a.2020

Verona, 14 aprile 2020

**Comunicato stampa**

**Epidemia e statistiche: quando i numeri fanno la differenza**

Intervista a Marco Minozzo, docente di Statistica in ateneo

 **L’aggiornamento sul numero di contagiati è entrato a far parte della nostra quotidianità scandita dall’evoluzione dell’emergenza da coronavirus. Ma quali sono i dati più significativi a cui fare riferimento e perché? Lo abbiamo chiesto a Marco Minozzo, docente di Statistica al dipartimento di Scienze economiche dell’università di Verona.**

**In quest’emergenza siamo sommersi di dati, percentuali e statistiche. Come imparare a interpretare correttamente i numeri che ci vengono forniti senza soccombere ad essi?**

Il numero totale dei **contagiati certificati** è forse il dato numerico che ha ricevuto più attenzione dall'inizio dell’epidemia. Questo numero è stato oggetto di innumerevoli analisi e modellizzazioni, ma in realtà non è molto rappresentativo del fenomeno nel suo complesso. Il numero totale di casi positivi non corrisponde al numero totale dei contagiati (certificati o meno). Esso dipende in larga misura dalla quantità di test effettuati e dalle caratteristiche delle sottopopolazioni di individui sottoposte al test. D’altra parte, anche l’analisi del numero dei **ricoverati** con sintomi o del numero di persone in terapia intensiva non è privo di controindicazioni. Infatti, questo numero dipende sia dalla capacità del sistema sanitario di intercettare gli ammalati, sia dalla saturazione dei reparti e dalle strategie di ospedalizzazione messe in atto dalle sanità regionali. Non è privo di problemi nemmeno il numero di **deceduti**. Proprio all’inizio di aprile l’Istat ha rilasciato i dati sulla mortalità dei primi mesi dell’anno, evidenziando come in certi Comuni e Province lombarde il tasso di mortalità sia stato ben superiore sia a quello degli ultimi anni, sia a quello previsto in base al numero di decessi da Covid-19 rilasciato dalla Protezione Civile.

Riguardo ai **test,** vale la pena di ricordare che i dati forniti si riferiscono al numero di tamponi effettuati e non al numero di soggetti testati. Quindi, stante il fatto che molti individui vengono sottoposti al tampone più di una volta, anche il semplice rapporto tra il numero di casi positivi e il numero di tamponi effettuati deve essere interpretato con una certa cautela.

Nelle ultime settimane di marzo si è assistito a un crescendo di interesse attorno alla data del cosiddetto picco epidemico, ma in un’epidemia ci sono diversi picchi. Nell’epidemia di Covid-19, il picco dei nuovi contagi, sia per il Veneto sia per l’Italia nel suo complesso, si è probabilmente avuto attorno al 10 marzo (in concomitanza con il lockdown nazionale), mentre il picco dei nuovi ammalati (ovvero di chi ha iniziato a manifestare i sintomi) si è verosimilmente avuto attorno o poco dopo il 20 marzo. Invece, il picco dei pazienti ospedalizzati ed in cura, di estremo interesse per valutare la sostenibilità del sistema sanitario, si è probabilmente avuto, almeno in alcune regioni del Nord, tra la fine di marzo e i primi giorni di aprile. Similmente per il picco dei decessi. Questi picchi riguardano, ovviamente, la prima ondata dell’epidemia di Covid-19. L’andamento delle successive ondate dipenderà dalla disponibilità di farmaci e vaccini, dalla più o meno acquisita immunità di gregge, oltre che dalle misure di sanità pubblica (distanziamento sociale, utilizzo di dispositivi di protezione individuale ecc.) che saremo in grado di implementare.

