

Statystyka matematyczna

w zastosowaniach

Wojciech Zieliński

STATYSTYKA: nauka poświęcona metodom badania (analizowania) zjawisk masowych; polega na systematyzowaniu obserwowanych cech ilościowych i jakościowych oraz przedstawianiu wyników w postaci zestawień tabelarycznych, wykresów, itp.; posługuje się rachunkiem prawdopodobieństwa.

STATYSTYKA MATEMATYCZNA: dział matematyki stosowanej oparty na rachunku prawdopodobieństwa; zajmuje się badaniem zbiorów na podstawie znajomości własności ich części.

Encyklopedia Popularna PWN, Warszawa 1982

Grupa	Płeć	Punkty	Ocena	Grupa	Płeć	Punkty	Ocena
1	1	0.576	3.0	2	0	0.406	2.0
1	1	0.537	3.0	2	1	0.365	2.0
1	0	0.573	3.0	2	0	0.497	2.0
1	0	0.772	4.0	2	0	0.372	2.0
1	0	0.462	2.0	2	0	0.449	2.0
1	0	0.348	2.0	2	1	0.678	3.5
1	1	0.675	3.5	2	1	0.566	3.0
1	1	0.613	3.5	2	0	0.617	3.5
1	0	0.827	4.5	2	1	0.678	3.5
1	1	0.650	3.5	2	0	0.595	3.0
1	1	0.554	3.0	2	1	0.619	3.5
1	1	0.551	3.0	2	1	0.619	3.5
1	1	0.710	4.0	2	1	0.548	3.0
1	1	0.562	3.0	2	1	0.397	2.0
1	1	0.625	3.5	2	1	0.588	3.0
1	1	0.646	3.5	2	0	0.469	2.0
1	1	0.741	4.0	2	1	0.702	3.5
1	1	0.650	3.5	2	0	0.679	3.5
1	0	0.632	3.5	2	1	0.681	3.5
1	1	0.584	3.0	2	1	0.499	2.0
1	1	0.629	3.5	2	1	0.555	3.0
1	1	0.617	3.5	2	0	0.620	3.5
1	0	0.626	3.5	2	1	0.577	3.0
1	1	0.711	4.0	2	1	0.696	3.5
1	1	0.542	3.0	2	0	0.437	2.0
1	1	0.653	3.5	2	0	0.625	3.5
1	0	0.667	3.5	2	0	0.435	2.0
1	0	0.387	2.0	2	0	0.584	3.0
1	0	0.510	3.0	2	1	0.688	3.5
1	1	0.600	3.0	2	1	0.431	2.0

Grupa	Płeć	Punkty	Ocena	Grupa	Płeć	Punkty	Ocena
3	0	0.485	2.0	4	0	0.774	4.0
3	0	0.276	2.0	4	0	0.549	3.0
3	1	0.595	3.0	4	0	0.638	3.5
3	1	0.514	3.0	4	1	0.737	4.0
3	0	0.308	2.0	4	0	0.610	3.5
3	0	0.332	2.0	4	1	0.581	3.0
3	0	0.366	2.0	4	1	0.623	3.5
3	1	0.547	3.0	4	1	0.583	3.0
3	0	0.591	3.0	4	1	0.658	3.5
3	0	0.482	2.0	4	1	0.544	3.0
3	0	0.512	3.0	4	1	0.586	3.0
3	0	0.541	3.0	4	0	0.692	3.5
3	0	0.574	3.0	4	0	0.632	3.5
3	1	0.660	3.5	4	1	0.590	3.0
3	1	0.435	2.0	4	0	0.525	3.0
3	0	0.533	3.0	4	1	0.699	3.5
3	1	0.484	2.0	4	0	0.661	3.5
3	1	0.448	2.0	4	1	0.551	3.0
3	0	0.469	2.0	4	0	0.535	3.0
3	0	0.444	2.0	4	1	0.603	3.0
3	0	0.460	2.0	4	1	0.524	3.0
3	0	0.641	3.5	4	0	0.516	3.0
3	0	0.362	2.0	4	0	0.561	3.0
3	0	0.393	2.0	4	0	0.630	3.5
3	0	0.580	3.0	4	0	0.739	4.0
3	0	0.453	2.0	4	0	0.641	3.5
3	1	0.547	3.0	4	0	0.791	4.0
3	1	0.513	3.0	4	0	0.768	4.0
3	1	0.344	2.0	4	0	0.606	3.0
3	0	0.403	2.0	4	1	0.732	4.0

