



UNIVERSITA' DEGLI STUDI DI PADOVA

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
"M.FANNO"**

DIPARTIMENTO DI SCIENZE STATISTICHE

CORSO DI LAUREA TRIENNALE IN ECONOMIA

PROVA FINALE

**AL DI LÀ DI $p < 0.05$
IMPLICAZIONI ECONOMICHE DI UN DIBATTITO STATISTICO**

RELATORE: TOMMASO DI FONZO

LAUREANDO: NICOLA MELLINATO

MATRICOLA N. 1113076

ANNO ACCADEMICO 2018 – 2019

INDICE

INTRODUZIONE	1
1. I DUBBI DI “THE AMERICAN STATISTICIAN”	3
1.1 La concezione di Fisher e i Big Data	5
1.2 Come muoversi in un contesto incerto.....	6
1.2.1 Accettare l’incertezza	6
1.2.2 Essere premurosi nella ricerca.....	7
1.2.3 Essere aperti alla trasparenza, al giudizio e alla comunicazione.....	10
1.2.4 Essere modesti	11
2 SOLUZIONI PRATICHE DEGLI AUTORI AL PROBLEMA.....	12
2.1 Goodman: riflessioni sulla statistica e i principali attori del sistema	12
2.1.1 Istituzioni	13
2.1.2 Giornali	13
2.1.3 Finanziatori	14
2.2 Un corso di pensiero statistico	15
2.2.1 Insegnare la saggezza statistica	15
2.2.2 La struttura del corso.....	15
2.2.3 Le sezioni del corso.....	16
2.2.4 Il progetto finale	17
2.2.5 Caratteristiche insolite del corso	18
2.2.6 Il p-value e i campioni a confronto	19
2.3 La Guinnessometria di Ziliak.....	20
2.3.1 I 10 principi “G”	20
3 PROCESSO DECISIONALE E RUOLO DEL MANAGER IN UN CONTESTO D’ INCERTEZZA.....	27
3.1 Fasi del processo decisionale.....	27
3.2 Dalla razionalità all’intuito.....	29
3.2.1 Esempio di razionalità: i modelli statistici.....	29
3.2.2 La razionalità limitata	30
3.3 Caratteristiche delle decisioni	31
3.3.1 Tipologie di decisione	31
3.3.2 Le condizioni delle decisioni	31
3.3.3 Stili ed errori decisionali	32
3.4 Il management oggi. Possibili soluzioni e supporto al processo decisionale	33
3.4.1 Learning organizations e decisioni in comune	34
3.4.2 Decision Support Systems.....	35

3.5 I manager ed il concetto ATOM	37
CONCLUSIONI.....	39
BIBLIOGRAFIA	40

INTRODUZIONE

Questo lavoro ripercorre le problematiche legate al livello di probabilità osservato (p-value) ed all'approccio inferenziale fondato sulla teoria delle decisioni. Sir Ronald Fisher, colui che ha stabilito i concetti di riferimento della statistica moderna, fu tra i primi a comprendere l'importanza del campionamento casuale per poter generalizzare i risultati; ma questo paradigma è ancora valido con l'avvento dei Big Data? E per la disciplina economica sociale?

Partendo da un concetto "totem", quale ormai è universalmente considerato il p-value, si approfondirà la nozione di incertezza che contraddistingue ogni tipo di ricerca e processo decisionale, a fronte di una moltitudine di variabili non totalmente gestibili dall'individuo. In particolar modo la rivista "The American Statistician" ha sollevato la problematica cercando di trovare delle soluzioni e modi d'essere più congrui nell'ambito della ricerca. Superare il concetto di "significatività" è il primo passo, ed il concetto "ATOM" sintetizza in quattro punti il nuovo approccio ad un mondo dove regole stringenti limitano il sapere e la curiosità.

La tesi proporrà delle soluzioni pratiche alle classiche concezioni, trovando nei contributi di Goodman (2018), Steel, Liermann, Guttorp (2018) e Ziliak (2018) esempi di come si possa avviare e, talvolta, superare approcci rigidi e ormai obsoleti, per migliorare la ricerca e la fase decisionale. Goodman (2018) offre una riflessione generale sull'uso del p-value dal punto di vista sociale, suggerendo anche delle possibili innovazioni che vadano ad "utilizzare" gli attori principali, istituzioni finanziatori e giornali, come veicolo di un buon processo di ricerca, attento a non tralasciare dettagli che oggi sono ormai irrinunciabili. Il trio di autori Steel, Liermann e Guttorp (2018) propone un corso universitario che va oltre le classiche lezioni di statistica, al fine di formare al meglio i nuovi ricercatori ed i nuovi statistici. Ziliak (2018) considera invece la problematica decisionale in un contesto maggiormente economico, rifacendosi a Gosset e alle sue teorie, sviluppate da questo autore classico durante il suo periodo di lavoro alla Guinness di Dublino, arrivando a suggerire il passaggio da "piccoli" p-value a grandi "g-value", e sintetizzando il tutto in 10 principi.

Dal punto di vista di uno studioso di processi e problemi economici gli strumenti quantitativi a supporto delle decisioni non possono limitarsi a quelli forniti da un corso base di statistica universitario, che è necessariamente legato ad una visione sperimentale basata sul principio del campionamento ripetuto. L'elemento comune tra le due materie è quella matrice di incertezza che verrà esplicitata nel capitolo tre, con riferimento alla figura del manager nell'aspetto più importante del suo ruolo, ovvero la "decisione". Il capitolo analizzerà

dapprima le fasi classiche del processo decisionale, spiegando poi come nella fase di scelta sia importante sia la razionalità sia l'intuito. Verranno introdotti i modelli statistici, evidenziandone limiti e opportunità, così come saranno enunciate tipologie, condizioni, errori e stili delle decisioni effettuate dal manager, il quale non potrà sempre affidarsi a schemi e convinzioni proprie per decidere. La dinamicità elevata del mercato attuale porta nuove sfide al decisore, il quale dovrà munirsi dei giusti mezzi per affrontare al meglio la fase di decision-making. Vengono così evidenziati due tipologie di soluzioni/strumenti, la learning organization e i DSS (decision support sistem), il primo è un modo d'essere dell'azienda, mentre il secondo sono sistemi informatici computer based per supportare la fase decisionale. Il capitolo poi concluderà con una riflessione ed un parallelismo tra la figura del manager e quella del ricercatore in relazione al concetto ATOM.

1. I DUBBI DI “THE AMERICAN STATISTICIAN”

1.1 The American Statistician



Fonte: ese-bookshelf.blogspot.com

Dietro al p-value ed ai test di significatività ci sono dei pericoli che non sono stati presi in considerazione per decenni. Ci sono delle credenze che devono essere “eliminate”, non si dovrà basare tutto sullo “statisticamente significativo” o che in ragione di quest’ultimo concetto sia possibile argomentare su quale sia la causa di un determinato effetto. Allo stesso modo può lasciare interdetti la conclusione secondo la quale l’assenza di un effetto sia dovuta al fatto che questo non sia significativo, così come la credenza che il p-value garantisca che un’ipotesi sia vera. Questo approccio non sarà quindi più sufficiente per prendere decisioni e stilare conclusioni scientifiche rilevanti.

Conoscere cosa non fare con il p-value è necessario, ma non risolutorio. Si è sempre cercato di distruggere i paletti della moderna ricerca scientifica, senza mai senza tentare di ricostruirla con solide basi. Ora abbiamo bisogno di fare un passo ulteriore.

Negli articoli pubblicati nel 2018 nella rivista “The American Statistician” (immagine 1.1) vengono proposte nuove idee, le quali meritano di essere prese in considerazione aprendo un dibattito. Si nota comunque, nelle pagine dell’edizione presa in esame, una dissonanza tra gli autori, la cosiddetta “statistic wars”. Si apre così una nuova era della statistica, e un mondo pronto ad avventurarsi al di là di “ $p < 0.05$ ”.

In questo studio vedremo come un p-value minore di 0.05 o un p-value maggiore di 0.05 non siano automaticamente in conflitto. Avventurandoci in questa direzione, ci sarà un minor numero di risultati non presi in considerazione, così come lo sviluppo di strategie statistiche maggiormente personalizzate. Il lavoro dei ricercatori deve essere valutato sulla base dell’effettiva qualità metodologica e di una corretta comunicazione della loro scienza, e non

basandosi esclusivamente sul p-value. Si passerà così maggiormente da un significato statistico ad un pensiero statistico.

Si vuole dare una soluzione pratica al problema, ma prima di enunciare ciò che si deve fare, dobbiamo spiegare ciò che non si deve fare. È importante innanzitutto smettere di utilizzare il termine “statisticamente significativo”, ed altre sue varianti come “significativamente diverso da...”, “ $p < 0.05$ ” e “non significativo”. L’intenzione originale di Edgeworth per la significatività statistica, era quella di utilizzare quest’ultima come strumento per indicare un risultato che meritasse un ulteriore controllo. Ma questa idea è stata irrimediabilmente persa. La rilevanza statistica non doveva implicare rilevanza scientifica, ma a causa del suo diffuso utilizzo i due concetti sono andati a confondersi, e così lo strumento è diventato “tiranno”. L’utilizzo di regole fisse per giustificare affermazioni scientifiche o conclusioni decisionali può portare a credenze errate e decisioni sbagliate, infatti un’etichetta di significatività statistica non aggiunge niente a ciò che è stato trasmesso dal p-value.

Nessun p-value può rivelare la plausibilità, la presenza, la verità o l’importanza di un’associazione o un effetto. Senza delle linee guida precise quindi diventa difficile, per esempio, dare delle interpretazioni in caso di differenze insignificanti nelle stime. La differenza tra “significativo” e “non significativo” diventa quindi essa stessa non significativa.

Questa divisione tra ciò che è giusto e ciò che è sbagliato porta a considerare i risultati sulla base della loro significatività. A causa di questa dicotomia secondo Amrhein, Trafimow e Greenland (2019), il problema va oltre la semplice pubblicazione, infatti gli autori che utilizzano p-value soglia per selezionare i risultati da discutere nei loro documenti, escluderanno inevitabilmente quelli che non rientrano nei limiti fissati. Tale attenzione selettiva quindi, non solo distorcerà la letteratura, ma metterà in luce i risultati prima citati, pubblicandoli e influenzando professionisti e pubblico in generale. Per l’integrità dell’editoria scientifica qualora una ricerca vada a superare il livello soglia p scelto arbitrariamente, questa non deve essere presa in considerazione, e ciò non dovrebbe accadere. Allo stesso modo, bisogna smettere di utilizzare gli intervalli di confidenza come un altro mezzo per dividere ciò che deve essere accettato e ciò che non deve essere accettato. Questo concetto, della non categorizzazione, deve essere di conseguenza applicato anche alle altre misure statistiche.

