Тема 2. Точечное оценивание, основные свойства и сравнение оценок. Точечное оценивание вероятностей и моментов. Линейная оценка среднего с наименьшей дисперсией.

1. Точечное оценивание и основные свойства оценок.

Во многих задачах математической статистики требуется построить оценку некоторой неизвестной величины, в качестве которой может выступать, например, вероятность некоторого события, начальный или центральный момент некоторой случайной величины, параметр некоторого распределения. Для построения оценок требуется статистическая информация, которая поступает в виде наблюдений, поэтому неформально построение оценки сводится к разработке способа (метода) обработки наблюдений, что формально соответствует понятию статистики.

Определение 2.1.

Статистикой называется функция наблюдений.

Рассмотрим основные положения теории точечного оценивания в частном случае, когда наблюдения образуют выборку, более общий случай постановки задачи точечного оценивания рассматривается в пособии [Ивченко, Медведев 1].

Пусть $\zeta_n = (\xi_1,..., \xi_n)$ является выборкой из распределения $F_\xi(x \mid \theta)$, где θ – неизвестный параметр (в общем случае вектор неизвестных параметров) из некоторого допустимого множества параметров Θ, и требуется построить оценку неизвестной величины $\tau(\theta)$, зависящей от параметра θ. Например, в качестве неизвестной величины $\tau(\theta)$ может выступать непосредственно сам параметр θ, $\tau(\theta) = \theta$, вероятность $F_\xi(x^* \mid \theta)$ при некотором фиксированном x^*, $\tau(\theta) = F_\xi(x^* \mid \theta)$, или математическое ожидание $m_1(\theta)$,

$$\tau(\theta) = m_1(\theta) = \int_{-\infty}^{\infty} x dF_\xi(x \mid \theta).$$

Предположим, что для оценки неизвестной величины $\tau(\theta)$ некоторым образом была построена статистика T, являющаяся в данном случае функцией выборки $T(\zeta_n)$. Оценка $T(\zeta_n)$ при фиксированном n является случайной величиной, а при изменении n оценка $T(\zeta_n)$ определяет некоторую последовательность по n случайных величин. Для исследования «качества» оценки $T(\zeta_n)$ определяют несколько свойств, среди которых особо выделяют основные свойства: несмещенность, состоятельность и оптимальность.

Определение 2.2.

Оценка $T(\zeta_n) = T(\xi_1,..., \xi_n)$ является *несмещенной оценкой* $\tau(\theta)$, если при всяком фиксированном n и любом допустимом значении параметра θ математическое ожидание оценки совпадает с оцениваемой величиной $\tau(\theta)$:

$$\forall n, \forall \theta \in \Theta : M_\theta[T(\xi_1,..., \xi_n)] = \tau(\theta),$$

где $M_\theta[T(\xi_1,..., \xi_n)] = \int_{R^n} T(x_1,..., x_n) dF_{\zeta_n}(x_1,..., x_n \mid \theta)$) и $F_{\zeta_n}(x_1,..., x_n \mid \theta)$ – функция распределения выборки $\zeta_n = (\xi_1,..., \xi_n)$.

Фактически, наличие свойства несмещенности означает, что каково бы ни было значение параметра θ, оценка $T(\zeta_n)$ «в среднем» окажется близкой к неизвестной величине $\tau(\theta)$.

Определение 2.3.

Оценка $T(\zeta_n)$ называется *состоятельной*, если при любом допустимом значении параметра θ последовательность по n случайных величин $T(\zeta_n)$ сходится по вероятности к оцениваемой величине $\tau(\theta)$ при возрастании n:

$$\forall\,\theta \in \Theta : T(\zeta_n) \xrightarrow{P_\theta} \tau(\theta) \text{ при } n \to \infty\,.$$

Из свойства состоятельности, согласно определению сходимости по вероятности (приложение 1), следует, что для всяких $\varepsilon > 0$ и $\delta > 0$ можно найти N такое, что при каждом $n \geq N$:

$$P_\theta\{|\,T(\zeta_n) - \tau(\theta)\,| < \varepsilon\} > 1 - \delta\,.$$

Фактически это означает, что с увеличением объема выборки n вероятность малого отклонения оценки $T(\zeta_n)$ от значения $\tau(\theta)$ оказывается близкой к 1, то есть при больших n с большой вероятностью (с вероятностью близкой к 1), значение $T(\zeta_n)$ окажется в ε-окрестности величины $\tau(\theta)$ при любом значении параметра θ.

Свойства несмещенности и состоятельности являются независимыми свойствами. Из несмещенности оценки вовсе не следует состоятельность оценки, и наоборот состоятельность оценки не гарантирует несмещенность оценки, в некоторых случаях из свойства состоятельности даже не следует существование математического ожидания оценки. Более того, даже если оценка состоятельна и имеет математическое ожидание, оно не обязательно совпадает с оцениваемой величиной.

Несмещенность и состоятельность характеризуют оценку с разных позиций. Несмещенность затрагивает свойства оценки при фиксированном n и каждом допустимом значении параметра θ, состоятельность напротив рассматривает оценку при каждом допустимом и фиксированном значении параметра θ, но переменном n, пробегающим все множество натуральных чисел, то есть рассматривает оценку как некоторую последовательность случайных величин.

Свойство состоятельности является обязательным свойством, оценки, не обладающие состоятельностью, не рассматриваются и не используются. Свойство несмещенности является желательным, более того, в некоторых случаях смещенные оценки оказываются «лучше» несмещенных, а исправление смещенности приводит к ухудшению оценки, поэтому в некоторых случаях предпочтение отдают смещенным оценкам, несмотря на отсутствие у них свойства несмещенности.

В общем случае для оценки величины $\tau(\theta)$ могут быть предложены несколько оценок, поэтому возникает вопрос о том, как сравнивать оценки и по какому критерию сравнения выбирать наилучшую в некотором смысле оценку из нескольких возможных оценок.

Рассмотрим вопрос сравнения оценок на следующем простом примере: пусть объем выборки n фиксирован и для оценки величины $\tau(\theta)$ используются две различные оценки $T_1(\zeta_n)$ и $T_2(\zeta_n)$, причем известно, что оценки являются несмещенными и что существует некоторое число $\varepsilon_0 > 0$, такое, что при всех $\varepsilon < \varepsilon_0$ имеет место следующее неравенство:

$$\exists\,\varepsilon_0 > 0\,\forall\,\varepsilon < \varepsilon_0 : P\{|\,T_1(\zeta_n) - \tau(\theta)\,| < \varepsilon\} < P\{|\,T_2(\zeta_n) - \tau(\theta)\,| < \varepsilon\}\,.$$

Вполне очевидно, что оценка $T_2(\zeta_n)$ лучше оценки $T_1(\zeta)$, поскольку оценка $T_2(\zeta_n)$ с большей вероятностью оказывается в ε-окрестности величины $\tau(\theta)$, чем оценка $T_1(\zeta_n)$. При однократном проведении эксперимента и вычислении оценок значение оценки $T_2(\zeta_n)$ с большей вероятностью окажется ближе к $\tau(\theta)$, чем значение оценки $T_1(\zeta_n)$. При многократном проведении эксперимента и вычислении оценок значения оценки $T_2(\zeta_n)$ группируются возле значения $\tau(\theta)$ с меньшем разбросом, чем значения оценки $T_1(\zeta_n)$.

Приведенный пример показывает, что оценки можно сравнивать по величине разброса значений оценки, который определяется «шириной» распределения. Если распределение «широкое», то разброс значений оценки оказывается большим, если распределение «узкое», то разброс значений оценки оказывается малым. К сожалению, в большинстве случаев получить вероятностные распределения оценок достаточно сложно, поэтому, как правило, используют не вероятностные распределения, а различные числовые характеристики, которые отражают величину «ширины» распределения, в частности дисперсию.

Предположим, что для оценки $T(\zeta_n)$ существует второй центральный момент (дисперсия) $D(\theta) = D_\theta[T(\zeta_n)]$, который в общем случае зависит от значения параметра θ,

$$D(\theta) = D_\theta[T(\zeta_n)] =$$
$$= \int_{R^n} \left(T(x_1,..., x_n) - M_\theta[T(\zeta_n)]\right)^2 dF_\zeta(x_1,..., x_n \mid \theta) .$$

По величине дисперсии $D(\theta)$ можно судить о мере «разброса» значений оценки $T(\zeta_n)$: оценки с большим «разбросом» могут с большей вероятностью принимать значения далекие от оцениваемой величины $\tau(\theta)$, чем оценки с малым «разбросом». Кроме того, оценка с малым «разбросом» оказывается «сосредоточенной» вокруг математического ожидания $M_\theta[T(\zeta_n)]$, которое в случае дополнительного свойства несмещенности совпадает с оцениваемой величиной $\tau(\theta)$, так что оценка фактически оказывается «сосредоточенной» вокруг величины $\tau(\theta)$. Отсюда следует, что предпочтение следует отдавать оценкам с малым «разбросом» (с малой величиной дисперсии).

Таким образом, при фиксированном объеме выборки n, используя величины дисперсий оценок, можно сформулировать критерий наименьшей дисперсии сравнения оценок – «из двух оценок лучше та оценка, у которой дисперсия меньше».

Предположим, что n фиксировано и U_τ есть класс несмещенных оценок величины $\tau(\theta)$ с конечной дисперсией:

$$\forall T \in U_\tau : M_\theta[T] = \tau(\theta) ,\ D_\theta[T] < \infty .$$

Пусть $T_1 \in U_\tau$ и $T_2 \in U_\tau$ две различных оценки из класса U_τ, в общем случае на основе дисперсий $D_\theta[T_1]$ и $D_\theta[T_2]$ не всегда удается указать, какая оценка «лучше», поскольку может оказаться (рисунок 2.1), что при одном значении параметра θ_1:

$$D_{\theta_1}[T_1] < D_{\theta_1}[T_2] ,$$

а при другом значении параметра θ_2 наоборот:

$$D_{\theta_2}[T_1] > D_{\theta_2}[T_2] .$$

В этом случае по принятому критерию наименьшей дисперсии невозможно указать, какая оценка лучше.

