

УДК 347.823.21

DOI 10.51955/2312-1327_2025_1_6

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ОТНОСИТЕЛЬНО КОНТРОЛИРУЕМЫХ ПАРАМЕТРОВ ИЗДЕЛИЙ

*Николай Сергеевич Херсонский,
orcid.org/0000-0003-1296-7131,
кандидат технических наук,
генеральный директор ООО «СОЮЗСЕРТ»,
ул. Викторенко, д. 7, корпус 30
Москва, 125167, Россия
hersn@yandex.ru*

*Людмила Геннадьевна Большедворская,
orcid.org/0000-0002-1425-7398,
доктор технических наук, профессор
Московский государственный технический
университет гражданской авиации,
Кронштадтский бульвар, д. 20
Москва, 125493, Россия
l.bolshedvorskaya@mstuca.aero*

Аннотация. Метод корреляционного анализа не ограничивается нахождением тесноты связи между исследуемыми параметрами. Иногда он дополняется составлением уравнений регрессии, которые представляют собой описание корреляционной зависимости между результативным и факторными признаками. Такой подход получил название метод корреляционно-регрессионного анализа.

В данной статье, посвященной исследованию процессов изготовления изделий промышленного назначения, показано расширение возможностей и практической применимости методов корреляционно-регрессионного анализа для установления взаимосвязи и взаимозависимости параметров технологических процессов с целью повышения контроля за качеством готовой продукции.

Ключевые слова: корреляционно-регрессионный анализ, качество готовых изделий, контролируемые параметры технологических процессов.

CORRELATION ANALYSIS OF TECHNOLOGICAL PROCESSES RELATIVE TO CONTROLLED PRODUCT PARAMETERS

*Nikolai S. Khersonsky,
orcid.org/0000-0003-1296-7131,
candidate of technical sciences,
General Director of SOYUZCERT LLC,
7, building 30, Viktorenko St.
Moscow, 125167, Russia
hersn@yandex.ru*

*Lmila G. Bolshedvorskaya,
orcid.org/0000-0002-1425-7398,
Doctor of Technical Sciences, Professor
Moscow State Technical University of Civil Aviation,
20, Kronshtadtsky blvd
Moscow, 125493, Russia
l.bolshedvorskaya@mstuca.aero*

Abstract. The correlation analysis method is not limited to finding the closeness of the relationship between the studied parameters. Sometimes it is supplemented by the compilation of regression equations, which are a description of the correlation relationship between the resulting and factorial features. This approach is called the correlation and regression analysis method.

This article devoted to the study of the processes of manufacturing industrial products shows the expansion of the possibilities and practical applicability of correlation and regression analysis methods to establish the relationship and interdependence of technological process parameters in order to increase quality control of finished products.

Keywords: correlation and regression analysis, quality of finished products, controlled parameters of technological processes.

Введение

С годами не снижается актуальность применимости методов корреляционного анализа, результаты которого не ограничиваются нахождением тесноты связи между исследуемыми параметрами. Чаще всего интерес для исследователей сводится к составлению уравнения регрессии, которое представляет собой не только описание корреляционной зависимости между результативным и факторными признаками, но и позволяет оценивать эффективность изучаемого процесса в целом. Сложность и актуальность решения такой задачи для промышленных предприятий обусловлены наличием значительного количества факторных признаков, например, оборудование может существенно различаться по производительности, надежности, срокам службы, величиной затрат на ремонт и обслуживание, требованиями к технологическому процессу и к квалификации обслуживающего персонала.

В связи с этим, результатом данного исследования, посвященного изучению процессов изготовления изделий промышленного назначения, является обоснование практической применимости методов корреляционно-регрессионного анализа для установления взаимосвязи и взаимозависимости параметров технологических процессов с целью повышения контроля за качеством готовой продукции.

Материалы и методы

Корреляционный анализ в виде уравнений линейной регрессии нашел широкое применение в различных отраслях промышленности: в авиационной, оборонной, химической и других [Венецкий и др., 1974; Виниченко, 2023; Гирилович и др., 2021; Калугин, 2020; Кумэ, 1990; Херсонский, 2011; Яшин и др., 2015]. Особенно это связано с оценкой зависимости инновационного развития и экономического состояния промышленных предприятий на основе

корреляционного анализа в период санкционных ограничений и наличия лимитирующих факторов в виде сырья, материальных и трудовых ресурсов.

Проведенное исследование результатов научных достижений в области применения статистических методов корреляционного анализа позволило по-новому взглянуть на область его применимости, поскольку в ряде исследований полагают, что «уравнение регрессии» целесообразно заменять более строгим термином «уравнение корреляционной связи». Существует и другая точка зрения, предполагающая исследование взаимозависимости относить к теории корреляции, а изучение зависимости – к теории регрессии.

