

16 / 17 maggio 2025

COMPLESSO SANTA CATERINA
VIA CESARE BATTISTI, 241 - PADOVA

EVENTO DIVULGATIVO/SCIENTIFICO
SULLA STATISTICA

TRA E CON I VINCITORI DEL PREMIO "OLIVIERO LESSI"

LESSI UNA TESI ...



PROGRAMMA

Venerdì 16 maggio 2025

Ore 14:30-14:40 **Inaugurazione dei lavori**

Ore 14:40-16:20 **Quattro interventi di premiati**

- **Giulia Risca** “*Capire la ricerca nelle malattie rare: sfide e nuove soluzioni per i trial clinici*”

Una malattia è considerata rara, in Europa, quando colpisce meno di 5 persone su 10.000. Proprio perché i pazienti sono così pochi, fare ricerca su queste malattie è particolarmente difficile. Una delle principali difficoltà è organizzare studi clinici (“*trial*”) per testare l'efficacia di nuove cure, tuttavia, trovare un numero sufficiente di pazienti con la stessa malattia è spesso complicato.

Per superare questi ostacoli, i ricercatori stanno sviluppando nuovi metodi per progettare i trial. Tra questi, i *Master Protocols*: studi che permettono di testare più trattamenti o più malattie all'interno di un unico progetto, risparmiando tempo e risorse. Inoltre, si sta usando sempre più spesso un approccio con metodi bayesiani: che permette di aggiornare i risultati man mano che si raccolgono nuovi dati, rendendo gli studi più flessibili e adattabili.

L'obiettivo di questo intervento è offrire una panoramica sul processo di ricerca clinica, illustrandone i passaggi principali e mostrando l'utilità di metodi innovativi, come i *Master Protocols*, per migliorare la cura delle persone con malattie rare.

- **Lorenzo Schiavon** “*Statistica e salute pubblica: strumenti di traduzione.*”

La comunicazione tra statistici e decisori in ambito sanitario può essere molto complessa. La sfida diventa ancora più ardua se si parla di inferenza bayesiana, escludendo quindi dalla discussione la lingua franca dei p-value. Eppure, questo sforzo di comprensione reciproca è cruciale per tradurre i risultati scientifici in azioni concrete. In questo contesto, verranno discussi alcuni possibili strumenti.

Verrà presentato FARSI (*Fast, Accessible, Reliable, Secure, Informative*): un quadro metodologico ideato per trasformare elementi come intervalli di credibilità, odds ratio e medie a posteriori, in informazioni immediatamente fruibili per i decisori pubblici. Attraverso un caso studio sulla stima della prevalenza di malattie croniche in Italia, vedremo come la “farsi”-ficazione possa facilitare la comunicazione dei risultati di modelli complessi, rendendoli più accessibili nel contesto della sanità pubblica, con esempi sia nazionali che locali.

Poiché la comunicazione è sempre bidirezionale, discuteremo anche del processo inverso: come tradurre quantitativamente l'arcano linguaggio dei documenti di salute pubblica. Esploreremo come le tecniche di *text mining* e *topic modeling* possano supportare questa operazione, facilitando sintesi, riconoscimento di pattern ricorrenti, o casi particolari all'interno di questi testi. La statistica viene talvolta definita come il linguaggio della scienza: in tal caso, alcuni strumenti di traduzione potrebbero tornare utili.

- **Daniele Zago** “*Metodi di ottimizzazione stocastica per problemi statistici complessi*”

L'intervento esplorerà l'utilizzo di tecniche di ottimizzazione stocastica per affrontare problemi statistici caratterizzati da elevata complessità strutturale e vincoli operativi. Verranno presentati approcci basati su simulazione e ricerca stocastica per la risoluzione di problemi analiticamente intrattabili, con un'attenzione particolare alla capacità dei metodi di adattarsi a contesti applicativi eterogenei.

