

СТАТИСТИЧЕСКИЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

ВОПРОСЫ

- 1. Точечные оценки и их свойства.**
- 2. Интервальные оценки.**

1. Точечные оценки и их свойства

Важной задачей выборочного метода является оценка параметров (характеристик) генеральной совокупности по данным выборки.

Параметры генеральной совокупности можно оценивать двумя видами оценок:

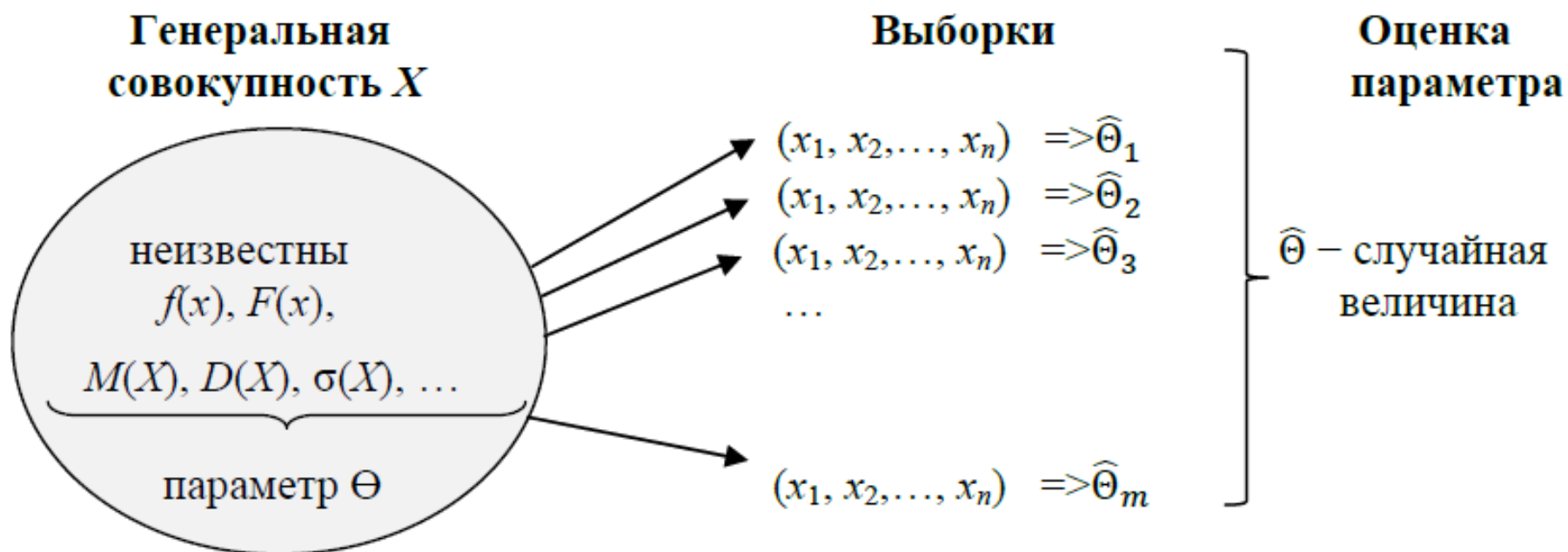
- точечные оценки;
- интервальные оценки.

Пусть Θ – оцениваемый параметр генеральной совокупности.

Определение. *Точечной оценкой $\tilde{\Theta}$*

неизвестного параметра Θ называется полученное по выборке приближенное значение этого параметра, выраженное одним числом.

Оценка $\hat{\Theta}$ есть функция элементов выборки, так как для разных выборок из одной и той же генеральной совокупности получают разные значения оценки, таким образом, *оценка сама является случайной величиной*.



Качество оценки, то есть ее близость к истинному значению параметра, характеризуется следующими свойствами.

а) Несмещенность

Определение. Оценка $\tilde{\Theta}$ параметра Θ называется **несмещенной**, если ее математическое ожидание равно оцениваемому параметру, т.е.

$$M(\tilde{\Theta}) = \Theta$$

В противном случае оценка называется смещенной.

Если это равенство не выполняется, то оценка, полученная по разным выборкам, будет в среднем либо завышать значение Θ (если $M(\tilde{\Theta}) > 0$), либо занижать его (если $M(\tilde{\Theta}) < 0$).

