Computazione per l'interazione naturale: fondamenti probabilistici

Corso di Interazione uomo-macchina II

Prof. Giuseppe Boccignone

Dipartimento di Informatica
Università di Milano

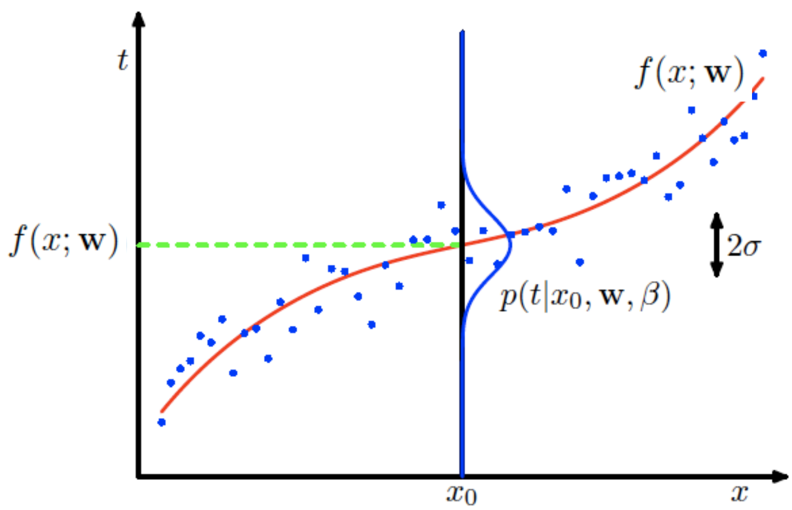
boccignone@di.unimi.it
http://boccignone.di.unimi.it/IUM2_2014.html

Regressione lineare //modelli con rumore

- Assumiamo una generazione di punti con un processo rumoroso

$$t = f(x; \mathbf{w}) + \varepsilon$$

come definiamo la precisione?



Probabilità
//assiomi:definizione formale

- L'insieme dei possibili valori osservati (risultati dell'osservazione di un processo casuale) definisce uno spazio campione

- Un evento A ⊆ Ω è un insieme di valori dallo spazio campione. La classe F dei possibili eventi soddisfa le proprietà seguenti:

1. $\emptyset \in \mathcal{F}$
2. $A \in \mathcal{F} \implies \Omega \backslash A \in \mathcal{F}$
3. $A_1, A_2, \ldots, A_n \in \mathcal{F} \implies \cup_{i=1}^{n} A_i \in \mathcal{F}$

- Su F è definita una misura di probabilità P : F → IR, che soddisfa le proprietà seguenti:

1. $\forall A \in \mathcal{F} : P(A) \geq 0$
2. $P(\Omega) = 1$
3. Se $A_1, A_2, \ldots, A_n$ sono eventi disgiunti (tali che $A_i \cap A_j = \emptyset, \forall i, j, i \neq j$) allora

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

Probabilità condizionata

- Sia B un evento a probabilità non nulla (P(B) > 0). La probabilità condizionata di un evento A dato B è definita come

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

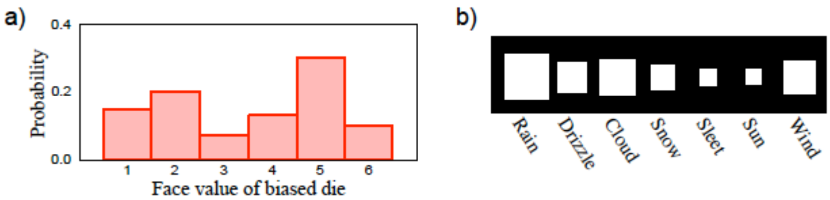
- Se P(A∩B) = P(A)P(B), e quindi P(A | B) = P(A), allora i due eventi sono detti indipendenti.

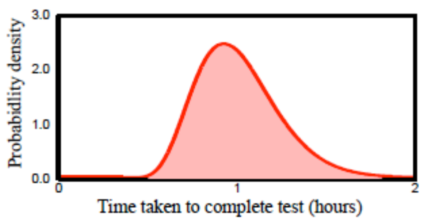
Variabili aleatorie

- Una variabile casuale (o random) è una funzione

$$X : \Omega \mapsto \mathbb{R}$$

- Se, dato $\omega \in \Omega, X(\omega)$ può assumere soltanto valori interi, allora X è una variabile casuale discreta, altrimenti viene detta continua.





