Statystyczna analiza danych

Kłamstwo, wierutne kłamstwo, statystyka.

⋆ ⋆ ⋆

Liczby nie kłamią, ale kłamcy liczą.

Charles H. Grosvenor

⋆ ⋆ ⋆

Statystyka jest bardziej sposobem myślenia lub wnioskowania niż pęczkiem recept na młócenie danych w celu odsłonięcia odpowiedzi.

C. R. Rao

⋆ ⋆ ⋆

... statystyka jest nauką o tym, jak wykorzystywać informacje do analizy i wytyczania kierunków działania w warunkach niepewności.

Vic Barnett

⋆ ⋆ ⋆

Nauka nie stara się wyjaśniać, a nawet niemal nie stara się interpretować, zajmuje się ona głównie budową modeli. Model rozumiany jest jako matematyczny twór, który, po dodaniu słownej interpretacji opisuje badane zjawiska. Jedynym i właściwym uzasadnieniem takiego tworu matematycznego jest oczekiwanie, że sprawdzi się on w działaniu.

John von Neumann

⋆ ⋆ ⋆

Prezentacje tabelaryczne

Prezentacje graficzne

Miary położenia

średnie, tendencji centralnej

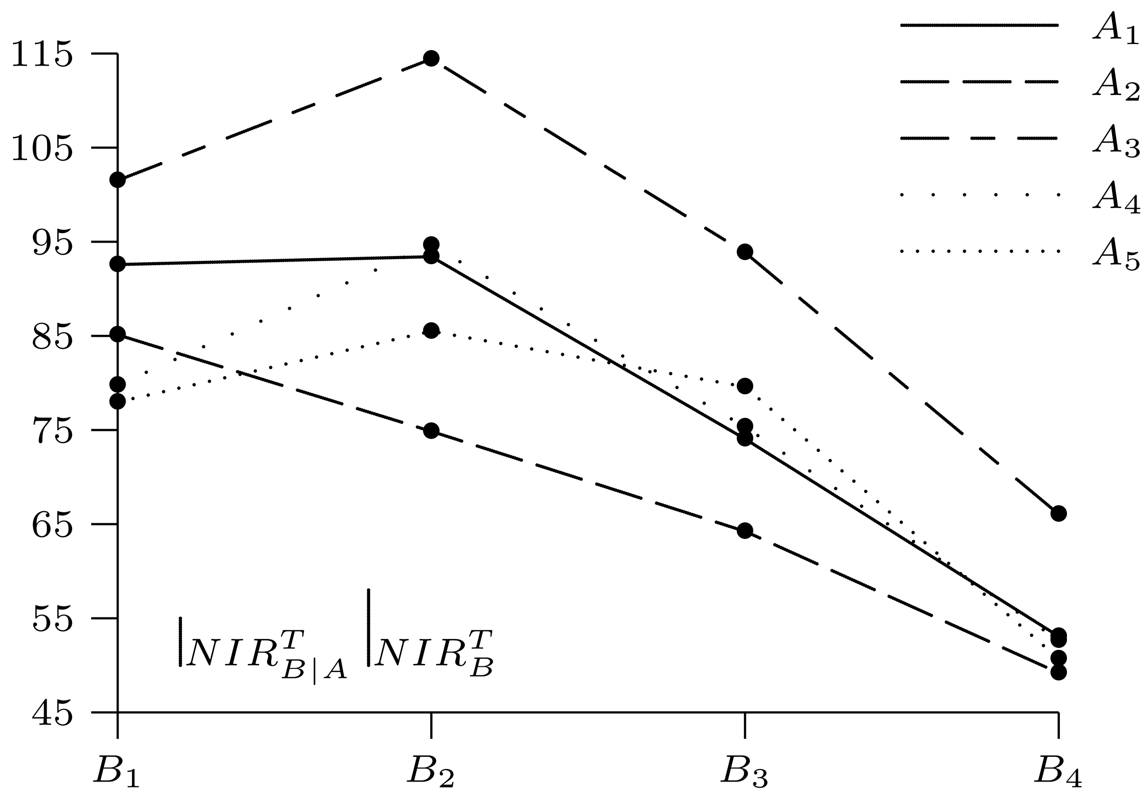
Miary rozproszenia

zmienności, zróżnicowania

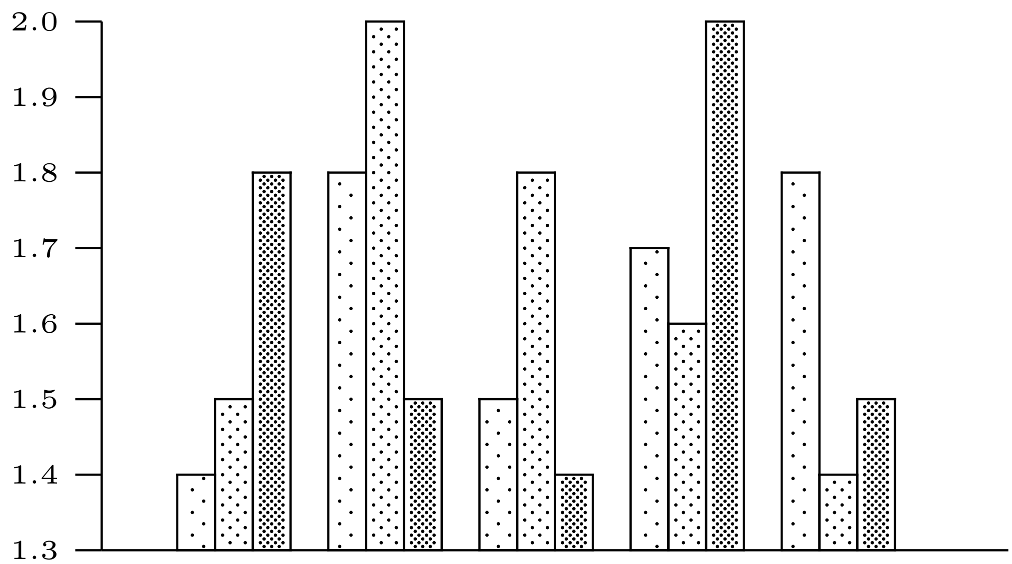
Miary asymetrii

Kurtoza

Grupa	Płeć	Ocena 2.0	3.0	3.5	4.0	4.5	Razem
1	K	3	2	3	1	1	10
	M		8	9	3		20
	Razem	3	10	12	4	1	30
2	K	7	2	4			13
	M	4	5	8			17
	Razem	11	7	12			30
3	K	13	6	1			20
	M	4	5	1			10
	Razem	17	11	2			30
4	K		6	7	4		17
	M		8	3	2		13
	Razem		14	10	6		30
5	K	6	6	3	3		18
	M		2	6	4		12
	Razem	6	8	9	7		30
6	K	2	3	9	4	1	19
	M	2	1	5	2	1	11
	Razem	4	4	14	6	2	30
7	K	6	6	5	1		18
	M	6	4	1		1	12
	Razem	12	10	6	1	1	30
8	K	1	8	6	3		18
	M	2	2	3	4	1	12
	Razem	3	10	9	7	1	30
9	K	2	4	6	2	1	15
	M	5	1	5	3	1	15
	Razem	7	5	11	5	2	30
10	K	6	4	4			14
	M	7	5	3	1		16
	Razem	13	9	7	1		30
Suma całkowita		76	88	92	37	7	300



A_1
A_2
A_3
A_4
A_5
115
105
95
85
75
65
55
45
$\big|NIR^T_{B|A}$
$\big|NIR^T_B$
B_1
B_2
B_3
B_4



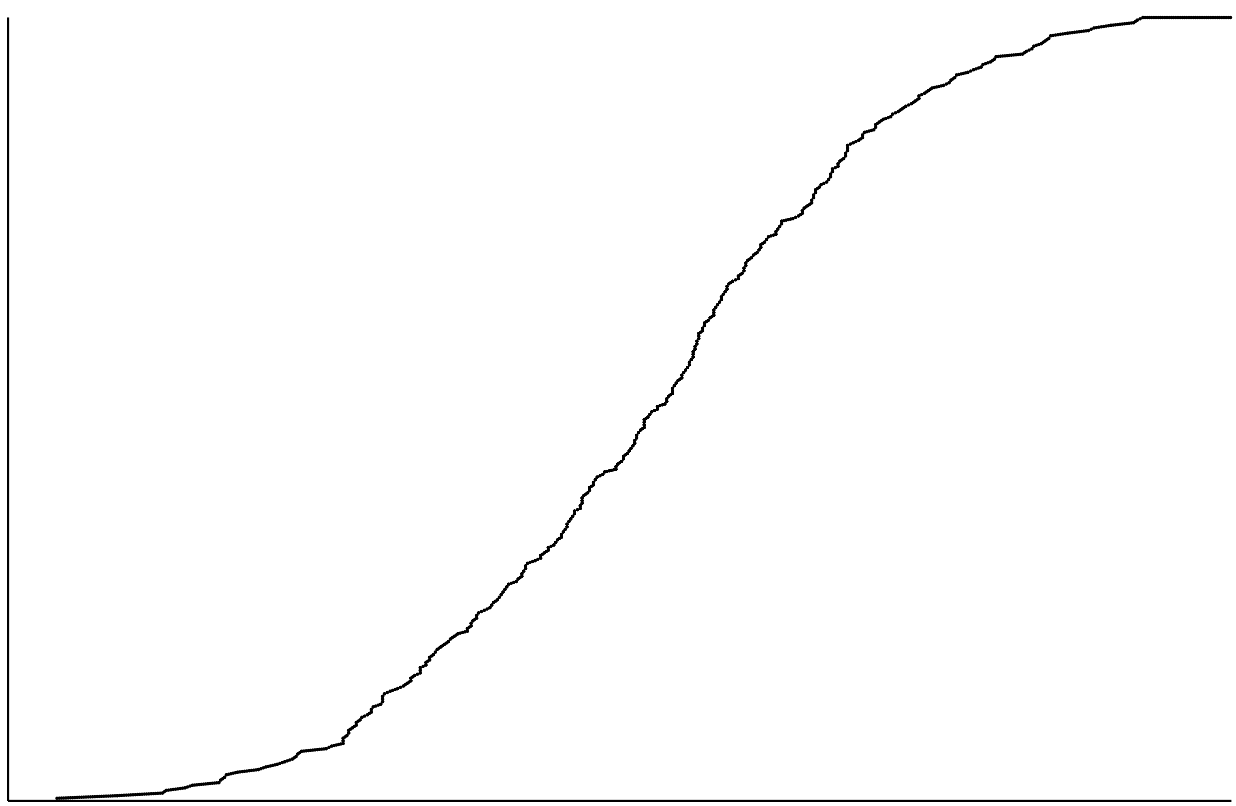
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3

