

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**
Федеральное государственное автономное
образовательное учреждение высшего образования
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Методические указания

по выполнению лабораторных работ по дисциплине

Математические методы в нефтегазовой геологии

Специальность
Специализация

21.05.02 Прикладная геология
Геология месторождений нефти и газа

Ставрополь, 2026

СОДЕРЖАНИЕ

Предисловие	
Лабораторная работа 1. Принципы и методы математического моделирования в геологии	
Лабораторная работа 2. Вероятностно-статистический анализ с применением ранжирования исходных данных	
Лабораторная работа 3. Распределение синкретных совокупностей	
Лабораторная работа 4. Распределение дискретных совокупностей	
Лабораторная работа 5. Корреляция геологических данных	
Лабораторная работа 6. Многофакторная корреляция в форме общей линейной модели	
Лабораторная работа 7. Группирование геологических объектов	
Лабораторная работа 8. Тренд-анализ геологических параметров	
Лабораторная работа 9. Дисперсионный анализ	

ПРЕДИСЛОВИЕ

Применение статистических методов тесно связано с построением математических моделей. Они позволяют распространить выводы, получаемые по ограниченному числу наблюдений (выборке), на значительно большую группу объектов – популяцию (генеральную совокупность). Применение подходящей статистической модели дает в руки исследователя аппарат, который усиливает надежность обобщений путем распространения статистических выводов, полученных по выборке, на всю совокупность.

Роль статистики в геологии заключается в разработке методов, позволяющих извлекать максимальную информацию из выборочных данных, точно так же, как физики и химики изучают законы строения материи, используемые при интерпретации геологических явлений.

Цель освоения дисциплины «Математические методы в нефтегазовой геологии» является формирование общекультурных компетенций будущего специалиста в области обучения, воспитания и развития, соответствующих целям ОП ВО по специальности 21.05.02 - «Прикладная геология».

ЛАБОРАТОРНАЯ РАБОТА 1

Принципы и методы математического моделирования в геологии

Цель: Данная лабораторная направлена на формирование компетенции, предусматривающей готовность обобщать, анализировать, воспринимать информацию, ставить цели и выбирать пути ее достижения.

Теоретическая часть. При построении статистических моделей исходные геологические данные используются в виде определенных наборов или совокупностей. По существу, такие геологические совокупности и являются предметом количественных исследований в геологии.

Совокупность характеризует изменчивость какого-либо геологического признака. По своей структуре совокупности могут быть одномерными, т. е. когда значение признака в одной точке характеризуется одним числом (например, глубина залегания пласта), или многомерными, когда значение признака характеризуется рядом чисел (например, коллекторские свойства характеризуются двумя числами – пористостью и проницаемостью).

В начале исследований необходимо сгруппировать исходные данные, распределив их по ранжированным равномерным интервалам. Количество интервалов n можно определить по формуле 1.1:

$$n = 1 + 3,322 \lg N, \quad (1.1)$$

где N – общее число элементов или объем исходной статистической совокупности.

В случае разбиения совокупности на равные интервалы их протяженность определяется *формулой Стерджесса* (1.2):

$$d = \frac{h_{\max} - h_{\min}}{n}, \quad (1.2)$$

где d – длина интервала; $h_{\max}$, $h_{\min}$ – значение изучаемого параметра, соответственно, максимальное и минимальное в объеме исходной статистической совокупности.

Границы j -го интервала определяются из выражений 1.3:

$$h_{N_j} = h_{\min} + d(j-1), \quad h_{K_j} = h_{\min} + dj, \quad (1.5)$$

где h_{N_j} , h_{K_j} – соответственно левая и правая граница j -го интервала. Необходимо учитывать, что левая граница принадлежит данному интервалу, а правая не принадлежит, исключая последний интервал.

Далее строим гистограммы распределения совокупности (рисунок 1.1) в координатах: 1) ось X – длина интервалов значений изучаемых параметров; 2) по оси Y откладываются две

величины: а) f_j – частота j -го интервала (число значений совокупностей, попавший в каждый интервал) и б) P_j – относительные частоты или частоты, вычисляемые по формуле 1.4:

$$P_j = \frac{f_j}{N}, \quad (1.4)$$

где P_j – относительная частота или частость, f_j – частота j -го интервала, N – объем совокупности.

Рисунок 1.1 – Пример построения гистограммы

Совокупность характеризуется рядом параметров, из которых наиболее важны средние значения и дисперсия.

Положение центра совокупности характеризует оценка *математического ожидания* или *выборочное среднее*, вычисляемое по формуле средней арифметической, где i -ые значения неранжированы, j -ые ранжированные (сгруппированные), вычисляется по формулам 1.5 и 1.6:

$$\bar{h} = \frac{1}{N} \sum_{i=1}^N h_i, \quad (1.5)$$

$$\bar{h} = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j, \quad (1.6)$$

для невзвешенных и сгруппированных данных соответственно.

Здесь $\bar{h}$ – математическое ожидание; f_j – частота j -го интервала; h_j – среднее значение признака в j -ом интервале, равное $h_j = (h_{Nj} + h_{Kj}) / 2$; n – число интервалов.

Размах значений совокупности характеризуется его *дисперсией* σ^2 , вычисляемой для равновероятных данных по формулам 1.7 и 1.8:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (h_i - \bar{h})^2, \quad (1.7)$$

$$\sigma^2 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j (h_j - \bar{h})^2, \quad (1.8)$$

для невзвешенных и сгруппированных данных соответственно.

Более нагляден так называемый *стандарт* или среднее *квадратическое отклонение от средней* (корень квадратный из центрального момента второго порядка по выборочным данным), равный корню квадратному по дисперсии 1.9 и 1.10:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h})^2}, \quad (1.9)$$

$$\sigma = \sqrt{\frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j (h_i - \bar{h})^2}, \quad (1.10)$$

для невзвешенных и сгруппированных данных соответственно.

Несмещенную точечную оценку дисперсии определяем по формулам 1.11 и 1.12:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (h_i - \bar{h})^2, \quad (1.11)$$

$$S^2 = \frac{1}{\sum_{j=1}^n f_j - 1} \sum_{j=1}^n f_j (h_i - \bar{h})^2, \quad (1.12)$$

для невзвешенных и сгруппированных данных соответственно.

Кумулятивные частоты вычисляют по формуле 1.13:

$$f_{kj} = \sum_{i=1}^j f_i, \quad (1.13)$$

Далее строим гистограмму распределения совокупности (рисунок 1.2) в координатах: ось X – длина интервалов значений изучаемых параметров, ось Y величина f_{kj} – кумулятивная частота j -го интервала (1.13).

Наиболее полная характеристика совокупности вытекает из анализа и сопоставления моментов распределения. Более того, значения параметров совокупности – средней и дисперсии – являются лишь частными случаями этих общих характеристик.

Оборудование и материалы. Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными для выполнения работы являются совокупности геологических параметров $h_1, \dots, h_n$, где h – значения пористости, толщин, коэффициентов заполнения ловушек, нефтеотдачи, амплитуд, площадей, объемов складок, площадей или линейных размеров залежей, их запасов и т. д., приведённых в Приложении А.

Указания по технике безопасности. Компьютер – высокотехнологичное технически хорошо продуманное устройство, но вместе с тем очень опасное. Иногда опасность реальна, а иногда, он незаметно воздействует на Ваше здоровье и психику.

Во избежание несчастного случая, поражения электрическим током, поломки оборудования, рекомендуется выполнять следующие правила:

1. Не входить в помещение, где находится вычислительная техника без разрешения старшего (преподавателя).

2. Не включать без разрешения оборудование.
3. При несчастном случае или поломке оборудования позвать старшего (преподавателя). Знать где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
4. Не трогать провода и разъемы (возможно поражение электрическим током).
5. Не допускать порчи оборудования.
6. Не работать в верхней одежде.
7. Не прыгать, не бегать (не пылить).
8. Не шуметь.

Задания.

Порядок выполнения работы:

В работе требуется определить:

- 1) количество интервалов n по формуле 1.1;
- 2) протяженность интервала d (1.2);
- 3) границы интервалов h_{N_j}, h_{K_j} (1.3).

Затем

4) строим гистограммы распределения совокупности в координатах: ось X – длина интервалов значений изучаемых параметров, ось Y – величина а) f_j – частота j -го интервала (число значений совокупностей, попавший в каждый интервал) и б) P_j – относительные частоты или частоты, определяемые по формуле 1.4.

Далее рассчитываются:

- 5) математическое ожидание или выборочное среднее по формулам 1.5 и 1.6;
- 6) дисперсия по формулам 1.7 и 1.8;
- 7) стандарт, или среднее квадратическое отклонение от средней по формулам 1.9 и 1.10;
- 8) несмещенная точечная оценка дисперсии по формулам 1.11 и 1.12;
- 9) кумулятивные частоты по формуле 1.13 и строится их гистограмма;

Содержание отчета.

В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Результаты расчетов привести в итоговой таблице 1.1. Защита работы проводится в устной форме.

Таблица 1.1 – Расчет основных характеристик совокупности геологических данных

№ интервала	Границы интервала	Середина интервала	Частота f_j	Частость, P_j	$f_j h_j$		$h_j - h_{cp}$	$f_j (h_j - h_{cp})$	$f_j (h_j - h_{cp})^2$
-------------	-------------------	--------------------	---------------	-----------------	-----------	--	----------------	----------------------	------------------------

1	2	3	4	5	6		7	8	9
			$\sum f_j$		$\sum f_j h_j$			$\sum f_j (h_j - h_{cp})$	$\sum f_j (h_j - h_{cp})^2$

Контрольные вопросы.

1. Какими параметрами характеризуются совокупности геолого-геофизических данных?
2. Каким образом рассчитывается количество интервалов n , протяженность интервала d и границы интервалов h_{N_j}, h_{K_j} ?
3. Какие графические построения позволяют представить частоты значений изучаемых признаков? Приведите необходимые для этого расчетные формулы.
4. Каким образом рассчитывается математическое ожидание и выборочное среднее?
5. Что характеризуют такие параметры совокупностей, как дисперсия и стандарт?
6. Чем отличаются расчёт величины дисперсии и несмещенной точечной оценки дисперсии? Каким образом определяются величины стандарт и среднее квадратического отклонения от средней.

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. /. М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.
5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с
6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.
7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.
8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 2

Вероятностно-статистический анализ с применением ранжирования исходных данных

Цель: Освоение методов расчета основных статистических показателей совокупностей и проверка статистических гипотез. Оценка статистических характеристик центра совокупности, расчёт моментов распределения и определение типа распределения геолого-геофизических данных.

Тема занятия:

1. Методы расчета основных статистических показателей совокупностей.
2. Проверка статистических гипотез.
3. Статистические характеристик центра совокупности.
4. Понятие моментов распределения. Асимметрия и эксцесс, связь их значений с формой кривой функции плотности распределения.

Лабораторная работа направлена на формирование компетенции , предусматривающей готовность обобщать, анализировать, воспринимать информацию, ставить цели и выбирать пути ее достижения.

Теоретическая часть. Размах значений исследуемой совокупности составляет от $h_{cp} - \sigma$ до $h_{cp} + \sigma$.

Распределение точечных оценок дисперсии используется для оценки доверительных интервалов. Однако в связи с асимметрией кривой распределения доверительные интервалы дисперсии неодинаковы.

Минимальное значение дисперсии определяется формулой:

$$\mu_1 = \frac{(N-1)S^2}{\chi_{1-\alpha/2, N-1}^2}, \quad (2.1)$$

а максимальное:

$$\mu_2 = \frac{(N-1)S^2}{\chi_{\alpha/2, N-1}^2}, \quad (2.2)$$

где χ^2 – доля площади, ограниченная кривой распределения, соответствующая выбранному уровню значимости α и числу степеней свободы m . Значения χ^2 табулированы и, кроме того, значения $\chi_{1-\alpha/2, m-1}^2$ и $\chi_{\alpha/2, m-1}^2$ с очень большой степенью точности можно рассчитать в программной среде EXCEL.

Истинное значение дисперсии, таким образом, с вероятностью $1-\alpha$ заключено в интервале

$$\mu_1 \leq \sigma^2 \leq \mu_2. \quad (2.3)$$

В различных геологических задачах часто используют *степенную среднюю*, определяемую по формуле:

$$\bar{h}_{cm} = \sqrt[k]{\frac{1}{N} \sum_{i=1}^N h_i^k}, \quad (2.4)$$

$$\bar{h}_{cm} = \sqrt[k]{\frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^k}, \quad (2.5)$$

для невзвешенных и сгруппированных данных соответственно.

При разных значениях k образуется ряд средних: средняя гармоническая ($k=-1$), средняя геометрическая ($k=0$), средняя арифметическая ($k=1$), средняя квадратическая ($k=2$) и средняя кубическая ($k=3$).

Средняя гармоническая используется в том случае, когда значения исходных данных представлены обратными величинами:

$$\bar{h}_{гарм} = \frac{N}{\sum_{i=1}^N \frac{1}{h_i}}, \quad (2.6)$$

$$\bar{h}_{гарм} = \frac{\sum_{j=1}^n f_j}{\sum_{j=1}^n \frac{f_j}{h_j}}, \quad (2.7)$$

для невзвешенных и сгруппированных данных соответственно.

Среднюю геометрическую вычисляют обычно для количественной оценки темпов изменения признака:

$$\lg \bar{h}_{\text{геом}} = \frac{1}{N} \sum_{i=1}^N \lg h_i, \quad (2.8)$$

$$\lg \bar{h}_{\text{геом}} = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j \lg h_j, \quad (2.9)$$

для невзвешенных и сгруппированных данных соответственно.

Значения *средней квадратической* рассчитывают по формулам:

$$\bar{h}_{\text{кв}} = \sqrt{\frac{1}{N} \sum_{i=1}^N h_i^2}, \quad (2.10)$$

$$\bar{h}_{\text{кв}} = \sqrt{\frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^2}, \quad (2.11)$$

для невзвешенных и сгруппированных данных соответственно.

Значения *средней кубической* рассчитывают по формулам:

$$\bar{h}_{\text{куб}} = \sqrt[3]{\frac{1}{N} \sum_{i=1}^N h_i^3}, \quad (2.12)$$

$$\bar{h}_{\text{куб}} = \sqrt[3]{\frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^3}, \quad (2.13)$$

для невзвешенных и сгруппированных данных соответственно.

