

ЛЕКЦИОННОЕ ЗАНЯТИЕ

Определение необходимого объема наблюдений

Как уже указывалось, под статистическим наблюдением понимают планомерный, научно организованный систематический сбор данных о явлениях и процессах путем регистрации заранее намеченных существенных признаков.

Статистическое наблюдение осуществляется в двух формах: представление отчетности и проведение специально организованных статистических наблюдений. При этом оно должно быть организовано таким образом, чтобы в результате были получены объективные, точные данные об изучаемом явлении, на основе которых можно правильно оценить и планировать развитие производства.

В зависимости от степени полноты охвата наблюдением изучаемого явления или объекта различают сплошное и несплошное статистическое наблюдение (рис. 3).

При сплошном наблюдении обследованию подвергаются все без исключения единицы изучаемой совокупности, а при несплошном – только часть их. Различают следующие основные виды несплошного наблюдения: выборочное наблюдение, монографическое обследование и метод основного массива. Наиболее совершенным, научно обоснованным способом несплошного наблюдения является выборочное наблюдение, при котором статистическому обследованию подвергаются не все единицы изучаемой совокупности, а лишь отобранные в определенном порядке и обеспечивающие получение данных, характеризующих всю совокупность в целом. К этому виду наблюдений прибегают в тех случаях, когда необходимо сэкономить силы и средства при проведении исследования, так как статистические исследования массовых совокупностей весьма трудоемки.



Рис. 3. Оценка статистического наблюдения

При выборочном наблюдении неизбежна некоторая свойственная ему погрешность, так как обследованию подвергается не вся совокупность, а только ее часть.

Ошибки, свойственные выборочному наблюдению, называются ошибками репрезентативности (представительства). Ошибки характеризуют размер расхождения между данными выборочного наблюдения и всей совокупности в целом.

Различают случайные и систематические ошибки репрезентативности. Случайные ошибки обусловлены тем, что выбранная совокупность недостаточно точно воспроизводит всю совокупность. Их размеры и пределы заранее можно вычислить на основании закона больших чисел, а главное – довести до незначительных размеров путем включения в выборку достаточного количества единиц совокупности.

Систематические ошибки возникают при нарушении принципа случайности отбора единиц совокупности для наблюдения. Их можно избежать, осуществляя строгий объективный отбор единиц совокупности, при котором каждая из них имела бы абсолютно одинаковую возможность попасть в выборку.

В статистической литературе принято называть совокупность единиц, из которой производится отбор, генеральной совокупностью, а ее численность обозначают буквой N . Часть единиц, попавшая в выборку, называется выборочной совокупностью, а их численность обозначается буквой n .

Обобщающими характеристиками генеральной и выборочной совокупностей являются средняя и дисперсия: $\bar{X}$ и σ^2 .

При формировании выборки применяют различные виды, схемы и способы отбора. По виду отбор единиц наблюдения подразделяют на индивидуальный (за прием отбирается одна единица), групповой (за один прием отбирается группа или серия единиц) и комбинированный.

В зависимости от участия отобранной единицы в дальнейшей выборке различают повторную и бесповторную схемы отбора. В первом случае отобранная однажды единица возвращается обратно в генеральную совокупность и снова участвует в выборке.

При бесповторной схеме отбора ранее отобранная единица не возвращается в генеральную совокупность и не может быть подвергнута повторному обследованию.

Бесповторный отбор дает более точные результаты по сравнению с повторным, так как при одном и том же объеме выборки наблюдение охватывает большее количество единиц изучаемой совокупности. Как правило, при выборочном наблюдении следует применять бесповторную схему отбора. И только в тех случаях, когда бесповторный отбор нельзя провести, применяется повторная схема отбора.

По способу формирования выборки различают, собственно-случайный, механический, типический, серийный и комбинированный отбор. При собственно-случайном способе формирования выборки включение единиц в

выборочную совокупность может осуществляться по схеме повторного или бесповторного отбора при помощи жеребьевки или таблицы случайных-чисел. Такой способ формирования выборки является наиболее простым, но он уступает другим способам с точки зрения репрезентативности и точности результатов, а также из-за сложности в организационном отношении.

При механическом способе формирования выборки отбор единиц из генеральной совокупности- производится механически через определенный интервал (например, выбирается каждая пятая, десятая и т. д. единица). Все единицы в изучаемой совокупности предварительно располагаются в определенном порядке (например, по алфавиту, местоположению и т. п.). Данный способ отбора является разновидностью собственно-случайного отбора, но имеет ряд организационных преимуществ и всегда бывает бесповторным. Поэтому средняя ошибка и необходимая численность механической выборки определяются по формулам собственно-случайной бесповторной выборки.

