О ФУНКЦИИ РАСПРЕДЕЛЕНИЯ ОШИБОК ШИРОТНЫХ НАБЛЮДЕНИЙ

И. В. Джунь

(Ровенский институт инженеров водного хозяйства, г. Ровно)

Ошибки систематических астрономических определений представляют собой некоторую временную последовательность случайных величин. В зависимости от различных причин — погоды, состояния инструмента или наблюдателя — ее математическое ожидание и дисперсия значимо изменяются [14]. При статистическом изучении такой последовательности следовало бы прежде всего найти ее закон распределения в любой фиксированный момент времени. Если предположить, что ошибки в момент t распределены нормально, то плотность распределения $f(x, t)$ имеет такой вид:

$$f(x, t) = \frac{1}{\sigma(t)\sqrt{2\pi}} e^{-\frac{[x-\mu(t)]^2}{2\sigma^2(t)}}, \tag{1}$$

где $\sigma^2(t)$ и $\mu(t)$ — соответственно изменяющиеся со временем дисперсия и математическое ожидание.

Для совокупности наблюдений, выполненных за время $\Delta\tau$, средняя плотность распределения такова:

$$f(x, T) = \frac{1}{\sqrt{2\pi\,\Delta\tau}} \int\limits_{T-\frac{\Delta\tau}{2}}^{T+\frac{\Delta\tau}{2}} \frac{1}{\sigma(t)} e^{-\frac{[x-\mu(t)]^2}{2\sigma^2(t)}} dt, \tag{2}$$

Полагаем t непрерывным в интервале $\Delta\tau$.

Выясним, действительно ли можно считать нормальным распределение ошибок широтных наблюдений, выполненных в интервале времени $\Delta\tau$, в течение которого условия наблюдений остаются практически постоянными.

Так как астрономические наблюдения выполняются на открытом воздухе и вследствие этого подвержены значительным влияниям внешней среды, приходится брать малые интервалы $\Delta\tau$. За это время не удается сделать достаточное количество наблюдений. Проверка гипотезы нормальности для малочисленных групп широт никогда не делалась, хотя существует простой способ такой проверки [11]. Суть этого способа в следующем: есть большое число малых независимых выборок одинакового объема; наша задача проверить гипотезу нормальности генеральных совокупностей, из которых взяты эти выборки, не предполагая, что параметры этих совокупностей тождественны. Иначе говоря, мы считаем устойчивой лишь форму распределения, а именно: предполагаем ее нормальной, не делая никаких допущений относительно центров и дисперсий генеральных совокупностей.

Пусть φ_{ki} — наблюдение, наудачу взятое из k-й выборки, $\bar{\varphi}_k$ и $\bar{S}_k$ — соответственно ее эмпирический центр и среднеквадратическое отклонение; $k = 1, 2, \ldots, m$, где m — число выборок одинакового объема. Величина

$$\varkappa_k = \frac{\varphi_{ki} - \bar{\varphi}_k}{\bar{S}_k}$$

имеет плотность распределения [4, 5]:

$$P(\varkappa)=\frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-1)\pi}\,\Gamma\left(\frac{n-2}{2}\right)}\left(1-\frac{\varkappa^2}{n-1}\right)^{\frac{n-4}{2}}\quad \text{при } |\varkappa|<\sqrt{n-1};$$

$$P(\varkappa)=0 \text{ при } |\varkappa|\geqslant\sqrt{n-1}.$$

Здесь $\Gamma\left(\frac{n-1}{2}\right)$, $\Gamma\left(\frac{n-2}{2}\right)$; Γ — функции, n — объем выборки. При $n=4$ плотность распределения $\varkappa$ равномерна, если исходные совокупности нормальны, хотя и имеют разные центры и дисперсии. Если $n\neq4$, проверка гипотезы осуществляется использованием величины:

$$\eta=\frac{\varkappa\sqrt{n-2}}{\sqrt{n-1-\varkappa^2}}.$$

При исходных гауссовских распределениях η подчиняется распределению Стьюдента с $n-2$ степенями свободы.

