Istogrammi e distribuzioni (1 di 8)

Di seguito sono riportate $N = 32$ osservazioni x_i ($i = 1, 2, ..., N$) di una grandezza fisica. Il valore della grandezza è espresso con 1 cifra decimale (dipende dalla risoluzione dello strumento di misura).

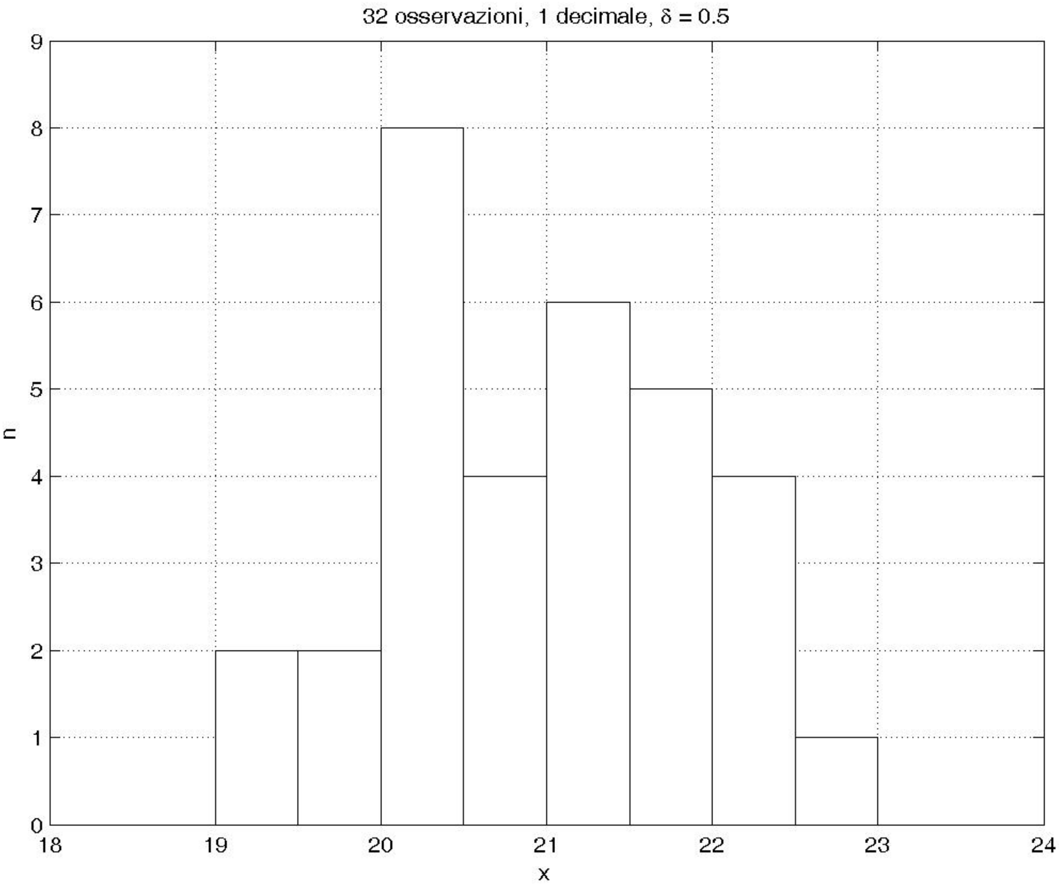
21.1	21.9	21.8	20.8	20.9	21.2	20.3	20.8
21.0	20.0	22.0	19.9	20.3	19.9	22.1	20.4
19.3	21.9	22.5	21.5	21.2	20.2	20.3	22.3
22.2	20.4	21.9	21.3	20.7	20.2	21.3	19.2

Riordiniamo le x_i

19.2	19.3	19.9	19.9	20.0	20.2	20.2	20.3
20.3	20.3	20.4	20.4	20.7	20.8	20.8	20.9
21.0	21.1	21.2	21.2	21.3	21.3	21.5	21.8
21.9	21.9	21.9	22.0	22.1	22.2	22.3	22.5

e contiamo le n_j ($j = 1, 2, ...M$) osservazioni che cadono negli intervalli di ampiezza $\delta = 0.5$ e di estremi 19.0, 19.5, 20.0, 20.5, 21.0, 21.5, 22.0, 22.5, 23.0 ($M = 8$ intervalli).