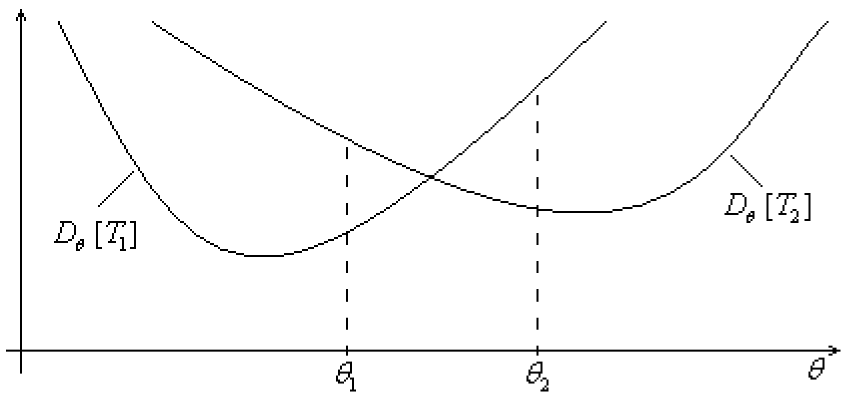

Рисунок 2.1. Дисперсии оценок.

Тем не менее, в некоторых случаях в классе U_τ есть оценка T^*, которая при каждом значении параметра θ имеет наименьшую дисперсию среди дисперсий оценок класса U (рисунок 1.3):

$$\forall \theta \in \Theta : D_\theta[T^*] = \inf_{T \in U_\tau} D_\theta[T] .$$

В этом случае, оценка T^*, очевидно, является наилучшей в классе U и её следует признать оптимальной.

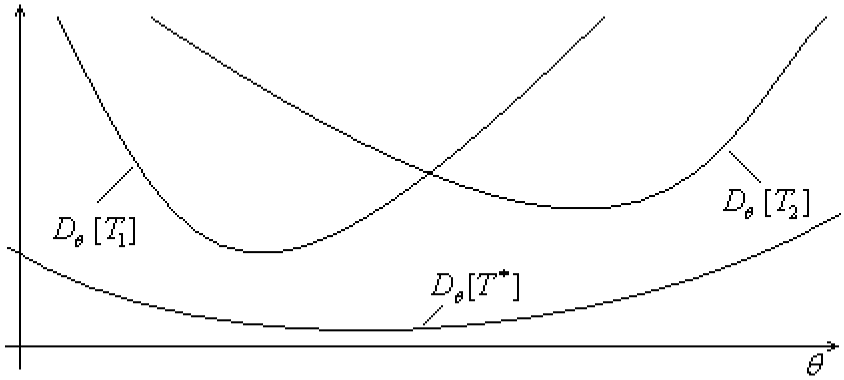

Рисунок 2.3. Дисперсия оптимальной оценки.

Определение 2.4.

Оценка $T^*(\zeta_n)$ называется *оптимальной в классе несмещенных оценок* U_τ, если

$$\forall\,\theta \in \Theta:\ D_\theta[T^*] = \inf_{T \in U_\tau} D_\theta[T]\,.$$

Утверждение 2.5.

Пусть U_τ класс несмещенных оценок $\tau(\theta)$, если в классе U_τ существует оптимальная оценка, то она единственна.

Доказательство:

Докажем утверждение от противного: пусть $T^* \in U_\tau$ оптимальная оценка, предположим, что $\tilde{T} \in U_\tau$ другая оптимальная оценка, которая не совпадает с T^*, $T^* \neq \tilde{T}$.

Образуем оценку T' следующим образом:

$$T' = \frac{T^* + \tilde{T}}{2}\,.$$

Вычислим математическое ожидание $M_\theta[T']$, учитывая, что $M_\theta[T^*] = \tau(\theta)$ и $M_\theta[\tilde{T}] = \tau(\theta)$, поскольку U_τ класс несмещенных оценок $\tau(\theta)$:

$$M_\theta[T'] = M_\theta\left[\frac{T^* + \tilde{T}}{2}\right] = \frac{M_\theta[T^*] + M_\theta[\tilde{T}]}{2} = \frac{\tau(\theta) + \tau(\theta)}{2} = \tau(\theta)\,,$$

таким образом, оценка T' также является несмещенной и поэтому $T' \in U_\tau$.

Вычислим дисперсию $D_\theta[T']$:

$$D_\theta[T'] = D_\theta\left[\frac{T^* + \tilde{T}}{2}\right] = \frac{D_\theta[T^*] + D_\theta[\tilde{T}] + 2\operatorname{cov}_\theta(T^*, \tilde{T})}{4} \qquad (2.1)$$

Согласно свойству ковариации $\left|\operatorname{cov}_\theta(T^*, \tilde{T})\right|^2 \leq D_\theta[T^*] \cdot D_\theta[\tilde{T}]$, поэтому:

$$\frac{D_\theta[T^*] + D_\theta[\tilde{T}] + 2\operatorname{cov}_\theta(T^*, \tilde{T})}{4} \leq \frac{D_\theta[T^*] + D_\theta[\tilde{T}] + 2\sqrt{D_\theta[T^*] D_\theta[\tilde{T}]}}{4}\,.$$

Поскольку T^* и $\tilde{T}$ оценки оптимальные, то их дисперсии совпадают, поскольку по определению оптимальной оценки:

$$D_\theta[T^*] = \inf_{T \in U_\tau} D_\theta[T] = D_\theta[\tilde{T}]\,,$$

тогда очевидно,

$$\frac{D_\theta[T^*] + D_\theta[\tilde{T}] + 2\sqrt{D_\theta[T^*] D_\theta[\tilde{T}]}}{4} = \frac{D_\theta[T^*] + D_\theta[T^*] + 2D_\theta[T^*]}{4} = D_\theta[T^*]\,,$$

таким образом,

$$D_\theta[T'] \le D_\theta[T^*].$$

С другой стороны, оценка T^* является оптимальной, поэтому по определению оптимальной оценки:

$$D_\theta[T^*] = \inf_{T \in U_\tau} D_\theta[T] \le D_\theta[T'].$$

Из двух полученных неравенств следует, что:

$$D_\theta[T^*] = D_\theta[T'].$$

С учетом полученного равенства и равенства $D_\theta[T^*] = D_\theta[\tilde{T}]$ из (1.1) получим:

$$D_\theta[T^*] = \frac{D_\theta[T^*] + D_\theta[T^*] + 2\operatorname{cov}_\theta(T^*, \tilde{T})}{4},$$

$$\operatorname{cov}_\theta(T^*, \tilde{T}) = D_\theta[T^*] = \sqrt{D_\theta[T^*] D_\theta[T^*]} \qquad (2.2)$$

По свойству ковариации полученное равенство возможно тогда и только тогда, когда T^* и $\tilde{T}$ связаны линейно:

$$T^* = k(\theta)\tilde{T} + a(\theta).$$

Вычислив математическое ожидание левой и правой частей, получим:

$$M_\theta[T^*] = k(\theta) M_\theta[\tilde{T}] + a(\theta),$$

$$\tau(\theta) = k(\theta)\tau(\theta) + a(\theta),$$

$$(1 - k(\theta))\tau(\theta) = a(\theta),$$

тогда,

$$T^* = k(\theta)\tilde{T} + a(\theta) = k(\theta)\tilde{T} + (1 - k(\theta))\tau(\theta),$$

$$T^* - \tau(\theta) = k(\theta)\tilde{T} + k(\theta)\tau(\theta) = k(\theta)(\tilde{T} - \tau(\theta)).$$

С учетом полученного равенства из (2.2) получим:

$$D_\theta[T^*] = \operatorname{cov}_\theta(T^*, \tilde{T}) = M_\theta[(T^* - M_\theta[T^*])(\tilde{T} - M_\theta[\tilde{T}])] = M_\theta[(T^* - \tau(\theta))(\tilde{T} - \tau(\theta))] =$$

$$= M_\theta[k(\theta)(\tilde{T} - \tau(\theta))(\tilde{T} - \tau(\theta))] = k(\theta) M_\theta[(\tilde{T} - \tau(\theta))(\tilde{T} - \tau(\theta))] = k(\theta) D_\theta[\tilde{T}] = k(\theta) D_\theta[T^*],$$

откуда $k(\theta) \equiv 1$, то есть:

$$T^* = k(\theta)\tilde{T} + a(\theta) = k(\theta)\tilde{T} + (1 - k(\theta))\tau(\theta) = \tilde{T} + (1-1)\tau(\theta) = \tilde{T}.$$

Полученное равенство $T^* = \tilde{T}$ противоречит исходному предположению о том, что оценки T^* и $\tilde{T}$ не совпадают.

Утверждение доказано.

Критерий наименьшей дисперсии может быть использован при сравнении несмещенных оценок. Попытка применить критерий наименьшей дисперсии к оценкам общего вида, не обладающих свойством несмещенности, приводит к неожиданному результату: для тривиальных оценок вида $C(\zeta_n) = C$, где C – некоторая постоянная, дисперсия $D_\theta[C(\zeta_n)] = D_\theta[C] = 0$ и по критерию наименьшей дисперсии тривиальные оценки являются наилучшими, что конечно неприемлемо.

При сравнении оценок общего вида может быть использован критерий наименьшей среднеквадратичной ошибки – «из двух оценок лучше та оценка, среднеквадратическая ошибка которой меньше».

Определение 2.6.

Среднеквадратической ошибкой статистики $T(\zeta_n)$, оценивающей величину $\tau(\theta)$, называется математическое ожидание квадрата отклонения:

$$M_\theta\left[\left|T(\zeta_n) - \tau(\theta)\right|^2\right].$$

Заметим, что в случае несмещенной оценки, $M_\theta[T(\zeta_n)] = \tau(\theta)$, среднеквадратическая ошибка становится равной дисперсии, поэтому для несмещенных оценок критерий наименьшей среднеквадратической ошибки совпадает с уже рассмотренным критерием наименьшей дисперсии (критерий наименьшей среднеквадратической ошибки является естественным продолжением критерия наименьшей дисперсии с множества несмещенных оценок на множество оценок общего вида, не обязательно являющихся несмещенными).

В еще более общем случае в рассмотрение вводится функция потерь $L(T(\zeta_n), \tau(\theta))$, которая используется при вычислении функции условного риска $R_\theta(T(\zeta_n))$:

$$R_\theta(T(\zeta_n)) = \int_{R^n} L(T(\zeta_n), \tau(\theta))\, dF_\zeta(x \mid \theta),$$

где $F_\zeta(x \mid \theta)$ функция распределения выборки $\zeta_n = (\xi_1, ..., \xi_n)$. Значения функции $R_\theta(T(\zeta_n))$ используются для сравнения оценок – «из двух оценок лучше та оценка, величина функции условного риска для которой меньше». В частности, если $L(T(\zeta_n), \tau(\theta)) = \left(T(\zeta_n) - \tau(\theta)\right)^2$, то функция условного риска $R_\theta(T(\zeta_n))$ совпадает с величиной среднеквадратической ошибки.

2. Состоятельность оценок.