После получения $\varkappa$ или η задача проверки гипотезы нормальности приобретает непараметрический характер, и для ее решения естественно воспользоваться законом распределения А. Н. Колмогорова. Сама процедура проверки описана в [11]. Заметим, однако, что в этой работе критерий А. Н. Колмогорова применен неточно [6, 7, 12]. Поэтому вместо таблиц, предназначенных для этой цели в [11], мы воспользовались таблицей критических значений для наибольшего отклонения эмпирического распределения от теоретического [12].

При уровне значимости q критическая область для проверки нашей гипотезы определялась неравенством:

$$\lambda_{max}=max\left|F(\eta_k)-\frac{(2k-1)}{2m}\right|>\lambda_q.$$

Здесь $F(\eta_k)$ — теоретическая функция распределения, подсчитанная для значений $\varkappa_k$ или η_k. При этом при $n=4$ для этого используется закон равномерной плотности, при $n\neq4$ — закон распределения Стьюдента; λ_q выбирается из таблиц [12] по двум параметрам: уровню значимости q и числу выборок m.

Материал нашего анализа — наблюдения в Мидзусаве на зенит-телескопах Бамберга и Куксона в 1940—1949 гг. [8, 9] и наблюдения по общей программе на зенит-телескопах Цейсса и Бамберга в Полтаве в 1949—1954,5 гг. [10]. Программы наблюдений на зенит-телескопах Куксона и Бамберга имеют 32 общие пары [9]. По этим наблюдениям были образованы разности широт:

$$y_{ijk}=\Delta\varphi_{ijk}=\varphi'_{ijk}-\varphi''_{ijk}, \tag{3}$$

где φ'_{ijk} — широта, полученная по i-й паре за j-й год в k-е сутки на зенит-телескопе Куксона, φ''_{ijk} — то же для зенит-телескопа Бамберга.

Разности широт для полтавских наблюдений вычислялись аналогично: широта, полученная на зенит-телескопе Цейсса, минус широта, полученная в те же сутки и по той же паре на зенит-телескопе Бамберга:

$$i=1,2,\ldots,l,$$
$$j=1,2,\ldots,s,$$
$$k=1,2,\ldots,r_j.$$

Здесь l — число разностей широт в выборке, s — продолжительность в годах ряда совместных наблюдений на двух инструментах, r_j — число ночей наблюдений в j-й годичной серии. Тогда число отдельных разностей в этой серии — lr_s, а общее число значений y_{ijk}

$$l = \sum_{j=1}^{s} r_j.$$

Гипотеза нормальности полтавских наблюдений проверялась по выборкам, полученным для звеньев и групп четырехгрупповой программы. Объем выборок зависел от числа пар в звене и группе. Для наблюдений в Мидзусаве серии выборок получены по парам, наблюдавшимся в пределах около 2 *ч*. Мы полагаем, что в течение этого времени условия наблюдений можно считать постоянными. Объемы выборок мы брали в зависимости от числа пар, наблюдавшихся близко одна за другой. При $n=4$ делали две проверки гипотезы, по первому и одному из последних экземпляров каждой выборки, если же $n=8$ — по трем. Каждая проверка осуществлялась по разностям y_{ijk}, принадлежавшим какой-либо одной паре в данной серии выборок, т. е. $\varkappa$ или η вычислялись по разностям y_{ijk} при фиксированном i. В табл. 1 приводится пример проверки гипотезы нормальности по наблюдениям пар 55, 57, 58 и 59 международной программы, принятой в 1935 г. [9, стр. 11]. В этой таблице x_k обозначает разность между значением y_{ijk}, полученным по паре 55, и средним из значений y_{ijk}, полученным по всем четырем указанным парам.

Максимальное значение λ_{max} в этой серии равно 0.167. В то же время по табл. 6,2 в [12] находим для $m=43$ при 10%-ном уровне значимости:

$$\lambda_{0.10} = 0.183.$$

Таким образом, можно считать, что выборки данной серии принадлежат нормальным генеральным совокупностям.

В табл. 2 дана сводка результатов проверки по всем сериям. Для 41 проверки табл. 2, в случае подтверждения гипотезы нормальности с 10%-ным уровнем значимости, можно ожидать всего четыре отрицательных исхода. Мы же имеем их 19, причем 16 превышают 5%-ный, а 11 — 1%-ный уровни значимости. Такой исход для имеющегося у нас объема материала (около 6 тыс. разностей y_{ijk}) нельзя считать случайным.