В тех случаях, когда генеральная совокупность неоднородна по показателям, подлежащим изучению, применяется типический способ отбора, при котором вся генеральная совокупность делится предварительно на группы по определенному типическому признаку, из которых в дальнейшем собственно-случайным или механическим способом формируется выборочная совокупность. Типический способ отбора также может быть как повторным, так и бесповторным. Из всех типических групп можно отобрать число единиц, пропорциональное их численности, и непропорциональное. В зависимости от этого различают пропорциональный и непропорциональный типический отбор.

При серийном способе отбора вместо случайного отбора единиц совокупности осуществляется отбор групп (серий), внутри которых производится сплошное наблюдение. Точность серийной выборки зависит от межсерийной дисперсии (дисперсии групповых средних), а не от величины общей дисперсии.

При планировании выборочного наблюдения возникает вопрос о необходимой численности выборки. Последнюю можно определить, исходя из допустимой ошибки при выборочном наблюдении, вероятности, с которой нужно гарантировать величину устанавливаемой ошибки, меры колеблемости изучаемого признака и способа отбора.

Необходимая численность выборки определяется на основе предельной ошибки выборки. Если предельную ошибку выборки обозначить буквой σ , то последнюю можно определить из выражения

$$\sigma = T\mu$$

где μ – средняя ошибка выборки;

T – коэффициент, зависящий от вероятности, с которой гарантируется ошибка выборки (коэффициент доверия).

Значения вероятности (P) для различных значений T приведены в таблице (табл. 3).

Предельная ошибка выборки зависит от трех факторов: степени колеблемости явления (σ^2), объема выборки (n) и от необходимой гарантированной вероятности (P).

Формулы для вычисления предельных ошибок выборки при различных способах отбора приведены в таблице (табл. 1).

Таблица 1

Способ отбора	Схема отбора	Предельная ошибка выборки
Собственно-случайный и механический отбор	Повторный	$\sigma = T \sqrt{\frac{\sigma^2}{n}}$
	Бесповторный	$\sigma = T \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
Типический отбор	Повторный	$\sigma = T \sqrt{\frac{\overline{\sigma_i^2}}{n}}$
	Бесповторный	$\sigma = T \sqrt{\frac{\overline{\sigma_i^2}}{n} \left(1 - \frac{n}{N}\right)}$
Серийный отбор	Повторный	$\sigma = T \sqrt{\frac{\overline{b^2}}{r}}$
	Бесповторный	$\sigma = T \sqrt{\frac{\overline{b^2}}{r} \left(1 - \frac{r}{R}\right)}$

В табл. 1 приняты следующие условные обозначения:

$\overline{\sigma_i^2}$ – средняя из групповых дисперсий;

$\overline{b^2}$ – межсерийная дисперсия;

r – число отобранных серий;

R – число серий в генеральной совокупности.

Величину $\overline{b^2}$ можно определить из выражения:

$$\overline{b^2} = \frac{\sum (\tilde{x}_i - \bar{X})^2}{r}$$

где $\tilde{x}_i$ – средняя в отдельных сериях;

$\bar{X}$ – общая средняя для всей совокупности.

Проведя необходимые вычисления и пользуясь расчетными формулами (табл. 2), определить необходимую численность выборки при различных способах и схемах отбора.

Таблица 2

Расчетные формулы для определения необходимой численности при различных способах и схемах отбора

Способ отбора	Схема отбора	Численность выборки
Собственно-случайный и механический отбор	Повторный	$n = \frac{T^2 \sigma^2}{\delta^2}$
	Бесповторный	$n = \frac{T^2 \sigma^2 N}{\delta^2 N + T^2 \sigma^2}$
Типический отбор	Повторный	$n = \frac{T^2 \bar{\sigma}_i^2}{\delta^2}$
	Бесповторный	$n = \frac{T^2 \bar{\sigma}_i^2 N}{\delta^2 N + T^2 \bar{\sigma}_i^2}$
Серийный отбор	Повторный	$n = \frac{T^2 \bar{b}^2}{\delta^2}$
	Бесповторный	$n = \frac{T^2 \bar{b}^2 N}{\delta^2 R + T^2 \bar{b}^2}$

Из приведенных в табл. 2 формул для определения необходимой численности выборки при различных способах и схемах отбора видно, что они зависят лишь от схемы отбора (повторный отбор или бесповторный).