В связи с этим, можно предположить, что повышение эффективности технологических процессов создания готовой продукции промышленного назначения может быть достигнуто посредством интеграции методов корреляционного анализа в формате двух задач: построение уравнения регрессии и разработка модели корреляционной связи посредством последовательного решения следующих задач. Во-первых, определить тесноту связи между исследуемыми переменными. Во-вторых, определить зависимости между изучаемыми признаками путем:

- выявления факторов, оказывающих наибольшее влияние на результативный признак;
- выявления неизученных ранее причинно-следственных связей;
- построения корреляционной модели с ее параметрическим анализом;
- исследования значимости параметров связи и их интервальной оценки;
- осуществления перехода к разработке «уравнения корреляционной связи».

Теснота линейной связи между X и Y характеризуется коэффициентом корреляции, который вычисляется по формуле:

$$r_{xy} = \frac{1}{n} \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}, \quad (1)$$

где: x_i, y_i – текущие значения параметров;

$\bar{x}, \bar{y}$ – средние арифметические значения параметров;

σ_x, σ_y – средние квадратические отклонения параметров.

Через центрированные величины коэффициент корреляции выразится в виде:

$$r_{xy} = \frac{1}{n} \frac{\sum_{i=1}^n \tilde{x}_i \tilde{y}_i}{\sigma_x \sigma_y}, \quad (2)$$

где: $\tilde{x}_i = x_i - \bar{x}$ и $\tilde{y}_i = y_i - \bar{y}$ – центрированные величины.

С учетом понятия ковариации случайных величин

$$\text{cov}(x, y) = \dots \frac{1}{n} \dots \sum_{i=1}^n \tilde{x}_i \tilde{y}_i \dots \quad (3)$$

можно выражение (2) переписать в виде:

$$r_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}. \quad (4)$$

Коэффициент корреляции является важнейшим средством количественной оценки степени и характера взаимосвязи тех или иных явлений.

Чтобы определить пределы, в которых находится величина коэффициента корреляции, запишем его формулу в виде:

$$r_{xy} = -1 + \frac{1}{2n} \sum_{i=1}^n \left(\frac{\tilde{x}_i}{\sigma_x} + \frac{\tilde{y}_i}{\sigma_y} \right) \quad (5)$$

или

$$r_{xy} = 1 + \frac{1}{2n} \sum_{i=1}^n \left(\frac{\tilde{x}_i}{\sigma_x} - \frac{\tilde{y}_i}{\sigma_y} \right). \quad (6)$$

Из приведенных уравнений видно, что коэффициент корреляции заключен в пределах от -1 до +1. При значениях $|r_{xy}| = 1$ между параметрами Y и X существует линейная функциональная связь.

Для измерения тесноты связи в процентах вычисляют коэффициент детерминации по формуле $R = r_{xy}^2$, который показывает, какая часть дисперсии σ_y^2 случайной величины Y приходится на долю слагаемого ax в уравнении прямой линии ($y = ax + b$).

Результаты

На практике довольно часто приходится анализировать статистические данные, когда одной из переменных соответствует несколько значений другой переменной [Бешелев и др., 1980; Ефремов, 2020; Закс, 1976; Терентьев, 1959; Терентьев, 1960].

Коэффициент корреляции такого комплекса определяется по уравнению:

$$r = \sqrt{\frac{\left(\sum (A_i \rho_i - \bar{y}^* A_i m_i) \right)^2}{\left(\sum (A_i m_i - \bar{Z}^*) \right) \left(\sum y_i^2 - H \right)}} \quad (7)$$

где: A_i – условные группы случайных переменных, которым присваивается номер 0, 1, 2, 3 ...12;

ρ_i – сумма значений в данной клетке комплекса;

m_i – число значений в клетке.

В формуле (7) числитель подкоренного выражения можно выразить в виде

$$(\sum(A_i \rho_i) - \frac{\sum m_i A_i}{n} \sum y_i)^2. \quad (8)$$

Актуальность решения подобных задач для оценки качества и надежности изделий авиапромышленного назначения обусловлена переходом к композитным материалам на основе углепластика и других компонентов [Тишанинов, 2019]. Предположим, что готовое изделие должно состоять из трех компонентов А, В и С, ограниченных их структурным содержанием в конечном продукте: А (31% - 36%), В (21% - 26%), С (42% - 46%). Необходимо установить влияние технологического процесса на качество готовой продукции в зависимости от изменения соотношения компонентов. Эксперимент проводился с использованием трех видов смесителей и анализа структурного состава компонентов в готовой продукции посредством расчета коэффициентов корреляции между массовыми долями компонентов: $r_{xy}(z)$, $r_{xz}(y)$, $r_{zy}(x)$ (табл. 1).

