ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»

ЧЕЛЯБИНСКИЙ ФИЛИАЛ

Кафедра экономики, финансов и бухгалтерского учета

КОМПЬЮТЕРНЫЙ ПРАКТИКУМ ПО ЭКОНОМЕТРИКЕ

Методические указания по проведению лабораторной работы «Описательная статистика» УДК 372.851 ББК 22.172 Д32

> Дза Демьянов Д.Г. Компьютерный практикум по эконометрике: методические указания по проведению лабораторной работы «Описательная статистика». Челябинск: Челябинский филиал РАНХиГС, 2017. – 30 с.

Компьютерный практикум по эконометрике содержит расчетно-графические работы по отдельным вопросам эконометрического моделирования в электронных таблицах MS Excel.

Учебно-методическое издание направлено на овладение компетенции ОПК-1 «Способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационнокоммуникационных технологий и с учетом основных требований информационной безопасности».

Для студентов направления подготовки 38.03.01 Экономика, очной и заочной форм обучения.

Введение

Профессиональная деятельность экономиста С связана необходимостью не только качественного описания существующих и возникающих тенденций развития экономических процессов, но и их количественной оценки, анализа и прогнозирования последствий на основе статистических данных. Решение этой задачи требует владения определенных компетенций в сфере информационно-коммуникационных обработку технологий, направленных на И анализ социальноэкономической информации.

Эконометрика, как научная дисциплина формирует определенную информационную культуру, позволяющую будущему экономисту сформировать цифровые навыки проведения статистического анализа и построения эконометрических моделей с использованием прикладных компьютерных программ.

Выбор электронных таблиц MS Excel обусловлен его широким распространением, доступностью и наличием большого набора инструментов для проведения статистического анализа и выполнения математических операций при решении большинства задач эконометрики.

Практикум содержит необходимые краткие теоретические сведения, содержание и этапы выполнения расчетно-графической работы, образец решения типовой задачи в MS Excel с необходимыми пояснениями выполняемых действий и описания диалоговых окон, контрольные вопросы по теме работы.

Содержательная часть вариантов

1. Оценить закон распределения и числовые характеристики генеральной совокупности – затраты на технологические инновации (млн руб.). Параметры выборки представлены в таблице 1.

 Оценить закон распределения и числовые характеристики генеральной совокупности – расхода бензина (литров) на 100 км.
 в городском цикле для автомобилей одной модели. Параметры выборки представлены в таблице 1.

 Оценить закон распределения и числовые характеристики генеральной совокупности – времени выполнения одной операции (мин.) рабочими. Параметры выборки представлены в таблице 1.

 Оценить закон распределения и числовые характеристики генеральной совокупности – суточных надоев молока (литров) от одной коровы. Параметры выборки представлены в таблице 1.

5. Оценить закон распределения и числовые характеристики генеральной совокупности – времени (мин.) затрачиваемого на поездку от дома до работы. Параметры выборки представлены в таблице 1.

6. Оценить закон распределения и числовые характеристики генеральной совокупности – месячной заработной платы (тыс. руб.) рабочих одной профессии. Параметры выборки представлены в таблице 1.

7. Оценить закон распределения и числовые характеристики генеральной совокупности – расход топлива (литров) на 100 км автомобилем «Урал». Параметры выборки представлены в таблице 1.

8. Оценить закон распределения и числовые характеристики генеральной совокупности – стоимости (тыс. руб.) квадратного метра жилья на вторичном рынке. Параметры выборки представлены в таблице 1.

9. Оценить закон распределения и числовые характеристики генеральной совокупности – количества клиентов (чел.) обслуживаемых за смену одним мастером. Параметры выборки представлены в таблице 1.

10. Оценить закон распределения и числовые характеристики генеральной совокупности – объема недельных продаж бензина АИ-92 (тонн). Параметры выборки представлены в таблице 1.

11. Оценить закон распределения и числовые характеристики генеральной совокупности – производительности труда (тыс. руб.) рабочих одной профессии. Параметры выборки представлены в таблице 1.

12. Оценить закон распределения и числовые характеристики генеральной совокупности – суточного объема продаж (тыс. руб.) цветочных киосков. Параметры выборки представлены в таблице 1.

13. Оценить закон распределения и числовые характеристики генеральной совокупности – уровня использования среднегодовой производственной мощности организаций по выпуску текстильной продукции (в процентах). Параметры выборки представлены в таблице 1.

14. Оценить закон распределения и числовые характеристики генеральной совокупности – производство основных видов продукции в натуральном выражении (тыс. т.). Параметры выборки представлены в таблице 1.

15. Оценить закон распределения и числовые характеристики генеральной совокупности – потребление электроэнергии в Уральском федеральном округе (млн кВт. час). Параметры выборки представлены в таблице 1.

16. Оценить закон распределения и числовые характеристики генеральной совокупности – действующих строительных организаций (тыс. ед.). Параметры выборки представлены в таблице 1.

17. Оценить закон распределения и числовые характеристики генеральной совокупности – средние потребительские цены (руб.) на отдельные товары и услуги. Параметры выборки представлены в таблице 1.