Próba prosta (dane indywidualne)

$$X_1, X_2, \ldots, X_n \qquad (X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n})$$

Dystrybuanta empiryczna:

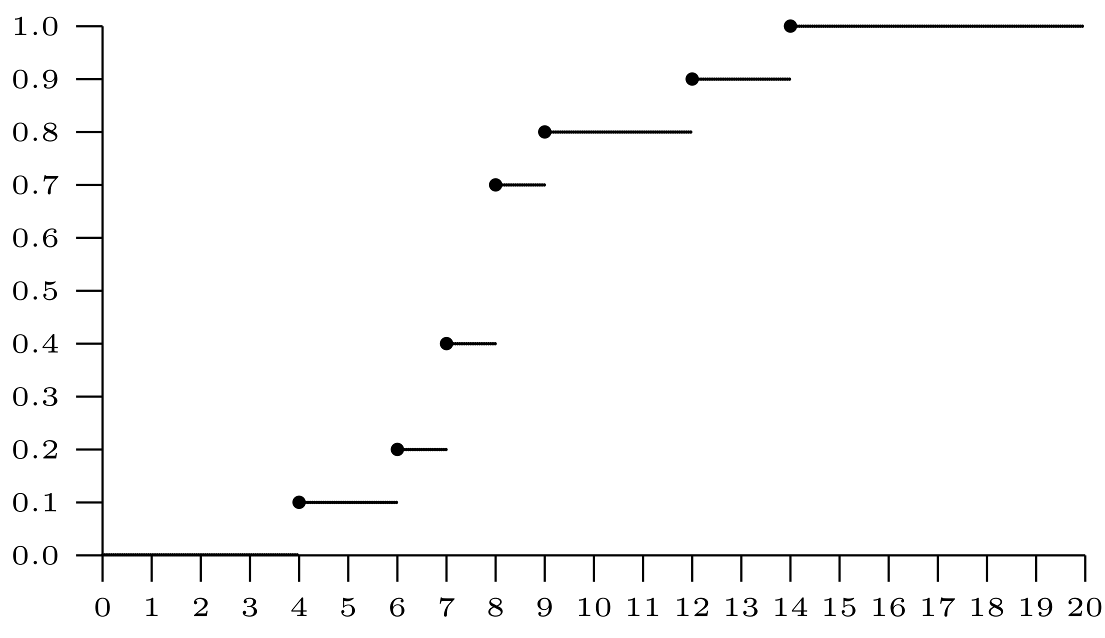
$$F_n(t) = \frac{\#\{1 \leq j \leq n : X_j \leq t\}}{n}$$



Próba: $14, 8, 12, 7, 8, 6, 8, 7, 4, 9$

$$F_{10}(t) = \frac{\#\{1 \leq j \leq n : X_j \leq t\}}{10} = \begin{cases} 0, & t < 4, \\ 1/10, & t < 6, \\ 2/10, & t < 7, \\ 4/10, & t < 8, \\ 7/10, & t < 9, \\ 8/10, & t < 12, \\ 9/10, & t < 14, \\ 1, & t \geq 14, \end{cases}$$

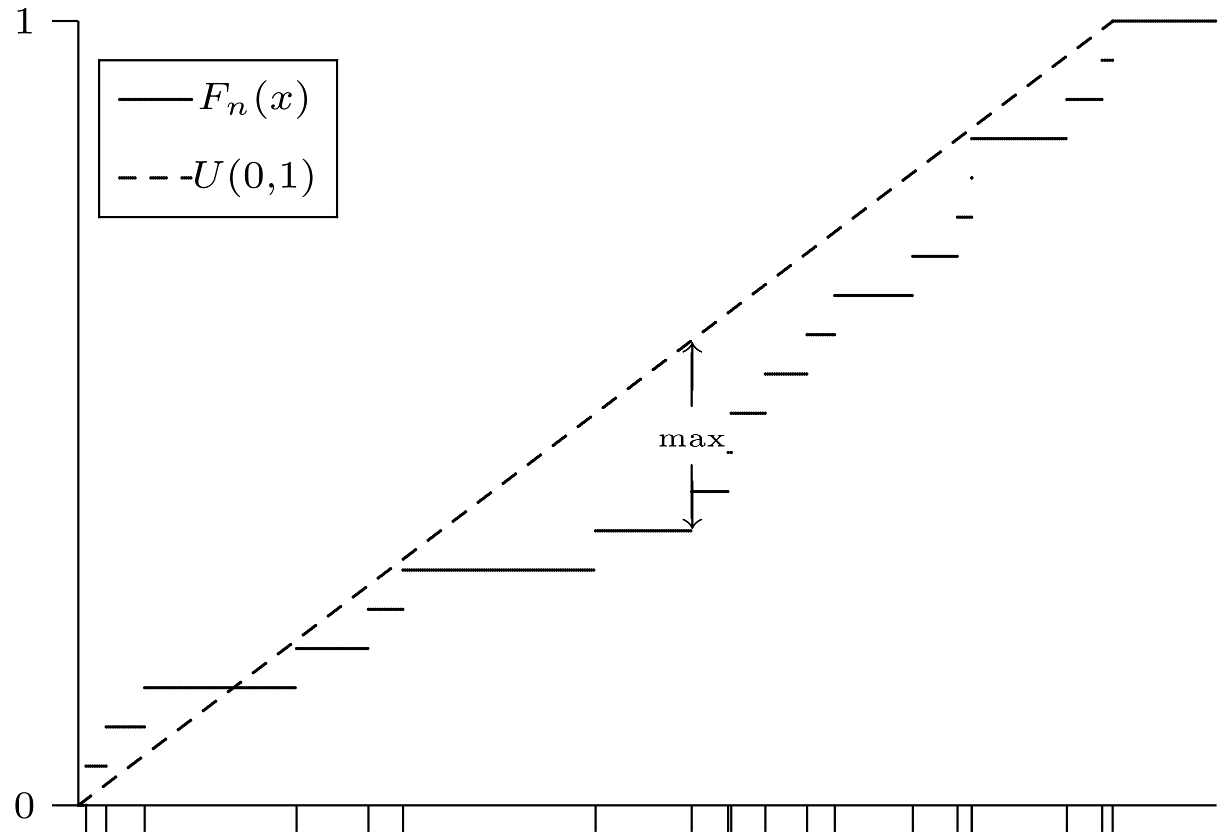
Dystrybuanta empiryczna dla próby o liczności 10



Test Kołmogorowa $H_0 : X \sim F$

Statystyka testowa

$$D_n = \max_{1\le i\le n}\left\{\max\left\{\left|F(X_{i:n}) - \frac{i-1}{n}\right|, \left|\frac{i}{n} - F(X_{i:n})\right|\right\}\right\}$$



$$P_F\{\sqrt{n}D_n + \frac{1}{6\sqrt{n}} \le y\} \approx \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2y^2} (= K(y))$$

Chi–kwadrat

$\nu = 0.58855333$	$D = 0.62679$ $p \approx 0.0000$

Wykładniczy

$\lambda = 0.58855333$	$D = 0.44674$ $p \approx 0.0000$

Jednostajny

$a = 0$ $b = 1$	$D = 0.35800$ $p \approx 0.0000$

Lognormalny

średnia $= -0.5502271$ wariancja $= 0.04209891$	$D = 0.06962$ $p \approx 0.1019$

Gamma

$\lambda = 0.02350998$ $\alpha = 25.034195$	$D = 0.05821$ $p \approx 0.2492$

Normalny

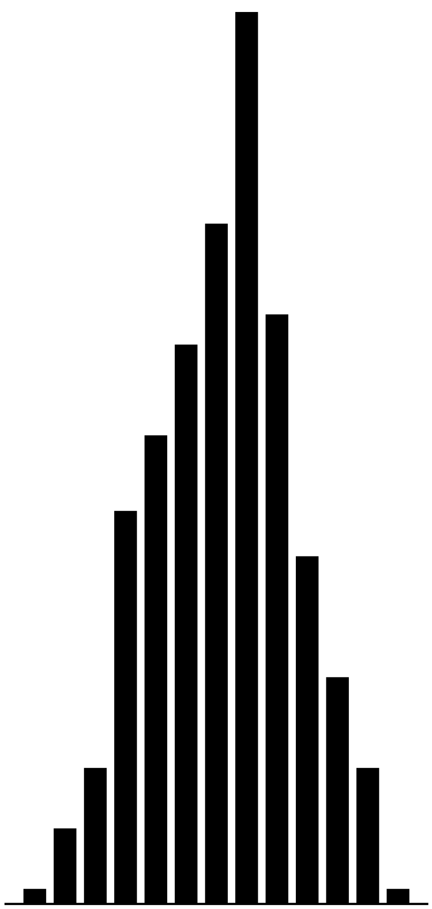
$\mu = 0.58855333$ $\sigma^2 = 0.01313309$	$D = 0.03349$ $p \approx 0.8773$

