Методология применения бинарной регрессии в точном земледелии

Буре В. М.

Введение

Методология точного земледелия предполагает широкое применение компьютерных технологий и автоматизированных систем анализа данных, непрерывного сбора и обработки статистической информации [1].

Методы анализа данных на основе бинарной регрессии находят применение и в точном земледелии [7-10].

7) Stan G. Daberkow, William D. McBride Farm and operator characteristics affecting the awareness and adoption of precision agriculture technologies in US // Precision agriculture, 4, 163-177, 2003.

8) Robertson, M. J., Llewellyn R. S., Mandel R., Lawes R., Bramley R. G. V., Swift L., Metz N., O'Callaghan C. Adoption of variable rate fertilizer application in the Australian grain industry: status, issues and prospects // Precision agriculture, 13, 181-199, 2012.

9) J. Anita Dille, David A. Mortensen, Linda J. Young Predicting weed species occurrence based on site properties and previous year's weed presence // Precision agriculture, 3, 193-207, 2002.

В работе предлагается новый подход на основе бинарной регрессии к задаче оценки вероятности достижения заданного порогового уровня урожайности на конкретном поле в рамках методологии точного земледелия.

1. Дискретные данные

Предположим, что зависимые переменные принимают дискретные значения, выражающие какие-либо качественные признаки. Объясняющие переменные могут быть как дискретными, так и непрерывными. Выделим несколько классов задач, в которых зависимые переменные принимают дискретные значения:

- Переменные — это решения «да» (1) или «нет» (0), т. е. выбор одной из двух альтернатив. Если имеется только две альтернативы, то результат наблюдения обычно описывается переменной, называемой бинарной. В общем случае при наличии k альтернатив результат выбора можно представить переменной, принимающей значения $1, \ldots, k$. Если альтернативы нельзя упорядочить, то их нумерация может быть произвольной. В этих случаях соответствующую переменную называют *номинальной*.
- Переменные — ранги. Например, «0» означает «категорически против», «1» — «против», «2» — «ни да, ни нет», «3» — «за», «4» — «полностью за». Такая система ранжирования может быть использована, например, при голосовании. Соответствующая переменная называется *порядковой, ординальной* или *ранговой*.

Дискретные данные

- Переменная — количественная целочисленная характеристика. Например, количество торговых точек, приносящих прибыль, количество частных предприятий, зарегистрированных в регионе и т.д.

Для моделей с дискретными зависимыми переменными, конечно же, возможно формальное применение метода наименьших квадратов, однако, получить удовлетворительные с содержательной точки зрения результаты удается, как правило, лишь для моделей с количественными целочисленными переменными. В случае порядковых переменных интерпретация оценок коэффициентов при объясняющих переменных значительно затруднена: увеличение на одну единицу порядковой переменной означает переход к следующей по рангу альтернативе, однако далеко не всегда переход от первой альтернативы ко второй численно эквивалентен переходу от второй к третьей. Если же зависимая переменная является номинальной, то результаты оценивания вообще теряют смысл в силу произвольности нумерации альтернатив. Таким образом, стандартная регрессионная схема нуждается в коррекции.

2. Модель линейной вероятности

Если следовать основным идеям регрессионного анализа, то будем считать, что на решение о покупке квартиры влияют также неучтенные факторы, совокупное влияние которых моделируется случайной компонентой. При различных предположениях о характере зависимости y от x, можно получить различные модели построения зависимости.

Пусть имеется выборка объема n наблюдений (x_i, y_i), $i = 1, \ldots, n$, где $x_i^T = (x_{i1}, \ldots, x_{ik})^T$, y_i — зависимая переменная, которая может принимать только два значения: ноль и единица. Рассмотрим стандартную модель линейной регрессии:

$$y_i = \beta^T x_i + \varepsilon_i, \tag{1}$$

где $\beta^T = (\beta_0, \ldots, \beta_k)$ — вектор неизвестных параметров, $\beta \in \mathbb{R}^k$, ε_i — случайная компонента. В предположениях регрессионного анализа считается, что случайная компонента подчиняется нормальному закону распределения с нулевым математическим ожиданием.

Модель линейной вероятности

Учитывая это, получаем, что

$$Ey_i = \beta^T x_i.$$

Так как y_i принимает значения 0 или 1, то для математического ожидания y_i имеем равенство:

$$Ey_i = 1 \cdot P\{y_i = 1\} + 0 \cdot P\{y_i = 0\} = P\{y_i = 1\}. \quad (2)$$

Таким образом, получаем равенство:

$$P\{y_i = 1\} = \beta^T x_i, \quad (3)$$

которое дало название модели линейной вероятности (linear probability model).

Следует отметить некоторые недостатки этой модели, которые не позволяют успешно применять метод наименьших квадратов для оценивания параметров β и построения прогнозов. Из (1) следует, что компонента ε_i в каждом наблюдении может принимать только два значения: $(1 - \beta^T x_i)$ с вероятностью $P\{y_i = 1\}$ и $(-\beta^T x_i)$ с вероятностью $1 - P\{y_i = 1\}$.

Модель линейной вероятности

Это, в частности, не позволяет считать случайную компоненту нормально распределенной случайной величиной или, подчиняющейся распределению, близкому к нормальному.

Проверим выполнение условия из первой группы предположений регрессионного анализа о равенстве дисперсий различных наблюдений. Вычислим дисперсию компоненты:

$$D\varepsilon_i = \beta^T x_i(1 - \beta^T x_i).$$

Получается, что дисперсия компоненты ε_i зависит от x_i. Известно, что оценка параметров β, полученная обычным методом наименьших квадратов, в этом случае не является эффективной.

Еще одним серьезным недостатком модели линейной вероятности является тот факт, что прогнозные значения $\hat{y}_i = \hat{\beta}^T x_i$, т. е. прогнозные значения вероятности $P\{y_i = 1\}$, могут лежать вне отрезка $[0, 1]$ (здесь $\hat{\beta}$ — оценка параметра β, полученная методом наименьших квадратов).

