Bollettino Unione Matematica Italiana

Emanuele Dolera

Su una proprietà asintotica della distribuzione finale

Bollettino dell'Unione Matematica Italiana, Serie 9, Vol. **6** (2013), n.3, p. 741–748.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2013_9_6_3_741_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2013.

Bollettino U. M. I.
(9) **VI** (2013), 741-748

Su una proprietà asintotica della distribuzione finale (*)

EMANUELE DOLERA

Il seminario tratta un problema di inferenza statistica in ambito bayesiano, ovvero quello di descrivere il comportamento della distribuzione finale in presenza di un grande numero di osserazioni. L'obiettivo principale è mettere in luce il ruolo della *frequenza*, valutata su n osservazioni, ai fini della valutazione della legge di probabilità di un segmento di osservazioni future rispetto alle n già considerate, impostando il problema solamente su risultati basilari del calcolo delle probabilità, come il teorema di Bayes. Come corollario, si hanno condizioni entro le quali la stima bayesiana concorda, in maniera più marcata al crescere di n, con una stima ottenuta tramite un principio di massima verosimiglianza, tipico di un'impostazione frequentistita.

Per chiarire tali concetti, si consideri sullo spazio $(\Omega, \mathscr{F}, \mathsf{P})$ una successione $\{\xi_n\}_{n\geq 1}$ di variabili aleatorie (v.a.) *scambiabili* a valori in uno spazio polacco $\mathbb{X}$, munito di una metrica ρ e relativa σ-algebra di Borel $\mathscr{X}$. Denotando con $\tilde{e}_n(\cdot, \omega) := \dfrac{1}{n}\sum_{i=1}^{n} \delta_{\xi_i(\omega)}(\,\cdot\,)$ il *processo empirico*, che riassume in sè il concetto di frequenza, un ben noto teorema di de Finetti [4] afferma che esiste una misura di probabilità (m.d.p.) aleatoria $\tilde{p}$ su $(\mathbb{X}, \mathscr{X})$ per cui $\lim_{n\to\infty} \mathrm{d}_P(\tilde{e}_n, \tilde{p}) = 0$ P-quasi certamente, dove d_P denota la metrica di Prokhorov su $\mathbb{M}$, lo spazio di tutte le m.d.p. su $(\mathbb{X}, \mathscr{X})$. La misura $\gamma(\,\cdot\,) := \mathsf{P}[\,\tilde{p} \in \cdot]$, definita sulla σ-algebra $\mathscr{M}$ dei boreliani di $\mathbb{M}$, è la *distribuzione iniziale* e vale la rappresentazione (di de Finetti)

$$\mathsf{P}[\boldsymbol{\xi}^{(n)} \in A_1 \times \ldots \times A_n, \tilde{p} \in B] = \int_B \Big[\prod_{j=1}^{n} p(A_j)\Big] \gamma(\mathrm{d}p)$$

per ogni n e $A_1, \ldots, A_n \in \mathscr{X}$, con $\boldsymbol{\xi}^{(n)} := (\xi_1, \ldots, \xi_n)$. In questo contesto, la m.d.p. (aleatoria) $\mathsf{P}[\,\tilde{p} \in \cdot \mid \boldsymbol{\xi}^{(n)}]$ su $(\mathbb{M}, \mathscr{M})$ viene detta *distribuzione finale*, la quale descrive la legge di $\tilde{p}$ condizionatamente ai dati osservati. Inoltre, fissata una specifica versione della finale, si indicherà con $\gamma_n(\cdot\,; \mathbf{x}^{(n)})$ la valutazione di tale versione in $\boldsymbol{\xi}^{(n)} = \mathbf{x}^{(n)} := (x_1, \ldots, x_n)$. Una notevole proprietà asintotica della finale è messa in evidenza da un celebre teorema di Doob [7], che afferma che

$$\lim_{n\to\infty} \mathrm{D}_P(\mathsf{P}[\,\tilde{p} \in \cdot \mid \boldsymbol{\xi}^{(n)}]\,; \delta_{\tilde{p}}(\,\cdot\,)) = 0 \qquad (\mathsf{P} - q.c.) \tag{1}$$