Grupa	Płeć	Punkty	Ocena	Grupa	Płeć	Punkty	Ocena
5	0	0.723	4.0	6	0	0.813	4.5
5	0	0.526	3.0	6	1	0.428	2.0
5	0	0.585	3.0	6	0	0.588	3.0
5	1	0.795	4.0	6	0	0.631	3.5
5	1	0.600	3.0	6	0	0.642	3.5
5	0	0.679	3.5	6	1	0.686	3.5
5	1	0.616	3.5	6	0	0.695	3.5
5	0	0.511	3.0	6	0	0.751	4.0
5	0	0.603	3.0	6	1	0.685	3.5
5	1	0.727	4.0	6	0	0.608	3.0
5	1	0.734	4.0	6	0	0.464	2.0
5	1	0.672	3.5	6	0	0.803	4.0
5	0	0.525	3.0	6	0	0.747	4.0
5	1	0.612	3.5	6	0	0.616	3.5
5	0	0.730	4.0	6	0	0.611	3.5
5	1	0.711	4.0	6	1	0.363	2.0
5	0	0.789	4.0	6	1	0.635	3.5
5	0	0.449	2.0	6	0	0.383	2.0
5	0	0.464	2.0	6	1	0.659	3.5
5	0	0.523	3.0	6	0	0.704	3.5
5	0	0.627	3.5	6	0	0.654	3.5
5	0	0.480	2.0	6	0	0.696	3.5
5	1	0.649	3.5	6	1	0.652	3.5
5	1	0.615	3.5	6	1	0.824	4.5
5	0	0.430	2.0	6	1	0.767	4.0
5	1	0.625	3.5	6	1	0.604	3.0
5	1	0.530	3.0	6	0	0.609	3.0
5	0	0.507	2.0	6	0	0.760	4.0
5	0	0.472	2.0	6	1	0.799	4.0
5	0	0.691	3.5	6	0	0.691	3.5

