Proposizione 5.6. *Siano ξ_n funzioni misurabili a valori reali. Allora $\sup_n \xi_n$, $\inf_n \xi_n$, $\limsup_n \xi_n$, $\liminf_n \xi_n$ sono misurabili a valori in $\bar{\mathbb{R}}$. Se $\lim_n \xi_n = \xi$ esiste finito per ogni ω in Ω, allora ξ è misurabile.*

Dimostrazione. [Corollary 8.1 (b) e (c) in [4], Sec. 1.2 di [1]] Dimostriamo solo la prima. $\{\omega : \sup_n \xi_n(\omega) \leqslant x\} = \cap_n \{\omega : \xi_n(\omega) \leqslant x\}$. Poiché le ξ_n sono misurabili, allora $\{\omega : \xi_n(\omega) \leqslant x\} \in \mathcal{B}(\bar{\mathbb{R}})$ e quindi intersezione numerabile di misurabili è misurabile. La tesi segue. ✠

Notazioni. Fino ad ora abbiamo usato la notazione $(\Omega, \mathcal{F})$ per indicare un generico spazio misurabile e le lettere $\mathbb{P}$ e μ per indicare una misura di probabilità e una misura definite su $\mathcal{F}$. Tradizionalmente, si indica con $(\Omega, \mathcal{F}, \mathbb{P})$ lo spazio di proabilità su cui viene definita una variabile aleatoria (spesso indicata con X), tuttavia nulla vieta, e in alcune circostanze sarà importante farlo, considerare una misura di probabilità denotata con un altro simbolo, ad esempio P, definita su uno spazio misurabile indicato con altre lettere, ad esempio $(\mathbb{X}, \mathcal{X})$. In questo caso non deve stupire se diremo che $(\mathbb{X}, \mathcal{X}, P)$ è uno spazio di probabilità. Gli assiomi e le proprietà di una misura di probabilità valgono ovviamente indipendentemente dai simboli usati!

Notazioni. Spesso in teoria della misura, invece di usare il simbolo $(\Omega, \mathcal{F})$ per lo spazio misurabile si usa $(E, \mathcal{E})$, così come una funzione misurabile da uno spazio $(E, \mathcal{E})$ in $(G, \mathcal{G})$ viene tipicamente indicata con le lettere f, g, h al posto delle lettere X, Y o ξ.

5.2. **Legge immagine per vettori aleatori e funzioni di ripartizione.** Sia $\mathbf{X}$ un vettore aleatorio definito su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, ossia $\mathbf{X} : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ misurabile.

Notazioni. Nel seguito scriveremo spesso $\mathbb{P}\{\mathbf{X} \in A\}$ per indicare ciò che in modo più corretto dovremmo scrivere come $\mathbb{P}\{\omega \in \Omega : \mathbf{X}(\omega) \in A\}$.

Inoltre, useremo talvolta $\mathbb{P}\{A, B\}$ al posto di $\mathbb{P}\{A \cap B\}$, ad esempio con queste convenzioni si ha

$$\mathbb{P}\{X_1 \in A, X_2 \in B\} = \mathbb{P}\{(\omega : X_1(\omega) \in A) \cap (\omega : X_2(\omega) \in B)\}$$

Proposizione 5.7. *Sia $\mathbf{X} = (X_1, \ldots, X_D)$ un vettore aleatorio, allora $B \mapsto P_{\mathbf{X}}(B) := \mathbb{P}\{\mathbf{X} \in B\}$ è una mdp su $\mathcal{B}(\mathbb{R}^D)$.*

Dimostrazione. [Theorem 8.5 in [4]] Chiaramente $P_{\mathbf{X}}(\mathbb{R}^D) = 1$ e $P_{\mathbf{X}}(B) \geqslant 0$ per ogni B in $\mathcal{B}(\mathbb{R}^D)$. Inoltre dati $B_1, \ldots, B_n, \ldots$ eventi a due a due incompatibili, allora

$$P_{\mathbf{X}}(\cup_j B_j) = \mathbb{P}\{\cup_j (\mathbf{X} \in B_j)\} = \sum_j \mathbb{P}(\mathbf{X} \in B_j) = \sum_j P_{\mathbf{X}}(B_j).$$

