Множественная регрессия

Пусть имеются случайные величины $\xi_1, \ldots, \xi_n$ с совместной плотностью $f(x_1, x_2, \ldots, x_n)$. Регрессией или условным математическим ожиданием случайной величины ξ_1 при фиксированных значениях случайных величин
$\xi_i,\ i = 2, \ldots, n$, называется функция, зависящая от переменных $x_2, \ldots, x_n$, вида

$$E(\xi_1 | \xi_2 = x_2, \ldots, \xi_n = x_n) = \frac{\int\limits_{-\infty}^{\infty} x_1 f(x_1, \ldots, x_n) dx_1}{\int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1}$$

Поверхность регрессии описывается уравнением
$x_1 = m_1(x_2, \ldots, x_n)$.

Для упрощения выводов будем считать $\mathrm{E}\xi_\mathrm{i} = 0$. Формулы для произвольного центра тяжести будут получаться простой заменой ξ_i на $\xi_\mathrm{i} - \mathrm{m}_\mathrm{i}$. Среднеквадратическая регрессия величины ξ_1 относительно $\xi_2, \ldots, \xi_\mathrm{n}$ определяется как гиперплоскость

$$\xi_1 = \beta_{12}\xi_2 + \ldots + \beta_{1\mathrm{n}}\xi_\mathrm{n}, \tag{1}$$

дающая наилучшую аппроксимацию математического ожидания квадрата разности

$$\mathrm{E}(\xi_1 - \beta_{12}\xi_2 - \ldots - \beta_{1\mathrm{n}}\xi_\mathrm{n})^2 \to \min,$$

т.е. правая часть (1) является наилучшей линейной оценкой ξ_1 величинами $\xi_2, \ldots, \xi_\mathrm{n}$ в смысле минимума (??). Коэффициенты $\beta_{1\mathrm{i}}$ называются частными коэффициентами регрессии.

Вычисление частных коэффициентов регрессии

Для их вычисления продифференцируем $\mathrm{E}(\xi_1 - \beta_{12}\xi_2 - \ldots - \beta_{1n}\xi_n)^2$ по β_{1i}.

$$\begin{cases} \frac{d}{d\beta_{12}}: & -2\mathrm{E}\xi_2(\xi_1 - \beta_{12}\xi_2 - \ldots - \beta_{1n}\xi_n) = 0 \\ & \ldots \\ \frac{d}{d\beta_{1n}}: & -2\mathrm{E}\xi_n(\xi_1 - \beta_{12}\xi_2 - \ldots - \beta_{1n}\xi_n) = 0 \end{cases}$$

Это равносильно при $\lambda_{ij} = \mathrm{E}\xi_i\xi_j$ системе уравнений

$$\begin{cases} \lambda_{22}\beta_{12} + \ldots + \lambda_{2n}\beta_{1n} = \lambda_{21} \\ \ldots \\ \lambda_{n2}\beta_{12} + \ldots + \lambda_{nn}\beta_{1n} = \lambda_{n1} \end{cases} \tag{2}$$

Определитель этой системы равен алгебраическому дополнению Λ_{11} матрицы вторых моментов

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \ldots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \ldots & \lambda_{2n} \\ & & \ldots & \\ \lambda_{n1} & \lambda_{n2} & \ldots & \lambda_{nn} \end{bmatrix}$$

По теореме Крамера, β_{12} равно отношению определителя

$$\begin{vmatrix} \lambda_{21} & \lambda_{23} & \dots & \lambda_{2\mathrm{n}} \\ & & \dots & \\ \lambda_{\mathrm{n}1} & \lambda_{\mathrm{n}3} & \dots & \lambda_{\mathrm{nn}} \end{vmatrix},$$

который равен $-\Lambda_{12}$, к определителю Λ_{11}, т.е.
$\beta_{12} = -\Lambda_{12}/\Lambda_{11}$.