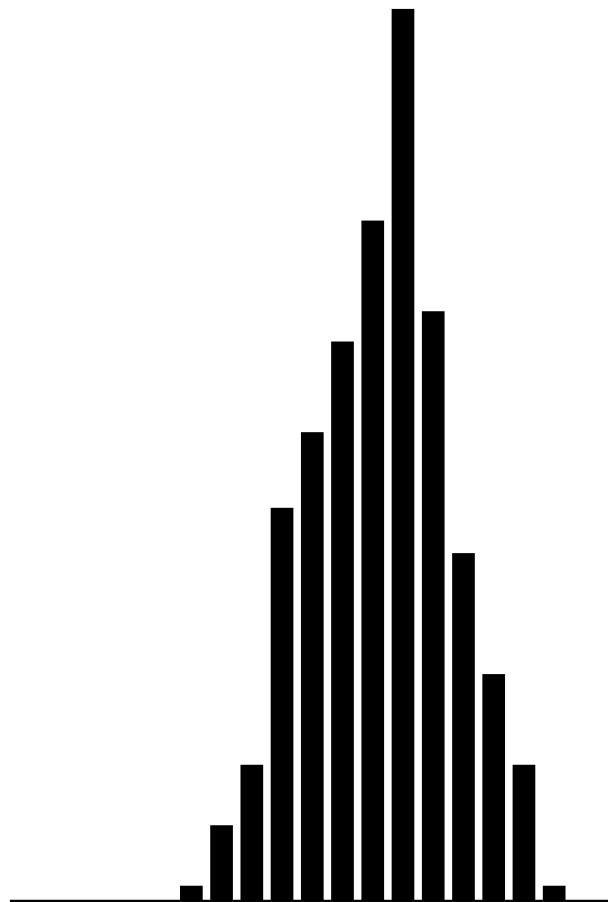
Grupa	Płeć	Punkty	Ocena	Grupa	Płeć	Punkty	Ocena
7	0	0.334	2.0	8	0	0.605	3.0
7	0	0.496	2.0	8	0	0.588	3.0
7	0	0.750	4.0	8	0	0.715	4.0
7	1	0.500	2.0	8	1	0.696	3.5
7	1	0.520	3.0	8	1	0.848	4.5
7	0	0.640	3.5	8	0	0.688	3.5
7	0	0.567	3.0	8	0	0.537	3.0
7	1	0.433	2.0	8	1	0.472	2.0
7	1	0.550	3.0	8	1	0.523	3.0
7	0	0.469	2.0	8	1	0.546	3.0
7	1	0.608	3.0	8	0	0.617	3.5
7	1	0.583	3.0	8	0	0.687	3.5
7	1	0.494	2.0	8	1	0.682	3.5
7	1	0.836	4.5	8	1	0.753	4.0
7	1	0.422	2.0	8	0	0.705	3.5
7	0	0.599	3.0	8	0	0.775	4.0
7	0	0.487	2.0	8	1	0.719	4.0
7	0	0.563	3.0	8	0	0.658	3.5
7	0	0.648	3.5	8	1	0.477	2.0
7	0	0.559	3.0	8	1	0.754	4.0
7	0	0.628	3.5	8	0	0.555	3.0
7	0	0.672	3.5	8	0	0.540	3.0
7	1	0.499	2.0	8	1	0.638	3.5
7	1	0.612	3.5	8	1	0.794	4.0
7	0	0.533	3.0	8	0	0.515	3.0
7	0	0.645	3.5	8	0	0.579	3.0
7	0	0.419	2.0	8	0	0.763	4.0
7	1	0.496	2.0	8	0	0.555	3.0
7	0	0.556	3.0	8	0	0.661	3.5
7	0	0.404	2.0	8	0	0.489	2.0

Grupa	Płeć	Punkty	Ocena	Grupa	Płeć	Punkty	Ocena
9	1	0.725	4.0	10	1	0.431	2.0
9	1	0.629	3.5	10	0	0.503	2.0
9	0	0.853	4.5	10	0	0.474	2.0
9	0	0.577	3.0	10	0	0.615	3.5
9	1	0.441	2.0	10	1	0.466	2.0
9	0	0.478	2.0	10	0	0.428	2.0
9	1	0.804	4.0	10	1	0.588	3.0
9	1	0.687	3.5	10	1	0.603	3.0
9	1	0.438	2.0	10	1	0.622	3.5
9	1	0.573	3.0	10	0	0.494	2.0
9	0	0.620	3.5	10	1	0.474	2.0
9	0	0.801	4.0	10	0	0.506	2.0
9	1	0.670	3.5	10	1	0.508	2.0
9	1	0.704	3.5	10	1	0.673	3.5
9	0	0.694	3.5	10	1	0.582	3.0
9	1	0.443	2.0	10	0	0.544	3.0
9	0	0.614	3.5	10	0	0.643	3.5
9	1	0.614	3.5	10	0	0.559	3.0
9	0	0.554	3.0	10	0	0.592	3.0
9	0	0.642	3.5	10	0	0.545	3.0
9	0	0.677	3.5	10	1	0.561	3.0
9	1	0.401	2.0	10	1	0.449	2.0
9	0	0.521	3.0	10	1	0.450	2.0
9	0	0.633	3.5	10	0	0.695	3.5
9	1	0.428	2.0	10	0	0.457	2.0
9	1	0.850	4.5	10	0	0.614	3.5
9	1	0.720	4.0	10	1	0.558	3.0
9	0	0.443	2.0	10	1	0.677	3.5
9	0	0.713	4.0	10	1	0.734	4.0
9	0	0.601	3.0	10	1	0.476	2.0

