



Выборочное наблюдение

- 
1. Выборочное наблюдение, его сущность и значение
 2. Ошибки выборочного наблюдения и определение необходимой численности выборки
 3. Методы, виды и способы отбора выборочных совокупностей
 4. Распространение результатов выборочного наблюдения на генеральную совокупность



Вопрос 1

- Выборочное наблюдение, его сущность и значение



Выборочное наблюдение - это такое несплошное наблюдение, при котором отбор подлежащих обследованию единиц осуществляется в случайном порядке, отобранная часть изучается, а результаты распространяются на всю изучаемую совокупность.

Преимущества выборочного наблюдения

- достижение большей точности результатов обследования благодаря сокращению числа ошибок регистрации и получение результатов допустимой степени точности;
- экономия трудовых и денежных средств и времени в результате уменьшения объема работы;
- возможность детального обследования каждой единицы наблюдения за счет расширения программы наблюдения;
- сведение к минимуму уничтожения и приведения в негодность обследуемых единиц совокупности (например, товаров при обследовании их качества);
- уточнение результатов сплошного наблюдения.

Принципы выборочного наблюдения

- случайность отбора единиц (при отборе каждой из единиц изучаемой совокупности обеспечивается равная возможность попасть в выборку);
- достаточное число отобранных единиц совокупности.

Этапы проведения выборочного наблюдения

Определение цели обследования



Установление границ генеральной совокупности



Составление программы наблюдения и программы разработки данных



Определение вида выборки, процента отбора и метода отбора



Отбор и регистрация наблюдаемых признаков у отобранных единиц



Расчет выборочных характеристик и их ошибок



Распространение полученных результатов на генеральную совокупность



Генеральная совокупность - это совокупность, из которой производится отбор единиц совокупности.

Выборочная совокупность – это совокупность отобранных в определенном порядке единиц, по которым собирается информация.



Доля выборки - это отношение численности выборочной совокупности (n) к численности генеральной совокупности (N).

Доля выборки определяется по формуле

$$k_v = \frac{n}{N}$$



Применяя выборочный метод наблюдения, обычно используют два вида обобщающих показателей:
среднюю величину количественного признака и относительную величину альтернативного признака (долю единиц, обладающую тем или иным признаком).

Обозначения основных характеристик параметров генеральной и выборочной совокупностей

Характеристика	Генеральная совокупность	Выборочная совокупность
Объем совокупности (численность единиц)	N	n
Численность единиц, обладающих исследуемым признаком	M	m
Доля единиц, обладающих исследуемым признаком	$p = \frac{M}{N}$	$w = \frac{m}{n}$
Средний размер признака	$\bar{x} = \frac{\sum x_i}{N}$	$\tilde{x} = \frac{\sum x_i}{n}$
Дисперсия количественного признака	$\sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{N}$	$\sigma_{\tilde{x}}^2 = \frac{\sum (x_i - \tilde{x})^2}{n}$
Дисперсия доли	$\sigma_p^2 = p(1 - p)$	$\sigma_w^2 = w(1 - w)$

Вопрос 2

- Ошибки выборочного наблюдения и определение необходимой численности выборки



При проведении выборочного наблюдения неизбежна некоторая свойственная ему погрешность (ошибки), обусловленная тем, что выборочная совокупность не может по всем параметрам в точности воспроизвести генеральную совокупность. Ошибки, присущие выборочному наблюдению, называются *ошибками репрезентативности*.



Ошибка репрезентативности (представительности) –

**это расхождение между выборочной
характеристикой и характеристикой
генеральной совокупности.**



Существует два вида ошибок репрезентативности:
систематические и случайные.

- **Систематические ошибки** возникают из-за нарушения принципов формирования выборочной совокупности. Они бывают *преднамеренными* и *непреднамеренными*.
- **Случайные ошибки** возникают в результате несплошного характера наблюдения. Случайные ошибки могут быть статистически оценены и учтены при распространении результатов выборочного наблюдения на генеральную совокупность. Оценка случайных ошибок основана на предельных теоремах теории вероятностей. К случайным ошибкам относятся *средняя ошибка выборки* и *предельная ошибка выборки*.

Средняя (стандартная) ошибка
выборки (μ) представляет собой такое
расхождение между средними
выборочной и генеральной
совокупностей ($\bar{x} - \bar{x}$), которое не
превышает $\pm\sigma$.

Средняя ошибка выборки зависит *от объема выборки*
(чем больше численность при прочих равных
условиях, тем меньше величина средней ошибки
выборки) и *степени варьирования признака* (чем
меньше вариация признака, следовательно, и
дисперсия, тем меньше ошибка выборки, и
наоборот).

