

10. Metody wykluczania dużych błędów.

Przy wielokrotnym mierzeniu tej samej wielkości zdarza się, że otrzymany wynik, który znacznie różni się od wszystkich otrzymanych wyników. Takie duże błędy mają małe prawdopodobieństwo pojawienia się, niemniej w wypadku otrzymania takiego wyniku należy zastosować odpowiednie testy, polegające na porównaniu tego wyniku z pozostałymi wynikami pomiaru. Stosowane są różne kryteria na odrzucanie pomiarów bardzo odległych od pomiarów pozostałych, w zależności od tego, czy znany względnie nieznany jest średni błąd kwadratowy pomiaru.

10.1. Metody wykluczania pomyłek przy znanym średnim błędzie kwadratowym.

Jeśli znany jest wcześniej średni błąd kwadratowy, to możemy uważać wynik x_p jakiegoś pomiaru jako pomyłkę tylko wtedy, jeżeli prawdopodobieństwo przypadkowego pojawienia się takiej wartości jest dostatecznie małe. Prawdopodobieństwo otrzymania wartości różniącej się od średniej arytmetycznej \bar{x} o więcej niż $3 \cdot 6$ równe jest 0,003 i wszystkie pomiary różniące się od \bar{x} o tę lub większą wielkość mogą zostać odrzucone, jako niezmiernie mało prawdopodobne. Jednakże w praktyce zazwyczaj liczba pomiarów jest mała i takim kryterium nie należy się posługiwać.

10.1.1. Wykluczanie pomyłek przez przyrównanie ich do znanej wartości współczynnika k .

W metodzie tej wykluczenie pomyłek odbywa się przy wykorzystaniu znanego współczynnika zależnego od ilości pomiarów, i tak:

- 1/ dla ilości pomiarów $n \leq 6$ nie można wykryć i odrzucić pomiaru x_p , podejrzanego o pomyłkę,
- 2/ dla ilości pomiarów $n > 6$ pomiar x_p możemy wykluczyć jeśli

$$\left| \frac{x_p - \bar{x}}{\sigma} \right| > k \quad 1.10$$

przy czym wartość średniej arytmetycznej \bar{x} i wartość średniego błędu kwadratowego σ oblicza się bez wartości x_p . Wartość współ-

t	β	t	β
2,5	0,01242	3,8	0,000145
2,6	0,00932	3,9	0,000096
2,7	0,00693	4,0	0,000063
2,8	0,00511	4,1	0,000041
2,9	0,00373	4,2	0,000027
3,0	0,00270	4,3	0,000017
3,1	0,00194	4,4	0,000011
3,2	0,00137	4,5	0,000008 0,0000068
3,3	0,00097	4,6	0,000004 0,0000041
3,4	0,00067	4,7	0,0000025 0,0000025
3,5	0,000465	4,8	0,0000016
3,6	0,000318	4,9	0,0000009
3,7	0,000216	5,0	0,0000006

Tabela 10.1. Dane do wyznaczania małych prawdopodobieństw /pomiarów istotności/.

Lp.	Poziomy małych prawdopodobieństw	Wykluczone są błędy, których prawdopodobieństwo powstania jest mniejsze niż
1.	5%	0,05
2.	1%	0,01
3.	0,1%	0,001

Tabela 10.2. Stosowane zazwyczaj pomiary małych prawdopodobieństw /poziomy istotności/.

czynnika k przyjmuje się w zależności od ilości przeprowadzonych pomiarów n :

$$\begin{aligned} k &= 4, \text{ jeśli } 6 < n \leq 100 \\ k &= 4,5 \text{ jeśli } 100 < n \leq 1000 \\ k &= 5 \text{ jeśli } 1000 < n \end{aligned}$$

10.1.2. Metoda wykluczania pomyłek przy przyjętym prawdopodobieństwie.

Jeśli średni błąd kwadratowy $\bar{\sigma}$ znany jest wcześniej, to możemy obliczyć wartość stosunku:

$$t = \frac{|x_p - \bar{x}|}{\bar{\sigma} \cdot \sqrt{\frac{n+1}{n}}} \quad 2.10$$

gdzie: x_p -- wynik pomiaru znacznie odbiegający od reszty pomiarów,
 \bar{x} -- średnia arytmetyczna z pozostałych pomiarów bez uwzględnienia wyniku pomiaru obciążonego dużym błędem x_p ,
 n -- ilość wszystkich pomiarów.