Punkty z klasówki

0.00:	0
0.05:	0
0.10:	0
0.15:	0
0.20:	0
0.25:	0
0.30:	1
0.35:	5
0.40:	9
0.45:	26
0.50:	31
0.55:	37
0.60:	45
0.65:	59
0.70:	39
0.75:	23
0.80:	15
0.85:	9
0.90:	1
0.95:	0
1.00:	0

Średnia	0.589
Kwartył dolny	0.508
Mediana	0.595
Kwartył górny	0.672



Stopnie z klasówki

2.0: 76

3.0: 88

3.5: 92

4.0: 37

4.5: 7

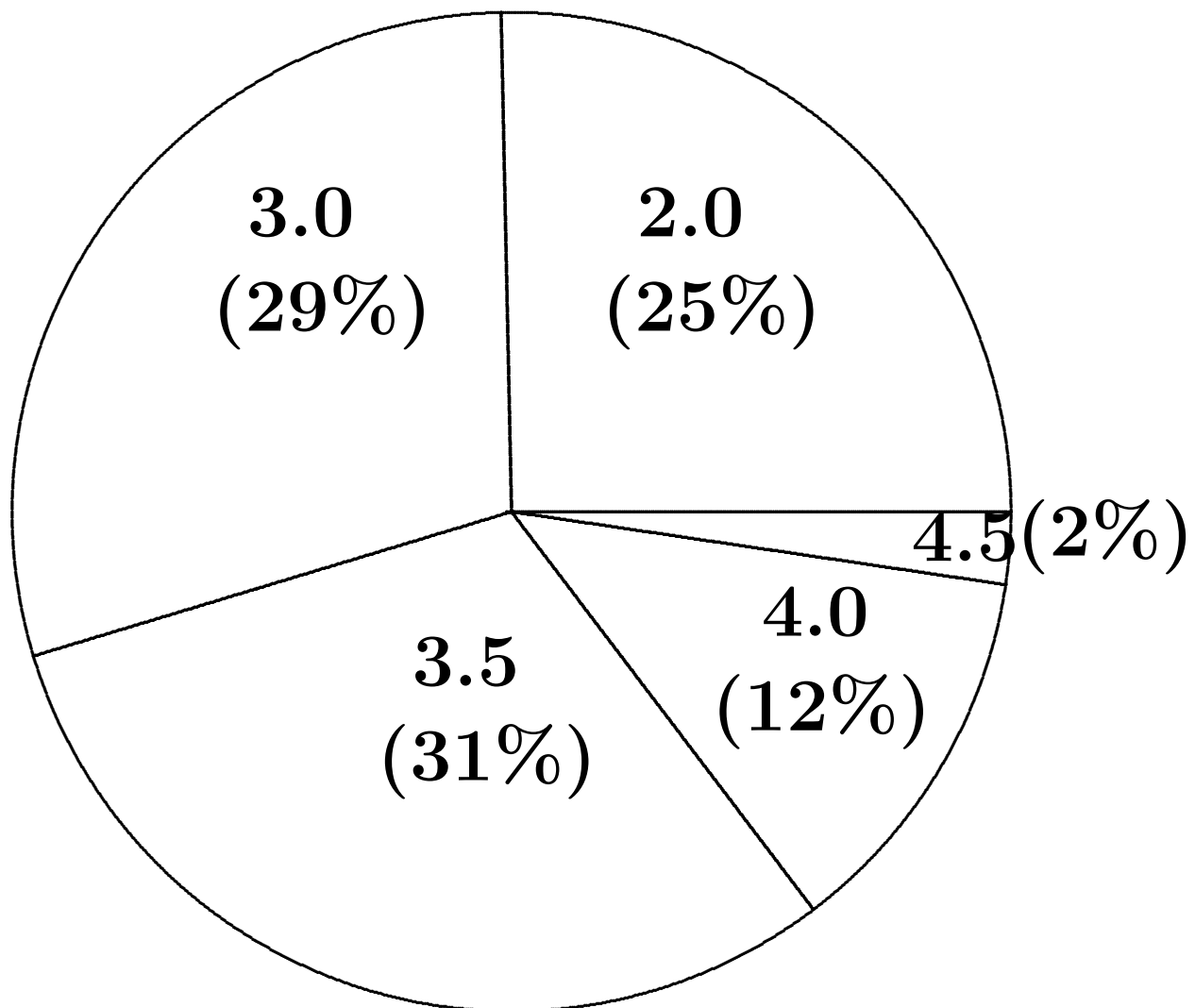
5.0: 0

