Corso di IDROLOGIA

Estratto dalle dispense di

STATISTICA APPLICATA ALL'IDROLOGIA

Redatte da Prof. Pierluigi Claps

Ing. Chiara Barberis

CONCETTI FONDAMENTALI DELLA TEORIA DELLE PROBABILITA'.

Esperimento aleatorio.

Spazio campionario o popolazione.

Esempi:

Esperimento	Popolazione	Tipo
Numero di giorni piovosi in un anno	$\{0,1,2,.....365\}$	Finito, numero
Numero di giorni non piovosi consecutivi	$\{0,1,2........\}$	Infinito, numero
Valori osservati della portata	$\{x; x \geq 0\}$	Infinito, non numero

- Evento aleatorio semplice A, B, C : ciascun elemento della popolazione (punto).
- Evento aleatorio composto A, B, C : insieme di due o più punti.
- Complemento dell'elemento $A : \overline{A}$: insieme dei punti che non appartengono ad A.
- Evento certo Ω : insieme di tutti i punti della popolazione.
- Evento nullo Φ : insieme vuoto.
- Unione di eventi A e B : $A \bigcup B$: insieme dei punti dei due eventi.
- Intersezione di $A \bigcap C$: insieme dei punti comuni ad A e a B

PROPRIETA' FONDAMENTALI DELLA PROBABILITA'.

Probabilità dell'evento A:

1. $0 \leq P[A] \leq 1$
2. $P[\Omega] = 1$
3. Se $B = A_1 \cup A_2 \cup A_3 \ldots.$ e se $A_1, A_2, A_3 \ldots..$ sono mutuamente escludentisi:

$$P[B] = P[A_1] + P[A_2] + P[A_3] + \ldots.$$

Esempi:

$$P[\overline{A}] = 1 - P[A]$$

$$P[\Phi] = 1 - P[\Omega] = 0$$

PROBABILITA' DELLE UNIONI DI EVENTI.

Esempio: In un istituto universitario vi sono 10 docenti (3 donne e 7 uomini) e 30 non docenti (10 donne e 20 uomini).

Qual è la probabilità che un membro dell'istituto scelto a caso sia un docente e/o una donna?

$$P[D \cup F] = P[D] + P[F] - P[D \cap F]$$

Eventi <u>mutuamente escludentisi:</u> $P[A \cap B] = 0$

PROBABILITA' CONDIZIONATA.

Esempio: Qual è la probabilità che un membro dell'istituto donna sia docente?

$$P[D|F] = \frac{P[D \cap F]}{P[F]}$$

Eventi statisticamente indipendenti: $P[A \cap B] = P[A] \cdot P[B]$

Eventi mutuamente escludentisi: $P[A|B] = 0$

TEOREMA DELLA PROBABILITA' TOTALE.

A evento qualsiasi.

$$B_1 B_2, B_3 \ldots . B_n \begin{cases} \textit{eventi mutuamente escludentisi.} \\ B_1 \cup B_2 \cup \ldots \ldots \cup B_n = \Omega \end{cases}$$

$(B_1 \cap A), (B_2 \cap A), (B_3 \cap A), \ldots ., (B_n \cap A)$ Altra serie di eventi mutuamente escludentisi.

$$(B_1 \cap A) \cup (B_2 \cap A) \cup (B_3 \cap A) \cup \ldots . \cup (B_n \cap A) = A$$

$$P[A] = P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + \ldots \ldots P[A|B_n]P[B_n]$$

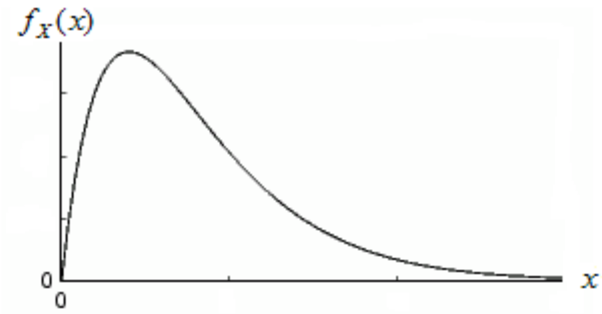
VARIABILI ALEATORIE CONTINUE.

