ГОСУДАРСТВЕННЫЙ СТАНДАРТ
СОЮЗА ССР

ПРИКЛАДНАЯ СТАТИСТИКА

ПРАВИЛА ПРОВЕРКИ СОГЛАСИЯ ОПЫТНОГО РАСПРЕДЕЛЕНИЯ С ТЕОРЕТИЧЕСКИМ

ГОСТ 11.006—74

Издание официальное

ГОСУДАРСТВЕННЫЙ КОМИТЕТ СТАНДАРТОВ
СОВЕТА МИНИСТРОВ СССР
Москва

РАЗРАБОТАН Всесоюзным научно-исследовательским институтом стандартизации (ВНИИС)

Директор института, д-р экон. наук, проф. **Гличев А. В.**
Научный руководитель, канд. техн. наук **Бендерский А. М.**
Исполнители: **Лосицкий О. Г., Диманштейн С. В., Шкотт З. Н., Лившиц И. Г., Шуленина Л. А., Шилова Л. С., Воробьева В. К., Чернобривко Л. К., Богатырев А. А., Пшеничникова Л. С.**

ВНЕСЕН Всесоюзным научно-исследовательским институтом стандартизации (ВНИИС)

Директор института, д-р экон. наук **Гличев А. В.**

ПОДГОТОВЛЕН К УТВЕРЖДЕНИЮ Всесоюзным научно-исследовательским институтом стандартизации (ВНИИС)

Директор, д-р экон. наук **Гличев А. В.**

УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Постановлением Государственного комитета стандартов Совета Министров СССР от 20 декабря 1974 г. № 2766

УДК 658.562.012.7(083.74) Группа Т59

ГОСУДАРСТВЕННЫЙ СТАНДАРТ СОЮЗА ССР

Прикладная статистика

ПРАВИЛА ПРОВЕРКИ СОГЛАСИЯ ОПЫТНОГО РАСПРЕДЕЛЕНИЯ С ТЕОРЕТИЧЕСКИМ

Applied statistics. The tests for goodness of fit distributions of empirical

ГОСТ 11.006—74

Постановлением Государственного комитета стандартов Совета Министров СССР от 20 декабря 1974 г. № 2766 срок действия установлен

с 01.01 1976 г.
до 01.01 1981 г.

Настоящий стандарт распространяется на продукцию всех отраслей народного хозяйства и устанавливает правила проверки согласия распределения случайной величины, полученной по результатам наблюдений с предполагаемым теоретическим распределением этой величины по критериям Колмогорова, χ^2, ω^2.

На основании настоящего стандарта устанавливают виды распределений показателей качества продукции и технико-экономических показателей, включаемых в нормативно-технические документы.

1. ОБЩИЕ ПОЛОЖЕНИЯ

1.1. Процедура проверки согласия опытного и теоретического распределений случайной величины x заключается в получении упорядоченного ряда результатов наблюдений этой величины

$$x_1 \leqslant x_2 \leqslant, \ldots, \leqslant x_n,$$

построении на основании их функции накопленных частостей и сравнении этой функции с заданной теоретической функцией.

Процедура осуществляется с целью установления функции распределения случайной величины x, применяемой для различных теоретико-вероятностных расчетов, статистического анализа, а также обоснования выбора планов статистического контроля, регулирования и испытаний качества продукции.

1.2. Наблюдения случайной величины x должны проводиться в практически одинаковых условиях, исследуемая совокупность должна быть однородной. Нарушение требований однородности может привести к ошибочным выводам.

1.3. Число наблюдений случайной величины x для проверки согласия опытного и теоретического распределений должно быть больше 100, если используют критерии Колмогорова и χ^2, и больше 50—если используют критерий ω^2.

1.4. Для наблюдений случайной величины x должны применяться средства измерения с ценой деления, не превышающей $^1/_5$ предполагаемой величины среднего квадратического отклонения исследуемого распределения.

1.5. В приложении 1 даны примеры проверки согласия опытного и теоретического распределений, в приложении 2—теоретическое обоснование стандарта, в приложении 3 — перечень справочной литературы.

2. КРИТЕРИЙ КОЛМОГОРОВА

2.1. Результаты наблюдений случайной величины x, полученные в специально поставленном эксперименте или на основании сбора статистических данных в условиях, соответствующих требованиям п. 1.2, располагают в порядке их возрастания

$$x_1 \leqslant x_2 \leqslant, \ldots, \leqslant x_n.$$

2.2. В первую графу табл. 1 записывают значения

$$x_1;\ x_1+\Delta;\ x_1+2\Delta;\ \ldots;x_r,$$

где Δ — цена деления применяемого средства измерения (приращение интервала).

Во вторую графу табл. 1 записывают частоты $m_1, m_2, \ldots, m_r$, с которыми встречаются величины x_1; $x_1+\Delta$; $x_1+2\Delta$; ...

Если величина $x_1+j\Delta$ не получена в результате наблюдений, то $m_{j+1}=0$.

2.3. Вычисляют среднее арифметическое значение и среднее квадратическое отклонение

$$\bar{x}=\frac{1}{n}\sum_{j=0}^{r} m_{j+1}(x_1+j\Delta); \qquad (1)$$

$$S=\sqrt{\frac{1}{n-1}\sum_{j=0}^{r} m_{j+1}(x_1+j\Delta-\bar{x})^2}, \qquad (2)$$

где $n=m_1+m_2+\ldots+m_r$;

r — количество интервалов, которое вычисляют по формуле $r=\frac{x_n-x_1}{\Delta}$. Данные вычисления располагают в третьей — шестой графах табл. 1.

2.4. В седьмую графу табл. 1 записывают величины

$$y_{j+1}=\frac{x_1+j\Delta-\overline{x}}{S}; \tag{3}$$

в восьмую графу — функцию опытного распределения

$$F_n(y_{j+1})=\frac{n_{j+1}}{n},$$

где $n_{j+1}=m_1+m_2+\ldots+m_{j+1}$, а в девятую графу — функцию теоретического распределения.

2.5. По данным табл. 1 строят график функций опытного и теоретического распределений $F_n(y_{j+1})$ и $F(y_{j+1})$.

По графику и данным табл. 1 определяют максимальное отклонение функции опытного распределения от функции теоретического распределения

$$D_n=\max\ |F_n(y_{j+1})-F(y_{j+1})| \tag{4}$$

и вычисляют величину $\lambda_n=D_n\sqrt{n}$.

2.6. Задаются доверительной вероятностью

$$\gamma=\text{Вер}\ \{\lambda_n\leqslant\lambda_n^*\} \tag{5}$$

того, что отклонение функции опытного распределения от теоретического будет меньше величины λ_{n*}, установленной для доверительной вероятности γ. В табл. 2 находят значение, соответствующее этой доверительной вероятности.

При необходимости строят доверительную область для теоретического распределения. С этой целью вычисляют

$$D_n^*=\frac{\lambda_n^*}{\sqrt{n}}, \tag{6}$$

и на график наносят доверительные границы

$$F_u\ (y)=F\ (y)+D_n^*; \tag{7}$$

$$F_L\ (y)=F(y)-D_n^*. \tag{8}$$

Если нанесенная на графике функция опытного распределения $F_n(y_{j+1})$ не выйдет за доверительные границы, то гипотеза о согласии принимается, в противном случае — отвергается.