Значения *средней четвертой степени* рассчитывают по формулам:

$$\bar{h}_4 = \sqrt[4]{\frac{1}{N} \sum_{i=1}^N h_i^4}, \quad (2.14)$$

$$\bar{h}_4 = \sqrt[4]{\frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^4}, \quad (2.15)$$

для невзвешенных и сгруппированных данных соответственно.

При правильном расчете степенных средних соблюдается контрольное равенство:

$$\bar{h}_{\text{гарм}} < \bar{h}_{\text{геом}} < \bar{h}_{\text{ср}} < \bar{h}_{\text{кв}} < \bar{h}_{\text{куб}}, \quad (2.16)$$

Наиболее полная характеристика совокупности вытекает из анализа и сопоставления моментов распределения. Более того, значения параметров совокупности – средней и дисперсии – являются лишь частными случаями этих общих характеристик.

Различают моменты начальные и центральные. *Начальные моменты* вычисляют по формулам:

$$\mu_1 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j \quad (2.17)$$

где μ_1 – начальный момент первого порядка, представляющий собой среднее арифметическое;

$$\mu_2 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^2 \quad (2.18)$$

где μ_2 – начальный момент второго порядка;

$$\mu_3 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^3 \quad (2.19)$$

где μ_3 – начальный момент третьего порядка;

$$\mu_4 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j h_j^4 \quad (2.20)$$

где μ_4 – начальный момент четвертого порядка.

Проверкой правильности расчетов служат равенства:

$$h_{cp} = \mu_1, \quad h_{kv} = \sqrt{\mu_2}, \quad h_{kub} = \sqrt[3]{\mu_3}, \quad h_4 = \sqrt[4]{\mu_4}. \quad (2.21)$$

Центральные моменты представляют собой средние значения степеней отклонений случайной величины от математического ожидания. Их вычисляют по формулам:

$$v_1 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j (h_j - \bar{h}) \quad (2.22)$$

где v_1 – центральный момент первого порядка, всегда равный нулю, так как суммы положительных и отрицательных отклонений от средней равны;

$$v_2 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j (h_i - \bar{h})^2 \quad (2.23)$$

где v_2 – центральный момент второго порядка, представляющий собой дисперсию;

$$v_3 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j (h_i - \bar{h})^3 \quad (2.24)$$

где v_3 – центральный момент третьего порядка;

$$v_4 = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j (h_i - \bar{h})^4 \quad (2.25)$$

где v_4 – центральный момент четвертого порядка.

На основе центральных моментов предлагается вычислить *нормированные моменты* по формуле:

$$r_k = \frac{v_k}{v_2^{\frac{k}{2}}} \quad (2.26)$$

где $k=1, 2, 3, 4$ и v_k – центральные моменты 1, 2, 3 и 4-го порядка соответственно. Отсюда $r_1 = 0$, $r_2 = 1$, $r_3 = v_3/s_3$, $r_4 = v_4/s_4$.

Нормированный момент третьего порядка служит для оценки асимметрии кривой плотности вероятности и называется *асимметрией*.

Ее расчет проводится по формуле:

$$r_3 = \frac{v_3}{\sqrt{v_2^3}} \quad (2.27)$$

Положительное значение асимметрии ($r_3 > 0$, рисунок 2.1) свидетельствует о смещении максимума кривой распределения относительно максимума гистограммы вправо (правая асимметрия), а при $r_3 < 0$ отмечается смещение максимума кривой влево (левая асимметрия). Естественно, что при $r_3 = 0$ совокупности симметричны.

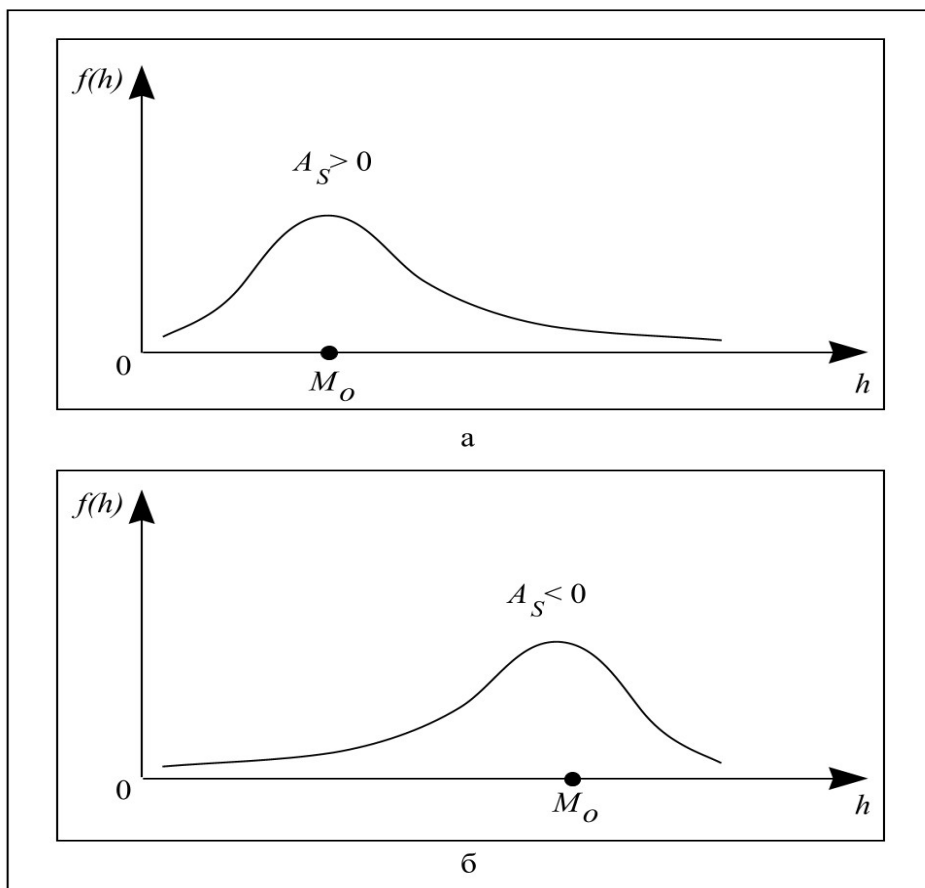


Рисунок 2.1. Оценки формы кривой распределения плотности вероятности по асимметрии.

Значение нормированного момента четвертого порядка используется для оценки формы кривой распределения плотности вероятности и называется *эксцессом*.

Расчет ведется по формуле:

$$E = r_4 - 3 = \frac{v_4}{\sqrt{v_2^4}} - 3 \quad (2.28)$$

При $E < 0$ распределение характеризуется плоской и широкой вершиной, при $E > 0$ – выпуклой и острой (рисунок 2.2.).

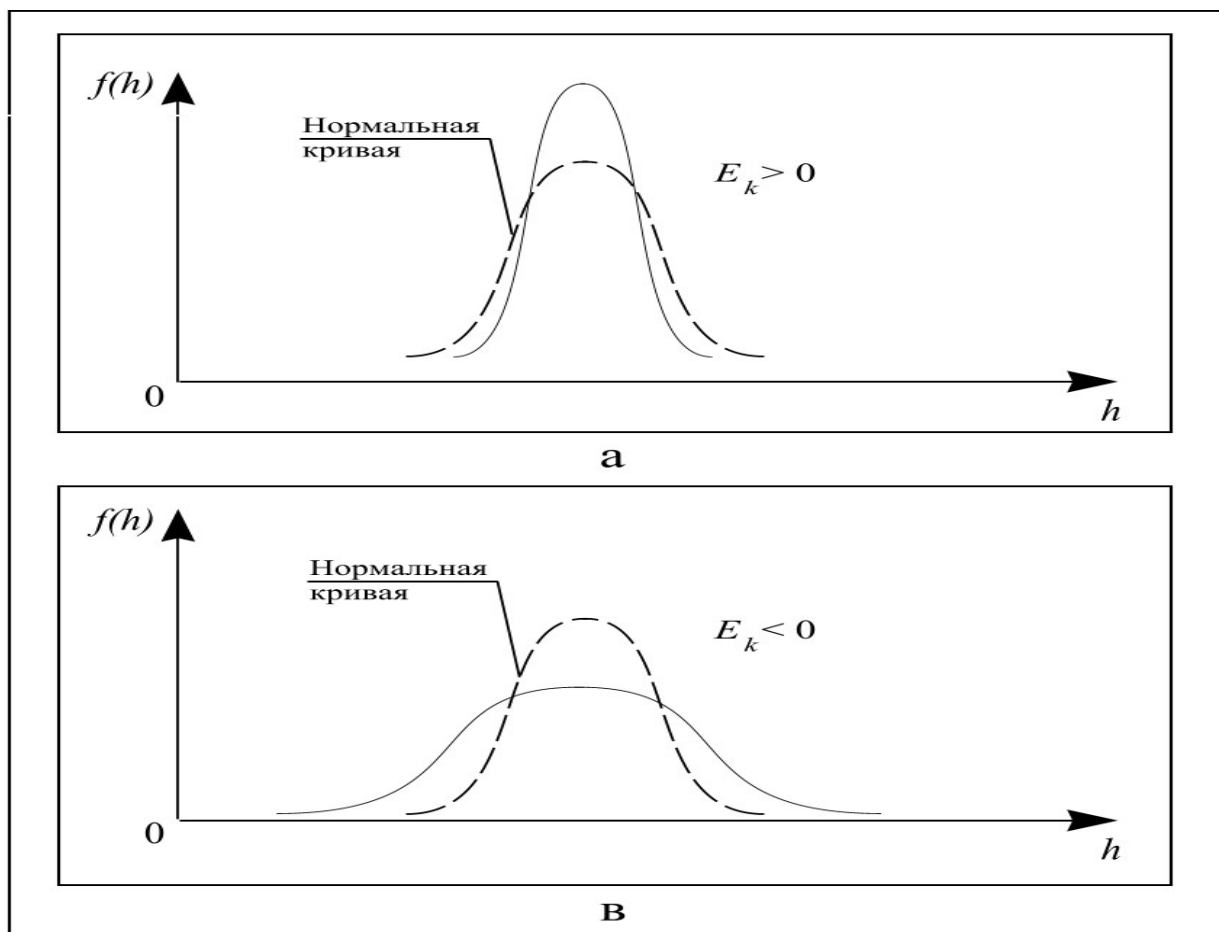


Рисунок 2.2. Оценки формы кривой распределения плотности вероятности по эксцессу.

При использовании математических методов для решения геологических задач значительная роль отводится проверке статистических гипотез о правильности выбранного закона распределения. Одна из наиболее распространенных задач – проверка соответствия изучаемой совокупности нормальному закону распределения. Возможность такой проверки вытекает из анализа моментов. Так показателем нормального распределения является его симметрия, т. е.

$$r_3 = 0, \quad (2.29)$$

и равенство

$$v_4 = 3 v_2^2. \quad (2.30)$$

$$E = 0 \quad (2.31)$$

Оборудование и материалы. Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Исходными данными для выполнения работы являются совокупности геологических параметров $h_1, \dots, h_n$, выданные в лабораторной работе 1.

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы

Определить:

1. Размах значений исследуемой совокупности, минимальное и максимальное значения дисперсии по формуле 2.1 и 2.2 проверить неравенство 2.3.

Затем рассчитать:

2. *Степенные средние* (2.4 – 2.5) и проверить неравенство 2.16;

3. По формулам 2.17 – 2.20 *начальные моменты* и по формуле 2.21 провести проверку правильности их расчета;

4. По формулам 2.22–2.25 *центральные моменты*;

5. По формуле 2.26 *нормированные моменты*.

Определить:

6. *Асимметрию* (2.27);

7. *Экцесс* (2.28).

Далее

8. Провести проверку статистических гипотез о правильности выбранного закона распределения по формулам 2.29–2.31;

Содержание отчета.

В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 2.1.

Таблица 2.1 – Расчет основных характеристик совокупности геологических данных

№ ин-тервала	Границы ин-тервала	Середина интервала	Частота f_j	Частость, P_j	$f_j h_j$	$h_j - h_{cp}$
1	2	3	4	5	6	7
			$\sum f_j$		$\sum f_j h_j$	

Продолжение таблицы 2.1

$(h_j - h_{cp})^2$	$(h_j - h_{cp})^3$	$(h_j - h_{cp})^4$	$f_j(h_j - h_{cp})$	$f_j(h_j - h_{cp})^2$	$f_j(h_j - h_{cp})^3$	$f_j(h_j - h_{cp})^4$
--------------------	--------------------	--------------------	---------------------	-----------------------	-----------------------	-----------------------

8	9	10	11	12	13	13
			$\sum f_j(h_j - h_{cp})$	$\sum f_j(h_j - h_{cp})^2$	$\sum f_j(h_j - h_{cp})^3$	$\sum f_j(h_j - h_{cp})^4$

Контрольные вопросы:

1. Каким образом определяется размах значений совокупности?
2. Какое распределение используется для оценки доверительных интервалов дисперсии? В связи с чем доверительные интервалы неодинаковы?
3. Приведите формулы для определения минимального и максимального значения дисперсии и интервала, в котором заключено истинное значение дисперсии.
4. Приведите формулы для оценки степенных средних и проверочное неравенство.
5. Каким образом определяются начальные моменты, и каким образом проводится проверка правильности их расчета?
6. Что должна включать в себя полная характеристика совокупности?
7. Каким образом определяются центральные и нормированные моменты?
8. Как изменяется форма кривой плотности вероятности нормального распределения в зависимости от знака асимметрии и эксцесса? Приведите расчётные формулы для определения асимметрии и эксцесса.
9. Какая гипотеза выдвигается в первую очередь при проверке статистических гипотез о законе распределения выборочной совокупности геолого-геофизических данных?
10. Приведите необходимые формулы и охарактеризуйте методику проверки соответствия изучаемой совокупности нормальному закону распределения.

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У., Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 3

Распределение синкретных совокупностей

Цель: Освоение методов расчета основных статистических показателей совокупностей и проверка статистических гипотез. Построение вероятностно-статистической модели одномерной геологической совокупности и определение типа распределения геолого-геофизических данных.

Теоретическая часть. При принадлежности исходной совокупности геолого-геофизических данных к нормальному закону следует рассчитать значения *функции плотности нормального распределения* для середины каждого интервала. Расчет ведется по формуле 3.1:

$$N_j(h_j, h_{cp}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(h_j - h_{cp})^2}{2\sigma^2}} \quad (3.1)$$

Плотность относительной частоты рассчитывается по формуле 3.2:

$$n_{jn} = \frac{f_j}{dN} \quad (3.2)$$

На гистограмму плотности относительных частот наносятся результаты расчета функции плотности нормального распределения в виде кривой плотности вероятности нормального распределения.

