

## 6. TESTY ZGODNOŚCI

Wynikiem działania testów statystycznych w STATISTICE są graniczne poziomy istotności  $p$ -value. Decyzję o **odrzućeniu** hipotezy  $H_0$  można podjąć, gdy:

założony poziom istotności  $\alpha$  **jest większy od** poziomu granicznego  $p$ -value.

O **braku podstaw** do odrzucenia hipotezy  $H_0$  świadczy:

poziom istotności  $\alpha$  **mniejszy od** granicznego poziomu istotności  $p$ -value.

W przypadku kilku testów nie ma możliwości określenia wartości poziomu  $\alpha$  (domyślnie przyjmowany jest poziom istotności  $\alpha = 0,05$ ).

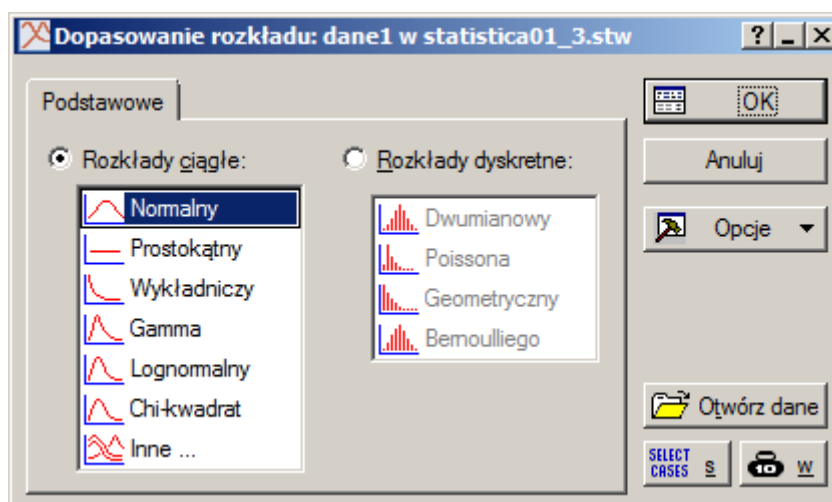
*Dla ułatwienia, wyniki testów, które dla ustalonego poziomu istotności  $\alpha$  wymagają odrzucenia hipotezy  $H_0$ , zaznaczane są na czerwono.*

W testach zgodności weryfikowane są hipotezy dotyczące zgodności próby z pewnym rozkładem teoretycznym. W STATISTICE dostępne są najczęściej stosowane testy zgodności: test  $\chi^2$  i test Kołmogorowa–Smirnowa.

### 6.1. Test zgodności $\chi^2$

Test  $\chi^2$  można przeprowadzić z poziomu okna **Dopasowanie rozkładu** (dostępne z menu głównego: **Statystyka/ Dopasowanie rozkładu**).

W pierwszym kroku należy wybrać rozkład teoretyczny do którego dane z próby będą porównywane.



Następnie należy określić zmienną, która zawiera analizowane dane (w oknie na kolejnym rysunku do testu została wybrana zmienna **długość**).

**Dopasowanie rozkładu ciągłego: dane1 w statistica01\_3.stw**

Rozkład: **Normalny**

**Zmienna** **długość**

Podstawowe Parametry Opcje

Liczba kategorii: **8** Ustaw domyślne

Dolna granica: **19** Klinij aby przywrócić domyślną liczbę kategorii, dolną i górną granicę i parametry rozkładu.

Górna granica: **23**

Średnia (M): **20,962157**

Wariancja: **,4815886**

Średnia obserwowana: 20,962157  
Wariancja obserwowana: ,4815889

Podsum Anuluj Opcje W Grupami

**Dopasowanie rozkładu ciągłego: dane1 w statistica01\_3.stw**

Rozkład: **Normalny**

**Zmienna** **długość**

Podstawowe Parametry Opcje

Test Kołmogorowa-Smimowa

Nie

Tak (skategoryzowany)

Tak (ciągły)

Test chi-kwadrat

Połączone kategorie

Jeżeli oczekiwana częstość w przedziale jest mniejsza lub równa 5, to przedziały będą łączone

Wykres

Wykres rozkładu

Rozkład licznosci

Dystrybujanta

Wykres licznosci lub %

Licznosci

Częstości (%)