* Lavoro svolto in collaborazione con E. Regazzini, Università degli Studi di Pavia.

essendo D_P la metrica di Prokhorov su $\mathbb{P}$, lo spazio di tutte le m.d.p. su $(\mathbb{M}, \mathscr{M})$. Dunque, considerando $\gamma_n(\cdot\ ; \mathbf{x}^{(n)})$ come una funzione definita su $(\mathbb{X}^\infty, \mathscr{X}^\infty)$, (1) si può rileggere come segue: esiste un $\mathbb{M}_* \in \mathscr{M}$ con $\gamma(\mathbb{M}_*) = 1$ tale che, per ogni p_* in $\mathbb{M}_*$, vale la formula

$$\lim_{n\to\infty} \mathrm{D}_P(\gamma_n(\cdot\ ; \mathbf{x}^{(n)})\ ; \delta_{p_*}(\,\cdot\,)) = 0 \tag{2}$$

per ogni successione $\mathbf{x}^{(\infty)} = (x_1, x_2, \dots)$ appartenente a $\mathbb{X}_*^\infty \in \mathscr{X}^\infty$, sottoinsieme di successioni tale che $p_*^\infty(\mathbb{X}_*^\infty) = 1$. Essendo il teorema di Doob solamente un risultato di esistenza, rimane aperto il problema di caratterizzare $\mathbb{M}_*$, o meglio di stabilire per quali p_* in $\mathbb{M}$ valga (2) per ogni successione $\mathbf{x}^{(\infty)} = (x_1, x_2, \dots)$ appartenente a $\mathbb{X}_*^\infty \in \mathscr{X}^\infty$, sottoinsieme di successioni tale che $p_*^\infty(\mathbb{X}_*^\infty) = 1$. Tale problema ha dato luogo alla seguente

DEFINIZIONE 1 (di consistenza bayesiana). – *Diciamo che $\gamma_n(\cdot\ ,\cdot)$ è consistente in $p_0 \in \mathbb{M}$ quando*

$$\lim_{n\to\infty} \mathrm{D}_P(\gamma_n(\cdot\ ; \mathbf{x}^{(n)})\ ; \delta_{p_0}(\,\cdot\,)) = 0 \tag{3}$$

per ogni successione $\mathbf{x}^{(\infty)} = (x_1, x_2, \dots)$ appartenente a $\mathbb{X}_0^\infty \in \mathscr{X}^\infty$, sottoinsieme di successioni tale che $p_0^\infty(\mathbb{X}_0^\infty) = 1$.

Tale importante definizione è dovuta a Diaconis e Freedman [5], i quali così ne sintetizzano la logica retrostante: "dopo aver assegnato una distribuzione iniziale, si generano immaginariamente successioni i.i.d. e, dopo aver calcolato la distribuzione finale, si valuta se quest'ultima riflette adeguatamente l'attuale conoscenza." Dopo i primi risultati di Freedman [8] in cui si mostrano esempi effettivi di inconsistenza, sono seguiti numerosissimi lavori sulla consistenza, tanto che essa rimane tuttora uno dei temi più studiati in ambito bayesiano. Tra i lavori pionieristici ricordiamo [5, 6, 7, 11] mentre tra i recenti [1, 2, 9, 12, 13].

