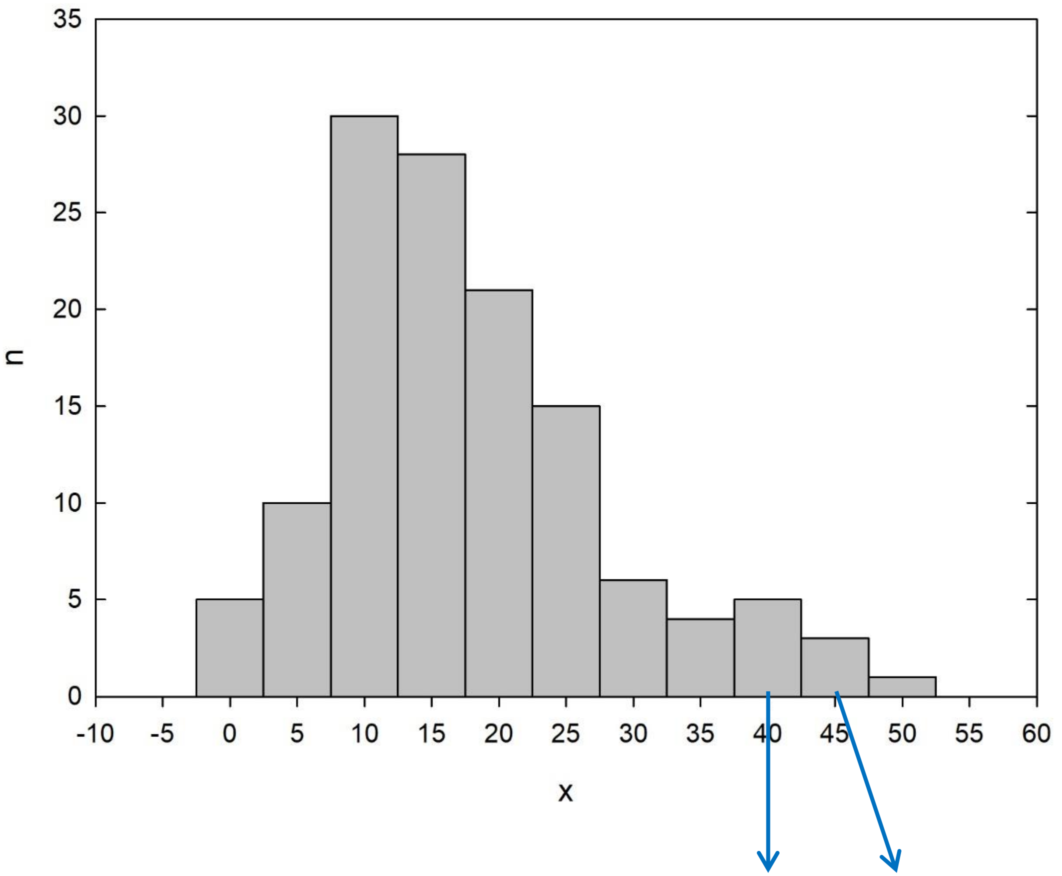
Histogram

- **histogram** – způsob, jak zjistit hustotu pravděpodobnosti z experimentálních dat



