Выборочное наблюдение

Понятие выборочного наблюдения

Выборочное наблюдение

это

Такое несплошное наблюдение, при котором статистическому наблюдению подвергаются не все единицы изучаемой совокупности , а лишь отобранные в определенном порядке

Статистическое наблюдение можно организовать как сплошное и несплошное. Сплошное наблюдение предусматривает обследование всех единиц изучаемой совокупности явлений, несплошное – лишь его части.

Цель выборочного наблюдения состоит в том, чтобы по характеристикам отобранной части единиц судить о характеристиках всей совокупности.

Преимущества выборочного наблюдения

- Достижение большей точности результатов обследования благодаря сокращению ошибок регистрации
- Экономия трудовых и денежных средств и времени в результате сокращения объема работы
- Возможность детального обследования каждой единицы наблюдения за счет расширения программы наблюдения
- Сведение к минимуму уничтожения и приведения в негодность обследуемых единиц совокупности
- Уточнение результатов сплошного наблюдения

Научные принципы выборочного наблюдения

Обеспечение случайности отобранных единиц
(при отборе каждой из единиц изучаемой совокупности обеспечивается равная возможность попасть в выборку)

Обеспечение достаточного числа отобранных единиц совокупности

Понятие генеральной и выборочной совокупности

- ***Генеральная совокупность (N)*** – это совокупность, из которой производится отбор единиц совокупности.

- ***Выборочная совокупность (n)*** – совокупность отобранных в определенном порядке единиц, по которым собирается информация.

Доля выборки

Отношение численности выборочной совокупности к численности генеральной совокупности

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

Генеральная средняя

Среднее значение признака всей совокупности

$$\bar{x} = \frac{\sum x_i}{N}$$

Выборочная средняя

Среднее значение признака у единиц, которые подверглись выборочному наблюдению

$$\widetilde{x} = \frac{\sum x_i}{N}$$

Генеральная доля

Доля единиц, обладающих тем или иным признаком в генеральной совокупности

$$p = \frac{M}{N}$$

где *M* – численность единиц, обладающих определенным признаком в генеральной совокупности.

Выборочная доля или частность

Доля единиц, обладающих тем или иным признаком в выборочной совокупности

$$\omega = \frac{m}{n}$$

где m – численность единиц, обладающих определенным признаком в выборочной совокупности.

Дисперсия количественного признака в генеральной совокупности

$$\sigma_{\bar{x}}^{2} = \frac{\sum\left(x_i - \bar{x}\right)^2}{N}$$

Дисперсия количественного признака в выборочной совокупности

$$\sigma_{\tilde{x}}^{2} = \frac{\sum \left(x_i - \tilde{x} \right)^2}{n}$$

Дисперсия доли признака в генеральной совокупности

$$\sigma_p^2 = p \cdot (1 - p)$$

Дисперсия доли признака в выборочной совокупности

$$\sigma_{\omega}^{2} = \omega \cdot (1 - \omega)$$

Понятие ошибки репрезентативности

Ошибка репрезента-тивности

- это

Расхождение между выборочной характеристикой и характеристикой генеральной совокупности

Ошибки репрезентативности

Систематические
(возникают в результате нарушения научных принципов отбора единиц совокупности)

- ***Преднамеренные***
- ***Непреднамеренные***

Случайные
(возникают в результате несплошного характера наблюдения)

- ***Средняя (стандартная) ошибка выборки***
- ***Предельная ошибка выборки***

Понятие средней (стандартной) ошибки выборки

Средняя ошибка выборки

- это

Такое расхождение между средними выборочной и генеральной совокупностями ($\tilde{x} - \bar{x}$), 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)

- **степени варьирования признака** (чем меньше вариация признака, а следовательно, и дисперсия, тем меньше ошибка выборки, и наоборот)

Формула для определения величины средней ошибки выборки для количественного признака

Средняя ошибка выборки - это

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma_{\tilde{x}}^{2}}{n}} = \frac{\sigma_{\tilde{x}}}{n} \quad (1)$$

