УДК 621.586

СТАТИСТИЧЕСКИЙ АНАЛИЗ И УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКИМ ПРОЦЕССОМ ПРОИЗВОДСТВА ИЗДЕЛИЙ

Асс. ***Мустафаев М.Г.,*** доц. ***Мустафаева Д.Г.***
Кафедра электронных приборов.
Северо-Кавказский горно-металлургический институт
(государственный технологический университет)

Рассмотрены вопросы статистического анализа технологического процесса производства интегральных элементов, на основе комплексного подхода с применением методов многомерной математической статистики, причинного анализа и принятия решений.

Статистический анализ технологического процесса (ТП) производства изделий осуществляется на основе комплексного подхода, при котором во взаимосвязи применяются методы многомерной математической статистики, причинного анализа и принятия решений.

При многомерном статистическом анализе производится:

– первичная статистическая обработка;

– установление статистических связей параметров;

– определение скрытых закономерностей технологической системы;

– построение регрессионных моделей связи выхода качественных изделий и режимами технологических операций (ТО);

– анализ устойчивости ТП.

Первичная статистическая обработка проводится для определения средних значений и дисперсий для параметров $\{x_i\}$, $i=\overline{1,n}$, с целью исключения результатов, выходящих за границы изменения x_i, проверяется вид закона распределения x_i.

Для статистических данных, распределенных по нормальному закону, проводится исключение аномальных наблюдений по критерию Граббса. При этом вычисляется расчетное значение критерия

$$G=\frac{\max\limits_{i=1,\dots,n}|x_i-\bar{x}|}{\bar{S}_x},$$

где $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ – среднее значение исследуемого параметра;

$\bar{S}_x^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-x)^2$ – дисперсия параметра;

n – количество наблюдений.

Если $G<G_{крит}=G_{таб}(\alpha,n)$,

где $G_{таб}$ – табличное значение критерия Граббса [1];

α – уровень значимости,

то в выборке нет аномальных значений.

При первичной статистической обработке исключаются неинформативные параметры. Для этого область изменения x_i разбивается на k интервалов. k выбирается в зависимости от физической природы параметра и может быть принято равным значению $[I+3{,}2\log n]$. Тогда энтропия по Шеннону, оценивающая информативность x_i, определяется как

$$H_i=-\sum_{j=1}^{k}p_{ij}\log p_{ij},$$

где p_{ij} – частота попадания x_i в j-й интервал.

Чем меньше значение H_i, тем менее информативным является x_i. Если $H_i<\gamma$, где γ – некоторое наперед заданное малое число, то x_i считается неинформативным и исключается из рассмотрения.

Существуют две причины информативности параметра x_i:

– x_i не характеризует изменение функционирования технологического процесса;

– методика измерения не отражает действительного изменения x_i.

Корреляционный анализ проводится для установления взаимосвязей между характеристиками ТП и выходом годных изделий [2]. Оценкой связи двух x_i и x_j служит парный коэффициент корреляции, вычисляемый по формуле

$$R_{ij} = \frac{\sum_{l=1}^{n}(x_{il} - \bar{x}_i)(x_{jl} - \bar{x}_l)}{\bar{S}_i \bar{S}_j}.$$

Значимость коэффициента корреляции проверяется с помощью статистики Стьюдента

$$t_{ij} = R_{ij}\sqrt{(n-2)(1-R^2 ij)}$$

с $k = n - 2$ степенями свободы.

Если $|t_{ij}| < t_{крит} = t_{таб}(\alpha, k)$,

где $t_{таб}$ – табличное значение распределения Стьюдента [3],

то R_{ij} считается незначимым, т. е. линейной зависимости в статистическом смысле нет.