Если изучаемое распределение геолого-геофизических параметров не принадлежит к нормальному, следует рассмотреть принадлежность изучаемой совокупности к логнормальному распределению.

Функция плотности вероятности имеет вид:

$$L_j(h_j, \theta, \gamma) = \frac{1}{h_j \theta \sqrt{2\pi}} e^{-\frac{(\ln h_j - \gamma)^2}{2\theta^2}} \quad (3.3)$$

где θ, γ – параметры распределения, которые оцениваются по формулам 3.4, 3.7:

– дисперсия логарифмов значений (3.4, 3.5)

$$\theta^2 = \frac{1}{N} \sum_{i=1}^N (\ln h_i - \gamma)^2, \quad (3.4)$$

$$\theta^2 = \frac{1}{\sum_{j=1}^n f_j - 1} \sum_{j=1}^n f_j (\ln h_j - \gamma)^2, \quad (3.5)$$

соответственно для неранжированных и ранжированных данных;

– *среднее значение логарифмов:*

$$\gamma = \frac{1}{N} \sum_{i=1}^N \ln h_i, \quad (3.6)$$

$$\gamma = \frac{1}{\sum_{j=1}^n f_j} \sum_{j=1}^n f_j \ln h_j, \quad (3.7)$$

соответственно для неранжированных и ранжированных данных.

Логнормальное распределение часто характеризуется такими привычными параметрами, как *среднее арифметическое*:

$$\bar{h}_L = e^{\gamma + \frac{1}{2}\theta^2}; \quad (3.8)$$

дисперсия:

$$\sigma_L^2 = e^{2\gamma + 2\theta^2} (e^{2\theta^2} - 1), \quad (3.9)$$

медиана, делящая совокупность на две равные части:

$$Me_L = e^{\gamma}; \quad (3.10)$$

мода, соответствующая наиболее распространенному значению:

$$Mo_L = e^{\gamma - \theta^2}. \quad (3.11)$$

Из сопоставления вычисленных значений *среднего арифметического, дисперсии, медианы и моды* должно выполняться неравенство:

$$Mo_L < Me_L < \bar{h}_L. \quad (3.12)$$

Оборудование и материалы. Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Исходными данными для выполнения работы являются совокупности геологических параметров $h_1, \dots, h_n$, выданные в лабораторной работе 1.

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Методика и порядок выполнения работы

В работе требуется рассчитать:

1) значения *функции плотности нормального распределения* для середины каждого интервала (3.1);

2) значения *плотности относительной частоты* (3.2).

Далее

3) строится гистограмма плотности относительных частот и на нее наносятся результаты расчета функции плотности нормального распределения в виде кривой плотности вероятности нормального распределения.

Если изучаемое распределение геолого-геофизических параметров не принадлежит к нормальному, изучается принадлежность изучаемой совокупности к логнормальному распределению.

Рассчитываются:

4) *среднее значение логарифмов* (3.6)

5) *дисперсия логарифмов значений* (3.4);

6) *функция плотности вероятности* (3.3);

7) параметры логнормального распределения *среднее арифметическое* (3.8), *дисперсия* (3.9), *медиана* (3.10), *мода* (3.11).

8) Проверяется неравенство (3.12).

Содержание отчета.

В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 3.1

Таблица 3.1 – Расчет основных характеристик логнормального распределения геологических данных

Среднее значение параметра в группе h_j	Частота, f_j	$\ln h_j$	$f_j \ln h_j$	$\ln h_j - \gamma$	$(\ln h_j - \gamma)^2$	$f_j (\ln h_j - \gamma)^2$
1	2	3	4	5	6	7
	$\sum_{j=1}^n f_j$		$\sum_{j=1}^n f_j \ln h_j$			$\sum_{j=1}^n f_j (\ln h_j - \gamma)^2$

Контрольные вопросы:

1. Приведите формулы для расчёта значений *функции плотности вероятности нормального распределения* для середины каждого интервала и *плотности относительной частоты*

2. Каким образом проводится построение и анализ кривой функции плотности вероятности нормального распределения? Какими эмпирическими графиками при этом пользуются?

3. Приведите необходимые формулы и охарактеризуйте методику проверки соответствия изучаемой совокупности нормальному закону распределения.

4. Приведите формулы для расчёта значений *функции плотности вероятности логнормального распределения*.

5. Приведите формулы для расчёта параметров логнормального распределения *среднего арифметического, дисперсии, медианы, моды*.

6. Изобразите на графике функцию плотности вероятности логнормального распределения, укажите расположение основных статистических характеристик центра совокупности. Какой закономерности они подчиняются?

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 4

Распределение дискретных совокупностей

Цель: Освоение методики анализа дискретных совокупностей геологических данных. Вышеизложенное позволяет сформировать компетенцию, предусматривающей готовность обобщать, анализировать, воспринимать информацию, ставить цели и выбирать пути ее достижения.

Теоретическая часть. Если производится несколько испытаний, причем вероятность события A в каждом испытании не зависит от исходов других испытаний, то такие испытания называют *независимыми относительно события A* .

В разных независимых испытаниях событие A может иметь либо различные вероятности, либо одну и ту же вероятность. Будем далее рассматривать лишь такие независимые испытания, в которых событие A имеет одну и ту же вероятность.

Ниже воспользуемся понятием *сложного события*, понимая под ним совмещение нескольких отдельных событий, которые называют *простыми*.

Пусть производится n независимых испытаний, в каждом из которых событие A может появиться либо не появиться. Условимся считать, что вероятность события A в каждом испытании одна и та же, а именно равна p . Следовательно, вероятность ненаступления события A в каждом испытании также постоянна и равна $q = 1 - p$.

Поставим перед собой задачу вычислить вероятность того, что при n испытаний событий A осуществится ровно k раз и, следовательно, не осуществится $n - k$ раз. Важно подчеркнуть, что не требуется, чтобы событие A повторилось ровно k раз в определенной последовательности. Например, если речь идет о появлении события A три раза в четырех испытаниях, то возможны следующие сложные события: $AAA\bar{A}$, $AA\bar{A}A$, $A\bar{A}AA$, $\bar{A}AAA$. Запись $AAA\bar{A}$ означает, что в первом, втором и третьем испытании событие A наступило, а в четвертом испытании оно не появилось, т.е. наступило противоположное событие $\bar{A}$; соответственный смысл имеют и другие записи.

Искомую вероятность обозначим $P_n(k)$. Например, символ $P_5(3)$ означает вероятность того, что в пяти испытаниях событие появится ровно три раза и, следовательно, не наступит два раза.

Поставленную задачу можно решить с помощью формулы Бернулли, вывод которой приводится ниже.

Вероятность одного сложного события, состоящего в том, что в n испытаниях событие A наступит k раз и не наступит $n - k$ раз, по теореме умножения вероятностей независимых событий равна $p^k q^{n-k}$. Таких сложных событий может быть столько, сколько можно составить сочетаний из n элементов по k элементов, т.е. C_n^k . Так как эти сложные события не-

совместны, то по теореме сложения вероятностей несовместных событий искомая вероятность равна сумме вероятностей всех невозможных сложных событий. Поскольку же вероятности всех этих сложных событий одинаковы, то искомая вероятность (появление k раз события A в n испытаниях) равна вероятности одного сложного события, умноженной на их число:

$$P_n(k) = C_n^k p^k q^{n-k} \quad (4.1)$$

или

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad (4.2)$$

Полученную формулу называют *формулой Бернулли*.

В геологии по этой формуле рассчитывают вероятность проявления единичного значения изучаемого признака в случае, когда дискретные совокупности могут быть описаны *биномиальным распределением*. В этой формуле p – параметр распределения, n – объем совокупности, k – текущее значение признака.

Стандартное отклонение совокупности

$$\sigma_B = \sqrt{np(1-p)} \quad (4.3)$$

Биномиальное распределение обладает единственным параметром p , а графическое изображение этих вероятностей представляет собой не непрерывную кривую, а лишь систему ординат, пропорциональных значениям вероятностей дискретных значений признака.

Оборудование и материалы. Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными для выполнения работы являются совокупности нефтегазовых промысловых данных, приведённых в таблице 4.1.

Таблица 4.1 - Результаты опробования карбонатно-глинистых отложений подсвиты Морозкинской балки (нижний майкоп) на территории нескольких смежных площадей

Вариант	n	p	p-1
1	53	0,87	0,13
2	42	0,36	0,64
3	27	0,70	0,30
4	36	0,39	0,61
5	41	0,66	0,34
6	23	0,78	0,22
7	45	0,29	0,71
8	21	0,76	0,24
9	63	0,62	0,38

10	47	0,49	0,51
11	72	0,64	0,36
12	65	0,26	0,74
13	56	0,66	0,34
14	52	0,58	0,42
15	40	0,30	0,70
16	57	0,49	0,51
17	64	0,17	0,83
18	49	0,35	0,65
19	78	0,21	0,79
20	71	0,20	0,80
21	46	0,17	0,83
22	33	0,27	0,73
23	57	0,79	0,21
24	29	0,41	0,59
25	83	0,23	0,77
26	77	0,14	0,86
27	48	0,56	0,44
28	37	0,51	0,49

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы:

Задача лабораторной работы состоит в определении оптимального количества скважин, необходимых для открытия месторождения УВ используя биномиальное распределение. Для этого:

1) Рассчитываются вероятности получения промышленного притока нефти в интересующем нас стратиграфическом интервале при некоторых k значениях (4.2); *Данные заносятся в соответствующую таблицу.*

2) рассчитывается значение стандарта (4.3).

3) строится гистограмма биномиального распределения вероятности вскрытия нефтесодержащей части карбонатно-глинистых отложений геологического подразделения;

4) определяется, какому введенному в разведку количеству скважин соответствует максимальная вероятность вскрытия продуктивной части изучаемого стратиграфического подразделения.

Пример решения задачи. На территории нескольких смежных площадей (Журавская, Искринская и т. д.) Ставропольского края при опробовании карбонатно-глинистых отложений подсвиты Морозкинской балки (нижний майкоп) в 38 скважинах было получено 12 промышленных притоков нефти.

Оценивая ориентировочно вероятность вскрытия нефтесодержащей части изучаемых отложений через $p \approx 0,32$, а непродуктивной – через $(1-p) \approx 0,68$, по формуле (4.2) проведем расчет вероятности получения промышленного притока нефти в интересующем нас стратиграфическом интервале (таблица 4.2) при некоторых k значениях.

На гистограмме распределений (рисунок 4.1) видно, что максимальная вероятность вскрытия продуктивной части изучаемого стратиграфического интервала соответствует моменту ввода в разведку 10 – 12 скважин. Причем «пик» максимальной вероятности достаточно острый – значение стандарта $\sigma_B \approx 2,87$.

Таблица 4.2 – Вероятность вскрытия нефтесодержащего интервала при опробовании карбонатно-глинистых отложений подсвиты Морозкинской балки (нижний майкоп)

Число скважин, k	Вероятность, $P_n(k)$	Число скважин, k	Вероятность, $P_n(k)$
4	0,001828887	14	0,1050817
6	0,01457048	16	0,05148386
8	0,05498119	18	0,01655798
10	0,1132157	20	0,003527143
12	0,1381242	22	0,0004,976435

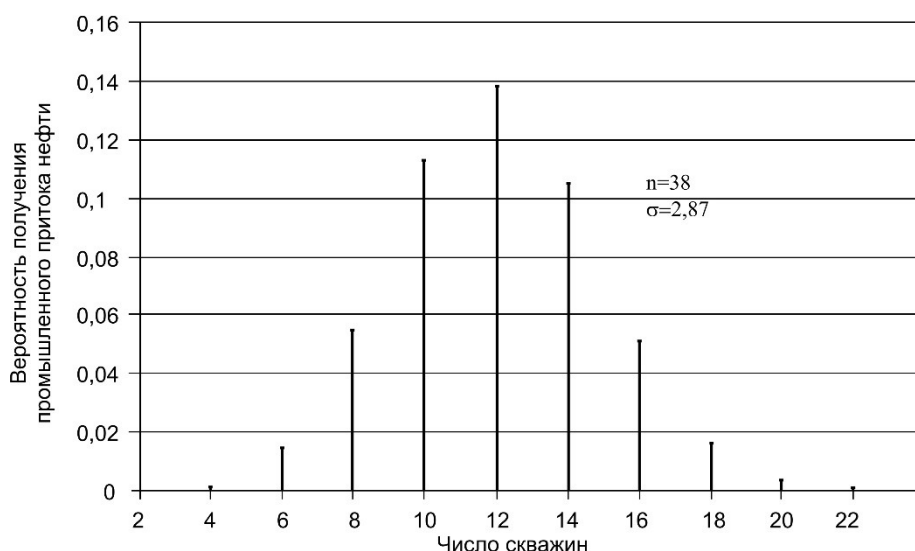


Рисунок 4.1 – Биноминальное распределение на примере изучения вероятности вскрытия нефтесодержащей части карбонатно-глинистых отложений подсвиты Морозкинской балки (нижний майкоп)

Содержание отчета. В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенной задачи. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 4.3.

Таблица 4.3 – Вероятность вскрытия нефтесодержащего интервала при карбонатно-глинистых отложениях подсвиты Морозкинской балки на территории изучаемой площади

Число скважин, k	Вероятность, $P_n(k)$	Число скважин, k	Вероятность, $P_n(k)$

Контрольные вопросы:

1. Какие совокупности являются дискретными совокупностями.? Приведите примеры дискретных геологических данных.

2. Что является единственным параметром биномиального распределения?

3. Охарактеризуйте условия применения распределения Пуассона.

4. Какие испытания называют независимыми?

5. Охарактеризуйте, какие события называются сложными, какие простыми.

Приведите примеры.

6. Что обозначает запись $AA\bar{A}A$?

7. Что обозначает запись $P_7(4)$, $P_{20}(4)$?
8. Что такое сочетание? Приведите примеры.
9. Чему равна вероятность одного сложного события?

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 5

Корреляция геологических данных

Цель: Освоение методики анализа дискретных совокупностей геологических данных, полученных в условиях малых значений вероятности события. Вышеизложенное позволяет сформировать компетенцию, предусматривающую готовность обобщать, анализировать, воспринимать информацию, ставить цели и выбирать пути ее достижения.