Podsum Anuluj Opcje W Grupami

Bazując na parametrach domyślnych można już na tym etapie przeprowadzić test zgodności (po naciśnięciu np. przycisku **Podsum**). Parametry domyślne można zmienić na zakładkach: **Parametry** i **Opcje**. Na powyższym rysunku zmienione zostały: **Liczba kategorii**, **Dolna granica** i **Górna granica**. W przypadku testowania zgodności z rozkładem normalnym można również określić parametry rozkładu: **Średnią** i **Wariancję**. Dla ułatwienia parametry te są wstępnie ustawiane na podstawie wartości obliczonych z próby (**Średnia obserwowana** i **Wariancja obserwowana**). Na zakładce **Opcje** można również wymusić jednoczesne wykonanie testu Kołmogorowa–Smirnowa, można też zrezygnować z domyślnego łączenia kategorii jeśli liczebność w przedziałach klasowych jest mniejsza lub równa 5.

Górna Granica	Obserw. Liczność	Skum. Obserw.	Procent Obserw.	Skum.% Obserw.	Oczekiw. Liczność	Skum. Oczekiw.	Procent Oczekiw.	Skumul.% Oczekiw.	Obserw. - Oczekiw.
<= 19,5	1	1	1,00000	1,0000	1,75606	1,7561	1,75606	1,7561	-0,75606
20,0	6	7	6,00000	7,0000	6,52424	8,2803	6,52424	8,2803	-0,52424
20,5	18	25	18,00000	25,0000	16,99137	25,2717	16,99137	25,2717	1,00863
21,0	29	54	29,00000	54,0000	26,90274	52,1744	26,90274	52,1744	2,09726
21,5	26	80	26,00000	80,0000	25,90943	78,0838	25,90943	78,0838	0,09057
22,0	12	92	12,00000	92,0000	15,17728	93,2611	15,17728	93,2611	-3,17728
22,5	6	98	6,00000	98,0000	5,40438	98,6655	5,40438	98,6655	0,59562
<niesk.	2	100	2,00000	100,0000	1,33450	100,0000	1,33450	100,0000	0,66550

Niezależnie od ustawienia opcji **Połączone kategorie** (na zakładce **Opcje**) wynikowa tabela testu zawiera wszystkie kategorie wynikające z ustawień parametrów z zakładki **Parametry** – łączenie wpływa na dopiero na wynikową wartość statystyki testowej, ilość stopni swobody rozkładu  $\chi^2$  oraz obliczaną wartość granicznego poziomu istotności, wyniki obliczeń wyświetlane są w tytule wynikowej tabeli testu w polach: **Chi-kwadrat**, **df** oraz **p**. Domyślny poziom istotności  $\alpha = 0,05$  jest **mniejszy od** otrzymanego granicznego poziomu istotności  $p\text{-value} = 0,72373$ , więc w rozważanym przypadku nie ma podstaw do odrzucenia hipotezy o zgodności rozkładu z próby z rozkładem normalnym.

## 6.2. Test zgodności $\lambda$ Kołmogorowa (Kołmogorowa – Smirnowa)

Test Kołmogorowa – Smirnowa można przeprowadzić z poziomu okna:

- **Dopasowanie rozkładu** (dostępne z menu: **Statystyka/Dopasowanie rozkładu**),
- **Dopasowanie rozkładów** (dostępne z menu: **Statystyka/Rozkłady i symulacja/Dopasuj rozkład**)
- **Statystyki opisowe** (dostępne z menu: **Statystyka/Statystyki podstawowe**),
- **Analiza zdolności procesu** (dostępne z menu: **Statystyka/Analiza procesu**).

### 6.2.1. Okno Dopasowanie rozkładu

Okno **Dopasowanie rozkładu** zostało omówione w punkcie poprzednim. Test Kołmogorowa–Smirnowa przeprowadzany jest w przypadku ustawienia na zakładce **Opcje** w grupie opcji **Test Kołmogorowa–Smirnowa** opcji: **Tak (skategoryzowany)** lub **Tak (ciągły)**.

Opcje **Tak (skategoryzowany)** lub **Tak (ciągły)** wpływają na sposób wyznaczania statystyki  $D$  liczonej jako maksymalna różnica pomiędzy dystrybucjami empiryczną i teoretyczną. Dla obliczeń skategoryzowanych obliczenia wykonywane są dla danych pogrupowanych, dla obliczeń ciągłych dane są sortowane a obliczenia są przeprowadzane dla każdej z wartości próbki. Okno wynikowe testu w każdym z przypadków zawiera te same obliczenia niezbędne dla przeprowadzenia testu  $\chi^2$ . Wyniki testu Kołmogorowa-Smirnowa wyświetlane są wyłącznie w nagłówku tabeli. Wyniki dla testu z obliczeniami



