

ГЕОЛОГИЧЕСКИЕ ПРИЛОЖЕНИЯ ОДНОМЕРНОЙ СТАТИСТИЧЕСКОЙ МОДЕЛИ

1. Точечная оценка погрешности среднего значения

Среднее значение \bar{x} из n независимых значений случайной величины x также является случайной величиной. Дисперсия x σ^2 , следовательно дисперсия \bar{x} δ^2 в n раз меньше:

$$\delta^2 = \sigma^2 / n \quad \text{или} \quad \delta = \sigma / \sqrt{n}$$

δ - абсолютная среднеквадратичная случайная погрешность среднего значения \bar{x} .

$$\tau = \frac{\delta}{\bar{x}} = \frac{\sigma}{\bar{x} * \sqrt{n}} = \frac{V}{\sqrt{n}} \quad \text{относительная погрешность}$$

V — коэффициент вариации. τ может быть выражена в долях единицы или в %.

Формулы δ и τ показывают, что погрешность среднего значения прямо пропорциональна изменчивости случайной величины и обратно пропорциональна корню квадратному из числа измерений.

Это позволяет решать 2 задачи:

- 1) оценивать абсолютную или относительную погрешность при известном числе наблюдений n ;
- 2) находить необходимое число измерений n для достижения заданной погрешности среднего значения.

Содержания меди (Сi , %)

Среднее	0.8742
<u>Стандартная ошибка</u>	<u>0.028882719</u>
Медиана	0.865
Мода	0.96
Стандартное отклонение	0.204231661
Дисперсия выборки	0.041710571
Эксцесс	-0.02813976
Асимметричность	0.38519861
Интервал	0.93
Минимум	0.47
Максимум	1.4
Сумма	43.71
Счет	50
Уровень надежности(95.0%)	0.058041995

Среднеквадратические погрешности:

среднего $\sigma_C = S / \sqrt{N} =$

0.02888

стандарта $\sigma_S = S / \sqrt{2 \cdot N} =$

0.02042

асимметрии $\sigma_A = \sqrt{6 / N} =$

0.346410162

эксцесса $\sigma_E = \sqrt{24 / N} = 2 \cdot \sigma_A =$

0.692820323

Пример В результате анализа 16 проб гранита рассчитано среднее содержание кремнезема $\bar{x} = 70,35 \%$ и среднеквадратичное отклонение

$\sigma = 3,20 \%$. Определить, чему равна среднеквадратичная погрешность среднего содержания и сколько дополнительно нужно взять проб, чтобы снизить относительную погрешность до 1 %.

Абсолютная среднеквадратичная случайная погрешность

$$\delta = 3,2 / \sqrt{16} = 0,80\% ; \text{ относительная случайная погрешность}$$

$$\tau = 0,80/70,35 = 1,14\%.$$

Если $\tau = 1\% = 0,01$, то из формулы $\tau = \frac{\delta}{\bar{x}} = \frac{V}{\sqrt{n}}$ получим

$$\delta = \tau * \bar{x} = 0,01 * 70,35 = 0,70 \quad \text{Из формулы } \tau = \frac{\delta}{\bar{x}} = \frac{V}{\sqrt{n}}$$

имеем
$$n = \sigma^2 / \delta^2 = 3,20^2 / 0,70^2 = 21$$

дополнительно нужно взять и проанализировать $21 - 16 = 5$ проб.

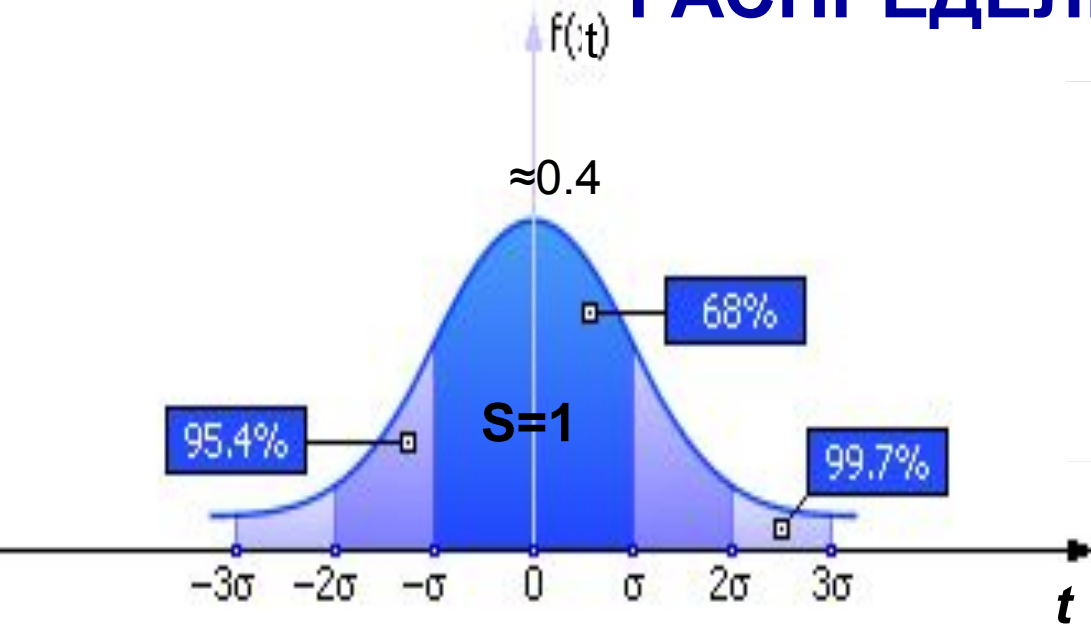
2. Интервальная оценка математического ожидания случайной величины

