УДК 528.1:519.281:521.1

Джунь Й. В., проф., д.ф.-м.н., (Міжнародний економіко-гуманітарний університет імені академіка Степана Дем’янчука)

ОБ ОДНОМ ОБОБЩЕНИИ ФУНДАМЕНТАЛЬНОГО ПРИНЦИПА МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ (М.Н.К.) В СВЯЗИ С ЭВОЛЮЦИЕЙ ПРЕДСТАВЛЕНИЙ О ЗАКОНЕ ОШИБОК НАБЛЮДЕНИЙ

Анотація. *Обоснована целесообразность замены фундаментального принципа М.Н.К. Гаусса принципом максимума информации по Фишеру. Показано, что такая замена обеспечивает более глубокое и естественное понимание сути М.Н.К., придает его процедурам более универсальный характер, позволяет проводить устойчивое оценивание.*
Ключевые слова: *фундаментальный принцип М.Н.К., информация по Фишеру.*

Annotation. *The expediency of substitution of Gauss' MLS fundamental principle for Fisher's information maximum principles is grounded. Such substitution guarantees better realization of MLS sense, makes its procedures more universal, allous permanent assessment.*
Key words: *MLS fundamental principle, information according to Fisher.*

Фундаментальным принципом классического М.Н.К. есть принцип наибольшего веса. Исходя из этого принципа Гаусс дал математическое обоснование М.Н.К. полагая, что ошибки наблюдений следуют закону e^{-h^2}. Однако кембриджский профессор Г. Джеффрис тщательно проанализировав результаты известного эксперимента К.Пирсона [10] пришел к выводу, что случайные независимые ошибки, если их число больше 500, при постоянных условиях наблюдений, следуют такому симметричному закону распределения вероятностей:

$$f(x)=\frac{\Gamma(m+1)}{\sqrt{2\pi(m-0{,}5)}\Gamma(m+0{,}5)}\cdot\frac{1}{\sigma}\left[1+\frac{0{,}5m^2}{(m-0{,}5)^3}\left(\frac{x-\lambda}{\sigma}\right)^2\right]^{-m}, \quad (1)$$

с показателем степени *m* в пределах

$$5\geq m\geq 3\ , \quad (2)$$

что соответствует таким значениям для эксцессов:

$$1{,}2\leq\varepsilon\leq 6\,. \quad (3)$$

В формуле (1): λ, σ, m - параметры закона плотности; $\Gamma(m)$ – гамма-функция.

Джеффрис делает вывод, что качество эксперимента, выполненного при стабильной метрологической ситуации, не вызывает сомнений, если эксцесс ошибок попадает в границы (3). Он называет (1) распределением Пирсона VII типа. Однако это не совсем отвечает истине – форму (1) он получил исходя из классической кривой Пирсона VII типа, которая имеет недиагональную информационную матрицу. Эту классическую кривую Джеффрис преобразовал к виду (1), который, как и закон Гаусса, имеет независимые параметры. Поэтому форму (1) правильнее было бы назвать распределением Пирсона-Джеффриса. Обладая необыкновенной научной скромностью, Джеффрис не дал особого названия распределению (1), которое он создал. Поэтому многие исследователи идентифицируют классическое распределение Пирсона VII типа и даже обобщенное распределение Коши, с формой (1), хотя этого делать нельзя, так как это разные распределения. Учитывая сказанное, распределение (1) можно назвать, чтобы избежать путаницы, законом ошибок Пирсона-Джеффриса или просто законом Пирсона-Джеффриса, подобно тому как в свое время был назван закон Гаусса.

С целью проверки идеи Джеффриса о соответствии эмпирических распределений ошибок кривой Пирсона VII типа, автором был проделан эксперемент [4]. В качестве рабочего поля для такой проверки использован график для идентификации типов распределений Пирсона в [1]. Каждое эмпирическое распределение ошибок на рис.1 характеризуется тремя координатами: эксцессом ε, квадратом асимметрии A^2 и объёмом выборки n. Нормальному распределению на рис.1 отвечает точка N – начало координат. На рис.1 почти все эмпирические распределения ошибок имеют практически нулевую асимметрию и $\varepsilon > 0$. Вертикальные прямые существенно сдвинуты вправо от точки N, соответствующей закону Гаусса, группируясь на линии кривой Пирсона VII типа вокруг точки $\varepsilon = 1$. Таким образом, закон Пирсона- Джеффриса (1) можно рассматривать в качестве новой, универсальной модели ошибок наблюдений большого объёма.