3. Логит и пробит модели бинарного выбора

Откажемся от предположения о линейной зависимости вероятности $P\{y_i = 1\}$ от β. Предположим, что

$$P\{y_i = 1\} = F(\beta^T x_i), \tag{4}$$

где $F(x)$ — некоторая функция, область значений которой лежит в отрезке $[0, 1]$. В частности, в качестве функции $F(x)$ можно рассмотреть функцию распределения некоторой случайной величины. Возможна следующая интерпретация предположения (4). Предположим, что существует некоторая количественная переменная y_i^*, связанная с независимыми переменными x_i линейным регрессионным уравнением:

$$y_i^* = \beta^T x_i + \varepsilon_i, \tag{5}$$

где случайные компоненты ε_i независимы и одинаково распределены с нулевым математическим ожиданием и дисперсией σ^2. Пусть $F(\cdot)$ — функция распределения нормированной случайной величины ε_i/σ.

Логит и пробит модели бинарного выбора

Переменная y_i^* является ненаблюдаемой (латентной), а решение, соответствующее значению $y_i = 1$, принимается тогда, когда y_i^* превосходит некоторое пороговое значение. Без ограничения общности, если константа включена в число независимых переменных модели, можно считать это пороговое значение равным нулю. Величину y_i^* можно также интерпретировать как разность полезностей альтернативы 1 и альтернативы 0.

Таким образом, получаем

$$y_i = \begin{cases} 1, & \text{если} \quad y_i^* \geqslant 0, \\ 0, & \text{если} \quad y_i^* < 0. \end{cases} \tag{6}$$

Тогда, предполагая, что случайные компоненты ε_i имеют одно и то же симметричное распределение с непрерывной функцией распределения $F(x)$, $F(-x) = 1 - F(x)$, получаем следующие равенства:

$$P\{y_i = 1\} = P\{y_i^* \geqslant 0\} = P\{\beta^T x_i + \varepsilon_i \geqslant 0\} = P\{\varepsilon_i \geqslant -\beta^T x_i\} = \\ = 1 - F\left(\frac{-\beta^T x_i}{\sigma}\right) = F\left(\frac{\beta^T x_i}{\sigma}\right), \tag{7}$$

что с точностью до нормировки совпадает с (4).

В модели (7) параметры β и σ участвуют только в виде отношения β/σ и не могут быть по отдельности идентифицированы (т. е. оценить можно только отношения β/σ). Поэтому, в данном случае, без ограничения общности можно считать, что $\sigma = 1$.

Наиболее часто в качестве функции $F(x)$ используют:

1. Функцию стандартного нормального распределения

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{z^2}{2}} dz,$$

   в этом случае модель принято называть *пробит моделью*.

2. Функцию логистического распределения

$$\Lambda(u) = \frac{e^u}{1 + e^u}, \tag{8}$$

   тогда модель принято называть *логит моделью*.

Логит и пробит модели бинарного выбора

Использование функции стандартного нормального распределения представляется естественным в рамках приведенной выше интерпретации. Применение функции логистического распределения объясняется простотой численной реализации процедуры оценивания параметров. У исследователя может возникнуть вопрос о том, какую из моделей (пробит или логит) использовать в конкретном случае. Точного ответа на этот вопрос нет, но можно дать некоторые рекомендации. Например, можно выбрать ту модель, для которой функция правдоподобия имеет большее значение. Также можно отметить, что для значений u, близких по модулю к нулю (в частности, при $u \in -[1.2; 1.2]$), функции $\Phi(u)$ и $\Lambda(u)$ ведут себя примерно одинаково, в то же время «хвосты» логистического распределения значительно «тяжелее» «хвостов» нормального распределения. Практический опыт показывает, что для выборок с небольшим разбросом независимых переменных и при отсутствии существенного преобладания одной альтернативы над другой выводы, получаемые с помощью пробит и логит моделей, будут, как правило, совпадать. Функция $F(x)$ нелинейна по параметрам β, и интерпретация этих параметров отличается от интерпретации подобных коэффициентов в линейной регрессионной модели.

Вероятностные распределения, используемые в пробит и логит моделях, — стандартное нормальное и логистическое распределения соответственно. Обе указанные функции распределения симметричны относительно 0 и имеют дисперсии, равные 1 и $\pi^2/3$ соответственно. Рассмотрим модифицированное логистическое распределение с функцией распределения следующего вида:

$$\Lambda_\delta(u) = \frac{e^{\delta u}}{1 + e^{\delta u}}. \tag{9}$$

Значение параметра δ может быть выбрано таким образом, чтобы значения функции (9) были бы достаточно близкими к значениям функции стандартного нормального распределения на большой части области определения. Рассмотрим, например, $\delta = 1.6$. В следующей таблице можно найти значения функций $\Phi(u)$ и $\Lambda_{1.6}(u)$ для различных u.

Таблица значений функций $\Phi(u)$ и $\Lambda_{1.6}(u)$

u	$\Phi(u)$	$\Lambda_{1.6}(u)$
0.0	0.5	0.5
0.1	0.5398	0.5399
0.2	0.5793	0.5793
0.3	0.6179	0.6177
0.4	0.6554	0.6548
0.5	0.6915	0.6900
0.6	0.7257	0.7231
0.7	0.7580	0.7540
0.8	0.7881	0.7824
0.9	0.8159	0.8085
1.0	0.8413	0.8320
2.0	0.9772	0.9608
3.0	0.9987	0.9918

Таблица показывает, что функции распределения «очень близки» около 0, но логистическое распределение имеет более тяжелые «хвосты».

Таблица значений функций $\Phi(u)$ и $\Lambda_{1.6}(u)$

Из-за близости двух распределений трудно идентифицировать тип распределения при наличии выборки небольшого объема. Таким образом, при построении модели бинарной регрессии не имеет большого значения, будет использоваться пробит или логит модель, исключая случаи, когда большое количество данных расположены в хвостах, что может быть обусловлено спецификой рассматриваемой проблемной области. В моделях множественного дискретного выбора пробит и логит модели отличаются гораздо более существенно.

Предположим, что найдены оценки $\hat{\beta}_\Phi$ и $\hat{\beta}_\Lambda$ для параметров пробит и логит моделей соответственно. Тогда, используя приближенное равенство $\Lambda_{1.6}(u) \simeq \Phi(u)$, тогда можно записать следующее приближенное равенство для оценок параметров:

$$1.6\hat{\beta}_\Phi \simeq \hat{\beta}_\Lambda. \tag{10}$$

Формула (10) может быть полезна как быстрый способ сравнения оценок параметров пробит и логит моделей.

