PROP: Data una coppia di v.a. continue (X, Y) con funzione di densità congiunta

$$f_{XY}(x,y) = g(x)\, h(y)$$

per cui il dominio di definizione "non lega X e Y" allora X e Y sono indipendenti.

COVARIANZA & CO.

- (X, Y) coppia di v.a.
- $g(X, Y)$

$$E[g(X,Y)] = \begin{cases} \sum_{x_i}\sum_{y_j} g(x_i, y_j)\, p_{XY}(x_i, y_j) \\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x_i, y_j)\, f_{XY}(x_i, y_j)\, dy\, dx \end{cases}$$

Def:

Data una coppia di v.a. (X, Y)

- si definisce momento di ordine (p, q)

$$E[X^p Y^q]$$

- se le v.a. hanno valore medio, si definisce momento centrale di ordine (p, q)

$$E[(X-E[X])^p (Y-E[Y])^q]$$

OSS.

① il momento di ordine $(1,0)$ è $E[X]$, cioè

$$E[X] = \iint x f_{XY}(x,y)\, dy\, dx$$

- analogamente l'ordine $(0,1)$ è $E[Y]$ cioè

$$E[Y] = \iint y f_{XY}(x,y)\, dy\, dx$$

② il momento centrale di ordine $(2,0)$ è $Var(X)$ e quello $(0,2)$ è $Var(Y)$

COVARIANZA

Def: Si definisce covarianza della coppia di v.a. (X, Y) il suo momento centrale di ordine $(1,1)$, cioè:

$$\mathrm{Cov}(X,Y) = \begin{cases} \sum_{x_i} \sum_{y_j} (x_i - E[X])(y_j - E[Y]) \, p_{XY}(x_i, y_j) \\ \iint (x - E[X])(y - E[Y]) \, f_{XY}(x,y) \, dy \, dx \end{cases}$$

PROP

$$\boxed{\mathrm{Cov}(X,Y) = E[XY] - E[X]\,E[Y]}$$

OSS.

- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
- **Def**: Due v.a. X e Y si definiscono non correlate se $\mathrm{Cov}(X,Y) = 0$

a) Se X e Y sono indip $\Rightarrow \mathrm{Cov}(X,Y) = 0$
b) Se $\mathrm{Cov}(X,Y) = 0 \not\Rightarrow X$ e Y indip.
c) Se $\mathrm{Cov}(X,Y) \neq 0 \Rightarrow X$ e Y non sono indip.

PROP

Se X e Y sono v.a. indip. allora:

$$E[XY] = E[X]\,E[Y]$$

Coefficiente di correlazione lineare

Def. Data una coppia di v.a. (X, Y) si definisce c. di corr.

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

- $\rho_{X,Y}$ è adimensionale
- se $a, b \in \mathbb{R}^+$ $(a>0, b>0)$ allora

$$\rho_{aX, bY} = \rho_{X,Y}$$

- ρ_{XY} se e solo se $Cov(X, Y) = 0$
- $|\rho_{X,Y}| \leq 1$. In particolare $|\rho_{XY}| = 1$ se e solo se $Y = aX + b$ con

$$\rho_{XY} = 1 \text{ se } a>0$$

$$\rho_{XY} = -1 \text{ se } a<$$

DISTRIBUZIONE NORMALE BIVARIATA

In questo caso se $Cov(X, Y) = 0 \Rightarrow X$ e Y indipendenti

$$f_{XY}(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} q(x,y)}$$

$(x, y) \in$

dove $\sigma_x > 0$, $\sigma_y > 0$, $|\rho| < 1$

e

$$q(x,y) = \left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)$$

con μ_x e $\mu_y \in \mathbb{R}$

- SI DIMOSTRA CHE $\mu_x = E[X]$ $\sigma^2_x = Var(X)$
$\mu_y = E[Y]$ $\sigma^2_y = Var(Y)$

$\rho = \rho_{X,Y}$

FUNZIONE GENERATRICE di MOMENTI

X v.a.

