Fondamenti di Probabilità

Teoria della Probabilità: Esperimenti Aleatori, Operazioni sugli Eventi, Definizioni di probabilità, Assiomi e proprietà derivate

Corso di Data Analyst e AI per Programmatori • Lezione 1 • Unità IA.2

Corrado Centrone

Docente del corso di Data Analyst e AI per Programmatori.

Corrado Centrone porta una vasta esperienza nel campo dell'analisi dei dati e dell'intelligenza artificiale, guidandovi attraverso i concetti fondamentali e le applicazioni pratiche che trasformeranno la vostra comprensione della probabilità e della statistica. Lasciatevi guidare nella a esplorare il mondo dei dati con una guida esperta.

Perché Probabilità e Statistica Insieme?

Il Contesto: L'Incertezza

La Probabilità e la Statistica si occupano di situazioni caratterizzate da **incertezza**, ovvero uno stato di conoscenza parziale sul fenomeno che stiamo esaminando. Nel mondo reale, molti eventi non possono essere previsti con certezza assoluta, e dobbiamo imparare a gestire questa incertezza in modo razionale e sistematico.

Esempi di fenomeni aleatori (il cui esito non possiamo prevedere con certezza):

- Il risultato del lancio di un dado
- Le previsioni meteorologiche per domani
- Il successo di un nuovo prodotto sul mercato
- Il tempo di attesa per un servizio clienti
- La durata di vita di un componente elettronico

Come Lavorano Insieme

La probabilità e la statistica non sono discipline separate, ma dovrebbero lavorare insieme secondo una **visione fusionista**:

Probabilità: Lo scopo della probabilità è misurare la nostra incertezza sul problema. Questa misura è spesso interpretata come il nostro grado di fiducia nel verificarsi di un evento. Ci fornisce gli strumenti matematici per quantificare l'incertezza.

Statistica: Interviene per cercare di ridurre parzialmente l'incertezza. Raccogliamo dati statistici (informazioni o campioni) e, sulla base di questi, aggiorniamo la nostra valutazione iniziale di probabilità. La statistica trasforma i dati in conoscenza.

Il Processo di Aggiornamento: Esempio Diagnostico

La probabilità condizionata funge da **trait d'union** tra probabilità e statistica, permettendoci di aggiornare le nostre credenze sulla base di nuove informazioni. Vediamo come questo processo si applica in un contesto medico:

01

Conoscenza Parziale/Incertezza

Il medico osserva i sintomi iniziali del paziente (febbre, mal di testa, tosse). Non c'è una relazione logica univoca tra sintomi e malattia: gli stessi sintomi possono indicare diverse patologie.

02

Valutazione Iniziale (Probabilità)

Sulla base dell'esperienza e dei sintomi osservati, il medico valuta le probabilità iniziali delle varie ipotesi di malattia. Ad esempio: influenza 60%, polmonite 25%, altra patologia 15%.

03

Raccolta Dati (Statistica)

Vengono raccolti nuovi dati attraverso esami diagnostici: analisi del sangue, radiografia del torace, test specifici. Questi dati forniscono informazioni aggiuntive oggettive.

04

Aggiornamento (Probabilità Condizionata)

Sulla base dei risultati degli esami, si affina la valutazione iniziale. Se la radiografia mostra infiltrati polmonari, la probabilità di polmonite aumenta drammaticamente. Lo strumento matematico che consente questo aggiornamento è la probabilità condizionata.

💡 **Concetto Chiave:** Questo processo di aggiornamento bayesiano è fondamentale nell'AI e nel machine learning, dove i modelli apprendono continuamente dai dati per migliorare le loro previsioni.

Perché studiare la Probabilità nel Data Analysis e nell'AI

"La probabilità è il linguaggio dell'incertezza."

Ogni dato racconta una storia — ma mai una verità assoluta. La probabilità ci insegna quanto possiamo fidarci di quella storia, trasformando l'incertezza in decisioni informate e basate su evidenze.

Elena, la Data Analyst

Scenario: Elena deve scegliere tra due versioni di una pagina e-commerce (A vs B) per massimizzare le conversioni.

Domanda: "La versione B vende davvero di più o l'aumento è dovuto al caso?"

Soluzione: Elena applica un test d'ipotesi statistico, calcolando la p-value per determinare la significatività dei risultati.

Risultato: Grazie alla probabilità, Elena passa da una decisione intuitiva a una basata su evidenza statistica, ottimizzando le performance del sito.

Sara, la Data Scientist

Scenario: Sara deve stimare il rischio di insolvenza per i clienti di una banca al fine di concedere prestiti in modo responsabile.

Domanda: "A quali clienti è più sicuro concedere un prestito e con quali condizioni?"

Soluzione: Costruisce un modello di scoring probabilistico che assegna a ogni cliente una probabilità di default.

Risultato: Sara trasforma l'incertezza in valore economico per la banca, minimizzando i rischi e massimizzando le opportunità di business.

Che cos'è la Probabilità?

Framework Matematico per Fenomeni Aleatori

La probabilità è un **framework matematico** rigoroso che ci permette di descrivere, analizzare e fare inferenze su fenomeni aleatori nel mondo che ci circonda. Per fenomeni aleatori intendiamo eventi o esperimenti il cui esito non possiamo prevedere con certezza, ma che possiamo studiare sistematicamente.

Questo framework ci fornisce:

- Un linguaggio preciso per parlare di incertezza
- Regole matematiche per combinare probabilità
- Strumenti per prendere decisioni razionali sotto incertezza
- Metodi per aggiornare le nostre credenze con nuovi dati

La teoria della probabilità è alla base di discipline moderne come machine learning, data science, statistica inferenziale, e teoria delle decisioni.

Interpretazione Frequentista

Basata sulla **frequenza relativa** osservata in un grande numero di ripetizioni dello stesso esperimento. Esempio: lanciando una moneta equilibrata molte volte, la proporzione di "testa" converge a 1/2.

Interpretazione Soggettivista

Quantificazione del **grado di credenza personale** razionale che qualcosa accadrà. Esempio: stimare la probabilità che piova oggi basandosi su nuvole, umidità e esperienza personale.

Esperimenti Aleatori e Spazio Campionario

Per costruire una teoria matematica della probabilità, dobbiamo prima definire con precisione gli elementi fondamentali con cui lavoreremo:

Esperimento Aleatorio

Un processo attraverso il quale osserviamo qualcosa di incerto. L'esperimento può essere ripetuto in condizioni simili, ma il risultato può variare da una ripetizione all'altra. Caratteristiche: ripetibilità, esito non predicibile, insieme di esiti ben definito.

