TEOREMA: De Moivre-Laplace

Siano $X_1, \dots, X_n$ Bernoulli di parametro $p \in (0,1)$ indipendenti.

Indicando $S_n = X_1 + \dots + X_n$, allora vale che per $a, b \in \mathbb{R}$, $a < b$:

$$\mathbb{P}\left[np + a\sqrt{np(1-p)} \le S_n \le np + b\sqrt{np(1-p)}\right] \longrightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

DIM.

- Supponiamo per semplificare che $n = 2m$, $p = 1/2$. Inoltre, si può assumere $a = 0$, $b > 0$. Se questo caso particolare fosse vero:
  - $a < 0 < b$, allora:

$$\left\{np + a\sqrt{np(1-p)} \le S_n \le np + b\sqrt{np(1-p)}\right\} =$$

$$= \left\{np + a\sqrt{np(1-p)} \le S_n \le np\right\} \sqcup \left\{np \le S_n \le np + b\sqrt{np(1-p)}\right\}$$

$(T_n = n - S_n)$

$$= \left\{np + a\sqrt{np(1-p)} \le n - T_n \le np\right\} \sqcup \left\{\dots\right\}$$

$$= \left\{n(p-1) + a\sqrt{np(1-p)} \le -T_n \le n(p-1)\right\} \sqcup \left\{\dots\right\}$$

$$= \left\{n(1-p) \le T_n \le n(1-p) - a\sqrt{np(1-p)}\right\} \sqcup \left\{\dots\right\}$$

cioè due eventi nella forma del caso particolare.

  - $0 < a < b$, allora:

$$\left\{np + a\sqrt{np(1-p)} \le S_n \le np + b\sqrt{np(1-p)}\right\} =$$

$$= \left\{np \le S_n \le np + b\sqrt{np(1-p)}\right\} \setminus \left\{np \le S_n \le np + a\sqrt{np(1-p)}\right\}$$

  - $a < b < 0$, analogamente al passo precedente, con $T_n = n - S_n$.

Quindi con queste semplificazioni, dimostriamo che:

$$\mathbb{P}\left[m \le S_n \le m + a\sqrt{\frac{m}{2}}\right] \xrightarrow{n \to +\infty} \int_0^a \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

- Stima per $1 + x = e^x e^{R(x)}$ per $|x| < 1/2$

So che $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$

$$\le x + \frac{x^2}{2}\left(1 + |x| + x^2 + \dots\right) = x + R(x)$$

In particolare:

$$|R(x)| \le \frac{x^2}{2}\left(1 + |x| + x^2 + \dots\right) = \frac{x^2}{2} \cdot \frac{1}{1 - |x|} \le x^2 \quad \forall |x| < 1/2$$

- Scrivo ora la probabilità cercata:

$$\mathbb{P}\left[m \le S_n \le m + a\sqrt{\frac{m}{2}}\right] = \sum_{k=m}^{m + \lfloor a\sqrt{m/2} \rfloor} \mathbb{P}[S_n = k] =: \sum_{k=m}^{m + \lfloor a\sqrt{m/2} \rfloor} P_{1/2}(2m, k)$$

$$= \sum_{k=0}^{\lfloor a\sqrt{m/2} \rfloor} P_{1/2}(2m, m+k)$$

In particolare, riconduco $P_{1/2}(2m, m+k)$ a $P_{1/2}(2m, m)$:

$$P_{1/2}(2m, m+k) = \binom{2m}{m+k}\frac{1}{2^{2m}} = \frac{m-k+1}{m+k}\binom{2m}{m+k-1}\frac{1}{2^{2m}}$$

$$= \dots = \frac{m-k+1}{m+k}\cdot\frac{m-k+2}{m+k-1}\cdot\dots\cdot\frac{m}{m+1}\binom{2m}{m}\frac{1}{2^{2m}}$$

$$= \frac{m-k+1}{m+k}\cdot\frac{m-k+2}{m+k-1}\cdot\dots\cdot\frac{m}{m+1}P_{1/2}(2m, m)$$

$$= \frac{\left(1-\frac{k-1}{m}\right)\cdot\left(1-\frac{k-2}{m}\right)\cdot\dots\cdot\left(1-\frac{1}{m}\right)}{\left(1+\frac{k}{m}\right)\left(1+\frac{k-1}{m}\right)\cdot\dots\cdot\left(1+\frac{1}{m}\right)}P_{1/2}(2m, m)$$

$$= \frac{e^{-\frac{k-1}{m}}\cdot e^{-\frac{k-2}{m}}\cdot\dots\cdot e^{-\frac{1}{m}}\cdot e^{R\left(-\frac{k-1}{m}\right)}\cdot\dots\cdot e^{R\left(-\frac{1}{m}\right)}}{e^{\frac{k}{m}}\cdot e^{\frac{k-1}{m}}\cdot\dots\cdot e^{\frac{1}{m}}\cdot e^{R\left(\frac{k}{m}\right)}\cdot\dots\cdot e^{R\left(\frac{1}{m}\right)}}P_{1/2}(2m, m) = *$$

