

ЛЕКЦИОННОЕ ЗАНЯТИЕ

Статистическая совокупность наблюдений. Сбор и формирование

Любое статистическое исследование выражается в следующих основных ступенях:

- массовое статистическое наблюдение, сбор информации;
- сводка и группировка статистических данных, формирование совокупности;
- построение статистических показателей и их анализ.

Статистическим наблюдением называется процесс получения данных о каких-либо явлениях путем регистрации их существенных признаков.

Наблюдения являются важнейшим звеном в статистическом исследовании, так как они дают материал, который в дальнейшем будет подвергнут обработке и анализу. Если материалы, собранные при статистическом наблюдении, будут неполными или неточными, это может повлечь за собой получение неправильного результата при анализе. Поэтому основными требованиями, предъявляемыми к наблюдению, являются правильность и безошибочность его проведения.

Организация наблюдения имеет две основные формы. Первая – регистрация отдельных единиц наблюдения и запись результатов на специальных бланках. Таким образом могут формироваться «журнал наблюдений» или «таблица наблюдений».

Статистическими данными можно назвать сведения о некотором числе объектов, обладающих теми или иными признаками. Метод исследования статистических данных называется статистическим. Характерной особенностью настоящего этапа развития естественных и технических наук является широкое применение статистических методов и математики во всех областях знания. Но статистические методы в применении к различным областям настолько своеобразны, что их невозможно объединить в одну науку.

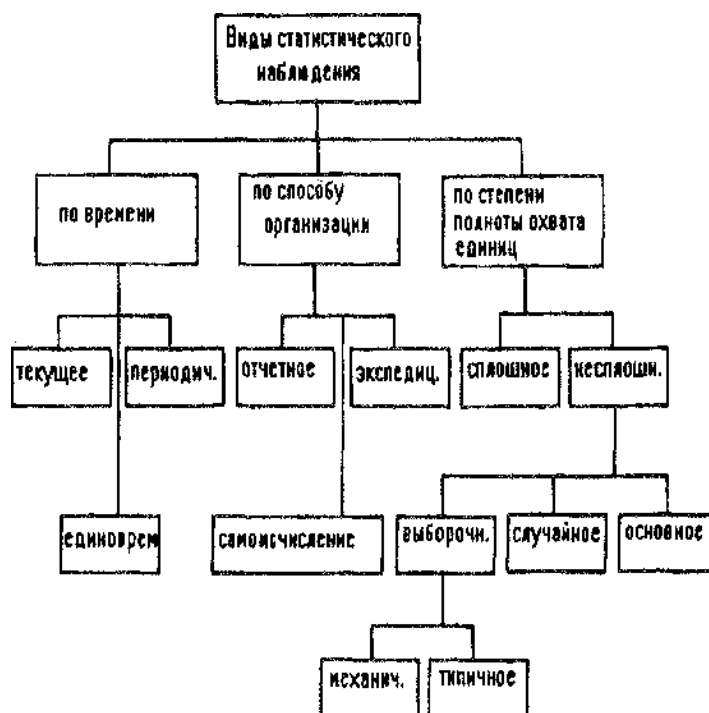


Рис. 1. Виды статистического наблюдения

Рассмотрим применение математических методов исследования статистических данных для решения вопросов горного дела, экономики, организации и управления горным производством.

Статистика находит все большее применение в технике. Однако нельзя переоценивать статистические методы. Правильное применение математического анализа не может быть сведено к одним математическим приемам, а требует, прежде всего, предварительного теоретического анализа, хорошего знания физической сущности явления. Внедрение математической статистики в производственную практику поможет найти дополнительные пути к повышению культуры производства, эффективности использования оборудования и росту производительности труда, снижению себестоимости и увеличению рентабельности.

Вопросы рентабельности, себестоимости и т.д. рассматриваются в разделах и экономической и математической статистики.

Общими чертами статистического метода в различных областях знаний являются: сведения к подсчету числа объектов, входящих в те или иные группы; рассмотрение распределения количественных признаков; применение выборочного метода, когда детальное исследование всех объектов обширной совокупности (например массы полезного ископаемого) затруднительно; использование теории вероятностей при оценке достаточного числа наблюдений. Для некоторых наук вопросы применения вероятностно-статистических методов и машинной математики более или менее разработаны, для других, например, горного дела, еще очень много надо сделать.

