UNIVERSITÀ
DI SIENA
1240

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE E SCIENZE MATEMATICHE

Corso di Laurea in Matematica e Statistica

La Giostra del Saracino: spunti e riflessioni di ordine probabilistico-statistico

Relatore:

Prof. Andrea Battinelli

Tesi di Laurea di:

Agnese Nocenti

Anno Accademico 2014 - 2015

Indice

Introduzione

Il presente lavoro di tesi si colloca nell'ambito della probabilità e della statistica, in particolare del calcolo combinatorio.

Ha una genesi che mi piacerebbe raccontare. Il tutto è partito dal mio amore per la Giostra del Saracino, unito alla mia viva curiosità riguardo all'apparente casualità degli eventi e alle loro possibili connessioni.

Riflettendo sugli avvenimenti della propria vita e leggendo la Storia, viene spesso da chiedersi in qual misura l'incertezza sia preponderante nelle scelte individuali e collettive e ci si meraviglia di quanto ogni evento, seppur apparentemente trascurabile, sia il nodo di una fitta trama. Ad esempio, è affascinante poter affiancare due accadimenti a prima vista incompatibili come l'eruzione del Laki nel 1783 e la Rivoluzione Francese, e film come "Sliding Doors" ci tengono incollati allo schermo.

Ogni giugno ed ogni agosto, durante l'Estrazione delle carriere, mentre attendevo trepidante di conoscere l'ordine di entrata in Piazza dei Quartieri, mi sono spesso chiesta se ci fosse un modo per sapere a priori quale pallina sarebbe uscita dall'urna: mi avrebbe risparmiato una gran quantità di ansia. Poi, studiando Probabilità e Statistica, ho ricevuto diverse risposte e sono scaturite nuove domande, che però esulano dal presente argomento di lavoro.

La teoria della probabilità ne rappresenta il nocciolo, dal quale si dipartono le varie riflessioni di carattere probabilistico-statistico, dopo una breve presentazione ed un excursus storico dell'argomento. Il contributo più importante viene dato dal lavoro con la distribuzione binomiale, unito alla sua approssimazione tramite la distribuzione normale e da qualche spunto ulteriore riguardo alla distribuzione multinomiale.

Ringraziamenti

Ringraziamenti ufficiali

Per me la Matematica è sempre stata come una montagna impervia da scalare. Proprio per questo motivo l'ho scelta come corso di laurea: a mo' di scalatrice, non per sfidare la montagna, ma per sfidare me stessa, imparando strada facendo, cercando di colmare le mie lacune e sviluppando pazienza, umiltà e autonomia. L'arrampicata è stata difficile: poche corde di sicurezza e pochi appigli; spesso sono scivolata, parecchie volte non ho avuto una presa salda e sono ruzzolata giù per molti metri e il baratro fa paura. Ma ogni volta mi sono rialzata; alla fine ce l'ho fatta, e la mano che mi ha tirato su durante l'ultimo sforzo è stata quella del professor Battinelli; per questo lo ringrazio: grazie per avermi garantito fiducia e sicurezza, nonché continua somministrazione di conoscenze, con la tranquillità e l'umorismo di chi sa ed ha padronanza dell'argomento. Un modus operandi molto simile a quello del professor Massimo Mirolli, al quale dedico questa tesi: l'immancabile sigaro in bocca, la placida bontà nei limpidi occhi cerulei e la magia dei discorsi a lezione sono vivi nei miei ricordi, anche se ci ha lasciato già da quasi tre anni.

Il mio amore per la Giostra del Saracino è grande, tanto che, nel presente lavoro di tesi, ho voluto porla nella giusta luce. Ringrazio quindi, per la disponibilità: Fabiana Peruzzi, archivista del Quartiere di Porta Santo Spirito; Rossella Capocasale dell'Ufficio Giostra del Saracino; Luca Berti, storico della Giostra; Gianfrancesco Chiericoni, figura storica della manifestazione.

Ringraziamenti ai familiari

Ringrazio:

mamma Silvia e babbo Massimo, per darmi sempre il meglio di tutto ciò che hanno;

il mio fratellino Andrea, perché mi vuole un bene dell'anima, anche se non lo dà a vedere;

le mie nonne e i miei nonni per la loro eredità morale: da Vineta ho appreso che la vera forza è la forza di volontà; da Aurora che bontà rivestita di gentilezza è la forma più appropriata di relazione con l'altro; da Raffaello che con la cultura non si è mai soli; da Attilio, che non ho conosciuto, dicono che ho preso alcuni buoni geni;

zia Chéché, perché è bello essere una colomba, ma è molto meglio essere una colomba che becca;

zio Nicola, per le dosi massicce di ottimismo che mi ha sempre dato;

per ultimo, ma di fatto per primo, ringrazio Dio, la cui esistenza, per me, è provata anche dalla Matematica, che evidenzia la perfezione del mondo e della mente umana: Lo ringrazio di essere come sono, nata e cresciuta in questa terra.

Ringraziamenti agli amici

Ringrazio:

Fabiana, la sorella che sono tanto contenta di aver trovato;

Valentina, Mariarosaria ed Elisa, le coinquiline acquisite e «il destino che ci ha reso amiche»;

Giulia per la costante cura dello spirito con la somministrazione delle foto di *Trono di Spade*;

Eleonora per la dolcezza: «non importa se vai avanti piano, l'importante è che non ti fermi»;

Elisa e Giulia per essere le mie "donne del donno";

tutto il gruppo del BTM per l'allegria pura del nostro stare insieme, a dispetto delle differenze, e del manifestarci simpatia (se un giraffino come Lorenzo ti chiama Fedora Saura, ti deve proprio voler bene);

le mie sgnacchere Eleonora e Chiara, il gruppo delle entusiasmanti, "Con Antico Ardore" e la gente di Sanfra (Veronica, Anne, Giulina, Simone, Nicola, Gioia, Claudia, Khaleesi, ...), per allietare le mie giornate con i messaggi su whatsapp e riempire i miei weekend di allegria;

Nilanthi, per condividere la passione per il fantasy;

la collega cinefila Linda, per condividere l'amore per i film e i libri;

Giada, Eleonora (prima amica) e Veronica per le cene "ricostituenti";

Ludovica, per l'intesa immediata (tra Scorpioni ci si intende con uno sguardo).

Ringraziamenti ad artisti del cuore

Grazie ai Led Zeppelin per avermi regalato, ad ogni ascolto, una scala per il Paradiso, carica ed evasione nei momenti di difficoltà, ed avermi inviato il messaggio, opportunamente recepito:

"Do What Thou Wilt So Mete It Be".

1. Premessa

È opportuno anticipare che il presente lavoro di tesi non si propone di trattare in maniera esaustiva l'approccio probabilistico-statistico alla Giostra del Saracino: ci sono altri ambiti, come, ad esempio, la correlazione tra ordine di estrazione ed effettiva vittoria che sarebbe interessante affrontare, così come la probabilità, ottenuta tramite la distribuzione gaussiana multidimensionale, che un Quartiere sia estratto in un certo intervallo.

Mi sono quindi proposta di indagare approfonditamente il meccanismo delle estrazioni delle carriere, fenomeno a prima vista semplice, ma che ha rivelato pieghe interessanti e spunti di riflessione.

2. La Giostra del Saracino

2.1 Breve storia della Giostra del Saracino

La Giostra del Saracino dell'età moderna è la manifestazione storica di un "torneamento" in uso già intorno all'anno Mille come esercizio militare preparatorio alla partenza per le crociate: il cavalcare lancia in resta contro un fantoccio (Buratto) con le fattezze di un saraceno, schivando il suo mazzafrusto[1], serviva ai giovani cavalieri per affinare le proprie abilità belliche.

Nel corso dei secoli, poi, ha assunto un connotato più "ludico" e celebrativo: in occasione di feste, matrimoni, visite di personalità importanti, il fior fiore della nobiltà si dava appuntamento in una piazza per assistere alle prove di abilità dei giovani rampolli, il tutto sotto gli occhi del popolo, che si godeva il doppio spettacolo dello sfarzo e del valore.

Di questi *hastiludia* in Arezzo abbiamo diverse testimonianze, tra le quali la più illustre (nonché molto discussa) è senza dubbio quella di Dante Alighieri, nelle prime terzine del XXII canto dell'*Inferno*:

« Io vidi già cavalier muover campo,
e cominciar stormo a far lor mostra,
e tal volta partir per loro scampo;
corridor vidi per la terra vostra,
o Aretini, e vidi gir gualdane,
fedir torneamenti e correr giostra;
quando con trombe, e quando con campane,
con tamburi e con cenni di castella,
e con cose nostrali e con istrane; »

Non mancano i documenti e i resoconti di giostre *ad burattum* nel Rinascimento e fino al XVIII secolo: nell'Archivio di Stato è conservato un regolamento di giostra del 1677 in cui fa la sua comparsa il tabellone (sorta di bersaglio su cui sono disposti i vari punteggi; tra questi, il più ambito è il pallino centrale di 4 cm di diametro, che, colpito, fa ottenere al cavaliere 5 punti).

Tra alterne vicende, si sono disputate diverse giostre durante il Settecento e l'Ottocento, ma è agli inizi del Novecento che si ha una vera e propria ripresa di interesse verso la manifestazione. Infatti, nei primi decenni del secolo scorso, all'interno dei circoli culturali della città venne maturando l'idea di riscoprire e restaurare il passato.

Con l'avvento del fascismo e la sua conseguente spinta alla ricostruzione delle tradizioni a fine di propaganda, la Giostra del Saracino rinasce in una veste moderna, seppure rispettosa della storicità: grazie al lavoro del giornalista Alfredo Bennati (che rinviene il regolamento del 1677, usato come base per il nuovo), del segretario federale del Partito nazionale fascista Antonio Cappelli e del podestà Pier Ludovico Occhini, viene corsa, il 7 agosto del 1931, una prima Giostra "di prova". Infatti, sia il percorso dei cavalieri, che la suddivisione della città in Rioni, che la tipologia di premio, erano ancora provvisori. L'anno successivo,

[1] Sorta di flagello con tre palle di cuoio, stretto nella mano destra.

dopo uno studio più preciso dei documenti cittadini e una ricostruzione storica più accurata, si corse la prima Giostra per come la conosciamo; questa fu l'apri fila di una lunga serie (interrotta per otto anni, a causa della Seconda Guerra Mondiale) che gloriosamente continua ancora oggi.

2.2 La cerimonia di Estrazione delle carriere

Chiunque capiti ad Arezzo durante giugno e settembre viene rapito dalla bellezza dei colori e dei suoni di una città che torna, due volte all'anno, come era settecento anni fa. Chiarine, tamburi, sbandieratori, dame che sembrano uscite dagli affreschi di Piero della Francesca sono splendide parti del quadro della manifestazione, che si articola in diversi eventi (Estrazione delle carriere, Bollatura dei cavalli, Bando, Giostra).

A tutto ciò partecipa, con grande fervore, ognuno dei quattro Quartieri nei quali è suddivisa la città: Porta Crucifera (settore nord-est, colori rosso e verde), Porta del Foro (settore nord-ovest, colori giallo e cremisi), Porta Sant'Andrea (settore sud-est, colori bianco e verde) e Porta Santo Spirito (settore sud-ovest, colori giallo e blu). Dal caloroso abbraccio dei quartieristi di ognuno di essi escono i due cavalieri che, in coppia, affronteranno il Buratto durante l'edizione della Giostra in notturna (penultimo sabato di giugno) e l'edizione in diurna, la prima domenica di settembre.

Per assicurare l'imparzialità nella sfida tra gli otto cavalieri, una settimana prima della giostra avviene la *cerimonia di Estrazione delle carriere,* i cui partecipanti indossano costumi trecenteschi: le rappresentanze armate dei quattro Quartieri accompagnano il proprio paggetto (un piccolo quartierista) fino a Piazza della Libertà, dove li attendono, tra gonfaloni, armigeri e Musici, il Sindaco e l'Araldo. Quest'ultimo chiama sul palco, seguendo l'ordine d'estrazione della Giostra precedente, il primo paggetto, che estrae da un sacchetto di cuoio una pallina delle quattro contenutevi; questa, una volta aperta dal Sindaco, sotto l'occhio vigile del Cancelliere e dei Rettori dei Quartieri, rivelerà i colori del Quartiere che scenderà nella Lizza per primo. L'operazione poi si ripete con i rimanenti tre paggetti, fino a che il sacchetto non sarà vuoto e l'ordine delle carriere completo. Durante la cerimonia, inoltre, ognuno dei quattro Capitani legge il proprio giuramento ed estrae, da un altro sacchetto, i numeri delle due lance da giostra per i propri cavalieri; in questo modo, ci si assicura l'imparzialità anche nella scelta delle aste: secondo il Regolamento, se il giostratore spezza la lancia nell'impatto col Buratto, il suo punteggio viene raddoppiato. Con il giuramento del Maestro di Campo (l'arbitro del "torneamento") prima, e la presentazione in Duomo della Lancia d'Oro (l'ambito premio) poi, si dichiara aperta la settimana del "pregiostra".

L'evento conclusivo sarà la giostra vera e propria: seguendo l'ordine deciso dalla "dea bendata", il primo cavaliere del primo Quartiere cavalcherà lancia in resta contro il Buratto; questa carriera verrà seguita da quelle degli altri tre cavalieri. Successivamente, sarà la volta del secondo cavaliere del primo Quartiere e così via, fino al secondo cavaliere del quarto Quartiere. Concluse queste due serie di quattro carriere ciascuna, verranno confrontate le somme dei

punteggi di ogni coppia di cavalieri e, in caso di parità, verranno corse una o più carriere di spareggio, fino ad ottenere la coppia vincitrice.

3. Breve storia della Teoria della Probabilità

3.1 Il percorso scientifico

Certamente, fin dagli albori della civiltà l'uomo ha avuto necessità di misurarsi con il concetto di incertezza e probabilità, ma è solo nel XVII secolo che ha iniziato a dargli una veste scientifica.

Non è facile fissare una data precisa, ma molti concordano nel vedere il 1654 come anno della nascita della teoria della probabilità, ad opera di due tra i maggiori matematici dell'epoca: Pierre **de Fermat** e Blaise **Pascal**. Antoine Gombaud, cavaliere de Méré, nonché arbiter elegantiarum e famoso giocatore d'azzardo, riscontrando notevoli discrepanze tra i propri calcoli probabilistici e le proprie fortune (o, meglio, sfortune) al gioco, si era rivolto a Pascal per ricevere spiegazioni al riguardo. Ne nacque una fitta corrispondenza tra questi e l'amico Fermat, densa di deduzioni scientifiche e spunti che funsero da base per la stesura, tre anni dopo, del *De ratiociniis in ludo aleae* di Christian Huygens, scienziato olandese insegnante di Leibniz.

Ben presto, forte della notevole attrattiva dei giochi aleatori, la teoria della probabilità divenne uno degli argomenti centrali dell'indagine matematica, grazie, tra gli altri, a Bernoulli e De Moivre.

Solo con Laplace si ha, comunque, un'emancipazione dal contesto ristretto dell'analisi dei giochi aleatori: nella sua opera *Théorie Analytique des Probabilités*, infatti, egli apre a nuovi ambiti di ricerca, applicando le "scoperte" probabilistiche a svariati problemi, anche pratici. Il matematico non si fermò qui, ma dette il proprio contributo alla formulazione della famosa distribuzione normale di Gauss (detta, infatti, anche distribuzione di Gauss-Laplace), considerata il mattone fondante della statistica moderna, il "braccio operativo" della teoria della probabilità.