✠

La misura di probabilità $P_{\mathbf{X}}$ è detta **legge di X** o **distribuzione di X** o **misura indotta da X**. Si noti che $P_{\mathbf{X}}$ è una misura di probabilità su $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D)$. Non si confondano i due spazi di probabilità

$$(\Omega, \mathcal{F}, \mathbb{P}) \qquad \text{e} \qquad (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D), P_{\mathbf{X}}).$$

Nel seguito della sezione considereremo il caso particolare $D = 1$, ossia ci concentreremo sui numeri aleatori. Come appena visto, dato un numero aleatorio, la sua legge è una misura di probabilità su $\mathcal{B}(\mathbb{R})$. Ci poniamo ora il problema di caratterizzare tutte le possibili probabilità su $\mathcal{B}(\mathbb{R})$.

Definizione 5.2. *Una funzione $F : \mathbb{R} \to [0, 1]$ si dice funzione di ripartizione se*

- *F è monotona non decrescente,*
- *F è continua da destra, ossia $\lim_{x \to a^+} F(x) = F(a)$ per ogni $a \in \mathbb{R}$,*
- *$\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to +\infty} F(x) = 1$.*

Proposizione 5.8.

(1) Sia F una funzione di ripartizione, allora esiste un'unica misura di probabilità P_F su $\mathcal{B}(\mathbb{R})$ tale che $P_F(-\infty, x] = F(x)$.

(2) Viceversa data una misura di probabilità P su $\mathcal{B}(\mathbb{R})$ allora la funzione $x \mapsto F_P(x) := P\{(-\infty, x]\}$ è una funzione di ripartizione.

Dimostrazione. [Theorem 7.2 in [4], Thm 1.14 in [1]] Dimostriamo per esteso la seconda implicazione. La monotonia è dovuta all'implicazione $(-\infty, a] \subset (\infty, a+h]$ se $h \geqslant 0$, la quale, unitamente alla monotonia della probabilità (cfr. 1.2), implica

$$F_P(a) = P((-\infty, a]) \leqslant P((-\infty, a+h]) = F_P(a+h)$$

per ogni a in $\mathbb{R}$ e $h > 0$. In conseguenza della monotonia, gli eventuali punti di discontinuità di F_P costituiscono un insieme numerabile (finito o numerabilmente infinito). Continuità da destra e i limiti sono conseguenza della continuità delle misure di probabilità (cfr. Teorema 2.4) e della monotonia di F_P. Infatti $\emptyset$ si può vedere come limite della successione $(-\infty, -n]$, $n = 1, 2, \dots$, per $n \to +\infty$. Quindi,

$$0 = P(\emptyset) = \lim_{n\to+\infty} P((-\infty, -n]) = \lim_{n\to+\infty} F_P(-n).$$

Per stabilire il limite a $-\infty$ lungo una generica successione, basta ricordare che F_P è monotona non decrescente, condizione che implica l'esistenza di $\lim_{x\to-\infty} F_P(x) = 0$. Analogamente, osservando che $\mathbb{R}$ si può vedere come limite della successione crescente $(-\infty, n]$, $n = 1, 2, \dots$. Pertanto, per continuità,

$$1 = P(\mathbb{R}) = \lim_{n\to+\infty} P((-\infty, n]) = \lim_{n\to+\infty} F_P(n)$$

e la tesi segue, ancora una volta, dalla monotonia di F_X. Per dimostrare la continuità da destra si applica la continuità delle misure di probabilità onde ricavare

$$F_P(a) = P((-\infty, a]) = \lim_{n\to+\infty} P((-\infty, a + \frac{1}{n}]) = \lim_{n\to+\infty} F_P(a + \frac{1}{n});$$

allora, poiché $\lim_{x\to a^+} F_P(x)$ esiste in virtù della monotonia di F_P, ricaviamo

$$F_P(a) = \lim_{n\to+\infty} F_P(a + \frac{1}{n}) = \lim_{x\to a^+} F_P(x).$$

Non dimostriamo il primo punto della proposizione, anche se nella sostanza è una conseguenza di Proposizione 3.2 e di Teorema 2.5. Lo studente interessato può guardare la dimostrazione di Theorem 7.2 in [4]. L'idea è definire una mdp sull'algebra $\mathcal{A}$ definita in Sezione 3.1.1 e poi mostrare che ammette un'estensione su $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R})$. Ricordando ancora una volta che se $A \in \mathcal{A}$ si può scrivere come $A = \cup_{i=1}^n (a_i, b_i]$ con $-\infty \leqslant a_1 < b_1 < a_2 < b_2$ (si veda in (3)) si pone