Possono assumere qualsiasi valore numerico reale in un dato intervallo.

Funzione di densità di probabilità.

$$f_X(x) = \lim_{\Delta x \to 0} \frac{P\left[x - \frac{\Delta x}{2} \le X \le x + \frac{\Delta x}{2}\right]}{\Delta x}$$

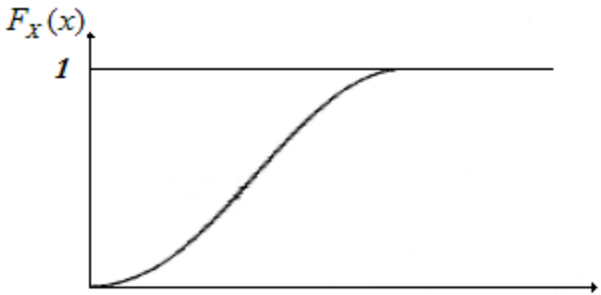
- $f_X(x) \ge 0$ $\qquad \int_{-\infty}^{+\infty} f_X(x)dx = 1$ $\qquad P[a \le X \le b] = \int_a^b f_X(x)dx$



Funzione di <u>distribuzione cumulata:</u>

$$F_X(x) = P[X \le x] = \int_{-\infty}^{+\infty} f_X(u)du$$

- $\dfrac{dF_X(x)}{dx} = f_x(x)$ solo per variabili assolutamente continue.
- $F_X(\infty) = 1;$ $\qquad F_X(-\infty) = 0$
- $F_X(x+\varepsilon) \ge F_x(x)$ per qualsiasi $\varepsilon > 0$; $\qquad F_X(x_2) - F_X(x_1) = P[x_1 \le X \le x_2]$



Per ogni tipo di variabile definita nell'intervallo $[a,b]$:

- $0 \le F_X(x) \le 1;$ $\qquad F_X(a) = 0;$ $\qquad F_X(b) = 1$

MOMENTI

MEDIA (VALORE SPERATO) di una variabile aleatoria discreta.

$$E[X] = \sum_{x_i} x_i P_x(x_i) = \mu$$

di una variabile aleatoria continua.

$$E[X] = \int_{-\infty}^{+\infty} x f_x(x) dx = \mu$$

di una funzione $g(x)$ di una v.a. continua o discreta:

$$E[g(x)] = \sum_{x_i} g(x_i) P_x(x_i)$$

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x) f_x(x) dx$$

r-esimo momento di X:

$$\left.\begin{aligned} \mu_r &= E[X^r] = \sum_{x_i} x_i^r P_x(x_i) \\ \mu_r &= E[X^r] = \int_{-\infty}^{+\infty} x^r f_x(x) dx \end{aligned}\right\} \mu_1 = \mu = E[X]$$

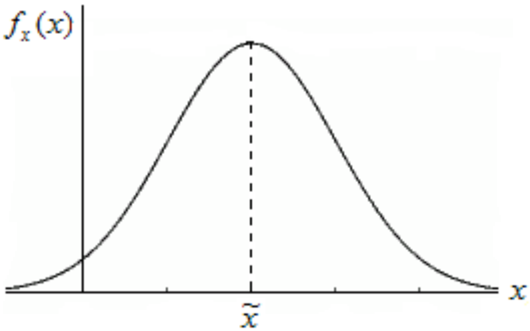
<u>r-esimo momento di centrale di</u> X :

$$\left.\begin{aligned} \mu_r' &= E\left[(X-\mu)^r\right] = \sum_{x_i}(x_i-\mu)^r P_x(x_i) \\ \mu_r' &= E\left[(X-\mu)^r\right] = \int_{-\infty}^{+\infty}(x-\mu)^r f_x'(x)dx \end{aligned}\right\} \mu_1' = 0; \mu_2' = \mathrm{var}[X] = \sigma^2 = E\left[X^2\right] - E^2[X] = \mu_2' - \mu^2$$

MISURA DI LOCAZIONE.