Таблица 1

Данные для построения функций опытного и теоретического распределений

$x_1+j\Delta$	m_{j+1}	$m_{j+1}(x_1+j\Delta)$	$x_1+j\Delta-\overline{x}$	$(x_1+j\Delta-\overline{x})^2$	$(x_1+j\Delta-\overline{x})^2\times$ $\times m_{j+1}$	$y_{j+1}=\frac{x_1+j\Delta-\overline{x}}{S}$	$F_n(y_{j+1})$	$F(y_{j+1})$
1	2	3	4	5	6	7	8	9
x_1								
$x_1+\Delta$								
.								
.								
.								
x_r								
		$\sum_{j=0}^{r} m_{j+1}(x_1+j\Delta)$			$\sum_{j=0}^{r} m_{j+1}(x_1+j\Delta-\overline{x})^2$			

Примечания:

1. Функция теоретического распределения $F(y)$ представляет собой вероятность того, что при принятом распределении случайной величины y будет выполняться условие $-\infty<y\leqslant y_k$. Значения этой функции приведены в специальных таблицах (см. [10] в приложении 3). При переносе данных в табл. 1 необходимо убедиться, что начало отсчета аргумента и масштаб его таковы, что математическое ожидание и дисперсия случайной величины y соответственно равны 0 и 1. Если это требование не выполнено, необходимо произвести перерасчет с изменением начала отсчета и масштаба.

2. Если $r=\frac{x_n-x_1}{\Delta}>60$, то допускается разбиение участка на равные интервалы длиной 2Δ, 3Δ, . . . таким образом, чтобы число интервалов было более 30. В этом случае m_{j+1} будет равно количеству результатов наблюдений, попавших в интервал $(j-1)h$, jh $(h=2\Delta, 3\Delta, \ldots)$.

Таблица 2

Предельные значения нормированных отклонений опытного распределения от значений теоретического распределения для заданных доверительных вероятностей

γ	λ_n^*	γ	λ_n^*
0,01	0,44	0,60	0,89
0,05	0,52	0,70	0,97
0,10	0,57	0,80	1,07
0,15	0,61	0,90	1,22
0,20	0,65	0,95	1,36
0,30	0,71	0,98	1,52
0,40	0,77	0,99	1,63

3. КРИТЕРИЙ χ^2

3.1. Результаты наблюдений случайной величины x в специально поставленном эксперименте или на основании сбора статистических данных в условиях, соответствующих требованиям п. 1.2, располагают в порядке возрастания

$$x_1 \leqslant x_2 \ldots \leqslant x_n.$$

3.2. Вычисляют размах $x_n - x_1$ и образуют r равных интервалов шириной

$$h = \frac{x_n - x_1}{r}. \tag{9}$$

Число интервалов r выбирают в зависимости от объема выборки n:

при $n=200$ $r=18 \div 20$;
при $n=400$ $r=25 \div 30$;
при $n=1000$ $r=35 \div 40$.

Примечание. При $100<n<200$ критерий χ^2 применяют в исключительных случаях с числом интервалов $r=15 \div 18$.

3.3. Результаты наблюдений x_j группируют по интервалам, подсчитывают частоты m_j величин x_j, попавшие в j-е интервалы. Эти частоты записывают во вторую графу табл. 3.

3.4. Вычисляют среднее арифметическое значение и среднее квадратическое отклонение

$$\bar{x} = \frac{1}{n} \sum_{j=1}^{r} \left(j - \frac{1}{2}\right) m_j; \tag{10}$$

$$S = \sqrt{\frac{1}{n-1} \sum_{j=1}^{r} \left\{\left(j - \frac{1}{2}\right) - \bar{x}\right\}^2 m_j} \tag{11}$$

Таблица 3

Данные для проверки согласия опытного и теоретического распределений по критерию χ^2

j	m_j	$\left(j-\frac{1}{2}\right)m_j$	$\left\{\left(j-\frac{1}{2}\right)-\bar{x}\right\}^2 m_j$	y_j	$F(y_j)$	p_j	np_j	m_j-np_j	$(m_j-np_j)^2$	$\frac{(m_j-np_j)^2}{np_j}$
1	2	3	4	5	6	7	8	9	10	11
1										
2										
3										
.										
.										
.										
r										

$$n=\sum_{j=1}^{r} m_j$$

3.5. Вычисляют величины

$$y_j = -\frac{\left(j-\frac{1}{2}\right)-\bar{x}}{S}, \qquad (12)$$

которые записывают в пятую графу табл. 3. В шестую графу табл. 3 переписывают из специальных таблиц значения функции проверяемого теоретического распределения $F(y_j)$.

В седьмую графу табл. 3 записывают вероятности попадания опытных данных в j-й интервал p_j.

$$p_1 = F(y_1);$$

$$p_j = F(y_j) - F(y_{j-1}),\ j=2, \ldots, r.$$

Примечание. Если в восьмой графе табл. 3 окажутся значения $np_j<10$, то следует объединить интервал, в котором ожидаемое число результатов наблюдений меньше десяти, с одним или несколькими соседними интервалами таким образом, чтобы в новом интервале ожидаемое число результатов наблюдений было не менее десяти.

3.6. Вычисляют данные, указанные в графах 9—11 табл. 3 и по ним

$$\chi^2 = \sum_{j=1}^{r} \frac{(m_j - np_j)^2}{np_j}. \qquad (13)$$

3.7. Задаются доверительной вероятностью

$$\gamma = \text{Вер}\{\chi^2 \leqslant (\chi^*)^2\}$$

того, что величина χ^2, полученная вследствие случайных отклонений частостей опытного распределения от соответствующих вероятностей теоретического распределения, будет меньше значения $(\chi^*)^2$, установленного для доверительной вероятности γ. В табл. 4 для доверительной вероятности и числа степеней свободы $k=r-1$ находят величину $(\chi^*)^2/k$, вычисляют $(\chi^*)^2$ и сравнивают с ним вычисленную по данным табл. 3 величину χ^2. Если χ^2 окажется меньше $(\chi^*)^2$, то для принятой доверительной вероятности гипотеза о согласии опытного и теоретического распределений принимается, в противном случае — отвергается.

4. КРИТЕРИЙ ω^2

4.1. Критерий ω^2 является более мощным, чем критерий χ^2 и Колмогорова, но его применение требует выполнения большого количества вычислительных операций. Критерий ω^2 может быть применен, если число наблюдений превышает 50. Его применение является обязательным, если число наблюдений меньше 200; если число наблюдений более 200, то его применение рекомендуется в случаях, когда результаты проверки по другим критериям не позволяют сделать безусловный вывод о согласии опытного и теоретического распределений, например, если при проверке согласия по критерию χ^2 гипотеза принята при уровне значимости 0,1 и отвергнута при уровне значимости 0,05, то следует дополнительно применить критерий ω^2.