Variabili aleatorie

- Una variabile casuale (o random) è una funzione

$$X : \Omega \mapsto \mathbb{R}$$

- Se, dato $\omega \in \Omega, X(\omega)$ può assumere soltanto valori interi, allora X è una variabile casuale discreta, altrimenti viene detta continua.

- Una distribuzione (cumulativa) di probabilità è una funzione

$$F_X : \mathbb{R} \mapsto [0, 1]$$

- tale che $F_X(x) = P(X \leq x)$. Per tale funzione valgono le seguenti

1. $0 \leq F_X(x) \leq 1$
2. $\lim_{x \to -\infty} F_X(x) = 0$
3. $\lim_{x \to \infty} F_X(x) = 1$
4. $x \leq y \implies F_X(x) \leq F_X(y)$

Distribuzione di probabilità

- Se X può assumere un insieme finito di possibili valori, è possibile definire la distribuzione (di massa) di probabilità

$$p_X : \Omega \mapsto \mathbb{R}$$

- tale che $p_X(x) = P(X = x)$. Valgono:

1. $0 \leq p_X(x) \leq 1$
2. $\sum_{x \in V(X)} p_X(x) = 1$, dove $V(X)$ è l'insieme dei possibili valori di X
3. $\sum_{x \in A} p_X(x) = P(X \in A)$

Densità di probabilità

- Se X è continua e F_X è derivabile ovunque, possiamo definire la densità di probabilità

$$f_X(x) = \frac{dF_X(x)}{dx}$$

- Dalla definizione di derivata:

$$P(x \leq X \leq x + \Delta x) \approx f_X(x)\Delta x$$

- Altre proprietà

1. $f_X(x) \geq 0$
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$
3. $\int_A f_X(x)dx = P(X \in A)$

Valore atteso

- X variabile casuale discreta con distribuzione $p_X(x)$, $g : \mathbb{R} \mapsto \mathbb{R}$ generica funzione, allora $g(X)$ è variabile casuale e possiamo definire il suo valore atteso come:

$$E[g(X)] = \sum_{x \in V(X)} g(x) p_X(x)$$

- X variabile casuale continua

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

Valore atteso: proprietà

1. $E[a] = a$ per ogni $a \in \mathbb{R}$
2. $E[af(X)] = aE[f(X)]$ per ogni $a \in \mathbb{R}$
3. $E[f(X) + g(X)] = E[f(X)] + E[g(X)]$
4. Se X è discreta, allora $E[1\{X = k\}] = P(X = k)$, dove $1\{X = k\}$ è la funzione che assume valore 1 per $X = k$ e 0 altrimenti