Таблица 1 – Результаты расчетов коэффициентов корреляции между массовыми долями компонентов смеси D

№ эксперимента	Смеситель № 1			Смеситель № 2			Смеситель № 3		
	Массовые доли компонентов смеси, в %			Массовые доли компонентов смеси, в %			Массовые доли компонентов смеси, в %		
	x_i	y_i	z_i	x_i	y_i	z_i	x_i	y_i	z_i
1	33,5	23,5	43,0	33,6	22,8	43,6	33,5	23,2	43,3
2	34,2	22,2	43,6	34,9	22,4	42,7	31,2	23,5	45,3
3	31,5	23,4	45,1	31,6	23,5	44,9	32,8	22,7	44,5
4	32,2	22,9	44,9	34,0	22,9	43,1	31,2	23,6	45,2
5	33,6	23,1	43,3	32,6	23,4	44,0	34,2	22,8	43,0
6	32,6	22,6	44,6	31,7	23,5	44,8	33,1	22,9	44,0
7	31,6	23,8	44,6	31,8	24,2	44,0	31,7	24,1	44,2
8	32,0	22,6	45,4	33,4	22,5	44,1	32,4	22,8	44,8
9	32,8	22,8	44,4	31,5	23,9	44,6	31,5	23,3	45,2
10	32,0	23,1	44,9	31,3	22,8	45,9	31,6	22,9	45,5
	r_{xy}	r_{yz}	r_{xz}	r_{xy}	r_{yz}	r_{xz}	r_{xy}	r_{yz}	r_{xz}
	0,522	0,069	-0,860	0,525	0,181	-0,883	-0,664	-0,045	0,774

Для проведения расчетов составим вспомогательные таблицы для каждого смесителя (табл. 2-4).

Таблица 2 – Результаты при использовании в технологическом процессе смесителя № 1

X_i^1	$(X_i^1)^2$	ρ_i	m_i	$m_i X_i^1$	$m_i (X_i^1)^2$	$X_i^1 \rho_i$
0	0	326	10	0	0	0
1	1	230	10	10	10	230
2	4	444	10	20	40	888
Смеситель № 1	$\sum \rho_i = 1000$	$\sum m_i = 30$	$\sum m_i x_i^1 = 30$	$\sum m_i (X_i^1)^2 = 50$	$\sum \rho_i x_i^1 = 1118$	

Таблица 3 – Результаты при использовании в технологическом процессе смесителя № 2

X_i^1	$(X_i^1)^2$	ρ_i	m_i	$m_i X_i^1$	$m_i (X_i^1)^2$	$X_i^1 \rho_i$
0	0	326	10	0	0	0
1	1	232	10	10	10	232
2	4	442	10	20	40	884
Смеситель № 2	$\sum \rho_i = 1000$	$\sum m_i = 30$	$\sum m_i x_i^1 = 30$	$\sum m_i (X_i^1)^2 = 50$	$\sum \rho_i x_i^1 = 1116$	

Таблица 4 – Результаты при использовании в технологическом процессе смесителя № 3

X_i^1	$(X_i^1)^2$	ρ_i	m_i	$m_i X_i^1$	$m_i (X_i^1)^2$	$X_i^1 \rho_i$
0	0	323	10	0	0	0
1	1	232	10	10	10	232
2	2	445	10	20	40	890
Смеситель № 3	$\sum \rho_i = 1000$	$\sum m_i = 30$	$\sum m_i x_i^1 = 30$	$\sum m_i (X_i^1)^2 = 50$	$\sum \rho_i x_i^1 = 1122$	

Алгоритм расчета коэффициента корреляции статистического комплекса представлен на примере смесителя № 1, для остальных смесителей порядок расчета аналогичен.

$$\begin{aligned} \bar{y}^* &= \frac{\sum \rho_i}{n} = \frac{1000}{30} = 33,3 \\ Z^* &= \frac{\sum m_i A_i}{n} \\ H- &= \frac{(\sum \rho_i)^2}{n} = \frac{1000^2}{30} = 33333,3 \\ (\sum y_i)^2 &= 35501,8 \\ r_1 &= \sqrt{\frac{(1122,0 - 33,3)^2}{(30 - 30)(35501,8 - 33333,3)}} = 0,585. \end{aligned}$$

Коэффициенты корреляции в зависимости от номера смесителя приведены в таблице 5.

Таблица 5 – Значения коэффициентов корреляции в зависимости от номера смесителя

№ смесителя	Коэффициент корреляции в зависимости от номера смесителя Γ_j ($j = 1, 2, 3$)	Коэффициент корреляции между массовыми долями компонентов x и y $\Gamma_{xy}(z)$	Коэффициент корреляции между массовыми долями компонентов x и z $\Gamma_{xz}(y)$
1	$r_1 = 0,542$	-0,99	-0,99
2	$r_2 = 0,540$	-0,911	-0,968
3	$r_3 = 0,566$	-0,792	-0,968

Дискуссия

Проведенные расчеты и полученные результаты подчеркивают тесную корреляционную связь этапов технологического процесса, применяемого оборудования (смесителей) и изменения структурного состава получаемой смеси [Григорьева и др., 2018].