Come è capitato in tante altre branche della Matematica, lo sviluppo della teoria della probabilità è stato stimolato dalla varietà delle sue applicazioni (genetica, economia, psicologia, ingegneria sono solo alcuni tra i settori che hanno beneficiato degli strumenti della statistica). Di converso, ogni progresso ha ampliato la portata dell'influenza della teoria della probabilità. Ecco spiegata la lunga lista di importanti matematici che si sono "passati il testimone" durante la corsa alla scoperta di nuovi aspetti della probabilità; solo per citarne alcuni: Chebyshev, Markov, von Mises e **Kolmogorov**.

Proprio quest'ultimo è protagonista di una svolta nella storia della teoria della probabilità. Infatti, una delle difficoltà principali nello sviluppo della teoria è stata lo stilare una definizione di probabilità che fosse abbastanza rigorosa e precisa per l'utilizzo in matematica, ma che, al tempo stesso, fosse sufficientemente flessibile per l'applicazione in un ampio raggio di fenomeni. La ricerca di una definizione accettabile richiese quasi tre secoli e fu segnata da molte controversie. Il problema fu sommariamente risolto nel 1933 dal matematico russo nella monografia *Fondamenti della teoria della probabilità*, attraverso la

trattazione su base assiomatica. Questa impostazione non dà della probabilità una definizione diretta, operativa, né fornisce indicazioni su come calcolarla, ma accetta qualunque approccio, purché questo rispetti le proprietà fondamentali, assunte come assiomi; da queste si deducono le altre proprietà come teoremi. Molto brevemente, la teoria assiomatica si fonda su tre momenti fondamentali: individuazione dei concetti fondamentali (prova, evento, spazio); enunciazione degli assiomi (o postulati) della probabilità (positività, certezza, unione); dimostrazione dei teoremi mediante i postulati e con l'ausilio della logica.

3.2 Italia, culla della Teoria della Probabilità

Scorrendo la breve storia della teoria della probabilità salta subito all'occhio il fatto che essa sia nata in tempi piuttosto recenti, "solo" quattro secoli fa. Diversi motivi, che spiegheremo alla fine del capitolo, hanno impaniato per molto tempo lo sviluppo scientifico dell'argomento.

Benché diversi studiosi siano del parere che sia dell'umanista de Fournival (1250) la paternità del poema latino *De vetula* (una delle prime opere in cui si abbozza una trattazione del concetto di aleatorietà e del calcolo combinatorio, attribuita per secoli ad Ovidio), gli italiani sembrerebbero essere i pionieri dello studio della probabilità.

Tema ricorrente nel *De vetula* è il gioco dei dadi. A lungo stigmatizzato dalla Chiesa non tanto per la sua casualità, quanto per l'entourage di vizi che lo accompagna, questo gioco ha riscosso un certo interesse nei letterati di varie epoche, proprio per i suoi aspetti di imprevedibilità, di rischio, di trasgressione, piuttosto affascinanti e densi di spunti per una mente fertile.

Dante è uno dei primi letterati italiani a far menzione di un gioco d'azzardo, nel VI canto del *Purgatorio*:

«Quando si parte il gioco della zara
Colui che perde si rimane dolente
Ripetendo le volte e tristo impara»,

dandone notevole pubblicità, tanto che successivamente fioccarono alcune opere al riguardo.

Nel XV secolo era già stata data una base piuttosto rigorosa all'argomento, con l'assunto che i dadi debbano essere perfettamente quadrati (e che quindi tutte le sei facce abbiano la stessa probabilità di venire), nonché con l'idea di associare coefficienti binomiali alle probabilità degli eventi realizzabili lanciando due o più dadi, al fine di calcolare la loro frequenza relativa di occorrenza.

Sebbene quest'ultimo passo pionieristico del *De vetula* non sembri noto agli scrittori occupatisi dell'argomento nei due secoli successivi, l'idea di enumerare i modi di avere un dato punteggio, quando venga tenuta presente la possibilità di ottenerlo con svariate permutazioni, deve essere stata riscoperta all'inizio del secolo XVI, in quanto il *De ludo aleae* (1526) di Cardano ne contiene i tratti essenziali.

Facciamo un piccolo doveroso passo indietro. Uno dei primi studiosi ad interessarsi al problema della divisione della posta (argomento preponderante nel sopracitato carteggio tra Pascal e Fermat, più di un secolo dopo) fu Fra Luca

Pacioli, la cui *Summa de Arithmetica, Geometria, Proportioni et Proportionalità*, pubblicata nel 1494, venne ampliamente studiata in Italia. Egli considerò una semplice versione del problema: A e B giocano ad un gioco equo (non dadi, ma "balla", verosimilmente un gioco di palla) e si accordano nel continuare sino a quando uno dei due vince 6 partite, ma la competizione deve essere interrotta quando A ha vinto 5 partite e B 3. Come dovrebbe essere ripartita la posta? Benché Pacioli faccia sì che il problema sembri più difficile di quanto non sia, la sua soluzione si compendia nel dire che le poste dovrebbero esser suddivise nella proporzione di 5 a 3. L'errore fu notato da diversi studiosi e dette un grande impulso alla proliferazione di scritti: tra gli altri, Tartaglia nel suo monumentale *General Trattato* del 1556 e **Peverone** in *Due brevi e facili trattati, il primo di Aritmetica e l'altro di Geometria*. In quest'ultimo, è interessante notare come Peverone sia pervenuto alla stessa soluzione di Pascal e Fermat, più di un secolo prima dei due. Ma, come purtroppo accade spesso nella storia della scienza, ciò non ha ricevuto un degno riconoscimento: nei tempi in cui scriveva il cuneese, l'uditorio scientifico era molto preso dalla rivoluzione copernicana prima e dallo scandalo galileiano poi; in breve, Peverone è stato troppo in avanti coi tempi.

A proposito di Galileo, anche quest'ultimo ha scritto di un problema di probabilità nel frammento *Sulla scoperta dei dadi*, dandone una soluzione completa con l'annotazione corretta di tutte le possibilità e scrivendo come se fosse una propria scoperta, non richiamando alcun precedente autore; nondimeno, queste idee erano correnti già da un secolo prima che Galileo scrivesse.

Si può quindi pensare che il calcolo delle probabilità non solo si sviluppò tardivamente (XIII secolo) ma che, una volta iniziato, progredì in modo estremamente lento. Prima di concludere, considerandone le ragioni, è doveroso chiudere il cerchio: il fatto che Francia e Italia si contendano la maternità della teoria della probabilità è da ricercarsi negli sviluppi storici del XIV secolo. L'invasione dell'Italia di Carlo VIII nel 1494, sebbene sia stata un fallimento politico e militare, viene generalmente ritenuta come una parte di uno sviluppo intellettuale incrociato. Senza alcun dubbio, una gran quantità di idee e di lavori artistici italiani trovarono la via per la Francia con quanto restava dell'esercito di Carlo. Vi furono, quindi, connessioni strette fra la Francia e l'Italia di carattere non solo politico, ma anche intellettuale.

Concludendo, la strana penuria, fin dai tempi antichi, di riferimenti scritti a problemi di calcolo delle probabilità non è dovuta a mancanza di un contemporaneo interesse, tanto più che la conoscenza del caso avrebbe valso non poco denaro, nei vari giochi. Sembra ci siano quindi diverse cause di ciò, non solo di carattere prettamente matematico.

Una delle prime sembrerebbe da ascrivere all'imperfezione dei dadi, ma non è la preponderante, dato che molti dadi antichi sono piuttosto ben fatti; nemmeno il lento sviluppo della notazione matematica sembra essere un motivo rilevante: le partizioni delle uscite delle facce dei dadi venivano contate senza difficoltà già nel X secolo. È, comunque, piuttosto interessante notare l'assenza di un'Algebra combinatoria fin dai Greci: solo con Leibniz (*De arte combinatoria*, 1660) è stata data una veste scientifica alle varie idee e ai vari metodi aritmetici usati per contare. Si può mettere sul piatto della bilancia anche la notevole superstizione

che ha da sempre aleggiato sui giochi aleatori, ma nemmeno questo sembra essere un ostacolo insormontabile: notoriamente, i grandi pensatori sono persone con una viva intelligenza, che esula dagli schemi precostituiti e dalle "paure" della gente comune, ed, anzi, attinge forza dalla sete di scoperta, dalla sfida dell'ignoto.

Spesso sfugge che la nozione di caso, l'idea di legge naturale, la possibilità che una proposizione possa essere vera e falsa in una data proporzione sono tutti concetti ben presenti nelle nostre comuni abitudini di pensiero, ma non era così per i nostri antenati. Anche se per i Greci e i Romani, in generale, il mondo era in buona parte dominato dal caso (la dea Fortuna, il Fato), la situazione venne radicalmente mutata dall'avvento del cristianesimo. Per i Padri della Chiesa il dito di Dio era ovunque. Alcune cause erano palesi, altre nascoste, ma nulla accadeva senza motivo; in tal senso, nel mondo non v'era posto per il caso. Ciò ha indirettamente condizionato la mentalità degli studiosi, fino all'avvento del Rinascimento e dell'Illuminismo.

Fare di tutta l'erba un fascio, comunque, è spesso deleterio, perché offusca l'occhio scientifico: autori come San Tommaso d'Aquino lasciavano comunque spazio a dubbi riguardo all'assenza del caso e i pensatori Greci erano spesso ingabbiati in una visione piuttosto rigida della Matematica.

Le cause della nascita ritardata di un pensiero scientifico sulla probabilità sono quindi tante e così intrecciate tra loro che è molto difficile districarsi, ma, in fondo, in qualunque ambito dello scibile umano accade così.

4. Concetti fondamentali

4.1 Elementi di probabilità

Il calcolo delle probabilità è uno strumento essenziale per la statistica, proprio perché è la risposta al problema inverso di quest'ultima: calcolando a priori la probabilità che un dato esperimento abbia un certo risultato, esso ci consente di far luce sui possibili risultati di ogni esperimento, una volta avvenuto.

Ciò è possibile poiché il concetto di probabilità è strettamente collegato con il concetto di incertezza. Di fatto, possiamo pensare che ogni azione umana venga intrapresa in condizioni di incertezza, cui si accompagnano forme di scelta di tipo rischioso. Tuttavia non agiamo nella completa ignoranza, ma cerchiamo di valutare tutti i possibili fattori che entrano in gioco e ci regoliamo di conseguenza, cioè facciamo, più o meno consapevolmente, delle valutazioni di probabilità. Legato al concetto di incertezza, quindi, c'è quello di evento.

Definizione. Ogni fenomeno o insieme di fenomeni (astratti o concreti; del passato, del presente o del futuro) è definito *evento* se è espresso da una proposizione dichiarativa (semplice o composta) accertabile vera o, in alternativa, falsa, senza dubbi da parte dei suoi assertori, i quali utilizzano una serie di protocolli adatti allo scopo.

Esempio. Durante l'estrazione delle carriere, il paggetto estrae una e una sola pallina dal sacchetto. I colori di quale Quartiere conterrà?

L'ambito del calcolo della probabilità è considerabile come uno dei ponti tra la Logica proposizionale e la Teoria degli Insiemi: queste, infatti, hanno diversi elementi di affinità. Nella fattispecie, possiamo dire che una proposizione dichiarativa si "comporta" in modo "simile" ad un insieme: i connettivi logici $\neg \vee \wedge$ valgono in modo analogo, rispettivamente, alle operazioni insiemistiche di complemento, unione ed intersezione. Non solo: se ci soffermiamo ad osservare l'attribuzione del significato di una FBF, notiamo che il meccanismo è affine a quello (che vedremo a breve) di una variabile aleatoria; in entrambi i casi, infatti, abbiamo una funzione che parte da un dominio (l'insieme delle FBF/l'insieme dei risultati di un esperimento aleatorio) e che associa un elemento di esso ad un valore di un insieme di numeri (anche se, nella Semantica del Calcolo Proposizionale questo è semplicemente $\{0, 1\}$, mentre nel secondo caso può essere anche un intervallo o un insieme infinito numerabile).

Quindi, grazie a questi punti di contatto, gli eventi della Teoria della Probabilità sono "camaleontici", possono passare da una veste logica ad una insiemistica e viceversa: sono espressi da una proposizione il cui valore di verità può essere 0 o 1 (vero o falso), ma sono anche elementi di un insieme, come stiamo per vedere.

Per studiare un evento, è essenziale che avvengano delle prove, delle attuazioni del fenomeno preso in esame: in parole povere, c'è bisogno di un esperimento aleatorio.

Definizione. Lo *spazio campionario* (o spazio *degli eventi*) Ω è l'insieme di tutti i possibili risultati di un esperimento aleatorio. Chiameremo *evento*

elementare ogni singolo elemento e di Ω; ogni sottoinsieme E di Ω sarà invece detto *evento composto.*

Esempio. Per l'estrazione delle carriere, $\Omega = \{PC, PDF, PSA, PSS\}$, laddove PC, PDF, PSA, PSS sono gli eventi (elementari).[2]

È importante fare una precisazione riguardo allo spazio campionario: dato che l'argomento del presente lavoro di tesi ci fa operare nel discreto, nei vari approcci che vedremo è presente l'ipotesi implicita che lo spazio degli eventi sia finito; ciò, tuttavia, rappresenta solo un "caso particolare", poiché ogni approccio è applicabile in qualsiasi tipo di spazio campionario. Laddove sarà auspicabile, verrà trattato anche il caso continuo.

Per come abbiamo definito Ω, infatti, verrebbe da pensare che la Teoria della Probabilità si appoggi solo sulla Teoria degli Insiemi e sulla Logica proposizionale, ma, in realtà, anche la Teoria della Misura ha la sua importanza. Con essa, infatti, riusciamo ad estendere i concetti di lunghezza, area, volume e quindi misurare l'insieme dei casi possibili, e dei casi favorevoli anche quando essi siano insiemi infiniti, di cardinalità numerabile o continua, e a utillizzare i concetti di funzione di densità e di distribuzione (che vedremo a breve).

Intanto accenniamo ad una definizione che ci servirà:

Definizione. Una classe F di sottoinsiemi di Ω viene detta *algebra (di parti)* se:

1. $\Omega \in F$
2. Se $A \in F$, allora $A^c \in F$ (chiusura per il complementare)
3. Se $\{A_i\}_{1 \leq i \leq n} \subseteq F$, allora $\bigcup_{i=1}^{n} A_i \in F$ (chiusura per unione finita)

Se aggiungiamo un'altra proprietà:

3 bis $\{A_n\}_{n \in \mathbb{N}} \subseteq F \Longrightarrow \bigcup_{n \in \mathbb{N}} A_n \in F$ (chiusura per unione numerabile)

la nostra algebra diventa una σ-algebra.

A questo punto, possiamo vedere la probabilità come una misura, una funzione che ha per dominio l'algebra delle parti di Ω ed è a valori reali:

Definizione. Data un'algebra F, viene detta *massa* la funzione

$$\mu : F \to [0, \infty[$$

tale che

1. $\mu(\emptyset) = 0$
2. $\{A_i\}_{1 \leq i \leq n} \subseteq F$ t. c. $\forall i, \forall j = 1, ..., n, i \neq j, A_i \cap A_j = \emptyset \Rightarrow \mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$ (additività finita).

[2] Per semplicità e brevità, d'ora innanzi utilizzeremo queste sigle per significare, rispettivamente: «La pallina estratta contiene i colori del Quartiere di Porta Crucifera (o di Porta del Foro, Porta Sant'Andrea, Porta Santo Spirito)».

