Note di Teoria della Probabilità.

In queste brevi note, si richiameranno alcuni risultati di Teoria della Probabilità, riguardanti le conseguenze elementari delle definizioni di probabilità e σ-algebra. Si ricordano le definizioni fondamentali.

1. Una σ-algebra $\mathcal{F}$ è una famiglia di sottoinsiemi di un insieme Ω che soddisfa le seguenti proprietà:
   (a) $\Omega \in \mathcal{F}$;
   (b) è chiusa rispetto al complemento: se $A \in \mathcal{F}$, allora $A^c \in \mathcal{F}$;
   (c) è chiusa rispetto all'unione numerabile: se $A_i \in \mathcal{F}$, $i = 1, 2, ...$, allora
   $$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}.$$
2. Uno spazio di probabilità è una terna $\mathcal{S} = \{\Omega, \mathcal{F}, P\}$, dove Ω, è un insieme generico denominato *spazio campione*, $\mathcal{F}$ è una σ-algebra di sottoinsiemi di Ω (denominati *eventi*), e P è una funzione definita sugli insiemi di $\mathcal{F}$ e a valori reali, che soddisfa i seguenti assiomi:
   (a) $P[\Omega] = 1$;
   (b) per ogni $A \in \mathcal{F}$, $P[A] \geq 0$;
   (c) per ogni sequenza numerabile di insiemi $A_i \in \mathcal{F}$, $i = 1, 2, ...$, a due a due disgiunti, $A_i \cap A_j = \emptyset$ per $i \neq j$, si ha (σ-additività)
   $$P[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} P[A_i].$$

Valgono i seguenti risultati, che possono essere dimostrati come semplici esercizi (nella dimostrazione, occorre utilizzare solamente le definizioni e le proprietà enunciate precedentemente). Nel seguito, denoteremo con $\mathcal{F}$ una σ-algebra di sottoinsiemi di un insieme Ω, e con $P[.]$ una funzione probabilità definita sugli elementi di $\mathcal{F}$. Nel caso $\Omega = \mathbb{R}$, denoteremo con $\mathcal{B}$ la σ-algebra di Borel, ovvero la *minima* σ-algebra che contiene le semirette $(-\infty, a]$, con $a \in \mathbb{R}$. La σ-algebra di Borel è minima, nel senso che una qualsiasi σ-algebra che ha come elementi le semirette, deve comprendere tutti gli insiemi che costituiscono la σ-algebra di Borel. Ad esempio, la classe di tutti i sottoinsiemi di $\mathbb{R}$ è ovviamente una σ-algebra che contiene le semirette (perché?): essa è tuttavia più ampia della σ-algebra di Borel, dato che è possibile descrivere sottoinsiemi di $\mathbb{R}$ che non appartengono a $\mathcal{B}$.

1. Dimostrare che $\emptyset \in \mathcal{F}$.

   Prova. Le proprietà 1.(a) e 1.(b) assicurano che $\emptyset = \Omega^c \in \mathcal{F}$.

2. Dimostrare che $\mathcal{F}$ è chiusa rispetto all'intersezione numerabile: se $A_i \in \mathcal{F}$, $i = 1, 2, ...,$ allora
$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}.$$

   Prova. Si ricorda, dalla teoria degli insiemi, la regola di De Morgan $(\bigcap_{i=1}^{\infty} A_i)^c = \bigcup_{i=1}^{\infty} A_i^c$. Si applicano poi le proprietà 1.(b) e 1.(c).