Esito (Outcome)

Il risultato singolo e specifico di un esperimento aleatorio. Ogni volta che eseguiamo l'esperimento, otteniamo esattamente uno degli esiti possibili. Gli esiti sono mutuamente esclusivi: non possono verificarsi contemporaneamente.

Spazio Campionario (S o Ω)

L'insieme completo e esaustivo di tutti i possibili esiti di un esperimento. Lo spazio campionario deve contenere ogni possibile risultato dell'esperimento, e nessun altro elemento. È la "base" su cui costruiamo la nostra teoria.

Esempi di Spazi Campionari

Lo spazio campionario dipende da come definiamo l'esperimento. La scelta dello spazio campionario appropriato è il primo passo cruciale nell'analisi probabilistica:

Esperimento Aleatorio	Spazio Campionario	Caratteristiche
Lancio di una moneta	S = {Testa, Croce}	Finito, 2 elementi, discreto
Lancio di un dado	S = {1, 2, 3, 4, 5, 6}	Finito, 6 elementi, discreto
Due lanci di moneta	S = {TT, TC, CT, CC}	Finito, 4 elementi, sequenziale
Tre lanci di moneta	S = {TTT, TTC, TCT, TCC, CTT, CTC, CCT, CCC}	Finito, $2^3 = 8$ elementi
Numero di clienti in un giorno	S = {0, 1, 2, 3, ...}	Infinito numerabile, discreto
Tempo di attesa (minuti)	$S = [0, +\infty)$	Infinito non numerabile, continuo
Temperatura giornaliera (°C)	$S = (-\infty, +\infty)$	Continuo, tutti i numeri reali

⚠️ **Importante:** Per esperimenti composti (come "3 lanci di moneta"), lo spazio campionario cresce esponenzialmente. Se ogni lancio ha n esiti e facciamo k lanci, lo spazio campionario ha n^k elementi.

Domande: Esperimenti Aleatori e Spazi Campionari

Mettiamo in pratica quanto appreso! Per ogni esperimento aleatorio descritto di seguito, prova a identificare lo spazio campionario (Ω) e a descriverne le caratteristiche.

Lancio di una moneta truccata

Se lanciamo una moneta una volta, e l'esito "Testa" ha una probabilità diversa da "Croce", qual è lo spazio campionario? Quali sono le sue caratteristiche principali?

Estrazione di una carta da un mazzo

Estraiamo una singola carta da un mazzo standard di 52 carte. Qual è lo spazio campionario che descrive tutti i possibili esiti? È finito, numerabile o continuo?

Misurazione della durata di una lampadina

Misuriamo il tempo (in ore) in cui una lampadina funziona prima di bruciarsi. Qual è lo spazio campionario di questo esperimento? Qual è la sua natura (discreta o continua)?

Somma dei risultati di due dadi

Lanciamo due dadi equi e sommiamo i punti ottenuti. Qual è lo spazio campionario degli esiti possibili per questa somma? Come lo descriveresti in termini di grandezza e tipo?

Soluzioni: Esperimenti Aleatori e Spazi Campionari

Ecco le soluzioni e le spiegazioni per gli esercizi sugli esperimenti aleatori e gli spazi campionari:

Lancio di una moneta truccata

Spazio Campionario (Ω): $\{\text{Testa}, \text{Croce}\}$ o $\{T, C\}$

Caratteristiche: Finito, con 2 elementi distinti. È discreto. Il fatto che la moneta sia "truccata" influisce solo sulla probabilità di ottenere Testa o Croce, ma non modifica l'insieme degli esiti possibili.

Estrazione di una carta da un mazzo

Spazio Campionario (Ω): L'insieme di tutte le 52 carte del mazzo (es. $\{\text{Asso di Picche}, 2 \text{ di Picche}, \ldots, \text{Re di Quadri}\}$).

Caratteristiche: Finito, con 52 elementi. È discreto, poiché ogni carta è un esito distinto e separato.

Misurazione della durata di una lampadina

Spazio Campionario (Ω): $[0, +\infty)$ (tutti i numeri reali non negativi).

Caratteristiche: Infinito e non numerabile. È continuo, poiché la durata può assumere qualsiasi valore all'interno di un intervallo di tempo.

Somma dei risultati di due dadi

Spazio Campionario (Ω): $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

Caratteristiche: Finito, con 11 elementi. È discreto, dato che la somma può assumere solo valori interi.

Comprendere come definire lo spazio campionario è fondamentale per la corretta impostazione di qualsiasi problema di probabilità.

Eventi e Operazioni Fondamentali

Cos'è un Evento?

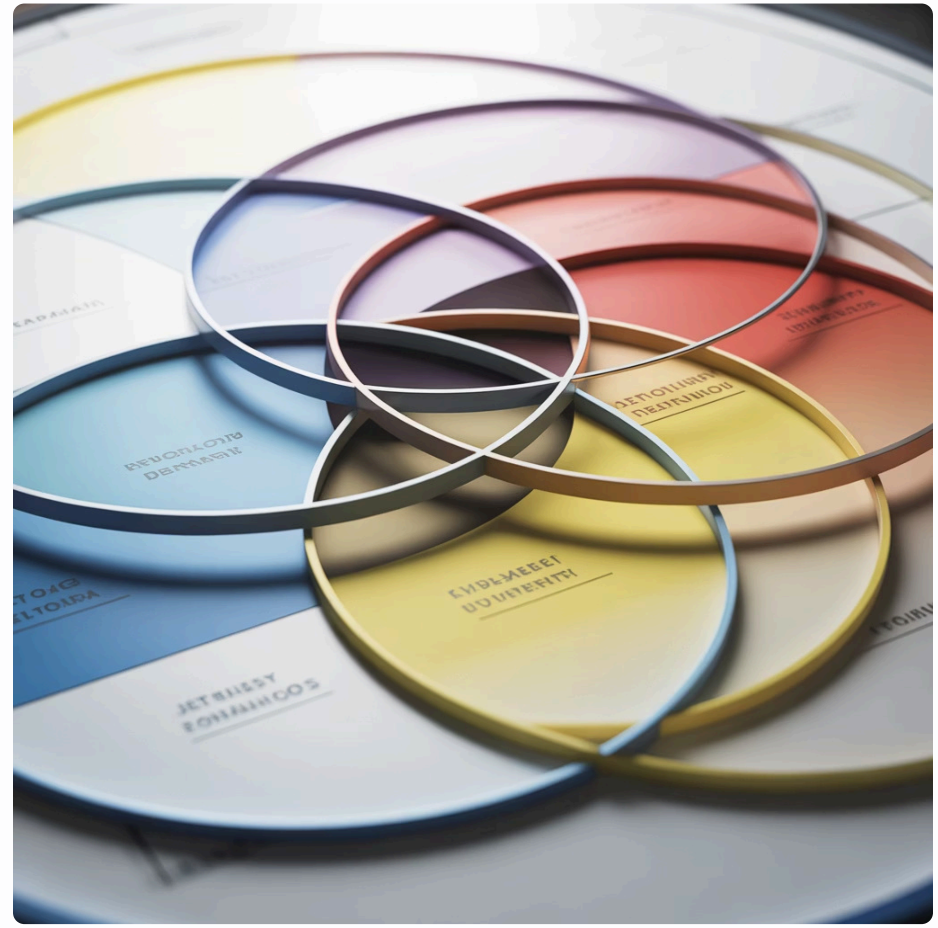
Un **evento** è un sottoinsieme dello spazio campionario S al quale possiamo assegnare una probabilità. Un evento è anche definito come un'affermazione o proposizione a cui possiamo assegnare univocamente un valore di verità (Vero o Falso) dopo aver osservato il risultato dell'esperimento.