Se a ciò aggiungiamo la proprietà di essere stabile per additività numerabili[3]:

2 bis $\{A_n\}_{n\in\mathbb{N}} \subseteq F$ t. c. $\forall i, \forall j \in \mathbb{N}, i \neq j, A_i \cap A_j = \emptyset \Rightarrow \mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$

la funzione μ è chiamata *misura*.

Se una misura soddisfa pure la condizione $\mu(\Omega) = 1$, abbiamo che $\mu := P$ è una *misura di probabilità* e quindi la terna (Ω, F, P) è uno spazio di probabilità.

La probabilità eredita, in questo modo, gli assiomi della misura, tra i quali la monotonia: se $E_1 \subseteq E_2$, allora $P(E_1) \leq P(E_2)$.

Alla luce di quanto detto finora, le seguenti definizioni acquistano un fondamento più solido.

Definizione. Un evento si dice *certo* se, sulla base delle informazioni a nostra disposizione, possiamo dedurre che la proposizione che lo definisce è vera; viene definito, invece, *possibile* se le informazioni di cui al momento disponiamo sono insufficienti e non possiamo dedurre con certezza se l'evento è verificato o meno; infine, un evento è detto *impossibile* quando possiamo dedurlo tale in base alle informazioni in nostro possesso, che invalidano la proposizione enunciante.

Definizione. Si definisce $\underline{E}$ (*indicatore* di un evento E) quel numero aleatorio che, al momento, è sconosciuto e che vale 0 se E è falso e 1 se E viene accertato vero.

Definizione. Dato un evento E, si dice *evento opposto* (*o complementare*) di E e si indica con E^c l'evento che è vero quando E è falso e viceversa. $\underline{E} = 1 - \underline{E^c}$.

Definizione. Dati due eventi E_1 e E_2, l'evento $E_1 \vee E_2$ si dice *somma logica di* E_1 e E_2 ed è vero quando almeno uno dei due componenti è vero. Volendolo guardare sotto una luce insiemistica, esso è un evento unione, formato da tutti gli eventi elementari che appartengono ad almeno uno dei due sottoinsiemi. L'*evento intersezione* $E_1 \wedge E_2$ è invece composto dagli eventi che appartengono ad entrambi i sottoinsiemi ed ha come equivalente logico il *prodotto logico* $E_1 \wedge E_2$, evento che risulta vero quando sia E_1 che E_2 sono veri (vale pertanto $\underline{E_1 \wedge E_2} = \underline{E_1} \cdot \underline{E_2}$). Inoltre, vale

$$\underline{E_1 \vee E_2} = \underline{E_1} + \underline{E_2} - \underline{E_1} \cdot \underline{E_2}$$

Ovviamente, tutte le varie proprietà legate alle operazioni insiemistiche si trasferiscono intatte nell'ambito degli eventi: abbiamo la proprietà commutativa, associativa e distributiva, oltre alle formule di De Morgan:

$$\begin{aligned} (E_1 \cup E_2)^c &= E_1^c \cap E_2^c \\ (E_1 \cap E_2)^c &= E_1^c \cup E_2^c \end{aligned}$$

.

E le altre relazioni, piuttosto immediate:

[3] Dato che, in generale, non si può estrarre un'algebra delle parti numerabile da un insieme infinito, è sempre opportuno preventivamente verificare che la famiglia di parti su cui definiamo μ sia chiusa per unione numerabile.

$$\begin{aligned} E \cup E &= E, E \cap E = E \\ (E^c)^c &= E \\ \Omega \cup E &= \Omega, \Omega \cap E = E \\ \emptyset \cup E &= E, \emptyset \cap E = \emptyset \end{aligned}$$

Definizione. Diremo che due eventi E_1 e E_2 sono *mutuamente esclusivi* (o *incompatibili*) se, detta I la classe degli eventi impossibili[4], avviene che $E_1 \wedge E_2 \in I$. Insiemisticamente, due eventi sono incompatibili se $E_1 \cap E_2 = \emptyset$, ovvero se non hanno eventi elementari in comune. È ovvio che, quindi, il realizzarsi di E_1 escluda che si verifichi, contemporaneamente, E_2 e viceversa. In aggiunta, è implicita l'assunzione che due eventi elementari distinti siano sempre incompatibili. Se E_1ed E_2 sono incompatibili, avremo anche che $\underline{E_1} \cdot \underline{E_2} = 0$ e $\underline{E_1 \vee E_2} = \underline{E_1} + \underline{E_2}$.

Adesso che abbiamo i "mattoni", passiamo in rassegna i diversi tipi di "cemento".

La probabilità è un concetto primitivo, innato nella natura umana, poiché ci permette di far fronte all'incertezza. Non è facile, quindi, dare una definizione soddisfacente di questo concetto: nessuna delle definizioni date durante lo studio della probabilità è stata esente da critiche.

L'approccio originario e più semplice alla definizione di probabilità è, senza dubbio, l'*approccio classico*, che dobbiamo a Laplace. Per calcolare la probabilità di un certo evento, prima si stabilisce qual è lo spazio degli eventi e si conta il numero dei suoi elementi (considerati equipossibili) e poi si valuta quanti sono i casi favorevoli tra quelli totali. Infine si considera il rapporto tra il numero dei casi favorevoli e quello dei casi possibili; tale rapporto indica numericamente la probabilità che ha quell'evento di verificarsi.

Definizione. Si definisce *probabilità di un evento* E, e si indica con $P(E)$, il rapporto $\frac{m}{n}$ tra il numero (m) di casi favorevoli al verificarsi di E e gli n casi possibili (giudicati egualmente possibili).

Si osservi che la probabilità è una frazione positiva, il cui numeratore non può essere maggiore del denominatore; è, perciò, un numero razionale compreso tra 0 e 1:

$0 \leq P(E) \leq 1$

Inoltre, se non esistono casi favorevoli, l'evento è impossibile e la probabilità nulla; al contrario, se tutti i casi possibili sono favorevoli, la probabilità è 1 e l'evento è certo:

$P(\emptyset) = 0$, $P(\Omega) = 1$.

È valida anche la proprietà di *additività semplice* (mutuata dal Teorema delle Probabilità Totali): comunque scelta una coppia di eventi incompatibili E_1 e E_2, si ha che:

$P(E_1 \cup E_2) = P(E_1) \cup P(E_2)$.

[4] (Durante questa trattazione sono stati tenuti presenti gli "Appunti complementari al corso di Probabilità e Statistica" del professor Battinelli).

La definizione appunto data è estendibile anche al caso non finito: la probabilità, come abbiamo visto, è considerabile come una misura e, in quanto tale, si estende anche ad F più che numerabili.

L'approccio classico alla definizione di probabilità viene anche detto *a priori*, dato che la probabilità è determinata sulla base di una previsione teorica e non sperimentale. Inoltre, essa è calcolata matematicamente come rapporto tra il numero di casi favorevoli e quello di casi possibili; di conseguenza, per poterla applicare, è indispensabile conoscere a priori non solo questi due valori, ma, soprattutto, sapere se i casi possibili siano tutti egualmente probabili.

Proprio da quest'ultimo concetto parte la critica dei *soggettivisti*: la probabilità di un evento casuale può risultare diversa a seconda di chi si accinge a calcolarla, dato che il giudicare quali eventi siano equipossibili è evidentemente un fatto soggettivo.

Definizione. La probabilità di un evento E, secondo l'individuo I, è il grado di fiducia che I ha che E sia vero. Essa viene misurata dalla quota p di scommessa coerente (ovvero, che non consente una perdita certa a priori per il banco o lo scommettitore), che un individuo equo (disposto, cioè, ad accettare la scommessa senza mutare la somma puntata nel caso in cui da scommettitore diventi banco e viceversa), in base alle informazioni in proprio possesso, giudica giusto pagare per riscuotere l'importo unitario se si verifica E e niente se si verifica E^c.

La critica a questo approccio proviene dal presupposto di carattere soggettivo della definizione: questa, infatti, ha il suo fondamento negli assiomi di coerenza ed equità[5], quindi la criticità sta nel l'assumere implicitamente che l'individuo che decide sia neutrale.

I *frequentisti* hanno tentato di fare un passo avanti, usando la frequenza come pietra portante della definizione di probabilità.

Definizione. Il rapporto $\frac{m}{n} = f_n$, tra il numero delle volte in cui l'evento E si è verificato e il numero delle prove effettuate, viene detto *frequenza* dell'evento E relativa alle n prove fatte.

Esempio. Prendiamo in esame le 14 giostre dal 1932 al 1938[6]. La frequenza dell'evento PSS al quarto posto è $\frac{2}{14} = 0.14$

[5] Riassumibili come una situazione in cui l'individuo valuta la probabilità nell'intervallo $[0, 1]$, attribuendo la probabilità 0 all'evento impossibile e 1 a quello certo (coerenza), e, nello stesso tempo, gli risulta indifferente assumere le parti dello "scommettitore" o del "banco" (equità).

[6] Le sigle in grassetto indicano il Quartiere vincitore.

Data Giostra	I carriera	II carriera	III carriera	IV carriera
domenica 7 agosto 1932	PDF	**PSS**	PSA	PC
domenica 18 settembre 1932	PSA	**PC**	PDF	PSS
domenica 6 agosto 1933	PC	PSA	PSS	**PDF**
domenica 24 settembre 1933	PC	PSS	**PSA**	PDF
domenica 10 giugno 1934	PC	PSA	**PSS**	PDF
domenica 5 agosto 1934	PC	PSA	PDF	**PSS**
domenica 9 giugno 1935	**PC**	PSS	PSA	PDF
mercoledì 7 agosto 1935	PC	PSS	PSA	**PDF**
domenica 14 giugno 1936	PSA	PSS	PC	**PDF**
domenica 9 agosto 1936	PDF	PSS	**PSA**	PC
domenica 13 giugno 1937	PDF	**PC**	PSA	PSS
domenica 8 agosto 1937	PDF	**PSS**	PSA	PC
domenica 12 giugno 1938	**PC**	PSA	PDF	PSS
giovedì 4 agosto 1938	PSA	PDF	PC	**PSS**

Quindi, per definire la probabilità, ci si riconduce ad un fatto sperimentale: se si osserva un gran numero di volte un evento ripetibile, la frequenza relativa di successo si stabilizza attorno ad un valore costante.

Definizione. La *probabilità di un evento* E è il limite a cui tende la frequenza relativa di successo, al divergere del numero delle prove:

$$P(E) = \lim_{n\to\infty} f_n$$

In sostanza, la teoria frequentista poggia sull'ipotesi che tutti i conteggi che si fanno sono relativi ad un solo esperimento, ottenendo una stringa di misure dello stesso evento. Proprio quest'ultimo punto ha originato critiche all'approccio: la misura che ottengo è basata su un troncamento della successione infinita di osservazioni ed i termini della stringa di troncamento, essendo un numero finito (benché, magari, molto grande) non hanno nulla a che fare con il limite. Il limite non può essere assimilato all'osservazione.

Si giunge all'*approccio assiomatico*, che si differenzia dai precedenti perché non è riferito semplicemente ad un protocollo di misure, ma rappresenta la sintesi, l'intersezione tra le teorie sulla probabilità, riassumendone ed esplicitandone le proprietà principali grazie a dei postulati indipendenti logicamente legati tra loro.

Definizione. La funzione $P : F \to \mathbb{R}$, che è una misura definita su una σ-algebra di parti di uno spazio campionario Ω, è detta *misura di probabilità* (o, più semplicemente, *probabilità*) su Ω. Essa soddisfa le seguenti proprietà:

1. $\forall E \in F, 0 \leq P(E) \leq 1$ (Assioma di positività)
2. $P(\Omega) = 1$ (Assioma di certezza)
3. Per ogni coppia di eventi E_1 e E_2 incompatibili, si ha: $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ (Assioma di unione).

4.2 Accenni di Calcolo Combinatorio

Nel caso in cui l'universo, o spazio campionario, sia finito e si giudichi che tutti i suoi n casi elementari siano equipossibili, si ha:

$$1 = \sum_{h=1}^{n} P(e_h) = n \cdot P$$

dove e_h è l'h-esimo evento elementare e P la sua probabilità.

Ovviamente, se ne ricava che $P = \frac{1}{n}$.

Ci si trova, insomma, nell'ambito della probabilità classica ed è opportuno introdurre una formula del Calcolo Combinatorio particolarmente utile per l'argomento dell'estrazione delle carriere.

Definizione. Si chiamano *permutazioni* di n elementi (e si indicano con M_n) le disposizioni di n elementi a gruppi di n, ossia il numero di modi diversi in cui può essere ordinato un insieme di n elementi. In parole più formali: M_n è il numero delle biiezioni da Z_n a Z_n, laddove Z_n è l'insieme dei numeri interi compresi tra 1 e n. Le permutazioni di n elementi sono $M_n = n!$.

Esempio. Il numero di casi possibili per l'estrazione delle carriere si calcola tramite la formula appena definita ed è $4! = 24$. Possiamo dire, quindi, che Ω ha 24 elementi.

(PC, PDF, PSA, PSS), (PC, PSA, PSS, PDF), (PC, PSS, PDF, PSA),
(PC, PSS, PSA, PDF), (PC, PDF, PSS, PSA), (PC, PSA, PDF, PSS),
(PDF, PC, PSA, PSS), (PDF, PSA, PSS, PC), (PDF, PSS, PC, PSA),
(PDF, PC, PSS, PSA), (PDF, PSA, PC, PSS), (PDF, PSS, PSA, PC),
(PSA, PC, PDF, PSS), (PSA, PDF, PSS, PC), (PSA, PSS, PC, PDF),
(PSA, PC, PSS, PDF), (PSA, PDF, PC, PSS), (PSA, PSS, PDF, PC),
(PSS, PC, PDF, PSA), (PSS, PDF, PSA, PC), (PSS, PSA, PC, PDF),
(PSS, PC, PSA, PDF), (PSS, PDF, PC, PSA), (PSS, PSA, PDF, PC).

4.3 Variabile aleatoria e funzione di distribuzione

Un approfondimento del mio discorso merita il concetto di variabile aleatoria: intuitivamente, un numero, esattamente determinato, ma che al momento corrente ci è ignoto. Per delinearne una definizione rigorosa, è opportuno, anche qui, prendere in prestito alcune nozioni di Teoria della Misura.

Definizione. Si dice *spazio probabilizzato* una terna (Ω, F, P), dove Ω è un insieme generico, A è una σ-algebra delle parti di Ω e P è una misura di probabilità su Ω. La coppia (Ω, A) (senza specificare una misura) si dice *spazio misurabile*.

Definizione. Se (Γ, G) ed (Υ, Y) sono spazi misurabili, una funzione $f : \Gamma \to \Upsilon$ si dice *misurabile* se per ogni $A \in Y$ si ha

$$f^{-1}(A) = \{x \in \Gamma \mid f(x) \in A\}$$

Detto questo, se X è una funzione misurabile da uno spazio probabilizzato (Ω, A, P) nello spazio (Γ, G), si dice *variabile aleatoria* a valori in Γ.

Quindi, a seconda dell'insieme di arrivo Γ, otteniamo le due tipologie di variabili aleatorie: se esso è un insieme numerabile (i.e. $\mathbb{N}$ o anche un insieme numerico finito), avremo che la variabile aleatoria è *discreta*; altrimenti, sarà *continua*, avendo come Γ un insieme più che numerabile ($\mathbb{R}$).

Una variabile aleatoria discreta può assumere solo un numero finito o numerabile di valori; in tal caso, ad ogni valore x viene associato il numero $f_X := P(X = x)$ e la funzione f_X è chiamata *funzione di (densità di) probabilità*. Essa si estende a tutti i valori reali, ma ha valore 0 al di fuori del range di X; soddisfa, inoltre, la condizione di normalizzazione $\sum_{X \in \Psi} f_X = 1$, dove Ψ è il range di X: in altre parole, la probabilità che X assuma almeno uno dei valori di Ψ è 1.