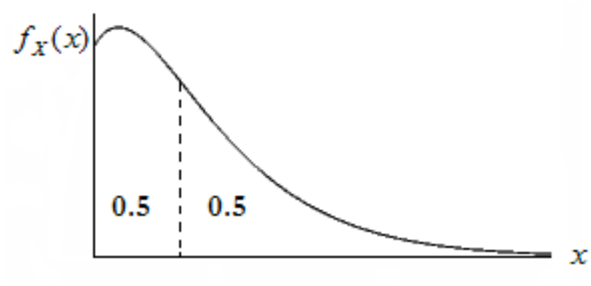
Moda: $\tilde{x}$

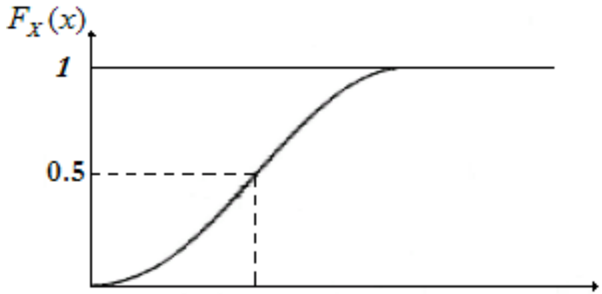
$f_x(\tilde{x}) = \max$



Mediana: $x_{0.5} = \breve{x}$

$F_x(\breve{x}) = 0.50$





Media: $\mu_x = E[x]$

Media geometrica: $M_g = \prod_k x_i k_i$

$\log M_g = E[\log x]$

MISURA DI DISPERSIONE.

Varianza: $\operatorname{var}[x] = E\left[(x-\mu)^2\right] = \sigma^2$

Scarto quadratico medio: $\sigma = \sqrt{\operatorname{var}[x]}$

Coefficiente di variazione: $Cv = \dfrac{\sigma}{\mu} = \gamma$

MISURA DI ASIMMETRIA.

Coefficiente di asimmetria: $\gamma_1 = Ca = \dfrac{\mu_3^{'}}{\sigma^3}$

MISURA DI APPIATTIMENTO O CURTOSI.

Curtosi: $k = \dfrac{\mu_4^{'}}{\sigma^4}$

Coefficiente di eccesso o di Curtosi: $\gamma_2 = k - 3 = \dfrac{\mu_4^{'}}{\sigma^4} - 3$

DISTRIBUZIONE NORMALE DEL CASO O DI GAUSS.

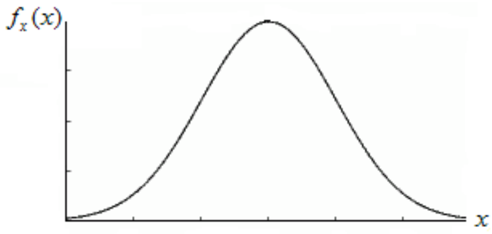
Funzione densità di probabilità:

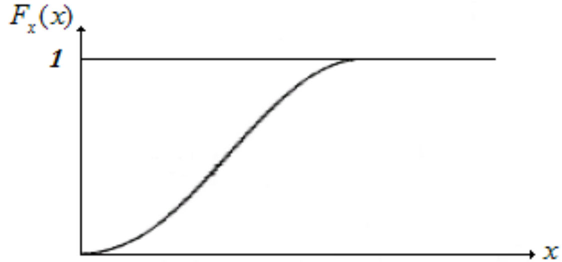
$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(x-\mu)}{\sigma}\right]^2} \qquad -\infty < x < +\infty$$

Funzione di distribuzione cumulata:

$$F_x(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(x-\mu)}{\sigma}\right]^2} dx$$

può essere calcolata numericamente per ogni μ e σ.