$$P_F(A) = \sum_{i=1}^n (F(b_i) - F(a_i)).$$

Si verifica facilmente che in questo modo si definisce una misura di probabilità finitamente additiva su $\mathcal{A}$, il punto delicato è, ancora una volta, mostrare che tale misura è anche σ-additiva (cosa che non faremo). L'unicità segue dal teorema di unicità sulle π-classi. ✠

Definizione 5.3. *Se X è un numero aleatorio la funzione di ripartizione di X è*

$$x \mapsto F_X(x) := P_X((-\infty, x]) = \mathbb{P}\{\omega : X(\omega) \leqslant x\}$$

[Sec. 2.1 [1]]

Si noti che la Definizione 5.2 prescinde da una variabile aleatoria, mentre la definizione di F_X in Definizione 5.3 assume l'esistenza di una una v.a. X. In effetti, data una v.a. X e definita F_X come sopra, la F_X è una funzione di ripartizione nel senso della Definizione 5.2, come dimostrato dalla seguente proposizione.

Proposizione 5.9. *Sia F_X la funzione di ripartizione di un numero aleatorio X. Allora:*

- *F_X è monotona non decrescente,*
- *F_X è continua da destra, ossia* $\lim_{x\to a^+} F_X(x) = F_X(a)$,
- *inoltre*

$$\lim_{x\to-\infty} F_X(x) = 0, \qquad \lim_{x\to+\infty} F_X(x) = 1.$$

Dimostrazione. La proposizione è una conseguenza immediata di punto 2 di Proposizione 5.8, una volta che si consideri $P(A) = P_X(A) = \mathbb{P}\{X \in A\}$. ✠

Si ha

$$P_X\{(a,b]\} = F_X(b) - F_X(a)$$

purché si convenga di porre $F_X(-\infty) = 0$. Inoltre,

$$P_X(a,+\infty) = 1 - F_X(a)$$

per $-\infty \leqslant a < +\infty$.

In definitiva, data una funzione di ripartizione, si possono fissare immediatamente le probabilità degli intervalli aperti a sinistra e chiusi a destra. D'altro canto, per la continuità di P_X, poiché $(a,b) = \lim_{n\to+\infty}(a, b-1/n]$ vale per ogni a, b per cui $-\infty \leqslant a < b < +\infty$, si ha $P_X(a,b) = \lim_{n\to+\infty} P_X(a, b-1/n]$, ovvero

$$P_X(a,b) = F_X(b^-) - F_X(a) \tag{14}$$

dove $f(x_0^-)$ indica $\lim_{x\to x_0^-} f(x)$ (purché il limite esista). Infatti, $((a, b-1/n])_{n\geqslant 1}$ costituisce una successione crescente di insiemi verso l'aperto (a,b) e, pertanto, in virtù della continuità di P_X si ha $F_X(b-1/n) - F_X(a) = P_X((a, b-1/n]) \to P_X(a,b)$ per $n \to +\infty$ e, inoltre, essendo F_X monotona non decrescente, $\lim_n F_X(b-1/n) = F_X(b^-)$. Vale anche $P_X\{[a,b]\} = F_X(b) - F_X(a^-)$ e, in particolare, per $a = b = x_0$

$$F_X(x_0) - F_X(x_0^-) = P_X\{x_0\}.$$

Ciò chiarisce che l'eventuale salto di F_X in x_0 coincide con la probabilità concentrata nel *singoletto* $\{x_0\}$. Chiaramente, $P_X\{x_0\} = 0$ se e solo se x_0 è un punto di continuità per F_X.

Riassumendo: sia F_X la funzione di ripartizione di una v.a. X , allora:

(1) $\mathbb{P}\{X > x\} = 1 - F_X(x)$.
(2) $\mathbb{P}\{x_1 < X \leqslant x_2\} = F(x_2) - F(x_1)$.
(3) $\mathbb{P}\{X < x\} = \lim_{y\to x^-} F_X(y) = F_X(x^-)$.
(4) $\mathbb{P}\{X = x_1\} = F(x_1) - \lim_{x\to x_1^-} F(x) =$ valore del salto di F in x_1.