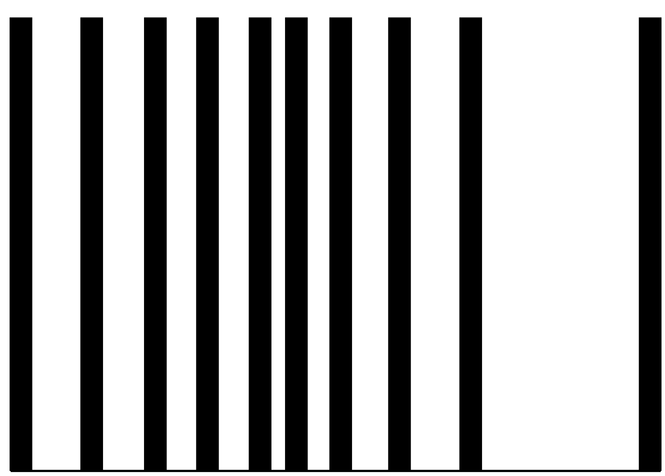
Szereg rozdzielczy (dane skumulowane)

Przedział klasowy	Liczebność	Liczebność skumulowana
$x_0 - x_1$	n_1	$n_{(1)}$
$x_1 - x_2$	n_2	$n_{(2)}$
$\vdots$	$\vdots$	$\vdots$
$x_{k-1} - x_k$	n_k	$n_{(k)}$

$$n_{(i)} = n_1 + n_2 + \cdots + n_{i-1}$$

$$n_{(1)} = n_1 \qquad n_{(k)} = n$$





Test chi–kwadrat zgodności $H_0 : X \sim F$

Klasa	Liczebność
1	n_1
2	n_2
$\vdots$	$\vdots$
k	n_k

$$\chi^2_{\mathrm{emp}} = \sum_{i=1}^{k} \frac{(n_i - n_i^t)^2}{n_i^t}$$

$$n_i^t = N p_i^t,\ N = \sum_{i=1}^{k} n_i,$$

$$p_i^t = P_F\{X \text{ przyjęła wartość z klasy } i\}$$

$$P_F\{\chi^2_{\mathrm{emp}} \le y\} \approx \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^y x^{\nu/2-1} e^{-x/2} dx$$

$$\nu = k - u - 1$$

Chi–kwadrat

$\nu = 0.58855333$	$\chi^2 = 2172.68757$ $p \approx 0.0000$

Wykładniczy

$\lambda = 0.58855333$	$\chi^2 = 896.28495$ $p \approx 0.0000$

Jednostajny

$a = 0$ $b = 1$	$\chi^2 = 123.82161$ $p \approx 0.0000$

Lognormalny

średnia $= -0.5502271$ wariancja $= 0.04209891$	$\chi^2 = 24.80685$ $p \approx 0.1019$

Gamma

$\lambda = 0.02350998$ $\alpha = 25.034195$	$\chi^2 = 17.87920$ $p \approx 0.2492$

Normalny

$\mu = 0.58855333$ $\sigma^2 = 0.01313309$	$\chi^2 = 7.43925$ $p \approx 0.8773$

Pracodawca przypuszcza, że liczba pracowników nieobecnych w różne dni tygodnia nie jest taka sama. W tym celu w ciągu pewnego okresu czasu zebrał następujące dane

Dzień	n_i
Poniedziałek	200
Wtorek	160
Środa	140
Czwartek	140
Piątek	100

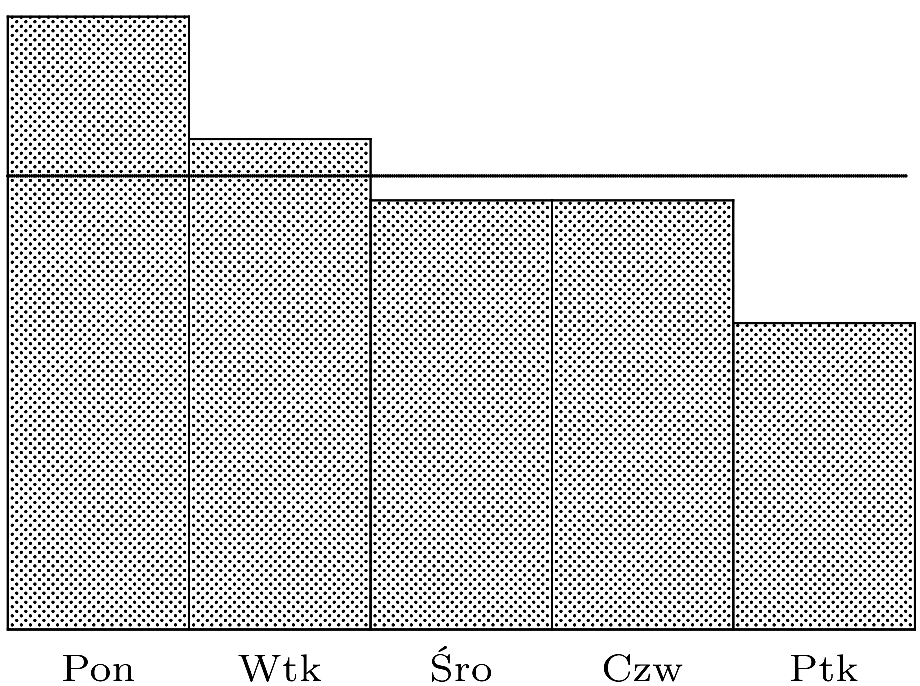
$$H_0 : X \text{ ma rozkład } \begin{array}{ccccc} \text{Pon} & \text{Wtk} & \text{Śro} & \text{Czw} & \text{Ptk} \\ \hline 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{array}$$

Test chi–kwadrat zgodności

Obliczenia

Dzień	n_i	p_i^t	n_i^t	$(n_i-n_i^t)^2/n_i^t$
Poniedziałek	200	1/5	148	$\frac{(200-148)^2}{148}=18.270$
Wtorek	160	1/5	148	$\frac{(160-148)^2}{148}=\ 0.973$
Środa	140	1/5	148	$\frac{(140-148)^2}{148}=\ 0.432$
Czwartek	140	1/5	148	$\frac{(140-148)^2}{148}=\ 0.432$
Piątek	100	1/5	148	$\frac{(100-148)^2}{148}=15.676$
	740			$\chi^2_{\mathrm{emp}}=35.676$

$$p \approx 0.0000 \qquad (\nu = 4)$$



Zgodność z rozkładem normalnym

$\dot{x}_i$	x_{i-1}	x_i	n_i	p_i^t	n_i^t	$(n_i - n_i^t)^2/n_i^t$
0.2875	$-\infty$	0.300	1	0.00712	2.13634	0.60443
0.3125	0.300	0.325	1	0.00550	1.65001	0.25607
0.3375	0.325	0.350	4	0.00885	2.65528	0.68101
0.3625	0.350	0.375	5	0.01361	4.08344	0.20573
0.3875	0.375	0.400	4	0.02000	6.00115	0.66730
0.4125	0.400	0.425	6	0.02809	8.42818	0.69957
0.4375	0.425	0.450	20	0.03771	11.31166	6.67341
0.4625	0.450	0.475	14	0.04836	14.50815	0.01780
0.4875	0.475	0.500	17	0.05927	17.78245	0.03443
0.5125	0.500	0.525	18	0.06943	20.82884	0.38420
0.5375	0.525	0.550	19	0.07772	23.31479	0.79852
0.5625	0.550	0.575	20	0.08313	24.93962	0.97836
0.5875	0.575	0.600	25	0.08498	25.49436	0.00959
0.6125	0.600	0.625	34	0.08302	24.90515	3.32125
0.6375	0.625	0.650	25	0.07750	23.25038	0.13166
0.6625	0.650	0.675	15	0.06914	20.74259	1.58984
0.6875	0.675	0.700	24	0.05895	17.68434	2.25553
0.7125	0.700	0.725	13	0.04803	14.40816	0.13762
0.7375	0.725	0.750	10	0.03739	11.21817	0.13228
0.7625	0.750	0.775	10	0.02782	8.34697	0.32736
0.7875	0.775	0.800	5	0.01978	5.93511	0.14733
0.8125	0.800	0.825	5	0.01344	4.03292	0.23190
0.8375	0.825	0.850	4	0.00873	2.61880	0.72846
0.8625	0.850	∞	1	0.01241	3.72313	1.99172

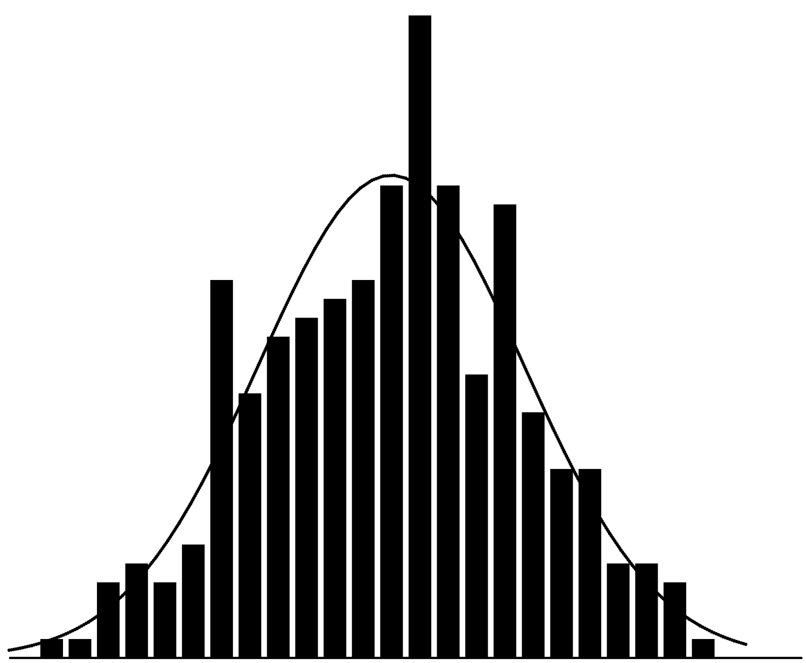
$$p_i^t = \Phi\left(\frac{x_i - \hat{\mu}}{\hat{\sigma}}\right) - \Phi\left(\frac{x_{i-i} - \hat{\mu}}{\hat{\sigma}}\right)$$