Таблица 4

k	Квантили χ^2-распределения $(\chi^*)^2/k$ при доверительной вероятности γ							
	0,001	0,005	0,010	0,025	0,05	0,10	0,20	0,30
15	0,232	0,307	0,349	0,418	0,484	0,570	0,687	0,781
16	0,246	0,321	0,363	0,432	0,498	0,582	0,697	0,789
17	0,260	0,335	0,377	0,445	0,510	0,593	0,706	0,796
18	0,272	0,348	0,390	0,457	0,522	0,604	0,714	0,802
19	0,285	0,360	0,402	0,469	0,532	0,613	0,722	0,808
20	0,296	0,372	0,413	0,480	0,543	0,622	0,729	0,813
22	0,317	0,393	0,434	0,499	0,561	0,638	0,742	0,823
24	0,337	0,412	0,452	0,517	0,577	0,652	0,753	0,831
26	0,355	0,429	0,469	0,532	0,592	0,665	0,762	0,838
28	0,371	0,445	0,484	0,547	0,605	0,676	0,771	0,845
30	0,386	0,460	0,498	0,560	0,616	0,687	0,779	0,850
35	0,420	0,491	0,529	0,588	0,642	0,708	0,795	0,862
40	0,448	0,518	0,554	0,611	0,663	0,726	0,809	0,872
45	0,472	0,540	0,576	0,630	0,680	0,741	0,820	0,880
50	0,494	0,560	0,594	0,647	0,695	0,754	0,829	0,886

Продолжение табл. 4

k	Квантили χ^2-распределения $(\chi^*)^2/k$ при доверительной вероятности γ										
	0,40	0,50	0,60	0,70	0,80	0,90	0,95	0,975	0,99	0,995	0,999
15	0,869	0,956	1,05	1,15	1,29	1,49	1,67	1,83	2,04	2,19	2,51
16	0,874	0,959	1,05	1,15	1,28	1,47	1,64	1,80	2,00	2,16	2,45
17	0,879	0,961	1,05	1,15	1,27	1,46	1,62	1,78	1,97	2,10	2,40
18	0,883	0,963	1,05	1,14	1,26	1,44	1,60	1,75	1,93	2,06	2,35
19	0,887	0,965	1,05	1,14	1,26	1,43	1,59	1,73	1,90	2,03	2,31
20	0,890	0,967	1,05	1,14	1,25	1,42	1,57	1,71	1,88	2,00	2,27
22	0,897	0,970	1,05	1,13	1,24	1,40	1,54	1,67	1,83	1,95	2,19
24	0,902	0,972	1,05	1,13	1,23	1,38	1,52	1,64	1,79	1,90	2,13
26	0,907	0,974	1,05	1,12	1,22	1,37	1,50	1,61	1,76	1,86	2,08
28	0,911	0,976	1,04	1,12	1,22	1,35	1,48	1,59	1,72	1,82	2,03
30	0,915	0,978	1,04	1,12	1,21	1,34	1,46	1,57	1,70	1,79	1,99
35	0,922	0,981	1,04	1,11	1,19	1,32	1,42	1,52	1,64	1,72	1,90
40	0,928	0,983	1,04	1,10	1,18	1,30	1,39	1,48	1,59	1,67	1,84
45	0,933	0,985	1,04	1,10	1,17	1,28	1,37	1,45	1,55	1,63	1,78
50	0,937	0,987	1,04	1,09	1,16	1,26	1,35	1,43	1,52	1,59	1,73

Таблица 5

Форма для вычисления величины

Номер наблюдения в вариационном ряду	$\frac{2j-1}{2n}$	$F(x_j)$	ln(3)	(2)·(4)	1–(2)	1–(3)
1	2	3	4	5	6	7
1	$\frac{1}{2n}$	$F(x_1)$	$\ln F(x_1)$	$\frac{1}{2n}\ln F(x_1)$	$1-\frac{1}{2n}$	$1-F(x_1)$
2	$\frac{3}{2n}$	$F(x_2)$	$\ln F(x_2)$	$\frac{3}{2n}\ln F(x_2)$	$1-\frac{3}{2n}$	$1-F(x_2)$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
j	$\frac{2j-1}{2n}$	$F(x_j)$	$\ln F(x_j)$	$\frac{2j-1}{2n}\ln F(x_j)$	$1-\frac{2j-1}{2n}$	$1-F(x_j)$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
n	$\frac{2n-1}{2n}$	$F(x_n)$	$\ln F(x_n)$	$\frac{2n-1}{2n}\ln F(x_n)$	$1-\frac{2n-1}{2n}$	$1-F(x_n)$

Продолжение

Номер наблюдения в вариационном ряду	ln(7)	(6)·(8)	(5)+(9)
1	8	9	10
1	$\ln[1-F(x_1)]$	$\left(1-\frac{1}{2n}\right)\ln[1-F(x_1)]$	$\frac{1}{2n}\ln F(x_1)+\left(1-\frac{1}{2n}\right)\ln[1-F(x_1)]$
2	$\ln[1-F(x_2)]$	$\left(1-\frac{3}{2n}\right)\ln[1-F(x_2)]$	$\frac{3}{2n}\ln F(x_2)+\left(1-\frac{3}{2n}\right)\ln[1-F(x_2)]$
.	.	.	.
.	.	.	.
.	.	.	.
j	$\ln[1-F(x_j)]$	$\left(1-\frac{2j-1}{2n}\right)\ln[1-F(x_j)]$	$\frac{2j-1}{2n}\ln F(x_j)+\left(1-\frac{2j-1}{2n}\right)\ln[1-F(x_j)]$
.	.	.	.
.	.	.	.
.	.	.	.
n	$\ln[1-F(x_n)]$	$\left(1-\frac{2n-1}{2n}\right)\ln[1-F(x_n)]$	$\frac{2n-1}{2n}\ln F(x_n)+\left(1-\frac{2n-1}{2n}\right)\ln[1-(Fx_n)]$

Примечание. Цифры в скобках в заголовке таблицы означают номера граф, из которых надо брать числа для вычисления, например, ln (3) означает, что надо вычислить натуральный логарифм числа, содержащегося в графе 3.

4.2. Вычисление по критерию ω^2 проводят в следующем порядке:

а) вычисляют значение величины Ω_n^2 по формуле

$$\Omega_n^2 = -n - 2\sum_{j=1}^{n} \left\{\frac{2j-1}{2n} \ln F(x_j) + \left(1 - \frac{2j-1}{2n}\right) \ln [1 - F(x_j)]\right\}, \qquad (14)$$

где x_j ($j=1,2,\ldots,n$) — результат наблюдений, имеющий j-й номер в вариационном ряду $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$;

$F(x_j)$ — значение функции теоретического распределения при значении аргумента, равном x_j.

Вычисления по формуле (14) рекомендуется сводить в табл. 5. Рекомендуется проводить вычисления с точностью до 5 значащих цифр, округляя окончательный результат до двух значащих цифр.

После заполнения таблицы суммируют значения, занесенные в графу 10. Значение величины Ω_n^2 получают по формуле (14);

б) в табл. 6 находят значение функции a, соответствующее вычисленному значению Ω_n^2. Функция a представляет собой функцию распределения величины Ω_n^2;

в) задают уровень значимости α. Рекомендуется выбирать значение α, равное 0,1 или 0,2;

г) если $a \geqslant (1-\alpha)$, то гипотезу о согласии эмпирического и теоретического распределений отвергают, если $a < (1-\alpha)$, то гипотезу принимают.