- Bisogna quindi trovare una stima per:

$$* = \underbrace{e^{-\frac{k}{m}-2\left(\frac{k-1}{m}+\dots+\frac{1}{m}\right)}}_{(1)}\cdot\underbrace{e^{R\left(-\frac{k-1}{m}\right)+\dots+R\left(-\frac{1}{m}\right)-R\left(\frac{k}{m}\right)-\dots-R\left(\frac{1}{m}\right)}}_{(2)}\cdot\underbrace{P_{1/2}(2m, m)}_{(3)}$$

1) Osservo che:

$$e^{-\frac{k}{m}-2\left(\frac{k-1}{m}+\dots+\frac{1}{m}\right)} = e^{-\frac{k}{m}-2\frac{k(k-1)}{2m}} = e^{-\frac{k^2}{m}}$$

2) Usando la stima precedente, indico $E(k) = R\left(-\frac{k-1}{m}\right)+\dots+R\left(-\frac{1}{m}\right)-R\left(\frac{k-1}{m}\right)-\dots-R\left(\frac{1}{m}\right)$

$$|E(k)| \leq \left|R\left(-\frac{k-1}{m}\right)\right|+\dots+\left|R\left(-\frac{1}{m}\right)\right|+\left|R\left(\frac{k-1}{m}\right)\right|+\dots+\left|R\left(\frac{1}{m}\right)\right|$$

$$\leq \left(-\frac{k-1}{m}\right)^2+\dots+\left(-\frac{1}{m}\right)^2+\left(\frac{k-1}{m}\right)^2+\dots+\left(\frac{1}{m}\right)^2$$

$$\leq 2\left[\left(\frac{k-1}{m}\right)^2+\dots+\left(\frac{1}{m}\right)^2\right] \leq 2k\left(\frac{k}{m}\right)^2 = \frac{2k^3}{m^2}$$

(Vale perché $0 \leq k \leq a\sqrt{\frac{m}{2}} \Rightarrow \frac{k}{m} \leq \frac{a}{\sqrt{2m}} \leq \frac{1}{2}$ per $m > \frac{2}{a^2}$)

$$\leq \frac{2}{m^2}\cdot\frac{a^3\cdot m^{3/2}}{2^{3/2}} = \frac{a^3}{\sqrt{2m}} \xrightarrow{m\to+\infty} 0$$

3) Usando Stirling:

$$P_{1/2}(2m, m) = \binom{2m}{m}\cdot\frac{1}{2^{2m}} = \frac{(2m)^{2m+\frac{1}{2}}\cdot e^{-2m}\sqrt{2\pi}}{m^{2m+1}\cdot e^{-2m}\cdot(2\pi)}\frac{1}{2^{2m}}(1+o(1))$$

$$= \frac{1}{\sqrt{2\pi}}\cdot\sqrt{\frac{2}{m}} = \frac{1}{\sqrt{\pi m}}(1+o(1)) \quad \text{per } m\to+\infty$$

- Ora, usando delle stime, confrontiamo il valore cercato. Dato che:

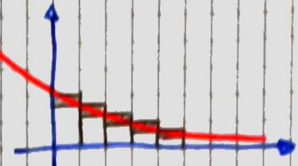
$$-\frac{a^3}{\sqrt{2m}} \leq -|E(k)| \leq |E(k)| \leq \frac{a^3}{\sqrt{2m}}$$

Posso scrivere:

$$\underbrace{e^{-\frac{a^3}{\sqrt{2m}}}P_{1/2}(2m, m)\cdot\sum_{k=0}^{\left[a\sqrt{\frac{m}{2}}\right]}e^{-\frac{k^2}{m}}}_{(A)} \leq \sum_{k=0}^{\left[a\sqrt{\frac{m}{2}}\right]}P_{1/2}(2m, m+k) \leq \underbrace{e^{\frac{a^3}{\sqrt{2m}}}P_{1/2}(2m, m)\sum_{k=0}^{\left[a\sqrt{\frac{m}{2}}\right]}e^{-\frac{k^2}{m}}}_{(B)}$$

Osservo che:

$$(A) \geq e^{-\frac{a^3}{\sqrt{2m}}} \cdot P_{1/2}(2m,m) \cdot \int_0^{\lfloor a\sqrt{\frac{m}{2}} \rfloor} e^{-\frac{x^2}{m}} dx$$
$$\geq e^{-\frac{a^3}{\sqrt{2m}}} \cdot P_{1/2}(2m,m) \cdot \int_0^a e^{-\frac{y^2}{2}} \cdot \sqrt{\frac{m}{2}}\, dy$$
$$\geq e^{-\frac{a^3}{\sqrt{2m}}} \cdot P_{1/2}(2m,m) \sqrt{\frac{m}{2}} \int_0^a e^{-y^2/2} dy$$
$$\xrightarrow{m \to +\infty} 1 \cdot \frac{1}{\sqrt{m\pi}} \cdot \sqrt{\frac{m}{2}} \int_0^a e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_0^a e^{-y^2/2} dy$$



A) trovo Σ come somma superiore discretizzata dell'integrale
B) trovo Σ come minorazione

Analogamente con (B) $\longrightarrow \frac{1}{\sqrt{2\pi}} \int_0^a e^{-y^2/2} dy$.