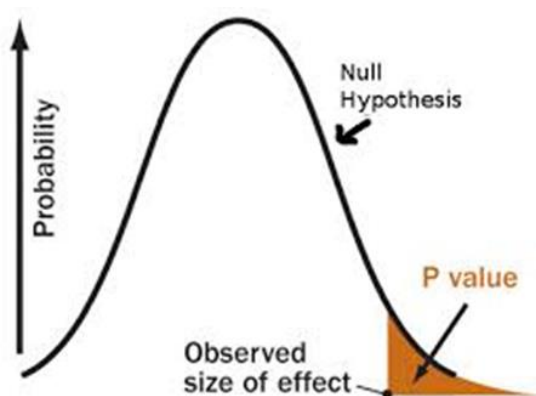
Il valore p non dovrà essere eliminato, ma invece dovrà essere utilizzato e riportato come una quantità continua, con una descrizione del significato del valore nel contesto scientifico preso in esame. Un p appropriato è tale solo se viene interpretato nella maniera corretta.

Per muoverci oltre “ $p < 0.05$ ” dobbiamo essere consapevoli che l’inferenza statistica non è uguale all’inferenza scientifica. Andare oltre la "significatività" apre ai ricercatori il vero significato della statistica, che è "la scienza dell'apprendimento dai dati e della misurazione, controllo e comunicazione dell'incertezza".

1.1 La concezione di Fisher e i Big Data

Tutta la concezione classica si basa sul paradigma coniato da Ronald Fisher (Londra, 17 febbraio 1890- Adelaide, 29 luglio 1962), colui che viene considerato il fondatore dei concetti di riferimento della statistica-matematica moderna. Egli fu tra i primi a comprendere l'importanza del campionamento casuale in opposizione ai campionamenti fatti secondo vari criteri di opportunità per poter generalizzare i risultati. Coniò inoltre i concetti di ipotesi nulla (H_0) e ipotesi sperimentale (H_1), affermando che nessun tipo di ricerca può dimostrare H_1 ma solo accettare o rifiutare l'ipotesi di partenza. Fisher associava ai test d'ipotesi livelli di significatività α , cioè la possibilità rilevante che compaia un determinato valore. Dato questo livello, un test statistico ci porterà a rifiutare o meno l'ipotesi nulla a favore dell'ipotesi alternativa. Con l'introduzione del p value notiamo una differenza, infatti questo valore non è una quantità fissa come α bensì è un valore variabile che misura l'evidenza dei dati contro H_0 , come possiamo notare dal grafico 1.2. Esso è dato dalla probabilità di osservare un valore della statistica test uguale o più estremo del valore ottenuto dal campione, sotto l'ipotesi nulla.

1.2 Grafico p-value



Fonte: sciencebasedmedicine.org

Come abbiamo detto in fase di introduzione del capitolo, l'utilizzo di questi metodi è stato utilizzato senza un vero criterio, non tenendo in considerazione le caratteristiche del contesto di riferimento. Fisher utilizzava un classico campionamento ripetuto per i suoi esperimenti, ed

il p-value veniva utilizzato con piccoli campioni. Nelle scienze e negli ambienti economici attuali i campionamenti sono solitamente *one shot* al contrario di ciò che veniva fatto in passato, ed inoltre siamo esposti ad una moltitudine di eventi e dati che rendono il contesto altamente variabile e ricco di informazioni, con la possibilità di utilizzare campioni enormi.

Vediamo dunque che la classica teoria di Fisher scricchiola con l'avvento dei Big Data, ovvero insiemi di informazioni estremamente grandi i quali possono essere analizzati attraverso computer per rivelare modelli e tendenze. La rapida diffusione delle nuove tecnologie permette un accumulo enorme di dati, con un ritmo che cresce ad un tasso annuo del 40%. Il fenomeno analizzato è presente nelle economie più avanzate, e la montagna di informazioni combinata con la potenza di calcolo, grazie all'utilizzo di software, garantiscono nuove interessanti alternative alle misurazioni economiche, opportunità che ovviamente non potrebbero essere colte con i classici metodi statistici. Vediamo dunque non solo un problema di natura tecnica nell'utilizzo della significatività e del p-value, ma una vera e propria obsolescenza soprattutto in ambito economico delle regole e delle norme che sorreggevano questi metodi.

1.2 Come muoversi in un contesto incerto

Dal lavoro pubblicato nel 2018 su “The American Statistician” non emerge una soluzione univoca, in grado di prevalere sulle altre. La comunità statistica infatti deve ancora convergere verso un unico paradigma per quanto riguarda l'inferenza e la ricerca scientifica. Detto ciò, non si può pensare di lavorare e muoversi in questo contesto senza delle linee guida di comportamento necessarie durante le fasi del processo di studio posto in essere, al fine di trovare risultati adeguati e credibili. Le raccomandazioni in merito sono sintetizzate attraverso il concetto ATOM: “Accept uncertainty. be Thoughtful, Open, and Modest.”

1.2.1 Accettare l'incertezza

L'incertezza è una costante nella ricerca, così come la variabilità, tanto è vero che l'esatta replicazione di uno studio risulta spesso difficile da ottenere. Introdurre l'incertezza nel contesto statistico significa considerare i risultati ottenuti con una maggiore cautela, pertanto ci serviranno metodi per quantificare, visualizzare ed interpretare il potenziale d'errore.

Molteplici sono i metodi per accettare l'incertezza, come accompagnare ogni punto stimato con uno standar error o introducendo intervalli di fiducia. Se i soggetti sono consapevoli della situazione andranno ad utilizzare misure migliori, progettazioni più attente, campioni più grandi (e non piccoli come quelli che considerava Fisher), incrementando così il rigore della ricerca. Sarà quindi importante trovare repliche ed evidenze degli studi, per esempio attraverso la meta-analisi, strumento il cui scopo è quello di riassumere i dati provenienti da diversi strumenti di ricerca al fine di ottenere un unico indice quantitativo di stima che permetta di trarre conclusioni più forti di quelle basate su ciascun singolo studio.

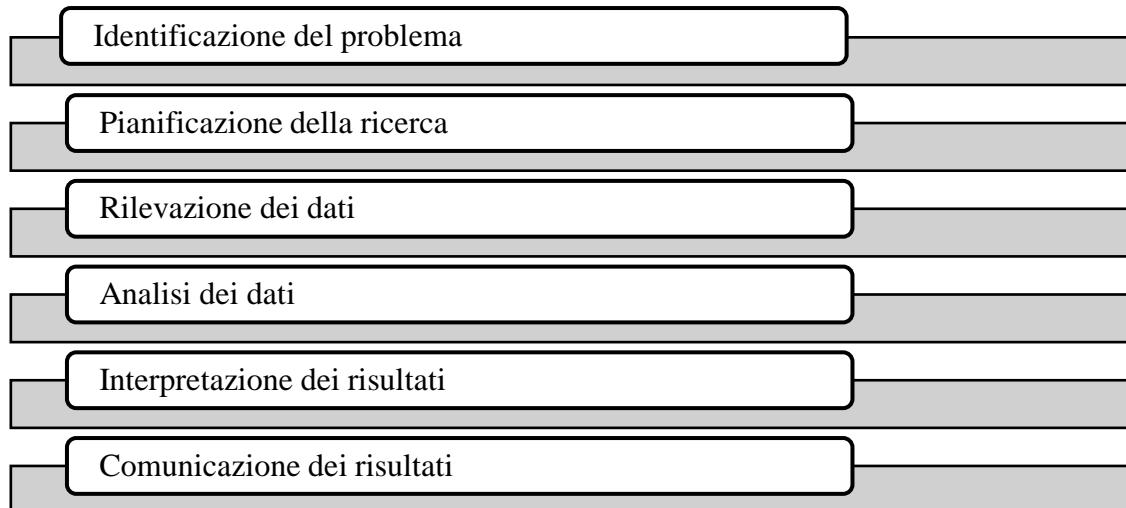
1.2.2 Essere premurosi nella ricerca

Un ricercatore statisticamente premuroso inizia i suoi studi con chiari obiettivi espressi, riconoscendo gli studi esplorativi da quelli pre-pianificati, investendo nella produzione di dati e considerando una moltitudine di tecniche di analisi delle informazioni.

La condotta e l'analisi di uno studio sono necessariamente flessibili, bisognerà così esser pronti a trovarsi di fronte a modelli inaspettati che andranno a generare nuove domande e nuove ipotesi. Il prezzo però da pagare consentendo questa flessibilità è che la validità di ogni inferenza statistica è compromessa, in quanto nessun risultato può essere certo di fronte a questo contesto variabile.

Una ricerca premurosa dà priorità alla produzione di dati, incentivando una rigorosa pianificazione ed esecuzione dello studio. Alcuni autori affermano addirittura che la valutazione delle ricerche dovrebbe essere basata sulle domande che ci si è posti ed il metodo utilizzato per rispondere a queste, piuttosto che sui risultati ottenuti. Congiuntamente poi le assunzioni dei modelli di ricerca dovrebbero essere sufficientemente documentate cosicché parti indipendenti possano replicare e criticare il lavoro.

Dalla ricerca fino ai processi decisionali d'azienda ci sono dei modelli classici di riferimento. Questi servono per dare un filo logico e coerente ai modi d'agire del ricercatore, nonostante sia evidente che si possano trovare degli inconvenienti durante i vari step del processo, indicati nella tabella 1.3. Se necessario quindi bisognerà intervenire con metodi correttivi rispetto al lavoro che si sta svolgendo.



Una ricerca ponderata andrà a considerare il contesto scientifico attuale e le prove precedenti affini ad essa, mentre i ricercatori dovrebbero chiedersi “Cosa sappiamo? Quanto certi siamo su quello che già è di nostra conoscenza?” e ancora “Quali differenze, odds ratio (dato statistico che misura grado correlazione tra due fattori, esempio: fattore di rischio-malattia) o altri effetti sono importanti?”. Tramite queste linee guida si potranno utilizzare prove esistenti dalla letteratura, le quali saranno importanti per i risultati chiave rispetto allo studio attuato.

La cosa più pericolosa e fuorviante è andare a giustificare risultati dopo che questi sono avvenuti, andando ad interpretare effetti banali considerandoli significativi. Sarà quindi importante verificare la validità dei meccanismi di ricerca, la progettazione e la qualità dei dati, i reali costi di implementazione dello studio rispetto ai benefici derivanti da esso, la reale novità della scoperta e altri fattori che variano da ricerca a ricerca. Un’attitudine del genere non darà priorità al p-value o ad altre misure simili, ma porterà il ricercatore ad utilizzare una “cassetta” di tecniche statistiche, lavorando con buonsenso e tenendo in considerazione lo sviluppo della statistica rispetto ai dati scientifici.

Come abbiamo detto verranno utilizzati più approcci durante la ricerca per risolvere i problemi. In particolar modo, soffermandoci sul p-value, c’è chi suggerisce di integrarlo o sostituirlo con valori che vadano a garantire maggior chiarezza ed evitando confusione nelle considerazioni finali. Un esempio di approccio alternativo è il p-value di seconda generazione di Blume (SPGV), che coincide essenzialmente con la proporzione di ipotesi nulle supportate da dati. È evidente che questo può rappresentare un passo avanti in ambienti ricchi di dati, come appunto quello economico. Per i sostenitori dell’SPGV, questo va ad eliminare i problemi classici della significatività statistica, riducendo il rischio di ottenere una falsa

scoperta e garantendo conclusioni con maggior probabilità di essere riprodotte in studi successivi. Tra i valori più rilevanti troviamo:

- $SPGV=1$ i dati supportano l'ipotesi nulla
- $SPGV=0$ i dati sono incompatibili con l'ipotesi nulla
- $0 < SPGV < 1$ nessuna conclusione definitiva

Una volta definiti i metodi di risoluzione della ricerca, converrà sempre fare un passo indietro dando una comunicazione dei propri studi e risultati evitando affermazioni fin troppo convincenti. Goodman (2018), per esempio, propone di accompagnare, ad ogni dichiarazione dello studio, la stima delle possibilità che le stesse dichiarazioni siano vere.