Величину средней ошибки выборки для количественного признака можно определить по формуле

$$\mu_{\tilde{x}} = \sqrt{\sigma_{\tilde{x}}^2 : n} = \sigma_{\tilde{x}} : \sqrt{n}$$

Величину средней ошибки выборки для альтернативного признака находят по формуле

$$\mu_w = \sqrt{\frac{w \times (1 - w)}{n}}$$

Предельная ошибка выборки (Δ) – максимально возможное расхождение выборочной и генеральной средних ($\tilde{x} - \bar{x}$), т.е. максимум ошибки при заданной вероятности ее появления.



О величине предельной ошибки можно судить с определенной вероятностью, на величину которой указывает **коэффициент доверия (t)** .

Значения вероятности (функция $F(t)$) при различных значениях t определяются на основе специально составленных таблиц.

Табличные значения коэффициента доверия следующие.

t	1,000	1,960	2,000	2,580	3,000
$F(t)$	0,683	0,950	0,954	0,990	0,997

Величина предельной ошибки выборки для количественного признака определяется по формуле

$$\Delta_{\tilde{x}} = \mu_{\tilde{x}} \times t$$

Величина предельной ошибки выборки для альтернативного признака определяется по формуле

$$\Delta_w = \mu_w \times t$$

Предельная ошибка выборки позволяет определять предельные значения характеристик генеральной совокупности при заданной вероятности и их доверительные интервалы.

Интервальную оценку генеральной средней находят по формуле

$$\bar{x} - \Delta_{\bar{x}} \leq \bar{x} \leq \bar{x} + \Delta_{\bar{x}}$$

Интервальную оценку генеральной доли определяют по формуле

$$w - \Delta_w \leq p \leq w + \Delta_w$$

При подготовке выборочного наблюдения с заранее заданным значением допустимой ошибки выборки важно правильно определить объем (численность) выборочной совокупности (n). Согласно одному из принципов выборочного наблюдения объем выборки должен быть достаточным для того, чтобы обеспечить репрезентативность выборки.

Для определения **необходимой численности выборки для средней** используют формулу

$$n = \frac{t^2 \times \sigma_{\tilde{x}}^2}{\Delta_{\tilde{x}}^2}$$

Необходимую численность выборки для доли можно определить по формуле

$$n = \frac{t^2 \times w \times (1 - w)}{\Delta_w^2}$$



Вопрос 3

- Методы, виды и способы отбора выборочных совокупностей



В теории выборочного наблюдения разработаны методы, способы и виды отбора единиц из генеральной совокупности.

По методам отбор делится на *повторный* и *бесповторный*.

При **повторном отборе** каждая единица, отобранная в случайном порядке, после ее обследования возвращается в генеральную совокупность и в последующем отборе может снова попасть в выборку. При таком отборе вероятность попасть в выборку для каждой единицы генеральной совокупности не меняется независимо от числа отобранных единиц.

При **бесповторном отборе** каждая единица, отобранная в случайном порядке, после ее обследования в генеральную совокупность не возвращается. Вероятность попасть в выборку для каждой единицы совокупности увеличивается по мере производства отбора.

Бесповторный отбор дает более точные результаты, чем повторный.



По видам отбор делится на *индивидуальный, групповой и комбинированный*.

При **индивидуальном отборе** в выборочную совокупность отбираются отдельные единицы генеральной совокупности.

При **групповом отборе** в выборочную совокупность отбираются качественно однородные группы или серии изучаемых единиц.

При **комбинированном отборе** сочетаются первый и второй виды отбора.



В статистике существует несколько **способов отбора**:

- собственно случайный;
- механический (систематический);
- типический;
- серийный (гнездовой);
- комбинированный.



Собственно-случайный отбор – это отбор, при котором наблюдению подвергается часть совокупности, отобранная из всей совокупности в случайном порядке, не зависящем ни от последовательности расположения единиц в совокупности, ни от значений их признаков.

Собственно-случайный отбор может быть повторным и бесповторным.

Механический отбор применяется в тех случаях, когда генеральная совокупность каким-либо образом упорядочена, т.е. имеется определенная последовательность в расположении единиц (например, номера домов, списки избирателей).

При проведении механического отбора устанавливается *шаг отсчета* ($h=N/n$), т.е. расстояние между отбираемыми единицами (h – величина, обратная доле выборки), и *начало отсчета* – номер единицы, которая должна быть обследована первой.

Механический отбор всегда бывает *бесповторным*. При этом отборе используются формулы, применяемые при собственно случайном бесповторном отборе.

Данный вид отбора имеет преимущество перед случайным отбором, его не только легче организовать, но при нем единицы выборочной совокупности равномернее распределяются в генеральной совокупности.

Типический отбор представляет собой отбор, при котором генеральная совокупность разбивается на качественно однородные типические группы, затем из каждой группы с помощью собственно-случайной или механической выборки проводится отбор единиц в выборочную совокупность.

Объем выборки из типической группы при отборе, пропорциональном численности единиц типических групп, определяется по формуле

$$n_i = n \times \frac{N_i}{N}$$

где n_i – объем выборки из i -й типической группы;
 N_i – объем i -й типической группы в генеральной совокупности.