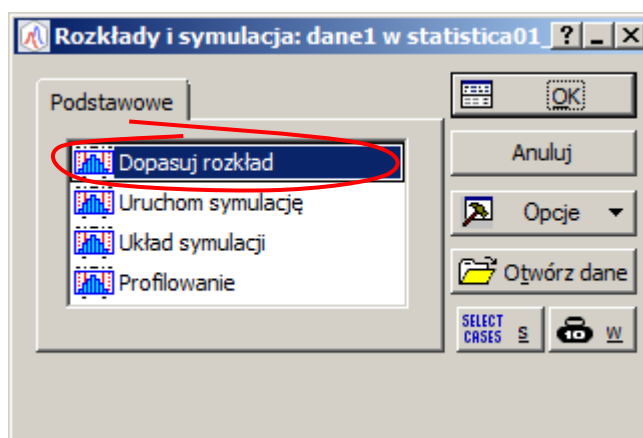
skategoryzowanymi i ciągłymi przedstawiono na kolejnych rysunkach. W pierwszym przypadku otrzymano wartość statystyki  $D = 0,01916$  w drugim  $D = 0,03410$ . Graniczny poziom istotności  $p$ -value podany został tylko w drugim przypadku jako  $p = \mathbf{n.i.}$  (w zasadzie zostały podane dwie wartości:  $p = \mathbf{n.i.}$  i  $p$  Lillieforsa =  $\mathbf{n.i.}$ , w przypadku testu normalności dokładniejszą wartością  $p$ -value jest wartość obliczana z testu Kołmogorowa–Smirnowa uwzględniająca poprawkę Lillieforsa, w rozważanym przykładzie obydwie wartości są  $\mathbf{n.i.}$  czyli nieistotne, tzn. dużo większe od poziomu istotności  $\alpha$ ). Brak wyróżnienia (na czerwono) i nieistotna wartość granicznego poziomu istotności oznaczają, że nie ma podstaw do odrzucenia hipotezy o zgodności rozkładu z próby z rozkładem normalnym.

The first screenshot shows the results of a Kolmogorov-Smirnov test for the variable 'długość'. The test statistic is  $D = 0,01916$  and the p-value is  $p = 0,72373$ . The second screenshot shows the same test results but with a different test statistic  $D = 0,03410$  and p-values for both Kolmogorov-Smirnov and Lilliefors tests marked as 'n.i.' (not significant).

Górna Granica	Obserw. Liczność	Skum. Obserw.	Procent Obserw.	Skum.% Obserw.	Oczekiw. Liczność	Skum. Oczekiw.	Procent Oczekiw.	Skum.% Oczekiw.	Obserw. - Oczekiw.
$\leq 19,5$	1	1	1,00000	1,0000	1,75606	1,7561	1,75606	1,7561	-0,75606
20,0	6	7	6,00000	7,0000	6,52424	8,2803	6,52424	8,2803	-0,52424
20,5	18	25	18,00000	25,0000	16,99137	25,2717	16,99137	25,2717	1,00863
21,0	29	54	29,00000	54,0000	26,90274	52,1744	26,90274	52,1744	2,09726
21,5	26	80	26,00000	80,0000	25,90943	78,0838	25,90943	78,0838	0,09057
22,0	12	92	12,00000	92,0000	15,17728	93,2611	15,17728	93,2611	-3,17728
22,5	6	98	6,00000	98,0000	5,40438	98,6655	5,40438	98,6655	0,59562
<niesk.	2	100	2,00000	100,0000	1,33450	100,0000	1,33450	100,0000	0,66550