18. Оценить закон распределения и числовые характеристики генеральной совокупности – уровень занятости в Уральском федеральном округе (в процентах). Параметры выборки представлены в таблице 1.

19. Оценить закон распределения и числовые характеристики генеральной совокупности – прирост высокопроизводительных рабочих мест (тыс. ед.). Параметры выборки представлены в таблице 1.

20. Оценить закон распределения и числовые характеристики генеральной совокупности – производство продуктов животноводства (тыс. штук). Параметры выборки представлены в таблице 1.

21. Оценить закон распределения и числовые характеристики генеральной совокупности – выбросы в атмосферу загрязняющих веществ (тыс. т.). Параметры выборки представлены в таблице 1.

22. Оценить закон распределения и числовые характеристики генеральной совокупности – степень износа основных фондов на конец года (в процентах). Параметры выборки представлены в таблице 1.

23. Оценить закон распределения и числовые характеристики генеральной совокупности – выплаты по договорам страхования, осуществленные страховщиками (млн руб.). Параметры выборки представлены в таблице 1.

24. Оценить закон распределения и числовые характеристики генеральной совокупности – убыток страховых организаций (млрд руб.). Параметры выборки представлены в таблице 1.

25. Оценить закон распределения и числовые характеристики генеральной совокупности – зарегистрированный уставный капитал действующих кредитных организаций (млрд руб.). Параметры выборки представлены в таблице 1.

Номер варианта	Параметр А размах выборочных данных	Параметр В	Параметр С	Объем выборки
1.	51 – 264	0,5	0,1	210
2.	7 – 12	0,5	0,1	190
3.	7 – 14	0,5	0,1	230
4.	5 – 21	0,5	0,1	250
5.	18 – 34	0,5	0,1	220
6.	15 – 29	0,5	0,1	210
7.	29 – 61	0,5	0,1	190
8.	19 – 69	0,5	0,1	230
9.	19 – 40	0,5	0,1	250
10.	8 – 19	0,5	0,1	220
11.	20 – 39	0,5	0,1	210
12.	9 – 17	0,5	0,1	190
13.	31 – 89	0,5	0,1	230
14.	75 - 153	0,5	0,1	250
15.	4206 – 98041	0,5	0,1	220
16.	115 – 268	0,5	0,1	210
17.	500 – 2100	0,5	0,1	190
18.	49 – 87	0,5	0,1	230
19.	15 – 89	0,5	0,1	250
20.	49 – 102	0,5	0,1	220
21.	134 – 512	0,5	0,1	210
22.	19 – 64	0,5	0,1	190
23.	961 – 57354	0,5	0,1	230
24.	1 – 37	0,5	0,1	250
25.	380 – 1463	0,5	0,1	220

Таблица 1 – Данные для расчетной части

Пример выполнения работы

Условие задачи

Оценить закон распределения и числовые характеристики генеральной совокупности – объем инновационных товаров (млн руб.). Если известно, что вариационный размах выборки составляет диапазон от 17 до 65 млн руб., параметры выборки C=0,5 и D=0,1.

Исходя из значений вариационного размаха выборки, сформировать выборочные данные.

По полученной выборке значений изучаемого признака генеральной совокупности:

1. Построить вариационный ряд.

2. На основе данных вариационного ряда построить интервальный вариационный ряд.

3. По данным интервального вариационного ряда построить кумулятивную кривую и гистограмму частот.

4. Рассчитать числовые характеристики по данным вариационного ряда.

5. На уровне значимости $\alpha = 0,05$ проверить гипотезу о нормальном распределении генеральной совокупности.

6. Написать заключение о проведенном исследовании свойств генеральной совокупности.

Порядок выполнения работы

В MS Excel создать книгу с названием **Фамилия_группа**, например, **Иванов_4-14**.

1. Генерирование выборочных данных

1. В ячейках первых двух строк прописываем заголовки «шапки» расчетной таблицы (рис. 1).

	Α	В	С	D	E	F	G	Н		1	J	К	
1				Вариационнь	ій ряд				Инте	Интервальный вариационный ряд			
2	Nº	Значения выборки	Фиксация выборки	Упорядоченная выборка	Частота	Шаг интервала	Границ интерва	ца Диапазо ила интервал	он 10в	Частота	Отностительная частота	Накопленная частота	
3	1	83,72	80,55	49,35	1	9,45	44,6	62 [44,62 - 54	1,07]	5	0,025	5	
							_						
								L		Μ	N	0	
										Расчет кр	итерия Пирсона хи	-квадрат	
							Γ	Относительные	Теоретические		Теоретические	Слагаемые	
								накопленные			частоты	критерия	
								частоты	Беро	лиости	Ider of Bi	Пирсона	
								0,025	0,	0107	2,1324	3,8561	

Рисунок 1 – «Шапка» расчетной таблицы

2. В ячейке АЗ вводим значение 1. В активной ячейке АЗ вызываем

функцию Прогрессия. Она расположена на вкладке Главная / Редактирование. В диалоговом окне устанавливаем значения

как на рисунке 2.