Allora per il Teorema dei Due Carabinieri:

$$\sum_{k=0}^{\lfloor a\sqrt{\frac{m}{2}} \rfloor} P_{1/2}(2m, m+k) \xrightarrow{m \to +\infty} \int_0^a \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

TEOREMA: Legge dei Grandi Numeri

Siano $X_1, \dots, X_n$ v. a. di Bernoulli $p \in (0,1)$ indipendenti.
Dato $S_n = X_1 + \dots + X_n$, vale che:

$$\mathbb{P}\left[\left| \frac{S_n}{n} - p \right| \geq \varepsilon \right] \xrightarrow{n \to +\infty} 0$$

· DIM.

Per De-Moivre-Laplace, fissato $\varepsilon > 0$, vale:

$$\mathbb{P}\left[np + a\sqrt{np(1-p)} \leq S_n \leq np + b\sqrt{np(1-p)} \right] =$$
$$= \mathbb{P}\left[\frac{a\sqrt{p(1-p)}}{\sqrt{n}} \leq \frac{S_n}{n} - p \leq \frac{b\sqrt{p(1-p)}}{\sqrt{n}} \right] \xrightarrow{n \to +\infty} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Sia $\alpha > 0$. $a = -\alpha$, $b = \alpha$.

$$\mathbb{P}\left[-\alpha\sqrt{\frac{p(1-p)}{n}} \leq \frac{S_n}{n} - p \leq \alpha\sqrt{\frac{p(1-p)}{n}} \right] = \mathbb{P}\left[\left| \frac{S_n}{n} - p \right| \leq \alpha\sqrt{\frac{p(1-p)}{n}} \right]$$

In particolare:

$$\mathbb{P}\left[\left| \frac{S_n}{n} - p \right| > \varepsilon \right] = 1 - \mathbb{P}\left[\left| \frac{S_n}{n} - p \right| \leq \varepsilon \right]$$

Per n abbastanza grande:

$$\left\{ \left| \frac{S_n}{n} - p \right| \leq \alpha \sqrt{\frac{p(1-p)}{n}} \right\} \subseteq \left\{ \left| \frac{S_n}{n} - p \right| \leq \varepsilon \right\}$$

Quindi

$$\mathbb{P}\left[\left| \frac{S_n}{n} - p \right| > \varepsilon \right] = 1 - \mathbb{P}\left[\left| \frac{S_n}{n} - p \right| \leq \varepsilon \right] \leq 1 - \mathbb{P}\left[\left| \frac{S_n}{n} - p \right| \leq \alpha\sqrt{\frac{p(1-p)}{n}} \right]$$
$$\leq 1 - \limsup_{n \to +\infty} \mathbb{P}\left[\left| \frac{S_n}{n} - p \right| \leq \alpha\sqrt{\frac{p(1-p)}{n}} \right]$$
$$\leq 1 - 2\int_0^\alpha \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 2\int_\alpha^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \xrightarrow{\alpha \to +\infty} 0$$

Momenti di una Gaussiana

- Sia $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Allora

$$\mathbb{E}[|Z|^k] = \int_{\mathbb{R}} |x|^k \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$
$$= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} x^k e^{-x^2/2} dx = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} -x^{k-1} (e^{-x^2/2})' dx$$
$$= \frac{2}{\sqrt{2\pi}} \left[-x^{k-1} e^{-x^2/2} \right]_0^{+\infty} + \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} (k-1) x^{k-2} e^{-x^2/2} dx$$
$$= \frac{2}{\sqrt{2\pi}} \left[-x^{k-1} e^{-x^2/2} \right]_0^{+\infty} + (k-1) \int_{\mathbb{R}} \frac{|x|^{k-2}}{\sqrt{2\pi}} e^{-x^2/2} dx$$
$$= 0 + (k-1)\mathbb{E}[|Z|^{k-2}] = (k-1)\mathbb{E}[|Z|^{k-2}]$$

Inoltre,

$$\mathbb{E}[|Z|] < +\infty \quad \text{perché} \quad \mathbb{E}[|Z|^2] = \mathbb{E}[Z^2] = 1$$

$\rightarrow$ Per induzione, ha tutti i momenti.

TLC debole

Siano $X_1, \dots, X_n$ var. alea. con stessa legge e tali che $\mathbb{P}[|X_1| \leq C] = 1$ per $C \in \mathbb{R}$. Allora X_i ha tutti i momenti e vale che:

$$\mathbb{P}\left[a \leq \frac{X_1 + \dots + X_n - \mathbb{E}[X_1 + \dots + X_n]}{sd(X_1 + \dots + X_n)} \leq b \right] \xrightarrow{n \to +\infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx$$

DIM.