$$\tilde{\mu} = \frac{1}{n}\sum_{i=1}^{k} \dot{x}_i n_i = 0.587833$$

$$\tilde{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{k} \left(\dot{x}_i - \tilde{\mu}\right)^2 n_i} = 0.114509$$

$$\hat{\mu} = 0.587119 \qquad \hat{\sigma} = 0.117139$$

$$\chi^2_{emp} = 23.0054 \quad p \approx 0.3437$$

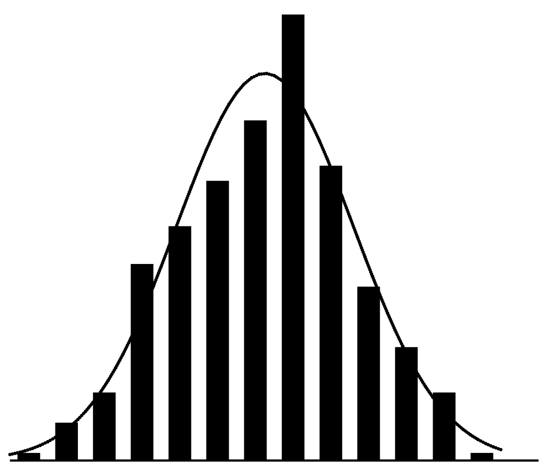


$\dot{x}_i$	x_{i-1}	x_i	n_i	p_i^t	n_i^t	$(n_i - n_i^t)^2/n_i^t$
0.275	$-\infty$	0.30	1	0.00692	2.07689	0.55838
0.325	0.30	0.35	5	0.01408	4.22519	0.14208
0.375	0.35	0.40	9	0.03318	9.95461	0.09154
0.425	0.40	0.45	26	0.06527	19.58244	2.10316
0.475	0.45	0.50	31	0.10722	32.16557	0.04224
0.525	0.50	0.55	37	0.14706	44.11750	1.14827
0.575	0.55	0.60	45	0.16843	50.52766	0.60472
0.625	0.60	0.65	59	0.16108	48.32268	2.35925
0.675	0.65	0.70	39	0.12863	38.59003	0.00436
0.725	0.70	0.75	23	0.08578	25.73333	0.29033
0.775	0.75	0.80	15	0.04776	14.32871	0.03145
0.825	0.80	0.85	9	0.02221	6.66179	0.82068
0.875	0.85	∞	1	0.01238	3.71361	1.98289

$$\tilde{\mu} = 0.587167 \qquad \tilde{\sigma} = 0.115659$$

$$\hat{\mu} = 0.587625 \qquad \hat{\sigma} = 0.116862$$

$$\chi^2_{emp} = 10.1793 \qquad p \approx 0.4249$$

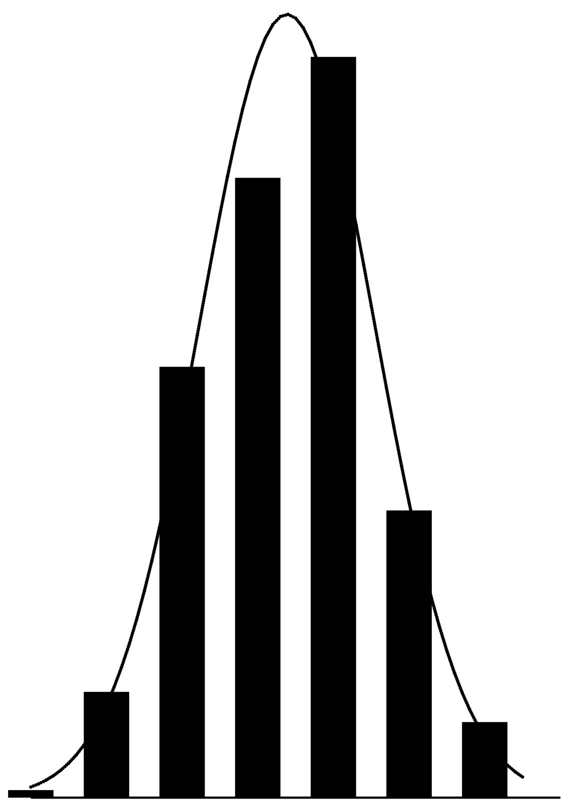


$\dot{x}_i$	x_{i-1}	x_i	n_i	p_i^t	n_i^t	$(n_i - n_i^t)^2 / n_i^t$
0.250	$-\infty$	0.30	1	0.00622	1.86708	0.40267
0.350	0.30	0.40	14	0.04497	13.49226	0.01911
0.450	0.40	0.50	57	0.17020	51.05911	0.69124
0.550	0.50	0.60	82	0.31778	95.33549	1.86536
0.650	0.60	0.70	98	0.29335	88.00614	1.13489
0.750	0.70	0.80	38	0.13385	40.15433	0.11558
0.850	0.80	∞	10	0.03362	10.08559	0.00073

$$\tilde{\mu} = 0.588667 \qquad \tilde{\sigma} = 0.117919$$

$$\hat{\mu} = 0.588639 \qquad \hat{\sigma} = 0.115492$$

$$\chi^2_{emp} = 4.2296 \qquad p \approx 0.3758$$

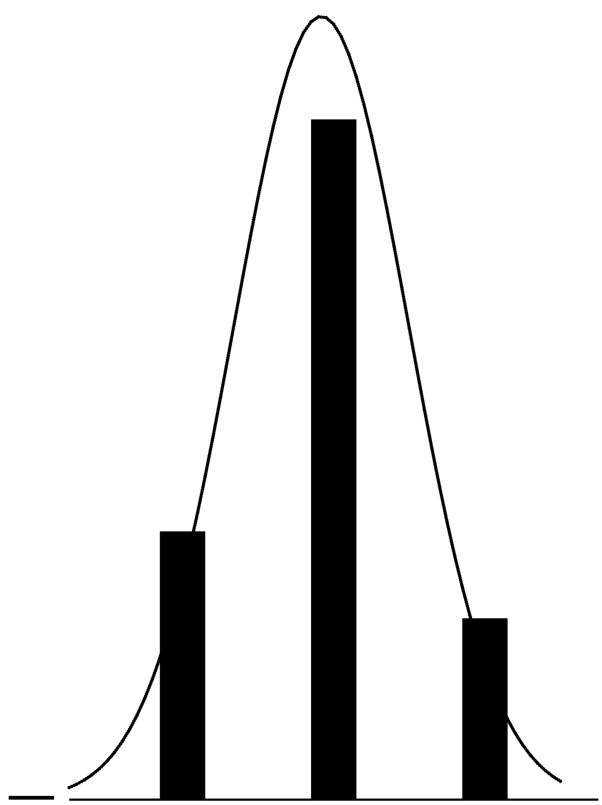


$\dot{x}_i$	x_{i-1}	x_i	n_i	p_i^t	n_i^t	$(n_i - n_i^t)^2/n_i^t$
0.200	$-\infty$	0.30	1	0.00704	2.11215	0.58560
0.400	0.30	0.50	71	0.22760	68.27994	0.10836
0.600	0.50	0.70	180	0.60861	182.58184	0.03651
0.800	0.70	∞	48	0.15675	47.02607	0.02017

$$\tilde{\mu} = 0.583333 \qquad \tilde{\sigma} = 0.126973$$

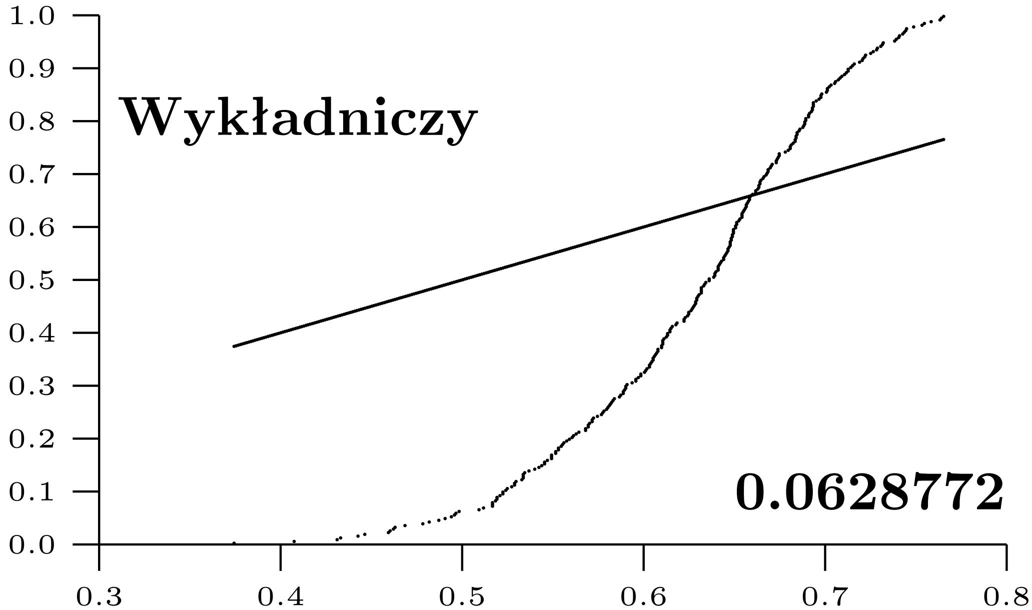
$$\hat{\mu} = 0.583585 \qquad \hat{\sigma} = 0.115504$$