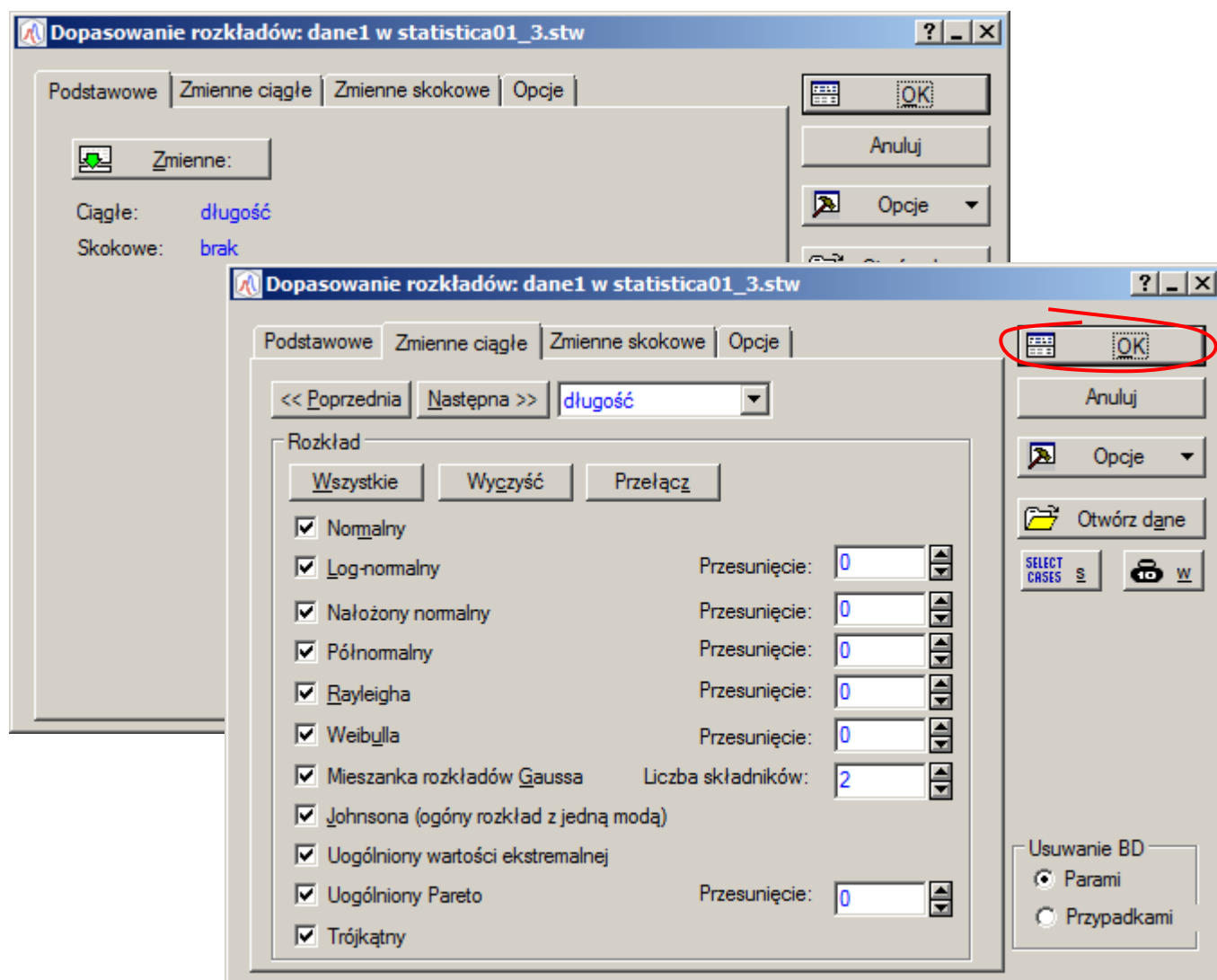
### 6.2.2. Okno Dopasowanie rozkładów

Zgodność danych z podstawowymi rozkładami teoretycznymi (ciągłymi i dyskretnymi) może zostać zbadana po wskazaniu opcji: **Statystyka/Rozkłady i symulacja/Dopasuj rozkład**.



Dopasowanie rozkładów można przeprowadzić po wskazaniu zmiennych na zakładce **Podstawowe**. Zakładka **Zmienne ciągle** pozwala na ustalenie parametrów ciągłych rozkładów teoretycznych, zakładka **Zmienne skokowe** rozkładów dyskretnych.