Es. Goodman:

1. Affermazione iniziale:” A questo tipo di farmaco è associato un elevato rischio di attacco di cuore.”
2. Affermazione probabilistica:” C'è un 80% di possibilità che questo farmaco aumenti il rischio di attacco di cuore, e c'è un 60% di possibilità che il rischio sia almeno raddoppiato.”

La probabilità è necessaria per prendere decisioni, soprattutto in un contesto aziendale. Si pensi ad una linea di produzione che venga organizzata senza un'adeguata pianificazione probabilistica delle vendite. Il responsabile si troverà in difficoltà nel gestire poi le turbolenze del mercato, incorrendo in elevati costi di gestione e probabilmente di magazzino, non riuscendo poi a soddisfare totalmente la domanda di mercato con rapidità ed efficienza. La statistica e i software relativi possono essere d'aiuto in questo caso, vediamo per esempio che tramite un open source come R, il quale contiene una miriade di pacchetti e funzioni statistiche, si possano ottenere buone previsioni di vendita sulla base di dati storici dell'azienda.

1.2.3 Essere aperti alla trasparenza, al giudizio e alla comunicazione

Un gran numero di autori, citati nella rivista americana, sono fautori di pratiche innovative quali: la pre-registrazione pubblica dei metodi di ricerca, la condivisione dei dati e dei codici, un'eventuale recensione a priori dello studio, la trasparenza e completezza dei report. Quest'ultima in particolare non implica solo la descrizione delle analisi effettuate, ma anche la presentazione di tutti i risultati ottenuti, senza rifarsi alla significatività statistica o a qualsiasi altro criterio.

Mettere a nudo il proprio lavoro, comporterà l'apertura al ruolo, ed al conseguente giudizio di esperti del settore, i quali andranno ad incidere sugli aspetti della pratica inferenziale statistica e nei processi decisionali. Questi soggetti aiuteranno i ricercatori nello sviluppo di modelli e nell'interpretazione dei risultati, oltre a correggere eventuali errori degli studi, dovuti per esempio a quella componente soggettiva (schemi, proprie idee...) dell'individuo durante le fasi del processo. Il ricercatore dovrà così basarsi sull'evidenza e su ragionamenti attenti, invece di concentrarsi su convinzioni che non per forza sono corrette.

Proseguendo il ragionamento d'apertura converrà comunicare il p-value dei propri studi; "novità" che destabilizza ovviamente la concezione classica dei ricercatori, ma se si proponesse un modello universale per la presentazione ed interpretazione dei valori in questione, verrebbero violati tutti i principi sopra citati. L'utilizzo del p-value dovrà quindi variare in base alla tipologia di studio posto in essere dai ricercatori, e ci saranno poi dei fattori che non saranno più subordinati a quest'ultimo, ma saranno dei veri e propri capisaldi di una buona ricerca, come la progettazione, la qualità dei dati, costi e benefici reali che determinano quali effetti sono scientificamente importanti.

L'utilizzo "sano" degli intervalli di confidenza non può che essere un plus del nostro ragionamento. Conviene discutere i limiti inferiori e superiori dell'intervallo, non dando particolare importanza se questo include o meno il valore nullo e ricordando che esso stesso è una stima soggetta ad errore e generalmente fornisce solo un valore approssimativo dell'incertezza.

Il non essere aperto nel report conduce a pubblicazioni con pregiudizi, sapendo inoltre che un unico studio è raramente sufficiente. Uno studio pionieristico dovrà quindi resistere ai ricercatori che andranno a valutarlo, permettendo loro, grazie alla fornitura di uno stock di informazioni, di eseguire significative analisi alternative, al fine di arrivare a risultati ottimi e conclusivi.

1.2.4 Essere modesti

I ricercatori con maggior successo sono modesti durante i loro studi, capendone e dichiarandone i limiti. Il problema maggiore è che pubblico e scienziati confondono la statistica con il mondo reale, ma l'inferenza è un esperimento mentale, e i modelli utilizzati sono necessariamente semplificati rispetto alla realtà studiata. I risultati derivanti dalla ricerca non possono descrivere la complessità ambientale; quando si crede che questi ultimi effettivamente rappresentino il “vero”, non è un problema causato dai metodi utilizzati, ma è un problema di interpretazione e comunicazione dei risultati. Essere modesti significa riconoscere che non esiste un modello statistico che possa risolvere ogni cosa, per questo è saggio andare a considerare una serie di possibili alternative.

Modestia significa riportare tutto, mentre allo stesso tempo non si deve concludere nulla di un singolo studio con ingiustificata certezza. A causa del forte desiderio di informare ed essere informati, c'è una richiesta inarrestabile di affermazione dei risultati senza un'adeguata motivazione. È quindi di fondamentale importanza accettare l'incertezza e la variabilità degli effetti, ed il fatto che un singolo studio è difficilmente definitivo, quindi conviene sponsorizzare, condurre e pubblicare repliche di quest'ultimo. Hubbard e Carrquiry (2018) suggeriscono ed incoraggiano la riproduzione della ricerca, proponendo uno stato di sottotitolo a coloro che replicano studi. I due autori vorrebbero vedere le versioni digitali dei documenti aggiornate dinamicamente per visualizzare "Riprodotta da ..." o "non ancora riprodotta" sotto i nomi degli autori di ricerca originali.

Se la replica non trova gli stessi risultati, questa non è necessariamente un elemento negativo, ma è parte di un processo naturale attraverso il quale la scienza evolve. L'obiettivo della metodologia scientifica dovrebbe essere quello di indirizzare questa evoluzione verso sempre più accurate descrizioni del mondo e di come “funziona”, e non verso una continua pubblicazione di inferenze, conclusioni o decisioni. Bisogna quindi seguire il concetto “un fenomeno, più studi” piuttosto che “tanti fenomeni, uno studio”.

Va invece riconosciuto che i lettori hanno volontà e scopi non omogenei rispetto alle analisi effettuate. Il che significa che si dovrebbe mantenere un punto di vista neutrale piuttosto che difendere una qualsiasi ipotesi, accettando che sia l'inferenza scientifica sia l'inferenza statistica sono di difficile comprensione e che nessuna conoscenza potrà essere ragionevolmente perseguita usando regole e procedure semplicistiche e meccaniche. L'incertezza ci sarà sempre. Siate premurosi, aperti e modesti. ATOM.

2 SOLUZIONI PRATICHE DEGLI AUTORI AL PROBLEMA

In questo capitolo vedremo come tre autori abbiano provato a fare un salto in avanti rispetto alla classica concezione inferenziale. Ognuno con le sue peculiarità pone l'accento su differenti aspetti della ricerca e dei processi decisionali e su come possano essere migliorati.

Nonostante anni e anni di critiche incisive alla significatività statistica, questa non venne mai scalzata come approccio dominante a livello scientifico. Oggi invece il momento del cambiamento sembra prendere sempre più piede. Il progresso sarà lento, e prima di vedere i suoi effetti ci vorrà pazienza e tempo, così come per tutti i movimenti sociali su larga scala.

Goodman, Greenland, Steel, Liermann e Guttorp sono solo alcune figure che hanno provato a trovare delle soluzioni a riguardo. Nel complesso, spaziando su più punti, il loro mix di idee e teorie risulta come un set di regole e consigli adatti per affrontare le esigenze dei nostri giorni, le quali saranno utili non solo per i professionisti di settori affini alla ricerca, ma anche per le organizzazioni, per i manager e perfino per gli studenti universitari.

2.1 Goodman: riflessioni sulla statistica e i principali attori del sistema

Steven Goodman è un professore americano di medicina presso la Stanford School of Medicine e le sue riflessioni partono dalla domanda sul perché eliminare il valore p sia così difficile. La spiegazione di base non è né filosofica né scientifica, ma sociologica: tutti lo usano. Quando tutti credono nel valore di qualcosa, questa viene utilizzata per cose reali; i valori p vengono presi in causa per richieste di conoscenza, pubblicazione, finanziamento e promozione. Non importa se il medesimo valore non significa ciò che la gente pensa che significhi, diventa prezioso a causa di ciò che acquista.

Il p -value fa parte di una struttura basata su regole rigide che funge da baluardo contro le pretese di conoscenza non legate dal supporto empirico. Possono essere cambiate, ma in primo luogo dobbiamo rispettare il motivo per cui esistono le procedure statistiche. Questo spiega in parte perché c'è così tanta resistenza agli approcci bayesiani, che sono spesso visti come un modo per reintrodurre la soggettività, che i metodi statistici tendono il più possibile ad evitare.

La nostra volontà è quella di arrivare ad una buona pratica scientifica che ci permetta di raggiungere i due obiettivi della statistica: misurare l'effetto e la sua incertezza correttamente. Per far ciò basterà eliminare i valori p ? In un mondo idealizzato, la risposta è sì, e molti statistici lo credono. Ma nel mondo reale, questa sicurezza viene meno.

Per Goodman conviene così collaborare, per dare uno slancio al cambiamento, con movimenti di riforma della scienza, riviste, agenzie di finanziamento e autorità di regolamentazione al fine di promuovere e premiare la scienza riproducibile e ridurre l'impatto della rilevanza statistica su pubblicazioni, finanziamenti e promozioni.

2.1.1 Istituzioni

Un rapporto del 2018 pubblicato dalla National Academy of Sciences intitolato "Promuovere l'integrità nella ricerca" ha ampliato la definizione di quest'ultima oltre le classiche categorie di falsificazione, in una varietà di pratiche inferenziali fuorvianti chiamate "pratiche di ricerca dannose". Queste, se non tutte, implicano una deviazione dalla progettazione, condotta, analisi e interpretazione statistica. Porre una pratica adeguata nel contesto scientifico, con strutture istituzionali per sorvegliarla, può fornire agli statistici la possibilità di influenzare il modo in cui le stesse istituzioni insegnano, monitorano e promuovono corrette regole scientifiche. Significa quindi mettere in atto strutture, sia di formazione sia di controllo, in cui è possibile introdurre input statistici al di fuori delle singole collaborazioni o dei corsi standard. Nuovi modelli di insegnamento che mettono in primo piano i problemi inferenziali possono anche insegnare e ispirare la prossima generazione di statistici e scienziati a implementare dei cambiamenti.

Bisogna così creare percorsi di promozione presso le istituzioni per valorizzare e premiare i docenti che sono efficaci in queste attività, oltre alla tradizionale ricerca metodologica. Con incentivi e ricompense professionali alcuni statistici potrebbero avere più tempo da dedicare al miglioramento della pratica statistica, e i loro collaboratori potrebbero provare meno pressione a "giocare" con gli strumenti d' inferenza.

2.1.2 Giornali

Le riviste specializzate sono un ottimo posto in cui possono essere dimostrati diversi approcci, poiché possono modellare le norme della disciplina. Gli esempi pubblicati che utilizzano

approcci inferenziali non di significatività dell'ipotesi nulla (NHST: null hypothesis significance testing) sono estremamente potenti. I giornali che utilizzano questi metodi, oltre ad illustrare come questi debbano essere implementati, saranno infatti fortemente presi in considerazione.

