

Un'introduzione
alla statistica asintotica
attraverso una generalizzazione
del Teorema Limite Centrale

Candidato:

Carmine Frascella

Università di Pisa

18 Settembre 2015

Relatore:

Maurizio Pratelli



Sezioni

1. Prime definizioni riguardo la statistica asintotica
2. Distanza e Trasformata di Hellinger
3. Esperimenti poissoniani e Poissonizzazione
4. Alcuni risultati preliminari
5. La generalizzazione del Teorema Limite Centrale

1-1. Esperimenti e confronto di rischi

Il contesto è quello del formalismo statistico decisionale.

1-1. Esperimenti e confronto di rischi

Il contesto è quello del formalismo statistico decisionale.

Tre passi:

1-1. Esperimenti e confronto di rischi

Il contesto è quello del formalismo statistico decisionale.

Tre passi:

- Si osserva il dato $x \in (\Omega, \mathcal{A}, P_\theta)$, con $\theta \in \Theta$ ignoto;

1-1. Esperimenti e confronto di rischi

Il contesto è quello del formalismo statistico decisionale.

Tre passi:

- Si osserva il dato $x \in (\Omega, \mathcal{A}, P_\theta)$, con $\theta \in \Theta$ ignoto;
- Si munisce lo spazio (Z, \mathcal{T}) di una probabilità ρ_x ;

1-1. Esperimenti e confronto di rischi

Il contesto è quello del formalismo statistico decisionale.

Tre passi:

- Si osserva il dato $x \in (\Omega, \mathcal{A}, P_\theta)$, con $\theta \in \Theta$ ignoto;
- Si munisce lo spazio (Z, \mathcal{T}) di una probabilità ρ_x ;
- Si compie una decisione $z \in (Z, \mathcal{T}, \rho_x)$, che comporta un **costo** $W_\theta(z)$ (funzione a valori in $[0, 1]$).

1-1. Esperimenti e confronto di rischi

Il contesto è quello del formalismo statistico decisionale.

Tre passi:

- Si osserva il dato $x \in (\Omega, \mathcal{A}, P_\theta)$, con $\theta \in \Theta$ ignoto;
- Si munisce lo spazio (Z, \mathcal{T}) di una probabilità ρ_x ;
- Si compie una decisione $z \in (Z, \mathcal{T}, \rho_x)$, che comporta un **costo** $W_\theta(z)$ (funzione a valori in $[0, 1]$).

Definizione (procedura decisionale)

È la funzione ρ che manda x in ρ_x (si richiede che sia un nucleo Markoviano).

Definizione (rischio)

Fissati W e θ , per ogni ρ è il seguente integrale:

$$R(\theta, \rho) \stackrel{\text{def}}{=} \int_{\Omega} \left(\int_Z W_{\theta}(z) \rho_x(dz) \right) P_{\theta}(dx)$$

Obiettivo: minimizzare il rischio.

Definizione (rischio)

Fissati W e θ , per ogni ρ è il seguente integrale:

$$R(\theta, \rho) \stackrel{\text{def}}{=} \int_{\Omega} \left(\int_Z W_{\theta}(z) \rho_x(dz) \right) P_{\theta}(dx)$$

Obiettivo: minimizzare il rischio.

Definizione (esperimento statistico)

È una famiglia di probabilità definite su (Ω, \mathcal{A}) e indicizzate mediante Θ , detto **spazio dei parametri**:

$$\mathcal{E} \stackrel{\text{def}}{=} \{P_{\theta} \mid \theta \in \Theta\}$$

Definizione (rischio)

Fissati W e θ , per ogni ρ è il seguente integrale:

$$R(\theta, \rho) \stackrel{\text{def}}{=} \int_{\Omega} \left(\int_Z W_{\theta}(z) \rho_x(dz) \right) P_{\theta}(dx)$$

Obiiettivo: minimizzare il rischio.

Definizione (esperimento statistico)

È una famiglia di probabilità definite su (Ω, \mathcal{A}) e indicizzate mediante Θ , detto **spazio dei parametri**:

$$\mathcal{E} \stackrel{\text{def}}{=} \{P_{\theta} \mid \theta \in \Theta\}$$

Dati \mathcal{E} e W , si definisce allora l'insieme di funzioni $\mathfrak{R}(\mathcal{E}, W)$. Una funzione $r : \Theta \rightarrow \mathbb{R}$ vi appartiene se esiste una procedura decisionale ρ tale che $R(\theta, \rho) \leq r(\theta)$. L'insieme $\mathfrak{R}(\mathcal{E}, W)$ è convesso.

1-2. Una distanza tra esperimenti

Supponiamo di dover effettuare un esperimento a scelta tra \mathcal{E} , definito su (Ω, \mathcal{A}) , e \mathcal{F} , definito su (Ω', \mathcal{B}) , ove lo spazio dei parametri Θ è comune.

