

Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования «Кузбасский государственный технический
университет имени Т. Ф. Горбачева»

Кафедра информационных и автоматизированных
производственных систем

Составители
В.В. Зиновьев, А.Н. Стародубов,
П.И. Николаев, И.С. Кузнецов

ДИСПЕРСИОННЫЙ АНАЛИЗ ПРИ МОДЕЛИРОВАНИИ СИСТЕМ

Методические указания к лабораторной работе

Рекомендовано учебно-методической комиссией направления подготовки
15.04.04 «Автоматизация технологических
процессов и производств» в качестве электронного издания для
использования в образовательном процессе

Кемерово 2022

Рецензенты:

Чичерин И. В. – к.т.н., доцент кафедры информационных и автоматизированных производственных систем

Зиновьев Василий Валентинович

Стародубов Алексей Николаевич

Николаев Петр Игоревич

Кузнецов Игорь Сергеевич

Дисперсионный анализ при моделировании систем: методические указания к лабораторной работе для обучающихся направления подготовки 15.04.04 Автоматизация технологических процессов и производств. / сост. В. В. Зиновьев, А. Н. Стародубов, П. И. Николаев, И. С. Кузнецов; Кузбасский государственный технический университет имени Т. Ф. Горбачева. – Кемерово, 2022. – Текст : электронный.

В данных методических указаниях разработанных по дисциплине «Математическое моделирование» приведены : цель и задача, общие сведения об изучаемом материале, задания для выполнения, вопросы для самопроверки. Рекомендуемая литература для самостоятельной подготовки приведена в рабочей программе дисциплины.

© Кузбасский государственный
технический университет имени
Т. Ф. Горбачева, 2022

© В.В. Зиновьев, А.Н. Стародубов,
П.И. Николаев, И.С. Кузнецов,
составление, 2022

Оглавление

1	Основные теоретические положения.....	4
1.1	Задачи дисперсионного анализа при имитационном моделировании систем	4
1.2	Технология дисперсионного анализа в MS Excel	6
2	Пример выполнения дисперсионного анализа	7
3	Задание на лабораторную работу	10
4	Порядок выполнения работы.....	11
5	Требования к отчету	11
6	Контрольные вопросы	12
7	Список рекомендуемой литературы	12

Цель работы – изучить технологию дисперсионного анализа для оценки чувствительности результатов моделирования к изменению входных параметров и изменению вида вероятностных распределений, используемых в модели.

В ходе выполнения лабораторной работы магистрант должен научиться формировать таблицу дисперсионного анализа при помощи программы табличных вычислений Microsoft Excel и выявлять зависимость результатов имитационных экспериментов от изменения входных параметров модели.

Примечание

Предполагается, что магистрант, знаком с основами работы с Microsoft Excel`2007 или более поздних версий, а также с основами имитационного моделирования систем в среде компьютерного моделирования GPSS World.

1 Основные теоретические положения

1.1 Задачи дисперсионного анализа при имитационном моделировании систем

Дисперсионный анализ (ANOVA–ANalysisOfVAriance) позволяет установить влияние отдельных факторов на изменчивость какого-либо отклика модели, значения которого могут быть получены путем проведения экспериментов с моделью.

Задачи дисперсионного анализа хотя и являются относительно простыми, но, тем не менее, часто используются на практике.

Например, стохастическая имитационная модель не может быть абсолютно точной, и необходимо как-то измерить степень этой неточности. Это измерение можно проделать (частично) путем анализа чувствительности отклика модели к изменению уровней факторов. Также важно определение чувствительности отклика к изменению вида вероятностных распределений, используемых в модели. Это делается для исключения эффектов, которые являются не более чем случайной флуктуацией. Более того, если при несущественных изменениях уровней факторов значение отклика модели меняется очень сильно, то необходимо проводить дополнительные исследования для получения более точных оценок. Если значение отклика при значительных изменениях уровней факторов изменяется не существенно, то даль-

нейшее экспериментирование нецелесообразно. Также анализ чувствительности посредством дисперсионного анализа может показать, как изменить (заменить) эти факторы, чтобы оптимизировать реальную систему.

Дисперсионный анализ широко используется при валидации модели путем выявления значимых факторов в системе и сравнении их с данными реальной системы или с экспертными оценками специалистов.

