

ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

Понятие о выборочном наблюдении

По охвату единиц изучаемой совокупности статистическое наблюдение подразделяется на сплошное и несплошное. *Сплошным* является такой вид наблюдения, при котором учету подвергаются все единицы изучаемой совокупности. *Несплошным* называется наблюдение, при котором учету подвергаются не все, а только часть единиц изучаемой совокупности, но часть эта должна быть достаточно массовой, чтобы обеспечить получение обобщающихся статистических показателей.

Несплошное наблюдение имеет ряд преимуществ перед сплошным: во-первых, оно требует меньших материальных и трудовых затрат; во-вторых, позволяет применять более совершенные способы учета; в-третьих, повышает оперативное значение статистических данных, так как проводится в более короткие сроки. Причем в некоторых случаях несплошное наблюдение является единственно возможным, если процесс наблюдения влечет за собой порчу или уничтожение наблюдаемых единиц.

Выборочное наблюдение является основной формой несплошного наблюдения. Сущность его состоит в том, что наблюдению подвергается часть единиц совокупности определенного вида. При этом данные, полученные на основании этой части, характеризуют всю совокупность. При выборочном наблюдении обеспечивается возможность всем единицам совокупности быть отобранными для проведения наблюдения. Совокупность единиц, из которых производится отбор, называется *генеральной*. Количество единиц, отобранных из генеральной совокупности для проведения выборочного наблюдения, составляет выборочную совокупность.

Результаты выборочного наблюдения не имеют самостоятельного значения. Они интересны только потому, что их можно распространить на всю генеральную совокупность, т.е. по ним можно установить обобщающие характеристики генеральной совокупности, не производя сплошного наблюдения. Так как сплошное наблюдение заменяется выборочным, то полученные средние также являются выборочными и обозначаются $\bar{x}$, поэтому они могут не совпадать со средними генеральными характеристиками, обозначаемыми $\bar{X}$.

Цель выборочного наблюдения – установить, с какой величиной отклоняется значение выборочной средней от средней генеральной, т.е. какова ошибка выборочного наблюдения. Эти ошибки называются *ошибками репрезентативности* или *представительности*. Ошибки выборочного наблюдения возникают потому, что обследуется не вся совокупность, а какая-то ее часть, притом эта часть отобрана случайно. Чем меньше величина отклонения, или ошибки, тем точнее выборочная средняя воспроизводит среднюю генеральную.

Ошибки выборочного наблюдения

Существуют средняя и предельная ошибки выборки, которые исчисляются для количественных и качественных признаков. Наиболее частой ошибкой является отождествление средней ошибки выборочной средней (количественный признак) и средней ошибки выборочной доли (качественный признак).

Средняя ошибка выборочной средней обозначается μ_x (мю) и определяется по вариации количественного признака ($x_1, x_2, \dots, x_n$) по следующим формулам:

- для повторного отбора: $\mu_x = \sqrt{\frac{\sigma_x^2}{n}}$;
- для бесповторного отбора: $\mu_x = \sqrt{\frac{\sigma_x^2}{n} \left(1 - \frac{n}{N}\right)}$;

где σ_x^2 - дисперсия признака; n - численность выборочной совокупности; N - численность генеральной совокупности.

Средняя ошибка выборочной доли определяется по показателям качественного или альтернативного признака. *Альтернативным* называется признак, которым обладают одни единицы совокупности и не обладают другие. Если n обозначает число единиц всей выборочной совокупности,

а m – число единиц из этой совокупности, обладающих определенным признаком, то их отношение

$\omega = \frac{m}{n}$ будет обозначать *выборочную долю, или частоту*.

Средняя ошибка выборочной доли определяется по следующим формулам:

для бесповторного отбора: $\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}}$; ра:

для повторного отбора: $\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n} \left(1 - \frac{n}{N}\right)}$.

