Studente 1

- Teorema di esistenza della distruzione invariante con un esempio di processo di quando non esiste

  - il **Teorema di esistenza**, stabilisce che se l'insieme degli stati E di un processo di Markov è finito, allora, esiste sempre almeno una distribuzione invariante.
    Tale distribuzione rappresenta un regime d'equilibrio in cui, le densità marginali del processo, rimangono costanti nel tempo.
    ESEMPIO DI NON ESISTENZA: L'ipotesi che l'insieme degli stati sia finito è strettamente necessaria.
    Se il processo si muove su un insieme infinito di stati, come i numeri naturali , la distribuzione invariante potrebbe non esistere per diverse ragioni:
    Processo di Deriva: il processo si muove continuamente verso lo stato successivo, questo di base non costituisce un problema, il problema nasce dal fatto che gli stati sono Infiniti, e dunque, non è possibile definire una somma della densità discreta che faccia 1, quindi non esiste la distribuzione invariante.
    Processo di Posson: ogni stato è transitorio e la probabilità viaggia progressivamente verso l'infinito, se si volgo i calcoli legati a $\pi L=0$ ci si accorge che rende impossibile soddisfare il vincolo che la somma delle probabilità sia unitaria.

- Condizioni per la stazionarietà di un processo ARIMA

  - un processo AR I MA e definito da:
    $$p(L)(1-L)^d X_t = q(L)W_t$$
    dove L e il ritardo, e questa formula per la stazionarietà ha bisogno di:
    - Assenza di integrazione (d=0), se questo non accade, il processo è ottenuto tramite integrazione la quale introduce instabilità (Si vede bene nella Passeggiata Aleatoria)
    - Avere però (d=0) non è sufficiente, devo avere che i coefficienti della parte AR, tali che le radici di p(z)=0 siano tutte in modulo maggiori di 1.

Studente 2

- É sempre vero che una matrice L è diagonalizzabile

  - Nei processi di Markov a Salti, la matrice L è la matrice delle intensità di salto.
    Mentre è sempre vero che una matrice di covarianza simmetrica sia diagonalizzabile, non è detto invece che la matrice L lo sia sempre, in quanto non possiede necessariamente la proprietà di simmetria.

- Dimostrare che la somma di gaussiane è gaussiana

  - Per dimostrare che la somma di variabili aleatorie gaussiane indipendenti è ancora una variabile gaussiana, il metodo più efficace è quello di utilizzare la **Funzione Generatrice dei Momenti (MGF)** [Funzione che "riassume" una variabile aleatoria trasformandola in un oggetto unico da cui, in linea di principio, puoi recuperare tutti i suoi momenti (media, varianza, e così via)].

- Proprietà della MGF su somme di variabili indipendenti:
Se abbiamo due variabili aleatorie indipendenti la MGF della loro somma è pari al prodotto delle singole MGF
- Forma della MGF per una variabile gaussiana:
una variabile aleatoria gaussiana ha una funzione generatrice dei momenti specifica:

- $MGF_x(t) = \exp\left(mt+\frac{\sigma^2}{2}t^2\right)$

  - Calcolo della MGF della somma:
  se prendiamo due funzioni generatrici, usando la regola del prodotto e sfruttando le proprietà dell'esponenziale sommando così gli esponenti abbiamo che:

$$MGF_{x+y}(t) = \exp\left((m_x + m_y)t + \frac{(\sigma_x^2+\sigma_y^2)}{2}t^2\right)$$

  Il risultato ottenuto ha esattamente la stessa struttura di una MGF per una variabile gaussiana.

Studente 3

- Enuncia la legge dei grandi numeri e spiega l'applicazione agli istogrammi
  - **Teorema della Legge dei Grandi Numeri:**
    Data una successione di variabili aleatorie non correlate (o indipendenti) che hanno lo stesso valore (m) atteso e la stessa varianza (σ^2) la loro media campionaria converge al valore atteso teorico al tendere del numero di osservazioni verso infinito.
    Per gli Istogrammi:
    La LGN assicura che aumentando il numero di campioni analizzati, e registrandoli sul grafico, l'istogramma diventa sempre di più una rappresentazione formale della densità teorica del processo, andando a descrivere la tipica forma a campana.

- Parlami del concetto di convergenza ad una costante
  - Il concetto di convergenza ad una costante, riguarda il comportamento di una successione di variabili aleatorie che, al crescere dei campioni analizzati perdono progressivamente la loro componente di incertezza tendendo a un valore deterministico preciso.
    Ne è un esempio proprio la legge dei grandi numeri sopra citata, questo fenomeno è ciò che permette ai sondaggi/alle misurazioni ripetute di fornire risultati affidabili.
    Aumentando il numero di osservazioni, l'errore (Varianza) si riduce fin quasi a svanire.

Studente 4

- Bilancio di flusso per Catene di Markov e processi di Markov a salti

- Il bilancio di flusso, è lo strumento fondamentale per capire se un processo ha raggiunto la Stazionarietà, infatti che un processo di Markov (sia esso una catena o a salti) è stazionario in senso stretto se e solo se la sua densità marginale al tempo iniziale è una distribuzione invariante che soddisfa i criteri di bilancio.
  Il concetto di Bilancio di Flusso: non è altro che una riformulazione delle equazioni matematiche che definiscono le Distribuzioni Invarianti, per i processi di Markov a stati discreti (Distribuzione è detta Invariante se: quando raggiunta, le densità marginali rimangono costanti nel tempo, indicando il raggiungimento di equilibrio statico).