Таблица 6

Значение Ω_n^2	Значение функции $a(\Omega_n^2)$ при втором знаке после запятой значения Ω_n^2									
	1	2	3	4	5	6	7	8	9	10
0,0	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
0,1	0,000	0,000	0,000	0,000	0,001	0,001	0,002	0,003	0,005	0,007
0,2	0,010	0,013	0,016	0,020	0,025	0,030	0,035	0,041	0,048	0,055
0,3	0,062	0,070	0,078	0,086	0,095	0,104	0,113	0,122	0,132	0,141
0,4	0,151	0,161	0,171	0,181	0,192	0,202	0,212	0,222	0,233	0,243
0,5	0,253	0,263	0,274	0,284	0,294	0,304	0,313	0,323	0,333	0,343
0,6	0,352	0,361	0,371	0,380	0,389	0,398	0,407	0,416	0,424	0,433
0,7	0,441	0,449	0,458	0,466	0,474	0,482	0,489	0,497	0,504	0,512
0,8	0,519	0,526	0,533	0,540	0,547	0,554	0,560	0,567	0,573	0,580
0,9	0,586	0,592	0,598	0,604	0,610	0,615	0,621	0,627	0,632	0,637

Продолжение

Значение Ω_n^2	Значение функции a (Ω_n^2) при втором знаке после запятой значения Ω_n^2									
	1	2	3	4	5	6	7	8	9	10
1,0	0,643	0,648	0,653	0,658	0,663	0,668	0,673	0,677	0,682	0,687
1,1	0,691	0,696	0,700	0,704	0,709	0,713	0,717	0,721	0,725	0,729
1,2	0,732	0,736	0,740	0,744	0,747	0,751	0,754	0,758	0,761	0,764
1,3	0,768	0,771	0,774	0,777	0,780	0,783	0,786	0,789	0,792	0,795
1,4	0,798	0,800	0,803	0,806	0,809	0,811	0,814	0,816	0,819	0,821
1,5	0,824	0,826	0,828	0,831	0,833	0,835	0,837	0,839	0,842	0,844
1,6	0,846	0,848	0,850	0,852	0,854	0,856	0,858	0,859	0,861	0,863
1,7	0,865	0,867	0,868	0,870	0,872	0,873	0,875	0,877	0,878	0,880
1,8	0,881	0,883	0,884	0,886	0,887	0,889	0,890	0,892	0,893	0,894
1,9	0,896	0,897	0,898	0,900	0,901	0,902	0,903	0,905	0,906	0,907
2,0	0,908	0,909	0,910	0,912	0,913	0,914	0,915	0,916	0,917	0,918
2,1	0,919	0,920	0,921	0,922	0,923	0,924	0,925	0,926	0,927	0,928
2,2	0,929	0,929	0,930	0,931	0,932	0,933	0,934	0,934	0,935	0,936
2,3	0,937	0,938	0,938	0,939	0,940	0,941	0,941	0,942	0,943	0,943
2,4	0,944	0,945	0,945	0,946	0,947	0,947	0,948	0,949	0,949	0,950

ПРИЛОЖЕНИЕ 1 к ГОСТ 11.006—74
Справочное

ПРИМЕРЫ ПРОВЕРКИ СОГЛАСИЯ ОПЫТНОГО И ТЕОРЕТИЧЕСКОГО РАСПРЕДЕЛЕНИЙ

Пример 1. Применение критерия Колмогорова

По результатам измерений диаметра 316 валиков требуется проверить гипотезу о нормальном распределении диаметра валика. Измерения проводились индикаторным прибором с ценой деления $\Delta=0{,}001$ мм.

Результаты измерений и данные их обработки, необходимые для построения функций опытного и теоретического распределений, приведены в табл. 1.

На чертеже приведен график, построенный по этим данным, согласно которому максимальное отклонение между функциями опытного и теоретического распределений при $y=-0{,}6218$ равно $D_{316}\approx 0{,}037$.

Величина

$$\lambda_{316}=D_{316}\sqrt{316}=0{,}65$$

вписывается в любые границы для доверительных вероятностей $\gamma>20$ (см. табл. 2 настоящего стандарта). Поэтому согласие данного опытного распределения с нормальным распределением при математическом ожидании диаметра валика 10,000 мм и среднем квадратическом отклонении 0,008 мм следует считать хорошим.

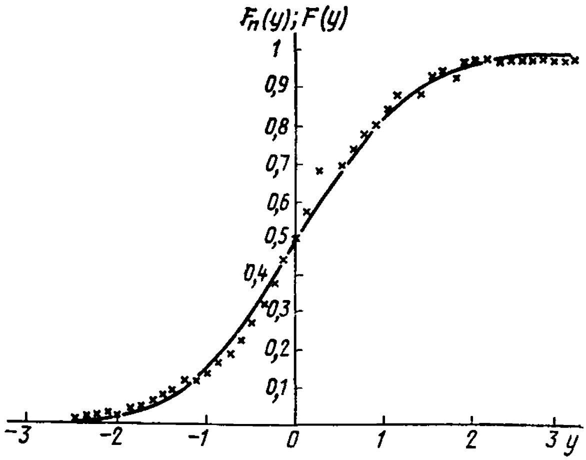
Таблица 1

$x_1+j\Delta$	m_{j+1}	$m_{j+1}(x_1+j\Delta)$	$x_1+j\Delta-\bar{x}$	$(x_1+j\Delta-\bar{x})^2$	$(x_1+j\Delta-\bar{x})^2 m_{j+1}$	$y_{j+1}=\dfrac{x_1+j\Delta-\bar{x}}{S}$	$F_n(y_{j+1})$	$F(y_{j+1})$
1	2	3	4	5	6	7	8	9
9,978	1	9,978	—0,022	0,000484	0,000484	—2,7363	0,00316	0,0031
9,979	0	0	—0,021	0,000441	0	—2,6119	0,00316	0,0045
9,980	2	19,960	—0,020	0,000400	0,000800	—2,4875	0,00949	0,0064
9,981	1	9,981	—0,019	0,000361	0,000361	—2,3631	0,01265	0,0091
9,982	1	9,982	—0,018	0,000324	0,000324	—2,2388	0,01582	0,0125
9,983	3	29,949	—0,017	0,000289	0,000867	—2,1144	0,0253	0,0174
9,984	0	0	—0,016	0,000256	0	—1,9900	0,0253	0,0233
9,985	4	39,940	—0,015	0,000225	0,000900	—1,8656	0,0380	0,0314
9,986	3	29,958	—0,014	0,000196	0,000588	—1,7412	0,0474	0,0409
9,987	6	59,922	—0,013	0,000169	0,001014	—1,6169	0,0664	0,0526
9,988	4	39,952	—0,012	0,000144	0,000576	—1,4925	0,0791	0,0681
9,989	3	29,967	—0,011	0,000121	0,000363	—1,3681	0,0886	0,0869
9,990	8	79,920	—0,010	0,000100	0,000800	—1,2437	0,1139	0,1075

