Экзамен по эконометрике 2014.

1. Регрессионная модель задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = (\beta_1, \beta_2, \beta_3)'$. Известно, что $\mathbb{E}(\varepsilon) = 0$ и $\mathrm{Var}(\varepsilon) = \sigma^2 \cdot I$. Известно также, что

$$y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Для удобства расчетов приведены матрицы

$$X'X = \begin{pmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ и } (X'X)^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}.$$

   (a) Найдите вектор МНК-оценок коэффициентов $\hat{\beta}$.

   (b) Найдите несмещенную оценку для неизвестного параметра σ^2.

   (c) Проверьте гипотезу $\beta_2 = 0$ против альтернативной о неравенстве на уровне значимости 5%

2. По данным о пассажирах Титаника оценивается логит-модель. Зависимая переменная `survived` равна 1, если пассажир выжил. Объясняющая переменная `sexmale` равна 1 для мужчин.

	Model 1
(Intercept)	1.92^{***}
	(0.28)
age	-0.01
	(0.01)
sexmale	-2.84^{***}
	(0.21)
AIC	633.45
BIC	646.80
Log Likelihood	-313.72
Deviance	627.45
Num. obs.	633

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Таблица 1: Statistical models

   (a) Оцените вероятность выжить для женщины 20 лет

   (b) Оцените предельный эффект увеличения возраста для женщины 20 лет

   (c) С помощью какого метода оценивается логит-модель? Каким образом при этом получаются оценки стандартных ошибок коэффициентов?

3. Теорема Гаусса-Маркова.

   (a) Аккуратно сформулируйте теорему Гаусса-Маркова для нестохастических регрессоров.

   (b) Поясните каждое из свойств оценок, фигурирующих в теореме.

   (c) Как меняются свойства оценок МНК при нарушении предпосылки теоремы о том, что дисперсия ε_i постоянна?

4. Рассмотрим временной ряд, описываемый MA(2) моделью,

$$y_t = \gamma + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2},$$

   где ε_t — белый шум с $\mathrm{Var}(\varepsilon_t) = \sigma^2$.

   (a) Является ли данный процесс стационарным? Что такое стационарный процесс?

   (b) Найдите автокорреляционную функцию данного процесса, $\rho(k) = \mathrm{Corr}(y_t, y_{t-k})$.

   (c) Выпишите функцию правдоподобия для данной модели в предположении нормальности ε_t.

5. Рассмотрите модель $y_i = \beta x_i + \varepsilon_i$. Предположим, что все предпосылки классической линейной регрессионной модели выполнены. Модель оценивается с помощью МНК и получается оценка $\hat{\beta}_{OLS}$. В условиях мультиколлинеарности для снижения дисперсии оценки $\hat{\beta}$ можно применять ряд методов, например, алгоритм «ridge regression». Он состоит в том, что при некотором фиксированном $\lambda \geq 0$ минимизируется по $\hat{\beta}$ величина

$$Q(\hat{\beta}) = \sum_i (y_i - \hat{\beta} x_i)^2 + \lambda \hat{\beta}^2$$

   (a) Как выглядит МНК оценка $\hat{\beta}_{OLS}$?

   (b) Как выглядит оценка методом «ridge regression», $\hat{\beta}_{RR}$?

   (c) Верно ли, что оценка $\hat{\beta}_{RR}$ является несмещенной только при $\lambda = 0$?

   (d) (*) Верно ли, что всегда найдется такое λ, что среднеквадратичная ошибка оценки $\hat{\beta}_{RR}$ будет меньше, т.е. $\mathbb{E}((\hat{\beta}_{RR} - \beta)^2) < \mathbb{E}((\hat{\beta}_{OLS} - \beta)^2)$?