  Per Catene di Markov:
  a tempi discreti, caratterizzati da una Matrice di Transizione (Q), un vettore riga (π) rappresenta una Distribuzione invariante, se soddisfa l'equazione π=πQ.
  $$\sum_{y \neq x} \pi_x Q_{x \to y} = \sum_{y \neq x} \pi_y Q_{y \to x}$$
  Sinistra: Flusso Uscente, Destra: Flusso Entrante.
  Il bilancio di flusso assicura quindi che la "massa" di probabilità in ogni stato rimanga invariata tra un istante temporale e il successivo.
  Per Markov a Salti: operano su tempi continui, la distribuzione invariante (π) è definita tramite la Matrice delle Intensità di Salto (L) e deve soddisfare l'equazione 0=πL.
  $$\sum_{y \neq x} \pi_x L_{xy} = \sum_{y \neq x} \pi_y L_{yx}$$
  L'espressione indica che la velocità con cui la probabilità "esce" da uno stato deve essere perfettamente bilanciata dalla velocità con cui la probabilità vi "entra" dagli altri stati.
  [In termini dinamici, questo equilibrio garantisce che la derivata temporale della densità marginale sia nulla, rispettando la cosiddetta **Equazione di Kolmogorov**]

- Parlami della teoria delle code nel caso M/M/1

  - **M/M/1:** Uno dei modelli base per l'analisi delle linee d'attesa, è modellizzato come un processo di Markov a salti su tempi e stati discreti.
    - La prima M: indica un processo di arrivo Markoviano (di Poisson) con intensità λ (chiamata tasso di ingresso).
    - La seconda M: indica il tempo di servizio per ciascun cliente, segue una legge esponenziale di parametro μ (detto tasso di uscita).
    - L'1: indica la presenza di un solo servente.
    Il sistema funziona spostandosi tra vari stati secondo le seguenti regole di transizione:
    -Arrivo di un Cliente: passo dallo stato n allo stato n+1 ($L_{n \to n+1} = \lambda$)
    -Completamento del servizio: passo dallo stato n allo stato n-1 ($L_{n \to n-1} = \mu$)
    -Permanenza nello stato: L'intensità di uscita dallo stato n è data da $L_{n->n} = -\lambda$ [se la coda è vuota] oppure $= -(\lambda + \mu)$ [se presenti clienti]

    Posso avere un regime di equilibrio la dove il tasso di arrivo è strettamente inferiore alla capacità di servizio

L'indicatore principale in regime stazionario è: Il numero medio di clienti nel Sistema (N), si calcola come: $E[N] = \frac{\lambda}{\mu-\lambda}$

Studente 5

- Stima parametri delle code M/M/0
  - Detto anche Processo di Poisson di Intensità $\lambda > 0$, è un processo in cui non ci sono serventi.
    - Dinamica del Modello: le uniche transazioni possibili avvengono dallo stato n allo stato n+1 con intensità : $(L_{n\to n+1} = \lambda)$.
    Dato che abbiamo solo ingressi e nessuna uscita, ogni stato è Transitorio e non esiste una distribuzione invariante (si va solo verso infinito).
    Tramite un conto del genere, calcolo solo l'intensità di arrivo, utile a capire in un dato periodo di tempo T, quanti avvenimenti/campioni, riesco ad analizzare, per fare ciò uso la stima di Massima Verosimiglianza (MLE) la quale dice che:
    Per stimare l'intensità di arrivo (λ) seguiamo una traiettoria, partendo da un istante iniziale, registra n arrivi successivi in un tempo totale di osservazione, seguendo questa formula: $L(\lambda;\gamma) = \lambda^n \exp(\ -\lambda T)$.
    Se massimizzo il logaritmo della verosimiglianza rispetto a (λ) si ottiene lo stimatore: $\lambda_{MLE} = \frac{n}{T}$.
    - Approccio Bayesiano (MAP) : Disponendo di informazioni a priori sul tasso di arrivo, possiamo utilizzare un approccio bayesiano, influenzando con l'informazione conosciuta a priori i calcoli svolti prima, correggendo così il calcolo (questo è molto utile solitamente se il tempo di osservazione è molto breve).

- Parlami della correlazione
  - Uno degli indicatori fondamentali della statistica e del calcolo delle probabilità, è utilizzato per quantificare la relazione e la dipendenza tra due o più variabili aleatorie.

    [Immagina la correlazione come il movimento di due ballerini in una sala: se si muovono sempre all'unisono nella stessa direzione, la loro correlazione è 1 (positiva massima). Se uno si muove a sinistra ogni volta che l'altro va a destra, la correlazione è -1 (negativa massima). Se invece ognuno balla per conto suo senza curarsi dell'altro, i loro movimenti sono non correlati (0). Tuttavia, potrebbero non scontrarsi mai perché seguono uno schema complesso non lineare: in quel caso sarebbero non correlati, ma non del tutto indipendenti.]