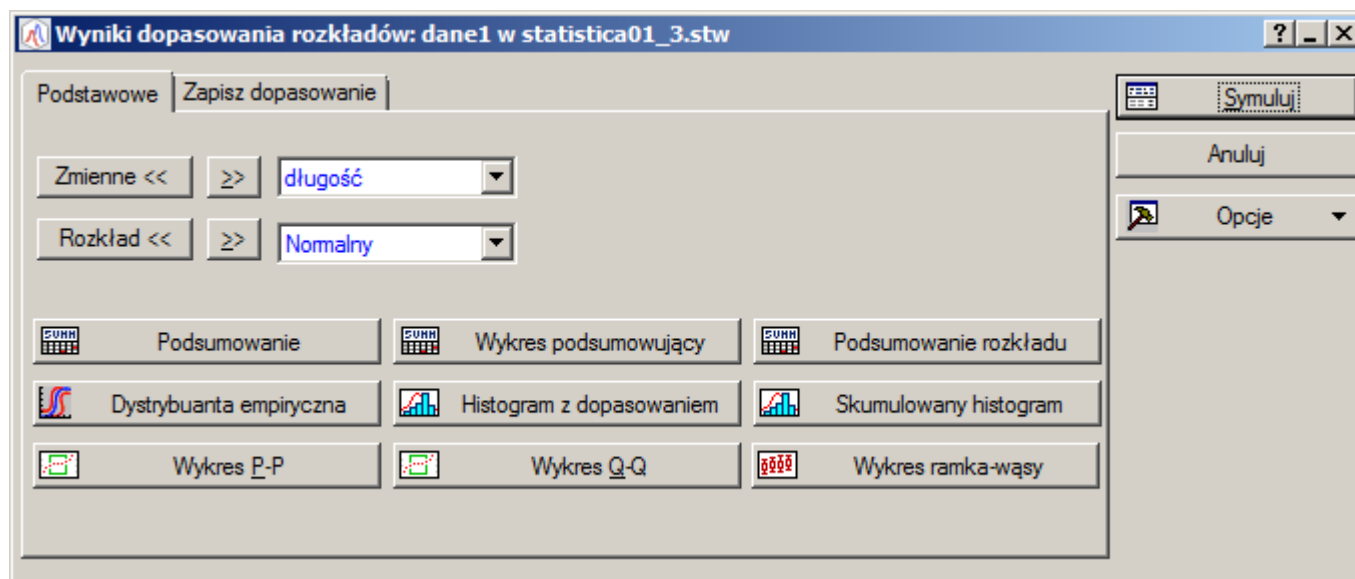




Wyniki dopasowania rozkładu danych do wskazanych rozkładów teoretycznych są wyświetlane po zaakceptowaniu okna **Dopasowania rozkładów**. Okno wyników na zakładce **Podstawowe** pozwala na analizę dopasowania danych do wybranego rozkładu teoretycznego.

Użytkownik może zobaczyć:

- wartości podstawowych statystyk wyznaczonych z próby (przycisk **Podsumowanie**),





- wyniki trzech testów zgodności (przycisk **Podsumowanie rozkładu**)  
 Kołmogorowa – Smirnowa (wartość statystyki  $D$  oraz  $p$ -value – kolumny: d K-S i K-S p),  
 Andersona–Darlinga (wartość statystyki AD oraz  $p$ -value – kolumny Stat. AD i p AD),  
 $\chi^2$  (wartość statystyki  $\chi^2$ ,  $p$ -value oraz ilość stopni swobody – kolumny: Chi<sup>2</sup>, Chi<sup>2</sup> p, Chi<sup>2</sup> df)  
 oraz dopasowane parametry rozkładu teoretycznego (na poniższym rysunku Param1 i Param2),

	d K-S	K-S p	Stat. AD	p AD	Chi <sup>2</sup>	Chi <sup>2</sup> p	Chi <sup>2</sup> df	P r z	Param1	Param2
Normalny (location, skala)	0,034096	0,999588	0,131388	0,999263	1,322798	0,723727	3,000000		20,96216	0,693966

Okno udostępnia również przyciski umożliwiające szybkie przygotowanie wykresów porównujących:

- dystrybuantę empiryczną i teoretyczną (przycisk **Dystrybuanta empiryczna**),
- histogram liczebności z nałożoną funkcją gęstości rozkładu teoretycznego (przycisk **Histogram z dopasowaniem**),
- histogram liczebności skumulowanych z nałożoną dystrybuantą rozkładu teoretycznego (przycisk **Skumulowany histogram**),
- wykres prawdopodobieństwo–prawdopodobieństwo (przycisk **Wykres P–P**),
- wykres kwantyl–kwantyl (przycisk **Wykres Q–Q**),
- wykres pudełkowy (przycisk **Wykres ramka–wąsy**),
- zestawienie wybranych wyników i wykresów (przycisk **Wykres podsumowujący**).