Коэффициент β_{13} равен отношению определителя

$$\begin{vmatrix} \lambda_{22} & \lambda_{21} & \lambda_{24} & \dots & \lambda_{2\mathrm{n}} \\ & & & \dots & \\ \lambda_{\mathrm{n}2} & \lambda_{\mathrm{n}1} & \lambda_{\mathrm{n}4} & \dots & \lambda_{\mathrm{nn}} \end{vmatrix}, \tag{3}$$

полученного в результате замены в матрице Λ_{11} второго столбца на столбец свободных членов системы (2), к определителю Λ_{11}. Поменяем местами первые два столбца. Определитель (3) равен

$$-\begin{vmatrix} \lambda_{21} & \lambda_{22} & \lambda_{24} & \dots & \lambda_{2\mathrm{n}} \\ & & & \dots & \\ \lambda_{\mathrm{n}1} & \lambda_{\mathrm{n}2} & \lambda_{\mathrm{n}4} & \dots & \lambda_{\mathrm{nn}} \end{vmatrix} = -\Lambda_{13}, \tag{4}$$

т.е.

$$\beta_{13} = -\frac{\Lambda_{13}}{\Lambda_{11}}.$$

Аналогично получаем, что

$$\beta_{1\mathrm{k}} = -\frac{\Lambda_{1\mathrm{k}}}{\Lambda_{11}}. \tag{5}$$

Простой перестановкой индексов получаем

$$\beta_{\mathrm{ik}} = -\frac{\Lambda_{\mathrm{ik}}}{\Lambda_{\mathrm{ii}}}. \tag{6}$$

Остатки и остаточная дисперсия

Ограничимся рассмотрением случая, когда $|\Lambda_{11}| \neq 0$ и все $E\xi_i = 0$. Случайная величина вида

$$\eta_{1\cdot 23\ldots n} = \xi_1 - \beta_{12}\xi_2 - \ldots - \beta_{1n}\xi_n \tag{7}$$

называется остатком величины ξ_1 относительно $\xi_2, \ldots, \xi_n$. Покажем, что остаток не коррелирован ни с одной из "вычитаемых" случайных величин. Из $\beta_{1k} = -\frac{\Lambda_{1k}}{\Lambda_{11}}$ и из того, что $1 = \Lambda_{11}/\Lambda_{11}$ имеем выражение остатка через линейную комбинацию случайных величин

$$\eta_{1\cdot 23\ldots n} = \frac{1}{\Lambda_{11}} \sum_{k=1}^{n} \Lambda_{1k}\xi_k. \tag{8}$$

Отсюда, с учетом того, что $E\xi_i = 0$, получаем

$$E\eta_{1\cdot 23\ldots n} = 0. \tag{9}$$

Остаточная дисперсия

$$\mathrm{E}\xi_i\eta_{1\cdot 23\ldots n} = \frac{1}{\Lambda_{11}}\sum_{k=1}^{n}\mathrm{E}\xi_i\xi_k\Lambda_{1k} =$$

$$= \frac{1}{\Lambda_{11}}\sum_{k=1}^{n}\lambda_{ik}\Lambda_{1k} = \begin{cases} \frac{\Lambda}{\Lambda_{11}}, & i = 1, \\ 0, & i \neq 1, \end{cases} \tag{10}$$

Так как $\sum_{k=1}^{n}\lambda_{1k}\Lambda_{1k} = \Lambda$, а $\sum_{k=1}^{n}\lambda_{ik}\Lambda_{1k} = 0$. Дисперсия остатка

$$\sigma^2_{1\cdot 23\ldots n} = \mathrm{E}(\eta^2_{1\cdot 23\ldots n}).$$

Заменим одну из η на $\eta_{1\cdot 23\ldots n} = \frac{1}{\Lambda_{11}}\sum_{k=1}^{n}\Lambda_{1k}\xi_k$.. В силу (10)

$$\sigma^2_{1\cdot 23\ldots n} = \mathrm{E}\left(\frac{1}{\Lambda_{11}}\sum_{i=1}^{n}\Lambda_{1i}\xi_i\eta_{1\cdot 23\ldots n}\right) = \mathrm{E}(\xi_1\eta_{1\cdot 23\ldots n}) = \frac{\Lambda}{\Lambda_{11}}. \tag{11}$$