$M(x)$ в генеральной совокупности обычно неизвестно. Его можно приближенно оценить с помощью выборочного среднего значения \bar{x} , которое является случайной величиной и имеет дисперсию δ^2 . Предполагается, что случайная величина имеет распределение, близкое к нормальному. Размах значений нормально распределенной величины составляет приблизительно $\pm 3\delta$. Где-то в этом интервале и заключено математическое ожидание $M(x)$. Наиболее вероятно, что оно совпадает со средним значением \bar{x} , которое является *точечной оценкой* математического ожидания. Менее вероятно, что $M(x)$ смещено в ту или иную сторону от среднего значения. Интервал возможных значений $M(x)$ зависит от вероятности $q = \Phi(t)$ и выражается через коэффициент вероятности t соотношением

$$\bar{x} - t\delta < M(x) < \bar{x} + t\delta$$

доверительный интервал или *интервальная оценка* математического ожидания

СТАНДАРТНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ



Из симметричности следует:

$$F(-t) = 1 - F(t)$$

Функция $F(t)$ для $t \geq 0$ нормированная функция Лапласа. Обозначается $\Phi_0(t)$ и имеет вид

$$\Phi_0(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt$$

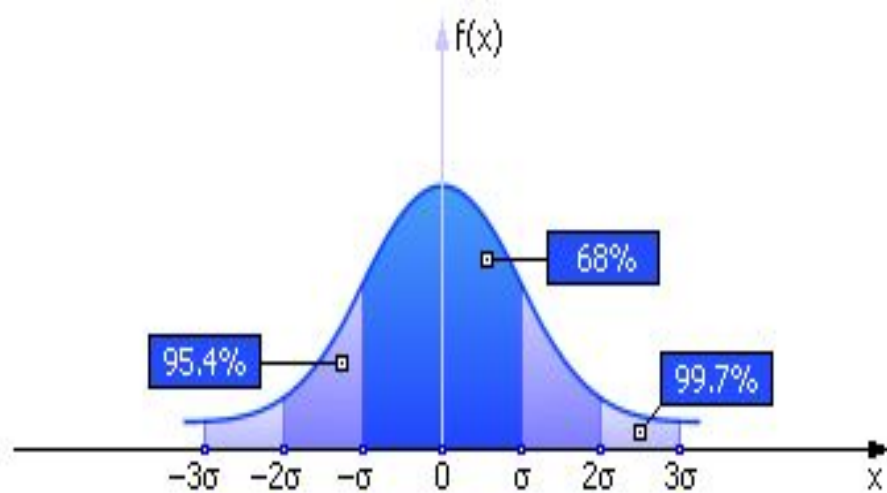
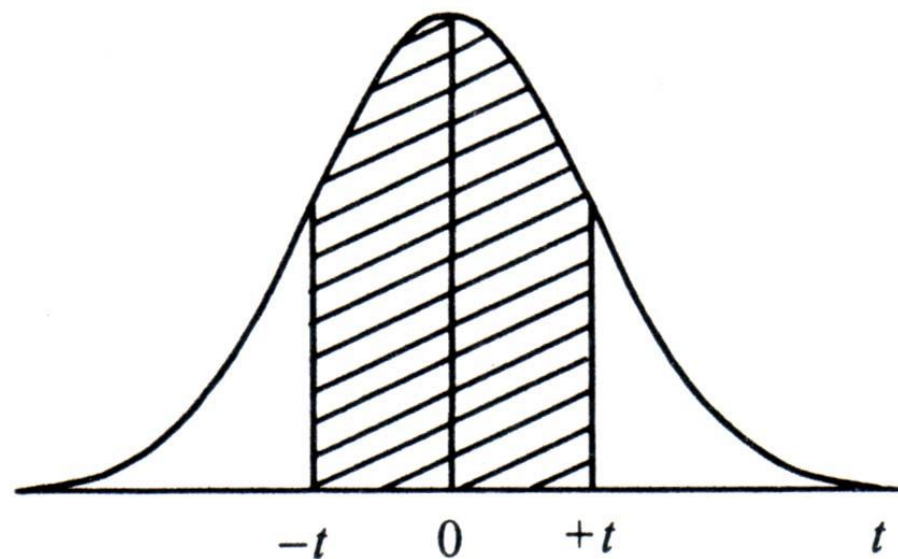
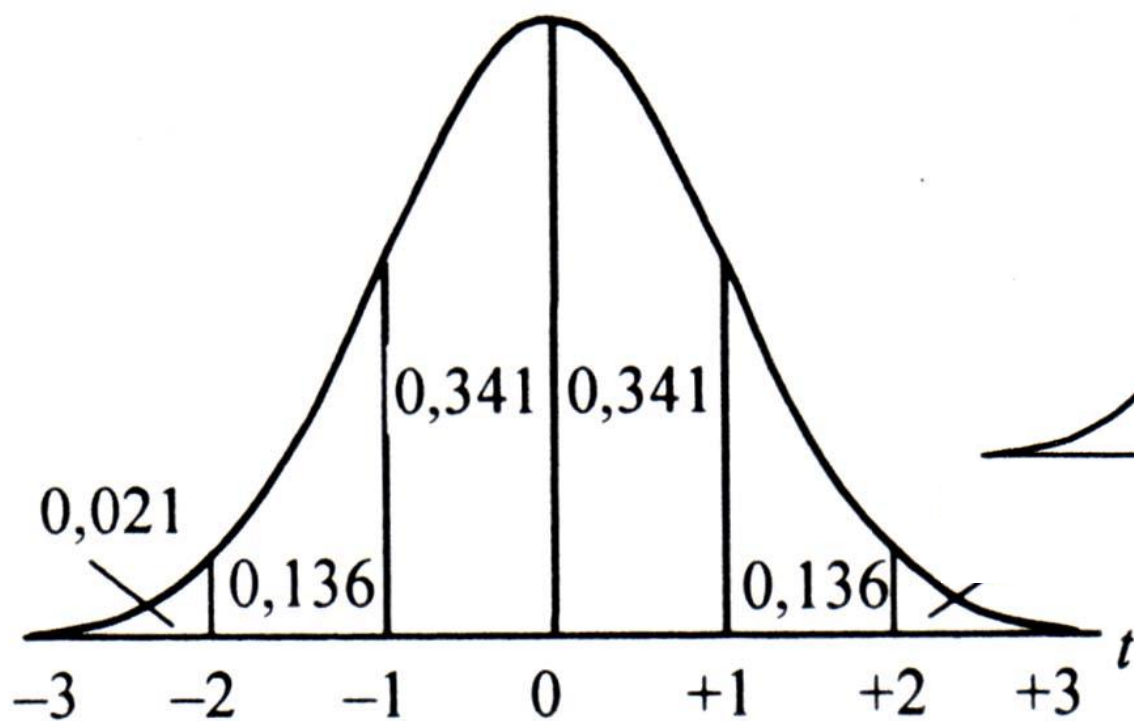
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-Mx)^2}{2\sigma^2}} dx;$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-Mx)^2}{2\sigma^2}}$$