Esempio Fondamentale: Nel lancio di un dado con S = {1, 2, 3, 4, 5, 6}:

- L'evento "numero pari" è E = {2, 4, 6}
- L'evento "numero maggiore di 4" è F = {5, 6}
- L'evento "numero primo" è G = {2, 3, 5}

Eventi Speciali:

- **Evento certo**: S (si verifica sempre)
- **Evento impossibile**: ∅ (non si verifica mai)
- **Eventi elementari**: singoli esiti, es. {3}



Convenzioni e Notazioni:

- Spazio campionario: **S** oppure **Ω**
- Evento: **A, B, C** ⊆ S
- Cardinalità: |**A**| = numero di esiti in A
- Complemento: **A^c** o **Ā**
- Unione: **A ∪ B**
- Intersezione: **A ∩ B**
- Probabilità condizionata: P(A|B)

Operazioni sugli Eventi

Come per gli insiemi matematici, possiamo combinare eventi usando operazioni logiche. Queste operazioni sono fondamentali per costruire eventi complessi da eventi semplici:

Unione: $A \cup B$

Significato: L'evento che si verifica se A o B (o entrambi) si verificano.

Parole chiave: "o", "almeno uno di", "oppure"

Esempio: Dado, $A = \{2,4,6\}$ (pari), $B = \{1,2,3\}$ (≤ 3)
$A \cup B = \{1,2,3,4,6\}$ (pari O ≤ 3)

Intersezione: $A \cap B$

Significato: L'evento che si verifica se A e B si verificano entrambi simultaneamente.

Parole chiave: "e", "entrambi", "contemporaneamente"

Esempio: Con A e B come sopra
$A \cap B = \{2\}$ (pari E ≤ 3)

Complemento: A^c

Significato: L'evento che si verifica se A non si verifica.

Parole chiave: "non", "contrario di", "negazione"

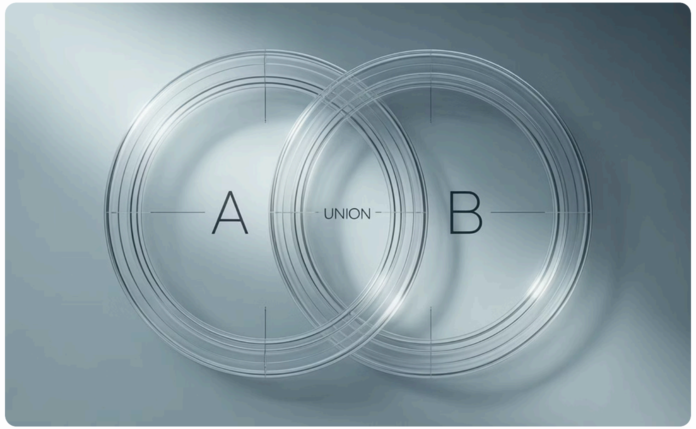
Esempio: Con $A = \{2,4,6\}$
$A^c = \{1,3,5\}$ (non pari = dispari)

Proprietà Importanti:

- **Eventi disgiunti:** $A \cap B = \emptyset$ (incompatibili)
- **Legge di De Morgan:** $(A \cup B)^c = A^c \cap B^c$
- **Complemento doppio:** $(A^c)^c = A$
- **Inclusione:** Se $A \subseteq B$, ogni esito di A è anche in B
- **Partizione:** Eventi disgiunti che coprono tutto S
- **Distributività:** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

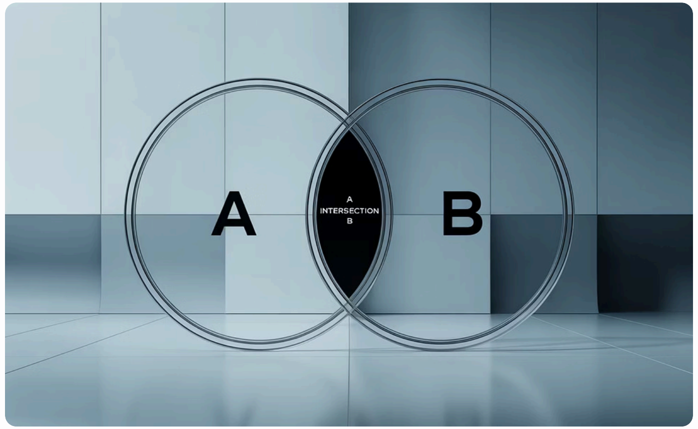
Visualizzazione delle Operazioni sugli Eventi

I diagrammi di Venn sono uno strumento eccellente per comprendere visivamente le operazioni logiche tra gli eventi nel contesto della probabilità.



Unione: A ∪ B

L'unione di A e B include tutti gli esiti che appartengono ad A, a B, o a entrambi. Si verifica se almeno uno dei due eventi accade.



Intersezione: A ∩ B

L'intersezione di A e B include solo gli esiti comuni ad A e B. Si verifica se entrambi gli eventi A e B accadono simultaneamente.

Complemento: Ac

Il complemento di A include tutti gli esiti nello spazio campionario S che non appartengono ad A. Si verifica se l'evento A non accade.

Queste visualizzazioni aiutano a chiarire le definizioni e le relazioni tra gli eventi, essenziali per il calcolo delle probabilità.