Теоретическая часть. При решении геологических задач используется еще одно дискретное распределение – *распределение Пуассона*. Особенностью этого распределения является возможность количественной оценки вероятности события при малых значениях p (при $p \rightarrow 0$ распределение теряет смысл).

Вероятность таких событий оценивается по асимптотической формуле Пуассона.

$$P(k\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (5.1)$$

где $P(k\lambda)$ – вероятность; λ – единственный параметр распределения, вычисляемый по формуле

$$\lambda = (\bar{h} + \sigma^2) / 2. \quad (5.2)$$

так, как

$$\lambda = \bar{h} = \sigma^2 = Np \quad (5.3)$$

Последнее обстоятельство является характерной чертой этого распределения: $\bar{h} = \sigma^2$ совокупность может быть описана формулой Пуассона.

Вычисленные по формуле Пуассона значения вероятностей могут быть использованы для расчета теоретических значений частот F_k по формуле:

$$F_k = P(k=j, \lambda) \sum_{j=0}^n f_j. \quad (5.4)$$

Оборудование и материалы.

Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными для выполнения работы являются совокупности нефтегазопромысловых данных - гистограммы частот нефтегазопроявлений из различных комплексов горных пород, приведённых в таблице 5.1.

Таблица 5.1 - Результаты опробования из различных комплексов горных пород на территории нескольких смежных площадей

Вариант	Продуктивный пласт	Частота нефтегазопроявлений, f_i
1	II	12
	V	45
	I	10
	IV	38
	VII	3
2	I	15
	II	33
	VI	40
	X	17
	IX	12
3	III	17
	V	54
	IV	26
	VII	19
	I	11
4	V	37
	III	42
	IV	30

	II	28
	IX	17
5	I	29
	VII	38
	II	45
	V	43
	VIII	46
6	X	42
	IV	34
	I	17
	III	20
	V	35
7	I	19
	II	23
	V	28
	VI	30
	IX	25
8	II	14
	III	27
	V	43
	VI	39
	X	23
9	III	30
	V	33
	VI	28
	VI I	47
	IX	19
10	I	20
	II	22
	IV	37
	V	40
	VII	18
11	I	26
	II	21
	III	37
	V	42
	VI	31
12	II	19
	V	37
	IV	42
	VII	38
	IX	20
13	I	27
	VIII	25
	V	34
	IX	46
	X	23
14	II	32
	IV	31
	VI	35

	VII	40
	IX	12
15а	III	13
	IV	40
	V	33
	VI	41
	VII	17
15б	I	15
	II	31
	III	38
	V	39
	VI	11
16	K _{1a}	54
	K ₂	70
	I ₂	67
	p-T	15
	P ₂	10

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы:

Задача лабораторной работы состоит в оценке возможности использования распределения Пуассона для определения вероятности нефтегазопроявлений из комплекса горных пород и проведении сравнения фактических и теоретических частот нефтегазопроявлений для исследуемого комплекса.

Для этого:

- 1) строится гистограмма распределения нефтегазопроявлений, полученных во время опробования отложений;
- 2) рассчитываются значения среднего и дисперсии, определяется - выполнение для изучаемой совокупности требования (5.3). Данные заносятся в соответствующую таблицу.
- 3) рассчитываются значения вероятностей $P(k\lambda)$ по формуле Пуассона (5.1);
- 4) рассчитываются теоретические значения частот F_k по формуле (5.4).
- 5) строится гистограмма распределения теоретических частот нефтегазопроявлений, проводится сопоставление фактических и теоретических частот и анализ возможности использования графиков для планирования дальнейших геолого-разведочных работ.

Пример решения задачи. Рассмотрим особенности распределения Пуассона на примере распределения нефтегазопроявлений, полученных во время опробования отложений баталпашинской свиты Ставропольского края (Александровская, Северо-Нагутская, Журавская, Искринская и др. площади). Анализ гистограммы (рисунок 5.1) позволил выделить пять

основных продуктивных пластов (III, IV, V, VI, VII), с которыми связано большинство нефтегазопроявлений, и в том числе практически промышленные притоки нефти и газа. Поэтому, если учесть, что остальные пласты и пропластки практически выполняют роль крышек-экранов, а нефтегазопроявления из них – случайные, нехарактерные в региональном плане проявления, можно составить ряд дискретных совокупностей, характеризующих нефтегазопроявления только в основных продуктивных пластах.

Вероятность нефтегазопроявлений в каждом из комплексов можно оценить по формуле Пуассона. Для этого необходимо удостовериться в том, что рассматриваемая совокупность подчиняется условию (5.3), т. е. в равенстве среднего и дисперсии. Следует отметить, что практически такое равенство при изучении совокупностей геолого-геофизических данных невыполнимо, но можно придерживаться условия $\bar{h} \approx \sigma^2$. Результаты расчета сведены в таблице (5.2).

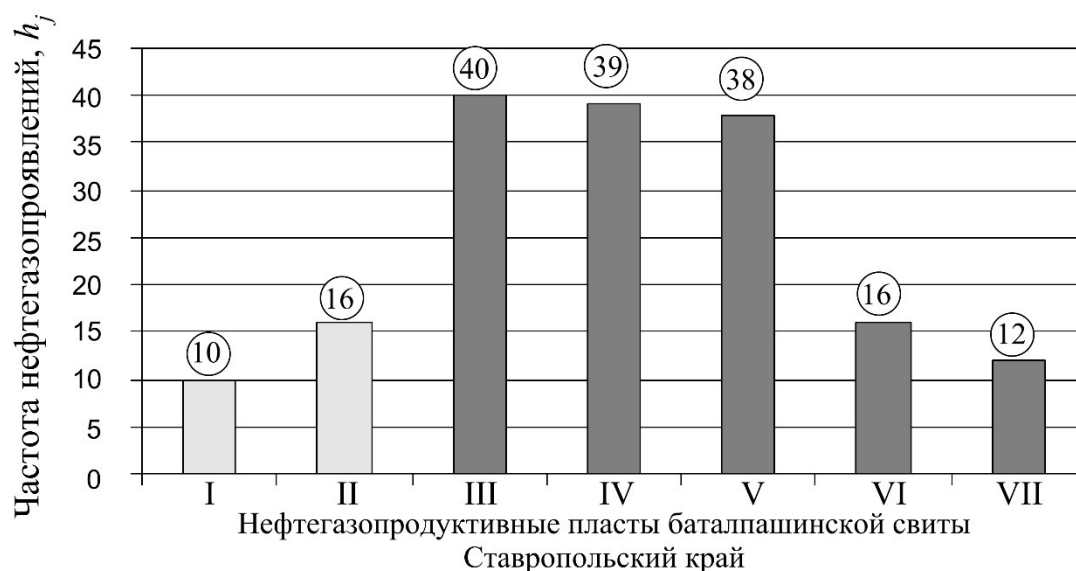


Рисунок 5.1 – Гистограмма нефтегазопроявлений при опробовании отложений баталпашинской свиты отдельных площадей Ставропольского края

Таблица 5.2 – Расчет параметров совокупности нефтегазопроявлений из отложений баталпашинской свиты (Ставропольский край)

Продуктивный пласт	Кодовый номер комплекса, h_j	Частота нефтегазопроявлений, f_j	$h_j f_j$	$h_j - \bar{h}$	$(h_j - \bar{h})^2$	$f_j (h_j - \bar{h})^2$
1	2	3	4	5	6	7
III	0	40	0	-0,7158	0,5124	20,4963
IV	1	39	39	0,2841	0,0807	3,1494

<i>V</i>	2	32	64	1,284 1	1,6490	52,7711
<i>VI</i>	3	16	48	2,284 1	5,2174	83,4791
<i>VII</i>	4	12	48	3,284 1	10,7857	129,4295
		$\sum f_j = N$	$\sum h_j f_j$			$\sum f_j (h_j - \bar{h})^2$
		139	199			289,3255

Отсюда значение средней $\bar{h} = \frac{199}{139} = 0,716$ и дисперсии $\sigma^2 = \frac{289,326}{139} = 1,041$. Будем считать их допустимо близкими, чтобы рассматриваемая совокупность могла быть описана распределением Пуассона. Вычислим значения вероятностей $P(k\lambda)$ по (5.1):

при $k=0$

$$P(0; 0,89) = 0,4107;$$

при $k=1$

$$P(1; 0,89) = 0,3655;$$

при $k=2$

$$P(2; 0,89) = 0,16264;$$

при $k=3$

$$P(3; 0,89) = 0,04825;$$

при $k=4$

$$P(4; 0,89) = 0,01074.$$

Вычисленные по формуле Пуассона значения вероятностей используются для расчета теоретических значений частот F_k по формуле (5.4):

Например, теоретическая частота нефтегазопроявлений в III пласте баталпашинской свиты ($k=0$) равна $F_0 = 0,411 \times 139 \approx 57$ при фактической частоте, равной 40.

Содержание отчета. В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенной задачи. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 5.3.

Таблица 5.3 – Расчет параметров совокупности нефтегазопроявлений из отложений баталпашинской свиты

Продуктивный пласт	Кодовый номер комплекса, h_j	Частота нефтегазопроявлений, f_j	$h_j f_j$	$h_j - \bar{h}$	$(h_j - \bar{h})^2$	$f_j (h_j - \bar{h})^2$
1	2	3	4	5	6	7
	0					
	1					
	2					
	3					

	4					
		$\sum f_j = N$	$\sum h_j f_j$			$\sum f_j (h_j - \bar{h})^2$

Контрольные вопросы:

1. Дайте характеристику условиям применения распределения Пуассона.
2. Как определяются теоретические и фактические частоты в случае использования распределения Пуассона?
3. Какие геологические задачи можно решить, используя распределение Пуассона?
4. Каким образом можно провести расчет вероятности получения промышленного притока нефти в стратиграфическом интервале при заданном преподавателем значении k ?

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 6

Многофакторная корреляция в форме общей линейной модели

Цель: Освоение методики оценки формы и тесноты парных связей между геологическими параметрами; освоение методики прогноза значений геологических параметров по уравнениям регрессии. Вышеизложенное позволяет сформировать , предусматривающей

готовность обобщать, анализировать, воспринимать информацию, ставить цели и выбирать пути ее достижения.

Формируемые компетенции. Данная работа направлена на формирование компетенции планировать и выполнять аналитические, имитационные и экспериментальные исследования, критически оценивать результаты исследований и делать выводы.

Теоретическая часть. Вычислительные модели дисперсионного анализа позволяют не только сравнивать совокупности, но и выявлять наличие или отсутствие зависимости изучаемого геологического признака от различных факторов. Изучение таких связей, количественная оценка их тесноты и моделирование их математическими функциями составляют предмет корреляционного анализа.

В зависимости от структуры изучаемой совокупности различают корреляции двухмерную (или парную), трехмерную и т. д. В случае парной изменения изучаемого признака связаны только с одним фактором. При многомерной корреляции может быть проведен многофакторный корреляционный анализ.

Большую роль в корреляционном анализе играет выбор функций, с помощью которых аппроксимируются изучаемые статистические связи. В простейшей форме часто используют уравнение прямой линии или плоскости. Для аппроксимации более сложных зависимостей применяют степенные полиномы, показательные, логарифмические, тригонометрические и другие функции. В общем виде корреляционная связь изучаемого признака с одним или несколькими факторами многомерной совокупности выражается в виде:

$$Y = M[Y] + e, \quad (6.1)$$

где Y – изучаемый признак; $M[Y]$ – его математическое ожидание; e – случайная величина, отображающая влияние неучтенных факторов.

В простейшей форме для двухмерной совокупности эта связь может быть представлена уравнением прямой линии

$$M[Y] = aX + b, \quad (6.2)$$

где $M[Y]$ – математическое ожидание изучаемого признака; X – фактор, с которым устанавливается связь; a и b – постоянные коэффициенты.

Особое место в линейной модели занимает определение значений a и b уравнения 6.2, наилучшим образом описывающих связь признаков в двухмерной совокупности. Этот метод определения значений a и b называется методом наименьших квадратов и реализуется следующим образом.

Обозначив $M[Y] = y$ и приняв $y = F(a; b) = \sum_{i=1}^N e_i^2$, выпишем условие минимума:

$$F(a; b) = \sum_{i=1}^N e_{i^2} \sum_{i=1}^N (Y_i - y_i)^2 \sum_{i=1}^N (Y_i - aX_i - b)^2 = \min, \quad (6.3)$$

из которого могут быть вычислены a и b .

Для этого найдем частные производные от сложной функции $F(a; b)$ по a и b и приравняем их нулю:

$$\left\{ \frac{\partial F}{\partial a} = \sum_{i=1}^N 2(Y_i - aX_i - b)(-X_i) = 0 \right. \quad (6.4)$$

$$\left. \frac{\partial F}{\partial b} = \sum_{i=1}^N 2(Y_i - aX_i - b)(-1) = 0 \right\} \quad (6.5)$$

Проведем алгебраические преобразования системы 6.5 – вынесем постоянный множитель 2 и « $\rightarrow$ » за знак суммы:

$$\left\{ -2 \sum_{i=1}^N (Y_i - aX_i - b)(X_i) = 0 \right. \quad (6.6)$$

Умножая правую и левую части уравнений системы 6.6 на $-\frac{1}{2}$, приведем систему к виду:

$$\left\{ \sum_{i=1}^N (Y_i - aX_i - b)(X_i) = 0 \right. \quad (6.7)$$

Продолжим далее алгебраические преобразования:

$$\left\{ \sum_{i=1}^N (Y_i X_i - aX_i^2 - bX_i) = 0 \right. \quad (6.8)$$

$$\begin{cases} \sum_{i=1}^N Y_1 X_i - a \sum_{i=1}^N X_i^2 - b \sum_{i=1}^N X_i = 0 \end{cases}, \quad (6.9)$$

$$\begin{cases} a \sum_{i=1}^N X_i^2 + b \sum_{i=1}^N X_i = \sum_{i=1}^N Y_1 X_i \end{cases}. \quad (6.10)$$

Полученную систему линейных уравнений 6.10 предлагается решить методом Крамера. Для этого заменим систему матричным равенством:

$$\begin{pmatrix} \sum_{i=1}^N X_i^2 & \sum_{i=1}^N X_i \\ \sum_{i=1}^N X_i & N \end{pmatrix} \times \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N X_i Y_i \\ \sum_{i=1}^N Y_i \end{pmatrix}, \quad (6.11)$$

Коэффициенты a и b находятся по формулам:

$$a = \frac{\Delta_a}{\Delta}, \quad b = \frac{\Delta_b}{\Delta} \quad (6.12)$$

В выражениях 6.12 Δ – определитель системы, а Δ_a и Δ_b – определители, полученные из определителя системы заменой первого и второго столбцов коэффициентов при определяемом неизвестном столбцом свободных членов.