Definizione. Si dice *funzione di distribuzione (cumulativa)* della variabile aleatoria X, la funzione $F_X(k) := P(X \leq x_k) = \sum_{i=1}^{k} P(x_i)$, la quale rappresenta la probabilità che X assuma un qualunque valore minore o uguale a x_k. Posto che (x_k) sia una successione crescente, F_X gode delle seguenti proprietà:

1. F_X è una funzione non decrescente
2. $\lim_{k \to +\infty} F_X(k) = 1$
3. $\lim_{k \to -\infty} F_X(k) = 0$
4. F_X è continua a destra, ovvero: $\lim_{k \to z^+} F_X(k) = F_X(z), \forall z \in \mathbb{R}$.

L'estrema importanza della funzione di distribuzione sta nel fatto che tramite essa si possono esprimere tutte le quantità riguardanti la probabilità di X, operando in modo unificato per le variabili discrete e continue (il cui range è tutta la retta reale o un'unione di intervalli sulla retta reale).

Esempio. Contiamo quante volte PSS è stato estratto per primo, nelle 112 estrazioni dal 1948 al 2015. È sufficiente indicare con
$X(\text{"}PSS\_è\_uscito\_per\_primo\text{"}) = 1$ e
$X(\text{"}PSS\_non\_è\_uscito\_per\_primo\text{"}) = 0$;
così, la nostra variabile aleatoria
$X : \{\text{"}PSS\_è\_uscito\_per\_primo\text{"}, \text{"}PSS\_non\_è\_uscito\_per\_primo\text{"}\} \to \{0, 1\}$
ha range $R = \{0, 1\}$.

Chiamando $X_1, X_2, ..., X_{113}$ ognuno degli "esperimenti" e sommandoli, otterremo la variabile aleatoria $X = \sum_{k=1}^{112} X_k$, che, in questo caso, è pari a 36, come illustra la *Tabella delle Estrazioni* a pagg. 45-46. La frequenza empirica di PSS, assimilabile molto grossolanamente, secondo l'approccio frequentista, alla probabilità $P(X = 1)$, è pari a $\frac{36}{112} = 0,32$.

Per il presente lavoro di tesi mi servirò maggiormente di variabili aleatorie discrete, tuttavia mi è d'obbligo fare un breve accenno alla controparte continua, visto che farò uso della cosiddetta approssimazione normale.

Qui non si parla più di conteggio, bensì calcoliamo la probabilità che una variabile aleatoria sia compresa in un certo intervallo; a questo scopo, una semplice somma di valori non è adatta: abbiamo bisogno di integrare. Questo

nonostante il fatto che ogni misura ha un proprio livello di precisione oltre il quale non è possibile andare, ci fornisce cioè un risultato che è un «multiplo intero dell'unità di misura minima» per la quale è tarata[7]. L'enorme insieme di informazioni che può derivare da un'indagine statistica deve essere trattato adeguatamente: è dunque opportuno considerare il range della nostra variabile aleatoria X come un sottoinsieme Z di $\mathbb{R}$, del quale conserva la densità. Ciò porta a dover ammettere due fatti: il primo, che il valore assunto da X sia un numero reale e che quindi l'evento "$X \in Z$" sia certo; il secondo, che l'evento "$X = x$" sia impossibile $\forall x \in Z$, eccetto che per al massimo un sottoinsieme numerabile di Z (ma qui ricadiamo nella condizione discreta di X). Perciò, nonostante la "doppia natura" discreta e continua di alcune variabili aleatorie, è comodo considerarle in senso continuo e il "ponte" che unisce i due fatti che abbiamo appena ammesso è la funzione di densità di probabilità.

La f*unzione di densità di probabilità* f_X è ovviamente definita nel range e non-negativa per ogni punto di esso ed è, per $x \in \Psi$:

$$P(\text{"}X \in \Psi\text{"}) = \int_{\Psi} f_X(x)dx$$

Soddisfa inoltre la condizione di normalizzazione $\int_{\Psi} f_X(x)dx = 1$.

Abbiamo qui implicitamente definito la variabile aleatoria continua: un numero al momento ignoto, associato ad una σ-algebra di parti di $\mathbb{R}$ e tale che, messo al posto di X, per ogni sottoinsieme Ψ appartenente alla σ-algebra, soddisfa quanto appena detto.

La funzione di distribuzione $F_X(x_k) = P(X \leq x_k)$ sarà quindi pari a $\int_{-\infty}^{x_k} f_X(y)dy$ e il valore atteso $E(X) \equiv \int_{-\infty}^{+\infty} x f_X(x)dx$, purché questo integrale sia finito.

[7] Da "Appunti complementari al corso di Probabilità e Statistica" del professor Battinelli.

5. La distribuzione binomiale

Vengono chiamati **bernoulliani** quegli esperimenti, ripetibili e indipendenti, tali che:

1. ognuno di essi sia passibile di soli due risultati (tra loro esclusivi): "successo", S, e "insuccesso", I
2. la loro probabilità rimanga la stessa dal primo all'ultimo esperimento.

Usualmente, denotiamo con p la probabilità di S e q la probabilità di I; ovviamente, p e q devono essere non-negative e tali che $q = 1 - p$.

Per quanto riguarda lo spazio campionario, esso è l'insieme delle 2^n stringhe di n caratteri, scelti arbitrariamente tra S e I; dato che ogni risultato di ognuno degli n esperimenti bernoulliani è indipendente, le probabilità si moltiplicano: ad esempio, $P(IISISSSI...IS) = qqpqpppq...qp$.

Lo schema bernoulliano di esperimenti è un modello teoretico, quindi solo l'esperienza e l'osservazione empirica possono mostrare in quali casi sia adeguato per la descrizione di specifiche osservazioni. In altre parole, fornisce uno standard ideale, impossibile da raggiungere pienamente: il mondo reale sembra comprovare l'indipendenza stocastica di certi esperimenti in successione e non ce ne dobbiamo sorprendere, dato che tutto è soggetto al cambiamento. Tuttavia, entro certi limiti, gli eventi tendono ad uniformarsi ad un modello e, perciò, è importante rilevare, fin dalla prima minima fase, eventuali allontanamenti dallo schema ideale e usarli come indicatori di anomalie.

5.1 Definizione e proprietà della distribuzione

Dato che è automatico notare il numero di successi in una serie di esperimenti, mentre spesso ne viene ignorato l'ordine, è interessante considerare che l'evento "n prove consistono in k successi e $n - k$ insuccessi" può accadere in tanti modi quante sono le possibili disposizioni di k lettere S in n collocazioni. In altre parole, l'evento considerato contiene $\binom{n}{k}$ punti ognuno dei quali, per definizione, ha probabilità $p^k q^{n-k}$. Ciò prova che:

Teorema. Detta $b_{n,p}(k)$ la probabilità che, in n esperimenti bernoulliani ognuno dei k successi abbia probabilità p e ognuno degli $n - k$ insuccessi probabilità q (tali che $p + q = 1$),

$$b_{n,p}(k) = \binom{n}{k} p^k q^{n-k}$$

.

Considerando p come una costante e chiamando S_n il numero di successi[8], avremo $b_{n,p}(k) = P(S_n = k)$;, possiamo ora definire la *distribuzione binomiale* di parametri (n,p):

$$B_{n,p} := (b_{n,p}(k))_{k \in \mathbb{N}_0}$$

[8] $S_n = X_1 + X_2 + ... + X_n$, dove X_i $(1 \leq i \leq n)$ è una variabile aleatoria indipendente che rappresenta un successo nella serie di prove bernoulliane.

Essa viene chiamata così perché rappresenta il k-esimo termine dell'espansione binomiale di $(p+q)^n$. Cioè:

$$\sum_{k\in\mathbb{N}_0} b_{n,p}(k) = (p+q)^n = 1$$

Se si parla di distribuzione binomiale, è naturale parlare anche di termine centrale e code[9].

Partendo dalla formula sopra, possiamo notare che, al variare di p, la distribuzione mantiene un comportamento "unimodale". Infatti, nel caso in cui $p = \frac{1}{2} = q$, le componenti di $B_{n,p}$ sono disposte simmetricamente rispetto a quella (o quelle) centrali: dato che $\forall k \in \mathbb{N}_0$, $\binom{n}{k} = \binom{n}{n-k}$, a meno del fattore di normalizzazione 2^{-n}, si origina una sequenza di coefficienti binomiali $\left(\binom{n}{k}\right)_{k\in\mathbb{N}_0}$:

$$M_{n,\frac{1}{2}} \equiv \max_{\mathbb{N}_0} b_{n,\frac{1}{2}}(k) = \begin{cases} b_{n,\frac{1}{2}}(j)=\binom{n}{j}\cdot 2^{-n} \text{ se } n=2j \\ b_{n,\frac{1}{2}}(j)=b_{n,\frac{1}{2}}(j+1)=\binom{n}{j}\cdot 2^{-n} \text{ se } n=2j+1 \end{cases}$$

Al variare di p e q all'interno di $[0,1]$, il rapporto

$$\begin{aligned} \beta(k) &\equiv \frac{b_{n,p}(k)}{b_{n,p}(k-1)} \\ &= \frac{\frac{n!}{(n-k)!k!}}{\frac{n!}{(n-k+1)!(k-1)!}} \cdot \frac{p^k q^{n-k}}{p^{k-1}q^{n-k+1}} \\ &= \frac{(n-k+1)}{k}\cdot\frac{p}{q} \\ &= \frac{(n+1)p - k(1-q)}{kq} \\ &= 1 + \frac{(n+1)p-k}{kq} \end{aligned}$$

è minore o maggiore di 1 a seconda del valore di k: se k è minore o maggiore di $(n+1)p$, di modo che, finché $\frac{(n+1)p-k}{kq}$ è positivo, $\beta(k) > 1$ e la distribuzione sarà crescente; altrimenti, $\beta(k)$ sarà maggiore di 1 e, avendo $b_{n,p}(k) < b_{n,p}(k-1)$, l'andamento della $B_{n,p}$ sarà decrescente.

Tutto dipende, più in particolare, dalle realzioni tra i due parametri n e p, il che origina tre casi, nella totalità dei quali

$$m_{n,p} = \lfloor (n+1)p \rfloor$$

(dove $\lfloor (n+1)p \rfloor$ sta per "parte intera inferiore di $(n+1)p$"), è il numero (appartenente all'insieme dei massimizzatori) che rende massimo il valore di $B_{n,p}$. Abbiamo, in sostanza, che, a seconda del "comportamento" di n e p,

[9] Le seguenti considerazioni sono state estratte dalle dispense del professor Andrea Battinelli: "Sulla distribuzione binomiale e le sue code".

$0 \leq m_{n,p} \leq n$ e qui notiamo che, a seconda se approcciamo la trattazione dalla parte di un parametro o di un altro, la strada si fa più o meno impervia.

Iniziamo con quella più "dolce": nel caso in cui sia molto basso il valore di p, $p < \frac{1}{n+1}$, abbiamo che il rapporto $\beta(k)$ tra componenti adiacenti è minore di 1. Possiamo considerarlo come primo caso, in cui $m_{n,p} = 0$ e la binomiale è uniformemente decrescente. Quando, invece, $p > \frac{n}{n+1}$, $\beta(k)$ è maggiore di 1 e la $B_{n,p}$ cresce fino al raggiungimento del valore soglia corrispondente a $m_{n,p} = n$; il terzo caso, una via di mezzo tra il primo e il secondo, prevede $\beta(k)$ sempre positivo, ma maggiore di 1 fino a $m_{n,p}$ (binomiale crescente), per poi diventare minore di 1 (binomiale descrescente) una volta oltrepassato il "picco"; è il caso in cui p è compreso tra $\frac{1}{n+1}$ e $\frac{n}{n+1}$.

Leggermente più laborioso è parlare di n. Anche qui i casi sarebbero tre, ma diventano sei, essendo necessario dividere la trattazione in due sottocasi: $p > q$ e $p < q$. Laddove, dunque, $n < \frac{1}{p} - 1 = \frac{q}{p}$, con $p > q$, rientriamo nel primo caso sopracitato; per $\frac{q}{p} < n < \frac{p}{q}$ nel terzo caso; quando, infine, n è maggiore di $\frac{p}{q}$, la binomiale risulta sempre crescente. Le cose cambiano se, invece, $p < q$: troviamo il primo caso nella porzione degli n compresi tra $\frac{p}{q}$ e $\frac{q}{p}$ (poiché qui, ovviamente, $\frac{p}{q} < \frac{q}{p}$), mentre il terzo caso si verifica quando $n > \frac{q}{p}$. Rimane da notare che, per quegli n minori di $\frac{p}{q}$, $B_{n,p}$ è uniformemente decrescente.

Spostando l'attenzione su $m_{n,p}$, se questi è un numero intero, avremo due "picchi" massimi di $B_{n,p}$, che formano una sorta di altopiano nel grafico e sono denominati $b_{n,p}(m_{n,p})$ e $b_{n,p}(m_{n,p}-1)$; altrimenti, la soglia massima è la componente con massimo indice inferiore a $(n+1)p$, ovvero $\lfloor (n+1)p \rfloor$. Per definizione di parte intera inferiore, $\lfloor (n+1)p \rfloor = (n+1)p$ se quest'ultimo è un numero intero; avremo, perciò

$$\lfloor (n+1)p \rfloor \leq (n+1)p \leq \lfloor (n+1)p \rfloor + 1$$

il che ci riporta a confermare che $m_{n,p} = \lfloor (n+1)p \rfloor$.

Dato che la distribuzione binomiale $B_{n,p}$ descrive una variabile aleatoria S_n che è somma di n variabili aleatorie indipendenti X di uguale legge di Bernoulli B_p, possiamo ricavare molte caratteristiche di S_n da quelle di X:

- Valore atteso:

$$\begin{aligned} E(S_n) &= \sum_{i=1}^{n} E(X_i) = nE(X) \\ &= n\sum_{k\in\mathbb{N}} \binom{n}{k} p^k q^{n-k} \\ &= n\sum_{k\in\mathbb{N}} \frac{n!}{(n-k)!k!} p^k q^{n-k} \\ &= \sum_{k\in\mathbb{N}} k\frac{n(n-1)!}{(n-k)!k(k-1)!} p^{(k-1)} pq^{(n-1)-(k-1)} \\ &= np\sum_{k\in\mathbb{N}} \binom{n-1}{k-1} p^{(k-1)} q^{(n-1)-(k-1)} \\ &= np(p+q)^{(n-1)} \\ &= np \end{aligned}$$

- Varianza:

$$\begin{aligned} Var(S_n) &= \sum_{i=1}^{n} Var(X_i) \\ &= nVar(X) \\ &= n[E(X^2) - E^2(X)] \\ &= n(p-p^2) \\ &= n[p(1-p)] \\ &= npq \end{aligned}$$

Tornando al valore "soglia", al termine centrale $m_{n,p}$, esso si può definire, in parole povere, come il numero di successi più probabile. Ciò, tuttavia, è difficilmente osservabile: se eseguiamo molti esperimenti (ovvero, se n è molto grande), tutti i $b_{n,p}(k)$ saranno molto piccoli.