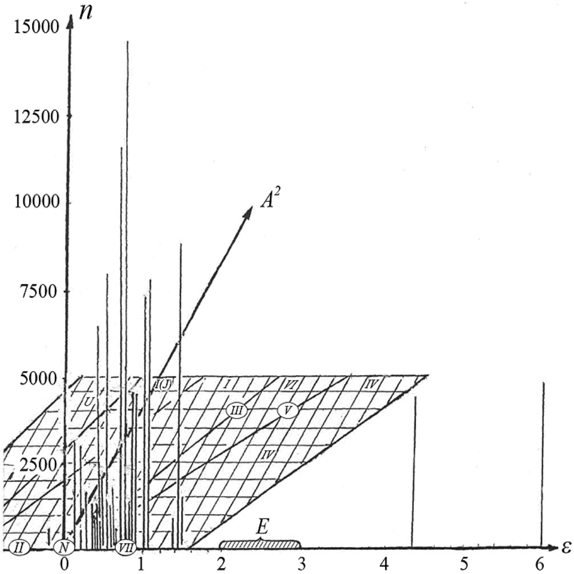

Рис. 1. Расположение эмпирических распределений ошибок астрономических, космических, гравиметрических, геодезических, экономических рядов (область E) на оси эксцесса (линия VII, правее точки *N*, соответствует семейству распределений Пирсона VII типа).

Значение закона Пирсона-Джеффриса (3) состоит еще и в том, что он имеет очень ценные математические свойства:

1. Является обобщением двух наиболее важных в теории ошибок распределений: Гаусса, в который форма (1) превращается при $m=\infty$, и распределения Стьюдента при значениях *m* в (1) кратных 0,5.

2. Параметр *m* закона Пирсона-Джеффриса (1) и число степеней свободы ν *t*-распределения связаны отношением:

$$\nu = 2m - 1. \tag{4}$$

3. Форма (1) является более гибкой при аппроксимации эмпирических распределений, чем *t*-распределение, так как она имеет промежуточные между

степенями свободы значения;

4. Параметр m - чувствительная мера уклонения закона Пирсона-Джеффриса от нормального закона. Гауссова форма распределения (1) постулирует значение m наперед известным и равным ∞. На самом же деле каждый ряд ошибок характеризуется своим значением m, которое очень далекое от ∞ и является его наиважнейшей характеристикой;

5. Информационная матрица закона Пирсона-Джеффриса диагональна, как и у закона Гаусса. Заметим, многочисленные распределения, которые предлагались для замены закона Гаусса, например, L_p - распределение, таким замечательным свойством не обладают;

6. Границы Рао-Крамера для закона Пирсона-Джеффриса определяются неравенствами:

$$\sigma_\lambda^2 \geq \frac{\sigma^2}{n} \cdot \frac{(m-0{,}5)^2(m+1)}{m^3}; \quad \sigma_\sigma^2 \geq \frac{\sigma^2}{2n} \cdot \frac{m+1}{m-0{,}5}; \tag{5}$$

$$\sigma_m^2 \geq \left\{ n\left[\psi'(m-0{,}5) - \psi'(m) - \frac{m+1}{2m^2(m-0{,}5)} \right] \right\}^{-1}, \tag{6}$$

где $\psi'(m)$ - тригамма функция. Из формул (5) видно, что при $m=\infty$ (закон Гаусса) σ_λ^2 и σ_σ^2 превращаются в классические формулы для определения дисперсий средней арифметической и среднего квадратического отклонения.

7. Главным достоинством закона Пирсона-Джеффриса является то, что ряди ошибок наблюдений, если их объёмы больше 500, почти всегда следует своей магической форме (1), рис.1. В астрономии, космических исследованиях, при уравнивании сплошных сетей одного класса точности, часто встречаются выборки объёма $n > 500$ [4].

Не следует думать,что Джеффрис пытался лишить нормальный закон его высокого статуса. Он только учел существующие реалии. Да и все варианты оценок, предложенные им, получают только после предварительного применения классических процедур оценивания по Гауссу.

