Dagli eventi alle osservabili fisiche

- Abbiamo spiegato gli assiomi di Kolmogorov, e abbiamo introdotto alcune definizioni e risultati di validita’ generale.
- Abbiamo parlato pero’ di probabilita’ di singoli eventi (definiti come sottoinsiemi di un insieme di riferimento S). Questi insiemi possono rappresentare eventi elementari o l’unione di vari possibili eventi elementari (ed es. considerando un mazzo di carte da gioco, l’evento "estrazione di una carta di cuori" e’ l’insieme di tutte le possibili carte di cuori – ciascuna delle quali rappresenta un evento elementare nello spazio S, l’insieme di tutte le carte).
- Molto spesso gli "eventi" di interesse in fisica consistono nell'aver compiuto una *misura* di una particolare *osservabile* X (usiamo in genere lettere maiuscole) e aver trovato un particolare valore x (appartenente a un insieme O_X finito o ~continuo). Una osservabile definisce una famiglia di sottoinsiemi di S, ognuno corrispondente a un certo specifico valore $x \in$ X.
- Si dice che l’osservabile X *e' una* ***osservabile aleatoria/ casuale/random***, se esiste ben definita (misurabile sperimentalmente) la Probabilita’ di ***<u>tutti</u>*** i suoi possibili *valori* - se discreta - o *insiemi di valori* (sottoinsiemi di O_X), (se continua).
- Allora si possono eseguire inferenze di tipo statistico a partire da osservazioni di X. [NB perche' gli *insiemi* di valori ?]

= se *ogni* Evento definibile da un sottoinsieme *[ehm, misurabile]* dei valori di X ha una probabilita' (misurata) ben definita:

$$\forall A \subset O_X \quad \exists P(x \in A) = \lim_{N \to \infty} \frac{n(x \in A)}{N}$$

- Non e' scontato che esista per il solo fatto che esiste per *alcuni* valori [es. orologio su faccia dado]

Variabili aleatorie e statistiche in matematica

- Il modello matematico di una osservabile aleatoria e' la cosiddetta ***variabile aleatoria(o casuale):*** definisco una funzione X da S in O_X tale che:
  $$\forall A \subset O_X \quad X^{-1}(A) \subset \Sigma$$
  $$\Longrightarrow \ \exists\, P(X^{-1}(A))$$
- X e' la *funzione* che associa a un particolare evento ('stato osservabile' del mondo) il valore di una certa particolare osservabile (es. il valore indicato da un certo strumento).
- NB la condizione di cui sopra e' banale se P e' definita sulle parti di S.
- Si noti che **se** una variabile aleatoria ammette una distribuzione di probabilita', **allora** lo stesso vale per *qualunque* sua funzione *[misurabile]*. Basta prendere l'immagine inversa della funzione composta $f \circ X$ e procedere come sopra per definire la $P_F(f)$ corrispondente a qualunque sottoinsieme *[misurabile]* nello spazio dei valori della funzione f
  $$\forall B \subset f(O_X) : P_F(f \in B) \equiv P(x \in f^{-1}(B))$$
- Questo vale anche per qualunque $f(x, y\ldots)$ di piu' variabili casuali - detta **statistica.** e' banalmente una variabile aleatoria essa stessa. Questa definizione matematica si trasferisce senza variazioni alle osservabili fisiche: una statistica calcolata sulle osservabili e' essa stessa una quantita' osservabile.

Distribuzioni di probabilita' di una osservabile

- La Probabilita' come l'abbiamo usata finora e' una *funzione di insieme*. Risulta spesso utile fare uso di concetti derivati, basati su funzioni 'puntuali' (il cui argomento non e' un insieme)
- Chiamiamo *distribuzione di probabilita'* di una osservabile discreta X, quella *funzione* (detta anche "***funzione di massa***", ***pmf***) definita da:

$$pmf(x) \equiv P(E : X = x)$$

- A volte in fisica si scrive semplicemente (un po' impropriamente) P(x) - ma quello che si intende e' quanto sopra
- La additivita' (finita) mi permette di derivare le P di qualunque sottoinsieme di S, come richiesto da Kolmogorov, a partire dalla funzione *pmf*(x).

$$P(E \subset O_X) = \sum_{x \in E} pmf(x)$$

- Lo spazio O_X dei risultati (outcomes) di una osservabile (o statistica) puo' essere discreto o continuo: {0,1}, {testa,croce}, N, R, R^N... la "somma" va definita opportunamente a seconda dell'insieme: e' una somma ordinaria soltanto se O_X e' discreto.

- Gli esempi di dadi e carte costituiscono un buon esempio di *pmf*, in particolare di Distribuzioni Uniformi

Densita' di Probabilita'

- In molti casi e' conveniente considerare osservabili sperimentali a valori in uno spazio continuo ($\mathscr{R}$, $\mathscr{R}^N$…) (anche se in realta' e' difficile concepire un esperimento in cui il risultato non sia in uno spazio di valori finito)
- In tali casi, la distribuzione di probabilita' (funzione di massa) $P_X(x)$ e' sempre 0 per ogni singolo *x*, e non e' utile a definire la situazione – per caratterizzare la distribuzione di probabilita' bisognerebbe specificarne il valore su TUTTI i sottoinsiemi di O_X a misura non nulla.
- Un concetto comodo e' allora quello di *densita' di probabilita'* ***(pdf)*** [NB terminologia diversa tra fisici e matematici]. Si definisce *densita' di probabilita'* una funzione reale positiva $p_X(x)$ tale che

$$P(x : x \in A) = \int_A p_X(x)dx$$

per ogni insieme A su cui la probabilita' e definita, e l'integrale e' nel senso di Lebesgue. Si generalizza in modo ovvio a *n* osservabili.