1-2. Una distanza tra esperimenti

Supponiamo di dover effettuare un esperimento a scelta tra \mathcal{E} , definito su (Ω, \mathcal{A}) , e \mathcal{F} , definito su (Ω', \mathcal{B}) , ove lo spazio dei parametri Θ è comune.

Definizione (deficit)

Si indica con $\delta(\mathcal{E}, \mathcal{F})$, ed è il minimo $\delta \in [0, 1]$ tale che valga:

$$\mathfrak{R}(\mathcal{F}, W) + \delta \subseteq \overline{\mathfrak{R}(\mathcal{E}, W)},$$

per ogni funzione di costo W .

1-2. Una distanza tra esperimenti

Supponiamo di dover effettuare un esperimento a scelta tra \mathcal{E} , definito su (Ω, \mathcal{A}) , e \mathcal{F} , definito su (Ω', \mathcal{B}) , ove lo spazio dei parametri Θ è comune.

Definizione (deficit)

Si indica con $\delta(\mathcal{E}, \mathcal{F})$, ed è il minimo $\delta \in [0, 1]$ tale che valga:

$$\mathfrak{R}(\mathcal{F}, W) + \delta \subseteq \overline{\mathfrak{R}(\mathcal{E}, W)},$$

per ogni funzione di costo W .

Osserviamo che $\delta(\mathcal{E}, \mathcal{F}) \neq \delta(\mathcal{F}, \mathcal{E})$.

Definizione (distanza)

Si indica con $\Delta(\mathcal{E}, \mathcal{F})$, ed è il massimo tra $\delta(\mathcal{E}, \mathcal{F})$ e $\delta(\mathcal{F}, \mathcal{E})$.

Osservazioni:

- Possono esistere esperimenti differenti a distanza 0: li chiameremo **dello stesso tipo, o equivalenti**;
- Diremo che \mathcal{E} è **migliore** di \mathcal{F} se $\delta(\mathcal{E}, \mathcal{F}) = 0$.

Osservazioni:

- Possono esistere esperimenti differenti a distanza 0: li chiameremo **dello stesso tipo, o equivalenti**;
- Diremo che \mathcal{E} è **migliore** di \mathcal{F} se $\delta(\mathcal{E}, \mathcal{F}) = 0$.

In effetti, se \mathcal{E} è migliore di \mathcal{F} , si possono prevedere le funzioni di rischio di \mathcal{F} analizzando quelle di \mathcal{E} . Se \mathcal{E} e \mathcal{F} sono poco distanti, o addirittura equivalenti, si possono studiare con buona approssimazione le funzioni di rischio di \mathcal{F} analizzando quelle di \mathcal{E} .

1-3. Rappresentazione canonica di Blackwell

Dato $\mathcal{E} = \{P_\theta \mid \theta \in \Theta\}$, con Θ di cardinalità finita k , sia $S = P_{\theta_1} + \dots + P_{\theta_k}$. Allora, per ogni $i = 1, \dots, k$, esiste la densità $f_i = dP_{\theta_i}/dS$.

1-3. Rappresentazione canonica di Blackwell

Dato $\mathcal{E} = \{P_\theta \mid \theta \in \Theta\}$, con Θ di cardinalità finita k , sia $S = P_{\theta_1} + \dots + P_{\theta_k}$. Allora, per ogni $i = 1, \dots, k$, esiste la densità $f_i = dP_{\theta_i}/dS$.

È allora definita una funzione $\mathbf{v} : \Omega \rightarrow U(\Theta)$:

$$\mathbf{v}(x) = [f_1(x) \quad \dots \quad f_k(x)]^t,$$

ove ogni coordinata è non negativa, e la somma delle coordinate è pari a 1. Dunque $v(x) \in U(\Theta)$ (**simplexso standard** di \mathbb{R}^k).

1-3. Rappresentazione canonica di Blackwell

Dato $\mathcal{E} = \{P_\theta \mid \theta \in \Theta\}$, con Θ di cardinalità finita k , sia $S = P_{\theta_1} + \dots + P_{\theta_k}$. Allora, per ogni $i = 1, \dots, k$, esiste la densità $f_i = dP_{\theta_i}/dS$.

È allora definita una funzione $\mathbf{v} : \Omega \rightarrow U(\Theta)$:

$$\mathbf{v}(x) = [f_1(x) \quad \dots \quad f_k(x)]^t,$$

ove ogni coordinata è non negativa, e la somma delle coordinate è pari a 1. Dunque $v(x) \in U(\Theta)$ (**simplexso standard** di \mathbb{R}^k).

Definizione (misura canonica di Blackwell)

Si indica con $m_{\mathcal{E}}$, ed è la misura immagine di S mediante \mathbf{v} , chiaramente definita su $U(\Theta)$.

Definizione (rappresentazione canonica di Blackwell)

Dato \mathcal{E} , e determinata la sua misura canonica di Blackwell $m_{\mathcal{E}}$, è l'esperimento $\mathcal{F} = \{Q_i \mid i = 1, \dots, k\}$ dove $Q_i(d\mathbf{u}) = u_i m_{\mathcal{E}}(d\mathbf{u})$.

