
	Стандарт ОАО «РЖД»	СТО 1.05.515.4
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

Утвержден
 распоряжением ОАО «РЖД»
 от «02» июня 2009 г.
 № 1150р


Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния

Перед использованием нормативного документа убедитесь в его актуальности	Дата ввода в действие:	Редакция:	Лист:
	01.07.2009 г.		1 из 20

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

Содержание

1 Область применения	3
2 Цель	3
3 Определение понятий	3
4 Схема осуществления деятельности	3
5 Описание применяемого метода.....	5
6 Принятые сокращения	9
7 Нормативные ссылки	9
Приложение А (обязательное) Контрольный листок	10
Приложение Б (рекомендуемое) Стратификация данных	14
Приложение В (рекомендуемое) Примеры диаграмм рассеяния	16
Приложение Г (рекомендуемое) Библиография	19

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

1 Область применения

Настоящий стандарт устанавливает особые требования по применению инструмента «Корреляционный анализ. Диаграмма рассеяния» в железнодорожной отрасли.

Настоящий стандарт должен применяться в процессах управления качеством и безопасностью в ОАО «РЖД» как один из методов при:

- реализации процессов предупреждающего и диагностического технического обслуживания и ремонта;
- разработке планов решения проблем;
- реализации процессов закупки, производства и обслуживания;
- проведении внутренних аудитов; мониторинге и измерении процессов;
- управлении несоответствующей продукцией;
- анализе данных;
- реализации процессов постоянного улучшения.

Требования данного документа распространяются на все департаменты, дирекции и филиалы ОАО «РЖД», а также на их структурные подразделения для анализа широкого круга проблем относящихся к любой сфере деятельности.

2 Цель

2.1 Метод «Корреляционный анализ. Диаграмма рассеяния» позволяет определить вид и степень связи между парами соответствующих переменных. Эти две переменные (x и y) могут относиться:

- к характеристике качества y и к влияющему на нее фактору x ;
- к двум различным характеристикам качества x и y ;
- к двум факторам x и y , влияющим на одну характеристику качества z .

2.2 Действия по применению инструмента «Корреляционный анализ. Диаграмма рассеяния» идентичны, независимо от места его применения.

3 Определение понятий

3.1 В настоящем стандарте применяются термины с соответствующими определениям по [ГОСТ Р 50779.10](#) и [ГОСТ Р 50779.11](#), а также приведенные ниже:

Вариабельность – изменчивость, неизбежные различия среди индивидуальных результатов процесса.


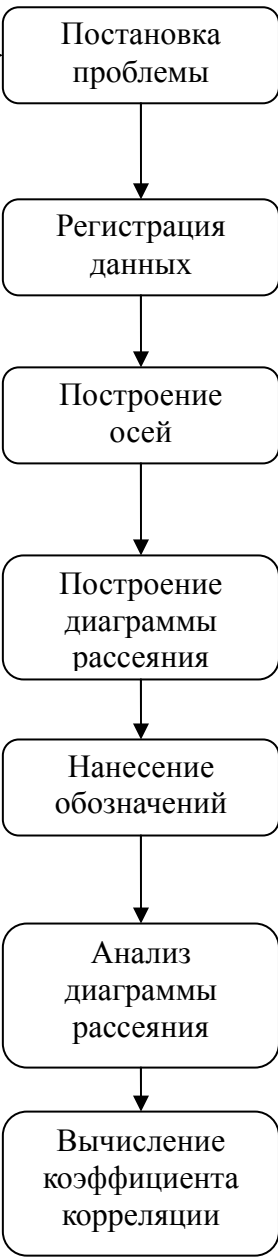
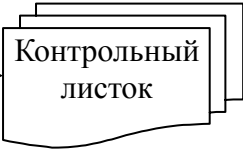
Диаграмма рассеяния (разброса) – графическое представление множества точек, координаты которых x и y в обычной прямоугольной системе координат – это значения признаков X и Y .


4 Схема осуществления деятельности

4.1 Схема осуществления деятельности по применению метода «Корреляционный анализ. Диаграмма рассеяния» приведена в [таблице 1](#).

Перед использованием нормативного документа убедитесь в его актуальности	Лист:
	3 из 20

Таблица 1 – Схема осуществления деятельности по применению метода «Корреляционный анализ. Диаграмма рассеяния»

Входные данные	Порядок выполнения	Выходные данные	Ответственность
			Специалисты соответствующих подразделений

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

5 Описание применяемого метода

5.1 Поэтапная процедура использования инструмента «Корреляционный анализ. Диаграмма рассеяния»

Примечание – Данная процедура позволяет построить диаграмму рассеяния без использования какого-либо программного обеспечения. Построение диаграммы рассеяния с помощью программных продуктов выходит за рамки данного стандарта.

