Тест

Задачи

1. Рассеянный исследователь Вовочка 555 дней замерял своё потребление шоколада и число решённых задач по эконометрике. Вовочка оценил по своим данным парную регрессию числа решенных задач на потребление шоколада (регрессию с константой), но потерял все результаты вычислений и не справится без вашей помощи!

   а) Вовочка запомнил, что 95%-ый доверительный интервал для коэффициента при шоколаде был от 1.72 до 8.28. Помогите ему восстановить оценку $\hat{\beta}_{choc}$ и оценку её стандартного отклонения.

   Ошибки в модели для этого и следующего пунктов считайте нормальными.

   б) Помогите Вовочке проверить значимость $\hat{\beta}_{choc}$ на 10% уровне значимости.

   в) Можно ли было бы считать полученные МНК-оценки коэффициентов несмещёнными и эффективными в случае равномерных от -5 до 5 ошибок? Почему?

   г) Можно ли было бы считать полученные МНК-оценки коэффициентов несмещёнными и эффективными в случае равномерных от 0 до 5 ошибок? Почему?

2. Рассмотрим уравнение линейной регрессии $Y_i = \beta X_i + u_i$. Все предпосылки теоремы Гаусса-Маркова выполнены.

   а) Найдите МНК оценку коэффициента $\hat{\beta}$.

   б) Проверьте, является ли эта оценка несмещенной.

   в) Выведите формулу для несмещённой оценки дисперсии этой оценки.

   г) По выборке оказалось, что $\hat{\beta} = 2.25$ и $se(\hat{\beta}) = 0.2$. Проинтерпретируйте значение оценки коэффициента.

   д) Выведите формулу оценки дисперсии для ошибки прогноза $\hat{Y}_{N+1}$.

3. Для модели $X_i = \beta_0 + \beta_1 Y_i + u_i$ известна МНК-оценка коэффициента $\hat{\beta}_1 = -1$. Также для данной регрессии известны $N = 102$, $\sum(Y_i - \bar{Y})^2 = 10$ и $TSS = 200$.

   а) Найдите коэффициент детерминации R^2 для этой модели.

   б) Найдите оценку дисперсии оценки коэффициента $\hat{\beta}_1$.

   в) Для регрессии $\hat{Y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 X_i$ найдите оценку $\hat{\alpha}_1$.

   г) Найдите выборочный коэффициент корреляции $\widehat{\mathbb{C}\text{orr}}(X, Y)$.

1. Решения

1. Вместо t_{553}-распределения можно использовать нормальное $\mathcal{N}(0;1)$.

   а) [2] Критическое значение равно $t_{crit} = 1.96$.
   Отсюда находим $\hat{\beta} = (8.28 + 1.72)/2 = 5$ и $se(\hat{\beta}) = (8.28 - 1.72)/(2 \cdot 1.96) = 1.67$.

   б) [1] $t_{obs} = 5/1.67 = 3$. Таблицы не нужны, достаточно помнить, что при уровне значимости $\alpha = 0.05$ критическое значение равно 1.96. При более высоком уровне значимости критическое значение падает. Значит H_0 отвергается.

   в) [1] Ожидание ошибки равно нулю, дисперсия постоянна, значит условия теоремы Гаусса-Маркова выполнены. Обе оценки являются несмещёнными и эффективными среди линейных несмещённых оценок.

   г) [2] Ожидание ошибки равно 2.5, дисперсия постоянна, значит условия теоремы Гаусса-Маркова нарушены. Однако при переносе 2.5 из ошибки в константу нарушение исчезает. Оценка наклона: несмещена и эффективна среди линейных несмещённых оценок.

   Оценка константы: смещена на 2.5, поэтому оценка не лежит в классе линейных несмещенных оценок, и говорить об её эффективности в этом классе бессмысленно. При этом дисперсия оценки константы равна дисперсии эффективной оценки.

2. Кратко:

   а) [1] $\hat{\beta} = \frac{\sum X_i Y_i}{\sum X_i^2}$

   б) [1] $\mathbb{E}(\hat{\beta}) = \beta$

   в) [2] $\widehat{\mathbb{V}\mathrm{ar}}(\hat{\beta}) = \frac{\hat{\sigma}_u^2}{\sum X_i^2}$

   г) [1] $t_{obs} = 11.25$, коэффициент значимо отличен от нуля. Зависимая переменная в среднем в 2.25 раз больше регрессора.

   д) [2] $\widehat{\mathbb{V}\mathrm{ar}}(\hat{Y}_{N+1} - Y_{N+1}) = \hat{\sigma}_u^2 \left(1 + \frac{X_{N+1}^2}{\sum X_i^2}\right)$

3. Обратите внимание, что Y_i является регрессором.

   а) [2] $\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (Y_i - \bar{Y})^2}$.
   Следовательно, $\sum (X_i - \bar{X})(Y_i - \bar{Y}) = -10$.
   Решаем одним махом г) и а)!

$$\widehat{\mathbb{C}\mathrm{orr}}(X, Y) = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{10 \cdot 200}} = -\sqrt{5}/10$$

   Отсюда:

$$R^2 = 5/100 = 1/20$$

   б) [2] $RSS = 0.95 \cdot 200 = 190$. Отсюда $\hat{\sigma}_u^2 = 190/100 = 1.9$ и $se^2(\hat{\beta}_1) = 1.9/10 = 0.19$.

   в) [2] В узких кругах широко известно, что корреляции по модулю равна среднему геометрическому оценок в прямой и обратной моделях.

$$R^2 = \hat{\alpha}_1 \cdot \hat{\beta}_1$$

   Следовательно, $\hat{\alpha}_1 = -1/20$.

г) [1] уже решили!