In questo seminario si discute un problema connesso a quello della consistenza, che resta però formulato in ambito genuinamente bayesiano. A proposito, si precisi una ben determinata versione della m.d.p. finale $\gamma_n(\cdot\ ,\cdot)$ e, in corrispondenza di una assegnata successione di osservazioni $\mathbf{x}^{(\infty)} = (x_1, x_2, \dots) \in \mathbb{X}^\infty$, si ponga $e_n(\,\cdot\,) := \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}(\,\cdot\,)$ per $n = 1, 2, \dots$. Emerge, in tale contesto, il seguente

PROBLEMA 1 (delle probabilità inverse). – *Supponendo che $\lim_{n\to\infty} \mathrm{d}_P(e_n\ ; p_0) = 0$ per una m.d.p. $p_0 \in \mathbb{M}$, trovare, per ogni $\varepsilon > 0$, un indice n_ε per cui, per ogni $n \geq n_\varepsilon$,*

$$\mathrm{D}_P(\gamma_n(\cdot\ ; \mathbf{x}^{(n)})\ ; \delta_{e_n}) \leq \varepsilon\ . \tag{4}$$

Si noti che il problema delle probabilità inverse (p.i.) possiede una natura finitaria e non fa cenno a "veri valori" del paramtro e ad insiemi eccezionali, ma considera solo successioni di frequenze con una tendenza a concentrarsi su valori dati. Pertanto, in accordo con la sua formulazione originaria, il problema delle p.i. equivale ad indagare il ruolo della frequenza, calcolata su un numero finito di prove, nella valutazione della probabilità. In questo spirito, pur sotto convinzioni radicalmente differenti, si sono mossi i grandi maestri del passato: Laplace, von Mises, de Finetti, Romanovski.

Così come il problema della consistenza, anche il problema delle p.i. può non avere soluzione, come mostra il seguente esempio. Si consideri $\mathbb{X} = \{0, 1\}$, cosicchè $\mathbb{M} = \{$leggi di Bernoulli di parametro $\theta\}$, omeomorfo a $[0, 1]$, con misura iniziale $\gamma(\mathrm{d}\theta) = w\delta_0 + (1 - w)\mathcal{U}_{[y,1]}(\mathrm{d}\theta)$, con $w, y \in (0, 1)$. Sulla succesione delle osservazioni si assuma che $e_n(\{1\}) =: f_n \to 0$ e $f_n > 0$ per qualche n. Perchè il problema delle p.i. abbia soluzione si dovrebbe avere $\gamma([0, s), \mathbf{x}^{(n)}) \to 1$ per $n \to \infty$ e ogni $s > 0$, mentre si dimostra che $\gamma([0, s), \mathbf{x}^{(n)}) \to 0$ per $n \to \infty$ e ogni $s \in (0, y)$, mostrando l'insolubilià del problema in questo caso.

Le tecniche impiegate per risolvere il problema differiscono a seconda della conplessità dello spazio $\mathbb{M}$. Si parla di "caso dominato" quando le realizzazioni di $\tilde{p}$ che appartengono al supporto di γ sono tutte assolutamente continue (a.c.) tra loro. L'esempio più cristallino si ha quando lo spazio delle osservazioni è finito, $\mathbb{X} = \{a_1, a_2, \dots, a_{k+1}\}$, dove $\mathbb{M} = \Big\{ \sum_{i=1}^{k+1} \theta_i \delta_{a_i} \mid \theta_i \geq 0$ e $\sum_{i=1}^{k+1} \theta_i = 1 \Big\}$, che risulta omeomorfo al simplesso $T_k := \Big\{ (\theta_1, \dots, \theta_k) \in \mathbb{R}^k \mid \theta_i \geq 0$ e $\sum_{i=1}^{k} \theta_i \geq 1 \Big\}$ tramite la mappa $\mathrm{t}_k\Big(\sum_{i=1}^{k+1} \theta_i \delta_{a_i} \Big) := (\theta_1, \dots, \theta_k)$. Posto $\mathbf{f}^{(n)} = (f_1^{(n)}, \dots, f_k^{(n)}) := (e_n(\{a_1\}), \dots, e_n(\{a_k\}))$ e $\mathbf{f} = (f_1, \dots, f_k) := (p_0(\{a_1\}), \dots, p_0(\{a_k\}))$, La condizione $e_n \Rightarrow p_0$ equivale a $f_j^{(n)} \to f_j$ per $j = 1, \dots, k$. Allora la distribuzione finale è data da