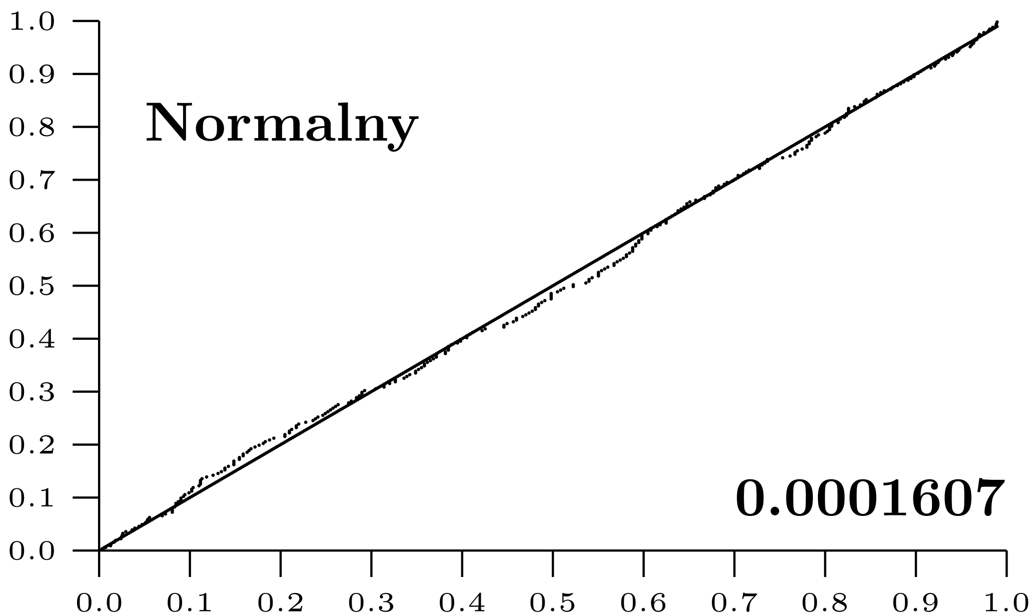
$$\chi^2_{emp} = 0.7506 \qquad p \approx 0.3863$$



Wykres $P-P$

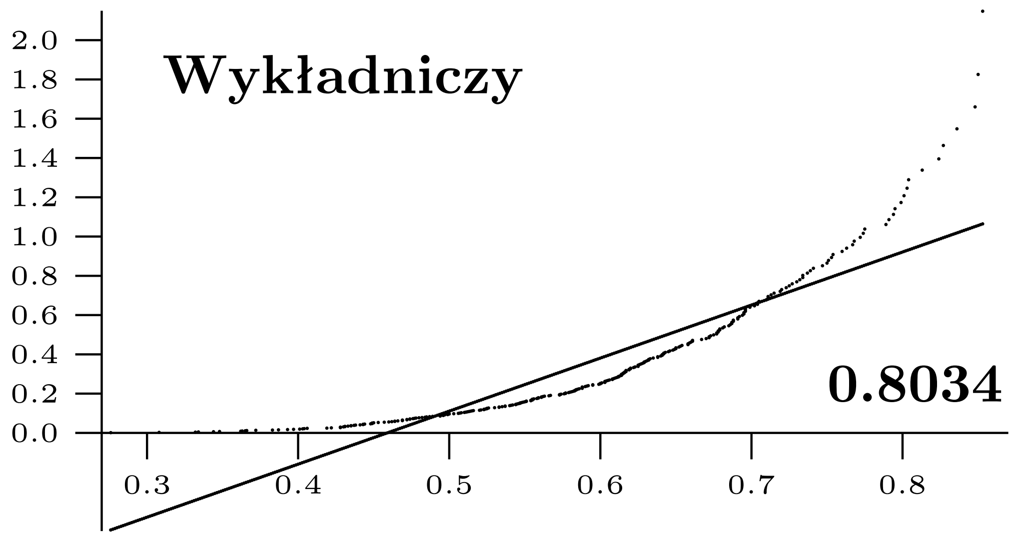
$$(F(X_{i:n}); F_n(X_{i:n}))\,, \quad i = 1, \ldots, n$$

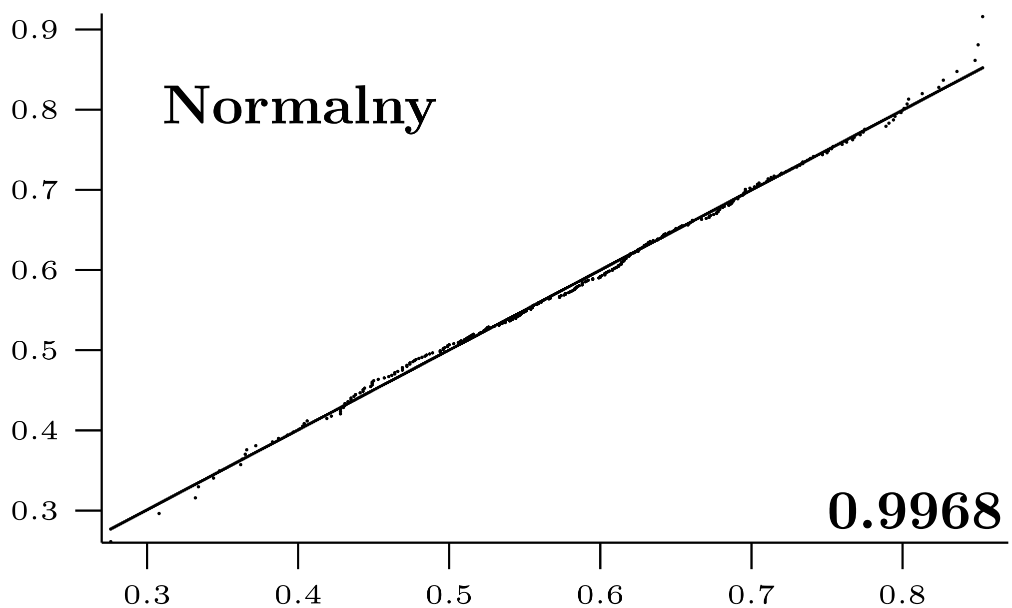




Wykres $Q - Q$

$$\left(X_{i:n}; F^{-1}\left(\frac{3i-1}{3n+1}\right)\right), \quad i = 1, \ldots, n$$





Wykres $P-P$

$$\frac{1}{n}\sum_{i=1}^{n}\left(F(X_{i:n})-F_n(X_{i:n})\right)^2$$

. . . .

Wykres $Q-Q$

$$\frac{\displaystyle\sum_{i=1}^{n}(\hat{Y}_{i:n}-\bar{Y})^2}{\displaystyle\sum_{i=1}^{n}(Y_{i:n}-\bar{Y})^2}$$

$$Y_{i:n}=F^{-1}\left(\frac{3i-1}{3n+1}\right); \qquad \bar{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_{i:n}$$

$$\beta_1=\frac{\sum_{i=1}^{n}\left(X_{i:n}-\bar{X}\right)\left(Y_{i:n}-\bar{Y}\right)}{\sum_{i=1}^{n}\left(X_{i:n}-\bar{X}\right)^2}; \beta_0=\bar{Y}-\beta_1\bar{X}$$

$$\hat{Y}_{i:n}=\beta_0+\beta_1X_{i:n}$$

Próba prosta (dane indywidualne)

$$X_1, X_2, \ldots, X_n \qquad (X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n})$$

Szereg rozdzielczy (dane skumulowane)

Przedział klasowy	Liczebność	Liczebność skumulowana
$x_0 - x_1$	n_1	$n_{(1)}$
$x_1 - x_2$	n_2	$n_{(2)}$
$\vdots$	$\vdots$	$\vdots$
$x_{k-1} - x_k$	n_k	$n_{(k)}$

$$n_{(i)} = n_1 + n_2 + \cdots + n_{i-1}$$

$$n_{(1)} = n_1 \qquad n_{(k)} = n$$

Niech $0 \leq p \leq 1$
x_p: początek przedziału z obserwacją o numerze $p \cdot n$
n_p: liczebność przedziału z obserwacją o numerze $p \cdot n$
h_p: długość przedziału z obserwacją o numerze $p \cdot n$
$n_{(p)}$: liczebność skumulowana przedziału poprzedzającego przedział o początku x_p
$\dot{x}_i = (x_{i-1} + x_i)/2$

Mierniki położenia (próba prosta)

średnia

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

mediana

$$Me = \begin{cases} X_{(n+1)/2:n} & n \text{ nieparzyste} \\ (X_{n/2:n} + X_{n/2+1:n})/2 & n \text{ parzyste} \end{cases}$$

dolny kwartyl

$$Q_1 = X_{[n/4]:n}$$

górny kwartyl

$$Q_3 = X_{[3n/4]:n}$$

dominanta (moda)

$$D = \text{ najczęściej występująca wartość}$$

minimum

$$Min = X_{1:n}$$

maksimum

$$Max = X_{n:n}$$

Mierniki położenia (szereg rozdzielczy)