$$t = \frac{x - Mx}{\sigma}$$

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{t^2}{2}} dt$$

$$\varphi(t) = Z(t) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}$$



Каждому значению вероятности q соответствует определенный коэффициент вероятности t и размер доверительного интервала:

Вероятность $q = \Phi(t)$	Коэффициент Вероятности t	Доверительный интервал
0.683	1	$\bar{x} - \delta < M(x) < \bar{x} + \delta$
0.954	2	$\bar{x} - 2\delta < M(x) < \bar{x} + 2\delta$
0.997	3	$\bar{x} - 3\delta < M(x) < \bar{x} + 3\delta$

Используя данные примера, где среднее содержание кремнезема в граните $\bar{x} = 70,35 \%$, и $\delta = 0,80 \%$, получаем доверительные интервалы:

Вероятность q	Доверительный интервал	Если \bar{x} или другая оцениваемая величина подчиняются не нормальному закону распределения, то вероятность q будет иная.
0.683	$69,55 < M(x) < 71,15$	
0.954	$68,75 < M(x) < 71,95$	
0.997	$67,95 < M(x) < 72,75$	

3. Выделение аномальных значений

Когда в однородную совокупность попадают единичные значения, значительно отличающиеся от среднего (аномальные или ураганные) происходит искажение статистических характеристик. Поэтому актуальной является задача

- о разделении неоднородной совокупности на однородные,
- о выделении из неоднородных совокупностей аномальных значений.

Наиболее распространенный способ выделения аномальных значений (если известен или задан закон распределения случайной величины) - правило «*трех сигм*» основан на том, что случайная величина при нормальном законе распределения практически полностью (на 99,7 %) заключена в пределах от $\bar{x} - 3\delta$ до $\bar{x} + 3\delta$.

Если значение случайной величины отличается от \bar{x} больше чем на 3δ , то оно является аномальным.

Испытуемое значение не должно участвовать в расчете \bar{x} и среднеквадратичного отклонения. Для удобства случайную величину нормируют по формуле

$$t = \frac{x - \bar{x}}{\sigma}$$

И правило «трех сигм» преобразуется:

если нормированное значение $|t| > 3$, то оно является аномальным.

Аномальные значения можно выявить на графике пробит-функции, с помощью критерия Титъена-Мура.

Пример.

Средняя зольность угля $\bar{x} = 6,5 \%$, среднеквадратичное отклонение

$\sigma = 2,1\%$. Определить, не является ли аномальной проба угля с зольностью 15% .

Найдем нормированное значение $t = (15 - 6,5) / 2,1 = 4,05$.

Поскольку $t > 3$, проба является аномальной и относится к другой совокупности.

На основе приведенных данных можно определить, какие вообще значения зольности являются аномальными.

Так как $\bar{x} - 3\sigma = 6,5 - 3 * 2,1 = 0,2 \%$;

$$\bar{x} + 3\sigma = 6,5 + 3 * 2,1 = 12,8 \%,$$

то аномальными являются значения зольности $< 0,2$ и $> 12,8 \%$.