Прогрессия		? ×	Γ						
Расположение По строкам По стодбцам По стодбцам	кая ©	ницы день рабочий день месяц год	:- *	Зставить /далить Формат ч Ічейки	Σ	ят (В <u>н</u> из Вправ <u>о</u>	²² ти и 1ить т		
Шаг: 1 Предельное з	начение: 2 ОК	200 Отмена	L		M 🔄	<u>в</u> верх В <u>л</u> ево По лис <u>т</u> ам Про <u>г</u> рессия В <u>ы</u> ровнять	P	Q	

Рисунок 2 – Параметры Прогрессии

В результате работы функции **Прогрессия** столбец А заполнится числами от 1 до 200.

2. Чтобы сформировать выборку данных, в ячейке ВЗ напишем формулу

```
=<mark>А</mark>*НОРМОБР(СЛЧИС();B;C)
```

где А, В, С – соответствуют номеру варианта (табл. 1).

Нажмите F9, убедитесь, что в ячейке B3 генерируются случайные значения.

Синтаксис функции <u>НОРМОБР</u>, <u>СЛЧИС</u> и <u>ОКРУГЛ</u>

Установим точность значений до двух знаков после запятой с помощью функции **ОКРУГЛ**.

Для заполнения последующих 199 ячеек выборочными данными воспользуемся инструментом Excel – Таблица данных (Таблица подстановки). Таблица данных относится к группе инструментов Анализ что если, которая размещена на ленте во вкладке Данные в блоке Работа с данными.

Выделяем диапазон ячеек АЗ:В202

Вызываем диалог Таблица данных

Таблица данных	? 🔀
Подставлять значения по ст <u>о</u> лбцам в:	
Подставлять значения по ст <u>р</u> окам в:	
ОК	Отмена

В поле **Подставлять значения по строкам** указать на любую пустую ячейку, нажать ОК. Столбец В заполнится случайными значениями из диапазона от 17 до 65.

Альтернативный вариант формирования выборки

В активной ячейке ВЗ захватить мышью в нижнем правом углу маркер, как показано на рисунке 3, и «протащить» формулу вниз, заполнив 200 ячеек столбца В.

A	А	В	С	D	E	
1				Вариационнь	ый ряд	
2	NՉ	Значения выборки	Фиксация выборки	Упорядоченная выборка	Частота	
3	1	36,08715116				
4	2					

Рисунок 3 – Копирование формулы

3. При выполнении каких-либо действий на рабочем листе, значения ячеек в столбце В постоянно пересчитываются. В силу этого каждый раз будут появляться новые выборки с различными числовыми характеристиками. Чтобы не потерять исходные выборочные данные их нужно зафиксировать.

Выделяем значения выборки в столбце В, для этого делаем ячейку В2 активной, переходим в конец выборки, удерживаем Shift и кликаем по последней ячейке выборки. Копируем выделенный диапазон (Ctrl+C). Переходим в ячейку СЗ и с помощью команды специальной вставки (Ctrl+Alt+V) вставляем скопированные значения выборки (рис. 4).

Специальная вставка	<u>१</u> ×					
Вставить						
🔘 вс <u>е</u>	🔘 с ис <u>х</u> одной темой					
© <u>ф</u> ормулы	🔘 без рам <u>к</u> и					
• значения	🔘 <u>ш</u> ирины столбцов					
форматы	🔘 форму <u>л</u> ы и форматы чисел					
примечания значения и форматы чисел						
🔘 условия на значения	🔘 все условные форматы объединения					
Операция						
<u>ө н</u> ет	умножить					
© сло <u>ж</u> ить	<u>р</u> азделить					
© в <u>ы</u> честь						
пропускать пустые ячейки	транспонировать					
Вставить связь	ОК Отмена					

Рисунок 4 – Диалоговое окно специальной вставки

2. Построение вариационного ряда

1. Вариационный ряд

Теперь когда выборочные данные зафиксированы, преобразуем их в вариационный ряд.

Выборочные данные, упорядоченные по возрастанию или убыванию, называются вариационным рядом. Различные значения исследуемого признака в выборке называются вариантами.

Для построения вариационного ряда скопируем выборочные данные столбца С в столбец D. Поскольку в ячейках столбца С находятся числовые значения, то при их копировании можно использовать сочетания Ctrl+C и Ctrl+V. После копирования выделяем выборочные данные столбца D и сортируем их по возрастанию.

2. Точечный вариационный ряд

Упорядоченная по возрастанию или убыванию последовательность вариант x_i с указанием частот n_i (или относительной частоты $\frac{n_i}{n}$) их повторения в выборке называется точечным вариационным рядом.

Для построения точечного вариационного ряда с частотами, в ячейке Е2 напишем формулу

=СЧЁТЕСЛИ(\$D\$2:\$D\$201;D2)

и «протянем» ее до конца диапазона. В результате станет понятно какие значения и сколько раз повторяются в выборке.