średnia

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} \dot{x}_i n_i$$

mediana

$$Me = x_{0.5} + \frac{h_{0.5}}{n_{0.5}}\left(\frac{n}{2} - n_{(0.5)}\right)$$

dolny kwartyl

$$Q_1 = x_{0.25} + \frac{h_{0.25}}{n_{0.25}}\left(\frac{n}{4} - n_{(0.25)}\right)$$

górny kwartyl

$$Q_3 = x_{0.75} + \frac{h_{0.75}}{n_{0.75}}\left(\frac{3n}{4} - n_{(0.75)}\right)$$

dominanta (moda)

$$D = x_D + h_D \frac{n_D - n_{D-1}}{2n_D - n_{D+1} - n_{D-1}}$$

minimum

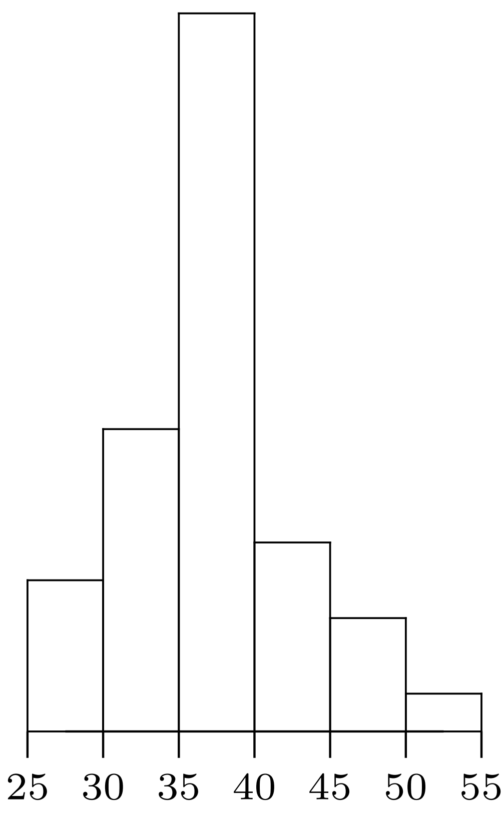
$$Min = x_0$$

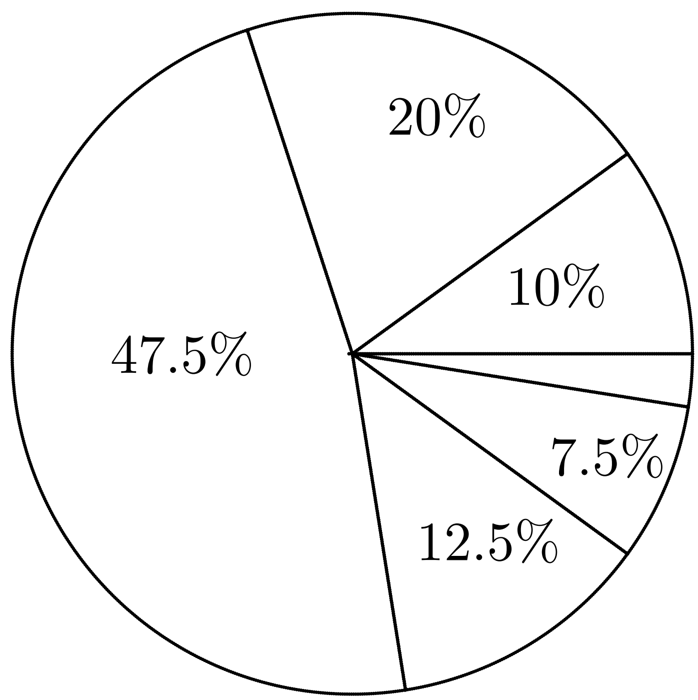
maksimum

$$Max = x_k$$

Przykład . Badano przebieg opon samochodowych wycofanych z eksploatacji.

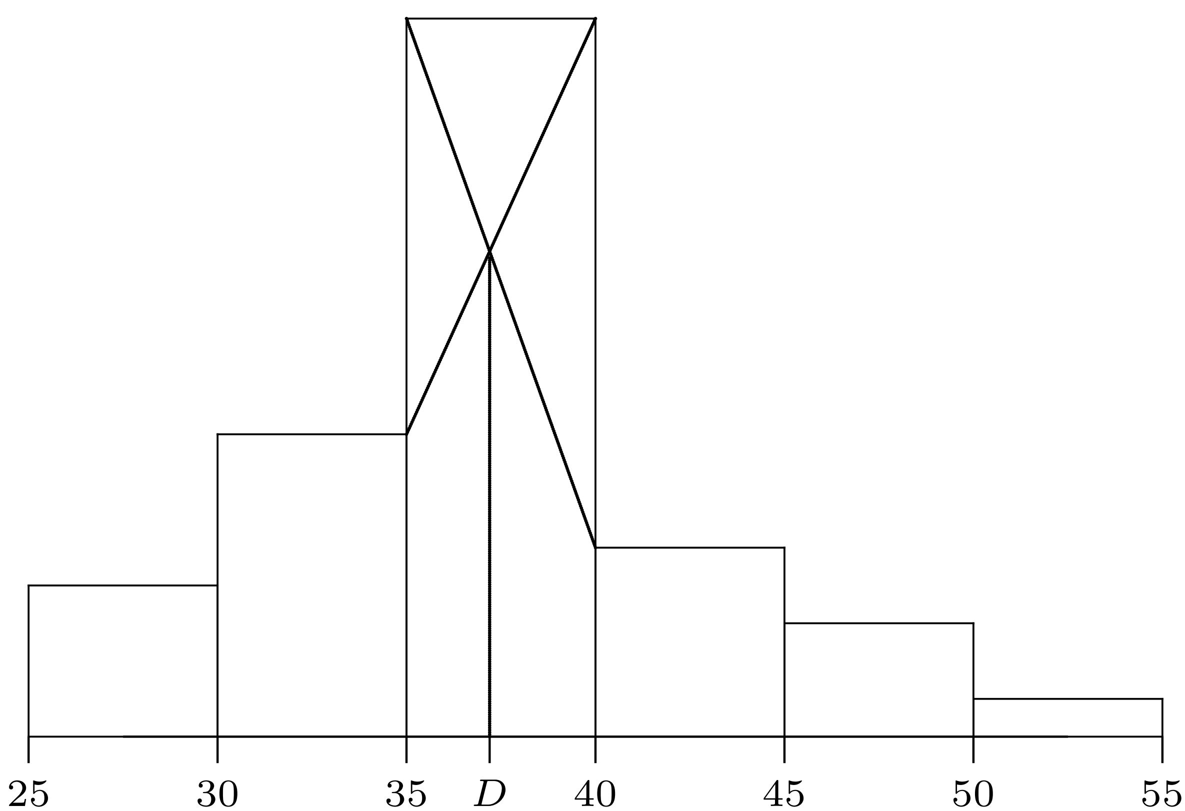
Przebieg $x_{i-1} - x_i$	Liczba n_i	$n_{(i)}$	Odsetek
$25 - 30$	20	20	10.00%
$30 - 35$	40	60	20.00%
$35 - 40$	95	155	47.50%
$40 - 45$	25	180	12.50%
$45 - 50$	15	195	7.50%
$50 - 55$	5	200	2.50%





Średni przebieg dwustu opon

$$\bar{x} = \frac{27.5 \cdot 20 + 32.5 \cdot 40 + \cdots + 52.5 \cdot 5}{200} = 37.25$$