3. Dimostrare che $\mathcal{F}$ è chiusa rispetto all'unione finita di insiemi: se $A_i \in \mathcal{F}$, $i = 1, 2, ..., N$, allora
$$\bigcup_{i=1}^{N} A_i \in \mathcal{F}.$$

   Prova. È sufficiente considerare la sequenza infinita di insiemi $B_i = A_i$ per $i = 1, ..., N$, e $B_i = A_N$, per $i > N$ (si ripete dunque l'ultimo insieme). Ovviamente, $B_i \in \mathcal{F}$ per ogni i, e si ha dunque, per la 1.(c),
$$\bigcup_{i=1}^{N} A_i = \bigcup_{i=1}^{\infty} B_i \in \mathcal{F}.$$

4. Dimostrare che $\mathcal{F}$ è chiusa rispetto all'intersezione finita di insiemi: se $A_i \in \mathcal{F}$, $i = 1, 2, ..., N$, allora
$$\bigcap_{i=1}^{N} A_i \in \mathcal{F}.$$

   Prova. È sufficiente usare la regola di De Morgan $(\bigcap_{i=1}^{N} A_i)^c = \bigcup_{i=1}^{N} A_i^c$, la proprietà precedente e la 1.(b).

5. Si consideri la σ-algebra di Borel $\mathcal{B}$. Si dimostri che essa contiene ad esempio, per $a, b \in \mathbb{R}$, gli insiemi del tipo $(a, +\infty)$, $(a, b]$, i punti isolati $\{a\}$, (a, b), $[a, b]$, l'insieme dei numeri interi, naturali, razionali, irrazionali.

   Prova. Essendo per definizione $\mathcal{B}$ la più piccola σ-algebra che contiene le semirette $(-\infty, a]$, essa contiene $(a, +\infty) = \{(-\infty, a]\}^c$. Contiene $(a, b] = (-\infty, b] \cap (a, +\infty)$. Contiene i punti isolati, intersezione di un'infinità numerabile di intervalli $\{a\} = \bigcap_n (a - 1/n, a]$. Contiene $(a, b) = (a, b] \cap \{b\}^c$. Contiene $[a, b] = (a, b] \cup \{a\}$. L'insieme dei numeri interi, dei razionali, dei naturali, sono unioni numerabili di punti isolati. Gli irrazionali sono l'insieme complementare dei razionali. Dunque $\mathcal{B}$ contiene tutti gli irrazionali in un intervallo, tutti gli interi negativi, le unioni di intervalli, ecc. Una classe estremamente ricca, anche se, come ricordato, possono essere descritti sottoinsiemi di $\mathbb{R}$ che non appartengono a $\mathcal{B}$.

6. Si consideri uno spazio di probabilità $\mathcal{S} = \{\Omega, \mathcal{F}, P\}$. Dimostrare che $P[\emptyset] = 0$. **Prova.** Nella dimostrazione, è necessario sfruttare i soli assiomi della probabilità.

   Si consideri la famiglia di insiemi $B_1 = \Omega$, $B_i = \emptyset$, $i > 1$. Ovviamente, $B_i \in \mathcal{F}$ e $B_i \cap B_j = \emptyset$, per $i \neq j$. Per la 2.(a) e la 2.(c), si può scrivere

$$1 = P[\Omega] = P[\bigcup_i B_i] = P[\Omega] + P[\emptyset] + P[\emptyset] + \ldots .$$

   Dunque

$$0 = P[\emptyset] + P[\emptyset] + \ldots .$$

   La somma di infiniti termini uguali maggiori o uguali a 0 (proprietà 2.(b)) può essere nulla solamente se tutti i termini sono nulli. Dunque $P[\emptyset] = 0$.

7. Si consideri uno spazio di probabilità $\mathcal{S} = \{\Omega, \mathcal{F}, P\}$. Dimostrare che la probabilità è semplicemente additiva, ovvero che, considerata la collezione finita di insieme $A_1, \ldots, A_N$, a due a due disgiunti, $A_i \cap A_j = \emptyset$ per $i \neq j$, si ha

$$P[\bigcup_{i=1}^{N} A_i] = \sum_{i=1}^{N} P[A_i].$$

   Prova. Si consideri la famiglia numerabile di insiemi $B_i = A_i$, $i = 1, \ldots, N$, e

   $B_i = \emptyset$, $i > N$. Si ha

$$P[\bigcup_{i=1}^{N} A_i] = P[\bigcup_{i=1}^{\infty} B_i] = P[B_1] + \ldots + P[B_N] + P[\emptyset] + \ldots = \sum_{i=1}^{N} P[A_i].$$