È interessante, quindi, tramite la funzione cumulativa, stimare la probabilità dell'evento "$S_n \geq r$", ossia "misurare" la coda destra di $B_{n,p,}$ laddove $r > np$. Ciò è possibile maggiorando le componenti $(b_{n,p}(k))_{k\in\mathbb{N}\cap(r-1)}$ tramite quelle appartenenti ad una particolare progressione geometrica, il rapporto tra le quali è un numero indipendente da k e strettamente compreso tra $\beta(k)$ e 1:

$$1 - \frac{r-np}{rq}$$

Dunque, $P(S_n \geq r)$ sarà uguale a:

$$\begin{aligned}\sum_{k=r}^{n} b_{n,p}(k) &< \sum_{k=r}^{n} b_{n,p}(r)(1-\frac{r-np}{rq})^{k-r} \\ &< \sum_{k=r}^{+\infty} b_{n,p}(r)(1-\frac{r-np}{rq})^{k-r} \\ &= b_{n,p}(r)\frac{rq}{np-r}\end{aligned}$$

A questo punto, manca da fare una stima di $b_{n,p}(r)$: essendo un valore maggiorato da tutte le componenti di $B_{n,p}$ aventi indice compreso tra $m_{n,p}$ e $r-1$,

$$1 > \sum_{k=m_{n,p}}^{r} b_{n,p}(k) > \sum_{k=m_{n,p}}^{r} b_{n,p}(r) = (r - m_{n,p} + 1)b_{n,p}(r)$$

perciò:

$$b_{n,p} < \frac{1}{r - m_{n,p} + 1} < \frac{1}{r - np}$$

infatti: $m_{n,p} - 1 < m_{n,p} - p = \lfloor (n+1)p \rfloor - p \leq (n+1)p - p = np$.
Allora:

$$P(S_n \geq r) < \frac{rq}{(np-r)^2}$$

Dualmente, la stima della coda sinistra, ovvero della probabilità di avere al massimo r successi, sarà:

$$P(S_n \leq r) < \frac{(n-r)p}{(np-r)^2}$$

con $r < np$.

5.2 La legge "debole" dei grandi numeri

È anacronistico vedere il mondo reale esclusivamente attraverso le lenti di una teoria matematica formale, anche se essa fornisce un modello teoretico a cui fare riferimento per spiegare il fenomeno studiato. Per questo, spesso è più utile conoscere il numero medio di successi in una successione di n prove bernoulliane, piuttosto che il semplice numero di successi. Considerando, infatti, la famiglia di numeri aleatori.$\left(\frac{S_n}{n}\right)_{n\in\mathbb{N}}$, essa è comunque descritta dalla famiglia di distribuzioni $(B_{n,p})_{n\in\mathbb{N}}$ ed è più facile confrontare fra loro le componenti di quest'ultima: infatti, il dominio di ciascuna $\frac{S_n}{n}$ è contenuto in $[0,1]$. Non solo: il valore atteso rimane invariato

$$E(\frac{S_n}{n}) = \frac{1}{n}E(S_n) = \frac{np}{n} = p$$

mentre la varianza è inferiore (oltre che infinitesima, al divergere di n):

$$Var(\frac{S_n}{n}) = \frac{1}{n^2}Var(S_n) = \frac{npq}{n^2} = \frac{pq}{n}$$

Un altro aspetto particolarmente interessante è che, osservando la forma della distribuzione, la maggior parte della probabilità si accumula nella parte centrale, scemando via via lungo le code. In pratica, se associamo al valore r precedentemente trattato la quantità $n(p+\varepsilon)$, posto $\varepsilon > 0$ un numero fissato e arbitrariamente piccolo, e vogliamo conoscere la probabilità che $\frac{S_n}{n} \geq (p+\varepsilon)$, scriveremo

$$P(\frac{S_n}{n} \geq (p+\varepsilon)) = P(S_n \geq n(p+\varepsilon)) < \frac{n(p+\varepsilon)q}{n^2\varepsilon^2} = \frac{(p+\varepsilon)q}{n\varepsilon^2}$$

quindi si può notare che essa è , oltre che uguale a $P(S_n \geq n(p+\varepsilon))$, anche maggiore di $\frac{1}{n\varepsilon^2}$. Ciò ci porta a dire che, per $n \to +\infty$,

$$P(S_n \geq n(p+\varepsilon)) \to 0$$

Dualmente, volendo stimare la probabilità della coda sinistra, avremo che:

$$P(\frac{S_n}{n} \leq (p-\varepsilon)) = P(S_n \leq n(p+\varepsilon)) < \frac{n(p-\varepsilon)p}{n^2\varepsilon^2} = \frac{(p-\varepsilon)p}{n\varepsilon^2}$$

e che:

$$\lim_{n\to+\infty} P(S_n < n(p-\varepsilon)) = 0$$

Spostando l'attenzione verso l'evento complementare alle due code, otteniamo che

$$P(\left|\frac{S_n}{n} - p\right| < \varepsilon) > 1 - \frac{(p+\varepsilon)q + (p-\varepsilon)p}{n\varepsilon^2} = 1 - \frac{p - (1-2p)\varepsilon}{n\varepsilon^2}$$

e, per $n \to +\infty$,

$$P(\left|\frac{S_n}{n} - p\right| < \varepsilon) \to 1$$

In parole povere, all'aumentare di n, la probabilità che il numero medio di successi devii da p più di un numero ε prefissato tende a 0. Questo discorso sta alla base della *Legge "debole" dei grandi numeri per la distribuzione binomiale*:

Proposizione. Sia $(E_n)_{n\in\mathbb{N}}$ una famiglia di eventi equipossibili e stocasticamente indipendenti, ciascuno con probabilità p. $\forall n \in \mathbb{N}$, il numero aleatorio $\frac{S_n}{n} \equiv \frac{1}{n}\sum_{k\in\mathbb{N}} E_k$ rappresenta la frequenza di successi relativi ai singoli eventi e la sua distribuzione è $B_{n,p} = (\binom{n}{k} \cdot p^k \cdot q^{n-k})_{n\in\mathbb{N}_\circ}$ sul dominio $D_n \equiv \left\{\frac{k}{n}\right\}_{k\in\mathbb{N}_0} \subseteq [0,1]$; la sua media è $E(\frac{S_n}{n}) = p$ e la sua varianza $Var(\frac{S_n}{n}) = \frac{pq}{n}$. Per un certo $\varepsilon \in (0, \min\{p,q\})$ fissato, la probabilità che $\frac{S_n}{n}$ non appartenga all'intorno aperto di centro p e raggio ε è infinitesima al divergere di n, ovvero

$$\lim_{n\to+\infty} P(\left|\frac{S_n}{n} - p\right| < \varepsilon) = 1$$

Questa legge viene definita "debole" perché ha in sè delle defaillances, che la rendono di interesse limitato: la successione delle frequenze converge in probabilità ad una costante, mentre è più utile verificare la convergenza quasi certa della successione stessa. A questo scopo è nata la legge "forte" dei grandi numeri.[10]

[10] **Legge "forte" dei grandi numeri**: $\forall \varepsilon > 0, \left|\frac{S_n}{n} - p\right| < \varepsilon$

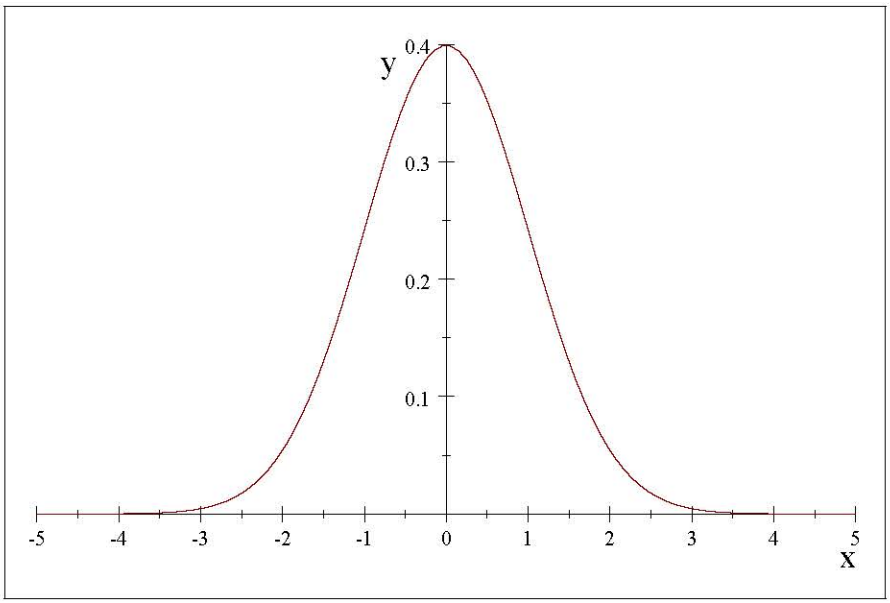
5.3 Approssimazione tramite la distribuzione normale

Tra i vari utilizzi della distribuzione normale troviamo l'approssimazione della distribuzione binomiale (nonché di quella di Poisson), che ne fornisce un supporto di notevole importanza teoretica e pratica. Prima di entrare nel nocciolo dell'argomento, è opportuno fissare alcuni concetti chiave:

Definizione. Si dice *funzione di densità normale* la funzione

$$\phi_{\mu,\sigma} : x \longmapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

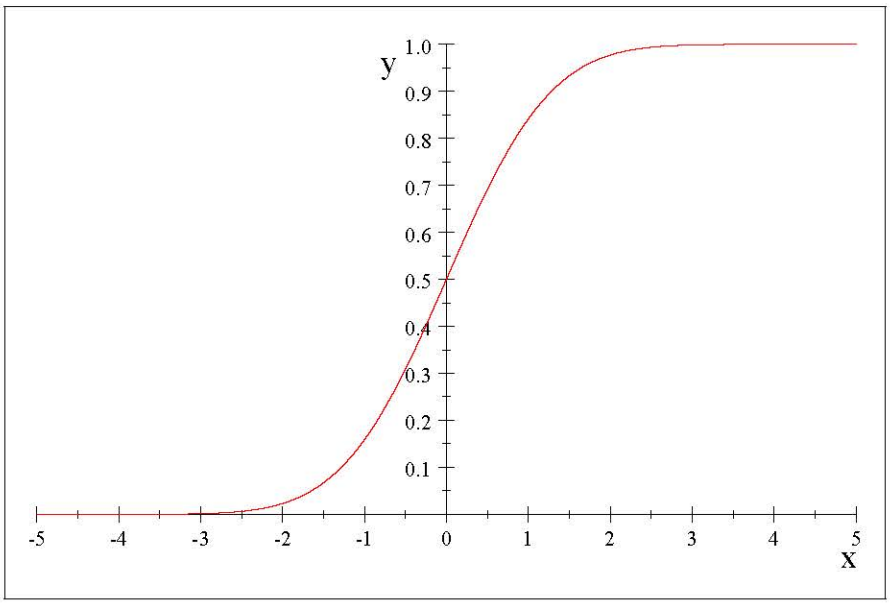
in cui μ è il valore atteso e σ^2 la varianza. Ecco qui sotto il grafico della funzione di densità normale standardizzata, avente $\mu = 0$ e $\sigma = 1$:



La primitiva di $\phi_{0,1}$ è

$$\Phi : x \longmapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy$$

e viene detta *funzione di distribuzione normale (standardizzata)*:



Il grafico di $\phi(x)$ è la famosa "campana" di Gauss, avente valore massimo per $\phi(x) = \frac{1}{2\pi}$, mentre quello di Φ è l'altrettanto famosa "curva ad esse", simmetrica rispetto al suo punto di intersezione con l'asse verticale, di coordinate $\left(0, \frac{1}{2}\right)$.

Teorema. L'area compresa tra il grafico di ϕ e l'asse delle ascisse è

$$\int_{-\infty}^{+\infty} \phi(x)dx = 1$$

Dimostrazione. Considerando

$$\begin{aligned} \left\{\int_{-\infty}^{+\infty} \phi(x)dx\right\}^2 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \phi(x)\phi(y)dxdy \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x^2+y^2)}dxdy \end{aligned}$$

e, passando alle coordinate polari,otteniamo la tesi:

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_0^{+\infty} e^{-\frac{1}{2}r^2} r\,dr &= \int_0^{+\infty} e^{-\frac{1}{2}r^2} r\,dr \\ &= \left[-e^{-\frac{1}{2}r^2}\right]_0^{+\infty} = 1 \end{aligned}$$

Quanto appunto verificato è importante per poter dire che $\Phi(x)$ è monotona crescente con immagine $[0,1]$;ϕ è inoltre una funzione pari, pertanto vale che:

$$\Phi(-x) = 1 - \Phi(x)$$

e questo in primo luogo spiega l'asserita simmetria del grafico di Φ, ed in secondo luogo evidenzia l'importanza della stima del valore $1 - \Phi(x)$ per x grande:

Teorema. Per $x \to +\infty$, $1 - \Phi(x) \sim x^{-1}\phi(x)$; laddove, $\forall x > 0$,

$$\left[x^{-1}x^{-3}\right]\phi(x) < 1 - \Phi(x) < x^{-1}\phi(x)$$

In generale, se una variabile aleatoria X_n è distribuita secondo la distribuzione binomiale e n è un numero molto grande, possiamo dire che la nostra X_n è distribuita approssimativamente tramite la distribuzione normale (standard). Questo è il significato del:

Teorema di DeMoivre-Laplace. Sia $X_n \sim B_{n,p}$. Il valore atteso di X_n è np, la varianza di X_n è npq, e per ogni coppia $(a, b) \in R^2$ fissata con $a < b$, vale

$$\lim_{n \to +\infty} P\left(a < \frac{X_n - np}{\sqrt{npq}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{z^2}{2}} dz$$

con a, b fissati.

È usuale considerare l'approssimazione gaussiana di una distribuzione binomiale con $p = \frac{1}{2}$, tuttavia, il caso specifico trattato in questa tesi ci spinge a considerare il problema nell'ottica di $p = \frac{1}{4}$: stiamo per lavorare su una distribuzione che si fonda su un consistente numero di prove in cui ciascuno dei quattro Quartieri ha la stessa probabilità di uscire, quindi $p = \frac{1}{4}$.

6. Un'indagine statistica sul Saracino

In un primo momento, ad un lettore distratto, può sembrare strano che la statistica sia accostata ad un "torneamento" medievale. Tuttavia, la Matematica è nata allo scopo di spiegare la realtà che ci circonda: spesso, sia i risultati delle indagini che le indagini stesse si rivelano astratti, devono quindi essere confrontati con l'esperienza per poter trovare affinità e contrasti; dopo vari passaggi e verifiche, un po' come delle serpentine di un alambicco, esce, come un elisir, la teoria, dalla quale poi è possibile ramificare gli utilizzi pratici.

Nel caso specifico, si è voluto applicare un modello teoretico statistico a 112 cerimonie di estrazione delle carriere, riducibili a una serie di esperimenti bernoulliani. Si è scelto di considerare le estrazioni dalla prima Giostra del Dopoguerra (settembre 1948) in poi per motivi di precisione scientifica: le numerose fonti consultate non hanno potuto fornire informazioni sicure riguardo ad alcune carriere nelle Giostre degli anni '30 e del 1940, perciò si è ritenuto opportuno non considerarle nella nostra indagine.

Va detto, comunque, che la cerimonia di estrazione delle carriere, nonostante le modifiche del canovaccio apportate durante gli anni, è stata sempre pubblica, oltre ad essere avvenuta sempre tramite operazioni di estrazione "imparziale". Secondo documenti e personaggi autorevoli, le prime estrazioni consistevano nel pescare da una teca di vetro un piccolo rotolo di pergamena sigillato, all'interno del quale, una volta aperto dal Sindaco, si poteva leggere il nome del Quartiere estratto; successivamente, si è fatto strada un nuovo metodo, che è quello che sopravvive tuttora: un paggetto estrae da un sacco di cuoio retto dal Cancelliere una pallina apribile, al cui interno il Sindaco scopre i colori del Quartiere designato dalla Fortuna. Ciò ci rassicura riguardo alla scientificità del campionamento che prendiamo in esame: siamo in presenza di 112 estrazioni casuali senza reimmissioni.