Состоятельность является важным свойством оценок, но установление свойства состоятельности непосредственно из определения (доказательство сходимости по вероятности) в некоторых случаях оказывается затруднительным, поэтому часто прибегают к использованию предельных теорем, в частности закона больших чисел в разных формах (теорема Бернулли, теорема Гливенко, теорема Хинчина, теорема Чебышева), «арифметических» свойств сходимости по вероятности и вспомогательных утверждений.

Теорема (Бернулли)

Пусть μ_n – количество появлений события A в n независимых испытаниях, тогда последовательность относительных частот $\frac{\mu_n}{n}$ сходится по вероятности к вероятности $P(A)$ события A, при $n \to \infty$:

$$\frac{\mu_n}{n} \xrightarrow{P} P(A), \text{ при } n \to \infty.$$

Теорема (Хинчин)

Пусть $\{\xi_n\}_{n=1}^{\infty}$ – последовательность независимых в совокупности случайных величин, имеющих одинаковую функцию распределения и конечное математическое ожидание m, тогда последовательность случайных величин $\frac{1}{n}\sum_{i=1}^{n} \xi_i$ сходится по вероятности к m, при $n \to \infty$:

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{P} m, \text{ при } n \to \infty.$$

Теорема (Чебышев)

Пусть $\{\xi_n\}_{n=1}^{\infty}$ – последовательность случайных величин таких, что:

1) ξ_n попарно независимы,

2) дисперсии $D[\xi_n]$ ограничены одной и той же постоянной C:

$$\forall n : D[\xi_n] \le C,$$

тогда,

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} M[\xi_i], \text{ при } n \to \infty.$$

Теорема (Марков)

Пусть $\{\xi_n\}_{n=1}^{\infty}$ – последовательность случайных величин таких, что:

$$\lim_{n\to\infty} \frac{1}{n^2} D\left[\sum_{i=1}^{n} \xi_i\right] = 0 ,$$

тогда

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} M[\xi_i], \text{ при } n \to \infty .$$

Утверждение (неравенство Чебышева)

Пусть случайная величина ξ имеет конечную дисперсию, $D[\xi] < \infty$, тогда:

$$P\{|\xi - M[\xi]| \geq \varepsilon\} \leq \frac{D[\xi]}{\varepsilon^2} .$$

Утверждение 2.7.

Пусть $\{\xi_n\}_{n=1}^{\infty}$ – последовательность случайных величин таких, что:

1) $\lim\limits_{n\to\infty} M[\xi_n] = m < \infty$,

2) $\lim\limits_{n\to\infty} D[\xi_n] = 0$,

тогда,

$$\xi_n \xrightarrow{P} m \text{ , при } n \to \infty .$$

Доказательство:

Согласно определению сходимости по вероятности последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ сходится по вероятности к величине m , если:

$$\forall \varepsilon > 0 \forall \delta > 0 \exists N(\varepsilon,\delta) \forall n > N(\varepsilon,\delta) : P\{| \xi_n - m |< \varepsilon\} > 1 - \delta . \qquad (2.3)$$

Пусть величины $\varepsilon > 0$ и $\delta > 0$ выбраны произвольным образом и зафиксированы, покажем, что действительно существует число $N(\varepsilon,\delta)$, удовлетворяющее (2.3).

Прежде всего, заметим, что при любом n справедливо неравенство:

$$| \xi_n - m |<| \xi_n - M[\xi_n] + M[\xi_n] - m |<| \xi_n - M[\xi_n] | + | M[\xi_n] - m | . \qquad (2.4)$$

Пусть события $A_n(\varepsilon)$ образованы элементарными исходами ω , при которых $|\xi_n - m| < \varepsilon$:

$$A(\varepsilon) = \{\omega :| \xi_n(\omega) - m |< \varepsilon\} ,$$

и события $B_n(\varepsilon)$ образованы элементарными исходами ω , при которых одновременно $| \xi_n - M[\xi_n] |< \frac{\varepsilon}{2}$ и $| M[\xi_n] - m |< \frac{\varepsilon}{2}$:

$$B(\varepsilon) = \left\{\omega :| \xi_n(\omega) - M[\xi_n] |< \frac{\varepsilon}{2}, | M[\xi_n] - m |< \frac{\varepsilon}{2}\right\} .$$

Легко видеть, что если при некотором элементарном исходе $\omega^* \in B_n(\varepsilon)$ выполняются неравенства:

$$\begin{cases} | \xi_n - M[\xi_n] |< \frac{\varepsilon}{2} \\ | M[\xi_n] - m |< \frac{\varepsilon}{2} \end{cases},$$

тогда в силу (2.4),

$$| \xi_n(\omega^*) - m |<| \xi_n(\omega^*) - M[\xi_n] | + | M[\xi_n] - m |< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon ,$$

$$|\,\xi_n(\omega^*) - m\,| < \varepsilon\,,$$

и следовательно $\omega^* \in A(\varepsilon)$. Таким образом, имеет место вложенность событий:

$$A_n(\varepsilon) \supseteq B_n(\varepsilon)\,,$$

откуда следует неравенство для вероятностей:

$$P\{A_n(\varepsilon)\} \ge P\{B_n(\varepsilon)\}\,.$$

Теперь покажем, что специальным выбором n можно сделать вероятности $P\{B_n(\varepsilon)\}$ больше, чем $1-\delta$.

Каждое событие $B_n(\varepsilon)$ можно представить в виде пересечения двух событий $C_{n,1}(\varepsilon)$ и $C_{n,2}(\varepsilon)$:

$$B_n(\varepsilon) = C_{n,1}(\varepsilon) \cap C_{n,2}(\varepsilon)\,,$$

$$C_{n,1}(\varepsilon) = \left\{\omega : |\,\xi_n(\omega) - M[\xi_n]\,| < \frac{\varepsilon}{2}\right\},$$

$$C_{n,2}(\varepsilon) = \left\{\omega : |\,M[\xi_n] - m\,| < \frac{\varepsilon}{2}\right\}.$$

По условию 1) теоремы:

$$\lim_{n\to\infty} M[\xi_n] = m\,,$$

откуда по определению предела для величины $\dfrac{\varepsilon}{2}$ существует число $N_2(\varepsilon)$ такое, что:

$$\forall n > N_2(\varepsilon) : |\,M[\xi_n] - m\,| < \frac{\varepsilon}{2}\,. \tag{2.5}$$

Заметим, что последнее неравенство выполняется при любом элементарном исходе ω, поскольку величина $|\,M[\xi_n] - m\,|$ не является случайной. Таким образом:

$$\forall n \ge N_2(\varepsilon) : C_{n,2}(\varepsilon) = \Omega\,,$$

где Ω – множество всех элементарных исходов и, следовательно:

$$\forall n \ge N_2(\varepsilon) : B_n(\varepsilon) = C_{n,1}(\varepsilon) \cap C_{n,2}(\varepsilon) = C_{n,1}(\varepsilon) \cap \Omega = C_{n,1}(\varepsilon)\,. \tag{2.6}$$

По условию 2) теоремы:

$$\lim_{n\to\infty} D[\xi_n] = 0\,,$$

откуда следует, что существует число N_D такое, что:

$$\forall n > N_D : D[\xi_n] < \infty\,.$$

Если дисперсия $D[\xi_n]$ случайной величины ξ_n конечна, тогда для случайной величины ξ_n справедливо неравенство Чебышева, согласно которому:

$$P\left\{\left|\xi_n - M[\xi_n]\right| \ge \frac{\varepsilon}{2}\right\} \le \frac{D[\xi_n]}{\left(\dfrac{\varepsilon}{2}\right)^2},$$

$$P\left\{\left|\xi_n - M[\xi_n]\right| < \frac{\varepsilon}{2}\right\} \ge 1 - \frac{D[\xi_n]}{\left(\dfrac{\varepsilon}{2}\right)^2}. \tag{2.7}$$

В силу условия 2) теоремы для величины $\left(\dfrac{\varepsilon}{2}\right)^2 \delta$ найдется число $N_1(\varepsilon, \delta)$ такое, что:

$$\forall n > N_1(\varepsilon,\delta) : D[\xi_n] < \left(\frac{\varepsilon}{2}\right)^2 \delta ,$$

тогда из (2.7):

$$\forall n > \max\{ N_1(\varepsilon,\delta), N_D(\varepsilon)\} :$$

$$P\left\{\left|\xi_n - M[\xi_n]\right| < \frac{\varepsilon}{2}\right\} \ge 1 - \frac{D[\xi_n]}{\left(\frac{\varepsilon}{2}\right)^2} > 1 - \frac{\left(\frac{\varepsilon}{2}\right)^2 \delta}{\left(\frac{\varepsilon}{2}\right)^2} = 1 - \delta . \tag{2.8}$$

Пусть теперь n больше, чем $N_1(\varepsilon,\delta)$, $N_2(\varepsilon)$ и N_D, тогда одновременно выполняются соотношения (2.6) и (2.8), тогда:

$$\forall n > \max\{ N_1(\varepsilon,\delta), N_2(\varepsilon), N_D(\varepsilon)\} : P\{B_n(\varepsilon)\} = P\{C_{n,1}(\varepsilon)\} > 1 - \delta .$$

Таким образом, для любых $\varepsilon > 0$ и $\delta > 0$ существует число $N(\varepsilon,\delta)$:

$$N(\varepsilon,\delta) = \max\{ N_1(\varepsilon,\delta), N_2(\varepsilon), N_D(\varepsilon)\}$$

такое, что:

$$\forall n > N(\varepsilon,\delta) : P\{|\xi_n - m| < \varepsilon\} = P\{A_n(\varepsilon)\} \ge P\{B_n(\varepsilon)\} > 1 - \delta ,$$

откуда по определению сходимости по вероятности:

$$\xi_n \xrightarrow{P} m \text{, при } n \to \infty .$$

Утверждение доказано.

3. Точечное оценивание вероятности события.

Пусть проводится n независимых испытаний, в каждом из которых может произойти некоторое событие A, имеющее вероятность p, которая не известна. Требуется построить оценку неизвестной вероятности p и исследовать свойства построенной оценки.

Пусть $\zeta_n = (\xi_1,..., \xi_n)$ – выборка, в которой каждая случайная величина ξ_i принимает значение равное единице, если в i-ом испытании произошло событие A, и значение равное нулю, если в i-ом испытании событие A не произошло:

$$\xi_i = \begin{cases} 1, & p \\ 0, & 1 - p \end{cases}$$

Случайная величина μ_n количества появлений события A в n испытаниях равна сумме ξ_i:

$$\mu_n = \sum_{i=1}^{n} \xi_i .$$

Возьмем в качестве оценки неизвестной вероятности p случайную величину относительной частоты:

$$T(\zeta_n) = \frac{\mu_n}{n} = \frac{1}{n}\sum_{i=1}^{n} \xi_i .$$

Легко видеть, что $T(\zeta_n)$ является несмещенной оценкой p, действительно:

$$\forall p \in [0;1] : M_p[T(\zeta_n)] = M_p\left[\frac{1}{n}\sum_{i=1}^{n} \xi_i\right] = \frac{1}{n}\sum_{i=1}^{n} M_p[\xi_i] = \frac{1}{n}\sum_{i=1}^{n} p = \frac{1}{n} np = p .$$

Согласно теореме Бернулли имеет место сходимость по вероятности случайной величины $\frac{\mu_n}{n}$ к вероятности p, отсюда следует, что оценка $T(\zeta_n)$ является состоятельной.

