1965 г. Март *Том 85, вып. 3*

УСПЕХИ ФИЗИЧЕСКИХ НАУК

НОВЫЕ ПРИБОРЫ И МЕТОДЫ ИЗМЕРЕНИЙ

519.281

ПОГРЕШНОСТЬ И НАДЕЖНОСТЬ ПРОСТЕЙШИХ ЭКСПЕРИМЕНТОВ

М. И. Корнфельд

Понятие погрешности эксперимента не имеет ясного смысла до тех пор, пока не указана его надежность, т. е. вероятность нахождения искомой величины в пределах погрешности.

Настоящая заметка посвящена оценке надежности самых простых экспериментов. В более сложных случаях читателю следует обратиться к математической теории обработки наблюдений (см. [1]).

1. СРАВНЕНИЕ ОБЪЕКТОВ

Отличить один физический объект от другого, значит, указать, который из них, A или B, обладает определенным признаком.

В силу неизбежных случайных помех отличительный признак может быть, однако, приписан не только объекту, действительно обладающему им, но и другому объекту. Иными словами, существует вероятность W_A, что признак будет приписан объекту A, и вероятность $W_B = 1 - W_A$, что признак будет приписан объекту B.

В соответствии с этим, степень отличимости объектов удобно характеризовать величиной

$$U = W_A - W_B, \tag{1,1}$$

равной нулю для неотличимых, и плюс или минус единице для предельно отличающихся объектов.

Пусть при экспериментальном определении этой величины (т. е. при сравнении объектов) выполнено n опытов (испытаний). Если в m_A случаях признак был приписан объекту A и в $m_B = n - m_A$ случаях объекту B, то степень отличимости приближенно равна

$$\widetilde{U} = \frac{m_A}{n} - \frac{m_B}{n}. \tag{1,2}$$

Задача состоит теперь в том, чтобы на основании приближенного значения степени отличимости $\widetilde{U}$ вынести суждение об истинной степени отличимости U.

Эта задача эквивалентна статистической задаче о нахождении вероятности события на основании ряда независимых испытаний (см., например, [2], стр. 136).

Ее решение показывает, что при $n \to \infty$, а практически уже при $n > 100$, с вероятностью p имеют место неравенства

$$\frac{\tilde{U} - \beta\sqrt{1-\tilde{U}^2+\beta^2}}{1+\beta^2} < U < \frac{\tilde{U} + \beta\sqrt{1-\tilde{U}^2+\beta^2}}{1+\beta^2}, \tag{1,3}$$

где

$$\beta = \frac{t}{\sqrt{n}},$$

а зависимость t от p дана в табл. I.

Таблица I

p	0	0,7	0,9	0,99	0,999
t	0	1,04	1,64	2,58	3,29

При $\beta \ll 1$ вместо неравенств (1,3) можно писать

$$U = \tilde{U} \pm \Delta U,$$

где

$$\Delta U = \beta\sqrt{1-\tilde{U}^2+\beta^2}. \tag{1,4}$$

Мы будем называть в дальнейшем ΔU погрешностью, а p — надежностью определения степени отличимости.

Таким образом, предыдущие формулы позволяют утверждать, с надежностью p, что разность $U - \tilde{U}$ не выходит за пределы погрешности $\pm \Delta U$.

В вопросах науки вряд ли можно удовлетвориться надежностью какого-либо утверждения, меньшей 0,9. Приняв эту цифру, т. е. примирившись с тем, что в одном случае из десяти наше утверждение окажется неверным, легко увидеть из (1,4), что для достижения мало-мальски приемлемой погрешности требуется очень большое число испытаний.

Для $\tilde{U}$, близкого к нулю, погрешности 0,1 отвечают двести пятьдесят испытаний, погрешности 0,01 — двадцать пять тысяч испытаний, погрешности 0,001 — два с половиной миллиона испытаний!

