**Формулы для задач**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№** | **Задание** | **Для линейной LPM** | **Probit** | **Logit** |
|  | Прогноз для **человека №Х** | **Z = СУММПРОИЗВ**  (1) Выписываем все регрессоры  (2) Выписываем все значения регрессоров для конкретного человека  (3) Делаем дорасчет (где нужно берем квадрат, где нужно логарифм. Для значения I(WA2) у людей пропущен этот показатель. Но это просто WA в квадрате -те его мы вычисляем сами, а не берем из таблички людей. FAMINC от него берез логарифм просто. Для человека константу берем как 1 (тк плюс константа, ее не нужно ни с чем перемножать).  (4) Прогноз вычисляем путем перемножения дорасчитанных характеристик с регрессорами из регрессии. | **Z = СУММПРОИЗВ**  **Прогноз** = (НОРМ.СТ.РАСП(Z;ИСТИНА) | **Z = СУММПРОИЗВ**  **Прогноз =** 1/(1+EXP(-Z))  (1) Ищем Z для женщины - перемножаем показатели регрессии с показателями женщины (СУММПРОИЗВ).  (2) Ищем Fzlogit(прогноз) по формуле 1/(1+EXP(-Z)) |
|  | **Совместная значимость** коэффициентов (втч. если возраст и возраст в квадрате) | **=F.ОБР.ПХ(B%;q;n-k-1)**  **q-** количество исключенных регрессоров  **K-** количество регрессоров без константы  **N –** количество наблюдений | LR(Тестовая) = 2\*(Log1-Log2)  Критическое =ХИ2.ОБР.ПХ  (1) Берем Log Likelihood для первой модели (это log 1).  (2) Выбираем модель, где тестируемого показателя нет и фиксируем ее Log Likelihood (это log 2).  (3) Далее вычисляем **тестовую статистику** по формуле LR = 2\*(Log1-Log2).  (4) Далее **критическое значение** по формуле ХИ2.ОБР.ПХ(% это уровень значимости из условия; количество тестируемых показателей это возраст/образование и тд). ХИ2ОБР - без точек внутри!! ХИ2ОБР или ХИ2.ОБР.ПХ это одно и то же.  (5) **Значимость** - сравниваем тестовую статистику и критическое значение. Если тестовое больше критического, то ЗНАЧИМО. Если тестовое меньше критического - НЕЗНАЧИМО. | |
|  | **Значимость влияния одного показателя (dep/married)** | Тестовая= Estimate / Std. Error  Критическое =  **Стьюдент обр 2х(%;n-k-1)**  (в данном случае в k константа не входит, поэтому n-k-1)  **N –** количество наблюдений  **Вывод о значимости: =ЕСЛИ (ABS(t расчетное))> ABS(t crit); “значимо»; «не значимо»**  При сравнении «t тест (расчетного)» с «t критическим» говорим, что либо влияния нет – либо наоборот, что оказывает статистически значимое влияние. | Тестовая= Estimate / Std. Error  Критическое= НОРМ.СТ.ОБР(1-%уровень значимости/2)  (1) Вычисляем **тестовую статистику** - делим значения для dep - Estimate на Std. Error. Отрицательный знак сохраняем и пишем в ответ.  (2) **Критическое значение** - НОРМ.СТ.ОБР(1-уровень значимости/2). На два делим, тк у колокола нормального распределения 2 хвоста.  (3) **Значимость**: сравниваем тестовую статистику (по модулю без отриц. знака) и критическое значение. Если тестовое больше критического, то ЗНАЧИМО. Если тестовое меньше критического - НЕЗНАЧИМО. | |
|  | **Значимость всей регрессии** | **=F.ОБР.ПХ(B%;q;n-k-1)**  **q-** все регрессоры  **K-** количество регрессоров без константы  **N –** количество наблюдений | Тестовая = 2\*(Log1-Log2)  Критическое=ХИ2.ОБР.ПХ  (1) Вычисляем тестовую статистику (2\*(LogL-LogL0).  (2) Критическое значение ХИ2.ОБР.ПХ(% это уровень значимости из условия; количество тестируемых показателей без константы).  (3) Значимость - сравниваем тестовую статистику и критическое. Если тестовое больше критического, то ЗНАЧИМО. Если тестовое меньше критического - НЕЗНАЧИМО. | |
|  | **Предельный эффект от «возраста»** для человека №Х | Это просто коэффициент из таблицы (датасета-регрессии для «возраста/образования» и тд.). | Предельный эффект =НОРМ.СТ.РАСП(Z;ЛОЖЬ)\*b  **(**1) Ищем Z для женщины 1 - перемножаем показатели регрессии с показателями 1 женщины (СУММПРОИЗВ).  (2) Ищем PZprobit по формуле НОРМ.СТ.РАСП(Z;ЛОЖЬ)  (3) **Предельный эффект** для 1 женщины по формуле PZprobit\*значение коэффициента (b) (значение возраста из регрессии (датасета), не характеристика женщины) | **Z=** **СУММПРОИЗВ**  **pz Logit = EXP(Z)/((1+EXP(Z))^2)**  **Предельный эффект = (b)\*pz Logit**  (b) - значение возраста из регрессии (датасета), не характеристика женщины |
|  | **Cox & Snell** | =1-EXP(-2\*(LogL - LogL0)/число наблюдений) | | |
|  | **Nagelkerke/Cragg & Uhler** | =Cox&Snel /(1-EXP(2\*LogL0/число наблюдений)) | | |
|  | **PseudoR2** | =1-(Log L/Log L0) | | |
|  | **PseudoR2.adj** | =1-(Log L- количество переменных в КОНКРЕТНОЙ регрессии)/Log L0 | | |

* Тестовую статистику берем по модулю для сравнения с критическим значением
* Pseudo – предпочтительнее самое больше значение, Акаике и тд – меньшее.