

# Come il Leicester ha vinto la Premier League?

Matteo Cavada, 5BS

20 giugno 2016

## Abstract

La storia del Leicester FC è forse una tra le più belle del calcio moderno. Una squadra, provinciale si direbbe, destinata secondo chiunque a retrocedere in seconda divisione nel giro di un paio di anni e che invece ribalta le sorti della Premier League e si classifica prima, in barba ai vari Manchester, Arsenal e Chelsea. Mi sono chiesto allora, quali sono i modelli utilizzati dai bookmaker per predire i risultati di una partita di calcio o l'esito di un campionato? E perché non hanno funzionato, in questo caso? Rispondendo a queste domande si scopre che il calcio, come la vita di ognuno di noi, è governato da eventi di difficile previsione e dalle conseguenze enormi, di cui puntualmente non teniamo conto: a questi fenomeni viene dato il nome di Cigni Neri.

## Parte I

# La storia del Leicester



Figura 1: Claudio Ranieri e Jamie Vardy che si abbracciano.

## 1 La città

Leicester è una piccola città di 330 mila abitanti posta a 140 km a nord di Londra. La sua storia non è particolarmente esaltante né costellata da eventi veramente significativi. Oggigiorno, è una delle città più multietniche dell'Inghilterra: consistenti minoranze irlandesi, polacche, indiane e est-asiatiche sono presenti sul territorio.

Per quanto riguarda lo sport, prima di quest'anno Leicester era famosa soprattutto per la sua squadra di Rugby, i Leicester Tigers, vincitori dal 1987 di quattro campionati inglesi. A partire dal 2015, tuttavia, la fama di Leicester è cresciuta soprattutto grazie all'incredibile impresa sportiva del Leicester City F.C.

### 1.1 Leicester FC, dalle origini al 2015

Il Leicester Fosse venne fondato 132 anni fa, nel 1884. Dopo una parentesi di discreto successo e dopo la prima guerra mondiale, il club fallì, ma venne rifinanziato e nel 1919 prese il nome che ancora oggi conserva, Leicester City Football Club. Nel 1928-29, raggiunse la seconda posizione nella First Division (l'equivalente della moderna Premier League), ad un solo punto di distanza dallo

Sheffield. Fino allo scoppio della seconda guerra mondiale, fu un andirivieni tra prima e seconda divisione, senza risultati particolarmente significativi.

Negli anni 1957, 61 e '63 il Leicester raggiunse la finale di FA Cup senza mai vincerla. Nella stagione '62-'63 furono i campioni d'inverno della Prima Divisione, salvo poi classificarsi quarti a fine stagione.

Nel 1978, Gary Lineker giocò le sue prime partite da professionista al Leicester, rimanendovi per sette anni; successivamente, verrà venduto all'Everton per poi finire a giocare nel Barcellona tra 1986 e 1989. Nel frattempo, il Leicester continuava ad oscillare tra prima e seconda divisione.

La svolta avvenne nel 1995-96, quando la squadra riuscì ad aggiudicarsi i playoff per l'ammissione all'allora neonata Premier League, in cui vi rimase per quattro stagioni; in questo periodo (1997 e 2000) il Leicester vinse per due volte la Coppa di Lega inglese, come già nel 1964. Risale a questo periodo un modo di dire, "fare il Leicester", che in termini calcistici indicava una squadra che, senza infamia né lode, si posizionava sempre a metà classifica senza particolari problemi.

Dal 2005 al 2014, il Leicester rimase una squadra di seconda categoria, costantemente afflitta da problemi di natura economica; il presidente di allora, Mandaric, contribuì a risanare le casse della società e a "riavviare" la squadra. Nel 2011, il Leicester venne venduto a Vichai Srivaddhanaprabha, l'uomo più ricco della Thailandia e proprietario di una delle più grandi catene di negozi duty-free al mondo.

Nel 2013, il Leicester aveva la concreta possibilità di tornare in Premier League grazie agli scontri play-off; in una clamorosa partita ribaltatasi nei minuti di recupero e conclusasi 3 a 1, il Watford rimandò i sogni di gloria delle Foxes di qualche anno: solo nella stagione 2013-2014 il Leicester trionfò in seconda divisione e si aggiudicò un posto nella Premier.

Le cose non furono affatto semplici allora. Nella rosa, tra titolari e riserve, vi erano tutti i nomi che, sebbene all'epoca fossero di giocatori sconosciuti, sarebbero diventati enormemente famosi solo due anni dopo: si parla di Drinkwater, del portiere Schmeichel, di Morgan e di Schlupp; ma anche (e soprattutto) di Jamie Vardy (che fino a cinque anni prima militava nell'ottava serie inglese) e di Riyad Mahrez (acquistato a sole 500.000 sterline!), i futuri campioni e team-men del Leicester del 2016.

