Точечное оценивание.

<u>Точечные оценки. Несмещенность и состоятельность оценок.</u>

Как правило, в статистики редко о проводимом эксперименте совсем ничего нельзя сказать. Обычно, во многих задачах, тип распределения известен заранее, но неизвестны параметры или часть параметров этого распределения. Например, ошибки измерения предполагаются распределенными по нормальному закону с математическим ожиданием равным нулю (если систематическая ошибка равна нулю) и неизвестной дисперсией, число покупателей в магазине в течении часа имеет распределение Пуассона с неизвестным параметром λ и т.д. В результате мы имеем задачу статистического оценивания параметров этого распределения на основе выборочных данных.

Определение. Класс распределений, целиком определяющийся значением параметра $\theta \in \Theta$ (одномерным либо векторным) будем называть параметрическим семейством распределений F_θ .

Например, $F_\theta = \Pi_\lambda$ - семейство распределений Пуассона с параметром $\lambda > 0$. Здесь $\theta = \lambda, \quad \Theta = (0, \infty)$.

Определение. Любая функция $\theta^* = \theta^*(X_1, X_2, \ldots, X_n)$ от выборки называется ***<u>статистикой</u>*** θ^*.

Статистика - случайная величина.

Определение. Статистика $\theta^* = \theta^*(X_1, X_2, \ldots, X_n)$, используемая для оценки неизвестного параметра θ , называется ***<u>оценкой параметра</u>*** θ .

Например, статистика $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ - выборочное среднее, может быть использована в качестве оценки математического ожидания, статистика $\overline{D} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2$ может быть использована для оценки дисперсии и т.д.

Определение. Статистика θ^* называется ***<u>несмещенной оценкой параметра</u>*** θ , если

$$M(\theta^*) = \theta, \quad \forall \theta \in \Theta. \qquad (2.1).$$

Свойство несмещенности означает отсутствия ошибки в среднем, при систематическом использовании оценки. Так выборочное среднее является несмещенной оценкой математического ожидания (1.14), выборочная несмещенная дисперсия s^2 является несмещенной оценкой дисперсии (1.16). Выборочная дисперсия $\overline{D}$ является смещенной оценкой дисперсии (1.16).

Определение. Статистика θ^* называется ***<u>асимптотически несмещенной оценкой параметра</u>*** θ, если при $n \to \infty$

$$M(\theta^*) \to \theta, \quad \forall \theta \in \Theta. \tag{2.2}$$

Так, выборочная дисперсия $\overline{D}$ является асимптотически несмещенной оценкой дисперсии.

Определение. Статистика θ^* называется ***<u>состоятельной оценкой параметра</u>*** θ, если при $n \to \infty$

$$\theta^* \xrightarrow{p} \theta, \quad \forall \theta \in \Theta. \tag{2.3}$$

То есть состоятельная оценка должна стремиться к параметру с ростом объема выборки. Выборочное среднее и дисперсии в соответствии с (1.15), (1.17) являются состоятельными оценками математического ожидания и дисперсии.

Требования несмещенности и состоятельности являются основными требованиями, предъявляемыми к "качественным" оценкам.

Выборочные среднее $\overline{X}$ и дисперсия s^2, как оценки математического ожидания и дисперсии произвольного распределения, не были выведены нами ни из каких математических формул, они просто разумны с практической точки зрения как аналоги характеристик генеральной совокупности, и, как видим, обладают неплохими свойствами, с точки зрения требований предъявляемых к оценке. Однако в общем случае необходим систематический и обоснованный подход к методам получения оценок. Существуют различные подходы.

<u>Метод моментов.</u>

Согласно методу моментов в качестве оценок начальных моментов генеральной совокупности принимаются их выборочные аналоги, т.е.

$$m_k^* = \overline{m_k} = \overline{X^k}. \tag{2.4}$$

Для оценки других параметров распределения по методу моментов, необходимо их выразить через начальные моменты случайной величины (в общем случае для этого необходимо знание закона распределения). Итак:

Пусть $X_1, X_2, \ldots, X_n$ выборка из семейства распределений F_θ, где $\theta \in \Theta$. Пусть $M\left(\xi^k\right) = h(\theta)$, где функция $h(x)$ имеет обратную. Тогда в качестве оценки θ^* параметра θ берем решение уравнения $\overline{X^k} = h(\theta^*)$. Другими словами:

Определение. Если параметр $\theta = h^{-1}(M(\xi^k))$, то

$$\theta^* = h^{-1}(\overline{X^k}) \tag{2.5}$$

называется **оценкой метода моментов** параметра θ по k -му моменту.