Definizione Classica della Probabilità

L'Approccio di Laplace (Combinatorio)

Quando si Usa

La definizione classica si applica quando tutti gli esiti dello spazio campionario sono **equiprobabili**, cioè hanno la stessa probabilità di verificarsi. Questa è la definizione più antica e intuitiva della probabilità.

Formula Fondamentale

La probabilità P(A) di un evento A è il rapporto tra il numero dei casi favorevoli e il numero dei casi possibili:

$$P(A) = \frac{\text{Casi favorevoli}}{\text{Casi possibili}} = \frac{|A|}{|S|}$$

dove |A| è il numero di esiti nell'evento A e |S| è il numero totale di esiti possibili.

Esempio Classico

Lancio di un dado equilibrato:

S = {1, 2, 3, 4, 5, 6}, quindi |S| = 6

Evento A = "esce 4" = {4}, quindi |A| = 1

$$P(A) = \frac{1}{6} \approx 0.167 = 16.7\%$$

Evento "numero pari":

B = {2, 4, 6}, quindi |B| = 3

$$P(B) = \frac{3}{6} = \frac{1}{2} = 0.5 = 50\%$$

Limiti della Definizione Classica

Il Problema della Circolarità

Sebbene intuitiva, la definizione classica presenta un grave limite logico: è formalmente circolare. Per definire la probabilità, essa richiede che i casi possibili abbiano tutti la stessa probabilità (equiprobabilità), che è proprio il concetto che stiamo cercando di definire!

📍 Il Caso della Gara Roma-Ostia

Immaginate una gara da Roma a Ostia. Io ho la **bicicletta (B)**, l'altra persona ha la **macchina (M)**.

Casi Possibili: 2 (vince B, vince M)

Applicando ingenuamente la definizione classica:

- P(vince B) = 1/2 = 50%
- P(vince M) = 1/2 = 50%

❌ Perché Questo Risultato è Assurdo

Nessuno, in condizioni di incertezza reale, attribuirebbe il 50% di probabilità a chi usa la bicicletta contro chi usa la macchina! Gli eventi **non sono equiprobabili**.

La realtà: P(vince M) $\gg$ P(vince B). La macchina è oggettivamente più veloce. Una stima ragionevole potrebbe essere P(vince M) = 0.95 e P(vince B) = 0.05.

Conclusione: Non possiamo usare la definizione classica in situazioni non stereotipate dove l'equiprobabilità non è garantita.

💡 **Lezione Importante:** La definizione classica funziona bene per dadi, monete, carte da gioco, ma fallisce per eventi complessi del mondo reale. Servono approcci più generali come le definizioni frequentista e soggettivista.

Esempi di Applicazione della Definizione Classica

Nonostante i suoi limiti, la definizione classica della probabilità è perfettamente valida e utile quando la condizione di equiprobabilità è intrinsecamente soddisfatta. Ecco alcuni scenari reali dove è ancora il modo più intuitivo e corretto per calcolare le probabilità:

Lancio di una Moneta Equilibrata

Quando lanci una moneta "equilibrata" (fair coin), ci sono solo due esiti possibili: Testa o Croce. Per sua natura, la moneta non ha alcuna preferenza per un lato rispetto all'altro.

$$P(\text{Testa}) = \frac{1}{2} = 50\%$$

Estrazione di una Carta da un Mazzo

Se estrai una singola carta da un mazzo standard di 52 carte ben mescolato, ogni carta ha la stessa identica probabilità di essere estratta. Il mescolamento garantisce l'equiprobabilità.

$$P(\text{Re di Cuori}) = \frac{1}{52}$$

Estrazione da un'Urna con Palline Identiche

Consideriamo un'urna con 10 palline, tutte della stessa dimensione e peso, di cui 3 rosse e 7 blu. L'atto di pescare una pallina rende ogni pallina ugualmente probabile.

$$P(\text{Pallina Rossa}) = \frac{3}{10} = 30\%$$

Questi esempi dimostrano che in contesti dove la simmetria e l'equilibrio sono garantiti, la definizione classica offre un metodo semplice e accurato per quantificare l'incertezza.

Definizione Frequentista

Probabilità come Frequenza Relativa

Quando si Usa

La definizione frequentista si applica quando l'esperimento è **ripetibile in condizioni analoghe**, ma i risultati non sono necessariamente equiprobabili. È l'approccio preferito nelle scienze sperimentali e nell'analisi statistica.

Concetto Fondamentale

La probabilità è valutata sulla base della frequenza relativa osservata nel passato in un gran numero di prove. All'aumentare del numero di ripetizioni, la frequenza relativa converge alla "vera" probabilità (Legge dei Grandi Numeri).

Formula di Valutazione

La probabilità P(A) è approssimata dalla frequenza relativa F(A):

$$P(A) \approx F(A) = \frac{K}{N}$$

dove K = numero di volte che A si verifica, N = numero totale di prove (con N grande)

Esempio: Moneta Truccata

Abbiamo una moneta di cui sospettiamo non sia equilibrata. Eseguiamo 1000 lanci:

- Testa (T): 610 volte
- Croce (C): 390 volte

Stima frequentista:

$$P(T) \approx \frac{610}{1000} = 0.61 = 61\%$$

$$P(C) \approx \frac{390}{1000} = 0.39 = 39\%$$

Conclusione: La moneta è sbilanciata verso Testa con circa 61% di probabilità.

Limiti dell'Approccio Frequentista

Sebbene oggettivo e verificabile sperimentalmente, questo approccio ha dei limiti importanti:

- **Non applicabile a eventi unici:** Non possiamo usarlo per calcolare la probabilità che piova domani in una specifica città (evento non ripetibile)
- **Richiede molte osservazioni:** Servono centinaia o migliaia di prove per ottenere stime affidabili
- **Condizioni identiche difficili:** Nella pratica è difficile ripetere un esperimento in condizioni davvero identiche
- **Eventi rari problematici:** Per eventi molto rari servirebbe un numero proibitivo di osservazioni

Esempi Utili della Definizione Frequentista

La definizione frequentista trova la sua massima utilità in contesti dove è possibile osservare un fenomeno ripetutamente nel tempo. È l'approccio ideale per stimare probabilità basandosi su dati storici e osservazioni empiriche.