5.1.1 Постановка проблемы. Определите, какие проблемы подлежат исследованию, определите две характеристики x и y , между которыми необходимо определить вид и степень зависимости

5.1.2 Регистрация данных. Разработайте контрольный листок ([приложение А](#)) для регистрации данных. Соберите парные данные (x , y) и занесите их в таблицу.

Примечание – Желательно собрать не менее 30 пар значений.

5.1.3 Построение осей

5.1.3.1 Найдите максимальное и минимальное значения x и для y .

5.1.3.2 Выберите шкалы на горизонтальной и вертикальной осях так, чтобы обе длины рабочих частей получились приблизительно одинаковыми.

5.1.3.3 Возьмите на каждой оси от 3 до 10 градаций и используйте для облегчения чтения круглые числа. Размах разметки осей не должен существенно превышать размах значений соответствующих переменных.

5.1.3.4 Если одна переменная – фактор, а вторая – характеристика качества, то выберите для фактора горизонтальную ось x , а для характеристики качества – вертикальную ось y .

Примечание – Нарушение правил построения осей, описанных в 5.1.3.1-5.1.3.4 может привести к ошибочной интерпретации информации.

5.1.4 Построение диаграммы рассеяния


5.1.4.1 Нанесите пары данных на размеченные оси.

Примечание – Если в разных наблюдениях получаются одинаковые значения, покажите эти точки, либо рисуя концентрические окружности (⊙), либо нанося вторую точку рядом с первой.

5.1.5 Нанесение необходимых обозначений. Нанесите на диаграмму необходимые обозначения. Убедитесь, что нижеперечисленные данные, отраженные на диаграмме, понятны:

- название диаграммы;
- интервал времени;
- число пар данных;
- названия и единицы измерения для каждой оси;
- имя (и прочее) человека, который делал эту диаграмму.

Перед использованием нормативного документа убедитесь в его актуальности	Лист:
	5 из 20

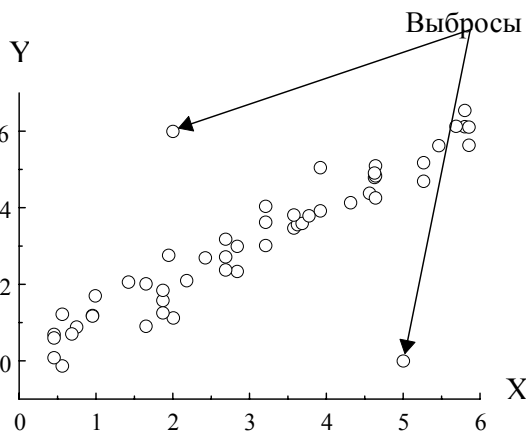
	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

5.1.6 Анализ диаграммы рассеяния

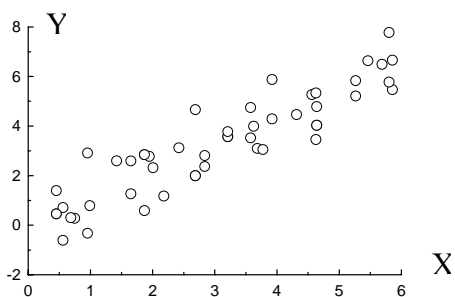
5.1.6.1 Выясните, есть ли на диаграмме выбросы ([рисунок 1а](#)). Выясните причины появления выбросов (ошибка измерения или записи данных, изменения в условиях работы, другие особые причины) и исключите их из дальнейшего корреляционного анализа.

5.1.6.2 Определите вид зависимости диаграммы рассеяния. Возможны следующие типичные варианты скопления точек:

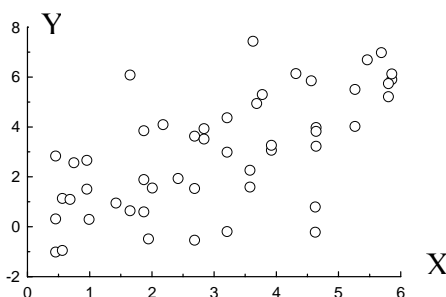
- сильная положительная корреляция ([рисунок 1б](#)). Ярко выраженная тенденция роста значения y с ростом значения x ;
- слабая положительная корреляция ([рисунок 1в](#)). Присутствует тенденция роста значения y с ростом значения x ;
- отсутствие корреляции ([рисунок 1г](#)). Между x и y не наблюдается никакой определенной зависимости;
- сильная отрицательная корреляция ([рисунок 1д](#)). Ярко выраженная тенденция уменьшения значения y с ростом значения x ;
- слабая отрицательная корреляция ([рисунок 1е](#)). Присутствует тенденция уменьшения значения y с ростом значения x ;
- криволинейная зависимость ([рисунок 1ж](#)). С ростом значения x величина y меняется криволинейно.