Istogrammi e distribuzioni (2 di 8)



Gli intervalli sono centrati in x_{cj}

j	**DA (≥)**	**A (<)**	n_j	x_{cj}	f_j
1	19.0	19.5	2	19.25	0.125
2	19.5	20.0	2	19.75	0.125
3	20.0	20.5	8	20.25	0.5
4	20.5	21.0	4	20.75	0.25
5	21.0	21.5	6	21.25	0.375
6	21.5	22.0	5	21.75	0.3125
7	22.0	22.5	4	22.25	0.25
8	22.5	23.0	1	22.75	0.0625

Istogrammi e distribuzioni (3 di 8)

Il rapporto $p_j = n_j / N$ stima la *probabilità* che un elemento x_i cada nell'intervallo centrato in x_{cj}, di ampiezza δ

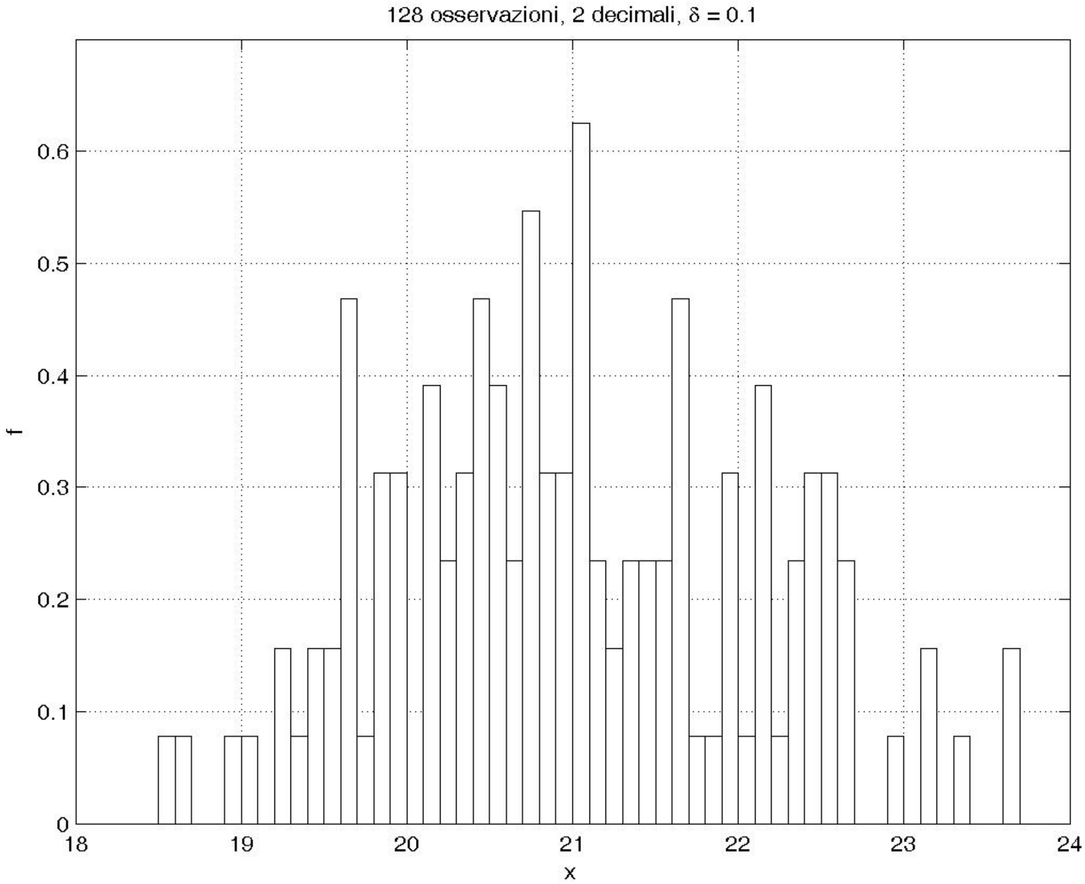
Si definisce *densità di probabilità* f_j la seguente grandezza

$$f_j = \frac{n_j}{N \cdot \delta} = \frac{p_j}{\delta} \qquad j = 1, 2, \ldots M$$

Il diagramma a barre (*istogramma*) può essere scalato in termini di f anziché di n, basta moltiplicare la scala verticale per $1/(N \cdot \delta)$ (cioè 1/16 nell'esempio)