Продолжение табл. 1

$x_1+j\Delta$	m_{j+1}	$m_{j+1}(x_1+j\Delta)$	$x_1+j\Delta-\bar{x}$	$(x_1+j\Delta-\bar{x})^2$	$(x_1+j\Delta-\bar{x})^2 m_{j+1}$	$y_{j+1}=\frac{x_1+j\Delta-\bar{x}}{S}$	$F_n(y_{j+1})$	$F(y_{j+1})$
1	2	3	4	5	6	7	8	9
9,991	0	0	—0,009	0,000081	0	—1,1194	0,1139	0,1314
9,992	7	69,944	—0,008	0,000064	0,000448	—0,9950	0,1361	0,1587
9,993	10	99,930	—0,007	0,000049	0,000490	—0,8706	0,1677	0,1922
9,994	8	79,952	—0,006	0,000036	0,000288	—0,7462	0,1930	0,2266
9,995	12	119,940	—0,005	0,000025	0,000300	—0,6218	0,2310	0,2676
9,996	14	139,944	—0,004	0,000016	0,000224	—0,4975	0,2753	0,3094
9,997	18	179,946	—0,003	0,000009	0,000162	—0,3731	0,3322	0,3557
9,998	16	159,968	—0,002	0,000004	0,000064	—0,2487	0,3829	0,4013
9,999	20	199,98	—0,001	0,000001	0,000020	—0,1243	0,4462	0,4522
10,000	19	190,000	0	0	0	0	0,5063	0,5000
10,001	21	210,021	0,001	0,000001	0,000021	0,1243	0,5727	0,5478
10,002	17	170,034	0,002	0,000004	0,000068	0,2487	0,6265	0,5987
10,003	6	60,018	0,003	0,000009	0,000054	0,3731	0,6455	0,6443
10,004	15	150,060	0,004	0,000016	0,000240	0,4975	0,6930	0,6879
10,005	14	140,070	0,005	0,000025	0,000350	0,6218	0,7373	0,7324
10,006	14	140,084	0,006	0,000036	0,000504	0,7462	0,7816	0,7734
10,007	8	80,056	0,007	0,000049	0,000392	0,8706	0,8069	0,8078
10,008	12	120,096	0,008	0,000064	0,000768	0,9950	0,8449	0,8389
10,009	13	130,117	0,009	0,000081	0,001053	1,1194	0,8860	0,8686
10,010	7	70,070	0,010	0,0001	0,000700	1,2437	0,9082	0,8925
10,011	0	0	0,011	0,000121	0	1,3681	0,9082	0,9147
10,012	8	80,096	0,012	0,000144	0,001152	1,4925	0,9335	0,9319
10,013	4	40,052	0,013	0,000169	0,000676	1,6169	0,9462	0,9474
10,014	5	50,070	0,014	0,000196	0,000980	1,7412	0,9620	0,9591
10,015	2	20,030	0,015	0,000225	0,000450	1,8656	0,9683	0,9693
10,016	4	40,064	0,016	0,000256	0,001024	1,9900	0,9810	0,9767
10,017	0	0	0,017	0,000289	0	2,1144	0,9810	0,9826
10,018	1	10,018	0,018	0,000324	0,000324	2,2388	0,9841	0,9875
10,019	1	10,019	0,019	0,000361	0,000361	2,3631	0,9873	0,9909
10,020	1	10,020	0,020	0,000400	0,000400	2,4875	0,9905	0,9936
10,021	0	0	0,021	0,000441	0	2,6119	0,9905	0,9955
10,022	0	0	0,022	0,000484	0	2,7363	0,9905	0,9968
10,023	2	20,046	0,023	0,000529	0,001058	2,8606	0,9968	0,9979
10,024	0	0	0,024	0,000576	0	2,9850	0,9968	0,9986
10,025	0	0	0,025	0,000626	0	3,1094	0,9968	0,9991
10,026	0	0	0,026	0,000676	0	3,2338	0,9968	0,9994
10,027	1	10,027	0,027	0,000729	0,000729	3,3582	1,0000	0,9996
		$\Sigma=3160,181$			$\Sigma=0,020377$			

$\bar{x}=10,000$ $S=0,008$



Пример 2. Применение критерия χ^2

В процессе специальных испытаний дизелей регистрировалась их наработка, после которой дизель подлежал направлению в капитальный ремонт. Полученные результаты были сгруппированы по интервалам продолжительностью $h=250$ моточасов. Данные о числе образцов, направленных в капитальный ремонт в первом, втором и т. д. интервалах наработки, приведены в первой и второй графах табл. 2. Требуется проверить гипотезу о том, что указанная наработка подчинена экспоненциальному распределению.

Обработка результатов наблюдений, необходимая для оценки этой гипотезы по критерию χ^2, приведена в графах с третьей по четырнадцатую табл. 2 (для удобства вычислений в табл. 2 по сравнению с табл. 3 настоящего стандарта введены дополнительные графы). В итоге получены среднее значение $\bar{x}=8{,}6$ и среднее квадратическое отклонение $S=8{,}5$ исследуемой наработки, измеренной числом интервалов (для перевода данных в моточасы надо $\bar{x}$ и S умножить на $h=250$).

Поскольку для экспоненциального распределения математическое ожидание равно среднему квадратическому отклонению, то начало отсчета наработки при определении нормированных величин y_j сдвинуто на $\bar{x}-S=0{,}1$.

В тринадцатом интервале $np_{13}=9{,}9<10$, в последующих интервалах $np_j<10$. Поэтому, при подсчете величин χ^2 объединены четырнадцатый и пятнадцатый, шестнадцатый и семнадцатый, восемнадцатый и девятнадцатый интервалы. Далее объединены двадцатый — двадцать второй, двадцать третий — двадцать шестой, двадцать седьмой—сороковой интервалы. При этом для всех расширенных интервалов значения

$$n(p_{14}+p_{15}),\ n(p_{16}+p_{17}),$$

$$n(p_{18}+p_{19}),\ n\ (p_{20}+p_{21}+p_{22}),$$

$$n(p_{23}+p_{24}+p_{25}+p_{26}),\ \text{и}$$

$$n(p_{27}+,\ .\ .\ .\ ,+p_{40})$$

Таблица 2

Данные для проверки согласия опытного распределения наработки дизелей до капитального ремонта с экспоненциальным распределением

j	m_j	$\left(j-\frac{1}{2}\right)m_j$	$\left(j-\frac{1}{2}\right)-\bar{x}$	$\left[\left(j-\frac{1}{2}\right)-\bar{x}\right]^2$	$\left[\left(j-\frac{1}{2}\right)-\bar{x}\right]^2 m_j$	$y_j=\frac{x_j-\bar{x}}{S}$	y_j-y_1	$F(y_j-y_1)=$ $=1-e^{-(y_j-y_1)}$	p_j	p_jn	m_j-p_jn	$(m_j-p_jn)^2$	$\frac{(m_j-p_jn)^2}{p_jn}$
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	35	17,5	—8,1	65,61	2296	—0,95	0,00	0,0000	0,1042	37,20	—2,20	4,84	0,128
2	32	48	—7,1	50,41	1613	—0,84	0,11	0,1042	0,1013	36,20	—4,20	17,64	0,488
3	30	75	—6,1	37,21	1116	—0,72	0,23	0,2055	0,0898	32,10	—2,10	4,41	0,137
4	28	98	—5,1	26,01	728	—0,600	0,35	0,2953	0,0797	28,60	—0,60	0,36	0,0126
5	23	103,5	—4,1	16,81	387	—0,48	0,47	0,3750	0,0707	25,30	—2,30	5,29	0,209
6	20	110	—3,1	9,61	192	—0,36	0,59	0,4457	0,0577	20,70	—0,70	0,49	0,0236
7	18	117	—2,1	4,41	79	—0,25	0,70	0,5034	0,0562	20,10	—2,10	4,41	0,219
8	17	127,5	—1,1	1,21	21	—0,130	0,82	0,5596	0,0498	17,82	—0,82	0,67	0,0376
9	17	144,5	—0,1	0,01	0,1	—0,012	0,94	0,6094	0,0441	15,80	1,20	1,44	0,0912
10	15	142,5	0,9	0,81	12	0,11	1,06	0,6535	0,0361	12,90	2,10	4,41	0,341
11	12	126	1,9	3,61	43	0,22	1,17	0,6896	0,0351	12,55	—0,55	0,30	0,0249
12	11	126,5	2,9	8,41	93	0,34	1,29	0,7247	0,0312	11,15	—0,15	0,02	0,0018
13	10	125	3,9	15,21	152	0,46	1,41	0,7559	0,0276	9,90	0,10	0,01	0,00101
14	10	135	4,9	24,01	240	0,58	1,53	0,7835	0,0225	8,05	1,95	3,80	0,473
											1,98	3,92	0,244
15	8	116	5,9	34,81	278	0,69	1,64	0,8060	0,0220	7,97	0,03	0,0009	0,000113
16	7	108,5	6,9	47,61	333	0,81	1,76	0,8280	0,0194	6,95	0,05	0,009	0,00288
											1,85	3,43	0,263
17	8	132	7,9	62,41	499	0,930	1,88	0,8474	0,0193	6,20	1,80	3,24	0,523
18	6	105	8,9	79,21	475	1,05	2,00	0,8647	0,0153	5,48	0,52	0,27	0,0494
											0,04	0,0016	0,0001
19	4	74	9,9	98,01	392	1,17	2,12	0,8800	0,0125	4,48	—0,48	0,23	0,0515