Osservo che per $Y_i = \frac{X_i - \mathbb{E}[X_i]}{sd(X_i)}$ vale:

$$Z = \frac{X_1 + \dots + X_n - \mathbb{E}[X_1 + \dots + X_n]}{sd(X_1 + \dots + X_n)} = \frac{Y_1 + \dots + Y_n}{\sqrt{n}}$$

Dato che $\mathbb{P}[|X_i| \leq C] = 1$ allora

$$\mathbb{E}[|X_i|] \leq \mathbb{E}[C] = C, \quad \mathbb{E}[|X_i|^k] \leq C^k$$

$\Rightarrow$ Per un certo C', $\mathbb{P}[|Y_i| < C'] = 1$, quindi Y_i ha tutti i momenti e in particolare: $\mathbb{E}[Y_i] = 0$, $\mathbb{E}[Y_i^2] = Var(Y_i) = 1$ in quanto Y_i normalizzazione di X_i.

Calcolo che: $\mathbb{E}[|Z|^k] = \mathbb{E}\left[\left| \frac{Y_1 + \dots + Y_n}{\sqrt{n}} \right|^k \right] \leq \mathbb{E}\left[\left(\frac{n|Y_1|}{\sqrt{n}} \right)^k \right] \leq (\sqrt{n}\, C)^k < +\infty$

Quindi calcolo i momenti:

- $\mathbb{E}[Z] = \mathbb{E}\left[\frac{Y_1 + \dots + Y_n}{\sqrt{n}} \right] = \sqrt{n}\, \mathbb{E}[Y_1] = 0$
- $\mathbb{E}[Z^2] = \mathbb{E}\left[\frac{(Y_1 + \dots + Y_n)^2}{n} \right] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[Y_i^2] + 0 = 1$

- $\mathbb{E}[Z^k] = \mathbb{E}\left[\frac{(Y_1+\dots+Y_n)^k}{\sqrt{n}^k}\right] = \frac{1}{n^{k/2}}\mathbb{E}[(Y_1+\dots+Y_n)^k]$

$$= \frac{1}{n^{k/2}} \sum_{\substack{i_1,\dots,i_k \\ \text{(indici ordinati)}}} \mathbb{E}[Y_{i_1} Y_{i_2} \dots Y_{i_k}]$$

$$= \frac{1}{n^{k/2}} \sum_{\ell=1}^{k} \sum_{\substack{j_1,\dots,j_\ell \\ \text{distinti}}} \sum_{m_1+\dots+m_\ell = k} \mathbb{E}[Y_{j_1}^{m_1} \cdots Y_{j_\ell}^{m_\ell}]$$

Ho allora 3 casi:

i) $\ell < \frac{k}{2}$: allora calcolo che:

$$\frac{1}{n^{k/2}} \sum_{\substack{j_1,\dots,j_\ell \\ \text{distinti}}} \sum_{m_1+\dots+m_\ell=k} \mathbb{E}[Y_{j_1}^{m_1} \cdots Y_{j_\ell}^{m_\ell}] \leq$$

$$\leq \frac{1}{n^{k/2}} \cdot n^\ell \cdot k^k \cdot C^k \xrightarrow{n\to+\infty} 0 \quad \text{perché } \ell < \frac{k}{2}.$$

ii) $\ell > \frac{k}{2}$: per Pidgeonhole, c'è un $m_i = 1$, in quanto la somma dev'essere k e ho ℓ termini $m_i > 0$.

Per tale serie di indici ed esponenti:

$$\mathbb{E}[Y_{j_1}^{m_1} \cdots Y_{j_\ell}^{m_\ell}] = \mathbb{E}[Y_{j_1}^{m_1}] \cdot \ldots \cdot \mathbb{E}[Y_{j_\ell}^{m_\ell}] = 0$$

iii) $\ell = \frac{k}{2}$ (cioè k pari) allora: se $\exists\, m_i > 2$, segue che per qualche indice $m_j = 1$ e il termine si annulla.

Restano quindi i termini $m_1 = m_2 = \dots = m_\ell = 2$. In tal caso:

$$\frac{1}{n^{k/2}} \cdot \sum_{\substack{j_1,\dots,j_{\frac{k}{2}} \\ \text{distinti}}} \mathbb{E}[Y_{j_1}^2 \cdots Y_{j_{\frac{k}{2}}}^2] =$$

$$= \frac{1}{n^{k/2}} \sum_{\substack{j_1,\dots,j_{\frac{k}{2}} \\ \text{distinti}}} \mathbb{E}[Y_{j_1}^2] \cdot \ldots \cdot \mathbb{E}[Y_{j_{k/2}}^2] = \frac{1}{n^{k/2}} \sum_{\substack{j_1,\dots,j_{\frac{k}{2}} \\ \text{distinti}}} 1$$

Conto allora tali combinazioni:

- ho $n\cdot(n-1)\cdot\ldots\cdot(n-\frac{k}{2}+1)$ modi per scegliere gli indici $j_1,\dots,j_{\frac{k}{2}}$ distinti
- in tutto, dati tali indici, ho lo stesso termine contato

$$\frac{k! \leftarrow \text{scelte per } Y_{j_1}\dots Y_{j_{\frac{k}{2}}}}{\underbrace{2^{k/2}}_{\text{coppie}} \cdot \underbrace{(\frac{k}{2})!}_{\text{tolgo "l'ordine" di } j_1,\dots,j_{\frac{k}{2}}}} \quad \text{volte}$$