Синтаксис функции СЧЁТЕСЛИ

3. Интервальный вариационный ряд

Поскольку выборные данные содержат большое количество вариант, по которым достаточно сложно судить об изменчивости их значений, то от точечного вариационного ряда перейдем к интервальному вариационному ряду.

Интервальный вариационный ряд упорядоченная — ЭТО последовательность интервалов с указанием частоты n_i каждого интервала. равной количеству выборочных попавших данных В рассматриваемый интервал.

Рекомендуемое количество интервалов *k* определяется как **натуральное** число по формуле Стерджесса

$$k \approx 1 + 3,222 \cdot \lg n,\tag{1}$$

где *n* – объем выборки.

Длина h интервала определяется по формуле

$$h = \frac{x_{\max} - x_{\min}}{k}, \qquad (2)$$

где x_{\max} , x_{\min} – наибольшее и наименьшее значения в выборке.

Границы интервалов ($C_{i-1}; C_i$) определяются следующим образом:

$$C_0 = x_{\min} - \frac{h}{2}, \ C_1 = C_0 + h, \ \dots, C_k = C_{k-1} + h$$
 (3)

Для построения интервального вариационного ряда определим шаг интервала используя формулу (2).

В ячейке F2 напишем формулу

значения которой будут определять границы интервалов.

Синтаксис функции <u>ЦЕЛОЕ, LOG10</u>

Диапазон интервалов разместим в столбце Н. Для записи интервалов в общепринятом математическом стиле воспользуемся оператором конкатенации, позволяющим объединить несколько данных в одну ячейку. Excel, в качестве оператора конкатенации, применяет символ амперсанда – &.

Сначала в столбце G сформируем значения левой границы интервала. Для этого в соответствии с выражением (3) в ячейке G3 напишем формулу

=ОКРУГЛ(D3-F3/2;2),

а в ячейке G4

Ячейку G4 делаем активной, захватываем мышкой нижний правый маркер и «протягиваем» формулу вниз до первого значения превосходящего максимального значения вариационного ряда.

В ячейке Н4 напишем формулу для первого диапазона

Далее делаем ячейку H4 активной и «протягиваем» формулу до предпоследнего значения (рис. 5).

	Α	В	C	D	E	F	G	н	I
1				Вариационнь	ій ряд			Инте	ервальнь
2	№ Значения выборки		Фиксация выборки	Упорядоченная выборка	Частота	Шаг интервала	Граница интервала	Диапазон интервалов	Частота
3	1	41,37	37,54	28,49	1	3,32	26,83	[26,83 - 30,15]	
4	2	34,35	40,07	30,03	1		30,15	[30,15 - 33,47]	
5	3	36,29	48,97	30,83	1		33,47	[33,47 - 36,79]	
6	4	47,13	45,84	30,87	1		36,79	[36,79 - 40,11]	
7	5	32,25	42,64	31,05	1		40,11	[40,11 - 43,43]	
8	6	42,3	38,17	31,26	1		43,43	[43,43 - 46,75]	
9	7	40,36	45,83	31,3	1		46,75	[46,75 - 50,07]	
10	8	41,58	39,27	32,12	1		50,07	[50,07 - 53,39]	
11	9	41,1	44,2	32,14	1		53,39	[53,39 - 56,71]	
12	10	38,24	44,98	32,62	1		56,71	Bcero	0
		22.0				l '			

Рисунок 5 – Интервальный вариационный ряд

Для определения частот (количество вариант, попадающих в соответствующий интервал) воспользуемся функцией **ЧАСТОТА**.



Функция **ЧАСТОТА** вводится как формула массива после выделения диапазона, это означает, что после ввода формулы необходимо вместо нажатия клавиши ENTER нажать сочетание клавиш **CTRL+SHIFT+ENTER**.

В столбце I выделяем диапазон ячеек равный количеству интервалов. Переходим в строку формул и записываем формулу (рис. 6)

после этого нажимаем сочетание клавиш CTRL+SHIFT+ENTER.

Важно! Первая граница интервала (ячейка G3) не включается в массив интервалов в формуле (4). В разных вариантах, границы диапазонов в формуле (4) могут отличаться.

C	умм	•	×v	<i>f_∞</i> =4A0	CTOTA(D	3:D202 <mark>;G4:G</mark>	612)		
A B C		С	D	E	F	G	н	I	
1				Вариационнь	ій ряд		Интервальны		
2	N≘	Значения выборки	Фиксация выборки	Упорядоченная выборка	Частота	Шаг интервала	Граница интервала	Диапазон интервалов	Частота
3	1	46,44	37,54	28,49	1	3,32	26,83	[26,83 - 30,15]	34:G12)
4	2	45,66	40,07	30,03	1		30,15	[30,15 - 33,47]	
5	3	37,07	48,97	30,83	1		33,47	[33,47 - 36,79]	
6	4	47,95	45,84	30,87	1		36,79	[36,79 - 40,11]	
7	5	43,67	42,64	31,05	1		40,11	[40,11 - 43,43]	
8	6	48,97	38,17	31,26	1		43,43	[43,43 - 46,75]	
9	7	46,55	45,83	31,3	1		46,75	[46,75 - 50,07]	
10	8	45,92	39,27	32,12	1		50,07	[50,07 - 53,39]	
11	9	28,25	44,2	32,14	1		53,39	[53,39 - 56,71]	
12	10	37,48	44,98	32,62	1		56,71	Bcero	0
13	11	46,71	43,37	32,82	1	'			

Рисунок 6 – Формирование диапазона частот

В качестве контроля промежуточных вычислений подсчитаем количество частот (рис. 7), проверяя, что все варианты вариационного ряда учтены. Сумма частот должна совпадать с объемом выборки.