Publicare su riviste disciplinari raggiunge anche molti insegnanti e professionisti i quali non sono esperti di statistica ma che lavorano all'interno di discipline che sono interessate alle procedure e agli sviluppi dei vari approcci. Un'idea potrebbe essere quella di scrivere e pubblicare commenti sulla metodologia di ricerca, sottolineando la presenza di una o più pratiche statistiche differenti, al fine di evidenziare errori commessi nelle pratiche condannate.

2.1.3 Finanziatori

Il focus dei finanziatori della ricerca dovrebbe essere incentrato non tanto sullo sviluppo di nuovi metodi, ma per formare con più strumenti la prossima generazione, per produrre codice user friendly per implementare alternative ai valori p, per sviluppare nuovi curricula e strumenti pratici di apprendimento e nuovi modi di divulgare e modellare la pratica inferenziale.

I finanziatori più influenti richiedono determinati livelli di rigore metodologico nella ricerca che sovvenzionano, ma la trasformazione delle raccomandazioni in politiche attuabili è difficile. Un primo esempio consiste nel presentare ogni studio alle riviste durante le fasi iniziali del progetto per essere considerato per la pubblicazione molti anni dopo. È chiaro che dal punto di vista operativo, c'è un forte consenso tecnico dietro di esso, l'attività target è definita, mentre le conseguenze sono esplicitate. Questa preregistrazione ha avuto l'effetto di aumentare la percentuale di risultati non significativi nelle pubblicazioni degli studi e, forse, di diminuire indirettamente l'importanza percepita dei risultati significativi come criterio per la presentazione o l'accettazione.

Quando diciamo che le agenzie di finanziamento hanno il potere di migliorare la qualità della ricerca, sia in termini tecnici sia in termini di inferenza, dobbiamo specificare come deve essere esercitato tale potere. Lo stesso finanziatore può avere delle lacune dal punto di vista conoscitivo per quanto riguarda ciò che sta andando a finanziare. In questo caso sarà quindi difficile individuare quali "violazioni" del processo siano da migliorare, non arrivando dunque con certezza ad una progettazione definitiva e certa di quest'ultimo.

2.2 Un corso di pensiero statistico

Gli statistici sono d'accordo nel riscontro di imperfezioni nell'applicazione della statistica nella scienza, un po' meno invece su come risolverle. L'idea di Steel (ricercatrice statistica presso "Pacific Northwest Research Station"), Liermann (ricercatore statistico presso "Northwest Fisheries Science Center") e Guttorp (professore presso "University of Washington") consiste nello sviluppo e nell'insegnamento di un corso di statistica universitario, capace di spiegare a pieno l'argomento in questione e di creare un collegamento col mondo della ricerca.

2.2.1 Insegnare la saggezza statistica

I risultati ottenuti da un processo di ricerca come abbiamo detto soffrono di alcuni errori, tra cui: difficoltà di replica, domande alla base della ricerca di scarso valore, utilizzo di un numero spropositato di test, il dar troppo peso ai valori p, mancanza di peer review. Per revisione paritaria (peer review) intendiamo la procedura di selezione degli articoli o dei progetti di ricerca proposti da membri della comunità scientifica, effettuata attraverso una valutazione di specialisti del settore che ne verificano l'idoneità alla pubblicazione su riviste specializzate.

Gli autori, per risolvere queste problematiche, si rifanno a Brown e Kass (2009). La statistica non può essere solamente un mero processo di calcolo, si dovrà così porre domande chiare all'inizio del processo scientifico, mentre alla fine sarà importante comunicare risultati e tutte le fasi attraverso cui si è passati.

L'idea è di fornire un corso universitario che sia in grado di spiegare opportunità ed insidie dell'applicazione della statistica all'interno dell'intero sistema. La "saggezza" nel capire quali siano gli errori deve essere insegnata nella maniera più coerente possibile col contesto di riferimento, al fine di formare nuovi statistici e ricercatori.

2.2.2 La struttura del corso

Il corso, come evidenziato dalla tabella 2.1, spazia da come "lavora" il cervello umano alla comunicazione statistica. Per affrontarlo non serviranno prerequisiti obbligatori, ma è naturale che per una miglior comprensione servirà conoscere alcune procedure statistiche (come il test

t o la regressione lineare). Elementi caratterizzanti sono la divisione dei componenti della classe in gruppi non omogenei (per esempio unendo studenti provenienti da diversi corsi) e le frequenti simulazioni dirette in aula (come per esempio estrazione di palline numerate da urne), con conseguente discussione dell'evento, accompagnando il tutto con un laboratorio informatico. Gli autori hanno poi constatato che gli studenti migliori non sono per forza quelli con molte conoscenze procedurali.

2.1 La struttura del corso

Topics	Detailed Understanding
Errors in thinking	How the wiring of the human brain leads to incorrect conclusions from data.
Clear questions	Quantifying the metric of interest in scientific questions.
Simulations	Using simulations to understand the behavior and power of statistical procedures.
Estimation	The advantages of estimation over testing for many common situations.
The p-value	Understanding what it is, what it tells us, and what it does not tell us.
Repeated experiments	Why results from a single study are inconclusive and how the scientific process is iterative.
The science process	Pitfalls and opportunities of the project life-cycle and peer-review process.
Study design	Principles of good study design and how poor design can lead to incorrect conclusions.
Sample size	Issues with small samples and with big data. How these both can lead to incorrect conclusions.
Communication	Principles of scientific communication and how they apply to statistics.

Fonte: tandfononline.com

2.2.3 Le sezioni del corso

Sezione 1-Making sense of data

La sezione uno incentra le sue attenzioni sul funzionamento del cervello umano e sulle distribuzioni campionarie (insieme di probabilità di una statistica). Verranno forniti agli studenti una serie di esempi per dimostrare come noi, esseri umani, siamo ingannati da schemi e coincidenze, al fine così di aumentare la consapevolezza dell'individuo, riducendo la possibilità d'errore. Si farà ampio ricorso a simulazioni, offrendo agli studenti una panoramica del software R, il quale oltre ad essere un software open scaricabile gratuitamente, è un linguaggio di programmazione e un ambiente di sviluppo specifico per l'analisi statistica dei dati.

Sezione 2-Using data for scientific understanding

L'obiettivo di questa sezione consiste nell'accumulo d'esperienza dello studente. Attraverso letture ed esercizi, aumentando così la propria comprensione rispetto al p-value, si insegnerà quando i test statistici sono realmente importanti, dando la possibilità ai ragazzi/e di crearne di propri per rispondere alla domanda d'interesse. Si andrà poi ad enunciare come la scienza passi da un esperimento in laboratorio fino a diventare una realtà accettata da tutti, e come il processo di peer review possa influenzare l'applicazione statistica.

Sezione 3-Using science in society

Il focus della sezione è rivolto all'interpretazione delle affermazioni scientifiche, la comunicazione del rischio (Definita da Giancarlo Sturloni come: "L'insieme degli scambi informativi tra i diversi portatori di interesse che prendono parte alla discussione pubblica sui rischi per la salute e per l'ambiente"), la comunicazione con e attraverso i media, le applicazioni di approfondimento statistico su temi delicati come i cambiamenti climatici e la capacità (o la non capacità) di fatti scientifici di manipolare la mente delle persone su argomenti controversi.

2.2.4 Il progetto finale

Durante la prima metà del corso vengono assegnati, come abbiamo già accennato in precedenza, letture ed esercizi con R, mentre nella seconda metà ci sarà un minor quantitativo di compiti, a favore di lavori di gruppo ed un progetto finale per verificare le nuove competenze e conoscenze acquisite dagli studenti rispetto alla conduzione di test statistici, alla capacità di trarre conclusioni dai dati e alla comunicazione dei risultati ottenuti.

Il progetto finale potrà essere una presentazione orale, un articolo di giornale/rivista, post sui social media. Questi lavori di gruppo verranno applicati ad una situazione reale per la quale si cerca realmente una risposta, e sebbene ognuno di questi sia concepito per coprire l'intera gamma di attività coinvolte nell'utilizzo dei dati per rispondere a una domanda, ogni progetto enfatizzerà un particolare elemento del corso, dal pensiero imperfetto fino alle sfide della comunicazione.

Durante la settimana finale verranno ordinati i progetti di gruppo in modo da fornire una revisione abbastanza organizzata dei messaggi e delle idee chiave del corso. Mentre l'ultimo

giorno verrà fornito un promemoria (tabella 2.2) su come condurre le analisi scientifiche, in coerenza con ciò che è stato spiegato durante il periodo di studi.

2.2 Come condurre l'analisi scientifica

Plot your data—early and often.
Understand your dataset as one of many possible sets of data that could have been observed.
Understand the context of your dataset—what is the background science and how were measurements taken.
Be thoughtful in choosing summary metrics.
Decide early which parts of your analysis are exploratory versus confirmatory and pre-register your hypotheses in your own mind.
If you are going to use p -values, which can be useful summaries when testing hypotheses, follow these principles:
<ul style="list-style-type: none"> • Report estimates and confidence intervals;
<ul style="list-style-type: none"> • Report the number of tests you conduct (formal and informal);
<ul style="list-style-type: none"> • Interpret the p-value in light of your sample size (and power);
<ul style="list-style-type: none"> • Don't use p-values to claim that the null hypothesis of no difference is true;
<ul style="list-style-type: none"> • Consider the p-value as one source of support for your conclusion not the conclusion itself.
Compute (and display) effect sizes and confidence intervals as an alternative to or in addition to statistical testing.
Consider creating customized, simulation-based statistical tests for answering your specific question with your particular dataset.
Use simulations to understand the performance of your statistical plan on datasets like yours and to test various assumptions.
Read with skepticism, remembering that pattern can easily occur by chance (especially with small samples), and that unexpected results based on small sample sizes are often wrong.

Fonte: tandfononline.com

2.2.5 Caratteristiche insolite del corso

Rispetto ad un tradizionale corso di statistica, il corso offerto in “statistical thinking” presenta notevoli differenze. Per prima cosa si basa sull’idea che le maggiori problematiche della statistica non derivino dai dettagli delle tecniche utilizzate (normalità, varianza...), ma dall’applicazione di strumenti statistici, dalla probabilità di trovare modelli validi solo per caso e la tendenza a sopravvalutare nuovi risultati. Il corso sarà multidisciplinare, e affronterà temi raramente trattati nelle lezioni classiche di statistica, come la psicologia umana, la comunicazione e la peer review.

C'è poi la volontà di creare un ponte tra conoscenza e applicazione statistica in un contesto scientifico, cercando di capire come quando e perché l'utilizzo degli strumenti funziona o meno. Questo grazie anche alla condivisione di esperienze e punti di vista tra studenti universitari, processo che garantirà un maggior apprendimento e migliori risultati ottenuti. Verranno condotti esperimenti sul p-value e tema importante sarà la discussione della crisi della replica dei risultati.