5. Оценивание параметров в логит и пробит моделях

Для нахождения оценок параметров β обычно используют метод максимального правдоподобия, предполагая, что наблюдения $y_1, \ldots, y_n$ независимы. Так как y_i может принимать значения 0 или 1, то функция правдоподобия примет следующий вид:

$$L(y_1, \ldots, y_n) = \prod_{i:y_i=0} (1 - F(\beta^T x_i)) \prod_{i:y_i=1} F(\beta^T x_i). \qquad (11)$$

Нетрудно заметить, что

$$L(y_1, \ldots, y_n) = \prod_{i=1}^{n} F^{y_i}(\beta^T x_i)(1 - F(\beta^T x_i))^{1-y_i}.$$

Рассмотрим логарифмическую функцию правдоподобия:

$$\ln L(y_1, \ldots, y_n) = \sum_{i=1}^{n} \left(y_i \ln F(\beta^T x_i) + (1 - y_i) \ln(1 - F(\beta^T x_i)) \right). \qquad (12)$$

Оценивание параметров в логит и пробит моделях

Дифференцируя равенство (12) по вектору β, получаем уравнение правдоподобия, записанное в векторной форме:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{y_i f(\beta^T x_i)}{F(\beta^T x_i)} - \frac{(1-y_i) f(\beta^T x_i)}{1 - F(\beta^T x_i)} \right) x_i = 0, \qquad (13)$$

где $f(x)$ — плотность распределения, соответствующая функции $F(x)$. Для логит модели уравнение (13) можно существенно упростить, если воспользоваться тождеством $\Lambda'(u) = \Lambda(u)(1 - \Lambda(u))$:

$$\sum_{i=1}^{n} (y_i - \Lambda(\beta^T x_i)) x_i = 0. \qquad (14)$$

Уравнение правдоподобия есть лишь необходимое условие локального экстремума. Можно показать, что для пробит и логит моделей логарифмическая функция правдоподобия (12) является вогнутой по β функцией и, значит, решение уравнения (13) дает оценку максимального правдоподобия параметра β.

Оценивание параметров в логит и пробит моделях

Гессиан для логит модели имеет следующий вид:

$$H = \frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} = -\sum_{i=1}^{n} \Lambda(\beta^T x_i)(1 - \Lambda(\beta^T x_i)) x_i x_i^T. \tag{15}$$

Заметим также, что гессиан в этом случае отрицательно определен, т. е. логарифмическая функция правдоподобия вогнута.

Для пробит модели логарифмическую функцию правдоподобия (12) можно записать в следующем виде:

$$\ln L = \sum_{i:y_i=0} \ln(1 - \Phi(\beta^T x_i)) + \sum_{i:y_i=1} \ln(\Phi(\beta^T x_i)). \tag{16}$$

Тогда условие (13) будет следующим:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i:y_i=0} \frac{-\varphi(\beta^T x_i)}{1 - \Phi(\beta^T x_i)} x_i + \sum_{i:y_i=1} \frac{\varphi(\beta^T x_i)}{\Phi(\beta^T x_i)} x_i,$$

где $\varphi(x) = \Phi'(x)$.

Оценивание параметров в логит и пробит моделях

Учитывая, что нормальное распределение, как и логистическое, симметрично, $1 - \Phi(\beta^T x) = \Phi(-\beta^T x)$, получаем:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{n} \frac{q_i \varphi(\beta^T x_i)}{\Phi(q_i \beta^T x_i)} x_i = \sum_{i=0}^{n} \lambda_i x_i = 0, \tag{17}$$

где $q_i = 2y_i - 1$, $\lambda_i = q_i \varphi(\beta^T x_i)/\Phi(q_i \beta^T x_i)$.

Для вычисления гессиана в модели пробит анализа будем использовать свойство стандартного нормального распределения: $d\varphi(u)/du = -u\varphi(u)$. Тогда для пробит модели получим следующее выражение для гессиана:

$$H = \frac{\partial^2 \ln L}{\partial \beta^T \partial \beta} = -\sum_{i=1}^{n} \lambda_i (\lambda_i + \beta^T x_i) x_i x_i^T. \tag{18}$$

Эта матрица также отрицательно определена.
Уравнения правдоподобия (14) и (17) являются системой нелинейных (относительно β) уравнений, аналитическое решение которой невозможно найти в явном виде в общем случае, поэтому при ее решении приходится прибегать к численным методам.

Будем использовать численные методы для того, чтобы найти оценку максимального правдоподобия параметра β, который является $(k+1)$-мерным вектором. Общая схема численных методов для нахождения оценки максимального правдоподобия имеет следующий вид. Сначала выбираем начальную точку $\beta^{(0)}$, на следующей итерации переходим в точку $\beta^{(1)}$ по следующему правилу:

$$\beta^{(1)} = \beta^{(0)} + \mu_0,$$

на t-ой итерации переходим в точку β^{t+1} по следующему правилу:

$$\beta^{(t+1)} = \beta^{(t)} + \mu_t.$$

Обычно μ_t выбирают в виде $\mu_t = D_t \partial \ln L / \partial \beta^{(t)}$, где $\partial \ln L / \partial \beta^{(t)}$ задает направление «спуска», т. е. направление изменения значений β, D_t — матрица «длин» шага. Итерационный процесс заканчивается на том шаге, на котором выполняется заранее определенное условие остановки. Такая итеративная процедура в результате определяет точку $\hat{\beta}$ (оценку максимального правдоподобия).

Задача — определить способ выбора вектора μ_t, поскольку от этого зависит скорость сходимости к искомой точке.

Метод градиентного спуска. Для этого метода предполагается, что $D_t = E$ для любого шага t. Выбор точки на $(t+1)$-ой итерации осуществляется по правилу:

$$\beta^{(t+1)} = \beta^{(t)} + \frac{\partial \ln L}{\partial \beta^{(t)}}.$$

В скалярном случае оценка будет возрастать, если градиент положителен, и уменьшаться, если отрицателен. Метод останавливается, когда значение производной будет «близко» к нулю.
К минусам этого метода можно отнести его «нечувствительность» к скорости изменения значений функции. Следующие три метода учитывают, как быстро логарифмическая функция правдоподобия меняется. Нельзя отдать предпочтение ни одному из ниже описанных методов, поскольку скорость работы алгоритмов зависит от обрабатываемых данных.