La stagione 2014-'15, sebbene partita sotto i migliori auspici (una serie di risultati utili nelle prime partite) si rivelò infine disastrosa: a 9 partite dalla fine, il Leicester era ultimo in classifica, dato già per spacciato dai maggiori bookmaker. Il caso volle che i Foxes, per qualche motivo, riuscirono a rimontare questo svantaggio e a salvarsi, vincendo 7 delle ultime 9 partite di campionato. Questo fu il primo miracolo calcistico di questa squadra, certamente meno significativo di quello della stagione successiva.

## 1.2 Claudio Ranieri e la stagione 2015-2016

Claudio Ranieri, quando venne chiamato ad allenare il Leicester, era reduce da un brutto esonero da CT della nazionale greca in seguito ad un pareggio

sofferto contro le piccole Isole Far Oer. Lo scetticismo sul tecnico italiano era abbastanza diffuso, la stampa inglese si dimostrò critica dell'acquisto e, come prevedibile, relegava il Leicester nelle retrovie della classifica.

I risultati però sconfessarono le attese: la squadra di Ranieri vince inaspettatamente sette delle prime nove partite; il Leicester era primo in classifica, ma tutti si aspettavano un repentino crollo, da un momento all'altro.

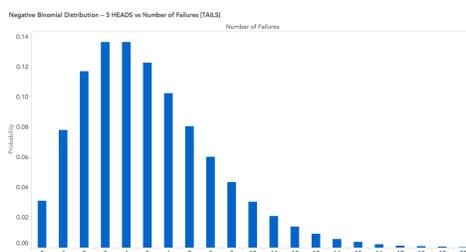
Il crollò non arrivò. Anzi, a Natale li ritroviamo ancora in cima alla classifica, dopo aver battuto il Chelsea per 2-1 (che dopo questa sconfitta esonererà Mourinho, un allenatore mai in buoni rapporti con Ranieri). Leggendo i giornali dell'epoca e gli articoli usciti in quel periodo sui principali quotidiani online inglesi, la sensazione è che in pochi, allora, credessero già ad una vittoria del Leicester: gli argomenti calcistici principali su cui si dibatteva erano appunto l'esonero di Mourinho e l'andamento sbilenco dello United fino a quel punto.

Il gioco dei Leicester, in effetti, era vincente ma non eccezionale: un 4-4-2 semplice e basato su rapidi e precisi contropiedi; le premesse per un calcio spettacolare, all'apparenza, non c'erano; e non ci sarebbero state senza l'aiuto di Jamie Vardie e di Mahrez, coppia spettacolare che ha dato consistenza (oltre che molti gol) alla squadra e che ha permesso di vincere un numero così inaspettatamente di partite.

Da gennaio in poi, il Leicester è *sempre* rimasto in prima posizione.

## Parte II

# Un modello matematico per il calcio



Una partita di calcio, come pressapoco qualsiasi cosa del nostro mondo, può essere vista come un insieme di moltissime informazioni o dati: il numero di passaggi riusciti, il numero di goal, i giocatori schierati, la media di gol subiti, il numero di espulsioni, l'umore della squadra... dati che entro certi limiti possono essere analizzati grazie ad un computer, e riutilizzati per osservare eventuali pattern. Quasi tutto è quantificabile, nel calcio: ciò, come vedremo, non significa che tutto è prevedibile.

Di seguito, descriverò i metodi e gli strumenti che ho utilizzato per scrivere un programma Python che possa simulare una stagione di Premier League basandosi sullo storico degli anni passati.

## 2 Tutto sui *data*

### 2.1 Dove e come ottenere i dati

Negli ultimi anni l'espressione *big data* ha iniziato ad essere utilizzata molto spesso nelle ricerche che consistono nell'analisi di un numero spaventosamente alto di generici "dati"; questi possono riguardare qualsiasi cosa: variazioni nelle portate dei fiumi, ricerca di pianeti abitabili nell'universo, termini di ricerca maggiormente utilizzati su Google, insorgenza di specifiche malattie e, ovviamente, risultati delle partite di calcio.

Purtroppo, per quanto riguarda le statistiche calcistiche non vi è grande abbondanza di fonti gratuite su internet: Wikipedia ed altri siti contengono molte informazioni su molti campionati europei, ad esempio, ma sono tabulati in modo tale da rendere difficile poterli normalizzare e, quindi, utilizzarli come input per un programma; similmente agiscono altri siti per le scommesse: abbiamo molti dati, ma sono difficili da utilizzare. Non ho certo intenzione di trascrivere in prima persona migliaia di partite in un file ordinato!

Dopo un po' di ricerche, sono incappato in un sito<sup>1</sup> dove gratuitamente si trovano i risultati di tutte le partite di campionato inglese dal 1880 al 2015 in formato .csv.

Ogni file contenente i risultati assume questa forma:

```
1 Date ,Team1 ,Team2 ,FT ,HT
2 2012-08-18 , Arsenal , Sunderland , 0-0 , 0-0
3 2012-08-18 , Fulham , Norwich , 5-0 , 2-0
4 2012-08-18 , Newcastle , Tottenham , 2-1 , 0-0
5 2012-08-18 , QPR , Swansea , 0-5 , 0-1
6 ...
```

Importare questi dati è molto semplice, vista la struttura estremamente lineare ed essenziale del file.