Вычислим дисперсию оценки $T(\zeta_n)$:

$$D_p[T(\zeta_n)] = D_p\left[\frac{1}{n}\sum_{i=1}^{n}\xi_i\right] = \frac{1}{n^2}D_p\left[\sum_{i=1}^{n}\xi_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}D_p[\xi_i] =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}p(1-p) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n},$$

где $D_p\left[\sum_{i=1}^{n}\xi_i\right] = \sum_{i=1}^{n}D_p[\xi_i]$, поскольку величины выборки $\xi_1, \ldots, \xi_n$ являются независимыми случайными величинами.

4. Точечное оценивание значений функции распределения.

Пусть $\zeta_n = (\xi_1,..., \xi_n)$ – выборка из распределения $F_\xi(x\,|\,\theta)$ с неизвестным параметром θ, $\theta \in \Theta$, и x^* некоторое фиксированное числовое значение, требуется построить оценку значения функции распределения – неизвестной величины $\tau(\theta) = F_\xi(x^*\,|\,\theta)$ (неизвестной в силу того, что параметр θ неизвестен) и исследовать свойства несмещенности и состоятельности построенной оценки.

Предположим, что в качестве оценки неизвестной величины вероятности $F_\xi(x^*\,|\,\theta)$ используется значение эмпирической функции распределения $F_n^*(x^*;\zeta_n)$,

$$T(\zeta_n) = F_n^*(x^*;\zeta_n) = \frac{\mu_n(x^*;\zeta_n)}{n},$$

где согласно определению эмпирической функции распределения 1.6 функция $\mu_n(x;\zeta_n)$ равна количеству случайных величин выборки $\zeta_n = (\xi_1,..., \xi_n)$ меньших x. Заметим, что функцию $\mu_n(x;\zeta_n)$ можно представить в виде суммы значений индикаторных функций от случайных величин выборки:

$$\mu_n(x;\zeta_n) = I(x;\xi_1) + I(x;\xi_2) + ... + I(x;\xi_n),$$

где $I(x;\xi_i)$ $(i = \overline{1,n})$ принимает значение 1 если $\xi_i < x$ и 0 в противном случае. Таким образом, каждая величина $I(x;\xi_i)$ является случайной величиной, принимающей лишь два значения: 1 с вероятностью $P\{\xi_i < x\}$ и 0 с вероятностью $1 - P\{\xi_i < x\}$:

$$I(x;\xi_i) = \begin{cases} 1, & P\{\xi_i < x\} \\ 0, & 1 - P\{\xi_i < x\} \end{cases}.$$

Поскольку $\zeta_n = (\xi_1,..., \xi_n)$ выборка из распределения $F_\xi(x\,|\,\theta)$, то в соответствии с определением выборки 1.1, все случайные величины имеют функцию распределения $F_\xi(x\,|\,\theta)$, отсюда следует, что $P\{\xi_i < x\} = F(x\,|\,\theta)$,

$$I(x;\xi_i) = \begin{cases} 1, & F_\xi(x\,|\,\theta) \\ 0, & 1 - F_\xi(x\,|\,\theta) \end{cases}$$

Таким образом, окончательно эмпирической функции распределения $F_n^*(x^*;\zeta_n)$ можно придать вид:

$$F_n^*(x^*;\zeta_n) = \frac{1}{n}\sum_{i=1}^{n}I(x^*;\xi_i) \qquad (2.9)$$

где $I(x^*;\xi_i)$ - случайные величины,

$$I(x^*;\xi_i) = \begin{cases} 1, & F_\xi(x^* \mid \theta) \\ 0, & 1 - F_\xi(x^* \mid \theta) \end{cases}.$$

Исследуем свойства оценки (2.9), покажем, что эмпирическая функция распределения является несмещенной оценкой $F_\xi(x^* \mid \theta)$, действительно, по свойству математического ожидания,

$$\forall\, \theta \in \Theta:\ M_\theta[F_n^*(x^*;\zeta_n)] = M_\theta\left[\frac{1}{n}\sum_{i=1}^{n} I(x^*;\xi_i)\right] = \frac{1}{n}\sum_{i=1}^{n} M_\theta\left[I(x^*;\xi_i)\right] =$$

$$= \frac{1}{n}\sum_{i=1}^{n} F_\xi(x^* \mid \theta) = \frac{1}{n} nF_\xi(x^* \mid \theta) = F_\xi(x^* \mid \theta).$$

Для исследования свойства состоятельности оценки (2.9) достаточно вспомнить теорему о сходимости по вероятности значений эмпирической функции распределения $F_n^*(x^*;\zeta_n)$ к значениям $F_\xi(x^* \mid \theta)$ при всяком фиксированном x^*. Отличие условий теоремы от рассматриваемого случая состоит лишь в том, что функция распределения $F_\xi(x \mid \theta)$ зависит от параметра θ. Тем не менее, при всяком допустимом и фиксированном значении параметра $\theta^* \in \Theta$ функция распределения $F_\xi(x \mid \theta^*)$ становится функцией только аргумента x и выполняются условия указанной выше теоремы, из которой следует сходимость по вероятности значения эмпирической функции распределения $F_n^*(x^*;\zeta_n)$ к значению функции распределения $F_\xi(x^* \mid \theta^*)$:

$$F_n^*(x^*;\zeta_n) \xrightarrow{P} F_\xi(x^* \mid \theta^*) \text{ при } n \to \infty .$$

Никаких ограничений на выбор значения θ^* параметра из множества допустимых значений Θ не накладывается, поэтому сходимость имеет место при любом допустимом значении параметра θ:

$$\forall\, \theta \in \Theta:\ F_n^*(x^*;\zeta_n) \xrightarrow{P} F_\xi(x^* \mid \theta) \text{ при } n \to \infty ,$$

откуда значение эмпирической функции распределения $F_n^*(x^*;\zeta_n)$ является состоятельной оценкой $F_\xi(x^* \mid \theta)$ при всяком фиксированном x^* по определению состоятельной оценки (2.9).

Таким образом, значение эмпирической функции распределения $F_n^*(x^*;\zeta_n)$ в любой точке x^* является несмещенной и состоятельной оценкой значения функции распределения $F_\xi(x^* \mid \theta)$ (в этой же точке x^*).

Вычислим дисперсию оценки (2.9):

$$D_\theta[F_n^*(x^*;\zeta_n)] = D_\theta\left[\frac{1}{n}\sum_{i=1}^{n} I(x^*;\xi_i)\right] = \frac{1}{n^2} D_\theta\left[\sum_{i=1}^{n} I(x^*;\xi_i)\right] = \frac{1}{n^2}\sum_{i=1}^{n} D_\theta\left[I(x^*;\xi_i)\right] =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} F_\xi(x^* \mid \theta)(1 - F_\xi(x^* \mid \theta)) = \frac{1}{n^2} nF_\xi(x^* \mid \theta)(1 - F_\xi(x^* \mid \theta)) =$$

$$= \frac{F_\xi(x^* \mid \theta)(1 - F_\xi(x^* \mid \theta))}{n},$$

где $D_\theta\left[\sum_{i=1}^{n} I(x^*;\xi_i)\right] = \sum_{i=1}^{n} D_\theta\left[I(x^*;\xi_i)\right]$ поскольку величины $I(x^*;\xi_i)$ являются независимыми случайными величинами, так как исходные величины ξ_1, …, ξ_n образуют выборку и по определению выборки являются независимыми случайными величинами.

5. Точечное оценивание математического ожидания и дисперсии.

Пусть $\zeta_n = (\xi_1,..., \xi_n)$ – выборка из распределения $F_\xi(x \mid \theta)$ с неизвестным параметром θ, принимающим значения из множества допустимых значений Θ. Требуется построить оценки первого начального момента (математического ожидания) $m_1(\theta)$ и второго центрального момента (дисперсии) $\mu_2(\theta)$ (при условии, что указанные моменты конечны):

$$\forall\, \theta \in \Theta : m_1(\theta) = \int_{-\infty}^{\infty} x dF_\xi(x \mid \theta) < \infty ,$$

$$\forall\, \theta \in \Theta : \mu_2(\theta) = \int_{-\infty}^{\infty} \left(x - m_1(\theta)\right)^2 dF_\xi(x \mid \theta) < \infty$$

и исследовать свойства построенных оценок.

Для построения оценок воспользуемся определениями моментов, приведенными выше, в которых неизвестную функцию распределения $F_\xi(x \mid \theta)$ заменим известным «приближением» – эмпирической функцией распределения $F_n^*(x;\zeta_n)$:

$$\hat{m}_1(\zeta_n) = \int_{-\infty}^{\infty} x dF_n^*(x;\zeta_n) ,$$
$$\hat{\mu}_2(\zeta_n) = \int_{-\infty}^{\infty} \left(x - \hat{m}_1(\zeta_n)\right)^2 dF_n^*(x;\zeta_n) \qquad (2.10)$$

Эмпирическая функция распределения $F_n^*(x;\zeta_n)$, вообще говоря, является случайной функцией и строгого определения интеграла по случайной функции не приводится. Тем не менее, интегралам, указанным выше, можно придать следующую интерпретацию: при всяком фиксированном элементарном исходе $\omega^* \in \Omega$ случайные величины $\xi_i(\omega)$ принимают числовые значения $x_i = \xi_i(\omega^*)$, и соответствующая реализация эмпирической функции распределения $F_n^*(x; x_1,..., x_n)$ является обычной детерминированной функцией, для которой указанные интегралы следует понимать как интегралы Стилтьеса. Поскольку реализация $F_n^*(x; x_1,..., x_n)$ является кусочно-постоянной функцией, с разрывами величины $1/n$ в точках x_i $(i = \overline{1,n})$, то в результате вычисления интегралов получим следующие значения:

$$\hat{m}_1(x_1,..., x_n) = \int_{-\infty}^{\infty} x dF_n^*(x; x_1,..., x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i ,$$

$$\hat{\mu}_2(x_1,..., x_n) = \int_{-\infty}^{\infty} \left(x - \hat{m}_1(x_1,..., x_n)\right)^2 dF_n^*(x; x_1,..., x_n) =$$
$$= \frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{m}_1(x_1,..., x_n))^2 .$$

Легко видеть, что приведенные вычисления можно провести для любого элементарного исхода $\omega \in \Omega$, поэтому для интегралов (2.10) получим следующие выражения:

$$\hat{m}_1(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \xi_i \qquad (2.11)$$

$$\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - \hat{m}_1(\zeta_n))^2 \qquad (2.12)$$

Определение 2.8.