Таким образом, *требование несмещенности гарантирует отсутствие систематических ошибок при оценивании.*

б) Состоятельность

Определение. Оценка $\tilde{\Theta}$ называется **состоятельной** оценкой параметра Θ , если при бесконечном увеличении объема выборки оценка стремится к самому параметру.

$$\tilde{\Theta} \xrightarrow{n \rightarrow \infty} \Theta$$

в) Эффективность

Определение. Несмещенная оценка $\tilde{\Theta}$ параметра Θ называется эффективной, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра Θ , вычисленных по выборкам одного и того же объема n .

Оценка математического ожидания и дисперсии

Обозначим:

x_i – значения признака (случайной величины X);

N и n – объемы генеральной и выборочной совокупностей;

N_i и n_i – число элементов генеральной и выборочной совокупностей со значением признака x_i .

Средние арифметические распределения признака в генеральной и выборочной совокупности называются соответственно *генеральной* и *выборочной средними*, а дисперсии этих распределений – *генеральной* и *выборочной дисперсиями*.

Наименование характеристики	Генеральная совокупность	Выборка
Средняя	$\bar{x}_0 = \frac{\sum_{i=1}^m x_i N_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^m x_i n_i}{n}$
Дисперсия	$\sigma^2 = \frac{\sum_{i=1}^m (x_i - \bar{x}_0)^2 N_i}{N}$	$s^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 n_i}{n}$

В качестве оценки математического ожидания генеральной совокупности возьмем выборочное среднее:

$$\bar{x} = \frac{\sum_{i=1}^m x_i n_i}{n}$$

Выборочное среднее является состоятельной, несмещённой и эффективной оценкой математического ожидания генеральной совокупности.

В качестве оценки дисперсии генеральной совокупности может служить выборочная дисперсия:

$$s^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 n_i}{n}$$

Выборочная дисперсия является состоятельной и смещённой оценкой дисперсии генеральной совокупности с недостатком.

Смещение можно устранить, умножив выборочную дисперсию на $n / (n - 1)$:

$$\hat{s}^2 = \frac{n}{n-1} s^2$$

Эта величина называется *исправленной выборочной дисперсией*.

Исправленная выборочная дисперсия является несмещённой и состоятельной оценкой генеральной дисперсии.

2. Интервальные оценки

Определение. *Интервальной оценкой параметра Θ называется числовой интервал (Θ_1, Θ_2) , который с заданной вероятностью γ накрывает неизвестное значение параметра Θ .*



Интервальная оценка параметра Θ определяется двумя числами, являющимися границами интервала, на который параметр Θ попадает с заданной довольно высокой вероятностью.

Интервал (Θ_1, Θ_2) , называется доверительным, а вероятность γ — доверительной вероятностью или надежностью оценки.

Доверительную вероятность γ обычно задают значениями близкими к единице: 0,9, 0,95 или 0,99.

Величина доверительного интервала существенно зависит:

- от объема выборки n (уменьшается с ростом n),
- от значения доверительной вероятности γ (увеличивается с приближением γ к единице).

Выборочные распределения отдельных оценок симметричны относительно параметра Θ , поэтому целесообразно рассматривать доверительный интервал симметричным относительно параметра Θ , т.е. $(\Theta - \Delta; \Theta + \Delta)$.

Наибольшее отклонение Δ несмещенной оценки от оцениваемого параметра Θ , которое возможно с заданной доверительной вероятностью γ , называется *предельной ошибкой выборки*.

Построение доверительного интервала для генеральной средней по большим выборкам

Теорема. Вероятность того, что отклонение выборочной средней от генеральной средней не превзойдет число $\Delta > 0$ (по абсолютной величине), равна:

$$P(|\bar{x} - \bar{x}_0| \leq \Delta) = \Phi(t) = \gamma,$$

$$\text{где } t = \frac{\Delta}{\sigma_{\bar{x}}},$$

$\Phi(t)$ - функция Лапласа.

Среднее квадратическое отклонение выборочной средней $\sigma_{\bar{x}}$ называется *средней квадратической (стандартной) ошибкой выборки*.

Формулы средних квадратических ошибок выборки

повторная выборка	бесповторная выборка
$\sigma_{\bar{x}} \approx \sqrt{\frac{s^2}{n}}$	$\sigma_{\bar{x}} \approx \sqrt{\frac{s^2}{n} \left(1 - \frac{n}{N}\right)}$

Следствие 1. При заданной доверительной вероятности γ предельная ошибка выборки равна t -кратной величине средней квадратической ошибки, где $\Phi(t) = \gamma$, т.е.

$$\Delta = t \sigma_{\bar{x}}$$

Следствие 2.

Интервальная оценка (доверительный интервал) для генеральной средней может быть найден по формуле :

$$\bar{x} - \Delta \leq \bar{x}_0 \leq \bar{x} + \Delta.$$