При $\tilde{U} \cong 1$ требуется значительно меньшее число испытаний. Нужно, однако, иметь в виду, что в этом случае начинает сказываться приближенный характер (1,3) и оценка числа испытаний становится неточной.

Для иллюстрации изложенных соображений приведем результаты исследования процедуры сравнения, в которой отличимость объектов устанавливалась невооруженным глазом (М. И. Корнфельд [3]).

Вычерченная тушью на белой бумаге окружность была сфотографирована и с полученного негатива сделаны отпечатки при разных увеличениях. Из этих отпечатков было образовано семь пар. На одном из отпечатков каждой пары круг имел диаметр 12 *мм*, на другом диаметр, отличающийся на $\Delta = 0{,}10$; 0,15; 0,20; 0,25; 0,30; 0,40; 1,00 *мм*.

Экспериментатор, «перетасовывая» с помощью специального приспособления данную пару отпечатков, должен был каждый раз указывать, на котором из них находится больший круг. Число верных (и ошибочных) заключений регистрировалось автоматически и отсюда по формуле (1,2) определялась приближенная степень отличимости кругов $\tilde{U}$.

К работе было привлечено 28 человек. Для каждой пары отпечатков производилось 1000 испытаний различными лицами (по 100 испытаний

Таблица II

Δ, *мм*	0,10	0,15	0,20	0,25	0,30	0,40	1,00
$\tilde{U}$	0,09	0,38	0,57	0,62	0,75	0,78	0,98

за один сеанс). Результаты приведены в табл. II и на рис. 1, где вертикальные черточки отвечают погрешности, вычисленной с помощью формулы (1,4) для надежности 0,9 и числа испытаний 1000.

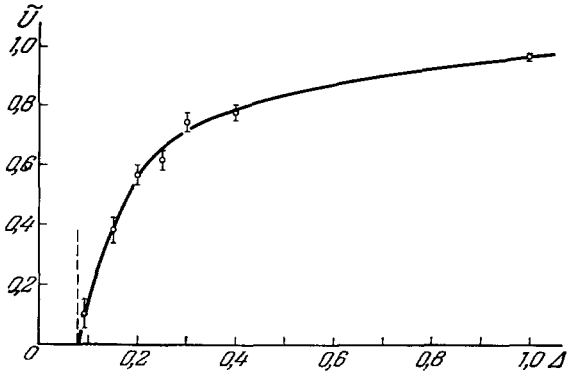

Рис. 1.

Кривая указывает прежде всего на наличие интервала нечувствительности, расположенного на рисунке слева от вертикального пунктира. Различия между кругами, лежащие в пределах этого интервала, не обнаруживаются глазом даже при чрезвычайно большом числе испытаний, что свидетельствует, по-видимому, о дискретности центров восприятия глаза.

Ход кривой за пределами интервала нечувствительности можно объяснить наложением случайных помех на действительное различие между кругами. Чем больше это различие, тем меньшую роль играют помехи и тем, следовательно, ближе к единице степень отличимости.

Полученные данные показывают, что разрешающая способность глаза, в обычном понимании этого термина, зависит от того, с какой уверенностью мы желаем отличать объекты друг от друга, т. е. какая доля ошибочных заключений считается нами допустимой. Например, при 30% ошибочных заключений разрешающая способность равна 0,15 *мм*, при 10% — 0,4 *мм*, при 1% — 1 *мм* и т. д.

2. ИЗМЕРЕНИЕ ВЕЛИЧИН

Вследствие неизбежных случайных помех повторение процедуры измерения величины приводит, как правило, каждый раз к другому результату. Задача состоит в том, чтобы на основании нескольких чисел, полученных при повторных измерениях, вынести суждение об истинной величине.

Эта задача может быть решена следующим образом (см. [4]).

Согласно принятым в предыдущем параграфе определениям при сравнении двух неотличимых объектов одинаково вероятно приписать отличительный признак каждому из них. В терминах операции измерения это значит, что при измерении величины объекта одинаково вероятно получить число большее или меньшее истинной величины.