$$\gamma(B; \mathbf{x}^{(n)}) = \frac{\int_B \Big[\prod_{j=1}^{k} p(\{a_j\})^{nf_j^{(n)}} \Big] \cdot \Big(1 - \sum_{j=1}^{k} p(\{a_j\}) \Big)^{nf_{k+1}^{(n)}} \gamma(\mathrm{d}p)}{\int_{\mathbb{M}} \Big[\prod_{j=1}^{k} p(\{a_j\})^{nf_j^{(n)}} \Big] \cdot \Big(1 - \sum_{j=1}^{k} p(\{a_j\}) \Big)^{nf_{k+1}^{(n)}} \gamma(\mathrm{d}p)}$$

$$= \frac{\int_{\mathrm{t}_k(B)} \Big[\prod_{j=1}^{k} \theta_j^{nf_j^{(n)}} \Big] \cdot \Big(1 - \sum_{j=1}^{k} \theta_j \Big)^{nf_{k+1}^{(n)}} \gamma \circ \mathrm{t}_k^{-1}(\mathrm{d}\boldsymbol{\theta})}{\int_{T_k} \Big[\prod_{j=1}^{k} \theta_j^{nf_j^{(n)}} \Big] \cdot \Big(1 - \sum_{j=1}^{k} \theta_j \Big)^{nf_{k+1}^{(n)}} \gamma \circ \mathrm{t}_k^{-1}(\mathrm{d}\boldsymbol{\theta})}$$

essendo B un elemento di $\mathscr{M}$. In questo contesto, vale il seguente

TEOREMA 1. – *Si assuma che* $\mathbf{f} \in \mathrm{supp}[\gamma \circ \mathrm{t}_k^{-1}]$ *e* $\gamma \circ \mathrm{t}_k^{-1}(\{\mathbf{f}\}) \neq 1$. *Sia poi* $B(\mathbf{f}, r)$ *una sfera euclidea di centro* $\mathbf{f}$ *e raggio* r *tale che* $\gamma \circ \mathrm{t}_k^{-1}(B(\mathbf{f}, r)^c \cap T_k) > 0$.

Allora, esistono due costanti positive α e β (esplicite) indipendenti da n tali che

$$\gamma(\mathrm{t}_k^{-1}(B(\mathbf{f},r)^c \cap T_k); \mathbf{x}^{(n)}) \leq \alpha e^{-\beta n} \tag{5}$$

in ciascuno dei casi seguenti

i) $\mathbf{f}$ *è interno a* T_k
ii) $\mathbf{f} \in \partial T_k$ *e* $\{B(\mathbf{f},\varepsilon) \cap \mathrm{supp}[\gamma \circ \mathrm{t}_k^{-1}]\} \setminus \{\mathbf{f}\} \neq \emptyset$ *per ogni* $\varepsilon > 0$.

Se (5) *vale, allora*

$$\mathrm{D}_P(\gamma_n(\cdot\ ; \mathbf{x}^{(n)})\ ; \delta_{e_n}) \leq ae^{-bn}$$

con opportune costanti positive a, b *che dipendono da* γ.

Una prima formulazione di questo teorema si trova in [6], mentre qui l'enunciato è stato migliorato considerando anche dati sul bordo del simplesso. La strategia dimostrativa coinvolge opportuni sviluppi di Taylor della verosimiglianza e semplici disuguaglianze, come già messo in evidenza da de Finetti [3].