Dla otrzymanej wartości t odczytujemy odpowiednie prawdopodobieństwo β . Z prawdopodobieństwem

$$p = 1 - \beta \quad 3.10$$

można uważać, że wynik x_p jest obciążony dużym błędem i należy go wykluczyć z dalszego opracowania wyników pomiaru. Z wzoru 3.10 widzimy, że jeśli odczytane z tablicy 10.1 prawdopodobieństwo β okaże się bardzo małe, to odbiegający wynik x_p zawiera duży błąd. Jeśli przyjąć zbyt niski poziom małych prawdopodobieństw /poziom istotności/ to nawet duże błędy mogą pozostać niewykluczone, jeśli zaś przyjąć poziom ten nieuzasadnienie duży, to można wykluczyć i wyniki z błędami przypadkowymi, które są niezbędne do prawidłowego opracowania wyników pomiaru.

W praktyce stosuje się zazwyczaj jeden z trzech poziomów istotności /poziom małych prawdopodobieństw/, przedstawionych w tabelce nr 10.2.

10.2. Metoda wykluczania pomyłek przy pomocy błędu standardowego.

Jeśli wartość $\bar{\sigma}$ nie jest znana, to ocenia się ją w przybliżeniu na podstawie wyników pomiarów, przyjmując empiryczny

n	t_n			
	$\beta=1\%$	$\beta=2.5\%$	$\beta=5\%$	$\beta=10\%$
3	1.414	1.414	1.412	1.408
4	1.723	1.710	1.699	1.645
5	1.953	1.917	1.869	1.791
6	2.130	2.067	1.996	1.894
7	2.265	2.182	2.093	1.974
8	2.374	2.273	2.172	2.041
9	2.464	2.349	2.237	2.097
10	2.540	2.414	2.294	2.148
11	2.606	2.470	2.343	2.190
12	2.663	2.519	2.387	2.229
13	2.714	2.562	2.426	2.264
14	2.759	2.602	2.461	2.297
15	2.800	2.638	2.493	2.328
16	2.837	2.670	2.523	2.354
17	2.871	2.701	2.551	2.380
18	2.903	2.729	2.577	2.404
19	2.932	2.754	2.600	2.428
20	2.959	2.778	2.623	2.447
21	2.984	2.801	2.644	2.467
22	3.008	2.823	2.664	2.486
23	3.030	2.843	2.683	2.504
24	3.051	2.862	2.701	2.520
25	3.071	2.880	2.717	2.537

Tabela 10.3. Wartości współczynników do wykluczania pomyłek przy pomocy błędu standardowego.

Ilość pomiarów n	prawdopodobieństwo pojawienia się dużego błędu P			
	0,99	0,98	0,95	0,90
3	0,988	0,972	0,941	0,886
4	0,889	0,846	0,765	0,679
5	0,760	0,729	0,642	0,557
6	0,698	0,644	0,560	0,482
7	0,637	0,586	0,507	0,434
8	0,590	0,543	0,468	0,399
9	0,555	0,510	0,437	0,370
10	0,527	0,483	0,412	0,349

Tabela 10.4. Wartości krytyczne α do odrzucania wyników ocie-
głych.

błąd standardowy s_n , obliczany wg wzoru 1.9 oraz wartość średniej arytmetycznej, uwzględniając w obu wypadkach wszystkie wyniki wraz z wynikiem odstającym. Następnie obliczamy wyrażenie:

$$t = \frac{\bar{x} - x_1}{s} \quad \text{lub} \quad t = \frac{x_n - \bar{x}}{s} \quad 4.10$$

gdzie: \bar{x} - średnia arytmetyczna wszystkich wyników,

x_1 - wynik odstający, zdecydowanie mniejszy od wyników pozostałych,

x_n - wynik odstający, zdecydowanie większy od wyników pozostałych,

t_n - współczynnik wyznaczony z tablicy nr 10.3 dla odpowiedniej ilości pomiarów.

Z tablicy nr 10.3 odczytujemy dla ilości pomiarów n i wybranego poziomu istotności β liczbę t_n . Jeśli między t wyznaczone ze wzoru 4.10 a t_n dla danej ilości pomiarów n i wybranym poziomem istotności będzie $t > t_n$, to wymiar x_1 względnie x_n należy odrzucić, traktując go jako zwykłą pomyłkę.