Коэффициент корреляции между остатками

$$\varrho_{12\cdot 34\ldots \mathrm{n}} = \mathrm{cor}(\eta_{1\cdot 34\ldots \mathrm{n}}\eta_{2\cdot 34\ldots \mathrm{n}}) = \frac{\mathrm{E}(\eta_{1\cdot 34\ldots \mathrm{n}}\eta_{2\cdot 34\ldots \mathrm{n}})}{\sqrt{\mathrm{E}(\eta_{1\cdot 34\ldots \mathrm{n}})^2\mathrm{E}(\eta_{2\cdot 34\ldots \mathrm{n}})^2}} \tag{12}$$

$$\mathrm{E}\eta^2_{1\cdot 34\ldots \mathrm{n}} = \mathrm{E}(\xi_1\eta_{1\cdot 34\ldots \mathrm{n}}) = \frac{\Lambda_{22}}{\Lambda_{22\cdot 11}},\ \mathrm{E}\eta^2_{2\cdot 34\ldots \mathrm{n}} = \frac{\Lambda_{11}}{\Lambda_{11\cdot 22}},$$

$$\mathrm{E}(\eta_{1\cdot 34\ldots \mathrm{n}}\eta_{2\cdot 34\ldots \mathrm{n}}) = \mathrm{E}(\xi_1\eta_{2\cdot 34\ldots \mathrm{n}}), \tag{13}$$

$$\text{так}\quad\text{как}\ \eta_{1\cdot 34\ldots \mathrm{n}} = \frac{1}{\Lambda_{22\cdot 11}}\sum_{\mathrm{k}=1,\mathrm{k}\neq 2}^{\mathrm{n}}\Lambda_{22\cdot 1\mathrm{k}}\xi_{\mathrm{k}},$$

$$\mathrm{E}\xi_{\mathrm{k}}\eta_{2\cdot 34\ldots \mathrm{n}}\begin{cases} = 0, & \mathrm{k} = 3, 4, \ldots, \mathrm{n} \\ \neq 0, & \mathrm{k} = 2 \\ \neq 0, & \mathrm{k} = 1 \end{cases}$$

Так как $\eta_{2\cdot 34\ldots n} = \frac{1}{\Lambda_{11\cdot 22}} \sum\limits_{k=2}^{n} \Lambda_{11\cdot 2k} \xi_k$, то (13) имеет вид

$$E(\xi_1 \eta_{2\cdot 34\ldots n}) = \frac{1}{\Lambda_{11\cdot 22}} \sum_{k=2}^{n} \Lambda_{11\cdot 2k} E\xi_1 \xi_k =$$

$$= \frac{1}{\Lambda_{11\cdot 22}} \sum_{k=2}^{n} \Lambda_{11\cdot 2k} \lambda_{1k} = -\frac{\Lambda_{12}}{\Lambda_{11\cdot 22}}.$$

Таким образом,

$$\varrho_{12\cdot 34\ldots n} = -\frac{\Lambda_{12}}{\sqrt{\Lambda_{11}\Lambda_{22}}} \tag{14}$$

Между частной корреляцией и соответствующим частным коэффициентом регрессии имеет место линейная зависимость. Согласно (14) и (5) имеем

$$\varrho_{12\cdot 34\ldots \mathrm{n}} = -\frac{\Lambda_{12}}{\sqrt{\Lambda_{11}\Lambda_{22}}} = -\frac{\Lambda_{12}}{\Lambda_{11}}\sqrt{\frac{\Lambda_{11}}{\Lambda_{22}}} = \beta_{12\cdot 34\ldots \mathrm{n}}\sqrt{\frac{\Lambda_{11}}{\Lambda_{22}}}. \qquad (15)$$