Метод Ньютона

В качестве матрицы D_t в методе Ньютона выбирается $\left(\frac{\partial^2 \ln L}{\partial \beta^{(t)T} \partial \beta^{(t)}}\right)^{-1}$.

Например, гессиан логарифмической функции правдоподобия с двумя параметрами $\beta^T = (\beta_0, \beta_1)$ выглядит следующим образом:

$$\frac{\partial^2 \ln L}{\partial \beta^T \partial \beta} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_0} & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} \\ \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_1} \end{pmatrix}.$$

Если $\partial^2 \ln L / \partial \beta_0 \partial \beta_0$ больше $\partial^2 \ln L / \partial \beta_1 \partial \beta_1$, то градиент меняется быстрее при возрастании β_0, чем при возрастании на то же самое число аргумента β_1.

Следующий элемент приближения в методе Ньютона выбирается по правилу:

$$\beta^{t+1} = \beta^{(t)} - \left(\frac{\partial^2 \ln L}{\partial \beta^{(t)T} \partial \beta^{(t)}}\right)^{-1} \frac{\partial \ln L}{\partial \beta^{(t)}},$$

где $\frac{\partial^2 \ln L}{\partial \beta^{(t)T} \partial \beta^{(t)}} = \left.\frac{\partial^2 \ln L}{\partial \beta^T \partial \beta}\right|_{\beta=\beta^{(t)}}$.

Метод scoring

В некоторых случаях математическое ожидание гессиана, известное как информационная матрица, проще вычислить, чем сам гессиан. Метод scoring использует информационную матрицу как матрицу D_t, и следующее приближение выбирается по правилу:

$$\beta^{t+1} = \beta^{(t)} + \left(E \left[\frac{\partial^2 \ln L}{\partial \beta^{(t)} \partial \beta^{(tT)}} \right] \right)^{-1} \frac{\partial \ln L}{\partial \beta^{(t)}}.$$

Метод BHHH

Метод BHHH (названный по первым буквам авторов E. Berndt, B. Hall, R. Hall, J. Hausman) — численный метод, имеющий широкое распространение в эконометрических задачах. Так как гессиан и информационную матрицу в некоторых случаях сложно вычислить, то в качестве матрицы D_t можно использовать произведение градиентов:

$$\sum_{i=1}^{n} \frac{\partial \ln L_i}{\partial \beta^{(t)}} \left(\frac{\partial \ln L_i}{\partial \beta^{(t)}} \right)^T,$$

где $\ln L_i$ — значение логарифмической функции правдоподобия, вычисленное в i-ом наблюдении.

Тогда уравнение перехода будет иметь вид:

$$\beta^{(t+1)} = \beta^{(t)} + \left(\sum_{i=1}^{n} \frac{\partial \ln L_i}{\partial \beta^{(t)}} \left(\frac{\partial \ln L_i}{\partial \beta^{(t)}} \right)^T \right)^{-1} \frac{\partial \ln L}{\partial \beta^{(t)}}.$$

Этот метод также называют модифицированным методом scoring.

Оценки параметров модели, найденные методом максимального правдоподобия, асимптотически подчиняются нормальному распределению:

$$\hat{\beta} \overset{a}{\sim} N(\beta, V(\hat{\beta})),$$

где $V(\hat{\beta})$ — ковариационная матрица вектора $\hat{\beta}$. Например, для модели с двумя независимыми переменными ковариационная матрица будет выглядеть следующим образом:

$$V\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} = \begin{pmatrix} V(\hat{\beta}_0) & Cov(\hat{\beta}_0, \hat{\beta}_1) & Cov(\hat{\beta}_0, \hat{\beta}_2)\\ Cov(\hat{\beta}_1, \hat{\beta}_0) & V(\hat{\beta}_1) & Cov(\hat{\beta}_1, \hat{\beta}_2)\\ Cov(\hat{\beta}_2, \hat{\beta}_0) & Cov(\hat{\beta}_2, \hat{\beta}_1) & V(\hat{\beta}_2)\end{pmatrix},$$

где $Cov(\hat{\beta}_i, \hat{\beta}_j)$ — ковариация оценок $\hat{\beta}_i$ и $\hat{\beta}_j$.
Вышеописанные вычислительные методы позволяют попутно получать оценку для ковариационной матрицы $V(\hat{\beta})$, которая будет необходима для проверки статистических гипотез о значении параметров регрессии.

По свойству асимптотической нормальности и асимптотической эффективности оценки максимального правдоподобия $\hat{\beta}$ ковариационная матрица $V(\hat{\beta})$ равна:

$$V(\hat{\beta}) \approx I^{-1}(\beta) = \left(-E \left[\frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} \right] \right)^{-1}, \tag{19}$$

т. е. ковариационная матрица равна обратному значению информационной матрицы (ожидаемое значение гессиана, взятое с противоположным знаком).

Выражение в правой части (19) вычисляется в точке β, но, поскольку мы знаем только оценку $\hat{\beta}$ параметра β, то можем найти только оценку для асимптотической ковариационной матрицы, т.е. производные и интегралы вычисляются при $\beta = \hat{\beta}$.

Можно использовать следующие три способа нахождения оценки $\hat{V}(\hat{\beta})$ асимптотической ковариационной матрицы $V(\hat{\beta})$:

1. $$\hat{V}_1(\hat{\beta}) = \left(-E\left[\frac{\partial^2 \ln L}{\partial \hat{\beta} \partial \hat{\beta}^T} \right] \right)^{-1}. \quad (20)$$

   Оценка (20) часто используется вместо с методом scoring, т.к. этот численный метод на каждой итерации вычисляет информационную матрицу.

2. $$\hat{V}_2(\hat{\beta}) = -\left(\sum_{i=1}^{n} \frac{\partial^2 \ln L_i}{\partial \hat{\beta} \partial \hat{\beta}^T} \right)^{-1}. \quad (21)$$

   Оценка (21) чаще всего используется вместе с численным методом Ньютона.

3. $$\hat{V}_3(\hat{\beta}) = \left(\sum_{i=1}^{n} g_i^2 x_i x_i^T \right)^{-1}, \quad (22)$$

   где $g_i = y_i - \Lambda(\beta x_i)$ для логит модели (см. (15)), и $g_i = q_i \varphi(\beta x_i)/\Phi(q_i \beta x_i)$, $q_i = 2y_i - 1$ (см. (17)) для пробит модели.