Oltre ad affrontare il disordine durante la fase di analisi dei dati, circa un 20% del corso sarà dedicato alla comunicazione statistica, tra le quali attività evidenziamo la valutazione dei comunicati stampa, visione di video di esperti di statistica che spiegano le esigenze di questo tipo di comunicazione, e spiegazione di come i social media aumentino la possibilità per gli scienziati di sopravvalutare le loro affermazioni.

2.2.6 Il p-value e i campioni a confronto

Le simulazioni come abbiamo detto sono uno strumento di forte utilità per la costruzione dell'intuito, per la creazione di nuovi approcci di valutazione e test personalizzati. Tramite queste si valuteranno i p-value in più situazioni, tanto è vero che, oltre a dare una breve panoramica storica del valore p durante la quarta settimana, per cercare di risolvere le problematiche legate a quest'ultimo, gli studenti condurranno come compito per casa un esperimento a scelta (5 possibilità). Gli obiettivi principali di questa attività sono quelli di intuire come il p-value si comporta quando l'ipotesi nulla è vera, distinguere tra benefici statistici e pratici, capire come la grandezza del campione influenzi la probabilità di rigetto di H_0 , la capacità di stimare la probabilità che H_0 sia vera quando c'è stato o quando non c'è stato rifiuto. Gli studenti riporteranno i risultati ottenuti in classe, e così facendo si potrà apprendere da tutti e cinque gli esperimenti proposti. Per finire ci sarà un'ampia discussione su come comunicare agli scienziati, ai giornalisti o ad altri utenti interessati, i risultati dell'analisi statistica in corrispondenza di p-value più o meno elevati.

È poi importante sottolineare come il corso preveda di limitare il confronto degli esperimenti a due campioni provenienti da due popolazioni, in quanto non c'è effettivamente troppo tempo per allargarne il numero (un semestre). Ma da svantaggio possibile diventa una caratteristica stessa delle lezioni in quanto verranno create simulazioni gestibili, capaci, tramite l'utilizzo di stime e grafici, di riscontrare parallelismi ed evidenziare differenze tra i campioni stessi.

2.3 La Guinnessometria di Ziliak

Stephen T. Ziliak è un professore americano di economia le cui ricerche e saggi spaziano dalla statistica fino alla produzione di birra. Al contrario degli autori citati in precedenza, Ziliak avvicina maggiormente la problematica decisionale ad un contesto economico. Gli attuali progetti comprendono la Guinnessometria, cioè un ripensamento all'ingrosso della filosofia sperimentale e della pratica econometrica.

Basandosi su ricerche d'archivio fatte alla Guinness di Dublino, Ziliak propone dieci "valori G" che egli ritiene tutti desiderino massimizzare nella ricerca. L'obiettivo è di ottenere grandi G-value e non piccoli p-value, valori sintetizzati nei 10 principi della "Guinnessometria", i quali derivano principalmente dalla disamina degli esperimenti condotti dallo statistico William S. Gosset (1876-1937), l'inventore della "t di Student", mentre lavorava come Head Brewer per la Guinness (immagine 2.3).

2.3 Guinness



Fonte: ambientebio.it

La Guinness è una birra scura ad alta fermentazione (tipo stout) prodotta dalla Arthur Guinness Son & Co., una fabbrica di birra irlandese fondata a Dublino nel 1759 da Arthur Guinness nella celebre St. James's Gate Brewery.

2.3.1 I 10 principi "G"

G-10: considera lo scopo dell'inchiesta e confrontala con le migliori pratiche

La falsificazione di un'ipotesi nulla non è l'obiettivo più importante di un esperimento o di uno studio. Fare soldi o birra, magari meglio delle migliori pratiche e della concorrenza, lo è! Così come la stima dei propri coefficienti rispetto ai risultati riportati dagli altri. In poche

parole, bisogna evitare di qualificare un'ipotesi con semplici no/sì, esiste/non esiste, significativo/non significativo.

G-9: stima la posta in gioco (quantità di perdita o guadagni)

La stima della grandezza degli effetti, e la dimostrazione dei loro significati, dovrebbe essere al centro delle inchieste. La posta in gioco di uno studio implica grandezze attese di guadagni e perdite espressi sotto forma di funzione, ed il loro effettivo ammontare è centrale nell'approccio Guinnessometrico. Il valore atteso di un esperimento o di un set di osservazioni non è affidabile, e non è una misura certa per valutare il comportamento umano, nonostante molti scienziati vadano ad utilizzarlo come se lo fosse. Quindi, tenendo conto della variabilità comportamentale dell'individuo, gli economisti e gli statistici tendono a prediligere il concetto di utilità attesa rispetto al valore atteso.

Es: lancio casuale di una moneta. Confronto tra Valore Atteso ed Utilità Attesa

Gioco 1: se esce testa, vincita \$100; se esce croce perdita \$0,50

Gioco 2: se esce testa, vincita \$200; se esce croce perdita \$100

Gioco 3: se esce testa, vincita \$20000; se esce croce perdita \$10000

Assumendo che la probabilità che esca testa sia uguale alla probabilità che esca croce, il valore atteso sarà dato dalla differenza della potenziale vincita e della potenziale perdita moltiplicate per la loro probabilità di accadimento, per cui avremo: $\frac{1}{2}(\$vincita) - \frac{1}{2}(\$perdita)$. Vediamo dunque che il gioco 1 ha il valore atteso più basso, ed è pari a \$49,75, a fronte di uno scenario di eventuale perdita di soli \$0,50. Il gioco 3 invece dovrebbe essere il preferibile, in quanto ha il valore atteso più elevato, pari a \$5000. Proprio qui però sta la critica rivolta a questo sistema decisionale, infatti non verrà tenuta in considerazione l'attitudine del soggetto. Prendendo ad esempio il gioco 3, questo garantirà una potenziale vincita elevata, ma c'è il rischio di uno scenario disastroso in caso di perdita (\$10000). Molti giocatori preferiranno quindi "rifugiarsi" nei primi due giochi.

Supponiamo ora che i soggetti vogliano massimizzare l'utilità piuttosto che il valore atteso. Questa teoria ha anch'essa dei limiti, ma offre qualcosa in più, ed è inoltre più sensata in un contesto decisionale. L'utilità prevista è una scala di valori che tiene conto, tra le altre cose,

dell'atteggiamento nei confronti del rischio e della diminuzione dell'utilità marginale del denaro. Supponiamo di essere una persona avversa al rischio con una funzione di utilità strettamente concava, come la funzione radice quadrata nell'accumulo di denaro ($EU = \sqrt{\text{denaro}}$).

Continuando l'esempio, e supponendo di avere un ammontare iniziale di soldi pari a \$10000, vogliamo sapere qual è l'utilità attesa dei tre giochi. La formula per il calcolo di quest'ultima corrisponde a: $\frac{1}{2}(\sqrt{\text{denaro iniziale} + \text{vincita}}) + \frac{1}{2}(\sqrt{\text{denaro iniziale} - \text{perdita}})$. Quindi l'utilità prevista per il primo gioco è massima, anche se il suo valore atteso è il più basso dei tre, ed è pari a $\frac{1}{2}(\sqrt{10.100}) + \frac{1}{2}(\sqrt{9999.50}) = \100.248 . L'utilità invece del gioco 3 è la più bassa, ed è pari a \$86.60, anche se il suo valore atteso è decisamente il più elevato.

Vediamo come ragionando con l'utilità attesa si faranno delle scelte decisamente diverse e maggiormente ponderate. Il gioco 1 è diventato così la scelta preferibile per un soggetto avverso al rischio che presta elevata cautela decisionale.

G-8: studio dei dati correlati e stratificazione della popolazione

Molti modelli di regressione assumono errori indipendenti ed identicamente distribuiti (iid), eppure la maggior parte dei dati nelle scienze sociali e della vita sono correlati da effetti sistematici non casuali, cioè non sono indipendenti.

IID-nella teoria della probabilità, una sequenza di variabili casuali è detta indipendente e identicamente distribuita se:

- le variabili hanno tutte la stessa distribuzione di probabilità
- le variabili sono tutte statisticamente indipendenti

es: a parità di condizioni una sequenza di risultati dei giri di una roulette è i.i.d. Da un punto di vista pratico, una conseguenza è che se la pallina della roulette cade sul rosso 20 volte consecutive, il successivo giro non ha maggiori o minori possibilità di cadere sul nero di qualsiasi altro giro.

Per correlazione in statistica intendiamo una relazione tra due variabili tale che a ciascun valore della prima corrisponda un valore della seconda, seguendo una certa regolarità. Per ovviare a questo problema gli statistici sanno che la stratificazione aumenta la precisione e l'efficienza di uno studio, rispetto al caso in cui quest'ultimo sia basato su una randomizzazione semplice. La stratificazione inoltre implica una minore varianza (e di

conseguenza un minor errore) all'interno di ogni strato, rispetto alla varianza complessiva della popolazione. Questo aumenterà la precisione della stima che si otterrà dallo studio del campione.

Randomizzazione semplice: estrazione di una certa quota di unità dalla popolazione attraverso un metodo che garantisce la casualità del campionamento.

Randomizzazione stratificata: popolazione suddivisa in strati (più gruppi) sulla base di un fattore capace di influenzare il carattere oggetto di studio. Successivamente selezione con metodo randomizzato di un campione all'interno di ciascuno strato.

Gosset utilizzò la stratificazione molto prima che venne coniato questo termine. Il problema era che il suolo su cui venivano effettuati gli esperimenti da parte della Guinness non era uniforme. Tramite la divisione del terreno in blocchi (strati) maggiormente omogenei e limitando l'analisi all'interno di essi, Gosset ottenne maggior precisione nei suoi studi.

Questa stratificazione torna utile, dal punto di vista economico e statistico, quando si conosce la popolazione, come nel caso dei terreni della Guinness, mentre la randomizzazione semplice ha senso quando il ricercatore sa poco o nulla dei possibili strati.

G-7: minimizza l'“Errore Reale” con le 3 R: rappresenta, replica, riproduci

Un test di significatività in un singolo set di dati è quasi privo di valore. La F di Fisher, la t di Student ed altri test, dovrebbero essere usati quando c'è una vera ripetizione dell'esperimento. Errore di misurazione, confusione e pregiudizio dei risultati auspicabili sono frequentemente presenti in tutte le prove scientifiche. La guinnessometria minimizza l'errore reale ripetendo le prove su unità indipendenti derivanti da un campionamento stratificato, controllando il più possibile gli effetti fissi locali.

Riproduzione: uno studio è riproducibile se è possibile prendere i dati originali e il codice computer utilizzato per analizzarli e riprodurre tutti i risultati numerici. Inizialmente questo potrebbe sembrare un compito banale, ma l'esperienza ha dimostrato che non è sempre facile raggiungere questo standard apparentemente minimo. Per Gosset la riproducibilità è la capacità di riprodurre la Guinness stout o ale in modo tale da gustare la stessa birra in tutto il mondo.

Replicabilità: questo è l'atto di ripetere un intero studio, indipendentemente dall'investigatore originale, senza l'uso di dati originali (ma generalmente usando gli stessi metodi). In 35 anni

di produzione e prove di qualità, Gosset e Guinness hanno commissionato esperimenti con orzo irlandese sparsi nelle diverse regioni in crescita del paese.