Статистические методы опираются на массовые, повторяющиеся явления, в которых изменчивость обуславливается рядом причин. При этом

задача статистики – установить характер этой изменчивости. При исследовании различных физических, химических, экономических, технических процессов часто встречаются явления, называемые случайными.

Как бы точно и подробно ни выполнялись условия отдельных экспериментов, невозможно достичь того, чтобы результаты полностью совпали. Результаты отдельных определений прочностных характеристик одной и той же породы еще больше будут отличаться друг от друга, если эти определения производятся в натуре, в различных участках горного массива, так как вариации факторов, сопровождающие эксперименты и порождающие различия результатов, будут еще более заметны.

Случайные отклонения неизбежно сопутствуют любому закономерному явлению, накладываясь на него.

В природе нет ни одного физического явления, в котором не присутствовали бы в той или иной мере случайности.

В ряде практических задач случайными элементами пренебрегают. При этом из бесчисленного множества признаков, влияющих на данное явление, выделяется главное, влиянием остальных, второстепенных, пренебрегают. Таким образом изучают некоторую модель явления. Затем применяется тот или иной аппарат (например, дифференциальные уравнения) для описания данного явления. Так можно выявить основную закономерность, свойственную данному явлению, и есть возможность предсказать результат при заданных условиях.

Чем больше учтено признаков, влияющих на явление, тем точнее получаемый результат.

Эта схема решения пригодна лишь для задач, где исход опыта зависит от основного количества главных признаков, остающихся постоянными от опыта к опыту.

Множество сопутствующих качеств можно приблизительно представить в виде схемы (рис. 2).

При изучении любого процесса, явления можно выделить совокупность однородных единиц. Такова совокупность отдельных объемов, блоков многоделимой массы горной породы. Такова совокупность единиц средств труда, с помощью которых осуществляется эксперимент: группа машин, приборов или деталей определенного типа.

Совокупность, состоящая из однородных единиц, обладающих количественной общностью, составляет *статистическую совокупность*.

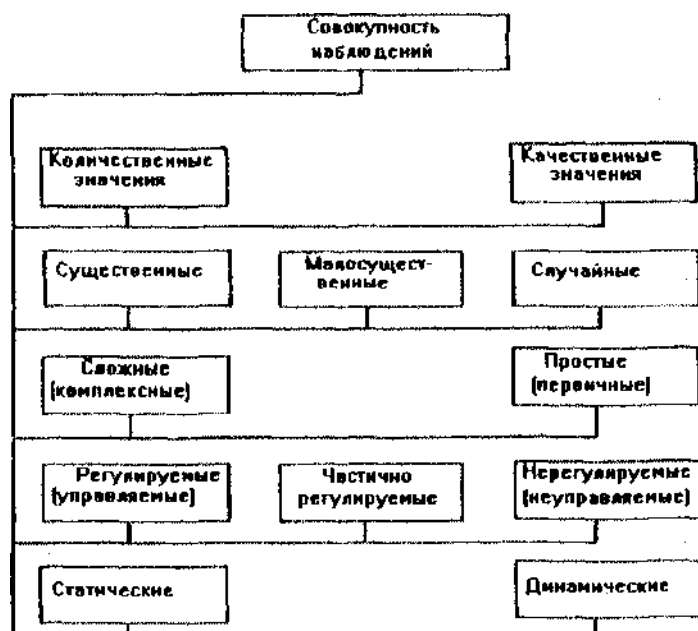


Рис. 2. Классификация наблюдений

Главное орудие статистики – обобщающие показатели, основанные на данных массового наблюдения. Для успешного применения методов математической статистики надо правильно разграничить совокупности, подлежащие обобщению. Такое разграничение производится методом группировки.

Группировка – важнейшее положение математической статистики. Необходимо отметить еще одно обстоятельство: математическая статистика тесно связана с теорией вероятностей. На предельных теоремах теории вероятностей базируется большинство выводов математической статистики.

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ

Выполнение расчетов в электронной таблице MS Excel

Задание: научиться пользоваться встроенными математическими и статистическими функциями MS Excel.

Ход работы:

1. Создать таблицу, приведенную на рисунке 1.
2. Ввести в столбец А данные, указав кавычками их текстовую принадлежность.
3. Записать в клетке В2 функцию =СЛЧИС(), возвращающую случайное число из диапазона $\{0,1\}$, и скопировать ее в ячейки С2:Е2.
4. Скопировать значения ячеек В2:Е2 в ячейки В3:Е3, используя специальную вставку *Значения*.