6.1 I Quartieri in posizione

Dissezioniamo attimo per attimo il procedimento dell'estrazione delle carriere. Poniamo, ad esempio, di tenere per il Quartiere di Porta Santo Spirito. Il primo paggetto sale sul palco, infila la mano nel sacchetto di cuoio ed estrae una pallina, che il Cancelliere porge al Sindaco; quest'ultimo la apre e i colori non sono giallo-blu. Quindi, se prima dell'estrazione la probabilità che il nostro Quartiere venisse estratto era pari ad $\frac{1}{4}$, adesso è diventata $\frac{1}{3}$: tolto un "avversario", sono rimasti tre posti liberi. Sale il secondo paggetto e la seconda pallina estratta è anch'essa di un altro Quartiere: la probabilità diventa $\frac{1}{2}$. Se poi il terzo paggetto estrae una pallina che non contiene i colori del Santo Spirito, abbiamo la certezza che il quarto paggetto estrarrà la nostra pallina. A parte il fatto che faremmo salti di gioia[11], avremmo anche notato l'andamento del valore

[11] Perché l'ultimo posto è il più ambìto, dato che, correndo per ultimi, i cavalieri possono regolare la propria strategia in base ai punteggi degli avversari precedenti.

della probabilità a seconda di una precisa concatenazione di eventi indipendenti. Abbiamo toccato con mano il concetto di probabilità condizionata.

Concetto che si fonda sulla dipendenza logica di un evento nei confronti di un altro ed esprime la probabilità che un evento A si verifichi, nell'ipotesi che un altro evento B sia avvenuto. In sostanza, B influenza il grado di fiducia che fissiamo riguardo all'avverarsi di A.

Vediamo la formula per calcolare la probabilità condizionata:

Definizione. Sia B un evento tale che $P(B) > 0$. Viene chiamato *probabilità condizionata di un evento A rispetto all'evento B* il rapporto

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Se A e B sono incompatibili (ovvero $A \cap B = \emptyset$), $P(A \mid B) = 0$, mentre, nel caso in cui $P(B) = 0$, $P(A \mid B)$ è indefinita (non a caso, «ex falso, quodlibet»).

Vediamo un'applicazione della formula appunto definita: se consideriamo l'evento "PSA alla prima estrazione" come l'evento B appena definito, la probabilità dell'evento A_1[12] "PSS alla seconda estrazione" sarà

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{\frac{1}{3} \cdot \frac{1}{4}}{\frac{1}{4}} = \frac{1}{3}$$

il che, oltretutto, conferma la nostra dissezione dell'estrazione.

Ciò ha ovviamente anche ripercussioni sul calcolo delle probabilità della distribuzione binomiale che possiamo applicare al fenomeno: l'estrazione di un qualsiasi Quartiere sottosta alla probabilità condizionata. Ogni volta che il paggetto infila la mano nel sacchetto e tira fuori una pallina, siamo in presenza di un processo bernoulliano: la pallina può rivelarsi dei colori del nostro Quartiere ("successo") oppure no ("insuccesso"); tuttavia, nello stesso tempo, entra in gioco anche qualcos'altro: questo "insuccesso" assume una precisa probabilità (pari a $\frac{3}{4}$ se siamo alla prima estrazione, a $\frac{2}{3}$ se siamo alla seconda e la nostra pallina è ancora nel sacchetto e così via), che dipende dagli eventi precedenti. La probabilità del "successo" è, naturalmente, complementare.

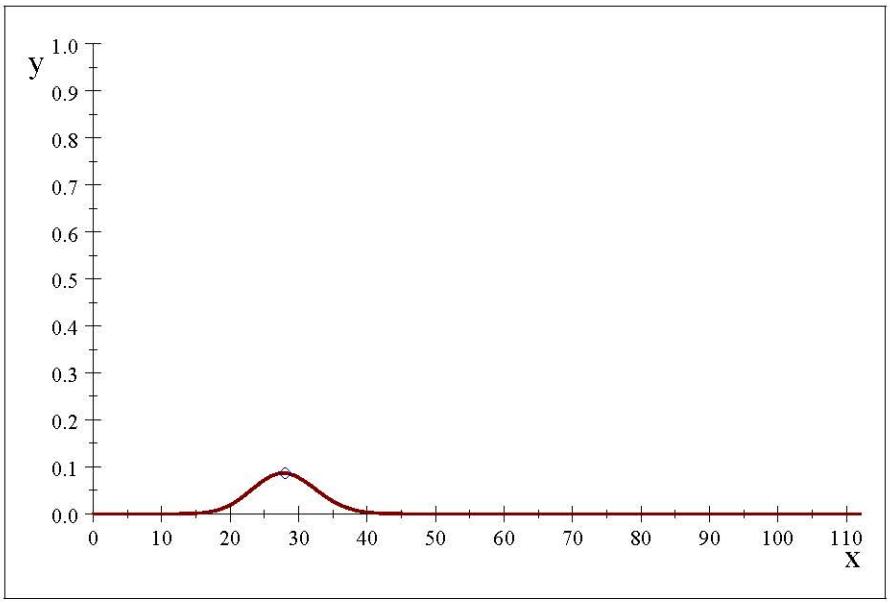
Così, il calcolo delle $b_{n,p}(k)$ avviene tenendo conto di questo: alla prima estrazione, abbiamo $\binom{n}{k} p^k q^{n-k}$ con $p = \frac{1}{4}$ e $q = 1 - \frac{1}{4} = \frac{3}{4}$; alla seconda, tenendo conto che dobbiamo moltiplicare la probabilità di non uscire alla prima estrazione $(\frac{3}{4})$ con la probabilità di uscire alla seconda estrazione $(\frac{1}{3})$: $p = \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}$ e, quindi, $q = \frac{3}{4}$. Discorso analogo per la terza estrazione: moltiplicare la probabilità di non uscire alla prima estrazione $(\frac{3}{4})$ per la probabilità di non uscire alla seconda $(\frac{2}{3})$ per la probabilità di uscire alla terza $(\frac{1}{2})$, ottenendo $p = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{4}$ e $q = \frac{3}{4}$. E così anche per la quarta estrazione: probabilità di non uscire alla prima estrazione $(\frac{3}{4})$ per la probabilità di non uscire alla seconda $(\frac{2}{3})$ per la probabilità di non uscire alla terza $(\frac{1}{2})$ per la probabilità di uscire alla quarta: $p = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{4}$ e, naturalmente, $q = \frac{3}{4}$.

[12] che è uno dei tre eventi elementari dell'evento A:"Viene estratto un Quartiere che non è Porta Sant'Andrea", di probabilità unitaria.

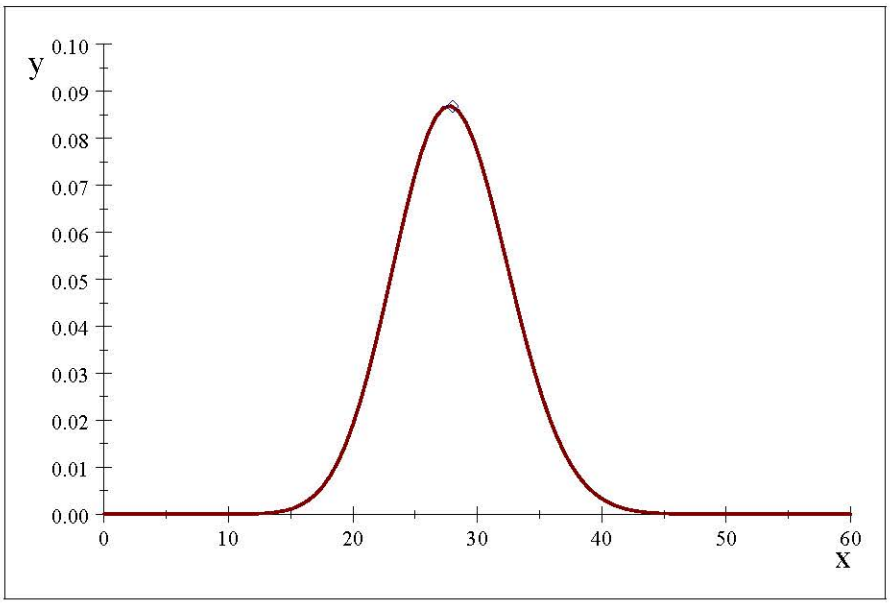
Osserviamo, dunque, la distribuzione di probabilità dei Quartieri in ognuna delle quattro posizioni, tramite la distribuzione binomiale. Abbiamo $p = \frac{1}{4}$ e $q = \frac{3}{4}$ e, dato che dal 1948 al 2015 si sono svolte 112 Giostre del Saracino (con relative estrazioni), $n = 112$.

$$B_{112,0.25}(k) = \sum_{k=0}^{112} \binom{112}{k} (0.25)^k (0.75)^{112-k}$$

Ecco il grafico della funzione di distribuzione per questi parametri:



Sull'asse delle ascisse è riportato il numero di prove, mentre in quello delle y la probabilità di "successo". Notiamo subito che il picco massimo di probabilità di "successo" $m_{112,\frac{1}{4}}$ è 28: in parole povere, è alla ventottesima prova che abbiamo la massima probabilità che un Quartiere sia estratto nella posizione considerata; ciò vale, ovviamente, per qualsiasi Quartiere e per qualunque posizione delle quattro disponibili. Notiamo, inoltre, che fino all'incirca alla quindicesima estrazione la probabilità di essere estratti è molto bassa, come anche accade dopo la quarantesima prova. Il comportamento crescente prima del "picco" e decrescente dopo il "picco" è evidente, restringendo la probabilità massima a 0.1.e le prove effettuate a 60:



Continuiamo la nostra riflessione procedendo in ordine alfabetico.

Osserviamo i "successi" del Quartiere di Porta Crucifera: potremmo calcolare la frequenza di uscita alla prima estrazione in base alle occorrenze raccolte in 76 anni di estrazioni e notare che è pari a $\frac{24}{112} = 0.21$, dato che PC si è verificato 24 volte; la binomiale, invece, ci fornisce una probabilità, che è pari a $b_{112,0.25}(24) = \binom{112}{24}(0.25)^{24}(0.75)^{112-24} = 0.061716$. Passiamo alla seconda estrazione: il "caso" ha voluto che il Quartiere sia stato estratto esattamente 24 volte, proprio la stessa quantità che abbiamo appena utilizzato. La cosa non continua a valere per la terza estrazione: in questo caso, $k = 33$, quindi $\frac{33}{112} = 0.29$, mentre $b_{112,0.25}(33) = 0.046451$. Per quanto riguarda la probabilità binomiale per la quarta estrazione, $b_{112,0.25}(31) = 0.068128$, essendo PC avvenuto per 31 volte su 112 (e quindi con frequenza $\frac{31}{112}$).

Passiamo ora al Quartiere di Porta del Foro. Avendo rilevato che è stato estratto, alla prima estrazione, 26 volte, calcoliamo la frequenza che è pari a 0.23 e la binomiale che invece vale 0.080769; a fronte dei 38 "successi"[13] alla seconda estrazione per PDF, avremo che $b_{112,0.25}(38) = 0.008583$, laddove $\frac{38}{112} = 0.34$. La terza estrazione vede $k = 23$, quindi la binomiale sarà pari a 0.049928 e la frequenza esattamente 0.2. Rimane, per Porta del Foro, la quarta estrazione: abbiamo $b_{112,0.25}(25) = 0.072414$, mentre $\frac{25}{112} = 0.22$.

Osserviamo l'andamento delle estrazioni del Quartiere di Porta Sant'Andrea. Partendo dalla prima, notiamo che ci sono stati 26 successi, quindi i calcoli che ne seguono sono identici a quelli per la prima estrazione di Porta del Foro. Rileviamo, invece, che PSA è avvenuto 29 volte alla seconda estrazione: $b_{112,0.25}(29) = 0.083784$ e $\frac{29}{112} = 0.26$; per quanto riguarda la terza estrazione, abbiamo frequenza pari a $\frac{30}{112} = 0.27$, mentre $b_{112,0.25}(30) = 0.077267$. I "suc-

[13] (il numero più alto tra i sedici valori di k raccolti sul "campo")

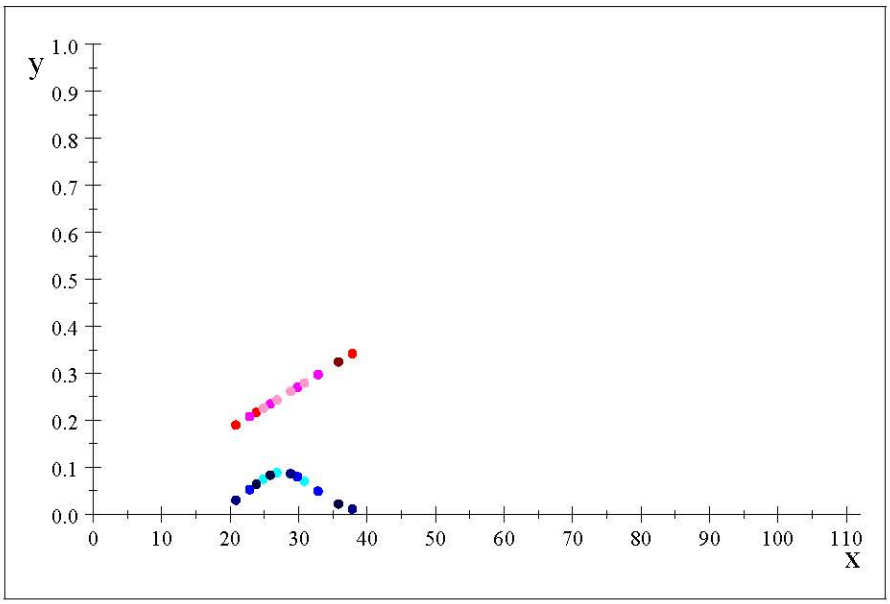
cessi" alla quarta estrazione sono stati 27, il che ci porta ad assegnare alla frequenza valore 0.24 e alla binomiale 0.085755.

Volgiamoci verso il Quartiere di Porta Santo Spirito: in 76 anni di cerimonie, PSS è avvenuto, alla prima estrazione, 36 volte. Potremmo dire che la frequenza osservata, data dal rapporto $\frac{k}{n}$, è dunque 0.32; un valore di probabilità più preciso ci viene fornito dalla binomiale $b_{112,0.25}(36) = \binom{112}{36}(0.25)^{36}(0.75)^{112-36} = 0.019054$. Passiamo alla seconda estrazione: rispetto a $\frac{21}{112} = 0.19$, abbiamo $b_{112,0.25}(21) = 0.027762$; invece, per la terza estrazione, la probabilità fornitaci dalla binomiale è pari a 0.080769, essendo stato estratto 26 volte (la frequenza osservata è quindi 0.23). Ci resta da calcolare la probabilità binomiale di uscire alla quarta estrazione, ovvero $b_{112,0.25}(29) = 0.083784$, laddove la frequenza si attesta su 0.26.