4. Выделение однородных совокупностей

Заключение о неоднородности совокупности лучше всего делать по гистограмме частот. Геологическая причина появления двух совокупностей заключается в том, что бедные руды возникли путем замещения алюмосиликатных пород, а богатые - карбонатных пород.

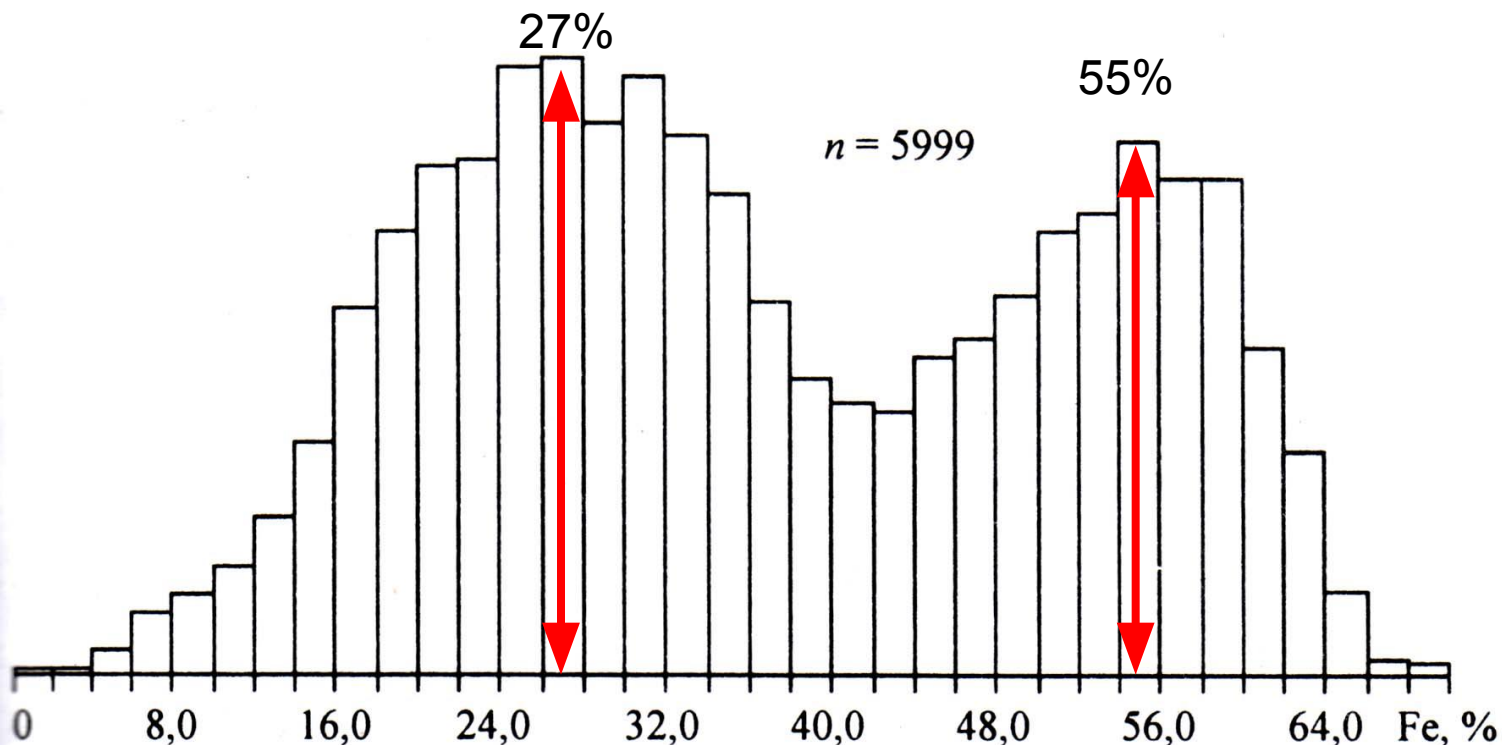


Рис.2.17. Гистограмма содержаний железа в рудах Качарского месторождения

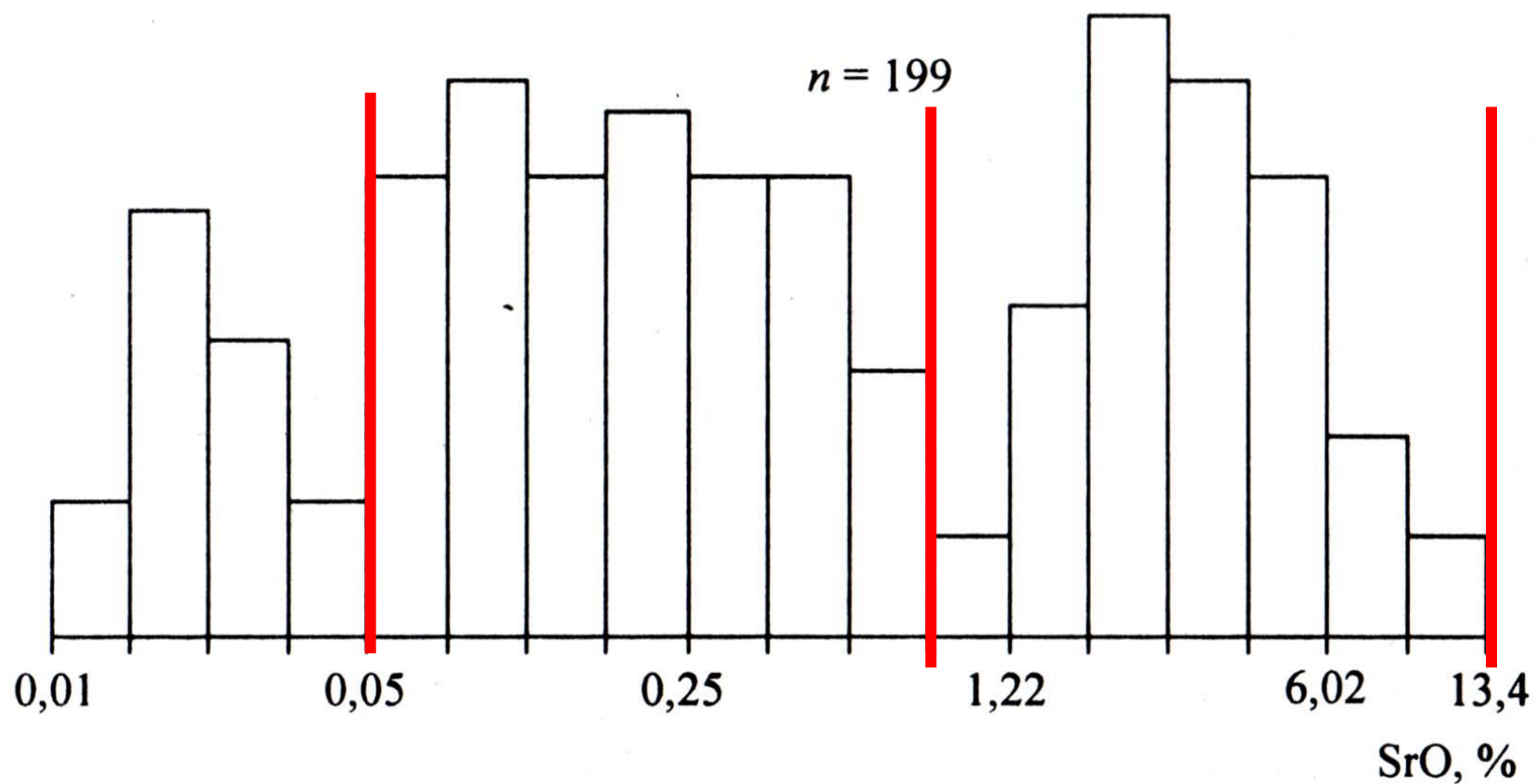
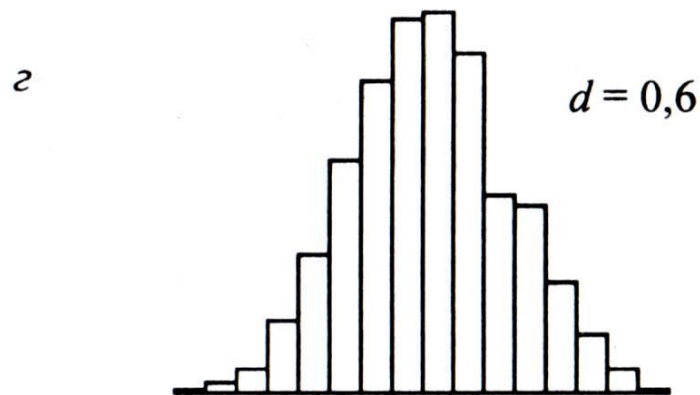
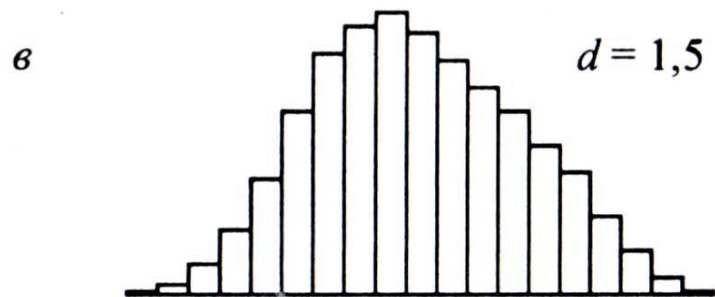
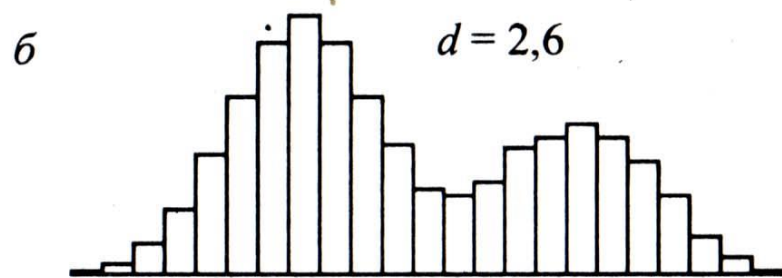
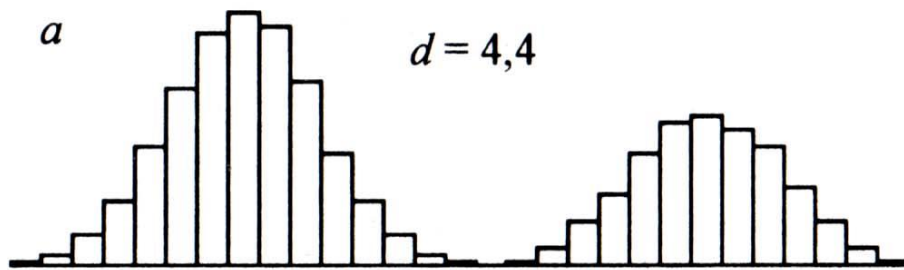