- Si definisce funzione generatrice dei momenti di X, (MGF)

$$M_X(t) = E\left[e^{tX}\right] = \begin{cases} \sum_{x_i} e^{tx_i} p_X(x_i) & \text{v.a. discreta} \\ \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx & \text{v.a. continua} \end{cases}$$

se è definita in $\underbrace{[-t_0, t_0]}_{\text{intorno di } t}$, con $t_0 > 0$.

- $M_X(0) = 1$

PROPRIETÀ

1. $M'_X(0) = E[X]$

in generale:

$$M_X^{\overbrace{(k)}^{\text{derivata } k}}(0) = E\left[X^k\right]$$

2. Se X e Y sono v.a. tali che

$$M_X(t) = M_Y(t) \quad \forall t \in [-t_0, t_0]$$

allora X e Y hanno stessa distribuzione

3. Se $Y = aX + b$ con $a, b \in \mathbb{R}$, $a \neq 0$

$$M_Y(t) = e^{tb} M_X(at)$$

4. Se X e Y sono v.a. indipendenti, allora

$$M_{X+Y}(t) = M_X(t)\, M_Y(t)$$

+ in generale, se $X_1, \dots, X_n$ sono v.a. indipendenti allora

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

$M_{aX}(t) = M_X(at)$

MGF per distribuzioni notevoli

a) $X \sim Be(p)$

$$p_X(k) = \begin{cases} 1-p & k=0 \\ p & k=1 \end{cases}$$

$$\Rightarrow M_X(t) = \sum_{x_i} e^{t x_i} p_X(x_i) = e^{t\cdot 0} p_X(0) + e^{t\cdot 1} p_X(1) + e^{t\cdot 1} p_X(1)$$

$$= (1-p) + pe^t$$

b) $X \sim B(n,p)$ $\quad p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k=0,1,\dots,n \\ 0 & \text{altrove} \end{cases}$

$$\Rightarrow M_X(t) = \sum_{x_i} e^{t x_i} p_X(x_i) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} =$$

$$= \sum_{k=0}^{n} \underbrace{(pe^t)^n}_{a} \underbrace{(1-p)^{n-k}}_{b} = (pe^t + (1-p))^n$$

$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a+b)^n$$

Oss.

Se $X_1, \dots, X_n$ sono v.a. indipendenti di tipo Bernoulli con lo stesso parametro p, allora

$$\sum_{i=1}^{n} X_i \sim B(n,p)$$

c) $X \sim Poiss(\lambda)$

$$p_X(k) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & k=0,1,2,\dots \\ 0 & \text{altrove} \end{cases}$$

$$\Rightarrow M_X(t) = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!}$$

ricorda che $\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$

$$\Rightarrow M_X(t) = e^{-\lambda} \cdot e^{\lambda e^t}$$

Oss.

Se $X_1, \dots, X_n$ sono v.a. indip. con $X_i \sim Poiss(\lambda_i)$ $i=1,\dots,n$

allora $\sum_{i=1}^{n} X_i \sim Poiss\left(\sum_{i=1}^{n} \lambda_i\right)$

! è così anche il contrario quindi:

$X \sim Poiss$ può essere visto come: $X = \sum_{i=1}^{n} X_i$ con

X_i v.a. indip. $\Rightarrow X_i \sim Poiss\left(\frac{\lambda}{n}\right)$

d) $X \sim N(0,1)$

$M_X(t) = e^{t^2/2}$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad x \in \mathbb{R} \Rightarrow M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx$$

Oss

Se $X \sim N(\mu, \sigma^2)$

- sappiamo che $Y = \frac{X - \mu}{\sigma} \sim N(0,1) \Rightarrow M_Y(t) = e^{t^2/2}$

ESPLICITO $X = \sigma Y + \mu$ $\Rightarrow$ (3ª proprietà) $M_X(t) = e^{t^2/2}$

$$M_X(t) = e^{\mu t} \cdot e^{\sigma^2 t^2/2}$$

Oss 2

Se $X_1, \dots, X_n$ sono v.a. indipendenti
con $X_i \sim N(\mu_i, \sigma_i^2)$
allora $\sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$

TEOREMI LIMITE

① Legge Debole dei grandi numeri (LDGN)

② Teorema limite CENTRALE (TLC)

• X popolazione

Un campione casuale di dimensione n estratto dalla popolazione X è una n-pla $X_1, \ldots, X_n$ di v.a. indipendenti e identicamente distribuite cioè aventi la stessa distribuzione della popolazione X.