Rischio di Incidenti Automobilistici

Le compagnie assicurative calcolano la probabilità di un incidente per stabilire i premi, basandosi sull'analisi di milioni di sinistri passati. Non c'è equiprobabilità, ma una vasta base di dati.

$$P(\text{Incidente}) \approx \frac{\text{Numero incidenti}}{\text{Numero polizze totali}}$$

Controllo Qualità Industriale

In una fabbrica, si può stimare la probabilità che un prodotto sia difettoso testando un grande campione dalla produzione. Questo aiuta a monitorare la qualità e a identificare problemi.

$$P(\text{Difettoso}) \approx \frac{\text{Numero prodotti difettosi}}{\text{Numero prodotti ispezionati}}$$

Statistiche Sportive

Le probabilità di vittoria di una squadra o di un atleta vengono calcolate analizzando le performance passate, i risultati storici degli incontri, le statistiche individuali e di squadra.

$$P(\text{Vittoria}) \approx \frac{\text{Numero vittorie}}{\text{Numero partite giocate}}$$

Questi esempi evidenziano come la definizione frequentista sia cruciale per prendere decisioni basate su dati reali, specialmente in assenza di simmetria o equiprobabilità intrinseca degli eventi.

Definizione Soggettivista

Probabilità come Grado di Fiducia

La definizione soggettivista, sviluppata principalmente da Bruno de Finetti, rappresenta l'approccio più generale e flessibile alla probabilità. È fondamentale per l'inferenza bayesiana e l'intelligenza artificiale.

Quando si Usa

In situazioni in cui non si applicano né la definizione classica (non c'è equiprobabilità), né quella frequentista (l'evento è unico o non ripetibile). Ideale per previsioni, decisioni strategiche, machine learning.

Concetto Fondamentale

La probabilità è la quantificazione del **grado di credenza personale** (o grado di fiducia) che un individuo razionale ripone nel verificarsi di un evento, sulla base delle informazioni disponibili.

Valutazione tramite Scommessa

Il grado di fiducia viene misurato attraverso la disponibilità a scommettere: quanto saresti disposto a pagare per vincere 1€ se l'evento si verifica?

Il Principio di Coerenza

Uno **scommettitore razionale e coerente** deve soddisfare questa condizione:

> La quota P che si è disposti a pagare per ricevere una somma S = 1€ se l'evento si verifica (e 0€ altrimenti) deve essere un numero compreso tra 0 e 1: **$0 \leq P \leq 1$**

La condizione di coerenza assicura che non si possa costruire un sistema di scommesse che garantisca una vincita o una perdita certa (no *Dutch Book*). È un principio di "buon senso" matematico.

Esempio Pratico

Partita di calcio: Juventus vs Inter

Sulla base di forma attuale, infortuni, statistiche storiche, un tifoso esperto valuta:

- P(vince Juventus) = 0.40
- P(pareggio) = 0.30
- P(vince Inter) = 0.30

Questo significa che sarebbe disposto a pagare 40 centesimi per vincere 1€ se vince la Juventus.

Vantaggi dell'Approccio

- Applicabile a **qualsiasi** evento
- Incorpora conoscenza pregressa
- Permette aggiornamento con nuovi dati
- Base dell'inferenza bayesiana
- Fondamentale per AI e machine learning

Quale Approccio Scegliere?

Mettiamo alla prova la tua comprensione delle diverse definizioni di probabilità. Per ciascuno scenario, individua quale approccio — classico, frequentista o soggettivista — sarebbe il più appropriato per determinare la probabilità.

Scenario 1: Il Lancio del Dado

Stai lanciando un dado perfettamente equilibrato a sei facce. Qual è la probabilità che esca un numero pari?

👉 Quale definizione useresti e perché?

Scenario 2: Lampadine Difettose

Un produttore ha testato migliaia di lampadine LED e ha registrato un tasso di difettosità del 0.5%. Qual è la probabilità che la prossima lampadina prodotta sia difettosa?

👉 Quale definizione useresti e perché?

Scenario 3: Successo del Nuovo Modello AI

La tua azienda ha sviluppato un nuovo modello di intelligenza artificiale, unico nel suo genere, per prevedere le tendenze di mercato. Qual è la probabilità che questo modello ottenga una precisione superiore al 90% nei prossimi 6 mesi?

👉 Quale definizione useresti e perché?

Riflettere su questi scenari ti aiuta a consolidare la comprensione di quando applicare ciascuna definizione.

Soluzioni: Quale Approccio Scegliere?

Vediamo insieme le risposte agli scenari proposti, analizzando il ragionamento dietro ogni scelta.

Scenario 1: Il Lancio del Dado

Definizione Classica

Il dado è perfettamente equilibrato, garantendo l'equiprobabilità di ogni faccia. Questo ci permette di calcolare la probabilità contando i casi favorevoli (3 numeri pari) rispetto ai casi totali (6 facce).

Scenario 2: Lampadine Difettose

Definizione Frequentista

La probabilità si basa su osservazioni ripetute e dati storici ("migliaia di lampadine testate"). Il tasso di difettosità del 0.5% è una frequenza relativa osservata, utilizzabile per prevedere l'esito della prossima lampadina.

Scenario 3: Successo del Nuovo Modello AI

Definizione Soggettivista

Il modello AI è "unico nel suo genere", non esistono dati storici o equiprobabilità. La probabilità è un grado di fiducia personale, basato sull'expertise, le informazioni disponibili e il giudizio, fondamentale per decisioni strategiche e l'AI.

Questi esempi dimostrano l'importanza di saper scegliere l'approccio più adatto a seconda della natura dell'evento e delle informazioni disponibili.

Gli Assiomi di Kolmogorov

La Fondazione Inconfutabile della Probabilità

Per superare i limiti delle definizioni classica, frequentista e soggettivista, e per dotare la probabilità di un fondamento matematico solido, Andrei Kolmogorov introdusse un approccio assiomatico. Ma perché è così importante avere una definizione "astratta" come questa?

Unificazione Universale

Forniscono un linguaggio e una struttura comuni per **tutti i tipi** di probabilità, integrandoli in un unico quadro matematico coerente. Non importa l'evento, gli assiomi valgono sempre.

Rigore Matematico

Elevano la probabilità a una disciplina matematica formale, consentendo la derivazione logica di tutte le sue proprietà e teoremi, senza ambiguità o dipendenza da interpretazioni specifiche.

Coerenza Infallibile

Garantiscono che qualsiasi assegnazione di probabilità, indipendentemente dalla sua origine, sia **internamente consistente** e non porti a contraddizioni logiche, eliminando i "paradossi" probabilistici.

Base per l'AI e il Data Science

Sono il pilastro teorico su cui si costruiscono tutte le teorie probabilistiche avanzate, essenziali per algoritmi complessi di apprendimento automatico, inferenza bayesiana e modelli predittivi moderni.