Для случая $\mathrm{n} = 3$ матрица вторых моментов Λ имеет вид

$$\Lambda = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\varrho_{12} & \sigma_1\sigma_3\varrho_{13} \\ \sigma_1\sigma_2\varrho_{12} & \sigma_2^2 & \sigma_2\sigma_3\varrho_{23} \\ \sigma_1\sigma_3\varrho_{13} & \sigma_2\sigma_3\varrho_{23} & \sigma_3^2 \end{bmatrix}$$

$$\varrho_{12\cdot 3} = -\frac{\Lambda_{12}}{\sqrt{\Lambda_{11}\Lambda_{22}}} = \frac{\sigma_1\sigma_2\varrho_{12}\sigma_3^2 - \sigma_1\sigma_3\varrho_{13}\sigma_2\sigma_3\varrho_{23}}{\sigma_2\sigma_3\sqrt{1-\varrho_{23}^2}\,\sigma_1\sigma_3\sqrt{1-\varrho_{13}^2}} =$$

$$= \frac{\varrho_{12} - \varrho_{13}\varrho_{23}}{\sqrt{1-\varrho_{23}^2}\sqrt{1-\varrho_{13}^2}}$$

ПРИМЕР

j	Годы	Продолж-ть жизни ξ_1	Нац.богатсво (млрд.дол.) ξ_2	Водка (бут.в год) ξ_3
1	1970	68.9	2372*	25.3
2	1975	68.1	2372*	28
3	1980	67.6	2489	30
4	1985	69.2	3379	23.5
5	1990	69.2	4130	18
6	1995	64.6	1171	38.4
7	1998	67	689	29.6

$$\varrho_{12} = 0.78,\ \varrho_{13} = -0.93,\ \varrho_{23} = -0.82.\ \sigma^2_{1\cdot 23} = \frac{\Lambda}{\Lambda_{11}} = 0.29.$$

$$\beta_{12} = -\frac{\Lambda_{12}}{\Lambda_{11}} = -\frac{-368}{13214148} = 2.8\cdot 10^{-5},\quad \varrho_{12\cdot 3} = -\frac{\Lambda_{12}}{\sqrt{\Lambda_{11}\Lambda_{22}}} = 0.032,$$

$$\beta_{13} = -\frac{\Lambda_{13}}{\Lambda_{11}} = -\frac{138823}{13214148} = -0.24,\quad \varrho_{13\cdot 2} = -\frac{\Lambda_{13}}{\sqrt{\Lambda_{11}\Lambda_{33}}} = -0.82,$$

Корреляция $\varrho_{12} = 0.78$ объясняется тем, что при снижении национального богатства увеличивается продажа водки, что влечет снижение продолжительности жизни.

Множественный коэффициент корреляции

Множественным коэффициентом корреляции $\varrho_{1(23\ldots n)}$ называется коэффициент корреляции между ξ_1 и $\xi_1^* = \beta_{12}\xi_2 + \ldots + \beta_{1n}\xi_n = \xi_1 - \eta_{1\cdot 23\ldots n}$. При $E\xi_i = 0$

$$\varrho_{1(23\ldots n)} = \frac{E(\xi_1\xi_1^*)}{\sqrt{E\xi_1^2 E(\xi_1^*)^2}}, \tag{16}$$

$$E(\xi_1\xi_1^*) = E(\xi_1(\xi_1 - \eta_{1\cdot 23\ldots n})) = E\xi_1^2 - E\xi_1\eta_{1\cdot 23\ldots n} = \lambda_{11} - \frac{\Lambda}{\Lambda_{11}},$$

$$E(\xi_1^*)^2 = E(\xi_1^2 - 2\xi_1\eta_{1\cdot 23\ldots n} + \eta_{1\cdot 23\ldots n}^2) = \lambda_{11} - \frac{\Lambda}{\Lambda_{11}}.$$

Следовательно,

$$\varrho_{1(23\ldots n)} = \frac{\sqrt{\lambda_{11} - \frac{\Lambda}{\Lambda_{11}}}}{\sqrt{\lambda_{11}}} = \sqrt{1 - \frac{\Lambda}{\Lambda_{11}\lambda_{11}}}. \tag{17}$$

Для численных вычислений оказывается полезной формула

$$\varrho^2_{1(23\ldots\mathrm{n})} = 1 - \frac{\sigma^2_{1\cdot 23\ldots\mathrm{n}}}{\sigma^2_1} = \frac{\sigma^2_1 - \sigma^2_{1\cdot 23\ldots\mathrm{n}}}{\sigma^2_1}, \tag{18}$$