Замечание 1. Как видно могут быть различные оценки (2.5) для одного и того же параметра (т.к. параметр может быть выражен через различные моменты).

Замечание 2. Как правило, априорно значения параметра принадлежат некоторому интервалу, т.е. $\theta \in \Theta$, а значение полученной оценки θ^* может и не попасть в этот интервал. Тогда в качестве оценки берут ближайшее к θ^* из Θ.

Замечание 3. Если θ - векторный параметр размерности l, то согласно методу моментов необходимо получить l уравнений для l различных моментов и разрешить их относительно координат вектора.

Замечание 4. Вообще говорю, в методе моментов можно использовать любое уравнение вида

$M(\varphi(\xi)) = h(\theta)$, тогда оценка $\theta^* = h^{-1}(\overline{\varphi(X)})$, где $\overline{\varphi(X)} = \dfrac{1}{n}\sum_{i=1}^{n}\varphi(X_i)$. Функция $\overline{\varphi(X)}$ в этом случае называется пробной функцией.

Согласно методу моментов в качестве оценок математического ожидания и дисперсии при любом законе распределения можно взять выборочное среднее $\overline{X}$ и дисперсия $\overline{D}$. (Действительно т.к. $M(X) = m_1$, следовательно $M^*(X) = \overline{m_1} = \overline{X}$, т.к. $D(X) = m_2 - m_1^2$, то $D^*(X) = \overline{m_2} - \overline{m_1}^2 = \overline{D}$).

Теорема. (состоятельность метода моментов). Пусть $\theta^* = h^{-1}(\overline{m}_k)$ - оценка параметра θ, полученная по методу моментов, причем функция $h^{-1}(x)$ непрерывна, тогда θ^* состоятельна.

Доказательство. Так как в соответствии с ЗБЧ $\overline{m}_k \xrightarrow{p} m_k$, а функция $\varphi(x)$ непрерывна, то $\varphi(\overline{m}_k) \xrightarrow{p} \varphi(m_k) = \theta$.

Пример. *Пусть* $X_1, X_2, \ldots, X_n$ *- выборка из генеральной совокупности равномерно распределенной на интервале* $[0, \theta]$, *где* θ *- неизвестный параметр. Требуется оценить* θ *по методу моментов.*

В данном случае параметр θ *может быть выражен через моменты любого порядка:*

$m_k = \dfrac{1}{\theta}\int_0^{\theta} x^k dx = \dfrac{1}{\theta}\dfrac{x^{k+1}}{k+1}\bigg|_0^{\theta} = \dfrac{\theta^k}{k+1}$. *Откуда* $\theta = \sqrt[k]{(k+1)m_k}$. *Следовательно* $\theta^* = \sqrt[k]{(k+1)\overline{m}_k}$.

Соответственно имеем различные оценки для θ: $\theta_1^* = 2\overline{m}_1$, $\theta_2^* = \sqrt{3\overline{m}_2}$, *и т.д. Проверим несмещенность оценок.*

$M(\theta_1^*) = 2M(\overline{m}_1) = 2m_1 = \theta$, *т.е.* θ_1^* *- несмещенная.*

$M(\theta_2^*) = \sqrt{3}M\left(\sqrt{\overline{m}_2}\right)$. Чтобы оценка θ_2^* была несмещенной, очевидно необходимо, чтобы выполнялось равенство: $M\left(\sqrt{\overline{m}_2}\right) = \sqrt{m_2}$. Т.к. $M\left(\overline{m}_2\right) = m_2$, то это равносильно выполнению условия $M\left(\sqrt{\overline{m}_2}\right) = \sqrt{M(\overline{m}_2)}$, что в свою очередь для случайной величины $\eta = \sqrt{\overline{m}_2}$, равносильно условию $M(\eta^2) = M^2(\eta)$. Последнее возможно лишь при условии, что $D(\eta) = 0$, следовательно, оценка θ_2^* смещенная. Аналогично смещенными являются и другие оценки θ_k^* при $k > 2$.

Таким образом, несмещенной является только оценка θ_1^*, все остальные оценки являются смещенными. Заметим, что в любом случае оценка не может быть меньше, чем наибольшее из выборочных значений. Следовательно, в качестве оценки θ следует принять (при использовании для оценивания первого момента) $\theta^* = \max(\theta_1^*, X_{\max})$.