10.3. Metoda wykluczania pomyłek przy nieznanym błędzie standardowym.

Wszystkie wyżej stosowane kryteria wykluczania dużych błędów wymagały znajomości wartości σ bądź też wartości średniego błędu standardowego s_n , co w praktyce wymaga wykonania wielu żmudnych obliczeń. Dlatego wygodnie jest stosować kryterium Q wyliczonego na podstawie wzoru:

$$Q_1 = \frac{x_2 - x_1}{R} \quad \text{lub} \quad Q_n = \frac{x_n - x_{n-1}}{R} \quad 5.10$$

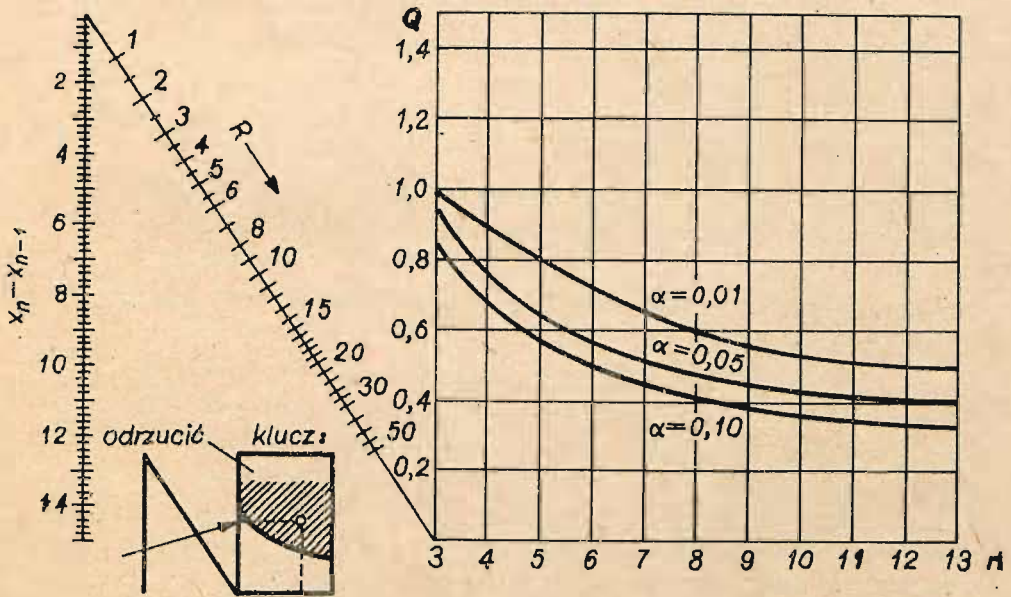
gdzie: $R = x_n - x_1$ jest rozpiętością pomiarów.

Wyniki pomiarów uszeregowane są w ten sposób, że:

$$x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n$$

Kryterium 5.10 stosujemy oczywiście zależnie od tego, czy podejrzewamy o zbyt wysoką wartość wynik ostatni, czy też o zbyt niską wartość wynik pierwszy. Obliczony wynik Q_1 lub Q_n porównujemy z wartością tablicową Q /tablica nr 10.4/ dla wybranego prawdopodobieństwa pojawienia się dużego błędu p i dla właściwej liczby pomiarów n . Jeśli

$$Q_1 > Q \quad \text{lub} \quad Q_n > Q$$



Rys.10.1. Wykres do odrzucania wyników odległych.

to wynik x_1 lub x_n odrzucamy z dalszych wyliczeń.

Test Q można używać przy odrzucaniu wyników zbyt odległych od pozostałych nawet w przypadku, kiedy liczba pomiarów jest nieduża, ale nie mniejsza niż 3. Wszystkie poprzednie kryteria oparte są na odchyleniu standardowym, obliczonym z sumy kwadratów, dlatego należy ich używać przy większej ilości pomiarów. Kryterium Q oparte jest na rozpiętości wyników i nadaje się dla małej ilości pomiarów $n \leq 10$.

Przykład. Podczas badania $n = 8$ próbek aluminiowych otrzymano następujące wyniki wytrzymałości na rozciąganie:

2675, 2707, 2707, 2709, 2718, 2720, 2723, 2742.

Wyniki $x_1 = 2675$ jak i wynik $x_n = 2742$ jako zbyt odstające sprawdzamy wg kryterium rozpiętości wyników.

Korzystając z wzoru 5.10 otrzymamy:

$$Q_1 = \frac{2707 - 2675}{2742 - 2675} = 0,478$$

W tabeli 10.4 dla $n = 8$ i dla prawdopodobieństwa pojawienia się dużego błędu 0,95 odczytujemy wartość $Q = 0,468$, a więc

$$Q_1 > Q$$

tzn. wynik 2675 odrzucamy z dalszych obliczeń. Dla wyniku $x_n = 2742$ otrzymamy:

$$Q_n = \frac{2742 - 2723}{2742 - 2675} = 0,284$$

W tym przypadku $Q_n < Q$, a to oznacza, że wynik $x_n = 2742$ bierzemy pod uwagę przy dalszych obliczeniach.

