Семинар 22.

Решение.

Рассмотрим применение тестов W (тест Вальда), LR (тест отношения правдоподобия) и LM (тест множителей Лагранжа) для тестирования гипотез о параметрах модели.

Пусть требуется протестировать систему ограничений относительно вектора неизвестных параметров

$$H_0 : \begin{cases} g_1(\theta) = 0 \\ g_2(\theta) = 0 \\ \dots \\ g_r(\theta) = 0 \end{cases}$$

где $g_i(\theta)$ — функция, которая задаёт i-ое ограничение на вектор параметров θ, $i = 1, \dots, r$.

$$\frac{\partial g}{\partial \theta^T} = \begin{pmatrix} \partial g_1/\partial\theta^T \\ \partial g_2/\partial\theta^T \\ \vdots \\ \partial g_r/\partial\theta^T \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial\theta_1} & \frac{\partial g_1}{\partial\theta_2} & \dots & \frac{\partial g_1}{\partial\theta_k} \\ \frac{\partial g_2}{\partial\theta_1} & \frac{\partial g_2}{\partial\theta_2} & \dots & \frac{\partial g_2}{\partial\theta_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_r}{\partial\theta_1} & \frac{\partial g_r}{\partial\theta_2} & \dots & \frac{\partial g_r}{\partial\theta_k} \end{pmatrix}$$

$$\frac{\partial g^T}{\partial \theta} = \begin{pmatrix} \frac{\partial g_1^T}{\partial\theta} & \frac{\partial g_2^T}{\partial\theta} & \dots & \frac{\partial g_r^T}{\partial\theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial\theta_1} & \frac{\partial g_2}{\partial\theta_1} & \dots & \frac{\partial g_r}{\partial\theta_1} \\ \frac{\partial g_1}{\partial\theta_2} & \frac{\partial g_2}{\partial\theta_2} & \dots & \frac{\partial g_r}{\partial\theta_2} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_1}{\partial\theta_k} & \frac{\partial g_2}{\partial\theta_k} & \dots & \frac{\partial g_r}{\partial\theta_k} \end{pmatrix}, \quad \frac{\partial \ell}{\partial\theta} = \begin{pmatrix} \frac{\partial\ell}{\partial\theta_1} \\ \frac{\partial\ell}{\partial\theta_2} \\ \vdots \\ \frac{\partial\ell}{\partial\theta_k} \end{pmatrix}$$

$$I(\theta) = -E\left(\frac{\partial^2\ell}{\partial\theta\partial\theta^T}\right) = -E\begin{pmatrix} \frac{\partial^2\ell}{\partial\theta_1\partial\theta_1} & \frac{\partial^2\ell}{\partial\theta_1\partial\theta_2} & \dots & \frac{\partial^2\ell}{\partial\theta_1\partial\theta_k} \\ \frac{\partial^2\ell}{\partial\theta_2\partial\theta_1} & \frac{\partial^2\ell}{\partial\theta_2\partial\theta_2} & \dots & \frac{\partial^2\ell}{\partial\theta_2\partial\theta_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2\ell}{\partial\theta_k\partial\theta_1} & \frac{\partial^2\ell}{\partial\theta_k\partial\theta_2} & \dots & \frac{\partial^2\ell}{\partial\theta_k\partial\theta_k} \end{pmatrix}$$ — информационная матрица Фишера

$\Theta_{UR} := \Theta$ — множество допустимых значений вектора неизвестных параметров без учёта ограничений

$\Theta_R := \{\theta \in \Theta : g(\theta) = 0\}$ — множество допустимых значений вектора неизвестных параметров с учётом ограничений

$\hat\theta_{UR} \in \Theta_{UR}$ — точка максимума функции ℓ на множестве Θ_{UR}

$\hat\theta_R \in \Theta_R$ — точка максимума функции ℓ на множестве Θ_R

Тогда для тестирования гипотезы H_0 можно воспользоваться одной из следующих ниже статистик:

$LR := -2(\ell(\hat\theta_R) - \ell(\hat\theta_{UR})) \overset{as.}{\sim} \chi^2_r$ — статистика отношения правдоподобия

$W := g^T(\hat\theta_{UR}) \cdot \left[\frac{\partial g}{\partial\theta^T}(\hat\theta_{UR}) \cdot I^{-1}(\hat\theta_{UR}) \cdot \frac{\partial g^T}{\partial\theta}(\hat\theta_{UR})\right]^{-1} g(\hat\theta_{UR}) \overset{as.}{\sim} \chi^2_r$ — статистика Вальда

$LM := \left[\frac{\partial\ell}{\partial\theta}(\hat\theta_R)\right]^T \cdot I^{-1}(\hat\theta_R) \cdot \left[\frac{\partial\ell}{\partial\theta}(\hat\theta_R)\right] \overset{as.}{\sim} \chi^2_r$ — статистика множителей Лагранжа

1. Рассмотрим модель бинарного выбора $P(y_i = 1) = F(\beta_1 + \beta_2 d_i)$, где d — фиктивная переменная (принимающая значения 0 и 1). Ниже представлены результаты 100 наблюдений:

	$y = 0$	$y = 1$
$d = 0$	20	32
$d = 1$	36	12

   (a) Оцените параметры β_1, β_2, используя logit-модель.