Определитель системы вычисляют следующим образом:

$$\Delta = \begin{vmatrix} \sum_{i=1}^N X_i^2 & \sum_{i=1}^N X_i \\ \sum_{i=1}^N X_i & N \end{vmatrix} = \sum_{i=1}^N X_i^2 \times N - \sum_{i=1}^N X_i \times \sum_{i=1}^N X_i \neq 0 \quad (6.13)$$

Значения коэффициентов вычисляют, как указывалось выше в 6,12, из отношений определителей по формулам:

$$a = \frac{\begin{vmatrix} \sum_{i=1}^N X_i Y_i & \sum_{i=1}^N X_i \\ \sum_{i=1}^N Y_i & N \end{vmatrix}}{\Delta}, \quad b = \frac{\begin{vmatrix} \sum_{i=1}^N X_i^2 & \sum_{i=1}^N X_i Y_i \\ \sum_{i=1}^N X_i & \sum_{i=1}^N Y_i \end{vmatrix}}{\Delta} \quad (6.14)$$

которые для удобства запишем в форме:

$$a = \frac{N \sum_{i=1}^N X_i Y_i - \sum_{i=1}^N X_i \sum_{i=1}^N Y_i}{N \sum_{i=1}^N X_i^2 - \left[\sum_{i=1}^N X_i \right]^2}, \quad b = \frac{\sum_{i=1}^N X_i^2 \sum_{i=1}^N Y_i - \sum_{i=1}^N X_i \sum_{i=1}^N X_i Y_i}{N \sum_{i=1}^N X_i^2 - \left[\sum_{i=1}^N X_i \right]^2}. \quad (6,15)$$

На рисунке 6.1 показана геометрическая интерпретация постоянных коэффициентов a и b уравнения 6.2. Коэффициент a , равный тангенсу угла наклона аппроксимирующей линии к оси OX (рисунок 6.1), называется *коэффициентом регрессии*. Значение b соответствует длине отрезка оси Y , отсекаемого при $X=0$.

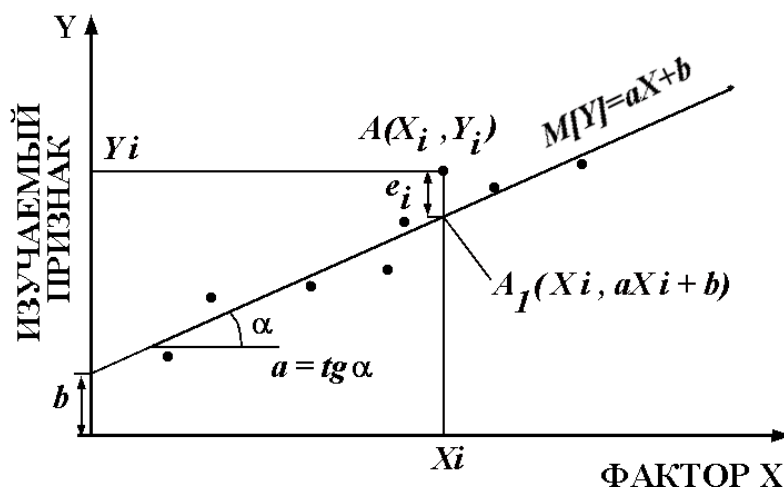


Рисунок 6.1 – Геометрическая интерпретация постоянных коэффициентов уравнения корреляционной связи

Вычисленные значения a и b используются для количественного описания изучаемой связи по формуле 6.2.

Оборудование и материалы.

Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными является данные по вскрытию несколькими разведочными и параметрическими скважинами стратиграфических горизонтов приведённые в Приложении В.

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы:

По данным бурения параметрических скважин выбираются две статистические совокупности (совокупности глубин залегания двух стратиграфических горизонтов), которые принимаются за X_i – факторы и Y_i – признаки. Далее необходимо:

- 1) рассчитать коэффициенты уравнения связи 6.2 a и b по 6.14;
- 2) построить график корреляционной связи, подставляя в уравнение 6.1 исходные величины признака.

Содержание отчета. В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 6.1.

Таблица 6.1 – Результаты расчетов при вычислении параметров уравнения регрессии

Номер точки	X_i	Y_i	$X_i Y_i$	X_i^2	Y_i^2	$Y_{рас} = aX + b$
1	2	3	4	5	6	7
	$\sum X_i$	$\sum Y_i$	$\sum X_i Y_i$	$\sum X_i^2$	$\sum Y_i^2$	

Контрольные вопросы:

1. Каким образом реализуется метод наименьших квадратов в моделировании геологических объектов и процессов?
2. Как решаются и где используются системы линейных уравнений в математическом моделировании геологических процессов?
3. Запишите условие минимума, используемое в методе наименьших квадратов, дайте геометрическую интерпретацию аргументов сложной функции $F(a; b)$.
4. Получите систему линейных уравнений для определения коэффициентов регрессии методом наименьших квадратов.
5. Охарактеризуйте основные действия над матрицами - сложение, произведение матрицы на число и на матрицу.
6. Каким образом определяются определители матрицы?

7. Дайте геометрическую интерпретацию постоянных коэффициентов уравнения корреляционной связи.

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 7

Группирование геологических объектов

Цель: Освоение методики оценки дисперсии, стандартного отклонения и коэффициента корреляции в многомерных статистических моделях; освоение методики прогноза значений геологических параметров по уравнениям регрессии.

Теоретическая часть. Вычисленные значения a и b используются для количественного описания изучаемой связи по формуле 6.1. В результате для каждого значения X_i могут быть вычислены математические ожидания (наиболее вероятные значения) признака Y . Однако, вследствие того, что уравнение 6.1 является только приближенным описанием связи, выборочные значения Y_i могут отличаться от их математических ожиданий y .

Дисперсию и стандартное отклонение разброса фактических значений Y_i относительно прямой $y = aX + b$ можно вычислить по формулам:

$$S^2 = \frac{1}{N-2} \sum_{i=1}^N (Y_i - aX_i - b)^2, \quad (7.1)$$

$$S = \sqrt{S^2}, \quad (7.2)$$

Невысокое значение дисперсии и стандартного отклонения свидетельствует о близости прямой к фактической зависимости. При больших значениях дисперсии изучаемая связь, описывается более сложным законом, чем 6.1, или изучаемый признак зависит не только от X , но и от других неучитываемых факторов.

Для количественной оценки тесноты связи в формуле 6.1 используют *коэффициент корреляции*, вычисляемый по формуле:

$$r = \frac{\sum_{i=1}^N (Y_i - \bar{Y})(X_i - \bar{X})}{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2 \sum_{i=1}^N (X_i - \bar{X})^2}}, \quad (7.3)$$

где $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$; $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$ – математические ожидания сопоставляемых параметров.

Коэффициент корреляции может изменяться от -1 до $+1$. При $r=0$ связь отсутствует, при $r < 0$ – обратная, при $r > 0$ – прямая.

Сила связи между коррелируемыми признаками может быть охарактеризована с помощью шкалы Чеддока (для $|r|$, таблица 7.1):

Таблица 7.1 - Сила связи между коррелируемыми признаками по шкале Чеддока

Показатели тесноты связи	0,1-0,3	0,3-0,5	0,5-0,7	0,7-0,9	0,9-1
Характеристика силы связи	слабая	умеренная	заметная	высокая	весьма высокая

Значение $r^2 \cdot 100$ % указывает на то, сколько процентов общей корреляции объясняется изменением факторного признака.

Уравнение регрессии используется для прогнозирования возможных ожидаемых значений результативного признака. Прогнозируемое значение результативного показателя (признака) получается подставлением в уравнение регрессии ожидаемой величины факторного признака. Например, если математическая модель корреляционной связи имеет вид $y = 6,11X - 15,77$, то при $x=12$ получаем $y = 6,11 \cdot 12 - 15,77 = 57,55$.

Оборудование и материалы. Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными является структурная схема стратиграфического горизонта вскрытого несколькими разведочными и параметрическими скважинами, выданные в лабораторной работе 6.

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы:

Рассчитываются:

- 1) дисперсия и стандартное отклонение по 7.1 и 7.2;
- 2) коэффициент корреляции по 7.3;
- 3) Далее дать качественную характеристику связи между фактором и признаком и оценить силу связи.
- 4) Установить величину влияния фактора (сколько процентов общей корреляции объясняется изменением факторного признака).

Содержание отчета. В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 7.2.

Таблица 7.2 – Результаты расчетов при вычислении параметров парной линейной корреляции

[illegible]

	ΣX_i	ΣY_i	$\Sigma X_i Y_i$	ΣX_i^2	ΣY_i^2				$\Sigma (Y_i - aX_i - b)^2$	$\Sigma (X_i - \bar{X})^2$	$\Sigma (Y_i - \bar{Y})^2$	$\Sigma (X_i - \bar{X})(Y_i - \bar{Y})$
			i									

Контрольные вопросы:

1. Что представляют собой дисперсия, стандартное отклонение? Каким образом определяются дисперсия и стандартное отклонение?
2. Что представляет собой коэффициент корреляции? Каким образом определяется коэффициент корреляции?
3. Какие параметры оценивают тесноту линейной связи между признаком и фактором?
4. Какова область значений коэффициента корреляции? Как охарактеризовать силу связи между коррелируемыми признаками?
5. Какую информацию можно получить, рассчитав величину $r^2 \cdot 100\%$?
6. Как строится график корреляционной связи? Каким образом проводится прогноз значений геологических параметров?

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. : Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб. пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. / М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 8

Тренд-анализ геологических параметров

Цель: Освоение методики доверительных интервалов коэффициентов уравнения общей линейной модели и доверительных интервалов дисперсии, освоение методики проверки гипотезы о наличии связи между признаком.

Теоретическая часть. Показатели корреляционной связи являются лишь оценками той или иной закономерности, так как в любом параметре сохраняется элемент полностью погасившейся случайности, присущей индивидуальным значениям признаков. Поэтому необходима статистическая оценка степени точности и надёжности параметров корреляции.

Приведем наиболее важные оценки интервалов.

Так, вычисленное по формуле 6.1 значение y_{oi} для $X_o = X_i$ при уровне значимости α заключено в интервале:

$$y_{oi} - t_{\alpha/2, N-2} \cdot S \sqrt{\frac{1}{N} + \frac{(X_{oi} - \bar{X})^2}{\sum_{i=1}^N (X_i - \bar{X})^2}} \leq Y \leq y_{oi} + t_{\alpha/2, N-2} \cdot S \sqrt{\frac{1}{N} + \frac{(X_{oi} - \bar{X})^2}{\sum_{i=1}^N (X_i - \bar{X})^2}}, \quad (8.1)$$

$$Y_{\min} = y_{oi} - t_{\alpha/2, N-2} \cdot S \sqrt{\frac{1}{N} + \frac{(X_{oi} - \bar{X})^2}{\sum_{i=1}^N (X_i - \bar{X})^2}} \quad (8.2)$$

$$Y_{\max} = y_{oi} + t_{\alpha/2, N-2} \cdot S \sqrt{\frac{1}{N} + \frac{(X_{oi} - \bar{X})^2}{\sum_{i=1}^N (X_i - \bar{X})^2}} \quad (8.3)$$

Доверительные интервалы для оценки коэффициента a :

$$a - t_{\alpha/2, N-2} \cdot \sqrt{\frac{S^2}{\sum_{i=1}^N (X_i - \bar{X})^2}} \leq A \leq a + t_{\alpha/2, N-2} \cdot \sqrt{\frac{S^2}{\sum_{i=1}^N (X_i - \bar{X})^2}}, \quad (8.4)$$

для оценки коэффициента b :

$$b - t_{\alpha/2, N-2} \cdot \sqrt{\frac{S^2 \sum_{i=1}^N X_i^2}{\sum_{i=1}^N (X_i - \bar{X})^2}} \leq B \leq b + t_{\alpha/2, N-2} \cdot \sqrt{\frac{S^2 \sum_{i=1}^N X_i^2}{\sum_{i=1}^N (X_i - \bar{X})^2}}, \quad (8.5)$$

Доверительный интервал для оценки дисперсии рассчитывается по

$$\frac{S^2(N-2)}{\chi_{\alpha/2, N-2}^2} \leq \sigma^2 \leq \frac{S^2(N-2)}{\chi_{1-\alpha/2, N-2}^2}, \quad (8.6)$$

Для проверки гипотезы о наличии связи между X и Y используется следующий статистический критерий

$$t = \frac{|r|}{\sqrt{1-r^2}} \sqrt{N-2} \quad (8.7)$$

Связь считается с вероятностью $1 - \alpha$ достоверной, если t превышает значение распределение Стьюдента $t_{\alpha, N-2}$.

Существуют и более простые способы оценки доверительных интервалов. Так, для оценки a может быть вычислена *дисперсия отклонений от коэффициента регрессии*:

$$S_a^2 = \frac{S_Y^2}{S_X^2} \left(\frac{1-r^2}{N-3} \right) \quad (8.8)$$

где S_Y^2 и S_X^2 – *общие средние квадратические отклонения признака X и Y соответственно*. Они определяются по формулам:

$$S_Y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2 \quad (8.9)$$

$$S_X^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2 \quad (8.10)$$

Оценка a считается надежной, если

$$a > 3S_a, \quad (8.11)$$

Для оценки r находим *дисперсию отклонений от значений коэффициента корреляции* по формуле:

$$S_r^2 = \frac{(1-r^2)^2}{N} \quad (8.12)$$

Оценка r считается надежной при

$$r > 3S_r, \quad (8.13)$$

Оборудование и материалы.

Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными является структурная схема стратиграфического горизонта вскрытого несколькими разведочными и параметрическими скважинами, выданные в лабораторной работе 6.

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы:

Рассчитать

1) граничные значения y_{oi} для X_{oi} для любого из X_i (фактические данные фактора) по 8.1;

2) *доверительные интервалы* для оценки коэффициентов a и b по 8.4 и 8.5;

3) *доверительные интервалы* для значений признака (Y_{pac}) по 8.2 и 8.3.