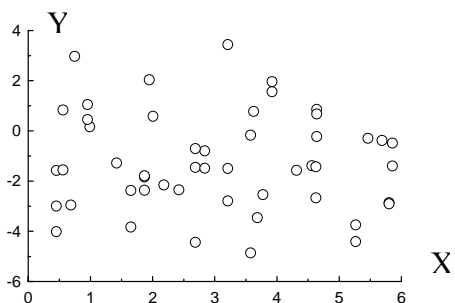
а) Наличие выбросов



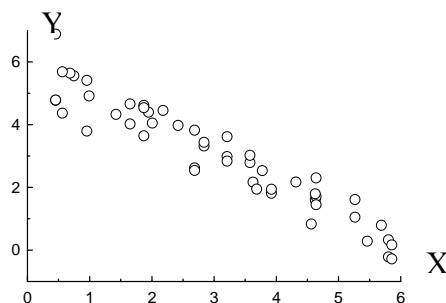
б) Сильная положительная корреляция



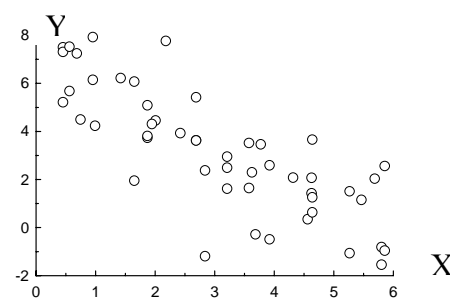
в) Слабая положительная корреляция



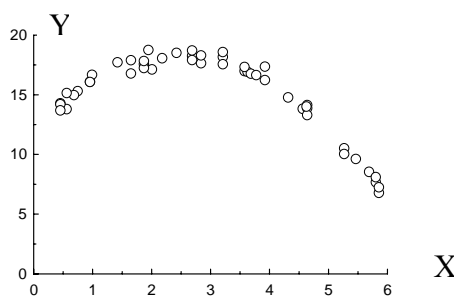
г) Отсутствие корреляции



д) Сильная отрицательная корреляция



е) Слабая отрицательная корреляция



ж) Криволинейная зависимость

Рисунок 1 – Типичные варианты диаграммы рассеяния

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

5.1.7 Вычисление коэффициента корреляции (r)

5.1.7.1 Коэффициент корреляции позволяет количественно определить силу линейной связи между x и y . Коэффициент корреляции вычисляется в соответствии со следующим определением (1):

$$r = \frac{S(xy)}{\sqrt{S(xx)S(yy)}}, \quad (1)$$

где

$$S(xx) = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{(\sum_{j=1}^n x_j)^2}{n};$$

$$S(yy) = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{(\sum_{j=1}^n y_j)^2}{n}$$

$$S(xy) = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \frac{(\sum_{j=1}^n x_j)(\sum_{j=1}^n y_j)}{n}.$$

5.1.7.2 Коэффициент корреляции (r) принимает значения из диапазона $-1 \leq r \leq 1$.

Примечания


- 1 Если абсолютное значение r окажется больше 1, то произошла ошибка, и нужно пересчитать результат. В случае сильной положительной корреляции достигается значение, близкое к плюс 1, а при сильной отрицательной корреляции достигается значение, близкое к минус 1. Значение r близкое к 0 указывает на слабую корреляцию или ее отсутствие. Когда коэффициент корреляции равен одному из указанных пределов, это означает, что существует точная линейная зависимость в серии парных наблюдений.
- 2 Когда между двумя переменными вычисляется коэффициент корреляции, иногда случайно проявляется сильная корреляция, которая или не подкреплена вовсе, или подкреплена слишком слабой причинно-следственной зависимостью между ними. Корреляция такого рода называется ложной корреляцией.

5.2 Степень корреляции между двумя переменными очень часто зависит от размахов этих переменных. Поэтому, не рекомендуется экстраполировать выводы анализа за пределами полученных данных без дополнительного исследования.

5.3 Следует сравнивать диаграммы рассеяния и коэффициенты корреляции, полученные в разные моменты времени. Изменения вида и степени корреляции с течением времени могут быть полезны в определении направления улучшения процесса.

5.4 Следует расслаивать (стратифицировать) диаграммы рассеяния для различных источников данных или условий при которых данные были получены (материалы, машины, операторы, условия производства, контроллеры, поставщики и т.д.) ([приложение Б](#)).