   (b) Проверьте гипотезу $H_0 : \beta_2 = 0$ с помощью LR, W и LM тестов.

Решение:
Рассмотрим модель $P(y_t = 1) = F(\beta_1 + \beta_2 d_t)$. Функция правдоподобия равна:

$$L = F(\beta_1)^{32} F(\beta_1 + \beta_2)^{12} (1 - F(\beta_1))^{20} (1 - F(\beta_1 + \beta_2))^{36}.$$

Обозначим $\gamma = \beta_1 + \beta_2$. Логарифмическая функция правдоподобия равна:

$$\ln L = 32 \ln F(\beta_1) + 12 \ln F(\gamma) + 20 \ln(1 - F(\beta_1)) + 36 \ln(1 - F(\gamma)).$$

Оценка максимального правдоподобия является решением системы уравнений:

$$\frac{\partial \ln L}{\partial \beta_1} = 32 \frac{f(\beta_1)}{F(\beta_1)} - 20 \frac{f(\beta_1)}{1 - F(\beta_1)} = 0,$$

$$\frac{\partial \ln L}{\partial \gamma} = 12 \frac{f(\gamma)}{F(\gamma)} - 36 \frac{f(\gamma)}{1 - F(\gamma)} = 0.$$

Решением является

$$F(\beta_1) = \frac{8}{13}, \quad F(\gamma) = \frac{1}{4}.$$

Для logit-модели $F(z) = \Lambda(z)$ получаем:

$$\hat{\beta}_1 = 0.4700, \quad \hat{\beta}_2 = -1.5686.$$

Проверим теперь гипотезу $\beta_2 = 0$ при помощи теста отношения правдоподобия, теста Вальда и теста множителей Лагранжа.

Сначала применим тест отношения правдоподобия:

$$LR = -2(\ln \tilde{L} - \ln \tilde{L}).$$

Значение функции правдоподобия для регрессии без ограничения равно

$$\ln \tilde{L} = 32 \ln\left(\frac{8}{13}\right) + 12 \ln\left(\frac{1}{4}\right) + 20 \ln\left(\frac{5}{13}\right) + 36 \ln\left(\frac{3}{4}\right) = -61.6386.$$

Оценим модель с ограничением $\beta_2 = 0$ или $\gamma = \beta_1$:

$$L = F(\beta_1)^{44}(1 - F(\beta_1))^{56},$$

$$\ln L = 44 \ln F(\beta_1) + 56 \ln(1 - F(\beta_1)),$$

$$\frac{\partial \ln L}{\partial \beta_1} = 44 \frac{f(\beta_1)}{F(\beta_1)} - 56 \frac{f(\beta_1)}{1 - F(\beta_1)} = 0.$$

Откуда получаем $F(\beta_1) = \frac{11}{25}$ и

$$\ln \tilde{L} = 44 \ln\left(\frac{11}{25}\right) + 56 \ln\left(\frac{14}{25}\right) = -68.5930.$$

Получаем $LR = 13.9088$, что превышает критическое значение $\chi^2_{0.05}(1) = 3.84$, т. е. гипотеза $\beta_2 = 0$ уверенно отвергается. Как видно из вычислений, этот результат не зависит от вида функции распределения $F(z)$ и, следовательно, одинаков для logit- и probit-моделей.

2. Методом максимального правдоподобия оценили logit–модель

$$\hat{y}_i^* = 2 + 3x_i - 5z_i.$$

(a) Оцените вероятность того, что $y = 1$ для $x = 15$, $z = 9$.

(b) Оцените предельный эффект увеличения x на единицу на вероятность того, что $y = 1$ для $x = 15$, $z = 9$.

(c) Рассчитайте отношение шансов для переменной x.

(d) При каком значении x предельный эффект увеличения x на единицу в точке $z = 9$ будет максимальным?

Решение:

Методом максимального правдоподобия оценили logit-модель:

$$\hat{y}_i^* = 2 + 3x_i - 5z_i.$$

Функция вероятности для logit-модели:

$$P(y=1) = \Lambda(\hat{y}^*) = \frac{e^{\hat{y}^*}}{1+e^{\hat{y}^*}}.$$

(a)

$$\hat{y}^* = 2 + 3 \cdot 15 - 5 \cdot 9 = 2 + 45 - 45 = 2,$$
$$P(y=1) = \frac{e^2}{1+e^2} = \frac{7.389}{1+7.389} \approx \frac{7.389}{8.389} \approx 0.881.$$

(b) Предельный эффект для logit-модели:

$$\frac{\partial P(y=1)}{\partial x} = \beta_x \cdot \Lambda(\hat{y}^*) \cdot (1 - \Lambda(\hat{y}^*)) = 3 \cdot 0.881 \cdot 0.119 \approx 0.314.$$

(c)

$$OR = e^{\beta_x} = e^3 \approx 20.086.$$

(d) Предельный эффект максимален при $P(y=1) = 0.5$:

$$\Lambda(\hat{y}^*) = 0.5 \implies \hat{y}^* = 0,$$
$$0 = 2 + 3x - 45 \implies 3x = 43 \implies x = \frac{43}{3} \approx 14.\overline{3}.$$