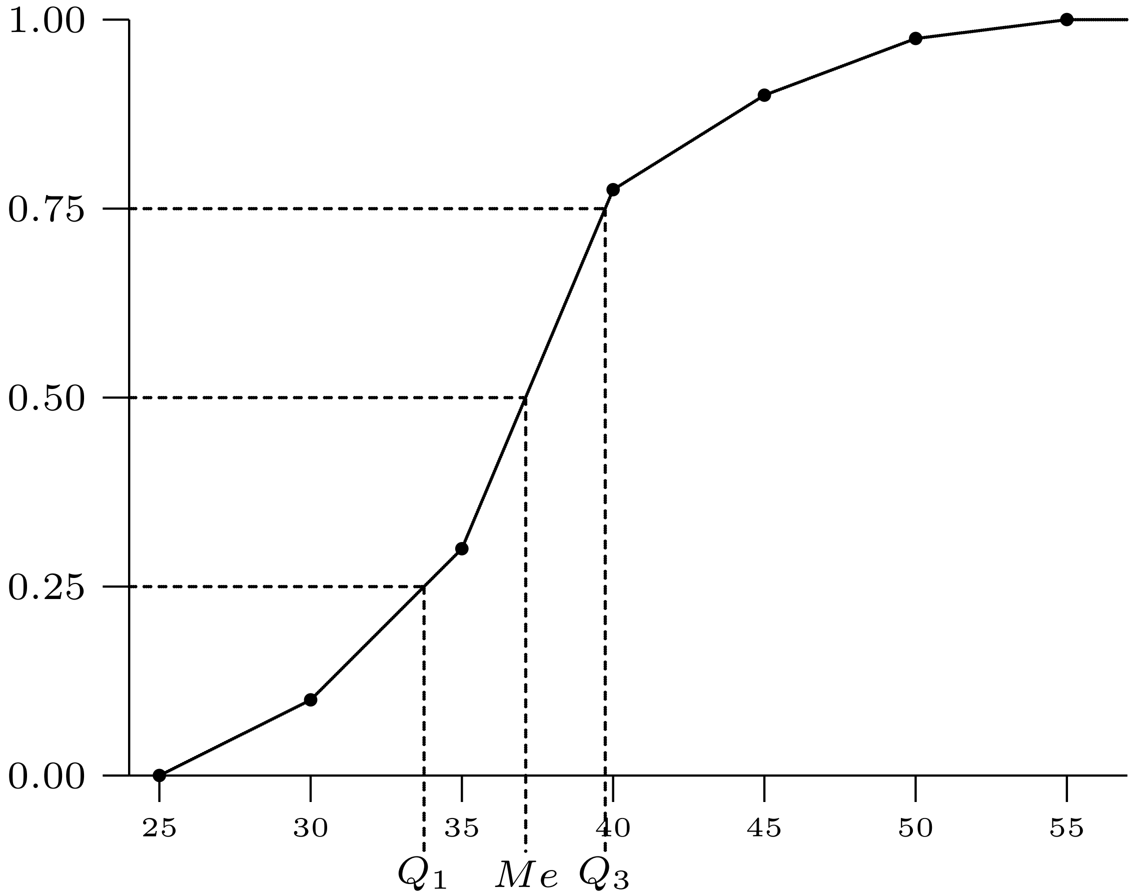


Dominanta przebiegu dwustu opon

$$x_D = 35 \quad h_D = 5$$

$$n_D = 95 \quad n_{D-1} = 40 \quad n_{D+1} = 25$$

$$D = 35 + 5 \cdot \frac{95 - 40}{2 \cdot 95 - 40 - 25} = 37.2$$



Mediana przebiegu dwustu opon

$$\frac{n}{2} = 100 \quad x_{0.5} = 35$$

$$h_{0.5} = 5 \quad n_{0.5} = 95 \quad n_{(0.5)} = 60$$

$$Me = 35 + \frac{5}{95} \cdot (100 - 60) = 37.11$$

Dolny kwartyl przebiegu dwustu opon

$$\frac{n}{4} = 50 \quad x_{0.25} = 30$$

$$h_{0.25} = 5 \quad n_{0.25} = 40 \quad n_{(0.25)} = 20$$

$$Q_1 = 30 + \frac{5}{40} \cdot (50 - 20) = 33.75$$

Górny kwartyl przebiegu dwustu opon

$$\frac{3n}{4} = 150 \quad x_{0.75} = 35$$

$$h_{0.75} = 5 \quad n_{0.75} = 95 \quad n_{(0.75)} = 60$$

$$Q_3 = 35 + \frac{5}{95}(150 - 60) = 39.74$$

Minimalny przebieg dwustu opon

$$Min = 25$$

Maksymalny przebieg dwustu opon

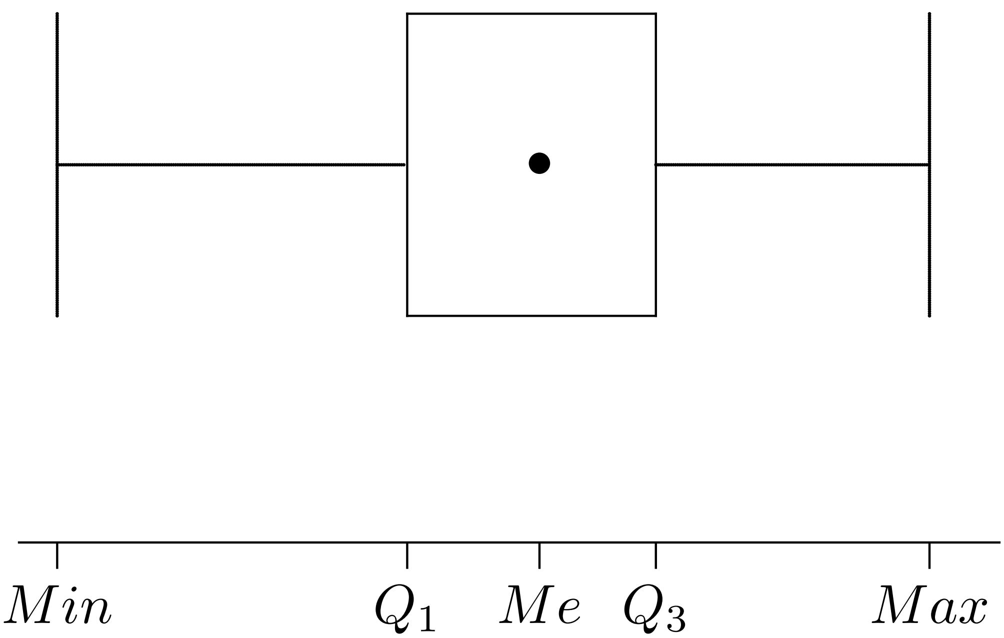
$$Max = 55$$

. . . .

$$n = 300, \qquad \bar{x} = 0.588553, \qquad D = 0.588\ \{4\}$$

$$Me = 0.595, \qquad Q_1 = 0.507, \qquad Q_3 = 0.672$$

$$Min = 0.276, \qquad Max = 0.853$$



Nieparametryczny przedział ufności dla p–tego kwantyla

Jeżeli założyć, że p–ty kwantyl x_p jest równy $X_{i:n}$, to prawdopodobieństwo uzyskania i obserwacji nie większych niż x_p wynosi

$$\binom{n}{i} p^i (1-p)^{n-i}$$

Przedział ufności dla x_p na poziomie ufności $1-\alpha$ ma postać

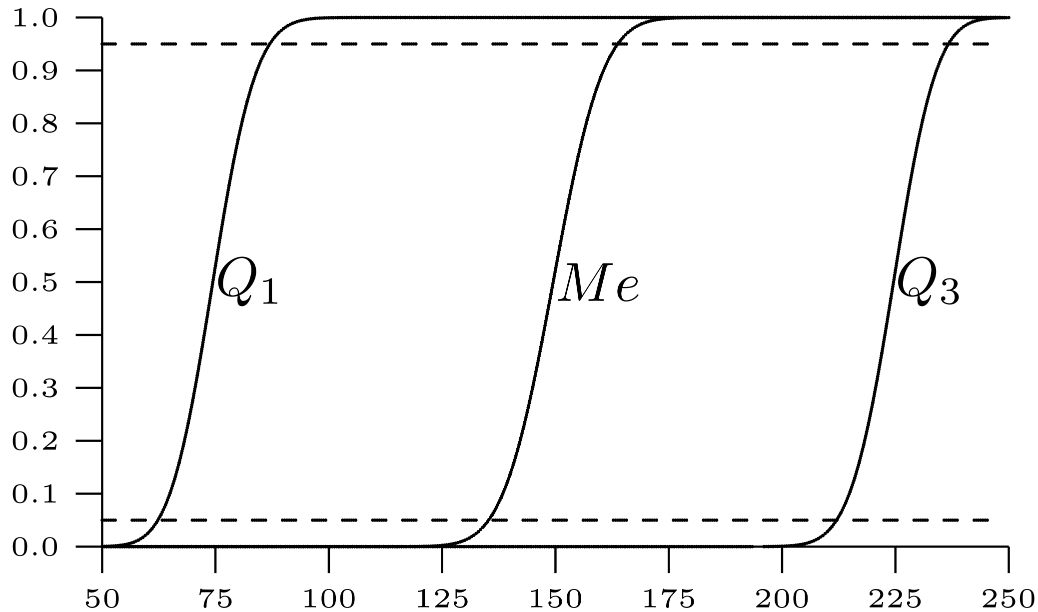
$$(X_{k_1:n}; X_{k_2:n})$$

gdzie $k_1 \leq k_2$ są takie, że

$$\sum_{j=k_1}^{k_2} \binom{n}{j} p^j (1-p)^{n-j} \geq 1-\alpha$$

oraz

$$\sum_{j=k_1+1}^{k_2-1} \binom{n}{j} p^j (1-p)^{n-j} \leq 1-\alpha$$



Przedziały ufności ($1 - \alpha = 0.9$)

Dla dolnego kwartyla: $k_1 = 63; k_2 = 87$

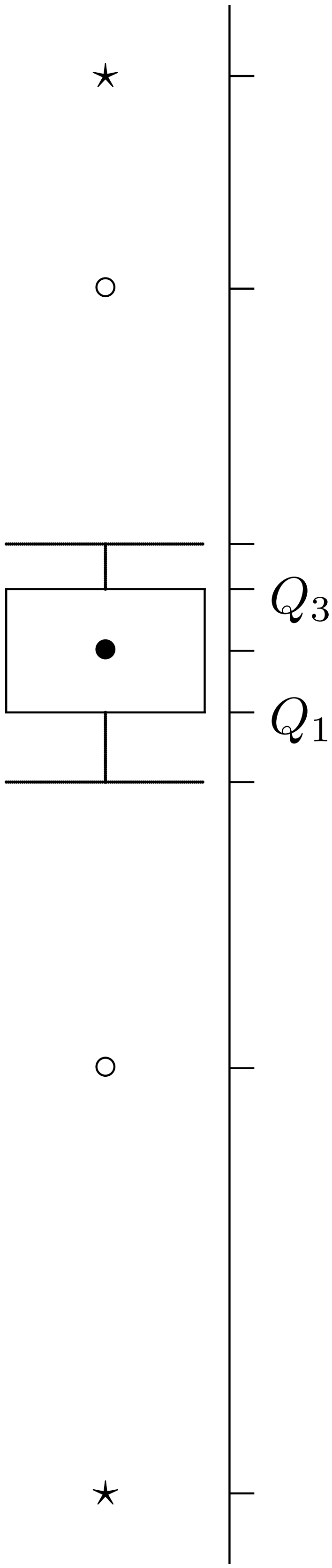
$$(0.487, 0.523)$$

Dla mediany: $k_1 = 134; k_2 = 166$

$$(0.582, 0.609)$$

Dla górnego kwartyla: $k_1 = 213; k_2 = 237$

$$(0.650, 0.685)$$



Obserwacje odstające

$$X > Q_3 + 1.5(Q_3 - Q_1)$$

$$X < Q_1 - 1.5(Q_3 - Q_1)$$

Wartości ekstremalne

$$X > Q_3 + 2 \cdot 1.5(Q_3 - Q_1)$$

$$X < Q_1 - 2 \cdot 1.5(Q_3 - Q_1)$$

Mierniki rozproszenia

wariancja

$$\mathcal{S}^2 = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{x})^2 \\ \frac{1}{n}\sum_{i=1}^{k} n_i(\dot{x}_i - \bar{x})^2 \end{cases}$$

odchylenie standardowe

$$\mathcal{S} = \sqrt{\mathcal{S}^2}$$

współczynnik zmienności

$$V = \frac{\mathcal{S}}{\bar{x}} 100\%$$

rozstęp

$$R = Max - Min$$

odchylenie przeciętne

$$d = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}|X_i - \bar{x}| \\ \frac{1}{n}\sum_{i=1}^{k} n_i|\dot{x}_i - \bar{x}| \end{cases}$$

odchylenie ćwiartkowe

$$Q = \frac{Q_3 - Q_1}{2}$$

Przykład cd. Przebieg opon
Wariancja

$$\mathcal{S}^2 = \frac{1}{200}(20(27.5 - 37.25)^2 + 40(32.5 - 37.25)^2 + \cdots + 5(52.5 - 37.25)^2) = 31.18$$