Формула для определения величины средней ошибки выборки для альтернативного признака

Средняя ошибка выборки

- это

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}} \quad (2)$$

Понятие предельной ошибки выборки

Предель-ная ошибка выборки

- это

Максимально возможное расхождение выборочной и генеральной средних ($\tilde{x} - \bar{x}$), т.е максимум ошибки при заданной вероятности ее проявления

О величине предельной ошибки можно судить с определенной вероятностью, на величину которой указывает ***коэффициент доверия t***. Табличное значение коэффициента следующее:

t	1,0	1,96	2,0	2,58	3,0
F(*t*)	0,683	0,950	0,954	0,990	0,997

Формула для определения величины предельной ошибки выборки

Предель-
ная
ошибка
выборки

-
эт
о

$\Delta_{\tilde{x}} = \mu_{\tilde{x}} \cdot t$, или $\Delta_{\tilde{x}} = \mu_{\tilde{x}} \cdot t$

где Δ – предельная ошибка выборки;

t – коэффициент доверия, зависящий от вероятности, с которой гарантируется предельная ошибка выборки

Формула для определения интервальной оценки генеральной средней

Интервальная оценка

$$\widetilde{x} - \Delta_{\bar{x}} \le \bar{x} \le \widetilde{x} + \Delta_{\widetilde{x}}$$

Формула для определения интервальной оценки генеральной доли

Интервальная оценка

$$\omega - \Delta_{\omega} \le p \le \omega + \Delta_{\omega}$$

Формула для определения необходимой численности выборки для средней

Числен-
ность
выборки

-
эт
о

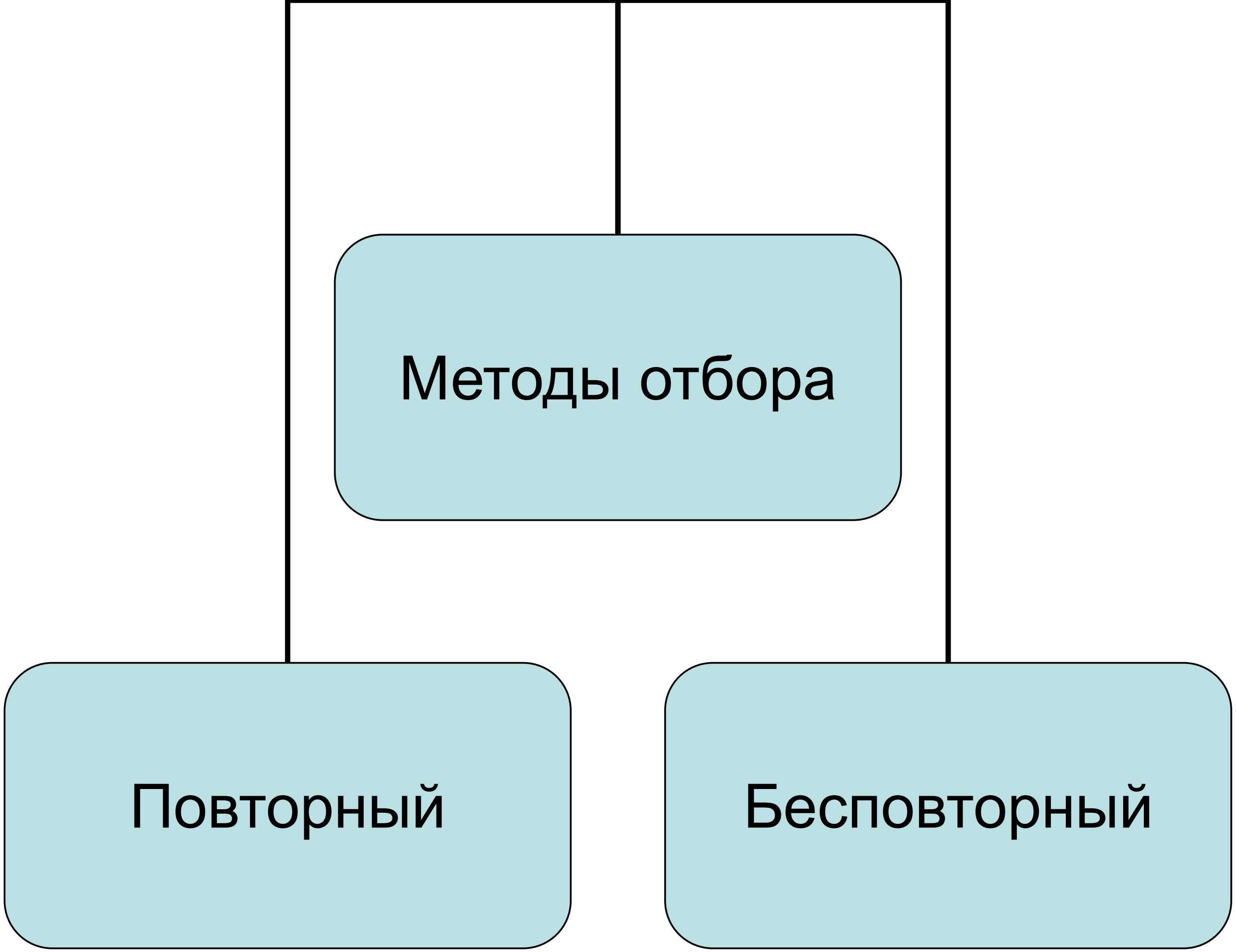
$$n = \frac{t^2 \cdot \sigma_{\tilde{x}}^2}{\Delta_{\tilde{x}}}$$