A	Α	В	С	D	E	F	G	н	I
1				Вариационнь	ій ряд			Инте	ервальный
2	N≘	Значения выборки	Фиксация выборки	Упорядоченная выборка Частота ин		Шаг интервала	Граница интервала	Диапазон интервалов	Частота
3	1	39,99	37,54	28,49	1	3,32	26,83	[26,83 - 30,15]	2
4	2	36,81	40,07	30,03	1		30,15	[30,15 - 33,47]	10
5	3	43,52	48,97	30,83	1		33,47	[33,47 - 36,79]	24
6	4	35,01	45,84	30,87	1		36,79	[36,79 - 40,11]	49
7	5	29,67	42,64	31,05	1		40,11	[40,11 - 43,43]	48
8	6	40,64	38,17	31,26	1		43,43	[43,43 - 46,75]	40
9	7	45,41	45,83	31,3	1		46,75	[46,75 - 50,07]	21
10	8	45,99	39,27	32,12	1		50,07	[50,07 - 53,39]	4
11	9	47,18	44,2	32,14	1		53,39	[53,39 - 56,71]	2
12	10	49,21	44,98	32,62	1		56,71	Bcero	200
13	11	39,9	43,37	32,82	1	'			

Рисунок 7 – Контроль учета вариант вариационного ряда

Полученный интервальный вариационный ряд дополним данными об относительных частотах, накопленных частотах и относительных накопленных частотах (рис. 8).

G	н	I	J	К	L
	Инте	ервальны	ій вариационный	ряд	
Граница интервала	Диапазон интервалов	Частота	Отностительная частота	Накопленная частота	Относительные накопленные частоты
26,83	[26,83 - 30,15]	2	0,01	2	0,01
30,15	[30,15 - 33,47]	10	0,05	12	0,06
33,47	[33,47 - 36,79]	24	0,12	36	0,18
36,79	[36,79 - 40,11]	49	0,245	85	0,425
40,11	[40,11 - 43,43]	48	0,24	133	0,665
43,43	[43,43 - 46,75]	40	0,2	173	0,865
46,75	[46,75 - 50,07]	21	0,105	194	0,97
50,07	[50,07 - 53,39]	4	0,02	198	0,99
53,39	[53,39 - 56,71]	2	0,01	200	1
56,71	Bcero	200	1		

Рисунок 8 – Интервальный вариационный ряд

Значения относительных частот определяются по формуле $\frac{n_i}{n}$, где n_i – частота *i* -го интервала, n – объем выборки.

Сумма относительных частот всегда равна единице.

Значения накопленных частот определяются по формуле

$$m_i = n_1 + n_2 + \ldots + n_i$$
 (5)

а относительные накопленные частоты интервалов – $\frac{m_i}{n}$.

4. Графический анализ вариационных рядов

Анализ вариационных рядов и определение формы распределения выборочных данных проводится на основе графического анализа полигона частот, гистограммы частот и кумулятивной кривой.

Вариационные ряды и их графические изображения представляют эмпирическое (выборочное) распределение исследуемого признака генеральной совокупности.

Гистограмма частот строится по интервальному вариационному ряду и представляет собой ступенчатую фигуру на плоскости, состоящую из прямоугольников основаниями которых служат интервалы, а высоты равны частотам или относительным частотам этих интервалов.

Кумулятивная кривая (кумулята) – это кривая относительных накопленных частот (или накопленных частот), строится по точеному вариационному ряду (или интервальному вариационному ряду) и представляет собой плавную линию, проходящую через середины интервалов с высотами равными относительным накопленным частотам этих интервалов.

На основе данных интервального вариационного ряда строим гистограмму частот и кумуляту распределения (рис. 9).





Рисунок 9 – Графическое представление вариационного ряда

3. Числовые характеристики выборки

Для одновременного нахождения всех числовых характеристик выборочных данных воспользуемся инструментом анализа Описательная статистика (рис. 10). Доступ к данному инструменту осуществляется по вкладкам:

Данные — Анализ — Анализ данных — Описательная статистика



Примечание

Если надстройка Анализ данных не подключена, необходимо выполнить следующие действия: Файл → Параметры → Надстройки → Управление (Перейти) → Пакет анализа данных → Ок

Инструменты анализа	_	
<u>encryymentor analysia</u>	_	OK
Двухфакторный дисперсионный анализ с повторениями 🖌		
Двухфакторный дисперсионный анализ без повторений 📃	01	гмена
Корреляция		
Ковариация		
Описательная статистика	<u> </u>	равка
Экспоненциальное сглаживание		
Двухвыборочный F-тест для дисперсии		
Анализ Фурье		
Гистограмма		
Скользящее среднее	-	

Рисунок 10 – Диалоговое окно «Анализ данных»

В диалоговом окне описательной статистики установить флажки в элементах управления окна, как показано на рисунке 11.