Ma $\frac{k!}{2^{k/2}(\frac{k}{2})!} = \frac{k!}{k!!} = (k-1)!!$

Mettendo assieme:

$$\frac{1}{n^{k/2}} \cdot n(n-1)\cdot\ldots\cdot(n-\frac{k}{2}+1) \cdot (k-1)!! \xrightarrow{n\to+\infty} (k-1)!!$$

Ma allora:

$$\mathbb{E}[Z^k] \xrightarrow{n\to+\infty} \begin{cases} 0 & \text{se } k \text{ dispari} \\ (k-1)!! & \text{se } k \text{ pari} \end{cases}$$

MODELLO ESPONENZIALE

- **DEF. Modello Esponenziale**

Sia $\Theta \subseteq \mathbb{R}$. Sia $(\Omega, \mathcal{F}, \{\mathbb{P}_\vartheta\}_{\vartheta\in\Theta})$ modello statistico e sia $(X_1, \dots, X_n)$ campione di legge $\{m_\vartheta\}_{\vartheta\in\Theta}$.

- CASO ASSOLUTAMENTE CONTINUO: esistono 2 funzioni:

$$g : \mathbb{R} \to \mathbb{R} \qquad T : \mathbb{R} \to \mathbb{R}$$

tali che:

$$m_\vartheta(x) = f_\vartheta(x) = c_\vartheta \cdot g(x) \cdot e^{\vartheta T(x)}$$

- CASO DISCRETO: esistono 2 funzioni:

$$g : \mathbb{N} \to \mathbb{R} \qquad T : \mathbb{N} \to \mathbb{R}$$

tali che:

$$m_\vartheta(k) = p_\vartheta(k) = c_\vartheta \cdot g(k) \cdot e^{\vartheta T(k)}$$

- **TEOREMA**

Dato un modello statistico $(\Omega, \mathcal{F}, \{\mathbb{P}_\vartheta\}_{\vartheta\in\Theta})$ e un campione $(X_1, \dots, X_n)$ estratto da $\{m_\vartheta\}_{\vartheta\in\Theta}$, supponiamo che:

1) il modello statistico non sia banale: per $\vartheta_1 \neq \vartheta_2$, $\mathbb{P}_{\vartheta_1} \neq \mathbb{P}_{\vartheta_2}$.
2) $\Theta \subseteq \mathbb{R}$ e' un insieme aperto
3) $\{m_\vartheta\}_{\vartheta\in\Theta}$ e' un modello esponenziale:
$$m_\vartheta(x) = f_\vartheta(x) = c_\vartheta \cdot g(x) \cdot e^{\vartheta T(x)}$$
4) la seguente funzione sia integrabile per ogni $\vartheta \in \Theta$:
$$x \longmapsto g(x) \cdot T(x)^2 e^{\vartheta (T(x))_+}$$
5) lo stimatore di massima verosimiglianza U_n esista per ogni $\vartheta \in \Theta$.

Allora $(U_n)_{n\geq 1}$ e' CONSISTENTE.

TEORIA ANALISI 2

Sia $F(t,x) = \int f(t,x)\,dx$ primitiva. Dato $J(t_0)$ intorno di t_0, se valgono le seguenti:

- $x \longmapsto f(t,x)$ e' integrabile per ogni $t \in J(t_0)$
- $t \longmapsto f(t,x)$ e' differenziabile in $J(t_0)$
- $\left|\frac{\partial}{\partial t} f(t,x)\right| \leq g(x)$ con g integrabile $\forall t \in J(t_0)$ e quasi ogni x

allora $$\frac{\partial}{\partial t} F(t,x) = \int \frac{\partial}{\partial t} f(t,x)\,dx$$

- **DIM.**

Per definizione di densita':

$$c_\vartheta^{-1} = \int_{\mathbb{R}} g(x)\, e^{\vartheta T(x)}\,dx$$

Studiamo allora:

$$\psi(\vartheta) = -\log c_\vartheta = \log \int_{\mathbb{R}} g(x) e^{\vartheta T(x)} dx$$

In particolare, considerando $f(t,x) = g(x) e^{t \cdot T(x)}$ su cui valgono le ipotesi di Derivazione Sotto Integrale (per 4):

$$\psi'(\vartheta) = -\frac{d}{d\vartheta} \log c_\vartheta = \frac{d}{d\vartheta} \log c_\vartheta^{-1} = c_\vartheta \cdot \frac{d}{d\vartheta}\left(\frac{1}{c_\vartheta}\right)$$

$$= c_\vartheta \cdot \frac{d}{d\vartheta} \int_{\mathbb{R}} g(x) e^{\vartheta T(x)} dx = c_\vartheta \cdot \int_{\mathbb{R}} \frac{d}{d\vartheta} g(x) e^{\vartheta T(x)} dx$$

$$= c_\vartheta \cdot \int_{\mathbb{R}} T(x) g(x) e^{\vartheta T(x)} dx = \int_{\mathbb{R}} T(x) f_\vartheta(x) dx = E_\vartheta[T(X_1)]$$