Внимание! Если копирование значений выполнено правильно, то после каждого пересчета таблицы, данные в ячейках В2: Е2 будут изменяться, а в остальных ячейках будут фиксированы.

5. Увеличить значения ячеек В3:Е3 в 1000 раз и разместить результаты в диапазоне В4:Е4.

6. Ввести функции, указанные в ячейках столбца А, в соответствующие клетки.

	А	В	С	Д	Е
1	Математические и статистические функции				
2	СЛЧИСЛ	0,047599	0,491069	0,306689	0,411324
3	"Случайные числа"	0,612148	0,889984	0,931162	0,367531
4	"Случайные числа"*1000	612,1485	889,9842	931,1624	367,531
5	ОКРУГЛ(В4;2)	612,15	889,98	931,16	367,53
6	ОКРУГЛ(В4;0)	612	890	931	368
7	ОКРУГЛ(В4;-1)	610	890	930	370
8	ОКРУГЛ(В4;-2)	600	900	900	400
9	ОКРУГЛВВЕРХ(В4;2)	612,15	889,99	931,17	367,54
10	ОКРУГЛВВНИЗ(В4;2)	612,14	889,98	931,16	367,53
11	ЧЁТН(В4)	614	890	932	368
12	НЕЧЁТ(В4)	613	891	933	369
13	ОТБР(В4;2)	612,14	889,98	931,16	367,53
14	МАКС(В4:Е4)	931,1624			
15	МИН(В4:Е4)	367,531			
16	СРЗНАЧ(В4:Е4)	700,2065			
17	РАНГ.СР(В4;В4:Е4;1)	2			
18	СРГЕОМ(В4:Е4)	657,1123			
19	СРГАРМ(В4:Е4)	610,5014			
20	ДИСП.В(В4:Е4)	69261,14			
21	ДИСП.РА(В4:Е4)	51945,85			
22	КВАДРОТКЛ(В4:Е4)	207783,4			

Рис. 1. Математические и статистические функции

7. Сохранить созданную таблицу в книге.

Описательная (дескриптивная) статистика в MS Excel

Задание: научиться пользоваться математическими инструментами для вычисления одномерных статистических показателей, дающих представление о совокупности.

Перед выполнением данной работы необходимо включить надстройку «Пакет анализа». Для этого открыть вкладку «Файл», нажать кнопку «Параметры» и выбрать категорию «Надстройки», далее в раскрывающемся списке «Управление» выбрать пункт «Надстройки Excel» и нажать кнопку «Перейти», далее в диалоговом окне «Надстройки» установить флажок «Пакет анализа», а затем нажать кнопку ОК.

Ход выполнения:

1. Сформировать таблицу исходных данных (рис. 2) используя значения для расчета температуры горных пород на глубинах 400-850 м (расчетный шаг 50 м), представленные в табл. 1. Номер варианта выбирается по последней цифре зачетной книжки.

Таблица 1

Значения для расчета температуры горных пород

№ варианта	Глубина нейтрального слоя h_n , м	Температура пород в нейтральном слое T_n , °C	Геотермический градиент района σ , °C/м
1	20	10...12	0,03
2	22	9...11	0,03
3	24	8...10	0,04
4	26	7...9	0,04
5	28	6...8	0,05
6	30	5...7	0,05
7	32	4...6	0,06
8	34	3...5	0,06
9	36	2...4	0,07
0	38	1...3	0,07

В ячейках В3:В12 указываются значения, полученные с помощью функции СЛЧИС()*($T_n^{max}-T_n^{min}$)+ T_n^{min} (T_n^{max} и T_n^{min} – максимальные и минимальные значения температуры пород в нейтральном слое, указанные в табл. 1).

Значения температуры пород, окружающих горные выработки, указанные в ячейках С3:С12, определяются по формуле

$$T_n = T_n + \sigma(H - h_n).$$

	А	В	С
1	Исходные данные		
	Глубина разработки, м	Температура пород в нейтральном слое, °C	Температура горных пород, °C
2			
3	400	=СЛЧИС()* (13-11)+11	=В3+0,05*(А3-20)
4	450	=СЛЧИС()* (13-11)+11	=В4+0,05*(А4-20)
5	500	=СЛЧИС()* (13-11)+11	=В5+0,05*(А5-20)
6	550	=СЛЧИС()* (13-11)+11	=В6+0,05*(А6-20)
7	600	=СЛЧИС()* (13-11)+11	=В7+0,05*(А7-20)
8	650	=СЛЧИС()* (13-11)+11	=В8+0,05*(А8-20)
9	700	=СЛЧИС()* (13-11)+11	=В9+0,05*(А9-20)
10	750	=СЛЧИС()* (13-11)+11	=В10+0,05*(А10-20)
11	800	=СЛЧИС()* (13-11)+11	=В11+0,05*(А11-20)
12	850	=СЛЧИС()* (13-11)+11	=В12+0,05*(А12-20)
13			