## 2.2 Cosa fare di questi dati

Avendo questi dati, si possono già effettuare le prime statistiche, le più basilari per così dire, quali il numero medio di gol segnati, il “fattore casa”, il rapporto gol segnati/subiti da ciascuna squadra e così via.<sup>2</sup>

```
1 Stats.MATCHES = 380
2 Stats.HOME_ADVG = 0.45263157894736844
3 Stats.TOT_GOALS = 975
4 Stats.AVG_GM_H = 1.4736842105263157
5 Stats.AVG_GM_T = 1.0921052631578947
6 Stats.AVG_GM_Tot = 2.5657894736842106
```

Nell'allegato 1 vi sono inoltre le statistiche squadra per squadra.

Gli acronimi utilizzati per questi dati sono combinazioni dei seguenti:

AVG (average) = media;  
GM (gol made) = gol fatti;  
GS = gol subiti;  
H (home) = in casa;  
T (transfert) = fuori casa.

MATCHES = il numero totale di match giocati

Si noti che vi è una certa differenza tra i gol segnati e subiti in casa e in trasferta, segno che è importante tener conto di quello che prima definivo come “fattore casa”, qualsiasi cosa esso sia.

---

<sup>1</sup><https://github.com/footballcsv/eng-england>

<sup>2</sup>Questi dati si riferiscono alla stagione 2014-2015.

### 3 La distribuzione di Poisson

#### 3.1 La distribuzione binomiale

La legge della distribuzione polinomiale, punto di partenza per lo studio delle variabili casuali discrete, è:

$$P_k = \binom{n}{k} p^k q^{n-k} \quad p + q = 1 \quad 0 \leq k \leq n$$

dove  $k$  indica il numero di casi favorevoli,  $n$  il numero totale di casi,  $p$  la probabilità che accada l'evento favorevole e  $q$  ( $1-p$ ).

La distribuzione binomiale è utile quando il numero di casi possibili ( $n$ ) si mantiene sotto un certo livello: è comodo utilizzarla quando, ad esempio, si lancia un dado 5, 10, 20, 30 volte; quando il numero di lanci però diventa più grande, questa distribuzione inizia a diventare sconveniente: se, ad esempio,  $n$  è 10000 e vogliamo conoscere la probabilità che il 6 esca solo 20 volte, dovremmo calcolare

$$P = \binom{10000}{20} \left(\frac{1}{6}\right)^{20} \left(\frac{1}{6}\right)^{9980}$$

... un calcolo non certo semplice.

#### 3.2 Il passaggio al limite

Per valori di  $n$  sufficientemente grandi, si può trovare una buona approssimazione, più semplice da calcolare, sfruttando un passaggio al limite.

Senza entrare nei particolari della dimostrazione, dopo una serie di calcoli si ottiene la seguente formula, detta *distribuzione poissoniana*:

$$P_k = e^{-\lambda} \frac{\lambda^k}{k!}$$

dove:

- $e$  è il numero di Nepero  $\approx 2,71828 18284\dots$
- $\lambda$  è il numero medio di eventi favorevoli nell'unità di tempo
- $k$  è il numero di casi favorevoli di cui si vuole la probabilità

#### 3.3 La distribuzione di Poisson applicata al calcio

Storicamente, il primo utilizzo della distribuzione di Poisson risale al calcolo di morti per cadute da cavallo nell'esercito prussiano. Oggigiorno, è più frequente vederla utilizzata, ad esempio, per prevedere il numero di chiamate che un call center riceverà in un determinato  $\Delta t$ , il numero di macchine che passano per un casello autostradale e così via.

Anche per il calcio è possibile sviluppare un modello basato su Poisson. Vediamo ora come simulare l'esito di una sola partita, attraverso l'algoritmo che ho elaborato.

### 3.3.1 Tutto il necessario

Per poter applicare Poisson, dobbiamo essere a conoscenza del parametro  $\lambda$  e con esso applicare la formula per vari valori di  $k$ :  $k$  è il numero di gol segnati durante la partita e  $P_k$  la probabilità che  $k$  gol vengano segnati.

Immaginiamo quindi di simulare una partita tra Leicester e Manchester United (evidenziate in blu nell'allegato 1), con il Leicester che gioca in casa.

### 3.3.2 Il calcolo di $\lambda$

Abbiamo definito  $\lambda$  come media degli eventi favorevoli nell'unità di tempo. Questo valore deve rappresentare, contemporaneamente, la forza d'attacco della squadra 1 e l'abilità in difesa della squadra 2. Definiamo quindi  $\lambda$  come

$$\lambda_1 = atk_1 \cdot dif_2 \cdot avg_{goal/campionato}$$

$$\lambda_2 = atk_2 \cdot dif_1 \cdot avg_{goal/campionato}$$

Che cosa sono, però, i parametri  $atk$  e  $dif$ ?