Так как в типическую выборку должны попасть представители всех групп, средняя ошибка типической выборки зависит только от средней из внутригрупповых дисперсий, а не от общей дисперсии.



Серийный отбор заключается в собственно случайном либо механическом отборе групп единиц (серий), внутри которых производится сплошное обследование. Единицей отбора при этой выборке является группа или серия, а не отдельная единица генеральной совокупности. При серийном отборе внутри отобранных групп обследуются все без исключения единицы.

Поскольку внутри серий обследуются все без исключения единицы, средняя ошибка выборки при отборе равновеликих серий зависит от величины только *межгрупповой дисперсии*:


$$\text{межгрупповая дисперсия средних } \delta_{\tilde{x}}^2 = \frac{\sum (\tilde{x}_i - \tilde{x})^2}{r},$$

где $\tilde{x}_i$ – средняя i -й серии;

$\tilde{x}$ – средняя по всей выборочной совокупности;

r – число отобранных серий;

$$\text{межгрупповая дисперсия доли } \delta_{\omega}^2 = \frac{\sum (\omega_i - \bar{\omega})^2}{r},$$

где ω_i – доля признака i -й серии;

$\bar{\omega}$ – общая доля признака во всей выборочной совокупности.

При разных способах и методах отбора используются следующие формулы расчета средней ошибки выборки

Способ выборки	Повторный отбор		Бесповторный отбор	
	для средней ($\mu_{\bar{x}}$)	для доли (μ_{ω})	для средней ($\mu_{\bar{x}}$)	для доли (μ_{ω})
Собственно-случайная и механическая	$\frac{\sigma}{\sqrt{n}}$	$\sqrt{\frac{w(1-w)}{n}}$	$\sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$	$\sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$
Типическая (при отборе пропорциональном объему групп)	$\sqrt{\frac{\bar{\sigma}^2}{n}}$	$\sqrt{\frac{w(1-w)}{n}}$	$\sqrt{\frac{\bar{\sigma}^2}{n} \left(1 - \frac{n}{N}\right)}$	$\sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$
Серийная (гнездовая)	$\sqrt{\frac{\mathcal{S}_x^2}{r}}$	$\sqrt{\frac{\mathcal{S}_{\omega}^2}{r}}$	$\sqrt{\frac{\mathcal{S}_x^2}{r} \left(\frac{R-r}{R-1}\right)}$	$\sqrt{\frac{\mathcal{S}_{\omega}^2}{r} \left(\frac{R-r}{R-1}\right)}$

Пример.

При случайном повторном отборе было установлено, что средний вес товара в выборочной совокупности, состоящей из 100 изделий, оказался равным 10 кг при среднем квадратическом отклонении 0,6 кг.

С вероятностью, равной 0,954, необходимо определить, в каких пределах находится средний вес товара в генеральной совокупности.



Вопрос 4

- Распространение результатов выборочного наблюдения на генеральную совокупность



Конечная цель любого выборочного наблюдения – распространение его характеристик на генеральную совокупность. Существуют два **способа распространения данных выборочного наблюдения на генеральную совокупность**: *прямого пересчета и поправочных коэффициентов.*

Способ прямого пересчета применяется в том случае, если цель выборочного наблюдения заключается в определении объема признака генеральной совокупности, когда известна лишь численность ее единиц.

Способ поправочных коэффициентов применяется в тех случаях, когда цель выборочного метода состоит в уточнении результатов сплошного наблюдения.

Пример. Способ прямого пересчета

При выборочном обследовании партии нарезных батонов в 2000 ед. доля нестандартных изделий в выборке составляет: $\omega = 0,1$ (10:100) при установленной с вероятностью $F(t) = 0,954$ предельной ошибке выборки $\Delta_{\omega} = \pm 0,06$.

На основе этих данных доля нестандартных изделий во всей партии составит: $p = 0,1 \pm 0,06$, или от 0,04 до 0,16.

Способом прямого пересчета можно определить пределы абсолютной численности нестандартных изделий во всей партии:

минимальная численность, шт.: $2000 \times 0,04 = 80$;

максимальная численность, шт.: $2000 \times 0,16 = 320$.

Пример. Способ поправочных коэффициентов

По данным годовых отчетов в сельскохозяйственных организациях имеется 127500 голов коров. В целях проверки данных сплошного учета провели контрольные обходы части обследованных предприятий и выявили, что если данные сплошного учета в организациях, попавших в выборку, показали 550 голов коров, то данные выборки в этих же организациях – 560 голов коров. Необходимо определить фактическую численность коров в сельскохозяйственных организациях.

Определим «процент недоучета» (коэффициент) при сплошном наблюдении:

$$\frac{560 - 550}{550} \times 100 = 1,82 \%$$

Количество голов коров необходимо умножить на этот коэффициент:

$$127500 \times 1,0182 = 129821 \text{ голов}$$

Это значит, что при сплошном учете недоучтена 2321 голова коров.