Рис. 2. Таблица исходных данных

После формирования таблицы исходных данных необходимо скопировать значения ячеек В3:С12 и вставить в них же, используя

специальную вставку *Значения*. После этого значения данных ячеек не будут изменяться при пересчете таблицы и будут фиксированы.

2. В ленте инструментов выбираем вкладку «Данные», группу «Анализ», средство анализа «Анализ данных», инструмент анализа «Описательная статистика». Для получения одномерного статистического отчета заполняем вызванное диалоговое окно «Описательная статистика» (рис. 3).

Описательная статистика

Входные данные

Входной интервал:

Группирование: ☒ по столбцам ☐ по строкам

☐ Метки в первой строке

Параметры вывода

☒ Выходной интервал:

☐ Новый рабочий лист:

☐ Новая рабочая книга

☒ Итоговая статистика

☒ Уровень надежности: %

☒ К-ый наименьший:

☒ К-ый наибольший:

OK Отмена Справка

Рис. 3. Диалоговое окно инструмента анализа «Описательная статистика»

3. Из полученных данных формируем электронную таблицу (рис. 4). В столбце встроенная функция используем формулы, представленные в табл. 2. Значения одномерных характеристик совокупности, полученные с помощью инструмента анализа «Описательная статистика» и встроенных функций должны совпадать.

	E	F	G
1	Разубоживание		
2	Характеристики	Показатели	Встроенные функции
3	Среднее	42,296138437206	=СРЗНАЧ(С3:С12)
4	Стандартная ошибка	2,42369511935716	=СТАНДОТКЛОН.В(С3:С12)/КОРЕНЬ(СЧЁТ(С3:С12))
5	Медиана	42,2362911802738	=МЕДИАНА(С3:С12)
6	Мода	#Н/Д	=МОДА.ОДН(С3:С12)
7	Стандартное отклонение	7,66439693100228	=СТАНДОТКЛОН.В(С3:С12)
8	Дисперсия выборки	58,7429803159571	=ДИСП.В(С3:С12)
9	Экссесс	-1,08401978461806	=ЭКССЕСС(С3:С12)
10	Асимметричность	-0,108538872102992	=СКОС(С3:С12)
11	Интервал	22,7657531252855	=G13-G12
12	Минимум	30,2728743400278	=МИН(С3:С12)
13	Максимум	53,0386274653133	=МАКС(С3:С12)
14	Сумма	422,96138437206	=СУММ(С3:С12)
15	Счет	10	=СЧЁТ(С3:С12)
16	Наибольший(1)	53,0386274653133	=НАИБОЛЬШИЙ(С3:С12;1)
17	Наименьший(1)	30,2728743400278	=НАИМЕНЬШИЙ(С3:С12;1)
18	Уровень надежности(95,0%)	5,48277927469285	=G4*СТЮДЕНТ.ОБР.2X(1-0,95;G15-1)

Таблица 2

Встроенные статистические функции MS Excel

Характеристики	Встроенная функция
Среднее	=СРЗНАЧ(В3:В12)
Стандартная ошибка	=СТАНДОТКЛОН.В(В3:В12)/КОРЕНЬ(СЧЁТ(В3:В12))
Медиана	=МЕДИАНА(В3:В12)
Мода	=МОДА.ОДН(В3:В12)
Стандартное отклонение	=СТАНДОТКЛОН.В(В3:В12)
Дисперсия выборки	=ДИСП.В(В3:В12)
Экссесс	=ЭКССЕСС(В3:В12)
Асимметричность	=СКОС(В3:В12)
Интервал	=G13-G12
Минимум	=МИН(В3:В12)
Максимум	=МАКС(В3:В12)
Сумма	=СУММ(В3:В12)
Счет	=СЧЁТ(В3:В12)
Наибольший(1)	=НАИБОЛЬШИЙ(В3:В12;1)
Наименьший(1)	=НАИМЕНЬШИЙ(В3:В12;1)
Уровень надежности(95,0%)	=G4*СТЮДЕНТ.ОБР.2Х(1-0,95;G15-1)

4. Вычисляем число интервалов используя формула Стерджесса

$$n = 1 + 3,322 \lg N,$$

где n – число групп;

N – число единиц совокупности.