Dopo l'invenzione del processo di Dirichlet, sono state introdotte incessantemente nuove distribuzioni iniziali non parametriche che non rientrano nel caso dominato (processi neutrali a destra, misure completamente aleatorie normalizzate, processi di Dirichlet gerarchici, misture nel senso di Antoniak, etc.). Si richiede in modo naturale un approccio generale al problema delle p.i. Uno studio pionieristico in questo senso è rappresentato da un articolo di Regazzini e Sazonov [10]. Gli autori considerano $\mathbb{X}$ totalmente limitato e introducono una successione di partizioni

$$\Pi_m := \{A_{m,1}, \dots, A_{m,k_m+1}\} \qquad (m = 1, 2, \dots)$$

dove ogni $A_{m,j} \in \mathscr{X}$. Posto $\varepsilon_m := \max_{1 \leq i \leq k_m+1} \mathrm{diam}(A_{m,i})$, si richiede che Π_{m+1} sia un raffinamento di Π_m per ogni m, $\mathscr{X}$ sia generata da $\bigcup_{m \geq 1} \Pi_m$ e $\varepsilon_m \downarrow 0$ per $m \to \infty$. Ora, per ogni p in $\mathbb{M}$ e m in $\mathbb{N}$ poniamo

$$p_{\varepsilon_m}(\,\cdot\,) := \sum_{i=1}^{k_m+1} p(A_{m,i})\delta_{a_{m,i}}$$

essendo gli $a_{m,i}$ punti di $A_{m,i}$. Si definiscono allora i nuovi elementi aleatori, da interpretarsi come "discretizzazione" delle osservazioni,

$$\xi_{j,m}(\omega) := a_{m,i} \qquad \text{se } \xi_j(\omega) \in A_{m,i}$$

per $j, m \in \mathbb{N}$ e $i = 1, \dots, k_m + 1$. Si pone anche $\xi_m^{(n)} := (\xi_{1,m}, \dots, \xi_{n,m})$. Si associa ad ogni $\mathbf{j} = (j_1, \dots, j_n) \in \{1, \dots, k_m + 1\}^n$ il rettangolo $R_{m,\mathbf{j}} := A_{m,j_1} \times \dots \times A_{m,j_n}$. Infine, senza ledere la generalità, si può assumere che $\int_{\mathbb{M}} p^n(R_{m,\mathbf{j}})\gamma(\mathrm{d}p) > 0$ per ogni $\mathbf{j}$. In questo contesto si ha

$$\mathsf{P}[\tilde{p} \in B \mid \xi_m^{(n)}](\omega) = \sum_{\mathbf{j}} 1\!\mathrm{l}_{R_{m,\mathbf{j}}}(\xi^{(n)}(\omega)) \frac{\int_B p^n(R_{m,\mathbf{j}})\gamma(\mathrm{d}p)}{\int_{\mathbb{M}} p^n(R_{m,\mathbf{j}})\gamma(\mathrm{d}p)}$$

e

$$\mathsf{P}[\tilde{p}_{\varepsilon_m} \in B \mid \xi_m^{(n)}](\omega)$$
$$= \sum_{\mathbf{j}} 1\!\mathrm{l}_{R_{m,\mathbf{j}}}(\xi^{(n)}(\omega)) \cdot \frac{\int_{\mathrm{t}_{k_m}(B)} \left[\prod_{i=1}^{k_m} \theta_i^{n_i(\mathbf{j})}\right] \cdot \left(1 - \sum_{i=1}^{k_m} \theta_i\right)^{n_{k_m+1}(\mathbf{j})} \gamma \circ \mathrm{t}_{k_m}^{-1}(\mathrm{d}\boldsymbol{\theta})}{\int_{T_{k_m}} \left[\prod_{i=1}^{k_m} \theta_i^{n_i(\mathbf{j})}\right] \cdot \left(1 - \sum_{i=1}^{k_m} \theta_i\right)^{n_{k_m+1}(\mathbf{j})} \gamma \circ \mathrm{t}_{k_m}^{-1}(\mathrm{d}\boldsymbol{\theta})}$$

dove $n_i(\mathbf{j}) := \sharp\{h \in \{1, \dots, n\} \mid j_h = i\}$ e $B \in \mathscr{M}$. Si giunge così alla seguente disuguaglianza, direttamente riconducibile al problema delle p.i.