Далее строится график корреляционной связи с учетом максимальных и минимальных граничных интервалов (по п.3).

Затем следует убедиться в наличии связи, для этого провести статистическую оценку степени точности и надёжности параметров корреляции.

Для этого определяем:

4) *доверительный интервал для оценки дисперсии* по 8.6 (значения критерия Пирсона (χ^2) при соответствующем уровне значимости определяется в программной среде EXCEL);

5) по 8.7 рассчитываем значение критерия Стьюдента и сравниваем его с табулированным значением (по EXCEL) Делаем вывод о подтверждении или отклонении гипотезы о наличии связи между X и Y ;

6) *дисперсию отклонений от коэффициента регрессии* по 8.8, 8.9 и 8.10. Оценку надёжности a проверяют по 8.11;

10) *дисперсию отклонений от значений коэффициента корреляции* по 8.12. Оценку надёжности r проверяют по 8.13.

Содержание отчета. В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Защита работы проводится в устной форме.

Исходные данные и результаты вычислений представляются в форме таблицы 8.1.

Таблица 8.1 – Результаты расчетов при вычислении *доверительных интервалов* признака

Номер точки	X_i	Y_i	$Y_{pac} = aX + b$	Y_{min}	Y_{max}
1	2	3	4	5	6
	$\sum X_i$	$\sum Y_i$			

Контрольные вопросы:

1. Приведите расчётную формулу для определения граничных значений y_{oi} для X_{oi} для любого из X_i .
2. Каким образом определяются оценки доверительных интервалов коэффициентов уравнения общей линейной модели?
3. Каким образом определяются оценки доверительных интервалов для значений признака?
4. Дайте геометрическую интерпретацию графика корреляционной связи с учетом максимальных и минимальных граничных интервалов признака.
5. Каким образом определяются оценки доверительных интервалов дисперсии уравнения общей линейной модели?
6. Как рассчитать значение критерия Стьюдента? С чем сравнивается его значение?
7. Каким образом проверяется гипотеза о наличии связи между признаком и фактором?
8. С какой целью проводится вычисления *дисперсии отклонений от коэффициента регрессии* и *дисперсии отклонений* от значений коэффициента корреляции?
9. Как определяется *дисперсия отклонений от коэффициента регрессии*? Как оценивается оценка надежности a ?
13. Как определяется *дисперсия отклонений* от значений коэффициента корреляции? Каким образом проводится оценка надежности r ?

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. / М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

ЛАБОРАТОРНАЯ РАБОТА 9

Дисперсионный анализ

Цель: Освоение методов построения дискриминантной функции для классификации геологических объектов, оценки информативности геологических показателей; использование методических приемов для классификации и прогноза геологических характеристик.

Теоретическая часть. Одним из приложений математических методов к геологическим задачам, получившим большое распространение в последние годы, является построение объективных геологических классификаций. Деление совокупности многомерных геологических объектов на классы – в общем случае довольно сложная задача, которая во многих геологических задачах обычно решалась на уровне качественного моделирования. Естественным, что при этом разбиение на классы осуществлялось условно, по одному-двум параметрам. Качественное разбиение на классы позволяло внести в этот процесс много субъективного, связанного с точкой зрения отдельных исследователей и научных школ. В настоящее время при построении объективных геологических классификаций получили широкое распространение статистические методы. С их помощью многомерные геологические объекты, описываемые большой совокупностью показателей, делятся на классы. Одним из распространенных статистических методов классификации геологических объектов является дискриминантный анализ. Математическая модель дискриминантного анализа основана на процедуре подбора так называемой дискриминантной функции (в нашем случае линейной), которая производит оптимальное разделение объектов на классы. Дискриминантная функция подбирается на некотором эталонном массиве исходных объектов, априорно отнесенных к разным классам. Вычисленные на эталонных данных параметры дискриминантной функции используются для прогноза классификации неизвестных объектов. Задача разделения (дискриминации) сводится к подбору параметров дискриминантной функции таким образом, чтобы разным классам соответствовали разные значения функции. На схеме показано схематическое разделение объектов на классы A и B (рисунок 9.1).

Для априорно заданных объектов вычислена дискриминантная функция $D=aX + b$, позволяющая в любой точке X_0, Y_0 вычислить граничные значения D_0 . Теперь оказывается возможным отнесение неизученных объектов к классам A и B . Например, объект 1 при $D_1 > D_0$ относят к классу A , а объект 2, для которого $D_2 < D_0$ – к классу B . Естественным, что на

практике чаще встречаются задачи разделения многомерных объектов, охарактеризованных не двумя, а тремя или более признаками.

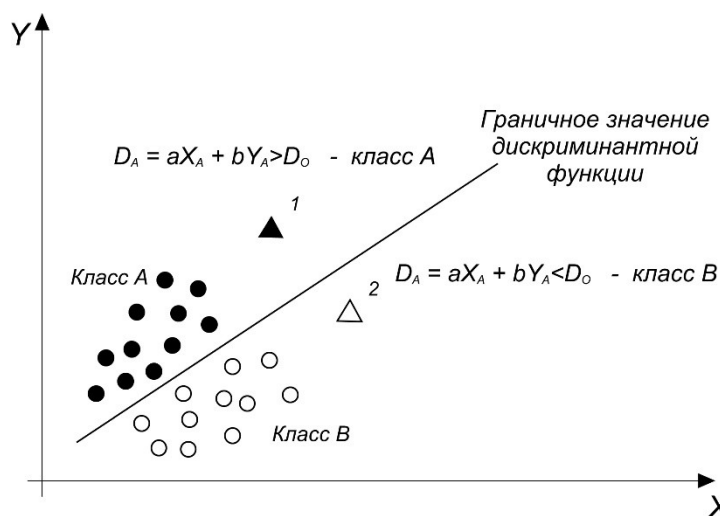


Рисунок 9.1 – Графическая интерпретация дискриминантной функции на плоскости

В наиболее общем случае геометрическая интерпретация дискриминантной функции представляет собой гиперплоскость в k -мерном признаковом пространстве, а каждый объект есть точка этого пространства. Необходимо провести в этом пространстве такую гиперплоскость, которая обеспечивала бы максимальное различие между множествами объектов, принадлежащих к разным классам, и сводила бы к минимуму рассеяние внутри каждого множества. Такую гиперплоскость называют дискриминантной.

Аналитически гиперплоскость в k -мерном пространстве имеет вид: $D = a_1X_1 + a_2X_2 + \dots + a_kX_k$.

Задача, следовательно, заключается в отыскании коэффициентов $a_1, a_2, \dots, a_k$, которые обеспечивали бы требуемые условия разделения.

Если дискриминантная функция найдена, то множество значений ее состоит из подмножеств, каждое из которых принадлежит одному из классов. Существует такое граничное значение, которое делит область значений дискриминантной функции на два подмножества. Отнесение неизученного объекта к тому или иному классу производится также как и в рассмотренном выше простейшем случае двухпризнакового объекта. Значения показателей подставляются как аргументы в построенную на эталонной выборке дискриминантную функцию. Если значение дискриминантной функции изучаемого объекта больше граничного значения D_0 , то функции изучаемого объекта относят к классу A , в противном случае – к классу B .

Представим, что исходные геологические данные позволяют выделить в многомерной совокупности два класса A и B , каждый из i объектов которых охарактеризован j значениями признаков. Представим эти исходные данные в матричной форме:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{1k} \\ A_{21} & A_{22} & A_{23} & A_{2k} \\ A_{N_1 1} & A_{N_1 2} & A_{N_1 3} & A_{N_1 k} \end{pmatrix}, \quad (9.1)$$

$$B = \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{1k} \\ B_{21} & B_{22} & B_{23} & B_{2k} \\ B_{N_2 1} & B_{N_2 2} & B_{N_2 3} & B_{N_2 k} \end{pmatrix}, \quad (9.2)$$

где N_1 – число объектов, входящий в класс A ; N_2 – число объектов, входящих в класс B ; k – число признаков, характеризующих каждый объект.

Следующий этап построения дискриминантной функции – составление *матриц центрированных сумм квадратов и смешанных произведений*:

$$S_A = \begin{pmatrix} \sum_{j=1}^k (A_{1j} - A_{1\bullet\delta})^2 & \sum_{j=1}^k (A_{1j} - A_{1\bullet\delta})(A_{2j} - A_{2\bullet\delta}) & \sum_{j=1}^k (A_{1j} - A_{1\bullet\delta})(A_{N_1 j} - A_{N_1\bullet\delta}) \\ \sum_{j=1}^k (A_{2j} - A_{2\bullet\delta})(A_{1j} - A_{1\bullet\delta}) & \sum_{j=1}^k (A_{2j} - A_{2\bullet\delta})^2 & \sum_{j=1}^k (A_{2j} - A_{2\bullet\delta})(A_{N_1 j} - A_{N_1\bullet\delta}) \\ \sum_{j=1}^k (A_{N_1 j} - A_{N_1\bullet\delta})(A_{1j} - A_{1\bullet\delta}) & \sum_{j=1}^k (A_{N_1 j} - A_{N_1\bullet\delta})(A_{2j} - A_{2\bullet\delta}) & \sum_{j=1}^k (A_{N_1 j} - A_{N_1\bullet\delta})^2 \end{pmatrix} \quad (9.3)$$

и

$$S_B = \begin{pmatrix} \sum_{j=1}^k (B_{1j} - B_{1\bullet\delta})^2 & \sum_{j=1}^k (B_{1j} - B_{1\bullet\delta})(B_{2j} - B_{2\bullet\delta}) & \sum_{j=1}^k (B_{1j} - B_{1\bullet\delta})(B_{N_2 j} - B_{N_2\bullet\delta}) \\ \sum_{j=1}^k (B_{2j} - B_{2\bullet\delta})(B_{1j} - B_{1\bullet\delta}) & \sum_{j=1}^k (B_{2j} - B_{2\bullet\delta})^2 & \sum_{j=1}^k (B_{2j} - B_{2\bullet\delta})(B_{N_2 j} - B_{N_2\bullet\delta}) \\ \sum_{j=1}^k (B_{N_2 j} - B_{N_2\bullet\delta})(B_{1j} - B_{1\bullet\delta}) & \sum_{j=1}^k (B_{N_2 j} - B_{N_2\bullet\delta})(B_{2j} - B_{2\bullet\delta}) & \sum_{j=1}^k (B_{N_2 j} - B_{N_2\bullet\delta})^2 \end{pmatrix} \quad (9.4)$$

Вычисляем *выборочную матрицу*:

$$M = \frac{1}{N_1 + N_2 - 2} (S_A + S_B). \quad (9.5)$$

Обращение этой матрицы позволяет вычислить коэффициенты дискриминантной функции:

$$C = M^{-1}. \quad (9.6)$$

Для обращения матрицы M найдем ее *детерминант* и убедимся, что:

$$D = \text{Det} M \neq 0. \quad (9.7)$$

Как известно, *обратная матрица* имеет вид:

$$M^{-1} = \begin{pmatrix} \frac{E_{11}}{\Delta} & \frac{E_{21}}{\Delta} & \frac{E_{k1}}{\Delta} \\ \frac{E_{12}}{\Delta} & \frac{E_{22}}{\Delta} & \frac{E_{k2}}{\Delta} \\ \frac{E_{1N}}{\Delta} & \frac{E_{2N}}{\Delta} & \frac{E_{kN}}{\Delta} \end{pmatrix}. \quad (9.8)$$

здесь E_{ij} – алгебраическое дополнение элементов матрицы M , определяемый по формуле:

$$E_{ik} = (-1)^{i+k} M_{ik}, \quad (9.10)$$

где M_{ik} – соответствующий алгебраическому дополнению минор.

Коэффициенты дискриминантной функции вычисляют по формуле:

$$a_p = \sum_{j=1}^k c_{pj} \left(X_j^A \ominus_{\delta} - X_j^B \ominus_{\delta} \right), \quad (9.11)$$

где c_{pj} – элементы обратной матрицы M^{-1} , X_{jcp} – средние значения соответствующего признака.

Отсюда дискриминантная функция имеет вид:

$$D = \sum_{j=1}^k a_p X_p, \quad (9.12)$$

где X_p – текущее значение p -го признака.

Оборудование и материалы. Для выполнения лабораторной работы рекомендуется использовать ЭВМ/ПК. Системные требования – Windows 2000, XP, программная среда EXCEL. Исходными данными для выполнения работы являются две многомерные совокупности геологических показателей, априорно отнесенные к двум различным классам (приведены в Приложении С).

Указания по технике безопасности смотреть в лабораторной работе 1.

Задания.

Порядок выполнения работы:

Для расчетов исходные данные следует представить в виде двух квадратных матриц.

Далее рассчитываются:

1) матрица центрированных сумм квадратов и смешанных произведений по (9.3) и (9.4);

2) выборочная матрица по (9.5);

3) обратная матрица по (9.8);

4) коэффициенты дискриминантной функции по (9.11);

5) дискриминантная функция по (9.12);

Исходные данные и результаты вычислений представляем в виде таблиц 9.1 и 9.2.

Содержание отчета. В отчете следует отразить: цель работы, краткое теоретическое обоснование и привести подробное решение предложенных задач. Защита работы проводится в устной форме.

