Probabilità e Inferenza

Corso di Laurea Triennale in **Scienze del Turismo ad Indirizzo Manageriale**
Corso di **Analisi dei dati**

Prof.ssa **Maria Spano**

Anno accademico 2023-2024

Probabilità - Recap Teorico

- La probabilità statistica è un campo matematico e statistico fondamentale che studia le leggi che governano l'incertezza nei fenomeni casuali e nei processi stocastici.

- Questa branca della statistica fornisce gli strumenti concettuali per quantificare e analizzare le incertezze nelle osservazioni, nei dati sperimentali e nei risultati dei processi casuali.

- Esamineremo i principi fondamentali della probabilità statistica, dalla definizione di probabilità alla sua applicazione in contesti reali e nella statistica inferenziale.

- Attraverso l'uso di concetti matematici rigorosi, esploreremo come la probabilità possa essere utilizzata per la modellizzazione di eventi casuali e per prendere decisioni informate in una vasta gamma di settori, dalla scienza ai mercati finanziari.

Definizione di Probabilità

- La probabilità è un concetto fondamentale nella teoria della probabilità statistica, che fornisce una base matematica per misurare e quantificare l'incertezza nei fenomeni casuali.
- In termini formali, la probabilità è una funzione matematica definita su un insieme di eventi, chiamato spazio campionario, che assegna un valore tra 0 e 1 a ciascun evento.
- La probabilità di un evento rappresenta la misura della sua verosimiglianza o la sua possibilità di verificarsi. Un valore di probabilità più alto indica una maggiore probabilità di successo dell'evento.
- In parole semplici, la probabilità è una misura della "percentuale" di fiducia che attribuiamo a un evento in un contesto dato.
- La somma delle probabilità di tutti gli eventi possibili in uno spazio campionario è sempre uguale a 1, il che significa che almeno uno degli eventi deve verificarsi.

Definizione di Probabilità

Scuola Classica:

La probabilità è il rapporto tra il numero dei casi favorevoli al verificarsi di un risultato e il numero totale dei possibili risultati, ammesso che questi siano egualmente possibili.

Questa definizione presuppone:
- la conoscenza di tutti i possibili casi che possano verificarsi, e che quindi il numero di questi casi sia finito
- la equiprobabilità di tutti i casi.

Alcune osservazioni critiche:
- la definizione di probabilità si fonda sul concetto di equiprobabilità (presupponendo quindi la conoscenza del concetto di probabilità che invece si intende definire)
- essa trova applicazione solo in esperimenti noti nelle loro caratteristiche (casi noti e finiti, ecc.) e i cui eventi siano equiprobabili

Definizione di Probabilità

Scuola Frequentista:

la probabilità è il limite della frequenza relativa di un evento ripetibile quando cresce, oltre ogni limite, il numero delle prove.

In altre parole la probabilità è pari alla frequenza relativa dei successi quando si ripete la prova all'infinito.

Questa definizione presuppone:

- la ripetibilità della prova all'infinito nelle stesse condizioni

Alcune osservazioni critiche:

- la prova non è sempre ripetibile per ragioni tecniche, economiche, ecc.

Scuola Soggettivista:

la probabilità rappresenta il grado di fiducia che un individuo coerente attribuisce al presentarsi di un evento, ovvero, per quantificare, come la somma p che è disposto a scommettere quando, verificandosi l'evento, vince 1.

Per individuo coerente (o razionale) si intende un individuo che, nell’ottica di poter vincere 1, è disposto a scommettere una somma p non superiore a 1.
Inoltre egli scommetterà una cifra maggiore quanto più alta sarà la fiducia che l’evento si verifichi.

Alcune osservazioni critiche:
- Tale definizione comporta l’individuazione di una misura di probabilità che, data una stessa prova, muta al variare dell’individuo considerato.

Definizione di Probabilità

Esempio:

- Nel lancio di un dado equo a sei facce, ogni faccia ha una probabilità di 1/6 di uscire poiché ci sono sei risultati possibili e sono tutti ugualmente probabili.

- Nel lancio di una moneta posso ottenere o testa o croce. I casi favorevoli sono 1 mentre i casi possibili sono 2 (le due facce della moneta), quindi:

$$p = \frac{1}{2} = 0{,}5 = 50\%$$

Esempio:

- Un sacchetto contiene 20 palline, 10 bianche , 6 rosse e 4 verdi. Calcoliamo la probabilità che, estraendo a caso una pallina, essa sia verde.

Le palline verdi sono 4 quindi ho 4 casi favorevoli mentre i casi possibili sono 20 (numero totale di palline), quindi:

$$p = \frac{4}{20} = \frac{1}{5} = 0{,}2 = 20\%$$

Concetti primitivi della Probabilità

Si introducono i Concetti Primitivi e la loro reciproca relazione:

‘’la Prova genera l'Evento con una certa Probabilità’’

Prova: è un esperimento soggetto a incertezza e può suddividersi in sottoprove.

