


ВЫБОРОЧНЫЙ МЕТОД И ОЦЕНКА ЕГО РЕЗУЛЬТАТОВ

Занятие 4

Выборочное наблюдение - вид несплошного

наблюдения, при котором отбор подлежащих обследованию единиц наблюдения из генеральной совокупности (population) осуществляется случайно, отобранная часть (выборка) (sample) подвергается обследованию, после чего результаты распространяются на всю исходную совокупность.

- 
- **Основное требование, предъявляемое к формированию выборки – случайность отбора единиц наблюдения из генеральной совокупности, при котором каждой единице наблюдения обеспечивается равная вероятность попадания в выборку (рандомизированный отбор) (randomization).**

Виды выборок:

• *В зависимости от способа отбора единиц наблюдения (от способа организации совокупности):*

- - **случайная:** отбор единиц наблюдения производится непосредственно из генеральной совокупности. Случайность отбора достигается путем применения жеребьевки или использования таблицы случайных чисел.
- Различают бесповторную выборку и повторную (после регистрации единицы вновь возвращаются в генеральную совокупность)

Виды выборок:

- механическая: генеральная
совокупность разбивается на равные части, из которых затем в заранее обусловленном порядке отбирают единицы наблюдения под определенным номером (например, каждую пятую), так, чтобы обеспечить необходимое число наблюдений.

Виды выборок:

- **типологическая (типическая):** генеральная совокупность разбивается на качественно однородные по изучаемому признаку группы, а затем из этих групп производят случайный отбор необходимого числа единиц наблюдения;
- объем выборки в каждой типической группе устанавливается пропорционально ее удельному весу в генеральной совокупности (пропорциональный отбор), а иногда и с учетом вариации в ней изучаемого признака (оптимальный отбор)

Виды выборок:

- **серийная (гнездовая):** отбору подлежат не отдельные единицы наблюдения, а целые их группы (серии или гнезда), в составе которых единицы наблюдения связаны определенным образом: территориально (районы, селения и др.) или организационно (студенческие группы, больницы, предприятия и др.) и которые отбираются из генеральной совокупности по принципу случайного или механического отбора. Внутри серии производится сплошной отбор единиц наблюдения.


Виды выборов:


- *По этапам отбора:*
 - - одноступенчатая
 - - многоступенчатая




Модификации видов выборок:

- - метод направленного отбора
- - метод парных выборок
- - когортный метод
- - метод моментного среза

- 
-
- **Репрезентативность** - это способность выборочной совокупности как количественно, так и качественно отражать свойства генеральной совокупности.

- 
-
- **Количественная репрезентативность** достигается достаточностью числа наблюдений, качественная - соответствием признаков единиц наблюдения в выборочной и генеральной совокупностях.

- 
-
- Любое значение параметра, вычисленное на основе ограниченного числа наблюдений, непременно содержит элемент случайности. Такое приближенное, случайное значение называется **оценкой параметра**.



Как правило проводят

- точечную и
- интервальную оценку параметра.

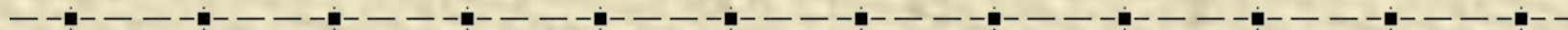


ОЦЕНКА РЕЗУЛЬТАТОВ ВЫБОРОЧНОГО МЕТОДА ИССЛЕДОВАНИЯ

- **1 этап - точечная оценка параметра**
- **2 этап – интервальная оценка параметра**

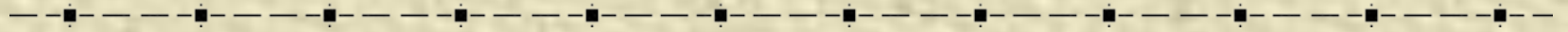
- Точечная оценка параметра выражается в ошибке репрезентативности (**standard error, m**), которая показывает на сколько отличаются обобщающие коэффициенты (показатели), полученные при выборочном исследовании, от тех коэффициентов, которые могли бы быть получены при сплошном исследовании.