По способу отбора единиц в выборочную совокупность выборка бывает повторной и бесповторной. *Повторной* называется выборка, при которой каждая отобранная единица возвращается в генеральную совокупность для последующего отбора и может повторно попасть в выборку. При этом численность генеральной совокупности остается неизменной. Обычно выборочное наблюдение проводится способом *бесповторного* отбора, при котором единица, попавшая в выборку, не возвращается в генеральную совокупность и дальнейший отбор производится без отобранных ранее единиц. При этом численность генеральной совокупности уменьшается на величину выборочной совокупности. Поэтому в формулы средней ошибки выборки для бесповторного

отбора вводится дополнительный множитель $\left(1 - \frac{n}{N}\right)$, который всегда меньше единицы. Отношение $\frac{n}{N}$ представляет собой относительную долю отобранных единиц, т.е. *процент выборки*.

Главным образом μ зависит от n и в малой степени от процента выборки $\frac{n}{N}$. Если это отношение

менее 5%, то поправкой $\left(1 - \frac{n}{N}\right)$ можно пренебречь и ошибка выборки находится по способу повторного отбора, даже если сама выборка была бесповторной.

Конечным этапом выборочного исследования является необходимость установления пределов с определенной степенью вероятности, за которые не выйдет величина ошибки выборочного наблюдения. Ошибка выборки, исчисленная с заданной степенью вероятности, называется *предельной ошибкой выборки*, обозначается знаком Δ (дельта) и определяется по общей формуле как для количественных, так и для качественных признаков: $\Delta = \mu_x t$.

Для количественного признака предельная ошибка выборки определяется по формулам:

- для повторного отбора: $\Delta_x = t \sqrt{\frac{\sigma^2}{n}}$;
- для бесповторного отбора: $\Delta_x = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$.

Для качественного признака предельная ошибка выборки определяется по формулам:

- для повторного отбора: $\Delta_{\omega} = t \sqrt{\frac{\omega(1-\omega)}{n}}$;
- для бесповторного отбора: $\Delta_{\omega} = t \sqrt{\frac{\omega(1-\omega)}{n} \left(1 - \frac{n}{N}\right)}$.

Способы распространения выборочных данных на генеральную совокупность

Сущность выборочного наблюдения для количественного признака раскрывается в следую-

щей формуле: $\bar{x} = \tilde{x} \pm \Delta_x$, которая определяет пределы генеральной средней величины $\bar{x}$. Для этого от величины средней выборочной $\tilde{x}$ отнимают и прибавляют величину предельной ошибки выборки. Генеральная средняя - это пределы средней величины какого-либо варьирующего признака, исчисляемой для всей генеральной совокупности. Средняя выборочная $\tilde{x}$ - это средняя величина этого же признака, исчисленная по выборочной совокупности.

Пределы генеральной доли качественного признака определяются по формуле $P = \omega \pm \Delta_{\omega}$, где P - генеральная доля или доля (удельный вес) какого-либо качественного признака.

По этой формуле определяются ее пределы P путем вычитания и прибавления к выборочной доле ω предельной ошибки выборки.

В практике в формулы предельной ошибки выборки как для качественных признаков, так и количественных подставляется не значение вероятностей, а соответствующий ему коэффициент доверия t . Значения t определяются по "Таблице значений вероятностей, вычисленных для различных значений t ". При вероятности, равной 0,954, значение t равно 2; при вероятности, равной 0,997, значение t равно 3.

Сущность выборочного наблюдения и его ошибки представлены на схеме 7.1.

Практика применения выборочного наблюдения

Пример 1. Продолжим анализировать пример 1 предыдущей лекции.

Для изучения естественной убыли произведено 5%-ное выборочное обследование партии хранящихся на базе товаров. В результате лабораторного анализа установлено следующее распределение полученных посредством механической выборки образцов (табл.1).

Таблица 1

Процент естественной убыли, x	Количество образцов, f
до 4	6
4-6	14
6-8	22
8-10	48
от 10 и выше	10
Итого	100

На основе показателей выборочной совокупности для всей партии товара, т.е. генеральной совокупности, определите:

- 1) с вероятностью 0,954 возможные пределы доли продукции с естественной убылью от 10% и выше, т.е. размер нестандартной продукции;
- 2) с вероятностью 0,997 возможные пределы среднего процента естественной убыли.