Def:

Si definisce media campionaria di $(X_1, \ldots, X_n)$ la v.a.

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

PROPIETÀ: ① Se $E[X] = \mu$ allora $E[\overline{X}_n] = \mu$

② $Var(X) = \sigma^2$ allora $Var(\overline{X}_n) = \frac{\sigma^2}{n}$

con il termine v.a. i.i.d. si intende v.a. indipendenti e identicamente distribuite.

LDGN

Data una successione $\{X_n\}$ con X_i v.a. i.i.d. con media finita μ, allora

$$\forall \varepsilon > 0 \quad P(\{|\overline{X}_n - \mu| > \varepsilon\}) \xrightarrow[n \to +\infty]{} 0$$

TLC

Sia $\{X_n\}$ una successione di v.a. i.i.d. con media μ e varianza σ^2 finite. Allora la v.a.

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}\,\sigma}$$

è tale che

$$\lim_{n\to\infty} F_{Z_n}(z) = \phi^*(z)$$

(Cioè Z_n converge in legge a una v.a. $N(0,1)$)
$(n \geq 30)$

conseguenze

① $\bar{X}_n \simeq N(\mu, \frac{\sigma^2}{n})$

② $\sum_{i=1}^{n} X_i \simeq N(n\mu, n\sigma^2)$

Osservazioni:

① Se $X \sim N(\mu, \sigma^2)$ per qualunque n si ha:

- $\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$

e

- $\bar{X}_n \sim N(\mu, \sigma^2/n)$

② il th. di De Moivre-Laplace è una conseguenza del TLC.

$X \sim B(n,p) \Rightarrow X = \sum_{i=1}^{n} X_i$ con $X_i \sim Be(p)$ indip. $\rightarrow$ soddisfa TLC

$\mu_{X_i} = p \qquad \sigma^2_{X_i} = p(1-p)$

$\Downarrow$ TLC

$$\sum_{i=1}^{n} X_i \simeq N(np, np(1-p))$$

③ Se $X \sim Poiss(\lambda)$

$X = \sum_{i=1}^{n} X_i$ con $X_i \sim Poiss(\frac{\lambda}{n})$ indip

$\mu = \frac{\lambda}{n}$ $\sigma^2 = \frac{\lambda}{n}$

$\underset{TLC}{\Rightarrow}$ $\sum_{i=1}^{n} X_i \approx N(n\frac{\lambda}{n}, n\frac{\lambda}{n})$

$X \approx N(\lambda, \lambda)$

$\lambda \geq 5$

STIMATORE

X popolazione, θ parametro da stimare

- $(X_1, \dots, X_n)$ campione casuale (iid)

Def. Si definisce stimatore del parametro θ la statistica $T(X_1, \dots, X_n)$ che usiamo per stimare θ

Oss. • T è una v.a.

$T(x_1, \dots, x_n) = \tau \leftarrow$ stima del parametro (è un numero)

Def.

- Uno stimatore T per il parametro θ è detto corretto (o non distorto) se

$$E[T] = \theta$$

- Si definisce distorsione dello stimatore:

$$D(T) = E[T] - \theta$$

- Si dice che uno stimatore è assintoticamente corretto se:

$$E[T] \xrightarrow[n \to +\infty]{} \theta$$

(o equivalentemente $D(T) \xrightarrow[n \to +\infty]{} 0$)

Es

La media campionaria:

$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ è stimatore corretto per il valore medio ($E[\bar{X}_n] = \mu_X$, $Var(\bar{X}_n) = \frac{\sigma_X^2}{n}$

esiste S. corretto per la varianza?