Продолжение табл. 2

j	m_j	$\left(j-\frac{1}{2}\right)m_j$	$\left(j-\frac{1}{2}\right)-\bar{x}$	$\left[\left(j-\frac{1}{2}\right)-\bar{x}\right]^2$	$\left[\left(j-\frac{1}{2}\right)-\bar{x}\right]m_j$	$y_j=\frac{x_j-\bar{x}}{S}$	y_j-y_1	$F(y_j-y_1)=1-e^{-(y_j-y_1)}$	p_j	p_jn	m_j-p_jn	$(m_j-p_jn)^2$	$\frac{(m_j-p_jn)^2}{p_jn}$
1	2	3	4	5	6	7	8	9	10	11	12	13	14
20	8	58,5	10,9	118,81	358	1,28	2,23	0,8925	0,0121	4,34	—1,34	1,79	0,242
											—0,65	0,42	0,036
21	8	61,5	11,9	141,61	425	1,40	2,35	0,9046	0,0108	3,87	—0,87	0,75	0,194
22	5	107,5	12,9	166,41	832	1,50	2,47	0,9154	0,0096	3,44	1,56	2,43	0,707
23	4	90	13,9	193,21	773	1,64	2,59	0,9250	0,0078	2,79	1,21	1,46	0,524
24	2	47	14,9	222,01	444	1,75	2,70	0,9328	0,0076	2,72	—0,72	0,51	0,187
											3,94	15,52	1,54
25	3	73,5	15,9	252,81	758	1,87	2,82	0,9404	0,0067	2,40	0,60	0,36	0,15
26	5	127,5	16,9	285,61	1428	1,99	2,94	0,9471	0,0060	2,15	2,85	8,12	3,78
27	2	53	17,9	320,41	641	2,11	3,06	0,9531	0,0049	1,76	0,24	0,05	0,0284
28	4	110	18,9	357,21	1429	2,22	3,17	0,9580	0,0048	1,72	2,28	5,19	3,05
29	3	85,5	19,9	396,01	1188	2,34	3,29	0,9628	0,0041	1,46	0,60	0,36	0,15
											5,22	27,25	1,62
30	0	0	20,9	436,81	0	2,46	3,41	0,9669	0,0038	1,36	—0,42	0,17	0,405
31	2	61	21,9	479,61	959	2,58	3,53	0,9707	0,0033	1,18	0,82	0,67	0,567
32	3	94,5	22,9	524,41	1573	2,70	3,65	0,9740	0,0030	1,07	1,93	3,72	3,48
33	1	32,5	23,9	571,21	571	2,82	3,77	0,9770	0,0024	0,86	0,14	0,02	0,0233
34	0	0	24,9	620,01	0	2,93	3,88	0,9794	0,0024	0,86	—0,86	0,73	0,85
35	1	34,5	25,9	670,81	671	3,05	4,00	0,9818	0,0020	0,72	0,28	0,08	0,111
36	2	71	26,9	723,61	1447	3,17	4,12	0,9838	0,0018	0,64	1,36	1,85	2,90
37	0	0	27,9	778,41	0	3,29	4,24	0,9856	0,0015	0,54	—0,54	0,29	0,536
38	2	75	28,9	835,21	1670	3,40	4,35	0,9871	0,0015	0,54	1,46	2,13	3,95
39	1	38,5	29,9	894,01	894	3,52	4,47	0,9886	0,0012	0,42	0,58	0,33	0,786
40	1	39,5	30,9	954,81	955	3,63	4,58	0,9898	0,0102	3,65	—0,65	7,02	1,92
	Σ=358	3071,5			Σ=25967								

$$\bar{x}=\frac{3071}{358}=8,6;\quad S=\sqrt{\frac{25967}{357}}=8,5;\quad \chi^2=5,42;\quad \frac{\chi^2}{k}=\frac{5,42}{18}=0,300.$$

больше 10. Эти интервалы в табл. 2 заключены в полужирную рамку. Соответственно вычисляются и все остальные величины, например,

$$\frac{\{m_{14}+m_{15}-n(p_{14}+p_{15})\}^2}{n\,(p_{14}+p_{15})}$$

и т. д.

После объединения указанных интервалов полное число неравновеликих интервалов составляет $13+6=19$ (вместо сорока равновеликих) и число степеней свободы при оценке χ^2 принимается равным $k=19-1=18$.

Окончательно получается

$$\frac{\chi^2}{k}=\frac{5,42}{18}=0,300,$$

что соответствует вероятности $\alpha=0,005$.

Это значит, что гипотеза о согласии наработки дизеля до капитального ремонта с экспоненциальным распределением принимается.

Пример 3. Применение критерия ω^2

Настоящий пример составлен при малом количестве данных с целью иллюстрации вычислительного процесса.

Исходные данные, упорядоченные по величине, приведены ниже в табл. 3.

Таблица 3

Номер наблюдения	Значение	Номер наблюдения	Значение	Номер наблюдения	Значение
1	15,61	6	24,14	11	27,88
2	20,71	7	24,59	12	28,74
3	21,68	8	26,18	13	29,34
4	22,28	9	26,23	14	30,86
5	23,22	10	27,59	15	32,08

Требуется проверить гипотезу о том, что выборка принадлежит нормально распределенной генеральной совокупности. Оценки параметров нормального распределения, вычисленные по исходным данным, равны соответственно $\overline{x}=25,340$; $S=4,319$. Принимаем эти оценки в качестве значений параметров нормального распределения.

Результаты дальнейших вычислений приведены в табл. 4.

Сумма значений, помещенных в графе 10, равна $-7,58828$. Тогда результат вычисления по формуле (14) будет $\Omega_n^2=-n-2(-7,58828)=0,17656\approx0,18$. Значение функции $a\ (\Omega_n^2)$ для $\Omega_n^2=0,18$ равно 0,003. Это значение очень мало; следовательно гипотеза о том, что выборка принадлежит нормально распределенной генеральной совокупности, не может быть отвергнута.