L'attacco deve rappresentare il rapporto tra la media di gol segnati per partita di una squadra e la media di gol di tutte le squadre: più alto è il rapporto, maggiori sono i gol segnati per partita rispetto alla media delle altre squadre. In breve:

$$atk = \frac{avg_{goal/squadra}}{avg_{goal/campionato}}$$

La difesa deve invece rappresentare l'efficacia del comparto difensivo di una squadra. Si misura similmente all'attacco, con la differenza che si tiene conto dei goal subiti, non di quelli segnati; più alto è il rapporto, maggiore è il numero di gol subiti dalla squadra rispetto alla media. In formula:

$$dif = \frac{avg_{goalsubiti/squadra}}{avg_{goalsubiti/campionato}}$$

Come ho accennato prima, bisogna sempre tener conto del "fattore campo". Se una squadra gioca in casa,  $atk$  sarà calcolato con la media dei goal segnati *in casa*, mentre  $dif$  con la media dei goal subiti *in trasferta*. Dobbiamo quindi tener conto di 4 variabili per quanto riguarda la partita Squadra1 - Squadra2:

$$\lambda_1 = atk_{1h} \cdot dif_{2t} \cdot avg_{goal_h/campionato}$$

$$\lambda_2 = atk_{2t} \cdot dif_{1h} \cdot avg_{goal_t/campionato}$$

Dove  $h$  sta per home (= in casa) e  $t$  sta per trasferta.

### 3.3.3 Una partita d'esempio: Leicester - Manchester Utd.

Calcoliamo i lambda per la partita in questione utilizzando i valori evidenziati nella tabella 1 (ricordiamo che in questo esempio il Leicester gioca in casa):

$$atk_{LFC_H} = \frac{1.47}{1.47368} \approx 0.99 \quad def_{LFC_H} = \frac{1.16}{1.09210} \approx 1.06$$

$$atk_{MAN_T} = \frac{1.11}{1.47368} \approx 0.75 \quad def_{MAN_T} = \frac{1.16}{1.09210} \approx 1.06$$

$$\lambda_{lfc} = 0.99 \cdot 1.06 \cdot 1.47368 = 1.54 \quad \lambda_{man} = 0.75 \cdot 1.06 \cdot 1.0921 = 0.89$$

L'allegato 2, di cui la figura 2 è un estratto, mostra quindi la tabella delle probabilità per la partita: il risultato più probabile è 1-0 per il Leicester!

Sembra strano che il risultato più probabile non sia in favore dello United; in realtà, questo può essere giustificato dal fatto che il Manchester, nella stagione 2014-2015, sia stato piuttosto debole nelle partite in trasferta: il rendiconto, infatti, è di cinque vittorie, nove pareggi e quattro sconfitte fuori casa.

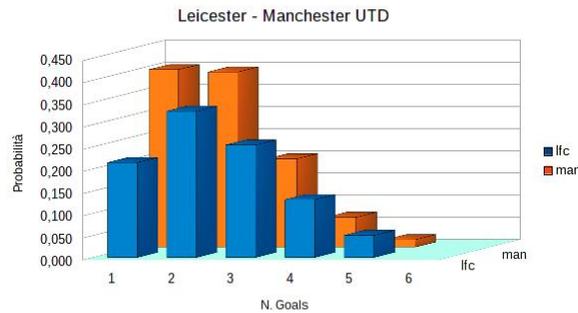


Figura 2: Il grafico della partita Leicester-Manchester UTD. Si può distinguere chiaramente la classica curva a campana che caratterizza le distribuzioni poissoniane.

### 3.3.4 Simulare un intero campionato

Simulare un intero campionato è molto semplice: basta ripetere il procedimento seguito fin'ora per tutte le partite del calendario.

Vi è una tuttavia una piccola modifica da apportare. Come ho premesso ad inizio capitolo, il calcio è anche casualità: un goal può nascere da un improvviso contropiede, da un tiro da lontano, dalla giocata di un singolo. Le partite non

finiscono tutte come sono pronosticate prima della partita: bisogna dare un velo di casualità ai risultati.

Per far ciò, ho scritto una funzione che permette, a partire dalle probabilità calcolate qui sopra, di ritornare un numero casuale che si uniforma alla distribuzione di Poisson. Qui di seguito, in pseudocodice:

```
1 definisci RandomNumberPoisson (lambda)
2     limite = e ^ (- lambda)
3     goal = 0
4     p = 1
5
6     while (p > limite):
7         goal = goal + 1
8         p = p * random(0,1)
9     return (k - 1)
```

In pratica:

- Questa funzione prende come input il lambda di una squadra e stabilisce come valore limite  $e^{-\lambda}$  (si noti che questo è un numero in pratica sempre minore di 1 e che è più piccolo quanto più grande è lambda)
- Si inizializza quindi un ciclo while, che per ogni iterazione moltiplica p per un numero casuale da 0.0 a 1.0 e aggiunge 1 al numero di goal
- Quando p sarà minore di  $e^{-\lambda}$ , allora la funzione ritornerà k-1, ovvero il numero di goal di quella partita.