Studente 6

- Teorema Ergodico con relativa dimostrazione
  - Rappresenta un raffinamento e un'estensione della Legge dei Grandi Numeri, applicata a contesti in cui le variabili aleatorie non sono necessariamente indipendenti (come nel caso dei processi stocastici) .
    L'idea fondamentale è quella di poter identificare il valore atteso calcolato sulla densità di probabilità degli stati.

- Teorema: Preso $(X_t)_{t=0}^{\infty}$ un processo stazionario, e sia (g : E -> R) un osservabile, se la correlazione tra istante iniziale e il futuro svanisce nel tempo, ovvero se:

$$\lim_{t \to \infty} Cov\big(g(X_0), g(X_t)\big) = 0$$

allora la media temporale $g(X)_T$ converge in media quadratica al valor medio teorico $E[g(X_0)] per\ T \to \infty$.

- Dimostrazione: La dimostrazione segue la logica analoga a quella della LGN, ma deve gestire la presenza di correlazione tra le variabili.
Poniamo per semplicità Y = $g(X_t)$, per ipotesi di stazionarietà il valor medio di $E[Y_t]$ è costante e pari a $E[Y_0]$ per ogni t.
Di conseguenza applicando la linearità al calore atteso si ottiene che il calor medio della media temporale è anch'esso costante.
La varianza della media temporale è data dalla somma di tutti i termini della matrice di covarianza tra i tempi s e t: $Var(\bar{Y}_T) = \frac{1}{T^2}\sum_{s,t=1}^{T} Cov(Y_s, Y_t)$.
Sfruttando la stazionarietà sappiamo che la Covarianza dipende solo dalla differenza temporale |t-s| quindi possiamo scriverla come Cov(0,|t-s|)
Poiché per ogni ritardo k esistono al più 2T coppie (s,t) tali che |t-s|=k, possiamo stimare la varianza come: $Var(\bar{Y}_T) \leq \frac{2}{T}\sum_{k=0}^{T} |C(0,k)|$.

Conclusione tramite il Lemma di Cesaro: per ipotesi sappiamo che la covarianza C(0,k) tende a zero per $k \to \infty$.
Un risultato classico dell'analisi garantisce che se una successione tende a zero anche la sua media aritmetica tende a zero.
Ciò implica che la varianza è infinitesima per T che tende a Infinito.
Avendo dimostrato che il valore atteso è costante e la varianza svanisce, per i criteri di convergenza delle variabili aleatorie concludiamo che la media temporale converge al valore atteso in media quadratica e in probabilità.

[**L'analogia** del "monitoraggio meteo" : Immagina di voler conoscere la temperatura media di una città. Il valor medio spaziale sarebbe la media delle temperature misurate in ogni singolo punto della città nello stesso istante. Il Teorema Ergodico ti dice che, invece di piazzare migliaia di sensori ovunque, puoi piazzarne uno solo in un punto e misurare la temperatura per un tempo molto lungo (media temporale). Se il "sistema meteo" è ergodico (ovvero se le condizioni di oggi non influenzano per sempre quelle di tra un anno), la media delle tue misurazioni nel tempo rifletterà perfettamente la media spaziale della città.]

Studente 7

- Matrice di covarianza: spiega le sue caratteristiche
  - La matrice di covarianza, rappresenta lo strumento fondamentale per descrivere la dispersione congiunta e le relazioni lineari tra le diverse componenti di un vettore aleatorio $X = (X_1, \dots, X_d) \in R^d$.

    Caratteristiche: La matrice di covarianza è una matrice quadrata di dimensione dxd, ogni sua entrata alla riga i e colonna j è definita come la covarianza tra le variabili $X_i\ e\ X_j$.

Sulla diagonale principale, gli elementi rappresentano le varianze delle singole componenti dato che la $Cov(X_i, X_i) = Var(X_i)$, gli elementi eventualmente presenti furi dalla diagonale indicano come le coppie di variabili “si muovano insieme”.

La matrice possiede caratteristiche matematiche precise che riflettono la sua natura statistica:
- Simmetria [$Cov(X_j, X_i) = Cov(X_i, X_j)$]
- Semi-Definita Positiva [gli autovalori della matrice sono tutti non negativi, confermando che la varianza di una qualsiasi combinazione lineare delle componenti non può mai essere negativa]

Grazie al teorema spettrale, ogni matrice di covarianza può essere diagonalizzata tramite una matrice ortonormale.
Dalla matrice di covarianza si ricava la matrice delle correlazioni, normalizzando ogni entrata per le rispettive deviazioni standard, il coefficiente di correlazione di Pearson è sempre compreso tra -1 e 1.
Ricordare che se due variabili sono indipendenti, la loro covarianza è nulla, tuttavia il viceversa non è sempre vero, tranne caso speciale variabili gaussiane vettoriali, la non correlazione implica necessariamente l'indipendenza.