<u>Метод максимального правдоподобия (ММП)</u>

В соответствии с ММП, в качестве оценок параметров $\theta_1, \theta_2, \ldots$ закона распределения величины ξ, берут такие значения $\theta_1^*, \theta_2^*, \ldots$, при которых вероятность получить данную выборку $X = \{X_1, X_2, \ldots, X_n\}$ из генеральной совокупности величины ξ максимальна.

Определим предварительно, что значит "вероятность получить данную выборку". Для ДСВ очевидно:

$$P(\xi = \{X_1, X_2, \ldots, X_n\}) = P(\xi = X_1) \cdot P(\xi = X_2) \cdot \ldots \cdot P(\xi = X_n).$$

Для НСВ вероятность отдельного значения случайной величины равна нулю, однако можно считать, что вероятность получить значение $\xi = x$, есть вероятность попасть в некоторый малый интервал Δx, содержащий x, т.е. $P(\xi = x) = f_\xi(x)\Delta x$. Тогда для НСВ:

$$P(\xi = \{X_1, X_2, \ldots, X_n\}) = f_\xi(X_1) \cdot f_\xi(X_2) \cdot \ldots \cdot f_\xi(X_n) \cdot (\Delta x)^n.$$

Обозначим:

$$f(x) = \begin{cases} P(\xi = x), & \text{если } \xi \text{ - дискретная случайная величина} \\ f_\xi(x), & \text{если } \xi \text{ - непрерывная случайная величина.} \end{cases} \qquad (2.6)$$

Тогда, вероятность получить выборку $X_1, X_2, \ldots, X_n$:

$$P(\xi=\{X_1,X_2,\cdot\ldots\cdot,X_n\})=\begin{cases}\prod_{i=1}^{n} f(X_i), & \text{для ДСВ,}\\ \prod_{i=1}^{n} f(X_i)\Delta x, & \text{для НСВ.}\end{cases}$$

Так как величина Δx является постоянной, то как для ДСВ, так и для НСВ эта вероятность зависит только от $\prod_{i=1}^{n} f(X_i)$.

Определение.

Функция

$$\Psi(\vec{X},\theta)=\prod_{i=1}^{n} f(X_i) \tag{2.7}$$

называется ***функцией правдоподобия.***

Функция

$$L(\vec{X},\theta)=\ln\left(\prod_{i=1}^{n} f(X_i)\right)=\sum_{i=1}^{n}\ln\left(f(X_i)\right) \tag{2.8}$$

называется ***логарифмической функцией правдоподобия***.

Как видим, функция правдоподобия - это и есть вероятность того, что случайная величина примет значения $X_1,X_2,\ldots X_n$ (в случае НСВ с точностью до постоянного множителя dx^n), причем эта вероятность зависит от неизвестного параметра распределения θ.

Определение. ***Оценкой параметра θ по методу МП*** называется значение θ^*, при котором функция максимального правдоподобия $\Psi(X,\theta)$ достигает наибольшего значения, как функция переменной θ, в области Θ.

Замечание. Т.к. функция $\ln x$ монотонна, то максимумы функции $\Psi(X,\theta)$ совпадают с максимумами функции $L(X,\theta)$, которая более проста для исследования на экстремумы.

Пример. *Пусть $X_1,X_2,\ldots X_n$ - выборка из распределения Пуассона с неизвестным параметром λ. Требуется найти оценку λ по методу МП.*

Для случайной величины распределенной по закону Пуассона $P(\xi=m)=\dfrac{\lambda^m e^{-\lambda}}{m!}$.

$$\Rightarrow\quad L\left(\vec{X},\lambda\right)=\sum_{i=1}^{n}\ln\left(\frac{\lambda^{X_i}e^{-\lambda}}{X_i!}\right)=\sum_{i=1}^{n}\left(X_i\ln\lambda-\lambda-\ln X_i!\right)=\ln\lambda\sum_{i=1}^{n}X_i-\lambda n-\sum_{i=1}^{n}\ln X_i!.$$

Ищем максимум функции, используя методы дифференциального исчисления.

Вычисляем $\frac{\partial L}{\partial \lambda} = \frac{1}{\lambda}\sum_{i=1}^{n} X_i - n$, *и приравнивая к нулю находим оценку для* λ.

$\frac{1}{\lambda}\sum_{i=1}^{n} X_i - n = 0 \quad \Rightarrow \quad \lambda^* = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$. *Легко убедиться, что полученная точка является точкой максимума.*