Указание. Для определения средней ошибки выборочной средней по данным ряда распределения были исчислены средний процент естественной убыли в выборке и средний квадрат отклонения или дисперсия признака (см. пример 6.1 и 6.4).

Решение

1. Определим численность генеральной совокупности. По условию примера выборочная совокупность (n) состоит из 100 образцов, что составляет 5% генеральной совокупности N . Составим пропорцию:

$$\begin{array}{ll} 100 \text{ образцов} & - 5\%, \\ N & - 100\%. \end{array}$$

$$\text{Отсюда } N = \frac{100 \cdot 100}{5} = 2000 \text{ образцов. Значение альтернативного признака } m=10. \text{ Значит, } 10$$

образцов имеют естественную убыль от 10% и выше, что соответствует нестандартной продукции. Численность выборочной совокупности $n=100$ образцам. Численность генеральной совокупности $N=2000$ образцов.

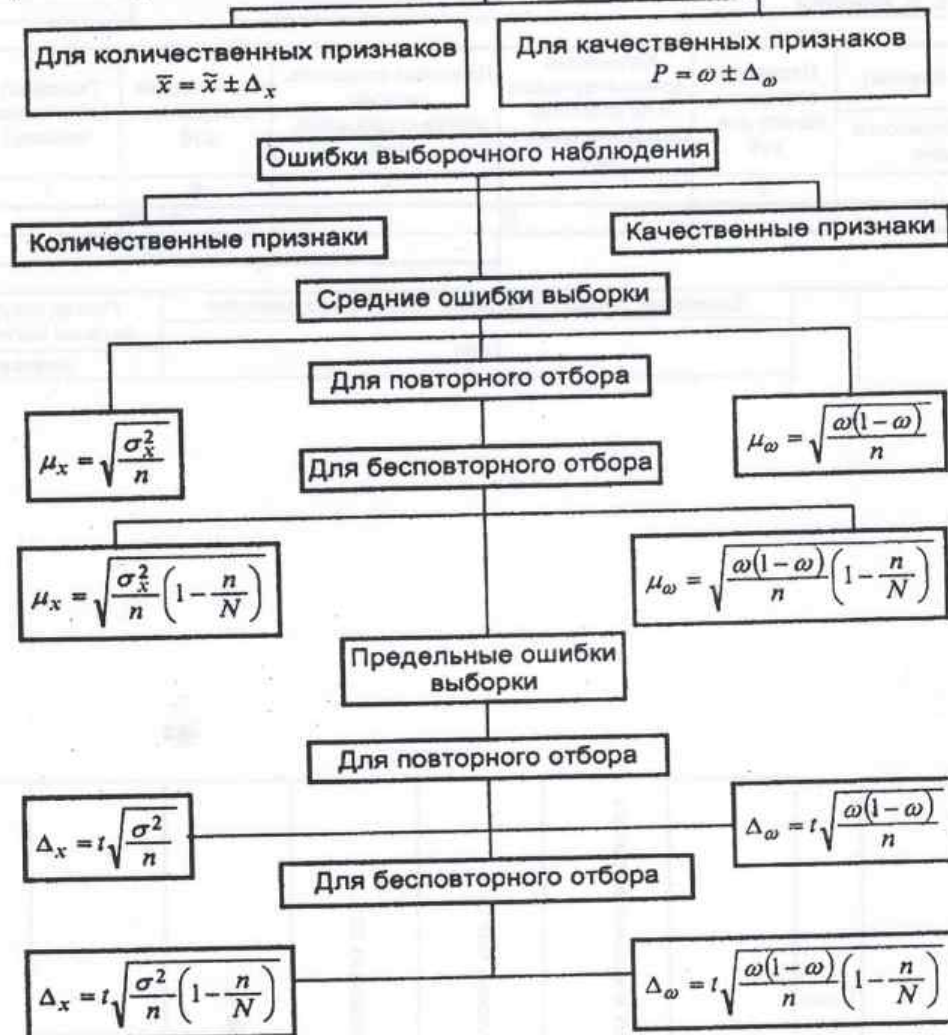
С вероятностью 0,954, которой соответствует коэффициент доверия $t=2$, необходимо определить пределы доли продукции с естественной убылью от 10% и выше, т.е. нестандартной продукции, в генеральной совокупности.