- Coefficiente di correlazione
  - Indice di Correlazione di Pearson: è un indicatore statistico utilizzato per quantificare la forza e la direzione della relazione lineare tra due variabili aleatorie reali X e Y.
    Rappresenta una versione normalizzata della covarianza, permettendo di confrontare il legame tra fenomeni espressi in unità di misura differenti.
    Il coefficiente di correlazione, indicato solitamente con: ρ_{xy}, è definito come il rapporto tra la covarianza delle due variabili e il prodotto delle loro deviazioni standard.

$$\rho_{xy} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

Proprietà:
Intervallo: Assume sempre valori compresi tra -1 e 1
Correlazione Perfetta: ($|\rho| = 1$) Se esiste una dipendenza lineare perfetta tra le due variabili
Assenza di Correlazione: ($\rho = 0$) Le variabili si dicono non correlate, indicando l'assenza di un legame lineare tra loro.
Correlazione Positiva: ($\rho > 0$) Le variabili tendono a variare nella stessa direzione.
Correlazione Negativa($\rho < 0$) Al crescere di una variabile l'altra tende a diminuire.
Distinzione tra indipendenza probabilistica e non correlazione:
- L'indipendenza implica sempre la non correlazione: Se due variabili sono indipendenti, la loro covarianza è necessariamente nulla.
- La non correlazione NON implica necessariamente l'indipendenza: Due variabili possono avere correlazione a zero ma essere comunque dipendenti attraverso relazioni non lineari.
- Caso speciale delle Gaussiane: Se un vettore aleatorio ha densità

gaussiana, la non correlazione tra le sue componenti è equivalente alla loro indipendenza.

- Regressione lineare semplice con modello lineare
  - Metodo statistico e probabilistico utilizzato per modellizzare la relazione tra due variabili: un **predittore** (o variabile esplicativa : X) e una **risposta** (o esito : Y).
    L'obiettivo è determinare una funzione g(X) in modo che Y sia molto vicina a g(X) in base alle osservazioni effettuate.

    Per fare ciò in questo modello assumiamo che la funzione g appartenga a una famiglia parametrizzata da un vettore u=(a,b) definita come: $g(x; a, b) = ax + b$.
    Con "a" che rappresenta il coefficiente angolare e "b" l'intercetta.

    Per stimare i parametri a partire da "n" osservazioni indipendenti, si usa il metodo:
    Metodo dei Minimi Quadrati (OLS), serve per minimizzare la somma dei minimi quadrati, dunque la discrepanza tra i valori osservati e quelli previsi dal modello.

    $$(a_{OLS}, b_{OLS}) \in \arg\min \sum_{i=1}^{n} \left(y_i - (ax_i + b)\right)^2$$

    Il Metodo dei Minimi Quadrati: è una tecnica geometrica utile anche nel calcolo delle probabilità, ipotizziamo: $Y = g(X; U) + W$.
    Con il residuo (W), variabile aleatoria con densità Gaussiana [N(0,v)], la stima dell'OLS coincide con la stima di Massima Verosimiglianza (MLE).
    Minimizzare i quadrati dei residui, equivale a massimizzare la probabilità di aver osservato quei dati sotto l'ipotesi di rumore gaussiano.
    Per misurare l'aderenza del modello ai dati si utilizza il coefficiente di determinazione.

Studente 8

- Parlami del modello M/M/infinito
  - Rappresenta: una situazione limite della teoria delle code in cui si dispone di un numero arbitrariamente grande, idealmente infinito, di serventi.
    In questo sistema, i clienti arrivano secondo un **Processo di Poisson** di intensità **λ**>0 e il tempo di servizio per ciascuno di essi è una variabile aleatoria esponenziale di parametro μ.

    Il sistema viene modellizzato come un processo di **Markov a salti** dove lo stato "n" indica il numero di clienti presenti nel sistema.

    Meccanismo di Ingresso/Uscita:
    - Ingresso: [n -> n+1] Avviene con intensità costante $L_{n \to n+1} = \lambda$ corrispondente all'arrivo di un nuovo cliente.
    - Uscita: [n -> n-1] Avviene con intensità $L_{n \to n-1} = n\mu$, l'intensità d'uscita deriva dal fatto che, essendoci infiniti serventi, tutti gli clienti presenti

vengono serviti contemporaneamente, ciò implica che il tempo di attesa affinché il primo tra loro termini il servizio è il minimo di "n" variabili esponenziali indipendenti, che è ancora una variabile esponenziale con parametro pari alla somma dei singoli parametri, ovvero "nμ" .

Distribuzione di Probabilità: la distribuzione che descrive il sistema a regime (invariante) è una densità di Poisson di parametro **λ**/μ.
Numero Medio : clienti nel sistema in regime stazionario è dato dal valore atteso della distribuzione di Poisson E[N] = **λ**/μ (il valore cresce linearmente al crescere di **λ** e non ha asintoti critici).

Stima dei Parametri:
Se osserviamo una traiettoria del sistema per un tempo totale , possiamo stimare i parametri attraverso il metodo della Massima Verosimiglianza (MLE).