-	Описательная статистика	4 8	? X
	Входные данные В <u>х</u> одной интервал: Группирование:	\$D\$3:\$D\$202	ОК Отмена <u>С</u> правка
	 <u>М</u>етки в первои строке Параметры вывода В<u>ы</u>ходной интервал: Новый рабочий <u>л</u>ист: Новая рабочая <u>к</u>нига 	SK\$52	
	✓ Итоговая статистика ✓ Уровень надежности: К-ый наименьший: К-ый наибольший:	95 % 1 1	

Рисунок 11 – Диалоговое окно «Описательная статистика»

В поле **Входной интервал** указывается диапазон ячеек с выборочными данными (столбец D).

В поле Выходной интервал указать любую ячейку ниже графиков гистограммы и кумуляты.

В результате расчетов инструмента анализа Описательная статистика появится таблица (рис. 12).

G	Н	I
Столбец1		
Среднее		41,2415
Стандартная ошибка		0,358147
Медиана		41,28
Мода		33,71
Стандартное отклонение		5,064965
Дисперсия выборки		25,65387
Эксцесс		-0,22188
Асимметричность		-0,0162
Интервал		26,52
Минимум		28,49
Максимум		55,01
Сумма		8248,3
Счет		200
Уровень надежности(95,0%)		0,706251

Рисунок 12 – Числовые характеристики выборки

Показатели столбца 1 в таблице описательной статистики (рис. 12) характеризуют:

1. Положение выборки

Это средние величины, определяющие положение выборки на числовой оси одним числом, вокруг которого концентрируются выборочные данные. Наиболее распространенными характеристиками положения являются:

- среднее – выборочная средняя;

- медиана – медиана выборки (значение признака, приходящее на середину вариационного ряда (выборочных значений, упорядоченных по

возрастанию), т.е. медиана выборки делит выборку на две части равные по частоте);

- мода – мода выборки (варианта с наибольшей частотой).

2. Вариацию выборки

Вариации (рассеяния) выборки описывают изменчивость значений изучаемого признака. Наиболее распространенными характеристиками вариации являются:

- дисперсия выборки – исправленная выборочная дисперсия;

- *стандартная ошибка* – среднее квадратическое отклонение выборочной средней;

- *стандартное отклонение* – исправленное выборочное среднее квадратическое отклонение;

- интервал – вариационный размах (разность между наибольшим и наименьшим значением в выборочных данных).

3. Форму распределения выборки

- эксцесс – показатель «крутости» распределения выборки по сравнению с нормальным распределением;

– асимметричность – мера отклонения распределения выборки от симметричного, при $\tilde{A} = 0$ распределение выборки (полигон частот, гистограмма частот) симметрично относительно прямой $x = \overline{x}$, при $\tilde{A} > 0$ $(\tilde{A} < 0)$ распределение выборки имеет более пологую правую (левую) часть).

4. Дополнительные показатели выборки

- минимум наименьшее значение в выборке данных;
- максимум наибольшее значение в выборке данных;
- сумма сумма всех выборочных данных;
- *счет* объем выборочных данных.

4. Проверка гипотезы о нормальном распределении генеральной совокупности по критерию согласия Пирсона

Критерий согласия Пирсона χ^2 (хи-квадрат) проверяет значимость расхождения эмпирических (наблюдаемых) и теоретических (ожидаемых) частот.

В соответствии с критерием проверяется нулевая гипотеза $H_0: X \sim N(a, \sigma)$ о нормальном распределении генеральной совокупности (случайной величины *X* – объем инновационных товаров). Параметры *а* (математическое ожидание) и σ (среднее квадратическое отклонение) Зa ИХ значения неизвестны. принимаются ИХ несмещенные И состоятельные оценки: выборочная средняя \overline{x} И исправленное выборочное среднее квадратическое отклонение $s = \sqrt{S^2}$. Таким образом, проверяется нулевая гипотеза $H_0: X \sim N(\overline{x}, s)$. Для проверки этой гипотезы используется статистика критерия

$$\chi^{2} = \sum_{i=1}^{k} \frac{(n_{i} - n \cdot p_{i})^{2}}{n \cdot p_{i}},$$
(6)

которая является мерой расхождения теоретического распределения $N(\overline{x},s)$ и эмпирического распределения представленного точечным или интервальным вариационным рядом.

В критерии (6) k – количество интервалов в интервальном вариационном ряду; n_i – частоты интервалов ; n – объем выборки; p_i – теоретические вероятности попадания случайной величины X в интервал (C_{i-1} ; C_i); $n \cdot p_i$ – теоретические частоты интервалов.