In sintesi, gli assiomi di Kolmogorov forniscono la grammatica e la sintassi fondamentali che permettono alla probabilità di essere un potente strumento universale per la scienza e la tecnologia.

Gli Assiomi di Kolmogorov

Le Fondamenta Universali della Probabilità

Indipendentemente da come si valuta la probabilità (classica, frequentista, soggettiva), essa deve rispettare delle regole matematiche universali. Queste regole sono i **tre assiomi di Kolmogorov**, che forniscono la sintassi matematica della teoria della probabilità:

Assioma 1: Non-Negatività

La probabilità di qualsiasi evento A non può essere negativa:

$$P(A) \geq 0$$

Significato: La probabilità è sempre un numero positivo o zero. Il valore minimo possibile è zero (evento impossibile). Questo assioma riflette il fatto che le probabilità rappresentano quantità di "credenza" o "frequenza", che non possono essere negative.

Assioma 2: Normalizzazione

La probabilità dello spazio campionario S (l'evento certo, Ω) è uguale a 1:

$$P(S) = P(\Omega) = 1$$

Significato: Uno degli esiti possibili si verificherà sicuramente. La somma di tutte le probabilità degli eventi elementari deve essere 100%. Questo assioma "normalizza" la scala delle probabilità tra 0 e 1.

Assioma 3: Additività

Se due eventi A e B sono disgiunti (incompatibili: $A \cap B = \emptyset$), allora:

$$P(A \cup B) = P(A) + P(B)$$

Significato: Se due eventi non possono verificarsi insieme, la probabilità che si verifichi almeno uno dei due è la somma delle probabilità individuali. Questo si estende a qualsiasi collezione finita o numerabile di eventi mutuamente disgiunti.

💡 **Insight Fondamentale:** Gli assiomi forniscono la **sintassi** (le regole formali) che la probabilità deve seguire, mentre le definizioni (classica, frequentista, soggettiva) si occupano della **semantica** (il significato di quel numero). Tutti i metodi di valutazione coerenti rispettano questi assiomi.

Applicazione degli Assiomi: Esempio Completo

Per comprendere meglio l'importanza e la logica degli assiomi di Kolmogorov, esploriamo come si applicano a scenari concreti, fornendo un fondamento matematico coerente a qualsiasi calcolo di probabilità.

Assioma 1: Non-Negatività

Considera il lancio di una moneta. La probabilità di ottenere "testa" è $P(\text{Testa}) = 0.5$, un valore non negativo. Non si può mai avere una probabilità negativa. Allo stesso modo, la probabilità di un evento impossibile, come ottenere "7" da un dado a sei facce, è $P(7) = 0$.

Assioma 2: Normalizzazione

Lanciando un dado a sei facce, l'evento certo è ottenere un numero qualsiasi tra 1 e 6. La probabilità di questo evento (l'intero spazio campionario Ω) è $P(\{1, 2, 3, 4, 5, 6\}) = 1$. Questo significa che uno degli esiti possibili si verificherà sempre.

Assioma 3: Additività

Se lanci un dado, qual è la probabilità di ottenere un "1" o un "2"? Gli eventi "ottenere 1" e "ottenere 2" sono disgiunti (non possono verificarsi contemporaneamente). Quindi, $P(1 \cup 2) = P(1) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$.

Questi esempi illustrano come i tre assiomi definiscono le regole fondamentali che ogni assegnazione di probabilità deve soddisfare, garantendo consistenza e rigore matematico.

Applicazione degli Assiomi: Esempio Completo

Elezione Presidenziale

Vediamo come gli assiomi si applicano in un caso pratico realistico. Quattro candidati si sfidano per la presidenza: A, B, C e D. Sulla base di sondaggi, analisi politica e dati storici, un analista stima le seguenti probabilità:

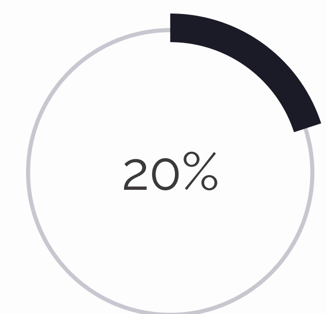
Probabilità Individuali Stimate

- P(A vince) = 0.20 (20%)
- P(B vince) = 0.40 (40%)
- P(C vince) = 0.15 (15%)
- P(D vince) = 0.25 (25%)

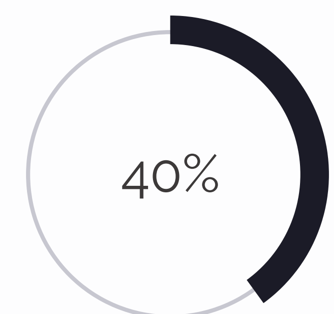
Verifica Assioma 2 (Normalizzazione):

$$P(A) + P(B) + P(C) + P(D) = 0.20 + 0.40 + 0.15 + 0.25 = 1.00 \checkmark$$

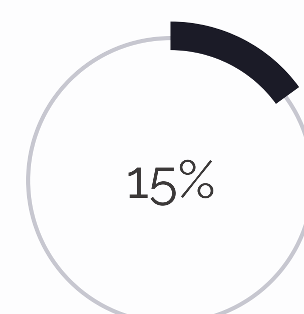
La somma è 1, quindi uno dei candidati vincerà sicuramente (evento certo).



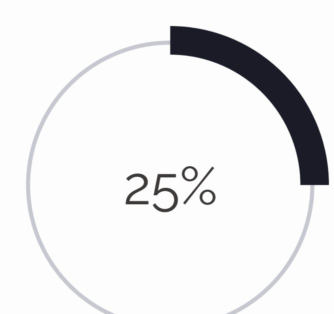
Candidato A



Candidato B



Candidato C



Candidato D

Domanda: Qual è la probabilità che A o B vinca l'elezione?

01

1. Identificare gli Eventi

Evento A = "candidato A vince" con P(A) = 0.20
Evento B = "candidato B vince" con P(B) = 0.40

02

2. Verificare che siano Disgiunti

A e B sono eventi disgiunti perché **non possono vincere entrambi** la stessa elezione. Solo uno può essere presidente. Quindi: A ∩ B = ∅

03

3. Applicare l'Assioma 3 (Additività)

Poiché A e B sono disgiunti, possiamo sommare le probabilità:

$$P(A \cup B) = P(A) + P(B) = 0.20 + 0.40 = 0.60$$

04

4. Interpretare il Risultato

C'è una probabilità del **60%** che vinca A o B. Di conseguenza, c'è una probabilità del 40% che vinca C o D.