Varianza Campionaria

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x_n})^2$$

! questo non è stimatore corretto ma solo assintoticamente corretto

OSS

dividiamo somme

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} \left(x_i^2 + \overline{x_n}^2 - 2\,\overline{x_n}\, x_i \right) =$$

non dipende da i

posso tirarlo fuori x motivo precedente

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \frac{1}{\cancel{n}} \cancel{n}\, \overline{x_n}^2 + (-2\,\overline{x_n}) \cdot \underbrace{\frac{1}{n} \sum_{i=1}^{n} x_i}_{\overline{x_n}} =$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \overline{x_n}^2 - 2 x_n^2 = \underbrace{\frac{1}{n} \sum_{i=1}^{n} x_i^2}_{\text{momento campionario di ordine 2}} - x_n^2$$

Def:

Si definisce momento campionario di ordine r.

$$M_r = \frac{1}{n} \sum_{i=1}^{n} x_i^r$$

PROPIETÀ *

Ogni momento campionario è uno stimatore corretto per il relativo momento (della popolazione) $\Rightarrow E[M_r] \overset{?}{=} E[X^r]$

PROPOSIZIONE:

La varianza campionaria è uno stimatore asintoticamente corretto per la varianza σ^2 della popolazione. $\Rightarrow E[S^2] = E[M_2] - E[\overline{X_n}^2]$

Dim

$$E[M_2] \overset{*}{=} E[X^2] \overset{\text{def. di varianza}}{=} \sigma^2 + \mu^2$$

$$E[\overline{X_n}^2] \underset{\text{def di varianza}}{=} \frac{\sigma^2}{n} + \mu^2$$

$$E[S^2] = \sigma^2 - \frac{\sigma^2}{n} \xrightarrow[n \to +\infty]{} \sigma^2$$

OSS

$$\frac{n}{n-1} E[S^2] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2 \cdot \frac{n}{n-1}$$ ← x ottenere σ^2

$$E\left[\frac{n}{n-1} S^2\right] = \sigma^2$$

↳ è stimatore corretto per σ^2

Varianza campionaria Corretta

$$\bar{S}^2 = \frac{n}{n-1} S^2 = \frac{n}{n-1} \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \Rightarrow$$

$$\bar{S}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$$

Def Si definisce ERRORE QUADRATICO MEDIO di uno stimatore $T(x_1, \dots, x_n)$ per il parametro θ.

$$EQM(T) = E[(T-\theta)^2]$$

OSS Se T_1 e T_2 sono 2 stimatori per θ, si dice che T_1 è PIÙ EFFICIENTE di T_2 se

$$EQM(T_1) < EQM(T_2)$$

PROPIETÀ

$$EQM(T) = Var(T) + D(T)^2$$

Def:

- Uno stimatore $T(X_1, \dots, X_n)$ per un parametro θ è detto consistente in probabilità se

$$\forall \varepsilon > 0 \quad P(|T(X_1, \dots, X_n) - \theta| > \varepsilon) \xrightarrow[n \to +\infty]{} 0$$

- Uno stimatore $T(X_1, \dots, X_n)$ per un parametro θ è detto consistente in media quadratica se

$$E[(T-\theta)^2] \xrightarrow[n \to +\infty]{} 0$$

Oss

$$\forall \varepsilon > 0 \quad P(|T-\theta| > \varepsilon) \leq \frac{E[(T-\theta)^2]}{\varepsilon^2}$$

PROP

- Se T è asintoticamente CORRETTO:

$$\left(D(T) \xrightarrow[n \to +\infty]{} 0\right)$$

- $Var(T) \xrightarrow[n \to +\infty]{} 0$

allora T è uno stimatore consistente $\Rightarrow EQM(T) = D(T)^2 + Var(T)$

Oss.

- Dalla LDGN, $\bar{X}_n$ è stimatore consistente in prob. per μ, ed anche in media quadratica.