Con questo piccolo accorgimento, posso iniziare a simulare stagioni calcistiche.

### 3.4 Simulazione della stagione 2015-2016

```
[ 'Chelsea ', 77]
[ 'Man City ', 72]
[ 'Arsenal ', 69]
[ 'Southampton ', 64]
[ 'Man United ', 62]
[ 'Liverpool ', 59]
[ 'Everton ', 59]
[ 'Stoke ', 57]
[ 'West Ham ', 56]
[ 'Tottenham ', 54]
[ 'Swansea ', 53]
[ 'West Brom ', 48]
[ 'Sunderland ', 47]
[ 'Watford ', 46]
[ 'Bournemouth ', 46]
[ 'Newcastle ', 42]
```

[ 'Crystal Palace ' , 38]  
[ 'Norwich ' , 37]  
[ 'Leicester ' , 35]  
[ 'Aston Villa ' , 33]

Secondo il simulatore, questo è il risultato della simulazione della stagione 2015-2016 di Premier League, basata sui dati dell'anno precedente.

Sebbene ci siano qualche sorprese “minori” (il Southampton in quarta posizione, ad esempio), almeno ai “piani alti” la simulazione sembra verosimile: ai primi posti ci sono i soliti Chelsea, Arsenal, City, Liverpool e United; col senno di poi, in ogni caso, la mia attenzione non può che ricadere sul penultimo posto del Leicester! È quasi imbarazzante, il nostro modello ha sbagliato di ben... 19 posizioni nel suo caso! E, guardando i dati reali della stagione 15/16 (allegato n.3), non è l'unico “grave” errore che ha commesso.

### Parte III

## Hume, Popper e il Leicester

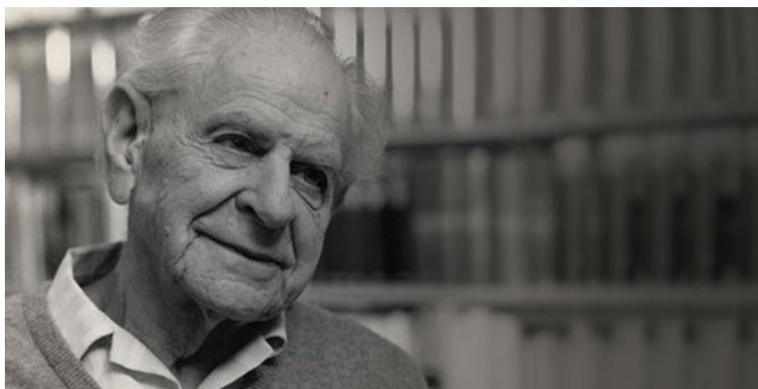


Figura 3: Karl Popper in un momento di compassione

### 4 Alla ricerca di un fondamento

Quegli strumenti matematici che fin'ora ho portato avanti ed utilizzato si sono rivelati fallaci! Il che è abbastanza preoccupante: gli scommettitori professionisti utilizzano modelli sì più complicati (certamente il numero di variabili studiate non si limita al numero di gol segnati/subiti, ma si estenderà al numero di passaggi riusciti, al possesso palla, alla forma fisica dei giocatori...) ma che, in ultima analisi, sono pur sempre modelli matematici, solo varianti più complesse di quello che ho utilizzato io. Se neppure costoro sono riusciti a tirar fuori, dalle loro simulazioni, una vittoria del Leicester, questo può voler dire una cosa: tutti i modelli si sono rivelati fallaci perché in realtà il problema di fondo è molto più nascosto, insidioso, sottile e per trovarlo è necessario ragionare “fuori dal sistema”. La domanda da porsi è quindi: su quali basi abbiamo fondato le nostre ricerche matematiche e statistiche? I nostri modelli rispecchiano il mondo che vogliamo studiare?

Popper, a questo punto, direbbe che le nostre teorie matematiche sulla previsione delle partite sono state falsificate poiché non sono state in grado di prevedere questa vittoria; purtroppo, non possiamo che dargli ragione. Ma cosa c'è di sbagliato nelle procedure che abbiamo utilizzato?

## 4.1 Mediocristan ed estremistan

Nel suo libro “Il Cigno Nero”, il filosofo N. T. Nassim formula la metafora di due paesi in cui gli eventi si comportano in maniera differente, il Mediocristan e l’Estremistan.