[L'analogia del self-service: Immagina un enorme buffet a self-service con infiniti vassoi e spazio illimitato, ogni persona che entra (intensità) inizia a mangiare immediatamente senza fare fila. Se ogni persona impiega mediamente un certo tempo per mangiare (intensità), il numero di persone che vedrai contemporaneamente nel locale dipenderà solo dal rapporto tra quanto velocemente entrano e quanto velocemente finiscono di mangiare. Anche se arrivasse una comitiva enorme, il sistema non si bloccherebbe mai, perché c'è sempre un "posto" pronto per tutti.]

- Coefficiente di correlazione
    - Risposta già data sopra

- Parlami della covarianza
    - La covarianza tra due variabili aleatorie reali e è definita matematicamente come il valore atteso del prodotto degli scarti di ciascuna variabile dalla propria media:
    $$Cov(X, Y) = E[(X - E[X])(Y - E[Y])$$

- Qual era più significativo nel caso in cui le stime fossero circa uguali
    - **Se le covarianze stimate risultano "circa uguali", la misura più significativa è il coefficiente di correlazione**.
    Motivazione:
    - La covarianza dipende dalle unità di misura (se cambi scala a una variabile, la covarianza cambia), quindi due covarianze simili non implicano automaticamente "stessa forza di legame".
    - La correlazione, invece, è normalizzata (scala fissa tra −1 e 1) e quindi è confrontabile tra coppie di variabili diverse.

Studente 9

- Densità di probabilità

- É lo strumento matematico che utilizziamo per descrivere come la probabilità si "distribuisce" o si "spalma" tra i possibili valori che una variabile aleatoria può assumere.
  Esistono quindi due tipi di densità:
  - Discreta: Se la variabile aleatoria assume valori "isolati" (come il lancio di un dado o il numero di successi in un esperimento), la densità discreta è semplicemente la funzione che associa a ogni possibile valore "x" la sua probabilità esatta P(X=x).
  - Continua: Quando i valori possibili si trovano su un intervallo continuo (come l'altezza di una persona o la temperatura), la situazione cambia, la probabilità di un singolo punto esatto è sempre zero.

  La densità è fondamentale perché: "riassume" l'intera legge di una variabile. Senza di essa, dovremmo calcolare la probabilità per ogni possibile sottoinsieme di valori, operazione che sarebbe quasi impossibile sul piano pratico. Inoltre, attraverso la densità possiamo calcolare indicatori chiave come il valor medio e la varianza.

- Densità di X^2
  - Per trovare la densità di X^2, non basta semplicemente elevare al quadrato la funzione di densità di X . Il calcolo richiede di individuare quali valori di X vengono "mandati" in un determinato valore "z" dalla funzione $g(x) = x^2$.
    La formula del cambio di variabile:
    Il teorema del cambio di variabile afferma che, se "X" ha una densità continua $p(X = x)$, la densità della trasformata "$g(X)$" , si ottiene pesando la densità originale con l'inverso della derivata della funzione di trasformazione.
    Per $g(x) = x^2$ la derivata $g'(x) = 2x$, applicando la formula generalizzata, la densità di X^2 in un punto "$z > 0$" è data dalla somma dei contributi delle due immagini inverse:
    $$\boldsymbol{p(X^2 = z) = \frac{p(X = -\sqrt{z} + p(Z = \sqrt{z})}{2\sqrt{z}}}$$
    [**L'analogia della nebbia**: Immagina la densità di come una nuvola di nebbia distribuita lungo una strada che va da -1 a 1. Quando applichi la funzione , è come se piegassi la strada a metà sopra lo zero. La nebbia che prima era nella parte negativa ora si sovrappone a quella nella parte positiva. La densità di misura quanto è diventata fitta la nebbia in ogni punto della strada ripiegata, poiché vicino allo zero la strada è "molto curva" (la derivata è piccola), la nebbia si accumula tantissimo, diventando molto densa]

- Processi di Markov a salti
  - Rappresentano l'estensione delle catene di Markov al caso in cui il tempo è continuo**,** ma l'insieme degli stati rimane discreto.
    Un processo a salti è "markoviano" se soddisfa la proprietà di Markov: il futuro e il passato del processo sono indipendenti tra loro, a patto di conoscere esattamente lo stato presente.
    A differenza delle catene a tempi discreti (che usano una matrice di transizione Q), i processi a salti si descrivono tramite la **matrice delle intensità di salto** (L). In tale matrice:

- Gli elementi **fuori diagonale** indicano la velocità o l'intensità con cui il processo "salta" da uno stato “x” a uno stato “y”.
- Gli elementi sulla **diagonale** sono negativi e pari alla somma degli altri elementi della riga cambiata di segno, garantendo che ogni riga della matrice sommi a zero.

- Distribuzioni invarianti
  - Una distribuzione invarianti (o stazionaria) per una catena/processo di Markov è una distribuzione di probabilità sugli stati che rimane uguale dopo l’evoluzione del processo (davvero utile se il processo è Egrodico).
    É una “fotografia statistica” tale che, se il sistema parte già con quella fotografia, dopo un passo (o dopo un po’ di tempo) la fotografia non cambia, descrive spesso il comportamento di lungo periodo: “quanto tempo, in media, il processo passa in ciascuno stato”.