Обозначим истинную величину через x и расположим полученные при n измерениях числа в порядке возрастания: $x_1, x_2, x_3, \ldots, x_n$

Тогда, в соответствии со сказанным, вероятность неравенства $x < x_1$, так же как и вероятность неравенства $x > x_n$, будет равна $(1/2)^n$.

Отсюда вероятность того, что не удовлетворяется ни одно из этих неравенств, т. е. вероятность неравенства $x_1 < x < x_n$, есть

$$p = 1 - 2(1/2)^n. \qquad (2,1)$$

Таким образом, с надежностью p можно утверждать, что x находится в пределах между x_1 и x_n. В дальнейшем вместо указания границ этого интервала мы будем писать

$$x = \tilde{x} \pm \Delta x,$$

где

$$\tilde{x} = \frac{1}{2}(x_1 + x_n), \qquad (2,2а)$$

$$\Delta x = \frac{1}{2}(x_n - x_1). \qquad (2,2б)$$

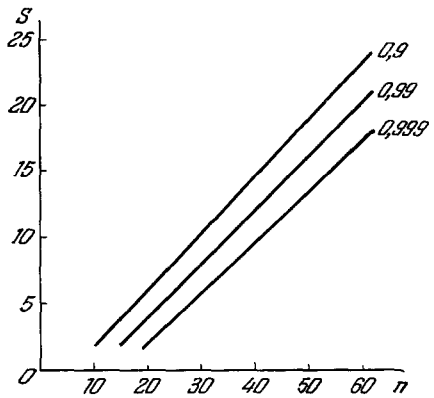

Рис. 2.

Первое из этих чисел будет называться приближенным значением величины, второе — погрешностью измерения. Когда число измерений велико и надежность, вычисленная с помощью (2,1), представляется нам излишне высокой, можно отбросить с начала и конца ряда $x_1, x_2, x_3, \ldots, x_n$ соответственно q и l чисел и тем самым уменьшить погрешность.

Для вычисления надежности полученного этим путем результата служит выражение

$$p = 1 - \frac{1}{2^n}\left(\sum_{k=0}^{q} \frac{n!}{k!\,(n-k)!} + \sum_{k=0}^{l} \frac{n!}{k!\,(n-k)!}\right), \qquad (2,3)$$

представляющее собою обобщение (2,1) на случай неравенства $x_{q+1} < x < x_{n-l}$.

Положив для простоты $q = l = s$, что отвечает уменьшению погрешности с $\frac{1}{2}(x_n - x_1)$ до $\frac{1}{2}(x_{n-s} - x_{s+1})$, и задавшись определенной надежностью, можно с помощью (2,3) найти соответствующее этой надежности число s. В табл. III и на рис. 2 представлена зависимость s от n при $p \cong 0,9$, 0,99 и 0,999.

Таблица III

n	s		
	$p \simeq 0,9$	$p \simeq 0,99$	$p \simeq 0,999$
10	2	—	—
20	6	4	2
30	10	8	6
40	14	12	10
50	19	16	13
60	23	20	17

Рассмотрим еще случай, когда одно из значений измеряемой величины резко отличается от всех остальных. Обычно такое значение

попросту отбрасывается. При этом, однако, снижается не только погрешность, но и надежность результата. Поэтому в каждом конкретном случае нужно сначала взвесить, чем лучше поступиться — погрешностью или надежностью, и в соответствии с этим оставить или отбросить «выскакивающее» измерение.

Положив в формуле (2,3) $q = 0$ и $l = 1$, получим

$$p = [1 - 2(1/2)^n] - \frac{n}{2^n}. \qquad (2,4)$$

Отсюда видно, что отбрасывание выскакивающего значения допустимо лишь при достаточно больших n.

Для иллюстрации изложенного метода оценки истинной величины воспользуемся выполненными Р. Милликеном 58 измерениями заряда электрона (в единицах 10^{-10} CGSE) (табл. IV).