**In che modo, secondo lei, i dati costituiscono uno strumento per combattere la battaglia al Coronavirus?**

È ovvio che senza una misurazione tempestiva e accurata dei fenomeni che ci circondano saremmo in completa balìa degli eventi. Sull’importanza della tempestività della raccolta di informazioni epidemiologiche basti solo ricordare come a ogni epidemiologo sia ben chiaro che nel momento in cui si registra anche un solo decesso per un virus sconosciuto, ci sono probabilmente nella popolazione già centinaia se non migliaia di persone infette. Non per nulla il medico cinese Li Wenliang, che a Wuhan nel dicembre del 2019 ha scoperto i primi sette casi di Covid-19, dopo non essere stato ascoltato dalle autorità locali, ha ritenuto di lanciare l’allarme su internet.

**Nell’analisi della crisi contano solo i numeri o ci sono altri fattori, più trascurati, che invece è importante considerare per avere una fotografia più veritiera di quanto stiamo vivendo?**

Ogni quantificazione di un qualsiasi fenomeno fornirà sempre una visione parziale, anche se a volte utilissima, dello stesso. In questo momento stiamo assistendo a un’esplosione di analisi e modellizzazioni che, oltre a non tenere in debito conto di come sono generati e raccolti i dati, cioè di quali sono le grandezze effettivamente misurate, sono completamente o quasi esclusivamente data-driven. Ossia cercano di “far parlare” i dati senza preoccuparsi troppo della peculiarità del fenomeno epidemiologico oggetto di studio. Questo è in linea con i tempi che stiamo vivendo, dominati in lungo e in largo dalla scienza dell’analisi dei dati in cui alcune tecniche e algoritmi statistici stanno dimostrando più di altri il loro potenziale in più di un ambito di applicazione. Tuttavia, nell’attuale contesto epidemiologico, molte di queste analisi e modellizzazioni hanno uno scarso valore, soprattutto predittivo, e sarebbe auspicabile focalizzare l’attenzione su modelli theory-driven, ovvero su modelli che diano conto sia di come sono generati e raccolti i dati sia delle specifiche conoscenze epidemiologiche.

**Cosa, in questo momento, è necessario calcolare o stimare e perché?**

Un parametro fondamentale è la proporzione di contagiati nella popolazione. Come molti altri parametri di questa epidemia, questa proporzione si può solo stimare attraverso modelli probabilistici e indagini campionarie. All’inizio di aprile gli studi più autorevoli sembravano indicare un numero di persone contagiate attorno ai 6 milioni, con un margine di incertezza tra 1 e 10 milioni. Una stima più affidabile si potrà avere solo con indagini campionarie della popolazione utilizzando sia test del tampone sia test sierologici (anticorpali) per individuare anche chi è stato infettato senza sviluppare sintomi. Uno studio sieroepidemiologico della popolazione permetterà di stimare la prevalenza per zona geografica, fasce di età, genere, condizione occupazionale e professionale, settore di attività economica ecc., e quindi di comprendere meglio le caratteristiche dell’epidemia per poter affinare le strategie di contrasto e di ripartenza per i prossimi mesi.

Un altro parametro di vitale importanza è il cosiddetto “numero di riproduzione di base” indicato con R0, ossia il numero medio di infezioni secondarie prodotte da ciascun individuo infetto. Attualmente si hanno solo stime preliminari sul valore di questo parametro per una popolazione completamente suscettibile, cioè mai venuta a contatto con il virus SARS-CoV-2. Fondamentalmente, le misure di distanziamento sociale fronteggiano la diffusione del virus riducendo R0. Esse non incidono sul livello di letalità dell’infezione, né sui tempi di guarigione o morte dei malati, se non indirettamente riducendo il numero di malati che necessitano di cure e quindi rendendo disponibili maggiori risorse (personale sanitario, ventilatori polmonari ecc.) e permettendo livelli di cura più elevati ed efficaci. L’obiettivo di tutte le misure di distanziamento sociale messe in atto è quello di portare R0 ad un valore inferiore ad uno, in un intervallo tra 0,75 e 0,5, cioè a un valore tale da soffocare l’epidemia.

**Università degli Studi di Verona**

**Ufficio Stampa e Comunicazione istituzionale**

Area Comunicazione

Direzione Comunicazione e Governance

Telefono: 045.8028015 - 8717

M. 349.1536099

Email: ufficio.stampa@ateneo.univr.it