Статистика $\hat{m}_1(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \xi_i$ называется *выборочным средним*.

Определение 2.9.

Статистика $\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - \hat{m}_1(\zeta_n))^2$ называется *выборочной дисперсией*.

Легко видеть, что оценка $\hat{m}_1(\zeta_n)$ является несмещенной, действительно:

$$M_\theta\left[\hat{m}_1(\zeta_n)\right] = M_\theta\left[\frac{1}{n}\sum_{i=1}^{n} \xi_i\right] = \frac{1}{n}\sum_{i=1}^{n} M_\theta\left[\xi_i\right] = \frac{1}{n}\sum_{i=1}^{n} m_1(\theta) = \frac{1}{n} nm_1(\theta) = m_1(\theta)\ .$$

Для доказательства состоятельности оценки $\hat{m}_1(\zeta_n)$ может быть использована теорема Хинчина: в данном случае величины ξ_1, ξ_2, …, ξ_n образуют выборку, поэтому они независимы в совокупности и имеют одинаковую функцию распределения, и, следовательно, одинаковое математическое ожидание $m_1(\theta)$, конечность которого гарантируется исходной постановкой рассматриваемой задачи оценки. Таким образом, к совокупности случайных величин ξ_1, …, ξ_n при $n \to \infty$ применима теорема Хинчина, и поскольку статистика $\hat{m}_1(\zeta_n)$ есть в точности $\frac{1}{n}\sum_{i=1}^{n} \xi_i$ получим:

$$\hat{m}_1(\zeta_n) \xrightarrow{P} m_1(\theta), \text{ при } n \to \infty\ ,$$

что и означает состоятельность оценки $\hat{m}_1(\zeta_n)$.

Вычислим $D_\theta[\hat{m}_1(\zeta_n)]$ непосредственно из определения статистики $\hat{m}_1(\zeta_n)$:

$$D_\theta\left[\hat{m}_1(\zeta_n)\right] = D_\theta\left[\frac{1}{n}\sum_{i=1}^{n} \xi_i\right] = \frac{1}{n^2} D_\theta\left[\sum_{i=1}^{n} \xi_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{cov}_\theta(\xi_i, \xi_j)$$

поскольку величины ξ_1, …, ξ_n образуют выборку, то согласно определению выборки они независимы в совокупности и, следовательно, попарно независимы, так что при $i \neq j$ $\operatorname{cov}_\theta(\xi_i, \xi_j) = 0$. При $i = j$, конечно, $\operatorname{cov}_\theta(\xi_i, \xi_i) = D_\theta\left[\xi_i\right]$, поэтому:

$$D_\theta\left[\hat{m}_1(\zeta_n)\right] = \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{cov}_\theta(\xi_i, \xi_i) = \frac{1}{n^2}\sum_{i=1}^{n} D_\theta\left[\xi_i\right] = \frac{1}{n^2} n\mu_2(\theta) = \frac{\mu_2(\theta)}{n} \qquad (2.13)$$

Исследуем свойства оценки $\hat{\mu}_2(\zeta_n)$, предварительно преобразовав статистику $\hat{\mu}_2(\zeta_n)$ к следующему виду:

$$\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - \hat{m}_1(\zeta_n))^2 = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta) + m_1(\theta) - \hat{m}_1(\zeta_n))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} \left\{(\xi_i - m_1(\theta))^2 + 2(\xi_i - m_1(\theta))(m_1(\theta) - \hat{m}_1(\zeta_n)) + (m_1(\theta) - \hat{m}_1(\zeta_n))^2\right\} =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 + 2\frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))(m_1(\theta) - \hat{m}_1(\zeta_n)) + \frac{1}{n}\sum_{i=1}^{n} (m_1(\theta) - \hat{m}_1(\zeta_n))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 - 2(\hat{m}_1(\zeta_n) - m_1(\theta))\frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta)) + \frac{1}{n} n(m_1(\theta) - \hat{m}_1(\zeta_n))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 - 2(\hat{m}_1(\zeta_n) - m_1(\theta))\left(\frac{1}{n}\sum_{i=1}^{n} \xi_i - \frac{1}{n}\sum_{i=1}^{n} m_1(\theta)\right) + (m_1(\theta) - \hat{m}_1(\zeta_n))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 - 2(\hat{m}_1(\zeta_n) - m_1(\theta))\left(\hat{m}_1(\zeta_n) - \frac{1}{n} nm_1(\theta)\right) + (m_1(\theta) - \hat{m}_1(\zeta_n))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 - 2(\hat{m}_1(\zeta_n) - m_1(\theta))^2 + (m_1(\theta) - \hat{m}_1(\zeta_n))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\xi_i - m_1(\theta))^2 - (\hat{m}_1(\zeta_n) - m_1(\theta))^2 \qquad (2.14)$$

Вычислим математическое ожидание $M_\theta[\hat{\mu}_2(\zeta_n)]$:

$$M_\theta[\hat{\mu}_2(\zeta_n)] = M_\theta\left[\frac{1}{n}\sum_{i=1}^{n}(\xi_i - m_1(\theta))^2 - (\hat{m}_1(\zeta_n) - m_1(\theta))^2\right] =$$

$$= \frac{1}{n}M_\theta\left[\sum_{i=1}^{n}(\xi_i - m_1(\theta))^2\right] - M_\theta\left[(\hat{m}_1(\zeta_n) - m_1(\theta))^2\right].$$

Поскольку $M_\theta[\hat{m}_1(\zeta_n)] = m_1(\theta)$, то

$$M_\theta\left[(\hat{m}_1(\zeta_n) - m_1(\theta))^2\right] = M_\theta\left[(\hat{m}_1(\zeta_n) - M_\theta[\hat{m}_1(\zeta_n)])^2\right] = D_\theta[\hat{m}_1(\zeta_n)],$$

тогда,

$$M_\theta[\hat{\mu}_2(\zeta_n)] = \frac{1}{n}\sum_{i=1}^{n}M_\theta\left[(\xi_i - m_1(\theta))^2\right] - D_\theta[\hat{m}_1(\zeta_n)] = \frac{1}{n}\sum_{i=1}^{n}D_\theta[\xi_i] - D_\theta[\hat{m}_1(\zeta_n)] =$$

(дисперсия $D_\theta[\hat{m}_1(\zeta_n)]$ была вычислена ранее – соотношение 2.13)

$$= \frac{1}{n}\sum_{i=1}^{n}\mu_2(\theta) - \frac{\mu_2(\theta)}{n} = \frac{1}{n}n\mu_2(\theta) - \frac{\mu_2(\theta)}{n} = \mu_2(\theta) - \frac{\mu_2(\theta)}{n} = \frac{n-1}{n}\mu_2(\theta)$$

Таким образом,

$$M_\theta[\hat{\mu}_2(\zeta_n)] = \frac{n-1}{n}\mu_2(\theta) \qquad (2.15)$$

и оценка $\hat{\mu}_2(\zeta_n)$ оказывается смещенной.

Для доказательства состоятельности оценки $\hat{\mu}_2(\theta)$ может быть использована теорема Хинчина и свойства сходимости по вероятности. Ранее было получено выражение (2.14) для оценки $\hat{\mu}_2(\zeta_n)$ в виде:

$$\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - m_1(\theta))^2 - (\hat{m}_1(\zeta_n) - m_1(\theta))^2.$$

К совокупности случайных величин $(\xi_i - m_1(\theta))^2$ применима теорема Хинчина: величины $(\xi_i - m_1(\theta))^2$ независимы и имеют одинаковую функцию распределения (поскольку величины ξ_i образуют выборку), кроме того, математическое ожидание каждой величины $M_\theta[(\xi_i - m_1(\theta))^2] = \mu_2(\theta)$ конечно (по условию постановки исходной задачи, приведенной в начале пункта). Таким образом, по теореме Хинчина:

$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i - m_1(\theta))^2 \xrightarrow{P} \mu_2(\theta), \text{ при } n \to \infty.$$

Оценка $\hat{m}_1(\zeta_n)$ является состоятельной, что по определению означает:

$$\hat{m}_1(\zeta_n) \xrightarrow{P} m_1(\theta), \text{ при } n \to \infty$$

откуда по определению сходимости по вероятности

$$\hat{m}_1(\zeta_n) - m_1(\theta) \xrightarrow{P} 0, \text{ при } n \to \infty.$$

Функция возведения в квадрат является непрерывной, и, как и для всякой непрерывной функции, в силу свойства сходимости по вероятности:

$$(\hat{m}_1(\zeta_n) - m_1(\theta))^2 \xrightarrow{P} 0^2, \text{ при } n \to \infty.$$

В силу свойства суммы двух пределов по вероятности,

$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i - m_1(\theta))^2 - (\hat{m}_1(\zeta_n) - m_1(\theta))^2 \xrightarrow{P} \mu_2(\theta) - 0, \text{ при } n \to \infty,$$

откуда следует, что

$$\hat{\mu}_2(\zeta_n) \xrightarrow{P} \mu_2(\theta), \text{ при } n \to \infty,$$

то есть оценка $\hat{\mu}_2(\zeta_n)$ в соответствии с определением является состоятельной.