Rappresentazione: copertura e stratificazione di tutte le fonti sistematiche di fluttuazione, indipendentemente dal fatto che la fonte sia una forza temporale, spaziale o di altra natura esogena (come il tempo: piovoso o non piovoso, o la qualità del suolo: argilloso o meno). Non si tratta di un errore casuale: questo è il punto principale. Si devono minimizzare e controllare gli errori sistematici, che sono di importanza maggiore e il numero di quelli che si suppone siano "errori di campionamento casuale".

Dal 1901 fino all'inizio della seconda guerra mondiale, Guinness (e quindi Gosset) ha investito molto nelle 3 R dei valori G: cioè ripetendo ogni anno una piccola serie di esperimenti indipendenti, rappresentativi ed equilibrati (che i sociologi chiamano ripetuti e stratificati). I risultati positivi della Guinness e degli investimenti irlandesi furono a dir poco sbalorditivi.

G-6: economizzare con "Less is More": piccoli campioni di esperimenti indipendenti

L'analisi e la distribuzione teorica di piccoli campioni hanno un'origine ed un fondamento economico: cambiare gli input per la birra su larga scala (per Guinness, enorme scala globale) è rischioso, tenendo conto del denaro messo in gioco. Ma piccoli campioni, come Gosset ha mostrato in decenni di esperimenti su luppolo ed orzo, non significano "meno di", e i Big Data non sono in ogni caso la soluzione di molti problemi.

I piccoli campioni sono legati alla scarsità delle risorse, in relazione alle spese per l'ottenimento di nuove informazioni su orzo, malto, luppolo e altri input di birra. Quindi, si può dire in generale che Gosset ha adottato un approccio economico alla logica dell'incertezza, partendo dalla dimensione del campione.

G-5: tieni gli occhi aperti sulle domande "importanza della dimensione" e "quanto"

Nonostante le distrazioni, le funzioni di perdita e profitto attese devono regolare sul da farsi, o almeno dovrebbero. I coefficienti di regressione o le differenze tra loro sono grandi o piccole? Rispetto a cosa?

Il consenso su una semplice direzione di effetto su o giù, positivo o negativo è abbastanza raro nella scienza. Dopo secoli e secoli di assistenza pubblica per i poveri negli Stati Uniti e

nell'Europa Occidentale, ad esempio, gli economisti non sono concordi sull'effetto esercitato sull'offerta di lavoro dai sussidi al reddito finanziati dalle tasse. Il disaccordo preso in esame, così come per la dimensione dell'effetto, è più una regola che un'eccezione.

Il punto cinque della guinessometria ci dice di tracciare una linea di effetto minimamente importante o grandezza minimamente importante di un coefficiente di regressione, da tenere come punto di riferimento, valutandone poi la dimensione.

G-4: visualizzare

L'incertezza dei parametri non è la stessa cosa dell'incertezza del modello. Il risultato colpisce? Lo studio mostra grandi quantità di effetti sull'intera distribuzione? Un aiuto in questo caso è dato dai progressi nei software di visualizzazione. Uno studio di Soyer e Hogarth (2011) ha testato la capacità predittiva di oltre 200 economisti usando modelli lineari. La previsione era più accurata quando agli esperti veniva fornita una linea di regressione teorica e un diagramma di dispersione dei dati.

La varianza può variare, i limiti di errore variano, il grado e la direzione dell'incertezza variano. Le immagini ci aiutano a capire meglio l'effettiva grandezza di queste variazioni. Come Gosset, dovremmo prestare maggiore attenzione alla distribuzione effettiva dei nostri dati, così come anche a simulazioni o altre conferme di questa, e non solo ai rifiuti della distribuzione normale.

G-3: considera anche il "prima" e il "dopo"

Il campione a portata di mano è raramente l'unica cosa che è "conosciuta". L'esperienza in materia è un importante contributo preliminare alla progettazione statistica che andrà a influenzare l'analisi dei risultati "posteriori". Ad esempio, Gosset alla Guinness è stato saggio a mantenere le metriche di assicurazione della qualità e il profitto di base al centro della sua indagine.

G-2: cooperare su, giù e attraverso (reti e catene di valore)

Dove sarebbero oggi i birrai senza la continua cooperazione con gli agricoltori? Forse si tornerebbe alla fattoria e non al birrifico. La scienza statistica è sociale e la cooperazione aiuta, tanto che la Guinness finanziò una grande parte della moderna teoria statistica inviando

birrai scientifici (e quindi non solo Gosset) in un periodo sabbatico per gli studi post-laurea in statistica, chimica e genetica.

Un altro esempio è stata la condivisione dei semi e del know-how tecnico con la fabbrica di birra Carlsberg, la quale ricambiò generosamente offrendo un orzo stellare chiamato "Prentice" alla Guinness, che si è rivelò la più importante tipologia in quel periodo in Europa.

G-1: rispondere alla domanda originale del birrificio ("Come dovresti impostare le probabilità? ")

Nessuna regola precisa di rilevanza statistica può rispondere alla domanda del birraio. Come disse Gosset nel lontano 1904, il modo in cui si impostano le probabilità dipende "dall'importanza dei problemi in gioco". Il grado di certezza da perseguire deve dipendere dal vantaggio pecuniario che si può ottenere seguendo il risultato dell'esperimento, rispetto all'aumento dei costi d'implementazione del nuovo metodo. Questo punto in particolar modo può essere un riferimento per tutte le imprese che intendono investire in progetti al fine di migliorare la propria attività. È quindi di fondamentale importanza valutare il trade off costi-benefici.

3 PROCESSO DECISIONALE E RUOLO DEL MANAGER IN UN CONTESTO D' INCERTEZZA

L'incertezza e la variabilità di cui abbiamo parlato nelle pagine precedenti possiamo ricondurle ad un contesto economico e in particolar modo al processo decisionale. Come già annotato nel corso della tesi il metodo statistico adottato a livello universitario è di scarsa applicazione nelle scienze economiche, ovvero in un contesto altamente mutevole dove servirà flessibilità e rapidità nelle scelte che incidono nella vita, negli obiettivi e nel vantaggio competitivo delle organizzazioni.

La figura chiamata a svolgere l'attività decisionale è quella del manager; prendere decisioni di diversa natura tra un ventaglio di possibilità è parte del suo lavoro quotidiano. Anche qui come per la statistica vediamo che per prendere una decisione definitiva si segue, in partenza, un percorso schematico non esente da critiche e limitazioni.

3.1 Fasi del processo decisionale

Introduciamo lo schema meccanico alla base del processo decisionale, che come vedremo, risulterà solo come punto di partenza per un sistema che raccoglierà molti più elementi. La similitudine con Fisher e la statistica classica è evidente, la variabilità e la rilevanza dei Big Data non possono che implicare dei cambiamenti ai metodi sistematici e obsoleti a favore della creatività e capacità decisionale all'interno di contesti turbolenti. Quindi come per la significatività statistica, lo schema del processo decisionale non verrà abbandonato, ma verrà arricchito con elementi di supporto che garantiranno maggiore adeguatezza e consapevolezza. Le fasi in questione sono:

1. Identificazione del problema

Risulta fin da subito difficile capire cosa sia un problema, infatti nella realtà non si presenta con immediatezza e certezza. Non manca l'elemento di soggettività, ciò che può essere considerato "problema" per un manager potrebbe non esserlo per un altro. Identificare un problema è importante, ma non facile.

2. Identificazione e attribuzione del peso ai criteri di decisione

È necessario indicare i criteri più importanti su cui basare una decisione. Questi variano anche dalla natura e dalla complessità della situazione che si deve risolvere. Una scelta di outsourcing esterno di natura strategica sarà sicuramente più delicata rispetto a contattare e ad esternalizzare le pulizie dell'azienda ad un'impresa specializzata (in questo caso potrei attenermi semplicemente a chi offre il prezzo più basso). Una volta selezionati i criteri bisognerà attribuirli un peso, cioè capire quali tra questi sono più rilevanti, al fine di fare una graduatoria privilegiando in fase di attuazione quelli con maggior rating.

3. Sviluppo delle alternative

Questa fase presuppone una lista di alternative strategiche perseguibili da parte del manager, che porteranno alla soluzione del problema. Fase di pura elencazione e non di valutazione, lascia spazio anche alla componente creativa.

4. Analisi delle alternative

Il manager dovrà valutare le alternative tramite i pesi dei criteri indicati nella fase tre. La capacità di analisi e di disamina di differenti opportunità di scelte e le conseguenze che ne derivano sono parte dell'attività quotidiana di un manager.

5. Scelta dell'alternativa

Scelta sulla base delle valutazioni effettuate nella fase precedente. La figura del manager è un fulcro decisionale, in quanto è una figura professionalizzata all'interno dell'azienda con elevate capacità non solo pratiche ma anche culturali e relazionali, pertanto è suo dovere prendere una decisione, nonostante non si possa essere certi del tutto del futuro e se realmente la scelta effettuata sia corretta.

6. Attuazione dell'alternativa scelta

Fase di implementazione della scelta effettuata da part del management e non solo. I collaboratori operanti all'interno dell'organizzazione sosterranno il processo nel caso in cui questi siano stati presi in causa in maniera attiva durante gli step decisionali. Sarà poi da tenere in considerazione il cambiamento continuo degli ambienti di riferimento durante la fase di attuazione, soprattutto per quanto riguarda decisioni di lungo periodo. In caso di mutazione significativa, la stessa decisione dovrà essere rivalutata ed eventualmente modificata in corso d'opera.

7. Valutazione dell'efficacia della decisione

La fase di valutazione porta a capire se l'alternativa scelta ha condotto alla soluzione del problema. E se il problema fosse ancora presente? Non è stato definito correttamente? Sono presenti errori nella fase di valutazione? O nella fase di attuazione? Potrebbe essere necessario così ricominciare l'intero processo.

3.2 Dalla razionalità all'intuito

L'idea diffusa è che le scelte effettuate dai manager siano razionali, logiche e coerenti rispetto al raggiungimento del massimo risultato, avendo a disposizione strumenti e tecniche che permettono loro di mantenere questa linea di condotta. In questo caso il problema da affrontare dovrebbe essere chiaro e non ambiguo, l'obiettivo da raggiungere dovrebbe essere ben formulato conoscendo tutte le possibili alternative e adottando quella che massimizza la probabilità di perseguirlo. Ma nella realtà questa razionalità manageriale è irrealistica. Un buon decisore deve adottare validi comportamenti decisionali valutando problemi e analizzando le alternative. Decisioni prese tramite un processo "intelligente" fanno sì che il manager possa essere considerato competente. Vediamo dunque che modelli razionali come quelli statistici potranno tornare utili ma non saranno sufficienti per prendere delle decisioni.

3.2.1 Esempio di razionalità: i modelli statistici

I modelli statistici, che rientrano nella sfera dei modelli matematici, sono in genere distinti tra modelli statistico inferenziali e modelli statistico descrittivi. Questi ultimi partendo da un ridotto numero di dati originari, hanno lo scopo di fornire il massimo grado di informazione possibile. Quelli inferenziali sono invece utilizzati per colmare le mancanze di dati utili per decisioni razionali, attraverso la formulazione di ipotesi.

L'applicazione dei modelli a problemi d'impresa ed economici in generale presenta dei limiti, e la sola osservazione statistica non basta a interpretare in modo chiaro i dati analizzati. È quindi ragionevole combinare al processo statistico il ragionamento individuale umano.