К сожалению, до сих пор геодезистами не оценено в полной мере значение закона ошибок Пирсона-Джеффриса. Они не используют его в практике обработки геодезических наблюдений, несмотря на его универсальность для больших выборок и на то, что он имеет превосходные математические и аппроксимирующие свойства. Ни в одном современном курсе теории ошибок измерений распределение (1) не упоминается, хотя "Теория вероятностей" Джеффриса, где убедительно показано неадекватность закона Гаусса при многократных наблюдениях объёма $n > 500$, переиздавалась в Великобритании 9 раз.

Что происходит, если ошибки геодезических измерений не укладываются в указанные Джеффрисом границы (3) для эксцессов, а имеют, например, отрицательный эксцесс или значимую асимметрию? Попробуем ответить на этот вопрос опираясь на фишеровскую теорию оценивания. Для этого все разнообразие возможных эмпирических распределений ошибок представим четырехпараметрическими семействами. «За счет подбора четырех параметров можно очень точно подогнать любой закон распределения к любым эмпирическим даным» [6]. С этой целью воспользуемся общим дифференциальным представлением семейств Пирсона [1]:

$$\frac{f'(x)}{f(x)} = -\frac{x + c_1}{c_0 + c_1 x + c_2 x^2}\ , \tag{7}$$

где началом отсчета для *x* служит среднее значение, а величины c_0, c_1, c_2 связаны простыми соотношениями с асимметрией A и эксцессом ошибок [1]:

$$c_0 = \frac{\sigma^2(4\beta_2 - 3\beta_1)}{b}\ ; c_1 = \frac{\sigma\sqrt{\beta_1(\beta_2 + 3)}}{b}\ ;\ c_2 = \frac{2\beta_2 - 3\beta_1 - 6}{b}\ , \tag{8}$$

здесь $\beta_1 = \mathrm{A}^2 = \mu_3^2\mu_2^{-3}$; $\varepsilon = \beta_2 - 3$; $\beta_2 = \mu_4\mu_2^{-2}$; $b = 2(5\beta_2 - 6\beta_1 - 9)$;

$\mu_r = \int\limits_{l_1}^{l_2} x^r f(x)dx;\ r = 2,3,4\ ; \mu_0 = 1;\ \mu_0 = 0;$ l_1 и l_2 - области определения плотности *f(x)*.

При многократных наблюдениях обычно ищут оценку параметра расположения плотности *f(x)* на оси X. При негауссовых ошибках для вычисления такой оценки используют функцию правдоподобия *L*, полученную для *n* наблюдений y_i, следующих закону плотности *f(x)*:

$$L = \prod_{i=1}^{n} f(x_i), \tag{9}$$

где $f(x_i)$ - ордината плотности вероятности в точке:

$$x_i = y_i - a\ . \tag{10}$$

Логарифмируя функцию L и полагая, что она зависит только от a, находим частную производную:

$$-\frac{\partial lnL}{\partial a}=\sum_{i=1}^{n}\frac{f'(x_i)}{f(x)}=0. \tag{11}$$

Для получения оценки параметра a умножим числитель и знаменатель формулы (11) на $y_i - a$. Затем, решая уравнение относительно a, имеем:

$$a=\frac{\sum y_i \cdot p(x_i)}{\sum p(x_i)}, \tag{12}$$

В формуле (12) весовая функция $p(x_i)$ имеет вид:

$$p(x_i)=\frac{f'(x_i)}{x_i f(x_i)}. \tag{13}$$

Выражение (13) впервые получили астрономы Х.Р. Хюльме и Л.С.Т. Симс в 1939 г. Но они не предложили аналитического выражения для вычисления весов $p(x_i)$ и находили их с помощью линейки и транспортира по сглаженной кривой эмпирического распределения ошибок.