Рис.2.18. Гистограмма содержания стронция в апатитах различных природных образований. Масштаб по горизонтальной оси логарифмический

Наиболее чистыми по содержанию стронция являются апатиты из гранитоидов, ультрабазитов и метаморфических пород. Средние по содержанию стронция - апатиты скарновых месторождений и некоторых массивов щелочных пород. Наиболее высокие содержания стронция наблюдаются в апатитах Хибинской группы месторождений.

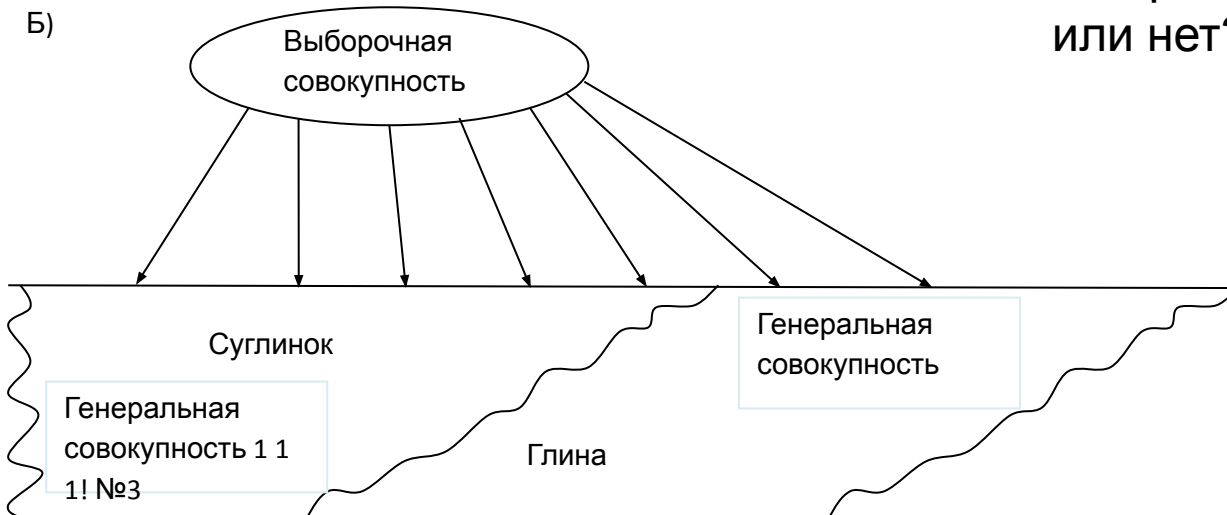
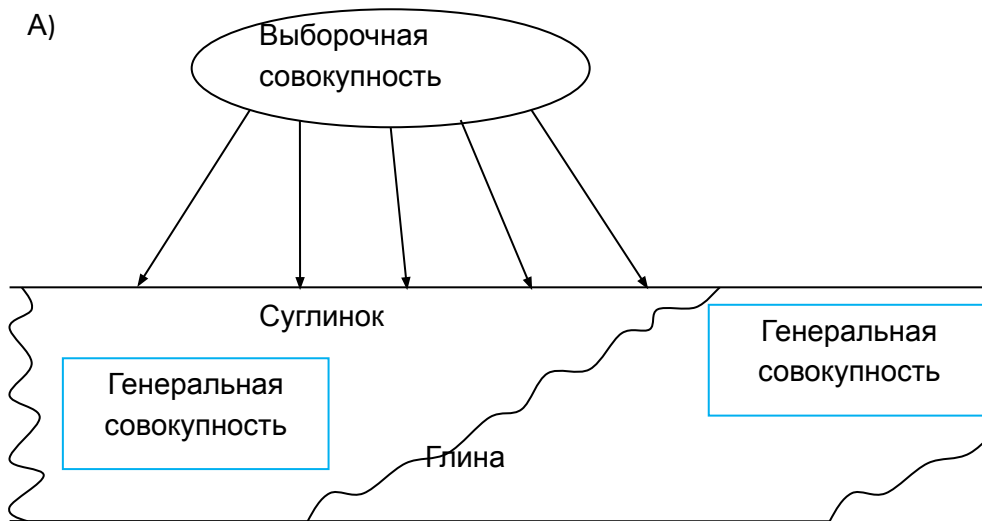


Однородные совокупности, входящие в смешанную совокупность, различаются средними значениями \bar{x} , \bar{y} и дисперсиями σ_x^2 , σ_y^2 .