8. Si consideri consideri un evento $A \in \mathcal{F}$. Dimostrare che $P[A^c] = 1 - P[A]$.

   Prova. Si ha $\Omega = A \cup A^c$ e $A \cap A^c = \emptyset$. Per la proprietà precedente,

$$1 = P[\Omega] = P[A] + P[A^c].$$

9. Si consideri consideri un evento $A \in \mathcal{F}$. Dimostrare che $P[A] \leq 1$.

   Prova. Per la 2.(b), si ha $P[A^c] \geq 0$. Dalla relazione $0 \leq P[A^c] = 1 - P[A]$, si deduce $P[A] \leq 1$.

10. Siano A e B due eventi con $B \supset A$. Dimostrare che $P[B] \geq P[A]$.

    Prova. Per note proprietà degli insiemi, si ha $B = A \cup (B \cap A^c)$ (ci si aiuti con i diagrammi di Venn per la rappresentazione degli insiemi). Essendo $A \cap (B \cap A^c) = \emptyset$, si può scrivere

$$P[B] = P[A] + P[B \cap A^c] \geq P[A].$$

    Nella relazione precedente, si è usata la 2.(b).

11. Dimostrare che la probabilità è una funzione continua, nel senso che, data una successione *crescente* di eventi A_i, $i = 1, 2, ...$, (tale cioè che $A_{i+1} \supset A_i$), posto

$$\lim_{i \to +\infty} A_i \triangleq \bigcup_{i=1}^{+\infty} A_i = A,$$

si ha

$$\lim_{i \to +\infty} P[A_i] = P[\lim_{i \to +\infty} A_i].$$

Allo stesso modo, si può dimostrare che, data una successione *decrescente* di insiemi A_i (tale cioè che $A_{i+1} \subset A_i$), posto in questo caso

$$\lim_{i \to +\infty} A_i \triangleq \bigcap_{i=1}^{+\infty} A_i = A,$$

si ha

$$\lim_{i \to +\infty} P[A_i] = P[\lim_{i \to +\infty} A_i].$$

Prova. Dimostriamo la proprietà nel caso di una sequenza crescente di insiemi. Si consideri la Fig. 1.(a), dove sono rappresentati gli insiemi A_i e il loro limite A. Si consideri poi la famiglia di insiemi B_i, definiti ponendo $B_1 = A_1$ e $B_i = A_i \cap A_{i-1}^c$ per $i > 1$, mostrata in Fig. 1.(b). Ovviamente, i B_i sono a due a due disgiunti, e si ha

$$A_n = \bigcup_{i=1}^{n} B_i.$$

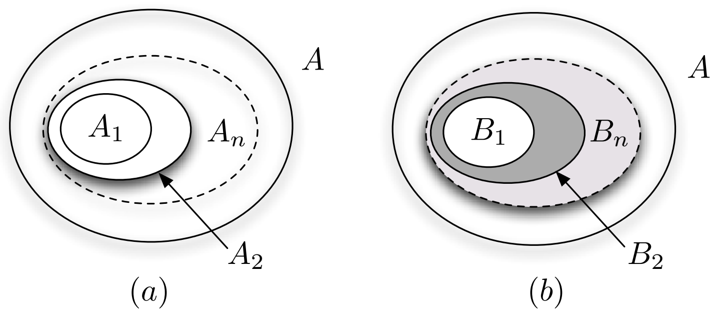

Figura 1: Sequenza crescente degli insiemi A_i (a), e successione B_i dei corrispondenti insiemi disgiunti (b).