Центральная проблема планирования эксперимента – выбор ограниченного набора комбинаций факторных уровней, которые будут фактически моделироваться при экспериментировании с моделью. Здесь также применяется дисперсионный анализ [1].

В отличие от экспериментов на реальных системах, эксперимент на имитационной модели является контролируемым. Можно варьировать любой фактор и судить о поведении модели по наблюдаемым откликам.

Дисперсионный анализ модели требует выполнения набора прогонов, при этом факторы в течение прогона моделирования изменяются от прогона к прогону. Каждый фактор имеет, по крайней мере, два уровня в эксперименте. Фактор может быть качественным (например, приоритетные правила или закон распределения). Дисперсионный анализ, основанный на статистической модели, заканчивается построением таблицы ANOVA, в которой анализируется влияние факторов А, В, взаимодействие между факторами АВ и случайные помехи наблюдения [2]. С помощью этой таблицы проверяется гипотеза об отсутствии влияния фактора. Если справедлива гипотеза об отсутствии влияния фактора, то считается, что все наблюдения получены из одной генеральной совокупности. Для проверки гипотезы используется F-распределение Фишера. Критерий Фишера определяет отношение двух выборочных дисперсий. Если фактор существенно влияет на отклик, то значения F-распределения принимает большие значения и F-статистика становится значимой. Таким образом, большие значения F приводят к отбрасыванию гипотезы об отсутствии влияния фактора, т. е. фактор является значимым [3].

Многие системы имитационного моделирования содержат встроенные библиотечные процедуры ANOVA. Также стандартные статистические методы обработки данных включены в со-

став программ табличных вычислений Microsoft Excel, Quatro Pro; в математические пакеты Mathcad, Mathlab, Mapl. Еще более мощными возможностями статистической обработки обладают специализированные пакеты STATISTICA, STATGRAPHICS, SPSS.

Для изучения и понимания технологии дисперсионного анализа, учитывая универсальность и распространенность табличного процессора MicrosoftExcel (с встроенным пакетом анализа), используем именно этот программный продукт.

Опустим из рассмотрения теорию дисперсионного анализа, т. к. ей посвящено множество литературы по математической статистике. Рассмотрим практическое приложение этого метода к моделированию систем и технологию дисперсионного анализа в MicrosoftExcel.

1.2 Технология дисперсионного анализа в MS Excel

В зависимости от числа оказывающих влияние факторов различают однофакторный и многофакторный (двухфакторный и т. д.) дисперсионный анализ.

Режим работы «Однофакторный дисперсионный анализ» служит для выяснения факта влияния контролируемого входного параметра модели (фактора) A на результат моделирования (отклик) Y на основе выборочных данных результатов имитационных экспериментов. Режимы работы «Двухфакторный дисперсионный анализ без повторений» и «Двухфакторный дисперсионный анализ с повторениями» служат для выяснения на основе выборочных данных результатов экспериментов факта влияния контролируемых факторов A и B на отклик модели Y . При этом в режиме «Двухфакторный дисперсионный анализ без повторений» каждому уровню факторов A и B соответствует только одна выборка результатов экспериментов, а в режиме «Двухфакторный дисперсионный анализ с повторениями» каждому уровню одного из факторов A (или B) соответствует более одной выборки данных. В последнем случае число выборок для каждого уровня должно быть одинаковым.

Так как задачи однофакторного дисперсионного анализа являются самыми простыми в своем классе, а проведение имитационных экспериментов – управляемый процесс, при котором не

составляет труда проводить множество прогонов модели (увеличивая тем самым статистическую значимость результатов), рассмотрим двухфакторный дисперсионный анализ с повторениями.

Для активизации дисперсионного анализа в MS Excel, необходимо зайти во вкладку «Данные», выбрать раздел «Анализ данных» рис.1.1, затем в зависимости от задачи из предложенного списка выбрать пункт «Однофакторный дисперсионный анализ» либо «Двухфакторный дисперсионный анализ» (рис.1.2).