Nel primo paese regna la monotonia, la prevedibilità, la normalità. Qui non ci sono singoli casi in grado di cambiare radicalmente la situazione generale: ad esempio, se prendiamo un gruppo di mille persone scelte a caso e per ognuna ne misuriamo l’altezza, osserveremo che sì, c’è chi è più basso e chi più alto, ma tutti si stabilizzano entro un range ben determinato di valori. Calcoliamo l’altezza media di queste mille persone,  $m_1$ ; se ora alle nostre mille persone aggiungessimo la persona più alta o più bassa al mondo e ricalcolassimo la media, otterremmo un valore  $m_2$  tale che

$$m_1 \approx m_2$$

Altri esempi di Mediocristan sono il peso di un essere umano, l’età media di una nazione, le probabilità di vittoria a poker e così via; fenomeni per cui un singolo risultato non può mai influenzare la media di tutti i risultati.

Il secondo regno è il regno dell’imprevedibile, dei fenomeni caotici, dell’irregolarità. Qui, al contrario, uno o pochi casi sono sufficienti a sovvertire una situazione generale: prendiamo lo stesso gruppo di mille persone di prima e per ognuna di loro annotiamone il reddito annuo. Osserveremo delle discrete differenze già da subito, e la media  $m_1$  già “faticherà” a rappresentare bene i redditi estremi che abbiamo registrato. Ma cosa succede, inoltre, se alle mille persone aggiungessimo il multi-miliardario Bill Gates? La nuova media  $m_2$  uscirebbe pesantemente influenzata da un solo dato molto rilevante; in altre parole,

$$m_1 \ll m_2$$

All’Estremistan assegniamo, ad esempio, la media di copie vendute di un libro, il denaro vinto giocando in borsa, il numero di visualizzazioni di un video su Youtube, e così via; fenomeni in cui un solo risultato può influenzare pesantemente la media di tutti i risultati.

## 4.2 Errori di classificazione

La divisione tra Mediocristan ed Estremistan inizialmente può apparire del tutto arbitraria, eppure ha un fondamento nella realtà: ci sono eventi prevedibili con buona approssimazione, e tra essi annoveriamo alcuni fenomeni anatomici, naturali, fisici, ed eventi fondamentalmente imprevedibili, tra cui annoveriamo eventi la cui natura è prettamente umana/sociale e, pertanto, composta da un gran numero di variabili difficilmente stimabili. E dove sta il problema?

Il problema, epistemologico in un certo senso, è che noi pretendiamo di voler ridurre tutto al Mediocristan, anche laddove un fenomeno presenti tratti caratteristici dell’Estremistan; ricerchiamo regolarità laddove regna il caos, tentiamo di platonizzare, di modellizzare, di trascrivere in termini matematici la realtà

quando la realtà stessa è lontanissima dalle eleganti formule matematiche che utilizziamo per descriverla.

Questo, che può sembrare un problema assai astratto, è in realtà molto pratico e tangibile: la crisi economica del 2008, evento evidentemente dell'Estremistan, è stata forse prevista e/o presa sul serio dai modelli di previsione utilizzati dagli economisti? Gli attentati dell'11 settembre 2001, prodotto umano e quindi dell'Estremistan, è stato previsto dagli esperti di geopolitica di uno qualsiasi dei governi occidentali? Il 7-1 tra Germania e Brasile dei mondiali di calcio del 2014 oppure la vittoria del Leicester, eventi prettamente umani e quindi dell'Estremistan, sono stato previsti da uno soltanto dei bookmaker calcistici? A quanto pare siamo molto bravi a stimare e prevedere gli esiti di eventi ripetitivi (grazie, ad esempio, a distribuzioni come quella di Poisson), ma perdiamo totalmente di vista gli eventi improbabili e dalle conseguenze più significative: è come se fossimo meteorologi in grado di prevedere per 364 giorni se ci sarà il sole o poverà ma incapaci di prevedere l'uragano Katrina il 365esimo giorno!

### 4.3 Come ci inganniamo

Hume giustificava questa nostra tendenza a ricercare regolarità “postulato della regolarità del corso naturale”: siamo abituati a pensare che la realtà sia regolare (e solo grazie a questo è possibile inferire il principio di causalità), ma in fin dei conti su cosa è basato questo assunto? Per Hume, nessun ragionamento può giustificare una simile legge; né tantomeno l'esperienza può: servirebbero infinite osservazioni di un fenomeno per poter confermare che, effettivamente, una regolarità esiste. Quest'idea sarà poi alla base della teoria falsificazionista proposta, quasi 200 anni dopo, dal viennese Popper, che per primo conierà il termine “Cigno Nero” riferendosi proprio all'evento imprevisto che spezza la regolarità che noi ci aspettiamo.

Insomma, quando vogliamo interpretare la realtà, che secondo i nostri standard deve essere regolare, tendiamo ad attribuirle un ordine, una regolarità che in realtà non ha alcun fondamento o riscontro: questo fenomeno è detto “narrativizzazione”, fenomeno studiato per la prima volta da Hume stesso e poi ripreso dagli psicanalisti del '900.