В некоторых случаях возможен альтернативный подход к оцениванию неизвестных параметров. Рассмотрим его на примере логит модели. Предположим, что для каждого набора факторов $x_l^T = (1, x_{1l}, \ldots, x_{nl})$ проведено несколько наблюдений, или все наблюдения сгруппированы таким образом, что внутри каждой группы значения факторов меняются мало. Тогда можно заменить различные наборы факторов внутри каждой группы некоторыми средними значениями и, соответственно, все наблюдения внутри группы рассматривать как наблюдения, соответствующие выбранным средним значениям факторов. В обоих случаях появляется возможность оценить эмпирическую вероятность появления единичного значения для соответствующего набора факторов — пусть это будет оценка $\hat{p}_l$, получаемая как относительная частота (доля) наблюдений, равных единице.

Тогда можно применить так называемое логит-преобразование $z_l = \ln(\hat{p}_l/(1-\hat{p}_l))$, после которого, учитывая модель (4) с логистической функцией распределения (8), получаем новую модель наблюдений:

$$z_i = \beta^T x_i + \zeta_i, \tag{23}$$

где относительно случайной компоненты ζ_i сделаем традиционные предположения о взаимной независимости наблюдений.

Модель (23) можно анализировать как модель наблюдений множественной регрессии, и для оценивания коэффициентов модели применить метод наименьших квадратов. Трудности применения метода наименьших квадратов к модели (23) связаны с тем, что дисперсии наблюдений не постоянны, и выражения для дисперсий случайных компонент носят приближенный характер. Аналогичный подход возможен и для случая пробит модели, где вместо логит-преобразования применяют обратную функцию к функции стандартного нормального распределения. Возникающие проблемы аналогичны выше рассмотренным.

Сравнивая два подхода к оценке неизвестных параметров модели (4), можно согласиться с тем, что метод максимального правдоподобия выглядит предпочтительнее. Каким бы способом ни оценивались параметры модели (4), о качестве построенной модели можно судить по ее способности правильно прогнозировать имеющиеся наблюдения. Подставляя в модель оценку $\hat{\beta}$ и значения факторов x_i, находим оценку вероятности появления единицы: $\hat{P}\{y_i = 1\} = F(\hat{\beta}^T x)$. Если наблюдение оказалось равным единице, то для правильного прогноза найденная вероятность должна принимать значение, большее 0.5. Перебирая имеющиеся наблюдения и определяя соответствие наблюдения вычисленной вероятности, можно оценить качество построенной модели.

Для логит и пробит моделей проверка гипотез о наличии ограничений на коэффициенты, в частности, гипотез о значимости одного или группы коэффициентов, может проводиться с помощью любого из трех критериев — Вальда, отношения правдоподобия, множителей Лагранжа.

Рассмотрим нулевую гипотезу в виде системы уравнений:

$$H_0: \quad Q\beta = r, \tag{24}$$

где $\beta^T = (\beta_0, \beta_1, \ldots, \beta_k)$, Q — матрица констант, r — вектор констант, которые формируются определенным образом в зависимости от того, какую гипотезу необходимо проверить. Например, рассмотрим пробит модель $P\{y = 1\} = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$. Для проверки нулевой гипотезы $H_0 : \beta_1 = 0$ система уравнений (24) примет следующий вид:

$$\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix}.$$

Для проверки гипотезы $\beta_1 = \beta_2 = 0$ система уравнений (24) примет следующий вид:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Сначала рассмотрим случай, когда $q = 1$, где q — число строк матрицы Q. Пусть в строке все элементы, кроме одного равного единице, равны нулю. В этом случае исследователю требуется проверить гипотезу о значении одного параметра модели, т. е. гипотезу H_0: $\beta_i = \beta^*$, где β^* — некоторое число (часто равное 0). Так как среднеквадратичное отклонение оценки $\hat{\beta}_i$ параметра β_i неизвестно, то для него может быть найдена оценка методами, описанными ранее.

Будем предполагать, что оценки неизвестных параметров асимптотически нормальны. В качестве статистики критерия рассмотрим функцию:

$$z = \frac{\hat{\beta}_i - \beta^*}{\hat{\sigma}(\hat{\beta}_i)}, \tag{25}$$

которая в случае справедливости гипотезы H_0 асимптотически подчиняется стандартному нормальному распределению, если оценки параметров в модели производятся методом максимального правдоподобия.

Алгоритм критерия

1. Выдвигаем нулевую гипотезу $H_0 : \beta_i = \beta^*$. Сформулируем альтернативную гипотезу: $H_1 : \beta_i \neq \beta^*$.
2. Задаем уровень значимости α.
3. Вычисляем значение статистики z по формуле (25).
4. Находим критическую область — это объединение интервалов $(-\infty; -u_{1-\frac{\alpha}{2}}) \cup (u_{1-\frac{\alpha}{2}}; \infty)$, где $u_{1-\alpha/2}$ — квантиль стандартного нормального распределения уровня $1 - \frac{\alpha}{2}$.
5. Если значение статистики (25) попадет в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований ее отвергнуть при уровне значимости α.

Замечание. В некоторых прикладных статистических пакетах в качестве асимптотического распределения статистики (25) используется распределение Стьюдента с числом степеней свободы, равным $n - k$. Кроме того, следует отметить, что критерий носит асимптотический характер, истинный уровень значимости близок к α.

Далее рассмотрим основные критерии, используемые для проверки значимости коэффициентов бинарной регрессии для случая, когда $q \geqslant 1$.

8. Критерий Вальда

Выдвинем нулевую гипотезу $H_0 : Q\beta = r$ при альтернативной гипотезе $H_1 : Q\beta \neq r$.
Пусть мы нашли оценку максимального правдоподобия $\hat{\beta}$ для неизвестного параметра β, и $\hat{V}(\hat{\beta})$ — состоятельная оценка для асимптотической ковариационной матрицы $V(\hat{\beta})$. Статистика критерия Вальда выглядит следующим образом:

$$W = (Q\hat{\beta} - r)^T (Q\hat{V}(\hat{\beta})Q^T)^{-1}(Q\hat{\beta} - r). \tag{26}$$

При справедливости нулевой гипотезы статистика (26) асимптотически подчиняется распределению χ^2 с числом степеней свободы, равным количеству тестируемых параметров, т. е. равным q .