10.4. Metoda graficzna wykluczania pomyłek.

Do wykluczania pomyłek można zastosować także wykres, oparty na teście Q , przedstawiony na rys.10.1.

Szeregując wyniki pomiarów wg wartości niemalejącej $x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n$, bierzemy różnicę $x_2 - x_1$ lub $x_n - x_{n-1}$ i odkładamy

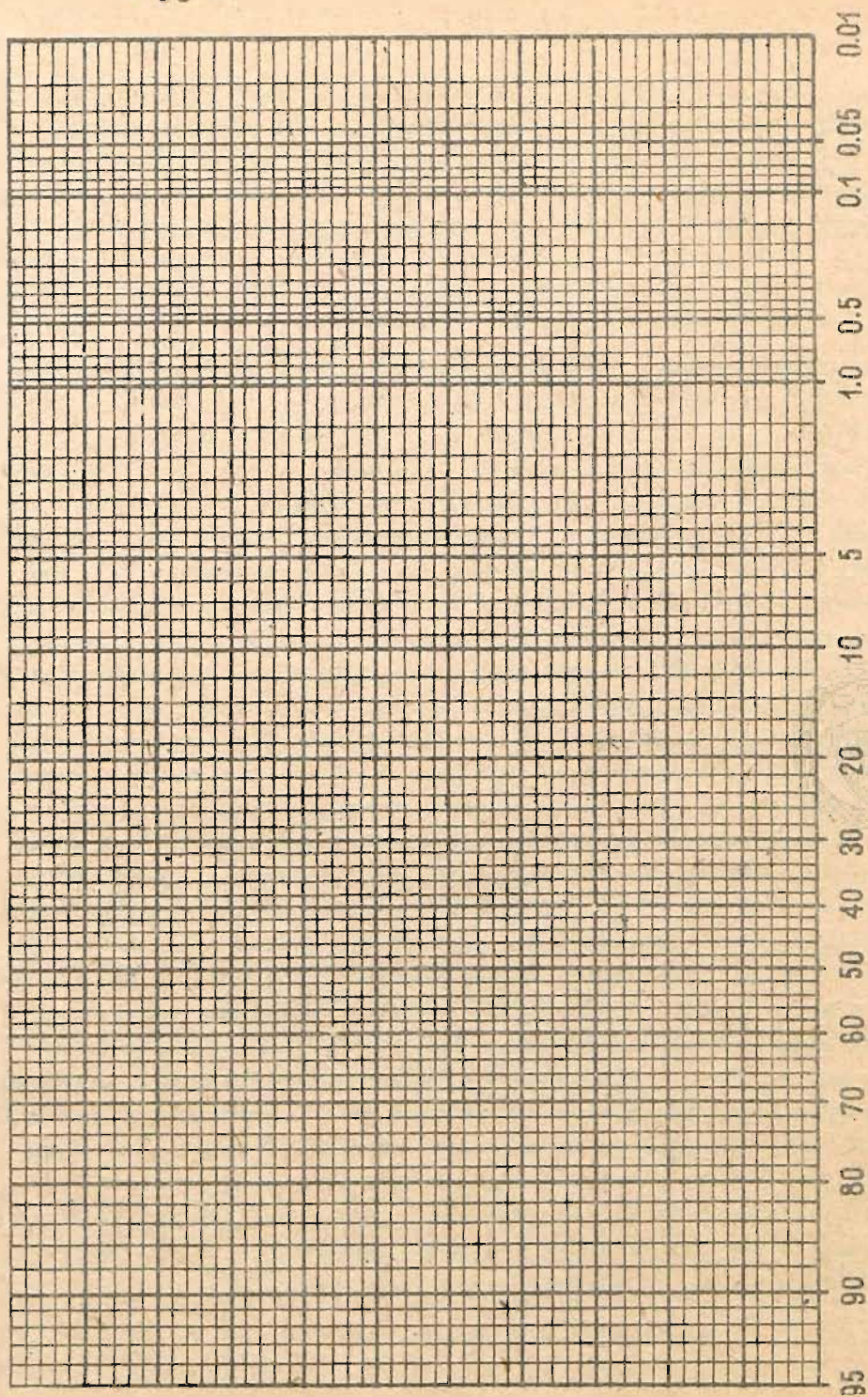
jej wartość na podziałce lewej. Następnie łączymy ją z wartością $R = x_n - x_1$ na podziałce ukośnej i przedłużamy odcinek łączący tak, aby zaznaczyć na podziałce prawej wartość Q . Tę porównujemy wówczas z wartością krytyczną dla danego n w prawej części wykresu. Jeśli wynik znajdzie się powyżej wybranego poziomu β , wówczas wynik x_1 , względnie x_n odrzucamy. Schemat postępowania

wykluczania dużych błędów jest pokazany w lewej dolnej części wykresu. Test ten jest dogodny ze względu na swoją prostotę. Wartości $x_2 - x_1$ oraz $x_n - x_{n-1}$ oraz R należy podawać w tych samych jednostkach i tak je dobrać, aby zmieścić się w podanym zakresie wartości różnic pomiarów i rozpiętości pomiarów R .