Таблица IV

I	II	III	IV	V	VI
4,781	4,764	4,777	4,809	4,761	4,769
4,795	4,776	4,765	4,790	4,792	4,806
4,769	4,771	4,785	4,779	4,758	4,779
4,792	4,789	4,805	4,788	4,764	4,785
4,779	4,772	4,768	4,772	4,810	4,790
VII	**VIII**	**IX**	**X**	**XI**	**XII**
4,775	4,789	4,801	4,791	4,799	4,777
4,772	4,764	4,785	4,788	4,799	4,749
4,791	4,774	4,783	4,783	4,797	4,781
4,782	4,778	4,808	4,740	4,790	—
4,767	4,791	4,771	4,775	4,747	—

Расположим эти числа в порядке возрастания:

4,740	4,747	4,749	4,758	4,761	4,764
4,764	4,764	4,765	4,767	4,768	4,769
4,769	4,771	4,771	4,772	4,772	4,772
4,774	4,775	4,775	4,776	4,777	4,777
4,778	4,779	4,779	4,779	4,781	4,781
4,782	4,783	4,783	4,785	4,785	4,785
4,788	4,788	4,789	4,789	4,790	4,790
4,790	4,791	4,791	4,791	4,792	4,792
4,795	4,797	4,799	4,799	4,801	4,805
4,806	4,808	4,809	4,810		

Зададимся надежностью $p = 0,9$. Из табл. III видно, что надежности 0,9 при $n = 58$ отвечает $s = 23$. Отбросив с обеих сторон приведенного только что ряда по 23 числа, находим крайние числа оставшейся части: 4,777 и 4,785.

Следовательно, искомое истинное значение заключено в пределах между 4,777 и 4,785, т. е. заряд электрона с надежностью 0,9 равен

$$(4,781 \pm 0,004) \cdot 10^{-10}.$$

Посмотрим далее, как выглядит оценка истинного значения заряда электрона при небольшом числе измерений, равном, например, пяти.

Пользуясь формулой (2,2), находим непосредственно из табл. IV для каждой из одиннадцати групп, отмеченных римскими цифрами, приближенное значение и погрешность. Эти результаты приведены в табл. V; они отвечают, согласно формуле (2,1), надежности $p = 0{,}94$.

Таблица V

Группа	Результат	Группа	Результат
I	4,782±0,013	VII	4,779±0,012
II	4,777±0,012	VIII	4,777±0,014
III	4,775±0,020	IX	4,789±0,019
IV	4,779±0,019	X	4,765±0,025
V	4,784±0,026	XI	4,773±0,026
VI	4,787±0,019		

Как видно из табл. V, уменьшение числа измерений приводит к увеличению погрешности.

3. ИССЛЕДОВАНИЕ СВЯЗЕЙ

Результат экспериментального исследования связи между двумя величинами дается таблицей:

$$\begin{array}{c|cccc} x & x_1 & x_2 & x_3 \ldots & x_N \\ \hline y & y_1 & y_2 & y_3 \ldots & y_N \end{array}, \tag{3,1}$$

где

$$x_1 < x_2 < x_3 < \ldots < x_N.$$

На основании этой таблицы нам надлежит вынести суждение об истинной функции $y = F(x)$.

Полагая, для простоты, что только игреки отягощены ошибками измерений и что эти ошибки не зависят от икса, изобразим данные таблицы на графике, отнеся каждой паре значений (x_i, y_i) соответствующую точку (рис. 3).

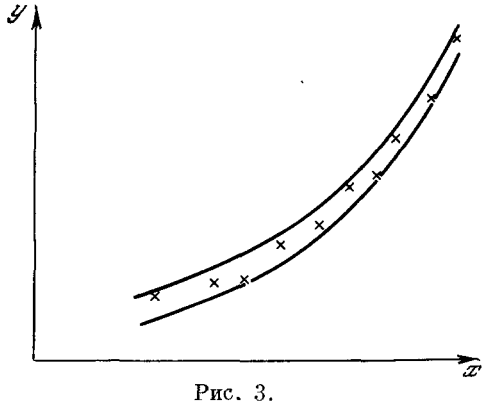

Рис. 3.