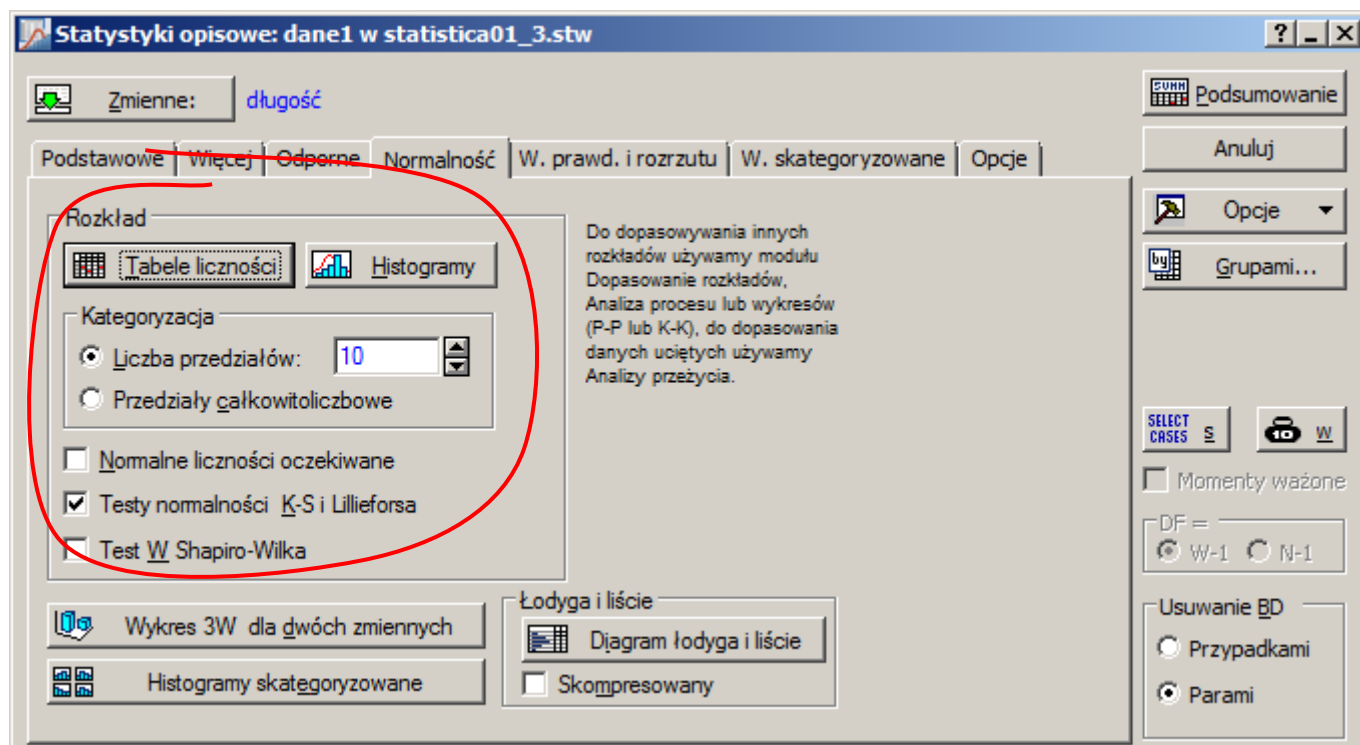
Zakładka **Zapisz dopasowanie** pozwala na analizę wyników dopasowania uzyskanych dla wszystkich rozważanych rozkładów teoretycznych.

Rozkład	d K-S	K-S p	Stat. AD	p AD	Chi-kwadrat	Chi
Johnson SU (typ, Gamm...	0,032319	0,999849	0,086619	0,999982	0,506198	0,4
Uogólniony wartości ekst...	0,034274	0,999548	0,147229	0,998703	0,715509	0,6
Log-normalny (skala, kształt...)	0,028586	0,999990	0,102308	0,999938	0,908069	0,8
Normalny (location, skala)	0,034096	0,999588	0,131388	0,999263	1,322798	0,7
Trójkątny (min, max, moda)	0,081717	0,491061	1,475839	0,182364	3,878467	0,1
Rayleigha (skala)	0,572317	0,000000	40,219927	0,000000	879,590183	0,0



### 6.2.3. Okno Statystyki opisowe

Z poziomu okna **Statystyki opisowe** można przeprowadzić test zgodności z rozkładem normalnym.