Definizione (rappresentazione canonica di Blackwell)

Dato \mathcal{E} , e determinata la sua misura canonica di Blackwell $m_{\mathcal{E}}$, è l'esperimento $\mathcal{F} = \{Q_i \mid i = 1, \dots, k\}$ dove $Q_i(d\mathbf{u}) = u_i m_{\mathcal{E}}(d\mathbf{u})$.

Dati \mathcal{E} e \mathcal{G} , vale (Φ_1 è un ben preciso sottoinsieme di funzioni misurabili):

$$\Delta(\mathcal{E}, \mathcal{G}) = \sup_{\phi \in \Phi_1} \left| \int_{U(\Theta)} \phi(\mathbf{u}) m_{\mathcal{E}}(d\mathbf{u}) - \int_{U(\Theta)} \phi(\mathbf{u}) m_{\mathcal{G}}(d\mathbf{u}) \right|$$

Definizione (rappresentazione canonica di Blackwell)

Dato \mathcal{E} , e determinata la sua misura canonica di Blackwell $m_{\mathcal{E}}$, è l'esperimento $\mathcal{F} = \{Q_i \mid i = 1, \dots, k\}$ dove $Q_i(d\mathbf{u}) = u_i m_{\mathcal{E}}(d\mathbf{u})$.

Dati \mathcal{E} e \mathcal{G} , vale (Φ_1 è un ben preciso sottoinsieme di funzioni misurabili):

$$\Delta(\mathcal{E}, \mathcal{G}) = \sup_{\phi \in \Phi_1} \left| \int_{U(\Theta)} \phi(\mathbf{u}) m_{\mathcal{E}}(d\mathbf{u}) - \int_{U(\Theta)} \phi(\mathbf{u}) m_{\mathcal{G}}(d\mathbf{u}) \right|$$

In particolare si ha $\Delta(\mathcal{E}, \mathcal{G}) \leq \|m_{\mathcal{E}} - m_{\mathcal{G}}\|_D$, ove $\|\mu\|_D = \sup_f \left| \int_{U(\Theta)} f d\mu \right|$, al variare delle funzioni f misurabili, 1-Lipschitziane con $\|f\|_{\infty} \leq 1$, è detta **norma dual-Lipschitziana** di μ .

1-4. Un'osservazione sulla convergenza di misure

Lemma

Siano $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, \mathcal{F} esperimenti aventi un comune spazio dei parametri Θ , di cardinalità finita k . Sono equivalenti:

1-4. Un'osservazione sulla convergenza di misure

Lemma

Siano $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, \mathcal{F} esperimenti aventi un comune spazio dei parametri Θ , di cardinalità finita k . Sono equivalenti:

- $(\mathcal{E}_n)_{n \in \mathbb{N}^+} \rightarrow_{\Delta} \mathcal{F}$;

1-4. Un'osservazione sulla convergenza di misure

Lemma

Siano $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, \mathcal{F} esperimenti aventi un comune spazio dei parametri Θ , di cardinalità finita k . Sono equivalenti:

- $(\mathcal{E}_n)_{n \in \mathbb{N}^+} \rightarrow_{\Delta} \mathcal{F}$;
- $(m_{\mathcal{E}_n})_{n \in \mathbb{N}^+} \rightarrow_{\|\cdot\|_D} m_{\mathcal{F}}$;

1-4. Un'osservazione sulla convergenza di misure

Lemma

Siano $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, \mathcal{F} esperimenti aventi un comune spazio dei parametri Θ , di cardinalità finita k . Sono equivalenti:

- $(\mathcal{E}_n)_{n \in \mathbb{N}^+} \rightarrow_{\Delta} \mathcal{F}$;
- $(m_{\mathcal{E}_n})_{n \in \mathbb{N}^+} \rightarrow_{\|\cdot\|_D} m_{\mathcal{F}}$;
- $(\mathbf{v}_n)_{n \in \mathbb{N}^+} \rightarrow_{\mathcal{L}} \mathbf{w}$, con $\mathbf{w}(x) = [dQ_{\theta_1}/dS' \quad \dots \quad dQ_{\theta_k}/dS']^t$.

1-4. Un'osservazione sulla convergenza di misure

Lemma

Siano $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, \mathcal{F} esperimenti aventi un comune spazio dei parametri Θ , di cardinalità finita k . Sono equivalenti:

- $(\mathcal{E}_n)_{n \in \mathbb{N}^+} \rightarrow_{\Delta} \mathcal{F}$;
- $(m_{\mathcal{E}_n})_{n \in \mathbb{N}^+} \rightarrow_{\|\cdot\|_D} m_{\mathcal{F}}$;
- $(\mathbf{v}_n)_{n \in \mathbb{N}^+} \rightarrow_{\mathcal{L}} \mathbf{w}$, con $\mathbf{w}(x) = [dQ_{\theta_1}/dS' \quad \dots \quad dQ_{\theta_k}/dS']^t$.