Пример. Для пробит модели $P\{y = 1\} = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$ проверим гипотезу $H_0 : \beta_1 = \beta^*$ при альтернативной гипотезе $H_1 : \beta_1 \neq \beta^*$ при заданном уровне значимости α. Найдем выражение для W из (26) для этого случая.

Критерий Вальда

Выражение $Q\hat{\beta} - r$ примет следующий вид:

$$\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} - \beta^* = \hat{\beta}_1 - \beta^*.$$

Выражение $(Q\hat{V}(\hat{\beta})Q^T)^{-1}$ в статистике (26) будет следующим:

$$\left[\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \hat{V}(\hat{\beta}) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right]^{-1} = \frac{1}{\hat{V}(\hat{\beta}_1)}.$$

Критерий Вальда

Получаем выражение для статистики (26):

$$W = \frac{(\hat{\beta}_1 - \beta^*)^2}{\hat{V}(\hat{\beta}_1)} = \left(\frac{\hat{\beta}_1 - \beta^*}{\hat{\sigma}(\hat{\beta}_1)} \right)^2, \tag{27}$$

где $\hat{\beta}_1$ — оценка максимального правдоподобия, асимптотически нормальная и сильно состоятельная, $\hat{\sigma}(\hat{\beta}_1)$ — состоятельная оценка среднеквадратического отклонения $\hat{\beta}_1$. О способах нахождения $\hat{\beta}_1$ и $\hat{\sigma}(\hat{\beta}_1)$ говорилось ранее. Статистика (27) асимптотически подчиняется χ^2-распределению с одной степенью свободы в случае справедливости гипотезы H_0. Нетрудно заметить, что статистика W — квадрат значения статистики z в (25). В рассматриваемом случае гипотезу H_0 следует отвергнуть при уровне значимости приближенно равном α в случае, когда $W \in [\chi^2_{1-\alpha,1}; \infty)$, где $\chi^2_{1-\alpha,1}$ — квантиль уровня $1 - \alpha$ распределения χ^2 с одной степенью свободы.

Критерий Вальда

Для пробит модели $P\{y = 1\} = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$ проверим гипотезу $H_0 : \beta_1 = \beta_2 = 0$ при альтернативной гипотезе H_1, которая говорит о том, что хотя бы один из параметров (β_1, β_2) отличен от нуля, при уровне значимости приближенно равном α. Гипотеза H_0 может быть записана в следующем виде:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

где матрица $Q\hat{\beta} - r$ и $(Q\hat{V}(\hat{\beta})Q^T)^{-1}$ приобретают вид:

$$Q\hat{\beta} - r = (\hat{\beta}_1, \hat{\beta}_2)^T,$$

$$(Q\hat{V}(\hat{\beta})Q^T)^{-1} = \left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \hat{V}(\hat{\beta}) \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \right]^{-1}.$$

Критерий Вальда

Предположив, что оценки $\hat{\beta}_1$, $\hat{\beta}_2$ не коррелируют (в общем случае это неверно), получаем равенства:

$$(Q\hat{V}(\hat{\beta})Q^T)^{-1} = \begin{pmatrix} \hat{\sigma}^2(\hat{\beta}_1) & 0 \\ 0 & \hat{\sigma}^2(\hat{\beta}_2) \end{pmatrix}^{-1} = \begin{pmatrix} 1/\hat{\sigma}^2(\hat{\beta}_1) & 0 \\ 0 & 1/\hat{\sigma}^2(\hat{\beta}_2). \end{pmatrix}$$

Тогда статистика Вальда (26) примет вид:

$$W = \sum_{i=1}^{2} \left(\frac{\hat{\beta}_i}{\hat{\sigma}(\hat{\beta}_i)} \right)^2 = z^2_{\hat{\beta}_1} + z^2_{\hat{\beta}_2}.$$

Здесь также статистика Вальда равна сумме квадратов статистик z, вычисленных для тестируемых параметров. В случае, когда оценки параметров коррелируют (что чаще всего встречается в реальных задачах), окончательный вид статистики W имеет более сложный вид.

Алгоритм критерия Вальда

1. Выдвигаем нулевую гипотезу $H_0 : Q\beta = r$. Сформулируем альтернативную гипотезу $H_1 : Q\beta \neq r$.
2. Задаем уровень значимости критерия α.
3. Находим оценку $\hat{\beta}$ для неизвестного параметра β и оценку $\hat{V}(\hat{\beta})$ для асимптотической ковариационной матрицы.
4. Вычисляем значение статистики W по формуле (26).
5. Находим критическую область — интервал $(\chi^2_{1-\alpha,q}; \infty)$, где $\chi^2_{1-\alpha,q}$ — квантиль уровня $1-\alpha$ распределения χ^2 с q степенями свободы.
6. Если численное значение статистики W попадет в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований ее отвергнуть при уровне значимости приближенно равном α.

Критерий Вальда носит асимптотический характер, и, поэтому, уровень значимости критерия должен быть близок к α при больших объемах наблюдений.

9. Критерий отношения правдоподобия

Часто для проверки адекватности пробит и логит моделей бинарных регрессий используют критерии, основанные на сравнении значений функции правдоподобия в случае, когда максимизация проводится по всем неизвестным параметрам, и при условии, что $Q\beta = r$. Пусть $\ln L_1$ — максимальное значение логарифмической функции правдоподобия (12) при условии, что максимизация производится по параметру β без ограничений на этот параметр; $\ln L_0$ — максимальное значение логарифмической функции правдоподобия (12) при условии, что $Q\beta = r$. Очевидно, что $\ln L_1 \geqslant \ln L_0$. Чем больше разность между значениями функций, тем более оправдано использование регрессионной пробит или логит модели. Статистика отношения правдоподобия (likelihood ratio) выглядит следующим образом:

$$LR = 2(\ln L_1 - \ln L_0), \tag{28}$$

которая при справедливости нулевой гипотезы асимптотически подчиняется распределению χ^2 с числом степеней свободы, равным q.