Вычислим дисперсию $D_\theta[\hat{\mu}_2(\zeta_n)]$ статистики $\hat{\mu}_2(\zeta_n)$, дополнительно предполагая, что существует четвертый центральный момент $\mu_4(\theta)$:

$$\forall\, \theta \in \Theta : \mu_4(\theta) = \int_{-\infty}^{\infty} (x - m_1(\theta))^4 \, dF_\xi(x;\theta) < \infty \, .$$

Подставляя выражение для $\hat{m}_1(\zeta_n)$ в (2.14) получим:

$$\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 - \left(\frac{1}{n}\sum_{i=1}^{n} \xi_i - m_1(\theta)\right)^2 = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))^2 - \left(\frac{1}{n}\sum_{i=1}^{n} (\xi_i - m_1(\theta))\right)^2 .$$

Введем центрированные случайные величины $\eta_i = \xi_i - m_1(\theta)$, тогда для $\hat{\mu}_2(\zeta_n)$ получим выражение:

$$\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \eta_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} \eta_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n} \eta_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} \eta_i\right)^2$$

Теперь, возводя в квадрат и преобразовывая, получим выражение:

$$\hat{\mu}_2(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \eta_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} \eta_i^2 + \sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k\right) = \frac{1}{n}\sum_{i=1}^{n} \eta_i^2 - \frac{1}{n^2}\sum_{i=1}^{n} \eta_i^2 - \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k =$$

$$= \frac{n-1}{n^2}\sum_{i=1}^{n} \eta_i^2 - \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k \, .$$

Представим дисперсию $D_\theta[\hat{\mu}_2(\zeta_n)]$ в виде:

$$D_\theta[\hat{\mu}_2(\zeta_n)] = M_\theta[(\hat{\mu}_2(\zeta_n))^2] - \left(M_\theta[\hat{\mu}_2(\zeta_n)]\right)^2 \qquad (2.16)$$

Вычислим $M_\theta[(\hat{\mu}_2(\zeta_n))^2]$:

$$M_\theta[(\hat{\mu}_2(\zeta_n))^2] = M_\theta\left[\left(\frac{n-1}{n^2}\sum_{i=1}^{n} \eta_i^2 - \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k\right)^2\right] =$$

$$= M_\theta\left[\frac{(n-1)^2}{n^4}\left(\sum_{i=1}^{n} \eta_i^2\right)^2 - 2\,\frac{n-1}{n^4}\sum_{i=1}^{n} \eta_i^2 \sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k + \frac{1}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n}\sum_{p=1}^{n}\sum_{\substack{q=1,\\q\neq p}}^{n} \eta_j\eta_k\eta_p\eta_q\right] =$$

$$= M_\theta\left[\frac{(n-1)^2}{n^4}\left(\sum_{i=1}^{n} \eta_i^4 + \sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j^2\eta_k^2\right) - 2\,\frac{n-1}{n^4}\sum_{i=1}^{n} \eta_i^2 \sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k + \frac{1}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n}\sum_{p=1}^{n}\sum_{\substack{q=1,\\q\neq p}}^{n} \eta_j\eta_k\eta_p\eta_q\right] =$$

$$= \frac{(n-1)^2}{n^4} M_\theta\left[\sum_{i=1}^{n} \eta_i^4\right] + \frac{(n-1)^2}{n^4} M_\theta\left[\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j^2\eta_k^2\right] - 2\,\frac{n-1}{n^4} M_\theta\left[\sum_{i=1}^{n} \eta_i^2 \sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} \eta_j\eta_k\right] +$$

$$+ \frac{1}{n^4} M_\theta\left[\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n}\sum_{p=1}^{n}\sum_{\substack{q=1,\\q\neq p}}^{n} \eta_j\eta_k\eta_p\eta_q\right] =$$

$$= \frac{(n-1)^2}{n^4}\sum_{i=1}^{n} M_{\theta}[\eta_i^4] + \frac{(n-1)^2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\ k\neq j}}^{n} M_{\theta}[\eta_j^2\eta_k^2] - 2\frac{n-1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{k=1,\\ k\neq j}}^{n} M_{\theta}[\eta_i^2\eta_j\eta_k] +$$

$$+\frac{1}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\ k\neq j}}^{n}\sum_{p=1}^{n}\sum_{\substack{q=1,\\ q\neq p}}^{n} M_{\theta}[\eta_j\eta_k\eta_p\eta_q].$$

Заметим, что величины η_i $(i=\overline{1,n})$ совместно независимы, поскольку независимы величины ξ_i $(i=\overline{1,n})$, так как ξ_i образуют выборку, и кроме того:

$$M_{\theta}[\eta_i] = M_{\theta}[\xi_i - m_1(\theta)] = M_{\theta}[\xi_i] - m_1(\theta) = \int_{-\infty}^{\infty} x dF_{\xi}(x;\theta) - m_1(\theta) = m_1(\theta) - m_1(\theta) = 0.$$

Рассмотрим слагаемые $M_{\theta}[\eta_i^2\eta_j\eta_k]$ в сумме $\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{k=1,\\ k\neq j}}^{n} M_{\theta}[\eta_i^2\eta_j\eta_k]$, зафиксируем i, и рассмотрим все возможные варианты для индексов j и k:

1) $j = i$,

а) $k = i$ – невозможно, поскольку $k \neq j$ и $j = i$;

а) $k \neq i$:

$$M_{\theta}[\eta_i^2\eta_j\eta_k] = M_{\theta}[\eta_i^2\eta_i\eta_k] = M_{\theta}[\eta_i^3]M_{\theta}[\eta_k] = M_{\theta}[\eta_i^3]\cdot 0 = 0;$$

2) $j \neq i$,

а) $k = i$:

$$M_{\theta}[\eta_i^2\eta_j\eta_k] = M_{\theta}[\eta_i^2\eta_j\eta_i] = M_{\theta}[\eta_i^3]M[\eta_j] = M_{\theta}[\eta_i^3]\cdot 0 = 0;$$

б) $k \neq i$:

$$M_{\theta}[\eta_i^2\eta_j\eta_k] = M_{\theta}[\eta_i^2]M_{\theta}[\eta_j]M_{\theta}[\eta_i] = M_{\theta}[\eta_i^2]\cdot 0\cdot 0 = 0.$$

Таким образом, все слагаемые $M_{\theta}[\eta_i^2\eta_j\eta_k] = 0$.

Рассмотрим слагаемые $M_{\theta}[\eta_j\eta_k\eta_p\eta_q]$ в сумме $\sum_{j=1}^{n}\sum_{\substack{k=1,\\ k\neq j}}^{n}\sum_{p=1}^{n}\sum_{\substack{q=1,\\ q\neq p}}^{n} M_{\theta}[\eta_j\eta_k\eta_p\eta_q]$, зафиксируем индексы j и $k \neq j$, и рассмотрим все возможные варианты для индексов p и q:

1) $p = j$:

а) $q = j$ – невозможно, поскольку $q \neq p$ и $p = j$;

б) $q = k$:

$$M_{\theta}[\eta_j\eta_k\eta_p\eta_q] = M_{\theta}[\eta_j\eta_k\eta_j\eta_k] = M_{\theta}[\eta_j^2\eta_k^2];$$

в) $q \neq j$, $q \neq k$:

$$M_{\theta}[\eta_j\eta_k\eta_p\eta_q] = M_{\theta}[\eta_j\eta_k\eta_j\eta_q] = M_{\theta}[\eta_j^2]M_{\theta}[\eta_k]M_{\theta}[\eta_q] = M_{\theta}[\eta_j^2]\cdot 0\cdot 0 = 0;$$

2) $p = k$:

а) $q = j$:

$$M_{\theta}[\eta_j\eta_k\eta_p\eta_q] = M_{\theta}[\eta_j\eta_k\eta_k\eta_j] = M_{\theta}[\eta_j^2\eta_k^2];$$

б) $q = k$ – невозможно, поскольку $q \neq p$ и $p = k$;

в) $q \neq j$, $q \neq k$:

$$M_{\theta}[\eta_j\eta_k\eta_p\eta_q] = M_{\theta}[\eta_j\eta_k\eta_k\eta_q] = M_{\theta}[\eta_j]M_{\theta}[\eta_k^2]M_{\theta}[\eta_q] = 0\cdot M_{\theta}[\eta_k^2]\cdot 0 = 0;$$

3) $p \neq j$ и $p \neq k$:

а) $q = j$:

$$M_{\theta}[\eta_j\eta_k\eta_p\eta_q] = M_{\theta}[\eta_j\eta_k\eta_p\eta_j] = M_{\theta}[\eta_j^2]M_{\theta}[\eta_k]M_{\theta}[\eta_p] = M_{\theta}[\eta_j^2]\cdot 0\cdot 0 = 0;$$

б) $q = k$:

$$M_\theta[\eta_j\eta_k\eta_p\eta_q] = M_\theta[\eta_j\eta_k\eta_p\eta_k] = M_\theta[\eta_j]M_\theta[\eta_k^2]M_\theta[\eta_p] = 0\cdot M_\theta[\eta_k^2]\cdot 0 = 0;$$

в) $q \neq j$, $q \neq k$:

$$M_\theta[\eta_j\eta_k\eta_p\eta_q] = M_\theta[\eta_j]M_\theta[\eta_k]M_\theta[\eta_p]M_\theta[\eta_q] = 0\cdot 0\cdot 0\cdot 0 = 0;$$

Заметим, что только в случае 1б) $p = j$, $q = k$ и в случае 2а) $p = k$, $q = j$ слагаемое $M_\theta[\eta_j\eta_k\eta_p\eta_q]$ может быть отлично от нуля, во всех остальных случаях слагаемое $M_\theta[\eta_j\eta_k\eta_p\eta_q] = 0$.

Таким образом, для $M_\theta[(\hat{\mu}_2(\zeta_n))^2]$ получим выражение:

$$M_\theta[(\hat{\mu}_2(\zeta_n))^2] = \frac{(n-1)^2}{n^4}\sum_{i=1}^{n} M_\theta[\eta_i^4] + \frac{(n-1)^2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} M_\theta[\eta_j^2\eta_k^2] +$$

$$+\frac{1}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n}(M_\theta[\eta_j\eta_k\eta_j\eta_k] + M_\theta[\eta_j\eta_k\eta_k\eta_j]) =$$

$$= \frac{(n-1)^2}{n^4}\sum_{i=1}^{n} M_\theta[\eta_i^4] + \frac{(n-1)^2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} M_\theta[\eta_j^2\eta_k^2] + \frac{1}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n}(M_\theta[\eta_j^2\eta_k^2] + M_\theta[\eta_j^2\eta_k^2]) =$$

$$= \frac{(n-1)^2}{n^4}\sum_{i=1}^{n} M_\theta[\eta_i^4] + \frac{(n-1)^2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} M_\theta[\eta_j^2\eta_k^2] + \frac{2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} M_\theta[\eta_j^2\eta_k^2] =$$

$$= \frac{(n-1)^2}{n^4}\sum_{i=1}^{n} M_\theta[\eta_i^4] + \frac{(n-1)^2+2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n} M_\theta[\eta_j^2\eta_k^2].$$

Поскольку $\eta_i = \xi_i - m_1(\theta)$, то

$$M_\theta[\eta_i^4] = M_\theta[(\xi_i - m_1(\theta))^4] = \int_{-\infty}^{\infty}(x - m_1(\theta))^4 dF_\xi(x;\theta) = \mu_4(\theta)$$

