

## Часть 1. Тест.

**Вопрос 1 ♣** В теореме Гаусса-Маркова предполагается, что ошибки

- ☐ А имеют ненулевое математическое ожидание и единичную дисперсию  
☐ Б имеют ненулевое математическое ожидание и неединичную дисперсию  
☐ В имеют нулевое математическое ожидание и единичную дисперсию  
☐ С Нет верного ответа.

**Вопрос 2 ♣** После применения МНК к модели  $y_i = \beta x_i + \varepsilon_i$  сумма  $\sum \hat{\varepsilon}_i^2$

- ☐ А не обязательно равна нулю  
☐ Б равна нулю

**Вопрос 3 ♣** При автокорреляции МНК-оценки коэффициентов являются

- ☐ А несмещёнными  
☐ Б смещёнными

**Вопрос 4 ♣** В случае мультиколлинеарности оценки дисперсий коэффициентов являются

- ☐ А в среднем заниженными  
☐ Б несмещёнными  
☐ С в среднем завышенными

**Вопрос 5 ♣** Двухшаговый метод наименьших квадратов — это стандартный способ борьбы с

- ☐ А гетероскедастичностью  
☐ Б автокорреляцией  
☐ С эндогенностью  
☐ Д Нет верного ответа.

**Вопрос 6 ♣** Нестрогая мультиколлинеарность — это одно из нарушений предпосылок теоремы Гаусса-Маркова

- ☐ А верно  
☐ Б неверно

**Вопрос 7 ♣** Нулевой гипотезой в тесте Уайта является гипотеза о

- ☐ А гомоскедастичности  
☐ Б наличии автокорреляции  
☐ С гетероскедастичности  
☐ Д отсутствии автокорреляции

**Вопрос 8 ♣** Тест Дарбина-Уотсона в регрессии без свободного члена

- ☐ А применим  
☐ Б неприменим

**Вопрос 9 ♣** Величина  $R_{adj}^2$  показывает, какую долю дисперсии зависимой переменной объясняют использованные регрессоры

- ☐ А неверно  
☐ Б верно

**Вопрос 10 ♣** При условной гетероскедастичности оценки коэффициентов

- ☐ А состоятельны  
☐ Б не состоятельны

## Часть 2. Задачи.

1. На основании опроса была оценена следующая модель:

$$\ln(wage_i) = \beta_1 + \beta_2 exper_i + \beta_3 exper_i^2 + \beta_4 married_i + \beta_5 educ_i + \beta_6 black_i + \varepsilon_i$$

где:

- $wage_i$  — величина заработной платы в долларах
- $exper_i$  — опыт работы в годах
- $educ_i$  — количество лет обучения
- $married_i$  — наличие супруга/супруги (1 — есть, 0 — нет)
- $black_i$  — принадлежность к негроидной расе (1 — да, 0 — нет)

Показатель	Значение
$R^2$	0.280
Скорректированный $R^2$	0.277
Стандартная ошибка регрессии	<b>B6</b>
Количество наблюдений	1271

Результаты дисперсионного анализа:

	df	SS	MS	F	P-значение
Регрессия	<b>B1</b>	<b>B4</b>	14.670	<b>B5</b>	0.000
Остаток	<b>B2</b>	188.951	0.149		
Итого	<b>B3</b>	262.301			

Коэффициент	Оценка	$se(\hat{\beta})$	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Константа	<b>B7</b>	0.110	42.019	0.000	<b>B8</b>	<b>B9</b>
$exper$	0.076	<b>B10</b>	7.081	0.000	0.055	0.097
$exper^2$	-0.002	0.000	-3.898	0.000	-0.003	-0.001
$married$	0.024	0.027	0.905	0.366	-0.028	0.076
$educ$	0.087	0.006	15.646	0.000	0.076	0.098
$black$	-0.229	0.025	-9.267	0.000	-0.277	-0.180

Найдите пропущенные числа **B1–B10**.

Ответ округляйте до 3-х знаков после запятой. Кратко поясняйте, например, формулой, как были получены результаты.