10.5. Sprawdzanie normalności rozkładu.

Prawo Gaussa rozkładu normalnego jest matematycznym modelem rozkładu wyników obciążonych tylko błędami przypadkowymi, przy założeniu, że każda pomierzona wartość pozostaje pod wpływem nieskończonej liczby małych błędów elementarnych, mogących się sumować lub wzajemnie kompensować. Wszystkie podane wartości średnich jak i wariancji są wyprowadzone przy założeniu, że błędy przypadkowe podlegają rozkładowi normalnemu i dlatego mogą być stosowane tylko wtedy, gdy wyniki doświadczenia nie przeczą temu założeniu.

Czasami jednak spotykamy się z różnymi odchyleniami od rozkładu normalnego, nawet w obszarze bliskim wartości średniej atymetycznej. Najważniejsze są w praktyce zakłócenia symetrii rozkładu, gdzie czasami można otrzymać rozkłady wielowierzchołkowe. Z rozkładem np. niesymetrycznym spotykamy się głównie wtedy, gdy z jakichś powodów pojawienie się błędu o danym znaku jest nieco prawdopodobniejsze, niż pojawienie się błędu o znaku przeciwnym. Małych odchyłeń od rozkładu Gaussa można oczekiwać od większości rzeczywistych rozkładów. Nie musi to być jednak powodem do obaw, że stosując wzory oparte na tym rozkładzie, dopuścimy się poważniejszych błędów. Jeśli jednak stwierdzimy większe odchylenie, zwłaszcza symetrię rozkładu, to możemy znaleźć taki model matematyczny, który lepiej odzwierciedla rzeczywisty rozkład. Możemy jeszcze użyć tzw. nieparametrycznych metod, które nie wychodzą z założenia jakiegokolwiek matematycznie definiowanego rozkładu. Dla sprawdzenia prawa rozkładu normalnego należy przeprowadzić dostatecznie dużą liczbę pomiarów i zastosować jeden z podanych niżej testów dla określenia stopnia zgodności rozkładu wartości empirycznych x_i z rozkładem normalnym.



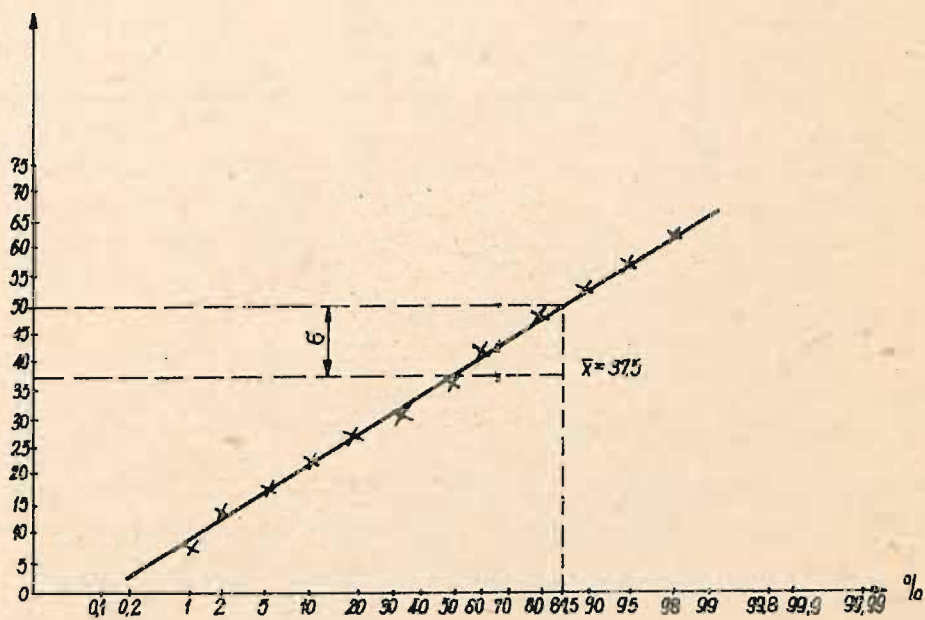
Rys. 10.2. Papier do sporządzania wykresów dla rozkładu normalnego.