Possiamo scrivere

$$P[A] \quad = \quad P[\lim_{i \to +\infty} A_i] = P[\cup_{i=1}^{+\infty} A_i]$$

$$
\begin{aligned}
&= P[\cup_{i=1}^{+\infty} B_i] = \sum_{i=1}^{+\infty} P[B_i] = \lim_{n\to+\infty} \sum_{i=1}^{n} P[B_i] \\
&= \lim_{n\to+\infty} P[\cup_{i=1}^{n} B_i] = \lim_{n\to+\infty} P[A_n]. \qquad (1)
\end{aligned}
$$

Si consideri ora il caso di una successione decrescente di insiemi, $A_{i+1} \subset A_i$,

$$\lim_{i\to+\infty} A_i \triangleq \bigcap_{i=1}^{+\infty} A_i = A.$$

Utilizzando la regola di De Morgan, possiamo scrivere $A^c = \cup_{i=1}^{+\infty} A_i^c$, con $A_{i+1}^c \supset A_i^c$. Gli insiemi complementari costituiscono dunque una successione di insiemi crescente con limite A^c, e, applicando il risultato appena dimostrato, si ha dunque

$$P[A^c] = \lim_{n\to+\infty} P[A_n^c].$$

Si ottiene pertanto

$$P[A] = 1 - P[A^c] = 1 - \lim_{n\to+\infty} P[A_n^c] = \lim_{n\to+\infty} (1 - P[A_n^c]) = \lim_{n\to+\infty} P[A_n].$$

Formalmente, una variabile aleatoria viene definita a partire da uno spazio di probabilità $\mathcal{S} = \{\Omega, \mathcal{F}, P\}$ come una funzione $x : \Omega \to \mathbb{R}$ *misurabile*, ovvero tale che, per ogni $a \in \mathbb{R}$, l'insieme $\{\omega : x(\omega) \leq a\}$ deve appartenere a $\mathcal{F}$. In altre parole, l'immagine inversa, tramite x, della semiretta $(-\infty, a]$, deve essere un insieme appartenente a $\mathcal{F}$. Non sarebbe troppo difficile dimostrare che la misurabilità di x garantisce che $\{\omega : x(\omega) \in B\}$ appartiene a $\mathcal{F}$ non solo per le semirette $B = (-\infty, a]$, ma anche se B è un generico insieme di Borel $B \in \mathcal{B}$. Risulta dunque possibile, a partire da x, definire lo spazio di probabilità *indotto*

$$\mathcal{S}_x = \{\mathbb{R}, \mathcal{B}, P_x\},$$

dove si pone, per ogni $B \in \mathcal{B}$, $P_x[B] = P[\{\omega : x(\omega) \in B\}]$, con $P[.]$ la funzione probabilità dello spazio di probabilità originario[1]. Per indicare $P_x[B]$, si usa spesso la notazione alternativa $P[x \in B]$. Ad esempio, per indicare $P_x[(-\infty, a]]$, possiamo usare la notazione $P[x \in (-\infty, a]]$ o anche $P[x \leq a]$.

La descrizione statistica di una v.a. consiste in sostanza nello specificare la funzione probabilità $P_x[B]$ per ogni insieme di Borel B. Fortunatamente, si può dimostrare[2], che P_x è univocamente determinata dai valori $P_x[(-\infty, a]]$. In altre parole, se conosciamo le probabilità $P[x \leq a]$, $a \in R$, in linea di principio possiamo calcolare la probabilità $P[x \in B]$ per ogni insieme di Borel B. La conoscenza della probabilità $F_x(a) \triangleq P[x \leq a]$, $a \in R$, è dunque sufficiente per la descrizione statistica di una v.a.: alla funzione $F_x(a)$ si dà il nome di *distribuzione* della v.a. x.

[1]Si potrebbe far vedere facilmente che lo spazio di probabilità indotto è ben definito e, in particolare, che $P_x[.]$ soddisfa effettivamente gli assiomi della funzione probabilità.

[2]La dimostrazione è data da un teorema del matematico greco Carathéodory (1873-1950).

Si noti che la descrizione statistica di x può in realtà essere data senza avere conoscenza dello spazio di probabilità originario $\mathcal{S}$, e senza neppure specificare esplicitamente la funzione $x(\omega)$, come nell'esempio che segue.

