**Часть I.**

**Вопрос 1.** Предпосылка об отсутствии систематической ошибки в модели означает, что

1. Да
2. Нет

**Вопрос 2.**  Стандартные ошибки в форме Уайта помогают в случае гетероскедастичности сделать оценки дисперсий коэффициентов состоятельными

1. Да
2. Нет

**Вопрос 3.** Незначимость всех коэффициентов регрессии обязательно свидетельствует о наличии мультиколлинеарности

1. Да
2. Нет

**Вопрос 4.** После применения МНК к модели сумма ESS+RSS

1. Равна TSS
2. Не равна TSS

**Вопрос 5.** Ошибки в измерении зависимой переменной делают оценки коэффициентов модели несостоятельными

1. Да
2. Нет

**Вопрос 6.** Если выполнены все предпосылки теоремы Гаусса-Маркова, но остатки модели не подчиняются нормальному закону распределения, то оценки коэффициентов регрессии являются

1. Несмещенными
2. Смещенными

**Вопрос 7.** Индексы вздутия дисперсии (VIF) в случае отсутствия проблемы мультиколлинеарности лежат в интервале от 0 до 1

1. Да
2. Нет

**Вопрос 8.** Нулевая гипотеза в тесте Дарбина-Уотсона состоит в наличии автокорреляции

1. Да
2. Нет

**Вопрос 9.** Если в модель добавили незначимый фактор, то коэффициент детерминации может

1. Упасть
2. Вырасти

**Вопрос 10.** Тест на автокорреляцию Бройша-Годфри применим при диагностике автокорреляции третьего порядка

1. Да
2. Нет

**Задача 1.**

На основании опроса была оценена следующая модель:

,

где

опыт работы;

– количество лет обучения;

– бинарная переменная (1 – в случае наличия супруга/супруги, 0 – иначе);

– бинарная переменная (1 – в случае принадлежности к негроидной расе, 0 – иначе).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Множественный R | 0.497 |  |  |  |  |  |
| R-квадрат | **В7** |  |  |  |  |  |
| Нормированный R-квадрат | 0.219 |  |  |  |  |  |
| Стандартная ошибка | **В6** |  |  |  |  |  |
| Наблюдения | **В2** |  |  |  |  |  |
|  |  |  |  |  |  |  |
| Дисперсионный анализ | |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Значимость F* |  |
| Регрессия | **В1** | 5.993 | 1.199 | **В5** | 0.000 |  |
| Остаток | 134 | 18.240 | 0.136 |  |  |  |
| Итого | **В3** | **В4** |  |  |  |  |
|  |  |  |  |  |  |  |
|  | *Коэффициенты* | *Стандартная ошибка* | *t-статистика* | *P-Значение* | *Нижние 95%* | *Верхние 95%* |
| Y-пересечение | 4.529 | 0.331 | 13.688 | 0.000 | 3.874 | 5.183 |
| exper | 0.090 | 0.037 | 2.419 | 0.017 | 0.016 | 0.164 |
| exper2 | -0.003 | 0.002 | -1.790 | 0.076 | -0.006 | 0.000 |
| married | 0.240 | 0.079 | 3.045 | 0.003 | **В8** | **В9** |
| educ | 0.078 | 0.017 | **В10** | 0.000 | 0.045 | 0.111 |
| black | 0.073 | 0.171 | 0.424 | 0.672 | -0.266 | 0.411 |

Вставьте пропущенные числа на месте пропусков (округляйте ответ до 3 знака после запятой):