[Corollary 7.1 in [4]]

Esempio 5.1 (Legge uniforme su $(0,1)$, v.a. uniforme su $(0,1)$)**.** *Si verifica immediatamente che*

$$F(x) = \begin{cases} 0 & x \leqslant 0 \\ x & x \in (0,1] \\ 1 & x > 1 \end{cases}$$

è una funzione di ripartizione. Per Proposizione 5.8 esiste un'unica misura su $\mathcal{B}(\mathbb{R})$ tale che $P((a,b]) = F(b) - F(a)$. In particolare per $0 \leqslant a < b \leqslant 1$, P soddisfa $P((a,b]) = P((a,b)) = b - a$. Inoltre per ogni A boreliano tale che $A \cap (0,1) = \emptyset$ si ha $P(A) = 0$. Tale misura è detta uniforme su $(0,1)$ poiché la misura di un subintervallo di $(0,1)$ coincide con la sua lunghezza e coincide con la misura di Lebesgue ristretta a $(0,1)$. Una variabile aleatoria U con funzione di ripartizione definita sopra è detta con legge uniforme su $(0,1)$, in simboli $U \sim \mathcal{U}(0,1)$.

Esempio 5.2 (Variabile aleatoria con legge esponenziale negativa)**.** *Sia $X = -\ln(U)\mathbb{I}\{U > 0\}$ con U con legge uniforme su $(0,1)$. Visto che $\mathbb{P}\{U > 0\} = 1$ tale definizione è equivalente a porre $X = -\ln(U)$, inoltre $\mathbb{P}\{X > 0\} = 1$. Calcoliamo F_X. Se $x \leqslant 0$, si ha subito $F_X(x) = 0$ visto che $\mathbb{P}\{X > 0\} = 1$. Per $x > 0$*

$$F_X(x) = \mathbb{P}\{X \leqslant x\} = \mathbb{P}\{-\ln(U) \leqslant x\} = \mathbb{P}\{U \geqslant e^{-x}\} = 1 - F_U\big((e^{-x})^-\big) = 1 - F_U(e^{-x}).$$

Dal momento che F_U è una funzione di ripartizione continua. Essendo $F_U(x) = x$ per $x \in (0,1)$ segue che

$$F_X(x) = \begin{cases} 0 & x \leqslant 0 \\ 1 - e^{-x} & x > 0. \end{cases}$$

Una variabile aleatoria con tale funzione di ripartizione è detta con legge esponenziale (negativa) di parametro 1, in simboli $X \sim \mathcal{E}(1)$.

Proposizione 5.10. *Sia F una funzione di ripartizione, allora si può definire uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e una variabile aleatoria X definita su $(\Omega, \mathcal{F}, \mathbb{P})$ a valori in $\mathbb{R}$ tale che $F_X = F$.*

Dimostrazione. Vediamo una prima "rappresentazione" che dimostra Proposizione 5.10, sebbene questa rappresentazione possa avere un sapore un po' tautologico. Consideriamo la costruzione seguente: si consideri $\Omega = \mathbb{R}$ e $\mathcal{F} = \mathcal{B}(\mathbb{R})$. Data F, da Proposizione 5.8, sappiamo che esiste una mdp P_F su $\mathcal{B}(\mathbb{R})$ tale che $P_F(-\infty, x] = F(x)$ per ogni $x \in \mathbb{R}$. Definiamo $X : (\Omega, \mathcal{F}, P_F) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ come $X(\omega) = \omega$. Tale funzione è ovviamente misurabile (è l'identità) e quindi è una variabile aleatoria. Inoltre $F_X(x) = P_F\{\omega : X(\omega) \leqslant x\} = P_F(-\infty, x] = F(x)$. ✠

Mentre la P_F di Proposizione 5.8 è unica, non è unica la scelta di $(\Omega, \mathcal{F}, \mathbb{P})$ e X in Proposizione 5.10. In aggiunta alla costruzione data durante la dimostrazione vediamo ora un'altra "rappresentazione" di X in Proposizione 5.10.