Чтобы найти аналитическое выражение для вычисления весов (13), подставим выражение (7) в формулу (13). В результате получим такое общее аналитическое представление весовой фукции четырехпараметрических семейств Пирсона:

$$p(x)=\frac{x+c_1}{x\left(c_0+c_1x+c_2x^2\right)}=\frac{1}{c_0+c_1x+c_2x^2}+\frac{c_1}{x\left(c_0+c_1x+c_2x^2\right)} \tag{14}$$

Выражение (14), определяет множество весовых функций, зависящих от значений c_0, c_1, c_2. Но в (14) есть очень опасный член c_1, величина которого, как это видно из формулы (8), определяется асимметрией распределения ошибок $\mathrm{A}=\sqrt{\beta_1}$. Член c_1 возникает, если $\mathrm{A}\neq 0$. При наличии асимметрии происходит то, что если $c_1 \neq 0$ и $x=0$, мы получаем по формуле (14) бесконечные веса. Сверхэффективные оценки недопустимы с точки зрения фишеровской теории оценок, да и в компьютерном моделировании. В этом случае оценка a в формуле (12) попадает в недопустимую область оценивания. Иными словами, асимметрия распределения погрешностей является сигналом того, что геодезический эксперимент находится в крайне опасной области, делающей невозможным эффективное оценивание. Обозначим эту область C_1. Заметим, что еще Д.И. Менделеев считал нулевую асимметрию погрешностей хорошим критерием согласованности ряда измерений и отсутствия в них систематических ошибок.

Перейдем теперь к рассмотрению симметричных семейств весовых функций $p(x)$. Полагая в формуле (14) $c_1 = 0$ получаем:

$$p(x) = \frac{1}{c_0 + c_2 x^2} = \frac{5\varepsilon + 6}{2\sigma^2\beta_2 + \varepsilon x^2}, \tag{15}$$

где эксцесс $\varepsilon = \beta_2 - 3$.

Фрагмент поверхности весовой функции (15) для симметричных семейств Пирсона представлен на рис. 2, взятом нами из работы [4].

При $\varepsilon = 0$ имеем $c_0 = \sigma^2, c_2 = 0$, т.е., для закона Гаусса все наблюдения имеют одинаковый вес и весовая функция приобретает вид константы:

$$p(x) = \sigma^{-2}. \tag{16}$$

На рис.2 прямая (16) делит поверхность весовой функции на две совершенно различные области: *B* – область плосковершинных симметричных распределений с $\varepsilon < 0$, в которой вес погрешности наблюдения возрастает с ее увеличением, и A – область закона ошибок Пирсона-Джеффриса (1), в которой вес погрешности x убывает с ее увеличением. Рассмотрим область *B*. Вопрос состоит в следующем: можно ли оценивать искомую величину по результатам наблюдений, которые имеют плосковершинное распределение погрешностей? Если да, то в этом случае мы должны согласиться с таким метрологическим абсурдом – чем больше ошибка наблюдения, тем больше её вес, что, собственно, демонстрируется на рис.2 поведением весовой функции в области *B*. Такая идея противоречит самой сути измерения, т.е. в случае $\varepsilon < 0$ мы попадаем в неприятную зону недопустимого оценивания *B*.

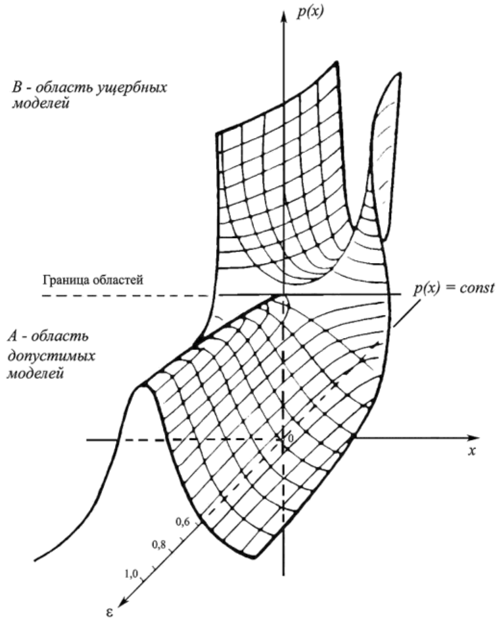

Рис.2. Фрагмент поверхности весовой функции.

Рассмотрим теперь область *A*, т. е. область джеффрисовых ошибок с эксцессами $\varepsilon \geq 0$. Как видно из рис.2, это благополучная область оценивания: в ней $p(x)$ либо постоянна (это область классического оценивания), либо неограниченно убывает при $x \to \infty$ (область устойчивого оценивания).