Analogamente:

$$\psi''(\vartheta) = \frac{d}{d\vartheta} \psi'(\vartheta) = \frac{d}{d\vartheta}\left(c_\vartheta \cdot \int_{\mathbb{R}} T(x) g(x) e^{\vartheta T(x)} dx \right)$$

$$= \frac{d}{d\vartheta}\left(e^{-\psi(\vartheta)} \cdot \int_{\mathbb{R}} T(x) g(x) e^{\vartheta T(x)} dx \right)$$

$$= -\psi'(\vartheta) e^{-\psi(\vartheta)} \int_{\mathbb{R}} T(x) g(x) e^{\vartheta T(x)} dx + e^{-\psi(\vartheta)} \int_{\mathbb{R}} T^2(x) g(x) e^{\vartheta T(x)} dx$$

$$= -\psi'(\vartheta)^2 + \int_{\mathbb{R}} T^2(x) c_\vartheta g(x) e^{\vartheta T(x)} dx$$

$$= -E[T(X_1)]^2 + E[T(X_1)^2] = Var_\vartheta(T(X_1)) \geq 0$$

In particolare allora $\psi(\vartheta)$ è convessa. Vogliamo vedere che $\psi''(\vartheta) > 0$.

Supponiamo per assurdo che $\exists \vartheta_0 \in \Theta$ tale che $\psi''(\vartheta_0) = 0$.

Ma allora: $Var_{\vartheta_0}(T(X_1)) = 0$ cioè $T(x)$ costante a meno di un insieme di misura nulla su f_{ϑ_0}.

Segue che $f_{\vartheta_0}(x) = 0$ quasi ovunque, sia:

$$A = \{x \in \mathbb{R} \mid f_{\vartheta_0}(x) = 0\}$$

Allora:

$$0 = f_{\vartheta_0}(x) = c_{\vartheta_0} g(x) e^{\vartheta_0 T(x)}$$

$$\Rightarrow g(x) e^{\vartheta_0 T(x)} = 0 \text{ quasi ovunque} \Rightarrow g(x) = 0 \text{ quasi ovunque}$$

Ma allora:

$$\int_A c_{\vartheta_0} g(x) e^{\vartheta_0 T(x)} dx = 0 = \int_A c_\vartheta g(x) e^{\vartheta T(x)} dx$$

che è assurdo perché il modello sarebbe banale ($P_\vartheta = P_{\vartheta_0}$).

Troviamo ora lo STIMATORE di MASSIMA VEROSIMIGLIANZA:

$$L_\vartheta(x_1, \dots, x_n) = f_\vartheta(x_1) \cdot \dots \cdot f_\vartheta(x_n)$$

$$\stackrel{!}{=} c_\vartheta^n \, g(x_1)\cdot\ldots\cdot g(x_n)\cdot e^{\vartheta(T(x_1)+\ldots+T(x_n))}$$

$$\log L_\vartheta(x_1,\ldots,x_n) = n\left[-\psi(\vartheta) + \vartheta\,\frac{T(x_1)+\ldots+T(x_n)}{n}\right] + \log(g(x_1)\cdot\ldots\cdot g(x_n))$$

Tale funzione e' strettamente concava (somma con $\psi(\vartheta)$ convessa) quindi la derivata ha un' unica radice:

$$\frac{d}{d\vartheta}\log L_\vartheta(x_1,\ldots,x_n) = n\left[\frac{T(x_1)+\ldots+T(x_n)}{n} - \psi'(\vartheta)\right]$$

cioe':

$$U_n = (\psi')^{-1}\left(\frac{1}{n}(T(x_1)+\ldots+T(x_n))\right)$$

Per la Legge dei Grandi Numeri allora

$$\frac{1}{n}(T_1(x_1)+\ldots+T(x_n)) \xrightarrow[\text{in prob}]{n\to+\infty} E_\vartheta[T(X_1)] = \psi'(\vartheta)$$

e passando per $(\psi')^{-1}$ continua:

$$U_n = (\psi')^{-1}\left(\frac{1}{n}(T_1(x_1)+\ldots+T_1(x_n))\right) \xrightarrow[\text{in prob}]{n\to+\infty} \vartheta$$

Modello a Rapporto di Verosimiglianza Crescente

- **DEF.** Modello a Rapporto di Verosimiglianza Crescente

Sia $(\Omega, \mathcal{F}, \{\mathbb{P}_\vartheta\}_{\vartheta\in\Theta})$ modello statistico con $\Theta \subseteq \mathbb{R}$. Sia $X_1, \dots, X_n$ campione. Il modello è a rapporto di verosimiglianza crescente rispetto alla variabile aleatoria T se $\forall\ \vartheta_1 < \vartheta_2$, $\vartheta_1, \vartheta_2 \in \Theta$:

$$\frac{L_{\vartheta_2}(X_1, \dots, X_n)}{L_{\vartheta_1}(X_1, \dots, X_n)} \quad \text{è strettamente crescente in funzione di } T$$