Per fare ciò introduciamo una nuova definizione. Data una funzione di ripartizione F il suo quantile (o inversa generalizzata) F^- si definisce come

$$F^-(u) = \inf\{x \in \mathbb{R} : \ F(x) \geqslant u\}.$$

Se F è strettamente monotona allora F^- coincide con l'inversa di F, ossia $F^- = F^{-1}$. Si noti che la definizione non presuppone che F sia una funzione di ripartizione di una variabile aleatoria X.

NON IN PROGRAMMA:

Proposizione 5.11. *Sia F una funzione di ripartizione. Se $U \sim \mathcal{U}(0,1)$ è definita su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, allora $X := F^-(U)$ ha funzione di ripartizione F, ossia*

$$F_X(x) = P\{X \leqslant x\} = \mathbb{P}\{F^-(U) \leqslant x\} = F(x).$$

Dimostrazione. Si verifica che $\{u \in (0,1) : F^-(u) \leqslant x\} = \{u \in (0,1) : u \leqslant F(x)\}$ (farsi un disegno!). La tesi segue. ✠

La precedente proposizione mostra che, disponendo di un una variabile aleatoria uniforme su uno spazio $(\Omega, \mathcal{F}, \mathbb{P})$, si può costruire sullo stesso spazio una variabile aleatoria X con funzione di ripartizione assegnata qualunque.

Si noti che scegliendo $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}((0,1))$ e $\mathbb{P} = P$ con P la misura uniforme su $(0,1)$ (Esempio 5.1), possiamo definire la variabile aleatoria $U(\omega) = \omega$ e verificare che tale variabile aleatoria ha legge uniforme su $(0,1)$. Infatti, $\mathbb{P}\{\omega : U(\omega) \leqslant x\} = x$ per ogni $x \in (0,1)$. Combinando questa osservazione con la precedente proposizione otteniamo un'altra dimostrazione di Proposizione 5.10.

Il concetto di funzione di ripartizione di un numero aleatorio X può essere esteso anche ad un vettore aleatorio $\mathbf{X}$. In questo caso la definizione diventa:

Definizione 5.4. *Dati un vettore aleatorio $\mathbf{X} = (X_1, \dots, X_D)$ e la sua distribuzione $P_{\mathbf{X}}$, si dice funzione di ripartizione di $\mathbf{X}$ la funzione $F_{\mathbf{X}} : \mathbb{R}^D \to [0,1]$ definita da*

$$F_{\mathbf{X}}(x) := \mathbb{P}\{X_1 \leqslant x_1, \dots X_D \leqslant x_D\} = P_{\mathbf{X}}\{(-\infty, x]\}.$$

Il caso di funzioni di ripartizione $D > 1$ è più complicato da trattare e ci limitiamo qui ad osservare che il Teorema 3.3 si può ri-enunciare come segue:

Teorema 5.12. *Se $F_{\mathbf{X}}$ e $F_{\mathbf{Y}}$ sono uguali allora $P_{\mathbf{X}} = P_{\mathbf{Y}}$.*

6. Variabili aleatorie discrete

6.1. **Variabili aleatorie discrete, densità, marginali.** Sia $(\mathbb{X}, \mathcal{X}) = (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ o $(\mathbb{X}, \mathcal{X}) = (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$.

Definizione 6.1. *Una variabile aleatoria $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{X}, \mathcal{X})$ è detta discreta se esiste un insieme $C \in \mathcal{X}$ numerabile tale che $\mathbb{P}\{X \in C\} = 1$.*

Se X è discreta in realtà non è molto rilevante chi sia $(\mathbb{X}, \mathcal{X})$, infatti ciò che conta è che gli insiemi immagine tramite X sono (a meno di insiemi di probabilità nulla) insiemi al più numerabili. Se P_X è la misura immagine di X

$$P_X(A) = P_X(A \cap C) + P_X(A \cap C^c) = P_X(A \cap C) \qquad \text{con } A' = A \cap C \in \mathcal{P}(C).$$