**В1=**

**В2=**

**В3=**

**В4=**

**В5=**

**В6=**

**В7=**

**В8=**

**В9=**

**В10=**

**Задача 2**

На основании данных по ценам на квартиры в Москве были построены модели для логарифма цены (в зависимости от общей площади (расстояния до метро ( мин), расстояния до центра города ( км), этажа, на котором расположена квартира (1 – если квартира не на первом и последнем этажах, 0 – иначе). Модели были оценены на 5 разных выборках, результаты представлены в таблице.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **1** | **2** | **3** | **4** | **5** |
|  | по всей выборке | для квартир в транспортной доступности от метро | для квартир в пешей доступности от метро | для 150 квартир с наименьшей общей площадью | для 150 квартир с наибольшей общей площадью |
| **totsp** | 0.0155\*\*\* | 0.0148\*\*\* | 0.0163\*\*\* | 0.0179\*\*\* | 0.0139\*\*\* |
| **metrdist** | -0.00858\*\*\* | -0.0169\*\*\* | -0.00566\*\* | -0.0108\*\*\* | -0.0077 |
| **dist** | -0.0267\*\*\* | -0.0186\*\*\* | -0.0253\*\*\* | -0.0150\*\*\* | -0.0350\*\*\* |
| **floor** | 0.0419\*\* | 0.0633\* | 0.0224 | 0.0225 | 0.0228 |
| **Constant** | 3.980\*\*\* | 3.926\*\*\* | 3.929\*\*\* | 3.719\*\*\* | 4.224\*\*\* |
| Observations | 460 | 145 | 315 | 150 | 150 |
| R-squared | 0.693 | 0.684 | 0.723 | 0.328 | 0.520 |
| RSS | 15.120 | 4.503 | 9.408 | 2.163 | 8.545 |

\*-значимость на 10%, \*\* - значимость на 5%, \*\*\* - значимость на 1%.

1. Для 1-ой модели (по всей выборке) проинтерпретируйте коэффициент при переменной .
2. Определите на 5%-ом уровне значимости, можно ли использовать одну модель для квартир, находящихся в пешей доступности от метро, и квартир, находящихся в транспортной доступности (выпишите H0, H1, найдите расчетную статистику, укажите ее распределение, найдите критическую статистику, сделайте выводы);
3. Исследователь предположил, что дисперсия ошибок модели возрастает с увеличением площади квартиры. Проверьте, есть ли в модели гетероскедастичность на 10% уровне значимости на основании соответствующего теста (выпишите H0, H1, найдите расчетную статистику, укажите ее распределение, найдите критическую статистику, сделайте выводы).

**Задача 3**

По ежемесячным данным (146 наблюдений) была оценена зависимость потребительских кредитов, выданных домашним хозяйствам РФ (), от ставки процента по кредитам () и индекса потребительских цен (

*.*

Известно, что

1. На 1%-ом уровне значимости проверьте гипотезу об адекватности регрессии (выпишите H0, H1, найдите расчетную статистику, укажите ее распределение, найдите критическую статистику, сделайте выводы).
2. Проведите тест Дарбина-Уотсона на 5% уровне значимости (выпишите H0, H1, найдите расчетную статистику, критические значения, сделайте выводы).
3. Кроме того, была оценена следующая модель:

Проведите тест Бройша-Годфри на 5% уровне значимости (выпишите H0, H1, найдите расчетную статистику, укажите ее распределение, найдите критическую статистику, сделайте выводы).

**Задача 4**

Домохозяйка Глаша очень любит читать романы Л.Н. Толстого и смотреть сериалы. Ее сын Петя учится на 3 курсе ВШЭ. Последние 30 дней он записывал, сколько Глаша прочитала страниц «Анны Карениной» () и серий «Доктора Хауса» посмотрела (). На основании этих наблюдений при помощи МНК он оценил следующую модель:

Оценка ковариационной матрицы коэффициентов выглядит следующим образом:

)

Оценка дисперсии ошибок

1. Сколько Глаша прочитает страниц романа, если посмотрит 10 серий «Доктора Хауса»?
2. Постройте 95%-ый доверительный интервал для E( - ожидаемой величины прочитанных страниц при 10 просмотренных сериях сериала.
3. Постройте 95%-ый доверительный интервал для конкретной величины прочитанных страниц при 10 просмотренных сериях сериала.

**III часть**

**Вопрос 1**

Опишите основную идею метода наименьших квадратов для парной регрессии: выпишите целевую функцию, систему нормальных уравнений, оценки коэффициентов, оценки дисперсий коэффициентов.

**Вопрос 2**

Сформулируйте теорему Гаусса-Маркова.

**Вопрос 3.**

Опишите тест Уайта