так как по (11) $\sigma^2_{1\cdot 23\ldots\mathrm{n}} = \frac{\Lambda}{\Lambda_{11}}$, а $\lambda_{11} = \sigma^2_1$.
Квадрат множественного коэффициента корреляции или коэффициент детерминации равен доле дисперсии, объясняемой регрессией, то есть влиянием случайных величин $\xi_2, \ldots, \xi_\mathrm{n}$ на ξ_1.
ПРИМЕР (продолжение). Коэффициент детерминации равен

$$\mathrm{R}^2 = \varrho^2_{1(23)} = \frac{\sigma^2_1 - \sigma^2_{1\cdot 23}}{\sigma^2_1} = \frac{2.31 - 0.29}{2.31} = 0.87.$$

Отношение дисперсий и частный коэффициент корреляции

$$\frac{\beta_{12}^2 \sigma_{2\cdot 34\ldots n}^2}{\sigma_{1\cdot 234\ldots n}^2} = \frac{\varrho_{12\cdot 34\ldots n}^2}{1 - \varrho_{12\cdot 34\ldots n}^2} \tag{19}$$

Равенство Якоби: $\Lambda \cdot \Lambda_{11\cdot 22} = \Lambda_{11} \cdot \Lambda_{22} - \Lambda_{12} \cdot \Lambda_{21}$

$$\beta_{12} = \frac{\Lambda_{12}}{\Lambda_{11}}, \quad \sigma_{1\cdot 23\ldots n}^2 = \frac{\Lambda}{\Lambda_{11}}, \quad \sigma_{2\cdot 34\ldots n}^2 = \frac{\Lambda_{11}}{\Lambda_{11\cdot 22}}, \quad \varrho_{12\cdot 34\ldots n}^2 = \frac{\Lambda_{12}^2}{\Lambda_{11}\Lambda_{22}}.$$

получаем левую часть в виде:

$$\frac{\beta_{12}^2 \sigma_{2\cdot 34\ldots n}^2}{\sigma_{1\cdot 234\ldots n}^2} = \frac{\Lambda_{12}^2}{\Lambda_{11}^2} \cdot \frac{\Lambda_{11}}{\Lambda_{11\cdot 22}} \cdot \frac{\Lambda_{11}}{\Lambda} = \frac{\Lambda_{12}^2}{\Lambda \cdot \Lambda_{11\cdot 22}} = \frac{\Lambda_{12}^2}{\Lambda_{11} \cdot \Lambda_{22} - \Lambda_{12} \cdot \Lambda_{21}}.$$

Правая часть совпадает с левой, так как

$$\frac{\varrho_{12\cdot 34\ldots n}^2}{1 - \varrho_{12\cdot 34\ldots n}^2} = \frac{\Lambda_{12}^2}{\Lambda_{11}\Lambda_{22}\left(1 - \frac{\Lambda_{12}^2}{\Lambda_{11}\Lambda_{22}}\right)}.$$

Равенство Якоби получается из произведения определителей $(\Lambda_{11}\Lambda_{22}-\Lambda_{12}\Lambda_{21})|\Lambda| = |\Lambda|^2\Lambda_{11\cdot 22}$ соответствующих матриц

$$\left[\begin{array}{cc|ccc} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & \dots & \Lambda_{1n} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} & \dots & \Lambda_{2n} \\ \hline \multicolumn{2}{c|}{\mathbb{O}} & \multicolumn{3}{c}{\mathbb{I}} \end{array}\right] \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2n} \\ \dots & \dots & \dots & \dots \\ \lambda_{n1} & \lambda_{n2} & \dots & \lambda_{nn} \end{bmatrix} =$$

$$= \left[\begin{array}{cc|ccc} |\Lambda| & 0 & 0 & \dots & 0 \\ 0 & |\Lambda| & 0 & \dots & 0 \\ \hline \lambda_{31} & \lambda_{32} & \lambda_{33} & \dots & \lambda_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_{n1} & \lambda_{n2} & \lambda_{3n} & \dots & \lambda_{nn} \end{array}\right]$$