Se $\Theta = 2$, il risultato si specializza: si dimostra infatti che è equivalente studiare la convergenza in legge dei rapporti di verosimiglianza (ciò tornerà utile più avanti).

2-1. Distanza di Hellinger

La norma dual-Lipschitziana, al pari della norma L^1 , non ha un buon comportamento quando si considerano esperimenti effettuati in sequenza.

Notazione: si indica con $\mathcal{E} \otimes \mathcal{F}$ l'esperimento ottenuto effettuando prima \mathcal{E} e poi \mathcal{F} .

2-1. Distanza di Hellinger

La norma dual-Lipschitziana, al pari della norma L^1 , non ha un buon comportamento quando si considerano esperimenti effettuati in sequenza.

Notazione: si indica con $\mathcal{E} \otimes \mathcal{F}$ l'esperimento ottenuto effettuando prima \mathcal{E} e poi \mathcal{F} .

Definizione (distanza di Hellinger)

Date due probabilità P e Q su (Ω, \mathcal{A}) :

$$h^2(P, Q) \stackrel{\text{def}}{=} \frac{1}{2} \int_{\Omega} \left(\sqrt{dP} - \sqrt{dQ} \right)^2 = \frac{1}{2} \int_{\Omega} \left(\sqrt{f} - \sqrt{g} \right)^2 d\mu ,$$

con μ misura dominante, $f = dP/d\mu$, $g = dQ/d\mu$ (essa è ben definita).

Definizione (affinità)

Date P e Q su (Ω, \mathcal{A}) :

$$\rho(P, Q) \stackrel{\text{def}}{=} \int_{\Omega} \sqrt{dP dQ} = \int_{\Omega} \sqrt{f g} d\mu$$

Definizione (affinità)

Date P e Q su (Ω, \mathcal{A}) :

$$\rho(P, Q) \stackrel{\text{def}}{=} \int_{\Omega} \sqrt{dP dQ} = \int_{\Omega} \sqrt{f g} d\mu$$

Consideriamo ora una successione di esperimenti binari $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, con $\mathcal{E}_n = \{P_{0,n}, P_{1,n}\}$ definito su $(\Omega_n, \mathcal{A}_n)$. Considerato allora $\mathcal{E} = \{\mathbf{P}_0, \mathbf{P}_1\}$, definito sullo spazio prodotto (Ω, \mathcal{A}) e contenente le probabilità prodotto, direttamente dal teorema di Tonelli si ha:

$$\rho(\mathbf{P}_0, \mathbf{P}_1) = \prod_{n=1}^{+\infty} \rho(P_{0,n}, P_{1,n}), \quad 1 - h^2(\mathbf{P}_0, \mathbf{P}_1) = \prod_{n=1}^{+\infty} (1 - h^2(P_{0,n}, P_{1,n}))$$

A priori, infatti, vale $\rho(P, Q) = 1 - h^2(P, Q)$.

2-2. Trasformata di Hellinger

Definizione (trasformata di Hellinger)

È il seguente operatore:

$$\mathcal{H}(\mathcal{E})(\alpha) \stackrel{\text{def}}{=} \int_{\Omega} \prod_{\theta \in \Theta} (dP_{\theta})^{\alpha_{\theta}} ,$$

ove $\alpha = \{\alpha_{\theta} \mid \theta \in \Theta\}$ è un vettore opportuno.

2-2. Trasformata di Hellinger

Definizione (trasformata di Hellinger)

È il seguente operatore:

$$\mathcal{H}(\mathcal{E})(\alpha) \stackrel{\text{def}}{=} \int_{\Omega} \prod_{\theta \in \Theta} (dP_{\theta})^{\alpha_{\theta}},$$

ove $\alpha = \{\alpha_{\theta} \mid \theta \in \Theta\}$ è un vettore opportuno.

Proprietà interessanti:

- $\mathcal{H}(\mathcal{E} \otimes \mathcal{F})(\alpha) = \mathcal{H}(\mathcal{E})(\alpha) \cdot \mathcal{H}(\mathcal{F})(\alpha)$ (analogia con la trasformata di Fourier);

2-2. Trasformata di Hellinger

Definizione (trasformata di Hellinger)

È il seguente operatore:

$$\mathcal{H}(\mathcal{E})(\alpha) \stackrel{\text{def}}{=} \int_{\Omega} \prod_{\theta \in \Theta} (dP_{\theta})^{\alpha_{\theta}},$$

ove $\alpha = \{\alpha_{\theta} \mid \theta \in \Theta\}$ è un vettore opportuno.