Wyniki testu Kołmogorowa-Smirnowa wyświetlane są po naciśnięciu przycisku **Tabele liczości** jeśli zaznaczona została opcja **Testy K-S i Lillieforsa** (wcześniej należy wybrać zmienną i ewentualnie zmienić domyślną liczbę przedziałów kategoryzacji danych).

statistica01\_3.stw - Tabela liczości: długość (dane1 w statistica01\_3.stw)

Tabela liczości: długość (dane1 w statistica01\_3.stw)  
K-S d=.03410, p> .20; Lilliefors p> .20

Klasa	Liczba	Skumulow. Liczba	Procent Ważnych	Skumul. % Ważnych	% ogółu Przypadki	Skumulow. % Ogółu
18,5<x<=19,0	0	0	0,00000	0,0000	0,00000	0,0000
19,0<x<=19,5	1	1	1,00000	1,0000	1,00000	1,0000
19,5<x<=20,0	6	7	6,00000	7,0000	6,00000	7,0000
20,0<x<=20,5	18	25	18,00000	25,0000	18,00000	25,0000
20,5<x<=21,0	29	54	29,00000	54,0000	29,00000	54,0000
21,0<x<=21,5	26	80	26,00000	80,0000	26,00000	80,0000
21,5<x<=22,0	12	92	12,00000	92,0000	12,00000	92,0000
22,0<x<=22,5	6	98	6,00000	98,0000	6,00000	98,0000
22,5<x<=23,0	2	100	2,00000	100,0000	2,00000	100,0000
Braki	0	100	0,00000		0,00000	100,0000

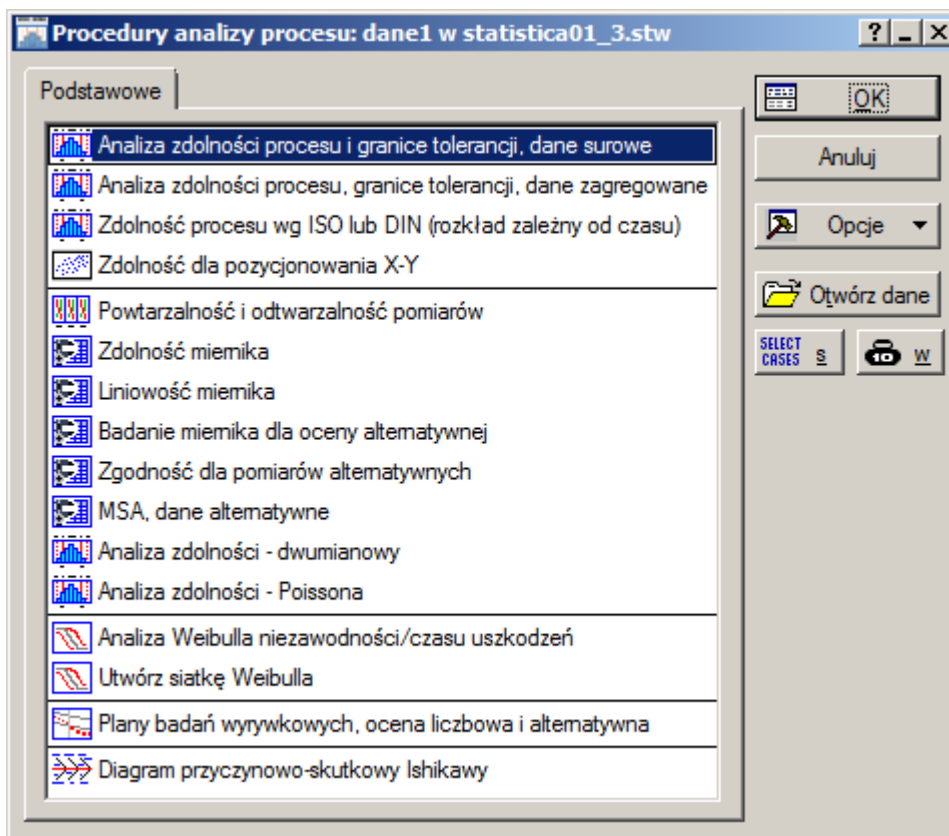
Tabela liczości: długość (dane1 w statistica01\_3.stw)

Podobnie jak poprzednio, wyniki wyświetlane są w nagłówku tabeli, tym razem wartość granicznego poziomu istotności określana jest jako:  $p > .20$ , więc nie ma podstaw do odrzucenia hipotezy o normalności rozkładu.