В соответствии с «гипотезой элементарных ошибок», введенной Хагеном и Бесселем, суммарная ошибка астрономического измерения рассматривается как сумма большого числа взаимно независимых элементарных ошибок. Тогда, согласно центральной предельной теореме, суммарная ошибка должна быть распределена приблизительно нормально, каким бы законам распределения ни были подчинены элементарные ошибки. И это приближение тем точнее, чем большее количество случайных величин суммируется. Однако это бывает так лишь в том случае, если среди отдельных ошибок нет такой, которая заметно преобладает над всеми другими. В противном случае закон распределения этой преобладающей ошибки наложит свое влияние на сумму. По-видимому, это происходит и в нашем случае. Тогда следующим этапом должен быть поиск таких превалирующих ошибок и их исключение из совокупности данных наблюдений. Если это удалось сделать, нужно затем повторить проверку закона распределения. Допустим, что такая проверка дала положительный результат. Тогда можно считать, что ошибки наблюдений обусловлены большим числом слабых влияний, и дальнейший поиск превалирующих влияний не имеет смысла. Если

Таблица 1

Номер выборки k	y_{ijk} в 0″.01				$\varkappa_k$	λ_{max}
	55	57	58	59		
1	1	—34	1	3	0.462	0.029
2	13	—31	—57	—14	1.198	.026
3	—12	10	—9	—50	0.129	.056
4	—103	—2	—51	—34	—1.314	.039
5	20	8	—5	—14	1.192	.005
6	—21	29	—25	4	—0.708	.111
7	—17	28	15	—1	—1.191	.033
8	37	38	—27	15	—0.970	.024
9	—73	39	—19	—44	—1.024	.007
10	—42	8	—36	—7	—0.958	.067
11	35	—18	29	—43	0.911	.039
12	1	—21	—20	—35	1.330	.058
13	3	72	—6	7	—0.447	.138
14	—50	17	—26	—31	—0.972	.002
15	—6	—2	—35	—20	0.651	.021
16	—40	25	10	—39	—0.866	.064
17	—30	—4	—21	7	—1.082	.013
18	—4	—46	—34	—5	0.866	.006
19	1	32	—20	—46	0.280	.035
20	—27	—10	—34	—32	—0.115	.056
21	99	13	—50	—31	1.376	.068
22	—74	28	—19	—10	—1.313	.063
23	22	11	5	—31	0.883	.024
24	—34	16	—35	—20	—0.660	.167
25	—37	48	60	—5	—1.176	.052
26	2	—1	—8	—29	0.787	.005
27	—11	33	11	—61	—0.099	.075
28	—28	33	—38	17	—0.698	.132
29	—25	—8	—17	24	—0.861	.086
30	11	—23	—11	—7	1.314	.039
31	—7	—33	—26	-49	1.249	.034
32	—26	—13	—25	0	—0.820	.097
33	—37	—1	22	—20	—1.104	.030
34	—41	—1	—1	—14	—1.149	.079
35	23	35	—13	10	0.450	.010
36	—1	48	—18	—32	—0.007	.072
37	—35	81	73	—23	—0.960	.046
38	—11	—18	—36	-20	0.970	.045
39	—7	2	20	—4	—0.806	.116
40	—11	3	—3	—13	—0.676	.149
41	14	—5	13	15	0.498	.042
42	—15	24	23	4	—1.300	.043
43	—18	6	—30	—29	1.493	.057

