VARIANZA DI UNA V.A.

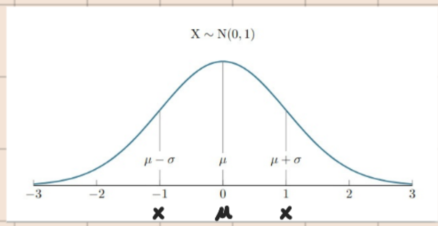
DEF: La varianza di una v.a. X è definita come:

$$\boxed{VAR(x) = E[(X - E[X])^2]} \quad , \text{ notazione: } \boxed{E[X] = \mu_x \in \mathbb{R}}$$

L'aspettazione non è altro che un numero.

Quindi riscrivendo:

$$VAR(x) = E[(X - \mu_x)^2] \begin{cases} \overset{\text{continua}}{=} \boxed{\int_{-\infty}^{+\infty} (x - \mu_x)^2 \cdot f_x(x) \cdot dx} \\ \overset{\text{discreta}}{=} \boxed{\sum_{a \in S_x} (a - \mu_x)^2 \cdot P_x(a)} \end{cases}$$



$$\underset{(\mu - x)}{x} \quad \mu \quad \underset{(x - \mu)}{x}$$

TEO: La varianza può essere calcolata come:

$$VAR(x) = \dots = E[X^2] - (\mu_x)^2 = \boxed{E[X^2] - (E[X])^2}.$$

DIM: sul quaderno.

Osservazione sulla positività:

$$\begin{cases} VAR(x) = 0 \iff X = \mu_x \iff \text{Distribuzione gaussiana} \\ VAR(x) \geq 0 \implies \text{altrimenti.} \end{cases}$$

X	μ
P_x	1

TEO: $\forall$ v.a. X e $\forall r, s \in \mathbb{R}$:

$$\boxed{E[r \cdot X + s] = \dots = r \cdot E[X] + s}$$

$$\boxed{VAR(r \cdot X + s) = \dots = r^2 \cdot VAR(X)}$$

DIM: segue dal teorema del cambio di variabile sul quaderno.

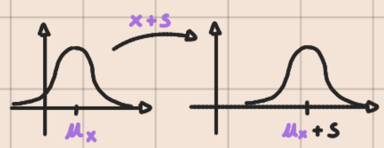
VARIABILE ALEATORIA STANDARDIZZATA

DEF: Siano: $E[X] = \mu_x$, $VAR(x) = \sigma^2 \implies \boxed{Z = \frac{X - \mu_x}{\sigma}}$ Var. alea. standardiz.

$$\implies E[Z] = E\left[\frac{X - \mu_x}{\sigma}\right] = \frac{1}{\sigma} \cdot \underbrace{E[X]}_{\mu_x} - \frac{\mu_x}{\sigma} = \boxed{0} \quad \text{MEDIA V.A.S.}$$

$$\implies VAR(Z) = VAR\left(\frac{X - \mu_x}{\sigma}\right) = \frac{1}{\sigma^2} \cdot \underbrace{VAR(x)}_{\sigma^2} = \boxed{1}. \quad \text{VARIANZA V.A.S}$$

- Si tratta di una definizione generale per tutte le variabili aleatorie.
- Non stiamo dicendo nulla riguardo alla distribuzione.



S implica una traslazione.

Esempio Varianza con Bernoulli

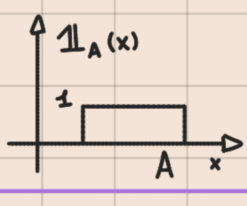
$X \sim Ber(p)$

$\boxed{\begin{array}{l}\text{Bernoulli} \\ E[x] = p\end{array}}$

- $E[x^2] = \sum_{a \in \{0,1\}} a^2 \cdot P_x(a) = 0^2 \cdot (1-p) + 1^2 \cdot p = p$
- $VAR(x) = \underbrace{E[x^2]}_{p} - (\underbrace{E[x]}_{p})^2 = p - p^2 = \boxed{p(1-p)}.$

Funzione Caratteristica su insieme A

- $$\mathbb{1}_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$
- $A \subseteq \mathbb{R}$



Deviazione Standard

$$\text{Deviazione standard} = \sqrt{VAR(X)} = \sigma$$

SIGNIFICATO delle FORMULE

Intuizione: Se l'**Aspettazione** (o Valore Atteso) è il "Baricentro" di una distribuzione probabilistica, allora **Varianza** e **Deviazione Standard** servono a dire quanta massa è sparsa o concentrata intorno al baricentro.

- Il **valore atteso** indica dove va messo il fulcro.
- La **varianza** spiega quanto questi pesi tirano verso sinistra o destra.
- La **deviazione standard** è la stessa informazione, ma nella scala originale dei dati.