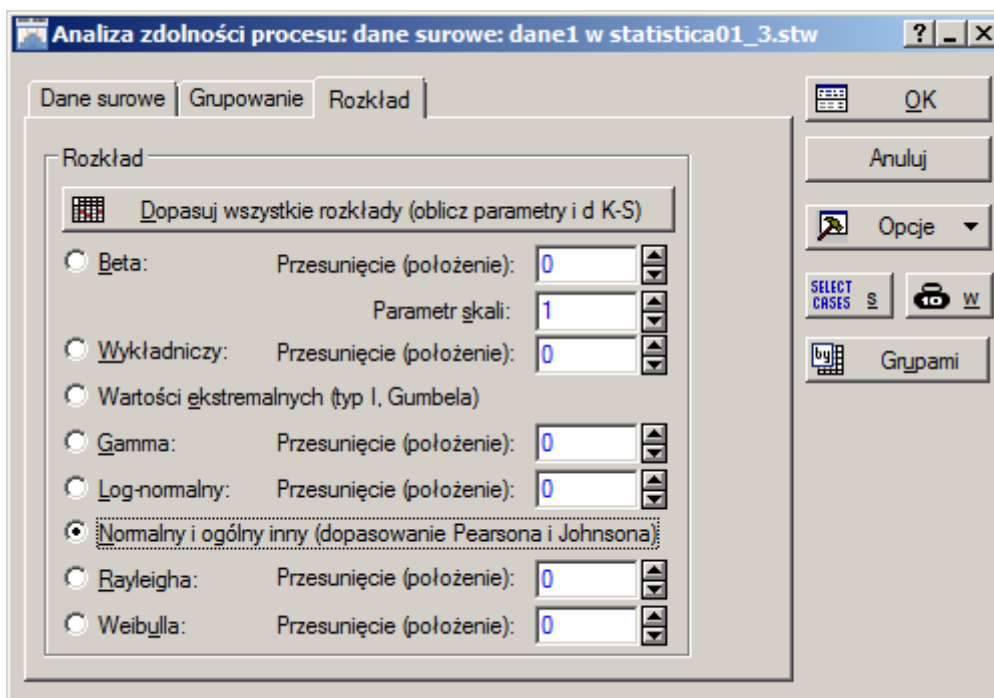


### 6.2.4. Okno Analiza zdolności procesu

Zgodność danych z podstawowymi rozkładami teoretycznymi może zostać zbadana po wskazaniu opcji



**Analiza zdolności procesu i granice tolerancji, dane surowe.** Na zakładce **Rozkład** można określić parametry rozkładów teoretycznych (w przypadku rozkładu normalnego, określane są na podstawie średniej i odchylenia standardowego z próby) i przeprowadzić test po naciśnięciu przycisku **Dopasuj wszystkie rozkłady (oblicz parametry i d K-S)**.



Wartość granicznego poziomu istotności dla rozkładu normalnego jest **n.i.**, więc mimo że wartość statystyki  $D$  dla rozkładu logarytmiczno normalnego ma mniejszą wartość (co oznacza że maksymalna





różnica pomiędzy dystrybuantami empiryczną i teoretyczną jest w tym przypadku najmniejsza) zgodnie z uwagą w nagłówku okna testu należy przyjąć że rozkład danych z próby jest rozkładem normalnym.

statistica01\_3.stw - Oceny parametrów dla wszystkich rozkładów (dane1 w statistica01\_3.stw)

Oceny parametrów dla wszystkich rozkładów  
(dane1 w statistica01\_3.stw) Zmienna: długość N = 100  
(wartości p przy założeniu znajomości parametrów a priori)  
Rozkłady są uporządkowane wg dobroci dopasowania do danych  
(u góry są najlepiej dopasowane).  
Wszystkie rozkłady z poziomem p testu Kołmogorowa-Smirnowa  
oznaczonym jako (w kolejności od najlepszej do najgorszej) n.i.,  
 $0,1 \leq p < 0,2$  lub  $0,5 \leq p < 0,1$  są dobrymi modelami danych.  
Uwaga: Jeśli rozkład normalny pasuje do danych, to należy go użyć.

Rozkład	Użytkow. Param. 1	Użytkow. Param. 2	Param. 1	Param. 2	d K-S	K-S p
Log-normalny (próg, skala, kształt)	0,00		3,04218	0,0330	0,028586	n.i.
Gamma( próg, skala, kształt)	0,00		0,02267	924,6606	0,030999	n.i.
Inny niż normalny (skośność, kurtoza)			0,21238	0,2050	0,032319	n.i.
<b>Normalny (położenie, skala)</b>			20,96216	0,6940	0,034096	n.i.
Wartości ekstrem.(położenie, skala)			20,62251	0,6559	0,071732	n.i.
Weibulla (próg, skala, kształt)	0,00		21,30178	29,6473	0,089967	n.i.
Rayleigha (próg, skala)	0,00		14,83052		0,572317	p<,01
Wykładniczy (próg, skala)	0,00		20,96216		0,602323	p<,01
Beta (próg, sigma, kształt, kształt)	0,00	1,000000				--

Oceny parametrów dla wszystkich rozkładów (dane1 w statistica...