Lp.	Prze- dział x_i	Ilość pomiarów w prze- dziale n_i	Ilość pomiarów zsuno- wana łącznie z danym przedziałem	Procent pomiarów zsuno- wany z poprzednimi pro- centami .
1.	2.	3.	4.	5.
1.	Δx_1	n_1	n_1	$\frac{n_1}{n} \cdot 100$
2.	Δx_2	n_2	$n_1 + n_2$	$\frac{n_1 + n_2}{n} \cdot 100$
3.	Δx_3	n_3	$n_1 + n_2 + n_3$	$\frac{n_1 + n_2 + n_3}{n} \cdot 100$
<hr/>				
k.	Δx_k	n_k	$\sum_{i=1}^k n_i = n$	$\frac{1}{n} \sum_{i=1}^k n_i = 100$

Tabela 10.5. Wzór tabeli do wpisywania danych pomiarowych.

Lp.	Przedział	Częstość	Częstość skumulowana	Procent skumulo- wany
1	2	3	4	5
1.	0 - 5	0	0	0
2.	5 - 10	2	2	1
3.	10 - 15	2	4	2
4.	15 - 20	6	10	5
5.	20 - 25	10	20	10
6.	25 - 30	20	40	20
7.	30 - 35	28	68	34
8.	35 - 40	32	100	50
9.	40 - 45	32	132	66
10.	45 - 50	28	160	80
11.	50 - 55	18	178	89
12.	55 - 60	12	190	95
13.	60 - 65	6	196	98
14.	65 - 70	4	200	100
15.	70 - 75	0	200	100

Tabela 10.6. Przykład tabeli częstości występowania i częstości skumulowanych pomiarów.



Rys.10.3. Wykres z naniesionymi punktami na papierze dla rozkładu normalnego /dane z tabeli 10.6/.

10.5.1. Sprawdzanie normalności rozkładu metodą graficzną

Dane pomiarowe x_i umieszczamy w odpowiednio dobranych przedziałach o długości Δx , przy czym ilość przedziałów dobieramy dowolnie, ale tak aby nie była mniejsza od 5-ciu. Wszystkie otrzymane w doświadczeniu punkty pomiarowe umieszczamy w odpowiednich przedziałach Δx_i . Tak więc w przedziale Δx_1 będzie n_1 wyników, w Δx_2 - n_2 wyników, w Δx_n - n_n wyników. oczywiście zachodzi równość $n_1 + n_2 + \dots + n_n = n$. Dane te umieszczamy w tabeli nr 10.5. Wartości podane w kolumnie 5 procentu skumulowanego należy nanieść na papier przeznaczony do sporządzania wykresów dla rozkładu normalnego, gdzie na osi odciętych podaje się wartość procentu skumulowanego, a na osi rzędnych punkty środkowe przedziałów Δx . Jeśli dane mają rozkład normalny to naniesione punkty będą układały się wzdłuż linii prostej. Jako największe odchylenie D punktów doświadczalnych od wykreślonej prostej przyjmuje się maksymalne odchylenie tych punktów liczone według osi rzędnych /pionowych/ siatki funkcyjnej. Sprawdzenie zgodności polega na wyznaczeniu nierówności

$$D \cdot \sqrt{n} \leq 1,0 \quad 6.10$$

gdzie: D - największe odchylenie naniesionych punktów od wykreślonej prostej, liczone według osi rzędnych /pionowych/
 n - liczba doświadczalnie otrzymanych punktów.

W przypadku, gdy spełniony jest związek 6.10 to nie ma podstaw do kwestionowania zgodności danego rozkładu empirycznego z założonym rozkładem teoretycznym. W przeciwnym wypadku hipotezę o zgodności wymienionego rozkładu należy odrzucić i z kolei przeprowadzić weryfikację dla innego rozkładu teoretycznego.

Warto zaznaczyć, że metoda ta pozwala na graficzne wyznaczenie wartości średniej arytmetycznej pomiarów oraz średniego błędu kwadratowego. Sposób wyznaczenia tych wielkości został podany w przytoczonym niżej przykładzie. Papier do sporządzenia wykresów dla rozkładu normalnego pokazano na rys.10.2.

Przykład. Pomiary w ilości $n = 200$ zostały umieszczone według tabeli nr 10.5 /tabela nr 10.6/. Sprawdzić czy otrzymane punkty mają rozkład normalny.