Если исследуются параметры $x_1, ..., x_m$, то связи между ними описываются корреляционной матрицей:

$$Q_m = \begin{bmatrix} 1 & R_{12} & \dots & R_{1m} \\ R_{21} & 1 & \dots & R_{2m} \\ \dots & \dots & \dots & \dots \\ R_{m1} & R_{m2} & \dots & 1 \end{bmatrix}.$$

В этом случае зависимости между x_i, $i = \overline{1, m}$ более многообразны и сложны, чем в случае двух отдельно-рассматриваемых параметров. Для исследования зависимости между x_i и x_j при фиксированном значении $x_1, \dots, x_l$, $l = m - 2$ вычисляется частный коэффициент корреляции

$$R_{ij \cdot 1, \dots, m} = q_{ij} / \sqrt{q_{ii} q_{jj}},$$

где q_{ij}, q_{ii}, q_{jj} – алгебраические дополнения к элементам R_{ij}, R_{ii}, R_{jj} корреляционной матрицы Q_m.

Значимость частного коэффициента корреляции проверяется со степенями свободы $k = n - l - 2$.

Для выявления скрытых закономерностей ТС используют методы факторного анализа [4] . При этом каждое наблюдение x_{ij}, $i = \overline{1, m}$, $j = \overline{1, n}$, представляется в виде:

$$x_{ij} = \sum_{r=1}^{k} l_{ir} f_{rj} + e_{ij},$$

где f_{rj} – r-й фактор в j-м наблюдении;

k – количество факторов (k значительно меньше m);

l_{ij} – нагрузка r-го фактора в i-й переменной;

e_{ij} – остаток, действующий только на x_{ij}.

Для вычисления матрицы факторных нагрузок определяются собственные значения и соответствующие собственные векторы корреляционной матрицы. Затем определяются собственные значения λ_1, $\lambda_2, \dots, \lambda_k$, большие либо равные значению заданной константы (обычно единицы). Факторные нагрузки для i-й переменной j-го фактора вычисляются по формуле

$$a_{ij} = v_{ij}\sqrt{\lambda_j}, \; i = \overline{1, m}, \; j = \overline{1, k},$$

где λ_j – отдельные собственные значения Q_m;

v_{ij} – элементы соответствующих нормативных векторов.

После выявления взаимосвязей характеристик ТП на этапе корреляционного анализа строятся регрессионные модели [5].

Выходная оценка y технологического процесса представляется с помощью линейной функции регрессии в виде:

$$y = b_0 + \sum_{j=1}^{n} b_j x_j + e,$$

где b_0 – свободный член;

b_j, $j = \overline{0, n}$ – неизвестные параметры;

e – ошибка аппроксимации y посредством функции регрессии.

Если производится анализ устойчивости ТП необходимо сравнить векторы средних по T^2-критерию Хотеллинга [6] .

Рост степени интеграции и уменьшение геометрических размеров элементов приводят к повышению роли методов контроля настройки и точности выполнения каждой технологической операции, так как возрастает число и влияние факторов, определяющих качество изделий. В роли характеристик уровня настройки технологической операции и точности ее выполнения используются статистические показатели, характеризующие положение центра распределения измеренных параметров и их разброс, а именно, средние значения $\overline{x}$ и среднеквадратические отклонения $\overline{S}_x$.

Интегральные показатели уровня настройки ТП определяют по параметрам функционального элемента, а уровень настройки и точность выполнения отдельных операций – по параметрам элементов физической структуры.

Для обеспечения высокой точности ТП производства используется метод статистического регулирования. Управления ТП при статистическом регулировании производится с помощью контрольных карт, на которые наносятся значения измеряемых параметров. Откладывая на контрольной карте текущие значения статистических характеристик, и определяя их положение относительно границ регулирования, отслеживают изменение точности ТП. Среди методов статистического регулирования наибольшее распространение получили: метод средних значений и среднеквадратических отклонений; метод средних значений и размахов; метод медиан и крайних значений [7].