Evento: è uno dei possibili risultati della prova e costituisce un insieme di descrizioni circa i possibili risultati dell'esperimento.

L'insieme di tutti i risultati possibili di una prova prende il nome **di Spazio Campionario.**

Spazio Campionario

- Nella teoria della probabilità, il concetto di "spazio campionario" svolge un ruolo cruciale.
- Lo spazio campionario rappresenta l'insieme di tutti i possibili risultati di un esperimento casuale o di un processo stocastico.
- È spesso indicato con il simbolo Ω (omega) ed è un insieme completo e esaustivo, cioè contiene tutti gli esiti possibili.
- La conoscenza dello spazio campionario è essenziale per definire e calcolare le probabilità degli eventi all'interno di quell'esperimento.
- Ad esempio, nel lancio di una moneta equa, lo spazio campionario consiste in due eventi: testa (T) e croce (C).

Esempio:

Lancio di una moneta equa:

$$Spazio\ Campionario\ (\Omega)\ =\ \{T, C\}$$

Lancio di un dado equo a sei facce:

$$Spazio\ Campionario\ (\Omega)\ =\ \{1, 2, 3, 4, 5, 6\}$$

Esempio:
Se stiamo studiando il comportamento di una particella subatomica, lo spazio campionario potrebbe essere un insieme infinito di punti nello spazio tridimensionale.

Tipologia e algebra degli eventi

- Gli "eventi" sono concetti chiave nella teoria della probabilità che rappresentano insiemi di risultati nello spazio campionario.

- Un evento è semplicemente una collezione di risultati che ci interessa. Ad esempio, nel lancio di un dado, l'evento "ottenere un numero pari" include i risultati {2, 4, 6}.

- L'evento complementare di un evento A, indicato come A', è costituito da tutti gli esiti nello spazio campionario che non fanno parte di A. In altre parole, rappresenta l'opposto di A.

- Ad esempio, se A è l'evento "ottenere un numero pari," allora A' è l'evento "ottenere un numero dispari.’’

- Gli eventi A e A' sono mutuamente esclusivi, il che significa che non possono verificarsi contemporaneamente.

- La somma delle probabilità di un evento e del suo complementare è sempre uguale a 1.

Esempio:

Lancio di un dado equo:

Evento A: Ottenere un numero pari.

$$A = \{2, 4, 6\}$$

Evento A': Ottenere un numero dispari. (evento complementare)

$$A' = \{1, 3, 5\}$$

$$P(A) + P(A') = 1$$

(poiché la somma delle probabilità di eventi complementari è 1).

Tipologia e algebra degli eventi

Tipologia di eventi:

- **Eventi compatibili**: si dicono compatibili due eventi che possono verificarsi contemporaneamente.
- **Eventi equiprobabili**: sono eventi che hanno pari probabilità di verificarsi.

Algebra degli eventi:
- **Unione (o somma logica) di due eventi A e B (simbolo U)**: è l’evento “si verifica A oppure B oppure entrambi”
- **Intersezione (o prodotto logico) di due eventi A e B (simbolo U capovolta)**: è l’evento “si verificano A e B contemporaneamente”
- **Negazione di un evento A**: è quell’evento che si verifica allorché non si verifica A

Teoria assiomatica della probabilità

Un **assioma** è un'affermazione che non si dimostra in quanto principio di base universalmente accettato.

La teoria assiomatica si fonda su tre momenti fondamentali:

- l'individuazione dei concetti primitivi di prova, evento e spazio così come definiti in precedenza;
- l'enunciazione degli assiomi (o postulati) della probabilità;
- la dimostrazione dei teoremi mediante i postulati e con l'ausilio della logica e della matematica.

Postulati della probabilità

➢ **Positività**:
La Probabilità di un evento A è un numero unico maggiore o uguale di 0: $P(A) \geq 0$

➢ **Certezza**:
La Probabilità dell'evento certo e quindi dello Spazio Campionario Ω è sempre 1: $P(I) = P(\Omega) = 1$
(dove con "I" si indica un evento certo)

➢ **Unione**:
Siano A e B due eventi incompatibili, allora la probabilità della loro unione è la somma delle singole probabilità di A e B.

NB.
Dal primo e secondo assioma si deduce che la probabilità di un evento A è sempre compresa tra 0 e 1.

- **Probabilità dell'evento impossibile**

La probabilità dell'evento impossibile è pari a 0.

