ГОДИШНИК НА МИННО-ГЕОЛОЖКИЯ УНИВЕРСИТЕТ "СВ. ИВАН РИЛСКИ", Том 48, Св. II, Добив и преработка на минерални суровини, 2005
ANNUAL OF THE UNIVERSITY OF MINING AND GEOLOGY "ST. IVAN RILSKI", Vol. 48, Part II, Mining and Mineral processing, 2005

ОТНОСНО НЕОБХОДИМИЯ БРОЙ ПРОБИ

Елена Демирева

Минно-геоложки университет "Св. Иван Рилски", 1700 София

РЕЗЮМЕ. Формулиран е проблемът за определяне на минималния брой проби и са определяни факторите, от които той зависи. Познатите методи на математическата статистика, позволяват да бъдат намерени обективни решения. На базата на архивни данни от 22 кариери е даден практически пример за определяне на необходимия брой проби. Принципът на този метод позволява да се приложи и в необходимия брой изпитвания, съобразно приложението на тестваните материали.

ABOUT NECESSARY NUMBER OF SAMPLES

Elena Demireva
University of Mining and Geology, 1700 Sofia

ABSTRACT. The problem about determination of minimum number of samples is formulated and the factors influencing this problem are stressed. The basic equations of mathematical statistics which allow possibly the most objective solution of the problem are set forth. The results of the statistical handling of archives data, gathered of 22 quarries are given and practical ways for determining the necessary number of samples are indicated. The principle of this method may also be applied in determining the necessary minimum number of tests in other fields of technology.

Въведение

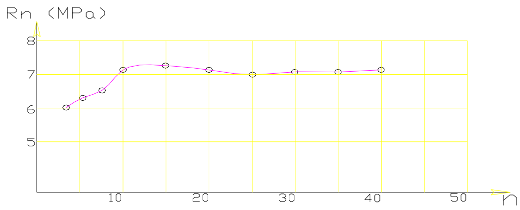
Проблемът за определяне на необходимия брой проби, достоверно отразяващи свойствата на материала, от който те за взети, е от първостепенно значение в земната механика.

Главните фактори, които оказват влияние върху числеността на пробите са:

1. Геоложки фактори:
   - издържаност на кариерата или земната основа в дълбочина и в ситуация;
   - еднородност на материала;
   - брой и размери на отделните находища (количество на материала);
   - вид на материала и др.
2. Технически фактори:
   - категория на съоръжението;
   - конструктивни особености на съоръжението;
   - техническа издържаност на проектното решение;
3. Икономически фактори:
   - капиталност на съоръжението;
   - икономичност на проектното решение;
   - разходи по вземането, транспортирането и изследването на пробите;
   - предписания относно допустимите разлики в сметната стойност (това при твърда договорена стойност е много важен фактор);
4. Организационни фактори:
   - капацитет на стационарните лаборатории;
   - необходимо време за провеждане на лабораторните изследвания;
   - точност на лабораторната работа;
   - методология на изследването;
   - обзаведеност на лабораторията и др.

Но между отделните фактори, както и между членовете на някои отделни групи фактори съществуват противоречия. Например, колкото по-голям е броят на взетите проби, толкова по-точни ще бъдат изчислените данни, т.е. по-точно ще определим сигурността на съоръжението. Явно е, обаче също така, че прекалено голям брой проби не би бил оправдан при еднородна и издържана кариера. Освен това големият брой проби не се оправдава от икономическа гледна точка. Нещо повече – ненужно големия брой проби не допринася за съществено уточняване на изчислителните данни (фиг. 1).

От една проба, независимо от това дали тя е нарушена или ненарушена, могат да се получат данни за няколко различни показатели. Дори и да бъдат направени на една проба няколко еднакви изследвания на един и същи показател, даже и резултатите да се различават помежду си, трябва да се счита, че средната стойност на еднаквите изследвания характеризира еднократно определения материал. Следователно, една проба, почти винаги се покрива с един определен материал, защото случаите, в които са необходими две или повече проби за установяване на един и същи показател са рядкост.