Если контролируемый параметр изделия Y корреляционно связан с двумя и более параметрами технологического процесса X и Z и все параметры связаны между собой, тогда корреляция между Y и Z и между X и Z будет влиять на корреляцию между X и Y .

Если величина контролируемого параметра Y в некоторой степени зависит от Z , то зависимость X от Y может частично отражать зависимость X от Z . Для того, чтобы устранить влияние Z , необходимо вычислить корреляцию между X и Y , когда Z – постоянно, т. е. определять частный коэффициент корреляции:

$$r_{xy(z)} = \frac{r_{xy} - r_{xz} r_{yz}}{\sqrt{1 - r_{xz}^2} \sqrt{1 - r_{yz}^2}}. \quad (9)$$

Для того, чтобы устранить влияние Y на XZ , необходимо определить частный коэффициент корреляции по формуле:

$$r_{xz(y)} = \frac{r_{xz} - r_{yx} r_{yz}}{\sqrt{1 - r_{yx}^2} \sqrt{1 - r_{yz}^2}}. \quad (10)$$

Для того, чтобы устранить влияние X на YZ , необходимо определить корреляцию между Y и Z , когда X постоянно:

$$r_{yz(x)} = \frac{r_{yz} - r_{yx} r_{xz}}{\sqrt{1 - r_{yx}^2} \sqrt{1 - r_{xz}^2}}. \quad (11)$$

Если величина контролируемого параметра изделия связана с четырьмя параметрами технологического процесса: X, Y, Z, T, тогда частный коэффициент корреляции между X и Y при условии, что Z и T будут постоянными, можно определить по следующей схеме:

Для этого составляется матрица вида:

$$\begin{matrix} X \\ Y \\ Z \\ T \end{matrix} \begin{vmatrix} 1 & r_{xy} & r_{xz} & r_{xt} \\ r_{yx} & 1 & r_{yz} & r_{yt} \\ r_{zx} & r_{zy} & 1 & r_{zt} \\ r_{tx} & r_{ty} & r_{tz} & 1 \end{vmatrix}, \quad (12)$$

из которой получают:

$$r_{xy(zt)} = \frac{r_{xy}(1-r_{zt}^2) - r_{xz}r_{yz} - r_{xt}r_{yt} + r_{zt}(r_{xz}r_{yt} + r_{xt}r_{yz})}{(1-r_{yt}^2 - r_{zt}^2 - r_{yz}^2 + 2r_{yz}r_{yz}r_{zt})^{\frac{1}{2}}(1-r_{xt}^2 - r_{zt}^2 - r_{xz}^2 + 2r_{xz}r_{xt}r_{zt})^{\frac{1}{2}}}. \quad (13)$$

Другие частные коэффициенты можно вычислить по такой же схеме.

В общем виде частные коэффициенты корреляции можно определять по общей формуле:

$$r_{ij} = r_{ji} = \frac{n\rho_{ij} - \rho_{i0}\rho_{0i}}{\sqrt{n\rho_{ii} - \rho_{i0}^2}\sqrt{n\rho_{jj} - \rho_{0j}^2}}, \quad (14)$$

где:

$$\begin{aligned} \sum x &= p_{01} = p_{10}; \sum x^2 = p_{11}; \sum xy = p_{12} = p_{21} \\ \sum y &= p_{02} = p_{20}; \sum y^2 = p_{22}; \sum xz = p_{13} = p_{31} \\ \sum z &= p_{03} = p_{30}; \sum z^2 = p_{33}; \sum xt = p_{14} = p_{41} \\ \sum t &= p_{04} = p_{40}; \sum t^2 = p_{44} \end{aligned}$$

Следует отметить, что как бы ни была сильна статистическая зависимость между различными параметрами технологического процесса и контролируемого параметра изделия, невозможно конкретизировать причинно-следственную связь.

С учетом результатов анализов смеси D, приведенных в таблице 1, были вычислены частные коэффициенты корреляции между массовыми долями компонентов $r_{xy}(z)$, $r_{xz}(y)$ на примере использования смесителя № 1.

$$r_{xy}(z) = \frac{-0,6641 - (-0,774)(-0,045)}{\sqrt{1 - (-0,774)^2}\sqrt{1 - (-0,045)^2}} = -0,99 \quad (15)$$

$$r_{xz}(y) = \frac{-0,744 - (-0,664)(-0,045)}{\sqrt{1 - (-0,664)^2} \sqrt{1 - (-0,045)^2}} = 0,99 \quad (16)$$

Расчеты показали, что между массовыми долями компонентов существуют сильные корреляционные связи, и это создает предпосылки о возможности прогнозирования качества смеси D.