Proprietà interessanti:

- $\mathcal{H}(\mathcal{E} \otimes \mathcal{F})(\alpha) = \mathcal{H}(\mathcal{E})(\alpha) \cdot \mathcal{H}(\mathcal{F})(\alpha)$ (analogia con la trasformata di Fourier);
- $\mathcal{H}(\mathcal{E}) \equiv \mathcal{H}(\mathcal{F}) \Leftrightarrow \Delta(\mathcal{E}, \mathcal{F}) = 0$;

2-2. Trasformata di Hellinger

Definizione (trasformata di Hellinger)

È il seguente operatore:

$$\mathcal{H}(\mathcal{E})(\alpha) \stackrel{\text{def}}{=} \int_{\Omega} \prod_{\theta \in \Theta} (dP_{\theta})^{\alpha_{\theta}},$$

ove $\alpha = \{\alpha_{\theta} \mid \theta \in \Theta\}$ è un vettore opportuno.

Proprietà interessanti:

- $\mathcal{H}(\mathcal{E} \otimes \mathcal{F})(\alpha) = \mathcal{H}(\mathcal{E})(\alpha) \cdot \mathcal{H}(\mathcal{F})(\alpha)$ (analogia con la trasformata di Fourier);
- $\mathcal{H}(\mathcal{E}) \equiv \mathcal{H}(\mathcal{F}) \Leftrightarrow \Delta(\mathcal{E}, \mathcal{F}) = 0$;
- $(\mathcal{E}_n)_{n \in \mathbb{N}^+} \rightarrow_{\Delta} \mathcal{F} \Leftrightarrow (\mathcal{H}(\mathcal{E}_n))_{n \in \mathbb{N}^+} \rightarrow \mathcal{H}(\mathcal{F})$ puntualmente su $U(\Theta)$.

3-1. Esperimenti poissoniani

Definizione (processo poissoniano di intensità λ)

In relazione a una funzione λ finitamente additiva, è un'applicazione Z a valori nella classe delle variabili aleatorie tale che:

- Se $A_1, A_2 \in \mathcal{A}$ sono disgiunti, allora $Z(A_1)$ e $Z(A_2)$ sono indipendenti, e $Z(A_1 \cup A_2) = Z(A_1) + Z(A_2)$;
- Per ogni $A \in \mathcal{A}$ $Z(A)$ ha legge poissoniana di parametro $\lambda(A)$.

3-1. Esperimenti poissoniani

Definizione (processo poissoniano di intensità λ)

In relazione a una funzione λ finitamente additiva, è un'applicazione Z a valori nella classe delle variabili aleatorie tale che:

- Se $A_1, A_2 \in \mathcal{A}$ sono disgiunti, allora $Z(A_1)$ e $Z(A_2)$ sono indipendenti, e $Z(A_1 \cup A_2) = Z(A_1) + Z(A_2)$;
- Per ogni $A \in \mathcal{A}$ $Z(A)$ ha legge poissoniana di parametro $\lambda(A)$.

Definizione (esperimento poissoniano)

Date delle funzioni finitamente additive μ_θ , con $\theta \in \Theta$, è l'esperimento $\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$, ove P_θ è la legge del processo poissoniano Z_θ di intensità μ_θ .

3-1. Esperimenti poissoniani

Definizione (processo poissoniano di intensità λ)

In relazione a una funzione λ finitamente additiva, è un'applicazione Z a valori nella classe delle variabili aleatorie tale che:

- Se $A_1, A_2 \in \mathcal{A}$ sono disgiunti, allora $Z(A_1)$ e $Z(A_2)$ sono indipendenti, e $Z(A_1 \cup A_2) = Z(A_1) + Z(A_2)$;
- Per ogni $A \in \mathcal{A}$ $Z(A)$ ha legge poissoniana di parametro $\lambda(A)$.

Definizione (esperimento poissoniano)

Date delle funzioni finitamente additive μ_θ , con $\theta \in \Theta$, è l'esperimento $\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$, ove P_θ è la legge del processo poissoniano Z_θ di intensità μ_θ .

Processi e esperimenti poissoniani sono molto usati in fisica, biologia, finanza.

3-2. Poissonizzazione

Notazione: $\mathcal{E}^k = \mathcal{E} \otimes \dots \otimes \mathcal{E}$, per $k \in \mathbb{N}^+$ (con \mathcal{E}^0 esperimento banale, avente tutte le probabilità uguali).

3-2. Poissonizzazione

Notazione: $\mathcal{E}^k = \mathcal{E} \otimes \dots \otimes \mathcal{E}$, per $k \in \mathbb{N}^+$ (con \mathcal{E}^0 esperimento banale, avente tutte le probabilità uguali).

Definizione (esperimento poissonizzato)

Dato \mathcal{E} , si indica con $\mathbf{P}(\mathcal{E})$, ed è l'esperimento (poissoniano) ottenuto considerando una variabile aleatoria N poissoniana di parametro 1, indipendente da tutto, ed effettuando \mathcal{E}^k se N assume il valore k .

3-2. Poissonizzazione

Notazione: $\mathcal{E}^k = \mathcal{E} \otimes \dots \otimes \mathcal{E}$, per $k \in \mathbb{N}^+$ (con \mathcal{E}^0 esperimento banale, avente tutte le probabilità uguali).