- **LEMMA**

Sia $(\Omega, \mathcal{F}, \{\mathbb{P}_\vartheta\}_{\vartheta\in\Theta})$ modello statistico, un campione $(X_1, \dots, X_n)$ e $\Theta = \{\vartheta_0, \vartheta_1\}$. Data la funzione di verosimiglianza

$$L_\vartheta(x_1, \dots, x_n) = f_\vartheta(x_1) \cdot \dots \cdot f_\vartheta(x_n)$$

definiamo per $c > 0$:

$$C = \{L_{\vartheta_0} < c \cdot L_{\vartheta_1}\}$$

e siano $\mathcal{H}_0 : \vartheta = \vartheta_0$ e $\mathcal{H}_1 : \vartheta = \vartheta_1$. Allora:

a) C è la regione critica di un test più potente di ogni altro test di livello $\mathbb{P}_{\vartheta_0}(C)$.

b) $\mathbb{P}_{\vartheta_0}(C) \leq \mathbb{P}_{\vartheta_1}(C)$

- DIM.

a) Sia C_0 la regione critica di un altro test. In particolare:

$$(\mathbb{1}_C - \mathbb{1}_{C_0})(L_{\vartheta_0} - c L_{\vartheta_1}) \leq 0$$

in quanto:

- se $\omega \in C$, $\mathbb{1}_C - \mathbb{1}_{C_0} \geq 0$, $L_{\vartheta_0} - c L_{\vartheta_1} \leq 0$
$$\Rightarrow (\mathbb{1}_C - \mathbb{1}_{C_0})(L_{\vartheta_0} - c L_{\vartheta_1})(\omega) \leq 0$$
- se $\omega \notin C$, $\mathbb{1}_C - \mathbb{1}_{C_0} \leq 0$, $L_{\vartheta_0} - c L_{\vartheta_1} \geq 0$
$$\Rightarrow (\mathbb{1}_C - \mathbb{1}_{C_0})(L_{\vartheta_0} - c L_{\vartheta_1})(\omega) \leq 0$$

In particolare, scrivendo:

$$(\mathbb{1}_C - \mathbb{1}_{C_0}) L_{\vartheta_0} \leq c(\mathbb{1}_C - \mathbb{1}_{C_0}) L_{\vartheta_1}$$

integrando su $\mathbb{R}^n$:

$$\mathbb{P}_{\vartheta_0}(C) - \mathbb{P}_{\vartheta_0}(C_0) \leq c\,(\mathbb{P}_{\vartheta_1}(C) - \mathbb{P}_{\vartheta_1}(C_0))$$

e dato che $\mathbb{P}_{\vartheta_0}(C) - \mathbb{P}_{\vartheta_0}(C_0) \geq 0$ (perché C_0 ha livello $\mathbb{P}_{\vartheta_0}(C)$):

$$\mathbb{P}_{\vartheta_1}(C) > \mathbb{P}_{\vartheta_1}(C_0) \qquad \text{cioè il test è più potente}$$

b) Analogamente, si considera la relazione:

$$(\mathbb{1}_C - \mathbb{P}_{\vartheta_0}(C))(L_{\vartheta_0} - c L_{\vartheta_1}) \leq 0 \qquad \text{(si dimostra come prima)}$$

da cui:

$$c\,\mathbb{1}_C L_{\vartheta_1} - c\,\mathbb{P}_{\vartheta_0}(C) L_{\vartheta_1} \geq \mathbb{1}_C L_{\vartheta_0} - \mathbb{P}_{\vartheta_0}(C) L_{\vartheta_0}$$

Integrando su $\mathbb{R}^n$ (sapendo che $\int_{\mathbb{R}^n} L_{\vartheta_0} = 1$):

$$c(P_{\vartheta_1}(C) - P_{\vartheta_0}(C)) \geq P_{\vartheta_0}(C) - P_{\vartheta_0}(C) = 0$$

da cui:

$$P_{\vartheta_1}(C) \geq P_{\vartheta_0}(C)$$ cioè la tesi.

- TEOREMA

Sia $(\Omega, \mathcal{F}, \{P_\vartheta\}_{\vartheta \in \Theta})$ modello statistico e $(X_1, \dots, X_n)$ campione. Supponiamo il modello sia a rapporto di verosimiglianza crescente rispetto alla variabile aleatoria T. Dato il test:

$$\mathcal{H}_0: \vartheta \leq \vartheta_0 \qquad \mathcal{H}_1: \vartheta > \vartheta_0$$

e data la regione critica $C = \{T > \delta\}$, allora:

a) $\sup_{\vartheta \leq \vartheta_0} P_\vartheta(C) = P_{\vartheta_0}(C)$

b) il test di regione critica C è quello più potente tra i test di livello $P_{\vartheta_0}(C)$

- DIM.

a) Siano $\vartheta_1 \leq \vartheta_2 \leq \vartheta_0$. Allora dato $C = \{T > \delta\}$, esiste $c > 0$ in modo che (per rapporto di verosimiglianza crescente):

$$C = \{L_{\vartheta_1} \leq c \cdot L_{\vartheta_2}\}$$

Applicando il lemma a $\Theta' = \{\vartheta_1, \vartheta_2\}$ ottengo $P_{\vartheta_1}(C) \leq P_{\vartheta_2}(C)$.

b) Sia $C_\star$ la regione critica di un test di livello $P_{\vartheta_0}(C)$, cioè:

$$\sup_{\vartheta \leq \vartheta_0} P_\vartheta(C_\star) \leq P_{\vartheta_0}(C)$$

Dato $\vartheta > \vartheta_0$, si può riapplicare il lemma a $\{\vartheta_0, \vartheta\}$ e il test di regione critica C risulta più potente di $C_\star$ in ϑ.

