Точечные оценки. Свойства оценок.
Методы построения оценок.

Предположим, что имеется выборка $X_{[n]} = (X_1, \ldots, X_n)$ из генеральной совокупности с функцией распределения $F(x)$, принадлежащей некоторому семейству распределений $\Im$. Пусть θ - параметр, однозначно определяемый по каждому распределению F из семейства $\Im$. Например, $\theta = \int\limits_{-\infty}^{\infty} x^r dF(x)$. Следовательно, $\theta = \theta(F)$ - это функционал распределения $F \in \Im$. Часто предполагают, что само семейство распределений $\Im$ определяется одним или несколькими параметрами. Тогда любая $F \in \Im$ есть функция распределения $F(x, \theta)$ или $F(x, \theta_1, \ldots, \theta_k)$, зависящая от одного или нескольких параметров. Такое семейство распределений называется *параметрическим.* В каждом из этих случаев задача оценивания параметра θ состоит в нахождении такой измеримой функции $\hat{\theta} = \hat{\theta}(X_{[n]})$ от выборки, которая в каком-либо смысле близка к параметру θ, если выборка взята из распределения F с $\theta(F) = \theta$. Измеримые функции от выборки называются *статистиками.* Таким образом, $\hat{\theta}$ является статистикой. Предполагается, что статистика $\hat{\theta}$, которая в дальнейшем будет называться *точечной оценкой* или просто *оценкой* параметра θ не зависит от значения оцениваемого параметра θ и других неизвестных параметров, от которых может зависеть функция распределения F.

В основе почти всех методов оценивания лежит следующий основной принцип построения точечных оценок, который называют *принцип подстановки эмпирического распределения.* Пусть, например, $\theta = \theta(F)$, тогда принцип подстановки предписывает в качестве оценки $\hat{\theta}$ взять статистику $\hat{\theta} = \hat{\theta}(F_n^*)$, где F_n^* - эмпирическая функция распределения, построенная по выборке $X_{[n]} = (X_1, \ldots, X_n)$, Напомним, что по определению

$F_n^*(x) = \dfrac{\nu(-\infty, x]}{n}$, где $\nu(-\infty, x]$ - число элементов выборки $X_{[n]} = (X_1, \ldots, X_n)$, попавших во множество $(-\infty, x]$, n - объем выборки. Принцип подстановки представляет собой весьма естественный подход к задаче оценивания параметра, так как по теореме Гливенко–Кантелли эмпирическая функция распределения неограниченно сближается с функцией распределения генеральной совокупности.

Определение 1. Оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ (вернее, последовательность оценок $\hat{\theta}$) называется *состоятельной* , если при $n \to \infty$ она сходится по вероятности к параметру θ:

$$\hat{\theta} \xrightarrow{p} \theta, \quad n \to \infty.$$

Оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ называется *сильно состоятельной,* если при $n \to \infty$ она сходится с вероятностью единица к параметру θ:

$$\hat{\theta} \xrightarrow{\text{ï.í.}} \theta, \quad n \to \infty.$$

Примером сильно состоятельной оценки может служить выборочный $r-$ый момент $a_r^* = \frac{1}{n}\sum_{i=1}^{n} X_i^r$, если конечен $r-$ый момент генеральной совокупности $a_r = \int_{-\infty}^{\infty} x^r dF(x)$, так как по *усиленному закону больших чисел* $a_r^*(X_{[n]}) \xrightarrow{п.н.} a_r$, при $n \to \infty$.

Определение 2. Оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ называется *несмещенной,* если при любом возможном θ и n:

$$E\hat{\theta} = \theta.$$

Оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ называется *асимптотически несмещенной,* если при любом возможном θ и при $n \to \infty$:

$$E\hat{\theta} \to \theta, \quad n \to \infty.$$

Примером несмещенной оценки может служить выборочный $r-$ый момент $a_r^* = \frac{1}{n}\sum_{i=1}^{n} X_i^r$, если конечен $r-$ый момент генеральной совокупности $a_r = \int_{-\infty}^{\infty} x^r dF(x)$, так как нетрудно проверить выполнение равенства:

$$Ea_r^* = a_r$$

Примером *асимптотически несмещенной* оценки может служить выборочная дисперсия $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$, при условии, что дисперсия генеральной совокупности $\sigma^2 = \int_{-\infty}^{\infty}(x-a_1)^2 dF(x)$ конечна. Действительно, как нетрудно убедиться, $Es^2 = \frac{n-1}{n}\sigma^2$, откуда следует асимптотическая несмещенность оценки $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Нетрудно исправить выборочную дисперсию s^2 так, чтобы исправленная выборочная дисперсия s_0^2 была несмещенной, для этого положим $s_0^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, т.е. $s_0^2 = \frac{n}{n-1}s^2$.

Свойство несмещенности имеет важное значение в экспериментальных науках, так как позволяет обобщать результаты экспериментальных исследований, проведенных в разных научных центрах.

Определение 3. Оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ называется *асимптотически нормальной (а.н.)* с коэффициентом рассеивания σ^2, если $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \varsigma \propto N(0,\sigma^2)$.