Przez naniesione punkty umieszczone na papier do rozkładu normalnego wykreślamy linię prostą /rys.10.3/. Na osi rzędnych odczytu-

jemy wartość średnią, która jest wyznaczona przez wykreśloną prostą dla wartości częstości skumulowanej równej 50%. Odchylenie standardowe wyznaczamy w ten sposób, że na osi rzędnych przez wykreśloną prostą dla częstości skumulowanej równej 84,5% odejmujemy wartość średnią.

W wypadku, gdyby punkty nie układały się wzdłuż prostej rozkład, rozkład nie może być rozkładem normalnym.

10.5.2. Przybliżona metoda sprawdzania rozkładu normalnego.

Dla przybliżonego sprawdzenia normalności rozkładu stosuje się metodę związaną z oceną centralnych momentów trzeciego i czwartego rzędu, porównanych do średnich błędów kwadratowych. W przypadku rozkładu normalnego powinny zachodzić następujące nierówności:

$$\left| \sqrt{\frac{6/n-1/}{/n+1//n+3/}} \right| > \left| w \frac{M_3}{s_n^3} \right| \quad 6.10$$

oraz

$$\left| \sqrt{\frac{24n /n-2//n-3/}{/n-1/2 /n+3//n+5/}} \right| > \left| w \left(\frac{M_4}{s_n^4} - 3 \right) \right| \quad 7.10$$

gdzie: w - współczynnik równy $w = 2 \div 3$,

s_n - błąd standardowy,

n - liczba pomiarów,

$$M_3 = \sum_{i=1}^k m_i (x_i - \bar{x})^3$$

$$M_4 = \sum_{i=1}^k m_i (x_i - \bar{x})^4$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i x_i$$

przy założeniu, że wynik x_1 występuje m_1 razy, x_2 występuje m_2 razy, a wynik x_k występuje m_k razy, przy czym $m_1 + m_2 + \dots + m_k = n$.

Jeśli choć jedna z nierówności 6.10 lub 7.10 co do bezwzględnej wartości nie jest spełniona, to normalność rozkładu należy uznać za wątpliwą i trzeba wtedy przeprowadzić bardziej wnikliwą analizę wyników doświadczenia np. zastosować test zgod-

t	Setne części t									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0040	0080	0120	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0,7	2580	2611	2642	2673	2703	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3437	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2,0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2,2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936

Tabela 10.7. Wartości funkcji $\Phi(t)$.

$k \backslash P$	0,80	0,90	0,95	0,98	0,99	0,995	0,998	0,999
4	5,99	7,78	9,49	11,67	13,28	14,9	16,9	18,5
5	7,29	9,24	11,07	13,39	15,09	16,3	18,9	20,5
6	8,56	10,64	12,59	15,33	16,8	18,6	20,7	22,5
7	9,80	12,02	14,07	16,6	18,5	20,3	22,6	24,3
8	11,03	13,36	15,51	18,2	20,1	21,9	24,3	26,1
9	12,24	14,68	16,9	19,7	21,7	23,6	26,1	27,9
10	13,44	15,99	18,3	21,2	23,2	25,2	27,7	29,6
11	14,63	17,3	19,7	22,6	24,7	26,8	29,4	31,3
12	15,8	18,5	21,0	24,1	26,2	28,3	31,0	32,9
13	17,0	19,8	22,4	25,5	27,7	29,8	32,5	34,5
14	18,2	21,1	23,7	26,9	29,1	31,3	34,0	36,1
15	19,3	22,3	25,0	28,3	30,6	32,7	35,6	37,7
16	20,5	23,5	26,3	29,6	32,0	34,2	37,1	39,3
17	21,6	24,8	27,6	31,0	33,4	35,7	38,6	40,8
18	22,8	26,0	28,9	32,3	34,8	37,2	40,1	42,3
19	23,9	27,2	30,1	33,7	36,2	38,6	41,6	43,8
20	25,0	28,4	31,4	35,0	37,6	40,0	43,1	45,3
22	27,3	30,8	33,9	37,7	40,3	42,7	45,9	48,3
24	29,6	33,2	36,4	40,3	43,0	45,5	48,7	51,2
26	31,8	35,6	38,9	42,9	45,6	48,2	51,5	54,1
28	34,0	37,9	41,3	45,4	48,3	51,0	54,3	56,9
30	36,3	40,3	43,8	48,0	50,9	53,7	57,1	59,7

Tabela 10.8. Krytyczne wartości χ^2 przy prawdopodobieństwie ufności p i liczbie stopni swobody k .

ności χ^2 /chi kwadrat/.