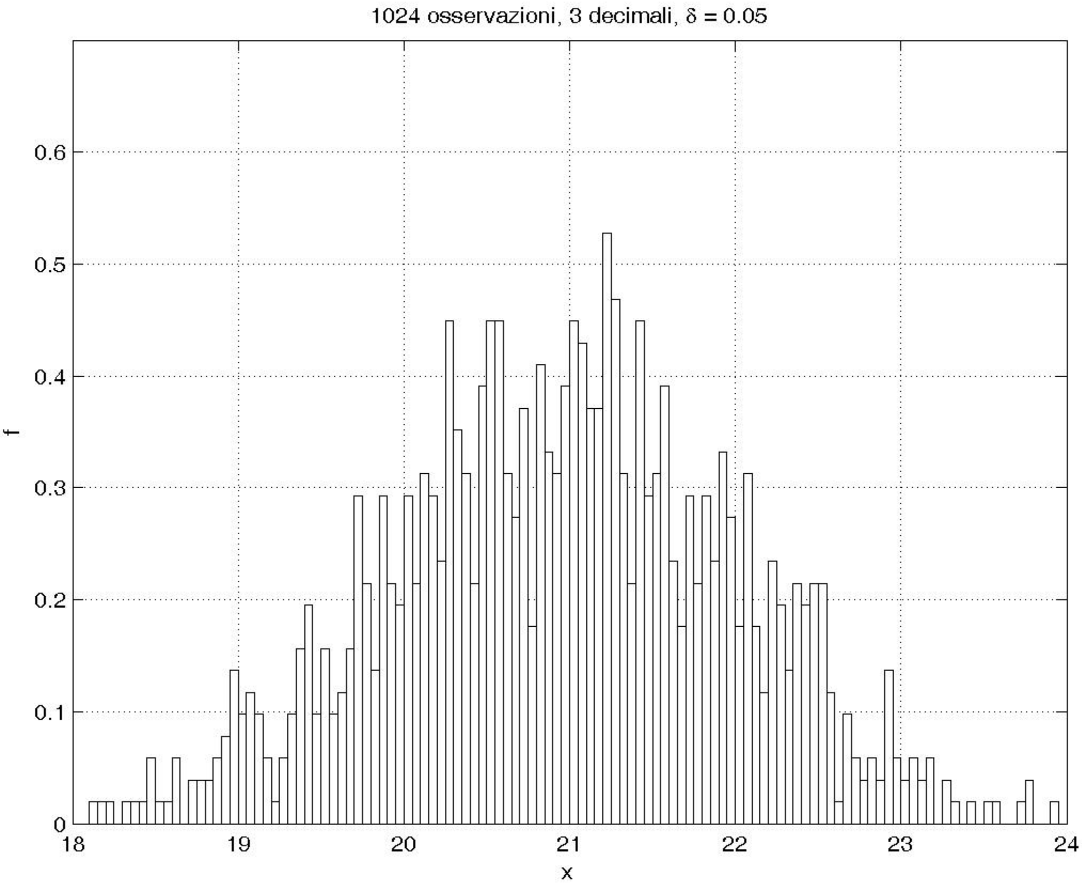
Aumentando il numero N di osservazioni, diminuendo l'ampiezza δ degli intervalli e disponendo di un numero di cifre decimali crescente si può far tendere l'istogramma ad una curva continua detta *densità di probabilità* o *distribuzione di probabilità* (ordinata espressa in termini di f)

Istogrammi e distribuzioni (4 di 8)



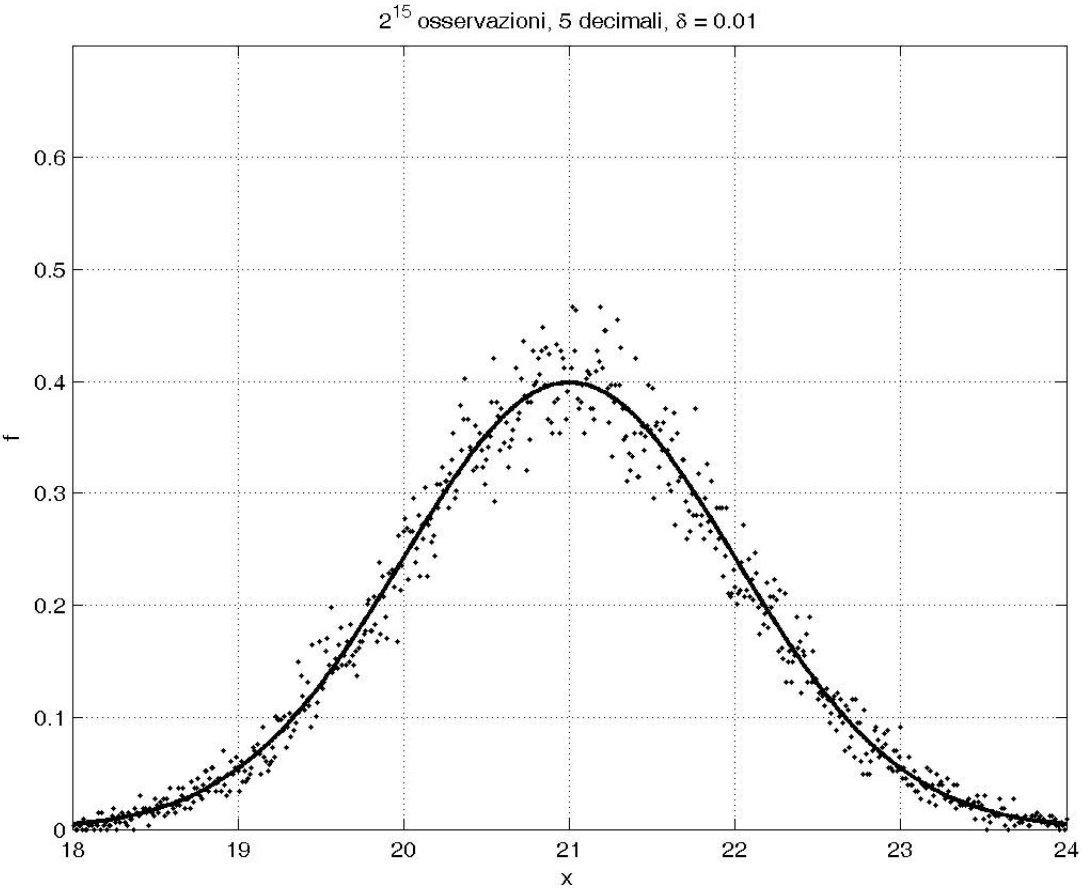
$N = 128$, 2 cifre decimali, $\delta = 0.1$