I parametri μ e σ.sono dati da:

$E[x] = \mu$ $\qquad$ $\text{var}[x] = \sigma^2$.

DISTRIBUZIONE NORMALE IN FORMA CANONICA

Variabile normale standardizzata o ridotta $u = \frac{(x-\mu)}{\sigma}$

$E[u] = 0$ $\mathrm{var}[u] = 1$

$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$ $F(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$ $F(-u) = 1 - F(u)$

Valori notevoli di u(F) e di F(u)

F(u)	u
0.025	-1.96
0.50	0.00
0.975	+1.96

u	F(u)
-2.0	0.0228
-1.0	0.1587
0.0	0.5000
1.0	0.8413
2.0	0.9772

Esempio: $X \cong N(\mu = 10, \sigma = 3)$

Valore di X corrispondente a $F(x) = 0.025$ ovvero a $T = \frac{1}{F(x)} = 40$

–> $X_{0.025} = \mu - 1.96\,\sigma = 10 - 1.96 \cdot 3 = 4.12$

PROPRIETA' DELLA DISTRIBUZIONE NORMALE

Probabilità che una variabile casuale normale cada in un intervallo $\left[\mu - k\sigma, \mu + k\sigma\right]$:

$$P\left[\mu - k\sigma \le X \le \mu + k\sigma\right] = P\left[-k \le \frac{x-\mu}{\sigma} \le k\right] = \Phi(k) - \Phi(-k) = 1 - \Phi(-k) - \Phi(-k) = 1 - 2\Phi(-k)$$

Coefficiente di asimmetria: $Ca = \dfrac{\mu_3^{'}}{\sigma^3} = \dfrac{E\left[(x-\mu)^3\right]}{\sigma^3} = 0$

Coefficiente di Curtosi (misura dell'appiattimento di una distribuzione): $k = \dfrac{\mu_4^{'}}{\sigma^4} = \dfrac{E\left[(x-\mu)^4\right]}{\sigma^4} = 3$

Somma di variabili normali $(X_i; i = 1,2,\ldots\ldots n)$ indipendenti e $N(\mu_i, \sigma_i^2)$.

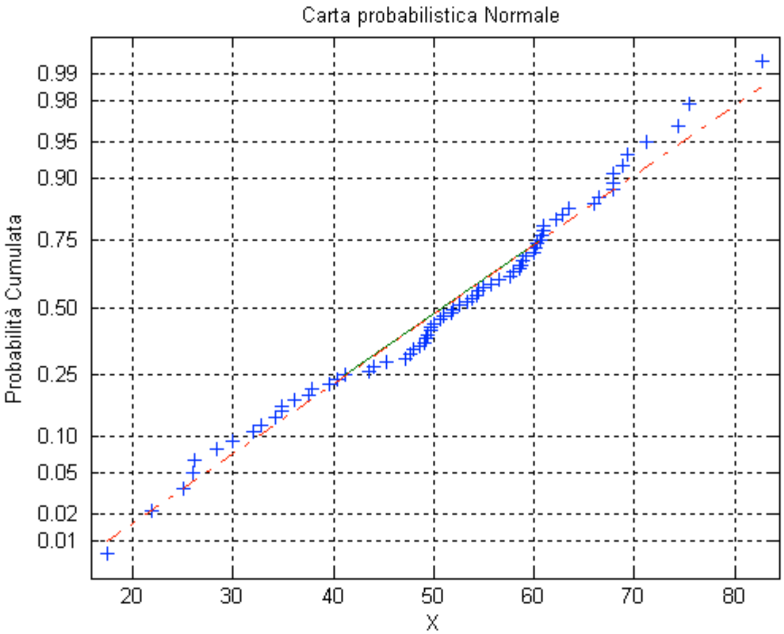
$$Z = \sum_{i=1}^{n} X_i \ \text{è}\ N(\mu_z, \sigma_z^2) \qquad \begin{aligned} \mu_z &= \sum \mu_i \\ \sigma_z^2 &= \sum \sigma_i^2 \end{aligned}$$

Carta probabilistica normale.