Алгоритм критерия правдоподобия

1. Выдвигаем нулевую гипотезу $H_0 : Q\beta = r$. Сформулируем альтернативную гипотезу $H_1 : Q\beta \neq r$.
2. Задаем уровень значимости α.
3. Находим значение функции правдоподобия $\ln L_1$ в точке $\hat{\beta}$, которая является оценкой максимального правдоподобия для неизвестного параметра β в задаче без ограничений и $\ln L_0$.
4. Вычисляем значение статистики по формуле (28).
5. Находим критическую область — интервал $(\chi^2_{q,1-\alpha}; \infty)$, где $\chi^2_{1-\alpha,q}$ — квантиль уровня $1-\alpha$ распределения χ^2 с q степенями свободы, где q — число строк матрицы Q.
6. Если численное значение статистики (28) попадет в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований ее отвергнуть при уровне значимости приближенно равном α.

Критерий отношения правдоподобия носит асимптотический характер, уровень значимости критерия должен быть близок к α при больших n.

10. Критерии адекватности моделей бинарной регрессии

Сумма квадратов остатков SSR вычисляется по формуле:

$$SSR = \sum_{i=1}^{n} (y_i - \hat{F}_i)^2, \tag{29}$$

где $\hat{F}_i = F(\hat{\beta}^T x_i)$. Значение SSR является часто используемой мерой, поскольку она используется для вычисления коэффициента детерминации R^2 в моделях линейной регрессии. Тем не менее, использование этой меры не может быть математически строго обосновано, поскольку модели бинарной регрессии не удовлетворяют условию равенства дисперсий. B. Efron предложил аналог R^2 следующего вида:

$$R^2_{Ef} = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{F}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{30}$$

где $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$.

Критерии адекватности моделей бинарной регрессии

Взвешенная сумма квадратов WSSR для моделей бинарной регрессии может быть вычислена по формуле:

$$WSSR = \sum_{i=1}^{n} \frac{(y_i - \hat{F}_i)^2}{\hat{F}_i(1 - \hat{F}_i)}. \tag{31}$$

Квадратичный коэффициент корреляции SCC вычисляется по формуле:

$$SCC = \frac{\left[\sum_{i=1}^{n}(y_i - \bar{y})\hat{F}_i\right]^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{F}_i - \bar{F})^2}, \tag{32}$$

где $\bar{F} = \sum_{i=1}^{n} \hat{F}_i / n$. В случае линейной регрессии коэффициенты (30) и (32) совпадают, но это не является верным для случая бинарной регрессии.

Критерии адекватности моделей бинарной регрессии

Существует еще одна мера адекватности моделей бинарной регрессии:

$$R^2_{BL} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i\hat{F}_i + (1-y_i)(1-\hat{F}_i)\right), \tag{33}$$

которая представляет собой среднюю вероятность правильного предсказания в соответствии с полученным правилом.

Существуют меры адекватности моделей бинарной регрессии, основанные на сравнении значений функции правдоподобия при различных ограничениях. Например, D. MacFadden предложил индекс отношения правдоподобия следующего вида:

$$LRI = 1 - \frac{\ln L(\hat{\beta})}{\ln L_0}, \tag{34}$$

где $\ln L_0$ — максимальное значение логарифмической функции правдоподобия при $\beta_1 = \cdots = \beta_k = 0$.

Критерии адекватности моделей бинарной регрессии

С практической точки зрения полезно составить таблицу 2×2 истинных и ложных прогнозов по следующему правилу:

$$\hat{y} = \begin{cases} 1, & \text{если } \hat{F} > F^\star, \\ 0, & \text{в противном случае.} \end{cases} \tag{35}$$

Обычно $F^\star$ выбирают равным 0.5, что означает, что мы прогнозируем 1 в случае, когда модель «говорит», что 1 более вероятна, чем 0. Например, в таблице 2

	$\hat{y}_i = 1$	$\hat{y}_i = 0$	Кол-во
$y_i = 1$	271	14	285
$y_i = 0$	17	58	75
	288	72	**360**

количество верных прогнозов равно $271 + 58 = 329$, т. е. модель, которой соответствует таблица 2, корректно прогнозирует 91% данных.

Критерии адекватности моделей бинарной регрессии

Различные скалярные меры адекватности моделей бинарной регрессии дают различные результаты. Оценка максимального правдоподобия, на которой основаны все выше перечисленные скалярные меры адекватности для моделей бинарной регрессии, не выбирается из условия максимизации критерия адекватности, в отличие от классической модели линейной регрессии (коэффициенты регрессии, найденные методом наименьших квадратов, максимизируют коэффициент детерминации R^2). В случае бинарной регрессии оценка максимального правдоподобия $\hat{\beta}$ максимизирует совместную плотность распределения наблюдаемых случайных величин. Возникает вопрос для исследователя: выбрать лучшую оценку параметров при возможно низком уровне достоверного прогноза или получить наилучшую оценку параметров, максимизирующую выбранную скалярную меру адекватности модели, которая чаще всего не будет являться оценкой максимального правдоподобия?

Спрогнозировать числовое значение урожайности на конкретном поле чрезвычайно трудно, даже имея самый исчерпывающий набор факторов.
Легче спрогнозировать возможность того, что урожайность превысит или не превысит некоторое фиксированное пороговое значение, определяющее хороший или допустимый уровень урожайности культуры для данного региона.
Любой набор факторов не может исчерпывающим образом описать все возможные взаимосвязи, существующие в природе и влияющие на урожайность культуры, поэтому любой прогноз в принципе может носить только вероятностный характер, что обосновывает применение стохастической методологии на основе использования бинарной регрессии.

Предположим, что поле разбито на несколько однородных зон.
Каждую из зон разобьем на элементарные участки равной площади, которые перенумеруем.
Зададим некоторый пороговый уровень урожайности.
Введем бинарную переменную y_i, (где $i-$ номер элементарного участка внутри данной зоны однородности), по следующему правилу: если на данном участке урожайность оказалась выше порогового значения, то будем считать, что бинарная переменная принимает значение равное единице, в противном случае бинарная переменная принимает значение ноль.
Таким образом, для каждого элементарного участка i внутри однородной зоны определено значение бинарной переменной y_i, равное единице или нулю. Выберем набор почвенных и климатических характеристик $x_1, x_2, \ldots, x_k$, определяющих по нашему мнению урожайность культуры.