$$\begin{aligned}
&\mathrm{D}_P(\gamma_n(\cdot\,; \mathbf{x}^{(n)})\,; \delta_{e_n}(\,\cdot\,)) \\
&\le 2\mathrm{D}_P(\delta_{p_0}(\,\cdot\,); \delta_{e_n}(\,\cdot\,)) + 2\mathrm{D}_P(\delta_{p_0}(\,\cdot\,); \delta_{p_{0,\varepsilon_m}}(\,\cdot\,)) \\
&+ \mathrm{D}_P(\delta_{e_{n,\varepsilon_m}}(\,\cdot\,)\,; \delta_{e_n}(\,\cdot\,)) + \mathrm{D}_P(\delta_{e_{n,\varepsilon_m}}(\,\cdot\,)\,; \mathsf{P}[\tilde{p}_{\varepsilon_m} \in \cdot \mid \xi_m^{(n)}](\omega)) \\
&+ \mathrm{D}_P(\mathsf{P}[\tilde{p} \in \cdot \mid \xi_m^{(n)}](\omega)\,; \mathsf{P}[\tilde{p}_{\varepsilon_m} \in \cdot \mid \xi_m^{(n)}](\omega)) \\
&+ \mathrm{D}_P(\mathsf{P}[\tilde{p} \in \cdot \mid \xi_m^{(n)}](\omega)\,; \gamma_n(\cdot\,; \mathbf{x}^{(n)}))
\end{aligned} \tag{6}$$

Ora, il primo termine è chiaramente infinitesimo in n, avendo supposto $e_n \Rightarrow p_0$. Il secondo e il terzo termine sono controllati da $2\varepsilon_m$ e ε_m rispettivamente, per le proprietà elementari delle distanza di Prokhorov. Il quinto termine è anche controllato da ε_m, per la Proposizione 1 in [10]. Il quarto termine è controllato da un'espressione del tipo ae^{-bn}, per il Teorema 1. Per concludere, bisogna prestare attenzione al fatto che a, b dipendono da m implicitamente. Una strategia efficace consiste nel cercare limitazioni sull'ultimo termine in (6) con maggioranti infinitesimi indipendenti da n. Un'ipotesi che asseconda tale scopo è che esista una versione specifica $\overline{\gamma}_n(\cdot\,;\cdot)$ della finale (determinabile a priori) per cui

$$\mathrm{D}_P(\overline{\gamma}_n(\cdot\,; \mathbf{y}^{(n)})\,; \overline{\gamma}_n(\cdot\,; \mathbf{x}^{(n)})) \le C_\alpha \max_{1 \le i \le n} \rho(x_i, y_i)^\alpha \tag{7}$$

per ogni $n \in \mathbb{N}$, $\mathbf{y}^{(n)}, \mathbf{x}^{(n)} \in \mathbb{X}^n$, essendo $\alpha \in (0, 1]$ e $C_\alpha > 0$ indipendenti da n, $\mathbf{y}^{(n)}, \mathbf{x}^{(n)}$. Infatti, si può affermare che

PROPOSIZIONE 1. – *Se vale* (7) *allora l'ultimo termine in* (6) *è limitato da* $C'_\alpha \varepsilon_m^{\alpha/2}$ *e il problema delle p.i. (relativamente alla versione* $\overline{\gamma}_n$*) ammette una soluzione esplicita.*

Tale proposizione riconduce il problema nell'individuare condizioni facilmente verificabili che garantiscano la validità di (7). A proposito si dà la seguente

PROPOSIZIONE 2. – *Si assuma che* $\mathbb{X}$ *sia un aperto limitato di* $\mathbb{R}^d$ *con bordo lipschitziano. Si supponga poi che* $A \mapsto \mathsf{P}[\xi^{(n)} \in A]$, *con* $A \in \mathscr{X}^n$, *sia a.c. rispetto alla misura di Lebesgue nd-dimensionale. Si ponga infine* $F_\phi : \mathbf{x}^{(n)} \mapsto \int_{\mathbb{M}} \phi(p)\gamma(\mathrm{d}p\ ; \mathbf{x}^{(n)})$. *Se esistono due costanti* $\tau > 1$ *e* C_τ *tali per cui*