Таблица 2

Выборка	m	$\lambda_{0.10}$	$\lambda_{0.05}$	λ_{max}
Наблюдения в Мидзусаве ($n=4$)				
10	61	0.154	0.171	0.129
11	61	0.154	0.171	0.356
1	40	0.189	0.210	0.133
2	40	0.189	0.210	0.244
45	43	0.183	0.203	0.125
47	43	0.183	0.203	0.260
55	43	0.183	0.203	0.167
57	43	0.183	0.203	0.303
37	44	0.181	0.201	0.291
43	44	0.181	0.201	0.305
43	77	0.137	0.152	0.260
43	77	0.137	0.152	0.173
Наблюдения в Полтаве ($n=4$)				
Звено	41	0.187	0.208	0.232
1950 г.	41	0.187	0.208	0.145
Звено	45	0.179	0.198	0.220
1951 г.	45	0.179	0.198	0.100
Звено	48	0.173	0.192	0.133
1953—1954 гг.	48	0.173	0.192	0.081
Звено	48	0.173	0.192	0.181
1953—1954 гг.	48	0.173	0.192	0.077
Звено	42	0.185	0.205	0.187
1950 г.	42	0.185	0.205	0.281
Звено	42	0.160	0.178	0.122
1950 г.	42	0.160	0.178	0.147
Наблюдения в Полтаве ($n=4$)				
Группа	48	0.173	0.192	0.117
1949—1951 гг.		0.173	0.192	0.138
		0.173	0.192	0.058
Группа	69	0.145	0.161	0.075
1951—1954 гг.		0.145	0.161	0.106
Группа	62	0.153	0.170	0.128
1950—1951 гг.		0.153	0.170	0.252
		0.153	0.170	0.079
Группа	74	0.140	0.155	0.236
1951—1954 гг.		0.140	0.155	0.110
		0.140	0.155	0.060
Группа	67	0.147	0.163	0.062
1950—1951 гг.		0.147	0.163	0.162
		0.147	0.163	0.070
Группа	74	0.140	0.155	0.289
1951—1954 гг.				

же ошибки все еще не удовлетворяют нормальному распределению, этот поиск нужно продолжить.

Вернемся к полученным нами разностям широт, наблюденных в Мидзусаве. Анализируя эти разности, Гатторн нашел их систематические смещения, обусловленные неточностью принятого значения цены оборота и температурным коэффициентом [8]. Он получил следующую формулу для исключения этих влияний:

$$\Delta y_{ijk} = \text{const} + 0.0105\, Z - 0.00011\, Z \cdot T,$$

где второй член обусловлен неточностью принятых значений цены оборота, третий учитывает влияние температуры инструмента во время наблюдений; Z — разность зенитных расстояний талькоттовской пары.

Исключим теперь из данного ряда наблюдений влияния, обнаруженные Гаттори, и посмотрим, подчиняются ли закону Гаусса остаточные уклонения.

Заметим, что поправки за температурный коэффициент и за неточность цены оборота мы получили несколько иначе, чем Гаттори. Чтобы исключить влияние на результат сезонных вариаций, мы воспользовались следующими величинами, полученными на основании (3):

$$\Delta y_{ijk} = y_{ijk} - y_{i+1,\, j,\, k}. \tag{4}$$

Эти разности, согласно проведенному нами дисперсионному анализу, не изменяются в зависимости от сезона, в то время как значения y_{ijk} изменяются и не представляют достаточно однородного материала для получения температурного коэффициента и поправки к принятым значениям цены оборота.

В результате мы получили:

$$\Delta y_{ijk} = -0.002941 + 0.00865\, Z + 0.000593\, Z \cdot T. \tag{5}$$

Таблица 3

i	m	$\lambda_{0.10}$	$\lambda_{0.05}$	λ_{max}
10	62	0.154	0.171	0.111
11	62	0.154	0.171	0.275
1	40	0.189	0.210	0.093
2	40	0.189	0.210	0.115
45	43	0.183	0.203	0.092
	43	0.183	0.203	0.124
55	43	0.183	0.203	0.086
57	43	0.183	0 203	0.099
37	44	0.181	0.201	0.156
43	44	0.181	0.201	0.182
43	77	0.137	0.152	0.158
45	77	0.137	0.152	0.107

Исключив по формуле (5) из значений y_{ijk} поправки Δy_{ijk}, мы по тем же сериям, которые приведены в табл. 2 для наблюдений в Мидзусаве, вновь проверили гипотезу нормальности. Результаты проверки приведены в табл. 3. В случае подтверждения гипотезы нормальности с 10%-ным уровнем значимости из 12 проверок, следовало бы ожидать всего один отрицательный исход. У нас же насчитывается три (в табл. 2 для этих наблюдений — восемь отрицательных исходов).