Не менее значимыми являются задачи оценки корреляции между качественными и количественными параметрами технологического процесса и готового изделия.

Вычисление таких корреляций необходимо, когда параметры технологических процессов и контролируемых параметров изделия не имеют количественных значений, а определяются по альтернативному вероятностному признаку «да» – «нет», «годен» или «не годен».

Оценить выход отдельных параметров технологических процессов за пределы, установленные технологическим регламентом, не представляется возможным. Оценка параметров изделия может осуществляться посредством проверки его геометрических размеров с помощью проходных и непроходных скоб, колец или контроля физических параметров, таких как герметичность, и др.

Особо следует остановиться на определении корреляционных связей между контролируемыми параметрами изделий при их испытаниях.

Рассмотрим случай двух качественных признаков, распадающихся на две группы, на примере контроля диаметра (параметр X) и диаметра (параметр Y) изделия с помощью проходных колец. Результаты контроля параметров диаметров изделия приведены в таблице 6.

Таблица 6 – Результаты контроля по альтернативному признаку контролируемых параметров диаметров изделия

Контролируемый параметр диаметра изделия – Y	Контролируемый параметр диаметра изделия – X		Всего
	Доля дефектных изделий по контролируемому параметру X (%)	Доля годных изделий по контролируемому параметру X (%)	
Доля дефектных изделий по контролируемому параметру Y (%)	$a_1 = 10,0$	$a_2 = 5,0$	15,0
Доля годных изделий по контролируемому параметру Y (%)	$a_3 = 70,0$	$a_4 = 15,0$	85,0
Всего	80,0	20,0	100,0

Коэффициент корреляции можно вычислить по формуле:

$$r = \frac{a_1 a_4 - a_2 a_3}{\sqrt{(a_1 + a_2)(a_3 + a_4)(a_1 + a_3)(a_2 + a_4)}} = -0,14. \quad (17)$$

Рассмотрим случай разделения каждого признака на три группы с целью изучения влияния нескольких вариантов технологического процесса (B_i) на качество изделия по контролируемому параметру, например, принадлежность изделия к высшей, первой или второй категории качества (A_i) (табл. 7).

Таблица 7 – Влияние нескольких вариантов технологического процесса (B_i) на качество изделия по контролируемому параметру

Варианты технологического процесса B_i	Варианты по качеству изделия B_i .		
	A_1	A_2	A_3
	Высшая категория	1 категория	2 категория
B_1	$A_1 B_1$	$A_2 B_1$	$A_3 B_1$
B_2	$A_1 B_2$	$A_2 B_2$	$A_3 B_2$
B_3	$A_1 B_3$	$A_2 B_3$	$A_3 B_3$

Для вычисления коэффициента корреляции такого статистического комплекса может быть использована формула:

$$r = \frac{n a_1 - a_x a_y}{\sqrt{a_x a_y (n - a_x)(n - a_y)}}, \quad (18)$$

где: a_1 – частота, стоящая на пересечении соответственного столбца (A) с соответствующей строкой (B);

n – сумма всех значений в клетках, т.е. $\sum a_i$;

a_x – сумма всех частот соответственного столбца (A);

$a_y \dots$ – сумма частот соответствующей строки (B).

Рассмотрим случай определения корреляционных зависимостей между качественными и количественными признаками контролируемого параметра технологического процесса и изделия.

Коэффициент корреляции в этом случае определится по формуле:

$$r = \frac{\bar{y} - \bar{y}_1}{\sigma} \sqrt{\frac{n_1}{n - n_1}}, \quad (19)$$

где: $\bar{y} \dots$ – среднее арифметическое значение параметра технологического процесса для всей совокупности изделий, у которых был проконтролирован параметр по альтернативному признаку;

$\bar{y}_1$ – среднее арифметическое значение параметра технологического процесса для той совокупности изделий, которые по контролируемому параметру обладают определенным признаком, например, являются дефектными;

σ – среднее квадратическое отклонение параметра технологического процесса, вычисленное для всей совокупности изделий, у которых был проконтролирован параметр по альтернативному признаку;

n – общая численность совокупности изделий;

n_1 – число изделий, которые обладают определенным признаком по контролируемому параметру, например, являются дефектными.

Не менее значимой и актуальной является задача установления корреляционных связей между количественными и качественными признаками контролируемой продукции и технологических процессов [Моделирование процесса..., 1998; Надгериева и др., 2009].

Предположим, что требуется измерить зависимость между скоростью изготовления детали V , (сек) и ее качеством. Результаты исследований приведены в таблице 8.