<p>Перед использованием нормативного документа убедитесь в его актуальности</p>	Лист:
	8 из 20

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

6 Принятые сокращения

6.1 В настоящем документе применяются следующие сокращения:

ДЗО – дочерние и зависимые общества;

ОАО «РЖД» – открытое акционерное общество «Российские железные дороги»;

СМК – система менеджмента качества;

СТК – стандарт по качеству;

СТО – стандарт организации.

6.2 В настоящем стандарте применяются следующие обозначения:

r – выборочный коэффициент корреляции ([ГОСТ Р 50779.10](#)).

\bar{x}, \bar{y} – среднее арифметическое ([ГОСТ Р 50779.10](#)).


7 Нормативные ссылки

В настоящем стандарте использованы ссылки на следующие нормативные документы:

ГОСТ Р ИСО 9000-2001 Системы менеджмента качества. Основные положения и словарь

ГОСТ Р 50779.10-2000 Статистические методы. Вероятность и основы статистики. Термины и определения

ГОСТ Р 50779.11-2000 Статистические методы. Статистическое управление качеством. Термины и определения

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

**Приложение А
(обязательное)
Контрольный листок**

А.1 Для принятия правильного и оптимального решения необходимо собрать наиболее полную информацию о проблеме. Кроме того, данные должны быть структурированы и удобны для дальнейшей обработки. Основным инструментом сбора информации для решения проблем являются контрольные листки.

Контрольный листок – это инструмент для сбора данных и автоматического их упорядочения для облегчения дальнейшего использования собранной информации.

А.2 Контрольные листки используются для:

- контроля контактов с посетителями;
- учета брака какой-либо детали или устройства на производстве.

Они сильно отличаются друг от друга и разрабатываются специально для записи информации определенного типа.

Когда цель сбора данных установлена, она становится основой для определения характера сравнения, которое надлежит произвести, и типа данных, которые нужно собрать. При этом следует учесть следующие моменты:

- необходимо определиться с временным периодом сбора информации, то есть выбрать, какой срок будет показательным для данных, которые мы собираем;
- разделить группу данных на подгруппы по определенному признаку;
- необходимо зафиксировать источник данных (день недели, когда собирались данные; оборудование, на котором производилась работа; рабочий, производивший операцию; партия использованных материалов и т.д.);
- данные необходимо регистрировать таким образом, чтобы их было легко фиксировать и использовать в дальнейшей обработке.

Перед использованием нормативного документа убедитесь в его актуальности	Лист:
	10 из 20


	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

Таблица А.1 – Контрольный листок для построения инструмента «Диаграмма рассеяния»

Контрольный листок для сбора данных																															
Форма №:																															
Дата начала:																															
Ответственный за сбор данных:																															
Дата завершения:																															
Переменная X:													Переменная Y:																		
Единицы измерения:													Единицы измерения:																		
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	
X																															
Y																															
	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	6	
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	
X																															
Y																															
Данные для построения осей										Вычисление коэффициента корреляции										Комментарии											
X_{\min} :										$S(xx) =$																					
X_{\max} :										$S(yy) =$																					
Y_{\min} :										$S(xy) =$																					
Y_{\max} :										$r =$																					

Примеры контрольных листков

СЕРТИФИКАТ № 518

ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО
 «РОССИЙСКИЕ ЖЕЛЕЗНЫЕ ДОРОГИ»
 (ОАО «РЖД»)
 ФИЛИАЛ
 «ЯРОСЛАВСКИЙ
 ЭЛЕКТРОВОЗРЕМОНТНЫЙ ЗАВОД
 ИМЕНИ Б.П. БЕЩЕВА»

выданного отделом технического контроля литейного цеха в том, что продукция в настоящей выписке

соответствует действующей НТД и отгружена в адрес

вагон № _____ Автомашина _____ " " _____ 2006 г.

"25" 12 2006

Наименование деталей (Гост, ТУ, чертёж)	Материал	Номер плавки	Количество деталей	Механические испытания.			Химический анализ.				
				Твёрдость ушка по Бринеллю	Твёрдость тела колодки по Бринеллю		C	Si	Mn	S	P
Колодка	СЧ	244	4 320		187		3,12	1,57	0,58	0,09	1,35
44-2890-00 СБ.					187		3,14	1,31	0,57	0,09	1,29
гост 28186-89					229		3,11	1,34	0,54	0,09	1,24
ССФЖТ RU ТМ01 Б.01425					197		2,95	1,35	0,60	0,09	1,38
с 29.12.03. по 29.12.06.					187		3,08	1,31	0,48	0,09	1,24
					217		3,04	1,40	0,55	0,09	1,17
					217		3,19	1,35	0,58	0,09	1,01
					229		3,12	1,36	0,60	0,09	1,24
					229		3,21	1,61	0,53	0,09	1,25