Проведем далее две плавные кривые, одну выше всех точек, другую ниже всех точек, таким образом, чтобы расстояние между кривыми (вдоль оси игреков) было минимальным и постоянным на всем протяжении.

Легко видеть, что расстояние между этими ограничивающими кривыми равно удвоенной погрешности, которую мы получили бы, измеряя игрек N раз при некотором фиксированном значении x_i.

Отсюда, в соответствии с (2,1), вероятность нахождения истинной величины игрека $F(x_i)$ между кривыми равна $1 - 2\,(1/2)^N$, а вероятность того, что истинная функция во всех точках $x_1, x_2, x_3, \ldots, x_N$ не выходит за пределы ограничивающей области, будет

$$p = [1 - 2\,(1/2)^N]^N. \tag{3,2a}$$

Заметим здесь же, что если в каждой точке выполнено по нескольку измерений игрека, то

$$p = \left[1 - 2\left(\frac{1}{2}\right)^{\mathfrak{N}}\right]^{N}, \qquad (3,2б)$$

где $\mathfrak{N}$ — общее число измерений.

Впрочем, в этом случае результат может быть представлен и другим способом. Вычислив для каждого x_i приближенное значение $\tilde{y}_i$ и погрешность Δy_i, мы наносим на графике вместо точки (x_i, y_i) вертикальную

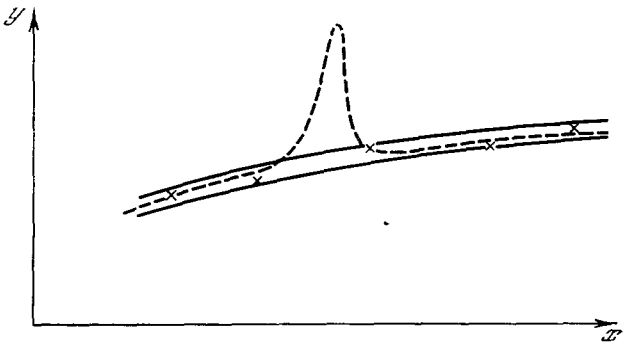

Рис. 4.

черточку, длина которой равна $2\Delta y_i$, а середина совпадает с $\tilde{y}_i$. Проведя ограничивающие кривые по верхним и, соответственно, по нижним концам черточек, имеем

$$p = p_1 p_2 p_3 \ldots p_N, \qquad (3,2в)$$

где $p_1, p_2, p_3, \ldots$ — надежности, отвечающие приближенным значениям $\tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \ldots$

Само собой разумеется, что ограничения, налагаемые предыдущими формулами, не относятся к промежуткам между соседними значениями x. При p сколь угодно близком к единице истинная функция может иметь, например, вид, изображенный на рис. 4.

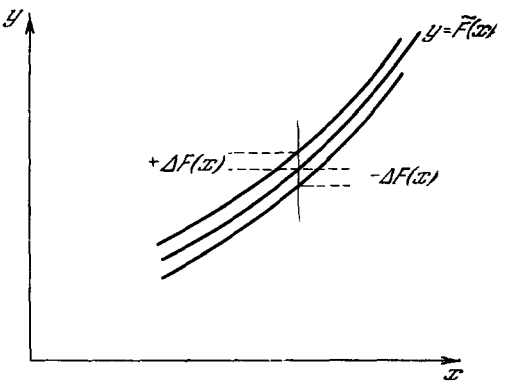

Рис. 5.

Но даже тогда, когда у нас есть основания полагать, что в «просветах» между измерениями функция не выходит за пределы ограничивающей области, ее нельзя определить вполне точно, так как внутри области заключено множество различных функций и неизвестно, которая из них является истинной.