Secondo il filosofo scozzese, nella nostra mente trasformeremmo ciò che è “*post hoc*” in “*propter hoc*”, daremmo ad eventi contigui nel tempo un rapporto di causa ed effetto, di necessità, che in realtà non esiste. Per essere più concreti, credo che basti riflettere sui nostri ricordi e sul modo in cui pensiamo al passato: si cercano, a posteriori, ragioni per cui le cose sono andate in un certo modo, si cercano cause che si adattino alla nostra visione degli eventi accaduti e li spieghino nel modo più esaustivo possibile; ma gli eventi non sono mai collegati da forti ed espliciti nessi causali, anzi, il mondo è molto più caotico di quanto non sembri a posteriori e compiere un'operazione di narrativizzazione significa, inevitabilmente, semplificare la realtà, riadattarla a nostro piacimento, riscrivere la storia.

La ricerca di ordine non è l'unica forza che ci spinge a creare narrazioni: ad esempio, dare nessi di causalità ci aiuta a ricordare meglio. È molto più semplice ricordare il numero 123456 invece che 832701: il primo lo possiamo definire attraverso nessi causali (sono numeri in ordine decrescente che differiscono di 1), il secondo, composto da numeri casuali, non segue apparentemente alcun nesso logico ed è, pertanto, più difficile da ricordare.

L'unione di questi fenomeni trova una perfetta sintesi nella vignetta qui di seguito<sup>3</sup>.



Figura 4: “Un generatore di numeri casuali pesati ha appena prodotto una serie di numeri.” “*Usiamoli per creare narrazioni!*”

Diabolicamente XKCD, l'autore della vignetta, definisce le telecronache sportive come “narrazioni costruite su numeri casuali pesati”. In parole povere: i risultati di una partita di calcio non sono altro che numeri, pseudocasuali (è questo che “numeri casuali pesati” significa), su cui ogni interpretazione è possibile e, talvolta, *difficilmente attaccabile*.

A proposito di ciò (il fatto che spesso interpretazioni narrative della realtà siano molto convincenti), D. F. Wallace scrisse nel suo monumentale *Infinite Jest*:

*“la validità logica di un ragionamento non ne garantisce la verità”*

Ogni narrazione può apparire o essere logicamente coerente e corretta, pur non essendo vera. Basti dare un'occhiata alla politica per una prova empirica di questo: partiti di destra e sinistra muovono interpretazioni anche opposte, logicamente coerenti ma non necessariamente vere, a partire dagli stessi dati.

<sup>3</sup><https://www.xkcd.com/904/>

## 5 Il Cigno Nero

### 5.1 Il Cigno Nero: perché ne siamo vulnerabili

Narrazioni, abitudine, common sense nonché l'uso di modelli platonizzanti sono tutte concause che hanno portato a credere, ad inizio stagione, che il Leicester non avrebbe mai potuto vincere il campionato. Alla luce di quanto detto, dei bias esposti nelle pagine precedenti, non poteva che andare così: d'altronde, nessuno poteva sinceramente dire che una tale idea fosse inverosimile, in quel momento! Ciononostante, pochi tifosi sfegatati, vuoi per fortuna vuoi per sincera passione, hanno puntato piccole cifre su questa piccola squadra; a maggio hanno avuto la meglio su schiere di commentatori, bookmaker, scommettitori, autorità del mondo calcistico che non avrebbero mai neppure pensato ad una tale eventualità.

Secondo la definizione di N. T. Nassim, Il Cigno Nero ha queste tre proprietà:

- è un evento isolato, che non rientra nel campo delle normali aspettative
- è un evento dall'impatto enorme nell'ambiente in cui si sviluppa
- è un evento di difficile comprensione: nonostante la sua fama di "evento imprevedibile", a posteriori ci ostiniamo a giustificarlo razionalmente e a cercare di renderlo spiegabile e prevedibile in futuro.

La storia del Leicester incarna bene la triplice definizione di Cigno Nero, nonché il fallimento di tutte le pratiche che ho descritto nelle sezioni precedenti: quale che sia il motivo per cui la squadra di Ranieri è riuscita ad arrivare fino in fondo, beh, rimane il fatto che la sua scalata ha, agli occhi di molti, dell'incredibile. Quando si semplifica troppo il mondo, d'altronde, prima o poi i problemi vengono a galla; il meccanismo è lo stesso delle previsioni meteorologiche, già citate: per poter prevedere che tempo farà, si utilizzano delle variabili i cui valori sono approssimati a  $n$  cifre decimali; queste previsioni, che sono buone entro i primi giorni, iniziano a manifestare ampie discrepanze con la realtà molto presto, siccome l'errore si propaga velocemente, e diventano inaffidabili in un tempo direttamente proporzionale alla precisione delle variabili suddette.