Studente 10

- Stima di Massima Verosimiglianza e Bayesiana della Media di “n”
  - La stima della media di una popolazione a partire da osservazioni indipendenti, è uno dei problemi centrali della statistica.
    Questo problema può essere affrontato attraverso due approcci principali: la **Massima Verosimiglianza** (MLE), che si basa esclusivamente sui dati osservati, e l'**Approccio Bayesiano**, che integra i dati con l'informazione già nota a priori.
    - Il metodo della Massima Verosimiglianza cerca di determinare il valore del parametro di media “m” che rende "più probabili" le osservazioni effettuate, si assume che le osservazioni provengano da variabili gaussiane indipendenti con la stessa media e varianza.
    Poiché le osservazioni sono indipendenti, la verosimiglianza congiunta è il prodotto delle singole densità gaussiane.

- Osservazioni Gaussiane indipendenti, fissata la Varianza
  - L'approccio Bayesiano considera la media non come un valore fisso ignoto, ma come una variabile aleatoria “M” . Questo metodo permette di aggiornare una conoscenza iniziale (a priori) con i nuovi dati osservati.
    [L'analogia del testimone e dell'esperto Immagina di dover stimare l'altezza di una torre. La stima MLE è come un testimone che guarda solo i dati del GPS: prende “n” misurazioni e ne fa la media.
    La stima Bayesiana è come un architetto (l'informazione a priori) che sa già quanto dovrebbe essere alta la torre in base ai progetti originali.
    Se hai solo una misurazione GPS poco precisa, ti fiderai più dell'architetto, ma se hai 10.000 misurazioni GPS precise (molto grande), l'evidenza dei dati schiaccerà l'opinione dell'architetto e ti fiderai quasi solo dei sensori.]

Studente 11

- Teorema Ergodico

- Sia $(X_t)_{t\geq 0}$ un processo stazionario sull'insieme degli stati E e sia $g: E \to R$ una Osservabile (quantità misurabile del processo), se la funzione di autocovarianza dell'osservabile tende a zero per tempi che tendono a infinito:
$$\lim_{t\to\infty} Cov\big(g(X_0), g(X_t)\big) = 0$$
allora la media temporale dell'osservabile sui primi T istanti:
$$g(X)_T = \frac{1}{T}\sum_{t=1}^{T} g(X_t)$$
converge in media quadratica al valor medio teorico dell'osservabile stessa:
$$\lim_{T\to\infty} g(X)_T = E[g(X_0)]$$

  Dimostrazione:
  La dimostrazione si basa sulla verifica delle condizioni di convergenza in media quadratica verso una costante, analizzando il valor medio e la varianza della media temporale.
  (Per i passaggi dettagliati si rimanda agli appunti del docente – Cap 8.3)

- Convergenza in probabilità
  - La convergenza in probabilità è una delle nozioni fondamentali per descrivere come una successione di variabili aleatorie si avvicini a un valore limite (o a un'altra variabile aleatoria) al crescere verso infinito.
  Definizione: Si dice che una successione X_n converte in probabilità verso X_∞ se per ogni valore $\varepsilon > 0$ la probabilità che la distanza tra X_n e il limite sia superiore a $\varepsilon \to 0$ quando $n \to \infty$.

- Convergenza in media quadratica
  - Si dice che una successione X_n converte in media quadratica verso X_∞ se il valore atteso del quadrato della loro differenza tende a zero quando n tende all'infinito:
$$\lim_{n\to\infty} E[|X_n - X_\infty|^2] = 0$$
Questa nozione implica che la "distanza media" (nel senso dell'energia dello scarto) tra la variabile e il suo limite diventi trascurabile.

- Teorema del limite centrale
  - È un raffinamento della legge dei grandi numeri, in cui si rende più precisa la convergenza delle medie empiriche, mostrando che le oscillazioni sono approssimabili tramite variabili gaussiane, qualsiasi fosse la distribuzione delle variabili di partenza.
  (Per i Teorema e Dimostrazione si rimanda agli appunti del docente – Cap 8.4.2)

Studente 12

- La CDF di una gaussiana
  - La Funzione di Ripartizione (CDF) di una variabile aleatoria gaussiana rappresenta la probabilità che la variabile assuma un valore minore o uguale a un certo valore "x".

Poiché la gaussiana è una variabile aleatoria continua, la sua CDF viene calcolata matematicamente come l'integrale della funzione di densità (la curva a campana) partendo da meno infinito fino al punto “x” desiderato.

- Comando R per calcolare la CDF della gaussiana
  - pnorm(x, mean = μ, sd = ϕ)

- Unicità della distribuzione invariante
  - Il fattore principale che garantisce l'unicità della distribuzione invariante è l'**irriducibilità** della matrice di transizione (o delle intensità di salto).
    Un sistema si dice **irriducibile** quando **ogni stato è raggiungibile da ogni altro stato** attraverso un cammino di probabilità positiva.
    Se questa condizione è soddisfatta su un insieme di stati **finito**, il teorema garantisce che **esiste una e una sola distribuzione invariante.**

- Fammi un esempio di catena non irriducibile
  - Prendiamo una catena formata da 3 nodi:
    {1,2,3,}
    - da 1: si va in 2
    - da 2: si va in 1 e in 3
    - da 3: si va solo in 3
    Lo stato 3 è una classe chiusa (assorbente).
    Gli stati 1 e 2 formano una classe che non è chiusa, perché da lì puoi uscire verso 3.
    Siccome non tutti gli stati sono raggiungibili l’uno dall’altro, la catena non è irriducibile.