Таким образом, единственной областью весовой функции (13), где возможно эффективное оценивание, является область *A*, соответствующая закону ошибок Пирсона-Джеффриса (1). Но Джеффрис ставит еще дополнительное условие: независимые погрешности наблюдений должны иметь положительный эксцесс в пределах границ (3), т.е. иметь эксцесс ε не меньше чем 1,2. Если же действительные ошибки наблюдений имеют эксцесс в пределах:

$$0 \leq \varepsilon \leq 1{,}2, \tag{17}$$

то они, хотя и находятся в допустимой области оценивания, но все же искажены неисключенными систематическими погрешностями. Иными словами, – если эксцесс находится в пределах (17), то это может быть свидетельством наличия в результатах наблюдений, хотя и слабых, но не исключенных систематических ошибок. Влияние последних в границах (17) будет тем больше, чем ближе положительный эксцесс находится к нулю. Различные случаи сверток распределений, подтверждающие такой вывод, подробно рассмотрены в работе [2]. Таким образом, если наблюдений больше 500, то нормальность ошибок наблюдений есть поводом для серьёзного беспокойства по поводу неисключенных постоянных погрешностей.

Чтобы проверить тот факт, что полученные результаты многократных наблюдений попадают в зону допустимого оценивания *A*, необходимо построить доверительные интервалы для оценок асимметрии и эксцесса, которые получены на основании несмещенных оценок моментов [5,с.386]:

$$\mathrm{A} = \frac{\sqrt{n(n-1)}}{n-2}\frac{m_3}{m_2^{-1.5}}; \tag{18}$$

$$\varepsilon = \frac{(n-1)(n^2-2n+3)}{n(n-2)(n-3)}\frac{m_4}{m_2^2} - \frac{3(n-1)(2n-3)}{n(n-2)(n-3)} - 3, \tag{19}$$

где m_r - выборочные центральные моменты порядка r, вычисленные для x_i:

$$m_r = n^{-1}\sum(x_i-\bar{x})^r; \quad \bar{x} = n^{-1}\sum x_i \tag{20}$$

Для построения доверитетельных интервалов для A и ε воспользуемся известными стандартными погрешностями этих статистик [5, с. 391]:

$$\sigma_{\mathrm{A}} = \sqrt{\frac{4\mu_2^2\mu_6 - 12\mu_2\mu_3\mu_5 - 24\mu_2^3\mu_4 + 9\mu_3^2\mu_4 + 35\mu_2^2\mu_3^2 + 36\mu_2^5}{4\mu_2^5 n}}; \tag{21}$$

$$\sigma_{\varepsilon} = \sqrt{\frac{\mu_2^2\mu_8 - 4\mu_2\mu_4\mu_6 - 8\mu_2^3\mu_3\mu_5 + 4\mu_4^3 - \mu_2^2\mu_4^2 + 16\mu_2\mu_3^2\mu_4 + 16\mu_2^3\mu_3^2}{\mu_2^6 n}}, \tag{22}$$

где μ_r – несмещенные оценки центральных моментов порядка r.

При значения $n > 500$ моменты μ_r в (20) и (21) можно заменить на выборочные моменты m_r, вычисляемые по формуле (20), так как в этом случае смещение на оценках σ_{A} и σ_{ε} особо не сказывается. Например, при нормальном законе, как это следует из (19), при $n = 500$ выборочный куртозис равен:

$$\frac{m_4}{m_2^2} = \frac{3n(n-2)(n-3)(n-1)^{-1} + 6n - 9}{n^2 - 2n + 3} = 3{,}000024,$$

т. е. практически не отличается от несмещенного.

Для закона Гаусса несмещенная оценка асимметрии вычисляется по формуле (20), эксцесс же получают из такого выражения [5, с. 423]:

$$\varepsilon = \frac{n-1}{(n-2)(n-3)}\left[(n+1)\left(\frac{m^4}{m_2^2} - 3\right) + 6\right], \tag{23}$$

а значения дисперсий оценок (18) и (23) из формул:

$$\sigma_A = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}};$$

$$\sigma_\varepsilon = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}} \tag{24}$$

Для симметричных распределений с эксцессом +1, а такими являются наиболее часто распределения ошибок (рис.1), формулы (21), (22) и (24) приводят практически к идентичным результатам. Поэтому используя оценки (18), (19) и (24) находим 90% доверительные интервалы для A и ε:

$$A \pm 1{,}645\sigma_A;\ \varepsilon \pm 1{,}645\sigma_\varepsilon, \tag{25}$$

где квантиль $t_{10\%} = 1{,}645$.