Per uno sguardo complessivo, possiamo creare una tabella di valori, divisa per ordine di estrazione, Quartiere, numero di "successi" in quella data estrazione, probabilità binomiale corrispondente e frequenza corrispondente:

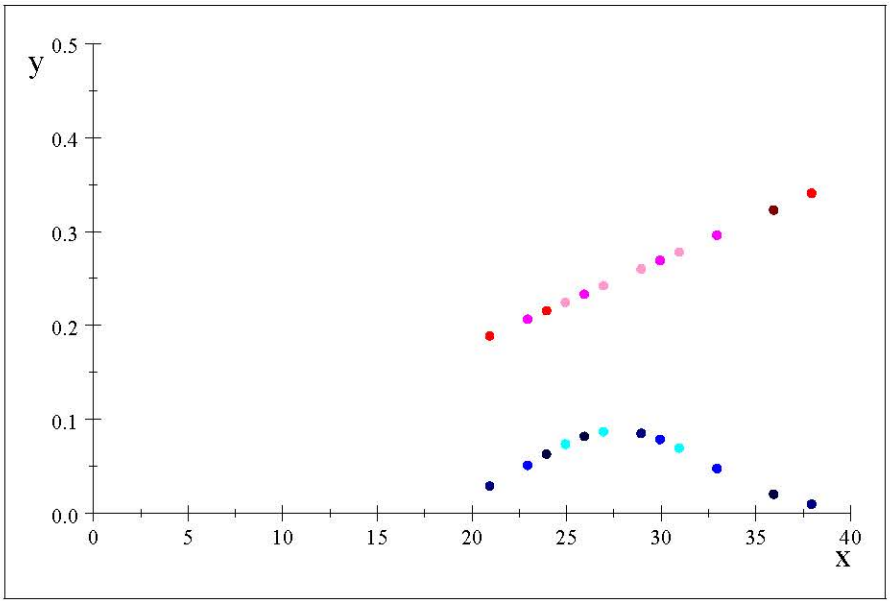
Estrazione	Quartiere	k	$b_{n,p}(k)$	$\frac{k}{n}$
I	PC	24	0.061716	0.21
	PDF	26	0.080769	0.23
	PSA	26	0.080769	0.23
	PSS	36	0.019054	0.32
II	PC	24	0.061716	0.21
	PDF	38	0.008583	0.34
	PSA	29	0.083784	0.26
	PSS	21	0.027762	0.19
III	PC	33	0.046451	0.29
	PDF	23	0.049928	0.2
	PSA	30	0.077267	0.27
	PSS	26	0.080769	0.23
IV	PC	31	0.068128	0.28
	PDF	25	0.072414	0.22
	PSA	27	0.085755	0.24
	PSS	29	0.083784	0.26

In base a questi valori, possiamo anche creare un grafico, nelle cui ascisse troviamo le prove effettuate, mentre le ordinate rappresentano il livello di probabilità corrispondente:

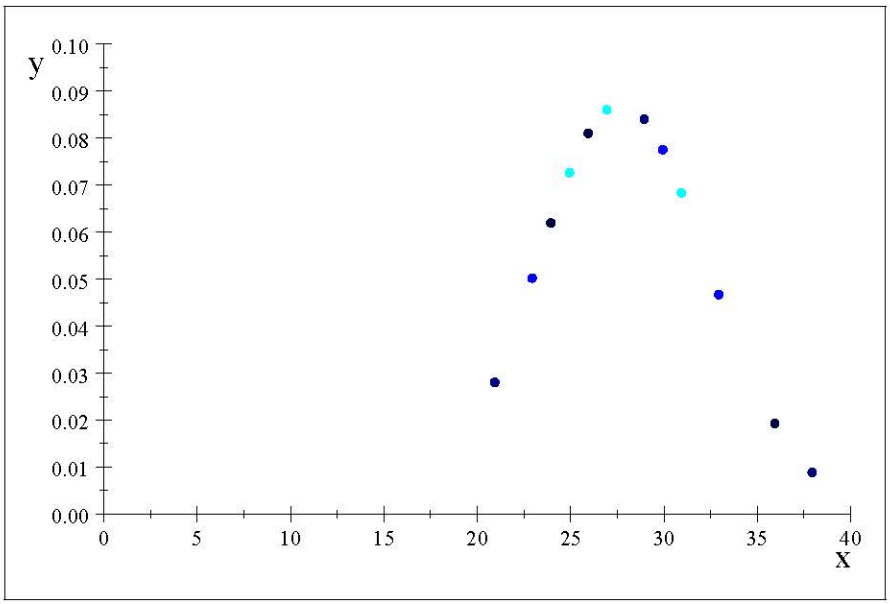


In rosa troviamo il grafico delle frequenze, che si allunga esattamente lungo la retta $y = \frac{x}{112}$; il grafico delle binomiali è invece quello in blu, che ricalca l'andamento già osservato teoricamente all'inizio del capitolo. Abbiamo infatti, anche qui, alla ventottesima "prova" il picco massimo di probabilità di essere estratti

Per uno sguardo più ravvicinato, tagliamo gli assi: a 40 per il numero di estrazioni, a 0.5 per la probabilità:



Prendendoci qualche licenza con la legge dei grandi numeri (112 è un numero irrisorio in confronto all'infinito), possiamo dire che, all'aumentare del numero delle prove, il grafico rosa collasserà come una nebulosa in una nana bianca di ascissa 28 e di ordinata 0.25 (la probabilità "a priori" $\frac{1}{4}$) e l'"ogiva" blu si acutizzerà mantenendo la "punta" sul valore massimo ottenuto con $m_{n,p}$. Ciò è già visibile restringendo l'asse delle y del grafico a $[0, 0.1]$:



6.2 La funzione cumulativa

Nella parte dei Concetti Fondamentali abbiamo affrontato la questione sulla probabilità che la variabile aleatoria presa in esame sia minore di un certo valore $r < np$. Ed abbiamo ottenuto che:

$$P(S_n \leq r) < \frac{(n-r)p}{(np-r)^2}$$

Vediamone un'applicazione pratica: calcoliamo la probabilità che S_n sia minore o uguale di $r = 20$.

$$P(S_n \leq 20) < \frac{(112-20)\cdot 0.25}{(28-20)^2} = 0.35938$$

Questo numero è la stima della coda sinistra della nostra binomiale: in sostanza, la probabilità che un certo Quartiere venga estratto, ad esempio, per primo in meno di 20 estrazioni è minore del 36%.

E se prendessimo un $r > np$ e volessimo vedere quant'è la probabilità che la variabile aleatoria sia maggiore o uguale di r?

In tal caso la formula che ci viene in aiuto è

$$P(S_n \geq r) < \frac{rq}{(np - r)^2}$$

Mettiamola all'opera, utilizzando $r = 92$.

$$P(S_n \geq 92) < \frac{92 \cdot 0.75}{(28 - 92)^2} = 0.01685$$

Valore molto basso, meno del 2%, ma che rispecchia la porzione di grafico della nostra binomiale.

Consideriamo adesso un intervallo di estrazioni: calcoliamo la probabilità che S_n sia compresa tra 20 e 92. Sapendo, dalle proprietà della funzione cumulativa, che

$$P(r \leq S_n \leq t) = P(S_n \leq t) - P\left(S_n \geq r\right)$$

con $r \leq t \leq n$,

$$P(20 \leq S_n \leq 92) = P(S_n \leq 92) - P\left(S_n \leq 20\right)$$

$$\begin{aligned} P(S_n \leq 92) &= \sum_{i=1}^{92} b_{n,p}\left(i\right) \\ &= 1 \end{aligned}$$

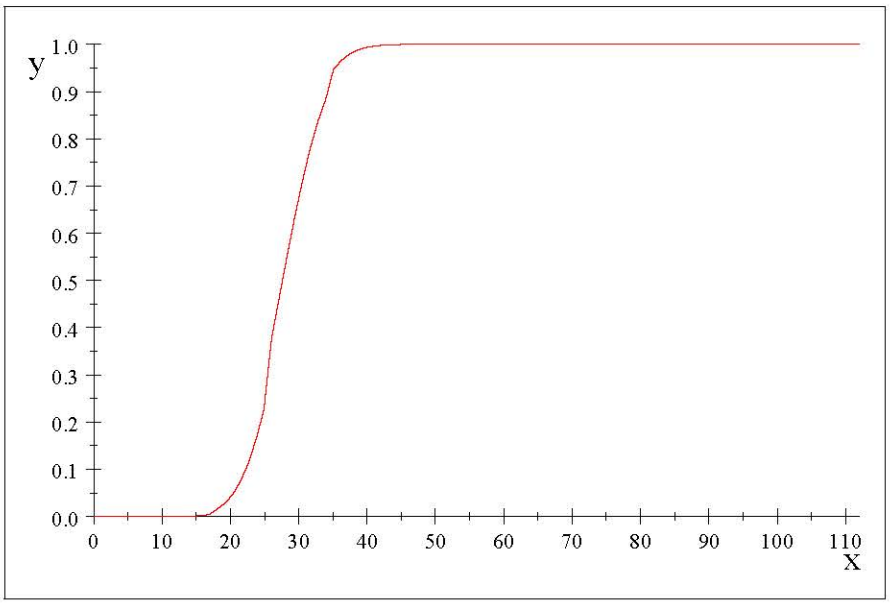
$$\begin{aligned} P\left(S_n \leq 20\right) &= \sum_{i=1}^{20} b_{n,p}\left(i\right) \\ &= 0.04715 \end{aligned}$$

La funzione cumulativa sarà:

$$\begin{aligned} P(20 \leq S_n \leq 92) &= 1 - 0.04715 \\ &= 0.95285 \end{aligned}$$

In sostanza, la probabilità che un Quartiere venga estratto secondo un preciso ordine è molto bassa fino all'incirca alla ventesima estrazione, per poi aumentare esponenzialmente fino a $x = 40$, attestandosi sulla certezza.

Il grafico della funzione cumulativa esplica quanto appunto scritto:



Ciò rafforza il fatto che $P(S_n \geq 92)$ sia minore del 2% e che la coda destra della binomiale si abbassi sempre più all'aumentare del numero di prove: è più probabile che tra la ventesima estrazione e la quarantesima ci sia il "successo"; già arrivati alla novantaduesima possiamo essere sicuri che l'evento si sia già avverato.

Restringiamo ancora di più l'intervallo: calcoliamo $P(20 \leq S_n \leq 40)$ e notiamo che è pari a

$$\begin{aligned} P(S_n \leq 40) - P(S_n \leq 20) &= 0.99578 - 0.04715 \\ &= 0.94863 \end{aligned}$$

un valore molto vicino a $P(20 \leq S_n \leq 92) = 0.95285$, nonostante le 52 prove di differenza: una riprova in più del fatto che oltre le 40 estrazioni abbiamo pressoché la certezza, la funzione cumulativa è costante.

Non ci resta che confrontare questi dati di aspettativa con quelli raccolti, utilizzando il test di verifica di ipotesi sui parametri della binomiale.

Innanzitutto, formuliamo le ipotesi: l'ipotesi nulla H_0 sarà che la probabilità di un Quartiere di uscire secondo un dato ordine sia sempre 0.25 e che quindi le frequenze osservate siano frutto del caso; l'ipotesi alternativa H_1, invece, smentisce quanto appunto detto.

$$\begin{aligned} H_0 &: p = 0.25 \\ H_1 &: p \neq 0.25 \end{aligned}$$

Come da convenzione, scegliamo un livello di significatività del 5%. Vediamo dunque fino a quale punto la funzione cumulativa assume un valore minore o uguale a 0.05:

$$P(S_n \leq 20) = 0.04715$$

Poiché già a 21: $P(S_n \leq 21) = 0.07491 > 0.05$.

Quindi la soglia critica, entro la quale viene rigettata H_0 si attesta nella coda sinistra della distribuzione binomiale è quindi 20 estrazioni.

Spesso e volentieri, nelle indagini statistiche, il livello di significatività del 5% la fa da padrone, ma non dobbiamo dimenticarci che fu creato da Fischer come la calibratura consigliata per quello strumento molto utile e pratico che è il p-value. In sostanza, 0.05 è un valore che il ricercatore fissa a priori, prima di fare la verifica di ipotesi, per poter vedere, una volta effettuata la verifica, se ciò che ha ipotizzato può essere considerabile come non dovuto al caso. Questo non significa che, ad esempio, una volta ottenuto 20 come valore soglia significativo per rigettare l'ipotesi "$p = 0.25$" si sia arrivati ad una conclusione, anzi: non è un punto di arrivo, ma di partenza per altre indagini più approfondite. Proviamo, quindi, a ripetere la verifica di ipotesi abbassando il livello di significatività.

Avendo le seguenti ipotesi:

$$\begin{aligned} H_0 &: p = 0.25 \\ H_1 &: p \neq 0.25 \end{aligned}$$

e fissando il livello di significatività al 4%, la funzione cumulativa assume un valore minore o uguale a 0.04 per $k = 19$. Infatti $P(S_n \leq 19) = 0.02814$.

Scendendo al 2%, rifiutiamo l'ipotesi nulla per $k = 18$: $P(S_n \leq 18) = 0.01588$; mentre, per $\alpha = 0.01$, troviamo 17 estrazioni come valore soglia: $P(S_n \leq 18) = 0.00844$.

Un'altra applicazione interessante della funzione cumulativa è anche questa: calcolare la probabilità di cadere nell'intervallo di occorrenze per ogni estrazione, centrato nella media.

Vediamo quanto detto con l'applicazione pratica: prendiamo in esame la prima estrazione, ovvero tutte e 112 le prime estrazioni. Notiamo che, divise per Quartiere, formano una stringa di quattro numeri: $24, 26, 26, 36$. Ovviamente, la loro somma è 112 e la loro media è $\frac{24+26+26+36}{4} = 28$. Il valore massimo della stringa è senza dubbio 36, quindi $36 - 28 = 8$; l'intervallo che ci interessa è perciò $(28 - 8, 28 + 8) = (20, 36)$. Calcoliamo la probabilità che S_n "cada" in $(20, 36)$:

$$\begin{aligned} P(20 \leq S_n \leq 36) &= P(S_n \leq 36) - P(S_n \leq 20) \\ &= \sum_{i=1}^{36} b_{n,p}(i) - \sum_{i=1}^{20} b_{n,p}(i) \\ &= 0.96542 - 0.04715 \\ &= 0.91827 \end{aligned}$$

Facciamo la stessa cosa con la seconda estrazione: la stringa ottenuta è $24, 38, 29, 21$, la cui media è, naturalmente 28. L'intervallo allora sarà

$$(28 - 10, 28 + 10) = (18, 38)$$

e quindi:

$$\begin{aligned} P(18 \leq S_n \leq 38) &= P(S_n \leq 38) - P(S_n \leq 18) \\ &= \sum_{i=1}^{38} b_{n,p}(i) - \sum_{i=1}^{18} b_{n,p}(i) \\ &= 0.98705 - 0.01588 \\ &= 0.97117 \end{aligned}$$

Continuiamo il calcolo con la terza estrazione, con una stringa più "uniforme": $33, 23, 30, 26$. L'intervallo centrato in 28 è $(28-5, 28+5) = (23, 33)$.e la probabilità di "caderci" è:

$$\begin{aligned} P(33 \leq S_n \leq 23) &= P(S_n \leq 33) - P(S_n \leq 23) \\ &= \sum_{i=1}^{33} b_{n,p}(i) - \sum_{i=1}^{23} b_{n,p}(i) \\ &= 0.88367 - 0.163\,12 \\ &= 0.72055 \end{aligned}$$

Terminiamo con la quarta estrazione: l'intervallo, ancora più ristretto degli altri tre, è $(28-3, 28+3) = (25, 31)$, dato che i quattro valori rilevati sono $31, 25, 27, 29$.

$$\begin{aligned} P(31 \leq S_n \leq 25) &= P(S_n \leq 31) - P(S_n \leq 25) \\ &= \sum_{i=1}^{31} b_{n,p}(i) - \sum_{i=1}^{25} b_{n,p}(i) \\ &= 0.77973 - 0.29725 \\ &= 0.48248 \end{aligned}$$

È piacevolmente interessante osservare che più un intervallo di occorrenze è ampio, maggiore sarà la probabilità di "caderci".