НАЧАЛЬНИК ОТК _____

НАЧАЛЬНИК ЦЕХА _____


КОНТРОЛЕР ОТК _____

Рисунок А.1 – Сертификат на тормозные колодки
(Данные для анализа твердости тела колодки)

Смазка пластичная Буксол					
КОНТРОЛЬНАЯ КАРТА ВХОДНОГО КОНТРОЛЯ					
№ /0254210300/					
Объем партии = _____ кг, кол-во бочек _____ шт. Отобрано проб _____					
<small>Примечание: отбирается 1 проба из каждой бочки, а при партии более 3 бочек отбирается 4 пробы</small>					
Изготовитель _____ <small>(наименование изготовителя)</small>					
Сопроводительная документация _____ <small>(наименование документа)</small>					
Паспорт № _____		Партия № _____			
Нормативный документ _____		ТУ 0254-107-01124328-01			
Инструментальный метод испытаний (ХТЛ)					
Наименование контролируемого параметра	Нормативное значение показателя	Фактическое значение			
		Проба № _____	Проба № _____	Проба № _____	Проба № _____
Массовая доля воды, %	Отсутствие				
Предел прочности при 50°С, Па, в пределах	300 – 700				
Массовая доля свободных органических кислот, выраженная в мг КОН на 1г смазки, не более	5,0				
Температура каплепадения, °С, не менее	180				
Массовая доля механических примесей, %	Отсутствие				
Пенетрация при 25°С с перегибом в 60 двойных тактов, мм ¹⁰ , в пределах	230-290				
Используемое испытательное оборудование					
№ п/п	Наименование	Тип	Заводской номер		
1	Прибор для определения массовой доли воды	АКОВ			
2	Аппарат комбинированный для определения предела прочности, термо- и влагопрочности пластичных смазок	АК-88			
3	Пенетрометр для нефтепродуктов (пластичных смазок)	ПН-1М			
4	Прибор для измерения температуры каплепадения	Каппа-1			
5	Спектроскоп	Спектроскан-V			
6	Колбонагреватель	ЛПВ-КН-500			
7	Весы	ВЛКТ-500Г-М			
По результатам испытаний продукцию признать _____ <small>(соответствующий, несоответствующий)</small>					
Исполнитель _____ <small>(Должность) (Фамилия, Инициалы) (Подпись) (дата)</small>					
Представитель отдела качества _____ <small>(Фамилия, Инициалы) (Подпись и инициалы) (дата)</small>					

Масло индустриальное И-40А					
КОНТРОЛЬНАЯ КАРТА ВХОДНОГО КОНТРОЛЯ					
№ /0253410028/					
Объем партии = _____ кг, кол-во бочек _____ шт. Отобрано проб _____					
<small>Примечание: отбирается 1 проба из каждой бочки, а при партии более 3 бочек отбирается 4 пробы</small>					
Изготовитель _____ <small>(наименование изготовителя)</small>					
Сопроводительная документация _____ <small>(наименование документа)</small>					
Паспорт № _____		Партия № _____			
Нормативный документ _____		ГОСТ 20799-88			
Инструментальный метод испытаний (ХТЛ)					
Наименование контролируемого параметра	Нормативное значение показателя	Фактическое значение			
		Проба № _____	Проба № _____	Проба № _____	Проба № _____
Кинематическая вязкость при 40°С, мм ² /с	61-75				
Кислотное число, мг КОН на 1г масла, не более	0,05				
Содержание механических примесей	Отсутствие				
Содержание воды	Следы				
Плотность, при 20°С, кг/м ³ , не более	900				
Температура вспышки определяемая в открытом тигле, °С, не ниже	220				
Используемое испытательное оборудование					
№ п/п	Наименование	Тип	Заводской номер		
1	Вязкозиметр	ВПКЖ-4			
2	Термостат	ЛТН-01М			
3	Прибор для определения массовой доли воды	АКОВ			
4	Весы	ВЛКТ-500Г-М			
5	Набор ареометров	АОН			
6	Прибор для определения температуры вспышки в открытом тигле	ТВО			
По результатам испытаний продукцию признать _____ <small>(соответствующий, несоответствующий)</small>					
Исполнитель _____ <small>(Должность) (Фамилия, Инициалы) (Подпись) (дата)</small>					
Представитель отдела качества _____ <small>(Фамилия, Инициалы) (Подпись и инициалы) (дата)</small>					

Рисунок А.2 – Контрольные листки для проведения контроля качества и сбора сведений для проведения статистического анализа

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

Приложение Б (рекомендуемое) Стратификация данных

Б.1 Стратификация – разделение полученных данных на отдельные группы (слои, страты) в зависимости от выбранного стратифицирующего фактора.