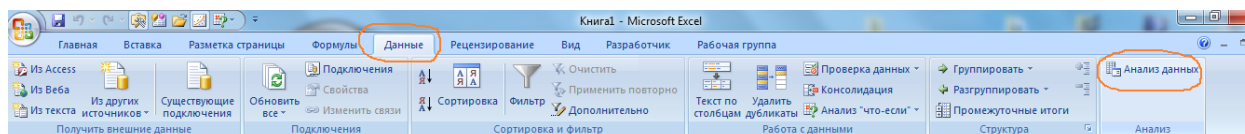


Рис. 1.1 Запуск пакета «Анализ данных»

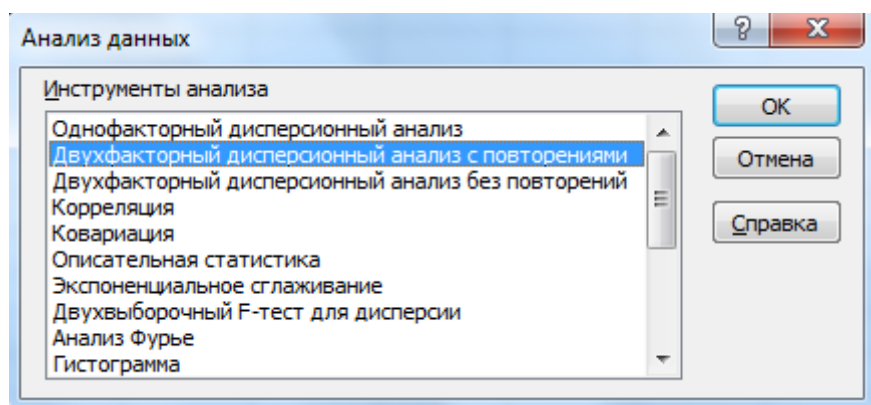


Рис. 1.2 Выбор «Двухфакторный дисперсионный анализ»

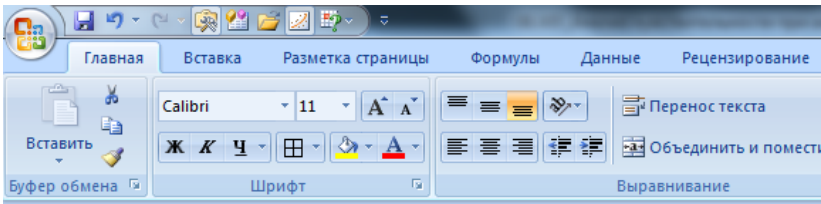
2 Пример выполнения дисперсионного анализа

Предположим, что на имитационной модели некоторой системы проведен ряд экспериментов. Для анализа чувствительности отклика модели Y к факторам A и B используем встроенный в Microsoft Excel режим «Двухфакторный дисперсионный анализ с повторениями». Для этого необходимо:

1. По результатам имитационных экспериментов сформировать на рабочем листе Microsoft Excel выборочные данные об отклике модели, при изменении факторов A и B на 4-х уровнях, используя результаты трех прогонов для каждого уровня.

2. Сформировать в Microsoft Excel соответствующую таблицу и при уровне значимости $\alpha=0,05$ выяснить, влияют ли на отклик Y факторы A и B .

Выборочные данные для отклика Y при различных уровнях факторов A и B , приведены на рис.2.1.



	A	B	C	D	E	F	G
1	Прогнозы	Фактор A	Фактор B				
2	модели		Уровень 1	Уровень 2	Уровень 3	Уровень 4	
3	Прогон 1	Уровень 1	21,4	20,9	19,6	17,6	
4	Прогон 2		21,2	20,3	18,8	16,6	
5	Прогон 3		20,2	19,8	16,4	17,5	
6	Прогон 1	Уровень 2	12	13,6	13	13,3	
7	Прогон 2		14,2	13,3	13,7	14	
8	Прогон 3		12,1	11,6	12	13,9	
9	Прогон 1	Уровень 3	13,5	14	12,9	12,4	
10	Прогон 2		11,9	15,6	12,9	13,7	
11	Прогон 3		13,4	13,8	12,1	13	
12	Прогон 1	Уровень 4	12,8	14,1	14,2	12	
13	Прогон 2		13,8	13,2	13,6	14,6	
14	Прогон 3		13,7	15,3	13,3	14	

Рис.2.1 Результаты моделирования при изменении факторов

Для выполнения задания необходимо запустить пакет анализа, встроенный в MSeXcel, и выбирать раздел «Двухфакторный дисперсионный анализ с повторениями» в соответствующем диалоговом окне (рис.2.2).