Использование одних лишь статистических взаимосвязей параметров не всегда дает надежные результаты при решении задач анализа, контроля и управления ТП. В большинстве статистических методов используются парные коэффициенты корреляции. При этом парный коэффициент корреляции R_{ij} двух переменных x_i и x_j является надежной оценкой их тесной статистической связи, когда на x_i и x_j не оказывают влияния другие переменные. В противном случае, используют частный коэффициент корреляции $R_{ij\cdot 1,\ldots,e}$, оценивающий связь x_i и x_j в системе переменных $\{x_S\}$, $S\ 1,\ldots,e$, при устранении воздействия других элементов системы. В случае, когда $R_{ij} \neq 0$, а $R_{ij\cdot 1,\ldots,e} = 0$, возникает явление, называемое ложной корреляцией x_i и x_j. При ложной корреляции связь x_i и x_j является "наведенной", возникающей вследствие их зависимости от других переменных.

Существенное различие матриц парных и частных коэффициентов корреляции исследуемых параметров дает основание для использования методов причинного анализа, позволяющих получить дополнительную информацию по сравнению со стандартными статистическими методами.

Причинный анализ является одним из направлений в статистических исследованиях [8], позволяющим решить две проблемы: выявление структуры причинно-следственных связей и количественная оценка выявленных связей. Для решения первой из них из совокупности статистических взаимосвязей параметров удаляются ложные корреляционные зависимости, оставшиеся связи интерпретируются как причинные. Разделение параметров на причины и следствия следует производить исходя из физической сущности параметров, соотношения времени их формирования и измерения. При этом x_i считается причиной x_j, если изменение x_i с высокой вероятностью сопровождается изменением x_j и не существует третьей переменной x_k, которую можно бы использовать для объяснения вероятностного соотношения между x_i и x_j.

После определения структуры причинно-следственных отношений и качественного характера выявленных причинных связей возникает задача их количественного оценивания. Среди методов количественного оценивания интенсивности причинных влияний известен путевой анализ, основные идеи которого разработаны Райтом. Данный метод основан на использовании линейных функциональных соотношений между параметрами. Использование линейных зависимостей между ними делает путевой анализ специальным случаем регрессивного анализа, в котором коэффициенты регрессии интерпретируются в терминах причинно-следственных отношений. Путевой коэффициент p_{ij}, измеряющий интенсивность причинного влияния x_i и x_j, определяется

$$p_{ij} = b_{ij}\frac{\overline{S}_i}{\overline{S}_j},$$

где $\overline{S}_i$ – среднеквадратическое отклонение x_i;

b_{ij} – коэффициент регрессии.

В этом случае предпочтительнее использовать метод, основанный на информационном подходе к определению мер причинного влияния – *IH*–анализ.

При анализе ТП производства интегральных элементов одновременно исследуется большое число параметров, причем каждый влияет на некоторое количество других. Отношения x_i образуют причинные системы, трудность изучения которых заключается в том, что исследуются причинно-следственные отношения, введенные в сеть других отношений.

При статистических исследованиях ТП производства интегральных элементов целесообразно применять одновременно методы причинного и факторного анализа для построения многоуровневых факторных моделей.

По результатам статистического и причинного анализа возникает задача принятия решений относительно управляющих воздействий на ТП производства изделий. Решение принимается с учетом полученных результатов и специфики наблюдений технологического процесса.

ЛИТЕРАТУРА

1. *Енюков И.С.* Методы, алгоритмы, программы многомерного статистического анализа. М.: Финансы и статистика, 1986.

2. *Бендат Д., Пирсол А.* Применение корреляционного и спектрального анализа. М.: Мир, 1983.

3. *Мюллер П., Нойман П., Шторм Р.* Таблицы по математической статистике. М.: Мир, 1982.

4. *Иберла К.* Факторный анализ. М.: Статистика, 1980.

5. *Себер Дж*. Нелинейный регрессивный анализ. М.: Мир, 1980.

6. *Афифи А., Эйзен С.* Статистический анализ. Подход с использованием ЭВМ. М.: Мир, 1982.

7. ГОСТ 15893-77. Статистическое регулирование технологических процессов при нормальном распределении контролируемого параметра.

8. *Хейс Д.* Причинный анализ в статистических исследованиях. М.: Финансы и статистика, 1981.