Definizione (esperimento poissonizzato)

Dato \mathcal{E} , si indica con $\mathbf{P}(\mathcal{E})$, ed è l'esperimento (poissoniano) ottenuto considerando una variabile aleatoria N poissoniana di parametro 1, indipendente da tutto, ed effettuando \mathcal{E}^k se N assume il valore k .

Se $\mathcal{H}(\mathcal{E}) \equiv \phi$, allora per α opportuno vale:

$$\mathcal{H}(\mathbf{P}(\mathcal{E}))(\alpha) = e^{\phi(\alpha)-1}$$

Definizione (esperimento infinitamente divisibile associato)

Data $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, e dato $\mathcal{F} = \bigotimes_{n=1}^{+\infty} \mathcal{E}_n$, si indica con $\mathbf{Aid}(\mathcal{F})$, ed è il prodotto delle versioni poissonizzate $(\mathbf{Aid}(\mathcal{F}) = \bigotimes_{n=1}^{+\infty} \mathbf{P}(\mathcal{E}_n))$.

Definizione (esperimento infinitamente divisibile associato)

Data $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, e dato $\mathcal{F} = \bigotimes_{n=1}^{+\infty} \mathcal{E}_n$, si indica con $\mathbf{Aid}(\mathcal{F})$, ed è il prodotto delle versioni poissonizzate $(\mathbf{Aid}(\mathcal{F}) = \bigotimes_{n=1}^{+\infty} \mathbf{P}(\mathcal{E}_n))$.

Proprietà interessanti:

- $\|\mathcal{H}(\mathbf{Aid}(\mathcal{F})) - \mathcal{H}(\mathcal{F})\|_{\infty}$ si controlla bene dall'alto;
- $\mathbf{Aid}(\mathcal{F})$, in molti casi, è una buona approssimazione di \mathcal{F} , nel senso della distanza Δ . A tal proposito...

Definizione (esperimento infinitamente divisibile associato)

Data $(\mathcal{E}_n)_{n \in \mathbb{N}^+}$, e dato $\mathcal{F} = \bigotimes_{n=1}^{+\infty} \mathcal{E}_n$, si indica con $\mathbf{Aid}(\mathcal{F})$, ed è il prodotto delle versioni poissonizzate $(\mathbf{Aid}(\mathcal{F}) = \bigotimes_{n=1}^{+\infty} \mathbf{P}(\mathcal{E}_n))$.

Proprietà interessanti:

- $\|\mathcal{H}(\mathbf{Aid}(\mathcal{F})) - \mathcal{H}(\mathcal{F})\|_\infty$ si controlla bene dall'alto;
- $\mathbf{Aid}(\mathcal{F})$, in molti casi, è una buona approssimazione di \mathcal{F} , nel senso della distanza Δ . A tal proposito...

Proposizione

Dati \mathcal{E}_1 e \mathcal{E}_2 aventi un insieme comune di parametri Θ di cardinalità k , esiste un modulo di continuità $\omega_k : [0, +\infty[\rightarrow [0, +\infty[$ tale che:

$$\|\mathcal{H}(\mathcal{E}_1) - \mathcal{H}(\mathcal{E}_2)\|_\infty \leq \varepsilon \Rightarrow \Delta(\mathcal{E}_1, \mathcal{E}_2) \leq \omega_k(\varepsilon)$$

4-1. La procedura di Lindeberg

Definizione (norma di Kolmogorov)

Date due probabilità P e Q su \mathbb{R} con funzioni di ripartizione F_P e F_Q , è la quantità $\|F_P - F_Q\|_\infty$.

4-1. La procedura di Lindeberg

Definizione (norma di Kolmogorov)

Date due probabilità P e Q su \mathbb{R} con funzioni di ripartizione F_P e F_Q , è la quantità $\|F_P - F_Q\|_\infty$.

In modo abbastanza elementare si dimostra quanto segue.

Teorema (di Lindeberg)

Siano X_1, \dots, X_n variabili aleatorie reali indipendenti, centrate e dotate di momento secondo (sia $\sigma_k^2 = \mathbb{E}[X_k^2]$). Siano $a = \max_{1 \leq k \leq n} \sigma_k$ e $s^2 = \sum_{k=1}^n \sigma_k^2$. Allora esiste una costante universale C tale che:

$$\|P - G\|_K \leq C \left(\frac{a}{s}\right)^{\frac{1}{4}},$$

ove P è la legge di $S = X_1 + \dots + X_n$, e $G \sim N(0, s^2)$.

Il teorema di Lindeberg è una generalizzazione del Teorema Limite Centrale.

Sia X una variabile aleatoria reale, con $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$, e siano Y_1, \dots, Y_n variabili aleatorie reali i.i.d., uguali in legge a $\frac{1}{\sqrt{n}}X$.

Il teorema di Lindeberg è una generalizzazione del Teorema Limite Centrale.

Sia X una variabile aleatoria reale, con $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$, e siano Y_1, \dots, Y_n variabili aleatorie reali i.i.d., uguali in legge a $\frac{1}{\sqrt{n}}X$.