Можно поэтому назвать кривую $y = \tilde{F}(x)$, проходящую посередине между ограничивающими кривыми, приближенной функцией, половину интервала между ограничивающими кривыми $\Delta F(x)$ — погрешностью, а вероятность p — надежностью исследования связи (рис. 5).

Данные таблицы (3,1) могут быть представлены и в аналитической форме. С этой целью, пользуясь графиком приближенной функции, мы

находим для равноотстоящих значений x: x_0, $x_0 + \omega$, $x_0 + 2\omega$ . . .— соответствующие им значения y: y_0, y_1, y_2, . . . Далее образуем первые разности: $\Delta_0^{(1)} = y_1 - y_0$, $\Delta_1^{(1)} = y_2 - y_1$, $\Delta_2^{(1)} = y_3 - y_2$, . . ., вторые разности: $\Delta_0^{(2)} = \Delta_1^{(1)} - \Delta_0^{(1)}$, $\Delta_1^{(2)} = \Delta_2^{(1)} - \Delta_1^{(1)}$, . . ., третьи разности: $\Delta_0^{(3)} = \Delta_1^{(2)} - \Delta_0^{(2)}$, . . . и т. д.:

$$
\begin{array}{ccccc}
x_0 & y_0 & & & \\
 & & \Delta_0^{(1)} & & \\
x_1 & y_1 & & \Delta_0^{(2)} & \\
 & & \Delta_1^{(1)} & & \Delta_0^{(3)} \\
x_2 & y_2 & & \Delta_1^{(2)} & \\
 & & \Delta_2^{(1)} & & \Delta_1^{(3)} \\
x_3 & y_3 & & \Delta_2^{(2)} & \\
 & & \Delta_3^{(1)} & & \Delta_2^{(3)} \\
x_4 & y_4 & & \Delta_3^{(2)} & \\
 & & \Delta_4^{(1)} & & \\
x_5 & y_5 & & & \\
\cdot & \cdot & & & \\
\cdot & \cdot & & &
\end{array}
$$

Если m-е разности постоянны, а $(m + 1)$-е равны нулю, то рассматриваемая функция представима полиномом $y = \tilde{a}_0 + \tilde{a}_1 x + \tilde{a}_2 x^2 + \ldots + \tilde{a}_m x^m$. Определив таким образом степень полинома, нетрудно затем вычислить его коэффициенты:

$$
m = 1 \begin{cases} \tilde{a}_1 = \dfrac{\Delta^{(1)}}{\omega}, \\ \tilde{a}_0 = y - x\tilde{a}_1; \end{cases} \tag{3,3а}
$$

$$
m = 2 \begin{cases} \tilde{a}_2 = \dfrac{\Delta^{(2)}}{2\omega^2}, \\ \tilde{a}_1 = \dfrac{\Delta^{(1)}}{\omega}, \\ \tilde{a}_0 = y - x\tilde{a}_1 - x^2\tilde{a}_2 \end{cases} \tag{3,3б}
$$

и т. д., где y, x, x^2, . . ., $\Delta^{(1)}$, $\Delta^{(2)}$, . . . отвечают значениям, усредненным по всем строкам таблицы разностей.

Нам остается теперь найти из графика погрешность $\Delta F(x)$ и записать полученный результат в виде

$$
y = \tilde{a}_0 + \tilde{a}_1 x + \tilde{a}_2 x^2 + \ldots + \tilde{a}_m x^m \pm \Delta F(x). \tag{3,4}
$$

Для иллюстрации выводов настоящего параграфа рассмотрим следующий пример.

Таблица VI

t° C	y, сек	t° C	y, сек
—10,0	2,60	9,4	0,94
—5,4	2,01	14,8	1,06
1,0	1,34	19,4	1,25
4,6	1,08		

Нормальной температурой эксплуатации хронометра считается $T = 15°$ С. П. Г. Штулькерц, исследовав влияние отклонений от этой температуры $t = (T - 15)$ на суточный ход хронометра, получил результаты, приведенные в табл. VI.

