Вопрос 3. Точечная оценка параметров. Несмещенность, состоятельность, оптимальность оценок

Определение. *Выборкой* называется набор $X_1, \ldots, X_n$ независимых одинаково распределенных случайных величин.

Как правило, распределение этих с.в. неизвестно, и его нужно каким-то образом оценить. Предположим, что известен *вид* распределения (функции распределения) с точностью до неизвестного параметра (-ов): $F(x, \theta) = \mathsf{P}_\theta(X_1 < x)$, $x \in \mathbb{R}$, $\theta = (\theta_1, \ldots, \theta_m) \in \Theta \subset \mathbb{R}^m$. При этом множество возможных значений неизвестного параметра Θ называется *параметрическим множеством.* В этом случае задача оценивания неизвестной функции распределения сводится к задаче оценивания одного или нескольких неизвестных параметров рассматриваемого распределения.

Определение. *Статистикой* называется любая измеримая функция от выборки, не зависящая от θ (важно, что значение статистики можно вычислить при каждой реализации выборки, не зная значения θ).

Всюду далее будем предполагать, что $m = 1$ (т.е. неизвестное распределение параметризовано всего одним параметром).

Определение. *Точечной оценкой* неизвестного параметра θ называется любая статистика. Однако, нам хотелось бы работать с "хорошими" оценками, которые в некотором смысле приближали бы значение неизвестного параметра θ. Далее рассматривается несколько разумных свойств, которыми могли бы обладать "хорошие" оценки:

1. **Несмещенность.** Статистика $T(X_1, \ldots, X_n)$ называется *несмещенной* оценкой параметра θ, если $\mathsf{E}_\theta T(X_1, \ldots, X_n) \equiv \theta$, $\forall \theta \in \Theta$.

   Здесь индекс θ у символа математического ожидания (м.о.) появляется из-за того, что ф.р., определяющая оператор м.о., зависит от θ, и поэтому получается, что среднее значение (м.о.) статистики T зависит от θ, несмотря на то, что сама функция T от θ не зависит. Напомним, что

   $$\mathsf{E}_\theta T(X_1, \ldots, X_n) = \int_{-\infty}^{+\infty} t\, dF_T(t, \theta),$$

   где $F_T(t, \theta) \equiv \mathsf{P}_\theta(T(X_1, \ldots, X_n) < t)$, $t \in \mathbb{R}$ — ф.р. случайной величины $T(X_1, \ldots, X_n)$, она зависит от θ, потому что распределение с.в. $X_1, \ldots, X_n$ зависит от θ.

   Несмещенных оценок может быть несколько, а может не существовать вообще. В случае, когда несмещенных оценок много, естественно отдать предпочтение той, которая имеет минимальный разброс. Последний характеризуется дисперсией (среднеквадратичным отклонением от среднего): $\mathsf{D}_\theta T = \mathsf{E}_\theta(T - \mathsf{E}_\theta T)^2$. Эти рассуждения лежат в основе следующего определения.

2. **Оптимальность.** Несмещенная оценка $T^*(X_1, \ldots, X_n)$ параметра θ называется *оптимальной*, если она имеет равномерно минимальную дисперсию в классе всех несмещенных оценок, т.е. для любой несмещенной оценки T с конечным вторым моментом имеем $\mathsf{D}_\theta T^* \leqslant \mathsf{D}_\theta T$ для всех $\theta \in \Theta$.

   Следующее свойство касается асимптотического поведения оценки с ростом объема выборки.

3. **Состоятельность.** Статистика $T(X_1, \ldots, X_n)$ называется *состоятельной* оценкой параметра θ, если она сходится по вероятности к θ:

$$\forall \varepsilon > 0 \quad \mathsf{P}_\theta(|T(X_1, \ldots, X_n) - \theta| > \varepsilon) \longrightarrow 0, \quad n \to \infty, \quad \forall \theta \in \Theta.$$

Теорема 1 *Оптимальная оценка единственна, т.е. если $T_1 = T_1(X_1, \ldots, X_n)$ и $T_2 = T_2(X_1, \ldots, X_n)$ — оптимальные оценки параметра θ, то $\mathsf{P}_\theta(T_1 = T_2) = 1$ для всех $\theta \in \Theta$.*

Теорема 2 *Если $T_n = T(X_1, \ldots, X_n)$ — несмещенная оценка параметра θ и $\mathsf{D}_\theta T_n \longrightarrow 0$, $n \to \infty$, $\theta \in \Theta$, то T_n — состоятельная оценка параметра θ.*

Данный конспект ответа далеко не полный, хотелось бы, чтобы студент еще мог проиллюстрировать сказанное на примерах, или решить какие-нибудь задачи по теме. Упомянутые примеры, решения типичных задач, а также доказательства сформулированных утверждений можно найти, например, в книге [1].

Список литературы

[1] Г. И. Ивченко, Ю. И. Медведев. *Математическая статистика.* — М., “Высшая школа”, 1984, 248 с.

[2] В. А. Ватутин, Г. И. Ивченко, Ю. И. Медведев, В. П. Чистяков. *Теория вероятностей и математическая статистика в задачах.* — М., “Дрофа”, 2005, 315 с.