Esempio. Si considerino due spazi di probabilità $\mathcal{S}_1 = \{\Omega_1, \mathcal{F}_1, P_1\}$ e $\mathcal{S}_2 = \{\Omega_2, \mathcal{F}_2, P_2\}$. Lo spazio $\mathcal{S}_1$ modella il lancio di una moneta. Indicando con T l'esito *testa* e con C l'esito *croce*, si può porre

$$\Omega_1 = \{T, C\},$$

$$\mathcal{F}_1 = \{\{T, C\}, \{T\}, \{C\}, \emptyset\},$$

$$P_1[\{T, C\}] = 1, \quad P_1[\{T\}] = P_1[\{C\}] = 0.5, \quad P_1[\emptyset] = 0.$$

Lo spazio $\mathcal{S}_2$ modella il lancio di una dado. Indicando con f_i l'esito corrispondente all'uscita della faccia i-esima, si può porre

$$\Omega_2 = \{f_1, f_2, ..., f_6\}.$$

La σ-algebra $\mathcal{F}_2$ può essere scelta uguale alla classe dei $2^6 = 64$ sottoinsiemi di Ω_2, ponendo, per ogni $A = \cup_{k=1}^{K} \{f_{i_k}\} \in \mathcal{F}_2$, $P_2[A] = K/6$ (ad esempio, la probabilità che nel lancio del dado si osservi una delle facce $\{f_1, f_4, f_5\}$ risulta $P_2[\{f_1, f_4, f_5\}] = 3/6$).

Si consideri la v.a. x_1 definita a partire da $\mathcal{S}_1$ e così definita

$$x_1(\omega) = \begin{cases} 1, & \omega = T, \\ 0, & \omega = C. \end{cases}$$

Si consideri poi x_2 definita a partire da $\mathcal{S}_2$ e così definita

$$x_2(\omega) = \begin{cases} 1, & \omega \in \{f_1, f_2, f_3\}, \\ 0, & \omega \in \{f_4, f_5, f_6\}. \end{cases}$$

Risulta immediato verificare che x_1 e x_2 hanno la stessa descrizione statistica e, in particolare, hanno distribuzione

$$F_x(a) = \begin{cases} 0, & a < 0, \\ 0.5, & 0 \leq a < 1, \\ 1, & a \geq 1. \end{cases}$$

Si possono dimostrare facilmente le seguenti proprietà della distribuzione.

1. La funzione $F_x(a)$ è non descrescente, ovvero $b \geq a \Rightarrow F_x(b) \geq F_x(a)$.

   Prova. Il risultato è una conseguenza del fatto che $b \geq a$ implica $(-\infty, b] \supset (-\infty, a]$ e delle proprietà della funzione probabilità.

2. Si ha

$$\lim_{a \to +\infty} F_x(a) = 1, \quad \lim_{a \to +\infty} F_x(a) = 0.$$

Prova. Per $a \to +\infty$, si ha $(-\infty, a] \to (-\infty, +\infty) = \mathbb{R}$ e dunque

$$\lim_{a \to +\infty} F_x(a) = \lim_{a \to +\infty} P_x[(-\infty, a]] = P_x[\lim_{a \to +\infty} (-\infty, a]] = P_x[\mathbb{R}] = 1.$$

Nella relazione precedente, si è sfruttata la continuità della funzione probabilità. Per $a \to -\infty$, si procede analogamente, osservando che $(-\infty, a] \to \emptyset$ e che $P[\emptyset] = 1$.

3. La $F_x(a)$ è una funzione continua da destra, ovvero

$$\lim_{h \to 0^+} F_x(a + h) = F_x(a),$$

dove la notazione $h \to 0^+$ indica che h tende a zero per valori positivi.

Prova. È sufficiente osservare che $(-\infty, a + h]$ tende a $(-\infty, a]$ quando h tende a zero per valori positivi.

È inoltre possibile dimostrare che se una funzione $F(a)$ soddisfa le proprietà indicate, allora è possibile costruire una v.a. che ha $F(a)$ come funzione distribuzione.