Quindi la probabilità immagine P_X può essere pensata come una misura di probabilità non su $(\mathbb{X}, \mathcal{X})$ ma semplicemente su $(C, \mathcal{P}(C))$.[3] In effetti al posto di P_X si può considerare senza perdita di generalità la sua traccia su $\mathcal{X}_C$, infatti

$$P_X(A) = P_X(A \cap C) = \frac{P_X(A \cap C)}{P_X(C)}.$$

In breve una v.a. discreta X è completamente caratterizzata dalla *densità discreta*

$$p_X(x) = P\{X = x\} \quad x \in C.$$

Infatti per ogni A in $\mathcal{P}(C)$ si ha che A è al più numerabile e

$$P\{X \in A\} = \sum_{x \in A} p_X(x)$$

[Theorem 4.1. in [4]]

Un vettore aleatorio $\mathbf{X} = (X_1, \dots, X_D)$ a valori in $\mathbb{R}^D$ discreto sarà caratterizzato da una densità discreta $p_{\mathbf{X}}$ definita a priori su un insieme discreto $C \subset \mathbb{R}^D$. Si può sempre considerare come C un insieme prodotto del tipo $C_1 \times \cdots \times C_D$ (in caso basta aggiungere un numero al più numerabile di punti). Ad esempio se $D = 2$ e $C = \{(0,1), (1,0), (1,1)\}$ potremo considerare come nuovo C l'insieme $\{0,1\} \times \{0,1\}$. Risulta ovvio che se $\mathbf{X}$ è discreto sono discrete anche le sue componenti X_i, $i = 1, \dots, D$ e i possibili sottovettori $(X_{i_1}, \dots, X_{i_k})$. Ogni X_i assume con probabilità 1 valori in C_i (definito sopra). Come conseguenza sono ben definite le densità delle componenti X_i, che si dicono *densità marginali* di $\mathbf{X}$.

Si verifica immediatamente che per ogni $x_i \in C_i$

$$p_{X_i}(x_i) = \sum_{y \in C : y_i = x_i} p_{\mathbf{X}}(y_1, \dots, y_D) = \sum_{y_1 \in C_1, \dots y_D \in C_D} p_{\mathbf{X}}(y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_D).$$

In modo analogo si possono considerare sottovettori, per cui, ad esempio

$$p_{(X_1, X_3)}(x_1, x_3) = \sum_{y_2 \in C_2} p_{\mathbf{X}}(x_1, y_2, x_3).$$

6.2. **Esempi.**

6.2.1. *Distribuzione binomiale.* Siano n un intero positivo e p un elemento fissato dell'intervallo $[0,1]$, X un numero aleatorio che prende valori in $C = \{0, 1, 2, \dots n\}$. La distribuzione di X si dice *binomiale* con parametro (n, p) [in simboli $Bi(n, p)$] se

$$\mathbb{P}\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{per } k = 0, 1, 2, \dots, n.$$

Si verifica facilmente che $\mathbb{P}\{X = k\}$ coincide con la probabilità che si verifichino esattamente k eventi su n, nell'ipotesi che gli eventi siano indipendenti e abbiano tutti probabilità p. Si veda (13).

6.2.2. *Distribuzione ipergeometrica.* Un numero aleatorio X a valori in $C = \{0, 1, 2, \dots n\}$ si dice con distribuzione *ipergeometrica* se

$$\mathbb{P}\{X = k\} = \begin{cases} \binom{n}{k} \frac{N\theta(N\theta-1)\dots(N\theta-k+1)(N-N\theta)(N-N\theta-1)\dots(N-N\theta-n+k+1)}{N(N-1)\dots(N-n+1)} \\ \quad \text{se } n \leqslant N,\ N\theta + n - N \leqslant k \leqslant N\theta \\ 0 \qquad \text{altrove.} \end{cases}$$

Si suppone $n \leqslant N$ e $\theta \in (0,1)$. In seguito indicheremo tale distribuzione con $\mathcal{H}(k; \theta, N, n)$. Si vede facilmente, che è la probabilità di avere k palline bianche in n estrazioni senza restituzione da un'urna che contiene N palline di cui $N\theta = R$ bianche, quando tutte le n–uple estraibili siano ritenute ugualmente probabili.

[3] Se consideriamo un $(\mathbb{X}, \mathcal{X})$ generico dobbiamo supporre che ogni $x \in C$ sia tale per cui $\{x\} \in \mathcal{X}$, cosa chiaramente soddisfatta se $(\mathbb{X}, \mathcal{X}) = (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$.