2. Туристическое агентство «Необыкновенные путешествия» в рамках программы импортозамещения продвигает новое направление пляжного отдыха — землю Франца-Иосифа. Недавно с отдыха вернулась первая партия из 254 туристов. Исследовательский отдел компании оценил модель полезности, которую туристы получили от отдыха:

$$\ln u_i = 1 + 2 \ln duration_i + 3 \ln weight_i + 4 \ln bears_i + \varepsilon_i$$

где:

- $u_i$  — полезность туриста в улыбках
- $duration_i$  — продолжительность отдыха в днях
- $weight_i$  — вес сухпайка в кг, выданного туристу по прилёту
- $bears_i$  — популяция белых медведей в радиусе 10 километров от пляжа

Выборка	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	RSS	$R^2$	N
A. Все вернувшиеся с отдыха	132.5*	9.7**	78.7**	1.7	972	0.64	254
B. Тур «На пляже с белыми медведями»	34.8**	5.8***	97.3**	-7.2**	234	0.73	67
C. Тур «Ultra All Inclusive в Заполярье»	139.7*	19.1**	0.7	1.9*	115	0.81	111
D. Тур «На айсберге в шезлонге»	97.2**	-11.1*	103.4*	5.8*	311	0.67	76
E. Жившие близко к базе полярников	267.1*	14.8**	50.9**	1.7*	467	0.48	100
F. Жившие далеко от базы полярников	3.8***	-5.8**	120.8*	-3.8**	112	0.84	100

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

- Для модели по всей выборке (выборка A) проинтерпретируйте коэффициент  $\hat{\beta}_2$
- Определите на 5%-ом уровне значимости, можно ли использовать общую модель для всех трёх туров фирмы (выборки B, C и D)
- Есть предположение, что удалённость от базы полярников влияет на дисперсию ошибок. Проверьте, верно ли это, на 10%-ом уровне значимости на основании соответствующего теста

При проверке гипотез: выпишите  $H_0$ ,  $H_a$ , найдите значение тестовой статистики, укажите её распределение, найдите критическое значение, сделайте выводы

3. Один из туристов каждый день в течение 146 дней отдыха измерял характеристики своего айсберга. По возвращении домой он построил модель зависимости высоты айсберга от объясняющих переменных:

$$\widehat{height}_t = 158 - 3.2temp_t - 1.8collisions_t + 0.4bears_t, TSS = 35578$$

где:

- $height_t$  — высота айсберга над уровнем моря в метрах
- $temp_t$  — температура воздуха по Цельсию
- $collisions_t$  — число столкновений с другими айсбергами
- $bears_t$  — число приплывших в гости в течение дня белых медведей

Известно, что  $\sum_{t=2}^{146} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2 = 9882$ ,  $\sum_{t=1}^{146} \hat{\varepsilon}_t^2 = 21089$ ,  $\sum_{t=2}^{146} |\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}| = 3617$ ,  $\sum_{t=1}^{146} |\hat{\varepsilon}_t| = 6382$ .

Кроме того, была оценена вспомогательная модель для остатков исходной модели:

$$\hat{\varepsilon}_t = 34.57 + 0.19temp_t - 0.71collisions_t + 0.24bears_t + 0.91\hat{\varepsilon}_{t-1} - 0.67\hat{\varepsilon}_{t-2} + 0.51\hat{\varepsilon}_{t-3}, R^2 = 0.37$$

- а) На 10%-ом уровне значимости проверьте гипотезу об адекватности исходной регрессии
- б) Проверьте наличие автокорреляции на 5%-ом уровне значимости при помощи теста Дарбина-Уотсона
- в) Проведите тест Бройша-Годфри на 1%-ом уровне значимости

При проверке гипотез: выпишите  $H_0$ ,  $H_a$ , найдите значение тестовой статистики, укажите её распределение, найдите критическое значение, сделайте выводы

4. Джон Сноу живёт на земле Франца-Иосифа и каждый день продаёт отдыхающим спойлеры к «Игре Престолов». После 40 дней, он установил, что выручка в тысячах долларов от продажи спойлеров,  $profit_t$ , определяется температурой в градусах Цельсия,  $temp_t$ :

$$\widehat{profit}_t = 10 - 2temp_t$$

Оценка ковариационной матрицы коэффициентов  $\widehat{\text{Var}}(\hat{\beta}) = \begin{pmatrix} 2 & -0.2 \\ -0.2 & 0.5 \end{pmatrix}$

Оценка дисперсии ошибок равна  $\hat{\sigma}^2 = 76$ .

Сегодня над Землёй Франца-Иосифа  $-20^\circ$ .

- а) Постройте точечный прогноз выручка от продажи спойлеров
- б) Постройте 95%-ый доверительный интервал для  $E(profit_t | temp_t = -20)$ , ожидаемой выручки от продажи спойлеров
- в) Постройте 95%-ый предиктивный интервал для фактической выручки от продажи спойлеров

### Часть 3. Теоретические вопросы

5. Дайте определение мультиколлинеарности и опишите основные методы её диагностики.
6. Дайте определение процесса стационарного в узком смысле. Приведите пример процесса авторегрессии первого порядка,  $AR(1)$ . Посчитайте для него математическое ожидание и дисперсию.
7. Опишите взвешенный МНК для случая известных дисперсий случайных ошибок: сформулируйте алгоритм получения оценок. Какие проблемы решает взвешенный МНК?