УДК 631:362.7

ОСОБЕННОСТИ ОБРАБОТКИ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ ПРОЦЕССА СУШКИ ЗЕРНА

П.С. Агеев, магистрант 1 года обучения;
С.А. Сутягин, кандидат технических наук, доцент;
Г.В. Карпенко, кандидат технических наук, доцент;
А.А. Павлушин, доктор технических наук, доцент;
В.И. Курдюмов, доктор технических наук, профессор
тел.: 89050359200, andrejpavlu@yandex.ru
ФГБОУ ВО Ульяновская ГСХА

Ключевые слова: методика обработки экспериментальных данных, критерий оптимизации, уравнения регрессии.

Рассмотрены основные этапы обработки экспериментальных данных. Охарактеризован метод наименьших квадратов.

Введение. Обработку данных, полученных по результатам выполненных исследований проводят, применяя корреляционно-регрессионный анализ.

При этом изучают и описывают зависимости между выходными характеристиками (результатами исследований) и входными параметрами (изменяемыми независимыми факторами). Кроме того, оценивают полученную функцию регрессии (математическую модель, характеризующую влияние основных независимых факторов процесса на критерий оптимизации) [1-3].

Материалы и методы исследований. Полученные функции регрессии (модельные уравнения регрессии) можно охарактеризовать функцией математического ожидания:

$$M(Y|X - x) - \varphi(x) - \beta_0 + \beta_1 x, \quad (1)$$

где *Y, X* – независимые случайные параметры, обладающие постоянной дисперсией и факторы наблюдаемого признака соответственно.

В указанной зависимости оценивали β_0 и β_1 – коэффициенты регрессии, а также значение остаточной дисперсии $\sigma^2_{ост}$.

За остаточную дисперсию принимают ту часть результативного признака, которую невозможно объяснить действием наблюдаемого признака. Это необходимо для выявления точности подбора функционального вида регрессии (модельного уравнения регрессии), полноты набора признаков, включённых в анализ полученных экспериментальных данных. Параметры функции регрессии находили, применяя способ наименьших квадратов, в основе которого лежит положение, что наилучший вид уравнения приближённой регрессии даёт та функция, для которой сумма квадратов отклонений имеет наименьшее значение [2]:

$$Q = \sum_{i=1}^{n}[y_i - f(x_i)]^2 \rightarrow min, \quad (2)$$

где $f(x_i)$ – принятый вид функции (модель уравнения регрессии).

При этом если $y = f(x,b_0,b_1,b_2,...,b_n)$ является дифференцируемой функцией и необходимо выявить численные значения коэффициентов входящих в её состав при ограничении, что

$$Q = \sum_{i=1}^{n}[y_i \quad f(x,b_0,b_1,b_2,...b_n)]^2 = min, \quad (3)$$

то важным условием минимума $\Phi = (b_0,b_1,b_2,...,b_n)$ служит соблюдение зависимости:

$$\frac{\partial Q}{\partial b_0} = 0; \frac{\partial Q}{\partial b_1} = 0; \frac{\partial Q}{\partial b_2} = 0 ... \frac{\partial Q}{\partial b_n} = 0, \quad (4)$$

или:

$$\begin{cases} \sum_{i=1}^{n} y_i \frac{\partial f(x_i)}{\partial b_0} - \sum_{i=1}^{n} f(x_i, b_0, b_1, b_2, ... b_n) \frac{\partial f(x_1)}{\partial b_0} = 0; \\ \sum_{i=1}^{n} y_i \frac{\partial f(x_i)}{\partial b_1} - \sum_{i=1}^{n} f(x_i, b_0, b_1, b_2, ... b_n) \frac{\partial f(x_1)}{\partial b_1} = 0; \\ \sum_{i=1}^{n} y_i \frac{\partial f(x_i)}{\partial b_2} \quad \sum_{i=1}^{n} f(x_i, b_0, b_1, b_2, ... b_n) \frac{\partial f(x_1)}{\partial b_2} = 0; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots; \\ \sum_{i=1}^{n} y_i \frac{\partial f(x_i)}{\partial b_n} \quad \sum_{i=1}^{n} f(x_i, b_0, b_1, b_2, ... b_n) \frac{\partial f(x_1)}{\partial b_n} = 0. \end{cases} \quad (5)$$

Полученная система является системой нормальных уравнений, так как количество зависимостей, входящих в её состав, соответствует количеству неопределённых коэффициентов уравнения регрессии ($b_0,b_1,b_2,...,b_n$).