Для эмпирического распределения, представленного интервальным вариационным рядом, теоретические вероятности вычисляются с помощью интегральной теоремы Муавра-Лапласа

$$p_{i} = P\left(C_{i-1} \le X \le C_{i}\right) \approx \Phi\left(\frac{C_{i} - \overline{x}}{s}\right) - \Phi\left(\frac{C_{i-1} - \overline{x}}{s}\right)$$
(7)

где $\Phi(x)$ – функция Лапласа.

Если вычисленное по выборке значение критерия χ^2 больше критического значения $\chi^2_{kp}(1-\alpha,k-r-1)$, то нулевая гипотеза $H_0: X \sim N(\overline{x},s)$ отвергается (гипотеза противоречит выборочным данным). Если вычисленное значение критерия $\chi^2 \leq \chi^2_{kp}(1-\alpha,k-r-1)$, то нулевая гипотеза $H_0: X \sim N(\overline{x},s)$, принимается на уровне значимости $\alpha = 0,05$ (гипотеза о нормальном распределении генеральной совокупности с параметрами $a = \overline{x}$ и $\sigma = s$ согласуется с выборочными данными).

Несмещенные точечные оценки математического ожидания \overline{x} и дисперсии генеральной совокупности *s* соответственно равны выборочной средней и исправленной выборочной дисперсии, значения которых берутся из таблицы числовых характеристик выборки (рис. 11). Эти параметры выделены желтым цветом.

Для вычисления значения статистики критерия χ^2 по формуле (6) используем интервальный вариационный ряд. Для этого к полученному ранее интервальному вариационному ряду добавим справа три столбца с заголовками: теоретические вероятности, теоретические частоты и слагаемые критерия Пирсона.

Расчет теоретических вероятностей выполним по формуле (7), для которой функция Лапласа в Excel задается формулой

$$= HOPM.CT.PAC\Pi() \tag{8}$$

Учитывая синтаксис формулы (8) в ячейке МЗ напишем формулу в соответствии с (7)

=HOPM.CT.PACΠ((G4-\$I\$54)/\$I\$58;1)-HOPM.CT.PACΠ((G3-\$I\$54)/\$I\$58;1)



Синтаксис функции НОРМ.СТ.РАСП

«Протягиваем» формулу по всем диапазонам интервального ряда. В итоговой ячейке определяем сумму теоретических вероятностей. Эта величина должна быть приближенно равна единице.

Аналогичным образом, определяем значения теоретических частот, умножая соответствующую теоретическую вероятность на объем выборки. В итоговой ячейке определяем сумму теоретических частот.

В столбце О (слагаемые критерия Пирсона) рассчитываем значения аргументов по формуле (6). Сумма рассчитанных значений, в итоговой ячейке, покажет значение критерия χ^2 .

Проверка гипотезы по критерию хи-квадрат требует расчета критического значения статистики $\chi^2_{kp}(1-\alpha, k-r-1)$ в зависимости от уровня значимости α и числа степеней свободы k-r-1 (параметр r для всех вариантов равен двум).

Расчет критического значения статистики χ^2_{kp} осуществим с помощью функции Excel

Синтаксис функции ХИ2.ОБР.ПХ

В ячейке О13 напишем формулу =ХИ2.ОБР.ПХ(0,05;6).

Сравнивая фактическое и критическое значения статистики приходим к выводу, что на уровне значимости $\alpha = 0,05$ нулевая гипотеза $H_0: X \sim N(\bar{x}, s)$ о нормальном распределении генеральной совокупности с оценками параметров $a = \bar{x} = 41,2415$ и $\sigma = s = 5,0650$ принимаются, поскольку выполняется условие $\chi^2 = 2,1666 < \chi^2_{kp} = 12,5916$.

5. Заключение о проведенном исследовании свойств генеральной совокупности

Проведенное исследование выборки объема инновационных товаров (млн руб.) позволяет сделать ряд выводов:

1. Гистограмма частот близка по форме к кривой нормального распределения, а кумулятивная кривая к графику функции распределения нормальной случайной величины.

2. Выборочная средняя $\overline{x} = 41,2415$, выборочная медиана $\overline{x}_{med} = 41,28$ и выборочная мода $\overline{x}_{mod} = 33,71$ имеют приближенно равные значения, следовательно, эмпирическое распределение симметрично относительно выборочного среднего, что свидетельствует в пользу нормального распределения объема инновационных товаров (млн руб.).

3. Выборочные коэффициент асимметрии $\tilde{A} = -0,0162$ и эксцесс $\tilde{E} = -0,2219$ имеют значения близкие к нулю, что свидетельствует о нормальном распределении исследуемого признака.

4. По критерию Пирсона χ^2 статистическая гипотеза о нормальном распределении генеральной совокупности, с математическим ожиданием равным 41,2415 и средне квадратическим отклонение равным 5,0650, согласуется с выборочными данными на уровне значимости 0,05, так как вычисленное значение статистики χ^2 , равное 2,1666, меньше критического, равного 12,5916.