- Anche $\bar{S}^2$ è uno stimatore ~~corre~~ consistente per la varianza (σ^2)

Metodo dei momenti

- $M_r = \frac{1}{n} \sum_{i=1}^{n} X_i^r$ ← momento campionario di ordine r

- X popolazione $X_1, \dots, X_n$ campione

$\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_n)$ ← parametro da stimare

- $E[X^r] = f_r(\theta_1, \theta_2, \dots, \theta_K)$ ← scrivere i momenti in funzione del parametro

RISOLVERE IL SISTEMA

$$\begin{cases} E[X] = M_1 = \overline{X_n} \\ E[X^2] = M_2 \\ \vdots \\ E[X^K] = M_K \end{cases} \qquad \hat{\theta}_M$$

(1) - $X \sim N(\mu, \sigma^2)$ $\quad \theta = (\mu, \sigma^2)$

$$\begin{cases} E[X] = \overline{X_n} \\ E[X^2] = M_2 \end{cases} \qquad \begin{cases} \mu = \overline{X_n} \\ \sigma^2 + \mu^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \end{cases} \Rightarrow$$

$\downarrow$

$\sigma^2 + \mu^2$

$$\begin{cases} \mu = \overline{X_n} \\ \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \overline{X_n}^2 = S^2 \end{cases}$$

$$\begin{cases} \hat{\mu}_M = \overline{X_n} \\ \hat{\sigma}^2_M = S^2 \end{cases}$$

② $X \sim Poiss(\lambda)$ $\quad \theta = \lambda$

$$E[X] = \overline{X}_n$$

$$\lambda = \overline{X}_n \implies \hat{\lambda}_M = \overline{X}_n$$

③ $X \sim Exp(\lambda)$ $\quad \theta = \lambda$

$$E[X] = \overline{X}_n = \frac{1}{\lambda} \implies \hat{\lambda}_M = \frac{1}{\overline{X}_M}$$

④ $X \sim U[0, \theta]$

$$E[X] = \overline{X}_n$$

$$\frac{\theta}{2} = \overline{X}_n \implies \hat{\theta}_M = 2\overline{X}_n$$

⑤ X con fz di densità

- $$f_X(x) = \begin{cases} \frac{2\theta+1}{2}(x^2)^\theta & -1 \leq x \leq 1 \\ 0 & \text{altrove} \end{cases}$$

- $$E[X] = \int_{-1}^{1} \frac{2\theta+1}{2} x^{2\theta} \cdot x \, dx = 0$$

✱ essendo il momento di ordine $M1 = 0$ devo calcolare $M2$

- $$E[X^2] = \int_{-1}^{1} \frac{2\theta+1}{2} x^{2\theta} \cdot x^2 \, dx = 2\int_0^1 \frac{2\theta+1}{2} x^{2\theta+2} dx$$

$$= \frac{2\theta+1}{2\theta+3} \implies \frac{2\theta+1}{2\theta+3} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \; M2$$

$$\implies 2\theta + 1 = (2\theta+3)M^2 \implies \theta = \frac{3M_2 - 1}{2(1-M_2)}$$

- $X \sim U[a-b, a+b]$ con $b > 0$

$$\left.\begin{array}{l} E[X] = a \\ E[X^2] = Var(X) + E[X]^2 = \frac{4b^2}{12} + a^2 \end{array}\right\} \Rightarrow$$

$$\begin{cases} a = \overline{X}_n \\ \frac{b^2}{3} + a^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \end{cases} \quad \frac{b^2}{3} = \underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}_n^2}_{S^2}$$

$$\begin{cases} a = \overline{X}_n \\ b^2 = 3S^2 \end{cases} \Rightarrow \quad \begin{array}{l} \hat{a}_M = \overline{X}_n \\ \hat{b}_M = \sqrt{3S^2} \end{array}$$

STIMATORE di Max verosimiglianza o stimatore MLE ($\hat{\Theta}_{MLE}$)

X popolazione $f_X(x; \Theta)$
fz. di densità che dipende da parametro

- $P(X = x_1, X = x_2, \dots, X_n = x_n) = \overset{p_X(x_1)\cdots p_X(x_n)}{P(X_1 = x_1) \cdots P(X_n = x_n)}$

$= \prod_{i=1}^{n} f_X(x_i; \Theta)$

- $L(\Theta) = \prod_{i=1}^{n} f_X(x_i; \Theta)$
↳ fz di verosimiglianza

-Lo stimatore MLE è $\hat{\Theta}$ che rende massimo $L(\Theta)$

- $\ell(\Theta) = \log L(\Theta)$ ⇒ viene fatto xkè trovare il max quando c'è un prodotto non è facile, trovarlo su una somma sì.
↳ fz di log-verosimiglianza