Формула для определения необходимой численности выборки для доли

Числен-
ность
выборки

-
эт
о

$$n = \frac{\omega \cdot (1-\omega)}{\Delta_{\omega}^{2}}$$


Методы отбора
Повторный
Бесповторный

Повторный отбор

- Каждая единица, отобранная в случайном порядке, после 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 не зависимо от числа отобранных единиц

Бесповторный отбор

- Каждая единица, отобранная в случайном порядке, после обследования в генеральную совокупность не возвращается.

Вероятность попасть в выборку для каждой единицы генеральной совокупности увеличивается по мере производства отбора

Виды отбора

- *Индивидуальный* (в выборочную совокупность отбираются отдельные единицы генеральной совокупности)
- *Групповой* (в выборочную совокупность отбираются качественно однородные группы или серии изучаемых единиц)
- *Комбинирован-ный* (происходит сочетание первого и второго видов отбора)

Способы отбора

- Собственно-случайный
- Механический
- Типический
- Серийный
- Комбинированный

Понятие собственно-случайного отбора

Собственно-случайный отбор - это отбор, при котором наблюдению подвергается часть совокупности, отобранная из всей совокупности в случайном порядке

Основные формулы, используемые при собственно-случайном отборе

Средняя ошибка выборки для средней

- Повторная выборка

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

- Бесповторная выборка

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma_{\tilde{x}}^{2}}{2}\left(1 - \frac{n}{N}\right)}$$

Основные формулы, используемые при собственно-случайном отборе

Средняя ошибка выборки для доли

- Повторная выборка

$$\mu_{\omega} = \sqrt{\frac{\omega \cdot (1-\omega)}{n}}$$

- Бесповторная выборка

$$\mu_{\omega} = \sqrt{\frac{\omega \cdot (1-\omega)}{n} \cdot \left(1 - \frac{n}{N}\right)}$$

Основные формулы, используемые при собственно-случайном отборе

Численность выборки при определении среднего размера признака

- Повторная выборка

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

- Бесповторная выборка

$$n = \frac{t^2 \cdot \sigma_{\tilde{x}}^2 \cdot N}{N \cdot \Delta_{\tilde{x}}^2 + t^2 \cdot \sigma_{\tilde{x}}^2}$$

Основные формулы, используемые при собственно-случайном отборе

Численность выборки при определении доли признака

- Повторная выборка

$$n = \frac{t^2 \cdot \omega \cdot (1 - \omega)}{\Delta_\omega^2}$$

- Бесповторная выборка

$$n = \frac{t^2 \cdot \omega \cdot (1 - \omega) \cdot N}{N \cdot \Delta_\omega^2 + t^2 \cdot \omega \cdot (1 - \omega)}$$

Понятие механического отбора

Механичес-кий отбор

- это

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

При механическом отборе устанавливается ***шаг отсчета***, т.е. расстояние между отбираемыми единицами (*N/n* – величина, обратная доле выборки) и ***начала отсчета*** – номер единиц, которая должна быть обследована первой.