Varianza

- La varianza di una variabile casuale X

$$Var[X] = E[(X - E[X])^2]$$

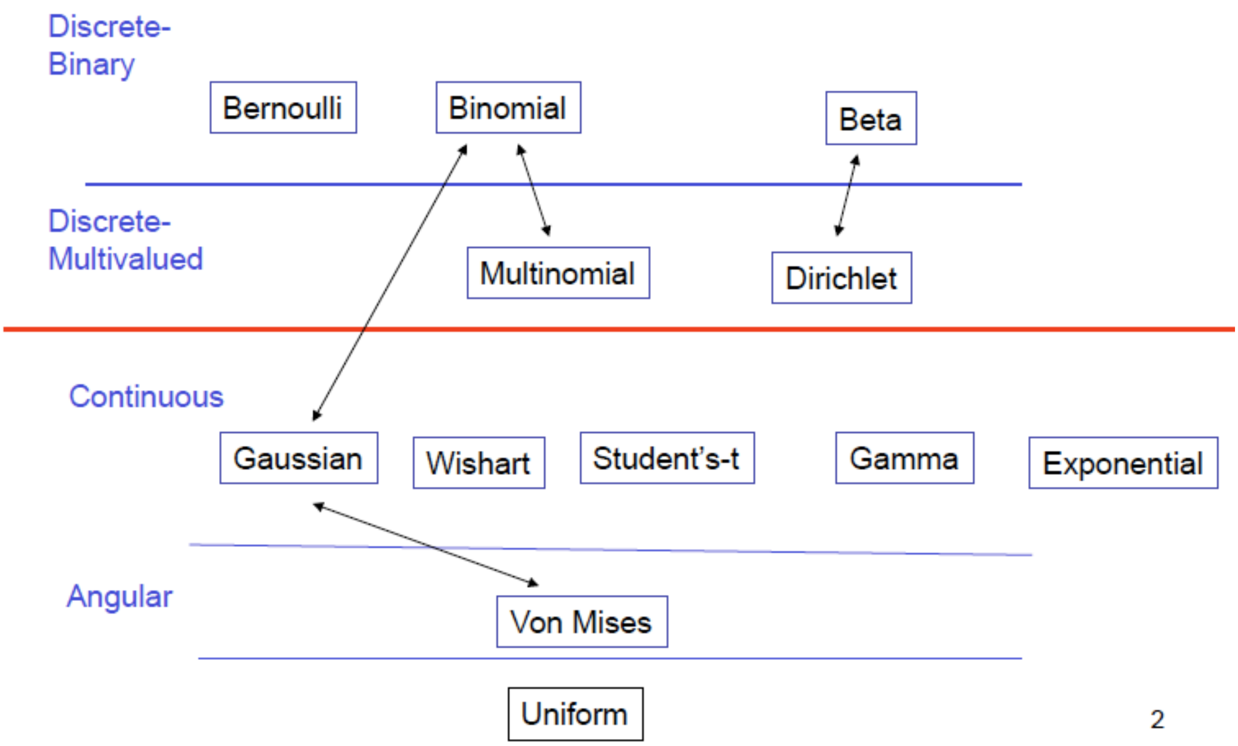
- Vale allora:

$$\begin{aligned} E[(X - E[X])^2] &= E[X^2 - 2E[X]X + E[X]^2] \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

- Inoltre

$$Var[a] = 0 \text{ per } a \in \mathbb{R}$$
$$Var[af(X)] = a^2 Var[f(X)] \text{ per } a \in \mathbb{R}$$

Distribuzioni utili



Distribuzioni discrete //Bernoulli


Jacob Bernoulli
1654-1705

$$X \sim Bernoulli(p) \ \ (0 \leq p \leq 1)$$

- Probabilità che, data una moneta con probabilità di “testa” uguale a p, un suo lancio abbia “testa” come esito.

$$p(x) = \begin{cases} p & \text{se esito} = \text{“TESTA”} \\ 1-p & \text{altrimenti} \end{cases} = p^x(1-p)^{1-x}$$

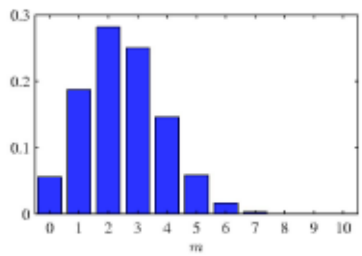
- Media
$$E[X] = p$$
- Varianza
$$Var[X] = p(1-p)$$

Distribuzioni discrete //Binomiale

$$X \sim Binomial(n,p) \ \ (0 \leq p \leq 1)$$

- probabilità che, data una moneta con probabilità di “testa” uguale a p, n suoi lanci indipendenti diano x volte “testa” come esito.

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$



n=10,
p=0.25

- Media
$$E[X] = np$$
- Varianza
$$Var[X] = np(1-p)$$

Distribuzioni discrete
//Poisson

$$X \sim Poisson(\lambda)\ (\lambda > 0)$$

- probabilità che, dato un evento che può avvenire in media lambda volte per unità di tempo, l'evento compaia x volte nell'unità di tempo.