Nella fase decisoria, alcune tecniche statistiche, possono aiutare il manager. In particolare, se le scelte fanno riferimento ad orizzonti temporali di lungo periodo, disporre di informazioni

ulteriori garantite dai modelli in questione, potranno migliorare il grado di attendibilità del verificarsi di un dato evento. Nonostante ciò la certezza perfetta non esiste, per cui bisognerà valutare costi-benefici dell'acquisizione di una determinata informazione. Secondo la ricerca operativa, infatti, bisogna analizzare il valore dell'informazione calcolando il massimo vantaggio previsto senza di essa e confrontandolo con il massimo vantaggio in sua presenza. Sarà così opportuno valutare se questa incrementi o meno l'attendibilità degli eventi futuri.

3.2.2 La razionalità limitata

Un concetto interessante, che delucida la questione, è quello della razionalità limitata, il quale afferma che il manager agisce razionalmente nei limiti della sua capacità di trattamento delle informazioni. Non conoscendo nel complesso tutto ciò che caratterizza le alternative (impossibile archiviare ed analizzare tutti i dati), il processo decisionale si concluderà con una scelta soddisfacente. Il soggetto sarà quindi razionale nei limiti della sua abilità di elaborare ed immagazzinare le informazioni.

Nonostante tecniche e metodi all'avanguardia, in alcune situazioni risulta complesso arrivare ad una decisione definitiva. Ciò che guida il manager di fronte ad un dubbio decisionale è l'intuito, che i ricercatori hanno suddiviso in cinque differenti aspetti:

- Decisioni basate su valori etici
- Decisioni basate sull'esperienza
- Decisioni basate sui sentimenti
- Decisioni basate sulla conoscenza
- Decisioni basate su processi mentali subconsci

L'intuito potrà così risultare un plus utile al processo decisionale. Prendiamo per esempio un manager che abbia elevata esperienza con un particolare tipo di problemi, questo potrà agire con elevata rapidità nonostante le poche informazioni in suo possesso. Queste ultime infatti saranno di rilevante importanza e garantiranno capacità decisoria, in quanto le situazioni in questione, cui è posto davanti il soggetto, sono già di sua conoscenza. Non verrà poi trascurata la parte emotiva dell'individuo, la quale può garantire maggior coinvolgimento ed impegno nella fase di decisione. Oggi non si potrà pensare di affidarsi alle vecchie concezioni totalmente lineari, ma si dovrà utilizzare anche quella parte "meno logica" di sé stessi.

3.3 Caratteristiche delle decisioni

Vediamo nei paragrafi che seguono come la fase decisoria sia influenzata da una serie di variabili. Le decisioni infatti non sono tutte uguali, sono più o meno semplici da prendere, possono essere prese in diverse situazioni, possono essere il risultato di un ragionamento soggettivo piuttosto che frutto di un processo razionale e possono presentare errori derivanti da convinzioni personali.

3.3.1 Tipologie di decisione

Ci sono delle decisioni che vengono definite programmate, che possono essere gestite secondo un processo routinario, poiché il problema è strutturato, chiaro e con informazioni definite e complete. Non servirà così fornirsi di un ventaglio di alternative in quanto la soluzione apparirà evidente. Per un problema strutturato ci si dovrà avvalere di adeguate procedure, le quali risulteranno sequenziali (es: procedura d'acquisto eseguita ogni nuovo ordine), attraverso regole e linee di condotta congrue a mantenere coerenza nell'agire.

Il problema reale è che la moltitudine di situazioni che si manifestano in ambito aziendale riguardano problemi non strutturati, ovvero problemi non usuali per i quali non si dispone di informazioni sufficienti e chiare (es: ampliare il mercato in un nuovo stato). Il manager dovrà così adottare una decisione non programmata (cioè unica per la risoluzione di un problema non ricorrente), tramite lo sviluppo di un processo che segua le fasi citate in precedenza ed arrivando ad una soluzione finale.

Le principali differenze tra programmate e non, sono la ripetitività ed il meccanismo della decisione. Le prime sono solitamente di competenza di manager di medio-basso livello, mentre le seconde sono di pertinenza dei top manager. Nella realtà sarà difficile definire una situazione totalmente univoca, molti problemi infatti richiedono soluzioni che contemplano sia decisioni programmate sia non programmate.

3.3.2 Le condizioni delle decisioni

La condizione più "rara" che si può trovare davanti un manager è quella di certezza, in questo caso si può assumere decisioni accurate in quanto il risultato di ogni alternativa è conosciuto

(es: tasso di interesse che la banca applica per un'operazione, il quale è noto, e si può basare così la decisione su un elemento di certezza).

Una situazione assolutamente più comune è quella di rischio, dove il manager riuscirà a stimare la probabilità solo di certe variabili. Mediante l'utilizzo di dati storici di precedenti esperienze e informazioni secondarie, verranno valutate differenti alternative.

Quando invece non si è certi dei risultati o non si possono effettuare stime di probabilità, si è in una condizione d'incertezza. Sarà disponibile un limitato numero d'informazioni e ci si affiderà all'intuito e agli aspetti psicologici del manager.

3.3.3 Stili ed errori decisionali

Il manager durante la fase di decisione può adottare due tipologie di stile: quello lineare è caratterizzato da una preferenza personale per l'utilizzo di risorse esterne, mentre i dati e i fatti verranno trattati attraverso un modello di pensiero razionale-logico al fine di prendere una decisione; lo stile non-lineare invece è caratterizzato dalla preferenza per le fonti interne di informazione (parte emotiva del soggetto), da cui deriveranno i dati che verranno trattati attraverso le conoscenze e l'intuito del manager per poi assumere delle decisioni e delle azioni conseguenti.

L'utilizzo di uno stile non è però sufficiente, il manager infatti può avvalersi di regole empiriche (basate sull'esperienza) approssimative per semplificare il processo. Queste regole non è detto siano affidabili, poiché possono condurre ad errori ed interferenze nel trattamento e nella valutazione delle informazioni, i quali sono evidenziati nella figura 3.1.

3.1 Errori ed interferenze



3.4 Il management oggi. Possibili soluzioni e supporto al processo decisionale

Una performance elevata nel breve periodo è possibile per molte aziende, ma solo poche di queste sono in grado di mantenerla per un lungo periodo. Per essere costanti e tenere sempre alta l'asticella, non si deve solo soddisfare il mercato esistente, ma occorre adattare le attività a causa di un continuo mutamento del contesto di riferimento. Più l'ambiente è instabile, più è importante una pianificazione strategica. Questa favorirà una visione più chiara, un maggior focus su ciò che è più importante e una migliore comprensione della rapidità del cambiamento ambientale.

Oggi all'azienda si richiede meno "burocrazia" e maggior flessibilità. Ciò garantirà rapidità di cambiamento di fronte ad ambienti turbolenti e dinamici, nonostante sia difficile mantenere costantemente una posizione competitiva. Un vantaggio sostenibile non si persegue seguendo ostinatamente un piano, centralmente gestito, ma con una serie concatenata di spinte strategiche di breve periodo. Questo significa che si dovrà passare da una strategia dominante ad un'altra di fronte a situazioni mutevoli, e proprio in questo caso starà al manager prendere le decisioni migliori possibili grazie a metodi e proprie skill.

Nonostante la "vittoria" dell'impresa possa essere legata a dei vantaggi competitivi, forti quote di mercato, know how di elevata qualità, questi non saranno più sufficienti a garantire un successo di lungo periodo, in quanto i mercati sono ad oggi altamente instabili.

I cambiamenti intervenuti nei mercati in seguito al cambio del mix delle competenze e di una maggiore dinamica strategica sono innumerevoli: la scomparsa della differenza tra servizi bancari e assicurativi nel mercato finanziario, una diversa segmentazione nel mercato automobilistico in termini di destinazione d'uso delle auto, nel campo della comunicazione è scomparsa la segmentazione del mercato in editoria, stampa, pubblicità e televisione...

Non ci saranno quindi strutture, strategie o comportamenti ottimali, ma piuttosto consone e coerenti con obiettivi, ambiente esterno e sistema. Ciascuna organizzazione dovrà quindi sapersi adattare e mutare con continuità, elemento vincente per una lunga sopravvivenza.

3.4.1 Learning organizations e decisioni in comune

Una via percorribile per garantire flessibilità e rapidità decisionale è quella di diventare una “Learning Organizations” cioè un’organizzazione con abilità di creazione, acquisizione e trasferimento di conoscenza e di modifica dei propri comportamenti.

Una learning organization è un’organizzazione, come mostra la figura 3.2, che incoraggia l’apprendimento, promuovendo lo scambio di informazioni ed il lavoro di gruppo tra le persone e creando una forza lavoro sempre più consapevole. Agire in questo modo permette di avere un’organizzazione flessibile dove i lavoratori saranno più portati ad accettare nuove idee, cambiamenti e richieste di adattamento proprio perché condivideranno una vision comune.

3.2 Learning organizations



Fonte: frontlinemanagement.wordpress.com

Questo tipo di soluzione risulta particolarmente valido per determinate attività: soluzione sistematica dei problemi, sperimentazione di nuovi approcci, apprendimento dalle proprie esperienze e da quelle altrui, trasferimento rapido ed efficiente delle conoscenze presso tutta l’organizzazione. La strategia verrà formulata grazie all’intervento non solo del top management, ma di tutte le persone presenti ai vari livelli d’azienda, le quali aiuteranno ad analizzare l’ambiente, suggeriranno cambiamenti per trarre vantaggio dal cambiamento, ed inoltre lavorando ed interagendo insieme si andranno a migliorare continuamente metodi di lavoro, procedure e tecniche di valutazione. Le imprese che sono disposte a sviluppare sperimentazioni e che sono in grado di apprendere dalla loro esperienza, sono destinate ad un

maggiore successo rispetto a quelle aziende che non hanno tali qualità. Avremo così a livello aziendale una visione e collaborazione comune, il cambiamento verrà spinto da un insieme diverso di punti di vista al fine di realizzare e prendere la decisione migliore.

3.4.2 Decision Support Systems

Il processo di decision making implica per sua natura differenti aspetti cognitivi, tra cui la raccolta delle informazioni, la valutazione di possibili scenari in cui il soggetto decisore si può trovare, la generazione e la selezione delle alternative, e la conseguente implementazione di una di esse. Come abbiamo visto in precedenza, il processo è incline ad errori, a causa dei giudizi, convinzioni e influenze del manager.

Il DSS (Decision Support Systems) è utilizzato per minimizzare gli errori cognitivi e per massimizzare la performance delle azioni. Per DSS intendiamo sistemi interattivi computer-based utilizzati dai soggetti decisori per la risoluzione di problemi non strutturati. Il supporto sarà quindi strettamente collegato all'interazione soggetto-macchina.

Un DSS è solitamente costituito da una base dati e da una base di modelli. La prima raccoglie informazioni interne ed esterne all'azienda utili al soggetto decisorio, mentre i modelli analizzano i dati in riferimento ad un preciso problema. Il DSS potrà così collezionare, elaborare e diffondere informazioni in supporto al manager.

Il DSS non è progettato una sola volta, ma in linea con il concetto di organizzazione dinamica, dovrà essere aggiornato continuamente per garantire effettivo supporto.