Поскольку величины η_i $(i = \overline{1,n})$ независимы, то $M_\theta[\eta_j^2\eta_k^2] = M_\theta[\eta_j^2]M_\theta[\eta_k^2]$, тогда:

$$M_\theta[\eta_j^2\eta_k^2] = M_\theta[\eta_j^2]M_\theta[\eta_k^2] = M_\theta[(\xi_j - m_1(\theta))^2]M_\theta[(\xi_k - m_1(\theta))^2] =$$

$$= \int_{-\infty}^{\infty}(x - m_1(\theta))^2 dF_\xi(x\mid\theta)\int_{-\infty}^{\infty}(x - m_1(\theta))^2 dF_\xi(x\mid\theta) = \mu_2(\theta)\mu_2(\theta) = \mu_2^2(\theta).$$

Таким образом,

$$M_\theta[(\hat{\mu}_2(\zeta_n))^2] = \frac{(n-1)^2}{n^4}\sum_{i=1}^{n}\mu_4(\theta) + \frac{(n-1)^2+2}{n^4}\sum_{j=1}^{n}\sum_{\substack{k=1,\\k\neq j}}^{n}\mu_2^2(\theta) =$$

$$= \frac{(n-1)^2}{n^4}n\mu_4(\theta) + \frac{(n-1)^2+2}{n^4}n(n-1)\mu_2^2(\theta) = \frac{(n-1)^2}{n^3}\mu_4(\theta) + \frac{((n-1)^2+2)(n-1)}{n^3}\mu_2^2(\theta).$$

Из (2.16) с учетом (2.15) получим выражение для дисперсии $D_\theta[\hat{\mu}_2(\zeta_n)]$:

$$D_\theta[\hat{\mu}_2(\zeta_n)] = \frac{(n-1)^2}{n^3}\mu_4(\theta) + \frac{((n-1)^2+2)(n-1)}{n^3}\mu_2^2(\theta) - \frac{(n-1)^2}{n^2}\mu_2^2(\theta) =$$

$$= \frac{(n-1)^2}{n^3}\mu_4(\theta) + \frac{(n-1)}{n^2}((n-1)^2 + 2 - n(n-1))\mu_2^2(\theta) =$$

$$= \frac{(n-1)^2}{n^3}\mu_4(\theta) + \frac{(n-1)}{n^2}(n^2 - 2n + 1 + 2 - n^2 + n)\mu_2^2(\theta) =$$

$$= \frac{(n-1)^2}{n^3}\mu_4(\theta) + \frac{(n-1)(n-3)}{n^2}\mu_2^2(\theta) =$$

$$= \frac{(n-1)^2}{n^3}\left[\mu_4(\theta) + \frac{n-3}{n-1}\mu_2^2(\theta)\right] \qquad (2.17)$$

Оценка $\hat{\mu}_2(\zeta_n)$ является состоятельной, но смещенной, тем не менее, смещение оценки $\hat{\mu}_2(\zeta_n)$ легко исправить, в результате приходим к новой оценке:

$$\tilde{\mu}_2(\zeta_n) = \frac{n}{n-1}\hat{\mu}_2(\zeta_n) = \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}(\xi_i - \hat{m}_1(\zeta_n))^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \hat{m}_1(\zeta_n))^2 .$$

Оценка $\tilde{\mu}_2(\zeta_n)$ является несмещенной, действительно,

$$M_\theta[\tilde{\mu}_2(\zeta_n)] = M_\theta\left[\frac{n}{n-1}\tilde{\mu}_2(\zeta_n)\right] = \frac{n}{n-1}M_\theta\left[\hat{\mu}_2(\zeta_n)\right] = \frac{n}{n-1}\frac{n-1}{n}\mu_2(\theta) = \mu_2(\theta) .$$

Определение 2.10.

Статистика $\tilde{\mu}_2(\zeta_n) = \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \hat{m}_1(\zeta_n))^2$ называется *исправленной выборочной дисперсией*.

Для доказательства состоятельности оценки $\tilde{\mu}_2(\zeta_n)$ могут быть использованы свойства сходимости по вероятности. В соответствии с определением оценки $\tilde{\mu}_2(\zeta_n)$:

$$\tilde{\mu}_2(\zeta_n) = \frac{n}{n-1}\hat{\mu}_2(\zeta_n) ,$$

где оценка $\hat{\mu}_2(\zeta_n)$ состоятельна, то есть $\hat{\mu}_2(\zeta_n) \xrightarrow{P} \mu_2(\theta)$ при $n \to \infty$, и числовая последовательность $a_n = \frac{n}{n-1}$ стремиться к 1 при $n \to \infty$. Последнее означает, что для любого $\varepsilon > 0$ всегда можно найти N такое, что при $n \ge N$:

$$|a_n - 1| < \varepsilon ,$$

отсюда, считая элементы последовательности a_n функциями $a_n(\omega) = \frac{n}{n-1}$, тождественно равные постоянным, получим:

$$P\{|a_n - 1| < \varepsilon\} = 1 ,$$

тогда, очевидно, что и для всякого $\delta > 0$:

$$P\{|a_n - 1| < \varepsilon\} > 1 - \delta .$$

Таким образом, числовая последовательность a_n сходится по вероятности к 1 при $n \to \infty$, $a_n \xrightarrow{P} 1$. Оценка $\tilde{\mu}_2(\zeta_n) = a_n\hat{\mu}_2(\zeta_n)$, отсюда по свойству сходимости по вероятности:

$$\tilde{\mu}_2(\zeta_n) = a_n\hat{\mu}_2(\zeta_n) \xrightarrow{P} 1 \cdot \mu_2(\theta) = \mu_2(\theta) \text{, при } n \to \infty .$$

Зная выражение для дисперсии $D_\theta[\hat{\mu}_2(\zeta_n)]$ (2.17), легко вычислить дисперсию $D_\theta[\tilde{\mu}_2(\zeta_n)]$ оценки $\tilde{\mu}_2(\zeta_n)$:

$$D_\theta[\tilde{\mu}_2(\zeta_n)] = D_\theta\left[\frac{n}{n-1}\hat{\mu}_2(\zeta_n)\right] = \frac{n^2}{(n-1)^2}D_\theta\left[\hat{\mu}_2(\zeta_n)\right] = \frac{n^2}{(n-1)^2}\frac{(n-1)^2}{n^3}\left[\mu_4(\theta) + \frac{(n-3)}{n-1}\mu_2^2(\theta)\right] =$$

$$= \frac{1}{n}\left[\mu_4(\theta) + \frac{n-3}{n-1}\mu_2^2(\theta)\right].$$

Заметим, что дисперсия оценки $\tilde{\mu}_2(\zeta_n)$ оказалась в $\frac{n^2}{(n-1)^2}$ раз больше дисперсии оценки $\hat{\mu}_2(\zeta_n)$, то есть исправление смещения оценки привело к увеличению дисперсии.

6. Точечное оценивание старших моментов.

Пусть, как и ранее, $\zeta_n = (\xi_1,..., \xi_n)$ – выборка из распределения $F_\xi(x \mid \theta)$ с неизвестным параметром θ, принимающим значения из множества допустимых значений Θ. Требуется построить оценки начальных $m_k(\theta)$ и центральных $\mu_k(\theta)$ моментов при $k \ge 1$, считая, что указанные моменты конечны:

$$\forall\, \theta \in \Theta : m_k(\theta) = \int_{-\infty}^{\infty} x^k dF_\xi(x \mid \theta) < \infty\ ,$$

$$\forall\, \theta \in \Theta : \mu_k(\theta) = \int_{-\infty}^{\infty} (x - m_1(\theta))^k\, dF_\xi(x \mid \theta) < \infty$$

и исследовать свойства построенных оценок.

При построении оценок используются статистики аналогичные введенным ранее статистикам (2.11) и (2.12):

$$\hat{m}_k(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \xi_i^k\ , \tag{2.18}$$

$$\hat{\mu}_k(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - \hat{m}_1(\zeta_n))^k\ , \tag{2.19}$$

$$k \ge 1\,.$$

Определение 2.11.

Статистика $\hat{m}_k(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \xi_i^k$ называется *выборочным моментом* k *-го порядка.*

Определение 2.12.

Статистика $\hat{\mu}_k(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} (\xi_i - \hat{m}_1(\zeta_n))^k$ называется *выборочным центральным моментом* k *-го порядка.*

Легко видеть, что статистики $\hat{m}_k(\zeta_n)$ являются несмещенными оценками $m_k(\zeta_n)$, действительно:

$$M_\theta[\hat{m}_k(\zeta_n)] = M_\theta\left[\frac{1}{n}\sum_{i=1}^{n} \xi_i^k\right] = \frac{1}{n}\sum_{i=1}^{n} M_\theta[\xi_i^k] = \frac{1}{n}\sum_{i=1}^{n} m_k(\theta) = m_k(\theta)\ .$$

Состоятельность оценок $\hat{m}_k(\zeta_n)$ может быть доказана с помощью теоремы Хинчина: случайные величины ξ_i^k независимы в совокупности, имеют одинаковое распределение и конечное математическое ожидание $m_k(\theta)$, тогда по теореме Хинчина:

$$\hat{m}_k(\zeta_n) = \frac{1}{n}\sum_{i=1}^{n} \xi_i^k \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} m_k(\theta) = m_k(\theta)\ ,$$

при $n \to \infty$.

Кроме того, если существует момент $m_{2k}(\theta)$, то дисперсия статистики $D_\theta[\hat{m}_k(\theta)]$:

$$D_\theta[\hat{m}_k(\zeta_n)] = D_\theta\left[\frac{1}{n}\sum_{i=1}^{n} \xi_i^k\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{cov}_\theta(\xi_i^k, \xi_j^k) = \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{cov}_\theta(\xi_i^k, \xi_i^k)\ ,$$

поскольку при $i \ne j$ в силу независимости ξ_i и ξ_j,

$$\operatorname{cov}_\theta(\xi_i^k,\xi_j^k) = M_\theta[(\xi_i^k - m_k(\theta))(\xi_j^k - m_k(\theta))] = M_\theta[\xi_i^k - m_k(\theta)]M_\theta[\xi_j^k - m_k(\theta)] = 0\cdot 0 = 0 .$$

Таким образом,

$$D_\theta[\hat{m}_k(\zeta_n)] = \frac{1}{n^2}\sum_{i=1}^n D_\theta[\xi_i^k] = \frac{1}{n^2}\sum_{i=1}^n D_\theta[\xi_i^k] = \frac{1}{n^2}\sum_{i=1}^n (m_{2k}(\theta) - m_k^2(\theta)) = \frac{m_{2k}(\theta) - m_k^2(\theta)}{n} .$$

Можно показать, что при некоторых условиях статистики $\hat{m}_k(\zeta_n)$ имеют асимптотически нормальное распределение $N(m_k(\theta),(m_{2k}(\theta) - m_k^2(\theta))/n)$.