В качестве стратифицирующего фактора могут быть выбраны любые параметры, определяющие особенности условий возникновения и получения данных:

- различное оборудование;
- операторы, производственные бригады, участки, цехи, предприятия и т.п.;
- время сбора данных;
- разные виды сырья и материалов;
- различие используемого оборудования, оснастки, инструмента, измерительных систем и т.д.

Б.2 При отсутствии учета стратифицирующего фактора (расслоения данных) происходит их объединение и обезличивание, затрудняющее установление действительной взаимосвязи между полученными данными и особенностями их возникновения.

Например, при анализе источника несоответствующей продукции, поставляемой предприятию несколькими сторонними поставщиками, целесообразно в качестве стратифицирующего фактора выбрать поставщиков и произвести стратификацию несоответствующей продукции по поставщикам.

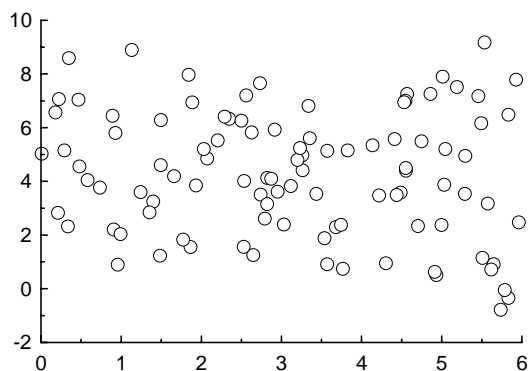
Применение стратификации данных

При практическом использовании метода стратификации рекомендуется действовать следующим образом:

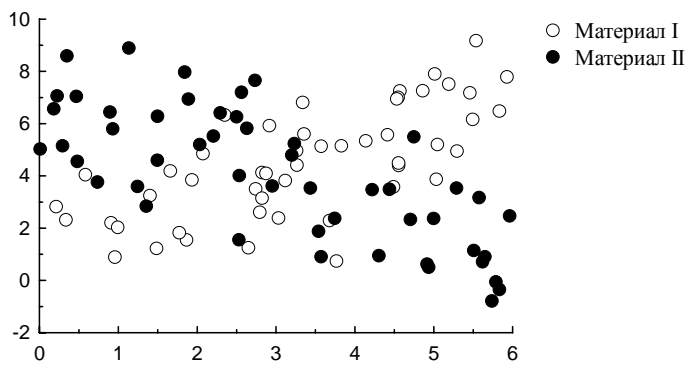
- 1) выберите данные, представляющие интерес для изучения;
- 2) выберите стратифицирующий фактор и категории (группы), на которые будут разделяться данные;
- 3) произведите группировку данных на основании выбранных категорий;
- 4) оцените результаты группировки по каждой из категорий;
- 5) соответствующим образом представьте полученные результаты;
- 6) проанализируйте необходимость дополнительного изучения данных;
- 7) спланируйте последующую работу для дополнительного подтверждения полученных результатов.

На [рисунке Б.1](#) приведены примеры стратификации данных на диаграмме рассеяния. Так, [на рисунке Б.1а](#) корреляция между двумя переменными отсутствует. Однако после расслоения данных для двух различных материалов, обнаруживается корреляция между параметрами для каждого материала в отдельности ([рисунк Б.1б](#)). Обратная ситуация наблюдается на [рисунках Б.1в-г](#): положительная корреляция между двумя параметрами, исчезает после стратификации данных по двум сменам.

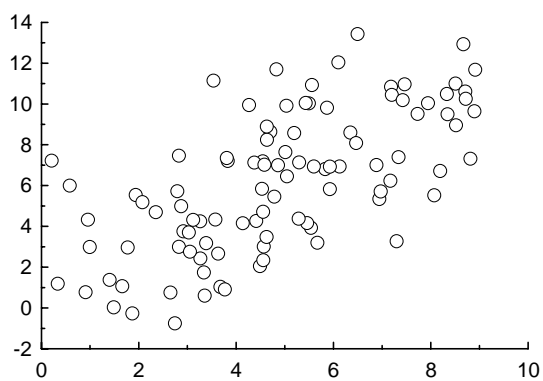
Перед использованием нормативного документа убедитесь в его актуальности	Лист:
	14 из 20



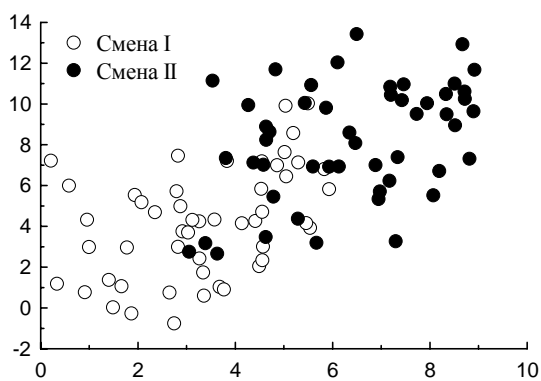
а)