L'evoluzione può essere analizzata attraverso il passaggio da una concezione classica, a studi più recenti che colgono gli aspetti immateriali e cognitivi dei modelli DSS, basati sulla gestione e creazione della conoscenza e delle informazioni.

Tale evoluzione ha portato alla distinzione di diversi sistemi di supporto sulla base dell'oggetto che li caratterizza:

- **Data Driven DSS:** forniscono accesso e manipolazione ad un'elevata quantità di dati. Attraverso dei meta-dati vengono creati dei "dizionari" che descrivono e integrano informazioni acquisite da fonti differenti. Per gestire questo tipo di sistema di supporto serve una ricezione di dati accurata, completa e tempestiva per poi condurre analisi multidimensionali.

- Document-Driven DSS: pongono il focus sulla gestione di problemi non strutturati, dando supporto attraverso l'archiviazione e la possibilità di consultazione di documenti. Questi ultimi possono essere di diverse tipologie: orali (interviste interne o esterne all'azienda), scritti (report, e-mail...) e video (servizi, pubblicità...).
- Model-Driven DSS: include sistemi che utilizzano modelli di contabilità e di ottimizzazione. Attraverso questi sistemi di supporto il decision-maker è aiutato nel processo di rappresentazione delle variabili dei problemi.
- Knowledge-Driven DSS: utilizza modelli che suggeriscono le azioni da implementare ai manager. Sono caratterizzati da una forte interazione uomo-macchina e pongono il loro accento sulle competenze legate al problem solving, sulla creazione e condivisione della conoscenza, sullo sviluppo di processi di apprendimento. Vengono anche chiamati sistemi esperti in quanto coniugano le funzionalità dei DSS di prima generazione con funzionalità più esperte e innovative.
- Communications-Driven DSS: i modelli sono incentrati sulla condivisione della conoscenza e sui canali distributivi dei dati. Il communications-driven affinisce alle modalità e agli strumenti di veicolazione posseduti dall'azienda per distribuire con efficienza ed efficacia le informazioni derivanti dal knowledge-driven per supportare i processi decisionali.
- Cognitive-Driven DSS: supporto ai manager in riferimento all'elaborazione e costruzione di mappe, per dare risalto alla conoscenza, attuale e futura. Modelli fondamentali per pianificazione strategica e per supporto a decisioni non strutturate.

Requisito fondamentale dei DSS deve essere la flessibilità, in quanto esistono problemi, decisioni, dati e utenti utilizzatori differenti. Tra gli esempi di attività invece che possono garantire questi sistemi troviamo: gestione della contabilità, organizzazione delle rimanenze di magazzino, elaborazione dati appartenenti ai clienti, ottimizzazione del capitale per investimenti su materiale ed attrezzature...

I DSS offrono quindi ai manager, ed in generale ai soggetti decisori, la possibilità di ridurre i tempi di decisione/azione di fronte a problemi complessi. I processi decisionali saranno così più credibili e tramite il supporto di questi modelli l'agire manageriale sarà maggiormente giustificato, facilitandone i consensi da parte degli stakeholders. Verrà poi stimolato un comportamento atto ad esplorare nuove possibili soluzioni risolutorie, e se gli strumenti

verranno utilizzati in modo integrato con l'intuito del soggetto decisore, questo potrà costituire un vero e proprio vantaggio competitivo.

I limiti del DSS emergono a fronte di problemi di elevata complessità, dove non è immediato valutare gli interventi da attuare, privilegiando così altre tipologie di strumenti. A volte non c'è invece "armonia" tra il soggetto decisore e lo strumento, a causa della scarsa utilità di quest'ultimo rispetto ad una determinata situazione complessa, o per vera e propria incapacità del soggetto all'utilizzo. Non esiste poi un "DSS universale", un'integrazione di più modelli sarebbe difficile e costosa, ragion per cui converrà utilizzare DSS multipli.

Nonostante un DSS vada a stimolare la creatività di un decision-maker, fornendogli la possibilità di generare modelli liberamente, il grado di integrazione tra lo strumento ed il soggetto dipende solo dalla propensione di quest'ultimo ad utilizzare modelli informatici differenti dalle sue convinzioni e dai propri schemi mentali.

3.5 I manager ed il concetto ATOM

Nel capitolo uno abbiamo enunciato un concetto che si esprime in un insieme di comportamenti e metodi per sviluppare e condurre una buona ricerca. Possiamo così andare a fare un parallelismo tra gli attori del mondo scientifico-statistico e i decisori d'azienda. Il manager se pur in termini diversi, con un'adeguata modellazione rispetto al mondo economico-organizzativo, potrà seguire le disposizioni che sono state enunciate nel corso della tesi. Essi non dovranno rifarsi totalmente a metodi razionali durante la loro vita lavorativa, ma dovranno combinare regole e capacità personali per prendere la miglior decisione possibile.

Il manager per prima cosa deve accettare l'incertezza, la quale oltre ad essere una costante nella ricerca, non può che esserlo anche nel processo decisionale d'azienda. Riprendendo il concetto di razionalità limitata, il manager ha effettivamente dei limiti rispetto alla gestione delle informazioni, pertanto sarà impossibile definire scenari certi. Accettando però questa situazione il decisore potrà assumere degli atteggiamenti e metodologie che possono garantire la maggior razionalità possibile, intervenendo se necessario in atti correttivi durante le fasi del processo decisionale o durante l'implementazione della strategia. Questa figura dovrà avere un elevato grado di discrezionalità, ma deve anche essere in grado di accettare critiche ed

errori. Essere modesti porterà a non focalizzarsi su determinati schemi o convinzioni, le quali potrebbero portare ad una scarsa flessibilità aziendale e a gravi conseguenze.

Due skill cruciali per il ruolo del manager sono la comunicazione e la capacità relazionale, le quali permettono di instaurare un rapporto con i collaboratori interni all'azienda, evitando così incomprensioni ed errori che rallentino il fare dell'organizzazione. Riuscire ad essere aperti valutando più punti di vista garantirà una strategia più accurata e attenta ai dettagli (come nel caso della "learning organizations"), un po' come le ricerche, se replicate ed analizzate da soggetti diversi queste risulteranno migliori e più rigorose. Non a caso vediamo un passaggio da una vecchia concezione di creazione strategica top-down, verso una concezione bottom-up, il cui intento è di modificare gli obiettivi e le azioni aziendali attraverso le conoscenze ottenute dall'attività quotidiana.

Il manager diventerà così una figura consapevole dei suoi limiti, che se compresi potranno essere un punto di partenza per sviluppare nuove tecniche decisionali garantendo un vantaggio competitivo per l'azienda. Dovrà essere flessibile, professionalizzato e con capacità di apprendimento frequente e continuo, non tralasciando dettagli fondamentali per prendere decisioni.

CONCLUSIONI

La gestione dell'incertezza sembra un tema fondamentale per la mente umana. A livello inconscio i nostri sistemi percettivi trasformano automaticamente l'incerto nel certo. È presente in quasi tutti noi una disposizione ad accettare acriticamente idee e convinzioni comuni che valgono da millenni. Le certezze illusorie sono parte integrante del nostro retaggio percettivo, emotivo e culturale; possono darci sia un'immagine utile, benché non sempre corretta del nostro ambiente, sia sensazioni di conforto e protezione.

Nell'ambito economico, così come nella ricerca, bisogna evitare questo irrigidimento intorno a veli di sicurezza. Ciò di cui abbiamo bisogno ora è di andare oltre alla discussione dei metodi classici e delle loro interpretazioni, verso proposte concrete per l'eliminazione di problemi ricorrenti.

Il p-value, tema cardine della tesi, è un esempio notevole della tendenza umana a soffermarsi su ciò che è dato, non affrontando la vera complessità con metodi più consoni e ragionevoli. La natura del cambiamento dovrà essere incrementale, notando che non tutti convergono verso un paradigma finale. La curiosità, differenti punti di vista, una moltitudine di tecniche ed il confronto faranno da capi saldi alla nuova era del post p-value.

L'incapacità di domare i Big Data e l'analfabetismo numerico, cioè l'ignoranza del rischio, la sua comunicazione scorretta o il pensiero annebbiato, diventano un problema non appena si passa dalla terra promessa della certezza ad un mondo in cui regna l'incertezza. I ricercatori e i manager ne devono essere consapevoli. "Al di là di $p < 0.05$ " risulta essere uno spaccato utile non solo per i professionisti, ma per chiunque si affacci a questo contesto, dettando regole tecniche ma anche comportamenti. La perfezione non sarà sempre raggiungibile, ma attraverso il supporto di nuove idee e nuovi strumenti, si potrà limitare il margine d'errore di ricerche e decisioni.

Questo è solo il primo passo, perché come insegna il percorso di studi in Economia, "del domani non v'è certezza". Si deve sempre essere pronti a cambiare, adattando strategie, ricerche e processi a ciò che è necessario in un contesto altamente mutevole. Il p-value era utile in un determinato periodo, oggi invece bisogna andare oltre.

Conteggio parole: 12707

BIBLIOGRAFIA

BERALDI, F., 2016. *Big Data e attività economica, una nuova frontiera* [online]. Disponibile su: <<https://www.quadrantefuturo.it/settori/big-data-e-attivita-economica,-una-nuova-frontiera.html>>. [Data accesso: giugno 2019].

BORRA, S., e DI CIACCIO, A., 2014. *Statistica: metodologie per le scienze economiche e sociali*. III edizione. (s.l.): McGraw-Hill Education.

BUTTIGNOL, L., 2017. *Sun Tzu L'arte della guerra. Riletto a uso dei manager*. IV edizione. Milano: Rizzoli.

CASERIO, C., 2015. *Modelli d'azienda per il supporto decisionale e la generazione della conoscenza*. I edizione. Torino: Giappichelli Editore.

GIGERENZER, G., 2003. *Quando i numeri ingannano. Imparare a vivere con l'incertezza*. I edizione. Milano: Raffaello Cortina Editore.

GOODMAN, S., 2018. *Why is getting rid of p-values so hard? Musings on science and statistics* [online]. Disponibile su: <<https://www.tandfonline.com/doi/full/10.1080/00031305.2018.1558111>>. [Data di accesso: aprile 2019].

SLACK, N., BRANDON-JONES, A., e JOHNSON, R., 2016. *Operations management*. VIII edizione. (s.l.): Pearson. Pagine 80-81.

STEEL, A., LIERMANN, N., e GUTTORP, P., 2018. *Beyond calculation: a course in statistical thinking* [online]. Disponibile su: <<https://www.tandfonline.com/doi/full/10.1080/00031305.2018.1505657?scroll=top&needAccess=true>>. [Data di accesso: luglio 2019].

UNIVERSITÀ DI CAGLIARI, 2015. *Dispensa Economia e Gestione delle imprese parte IV* [online].

WASSERSTEIN, R., SCHIRM, A., e LAZAR, N., 2019. Moving to a world beyond “ $p < 0.05$ ”. *The American Statistician*., 73 (sup1), 1-19.

ZILIAK, S., 2018. *How large are your G-values? Try Gosset's Guinnessometrics when a little “p” is not enough* [online]. Disponibile su: <https://www.tandfonline.com/doi/full/10.1080/00031305.2018.1514325?af=R#_i9>. [Data di accesso: aprile 2019].