Способ отбора влияет следующим образом; в формуле для необходимой численности выборки при собственно-случайном или механическом способе вместо общей дисперсии σ^2 используют среднюю из внутригрупповых дисперсий $\bar{\sigma}_i^2$ при типическом способе отбора и межсерийную дисперсию $\bar{b}^2$ при серийном способе. При этом в последнем случае вместо числа единиц в генеральной совокупности N используют число серий в генеральной совокупности R .

В связи с вышеизложенным в лабораторной работе при типическом и серийном способах отбора предварительно определяются значения $\bar{\sigma}_i^2$ или $\bar{b}^2$.

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ

Линейное программирование в MS Excel

Задание: научиться пользоваться аналитическими инструментами для решения задач оптимизации при определенных условиях.

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

Таблица 3

Исходные данные

№ варианта	Плановое качество руды, %	Потери руды по системе разработки, %		Объем добычи по руднику, тыс. т
		I	II	
1	5,4-6,9	18	12	400
2	6,2-8,3	3	10	600
3	7,5-9,0	24	16	550
4	5,1-6,4	14	22	400
5	7,9-9,2	8	5	350
6	3,7-6,0	28	19	650
7	7,4-10,3	16	17	600
8	4,8-7,1	10	14	550
9	6,0-8,5	15	11	600
0	7,4-9,3	5	9	450

Коэффициент извлечения руды из недр определяется по формуле

$$K_{изв} = 1 - П,$$

где $П$ – это потери руды, д.е.

Пример. Рассмотрим рудник, состоящий из двух участков, на каждом из которых с учетом горно-геологических условий могут применяться две системы разработки. Плановое качество руды определено содержанием полезного компонента в пределах 6,8-7,2%, а объем добычи должен быть не менее 500 тыс. т. Коэффициенты извлечения руды из недр для систем разработки равны 0,6 и 0,7. Остальные данные по участкам и системам разработки указаны в табл. 4.

Таблица 4

Исходные данные

№ участка	Содержание полезного компонента, %	Затраты на добычу руды (руб/т) при системах разработки		Минимальная добыча руды, тыс. т	Максимальная добыча руды, тыс. т
		I	II		
1	6	3,0	2,5	180	320
2	8	2,0	1,5	200	360

Какими системами разработки на участках можно обеспечить выполнение плановых и технологических условий с минимальными суммарными затратами на добычу?

Решение: обозначим через x_i искомый объем добычных работ на i -м участке при использовании j -й системы разработки ($i = \overline{1,2}, j = \overline{1,2}$).

Запишем математическую модель задачи.

Целевая функция – суммарные затраты на добычу

$$F(x) = 3,0x_{11} + 2,5x_{12} + 2,0x_{21} + 1,5x_{22} \rightarrow \min.$$

При наличии ограничивающих условий:

а) по плановой добыче

$$x_{11} + x_{12} + x_{21} + x_{22} \geq 500;$$

б) по качеству руды

$$6,8 \leq \frac{6 \times 0,6x_{11} + 6 \times 0,7x_{12} + 8 \times 0,6x_{21} + 8 \times 0,7x_{22}}{0,6x_{11} + 0,7x_{12} + 0,6x_{21} + 0,7x_{22}} \leq 7,2;$$

в) по общему объему добычи участков с учетом запасов

$$180 \leq x_{11} + x_{12} \leq 320;$$

$$200 \leq x_{21} + x_{22} \leq 360;$$

г) по неотрицательности искоемых объемов добычи

$$x_{ij} \geq 0, i = \overline{1,2}, j = \overline{1,2}.$$

После необходимых преобразований ограничивающее условие «б» запишем в виде двух неравенств:

$$\begin{aligned} &6,8(6 \times 0,6x_{11} + 6 \times 0,7x_{12} + 8 \times 0,6x_{21} + 8 \times 0,7x_{22}) \\ &\quad - (6 \times 0,6x_{11} + 6 \times 0,7x_{12} + 8 \times 0,6x_{21} + 8 \times 0,7x_{22}) \leq 0; \\ &(6 \times 0,6x_{11} + 6 \times 0,7x_{12} + 8 \times 0,6x_{21} + 8 \times 0,7x_{22}) \\ &\quad - 7,2(6 \times 0,6x_{11} + 6 \times 0,7x_{12} + 8 \times 0,6x_{21} + 8 \times 0,7x_{22}) \leq 0 \end{aligned}$$