(In pratica *ogni* osservabile fisica reale ha solo un insieme discreto di valori possibili, per cui le sottigliezze riguardanti la definizione di misura e l'esistenza di insiemi non misurabili non hanno vero senso fisico. La rappresentazione su $\mathscr{R}$ e' soltanto una comoda astrazione)

- Si noti che se si accetta che $p_X(x)$ possa essere piu' in generale una "distribuzione" (es. delta di Dirac), si puo' *sempre* utilizzare una densita', anche nel caso di distribuzioni discrete o miste.
- Si puo' definire la $\delta(x_0)$ come quella densita' che fornisce una distribuzione di probabilita' che e' =1 *sse* l'insieme in questione contiene il punto x_0; altrimenti P=0.

Distribuzioni di osservabili multiple

- Supponiamo di avere DUE osservabili X, Y. Posso considerare $P(X=x \cap Y=x)$ come *funzione* $P_{XY}(x,y)$ al variare di *x* e *y* nei rispettivi insiemi di esistenza. Questa si chiama distribuzione(/ funzione) di probabilita' *congiunta* di X e Y. Spesso si omette il pedice scrivendo semplicemente P(x,y), anche se questa e' ora una funzione definita su O1*O2, non una Probabilita' (funzione di insieme su S). [NB la virgola prende il posto del ∩]
- Generalizza a qualunque numero di osservabili:

$$P(x_1, x_2, x_3, ..., x_N)$$

- Abbiamo gia' definito l'indipendenza di due (o piu') eventi casuali: $P(A \cap B) = P(A)P(B)$. Due DISTRIBUZIONI di probabilita'definite sullo stesso S (finito) si dicono *indipendenti* se per **<u>ogni</u>** A, B si ha che

$$P((x \in A) \cap (y \in B)) = P(x \in A) \cdot P(y \in B)$$

da cio' segue in particolare che $P_{XY}(x,y)=P_X(x)P_Y(y)$ (vedremo in seguito perche' puo' non essere sufficiente)

- Si ha inoltre la Distribuzione di probabilita' Condizionata:

$$P_X(x \mid y)P_Y(y) = P_{XY}(x, y)$$

che si estende banalmente a *n* variabili:

$$P_{X^M}(x_1, \ldots x_m \mid y_1, \ldots y_n)P_{Y_1}(y_1 \mid y_2 \ldots y_n) = P_{X^M, Y_1}(x_1, \ldots x_m, y_1 \mid y_2 \ldots y_n)$$

- Considerata una partizione $\{Y_i\}$ di O_Y si ha

$$P(X = x) = \sum_i P((X = x) \mid (Y \in Y_i)) \cdot P(Y \in Y_i) = \sum_i P((X = x) \cap (Y \in Y_i))$$

- Se Y ha valori in uno spazio discreto, si puo' prendere in particolare $Y_i = \{y_i\}$, e allora si ha:

$$P_X(x) = \sum_i P_{XY}(x, y_i) = \sum_i P_X(x \mid y_i)P_Y(y_i)$$

- Questa operazione si chiama *marginalizzazione*, e la $P_X(x)$ e' chiamata *distribuzione marginale* di $P_{XY}(x,y)$ rispetto a *y*.
  Si noti che la seconda uguaglianza corrisponde alla legge della probabilita' totale.

Densita', Probabilita', e Cumulante

- Si noti che la funzione densita' (*pdf*) **non** soddisfa gli assiomi di Kolmogorov per una probabilita'.
  - Es. risulta $p(x) >= 0$ ma non necessariamente $p(x) < 1$.
  - Inoltre dimensioni fisiche di una densita' di x sono $[x]^{-N}$
- Si ha sempre $\int p(x)dx = P(X \in O_X) = 1$ (indifferentemente somma discreta o continua)
- Si puo' dimostrare che l'espressione della probabilita' condizionata e la formula di Bayes si estendono direttamente alle funzioni densita':

$$p(x|y) = p(x,y)/p(y) = p(y|x)*p(x)/p(y)$$

  e per osservabili indipendenti: $p(x,y) = p(x)p(y)$
- L'operazione di *marginalizzazione* si estende come:

$$p_X(x) = \int_{y \in O_Y} p_{XY}(x,y)dy$$

- In 1 dimensione (es. R) si puo' definire utilmente la funzione *CUMULANTE* (CDF):

$$F_X(x) \equiv P(X \in (-\infty, x]) = \int_{-\infty}^{x} p_X(x')dx'$$

- E' facile vedere che $F(x)$ e' monotona crescente, che $0 \leq F(x) \leq 1$ e che $p(x) = dF(x)/dx$
- Il concetto di cumulante ha molta utilita' anche pratica, poiche' non di rado risulta piu' facile da calcolare rispetto alla densita'.