In diagramma cartesiano con ascissa X ed ordinata u e $\Phi(u)$ la funzione di probabilità cumulata $F_X(x)$ sarà rappresentata dalla retta: $x = \mu + u\sigma$

Rappresentazione in carta normale

$$F_i = F(x_i) = \frac{i-\alpha}{N-2\alpha+1} \qquad \alpha = \frac{3}{8} \cong \frac{1}{2}$$



Media di variabili normali:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \text{ è } N\left(\mu_z, \frac{\sigma_z^2}{n^2}\right)$$

Variabili X_i indipendenti e identicamente distribuite $N(\mu_z, \sigma_z^2)$.

$$E\left[\overline{X}\right] = \frac{n\mu}{n}$$

$$\overline{X} \text{ è } N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \mathrm{var}\left[\overline{X}\right] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\sigma\left[\overline{X}\right] = \frac{\sigma}{\sqrt{n}}$$

DISTRIBUZIONI DERIVATE.

Funzione $Y = g(x)$ strettamente monotona crescente e derivabile di una v.a. continua X

Esempio: $Y = \log(x)$ oppure $x^{1/2}$ oppure $x^{1/3}$ $x \geq 0$

$$F_Y(y) = P[Y \leq y] = P[g(X) \leq g(x)] = P[X \leq x] = F_X(x)$$

$$dF_Y(y) = dF_X(x) \Rightarrow f_Y(y)dy = f_X(x)dx$$

$$f_Y(y) = f_X(x)\frac{dx}{dy} \qquad ovvero: \qquad f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|$$

Quando si conosce la distribuzione della Y e si ricerca quella della X vale, ovviamente

$$f_X(x) = f_Y(y)\left(\frac{dy}{dx}\right) = f_Y(g(x))\left|\frac{d(g(x))}{dx}\right|$$

Esempio: variabile normale standard $u = \dfrac{(x-\mu)}{\sigma} \Rightarrow x = \mu + u\sigma$

$$f_U(u) = \frac{dx}{du} f_X(\mu + u\sigma) = \sigma \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(\mu+u\sigma-\mu)}{\sigma}\right]^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

Vale anche: $F_x(x) = P[X \leq x] = P\left[u \leq \left(\dfrac{x-\mu_x}{\sigma_x}\right)\right] = F_u(u))$

MEDIA DI UNA VARIABILE FUNZIONE DI UN' ALTRA.

Se c è una costante:

$$E[c] = \int_{-\infty}^{+\infty} c f_x(x) dx = c \int_{-\infty}^{+\infty} f_x(x) dx = c$$

Similmente:

$$E[cX] = cE[X]$$

$$E[a + bX] = a + bE[X]$$

$$E[g_1(x) + g_2(x)] = E[g_1(x)] + E[g_2(x)]$$

In generale:

$$E[g(x)] \neq g(E[g(x)])$$

Esempio:

$$E\left[\frac{1}{X}\right] \neq \frac{1}{E[X]}$$

$$E[X^2] = \{E[X]\}^2 + \sigma_x^2$$

VARIANZA DI UNA VARIABILE FUNZIONE.

$$\mathrm{var}[x] = E[(x-\mu)^2] = E[x^2 - 2\mu x + \mu^2] = E[x^2] - 2\mu E[x] + \mu^2 = E[x^2] - \mu^2$$

$$\mathrm{var}[c] = 0$$

$$\mathrm{var}[cx] = c^2 \,\mathrm{var}[x]$$

$$\mathrm{var}[a + bx] = b^2 \,\mathrm{var}[x]$$

VARIABILE STANDARDIZZATA.

$$u = \frac{x-\mu}{\sigma} \qquad E[u] = 0 \qquad \mathrm{var}[u] = 1$$