Заполняем таблицу в программе MS Excel по следующему алгоритму:

1. В ячейках A2:D2 указываем начальные значения величин x_{ij} (нули);
2. В ячейках A3:D3 разместим коэффициенты при неизвестных из левых частей ограничений;
3. В ячейки A4:D5 вводим значения условий по качеству руды (по содержаниям полезного компонента и по коэффициентам извлечения руды из недр);
4. В ячейки A6:B6 вводим величины минимальной добычи;
5. В ячейки A7:B7 - величины максимальной добычи;
6. В ячейках A8:D8 указываем затраты на добычу 1 т руды;
7. В ячейки E4:E5 вводим формулы для упрощения расчета левых частей неравенств:

$$=\text{СУММПРОИЗВ}(\$A\$2:\$D\$2;A4:D4);$$

$$=\text{СУММПРОИЗВ}(\$A\$2:\$D\$2;A4:D4;A5:D5);$$

8. В ячейку H2 записываем целевую функцию

$$=\text{СУММПРОИЗВ}(A2:D2;A8:D8);$$

9. В ячейки F3:F7 вводим формулы для расчета левых частей неравенств:

$$=\text{СУММПРОИЗВ}(\$A\$2:\$D\$2;A3:D3);$$

$$=6,8*E4-E5;$$

$$=E5-7,2*E4;$$

$$=\text{СУММ}(A2:B2);$$

$$=\text{СУММ}(C2:D2);$$

10. В ячейки G3:G5 записываются ограничивающие условия (правые части неравенств)

	A	B	C	D	E	F	G	H
1	x ₁₁	x ₁₂	x ₂₁	x ₂₂	Формулы для расчета левых частей неравенств	Левые части неравенств	Правые части неравенств	F(x)
2	0	200	0	300		=СУММПРОИЗВ(\$A\$2:\$D\$2;A3:D3)	500	=СУММПРОИЗВ(A2:D2;A8:D8)
3	1	1	1	1				
4	0,6	0,7	0,6	0,7	=СУММПРОИЗВ(\$A\$2:\$D\$2;A4:D4)	=6,8*E4-E5	0	
5	6	6	8	8	=СУММПРОИЗВ(\$A\$2:\$D\$2;A4:D4;A5:D5)	=E5-7,2*E4	0	
6	180	200				=СУММ(A2:B2)		
7	320	360				=СУММ(C2:D2)		
8	3	2,5	2	1,5				
9								

Заполнение таблицы исходных данных

11. Выполним команду «Данные→Анализ→Поиск решения» – откроется диалоговое окно «Параметры поиска решения» (предварительно установить в «Надстройках Excel»);

Параметры поиска решения

Оптимизировать целевую функцию:

До: ☐ Максимум ☒ Минимум ☐ Значения:

Изменяя ячейки переменных:

В соответствии с ограничениями:

Добавить
Изменить
Удалить
Сбросить
Загрузить/сохранить

☒ Сделать переменные без ограничений неотрицательными

Выберите метод решения:

Параметры

Метод решения

Для гладких нелинейных задач используйте поиск решения нелинейных задач методом ОПГ, для линейных задач - поиск решения линейных задач симплекс-методом, а для негладких задач - эволюционный поиск решения.

Справка

Заполнение диалогового окна «Параметры поиска решения»

12. В поле «Оптимизировать целевую функцию» мышью укажем ячейку, содержащую оптимизируемое значение (H2).

	A	B	C	D	E	F	G	H
1	x_{11}	x_{12}	x_{21}	x_{22}	Формулы для расчета левых частей неравенств	Левые части неравенств	Правые части неравенств	F(x)
2	0	200	0	300				950
3	1	1	1	1		500	500	
4	0,6	0,7	0,6	0,7	350	-140	0	
5	6	6	8	8	2520	0	0	
6	180	200				200		
7	320	360				300		
8	3	2,5	2	1,5				
9								

Решение задачи оптимального применения систем разработки

Оптимизация позволила выявить, что и на первом и на втором участках следует применять только II систему разработки, добывая соответственно по участкам 200 и 300 тыс. т руды. При таком планировании минимальные затраты на добычу составят 950 тыс. руб.

Значения в ячейках F3 = 500 и F5 = 0 показывают, что первое и третье ограничения выполняются как равенства, т.е. суммарная добыча составит 500 тыс.т, а содержание полезного компонента 7