Odchylenie standardowe

$$S = \sqrt{\mathcal{S}^2} = 5.58$$

Współczynnik zmienności

$$V = \frac{5.58}{37.25} \cdot 100\% = 14.99\%$$

Rozstęp

$$R = 55 - 25 = 30$$

Odchylenie przeciętne

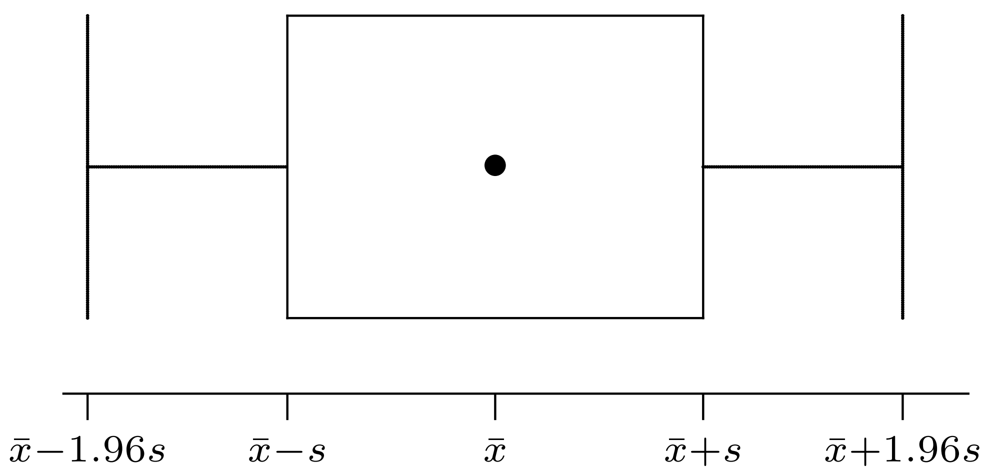
$$d = \frac{1}{200}(20|27.5 - 37.25| + 40|32.5 - 37.25| + \cdots + 5|52.5 - 37.25|) = 3.85$$

Odchylenie ćwiartkowe

$$Q = \frac{39.74 - 33.75}{2} = 2.99$$

. . . .

$$n = 300,\ R = 0.577,\ S^2 = 0.013133,\ S = 0.1146$$



Przedziały tolerancji

Niech $p, \alpha \in (0,1)$.
Dwustronny przedział tolerancji:

$$(T_1, T_2): \quad P_F\{F(T_2) - F(T_1) \geq p\} \geq 1 - \alpha$$

Jednostronny przedział tolerancji:

$$(T_1', \infty): \quad P_F\{F(T_1') \leq 1 - p\} \geq 1 - \alpha$$
$$(-\infty, T_2'): \quad P_F\{F(T_2') \geq p\} \geq 1 - \alpha$$

Jeżeli F jest rozkładem normalnym, to

$$T_1 = \bar{X} - K_n(p, \alpha)S; \quad T_2 = \bar{X} + K_n(p, \alpha)S$$

oraz

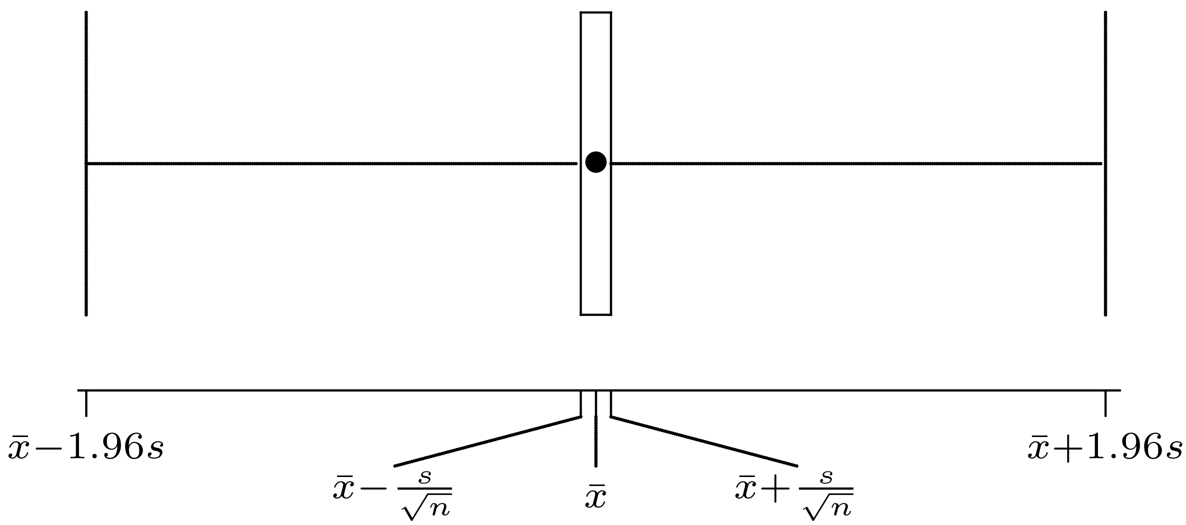
$$T_1' = \bar{X} - k_n(p, \alpha)S; \quad T_2' = \bar{X} + k_n(p, \alpha)S$$

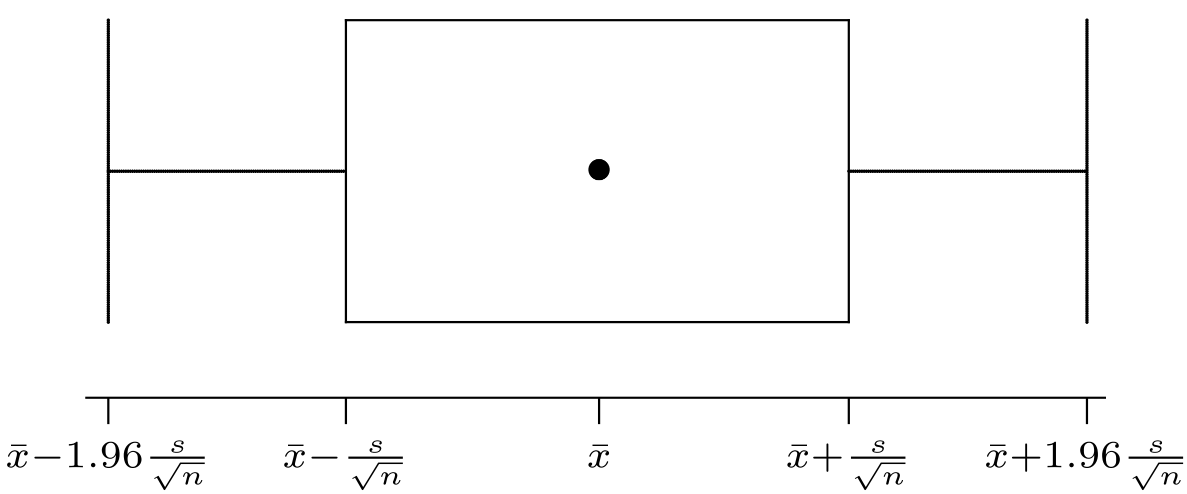
Jeżeli F jest dowolnym rozkładem ciągłym, to

$$T_1 = X_{r:n}; \quad T_2 = X_{n-s+1:n}$$

oraz

$$T_1' = X_{m:n}; \quad T_2' = X_{n-m+1:n}$$


$\bar{x} - 1.96s$
$\bar{x} - \frac{s}{\sqrt{n}}$
$\bar{x}$
$\bar{x} + \frac{s}{\sqrt{n}}$
$\bar{x} + 1.96s$


$\bar{x} - 1.96\frac{s}{\sqrt{n}}$
$\bar{x} - \frac{s}{\sqrt{n}}$
$\bar{x}$
$\bar{x} + \frac{s}{\sqrt{n}}$
$\bar{x} + 1.96\frac{s}{\sqrt{n}}$

Mierniki asymetrii

współczynnik asymetrii

$$g_1 = \frac{e_3}{S^3}$$

pozycyjny współczynnik asymetrii

$$\frac{Q_3 - 2Me + Q_1}{2Q}$$

współczynnik skośności

$$A_3 = \frac{\bar{x} - D}{S}$$

Przykład cd. Przebieg opon
Współczynnik asymetrii

$$g_1 = \frac{73.406}{5.58^3} = 0.4215$$

Pozycyjny współczynnik asymetrii

$$\frac{39.74 - 2 \cdot 37.11 + 33.75}{2 \cdot 2.99} = -0.121$$

Współczynnik skośności

$$A_3 = \frac{37.25 - 37.2}{5.58} = 0.004$$

Kurtoza

$$g_2 = \frac{3280.457}{5.58^4} = 3.3727$$

Rozkład Pearsona typu I o gęstości

$$f(x) \propto \left(1 + \frac{x}{\lambda_1}\right)^{\alpha_1} \left(1 - \frac{x}{\lambda_2}\right)^{\alpha_2}$$

. . . .

$$\bar{x} = 0.58855333, \qquad s^2 = 0.01308931$$

$$m_3 = -0.00010684, \qquad m_4 = 0.00044834$$

$$g_1 = \frac{m_3}{s^3} = -0.0713, \qquad g_2 = \frac{m_4}{s^4} = 2.6168$$

Rozkład Pearsona typu I o gęstości

$$f(x) \propto \left(1 + \frac{x}{\lambda_1}\right)^{\alpha_1} \left(1 - \frac{x}{\lambda_2}\right)^{\alpha_2}$$

lub typu II o gęstości

$$f(x) \propto \left(1 - \frac{x^2}{\lambda^2}\right)^{\alpha}$$

g_1 :	$\alpha = 0.05$		$\alpha = 0.01$	
	0.230		0.329	
g_2 :	$\alpha = 0.99$	$\alpha = 0.95$	$\alpha = 0.05$	$\alpha = 0.01$
	2.46	2.59	3.47	3.79