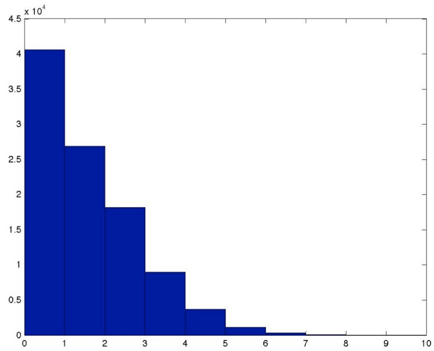
$$p(x) = e^{-\lambda}\frac{\lambda^x}{x!}$$

- Media

$$E[X] = \lambda$$

- Varianza

$$Var[X] = \lambda$$



MATLAB

```
lambda = 2;
x = poissrnd(lambda,1,1000);
hist(x)
```

Distribuzioni continue
//Uniforme

$$X \sim Uniform(a, b)\ (a < b)$$

- probabilità costante tra a e b, 0 altrove.

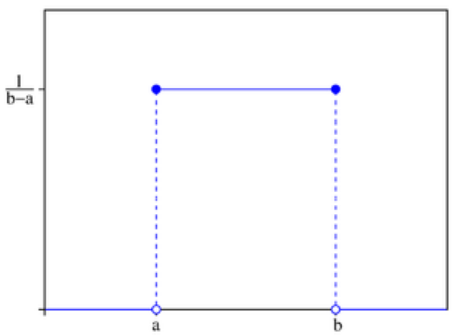
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{se } a \leq x \leq b \\ 0 & \text{altrimenti} \end{cases}$$

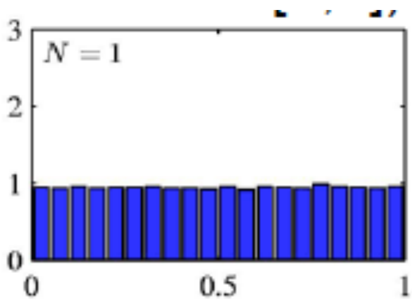
- Media

$$E[X] = \frac{a+b}{2}$$

- Varianza

$$Var[X] = \frac{(b-a)^2}{12}$$





Distribuzioni continue
//Esponenziale

$$X \sim Exponential(\lambda) \ (\lambda > 0)$$

- probabilità che l'intervallo tra due eventi successivi in un processo di Poisson con parametro λ abbia lunghezza x

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \leq 0 \\ 0 & \text{altrimenti} \end{cases}$$