Таким образом, распределение разностей y_{ijk} после исключения по формуле (5) ошибки, обуславливающей неточность принятого масштаба, приблизилось к нормальному. Но его еще нельзя считать нормальным; а это значит, что указанная ошибка не единственная превалирующая ошибка рассмотренных рядов наблюдений.

Распределение ошибок длительных широтных наблюдений, по-видимому, близко к распределению Пирсона типа VII [1, 2, 14]. При некоторых ограничительных условиях функция (2) также стремится к плотности распределения Пирсона типа VII.

ЛИТЕРАТУРА

1. Hulme H. R. and Symms L. S. T. — MN, Roy. Astr. Soc., **99**, 8, 1939.
2. Jeffreys H. — MN, **99**, 8, 1939.
3. Jeffreys H. Theory of Probability. Sec. ed. Oxford. 1940.
4. Thompson W. R. Annales of Mathematical Statistics, N 6, 214, 1935.
5. Arley N. K. Danske Vid. Selsk. Mat.-fys. Medd., **18**, 3, 1940.
6. Birnbaum Z. W. — Journal of the Amer. Stat. Association, N 47, 425, 1952.
7. Miller L. — Journal of the Amer. Stat. Association, N 51, 111, 1956.
8. Publications of the International latitude Observatory of Mizusawa, **1**, 2, 1953.
9. Publications of the International latitude Observatory of Mizusawa, **1**, 1, 1951.
10. Федоров Е. П. О задачах и программе Международной службы широты. М., 1954.
11. Дунин-Барковский И. В. и Смирнов Н. В. Теория вероятностей и математическая статистика в технике. М., 1955.

12. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики, «Наука». М., 1965.
13. Вентцель Е. С. Теория вероятностей, М., 1962.
14. Джунь И. В. — В кн.: Вопросы астрометрии, 4. «Наукова думка», К., 1968.

ON DISTRIBUTION FUNCTION OF THE ERRORS IN LATITUDE OBSERVATIONS

I. V. DZHUN

Summary

The method of small samples is used to test a hypothesis of the normal distribution of latitude differences obtained from concurrent observations with two zenith telescopes of Poltava and Mizusawa. It is shown that due to systematic errors the sample distribution function differs significantly from the normal one.

ОБ ОЦЕНКЕ СПЕКТРА ОШИБОК ШИРОТНЫХ НАБЛЮДЕНИЙ В ОБЛАСТИ ВЫСОКИХ ЧАСТОТ

В. Н. Корогвич, Я. С. Яцкив

(ГАО АН УССР, г. Киев)

Постановка задачи

Наиболее полной характеристикой ошибок наблюдений широты, как временной случайной последовательности, является ее спектр мощности во всей области частот от 0 до π [1]. Вследствие неравномерности распределения наблюдений во времени этот спектр можно оценить только на отдельных неперекрывающихся интервалах частот. В работе [2] эти интервалы обозначены A, B, C и D, подробно изучен спектр ошибок по уклонениям нормальных значений от сглаженной кривой изменения широты (интервал C) в области периодов от 2 до 6 недель. Оценки спектра в интервале D, который соответствует неполярным изменениям широты, получены в [3—5]. В каждом из двух других интервалов спектральная плотность предполагалась постоянной [2]

$$S(\omega) = \text{const}.$$

Следовательно, ошибки наблюдений рассматривались как независимые случайные величины.

Существенно упрощая дальнейшие расчеты, такое предположение, если оно противоречит реальности, в значительной мере уменьшает количество информации, которое содержится в выполненных наблюдениях. Например, отвергается гипотеза о наличии каких-либо преобладающих источников ошибок, о значимости расхождений характеристик изменений широты в течение рассматриваемых интервалов времени и т. п.

Чтобы частично восполнить этот пробел, в настоящей работе на примере пулковских наблюдений широты по расширенной программе рассмотрена методика оценки и изучения спектра ошибок широтных наблюдений в интервале A (частоты от 4 до 52.5 цикла в сутки). При этом большое внимание уделяется аппроксимации суточных изменений широты.

Пусть $\varphi_1, \varphi_2, \ldots, \varphi_n$ — последовательность наблюденных в течение ночи широт в среднем через промежуток времени Δt. В общем случае