Схема 7.1

Выборочное наблюдение

Цель выборочного наблюдения – установить отклонения выборочных средних от средних генеральных – для количественных признаков и выборочной доли от генеральной доли – для качественных признаков, т.е. определить ошибки репрезентативности – ошибки выборочного наблюдения.

Сущность выборочного наблюдения – определить пределы генеральной средней величины ($\bar{x}$) или генеральной доли (P).



Итак, для ответа на первое задание примера запишем, что дано в условии:

$m=10$ образцов - численность нестандартной продукции;

$n=100$ образцов - численность выборочной совокупности;

$N=2000$ образцов - численность генеральной совокупности;

$t=2$ - коэффициент доверия, соответствующий вероятности 0,954

$\Delta_\omega = ?$

Так как необходимо определить пределы доли качественного признака, используем соответ-

ствующие формулы: $P = \omega \pm \Delta_{\omega}$.

Генеральная доля равна выборочной доле плюс-минус предельная ошибка выборки. Выборочная доля $\omega = \frac{m}{n} = \frac{10}{100} = 0,1$.

Если коэффициент 0,1 умножим на 100%, то получим 10%, т.е. в выборочной совокупности количество образцов, соответствующих нестандартной продукции (с естественной убылью от 10% и выше), составляет 10%. По формуле предельной ошибки выборки определим ее значение. Причем, если при исследовании применялся механический отбор, то для расчетов используется формула для бесповторного отбора:

$$t \sqrt{\frac{\omega(1-\omega)}{n} \left(1 - \frac{n}{N}\right)} = 2 \sqrt{\frac{0,1 \cdot 0,9}{100} \left(1 - \frac{100}{2000}\right)} = 2 \sqrt{0,00085} = 2 \cdot (\pm 0,029) = \pm 0,058.$$

Полученные значения ω и Δ_{ω} подставляем в формулу $P = \omega \pm \Delta_{\omega}$, и получаем: $0,1 \pm 0,058$ или результаты выразим в процентах: $10\% \pm 5,8\%$.

Следовательно, с вероятностью 0,954 продукция с естественной убылью от 10% и выше, т.е. нестандартная, в генеральной совокупности будет составлять от 4,2 до 15,8%.

2. Воспользуемся расчетными данными $\bar{x}$ и σ^2 , полученными на основании решения примера 1 (пред. лекция).

По условию нашего примера:

$n=100$ образцов - численность выборочной совокупности;

$N=2000$ образцов - численность генеральной совокупности;

$\bar{x} = 7,84\%$ - средний процент естественной убыли в выборочной совокупности;

$\sigma^2 = 4,36\%$ - дисперсия процента естественной убыли в выборочной совокупности;

$t=3$ - коэффициент доверия, соответствующий вероятности 0,997

$$\Delta_x = ?$$

Полученные значения $\bar{x}$ и σ^2 подставим в формулу предельной ошибки выборки для количественных признаков:

$$\Delta_x = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} = 3 \sqrt{\frac{4,36}{100} \left(1 - \frac{100}{2000}\right)} = 3 \sqrt{0,0436 \cdot 0,95} = 3 \sqrt{0,0414} = 3 \cdot (\pm 0,2) = \pm 0,6\%$$

Затем значения $\bar{x}$ и Δ_x подставим в формулу: $\bar{x} = \bar{x} \pm \Delta_x$ и получим $7,84\% \pm 0,6\%$. Следовательно, с вероятностью 0,997 можно утверждать что средний процент естественной убыли в генеральной совокупности будет заключаться в пределах от 7,24 до 8,44%.

Задача выборочного исследования решена. Мы определили средние характеристики в генеральной совокупности ($\bar{x}$) по результатам исследования выборочной совокупности ($\bar{x}, \sigma^2$), не производя сплошного наблюдения и используя соответствующие формулы ошибок выборочного наблюдения.