Механический отбор всегда бывает бесповторным. При этом отборе применяются те же формул, что и при собственно-случайном бесповторном отборе.

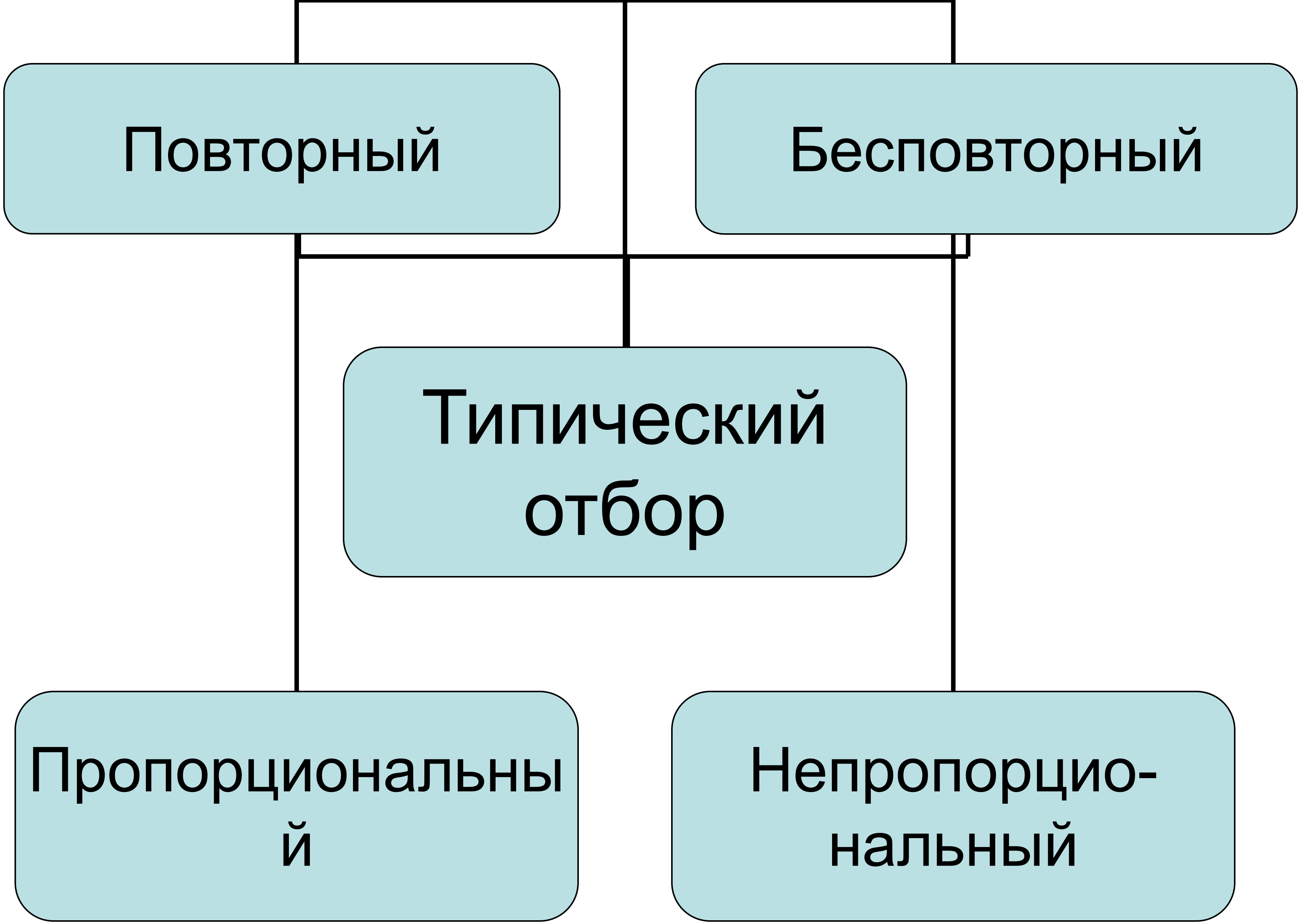
Механический отбор имеет преимущество перед случайным отбором, его не только легче организовать, но при нем единицы выборочной совокупности равномернее распределяются в генеральной совокупности.