Таблица 4

Результаты вычислений по критерию ω^2

Номер наблюдения в вариационном ряду	$\frac{2j-1}{2n}$	$F(x_j)$	ln(3)	(2)·(4)	1−(2)	1−(3)	ln(7)	(6)·(8)	(5)+(9)
1	2	3	4	5	6	7	8	9	10
1	0,033	0,0121	—4,41454	—0,14715	0,966	0,987870	—0,01217	—0,01176	—0,15892
2	0,100	0,142085	—1,95122	—0,19512	0,900	0,857915	—0,15327	—0,13794	—0,33307
3	0,166	0,198498	—1,61697	—0,26949	0,833	0,801502	—0,22127	—0,18439	—0,45389
4	0,233	0,239473	—1,42920	—0,33348	0,766	0,760527	—0,27378	—0,20990	—0,54338
5	0,300	0,382089	—0,96207	—0,28862	0,700	0,617911	—0,48143	—0,33700	—0,62562
6	0,366	0,390506	—0,94033	—0,34479	0,633	0,609494	—0,49512	—0,31358	—0,65836
7	0,433	0,430933	—0,84188	—0,36481	0,566	0,569067	—0,56370	—0,31942	—0,68424
8	0,500	0,485243	—0,72319	—0,36160	0,500	0,514757	—0,66397	—0,33199	—0,69358
9	0,566	0,581605	—0,54197	—0,30712	0,433	0,418395	—0,87132	—0,37757	—0,68469
10	0,633	0,698817	—0,35839	—0,22698	0,366	0,301183	—1,19998	—0,43999	—0,66697
11	0,700	0,721734	—0,32615	—0,22831	0,300	0,278266	—1,27905	—0,38372	—0,61202
12	0,766	0,784359	—0,24284	—0,18618	0,233	0,215641	—1,53433	—0,35801	—0,54419
13	0,833	0,822777	—0,19504	—0,16253	0,166	0,177223	—1,73048	—0,28841	—0,45095
14	0,900	0,899375	—0,10603	—0,09427	0,100	0,100625	—2,29660	—0,22966	—0,32509
15	0,966	0,940601	—0,06124	—0,05920	0,033	0,059380	—2,82346	—0,09412	—0,15331

ПРИЛОЖЕНИЕ 2 к ГОСТ 11.006—74
Справочное

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СТАНДАРТА

1. **Критерий А. Н. Колмогорова.** В 1933 г. А. Н. Колмогоров доказал важную теорему (см. [1] в приложении 3): если функция $F(x)$ непрерывна, то при $n \longrightarrow \infty$ вероятность

$$\text{Вер}\,\{\sqrt{n}\, D_n < Z\} = \begin{cases} 0 & \text{при } z \leqslant 0 \\ \sum\limits_{k=-\infty}^{\infty} (-1)^k e^{-2k^2z^2} & \text{при } z > 0 \end{cases} \qquad (1)$$

Теорема А. Н. Колмогорова привела к естественному критерию согласия опытного и теоретического распределений. Пусть в результате наблюдений случайная величина x получила значения

$$x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n.$$

Пусть D_n означает фактически найденное максимальное значение разности накопленной частоты $F_n(x)$ и вероятности $F(x)$

$$|F_n(x) - F(x)| \text{ и } \lambda^* = \sqrt{n}\, D^*.$$

Если разность

$$1 - k(\lambda_n^*) = \text{Вер}\{\sqrt{n}\, D_n \geqslant \lambda_n^*\} \qquad (2)$$

мала, то осуществилось маловероятное событие и расхождение между $F_n(x)$ и $F(x)$ следует считать существенным, уже необъяснимым случайностью наблюдаемых значений. Если же разность $1 - k(\lambda^*_n)$ велика, то расхождение между $F_n(x)$ и $F(x)$ надо считать несущественным и гипотезу о функции распределения можно признать согласованной с опытом.

За истекшие сорок лет были проведены многочисленные исследования, связанные с применением теоремы А. Н. Колмогорова. Особенно важные результаты в этом направлении были получены Н. В. Смирновым (см. [2], [3] в приложении 3), Б. В. Гнеденко (см. [4], [5] в приложении 3) и В. Феллером (см. [6] в приложении 3).

Применение критерия согласия А. Н. Колмогорова в практике осложняется следующими обстоятельствами. Для сопоставления функций опытного и теоретического распределений необходимо знать значения математического ожидания и среднего квадратического отклонения наблюдаемой случайной величины, которые обычно неизвестны.

Использование вместо этих неизвестных значений их выборочных оценок $\bar{x}$ и S при малом числе наблюдений (или при большом числе наблюдений, сгруппированных в малом числе интервалов) может привести к ошибочным выводам при проверке согласия опытного и теоретического распределений. В случае проверки нормальности распределения указанную трудность можно обойти сопоставлением функции опытного распределения величин $(x_j - \bar{x})/S$ с функцией распределения Стьюдента. Однако при малом числе наблюдений мощность применяемого критерия будет невысокой, а при большом числе наблюдений распределение Стьюдента сходится к нормальному, причем $\bar{x}$ и S становятся достаточно близкими к их математическим ожиданиям. Таким образом, при большом количестве наблюдений применение критерия А. Н. Колмогорова не вызывает каких-либо осложнений.

Аналогично обстоит дело при проверке гипотезы о других функциях распределения случайных величин, причем в отличие от нормального распределения для многих из них нет табулированных функций непараметрических статистик, подобных распределению Стьюдента при проверке гипотезы нормальности.

Необходимо отметить, что группировка большого числа результатов наблюдений по малому числу интервалов, которая равнозначна применению средств измерений невысокой точности, приводит в ряде случаев к неправильным выводам при проверке согласия опытного и теоретического распределений. Поэтому в стандарте введены ограничения минимального числа наблюдений и максимальной цены деления средств измерений.

2. **Критерий χ^2.** Пусть m_j и p_j суть соответственно частости и вероятности появления события A_j $(j=1, \ldots, r)$ в n независимых испытаниях.

Тогда при больших n величина

$$\chi^2 = \sum_{j=1}^{r} \frac{(m_j - np_j)^2}{np_j} \tag{3}$$

асимптотически подчинена распределению χ^2 с числом степеней свободы $k=r-1$. Если в качестве события рассматривать попадание j-го результата наблюдений над случайной величиной в интервал $(j-1)\,h,\ jh$, то при

$$p_j = F(y_j) - F(y_{j-1}) \tag{4}$$

приведенное выражение χ^2 можно рассматривать как критерий согласия опытного и теоретического распределений.

Для того чтобы можно было вычислить χ^2, нужно знать математические ожидания np_j, а для этого нужно найти оценки $\bar{x}$ и S для среднего значения квадратического отклонения проверяемого теоретического распределения, которые обычно не известны.

В [7] приложения 3 рекомендуется вычислять $\bar{x}$ и S по группированным значениям x и затем для S использовать поправку Шеппарда. При этом все x из интервала $(j-1)h,\ jh$ нужно считать сконцентрированными в средней точке этого интервала $\left(j - \frac{1}{2}\right)h$. С помощью таких модифицированных значений нужно вычислять $\bar{x}$ и S. Для того чтобы можно было применять поправку Шеппарда, нужно, чтобы все интервалы имели одинаковую длину h.

Если имеется очень много интервалов и их середины расположены очень близко друг к другу, то необходимость в применении поправок Шеппарда отпадает.

Второе условие, с которым связано применение критерия χ^2 — назначение числа степеней свободы. Строго говоря, распределение χ^2 с $r-1$ степенями свободы имело бы место лишь в том случае, если бы выражение

$$\chi^2 = \sum_{j=1}^{r} \frac{(m_j - np_j)^2}{np_j} \tag{5}$$

было вычислено с помощью истинных значений $\bar{x}$ и σ, которые неизвестны.

На практике (см. [8] в приложении 3) часто в этих случаях пользуются распределением χ^2 с числом степеней свободы, равным $r-3$. Вместе с тем в [9] приложения 3 показано, что такой выбор числа степеней свободы может привести к серьезному преуменьшению вероятности отклонения правильной гипотезы о функции распределения. В свою очередь применение распределения χ^2 с $r-1$ степенью свободы в некоторых случаях может привести к заметной потере мощности критерия.