В общем виде решить полученную систему уравнений проблематично. Для нахождения решения необходимо задать функцию *f*. Так, при определении коэффициентов уравнения вида $Y = b_0 + b_1 x_1$, при выборке *n*, система нормальных уравнений будет иметь вид [4]:

$$\begin{cases} \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} (b_0 + b_1 x_i) - 0; \\ \sum_{i-1}^{n} y_i x_i - \sum_{i-1}^{n} (b_0 + b_1 x_i) x_i = 0 \end{cases}, \tag{6}$$

или

$$\begin{cases} \sum_{i=1}^{n} y_i = n b_0 + b_1 \sum_{i=1}^{n} x_i; \\ \sum_{i-1}^{n} y_i x_i = b_0 \sum_{i-1}^{n} x_i + b_1 \sum_{i-1}^{n} x_i^2 \end{cases} \tag{7}$$

а коэффициенты b_0 и b_1 определяют из решения этой системы уравнений.

В случае нелинейной регрессии, когда зависимость действующих факторов нелинейна по переменным *x*, и описана видом:

$$M(Y|X - x) - \varphi(x) - \beta_0 + {}^{\beta_1}/_{x}, \tag{8}$$

оценкой выражения (8) служит уравнение регрессии:

$$y(x) - b_0 + b_1/x, \tag{9}$$

где b_0 и b_1 - оценки коэффициентов регрессии β_0 и β_1.

Используя указанный выше способ наименьших квадратов для нахождения коэффициентов, получим [5]:

$$Q - \sum_{i=1}^{n} \left| y_i - b_0 - {}^{b_1}/_{x_i}) \right|^2 - Q_{min}. \tag{10}$$

Дифференцируя равенство (4.10) по b_0 и b_1 и приравняв правые части полученных уравнений нулю, получим систему нормальных уравнений вида:

$$\begin{cases} \sum_{i-1}^{n} b_0 + \sum_{i-1}^{n} {}^{b_1}/_{x_i} = \sum_{i-1}^{n} y_i, \\ \sum_{i-1}^{n} {}^{b_0}/_{x_i} + \sum_{i-1}^{n} {}^{b_1}/_{x_i^2} = \sum_{i-1}^{n} {}^{y_i}/_{x_i}. \end{cases} \tag{11}$$

После получения и определения вида уравнения регрессий важным моментом является проверка значимости коэффициентов, входящих в её состав. Определяют, достаточна ли величина оценки для статистически аргументированного заключения о том, что тот или иной коэффициент регрессии отличен от нуля.

С этой целью проверяют гипотезу о равенстве нулю коэффициента регрессии, соблюдая предпосылки «нормальной регрессии». В этом случае вычисляемая для проверки нулевой гипотезы H_0: β = 0 величина

$$t = \left| {}^{b}/_{S_b} \right|$$

имеет распределение Стьюдента с $k = n$ -2 степенями свободы (b - оценка коэффициента регрессии, S_b - оценка среднего квадратического отклонения коэффициента регрессии - стандартная ошибка оценки). По уровню значимости α числу степеней свободы k находят по таблицам распределения Стьюдента критическое значение $t_{\alpha,k}$, удовлетворяющее условию [6]

$$P(|t| \geq t_{\alpha,k}) \geq t_{\alpha,k}. \tag{12}$$

Если $|t| \geq t_{\alpha,k}$, то нулевую гипотезу о равенстве нулю коэффициентов регрессии отвергали и коэффициенты считают значимыми. В случае $|t| \geq t_{\alpha,k}$ принимают нулевую гипотезу (коэффициенты регрессии незначимы).