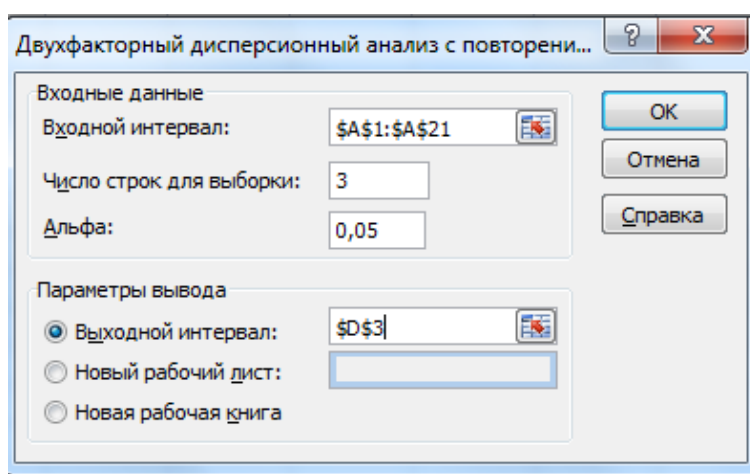


Рис. 2.2 Диалоговое окно дисперсионного анализа

Затем надо заполнить поля задав следующие параметры:

- «Входной интервал» – выделяются результаты моделирования, захватывая значения уровней факторов (рис. 2.1).
- «Число строк для выборки» – вводится число выборок, приходящихся на каждый уровень одного из факторов.
- «Альфа» – вводится уровень значимости α , равный вероятности возникновения ошибки первого рода (отвержение нулевой гипотезы).
- «Выходной интервал/Новый рабочий лист/Новая рабочая книга» – выбирается область вывода результатов дисперсионного анализа.

После заполнения указанных полей и запуска дисперсионного анализа путем нажатия кнопки «ОК», будет выведена таблица двухфакторного дисперсионного анализа с повторениями (рис. 2.3).

51	Дисперсионный анализ						
52	Источник вариации	SS	df	MS	F	P-Значение	F критическое
53	Выборка	310,135	3	103,3783333	124,7087208	9,80794E-18	2,901119588
54	Столбцы	10,01166667	3	3,337222222	4,025802128	0,015450891	2,901119588
55	Взаимодействие	25,98333333	9	2,887037037	3,48272877	0,004232014	2,188765768
56	Внутри	26,52666667	32	0,828958333			
57							
58	Итого	372,6566667	47				

Рис. 2.3 Результаты двухфакторного дисперсионного анализа

Как видим, расчетное значение F -критерия фактора $A - F_p^A = 124,7$, а критическая область образуется правосторонним интервалом $(2,90; +\infty)$. Так как F_p^A попадает в критическую область, то нулевую гипотезу H_A о равенстве выборочных средних отвергаем, т.е. считаем, что фактор A влияет на отклик Y .

Выборочный коэффициент детерминации для фактора A

$$\tilde{\rho}_A^2 = \frac{\tilde{\sigma}_A^2}{\tilde{\sigma}_Y^2} = \frac{310,135}{372,656} \approx 0,83$$

показывает, что 83 % общей выборочной вариации отклика Y связано с влиянием фактора A .

Расчетные значения F -критерия фактора B – $F_p^B=4,02$, а критическая область образуется правосторонним интервалом $(2,90; +\infty)$. Так как F_p^B попадает в критическую область, то нулевую гипотезу относительно второго фактора H_B отвергаем, т. е. считаем, что фактор B влияет на отклик Y .

Выборочный коэффициент детерминации для фактора B

$$\tilde{\rho}_B^2 = \frac{\tilde{\sigma}_B^2}{\tilde{\sigma}_Y^2} = \frac{10}{372,656} \approx 0,026$$

показывает, что только около 3 % общей выборочной вариации отклика Y связано с влиянием фактора B .

Значимость фактора взаимодействия F_p^{AB} ($F_p^{AB}=3,48$ и попадает в критический интервал $(2,18; +\infty)$) указывает на то, что фактор A варьируется при различных уровнях фактора B .