Правильный выбор числа степеней свободы зависит от объема выборки n и числа интервалов r. Этот вопрос исследован в ряде работ (см. [9] в приложении 3). Согласно исследованиям число интервалов должно быть таким, чтобы в каждый интервал в среднем попадало при $n=200$—12 наблюдений, при $n=400$—20 наблюдений и при $n=1000$—30 наблюдений. При этом число интервалов n оказывается таким, что результаты проверки согласия опытного и теоретического распределений при числе степеней свободы $r-1$ и $r-3$ получаются практически одинаковыми.

Необходимо отметить, что теория применения критерия χ^2 основана на том, что величины (m_j, np_j) приближенно распределены нормально. Это имеет место, когда величины $np_j > 10$ (см. [7] в приложении 3, стр. 457). Если некоторые значения $np_j < 10$, то целесообразно объединить маленькие группы, чтобы каждая из них после объединения содержала по крайней мере 10 ожидаемых результатов. Если наблюдений так мало, что этого сделать нельзя, то не следует пользоваться критерием χ^2.

3. **Сравнение критериев Колмогорова и χ^2.** В [12] (стр. 613—616) приложения 3 критерий А. Н. Колмогорова и χ^2 сравнивались по значениям $\Delta = \sup\limits_x |F_1(x) - F_0(x)|$, при которых асимптотические мощности критериев А. Н. Колмогорова и χ^2 достигают 0,5. При этом $F_0(x)$ — функция распределения, подлежащая проверке, а $F_1(x)$ — другая функция распределения, имеющая место при альтернативной гипотезе. Под мощностью критерия понимается вероятность $(1-\beta)$, где β — вероятность ошибок второго рода, т. е. вероятность принятия альтернативной гипотезы о наличии функций распределения $F_1(x)$ в то время, когда имеет место нулевая гипотеза о функции распределения $F_0(x)$. Для этих условий вероятностью ошибок первого рода α является вероятность отклонения нулевой гипотезы $F_0(x)$ в то время, когда она верна. Сравнение проводилось для $\alpha=0{,}01$ и $\alpha=0{,}05$. В первом случае критерии оказались очень близки по своим качествам, во втором случае при мощности 0,5 критерий А. Н. Колмогорова отличается от критерия χ^2 отклонением Δ примерно вдвое меньшим. С этой точки зрения критерий А. Н. Колмогорова предпочтительнее критерия χ^2. С практической стороны правильное применение критерия χ^2 требует выбора не равных интервалов, обеспечивающих $np_j > 10$, и, таким образом, является более сложным, чем применение критерия А. Н. Колмогорова. Тем не менее критерий χ^2 является весьма распространенным и при правильном применении дает результаты весьма близкие к критерию А. Н. Колмогорова.

4. **Критерий ω^2.** В отличие от критерия А. Н. Колмогорова, в котором расхождение между эмпирической и теоретической функциями распределения измеряется максимумом абсолютной величины разностей этих функций, критерий ω^2 использует статистику, представляющую собой взвешенную сумму квадратов таких разностей:

$$\omega_n^2 = \int\limits_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \psi[F(x)]\, dF(x),$$

где $F(x)$ — теоретическая функция распределения;

$F_n(x)$ — эмпирическая функция распределения;

$\psi[F(x)]$ — весовая функция, областью определения которой является область значений функции $F(x)$.

Конкретный вид статистики ω_n^2 зависит от вида весовой функции. Обычно используют весовые функции двух видов: $\psi(F)=1$, при которой все значения функции распределения обладают одинаковым весом, и $\psi(F)=\frac{1}{F(1-F)}$, при которой увеличивается вес наблюдений на «хвостах» распределений. В стандарте применяется весовая функция второго вида, поскольку на практике различия между распределениями наиболее отчетливо выступают в области крайних значений случайной величины. Однако, почти всегда имеется мало наблюдений именно в области крайних значений. Поэтому целесообразно придать этим наблюдениям больший вес.

Если принять весовую функцию указанного вида, статистика ω_n^2 после выполнения интегрирования имеет вид

$$\omega_n^2=-n-2\sum_{j=1}^{n}\left\{\frac{2j-1}{2n}\ln F(x_j)+\left(1-\frac{2j-1}{2n}\right)\ln\left[1-F(x_j)\right]\right\},$$

где $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$ — результаты наблюдений, упорядоченные по величине.

В работах [14] и [15] показано, что статистика ω_n^2 подчиняется асимптотическому (при $n \longrightarrow \infty$) распределению.

$$p\{\omega_n^2<x\}=a(x)=\frac{\sqrt{2\pi}}{x}\sum_{j=0}^{\infty}(-1)^j\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(j+1)}(4j+1)\,e^{-\frac{(4j+1)^2\pi^2}{8x}}\times$$

$$\times\int_0^{\infty}e^{\frac{x}{8(y^2+1)}-\frac{(4j+1)^2\pi^2y^2}{8x}}\,dy,$$

значения которого приведены в табл. 6 настоящего стандарта.

ПРИЛОЖЕНИЕ 3 к ГОСТ 11.006—74
Справочное

ПЕРЕЧЕНЬ ЛИТЕРАТУРЫ, ИСПОЛЬЗУЕМОЙ ПРИ РАЗРАБОТКЕ СТАНДАРТА

1. **Гнеденко Б. В.** Курс теории вероятностей. М-Л., ГИТТЛ, 1950.
2. **Смирнов Н. В.** Оценка расхождения между эмпирическими кривыми в двух независимых выборках. Бюллетень МТУ, т. 2, вып. 2, 1939.
3. **Смирнов Н. В.** Приближение законов распределения случайных величин по эмпирическим данным. Усп. матем. наук, т. X, 1944.
4. **Гнеденко Б. В., Королюк В. С.** О максимальном расхождении двух эмпирических распределений. ДАН СССР, 1951, т. 80, с. 525—528.
5. **Гнеденко Б. В., Рвачева Е. Л.** Об одной задаче сравнения двух эмпирических распределений. ДАН СССР, 1952, т. 82, с. 513—516.
6. **Феллер В.** Введение в теорию вероятностей и ее приложения, М., «Мир», 1967.
7. Feller W. On the Kolmogorov-Smirnov Limit Theorems for empirical distribution Annals of math. Statist, v. XIX, 2, 1948.
8. **Крамер Г.** Математические методы статистики. М., ИИЛ, 1948.
9. **Шор Я. Б., Кузьмин Ф. И.** Таблицы для анализа и контроля надежности. М., «Сов. радио», 1968.
10. **Ван-дер Варден Б. Л.** Математическая статистика. М., ИИЛ, 1960.
11. **Большев Л. Н., Смирнов Н. В.** Статистические таблицы. М., «Наука» 1968.
12. Таблицы нормального интеграла, нормальной плотности и ее нормированных производных. Под ред. Н. В. Смирнова. М., Изд-во АН СССР, 1960.
13. **Кендалл М. Дж, Стюарт А.** Статистические выводы и связи. М., «Наука», 1973.
14. **Смирнов Н. В.** О распределении ω^2-критерия Мизеса. Математический сборник. 1937, 2, вып. I, с. 973—993.
15. **Anderson T. W., Darliny D. A.** Asimptotic Theory „goodness of fif“ Criteria Based on Stochastic Processes Annals of Math. 1952, 23, c. 193—212.

Редактор *И. И. Топильская*
Технический редактор *В. Н. Солдатова*
Корректор *Е. И. Евтеева*

Сдано в наб. 30.12.74 Подп. в печ. 19.03.75 1,5 п. л. Тир. 16 000 Цена 8 коп.

Издательство стандартов. Москва, Д-22, Новопресненский пер., 3
Тип. «Московский печатник». Москва, Лялин пер., 6. Зак. 183