- Для количественных признаков:




$$m_x = \pm \frac{\sigma}{\sqrt{n}}$$

- Для номинальных и порядковых признаков:



$$m_p = \pm \sqrt{\frac{pq}{n}}$$

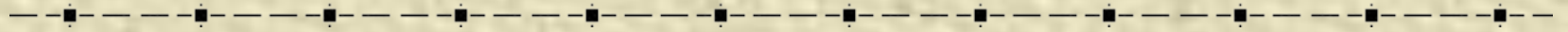
- 
-
- Способы, уменьшающие ошибку репрезентативности:
 - - увеличение числа наблюдений
 - - уменьшение вариабельности признака

Интервальная оценка параметра

-
- На практике определяют пределы возможных ошибок выборки или **предельную ошибку выборки (Δ)**.
 - Т.к. предельная ошибка может быть как в сторону увеличения, так и в сторону уменьшения, то говорят о **доверительном интервале** или **доверительных границах (confidence interval, CD)**, в пределах которых будет находиться показатель генеральной совокупности на основании данных выборочного исследования

- **Предельная ошибка выборки**

$$\Delta = tm_x$$




- **Определение доверительного интервала**


$$x \pm \Delta; x \pm tm;$$


$$X_{\text{выб}} - tm < X_{\text{ген}} < X_{\text{выб}} + tm$$


Существует таблица соответствия интеграла вероятности и значения “t” (таблица Стьюдента).


- **При числе наблюдений больше 30 можно пользоваться следующей закономерностью: вероятности безошибочного прогноза**
- **$p = 0,68$ приблизительно соответствует $t \cong 1$;**
- **при $p = 0,95$ $t \cong 2$,**
- **при $p = 0,99$ $t \cong 3$.**

- 
-
- Способы, уменьшающие ошибку репрезентативности:
 - - увеличение числа наблюдений
 - - уменьшение вариабельности признака

- 
-
- При малом числе наблюдений (не более 30) оперируют значениями t , приведенными в таблице Стьюдента.

- 
-
- Для большинства медицинских исследований допускают $p = 0,05$ или 5%. В этом случае вероятность выхода результата выборочного исследования за границы доверительного интервала, т.е. вероятность ошибки составляет 0,05 или 5%.
 - Поэтому говорят, что результат исследования получен с уровнем значимости 0,05 ($p=0,05$).

- 
-
- Поэтому говорят, что результат исследования получен с уровнем значимости 0,05 ($p=0,05$).

- 
-
- При необходимости более строгой оценки $p=0,99$ (99%), вероятность ошибки составит 0,01 (1%) и следовательно уровень значимости будет $p=0,01$


- **Определение доверительных границ
(доверительного интервала):**

- для количественных признаков

$$\bar{x} \pm \bar{x} \pm t \cdot m_x; \hat{x} - t \cdot m_x \leq \bar{x} \leq \hat{x} + t \cdot m_x;$$

-
- - для номинальных и порядковых признаков

$$\bar{p} = \tilde{p} \pm t \cdot m_p; \quad \tilde{p} - t \cdot m_p \leq \bar{p} \leq \tilde{p} + t \cdot m_p.$$



- **Определение объема выборки**

1. По специальным таблицам с заданной предельной ошибкой (таблица Боярского)

1. По формулам:


- а) для количественных признаков:

$$n = \frac{\sigma^2 \cdot t^2}{\Delta^2};$$

-
- для номинальных и порядковых признаков

$$N = \frac{P \cdot Q \cdot t^2}{\Delta K};$$

- Расчет объема выборки производят на начальных этапах исследования, когда неизвестны параметры распределения и предельная ошибка.
- Источниками для определения σ и Δ служат:
 - - результаты пилотажного исследования;
 - - данные литературы (аналогичные исследования);

- 
-
- Как крайний вариант определения этих параметров для альтернативного распределения можно взять максимально возможное значение $p \times q = 0.5 \times 0.5 = 0,25$, что существенно увеличивает необходимый объем наблюдений.

-
- Для количественных признаков приблизительное значение σ можно определить как

$$\sigma = \frac{X_{max} - X_{min}}{6},$$