⚠️ **Attenzione:** Possiamo sommare le probabilità SOLO perché gli eventi sono disgiunti. Se A e B potessero verificarsi insieme, dovremmo usare la formula dell'unione: P(A ∪ B) = P(A) + P(B) - P(A ∩ B)

Proprietà Derivate: Complemento ed Evento Impossibile

Dai tre assiomi fondamentali possiamo derivare molte proprietà utili. Vediamo le prime due proprietà fondamentali:

1

Probabilità dell'Evento Impossibile

$$P(\emptyset) = 0$$

L'insieme vuoto (evento impossibile) ha probabilità zero. Nessun esito appartiene all'evento impossibile, quindi non può mai verificarsi.

Dimostrazione rapida: ∅ e S sono disgiunti, ∅ ∪ S = S. Per l'Assioma 3: P(S) = P(∅) + P(S). Quindi P(∅) = 0.

2

Probabilità del Complemento

$$P(A^c) = 1 - P(A)$$

La probabilità che A non si verifichi è uno meno la probabilità di A. Questa è una delle formule più utili nel calcolo delle probabilità.

Esempio: Se P(piove) = 0.30, allora P(non piove) = 1 - 0.30 = 0.70 = 70%

Dimostrazione Completa: P(A^c) = 1 - P(A)

1

Passo 1: Osserviamo che A e A^c sono eventi disgiunti (non si sovrappongono):

$$A \cap A^c = \emptyset$$

2

Passo 2: Inoltre, A e A^c insieme formano tutto lo spazio campionario:

$$A \cup A^c = S$$

3

Passo 3: Per l'Assioma 2, sappiamo che:

$$P(S) = 1$$

quindi:

$$P(A \cup A^c) = 1$$

4

Passo 4: Poiché A e A^c sono disgiunti, per l'Assioma 3:

$$P(A \cup A^c) = P(A) + P(A^c)$$

5

Passo 5: Quindi:

$$1 = P(A) + P(A^c)$$

Da cui:

$$P(A^c) = 1 - P(A) \quad \blacksquare$$

Esempi Applicativi Completi delle Proprietà Derivate

Vediamo come le proprietà derivate dagli assiomi di Kolmogorov si manifestano in situazioni pratiche, fornendo strumenti essenziali per l'analisi probabilistica nel mondo reale.

Probabilità dell'Evento Impossibile: Il Quadrifoglio a Cinque Foglie

In natura, un quadrifoglio è definito dall'avere quattro foglie. La probabilità di trovare un quadrifoglio con esattamente cinque foglie, non per una mutazione ma come "evento regolare" in un campo, è considerata impossibile.

$$P(\text{Quadrifoglio con 5 foglie}) = 0$$

Questo riflette la regola $P(\emptyset) = 0$: per quanto si cerchi, un evento che per sua natura non può verificarsi ha probabilità zero.

Probabilità del Complemento: Arrivo dei Voli

Supponiamo che una compagnia aerea affermi che la probabilità che uno dei suoi voli arrivi in orario (Evento O) sia del 85%.

$$P(O) = 0.85$$

Vogliamo sapere qual è la probabilità che un volo non arrivi in orario (Evento O^c), il che include ritardi o cancellazioni. Utilizzando la proprietà del complemento:

$$P(O^c) = 1 - P(O) = 1 - 0.85 = 0.15$$

Questo significa che c'è una probabilità del 15% che un volo di questa compagnia non rispetti l'orario previsto, dimostrando come $P(A^c) = 1 - P(A)$ sia cruciale per calcolare la probabilità di "non accadimento" di un evento.

Proprietà Derivate: Monotonia e Unione

Proprietà 3: Monotonia

$$Se\, A \subseteq B,\, allora\, P(A) \leq P(B)$$

Se l'evento A è contenuto nell'evento B (ogni esito di A è anche in B), allora la probabilità di A non può superare quella di B. Questa proprietà riflette l'intuizione che eventi "più grandi" hanno probabilità maggiore o uguale.

Esempio: Lancio di dado. A = {6} (esce 6), B = {4,5,6} (esce ≥4). Chiaramente A ⊆ B, e infatti P(A) = 1/6 ≤ P(B) = 3/6 = 1/2. ✓

Proprietà 4: Formula dell'Unione

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Questa è la formula generale dell'unione, valida per eventi **non necessariamente disgiunti**. Sottraiamo P(A ∩ B) per evitare di contare due volte gli esiti che appartengono sia ad A che a B.

Caso speciale: Se A e B sono disgiunti (A ∩ B = ∅), allora P(A ∩ B) = 0 e ritorniamo all'Assioma 3: P(A ∪ B) = P(A) + P(B).

Esempio Completo: Lancio di un Dado

Problema: Nel lancio di un dado equo, qual è la probabilità di ottenere un numero pari O un numero primo?

Dati del Problema

- S = {1, 2, 3, 4, 5, 6}
- A = {2, 4, 6} (numeri pari)
- B = {2, 3, 5} (numeri primi)
- A ∩ B = {2} (pari E primo)

Probabilità individuali:

- P(A) = |A|/|S| = 3/6 = 1/2
- P(B) = |B|/|S| = 3/6 = 1/2
- P(A ∩ B) = 1/6

Soluzione

Applicando la formula dell'unione:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{1}{2} + \frac{1}{2} - \frac{1}{6}$$

$$= \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \frac{5}{6}$$

Risultato: P(pari O primo) = 5/6 ≈ 0.833 = 83.3%

Verifica: A ∪ B = {2, 3, 4, 5, 6}, che ha 5 elementi, quindi 5/6. ✓

Esempi Applicativi Completi

Vediamo quattro esempi che illustrano l'applicazione pratica degli assiomi e delle proprietà derivate in contesti diversi:

1. Lancio di un Dado

Problema: Qual è la probabilità di ottenere un numero pari?

Soluzione: $S = \{1,2,3,4,5,6\}$, $A = \{2,4,6\}$. Usando l'equiprobabilità: $P(A) = |A|/|S| = 3/6 = 1/2 = 0.5 = 50\%$

2. Elezioni Presidenziali

Problema: Quattro candidati. $P(A)=0.20$, $P(B)=0.40$, $P(C)=0.15$. Qual è $P(D)$?

Soluzione: Eventi disgiunti ed esaustivi. Usando l'Assioma 2 e 3: $P(A)+P(B)+P(C)+P(D)=1$, quindi $P(D)=1-0.75=0.25$

3. Estrazione da Mazzo

Problema: Probabilità di estrarre una carta rossa O una figura?