- PROPIETÀ di invarianza:

Se g è una fz. continua e monotona allora, se $\hat{\Theta}_{MLE}$ è lo stimatore MLE per Θ si ha che $g(\hat{\Theta}_{MLE})$ è lo stimatore MLE per $g(\Theta)$

Es

1. $X \sim Be(p)$ $\quad p_X(k) = \begin{cases} 0 & 1-p \\ 1 & p \end{cases}$

- $p_X(k) = p^k (1-p)^{1-k} \quad k = 0,1$

$L(p) = \underbrace{p^{x_1}(1-p)^{1-x_n}}_{p_X(x_1)} \cdot \underbrace{p^{x_2}(1-p)^{1-x_2}}_{p_X(x_2)} \cdots \underbrace{p^{x_n}(1-p)^{1-x_n}}_{p_X(x_n)}$

raccolgo p e (1-p)

$= p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i}$ ⇒

$$\log(L(p)) = \sum_{i=1}^{n} x_i \log p + \left(n - \sum_{i=1}^{n} x_i\right)\log(1-p)$$

ora deriviamo per trovare MAX

$$\frac{d}{dp} \log L(p) = \frac{1}{p}\sum_{i=1}^{n} x_i - \left(n - \sum_{i=1}^{n} x_i\right)\frac{1}{1-p} = 0$$

$$(1-p)\sum_{i=1}^{n} x_i - p\left(n - \sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} x_i - \cancel{p\sum_{i=1}^{n} x_i} - np + \cancel{p\sum_{i=1}^{n} x_i} = 0$$

$$\Rightarrow p_{MLE} = \underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i}_{\overline{x_n}} \Rightarrow \hat{p}_{MLE} = \overline{x_n}$$

② $X \sim Poiss(\lambda)$

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad k = 0, 1, 2, \dots$$

- $L(\lambda) = e^{-\lambda}\dfrac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda}\dfrac{\lambda^{x_2}}{x_2!} \cdot \dots e^{-\lambda}\dfrac{\lambda^{x_n}}{x_n!} =$

$$= \frac{e^{-\lambda n}\,\lambda^{\sum_{i=1}^{n} x_i}}{x_1!\, x_2! \dots x_n!}$$

- $\ell(\lambda) = \log L(\lambda) = -\lambda n + \sum_{i=1}^{n} x_i \log\lambda + \underbrace{\log\left(\frac{1}{x_1! \dots x_n!}\right)}_{\text{è indip. dal } \lambda \text{ quindi derivata} = 0}$

$$\frac{d}{d\lambda}\log(\lambda) = -n + \frac{1}{\lambda}\sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \lambda_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i \Rightarrow \hat{\lambda}_{MLE} = \overline{x_n}$$

③ • $X \sim N(\mu, \sigma^2)$

- $f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-(x-\mu)^2/2\sigma^2} \quad x \in \mathbb{R}$

$\sqrt{2\pi} \cdot \sigma \to \sqrt{2\pi\sigma^2}$

$\theta = (\mu, \sigma^2)$

- $L(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1-\mu)^2/2\sigma^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \cdots \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$

$L(\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2}$

- $\ell(\theta) = \log L(\theta) = \underbrace{-\frac{n}{2} \log(2\pi\sigma^2)}_{-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2} - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2$

$$\begin{cases} \frac{\partial \ell(\theta)}{\partial \mu} = 0 \\ \frac{\partial \ell(\theta)}{\partial \sigma^2} = 0 \end{cases} \qquad \begin{cases} \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu) = 0 \\ -\frac{n}{2} \cdot \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i-\mu)^2 = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} x_i - n\mu = 0 \\ -n\sigma^2 + \sum_{i=1}^{n} (x_i-\mu)^2 = 0 \end{cases} \qquad \begin{cases} \mu = \frac{1}{n} \sum_{i=1}^{n} x_i \\ \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i-\mu)^2 \end{cases}$$

$\hat{\mu}_{MLE} = \bar{X}_n$

$\hat{\sigma}^2_{MLE} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 = S^2$