б)



в)



г)

Рисунок Б.1 – Примеры стратификации диаграмм рассеяния

Приложение В
(рекомендуемое)
Примеры диаграмм рассеяния

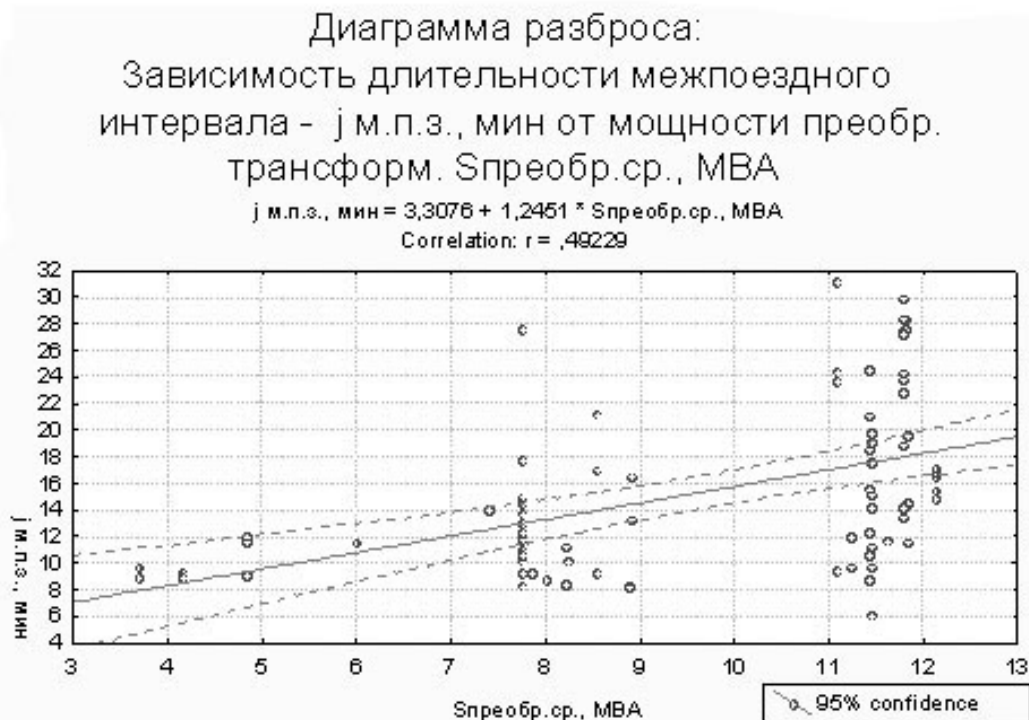


Рисунок В.1 – Диаграмма рассеяния значений межпоездного интервала от мощности преобразовательного трансформатора [\[1\]](#)