$$\left| \int_{\mathbb{X}^n} \partial_j \zeta(\mathbf{x}^{(n)}) F_\phi(\mathbf{x}^{(n)}) \mathrm{d}\mathbf{x}^{(n)} \right| \le \frac{C_\tau}{n} \|\zeta\|_{L^{nd\tau'}(\mathbb{X}^n)} \tag{8}$$

per ogni $n \in \mathbb{N}$, $\zeta \in C_c^\infty(\mathbb{X}^n)$ *e* $j = 1, \ldots, nd$, *dove* τ' *è l'esponente coniugato di* τ, *allora esiste una versione* $\overline{\gamma}_n$ *della finale che soddisfa* (7).

La dimostrazione di questo fatto è basata su un'applicazione di una celebre disuguaglianza di Morrey. In ogni caso, siccome la condizione (8) non è di agevole verifica, si può ricorrere al metodo delle partizioni e scrivere

$$\int_{\mathbb{X}^n} \partial_h \zeta(\mathbf{x}^{(n)}) F_\phi(\mathbf{x}^{(n)}) \mathrm{d}\mathbf{x}^{(n)} = \lim_{m \to \infty} \sum_{\mathbf{j}} L_{\mathbf{j}}[\phi] \int_{R_{m,\mathbf{j}}} \partial_h \zeta(\mathbf{x}^{(n)}) \mathrm{d}\mathbf{x}^{(n)} \tag{9}$$

dove

$$L_{\mathbf{j}}[\phi] := \frac{\int_{T_{k_m}} \phi\Big(\sum_{i=1}^{k_m+1} \theta_i \delta_{a_i}\Big) \Big[\prod_{i=1}^{k_m} \theta_i^{n_i(\mathbf{j})}\Big] \cdot \Big(1 - \sum_{i=1}^{k_m} \theta_i\Big)^{n_{k_m+1}(\mathbf{j})} \gamma \circ \mathrm{t}_{k_m}^{-1}(\mathrm{d}\boldsymbol{\theta})}{\int_{T_{k_m}} \Big[\prod_{i=1}^{k_m} \theta_i^{n_i(\mathbf{j})}\Big] \cdot \Big(1 - \sum_{i=1}^{k_m} \theta_i\Big)^{n_{k_m+1}(\mathbf{j})} \gamma \circ \mathrm{t}_{k_m}^{-1}(\mathrm{d}\boldsymbol{\theta})}$$

e ϕ è una funzione limitata e lipschitziana, con norma lipschitz minore o uguale a 1 rispetto alla metrica di Wasserstein. Si enuncia allora la seguente

PROPOSIZIONE 3. – *La disuguaglianza* (8) *resta verificata se esiste una costante* $C > 0$ *tale che*

$$|L_{\mathbf{j}}[\phi] - L_{\mathbf{j}'}[\phi]| \le C \frac{\varepsilon_m}{n} \tag{10}$$

per ogni $n, m \in \mathbb{N}$ *e per ogni* ϕ *limitata e lipschitziana, con norma lipschitz minore o uguale a 1 rispetto alla metrica di Wasserstein, essendo*

$$\mathbf{j} = (j_1, j_2, \ldots, j_l, \ldots, j_n)$$
$$\mathbf{j}' = (j_1, j_2, \ldots, j_l', \ldots, j_n)$$

e A_{m,j_l} *contiguo ad* $A_{m,j_l'}$.