Показателем, определяющим возможность аналитического разделения смешанных совокупностей при условии нормального их распределения, является *раздвиг* распределений:

$$d = \frac{|\bar{x} - \bar{y}|}{\sqrt{\sigma_x^2 + \sigma_y^2}}$$

1. очень большой ($d > 4$)
2. большой ($d = 2 \div 4$)
3. малый ($d = 0,7 \div 2$)
4. незначительный ($d < 0,7$)



С математической точки зрения инженерно-геологический элемент можно принять за генеральную совокупность. Тогда выделения ИГЭ сводятся к установлению, относится ли выборочная совокупность к одной генеральной совокупности или нет?

Рис. 5.3. Схема формирования выборочной совокупности из:

А) одной генеральной совокупности;

Б) из двух генеральных совокупностей одновременно.



Рис. 5.4. Вид гистограммы выборочной совокупности

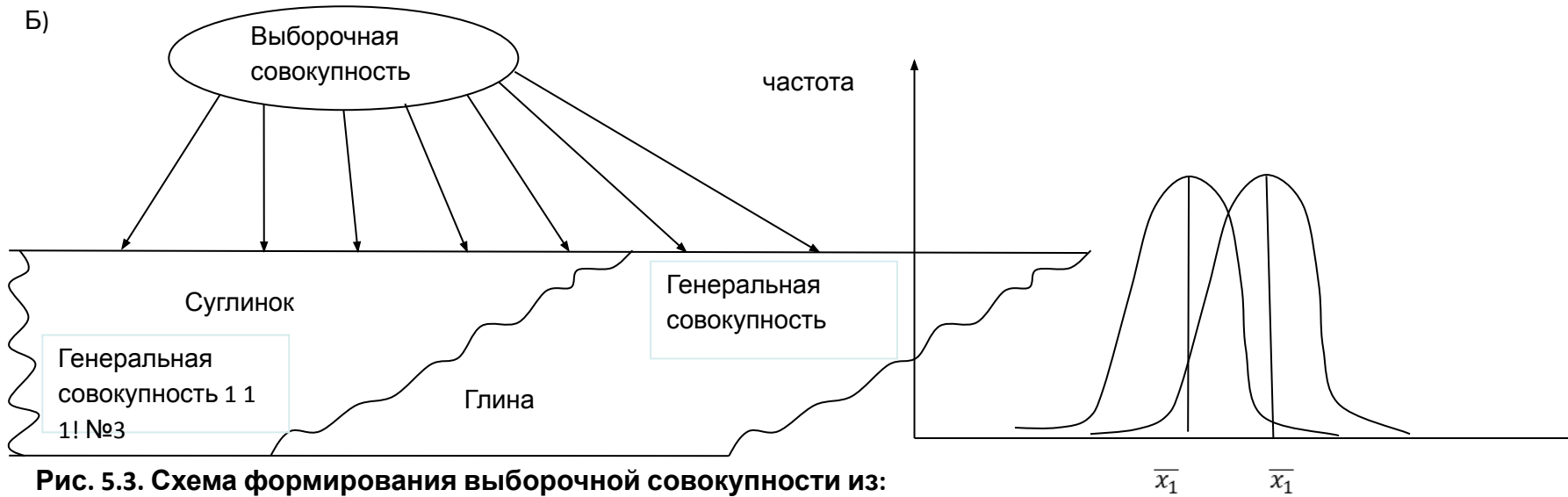


Рис. 5.3. Схема формирования выборочной совокупности из:

А) одной генеральной совокупности;