Оценки среднего квадратического отклонения коэффициентов регрессии определяют по следующим зависимостям:

$$S_{b_0} = {}^{S_{ост}}/_{\sqrt{n-2}}, \quad S_{b_1} = \frac{S_{ост}}{S_x\sqrt{n-2}}, \tag{13}$$

где $S_x = \sqrt{\sum_{i=1}^{n} {}^{(x_i - x)^2}/_{n}}$.

При этом доверительный интервал для значимых параметров строят исходя из условия

$$P(-t_{\alpha,k} < {}^{(\beta - b)}/_{S_b}) < t_{\alpha,k}) = 1 - \alpha, \tag{14}$$

где α - уровень значимости.

Тогда

$$b - t_{\alpha,k}S_b < \beta < b + t_{\alpha,k}S_b. \tag{15}$$

Следует отметить, что полученная регрессия характеризует изменение условного математического ожидания результативного признака от вариации остальных признаков. Точечной же оценкой условного математического ожидания $M(Y|X = \chi)$ служит условное среднее $\bar{y}(x)$. Кроме точечной оценки для $M(Y|X = \chi)$ можно оценить доверительный интервал в точке $x = x_0$.

Если известно, что зависимость $(\bar{y}(x) - M(Y|X = x))/S_{y(x)}$ имеет распределение Стьюдента с $k = n - 2$ степенями свободы, то, определив среднее квадратическое отклонение для условного среднего, можно построить доверительный интервал для условного математического ожидания $M(Y|X = \chi)$ в точке $x = x_0$ [4].

После получения уравнения регрессии оценивают его значимость - устанавливают, соответствует ли полученная математическая модель, характеризующая зависимость между Y и X экспериментальным данным. Для оценки значимости в предпосылках «нормальной регрессии» проверяли гипотезу H_0: β = 0. Если она отвергается, то считают, что между Y и X нет связи (или связь нелинейная). Для проверки нулевой гипотезы используют положение дисперсионного анализа о разбиении суммы квадратов на слагаемые.

Для проверки нулевой гипотезы вычисляют статистику $F = (Q_1/Q_{\text{ост}})(k_2/k_1)$, которая имеет распределение Фишера - Снедекора с $k_1 = 1$, $k_2 = n - 2$ степенями свободы (n – число наблюдений).

По уровню значимости α и числу степеней свободы k_1 и k_2 находят по таблицам F - распределение критическое значение $F(\alpha,k_1,k_2)$, удовлетворяющее условию:

$$P\left(F > F_{(\alpha,k_1,k_2)}\right) \geq \alpha . \tag{16}$$

Если $F > F_{(\alpha,k_1,k_2)}$, нулевую гипотезу отвергают, полученное уравнение регрессии считают значимым. В случае $F < F_{(\alpha,k_1,k_2)}$ уравнение регрессии статистически незначимо.

В случае если изменения результативного признака характеризуются действием совокупности других признаков, имеет место многомерный регрессионный анализ.

Так, если Y - результативный признак, а $(\chi_1, \chi_2, ..., \chi_m)$ - независимые признаки, тогда для многомерного случая предпосылки регрессионного анализа можно сформулировать следующим образом.

Y - независимые случайные величины со средним $M(Y|X = \chi_1, \chi_2, ..., \chi_m)$ и постоянной дисперсией $\sigma^2_{\text{ост}}$, $(\chi_1, \chi_2, ..., \chi_m)$ - линейно независимые векторы $\chi(\chi_{11}, \chi_{12}, ..., \chi_n), ... \chi_m(\chi_{m1}, \chi_{m2}, ..., \chi_{mn})$.

Рассмотрим модель вида

$$M(Y|X = x_1, x_2, \dots x_m) = \beta_0 + \beta_1 x_1 + \cdots \beta_m x_m. \quad (17)$$

Оценке при этом подлежат параметры β_0, β_1, ..., βm и остаточная дисперсия. Заменив параметры их оценками, запишем уравнение регрессии

$$\overline{y}(x) = b_0 + b_1 x_1 + \cdots + b_m x_m. \quad (18)$$

Коэффициенты в этом выражении находят методом наименьших квадратов [3].