- **Anna Zanin** “*Automazione dei processi di controllo di gestione e integrazione con la business intelligence*”

La digitalizzazione dei processi aziendali sta assumendo un ruolo centrale nelle strategie delle organizzazioni moderne, specialmente in contesti caratterizzati da una crescente disponibilità di dati e dalla crescente complessità dei processi operativi. Il progetto che verrà presentato si concentra sull'automazione dei processi di controllo di gestione all'interno di una multinazionale, articolata in diverse linee di business, ciascuna con le proprie specificità operative. In questo contesto, le attività di chiusura mensile per il bilancio economico risultavano particolarmente onerose, sia in termini di tempo che di risorse, a causa dell'elevato volume di inserimenti manuali e del conseguente rischio di errori.

Lo studio si è focalizzato sull'automazione di tali processi - preceduta da una fase iniziale di integrazione dei dati provenienti da molteplici sistemi e fonti aziendali, sottoposti a rigorosi controlli di qualità per garantirne l'affidabilità e la coerenza. Un elemento chiave del progetto è stato lo sviluppo di un sistema di *Business Intelligence*, in grado di trasformare i dati in informazioni strategiche a supporto delle decisioni, accessibili in tempo reale e strutturate coerentemente con le esigenze dei diversi stakeholder. Questo nuovo approccio ha permesso di far emergere aspetti dell'andamento aziendale prima non evidenti, migliorando significativamente l'accessibilità e la comprensione delle informazioni.

Ore 16:20-16:50 **Coffee break**

Ore 16:50-17:40 **Due interventi di premiati**

- **Nicolas Bianco** “*Modelli statistici per comprendere un mondo in continuo cambiamento.*”

La selezione del modello in processi non stazionari è cruciale per comprendere e interpretare dati complessi, in un mondo in continuo cambiamento. La presentazione illustrerà due approcci metodologici distinti, che in ultima analisi rispondono alla medesima esigenza di affrontare il problema. Il primo filone di ricerca riguarda la selezione delle variabili in modelli a parametri variabili, guidata da dipendenze spaziali e temporali. Il secondo approccio, invece, si concentra sulla rilevazione di punti di cambiamento in serie temporali, caratterizzate da comportamenti ciclici e oscillatori. Questi approcci metodologici sono supportati e giustificati da applicazioni a dati reali che spaziano dall'ambito dell'economia allo studio dei fattori climatici, fino al rilevamento dell'attività cerebrale nel sonno. Nonostante le differenze metodologiche, entrambi gli approcci mirano a semplificare la narrazione dei dati, offrendo strumenti più efficaci per interpretare e comprendere fenomeni complessi e in continua evoluzione.

- **Filippo Ascolani** “*Video consigliati e altre storie: quando la Statistica incontra l'Informatica*”

Negli ultimi decenni, abbiamo imparato a interagire sempre più spesso con modelli statistici nella nostra vita quotidiana. Un esempio banale viene da Youtube, Spotify e altre piattaforme di intrattenimento: ogni volta che visualizziamo un contenuto, ce ne viene suggerito un altro. Tale procedura si basa sui dati riguardanti ciò che noi e gli altri utenti abbiamo guardato, sentito o letto. Tuttavia, ogni giorno vengono girati nuovi video, registrate nuove canzoni e nuovi utenti si iscrivono alla piattaforma: quindi gli algoritmi, basati su appropriati modelli statistici, devono essere aggiornati con nuove informazioni in tempo rapido. Con questo intervento si illustreranno brevemente alcuni risultati ottenuti - nell'ambito di questo studio - nel campo della Statistica Computazionale, all'intersezione tra statistica, matematica e informatica.