Исследование свойства несмещенности оценок $\hat{\mu}_k(\zeta_n)$ сопряжено с определенными трудностями, тем не менее, достаточно легко убедиться в состоятельности оценок $\hat{\mu}_k(\zeta_n)$. Заметим, что статистики $\hat{\mu}_k(\zeta_n)$ выражаются через статистики $\hat{m}_p(\zeta_n)$, где $1\le p\le k$:

$$\hat{\mu}_k(\zeta_n) = \frac{1}{n}\sum_{i=1}^n (\xi_i - \hat{m}_1(\zeta_n))^k = \frac{1}{n}\sum_{i=1}^n\sum_{p=0}^k C_k^p \xi_i^p(-\hat{m}_1(\zeta_n))^{k-p} = \sum_{p=0}^k C_k^p(-\hat{m}_1(\zeta_n))^{k-p}\frac{1}{n}\sum_{i=1}^n \xi_i^p =$$

$$= \sum_{p=0}^k C_k^p(-\hat{m}_1(\zeta_n))^{k-p}\hat{m}_p(\zeta_n) .$$

Таким образом, статистики $\hat{\mu}_k(\zeta_n)$ являются непрерывными функциями от состоятельных оценок $\hat{m}_p(\zeta_n)$, которые сходятся по вероятности к значениям $m_p(\theta)$, отсюда в силу свойства сходимости по вероятности статистика $\hat{\mu}_k(\zeta_n)$ сходится по вероятности к $\sum_{p=0}^k C_k^p(-m_1(\theta))^{k-p}m_p(\theta)$:

$$\hat{\mu}_k(\zeta_n)\xrightarrow{P}\sum_{p=0}^k C_k^p(-m_1(\theta))^{k-p}m_p(\theta),\ \text{при } n\to\infty .$$

Заметим, что справа от предела стоит в точности $\mu_k(\theta)$:

$$\mu_k(\theta) = \int_{-\infty}^{\infty}(x - m_1(\theta))^k dF_\xi(x\mid\theta) = \int_{-\infty}^{\infty}\sum_{p=0}^k C_k^p x^p(-m_1(\theta))^{k-p}dF_\xi(x\mid\theta) =$$

$$= \sum_{p=0}^k C_k^p(-m_1(\theta))^{k-p}\int_{-\infty}^{\infty}x^p dF_\xi(x\mid\theta) = \sum_{p=0}^k C_k^p(-m_1(\theta))^{k-p}m_p(\theta)$$

Таким образом, статистика $\hat{\mu}_k(\zeta_n)$ сходится по вероятности к $\mu_k(\theta)$:

$$\hat{\mu}_k(\zeta_n)\xrightarrow{P}\mu_k(\theta),\ \text{при } n\to\infty ,$$

отсюда статистика $\hat{\mu}_k(\zeta_n)$ является состоятельной (по определению).

7. Линейная оценка среднего с минимальной дисперсией при разноточных измерениях.

Задана совокупность случайных величин $(\xi_1,\xi_2,\ldots,\xi_n)$, которые имеют следующие свойства:

1) $M[\xi_i] = m$, $i = \overline{1,n}$;
2) $D[\xi_i] = \sigma_i^2$, $i = \overline{1,n}$;
3) $\operatorname{cov}(\xi_i,\xi_j) = 0$, $i\ne j$, $i,j = \overline{1,n}$;

где m неизвестное числовое значение, а значения σ_i^2 известны.

Случайным величинам ξ_i можно придать следующую интерпретацию: в результате независимого измерения некоторой неизвестной величины m различными способами получены величины ξ_i, поскольку способы измерения различны, то «точности» σ_i могут различаться.

Требуется построить оценку $\hat{m}$ неизвестной величины m, такую что:

1) Оценка $\hat{m}$ линейная:

$$\hat{m} = \sum_{i=1}^{n} \hat{\alpha}_i \xi_i \,,$$

2) Оценка $\hat{m}$ является несмещенной оценкой m :

$$M[\hat{m}] = m$$

3) Оценка $\hat{m}$ имеет наименьшую дисперсию в классе линейных оценок:

$$D[\hat{m}] = \min_{\alpha_1,\ldots,\alpha_n} D\left[\sum_{i=1}^{n} \alpha_i \xi_i\right].$$

Вычислим математическое ожидание оценки $\hat{m}$:

$$M[\hat{m}] = M\left[\sum_{i=1}^{n} \hat{\alpha}_i \xi_i\right] = \sum_{i=1}^{n} \hat{\alpha}_i M[\xi_i] = \sum_{i=1}^{n} \hat{\alpha}_i m = m\sum_{i=1}^{n} \hat{\alpha}_i \,.$$

Поскольку $\hat{m}$ должна быть несмещенной оценкой, то нужно потребовать чтобы $M[\hat{m}] = m$:

$$M[\hat{m}] = m\sum_{i=1}^{n} \hat{\alpha}_i = m \,.$$

$$\sum_{i=1}^{n} \hat{\alpha}_i = 1 \,.$$

Вычислим дисперсию оценки $\hat{m}$, принимая во внимание, что $\mathrm{cov}(\xi_i, \xi_j) = 0$ при $i \neq j$:

$$D[\hat{m}] = D\left[\sum_{i=1}^{n} \hat{\alpha}_i \xi_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{cov}(\hat{\alpha}_i \xi_i, \hat{\alpha}_j \xi_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} \hat{\alpha}_i \hat{\alpha}_j \,\mathrm{cov}(\xi_i, \xi_j) =$$

$$= \sum_{i=1}^{n} \hat{\alpha}_i^2 \,\mathrm{cov}(\xi_i, \xi_i) = \sum_{i=1}^{n} \hat{\alpha}_i^2 D[\xi_i] = \sum_{i=1}^{n} \hat{\alpha}_i^2 \sigma_i^2 \,.$$

Таким образом, приходим к задаче минимизации функции $\sum_{i=1}^{n} \alpha_i^2 \sigma_i^2$ по неизвестным α_i при условии, что $\sum_{i=1}^{n} \alpha_i = 1$. Для решения задачи нахождения условного экстремума воспользуемся методом множителей Лагранжа с функцией Лагранжа $\Lambda(\alpha_1,\ldots, \alpha_n, \lambda)$, которая имеет вид:

$$\Lambda(\alpha_1,\ldots,\alpha_n,\lambda) = \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 + \lambda\left(\sum_{i=1}^{n} \alpha_i - 1\right).$$

Вычисляя значения частных производных $\Lambda(\alpha_1,\ldots, \alpha_n, \lambda)$ в точке экстремума $(\hat{\alpha}_1,\ldots, \hat{\alpha}_n, \hat{\lambda})$, получим систему:

$$\begin{cases} \left.\dfrac{\partial \Lambda}{\partial \alpha_i}\right|_{(\alpha_1,\ldots,\alpha_n,\lambda)=(\hat{\alpha}_1,\ldots,\hat{\alpha}_n,\hat{\lambda})} = 0 \\ \left.\dfrac{\partial \Lambda}{\partial \lambda}\right|_{(\alpha_1,\ldots,\alpha_n,\lambda)=(\hat{\alpha}_1,\ldots,\hat{\alpha}_n,\hat{\lambda})} = 0 \end{cases} \Leftrightarrow \begin{cases} 2\hat{\alpha}_i \sigma_i^2 + \hat{\lambda} = 0 \\ \displaystyle\sum_{i=1}^{n} \hat{\alpha}_i - 1 = 0 \end{cases} \Leftrightarrow \begin{cases} \hat{\alpha}_i = -\dfrac{\hat{\lambda}}{2\sigma_i^2} \\ -\displaystyle\sum_{i=1}^{n} \dfrac{\hat{\lambda}}{2\sigma_i^2} - 1 = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \hat{\alpha}_i = \dfrac{\dfrac{1}{\sum_{i=1}^{n} \dfrac{1}{2\sigma_i^2}}}{2\sigma_i^2} = \dfrac{1}{\sigma_i^2} \dfrac{1}{\sum_{i=1}^{n} \dfrac{1}{\sigma_i^2}}. \\ \hat{\lambda} = -\dfrac{1}{\sum_{i=1}^{n} \dfrac{1}{2\sigma_i^2}} \end{cases}$$

Таким образом, искомая оценка $\hat{m}$ имеет вид:

$$\hat{m} = \sum_{i=1}^{n} \hat{\alpha}_i \xi_i = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}} \sum_{i=1}^{n} \frac{\xi_i}{\sigma_i^2},$$

при этом дисперсия оценки $D[\hat{m}]$:

$$D[\hat{m}] = \sum_{i=1}^{n} \hat{\alpha}_i^2 \sigma_i^2 = \sum_{i=1}^{n} \left(\frac{1}{\sigma_i^2} \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}} \right)^2 \sigma_i^2 = \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \right)^2} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}.$$

Обозначим $c = \dfrac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$, тогда $\hat{\alpha}_i = \dfrac{1}{\sigma_i^2} c$ и отсюда следует, что чем меньше σ_i^2, тем больше коэффициент $\hat{\alpha}_i$, то есть чем более «точным» является измерение ξ_i, тем с большим «весом» оно входит в сумму оценки $\hat{m} = \sum_{i=1}^{n} \hat{\alpha}_i \xi_i$. Например, если i-ое измерение «точнее» j-го в 3 раза, то есть $\dfrac{\sigma_j}{\sigma_i} = 3$, тогда $\dfrac{\hat{\alpha}_i}{\hat{\alpha}_j} = \dfrac{\frac{1}{\sigma_i^2} c}{\frac{1}{\sigma_j^2} c} = \dfrac{\sigma_j^2}{\sigma_i^2} = 9$, то есть «вес» измерения ξ_i в сумме оценки $\hat{m}$ в 3^2 раз больше «веса» измерения ξ_j.

Если все измерения ξ_i одинаковы по точности, то есть все величины σ_i равны одной и той же величине σ, тогда все коэффициенты $\hat{\alpha}_i$ также равны между собой:

$$\hat{\alpha}_i = \frac{1}{\sigma_i^2} \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}} = \frac{1}{\sigma^2} \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma^2}} = \frac{1}{\sigma^2} \frac{1}{n \frac{1}{\sigma^2}} = \frac{1}{n}.$$

В этом случае линейная оценка среднего с минимальной дисперсией $\hat{m}$ имеет вид:

$$\hat{m} = \sum_{i=1}^{n} \hat{\alpha}_i \xi_i = \frac{1}{n} \sum_{i=1}^{n} \xi_i,$$

и совпадает с выборочным средним $\hat{m}_1(\xi_1,..., \xi_n)$. Откуда следует, что выборочное среднее $\hat{m}_1(\xi_1,..., \xi_n)$ является оптимальной оценкой в классе несмещенных и линейных по наблюдениями ξ_i оценок.