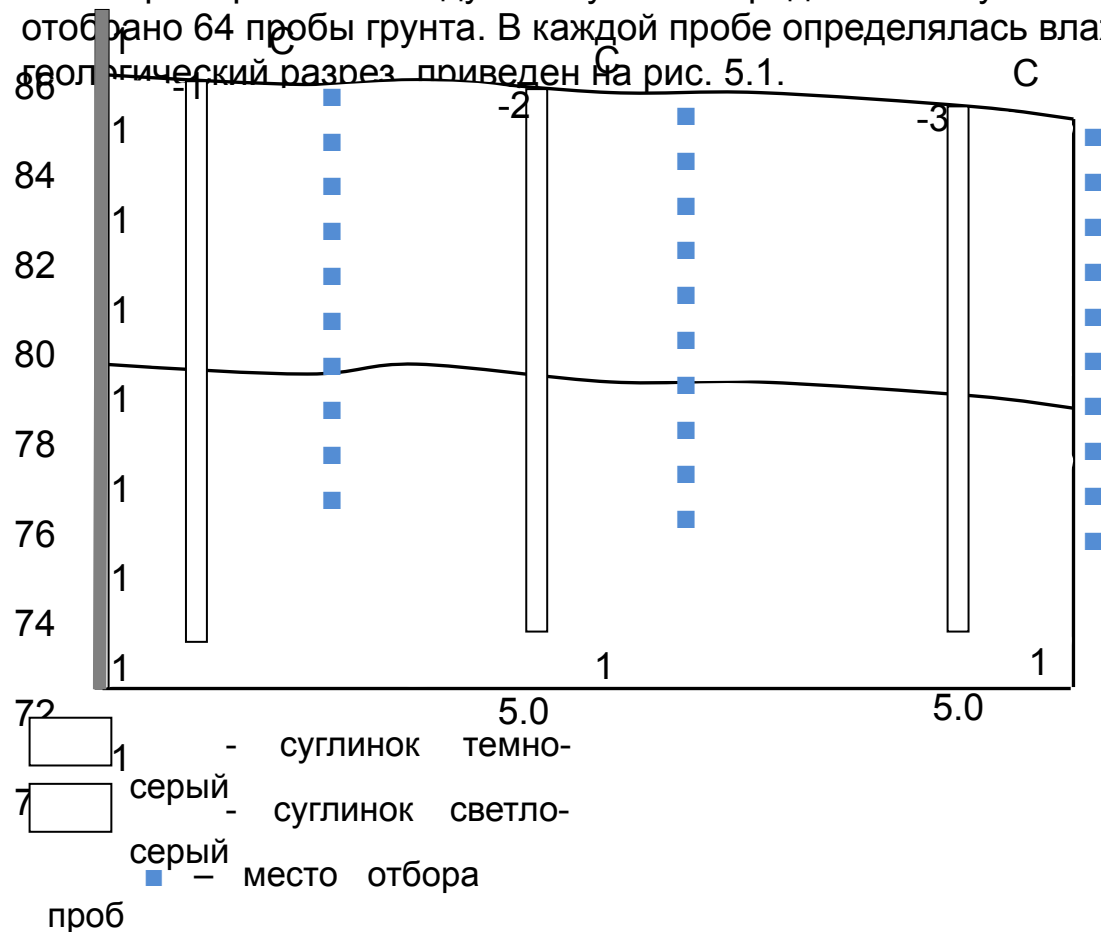
Б) из двух генеральных совокупностей одновременно.

Рис. 5.5. Схема размещения гистограмм одной совокупности представляющих две генеральные совокупности.

- $\overline{x_1}$

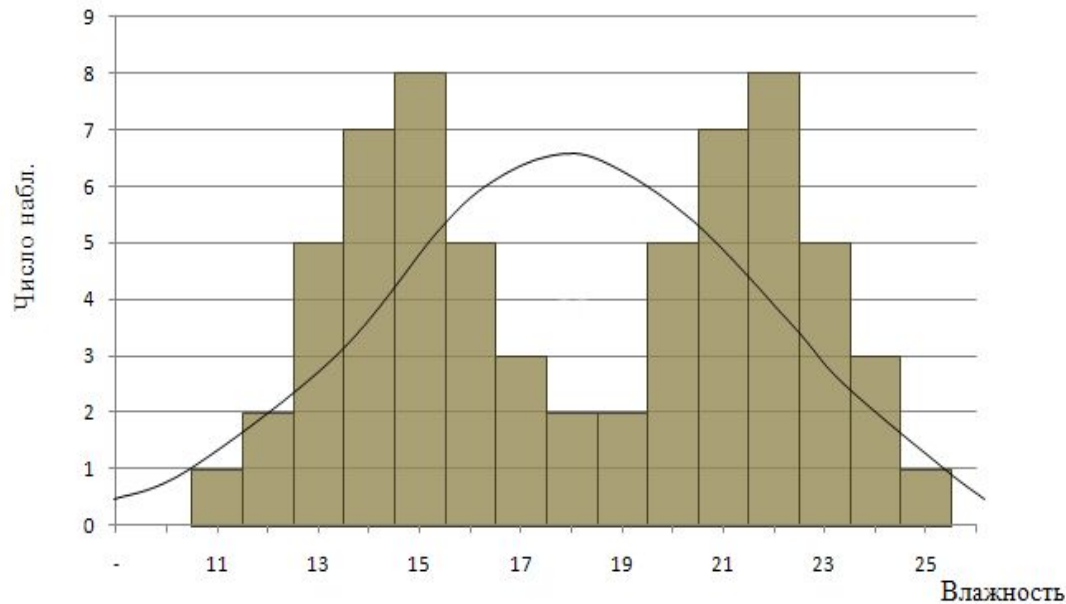
Методика выделения инженерно-геологических элементов статистическим методом.

Пример 5.1. Исследуемый участок представлен суглинком светло-темно-серого цвета. Всего отобрано 64 пробы грунта. В каждой пробе определялась влажность грунта. Схематический геологический разрез приведен на рис. 5.1.



Необходимо определить, геологическое тело представлено одним или несколькими инженерно-геологическими элементами.

Алгоритм решения:



Из рис. 5.2. видно, что в исследуемой выборке наблюдаются два «горба». Отсюда мы можем предположить, что данная выборка представляет две генеральные совокупности (два ИГЭ). Анализ распределения влажности по разрезу показал, что в верхней части разреза суглинков маловлажный и имеет светло-серый цвет, ниже залегает суглинок более влажный темно-серого цвета. Граница между ними установлена на глубине 9м. Используя в качестве классификационных признаков цвет суглинка и граничную влажность $w \approx 18,5\%$, исследуемая выборка разделена на две новые. В каждой выборке имеем по 32 наблюдения, т.е. $n_1=n_2=32$.

$$N=45$$

$$T=140.3\text{град.}$$

$$\sigma = 42$$