Ore 17:40-18:30 **Evento pubblico**

- **Saluti delle Autorità**
- Intervento della **Prof.ssa Francesca Parpinel**, studentessa di Oliviero Lessi
- Intervento del **Dott. Pompeo Volpe**, amico di Oliviero Lessi

Ore 18:30-19:30 **Aperitivo con accompagnamento musicale dal vivo**

Sabato 17 maggio 2025

Ore 9:00-10:40 **Quattro interventi di premiati**

- **Rocco Caprio** “*Nuove prospettive sull’algoritmo EM*”

L'algoritmo di aspettazione-massimizzazione (*expectation-maximization*, EM) è uno strumento centrale nell'analisi statistica in presenza di dati incompleti o variabili latenti. I suoi ambiti di applicazione sono estremamente vasti: dalla fisica all'ingegneria, dalla finanza al machine learning, fino alle più recenti applicazioni nell'intelligenza artificiale generativa. In questo intervento si ripercorreranno l'intuizione di base dell'algoritmo EM e verranno presentate alcune nuove direzioni di ricerca che propongono modi alternativi di interpretarlo, comprenderne meglio il comportamento e valutarne le potenzialità in applicazioni complesse.

- **Andrea Pedicone** “*Sul legame tra i moti aleatori a velocità finita ed il moto browniano*”

I moti aleatori a velocità finita costituiscono una classe di processi stocastici che descrivono il movimento di una particella che si muove a velocità finita e cambia direzione in presenza di un evento aleatorio. L'esempio più significativo di questo tipo di moto è il processo del telegrafo, la cui legge di probabilità soddisfa l'equazione del telegrafo, la stessa che regola la propagazione dei segnali nelle linee telegrafiche.

È di particolare interesse determinare le condizioni in cui il moto browniano emerge come limite di un moto aleatorio a velocità finita: all'aumentare della velocità della particella e del numero di cambiamenti di direzione, la traiettoria del processo del telegrafo tende ad assumere le caratteristiche di un moto browniano.

Questo comportamento limite si manifesta analiticamente nella convergenza dell'equazione del telegrafo verso l'equazione del calore, che descrive la diffusione del calore da una regione calda a una più fredda. Questo legame nasce dal fatto che la funzione che governa l'aleatorietà del processo è una soluzione dell'equazione del calore.

- **Elena Bortolato** “*Un framework dinamico per modelli grafici bayesiani*”

In molti *dataset* reali, specialmente in ambito economico e finanziario, il numero e l'identità delle unità osservate cambiano nel tempo: le imprese entrano ed escono dal mercato, le tecnologie si trasformano e l'intero sistema di dipendenze muta nel tempo. I modelli grafici tradizionali faticano a cogliere elementi dinamici, specialmente in contesti *high dimensional*. In questo intervento propongo un modello grafico bayesiano in grado di gestire dati mancanti di tipo strutturale, legati a dinamiche di "nascita e morte". Il modello si basa su una decomposizione a bassa dimensionalità della matrice di precisione e integra una catena di Markov nascosta per rilevare shock che alterano le dipendenze, adattandosi coerentemente alla variabilità strutturale del sistema nel tempo.

- **Simone Panzeri** “*Metodi non parametrici per dati spazio-temporali su regioni spaziali non-standard*”

Il lavoro si occupa dell'analisi di dati spazio-temporali osservati su domini spaziali non-standard, quali regioni planari irregolari, superfici curve e grafi, come ad esempio aree amministrative, porzioni della superficie terrestre e reti stradali o fluviali. Geometrie non banali, indotte da confini fisici e geografici, emergono in numerosi contesti applicativi, quali l'epidemiologia, l'ecologia e le scienze sociali. In tali applicazioni, le regioni spaziali spesso influenzano i fenomeni osservati e possono compromettere l'efficacia delle metodologie statistiche tradizionali - che trascurano la struttura del dominio fisico.