Наиболее желательной есть такая форма распределения погрешностей, когда доверительный интервал для A обязательно накрывает ноль, а доверительный интервал для ε хотя бы касается левого конца интервала (3). В этом случае можно считать, что геодезический эксперимент поставлен хорошо. Систематические погрешности подавлены и не оказывают ощутимого влияния на результаты оценивания.

Допустимой следует считать такую форму распределения, погрешностей, когда доверительный интервал для коэффициента асимметрии A накрывает нуль, а доверительный интервал для эксцесса ε находится внутри границ (17) или, хотя бы, одним своим концом касается нуля.

Недопустимой при оценивании искомых параметров следует считать такую форму распределения погрешностей, которая характеризуется хотя бы одной из следующих ситуаций:

- доверительный интервал для асимметрии ошибок не накрывает ноль;
- весь доверительный интервал для ε находится в области отрицательных эксцессов.

Такие две ситуации можно определить как области сингулярности весовой функции. Они соответствуют крайне нежелательным, образно сказать, патологическим случаям оценивания. Это наблюдается, когда в результате уравнивания геодезической сети остаточные погрешности попадают в недопустимую область сверхэффективных оценок C_1, либо в метрологически абсурдную область оценивания *B*.

Действительные распределения ошибок наблюдений, в зависимости от метрологической ситуации и числа наблюдений могут иметь самую разнообразную форму. Если условия наблюдений постоянны, а система инструмента хорошо отлажена, то геодезисты часто полагают, что результатом этого будет нормальность закона распределения ошибок. Но вследствие «парадокса Эльясберга - Хампеля» любая гипотеза о виде распределения вероятностей будет отвергнута при достаточно большом числе членов статистического ряда [7,8]. Если исходить из этого парадокса, то следует признать, что теория тех или иных методов обработки данных, включая М.Н.К., является корректной только в пределах определенных объёмов информации. Впервые всю серьёзность такой ситуации оценил Г. Джеффрис, показав теоретическую и практическую несостоятельность закона ошибок Гаусса, если число наблюдений $n > 500$. Таким образом, по Джеффрису, классический М.Н.К. и процедуры отбраковки являются состоятельным, если число многократных наблюдений $n \leq 500$. Он пишет в [9]: «Решающим вопросом в комбинации наблюдений является знание того, действительно ли распределения следует нормальному закону, если это не так, то должны быть придуманы другие методы, присущи даному закону».

Предложив более адекватный закон ошибок (1), Джеффрис на этом остановился и не предпринимал попыток эволюции основ М.Н.К. На наш взгляд, наиболее полезным для такой эволюции является понятие информации Фишера, удовлетворяющее таким требованиям:

1. Информация, имеющаяся в наблюдениях, должна увеличиваться пропорционально их числу.

2. Информация должна быть связана исключительно с задачами данного геодезического эксперимента.

3. Информация должна быть связана с точностью: чем выше точность наблюдений – тем больше информации мы получаем.

Всеми этими, ценными для анализа геодезических данных свойствами, обладает информация по Фишеру. Чтобы получить для нее математическое выражение, обозначим через $y(x_i;\theta_1,\theta_2,\ldots)$ – плотность вероятности, зависящую от *k* параметров θ; x_i обозначает вектор, содержащий n ошибок наблюдений. Функция максимального правдоподобия для значений x_i будет иметь вид:

$$L = \prod_{i=1}^{n} y(x_i;\theta_1,\theta_{2,}\dots) \tag{26}$$

По определению Фишера, объём информации об искомом параметре θ, которая содержится в одном наблюдении x, выражается формулой:

$$I_1(\theta) = \int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial\theta}\right)^2 y(x;\theta_1,\theta_{2,}\dots)dx\,. \tag{27}$$

Если учесть, что:

$$E\left(\frac{\partial \ln L}{\partial\theta}\right)^2 = -E\left(\frac{\partial^2 \ln L}{\partial\theta^2}\right), \tag{28}$$