šířka binu: $\Delta_i = x_{i+1} - x_i$

plocha histogramu: $\sum_{i=1}^{m} n_i \Delta_i$

normalizovaný histogram: $n_i \rightarrow x_i$

$$\xi_i = \frac{n_i}{\Delta_i N} \qquad N = \sum_{i=1}^{m} n_i$$

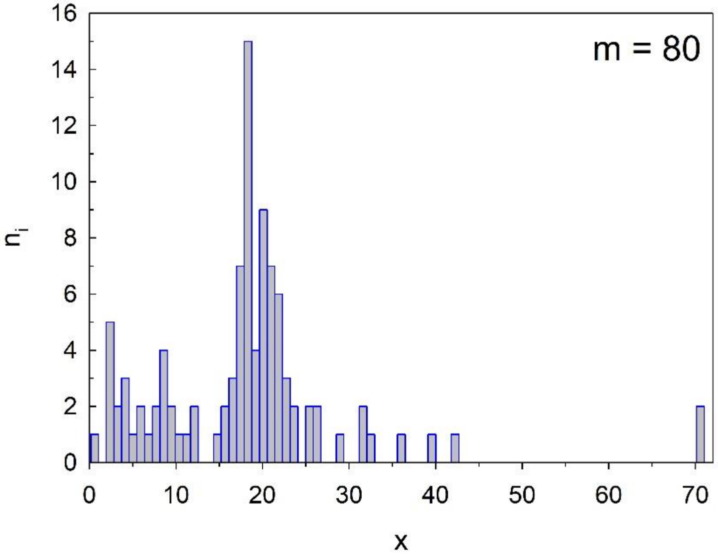
plocha normalizovaného histogramu:

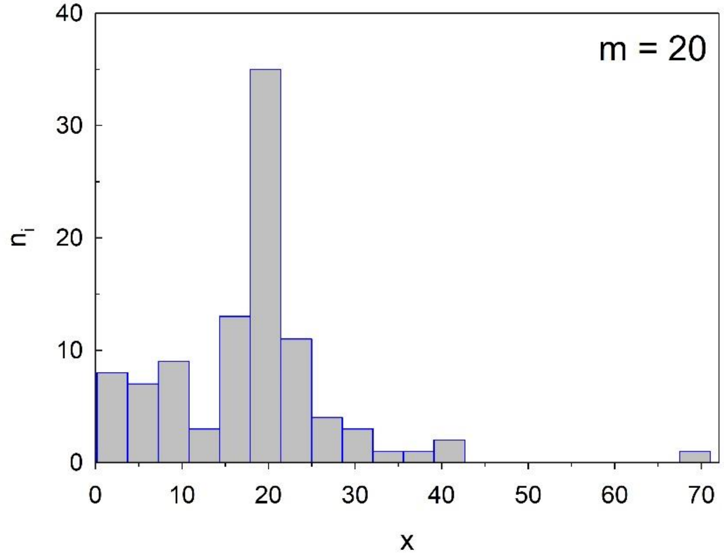
$$\sum_{i=1}^{m} \xi_i \Delta_i = 1$$

hustota pravděpodobnosti:

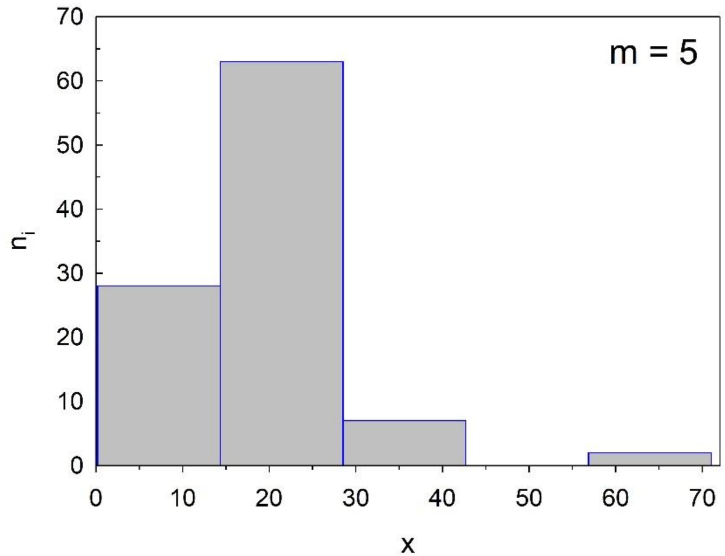
$$f(x) = \lim_{\substack{\Delta_i \to 0 \\ N \to \infty}} \xi_i = \lim_{\substack{\Delta_i \to 0 \\ N \to \infty}} \frac{n_i}{\Delta_i N}$$