3 Задание на лабораторную работу

Описание объекта моделирования:

Сегмент транспортной системы предприятия состоит из n индуктивных тележек, которые перемещают отливочные блоки от одного пункта до другого и двух накопителей – основного и резервного. В основном накопителе одновременно могут находиться не более двух отливочных блоков. Блоки перемещаются каждой индуктивной тележкой за случайное время, распределенное по экспоненциальному закону в среднем за x секунд. Если в накопителе уже находятся 2 отливочных блока, то очередной блок попадает в резервный накопитель. Интервалы поступления отливочных блоков распределены экспоненциально с интенсивностью – λ блоков в минуту.

Значения n , λ и x по вариантам представлены в таблице 1.

Таблица 1

Значения количества индуктивных тележек (n), интенсивности поступления отливочных блоков (λ) и времени их транспортировок

В а р и а н т														
1			2			3			4			5		
n	λ	x	n	λ	x	n	λ	x	n	λ	x	n	λ	x
1	2	20	2	5	30	1	3	30	2	3	30	1	5	30

Задание:

1. Отобразите работу транспортной системы, используя среду компьютерного моделирования GPSSWorld.
2. Определите число отливочных блоков, попавших в резервный накопитель, при общей транспортировке 100 блоков.
3. Проведите имитационные эксперименты и сформируйте на рабочем листе Microsoft Excel выборочные данные о числе отливочных блоков, попавших в резервный накопитель, при различных значениях интенсивности поступления блоков (2 и 5 заготовок в минуту) и емкости основного накопителя (1 и 2 блоков).
4. Используя режим «Двухфакторный дисперсионный анализ с повторениями» (4 повторения для каждого эксперимента) сформируйте в Microsoft Excel соответствующую таблицу и при уровне значимости $\alpha=0,05$ выясните, влияют ли на число отливочных блоков, попавших в резервный накопитель, интенсивность поступления этих блоков, и емкость основного накопителя.

4 Порядок выполнения работы

1. Ознакомиться с методическими указаниями по выполнению данной лабораторной работы.
2. Получить у преподавателя вариант задания.
3. Отобразить работу транспортной системы в среде компьютерного моделирования GPSSWorld и определить число отливочных блоков, ушедших на резервный канал.
4. Провести необходимые имитационные эксперименты, по результатам которых сформировать в Microsoft Excel таблицу дисперсионного анализа и выяснить влияние на число отливочных блоков, ушедших на резервный канал, емкости основного накопителя, интенсивности поступления заготовок.
5. Сделать выводы.

5 Требования к отчету

Отчет о работе должен содержать:

1. Задание и исходные данные по заданному варианту лабораторной работы.

2. Модель транспортной системы в среде компьютерного моделирования GPSSWorld (файл .gps).

3. Таблицу дисперсионного анализа, сформированную по результатам имитационных экспериментов, в Microsoft Excel (файл .xls).

4. Анализ полученных результатов и выводы по работе.

6 Контрольные вопросы

1 Для чего предназначен дисперсионный анализ при имитационном моделировании систем?

2 Как дисперсионный анализ помогает в оценке чувствительности при имитационном моделировании систем?

3 Перечислите задачи, возникающие в процессе имитационного моделирования систем, в которых применяется ANOVA.

4 Для чего применяют инструмент «Дисперсионный анализ» встроенный в табличный процессор Microsoft Excel?

5 Какие виды дисперсионного анализа можно проводить в Microsoft Excel и в чем их различие?

7 Список рекомендуемой литературы

1. Шеннон Р. Имитационное моделирование систем – искусство и наука. – М.: Мир, 1978.

2. Томашевский В. Имитационное моделирование в среде GPSS. Серия «Факультет» / В. Томашевский, И. Жданова. – Москва : Бестселлер, 2003.

3. Макарова Н. В. Статистика в Excel: Учеб. пособие / Н. В. Макарова, В. Я. Трофимец. – Москва : Финансы и статистика, 2002.

4. Зиновьев В. В. Моделирование процессов и систем учебное пособие / В. В. Зиновьев, А. Н. Стародубов, П. И. Николаев // – КузГТУ : Кемерово, 2016 – 146с.

5. Руководство пользователя «Расширенный редактор GPSSWorld».