- Media

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

- Varianza

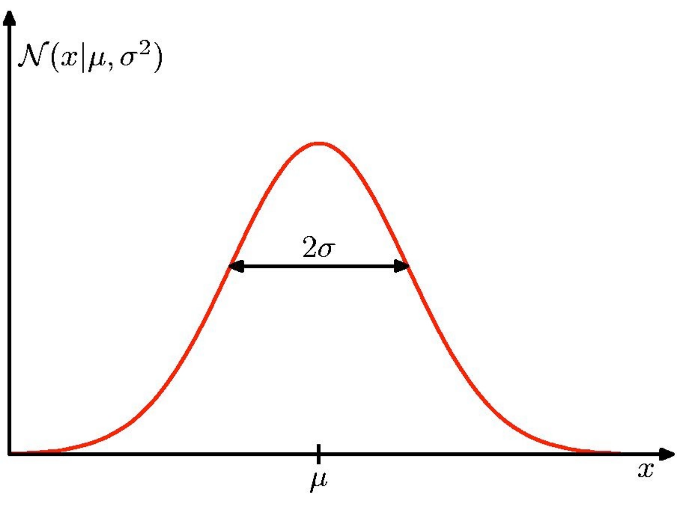
$$Var[X] = \frac{1}{\lambda^2}$$

Distribuzioni continue
//Normale o Gaussiana

Carl Friedrich Gauss
1777-1855

$$X \sim Normal(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$



- Media

$$E[X] = \mu$$

- Varianza

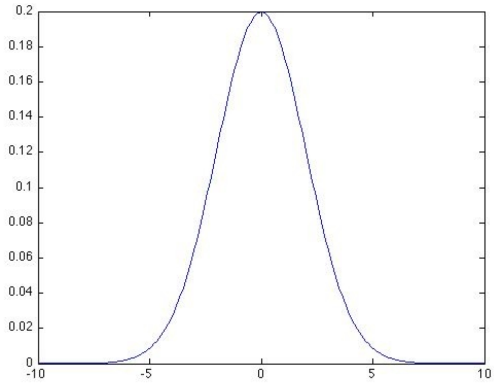
$$Var[X] = \sigma^2$$

Distribuzioni continue //Normale o Gaussiana

Carl Friedrich Gauss
1777-1855

$$X \sim Normal(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$



- Media

$$E[X] = \mu$$

- Varianza

$$Var[X] = \sigma^2$$

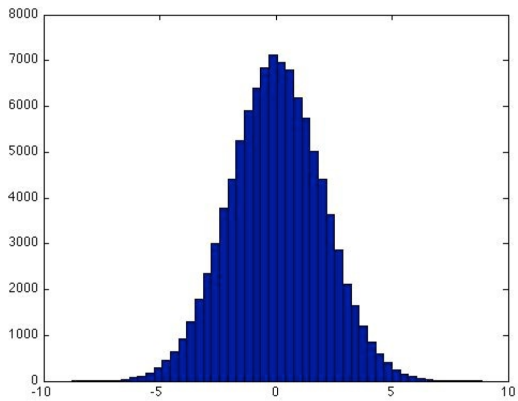
```
>> mu=0;
>> sigma=2;
>> x=-10:0.1:10;
>> y = normpdf(x,mu,sigma);
>> plot(x,y)
```

Distribuzioni continue //Normale o Gaussiana

Carl Friedrich Gauss
1777-1855

$$X \sim Normal(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$



- Media

$$E[X] = \mu$$

- Varianza

$$Var[X] = \sigma^2$$

```
>> mu=0;
>> sigma=2;
>> x = normrnd(mu,sigma,1,100000);
>> hist(x)
```

Distribuzioni continue
//Beta

- Usata come a priori sul parametro della Binomiale

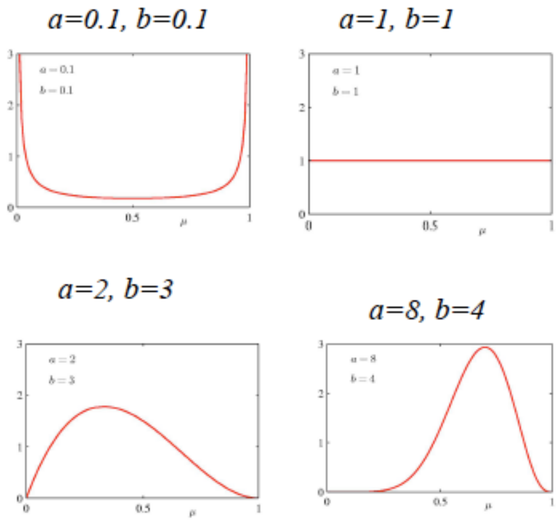
$$Beta(\mu \mid a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}$$

- dove $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du$ è la funzione Gamma
- Media

$$E[\mu] = \frac{a}{a+b}$$

- Varianza

$$\mathrm{var}[\mu] = \frac{ab}{(a+b)^2(a+b+1)}$$



Distribuzione Beta
in funzione degli iperparametri a,b

Distribuzioni continue
//Gamma

- Usata come a priori sulla varianza della Normale o sulla Poisson

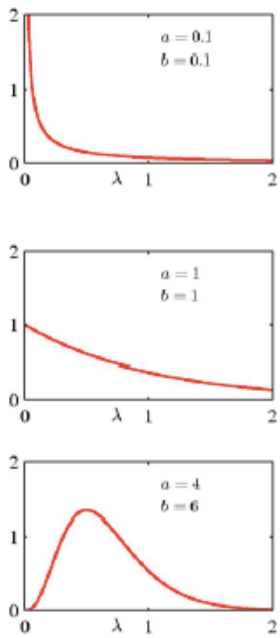
$$Gam(\lambda \mid a,b) = \frac{1}{\Gamma(a)} b^a \lambda^{a-1} \exp(-b\lambda)$$

- dove $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du$ è la funzione Gamma
- Media

$$E[\lambda] = \frac{a}{b}$$

- Varianza

$$\mathrm{var}[\lambda] = \frac{a}{b^2}$$



Distribuzione Gamma
in funzione degli iperparametri a,b