Понятие типического отбора

Типический отбор

- это

отбор, при котором генеральная совокупность разбивается на качественно однородные типические группы, затем из каждой группы при помощи собственно-случайной или механической выборки проводится отбор единиц в выборочную совокупность


Повторный
Бесповторный
Типический отбор
Пропорциональный
Непропорцио-нальный

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

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

где n_i – объем выборки из i-й типической группы;

N_i – объем i-й типической группы в генеральной совокупности

Разбивка на типические группы дает возможность избежать влияния межгрупповой вариации на точность выборки.

Так как в типическую выборку должны попасть представители всех групп, средняя ошибка типической выборки зависит только от *средней из внутригрупповых дисперсий* $\overline{\sigma}_i^2$, или $\overline{\omega\cdot(1-\omega)}$, а не от общей дисперсии $\overline{\sigma}_x^2$, или

· $\omega\cdot(1-\omega)$

Основные формулы, используемые при типическом отборе

Средняя ошибка выборки для средней

- Повторная выборка

$$\mu_{\tilde{x}} = \sqrt{\frac{\overline{\sigma}_i^{\,2}}{n}}$$

- Бесповторная выборка

$$\mu_{\tilde{x}} = \sqrt{\frac{\overline{\sigma}_i^{\,2}}{2}\left(1 - \frac{n}{N}\right)}$$

Основные формулы, используемые при типическом отборе

Средняя ошибка выборки для доли

- Повторная выборка

$$\mu_{\omega} = \sqrt{\frac{\omega \cdot (1 - \omega)}{n}}$$

- Бесповторная выборка

$$\mu_{\omega} = \sqrt{\frac{\omega \cdot (1 - \omega)}{n} \cdot \left(1 - \frac{n}{N}\right)}$$

Основные формулы, используемые при типическом отборе

Численность выборки при определении среднего размера признака

- Повторная выборка

$$n = \frac{t^2 \cdot \overline{\sigma}_i^2}{\Delta_{\tilde{x}}}$$

- Бесповторная выборка

$$n = \frac{t^2 \cdot \overline{\sigma}_i^2 \cdot N}{N \cdot \Delta_{\tilde{x}}^2 + t^2 \cdot \overline{\sigma}_i^2}$$

Основные формулы, используемые при типическом отборе

Численность выборки при определении доли признака

- Повторная выборка

$$n = \frac{t^2 \cdot \overline{\omega \cdot (1-\omega)}}{\Delta_\omega^2}$$

- Бесповторная выборка

$$n = \frac{t^2 \cdot \overline{\omega \cdot (1-\omega)} \cdot N}{N \cdot \Delta_\omega^2 + t^2 \cdot \omega \cdot (1-\omega)}$$

Понятие серийного отбора

Серийный отбор

-

это

такой отбор, когда в случайном порядке отбираются не единицы, подлежащие обследованию, а группы единиц (серии, гнезда).

Внутри отобранных серий обследованию подвергаются все единицы, т.е. применяется сплошное наблюдение.