Ora, è evidente che finché si parla di calcio, questi Cigni Neri sono perlopiù innocui: l'1-7 di Brasile-Germania nel 2014 così come il Leicester quest'anno non hanno causato veri danni (se non forse alle agenzie di scommesse); ma come ho accennato in precedenza, avere la presunzione di conoscere ciò che è imprevedibile e ignoto porta ad essere vulnerabilissimi agli effetti del Cigno Nero: ancora una volta, a metà anni '90 qualcuno immaginò per caso la nascita di un'azienda come Google, destinata a diventare egemone nei più disparati campi dell'informatica? No, eppure la sua fondazione è certamente uno dei più importanti avvenimenti degli ultimi 30 anni.

Questo è solo un esempio, ciò che è interessante è che questo concetto si può estendere più o meno a qualsiasi tecnologia, recente o antica: è impossibile determinare il futuro perché esso è indissolubilmente legato allo sviluppo

tecnologico dell'umanità; solo se (ipotesi assurda) conoscessimo già come si evolverà la tecnologia nei prossimi anni sapremmo dire, a grandi linee, come sarà il futuro: ma conoscere la tecnologia futura significa essere già nel futuro!

In definitiva: non si può prevedere un evento futuro, che si parli di tecnologia, di politica, di economia o di calcio, perché tutte le variabili che concorrono a generarlo sono, anche esse, nel futuro; prevederle significherebbe conoscere tutto il futuro. Parafrasando una variazione alla Legge di Murphy,

L'imprevedibilità di un Cigno Nero è più alta di quanto ti aspetti, anche tenendo conto di questa legge.

## 5.2 Il Cigno Nero del Leicester

In ultima analisi, dare così tanta importanza alla vittoria del Leicester ricamandola con considerazioni quali "è merito del coraggio dei giocatori", "è solo grazie alla loro determinazione che ce l'hanno fatta" e così via, è nient'altro che creare una narrazione su numeri pseudocasuali: significa, in pratica, cercare a posteriori di riordinare e causalizzare, compiere gli stessi errori che fin'ora abbiamo tanto contestato.

Dopo aver riflettuto a lungo, dopo aver discusso con chi conosce il calcio più di me, dopo aver avuto accese discussioni e litigi con i miei amici più attenti e critici sulla squadra di Ranieri, sono giunto alla conclusione che capire questa vittoria del Leicester, ovvero come una squadretta di provincia sia riuscita a battere squadre di campioni possedute da sceicchi pieni di soldi, è fin troppo complicato per me. Non pretendo di avere una risposta sul perché il Leicester abbia vinto. Non posso esimersi però dal dare a questa storia un'interpretazione, forse un po' romantica e che non vuole essere nulla se non un'osservazione di chi è rimasto meravigliato dagli eventi dell'ultimo anno calcistico.

Questo Leicester è un esempio magnifico di un sogno apparentemente impossibile che alla fine si realizza. I ragazzi di Ranieri quest'anno hanno saputo dare grandissime emozioni anche a chi di calcio non è davvero un appassionato, ed è questo infine che si ricerca in uno sport, emozioni, passione e, ovviamente, divertimento.

La storia del Leicester dimostra che i sogni si possono coronare, a prescindere da ciò che qualcuno, chiunque! ti potrebbe far credere. È la fiaba da raccontare ai bambini prima di andare a dormire, il ragazzino che sconfigge il drago e torna a casa da eroe, è Davide contro Golia. La dimostrazione che vale la pena credere in qualcosa, non importa quanto utopistico questo "qualcosa" sia, perché nessuno sa come sarà il futuro, e nessuno lo può sapere, e nessuno può impedirci, mai, di sognare.

E proprio dei sogni parla questo aforisma di Trapattoni, con il quale voglio concludere questa mia tesina:

*I sogni sono sogni e non peccati, ma la realtà è diversa, con i se e i ma non si va lontano e più che l'acqua santa servirebbero i santi...*

## Bibliografia

- Amaldi U., *La fisica del caos. Dall'effetto farfalla ai frattali*, Bologna, Zanichelli editore, 2011
- Baroncini P, Manfredi R., Fragni I., *Lineamenti.MATH BLU vol.5*, Novara, Ghisetti e Corvi, 2014
- Cioffi F., *Popper e le epistemologie contemporanee*, in AA.VV., *Il Discorso Filosofico. Vol.3B*, Milano, Edizioni Scolastiche Mondadori, 2011
- J. A. Gardner, *Modeling and Simulating Football Results*, Università di Leeds, 2011
- N. N. Taleb, *Il Cigno Nero. Come l'improbabile governa la nostra vita*, Milano, Il Saggiatore, 2013

## Sitografia

- <https://github.com/mattyonweb/poisson-football> - Il repository in cui è presente il codice sorgente del mio programma
- <https://github.com/footballsv/eng-england> - Repository in cui sono contenuti i dati di tutte le partite del campionato inglese, dal 1880 fino al 2014
- <https://www1.maths.leeds.ac.uk/~voss/projects/2010-sports/JamesGardner.pdf> - Link alla ricerca di Gardner (cfr. Bibliografia) su un algoritmo di previsione per le partite di calcio