Allora:

- Fissato $n \in \mathbb{N}^+$, ogni variabile Y_k ha varianza $1/n$, dunque $a = 1/\sqrt{n}$;
- Per ogni $n \in \mathbb{N}^+$ si ha $s^2 = 1$, dunque $s = 1$.

Il teorema di Lindeberg è una generalizzazione del Teorema Limite Centrale.

Sia X una variabile aleatoria reale, con $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$, e siano Y_1, \dots, Y_n variabili aleatorie reali i.i.d., uguali in legge a $\frac{1}{\sqrt{n}}X$.

Allora:

- Fissato $n \in \mathbb{N}^+$, ogni variabile Y_k ha varianza $1/n$, dunque $a = 1/\sqrt{n}$;
- Per ogni $n \in \mathbb{N}^+$ si ha $s^2 = 1$, dunque $s = 1$.

Per il teorema di Lindeberg, se P è la legge di S e $G \sim N(0, 1)$, si ha dunque:

$$\|P - G\|_K \leq C \left(\frac{1}{n}\right)^{\frac{1}{8}}$$

Operando il limite per $n \rightarrow +\infty$ si ottiene quindi la tesi (la convergenza in norma di Kolmogorov implica quella in norma di Lévy).

4-2. Shift-compattezza

Definizione (shift-compattezza)

Una successione $(X_n)_{n \in \mathbb{N}^+}$ si dice **shift-compatta** se esiste una successione di costanti $(b_n)_{n \in \mathbb{N}^+}$ tale che la successione delle leggi di probabilità delle variabili $Y_n = X_n - b_n$ è relativamente compatta, ossia tesa.

È una generalizzazione della nozione di relativa compattezza.

4-2. Shift-compattezza

Definizione (shift-compattezza)

Una successione $(X_n)_{n \in \mathbb{N}^+}$ si dice **shift-compatta** se esiste una successione di costanti $(b_n)_{n \in \mathbb{N}^+}$ tale che la successione delle leggi di probabilità delle variabili $Y_n = X_n - b_n$ è relativamente compatta, ossia tesa.

È una generalizzazione della nozione di relativa compattezza.

Si può dimostrare che:

- $(X_n)_{n \in \mathbb{N}^+}$ è shift-compatta se e solo se $(X_n - X'_n)_{n \in \mathbb{N}^+}$ è tesa (ove X_n e X'_n sono i.i.d.);
- $(X_n)_{n \in \mathbb{N}^+}$ è shift-compatta se e solo se $(X_n - \mu_n)_{n \in \mathbb{N}^+}$ è tesa, dove μ_n è una mediana di X_n .

4-3. Differenziabilità secondo Peano

Definizione (differenziabilità secondo Peano)

Una funzione $f : S \subseteq \mathbb{R} \rightarrow \mathbb{R}$ (supponiamo che $0 \in \overset{\circ}{S}$) lo è una volta in 0 se:

$$f(x) = f(0) + ax + o(x) , x \rightarrow 0 ,$$

e due volte se:

$$f(x) = f(0) + ax + \frac{1}{2}bx^2 + o(x^2) , x \rightarrow 0 .$$

Notazione: $P[0, 2]$ è la famiglia delle funzioni nulle in 0 e differenziabili due volte secondo Peano.

5-1. Array infinitesimali ed enunciato del teorema

Definizione (array infinitesimale)

Un array doppio $(X_{n,j})_{n \in \mathbb{N}^+}$, con la condizione che per ogni $n \in \mathbb{N}^+$ $(X_{n,j})_{j \in \mathbb{N}^+}$ è una successione di variabili indipendenti, si dice **infinitesimale** se per ogni $\varepsilon > 0$ vale:

$$\lim_{n \rightarrow +\infty} \sup_{j \in \mathbb{N}^+} \mathbb{P}(|X_{n,j}| > \varepsilon) = 0$$

5-1. Array infinitesimali ed enunciato del teorema

Definizione (array infinitesimale)

Un array doppio $(X_{n,j})_{n \in \mathbb{N}^+, j \in \mathbb{N}^+}$, con la condizione che per ogni $n \in \mathbb{N}^+$ $(X_{n,j})_{j \in \mathbb{N}^+}$ è una successione di variabili indipendenti, si dice **infinitesimale** se per ogni $\varepsilon > 0$ vale:

$$\lim_{n \rightarrow +\infty} \sup_{j \in \mathbb{N}^+} \mathbb{P}(|X_{n,j}| > \varepsilon) = 0$$

Notazione:

- $Y_{n,j} = X_{n,j} - m_{n,j}$, con $m_{n,j} = \mathbb{E}[U'_{n,j}]$, con $U'_{n,j}$ avente la legge di X condizionata all'evento $|X| \leq \tau$, con $\tau > 0$;
- $G_{n,j}$ è la probabilità su \mathbb{R} tale che $Y_{n,j} \sim G_{n,j}$.