По формуле (11) определим значение коэффициента корреляции между скоростью изготовления детали и ее качеством:

$$r = \frac{37,1 - 30,97}{5,66} \sqrt{\frac{90}{510 - 90}} = 0,5. \quad (20)$$

Можно сделать вывод, что между скоростью изготовления детали и ее качеством существует средняя теснота связи: причем, чем выше скорость изготовления детали V , тем выше ее качество.

Таблица 8 – Результаты исследований зависимости между скоростью изготовления детали V , (сек) и ее качеством

V, (сек)	Количество деталей, штук		
	годных	дефектных	всего
20-25	52	10	62
25-30	111	22	133
30-35	147	40	187
35-40	62	14	76
40-45	48	4	52
Всего	$n - n_1 = 420$	$n_1 = 90$	$n = 510$

При наличии нелинейной зависимости между параметрами технологического процесса и контролируемых параметров изделия

коэффициент может дать ошибочное представление о тесноте связи. В подобных случаях вместо коэффициента корреляции вычисляется корреляционное отношение – η .

Корреляционное отношение η – величина всегда положительная, колеблющаяся в пределах от 0 до 1 и дающая представление лишь о величине связи, причем характер этой связи в основном определяется формулой линии регрессии.

Корреляционное отношение вычисляется по формуле:

$$\eta = \sqrt{\frac{\sigma_y^2 - \sigma_{\bar{y}}^2}{\sigma_y^2}}, \quad (21)$$

$$\text{где: } \sigma_y^2 = \frac{1}{n} \sqrt{(y - \bar{y})^2} \quad \sigma_{\bar{y}}^2 = \frac{1}{n} \sqrt{\bar{y} - \bar{\bar{y}})^2} \quad \text{или } r^2 = 1 - \frac{\eta_y^2}{\sigma_y^2}.$$

В общем виде корреляционное отношение можно представить так:

$$\eta^2 = 1 - \frac{\sigma_{\bar{y}}^2}{\sigma_y^2} \quad \text{или } \eta^2 = \frac{\sigma_y^2 - \sigma_{\bar{y}}^2}{\sigma_y^2}, \quad (22)$$

где n – число измерений КП изделия или параметров технологического процесса.

Считается, что регрессия линейна, если $n(\eta^2 - r^2) < 11,37$.

Корреляционное отношение для статистического комплекса определяется уравнением:

$$\eta = \sqrt{\frac{\sum \frac{\rho_{ij}^2}{m_{ij}} - \frac{(\sum y_i)^2}{n}}{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}}, \quad (23)$$

где: ρ_{ij} – сумма значений клетки;

m_{ij} – количество значений клетки;

n – общее количество значений данного фактора.

Корреляционное отношение можно выразить через показатель достоверности: Q_η

$$\eta = \sqrt{\frac{Q_\eta(a-1)}{Q_\eta(a-1) + (n-a)}}, \quad (24)$$

где: a – число групп при числе степеней свободы $k_1 = a-1$, $k_2 = n-a$.

$$Q_{\eta} = \frac{(\sum \frac{\rho_{ij}^2}{m_{ij}} - \frac{(\sum y_i)^2}{n})(n-a)}{(\sum y_i^2 - \sum \rho_{ij}^2)(a-1)}. \quad (25)$$

На примере влияния различных видов оборудования и материала на стойкость режущего инструмента покажем вычисление корреляционного отношения. Для этого необходимо вычислить корреляционные отношения для групп материалов В₁ и В₂. Результаты экспериментов приведены в таблице 9.

Таблица 9 – Результаты экспериментов по оценке влияния различных видов оборудования и материала на стойкость режущего инструмента

Тип оборудования/ вид материала	A1	A2	A3	A4
B1	9 13 11	12 13 15	13 12 16	14 15 14
	14 13 13	12 10 11	15 16 12	12 15 16
	9 12 12	7 12 12	10 12 11	11 13 16
B2	9 11 11	8 15 12	14 16 11	14 10 13
	10 9 12	9 10 11	9 16 15	13 15 14
	10 9 9	7 8 10	11 11 10	12 15 12

Для сумм значений ρ_{ij} и m_{ij} составим новую таблицу.

Таблица 10 – Суммы значений ρ_{ij} и m_{ij}

Тип оборудования/ вид материала	A1	A2	A3	A4
B1	106	104	117	126
	9	9	9	9
B2	90	90	113	118
	9	9	9	9

По данным таблицы 10 определим коэффициент достоверности Q_{η} для групп В₁ по формуле (25).

$$Q_{\eta_{B_1}} = \frac{(5735 - \frac{205209}{36})}{(5861 - 5735)3} = 2,96.$$

По формуле (24) определим корреляционное отношение:

$$\eta_{B_1} = \sqrt{\frac{2,96(4-1)}{2,96(4-1) + (36-4)}} = 0,46.$$

Проведя аналогичные расчеты для групп В₂, получим:

$$Q_{\eta_{B_2}} = 5,76; \dots \eta_{B_2} = 0,59.$$

По результатам расчетов можно сделать вывод, что вид оборудования и материал оказывают влияние на стойкость режущего инструмента.