Фиг. 1.

Поради това и за удобство няма да се прави разлика между показател и проба и ще се говори за проба, като се има предвид, че винаги се предписва броя и вида на пробите, които трябва да се вземат при проучването, а не броя на показателите, които трябва да се изследват.

Като се вземе предвид казаното дотук възниква съвсем естествено въпросът как да се определи минималния брой проби, които трябва да се изследват?

Задоволителен отговор на този въпрос може да се получи само с познатите методи на математическата статистика, тъй като взаимодействието на факторите определящи броя на пробите е сложно, многообразно, а често и противоречиво. На практика сме принудени да наблюдаваме сумарния ефект от влиянието на отделните фактори и нашата безпомощност да разчленим многостранните им връзки.

Достоверни резултати биха били тези, които имат устойчиви, многократно повтарящи се величини т.е. сумарни статистически характеристики.

Теоретични постановки

Опитът показва, че поведението на редица многократно повтарящи се операции, като например технически измервания, стандартни лабораторни изпитвания е такова като че ли те се случват при устойчиви условия, или съществува само една променлива, която трябва да се измерва при дадена операция. При тези условия експерименталните данни се възпроизвеждат добре от теоретичната крива на нормално разпределение (Гаусова крива), при която колкото е по-голям броя на данните толкова това съответствие е по-добро.

Добре познатите

$$\overline{X} = \frac{\sum_{1}^{n} x_i}{n} \qquad \sigma = \frac{\sqrt{\sum \left(x_i - \overline{X}\right)^2}}{n-1} \qquad (1)$$

(1)

средно аритметична стойност, като представител на редицата и средноквадратично отклонение, което показва до каква степен отделните стойности се различават едно от друго.

Обикновено постановката на задачата при използване на методите на математичната статистика за обработка на данни се състои в следното: на една съвкупност от N елемента $x_1, x_2, x_3, \ldots x_N$, които имат характер на случайни величини да се установи точността ε, с която е определена средната аритметична стойност , т.е. при зададена вероятност α, да се определи т.н. доверителен интервал ($\overline{X} - \varepsilon, \overline{X} + \varepsilon$), в чиито граници може да се изменя $\overline{X}$; да се определи също така и средното квадратично отклонение σ, както и някои други статистически показатели, с чиято помощ да се прецени съвкупността.

В нашия случай е по-интересна задачата чрез разглеждане на само част от елементите означени с n, взети по пътя на случайния подбор от общата съвкупност от N елемента, да бъдат установени обобщените статистически показатели на цялата общност. С други думи: с помощта на ограничен, но правилно избран отбор на малък брой елементи n, чиято средно аритметична може лесно да се изчисли, да се прецени генералната средноаритметична стойност на общата съвкупност от N елемента. Трябва да се отбележи, че в аспекта на прилагане на метода на ограничения отбор в техниката ние почти винаги боравим с общи съвкупности, чиито признаци се подчиняват на общия закон за разпределение.[1]

По-детайлно въпросът може да се проследи в специализираната литература (Dlin и Heinhold). Там се извежда следното уравнение, което дава ключа за решаване на проблема

$$\left|\overline{x} - u\right| = \frac{\sigma}{\sqrt{n}}.t. \qquad (2)$$

или

$$\left|\varepsilon\right| = \frac{\sigma}{\sqrt{n}}.t. \qquad (3)$$

[1] Опитът показва, че дори разпределението да се отличава до някаква степен от нормалното, преценката на генералната средноаритметична по указания по-нататък начин е достатъчно добра.

където t е така нареченото нормирано отклонение, чиято големина се подчинява на закона за разпределение на Student.

При $k = n - 1$ степени на свобода, t зависи само от n и значенията на стойностите $x_1, x_2, x_3, \ldots x_n$, отбрани от общата съвкупност, а не зависи от нейното генерално, неизвесно средно квадратично отклонение σ.