Требуется оценить истинную функцию $y = F(t)$.

Нанесем на миллиметровке данные таблицы и проведем ограничивающие кривые (рис. 6). Согласно формуле (3,2а) при $N = 7$ $p \cong 0{,}9$. Тем

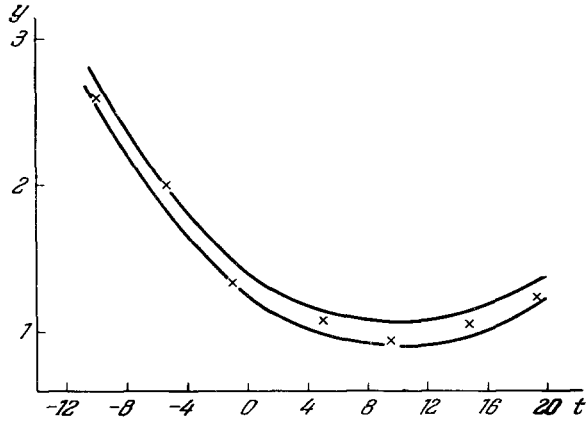

Рис 6

самым можно утверждать, с этой надежностью, что истинная функция заключена между ограничивающими кривыми.

Представим теперь результат в аналитической форме. Для срединной кривой (не показанной на рисунке), находим значения y, отвечающие равноотстоящим значениям t, и строим табл. VII.

Таблица VII

t	y	$\Delta^{(1)}$	$\Delta^{(2)}$	t^2
—10	2,62			100
		—0,32		
—8	2,30		0,02	64
		—0,30		
—6	2,00		0,04	36
		—0,26		
—4	1,74		0,02	16
		—0,24		
—2	1,50		0,06	4
		—0,18		
0	1,32		0,06	0
		—0,12		
2	1,20		0,02	4
		—0,10		
4	1,10		0,04	16
		—0,06		
6	1,04		0,02	36
		—0,04		
8	1,00		0,02	64
		—0,02		
10	0,98		0,00	100
		+0,02		
12	1,00		0,02	144
		+0,04		
14	1,04		0,02	196
		+0,06		
16	1,10		0,04	256
		+0,10		
18	1,20		0,02	324
		+0,12		
20	1,32			400
5,00	1,403	—0,0866	0,0285	110

Поскольку, как видно из таблицы, вторые разности примерно постоянны, приближенная функция может быть аппроксимирована

полиномом

$$y = \tilde{a}_0 + \tilde{a}_1 t + \tilde{a}_2 t^2.$$

Обратившись к (3,3б), вычислим коэффициенты:

$$\tilde{a}_2 = \frac{0{,}0285}{2 \cdot 4} = 0{,}0036,$$

$$\tilde{a}_1 = -\frac{0{,}0866}{2} - 2 \cdot 5{,}00 \cdot 0{,}0036 = -0{,}079,$$

$$\tilde{a}_0 = 1{,}403 + 5{,}00 \cdot 0{,}079 - 110 \cdot 0{,}0036 = 1{,}40.$$

Из графика рис. 6 находим погрешность $\Delta F(x) = 0{,}08$. Таким образом, окончательно имеем

$$y = 1{,}40 - 0{,}079t + 0{,}0036t^2 \pm 0{,}08.$$

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. Ю. В. Линник, Метод наименьших квадратов и основы теории обработки наблюдений, М., Физматгиз, 1962.
2. Н. В. Смирнов, И. В. Дунин-Барковский, Краткий курс математической статистики, М., Физматгиз, 1959.
3. М. И. Корнфельд, ЖТФ 24, 448 (1956).
4. М. И. Корнфельд, ДАН СССР 103, 213 (1955).