Таблица 9.1 – Расчетная таблица для составления *матриц центрированных сумм квадратов и смешанных произведений* класса *A*

Номер обра- зца	X_{1j}^A	$X_{1j}^A - X_{1\bullet\bullet}^A$	X_{2j}^A	$X_{2j}^A - X_{2\bullet\bullet}^A$	X_{3j}^A	$X_{3j}^A - X_{3\bullet\bullet}^A$	$(X_{1j}^A - X_{1\bullet\bullet}^A)$	$(X_{2j}^A - X_{2\bullet\bullet}^A)$	$(X_{3j}^A - X_{3\bullet\bullet}^A)$	$(X_{1j}^A - X_{1\bullet\bullet}^A) \cdot (X_{2j}^A - X_{2\bullet\bullet}^A)$	$(X_{1j}^A - X_{1\bullet\bullet}^A) \cdot (X_{3j}^A - X_{3\bullet\bullet}^A)$	$(X_{2j}^A - X_{2\bullet\bullet}^A) \cdot (X_{3j}^A - X_{3\bullet\bullet}^A)$
1												
2												
3												
	$\sum_{j=1}^k X_{1j}^A$	$\sum_{j=1}^k (X_{1j}^A - X_{1\bullet\bullet}^A)$	$\sum_{j=1}^k X_{2j}^A$	$\sum_{j=1}^k (X_{2j}^A - X_{2\bullet\bullet}^A)$	$\sum_{j=1}^k X_{3j}^A$	$\sum_{j=1}^k (X_{3j}^A - X_{3\bullet\bullet}^A)$	$\sum_{j=1}^k (X_{1j}^A - X_{1\bullet\bullet}^A)$	$\sum_{j=1}^k (X_{2j}^A - X_{2\bullet\bullet}^A)$	$\sum_{j=1}^k (X_{3j}^A - X_{3\bullet\bullet}^A)$	$\sum_{j=1}^k (X_{1j}^A - X_{1\bullet\bullet}^A) \cdot (X_{2j}^A - X_{2\bullet\bullet}^A)$	$\sum_{j=1}^k (X_{1j}^A - X_{1\bullet\bullet}^A) \cdot (X_{3j}^A - X_{3\bullet\bullet}^A)$	$\sum_{j=1}^k (X_{2j}^A - X_{2\bullet\bullet}^A) \cdot (X_{3j}^A - X_{3\bullet\bullet}^A)$

Таблица 9.2 – Расчетная таблица для составления *матриц центрированных сумм квадратов и смешанных произведений* класса *B*

Номер обра- зца	X_{1j}^B	$X_{1j}^B - X_{1\bullet\bullet}^B$	X_{2j}^B	$X_{2j}^B - X_{2\bullet\bullet}^B$	X_{3j}^B	$X_{3j}^B - X_{3\bullet\bullet}^B$	$(X_{1j}^B - X_{1\bullet\bullet}^B)$	$(X_{2j}^B - X_{2\bullet\bullet}^B)$	$(X_{3j}^B - X_{3\bullet\bullet}^B)$	$(X_{1j}^B - X_{1\bullet\bullet}^B) \cdot (X_{2j}^B - X_{2\bullet\bullet}^B)$	$(X_{1j}^B - X_{1\bullet\bullet}^B) \cdot (X_{3j}^B - X_{3\bullet\bullet}^B)$	$(X_{2j}^B - X_{2\bullet\bullet}^B) \cdot (X_{3j}^B - X_{3\bullet\bullet}^B)$
1												
2												
3												
	$\sum_{j=1}^k X_{1j}^B$	$\sum_{j=1}^k (X_{1j}^B - X_{1\bullet\bullet}^B)$	$\sum_{j=1}^k X_{2j}^B$	$\sum_{j=1}^k (X_{2j}^B - X_{2\bullet\bullet}^B)$	$\sum_{j=1}^k X_{3j}^B$	$\sum_{j=1}^k (X_{3j}^B - X_{3\bullet\bullet}^B)$	$\sum_{j=1}^k (X_{1j}^B - X_{1\bullet\bullet}^B)$	$\sum_{j=1}^k (X_{2j}^B - X_{2\bullet\bullet}^B)$	$\sum_{j=1}^k (X_{3j}^B - X_{3\bullet\bullet}^B)$	$\sum_{j=1}^k (X_{1j}^B - X_{1\bullet\bullet}^B) \cdot (X_{2j}^B - X_{2\bullet\bullet}^B)$	$\sum_{j=1}^k (X_{1j}^B - X_{1\bullet\bullet}^B) \cdot (X_{3j}^B - X_{3\bullet\bullet}^B)$	$\sum_{j=1}^k (X_{2j}^B - X_{2\bullet\bullet}^B) \cdot (X_{3j}^B - X_{3\bullet\bullet}^B)$

Контрольные вопросы:

1. Что представляет собой дискриминантная функция?
2. Как рассчитывается матрица центрированных сумм квадратов и смешанных произведений?
3. Как рассчитывается *выборочная матрица*?
4. Как рассчитывается определитель матрицы?
5. Что такое минор алгебраического дополнения?
6. Как рассчитывается *обратная матрица*?
7. *Каким образом определяются коэффициенты дискриминантной функции?*
8. Напишите формулу *дискриминантной функции*.

Список литературы: 1. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

2. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

3. Математические методы и ЭВМ в поисково-разведочных работах. Учебное пособие для вузов / М. С. Арабаджи, Э. А. Бакиров и др. /. М.: Недра, 1984. 264 с.

4. Дементьев Л. Ф. Математические методы и ЭВМ в нефтегазовой геологии: Учебное пособие для вузов. М.: Недра, 1983. 189 с.

5. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

6. Харбух Дж. У. , Давтон Дж. Х., Девис Дж. К.. Применение вероятностных методов в поисково-разведочных работах. М.: Недра, 1981. 246 с.

7. Горелик. А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. М.: Высшая школа, 1989. 232 с.

8. Справочник по теории вероятности и математической статистики / В. С. Корелюк, Н. И. Портенко, А. В. Скороход и др. /М.: Наука, 1985. 640 с.

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕ-
ДЕРАЦИИ**
Федеральное государственное автономное образовательное учреждение
высшего образования
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

**ПО ОРГАНИЗАЦИИ И ПРОВЕДЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ПО ДИСЦИПЛИ-
НЕ**

Математические методы в нефтегазовой геологии

Специальность
Специализация

21.05.02 Прикладная геология
Геология месторождений нефти и газа

СОДЕРЖАНИЕ

- 1** Общая характеристика самостоятельной работы студентов при изучении дисциплины «Математические методы в нефтегазовой геологии»
 - 2** План-график выполнения самостоятельной работы
 - 3** Контрольные точки и виды отчетности по ним
 - 4** Методические рекомендации по изучению теоретического материала
 - 5** Методические рекомендации по организации работы с литературой
 - 6** Методические указания по подготовке к экзамену
 - 7.** Методические указания по выполнению курсовой работы
- Список рекомендуемой литературы

ВВЕДЕНИЕ

Современные стандарты подготовки специалистов предусматривают значительный объем внеаудиторной работы. Освоение программы курса предполагает получение, как теоретических знаний, так и формирование навыков практической работы. Поэтому самостоятельная работа в рамках курса ориентирована как на теоретический, так и на практический аспект.

Самостоятельная работа студентов – это выполнение теоретических и практических заданий студентами по усвоению изучаемой дисциплины «Математические методы в нефтегазовой геологии».

Дидактические цели самостоятельных занятий:

- формирование профессиональных умений;
- формирование умений и навыков самостоятельного умственного труда;
- мотивирование регулярной целенаправленной работы по освоению специальности;
- развитие самостоятельности мышления;
- формирование убежденности, волевых черт характера, способности к самоорганизации;
- овладение технологическим учебным инструментом.

Самостоятельная работа включает те разделы курса, которые не получили достаточного освещения на лекциях по причине ограниченности лекционного времени и большого объема изучаемого материала.

Методическое обеспечение самостоятельной работы дисциплине состоит из:

- 1) Определения учебных вопросов, которые студенты должны изучить самостоятельно;
- 2) Подбора необходимой учебной литературы, обязательной для проработки и изучения;
- 3) Поиска дополнительной научной литературы, к которой студенты могут обращаться по желанию, если у них возникает интерес в данной теме;
- 4) Определения контрольных вопросы, позволяющих студентам самостоятельно проверить качество полученных знаний;
- 5) Организации консультаций преподавателя со студентами для разъяснения вопросов, вызвавших у студентов затруднения при самостоятельном освоении учебного материала.

Самостоятельная работа студента – это особым образом организованная деятельность, включающая в свою структуру такие компоненты, как:

- уяснение цели и поставленной учебной задачи;
- четкое и системное планирование самостоятельной работы;
- поиск необходимой учебной и научной информации;
- освоение собственной информации и ее логическая переработка;
- использование методов исследовательской, научно-исследовательской работы для решения поставленных задач;
- выработка собственной позиции по поводу полученной задачи;
- представление, обоснование и защита полученного решения;
- проведение самоанализа и самоконтроля

1. Общая характеристика самостоятельной работы студентов при изучении дисциплины «Математические методы в нефтегазовой геологии»

Цель данной дисциплины:

- систематическое изучение методов моделирования реальных процессов и объектов нефтегазовой геологии;
- обучение студентов математическому аппарату, используемому в расчетах при создании количественных геологических моделей.

Основными видами самостоятельной учебной деятельности студентов являются:

- 1) предварительная подготовка к аудиторным занятиям, в том числе и к тем, на которых будет изучаться новый, незнакомый материал.
- 2) своевременная доработка конспектов лекций;
- 3) подбор, изучение, анализ и при необходимости – конспектирование рекомендованных источников по учебным дисциплинам;
- 4) выяснение наиболее сложных, непонятных вопросов и их уточнение во время консультаций;
- 5) подготовка к контрольным занятиям и экзамену;
- 6) систематическое изучение периодической печати, научных монографий, поиск и анализ дополнительной информации по учебным дисциплинам.

На самостоятельную работу по дисциплине «**Математические методы в нефтегазовой геологии**» специальности 21.05.02 Прикладная геология, специализация отводится 72 академических часа.

Виды самостоятельной работы по дисциплине «**Математические методы в нефтегазовой геологии**» могут быть разделены на основные (аудиторные) и дополнительные (внеаудиторные).

Основные виды самостоятельной работы выполняются в обязательном порядке с последующим контролем результатов преподавателем, который проводит практические занятия в студенческой группе.

Дополнительные (внеаудиторные) виды самостоятельной работы выполняются по заданию преподавателя без его непосредственного участия.

Дополнительные виды самостоятельной работы по дисциплине «**Математические методы в нефтегазовой геологии**» рекомендуются тем студентам, которые наиболее заинтересованы в изучении этой дисциплины.

К основным (обязательным) видам самостоятельной работы студентов при изучении дисциплины «**Математические методы в нефтегазовой геологии**» относится:

- а) самостоятельное изучение теоретического материала;
- б) подготовка к лабораторным работам;
- в) подготовка к практическим занятиям работам;
- г) подготовка к дифференцированному зачету.

2. План-график самостоятельной работы

Виды и содержание самостоятельной работы; формы их контроля

Виды и содержание самостоятельной работы, формы их контроля					
Вид деятельности студентов	Итоговый продукт самостоятельной работы	Средства и технологии оценки	Объем часов, в том числе		
			СРС	Контактная работа с преподавателем	Всего
5 семестр					
Подготовка к лабораторным работам	Отчет по лабораторным работам	Устный опрос	8	1	9
Подготовка к практическим занятиям	Отчет по практическим занятиям	Устный опрос	-	-	-

	занятиям				
Самостоятельное изучение литературы по темам № 1-9	Конспект	Устный опрос	8	1	9
Итого за 5 семестр			16	2	18
Итого			16	2	18

3 Контрольные точки и виды отчетности по ним

№ п/п	Вид деятельности студентов	Сроки выполнения	Количество баллов
5 семестр			
1.	Лабораторная работа № 3 Построение гистограмм распределений и кривых функций плотности вероятности распределения геолого-геофизических данных.	5-6 неделя	15
2.	Лабораторная работа 7. Оценка дисперсии, стандартного отклонения и коэффициента корреляции в многомерных статистических моделях	12-13 неделя	15
3.	Лабораторная работа 10. Определение граничного значения дискриминантной функции при прогнозе нефтегазоносности методом распознавания образов.	15-16 неделя	20
Итого за 5 семестр			55
Итого			55

4 Методические рекомендации по изучению теоретического материала

Самостоятельное изучение теоретического материала по дисциплине «**Математические методы в нефтегазовой геологии**» предусмотрено на всём протяжении курса. Такая работа сопровождает лекционные и практические занятия, промежуточный контроль, и в то же время является отдельным видом самостоятельной работы студента.

Источниками для самостоятельного изучения теоретического курса по дисциплине «**Математические методы в нефтегазовой геологии**» выступают:

- учебники по предмету;
- курсы лекций по предмету;
- учебные пособия по отдельным темам;
- научные статьи в периодической геологической печати и рекомендованных сборниках;
- научные монографии.

В ходе лекций раскрываются основные вопросы в рамках рассматриваемых тем, делаются акценты на наиболее сложные и интересные положения изучаемого материала, которые должны быть приняты студентами во внимание. Материалы лекций являются основой для подготовки студентов к практическим занятиям.

Работа с конспектом лекций

Просмотрите конспект сразу после занятий. Отметьте материал конспекта лекций, который вызывает затруднения для понимания. Попытайтесь найти ответы на затруднительные вопросы, используя предлагаемую литературу. Если самостоятельно не удалось разобраться в материале, сформулируйте вопросы и обратитесь на текущей консультации или на ближайшей лекции за помощью к преподавателю.

Каждую неделю отводите время для повторения пройденного материала, проверяя свои знания, умения и навыки по контрольным вопросам и тестам. Умение студентов быстро и правильно подобрать литературу, необходимую для выполнения учебных заданий и научной работы, является залогом успешного обучения. Самостоятельный подбор литературы осуществляется при подготовке к семинарским, лабораторным занятиям, при написании курсовых работ, научных рефератов.

Выполнение СРС на занятиях с проверкой результатов преподавателем позволяет студентам усваивать изучаемый материал более глубоко, меняется отношение к лекциям, так как без понимания теории предмета, без хорошего конспекта трудно рассчитывать на успех в решении заданий. Это улучшает посещаемость как практических, так и лекционных занятий.

При подготовке к лекциям-беседам тема сообщается студентам заранее. Участники беседы делятся на группы по 8-10 человек, в течение 10—15 мин подготавливают свой вариант решения проблемы. Следующим этапом дискуссии является выступление представителей от каждой группы и защиты своего варианта решения. В конце общего обсуждения ведущий подводит итоги беседы, дает оценку совместно найденного варианта решения.

5. Методические рекомендации по организации работы с литературой

Изучение и анализ литературных источников является обязательным видом самостоятельной работы студентов. Изучение литературы по избранной теме имеет своей задачей проследить характер постановки и решения определенной проблемы различными авторами, аргументацию их выводов и обобщений, провести анализ и систематизировать полученный материал на основе собственного осмысления с целью выяснения современного состояния вопроса.

Проработка отобранного материала обязательно должна идти с одновременным ведением записей прочитанного и своих замечаний. Запись может иметь как форму конспекта, так и выписок, а также картотеку положений, тезисов, идей, методик, что в дальнейшем облегчит классификацию и систематизацию полученного материала. Такого рода записи являются лучшим способом накопления и первичной обработки материала, одной из обязательных форм организации умственного труда.