Soluzione: A=rosse (26/52), B=figure (12/52), $A \cap B$=figure rosse (6/52). Formula unione: $P(A \cup B)=26/52+12/52-6/52=32/52\approx0.615$

4. Complemento

Problema: In un dado, qual è la probabilità di NON ottenere 6?

Soluzione: $A=\{6\}$, $P(A)=1/6$. Usando la proprietà del complemento: $P(A^c)=1-P(A)=1-1/6=5/6\approx0.833$

💡 **Osservazione:** Tutti questi calcoli derivano rigorosamente dai tre assiomi fondamentali della probabilità. La matematica della probabilità è elegante e coerente!

Riepilogo dei Concetti Chiave

Sintesi della Lezione

1 Framework Matematico Rigoroso

La probabilità è un framework matematico rigoroso per descrivere e analizzare fenomeni aleatori. È applicabile con interpretazione frequentista (basata su frequenze osservate) o soggettivista (basata su gradi di credenza razionali). Entrambe le interpretazioni rispettano gli stessi assiomi matematici.

2 Esperimenti, Esiti e Spazi Campionari

Ogni esperimento aleatorio genera esiti appartenenti a uno spazio campionario S, che contiene tutti i possibili risultati. Gli eventi sono sottoinsiemi di S ai quali assegniamo probabilità. La scelta corretta dello spazio campionario è fondamentale per l'analisi.

3 I Tre Assiomi di Kolmogorov

Tutta la teoria della probabilità si costruisce su tre assiomi fondamentali: **(1) Non-negatività** $P(A) \geq 0$, **(2) Normalizzazione** $P(S)=1$, **(3) Additività** per eventi disgiunti. Questi assiomi forniscono le regole sintattiche universali che ogni probabilità deve rispettare.

4 Proprietà Derivate Fondamentali

Dagli assiomi derivano proprietà utili: $P(\varnothing)=0$, $P(A^c)=1-P(A)$, Monotonia ($A \subseteq B \Rightarrow P(A) \leq P(B)$), Formula dell'Unione $P(A \cup B)=P(A)+P(B)-P(A \cap B)$. Queste proprietà semplificano enormemente i calcoli pratici.

5 Operazioni sugli Eventi

Gli eventi si combinano usando operazioni insiemistiche: **Unione** ($A \cup B$ = "A o B"), **Intersezione** ($A \cap B$ = "A e B"), **Complemento** (A^c = "non A"). Queste operazioni permettono di costruire eventi complessi da eventi semplici e sono alla base di tutti i calcoli probabilistici.

Mettiti alla Prova: Domande e Problemi

Rafforza la tua comprensione con questi semplici problemi che riassumono i concetti chiave discussi finora.

1

Eventi in un Mazzo di Carte

Da un mazzo di 52 carte da gioco, estrai una carta. Qual è la probabilità che la carta sia di Fiori O una Regina?

2

Rischio di Ritardo in un Progetto

In un progetto, la probabilità che il lancio di un nuovo prodotto (Evento A) subisca un ritardo è del 15% ($P(A)=0.15$). La probabilità che la campagna marketing (Evento B) subisca un ritardo è del 10% ($P(B)=0.10$). Se la probabilità che entrambi gli eventi subiscano un ritardo contemporaneamente è del 5% ($P(A \cap B)=0.05$), qual è la probabilità che almeno uno dei due eventi subisca un ritardo?

Soluzioni: Mettiti alla Prova

1. Eventi in un Mazzo di Carte

Applichiamo la formula dell'unione per risolvere il problema dell'estrazione di carte.

Definizione degli Eventi

- **Spazio Campionario (S):** Un mazzo standard di 52 carte. Quindi, |S| = 52.
- **Evento A:** Estrarre una carta di Fiori (♣). Ci sono 13 carte di Fiori nel mazzo. P(A) = 13/52.
- **Evento B:** Estrarre una Regina (Q). Ci sono 4 Regine nel mazzo (Q♣, Q♦, Q♥, Q♠). P(B) = 4/52.
- **Intersezione (A ∩ B):** Estrarre una carta che sia di Fiori E una Regina. C'è solo una Regina di Fiori (Q♣). P(A ∩ B) = 1/52.

Calcolo della Probabilità

Poiché gli eventi "estrarre Fiori" e "estrarre una Regina" non sono disgiunti (possono verificarsi contemporaneamente, ovvero estraendo la Regina di Fiori), usiamo la formula generale dell'unione:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(\text{Fiori o Regina}) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52}$$

$$= \frac{13+4-1}{52} = \frac{16}{52}$$

$$= \frac{4}{13} \approx 0.3077$$

Risultato: La probabilità di estrarre una carta di Fiori o una Regina è di 16/52, che semplificato è 4/13, ovvero circa il 30.77%.

2. Rischio di Ritardo in un Progetto

Applichiamo la formula dell'unione per calcolare la probabilità che almeno uno dei due eventi di ritardo si verifichi.

Dati del Problema

- **Evento A:** Lancio del prodotto subisce ritardo. P(A) = 0.15 (15%)
- **Evento B:** Campagna marketing subisce ritardo. P(B) = 0.10 (10%)
- **Intersezione (A ∩ B):** Entrambi gli eventi subiscono ritardo. P(A ∩ B) = 0.05 (5%)

Vogliamo trovare la probabilità che almeno uno dei due eventi subisca un ritardo, ovvero P(A ∪ B).

Calcolo della Probabilità

Poiché i due eventi non sono disgiunti (possono verificarsi contemporaneamente), utilizziamo la formula generale dell'unione:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(\mathrm{A\ o\ B}) = 0.15 + 0.10 - 0.05$$

$$= 0.25 - 0.05 = 0.20$$

Risultato: La probabilità che almeno uno dei due eventi (lancio del prodotto o campagna marketing) subisca un ritardo è del **20%**.

Prossima Lezione: Probabilità Condizionale e Teorema di Bayes

01

Probabilità Condizionale P(A|B)

Come cambia la probabilità di A quando sappiamo che B si è verificato? Definizione, significato, interpretazione.

02

Regola del Prodotto

Come calcolare $P(A \cap B)$ usando probabilità condizionate: $P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$

03

Teorema di Bayes

Lo strumento fondamentale per "invertire" probabilità condizionate e aggiornare credenze con nuovi dati. Base dell'inferenza bayesiana.

04

Applicazioni Pratiche

Case study Simulazioni in Python.

📚 **Preparazione:** Rivedete gli assiomi e le proprietà derivate di questa lezione. Nella prossima sessione partiremo con simulazioni Python per applicare i concetti teorici a dati reali di business.