Per studiare l'evoluzione temporale di dati georeferenziati, si propongono metodi avanzati nell'ambito della classe dei modelli *Physics-Informed Statistical Learning*, che integrano nella metodologia statistica la conoscenza del fenomeno osservato derivante dalla sua descrizione fisica. In questa linea di ricerca, il mio lavoro si concentra sulla stima di densità spazio-temporali a partire da processi di punto spazio-temporali. L'approccio adottato regolarizza l'aderenza ai dati penalizzando la massima verosimiglianza. Il termine di penalizzazione è formulato mediante operatori differenziali opportunamente definiti sul dominio spazio-temporale di interesse e codifica la regolarità della stima ottenuta. In collaborazione con il gruppo di ricerca, sono state stabilite importanti proprietà teoriche relative al problema di stima considerato e allo stimatore proposto. Lo sviluppo metodologico ha richiesto l'uso di tecniche avanzate di analisi numerica e di algoritmi di ottimizzazione a basso costo computazionale, garantendo flessibilità ed efficienza, anche in presenza di domini estesi o grandi quantità di dati. I metodi proposti si sono dimostrati capaci di catturare segnali anisotropi e non stazionari con un elevato grado di accuratezza, mostrando vantaggi significativi rispetto alle tecniche esistenti in letteratura. L'efficacia di tali metodi è stata dimostrata tramite diverse applicazioni a dati reali riguardanti infezioni epidemiologiche, eventi sismici e incidenti stradali.

Ore 10:40-11:10 **Coffee break**

Ore 11:10-12:30 **Tre interventi di premiati**

- **Sofia Curzio** “*Dall’università al mondo corporate: la mia esperienza in Allianz SE*”

Nel contesto internazionale e dinamico del Global P&C Pricing Team di Allianz SE, ha avuto l'opportunità di lavorare a progetti stimolanti e ad alto impatto. L'intento di questo intervento è di raccontare, in linea generale, le attività principali a cui si è dedicata, mettendo in luce ciò che più l'ha arricchita in termini di competenze, responsabilità e metodo di lavoro.

- **Giovanni Romanò** “*Statistica (bayesiana) per svelare complessi fenomeni demografici*”

La grande mole di dati oggi a disposizione richiede strumenti adeguati per estrarre l'informazione in essi contenuta. Anche per questo, la ricerca in ambito demografico ha visto un incremento dell'utilizzo di metodi di natura prettamente statistica. Il lavoro di dottorato, con il gruppo di ricerca del Prof. D. Durante, si inserisce in questo filone di studio con molteplici obiettivi: (i) individuare quesiti aperti, la cui risoluzione possa avere un impatto reale, (ii) trasformarli in domande di ricerca in ambito statistico, (iii) proporre un modello che possa rispondere ai suddetti interrogativi, e (iv) portare un'innovazione metodologica esportabile anche in altri contesti.

Un'interessante dimostrazione di quest'approccio è lo studio delle relazioni tra i diversi fattori che portano alla morte di un individuo: la scoperta di legami attualmente non noti potrebbe avere un notevole impatto in ambito medico. In particolare, l'analisi distingue tra le cause che portano una persona al decesso, in funzione del ruolo che hanno avuto nella serie di eventi che l'hanno condotta alla morte: l'OMS definisce sottostante (*underlying*) la causa che ha avviato la sequenza di eventi morbosi, e multiple (*multiple*) tutti gli altri fattori che hanno contribuito al decesso. Nello specifico, il nostro obiettivo è individuare insiemi di cause sottostanti che condividono simili interazioni con gruppi di cause co-occorrenti, e viceversa, e come queste relazioni evolvono in diverse fasce d'età.

Per evidenziare la forte natura relazionale di questo fenomeno, proponiamo un'innovativa rappresentazione dei dati tramite una rete sociale, in cui le diverse cause di morte agiscono come entità interagenti. Sviluppiamo perciò un nuovo modello statistico per reti che tenga conto del nostro interesse per la sopracitata distinzione tra sottostante e co-occorrente - che possono essere più generalmente interpretati come mittente/destinatario - e l'evoluzione delle interazioni in diverse fasce di età.

- **Matteo Gasparin** “*Quantificazione dell’incertezza per algoritmi black-box*”

Gli algoritmi “*black box*”, come le reti neurali, sono ampiamente utilizzati per fare previsioni, ma spesso è difficile valutarne l'affidabilità. La *conformal prediction* è una tecnica che, oltre a fornire una previsione, assegna un livello di confidenza che aiuta a valutare l'incertezza dei risultati. Questo approccio consente di prendere decisioni più informate, soprattutto in contesti in cui la precisione è cruciale.

Ore 12:30-13:00 **Conclusione e saluti**

Ore 13:00-14:00 **Light lunch**