Histogram – šířka binu





šířka binu: $$\Delta = \frac{x_{\max} - x_{\min}}{m}$$



H. A. Sturges, J. American Statistical Association, 65–66 (1926).

$$m_{opt} = \frac{\log N}{\log 2} + 1$$

$$N = 100$$

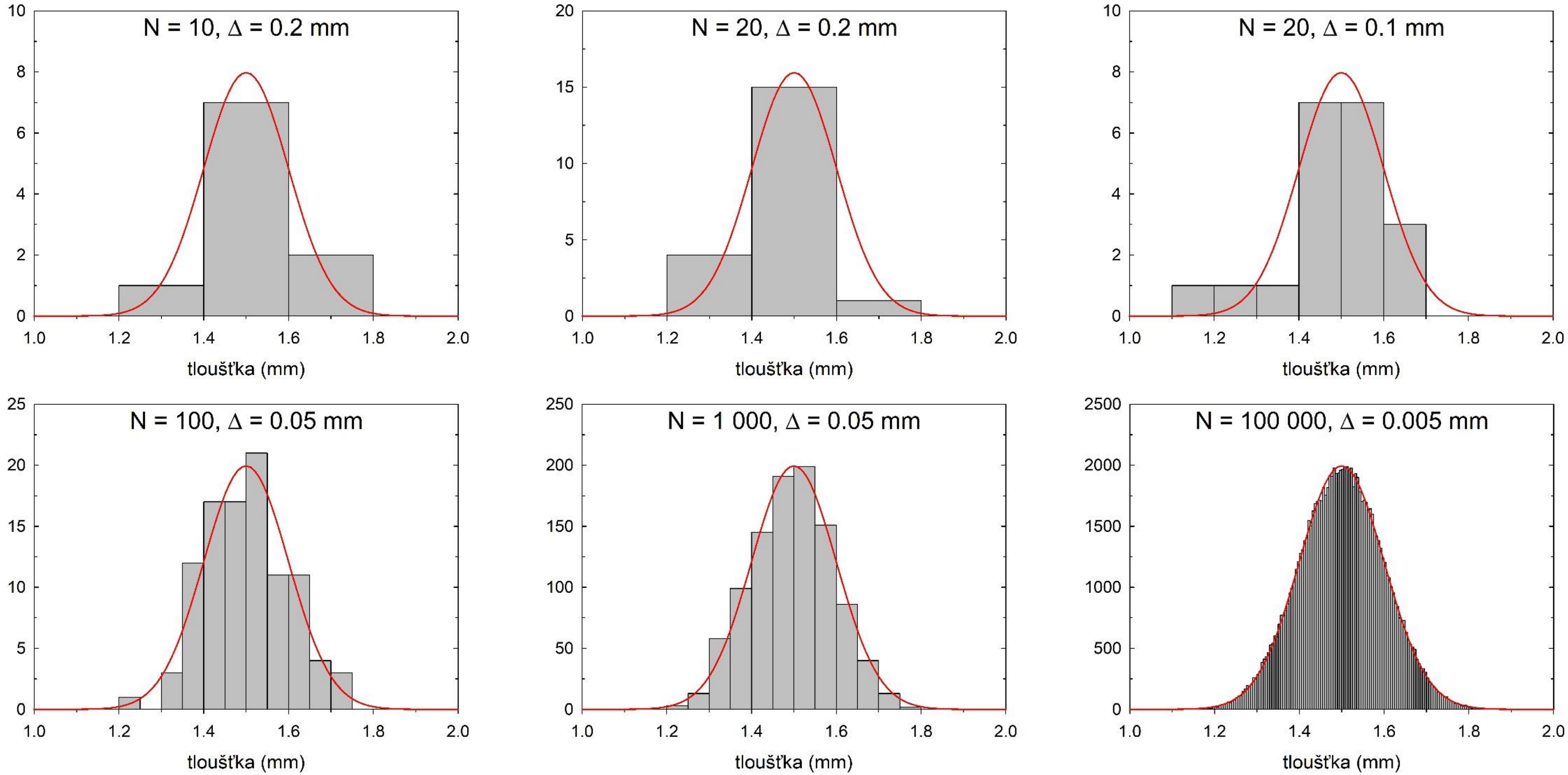
		26.23743
16.17759	21.52645	2.10586
22.36831	18.75516	8.60201
3.29369	41.95208	3.18462
17.96900	19.23135	15.66299
18.52658	8.88075	21.10663
17.63568	32.60371	2.28124
17.79473	4.27135	16.14332
39.80907	18.43469	35.88762
18.25682	23.99716	28.72841
20.63264	18.94920	0.17358
25.89910	8.22661	31.91945
17.57289	17.88642	70.80681
18.74632	17.96704	9.47664
8.46536	20.07927	23.20253
21.63599	7.04639	6.16414
31.43157	12.39286	15.65710
2.71104	18.06331	7.47195
9.89574	17.36080	20.18533
18.16503	17.95492	1.98676
20.18927	7.71726	17.71942
11.27086	20.49528	21.70207
2.49163	21.00411	21.28737
11.77613	25.37069	16.99344
0.25810	21.77872	18.19663
4.53349	24.99534	19.87326
21.22557	21.43774	5.84716
20.04356	10.56477	71.01371
18.79175	4.50194	18.09185
20.86614	23.01736	21.75327
17.80408	20.48741	17.09857
18.29748	20.42592	15.19833
18.08830	20.22713	19.04226
4.65786	21.35032	22.89348