Обозначения

R – общее число серий;

r – число отобранных серий;

$\delta^2_{\tilde{x}}$ - межгрупповая дисперсия средних, определяемая по формуле

$$\delta^2_{\tilde{x}} = \frac{\sum\left(\tilde{x}_i - \tilde{x}\right)^2}{r}$$

$\tilde{x}_i$ - средняя i –ой серии;

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

δ^2_{ω} - межгрупповая дисперсия доли, определяемая по формуле

$$\delta^2_{\omega} = \frac{\sum\left(\omega - \overline{\omega}\right)^2}{n}$$

ω_i - доля признака i –ой серии;

$\overline{\omega}$ - общая доля по всей выборочной совокупности.

Основные формулы, используемые при серийном отборе

Средняя ошибка выборки для средней

- Повторная выборка

$$\mu_{\tilde{x}} = \sqrt{\frac{\delta_{\tilde{x}}^{2}}{r}}$$

- Бесповторная выборка

$$\mu_{\tilde{x}} = \sqrt{\frac{\delta_{\tilde{x}}^{2}}{r}\left(1 - \frac{r}{R}\right)}$$

Основные формулы, используемые при серийном отборе

Средняя ошибка выборки для доли

- Повторная выборка

$$\mu_{\omega} = \sqrt{\frac{\delta_{\omega}^{2}}{r}}$$

- Бесповторная выборка

$$\mu_{\omega} = \sqrt{\frac{\delta_{\omega}^{2}}{r}} \cdot \left(1 - \frac{r}{R}\right)$$

Основные формулы, используемые при серийном отборе

Численность выборки при определении среднего размера признака

- Повторная выборка

$$r = \frac{t^2 \cdot \delta_{\tilde{x}}^2}{\Delta_{\tilde{x}}^2}$$

- Бесповторная выборка

$$r = \frac{t^2 \cdot \delta_{\tilde{x}}^2 \cdot R}{R \cdot \Delta_{\tilde{x}}^2 + t^2 \cdot \delta_{\tilde{x}}^2}$$

Основные формулы, используемые при типическом отборе

Численность выборки при определении доли признака

- Повторная выборка

$$r = \frac{t^2 \cdot \delta_{\omega}^2}{\Delta_{\omega}^2}$$

- Бесповторная выборка

$$r = \frac{t^2 \cdot \delta_{\omega}^2 \cdot R}{R \cdot \Delta_{\omega}^2 + t^2 \cdot \delta_{\omega}^2}$$

Способы распространения данных выборочного наблюдения на генеральную совокупность

Способ прямого пересчета

Применяется, если целью выборочного наблюдения является определение объема признака генеральной совокупности, когда известна лишь численность ее единиц

Способ поправочных коэффициентов

Применяется в тех случаях, когда целью выборочного метода является уточнение результатов сплошного наблюдения

Понятие малой выборки

Малая выборка

- это

Несплошное статистическое обследование, численность единиц которого не превышает 30

Для определенного способа отбора единиц величина стандартной ошибки зависит от объема выборки и степени колеблемости изучаемого признака в генеральной совокупности.

Чем меньше объем выборки, тем большую величину стандартной ошибки следует ожидать, а это снижает точность оценки параметров генеральной совокупности.

Для оценки возможных пределов ошибки малой выборки применяется отношение Стьюдента:

$$t = \frac{\widetilde{x} - \bar{x}}{\mu_{\grave{I}\ .\hat{A}}}$$

где $\mu_{\grave{I}\ .\hat{A}}$ - величина среднего квадратического отклонения малой выборки:

$$\mu_{\grave{I}\ .\hat{A}} = \sqrt{\frac{\sigma^2}{n-1}} = \frac{\sigma}{\sqrt{n-1}}$$

Величина σ вычисляется на основе данных выборочного наблюдения:

$$\sigma = \sqrt{\frac{\sum (x_i - \widetilde{x})^2}{n}}$$

Таким образом, для теоретического распределения отношения Стьюдента *t* имеются величины, определяемые непосредственно по данным выборки.

Для определенных значений t и n доверительную вероятность малой выборки находят по специальным таблицам Стьюдента.

Предельная ошибка малой выборки:

$$\Delta_{M.B} = t \cdot \mu_{M.B}$$