Ora la disuguaglianza (10) ha il pregio di riguardare solo di distribuzioni finali di dimensione finita, che possono spesso essere note. Non vi è poi nessun problema a livello di scelta delle versioni. Inoltre, (10) può essere verificata a mano in alcuni casi notevoli, oppure mediante la

PROPOSIZIONE 4. – *La disuguaglianza* (10) *resta verificata se esistono due costanti* $C_1, C_2 > 0$ *tale che*

$$\mathrm{d}_W\Big(\sum_{i=1}^{k_m+1} \overline{\theta}_i \delta_{a_{m,i}} \; ; \sum_{i=1}^{k_m+1} \overline{\theta}_i^{'} \delta_{a_{m,i}}\Big) \leq C_1 \frac{\varepsilon_m}{n} \tag{11}$$

$$Var[\mathcal{L}] + Var[\mathcal{L}^{'}] \leq C_2 \frac{1}{n} \tag{12}$$

essendo $\mathcal{L}, \mathcal{L}^{'}$ *le leggi finali finito dimensionali con campioni individuati, rispettivamente da* $\mathbf{j}$ *e* $\mathbf{j}^{'}$, *e* $(\overline{\theta}_1, \ldots, \overline{\theta}_{k_m+1}), (\overline{\theta}_1^{'}, \ldots, \overline{\theta}_{k_m+1}^{'})$ *i relativi valori attesi.*

Per concludere la trattazione dell'argomento, utilizzando i precedenti enunciati si sono analizzate diverse classi di misure di probabilità non-parametriche, comunemente utilizzate nella pratica, quali le misure normalizzate di James, Lijoi e Prünster, il processo di Ongaro e il processo Beta-stacy di Walker e Muliere, ottenendo risultati di consistenza legati al problema delle p.i. finora non noti in letteratura.

REFERENCES

[1] BARRON A., SCHERVISH M.J. e WASSERMAN L., *The consistency of posterior distribution in nonparametric problems.* Ann. Statist. **27** (1999), 536-561.

[2] CHOI T. e RAMAMOORTHI R.V., *Remarks on consistence of posterior distributions. Bertrand Clarke and Subhashis Ghoshal, eds., Pushing the limits of contemporary statistics: Contributions in honor of Jayanta K. Ghosh* (IMS) (2008), 170-186.

[3] DE FINETTI B., *Funzione caratteristica di un fenomeno aleatorio.* Atti Reale Accademia Nazionale dei Lincei. **4** (1930), 86-133.

[4] DE FINETTI B., *La prévision: ses lois logiques, ses sources subjectives.* Ann. Inst. H. Poincaré. **7** (1939), 1-68.

[5] DIACONIS P. e FREEDMAN D.A., *On the consistency of Bayes estimates (with discussion)* Ann. Statist. **14** (1986), 1-26.

[6] DIACONIS P. e FREEDMAN D.A., *On the uniform consistency of Bayes estimates for multinomial probabilities*. Ann. Statist. **18** (1990), 1317-1327.

[7] DOOB J.L., *Application of the theory of martingales*. In *Le calcul des Probabilités et ses Applications* 23-27. Colloques Internationaux du Centre National de la Recherche Scientifique, 1949, Paris.

[8] FREEDMAN D.A., *On the asymptotic behavior of Bayes' estimates in the discrete case.* Ann. Math. Statist. **34** (1963), 1194-1216.

[9] GHOSHAL S., GHOSH J.K. e VAN DER VAART A.W., *Convergence rates of posterior distributions.* Ann. Statist. **28** (2000), 500-531.

[10] Regazzini E. e Sazonov V., *Approximation of laws of random probabilities by mixtures of Dirichlet distributions with applications to nonparametric bayesian inference*. Theory Probab. Appl. **45** (2001), 93-110.
[11] Schwartz L., *On Bayes procedures*. Z. Wahrsch. Verw. Gebiete. **4** (1965), 10-26.
[12] Shen X. e Wasserman L., *Rates of convergence of posterior distributions*. Ann. Statist. **29** (2001), 666-686.
[13] Walker S., *New approaches to Bayesian consistency*. Ann. Statist. **32** (2004), 2028-2043.

Emanuele Dolera
Dipartimento di Matematica Pura e Applicata "Giuseppe Vitali"
Università degli Studi di Modena e Reggio Emilia
Via Campi 213/b, 41125 Modena, Italy
E-mail: emanuele.dolera@unimore.it emanuele.dolera@unipv.it

Received December 1, 2012 and in revised form December 1, 2013