Significato Formule:

- **Valore atteso:** è la media pesata dei valori. Ti dice dove cade l'equilibrio.

$$E[X] = \boxed{\sum_i x_i \cdot p_i}$$

- **Varianza:** l'aspettazione del quadrato della distanza.

$$Var(X) = E[(X - E[X])^2] = \boxed{\sum_i (x_i - E[X])^2 \cdot p_i}$$

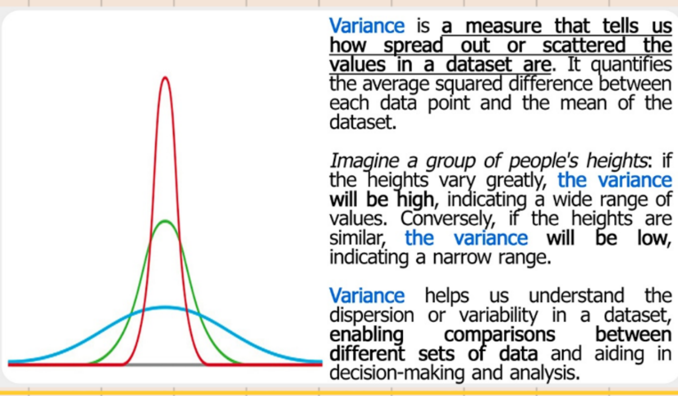
Dentro la formula:

- $x_i - E[X]$: quanto ogni valore si discosta dal baricentro (lo "spostamento").
- $(x_i - E[X])^2$: misura il quadrato dello spostamento:
  - le distanze a sinistra e a destra non si cancellano (eviti somme che fanno 0).
  - penalizzi i valori lontani (grandi scarti pesano molto).
- p_i : probabilità, si sta facendo ancora una media pesata.

- **Deviazione standard:** la varianza è ad unità quadrate (se misuro metri $\longrightarrow$ ottengo metri^2);

$$\boxed{\sigma = \sqrt{VAR(X)}}$$

la deviazione standard riporta tutto all'unità originale (in metri).



Riassunto Concettuale:

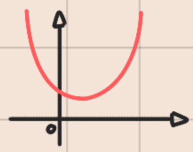
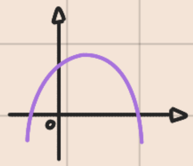
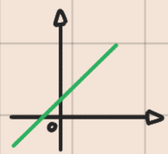
- **Valore atteso** $\longrightarrow$ posizione media / baricentro;
- **Varianza** $\longrightarrow$ sparsezza dei valori (media scarti quadratici);
- **Deviazione standard** $\longrightarrow$ sparpagliamento nella stessa scala dei dati.

DISUGUAGLIANZA di JENSEN

SAPPIAMO CHE: $E[X^2] \geq (E[X])^2$ PERCHÉ $VAR(X) = E[X^2] - (E[X])^2 \geq 0$

ALLORA IN QUALI CASI SI HA CHE: $\boxed{E[g(x)] \substack{\leq \\ \geq \\ =} g(E[X])}$?

SIA $g: S_X \to \mathbb{R}$ UNA FUNZIONE E SIA X UNA V.A. CON SUPPORTO S_X TALE CHE $E[g(x)] < \infty$:

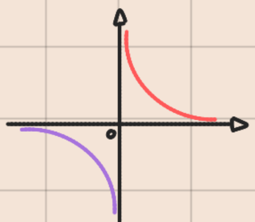
- $g(x)$ CONVESSA: $E[g(x)] \geq g(E[X]) \Rightarrow$  $g''(x) \geq 0 \quad \forall x \in S_X$
- $g(x)$ CONCAVA: $E[g(x)] \leq g(E[X]) \Rightarrow$  $g''(x) \leq 0 \quad \forall x \in S_X$
- $g(x)$ CONCAVA e CONVESSA: $E[g(x)] = g(E[X]) \Leftrightarrow g(x) = \alpha x + \beta$ = RETTA $\Rightarrow$  $g''(x) = 0 \quad \forall x \in S_X$

Esempio Disuguagilianza di Jensen

ESPONENZIALE: $E[X] = \frac{1}{\lambda}$

SIA $X \sim Exp(3)$ E SIA $Y \sim N(1,4)$, COLLOCARE $E[\frac{1}{X}]$, $E[\frac{1}{Y}]$:

- POTREMMO CALCOLARE: $E[\frac{1}{X}] = \int_0^{+\infty} \frac{1}{x} \cdot 3e^{-3x} \cdot dx$ MA È UN INTEGRALE COMPLICATO DA RISOLVERE.
- ALLORA APPLICHIAMO JENSEN: $g(x) = \frac{1}{x}$ $\quad E[\frac{1}{x}] = E[\frac{1}{3}] = 3 \quad \frac{1}{E[X]} = \frac{1}{E[3]} = \frac{1}{\frac{1}{3}} = 3$



$S_X = (0, +\infty) \Rightarrow g(x)$ CONVESSA IN $(0,+\infty) \Rightarrow E[g(x)] \geq g(E[X]) \Rightarrow \boxed{E[\frac{1}{x}] \geq 3}$

$S_X = (-\infty, 0) \Rightarrow g(x)$ CONCAVA IN $(-\infty, 0) \Rightarrow E[g(x)] \leq g(E[X])$

- $Y \sim N(1,4)$, $g(x) = \frac{1}{y}$, $E[\frac{1}{Y}]$ CON $S_Y \subseteq \mathbb{R} \Rightarrow$ NON POSSO APPLICARE JENSEN.