Studente 13

- Date delle estrazioni casuali, come può valutare la densità della sorgente
  - Per valutare la densità della sorgente a partire da un insieme di estrazioni casuali, le fonti indicano due percorsi metodologici principali:

    - La Massima Verosimiglianza (MLE) : Questo metodo determina il valore del parametro che rende "più probabili" le estrazioni effettivamente osservate, si calcola come il rapporto tra eventi avvenuti e tempo di osservazione.

    - La Statistica Bayesiana : Non fornisce un singolo valore, ma una densità di probabilità aggiornata sul parametro, permettendo di integrare una conoscenza iniziale (a priori) con i dati delle estrazioni/osservazioni
    [Ad Esaemio: Ogni nuova estrazione modifica la fiducia del robot, la formula è data come: $p(parametro|dati) \propto p(parametro) \times L(parametro: dati)$].

    Reminder:
    { Proporzionale (∝) : un valore (A) è proporzionale a un altro valore (B) che viene moltiplicato per la costante di proporzionalità (k) [A = kB]. }

- Stima di gaussiana con MLE
    - **La MLE per una variabile aleatoria gaussiana** ha l'obiettivo di determinare i parametri del modello (media “m” e varianza “v”) che rendono i dati effettivamente osservati i "più probabili" secondo il modello matematico.
      Si distinguono 3 tipologie di stima:
      - **Stima da una singola osservazione:** Con una singola osservazione la MLE della media coincide con l’osservazione stessa, importante però ricordare che, se anche la varianza è ignota, non esiste una MLE finita, la verosimiglianza, cresce, facendo tendere la varianza a zero.
      - **Stima da “n” osservazioni indipendenti** (Campione): Con “n” osservazioni indipendenti , la verosimiglianza complessiva è il prodotto delle singole densità gaussiane, la minimizzazione del logaritmo della verosimiglianza porta a definire gli stimatori classici della statistica (Media Campionaria e Varianza Campionaria) .
      - **Caso Vettoriale**: Per variabili gaussiane vettoriali con parametri (m, Σ), dove Σ è la matrice di covarianza, la stima MLE estende i risultati del caso “Campione”:
      > Vettore dei valori medi: vettore delle medie campionarie di ciascuna componente.
      > Matrice di covarianza campionaria: data dalla media dei prodotti esterni degli scarti.

- Parlare del processo AR I MA : con AR(1)
    - Il processo **ARIMA** (*AutoRegressive Integrated Moving Average*) rappresenta una famiglia generale di modelli stocastici a stati continui e tempi discreti, ampiamente utilizzati per l'analisi e la previsione delle serie storiche. La sua struttura è composta da tre componenti fondamentali: la parte **AutoRegressiva (AR)**, il procedimento di **Integrazione (I)** e la componente a **Media Mobile (MA)**.
      Un processo è definito **AR(1)** (AutoRegressivo di ordine 1) quando il valore della variabile al tempo (t) è stimato a partire dal valore assunto dalla variabile stessa all'istante immediatamente precedente (t-1).
      L’equazione fondamentale è:
      $$X_t = \alpha_1 X_{t-1} + W_t$$
      con W_t: rumore bianco gaussiano (Rumore con valori gaussiani e “senza memoria”, non c’è correlazione tra tempi diversi).

- Definizione di processo stazionario
    - Un processo stocastico si definisce stazionario (spesso indicato come stazionario in senso stretto) se la sua legge congiunta rimane invariata a seguito di una traslazione dei tempi.
      [**L'analogia del nastro trasportatore:** Possiamo immaginare un processo stazionario come un nastro trasportatore in movimento costante su cui sono appoggiati degli oggetti in modo casuale, ma con una densità uniforme. Se scattiamo una fotografia a una sezione del nastro in questo momento e un'altra fotografia alla stessa sezione tra dieci minuti, le posizioni esatte degli oggetti saranno diverse (dinamica del processo), ma le proprietà statistiche (come il numero medio di oggetti per metro o la distanza media tra di essi) saranno

identiche, rendendo i due scatti statisticamente indistinguibili.]

- Dati i valori [0,1,2,3,4,5], supponendo che questi provengano da un processo AR(1), come stimare i parametri
  - Si utilizza solitamente il metodo della **Massima Verosimiglianza (MLE)**, che in questo contesto si riduce alla minimizzazione della somma dei quadrati dei residui.

$$\widehat{\alpha_{MLe}} = \frac{\sum_{t=1}^{n} x_t x_{t-1}}{\sum_{t=1}^{n} {x^2}_{t-1}}$$

    [Nel es: Numeratore: (1x0)+(2x1)+(3x2)+(4x3)+(5x4)=40, Denominatore: $0^2+1^2+2^2+3^2+4^2=30$, $\widehat{\alpha_{MLe}} = \frac{40}{30} + 1.33$].