Последнее соотношение означает сходимость по распределению к случайной величине распределенной нормально с параметрами $(0,\sigma^2)$ и может читаться

также следующим образом: оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ асимптотически нормальна с параметрами $(0,\sigma^2)$.
Коэффициент рассеивания σ^2 может зависеть от параметра θ, т.е. $\sigma^2 = \sigma^2(\theta)$.
Для проверки асимптотической нормальности оценки можно воспользоваться центральными предельными теоремами.

Определение 4. Оценка $\hat{\theta} = \hat{\theta}(X_{[n]}) \in K$ называется *эффективной* в классе K, если для любой другой оценки $\tilde{\theta} \in K$ при всех возможных θ выполняется
$\mathring{A}(\hat{\theta} - \theta)^2 \geq E(\tilde{\theta} - \theta)^2$.
Особую роль играет класс K_0 несмещенных оценок. Эффективные оценки в классе $K_0 = \{\tilde{\theta} : E\tilde{\theta} = \theta\}$несмещенных оценок называются просто *эффективными.* При сравнении эффективности асимптотически нормальных оценок применяют асимптотический подход, основанный на следующем определении.

Определение 5. Асимптотически нормальная оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ с коэффициентом рассеивания $\sigma_1^2(\theta)$ называется *асимптотически эффективной,* если для любой другой асимптотически нормальной оценки $\tilde{\theta}$ с коэффициентом рассеивания $\sigma_2^2(\theta)$ при всех возможных θ выполняется
$\sigma_1^2(\theta) \leq \sigma_2^2(\theta)$.

Метод моментов

Пусть имеется выборка $X_{[n]} = (X_1,\ldots,X_n)$ из генеральной совокупности с функцией распределения $F(x,\theta_1,\ldots,\theta_k)$. Предположим, что все моменты $a_r(\theta_1,\ldots,\theta_k) = \int\limits_{-\infty}^{\infty} x^r dF(x,\theta_1,\ldots,\theta_k),\ r = 1,\ldots,k,$ конечны и что система уравнений:

$$a_r(\theta_1,\ldots,\theta_k) = a_r^*, \quad r = 1,\ldots,k$$

однозначно разрешима относительно $\theta_1,\ldots,\theta_k$, где $a_r^* = \frac{1}{n}\sum_{i=1}^{n} X_i^r$, тогда оценки $\hat{\theta}_r = g_r(a_1^*,\ldots,a_k^*),\quad r = 1,\ldots,k$, получаемые как решение системы уравнений называются *оценками по методу моментов.*
Нетрудно доказать сильную состоятельность оценок по методу моментов при выполнении сформулированных выше условий, если функции $g_r(z_1,\ldots,z_k), r = 1,\ldots,k$, непрерывны. Для доказательства достаточно заметить, что выборочные моменты являются сильно состоятельными оценками соответствующих моментов генеральной совокупности $a_r, r = 1,\ldots,k$.

Пример 1. Рассмотрим параметрическое семейство нормальных распределений $N(a,\sigma^2)$. Оценки метода моментов для параметров a,σ^2 имеют следующий вид $\hat{a} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

Метод максимального правдоподобия.

Пусть имеется выборка $X_{[n]} = (X_1, \dots, X_n)$ из генеральной совокупности с функцией распределения $F(x, \theta_1, \dots, \theta_k)$. Предположим, что функция распределения $F(x, \theta_1, \dots, \theta_k)$ имеет плотность распределения $f(x, \theta_1, \dots, \theta_к)$, тогда функция правдоподобия выборки имеет вид

$$L(X_1, \dots, X_n, \theta_1, \dots, \theta_к) = \prod_{i=1}^{n} f(X_i, \theta_1, \dots, \theta_k).$$

Оценками по методу максимального правдоподобия $\hat{\theta}_i = \hat{\theta}_i(X_{[n]}), i = 1, \dots, k,$ называются оценки, которые обращают в максимум функцию правдоподобия:

$$L(X_{[n]}, \hat{\theta}_1, \dots, \hat{\theta}_k) = \max_{\theta_1, \dots, \theta_k} L(X_{[n]}, \theta_1, \dots, \theta_k)$$

Аналогично определяются оценки по методу максимального правдоподобия в дискретном случае. При нахождении оценок максимального правдоподобия иногда бывает удобно логарифмировать функцию правдоподобия, после чего вычислять производные по параметрам и приравнивать их к нулю, получая так называемое *уравнение правдоподобия,* однако следует помнить, что это не более чем прием решения задачи максимизации функции правдоподобия, который далеко не всегда приводит к успеху.

Пример 2. Рассмотрим параметрическое семейство нормальных распределений $N(a, \sigma^2)$. Оценки метода максимального правдоподобия для параметров a, σ^2 имеют следующий вид $\hat{a} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Они легко находятся как решение уравнения правдоподобия.

Пример 3. Рассмотрим параметрическое семейство определяемое плотностью следующего вида:

$$f(x, \theta) = \begin{cases} \theta^{-1}, & x \in [0, \theta] \\ 0, & x \notin [0, \theta] \end{cases}$$

Нельзя найти оценку по методу максимального правдоподобия в этом примере решая уравнение правдоподобия. Однако нетрудно построить график функции правдоподобия, как функции параметра θ, для фиксированной выборки. График функции имеет разрыв. Максимум достигается в точке $\hat{\theta} = \max\{X_1, \dots, X_n\}$.