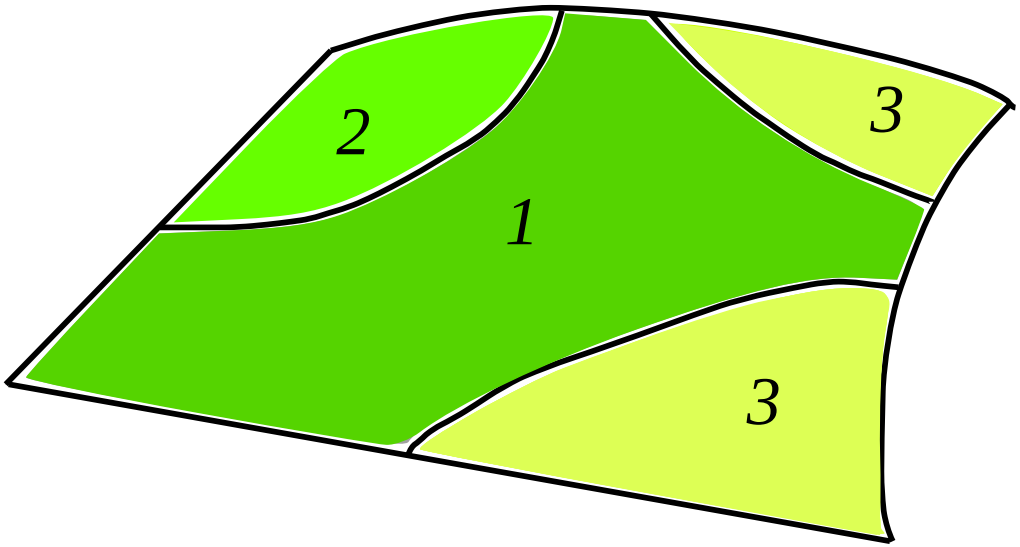

Рис. 1. Зоны однородности

Выбор факторов, оказывающих значимое влияние на урожайность культуры, не полностью очевиден. Может оказаться, что среди выбранных факторов некоторые не оказывают большого влияния на урожайность, и поэтому могут быть исключены из рассмотрения, тем более в ситуации, когда измерение этих факторов вызывает затруднения. В связи с этим представляет интерес задача определения минимально достаточного набора факторов, по которым возможно удовлетворительное прогнозирование урожайности культуры.

Для каждого элементарного участка с номером i, в результате проведения соответствующих измерений на элементарном участке получаем сопоставление набора объясняющих переменных $x_{i1}, x_{i2}, \ldots, x_{ik}$ и результирующей переменной y_i.
Расположенные рядом элементарные участки вряд ли могут считаться статистически независимыми. Однако, при наличии большого количества таких участков внутри зоны однородности, можно произвести «прореживание» элементарных участков так, чтобы исключить наличие соседних участков в «прореженной» совокупности.
В результате приходим к массиву данных, состоящему из наблюдений, которые можно с некоторым приближением считать статистически независимыми.
По полученному после «прореживания» массиву данных можно построить бинарную регрессию (логит или пробит регрессию).
Построенная эмпирическая модель позволит оценить вероятность превышения порогового уровня урожайности при данных значениях объясняющих переменных.

Для действительно однородной зоны значения характеристик $x_1, x_2, \ldots, x_k$ должны меняться незначительно и, следовательно, при изменении характеристик внутри зоны однородности вероятность, рассчитываемая по построенной модели, также должна меняться мало.

Однако на практике значения характеристик могут быть подвержены существенным изменениям, при сильном варьировании характеристик внутри зоны вариабельность вероятности может оказаться большой, что будет означать отсутствие однородности данной зоны.

Зная пределы изменения объясняющих переменных внутри данной зоны однородности, можно будет изучить вариабельность эмпирической вероятностной модели внутри этой зоны. Кроме того, можно будет получить некоторое усредненное значение для этой вероятности, характеризующие данную зону однородности в целом.

Подобное исследование должно быть проведено для каждой зоны однородности, после чего можно будет провести сравнение построенных эмпирических моделей.
Такое сравнение позволит оценить степень изменчивости вероятности урожайности между разными зонами однородности.
В принципе, подобная модель может быть построена для всего поля в целом, однако точность такой эмпирической модели может оказаться не достаточно высокой.
Построение отдельных моделей для каждой из зон должно привести к повышению точности и достоверности результатов.

Литература

1) Якушев В. П. , Якушев В. В. Информационное обеспечение точного земледелия. Рос. акад. с.-х. наук, Агрофиз. науч.-исслед. ин-т. - Санкт-Петербург : Изд-во ПИЯФ РАН, 2007.
2) Greene W. H. Econometric Analysis, 5th edition. – New Jearsey: Pearson Education, 2003.
3) Amemiya T. Advanced Econometrics. – Cambridge: Harvard University Press, 1985.
4) Ben-Akiva M., Lerman S. Discrete choice analysis – The MIT Press, Cambridge Massachusetts, 1985.
5) Буре В. М., Парилина Е. М. Теория вероятностей и математическая статистика-СПб.: Изд-во «Лань», 2013.
6) Буре В. М. Методология статистического анализа данных – СПб.: Изд-во С.- Петерб. Ун-та, 2007.

Литература

7) Stan G. Daberkow, William D. McBride Farm and operator characteristics affecting the awareness and adoption of precision agriculture technologies in US // Precision agriculture, 4, 163-177, 2003.
8) Robertson, M. J., Llewellyn R. S., Mandel R., Lawes R., Bramley R. G. V., Swift L., Metz N., O'Callaghan C. Adoption of variable rate fertilizer application in the Australian grain industry: status, issues and prospects // Precision agriculture, 13, 181-199, 2012.
9) J. Anita Dille, David A. Mortensen, Linda J. Young Predicting weed species occurrence based on site properties and previous year's weed presence // Precision agriculture, 3, 193-207, 2002.
10) Dale R. Walters, Anna Avrova, Ian J. Bingham, Fiona J. Burnett, James Fountain, Neil D. Havis, Stephen P. Hoad, Gareth Hughes, Mark Loosely, Simon J. P. Oxley, Alan Renwick, Cairistiona F. E. Topp, Adrian C. Newton Control of foliar diseases in barley: towards an integrated approach // Eur. J. Plant Pathol, 133, 33-73, 2012.

СПАСИБО ЗА ВНИМАНИЕ!