Таким образом, объем инновационных товаров (млн руб.) в РФ имеет нормальное распределение, в среднем ежегодный объем инновационных товаров составляет 41,2415 млн руб.



Функция НОРМОБР или НОРМ.ОБР в зависимости от версии Excel

Описание. Возвращает обратное нормальное распределение для указанного среднего и стандартного отклонения.

Синтаксис НОРМОБР(вероятность;среднее;стандартное_откл)

Аргументы функции:

Вероятность – вероятность, соответствующая нормальному распределению.

Среднее – среднее арифметическое распределения.

Стандартное_откл – стандартное отклонение распределения.

вернуться назад

Функция СЛЧИС()

Описание. Возвращает равномерно распределенное случайное вещественное число из отрезка [0; 1].

Синтаксис СЛЧИС()

У функции СЛЧИС нет аргументов.

Чтобы получить случайное вещественное число в диапазоне между а и b, можно использовать следующую формулу: a+(b-a)*СЛЧИС()

вернуться назад

Функция ОКРУГЛ

Описание. Округляет число до указанного количества дробных разрядов.

Синтаксис ОКРУГЛ(число;число_разрядов)

Аргументы функции:

Число – округляемое числовое значение.

Число_разрядов – количество дробных разрядов, до которого требуется округлить число.

вернуться к тексту

Функция СЧЁТЕСЛИ

Описание. Подсчитывает количество ячеек, отвечающих определенному условию. Синтаксис СЧЁТЕСЛИ(где нужно искать;что нужно найти)

вернуться к тексту

Функция ЦЕЛОЕ

Описание. Округляет число до ближайшего меньшего целого.

Синтаксис ЦЕЛОЕ(число)

Аргументы функции:

Число – вещественное число, округляемое до ближайшего меньшего целого.

вернуться к тексту

Функция LOG10

Описание. Возвращает десятичный логарифм числа.

Синтаксис LOG10(число)

Аргументы функции:

Число – положительное вещественное число, для которого вычисляется десятичный логарифм.

вернуться к тексту

Функция ЧАСТОТА

Описание. Вычисляет частоту появления значений в интервале значений и возвращает массив чисел.

Синтаксис ЧАСТОТА(массив_данных;массив_интервалов)

Аргументы функции:

Массив_данных – множество числовых данных, для которых вычисляются частоты.

Массив_интервалов – множество интервалов, в которые группируются значения аргумента «массив_данных».

вернуться к тексту

Функция НОРМ.СТ.РАСП в зависимости от версии Excel

Описание. Предназначена для расчета значений плотности вероятности или значений интегральной функции распределения вероятностей по нормированным данным. Синтаксис HOPM.CT.PACП(z;интегральная)

Аргументы функции:

Z- значение, для которого строится распределение.

Интегральная – логическое значение, определяющее форму функции.

Если аргумент «интегральная» имеет значение ИСТИНА (или 1), функция НОРМ.СТ.РАСП возвращает интегральную функцию распределения; если этот аргумент имеет значение ЛОЖЬ (или 0), возвращается весовая функция распределения.

Важно!

В младших версиях Excel используется функция **НОРМСТРАСП(z)**, которая возвращает функцию распределения стандартной нормальной величины.

вернуться к тексту

Функция ХИ2.ОБР.ПХ или ХИ2ОБР в зависимости от версии Excel

Синтаксис ХИ2.ОБР.ПХ(вероятность;степени_свободы)

Описание. Рассчитывает критическое значение критерия при заданной вероятности справа. Функция по сути дублирует функцию ХИ2.ОБР. Но для ХИ2.ОБР.ПХ можно сразу указывать уровень α , а не вычитать его из 1.

Аргументы функции:

Вероятность – вероятность, связанная с распределением хи-квадрат.

Степени_свободы – число степеней свободы.

вернуться к тексту

Контрольные вопросы

1. Сформулируйте понятия генеральной совокупности и выборки.

2. В чем заключается суть выборочного метода?

3. Какую выборку называют репрезентативной?

4. Что представляет собой точечный вариационный ряд?

5. Как строится интервальный вариационный ряд?

6. Что понимается под эмпирическим распределением и с помощью чего оно может быть представлено?

7. Как определяется эмпирическая функция распределения?

8. Как строится гистограмма частот?

9. Приведите числовые характеристики положения выборки.

10. Сформулируйте понятия выборочной моды и медианы.

11. Приведите числовые характеристики вариации выборки.

12. Как находятся выборочные дисперсия и среднее квадратическое отклонение?

13. Какие бывают оценки параметров распределения генеральной совокупности?

14. В чем заключаются свойства несмещенности, состоятельности и эффективности точечных оценок?

26. Сформулируйте понятие доверительного интервала и доверительной вероятности.

27. Как изменяется доверительный интервал с увеличением доверительной вероятности?

28. Какие могут быть ошибки при проверке статистических гипотез?

29. Как определяется мера расхождения теоретического (предполагаемого) и эмпирического распределения в критерии согласия Пирсона?

30. Для проверки каких гипотез применяются критерии согласия?