Для оценки значимости коэффициента корреляции для заданного числа измерений n КП изделия или параметров ТП и доверительной вероятности γ (обычно в технике $\gamma = 0,95$) вычисляют значение статистики Стьюдента по формуле:

$$t = \frac{|r_{xy}| \sqrt{n}}{1 - r_{xy}^2}. \quad (26)$$

Далее величина статистики t сравнивается с ее критическим значением $t_{кр}$, которое выбирается из таблицы работ в зависимости от величины γ и числа степеней свободы $\nu = n - 1$.

Если $t < t_{кр}$, то коэффициент корреляции считается несущественным и его появление объясняется случайностями выборки;

при $t > t_{кр}$, коэффициент корреляции можно считать существенным, а связь между соответствующими случайными параметрами X и Y достоверной.

В технике довольно часто используется следующая классификация тесноты корреляционных связей:

Если величина $|r_{xy}| < 0,2$ – между параметрами X и Y нет связи, при $0,2 > |r_{xy}| < 0,5$ – связь слабая.

Если величина $0,5 > |r_{xy}| < 0,75$ – связь средняя, при $0,75 > |r_{xy}| < 0,95$ – связь сильная.

Считается, что при $|r_{xy}| > 0,95$ между параметрами X и Y существует функциональная связь.

Для того, чтобы обеспечить заданную надежность определения коэффициента корреляции или корреляционного отношения, необходимо уметь определить число измерений контролируемых параметров или параметров технологического процесса, которое может определяться по формулам:

– при известном коэффициенте корреляции:

$$n^1 = \left(\frac{1 - r_{xy}^2}{r_{xy}^2} \right) (Q^1 + 2); \quad (27)$$

– при известном корреляционном отношении:

$$n^2 = \left(\frac{1 - \eta^2}{\eta^2} \right) (a - Q^1) + a, \quad (28)$$

где: Q^1 – величина, соответствующая определенной доверительной вероятности ($\gamma = 0,95$, $\gamma = 0,99$, $\gamma = 0,999$), выбирается из таблицы работы.

Следует отметить, что если связь между параметрами X и Y носит нелинейный характер, то коэффициент корреляции не может дать оценки зависимости между ними.

Если даже $r_{xy} = 0$, то это еще не свидетельствует о том, что между параметрами X и Y нет связи.

Выводы

Достоинством данного исследования является решение значимой и актуальной задачи установления корреляционных связей между количественными и качественными признаками контролируемой продукции и технологических процессов.

Результаты проведенного исследования подчеркивают новые, ранее не используемые возможности применения методов корреляционного анализа для исследования различных технологических процессов авиационной, оборонной, химической и других отраслей промышленности и их влияние на качество готовой продукции и изделий.

Библиографический список

- Бешелев Ф. Г. Математико-статистические методы экспертных оценок. 2 изд., перераб. и доп. / С. Д. Бешелев, Ф. Г. Гурвич. М.: Статистика, 1980. 263 с.
- Венецкий И. Г. Основные математико-статистические понятия и формулы в экономическом анализе / И. Г. Венецкий, В. И. Венецкая. М.: Статистика. 1974. 61 с.
- Виниченко А. В. Корреляционная матрица для сложноструктурированных поведенческих факторов и параметров технологического процесса // Инновационное приборостроение. 2023. Т. 2, № 4. С. 88–92. DOI 10.31799/2949-0693-2023-4-88-92. EDN BNTJOA.
- Гирилович Н. В. Применение статистических методов при анализе несоответствий несоответствующей продукции в процессе производства / Н. В. Гирилович, Г. В. Довгополая // Литье и металлургия. 2021. № 3. С. 40-45. DOI 10.21122/1683-6065-2021-3-40-45. EDN POAICU.
- Григорьева Т. А. Корреляционно-регрессионный анализ технологических параметров / Т. А. Григорьева, В. Н. Толубаев // Системы. Методы. Технологии. 2018. № 3 (39). С. 57–61. DOI 10.18324/2077-5415-2018-3-57-61. EDN YAUXHN.
- Ефремов Н. Ю. Применение регрессионного анализа для исследования влияния технологических параметров на показатели качества наполнения полимеров // Известия Самарского научного центра Российской академии наук. 2020. Т. 22. № 4 (96). С. 81-85. DOI 10.37313/1990-5378-2020-22-4-81-85. EDN RSBRVV.
- Закс Л. Статистическое оценивание. М.: Статистика. 1976. 598 с.
- Калугин Ю. Б. Использование корреляционного анализа для определения величины шага контроля при управлении технологическими процессами // Специальная техника и технологии транспорта. 2020. № 7 (45). С. 29-32. EDN ZQLTZI.
- Кумэ Х. Статистические методы повышения качества. М.: Финансы и статистика, 1990. 304 с.
- Надгериева Д. А. Обработка информации о количестве неисправностей электроэнергетических систем с применением регрессионного анализа / Д. А. Надгериева, А. Э. Дзгоев // Инженерные решения. 2009. № 6 (16). С. 18-23.
- Терентьев В. П. Метод корреляционных плеяд // Вестник ЛГУ. 1959. № 9. С. 137-141.
- Терентьев П. В. Дальнейшее развитие метода корреляционных плеяд // Применение математических методов в биологии. Л.: Изд-во ЛГУ, 1960. С. 27-36.

