.099	2.672	.831	8.585
	[ses=2]	1.274	.511	6.214	1	.013	3.575	1.313	9.736
	[ses=3]	0 ^b	.	.	0

a. The reference category is: academic.

b. This parameter is set to zero because it is redundant.