Hustota pravděpodobnosti – měření tloušťky vzorku

- hustota pravděpodobnosti
- počet naměřených hodnot N
- šířka binu Δ

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad \mu = 1.5\ \mathrm{mm},\ \sigma = 0.1\ \mathrm{mm}$$

$$f(x_i) \approx \frac{n_i}{N\Delta} \qquad N \to \infty, \Delta \to 0$$



Momenty

- operátor **střední (očekávané) hodnoty** $\mu = E[x]$

  diskrétní náhodná proměnná: $\mu = E[x] = \sum_i x_i P_i$

  spojitá náhodná proměnná: $\mu = E[x] = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x$

- operátor **rozptylu (variance)** $\sigma^2 = V[x]$

  standardní odchylka: $\sigma = \sqrt{V[x]}$

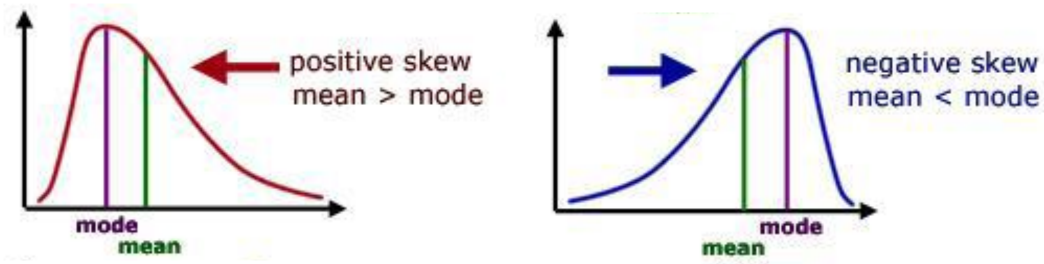
  druhý centrální moment: $$\sigma^2 = V[x] = \mu_2' = E[(x-\mu)^2] = E[x^2] - (E[x])^2$$

- **n-tý moment** $\mu_n = E[x^n]$ $\mu_1 = \mu$
- **n-tý centrální moment** $\mu_n' = E[(x-\mu)^n]$ $\mu_1' = 0$ $\mu_2' = \sigma^2$

Momenty vyšších řádů

- **šikmost** (skewness)

$$\gamma_3 = \frac{\mu_3'}{\sigma^3} = \frac{E[(x-\mu)^3]}{(E[(x-\mu)^2])^{\frac{3}{2}}}$$



- **dodatečná špičatost** (excess kurtosis)

$$\gamma_4 - 3 = \frac{\mu_4'}{\sigma^4} - 3 = \frac{E[(x-\mu)^4]}{(E[(x-\mu)^2])^2} - 3$$