Планом удобно пользоваться при подготовке к устному выступлению по выбранной теме. Каждый пункт плана должен раскрывать одну из сторон избранной темы, а весь план должен охватывать ее целиком.

Тезисы предполагают сжатое изложение основных положений текста в форме утверждения или отрицания. Они являются более совершенной формой записей и представляют основу для дискуссии. К тому же их легко запомнить.

Аннотация – краткое изложение содержания – дает общее представление о работе.

Резюме кратко характеризует выводы, главные итоги источника.

Конспект является наиболее распространенной формой ведения записей. Основную ткань конспекта составляют тезисы, дополненные доказательствами и рассуждениями. Конспект может быть текстуальным, свободным или тематическим. Текстуальный представляет собой цитатник с сохранением логики работы и структуры текста. Свободный конспект основан на изложении материала в том порядке, который более удобен автору. В этом смысле конспект представляет собирание воедино мыслей, разбросанных по всей книге. Тематический конспект может быть составлен по нескольким источникам, где за основу берется тема, интерпретируемая по-разному.

Экономии времени дает использование при записях различного рода сокращений, аббревиатуры и т.д. многие используют для регистрации исследуемых тем систему карточек. Преимущество карточек в том, что тема там излагается очень сжато, и они очень удобны в использовании, т.к. их можно разложить на столе, перегруппировать и без труда найти искомую тему.

6. Методические указания по подготовке к экзамену

Перед экзаменом студенту необходимо полностью завершить и защитить практические занятия. При наличии задолженностей по текущей аттестации по данной дисциплине студент к экзамену не допускается. Экзамен по дисциплине предусмотрен в устной форме по билетам.

Вопросы для проверки уровня обученности. Курсивом выделены вопросы повышенного уровня

Знать

1. Моделирование в геологии. Специфика геологических образований и процессов как объект изучения.

2. Виды моделей в геологии.

3. Качественные и количественные геологические модели. Теоретическое и практическое значение применения математических моделей в геологии в современный период.

4. Материальные системы как объект изучения. Типы геолого-математических моделей по принципу построения, по характеру связи между параметрами и свойствами изучаемых объектов, по типу решаемых задач.

5. Основные понятия теории вероятностей и математической статистики, используемые в геологических статистических моделях.

6. Виды и параметры совокупностей. Основные числовые характеристики совокупностей геолого-геофизических данных. Математическое ожидание. Дисперсия. Среднеквадратическое отклонение.

7. Степенные средние, мода, медиана как характеристики совокупностей геолого-геофизических данных.

8. Понятие о выборке и генеральной совокупности геолого-геофизических данных. Понятие точечной оценки.

9. Требования массовости, однородности, случайности и независимости к выборочной совокупности

10. Случайные события и случайные величины. Вероятность и частость.

11. Условие состоятельности, несмещенности и максимальной эффективности точечных оценок числовых характеристик случайной величины

12. Статистические распределения. Теоретические функции распределения (интегральная, дифференциальная) и формы их изображения.

13. *Геологические факторы, определяющие вид распределения.*

14. Использование основных теоретических законов распределения случайной величины при аппроксимации эмпирических распределений свойств геологических объектов

15. Нормальное распределение. Функция плотности нормального распределения.

16. Понятие моментов нормального распределения.

17. Логнормальное распределение геолого-геофизических данных. Основные параметры логнормального распределения.

18. Распределение дискретных совокупностей. Биноминальное распределение. Параметры биномиального распределения.

19. Распределения дискретных совокупностей. Распределение Пуассона и его применение в геологическом моделировании.

20. Распределение точечных оценок. Распределение доверительных интервалов выборочных оценок средней нормального и логнормального распределения (распределение Стьюдента).

21. Распределение точечных оценок. Распределение доверительных интервалов выборочных оценок дисперсии нормального и логнормального распределения (распределение χ^2).

22. Проверка статистических гипотез о законе распределения.

23. Проверка гипотез о близости совокупностей геолого-геофизических данных.

24. Общая линейная модель. Линейная модель в трехмерном пространстве.

25. Геометрическая интерпретация корреляционного анализа в виде общей линейной

модели.

26. Корреляция геологических данных. Простая линейная модель. Основные расчетные формулы для определения коэффициентов уравнения регрессии. Дисперсия, стандартное отклонение, коэффициент корреляции.

27. Понятие производной и частной производной.

28. Применение метода наименьших квадратов в моделировании геологических объектов и процессов.

29. Решение систем линейных уравнений. Применение систем линейных уравнений в математическом моделировании геологических процессов.

30. Основные действия над матрицами. Определители матрицы. Сложение, произведение матрицы на число и на матрицу. Транспонирование матрицы. Нахождение обратной матрицы.

31. *Полиномиальная модель в корреляционном анализе. Полиномиальная модель II-ой степени. Основные расчетные формулы. Дисперсия и стандартное отклонение. Корреляционное отношение.*

32. *Определение коэффициентов уравнения полиномиальной регрессии II-го порядка методом наименьших квадратов при соблюдении условия минимума квадрата относительной погрешности.*

33. *Использование общей линейной модели для корректировки исходной качественной модели и оценки относительной геологической информативности отдельных факторов.*

34. *Многофакторная корреляция в форме общей линейной модели. Основные расчетные формулы. Определение коэффициентов уравнения. Коэффициенты корреляции в случае многофакторной модели.*

35. *Использование методов математической корреляции в структурных и палеотектонических построениях.*

36. Понятие о тренде. Выявление линейных тенденций изменения геологического параметра.

37. *Тренд-анализ II-го порядка. Его применение в структурно-тектоническом районировании. Определение коэффициентов уравнения тренд-анализа II-го порядка методом наименьших квадратов при соблюдении условия минимума квадрата относительной погрешности.*

38. *Тренд-анализ III-го порядка. Его применение в структурно-тектоническом районировании. Определение коэффициентов уравнения тренд-анализа III-го порядка методом наименьших квадратов при соблюдении условия минимума квадрата относительной погрешности.*

39. Роль значений величин математического ожидания, дисперсии, среднеквадратического стандартного отклонения в тренд-анализе.

40. Оценка степени приближения поверхности тренда к наблюдаемым результатам с помощью дисперсионного анализа.

41. Разделение изучаемых компонент при тренд-анализе геологических данных на региональную и локальную компоненты.

42. *Применение тренд анализа для решения различных геологических задач. Тренд-анализ в виде полиномиальной модели.*

43. Аппроксимация геологических данных функциями плоскостей в пространстве как основа тренд-анализа.

44. *Использование тренд-анализа для решения палеогеографических задач.*

45. Основные признаки и факторы геологических объектов и применение метода распознавания образов в геологии нефти и газа.

46. Математический аппарат дискриминантного анализа. Построение дискриминантной функции.

47. Понятие о кластерном анализе

48. Сглаживание исходных данных скользящими статистическими окнами (горно-

геометрический метод П.К. Соболева)

49. Методы сглаживания временных рядов, используемые для выявления цикличности осадконакопления.

50. Дисперсионный анализ. Сопоставление геологических совокупностей.

51. Понятие о гармоническом анализе.

52. Компьютерный анализ геоинформации.

Владеть и уметь

1. Вероятностно-статистический анализ с применением ранжирования исходных данных. Необходимость ранжирования. Математический аппарат ранжирования исходных геолого-геофизических данных.

2. Ранжирование элементов совокупностей геолого-геофизических данных. Цель ранжирования. Гистограммы распределений геолого-геофизических данных.

3. Использование эмпирических графиков функции плотности распределения (гистограмм, полигонов) и функций распределения (кумулят) непрерывных случайных величин в геологии.

4. Использование эмпирических графиков функции плотности распределения (гистограмм) и функций распределения (кумулят) дискретных случайных величин в геологии.

5. Построение кривой функции плотности нормального распределения (кривая Гаусса).

6. Асимметрия и эксцесс, связь их значений с формой кривой функции нормального распределения.

7. Построение кривой функции плотности логнормального распределения.

8. Применение линейной модели для изучения геологических процессов.

9. Применение метода наименьших квадратов при определении коэффициентов уравнения регрессии.

10. Применение дифференциального исчисления для определения коэффициентов уравнений связей в корреляционном анализе геолого-геофизических данных.

11. Оценки доверительных интервалов коэффициентов уравнения общей линейной модели и доверительных интервалов дисперсии.

12. Проверка гипотезы о наличии связи между признаком и фактором.

13. Тренд-анализ геологических данных в форме общей линейной модели (тренд I-го порядка). Определение коэффициентов уравнения поверхности тренда методом наименьших квадратов при соблюдении условия минимума относительной погрешности.

14. Разделение изучаемых компонент при тренд-анализе геологических данных на региональную и локальную компоненты.

15. Группирование геологических объектов. Математическое моделирование и прогноз методом распознавания образов (дискриминантный анализ). Геометрическая интерпретация метода распознавания образов.

16. Математический аппарат дискриминантного анализа. Построение дискриминантной функции.

17. Действия над матрицами при построении дискриминантной функции (метод распознавания образов).

18. Определение коэффициентов и граничного значения дискриминантной функции.

1. Критерии оценивания компетенций

Оценка «отлично» выставляется студенту, если теоретическое содержание курса освоено полностью, без пробелов, необходимые практические компетенции сформированы, исчерпывающе, последовательно, четко и логически стройно излагает материал курса, умеет увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе дополнительный материал, все предусмотренные программой обучения

учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному.

Оценка «хорошо» выставляется студенту, если теоретическое содержание курса освоено полностью, необходимые практические компетенции в основном сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения достаточно высокое. Студент твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.

Оценка «удовлетворительно» выставляется студенту, если теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические компетенции в основном сформированы, большинство предусмотренных программой обучения учебных задач выполнено, но в них имеются ошибки, при ответе на поставленный вопрос студент допускает неточности, недостаточно правильные формулировки, наблюдаются нарушения логической последовательности в изложении программного материала.

Оценка «неудовлетворительно» выставляется студенту, если он не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы, необходимые практические компетенции не сформированы, большинство предусмотренных программой обучения учебных заданий не выполнено, качество их выполнения оценено числом баллов, близким к минимальному.

2. Описание шкалы оценивания

Промежуточная аттестация в форме экзамена предусматривает проведение обязательной экзаменационной процедуры и оценивается 40 баллами из 100. Минимальное количество баллов, необходимое для допуска к экзамену, составляет 33 балла. Положительный ответ студента на экзамене оценивается рейтинговыми баллами в диапазоне от **20** до **40** ($20 \leq S_{\text{экз}} \leq 40$), оценка **меньше 20** баллов считается неудовлетворительной.

Шкала соответствия рейтингового балла экзамена 5-балльной системе

Рейтинговый балл по дисциплине	Оценка по 5-балльной системе
35 – 40	Отлично
28 – 34	Хорошо
20 – 27	Удовлетворительно

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и опыта деятельности, характеризующих этапы формирования компетенций

Процедура проведения **экзамена** осуществляется в соответствии с Положением о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по образовательным программам высшего образования в СКФУ.

В экзаменационный билет включаются три вопроса, характеризующие этапы формирования компетенций. Как минимум один из вопросов билета должен оценивать знания, умения или навыки повышенного уровня.

Для подготовки по билету отводится 20 минут.

Список рекомендуемой литературы Основная литература

1. Брагин, Юрий Иванович. Нефтегазопромысловая геология [[Текст] :] : статическое геологическое моделирование залежей углеводородов / Ю. И. Брагин, Г. П. Кузнецова. - Москва : Недра, 2013. - 109 с. : ил., табл., цв. ил. ; 24 см. - ISBN 978-5-8365-0407-6.

2. Виноградов, Юрий Борисович. Математическое моделирование в гидрологии [[Текст] :] : учебное пособие для студентов высших учебных заведений / Ю. Б. Виноградов, Т. А. Виноградова. - Москва : Академия, 2010. - 297, [1] с. : ил., табл. ; 22 см. - (Высшее профессиональное образование. Естественные науки). Библиогр.: с. 292-294. - ISBN 978-5-7695-6785-8.

3. Самарский, А. А. Математическое моделирование : Идеи. Методы. Решения. / А.А. Самарский, А.П. Михайлов. - 2-е изд., испр. - М. :Физматлит, 2005.

Дополнительная литература

1. Арабаджи, М. С. Математические методы и ЭВМ в поисково-разведочных работах . Учебное пособие для ВУЗов / М. С. Арабаджи, Э. А. Бакиров и др. - М. : Недра, 1984. 264 с.

2. Дементьев, Л.Ф. Математические методы и ЭВМ в нефтегазовой геологии / Л.Ф. Дементьев: учебное пособие для вузов. – М: Недра, 1983. 189 с.

3. Радионов, Д. А. Справочник по математическим методам в геологии. / Д. А. Радионов, Р. И. Коган, В. А. Голубева и др. - М. : Недра, 1987. 335 с

4. Харбух, Дж. У. Применение вероятностных методов в поисково-разведочных работах. / Дж. У. Харбух, Дж. Х., Девис, Дж. К. Давтон - М. : Недра, 1981. 246 с.

5. Дж.С. Дэвис Статистический анализ данных в геологии, Кн. 1, 2.//Пер. с англ. В.А.Голубевой.-М.:Недра,1990. - 319с., 427с.

6. Боровко Н.Н. Статистический анализ пространственных геологических закономерностей. – Л.: Недра, 1971. – 174 с.

7. Гутман И.С., Балабан И.Ю. Геоestatистика в промыслово-геологических исследованиях Учебное пособие.— М.: Издательский центр РГУ нефти и газа имени И.М. Губкина, 2011.—154 с.

8. Математические методы моделирования в геологии. Часть I: Учебное пособие Автор/создатель: Мартыанова А.Е. Год: 2008

9. Гуськов, О. И. Математические методы в геологии : сборник задач : учеб.пособие для вузов / О. И. Гуськов, П. И. Кушнарев, С. М. Таранов. - М. : Недра, 1991. - 205 с. : ил. - Библиогр. с. 166. - Прил.: с. 167-203. - ISBN 5-247-01017-5

Методическая литература:

1. Методические указания к лабораторным работам по дисциплине "Математические методы в геологии" для студентов специальности 130101 "Прикладная геология" / сост. А. В. Лялин, Н.В. Еремина, М.В. Нелепов, Т.В. Гнедковская, Ставрополь :СКФУ, 2014.