где E – оператор математического ожидания, то, при наличии k параметров распределения, выражение (27) превращается в информационную матрицу размера $k\times k$, элемент которой имеет вид:

$$I(\theta_i\theta_j) = -\int_{-\infty}^{\infty} \left(\frac{\partial^2 \ln L}{\partial\theta_i\partial\theta_j}\right) y(x;\theta_1,\theta_{2,}\dots)dx\,, \tag{29}$$

Если переменная x распределена нормально с известной дисперсией σ^2 и неизвестным математическим ожиданием a, т.е.

$$y(x;a,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}\,, \tag{30}$$

то, с учетом формулы (27), получаем для закона Гаусса информацию Фишера, которая содержится в одном наблюдении:

$$I_1(a) = -\int_{-\infty}^{\infty} \frac{\partial^2}{\partial a^2}\left(-\ln\sigma\sqrt{2\pi} - \frac{(x-a)^2}{2\sigma^2}\right) y(x;\sigma,a)dx = \frac{1}{\sigma^2}\,. \tag{31}$$

Таким образом, для случая гауссовых ошибок, вес наблюдения – это не абстрактное значение обратной дисперсии, смысл которой не совсем понятен геодезисту, а мера количества информации по Фишеру.

В случае n наблюдений имеем для распределения Гаусса (30):

$$L = \prod_{i=1}^{n} y(x_i;\sigma,a)\,.$$

В соответствии с (28), получим:

$$I_n(a)=\int\limits_{-\infty}^{\infty}\frac{\partial^2}{\partial a^2}\left[\sum_{i=1}^{n}\left(-\ln\sigma\sqrt{2\pi}-\frac{(x-a)^2}{2\sigma^2}\right)\right]f(x;\sigma,a)dx=n\cdot I_1(a)=\frac{n}{\sigma^2}. \qquad (32)$$

Как видим, фундаментальный принцип М.Н.К. означает ни что иное, как максимизацию информации по Фишеру.

Для закона ошибок Пирсона-Джеффриса (1) можно получить, такое значение фишеровской информации для случая *n* наблюдений, предполагая, что σ и *m* известны:

$$I_n(\lambda)=-E\left(\frac{\partial^2 \ln L}{\partial\lambda^2}\right)=\frac{n}{\sigma^2}\cdot\frac{m^3}{(m-0{,}5)^2(m+1)}, \qquad (33)$$

где значение параметра *m* обратно зависимо от величины эксцесса. При $m=\infty$, когда закон ошибок Пирсона-Джеффриса становится нормальным, соотношение (33) идентично выражению (32), но, в отличие от последнего, формула (33) демонстрирует один парадокс: информация Фишера растет не только с увеличением *n* и уменьшением σ^2 она возрастает и с увеличением уклонения распределения (1) от закона Гаусса, увеличиваясь с уменьшением *m*.

Учитывая сказанное, принцип максимума информации по Фишеру, можно принять в качестве нового, универсального принципа М.Н.К., который придает ему новую динамику развития. Для закона ошибок Пирсона-Джеффриса, который более адекватен практике наблюдений, этот принцип можно выразить более общей формулой (33), которая учитывает меру уклонения действительного распределения от закона Гаусса.

Подводя итоги нашего исследования можно сделать такие выводы:

1. Оценки эксцесса $\mathcal{E}$ и асимметрии А, вычисленные для ошибок наблюдений или остаточных погрешностей, являются наиболее важными характеристиками, с помощью которых оценивается влияние ненормальности, так как гарантии нормальности нельзя дать практически никогда. Они играют решающую роль при обработке наблюдений и диагностике математического моделирования, а также при уравнительных процедурах в геодезии, если число многократных наблюдений $n > 500$. Наиболее желательными характеристиками независимых погрешностей являются границы (3) для эксцессов при условии, что доверительный интервал для коэффициента асимметрии накрывает ноль.

2. Допустимой можно считать такую форму распределения погрешностей, когда доверительный интервал для эксцесса находится внутри интервала (17) или накрывает ноль при условии, что доверительный интервал для асимметрии также накрывает ноль. Практически это означает, что действительные распределения погрешностей должны следовать закону ошибок Пирсона-Джеффриса (1) с эксцессами в границах (17). Во всех остальных случаях геодезический эксперимент, при числе наблюдений $n > 500$, следует считать некорректным, так как оценивание производится в недопустимых областях.