Średnia 3.06

Kwartył dolny 2.00

Mediana 3.00

Kwartył górny 3.50

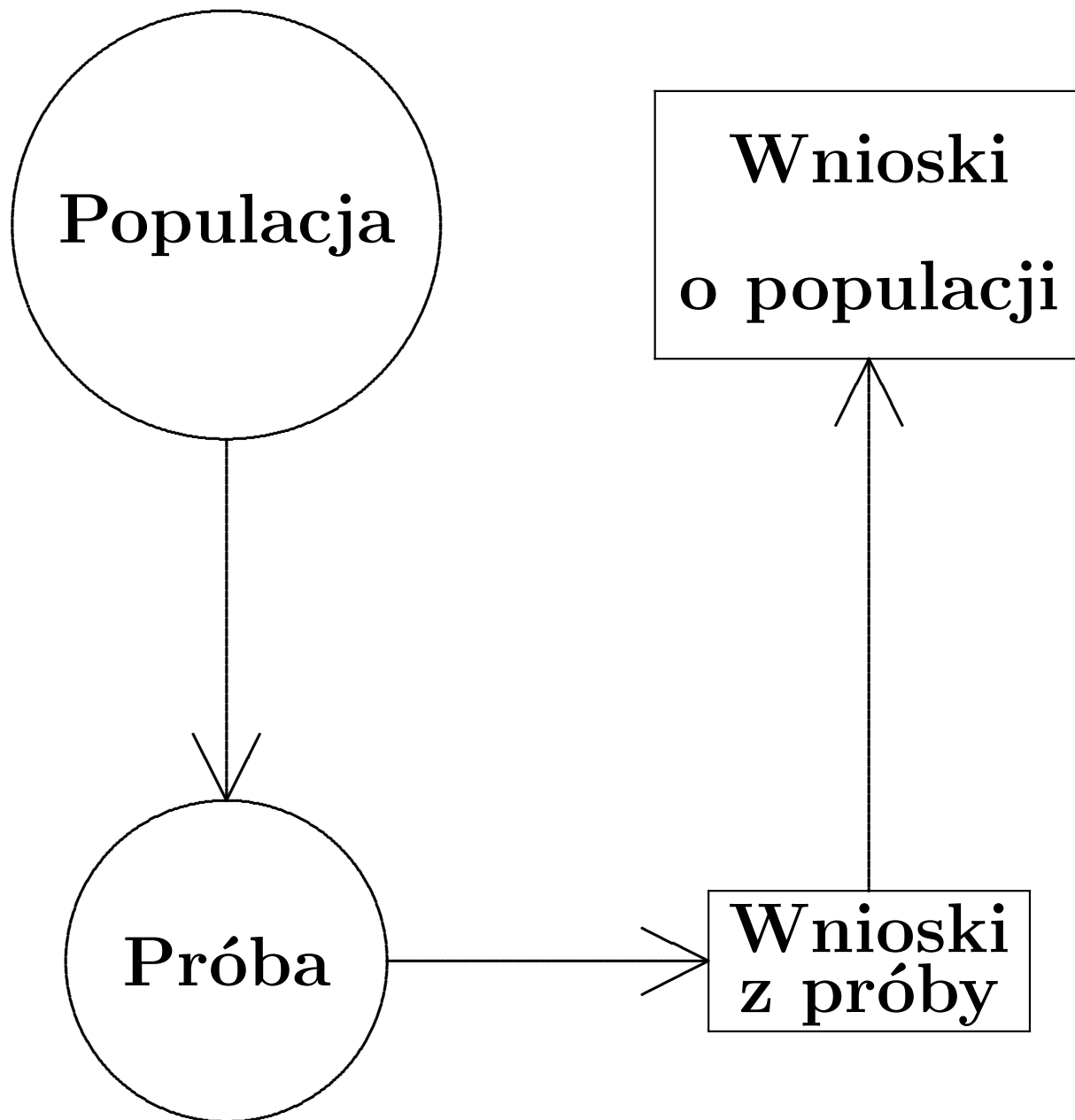


Oceny z klasówki

	Negatywne	Pozytywne	Razem
Ogółem	76 (25%)	224 (75%)	300
Kobiety	46 (28%)	116 (72%)	162
Mężczyźni	30 (22%)	108 (78%)	138
Grupa 1	3 (10%)	27 (90%)	30
Grupa 2	11 (37%)	19 (63%)	30
Grupa 3	17 (57%)	13 (43%)	30
Grupa 4	0 (0%)	30 (100%)	30
Grupa 5	6 (20%)	24 (80%)	30