10.5.3. Test zgodności χ^2 /chi kwadrat/.

Rozkład ten pod nazwą χ^2 /chi kwadrat/ służy do sprawdzania zgodności błędów losowych wyników pomiarowych z prawem rozkładu normalnego.

Wyniki pomiarów grupuje się wg przedziałów w taki sposób, aby pokrywały się one z całą osią $[-\infty, +\infty]$ i aby liczba danych w każdym przedziale była dostatecznie duża /nie mniejsza od pięciu, a najlepiej gdy będzie większa od dziesięciu/. Dla każdego przedziału $[x_{i-1}, x_i]$ oblicza się liczby m_i wyników pomiarów, które znalazły się w tym przedziale. Dla tak ułożonych wyników oblicza się sumę

$$\chi^2 = \sum_{i=1}^l \frac{(m_i - np_i)^2}{n \cdot p_i} \quad 8.10$$

gdzie: l - liczba wszystkich przedziałów $[-\infty, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$... $[x_{l-1}, +\infty]$,

n - liczba wszystkich wyników pomiarów $n = m_1 + m_2 + \dots + m_l$.

Z prawa rozkładu normalnego obliczamy dla każdego i -tego przedziału prawdopodobieństwo p_i , że dany wynik pochodzi z przedziału o wskaźniku i

$$p_i = \Phi\left(\frac{x_i - \bar{x}}{s}\right) - \Phi\left(\frac{x_{i-1} - \bar{x}}{s}\right)$$

gdzie: \bar{x} - średnia arytmetyczna wyników pomiaru,

s_n - empiryczny błąd standardowy.

Jako wartości x_i przyjęte są prawe końce przedziału. Przyjmuje się także, że

$$\Phi(-t) = -\Phi(t)$$

$$\Phi(+\infty) = 0,5$$

$$\Phi(-\infty) = -0,5$$

wartości funkcji $\Phi(t) = \Phi\left(\frac{x_i - \bar{x}}{s}\right)$ podane są w tabelicy nr 10.7.

Jeśli wartość sumy 8.10 jest większa od wartości krytycznej χ^2 odczytanej z tabelicy nr 10.8 przy z góry założonym prawdopodobieństwie ufności P i tzw. liczbie stopni swobody k ,

to z prawdopodobieństwem ufności P można uważać, że rozkład prawdopodobieństwa błędów przypadkowych w rozpatrywanej serii pomiarów różni się od rozkładu normalnego. W przeciwnym przypadku przyjmujemy hipotezę o normalnym rozkładzie błędów przypadkowych pomiaru. Liczbę stopni swobody k dla liczby przedziałów l przyjmuje się:

$k = l - 3$ dla przypadku pomiarów, gdy zamiast wartości rzeczywistej a i wartości σ przyjmuje się doświadczalnie wartość średniej arytmetycznej \bar{x} i empiryczny błąd standardowy s ,

$k = l - 2$ dla przypadku, gdy wartość rzeczywista a jest dokładnie znana /np. przy pomiarach wzorca/,

$k = l - 1$ dla przypadku, gdy znana jest wartość rzeczywista a oraz średni błąd kwadratowy σ .

W praktyce spotyka się raczej sytuacje, że zarówno wartość średniej arytmetycznej \bar{x} jak i empiryczny błąd standardowy s_n oblicza się na podstawie wyników pomiaru, stąd liczba stopni swobody $k = l - 3$. Dlatego też, aby otrzymać liczbę swobody nie mniejszą od pięciu należy brać liczbę przedziałów l nie mniejszą niż osiem.

11. Ilość niezbędnych pomiarów.

W celu zmniejszenia błędu losowego pomiarów można stosować dwa sposoby:

a/ zwiększenie dokładności poszczególnego pomiaru tzn. zmniejszenie wielkości σ ,

b/ zwiększenie liczby pomiarów tzn. skorzystanie z wzoru

$$\sigma_x = \frac{\sigma}{\sqrt{n}}$$

przy założeniu dodatkowym, że wszystkie możliwości udoskonalenia samej techniki pomiarów zostały już wykorzystane w celu poprawienia dokładności pomiaru.

Jeśli błąd systematyczny, określony w jakikolwiek sposób ma wartość Δa , to celowe jest zmniejszenie błędu losowego tylko do takich rozmiarów, aby ogólny błąd pomiarów określony był całkowicie błędem systematycznym. Niezbędne jest więc, aby przedział ufności zbudowany dla ustalonego współczynnika ufności był zdecydowanie mniejszy od wielkości błędu systematycznego tzn. aby