Диаграмма рассеивания значений
времени хода по перегону от длины перегона

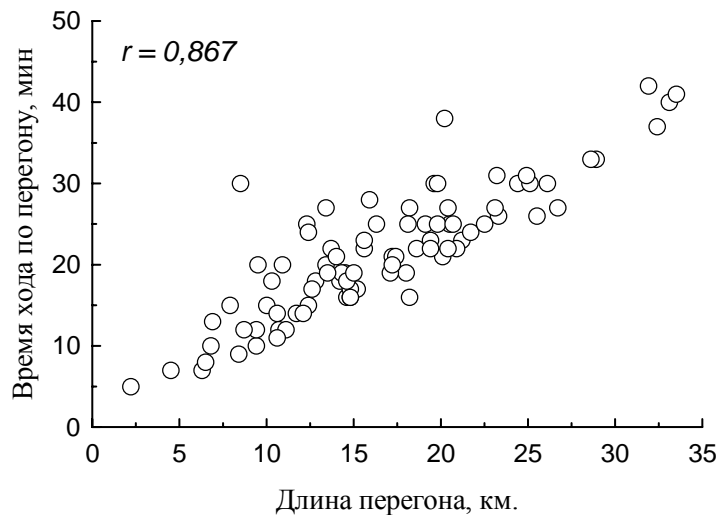


Диаграмма рассеивания значений
пропускной способности перегона
от длины перегона

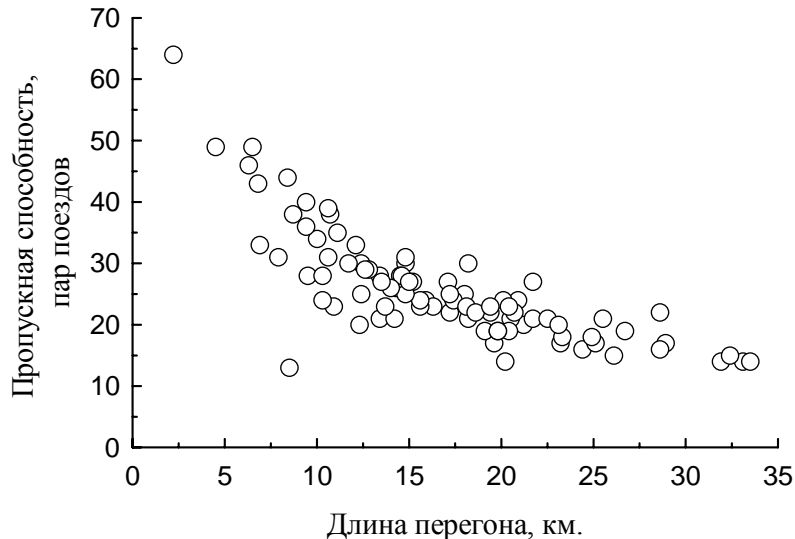


Рисунок В.2 – Диаграммы рассеяния времени хода по перегону и пропускной способности от длины перегона [\[2\]](#)

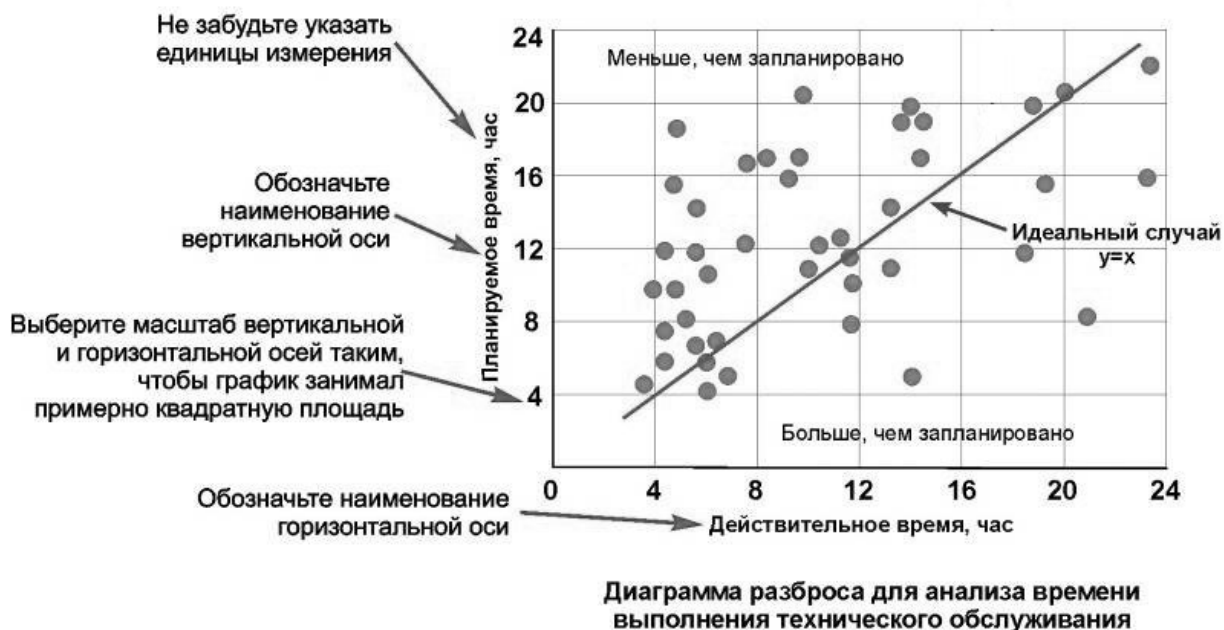



Рисунок В.3 – Процедура построения диаграммы рассеяния на примере исследования связи между планируемым и действительным временем выполнения технического обслуживания

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

**Приложение Г
(рекомендуемое)
Библиография**

- [1] Функциональный проект начальника службы электрификации и электроснабжения Куйбышевской железной дороги И.А.Крестовникова.
- [2] Функциональный проект начальника Комсомольского отделения Дальневосточной железной дороги Н.А.Дмитрука.

	Стандарт ОАО «РЖД»	СТО 1.05.515.4-2009
	Методы и инструменты улучшений. Корреляционный анализ. Диаграмма рассеяния	

Изменения

Версия	Дата изменения	Разработчик изменений	Краткое описание изменения