Studente 14

- Dato un processo di Markov, invertiamo i tempi, faccio un video e poi lo riproduco al contrario, rimane un processo di Markov? È omogeneo?
  - **Sì, rimane un processo di Markov.**
    La ragione risiede nella definizione stessa della proprietà di Markov: essa afferma che, dato il "presente", il "futuro" e il "passato" sono condizionatamente indipendenti, in questa definizione, passato e futuro hanno un ruolo simmetrico. Poiché l'indipendenza probabilistica è una relazione simmetrica, se il futuro non dipende dal passato (dato il presente), allora, guardando all'indietro, il passato non dipende dal futuro (dato il presente), pertanto, la struttura logica del processo non si rompe invertendo la freccia del tempo.
    Non necessariamente; dipende dalla stazionarietà.
    Un processo si dice omogeneo se le sue probabilità di transizione dipendono solo dalla distanza temporale e non dall'istante esatto in cui ci si trova, quando invertiamo il tempo, le nuove probabilità di transizione "all'indietro" sono regolate dalla legge di Bayes.
    Se il processo originale è stazionario: sarà omogeneo.
    Se il processo originale non ha ancora raggiunto l'equilibrio: sarà non omogeneo.
    In pratica:
    Se guardi il video di un sistema in pieno equilibrio (stazionario), non potresti distinguere se il tempo scorre in avanti o all'indietro dalla sola struttura delle transizioni: il processo rimane di Markov e omogeneo. Se invece il video mostra un sistema che si sta ancora evolvendo verso l'equilibrio, il processo invertito resterà di Markov, ma "sentirà" il cambiamento delle probabilità nel tempo, perdendo l'omogeneità.

- Parlare di M/M/a
  - Rappresenta una famiglia di modelli della **teoria delle code** basati su processi di Markov a salti con stati discreti e tempi continui.
    M (Arrivals): Il processo di arrivo dei clienti è Markoviano, specificamente un processo di Poisson con tasso di ingresso λ.
    M (Service): Il tempo di servizio per ogni cliente segue una legge esponenziale

di parametro μ.
a (o “c” Servers): Indica il numero di serventi (o canali di servizio) disponibili contemporaneamente.

[Nelle domande precedenti sono già stati analizzati i casi con serventi, a 1 a Infinito e a 0, per maggiori delucidazioni guardare gli appunti del professore, inerenti alle code, Cap 6.6]

- Può non esistere una situazione invariante
  - Sì, una distribuzione invariante può non esistere, ma ciò accade solo sotto determinate condizioni legate alla natura dell'insieme degli stati e alla dinamica del processo.
    I casi principali in cui il sistema non riesce a raggiungere una situazione di equilibrio:
    -Il vincolo dello spazio degli stati finito: Per ogni sistema con un numero limitato di configurazioni possibili si ha sempre l’esistenza, tuttavia, quando lo spazio degli stati è infinito, l'esistenza non è più garantita.
    -Processi che "scappano" all'infinito: caso tipico di non esistenza si verifica quando il processo tende ad allontanarsi progressivamente senza mai stabilizzarsi.
    -Squilibrio tra Tassi (per la Teoria delle Code): Se il tasso di ingresso è maggiore o uguale a quello di uscita, il numero di persone in coda cresce illimitatamente.
    In questo scenario, la serie geometrica necessaria a definire la distribuzione non converge e non esiste una distribuzione invariante; il sistema è destinato a "esplodere" o accumulare ritardo all'infinito.

- Stima del parametro λ tramite osservazioni
  - La stima del parametro λ (solitamente riferito al tasso di intensità di un processo di **Poisson** o al parametro di una distribuzione **esponenziale**) può essere effettuata attraverso due metodologie principali: la **Massima Verosimiglianza (MLE)** e l'**approccio Bayesiano**.

    [L’argomento è giá stato trattato in domande precedenti per maggiori info guardare gli appunti del docente, Cap 5.3]

Studente 15

- La $MGF_x(t)$ è sempre ben definita
  - No, per $E[e^{tX}]$ con X che assume valori grandi, viaggia a infinito (per t>0) e a -infinito (per t<0).

- Data una variabile gaussiana standard X, trovare la densità di X^n
  - Per trovare la densitá variabile aleatoria dove $Y = X^n$, con X Gaussiana Standard $[N(0,1)]$, dobbiamo applicare il teorema del cambio variabile per variabili aleatorie continue.

$$p(X = x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

Il procedimento per calcolare la densità dipende dalla natura dell'esponente “n” (pari o dispari), poiché questo influenza l'invertibilità della funzione.
Se “n” Pari: Non è invertibile globalmente perché assegna lo stesso valore “z” sia a “x” che a “-x” . In questo scenario, dobbiamo usare la formula generalizzata che somma i contributi di tutti i punti della controimmagine.
Se “n” Dispari: funzione è strettamente crescente e invertibile su tutto l'asse reale.

- Enuncia e dimostra il teorema limite (TLC)
  - Descritto già in domande precedenti.
    (Appunti docente Cap 8 – Paragrafo 8.4)

Studente 16

- Dimmi tutto ciò che sai sulla Legge dei grandi numeri
  - Descritto già in domande precedenti.