Тишанинов Н. П. Метод определяющих критериев качества технологического процесса / Н. П. Тишанинов, А. В. Анашкин // Наука в центральной России. 2019. № 3 (39). С. 48-55. EDN LEECHB.

Херсонский Н. С. Статистические методы в задачах менеджмента разработки, проектирования, производства и обслуживания изделий различного назначения. М.: «Эко-Пресс». 2011. 336 с.

Чистяков С. П. Моделирование процесса принятия решений при выборе методов статистического анализа. Отчет о НИР № 97–01–00554 / С. П. Чистяков, Ю. Л. Павлов, С. В. Стафеев, В. Н. Харин // Российский фонд фундаментальных исследований. 1998. EDN AQMEWF

Яшин С. Н. Оценка зависимости устойчивости инновационного развития и экономического состояния промышленных предприятий на основе корреляционного анализа / С. Н. Яшин, Ю. С. Солдатова // Финансовая аналитика: проблемы и решения. Риски, анализ и оценка. 2015. С. 10-18. EDN TSNJUZ.

References

Beshelev F. G., Gurvich F. G. (1980). Mathematical and statistical methods of expert assessments. 2nd ed., pererab. and add. Moscow: *Statistics*, 1980. 263 p. (In Russian)

Chistyakov S. P., Pavlov Yu. L., Stafeev S. V., Kharin V. N. (1998). Modeling the decision-making process when choosing statistical analysis methods. Research Report No. 97-01-00554. *Russian Foundation for Basic Research*. 1998.

Efremov N. Yu. (2020). Application of regression analysis to study the effect of technological parameters on polymer filling quality indicators. *Izvestia Samara Scientific Center of the Russian Academy of Sciences*. 22-4(96): 81-85. (In Russian)

Girilovich N. V., Dovgopolaya G. V. (2021). The use of statistical methods in the analysis of inconsistencies of inappropriate products in the production process. *Casting and metallurgy*. 3: 40-45.

Grigorieva T. A., Tolubaev V. N. (2018). Correlation and regression analysis of technological parameters. *Systems. Methods. Technology*. 3(39): 57-61. (In Russian)

Kalugin Yu. B. (2020). The use of correlation analysis to determine the size of the control step in process control. *Special equipment and technologies of transport*. 7(45): 29-32. (In Russian)

Kherson N. S. (2011). Statistical methods in the tasks of management of the development, design, production and maintenance of products for various purposes. Moscow: "Eco-Press", 2011. 336 p. (In Russian)

Kume H. (1990). Statistical methods for quality improvement. Moscow: *Finance and Statistics*. 1990. 304 p. (In Russian)

Nadgerieva D. A., Dzgoev A. E. (2009). Processing information on the number of failures of electric power systems using regression analysis. *Engineering solutions*. 6(16): 18-23. (In Russian)

Sachs L. Statistical evaluation. Moscow: *Statistics*. 1976. 598 p. (In Russian)

Terentyev P. V. (1960). Further development of the correlation galaxy method. *Application of mathematical methods in biology*. 27-36. (In Russian)

Terentyev V. P. (1959). Method of correlation pleiads. *Bulletin of Leningrad State University*. 9: 137-141.

Tishaninov N. P., Anashkin A. V. (2019). Method of determining the quality criteria of the technological process. *Science in Central Russia*. 3(39): 48-55.

Venetsky, I. G., Venetskaya V. I. (1974). Basic mathematical and statistical concepts and formulas in economic analysis. Moscow: *Statistics*. 1974. 61 p. (In Russian)

Vinichenko, A. V. (2023). Correlation Matrix for Complexly Structured Behavioral Factors and Process Parameters. *Innovative Instrumentation*. 2(4): 88-92. (In Russian)

Yashin S. N., Soldatova Yu. S. (2015). Assessment of the dependence of the sustainability of innovative development and the economic state of industrial enterprises based on correlation analysis. *Financial analytics: problems and solutions. Risks, analysis and assessment*. 10-18. (In Russian)