3. Предложеный метод может быть использован для диагностики ошибок ответственных геодезических экспериментов высокого научного и технического уровня. При выполнении массовых производственных измерений рассмотренная рафинированная диагностика формы распределения остаточных ошибок наблюдений, возможно, в большинстве случаев, не обязательна.

4. Ограниченность фундаментального постулата М.Н.К. – принципа наибольшего веса в том, что он не учитывает наиболее существенную особенность действительных распределений ошибок большого объёма – их положительный эксцесс. Если эксцесс $\varepsilon > 0$, то на практике это означает избыток больших ошибок, которые существенно увеличивают дисперсию и уменьшают информацию по Фишеру. Гауссова форма распределения (1) постулирует параметр m наперед известным и равным ∞. Такой постулат некорректен, так как он неадекватен практике многократных наблюдений. Этим блокируется дальнейшая эволюция процедур М.Н.К. На самом же деле, каждый ряд ошибок имеет свое уклонение от закона Гаусса, т.е. свое значение параметра m, далекое от ∞. Это значение и является определяющей характеристикой любого ряда ошибок большого объёма и именно его нужно учитывать при обработке данных.

5. Замена абстрактного принципа максимума обратной дисперсии на принцип максимума фишеровской информации обеспечивает более глубокое и естественное понимание действительной сути М.Н.К. – как научного метода максимизирующего получение информации при математической обработке данных. Такая замена придает установкам М.Н.К. более общий, понятный и целесообразный смысл.

6. Главный вывод в том, что переход на базовый принцип максимума информации по Фишеру, можно рассматривать как очередной этап эволюции М.Н.К. Мы показали, что он открывает новые возможности для усовершенствования методов обработки геодезических наблюдений с целью получения более эффективных оценок на основе дополнительного

взвешивання минимизированных классическим М.Н.К. остаточных погрешностей υ_i.

7. Нынешний этап развития геодезии, отличающийся все большими объёмами исследований и числом наблюдений, требует использования качественно более высоких технологий М.Н.К. Их может обеспечить обобщенный подход к фундаментальному принципу М.Н.К., в даном случае - принцип максимума информации по Фишеру.

8. С увеличением числа наблюдений остаточные погрешности υ_i приобретают все более отчетливо выраженные особенности закона Пирсона-Джеффриса (1). Иными словами, с возрастанием объёма выборок, эмпирические распределения ошибок все более и более уклоняются от закона Гаусса. Погрешности будут, как правило, следовать распределению Пирсона-Джеффриса и иметь на порядок различающиеся веса. Поэтому, значение эволюционных процедур М.Н.К. будет возрастать вместе с ростом объёмов многократных наблюдений в геодезических экспериментах.

Литература:

1. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: ВЦ АН СССР, 1968. – 476 с.
2. Бородачев Н. А. Основные вопросы теории точности производства. Под ред. А. Н. Колмогорова. – М. – Л.: Изд. АН СССР, 1950. – 360 с.
3. Гаусс К.Ф. Избранные геодезические сочинения. Том 1. Способ наименьших квадратов. Под ред. Г.В. Багратуни. Пер с лат. и нем. – М.: Изд. геод. лит., 1957. – 234 с.
4. Джунь И. В. Математическая обработка астрономической и космической информации при негауссовых ошибках наблюдений. Автореферат дис. докт. физ.-мат. наук.- К.: ГАО АНУ. – 1992 – 46 с.
5. Крамер Г. Математические методы статистики. М.: Мир, 1975 – 648 с.
6. Тутубалин В. Н. Теория вероятностей. Краткий курс и научно-методические замечания. М.: Изд-во Моск. ун-та, 1972 -230 с.
7. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. Пер. с англ. М.: Мир, 1989 – 512 с.
8. Эльясберг П. Е. Измерительная информация: сколько её нужно? Как обрабатывать. М.: Наука. 1983. – 208 с.
9. Jeffreys H. The Law of Error in the Greenwich Variation of Latitude Observations. // Mon. Notic. of the RAS, 1939, vol. 99, № 9, p. 703-709.
10. Pearson K. On the mathematical theory of errors of judgment with special reference to the personal equation. // Phil. Trans. Roy. Sos. London, A. – 1902. – 198. – P. 235 – 29