Следователно с вероятността α (която обикновено се приема от 0,95 до 0,99 за практически изследвания) е валидно следното неравенство

$$|\varepsilon| \langle \frac{\sigma}{\sqrt{n}} . t_\alpha^k . \qquad (4)$$

Изразено с думи вероятността, че нормалното отклонение лежи в интервала $-t_\alpha^k \langle t \langle + t_\alpha^k$ е равна на α.

В това неравенство t_α^k е възможно най-голямото отклонение на часната средноаритметична стойност $\overline{x}$ от генералната средна $\overline{X}$. То се отчита в зависимост от к и α от таблици дадени в литературата $[1,2]$. С това в същност задачата в главни линии е решена.

В случай, че от предварителни изследвания ни е извесна някоя стойност на поправеното средноквадратично отклонение можем да установим какъв е необходимия брой проби при зададена грешка $\pm E^2$ и вероятност α. Това с голяма вероятност можем да твърдим, ако n се избере толкова голямо, че да е изпълнено неравенството

$$\frac{\sigma}{\sqrt{n}} . t_\alpha^k \langle |E|. \qquad (5)$$

$$\frac{t_\alpha^k}{\sqrt{n}} \langle \frac{|E|}{\sigma}. \qquad (6)$$

Ако означим с $\chi = \frac{|E|}{\sigma}.$ и $\tau_\alpha = \frac{t_\alpha^k}{\sqrt{n}}.$ ще получим в крайна сметка условието $\tau_\alpha \langle \chi$.

С цел да се опрости решението на задачата е съставена специална таблица, която свежда изчислителната работа до минимум. След като се определи χ от таблицата веднага може да се определи за прието α онази стойност на n, за която уравнението (5) все още изпълнено. Именно тази стойност на n представлява минималния брой проби, които трябва да бъдат взети и изследвани.

Практическо приложение

Ако разгледаме неравенство 5 и/или 6, първата част на неравенството се отчита директно по таблицата, а втората съдържа грешката и средното квадратично отклонение т.е. геоложките фактори се съдържат в стандартното отклонение, а техническите се вземат под внимание в зададената грешка. Икономическите фактори се вземат предвид както при определянето на грешката, така и в това че се определя минималния брой проби, за които е удовлетворено неравенството 5. Организационните фактори намират израз именно при определяне на минимално необходимия брой проби и при подбора на една от определените стойности на σ. Върху избора на α оказват влияние както техникоикономическите, така и организационните фактори.

Проблемът явно се свежда до фиксиране на грешката и средното квадратично отклонение. Като пример за работа върху определяне на минималния брой проби е приложена таблица с архивни данни от кариери за добив на скален материал. С курсив са отбелязани минималните и максимални стойности, които определят т.н. размах на стойностите.

Данните в таблицата са подбрани в съответствие с необходимостта от оценка на качеството на работата по уплътняване на материалите от кариерите като насипна основа на пътища, диги и др.

Примерно изчисление:
Необходимо е да бъде определен минималния брой проби, които трябва да се вземат и изследват, с оглед да се организира правилно качественния контрол на обекта по отношение на w.

В литературата са описани (Davis, Heinhold, Linder, Пилгунов и др.) значително количество данни, които показват, че стандартното отклонение се движи в границите от $\pm 0{,}5\%$ до $\pm 3\%$, в зависимост от условията в кариерите. Така че, ако приемем въз основа на тези данни $\sigma = \pm 2\%$, то с помощта на Таблица 2 се получава $\chi = \frac{|E|}{\sigma}.$=0,645. От таблица 1 се отчита, че за средни условия неравенството $\tau_\alpha \langle \chi$. е изпълнено, ако се звемат и изследват за най малко n = 12 проби $(\alpha = 0{,}95)$.

По аналогичен път, ако се приеме, че средната допустима грешка за ъгъла на вътрешно триене е 1 градус се получава, че при $(\alpha = 0{,}95)$ са необходими 4 проби, а при обезпеченост 0,99 - 6 проби.