Оценкой остаточной дисперсии $\sigma^2_{\text{ост}}$ является

$$\sigma^2_{\text{ост}} = \{\textstyle\sum_{i=1}^{n}[y_i - y(x_i)]^2\}/(n-m), \quad (19)$$

где y_i - измеренное значение результативного признака; $y(x_i)$ - значение результативного признака, определённое из уравнения регрессии.

Если выборка получена из нормально распределённой генеральной совокупности, то можно проверить значимость оценок коэффициентов регрессии. В этом случае статистику $t = \left|b_j / S_{b_j}\right|$ вычисляем для каждого *j* - го коэффициента регрессии ($j = 1, 2, ..., m$):

$$S_{b_0} = S_{\text{ост}} \sqrt{1/n + \textstyle\sum_m x_i\, x_k C_{ij}^{-1}}, \quad S_{b_j} = S_{\text{ост}} \sqrt{C_{ij}^{-1}}, \quad i = 1, 2, \dots, n,$$

где C_{ij}^{-1} - элемент обратной матрицы, стоящий на пересечении *i* - й строки и *j* - го столбца; C_{ij}^{-1} - диагональный элемент обратной матрицы [6].

Заключение. Таким образом, при заданном уровне значимости α и числе степеней свободы k = n - m по соответствующим таблицам находят значение tα,k. Если расчётное значение $t > t_{\alpha,k}$, то оценку коэффициента принимают значимой. Подобную проверку осуществляют последовательно для каждого коэффициента регрессии. В случае если $t < t_{\alpha,k}$ - оценку коэффициента регрессии считают незначимой.

Библиографический список

1. Карпенко Г.В. Обоснование теплофизических параметров установки для сушки зерна контактного типа / Г.В. Карпенко, В.И. Курдюмов, А.А. Павлушин, М.А. Карпенко // Научное обеспечение устойчивого функционирования и развития АПК материалы Всероссийской научно-практической конференции с международным участием (в рамках XIX Международной специализированной выставки «АгроКомплекс-2009»). 2009. С. 84…87.
2. Курдюмов В.И. Совершенствование средств механизации переработки птичьего помета / В.И. Курдюмов, Н.Н. Аксенова, А.А. Павлушин, Е.В. Спирина // Аграрная наука и образование на современном этапе развития: опыт, проблемы и пути их решения Материалы IV Международной научно-практической конференции. 2012. С. 80…83.
3. Курдюмов В.И. Тепловая обработка зерна при подготовке комбикорма для поросят // В.И. Курдюмов, А.А. Павлушин, Г.В. Карпенко, С.А. Сутягин // Вестник Всероссийского научно-исследовательского института механизации животноводства. 2012.№ 3 (7). С. 102-107.
4. Курдюмов В.И. Тепловая обработка зерна в установках контактного типа // В.И. Курдюмов, А.А. Павлушин, Г.В. Карпенко, С.А. Сутягин: монография. – Ульяновск: УГСХА им. П.А. Столыпина, 2013. – 290 с.
5. Курдюмов В.И. Энергозатраты на процесс сушки зерна / В.И. Курдюмов В.И., А.А. Павлушин, С.А. Сутягин // Вестник ВИЭСХ. 2012. Т. 2. № 7. С. 52-54.
6. Курдюмов В.И. Теоретические и экспериментальные аспекты контактного способа передачи теплоты при сушке зерна / В.И. Курдюмов, А.А. Павлушин // Вестник Ульяновской государственной сельскохозяйственной академии. - 2011. - № 3. - С. 106-110.

INSTRUCTIONS RESULTS OF EXPERIMENTAL STUDIES OF GRAIN DRYING PROCESSES

Ageev P.S., Sutyagin S.A., Karpenko G.V., Pavlushin A.A., Kurdyumov V.I.

Key words: methods of experimental data processing, the optimization criterion, the regression equation.

The main stages of experimental data processing. Characterized by the method of least squares.