Grupa 6	4 (13%)	26 (87%)	30
Grupa 7	12 (40%)	18 (60%)	30
Grupa 8	3 (10%)	27 (90%)	30
Grupa 9	7 (23%)	23 (77%)	30
Grupa 10	13 (43%)	17 (57%)	30

Pytania

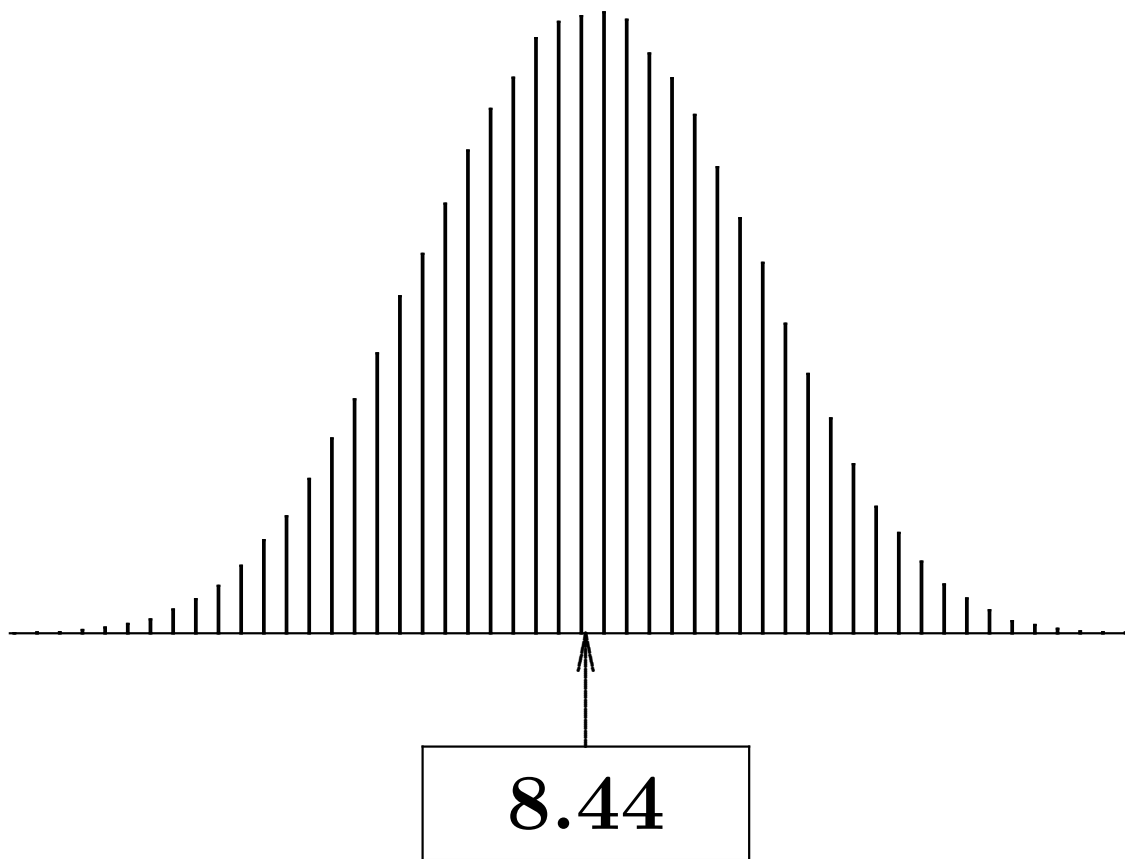
1. Na ile „dokładne” są podane wyniki?
2. Na ile wyniki odzwierciedlają stan wiedzy?
3. Jakich wyników można oczekiwać na następnej klasówce?
4. Jakich wyników można oczekiwać na egzaminie?