Длина интервала вычисляется по формуле

$$i = \frac{x_{\max} - x_{\min}}{n},$$

где $x_{\max}$ и $x_{\min}$ – это максимальное и минимальное значения совокупности.

5. Формируем электронную таблицу интервалов (рис. 5). Для начального интервала выбираем минимальное значение совокупности ($x_{\min}$), затем к данному значению прибавляем длину интервала (i) и так далее пока число интервалов не превысит максимального значения совокупности ($x_{\max}$).

Для фиксации ячеек в MS Excel используют символ – доллар (\$). Данный символ можно ввести вручную в строке формул используя фиксацию одновременно по вертикали и горизонтали (\$A\$1), по вертикали (\$A1), по горизонтали (A\$1). Для автоматической фиксации ячейки необходимо нажать на нее, в строке формул выделить ссылку на ту ячейку, которую необходимо зафиксировать и затем нажать клавишу F4 на клавиатуре. Если еще раз нажать на клавишу F4, то зафиксируется только столбец (по вертикали), еще раз – только строка (по горизонтали), еще раз – все вернется к первоначальному виду.

6. Определяем частоту используя функцию =ЧАСТОТА(массив_данных;массив_интервалов). Поскольку данная функция возвращает массив, ее необходимо вводить как формулу массива.

Формула массива – это формула, с помощью которой можно выполнить несколько расчетов для одного или нескольких элементов в массиве. Для того, чтобы редактор воспринял вводимую формулу как формулу массива нажимаем на сочетание клавиш **Ctrl + Shift + Enter**.

	В	С
14	n	=1+3,322*LOG(СЧЁТ(А3:А12))
15	i	=G11/C14
16		
17	Интервалы	Частота
18	=G12	=ЧАСТОТА(С3:С12;В18:В23)
19	=B18+\$C\$15	=ЧАСТОТА(С3:С12;В18:В23)
20	=B19+\$C\$15	=ЧАСТОТА(С3:С12;В18:В23)
21	=B20+\$C\$15	=ЧАСТОТА(С3:С12;В18:В23)
22	=B21+\$C\$15	=ЧАСТОТА(С3:С12;В18:В23)
23	=B22+\$C\$15	=ЧАСТОТА(С3:С12;В18:В23)

Рис. 5. Таблица интервалов

7. По полученным данным строим статистическую диаграмму, по горизонтальной оси – интервалы, по вертикальной – частота.

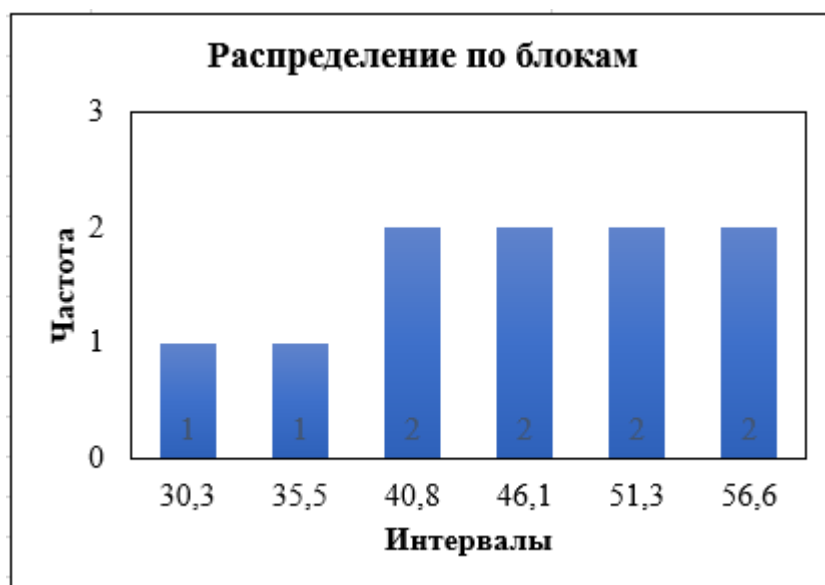


Рис. 6. Гистограмма распределения разубоживания по блокам