- **Probabilità dell'evento negazione**

Dato un evento A, la probabilità dell'evento negazione di A è pari al complemento a 1 della probabilità di A.

- **Probabilità totali**

Dati due eventi A e B, la probabilità dell'unione di A e B è pari alla somma delle singole probabilità dei due eventi meno la probabilità dell'intersezione.

Quest'ultimo teorema generalizza il concetto dell'unione per eventi compatibili.

Probabilità Condizionata

Si definisce probabilità condizionata la probabilità di un evento B condizionata al verificarsi di un evento A.

$P(B|A)$ "Probabilità dell'evento B dato che si è verificato l'evento A"

Teorema della probabilità condizionata
La probabilità condizionata dell'evento B dato A è pari al rapporto tra la probabilità dell'intersezione di A e B e la probabilità dell'evento A.

Due eventi A e B si dicono **stocasticamente indipendenti** quando il verificarsi dell'uno non modifica la probabilità del verificarsi dell'altro evento.

Formalmente si definiscono indipendenti due eventi per cui:

$$P(A\ intersezione\ B) = P(A) * P(B)$$

oppure

$$P(B|A) = P(B)$$

Variabile Casuale

Una Variabile Casuale (v.c.) è una regola (funzione reale) che associa ad E (evento elementare di Ω) uno ed un solo numero reale.
Spesso una v.c. non è altro che la traduzione numerica immediata degli eventi elementari, ma può essere anche una funzione più complessa definita sugli eventi di Ω.

Si pensi ad esempio alla prova “lancio di un dado”.
Per convezione, attribuiamo un numero ad ognuna delle facce del dado.
Oppure nelle carte da gioco, dove associamo un numero ad ognuna delle carte del mazzo (l’asso vale uno, il re vale 10, ecc.)

Funzione di probabilità e Supporto

Ad ogni valore assunto dalla variabile casuale è possibile associare una probabilità utilizzando la misura di probabilità definita sui sottoinsiemi dello spazio campionario Ω.

"Si verifica l'evento E con probabilità P(E)’’
da cui
"La v.c. X assume il valore x con probabilità P(x)’’

La misura di probabilità definita sui valori di X prende il nome di funzione di probabilità della v.c. X.

Si indica invece come supporto della v.c. X l’insieme dei valori che X può assumere con una certa probabilità positiva in una prova specifica.

Tipologia di variabili casuali

Una variabile casuale può essere:

- ***discreta***

è definita su uno spazio campionario Ω discreto, cioè composto da un numero di eventi finito e numerabile. La definizione della funzione di probabilità della v.c. è agevole in quanto, essendo Ω discreto, è possibile enumerare tutti i possibili eventi, associare ad ognuno di essi un numero e quindi la probabilità che la v.c. assuma tale numero.
Alcune prove sono naturalmente discrete (numero di figli, lancio della moneta, ecc.) altre invece possono essere rese discrete da un particolare partizionamento di Ω (reddito superiore a 50.000€, si o no?, durata del percorso Napoli-Roma, meno di 2 ore, si o no?, ecc.)

- ***continua***

È invece definita su uno spazio campionario Ω continuo, dove il numero degli eventi non è finito.
In questo caso la definizione di una funzione di probabilità è possibile solo se questa è misurabile.

Inferenza Statistica

L'inferenza statistica è uno dei pilastri fondamentali della teoria della probabilità statistica.
Questo processo analitico ci consente di fare affermazioni informative e prendere decisioni basate sui dati, utilizzando informazioni tratte da campioni o dati osservati.

Statistica Descrittiva

La statistica descrittiva coinvolge la sintesi e la presentazione dei dati osservati in modo comprensibile. Utilizza metodi come la media, la mediana, la deviazione standard e gli istogrammi per riassumere le caratteristiche dei dati.

Scopo principale: fornire una visione panoramica dei dati, senza fare affermazioni generalizzate sulla popolazione.

Statistica Inferenziale

La statistica inferenziale va oltre la descrizione dei dati e cerca di estrapolare informazioni significative sulla popolazione da cui provengono i dati.
Utilizza tecniche come l'estrazione di campioni, i test di ipotesi, l'intervallo di confidenza e la regressione per trarre conclusioni basate su campioni limitati.

Scopo principale: fare affermazioni o previsioni sulla popolazione sulla base dei dati campionari.

La statistica inferenziale si basa ampiamente sulla probabilità statistica per effettuare previsioni e prendere decisioni informate.
Un esempio classico è il test di ipotesi, in cui vengono formulate ipotesi sulla popolazione e viene utilizzata la probabilità statistica per valutare la loro validità.

L'inferenza statistica è ampiamente utilizzata in campi come la scienza, l'economia, la medicina e l'ingegneria per trarre conclusioni basate sui dati raccolti.