F 5 27	F 2 29	F 4 12	F 1 8	F 5 33
M 8 40	M 8 52	M 6 33	F 6 38	F 8 22
M 9 35	M 10 73	F 7 30	M 11 50	M 9 67
M 14 68	M 12 75	F 8 40	M 14 64	M 11 69
F 9 54	F 8 40	M 11 51	M 10 55	M 15 66

Próba 1: 5 2 4 1 5 Średnia z próby: 3.40
 Próba 2: 8 8 6 6 8 Średnia z próby: 7.20
 Próba 3: 9 10 7 11 9 Średnia z próby: 9.20
 Próba 4: 14 12 8 14 11 Średnia z próby: 11.80
 Próba 5: 9 8 11 10 15 Średnia z próby: 10.60

Średnia populacji: 8.44



Pytania

Czy mając do dyspozycji tylko jedną próbę można ocenić na ile dobrze średnia z tej próby przybliża prawdziwą średnią?

Co zrobić, by być „pewniejszym” wyniku?

Populacja

Zbiór obiektów z wyróżnioną cechą (cechami)

Próba

Wybrana część populacji podlegająca badaniu

Cecha

Wielkość losowa charakteryzująca obiekty danej populacji

Cecha jakościowa

Cecha przyjmująca wartości nie będące liczbami (np. *kolor, płeć, smakowitość*)

Cecha (ilościowa) skokowa

Cecha przyjmująca pewne wartości liczbowe i nie przyjmująca wartości pośrednich (np. *ilość bakterii, ilość pracowników, ilość pasażerów*). Cechy te nazywane są również **dyskretnymi**.

Cecha (ilościowa) ciągła

Cecha przyjmująca wartości z pewnego przedziału liczbowego (np. *wzrost, waga, plon*)

Jakość wnioskowania statystycznego

Oceniamy parametr θ cechy na podstawie próby X_1, X_2, \dots, X_n . Niech $\hat{\theta}(X_1, X_2, \dots, X_n)$ będzie „jakaś” oceną parametru θ

Nieobciążoność

Jeżeli średnia wartość oceny $\hat{\theta}$ jest równa wartości parametru θ , to ocenę $\hat{\theta}$ nazywamy nieobciążoną

Minimalna wariancja

Z dwóch różnych nieobciążonych ocen $\hat{\theta}$ oraz $\hat{\hat{\theta}}$ tego samego parametru θ za lepszą uznajemy tę, która „średnio” przyjmuje wartości bliższe parametrowi θ

Minimalny błąd średniokwadratowy

Jeżeli ocena $\hat{\theta}$ nie jest nieobciążona, to wówczas jako miernik jakości stosuje się błąd średniokwadratowy. Jest to „uśrednienie” obciążenia oraz wariancji