Definizione (versione poissonizzata)

Data μ su \mathbb{R} , si indica con $\pi(\mu)$. Data $X \sim \mu$, è la legge di $\sum_{j=1}^N X_k$, ove N, X_1, X_2, \dots sono variabili indipendenti, tutte di legge μ eccetto N (di legge poissoniana di parametro 1).

Definizione (versione poissonizzata)

Data μ su \mathbb{R} , si indica con $\pi(\mu)$. Data $X \sim \mu$, è la legge di $\sum_{j=1}^N X_k$, ove N, X_1, X_2, \dots sono variabili indipendenti, tutte di legge μ eccetto N (di legge poissoniana di parametro 1).

Definizione (convoluzione)

Date μ, ν definite su \mathbb{R} , e date $X \sim \mu$ e $Y \sim \nu$ indipendenti, $\mu * \nu$ è la legge di $X + Y$.

Teorema (prima generalizzazione del T.L.C.)

Sia $(X_{n,j})_{n,j \in \mathbb{N}^+}$ un array infinitesimale, e supponiamo che $(S_n)_{n \in \mathbb{N}^+}$ sia shift-compatta ($S_n = X_{n,1} + X_{n,2} + \dots$). Se ora definiamo:

- \mathbf{G}_n , la legge di $\sum_{j=1}^{+\infty} Y_{n,j}$:

$$\mathbf{G}_n \stackrel{\text{def}}{=} \bigstar_{j=1}^{+\infty} G_{n,j} ;$$

- $\mathbf{Aid}(n)$, la legge di $\sum_{j=1}^{+\infty} W_{n,j}$, con $W_{n,j} \sim \pi(G_{n,j})$:

$$\mathbf{Aid}(n) \stackrel{\text{def}}{=} \bigstar_{j=1}^{+\infty} \pi(G_{n,j}) ,$$

allora:

$$\lim_{n \rightarrow +\infty} \|\mathbf{G}_n - \mathbf{Aid}(n)\|_D = 0$$

5-2. Conclusioni

Il risultato può condurre ad una generalizzazione riguardante gli esperimenti binari, tramite l'uso dei **tre lemmi di Le Cam**. Sia quindi $(\mathcal{E}_{n,j})_{n,j \in \mathbb{N}^+}$ un array doppio di esperimenti binari. Per ognuno, definiamo:

$$Y_{n,j} \stackrel{\text{def}}{=} \sqrt{\frac{dP_{0,n,j}}{dP_{1,n,j}} - 1}$$

Il logaritmo del rapporto di verosimiglianza è uguale a $f(Y_{n,j})$, con $f(y) = 2 \ln(1 + y) \in P[0, 2]$: quest'osservazione tornerà utile tra poco.

5-2. Conclusioni

Il risultato può condurre ad una generalizzazione riguardante gli esperimenti binari, tramite l'uso dei **tre lemmi di Le Cam**. Sia quindi $(\mathcal{E}_{n,j})_{n,j \in \mathbb{N}^+}$ un array doppio di esperimenti binari. Per ognuno, definiamo:

$$Y_{n,j} \stackrel{\text{def}}{=} \sqrt{\frac{dP_{0,n,j}}{dP_{1,n,j}} - 1}$$

Il logaritmo del rapporto di verosimiglianza è uguale a $f(Y_{n,j})$, con $f(y) = 2 \ln(1 + y) \in P[0, 2]$: quest'osservazione tornerà utile tra poco.

Detta poi $h_{n,j}$ la distanza di Hellinger $h(P_{0,n,j}, P_{1,n,j})$, supponiamo che:

- La quantità $\sum_{j=1}^{+\infty} h_{n,j}^2$ sia limitata superiormente;
- $\lim_{n \rightarrow +\infty} \sup_{j \in \mathbb{N}^+} h_{n,j}^2 = 0$.

Notazione: se $X \sim \mu$ e f è una funzione, $f(\mu)$ è la legge di $f(X)$.

Teorema (seconda generalizzazione del T.L.C.)

Siano soddisfatte le due ipotesi enunciate poc'anzi, e sia $f \in P[0, 2]$ a valori reali e misurabile. Considerate per ogni $n \in \mathbb{N}^+$:

- \mathbf{G}_n , la legge di $\sum_{j=1}^{+\infty} f(Y_{n,j})$:

$$\mathbf{G}_n \stackrel{\text{def}}{=} \bigstar_{j=1}^{+\infty} f(G_{n,j}) ;$$

- $\mathbf{Aid}(n)$, la legge di $\sum_{j=1}^{+\infty} W_{n,j}$, con $W_{n,j} \sim \pi(f(G_{n,j}))$:

$$\mathbf{Aid}_n \stackrel{\text{def}}{=} \bigstar_{j=1}^{+\infty} \pi(f(G_{n,j})) ,$$

allora:

$$\lim_{n \rightarrow +\infty} \|\mathbf{G}_n - \mathbf{Aid}(n)\|_D = 0$$