Popolazioni Gaussiane

- **Lemma**

Sia $(X_1, \dots, X_n)$ popolazione gaussiana $\mathcal{N}(0,1)$ di variabili aleatorie indipendenti e sia A matrice ortogonale.
Allora la popolazione $A \cdot (X_1, \dots, X_n)$ è ancora di gaussiane standar indipend.

- Dim.

Sia $\underline{X} = (X_1, \dots, X_n)$. Allora la densità congiunta sarà:

$$f_{\underline{X}}(y_1, \dots, y_n) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(y_1^2 + \dots + y_n^2)} = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}|\underline{X}|^2}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}|A\underline{X}|^2} = f_{A\underline{X}}(x_1, \dots, x_n)$$

cioè la densità di una popolazione gaussiana.

- **Teorema**

Siano $(X_1, \dots, X_n)$ variabili aleatorie indipendenti con distribuzione $\mathcal{N}(m, \sigma^2)$.
Posti: $\overline{X} = \frac{1}{n}(X_1 + \dots + X_n)$ $\qquad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$
si ha che:

1) $\overline{X}$, S^2 indipendenti
2) $\overline{X}$ ha distribuzione $\mathcal{N}(m, \frac{\sigma^2}{n})$
3) $(n-1)\frac{S^2}{\sigma^2}$ ha distribuzione $\chi^2(n-1)$
4) $\sqrt{n}\,\frac{\overline{X} - m}{S}$ ha distribuzione Student con $n-1$ gradi di libertà

- Dim.

1/3) Consideriamo $X_i = m + \sigma Z_i$, con $Z_i \sim \mathcal{N}(0,1)$.
Segue che $\overline{X} = m + \sigma \overline{Z}$ e che:

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} (Z_i - \overline{Z})^2 =: U^2$$

Ma allora: $\overline{X}, S^2$ indipendenti $\iff$ $\overline{Z}, U^2$ indipendenti, perché
$\overline{X} = m + \sigma \overline{Z}$ e $S^2 = \frac{\sigma^2}{n-1} U^2$

Definiamo ora la base $\mathcal{D} = (f_1, \dots, f_n)$ ortonormale tale che

$$f_1 = \frac{1}{\sqrt{n}}(1, \dots, 1)$$

Prendiamo la matrice di cambio base $A = \mathcal{M}^{C}_{\mathcal{D}}(id)$ con C base canonica. A è ortogonale per costruzione. Allora:

- $[A(Z_1, \dots, Z_n)]_1 = \frac{1}{\sqrt{n}}(Z_1 + \dots + Z_n) = \sqrt{n}\,\overline{Z}$
- Calcolo la norma delle altre $n-1$ componenti.

$$|[A\cdot(Z_1,\dots,Z_n)]_2|^2+\dots+|[A\cdot(Z_1,\dots,Z_n)]_n|^2 =$$
$$= |A\cdot(Z_1,\dots,Z_n)|^2 - |[A\cdot(Z_1,\dots,Z_n)]_1|^2$$
$$= |(Z_1,\dots,Z_n)|^2 - |\sqrt{n}\,\bar{Z}|^2 = \sum_{i=1}^{n} Z_i^2 - n\bar{Z}^2$$
$$= \sum_{i=1}^{n} Z_i^2 - 2n\bar{Z}^2 + n\bar{Z}^2$$
$$= \sum_{i=1}^{n} Z_i^2 - 2\bar{Z}\sum_{i=1}^{n} Z_i + \sum_{i=1}^{n}\bar{Z}^2$$
$$= \sum_{i=1}^{n}\left(Z_i^2 - 2Z_i\bar{Z} + \bar{Z}^2\right) = \sum_{i=1}^{n}(Z_i-\bar{Z})^2 = U^2$$

Ma allora per il lemma, X_1 e $(X_2,\dots,X_n)$ indipendenti implica che $\sqrt{n}\bar{Z}$ e U^2 indipendenti. Segue il punto 1.
Inoltre, dato che ho ricondotto U^2 alla somma di $(Z_2,\dots,Z_n)$, in particolare U^2 somma di $n-1$ gaussiane indipendenti al quadrato.
$\Rightarrow U^2 = \frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$ segue il punto 3.

2) Segue dal Teorema di Somme di n Gaussiane.
4) Dati i risultati 1/2/3, $\sqrt{n}\,\frac{\bar{X}-m}{S}$ verifica la definizione della distribuzione di Student. ∎