Когато, обаче целта е да се определи минималният брой проби, необходими за определяне на якостни характеристики е уместно да бъде използван за контрол познатия метод на Пилгунов, при който при зададена обезпеченост 0,99, необходимият минимален брой проби е 10. В противен случай трябва да се пределят лабораторно минималните стойности за $tg\varphi$ и с, което явно е икономически неизгодно.

Задачата може да бъде формулирана и обратно – каква вероятна грешка бихме получили, ако изследваме определен брой проби по отношение на даден показател.

Ако преработим неравенството ще получим

$|E|\langle \sigma . \tau_{\alpha}$

Например, каква вероятна грешка ще получим, ако обработим 5 проби при изследване на γ . От Табл.1 отчитаме за $(\alpha = 0{,}95)$ $\tau_{0,95}$ = 1,241. Следователно със сигурност 95% можем да твърдим, че грешката ще бъде от порядъка на $\pm$ 0,078.1,241 = $\pm$ 0,096 .

От изложеното до тук се вижда, че при направените предпоставки, математическата статистика предлага сравнително лесен метод за определяне на необходимия брой проби, но към него не бива да се гледа схематично. Необходим е предварителен анализ на основните предпоставки и чак тогава използване на математическите правила.

Таблица 1.

	$\tau_{0,95}$	$\tau_{0,99}$
1	-	-
2	8,985	45,013
3	2,484	5,730
4	1,591	2,920
5	1,241	2,059
6	1,050	1,646
7	0,925	1,401
8	0,836	1,237
9	0,769	1,118
10	0,715	1,028
11	0,672	0,955
12	0,635	0,897
13	0,604	0,847
14	0,577	0,805
15	0,554	0,769
16	0,533	0,737
17	0,514	0,708

Препоръчана за публикуване от
Катедра "Подземно строителство", МТФ

Таблица 2.

Обект	n	Средно квадратично отклонение				
		γ_d	γ	φ	c	W
Кариера 1	9-14	0,072	*0,042*	0,032	0,049	0,038
Кариера 2	14-16	0,063	0,065	0,032	0,048	0,019
Кариера 3	21	0,076	0,054	*0,028*	0,104	0,030
Кариера 4	18-28	0,052	0,063	0,066	0,113	0,021
Кариера 5	6-14	0,147	0,125	0,048	0,054	*0,060*
Кариера 6	12-18	0,092	0,078	0,067	0,067	0,039
Кариера 7	5-9	0,128	0,089	0,034	0,061	0,040
Кариера 8	8-25	0,121	0,087	0,056	0,083	0,057
Кариера 9	22	-	-	0,046	0,070	-
Кариера 10	16-19	0,085	0,096	0,048	0,075	*0,018*
Кариера 11	15-22	*0,159*	*0,156*	0,040	0,070	0,039
Кариера 12	22	-	-	0,057	0,082	-
Кариера 13	9-14	0,052	0,069	*0,086*	*0,150*	0,020
Кариера 14	6	0,147	*0,153*	0,054	0,075	0,022
Кариера 15	7-14	*0,049*	*0,021*	0,075	0,067	0,046
Кариера 16	9	0,09	0,063	0,031	*0,031*	0,028
Кариера 17	8-14	0,050	0,082	0,047	0,081	0,025
Кариера 18	5-12	0,102	0,132	0,051	0,061	0,032
Кариера 19	11-15	0,063	0,071	0,057	-	0,025
Кариера 20	8-10	0,068	0,066	0,072	0,130	0,021
Кариера 21	13-17	0,075	0,107	0,044	0,071	0,021
Кариера 22	12-15	0,072	0,075	0,043	0,083	0,022
Ср. стойности	-	**0,088**	**0,078**	**0,051**	**0,077**	**0,031**

Литература

GoldsteIn, M. The Teory of Probability and Statistics in Relation to the Reology of Soils, Proceedings of the Fifth International Conference of Soil Mechanics, Paris, 1961.

Davis, F. Quality Control of Earth Embankments, Proceedings of the Jhird International Conference of Soil Mechanics, Zurich.