Istogrammi e distribuzioni (5 di 8)



$N = 1024$, 3 cifre decimali, $\delta = 0.05$

Istogrammi e distribuzioni (6 di 8)



$$N = 2^{15} = 32\,768,\ 5 \text{ cifre decimali},\quad \delta = 0.01$$

Per chiarezza di rappresentazione si sono usati dei punti al posto delle barre. La curva continua rappresenta il limite per numero infinito di osservazioni, numero illimitato di decimali (infinito potere risolutivo), ampiezza infinitesima degli intervalli. Nell'esempio la curva limite è la densità di probabilità *normale* o *gaussiana* $f(x)$ di parametri x_μ e σ (con $x_\mu = 21$ e $\sigma = 1$)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(x - x_\mu\right)^2}{2\sigma^2} \right)$$

Istogrammi e distribuzioni (7 di 8)

Torniamo al caso di $N = 32$ osservazioni

Media

$$x_m = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Nell'esempio $x_m = 20.9625$, approssimando

$$x_m \approx \frac{1}{N}\sum_{j=1}^{M} x_{cj} n_j = \sum_{j=1}^{M} x_{cj} p_j = \sum_{j=1}^{M} x_{cj} f_j \delta$$

Con l'approssimazione si trova $x_m \approx 20.96875$. Passando al limite

$$x_\mu = \int_{-\infty}^{+\infty} x\, f(x)\,\mathrm{d}x$$

Si osservi poi che $\frac{1}{N}\sum_{j=1}^{M} n_j = \sum_{j=1}^{M} p_j = \sum_{j=1}^{M} f_j \delta = 1$, quindi passando al limite

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$$

La probabilità che x si compresa fra x_A e x_B è

$$\mathrm{Prob}\{x_A \le x \le x_B\} = \int_{x_A}^{x_B} f(x)\,\mathrm{d}x$$

x_A può essere eventualmente $-\infty$, e x_B $+\infty$.

Istogrammi e distribuzioni (8 di 8)

Varianza (scarto tipo) e valore quadratico medio

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - x_m\right)^2$$

Nell'esempio $s^2 = 0.785$. Approssimando

$$s^2 \approx \sum_{j=1}^{M}\left(x_j - x_m\right)^2 f_j \delta$$

Approssimando $s^2 \approx 0.827$. Sia con le formule esatte che con le formule approssimate si ottiene per media $x_m = 21.0$ e scarto tipo della media $s_m = 0.2$. Passando al limite si ottiene la *varianza* σ^2

$$\sigma^2 = \int_{-\infty}^{+\infty}\left(x - x_\mu\right)^2 f\left(x\right)\mathrm{d}x$$

Si definisce *valore quadratico medio* x_{rms}^2 la seguente quantità

$$x_{rms}^2 = \int_{-\infty}^{+\infty} x^2 f\left(x\right)\mathrm{d}x$$

Vale che $x_{rms}^2 = x_\mu^2 + \sigma^2$. Se $x\left(t\right)$ è un segnale che varia casualmente nel tempo t allora

$$x_{eff}^2 = \lim_{T\to\infty}\frac{1}{T}\int_{\tau}^{\tau+T} x^2\left(t\right)\mathrm{d}t$$

x_{eff} è il *valore efficace* del segnale. Vale che $x_{rms} = x_{eff}$.