6.3 L'approssimazione normale

Sappiamo che la distribuzione normale si rivela essere una buona approssimazione della distribuzione binomiale per n grandi; ribadendo che 112 non è un numero considerevolmente grande, proviamo a verificare che l'approssimazione sia comunque soddisfacente.

Come introdotto nel capitolo dei Concetti Fondamentali, abbiamo visto, grazie al Teorema di DeMoivre-Laplace, che $X_n \sim B_{n,p}$. Abbiamo dunque che

$$b_{n,p}(k) \approx \phi_{np,npq}(k)$$

Conoscendo il valore di n, p, q, possiamo calcolare il valore atteso, pari a np e la varianza npq, questi ultimi due numeri possono quindi essere assimilati, rispettivamente, al valore atteso e alla varianza della distribuzione normale.

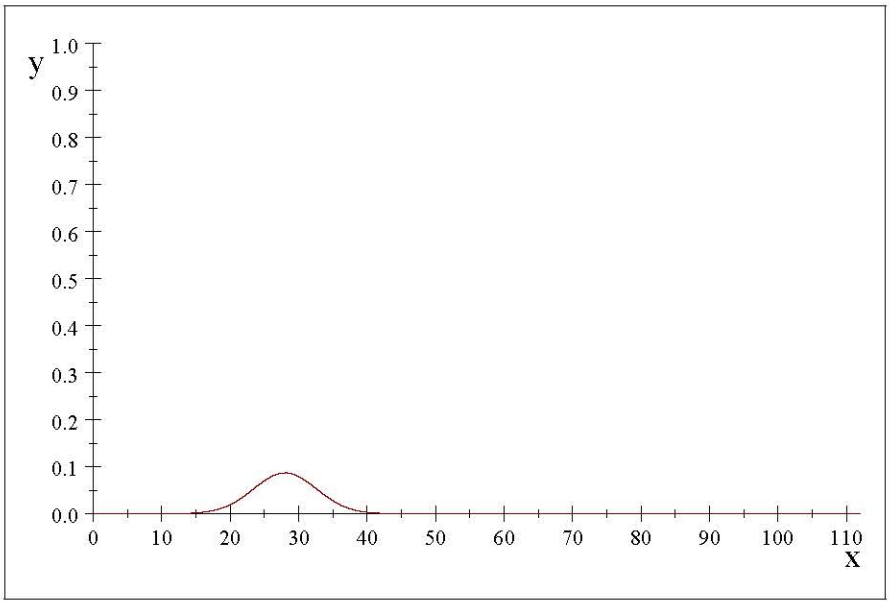
Operiamo dunque con i valori delle estrazioni delle carriere:

$$\begin{aligned} np &= 112 \cdot 0.25 = 28 \\ npq &= 112 \cdot 0.25 \cdot 0.75 = 21 \end{aligned}$$

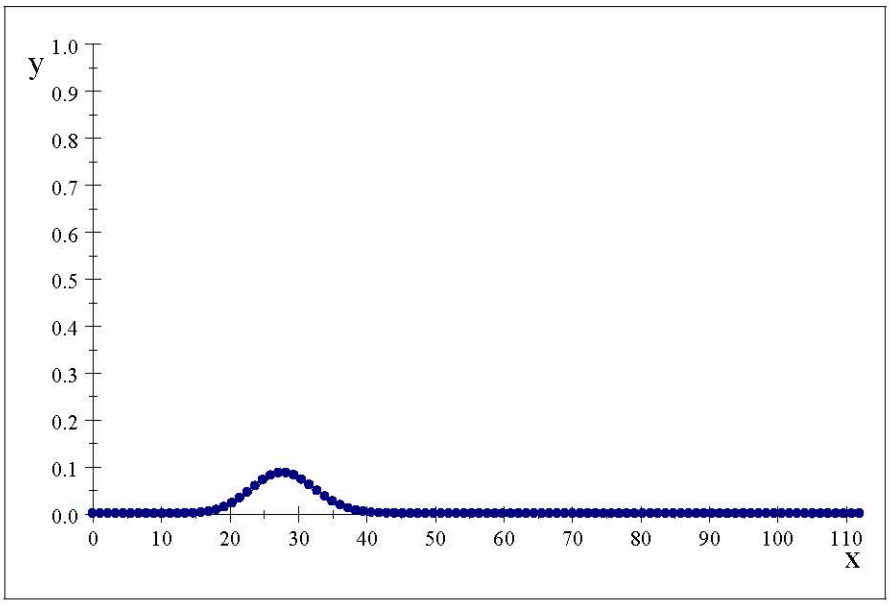
immettendoli nella formula della funzione di densità normale:

$$\phi_{np,npq}(k) = \frac{1}{\sqrt{21}\sqrt{2\pi}} e^{-\frac{(k-28)^2}{2 \cdot 21}}$$

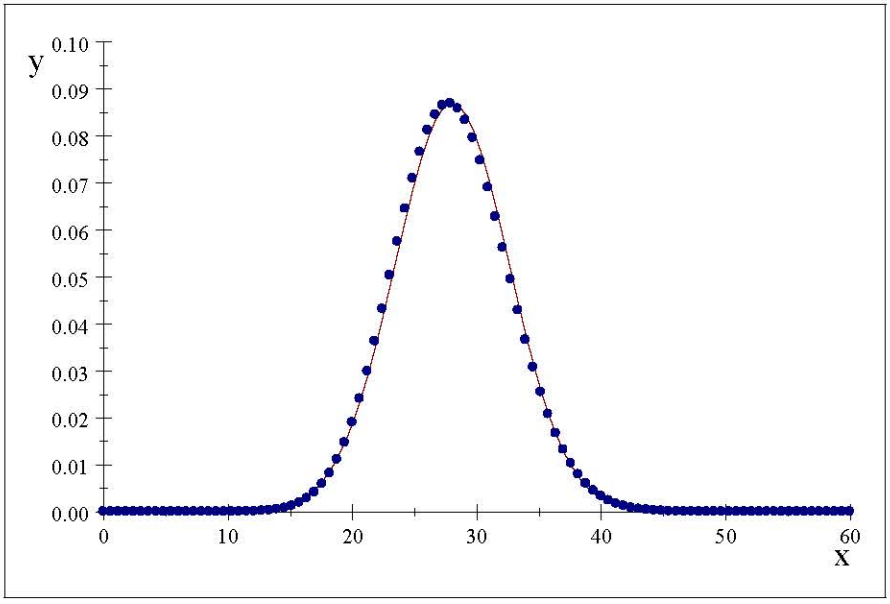
Della quale possiamo disegnare il grafico:



Notiamo che è sovrapponibile al grafico della distribuzione binomiale che abbiamo disegnato nel capitolo 6.1:



Restringiamo per notarlo meglio:



Possiamo infatti notare come il punti blu della binomiale seguano l'andamento della linea rossa, rappresentante la normale sul piano cartesiano.

La caratteristica forma ogivale in questo caso è piuttosto stretta: ciò significa che i valori sono raggruppati intorno alla media e quindi la varianza è bassa. In altre parole, ci suggerisce ciò che abbiamo detto nel capitolo precedente: intorno

alla ventottesima estrazione abbiamo il picco di probabilità che un Quartiere sia estratto alla posizione considerata, picco che si attesta intorno a 0.086776.

L'osservazione empirica va naturalmente d'accordo con questo valore: calcolando la media delle quattro occorrenze osservate per la prima estrazione $(24, 26, 26, 36)$, notiamo che questa è pari a 28; stessa cosa vale per la seconda estrazione, la terza e la quarta.

Il coefficiente di variazione $V = \frac{\sigma}{\mu}$, che sintetizza il rapporto tra la deviazione standard e la media, ci permette di quantificare la dispersione dei dati osservati, utilizzando il valore atteso come unità di misura. Vediamo quanto vale nel nostro caso:

$$V = \frac{\sqrt{npq}}{np} = \frac{\sqrt{21}}{28} = 0.16366$$

Essendo $0.16366 < 0.5$, possiamo dire che la media (28), benché non corrisponda ad alcun valore realmente osservato (come si può notare nella Tabella 1), è un indicatore corretto.

6.4 La distribuzione quadrinomiale

Un'interessante generalizzazione della distribuzione binomiale è la *distribuzione multinomiale*: ad ognuna delle n prove indipendenti corrisponde un insieme di possibili risultati, tali che la somma delle loro realizzazioni sia 1. In altre parole, chiamando $R_1, ..., R_s$ gli s risultati di ogni prova e p_i $(i = 1, ..., s)$ le corrispettive probabilità di successo, avremo che $p_1 + ... + p_s = 1$.

La probabilità che R_1 avvenga k_1 volte, R_2 k_2 volte, ecc... sarà quindi[14]:

$$\begin{aligned}\binom{n}{k_1 ... k_s} p^{k_1+...+k_s} &= \frac{n!}{k_1!(n-k_1)!}\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdots \\ &\quad \cdots \frac{(n-k_1-...-k_{s-2})!}{k_{s-1}!k_s!} p^{k_1} p^{k_2} ... p^{k_s} \\ &= \frac{n!}{k_1!k_2!...k_s!} p^n\end{aligned}$$

Ovviamente, il caso con $s = 2$ è quello della distribuzione binomiale, ma, insieme a quest'ultima, per il presente lavoro di tesi, ci farà comodo anche la distribuzione quadrinomiale:

$$\frac{n!}{k_1!k_2!k_3!k_4!} p^n$$

che possiamo utilizzare per analizzare ancora più da vicino l'estrazione delle carriere.

Nel caso speicifico, infatti, i possibili risultati di ognuno dei quattro eventi PC, PDF, PSA, PSS sono quattro, dato che ognuna delle quattro palline può essere estratta per prima, seconda, terza o quarta. La distribuzione quadrinomiale ci aiuta quindi a calcolare quante volte avviene PC, quante volte PDF,

[14] Ovviamente, $k_1 + ... + k_s = n$.

quante PSA e quante PSS alla prima estrazione, alla seconda, alla terza e alla quarta. Viene fuori che la nostra S_4 è un vettore formato dalle variabili aleatorie R_1, R_2, R_3, R_4 corrispondenti alle altrettante posizioni in cui possono cadere le palline e k_1, k_2, k_3, k_4 sono le rispettive quantità di "successi".

Calcoliamo quindi la probabilità cumulativa di uscire alla prima estrazione in 112 "prove":

$$\frac{112!}{24!26!26!36!}(0.25)^{112} = 0.0001951$$

E così per quanto riguarda la seconda estrazione, la terza e la quarta:

$$\begin{aligned}
\frac{112!}{24!38!29!21!}(0.25)^{112} &= 0.0000499 \\
\frac{112!}{33!23!30!26!}(0.25)^{112} &= 0.0003049 \\
\frac{112!}{31!25!27!29!}(0.25)^{112} &= 0.0005964
\end{aligned}$$

Se volessimo esprimere le discrepanze tra queste quadrinomiali mediante un grafico, il piano cartesiano non ci basterebbe: avremmo bisogno di tre dimensioni. Un simplesso originato da tre assi, ciascuno dei quali va da 0 a 112: un tetraedro, i cui vertici corrispondono alle quattro posizioni nell'estrazione.

Tabella delle Estrazioni

Data Giostra	*I carriera*	*II carriera*	*III carriera*	*IV carriera*
domenica 12 settembre 1948	**PSS**	PDF	PC	PSA
domenica 7 agosto 1949	**PC**	PDF	PSS	PSA
domenica 4 settembre 1949	PSA	PSS	**PC**	PDF
domenica 4 giugno 1950	PC	PDF	PSA	**PSS**
sabato 2 settembre 1950	PC	PSS	PSA	**PDF**
domenica 3 giugno 1951	PC	PSA	PDF	**PSS**
domenica 2 settembre 1951	PDF	PSA	**PC**	PSS
domenica 8 giugno 1952	PDF	PSA	**PC**	PSS
domenica 7 settembre 1952	PSS	**PC**	PDF	PSA
domenica 6 settembre 1953	PDF	PSS	PSA	**PC**
domenica 5 settembre 1954	**PC**	PSS	PSA	PDF
domenica 4 settembre 1955	PSS	**PDF**	PC	PSA
domenica 2 settembre 1956	PDF	**PC**	PSA	PSS
domenica 1 settembre 1957	**PSA**	PDF	PC	PSS
domenica 7 settembre 1958	**PC**	PSA	PSS	PDF
domenica 13 settembre 1959	PDF	**PSS**	PSA	PC
domenica 28 agosto 1960	PSA	**PDF**	PC	PSS
domenica 4 settembre 1960	PC	PSA	PDF	**PSS**
domenica 3 settembre 1961	PC	PSA	**PDF**	PSS
domenica 2 settembre 1962	PC	PDF	**PSS**	PSA
domenica 1 settembre 1963	PC	**PSA**	PSS	PDF
domenica 13 settembre 1964	**PDF**	PC	PSA	PSS
domenica 5 settembre 1965	PC	PDF	**PSS**	PSA
domenica 4 settembre 1966	PSS	PDF	PSA	**PC**
domenica 3 settembre 1967	PDF	PSA	**PSS**	PC
domenica 1 settembre 1968	PC	PSS	**PSA**	PDF
domenica 7 settembre 1969	**PSS**	PC	PDF	PSA
domenica 6 settembre 1970	PSS	**PSA**	PDF	PC
domenica 5 settembre 1971	**PSS**	PDF	PSA	PC
domenica 3 settembre 1972	**PSS**	PSA	PC	PDF
domenica 2 settembre 1973	PDF	PSS	**PC**	PSA
domenica 1 settembre 1974	PSS	**PDF**	PC	PSA
domenica 7 settembre 1975	PSS	**PSA**	PDF	PC
sabato 28 agosto 1976	PDF	**PC**	PSA	PSS
domenica 5 settembre 1976	PSS	PC	**PSA**	PDF
domenica 4 settembre 1977	PSS	PSA	PC	**PDF**
venerdì 23 giugno 1978	PSS	**PSA**	PC	PDF
domenica 3 settembre 1978	**PSS**	PSA	PC	PDF
domenica 2 / sabato 15 settembre 1979	PDF	PC	PSA	**PSS**
sabato 30 agosto 1980	PSA	PSS	**PC**	PDF
domenica 7 settembre 1980	PSA	**PC**	PDF	PSS
domenica 6 settembre 1981	**PSS**	PDF	PC	PSA
domenica 5 settembre 1982	PSS	PSA	**PDF**	PC
sabato 11 settembre 1982	PSS	PC	**PDF**	PSA
domenica 4 settembre 1983	PSS	PDF	**PSA**	PC
sabato 10 settembre 1983	PSA	**PDF**	PSS	PC
sabato 7 luglio 1984	PSS	**PSA**	PDF	PC
domenica 2 settembre 1984	**PSS**	PC	PSA	PDF
sabato 29 settembre 1984	**PSA**	PSS	PC	PDF
sabato 29 giugno 1985	**PDF**	PC	PSA	PSS
domenica 1 settembre 1985	PSS	**PDF**	PC	PSA

domenica 31 agosto 1986	PSA	**PC**	PSS	PDF
domenica 7 settembre 1986	**PC**	PSS	PSA	PDF
domenica 30 agosto 1987	PSS	PDF	PC	**PSA**
domenica 6 settembre 1987	**PSA**	PDF	PC	PSS
domenica 28 agosto 1988	PSS	PSA	**PC**	PDF
domenica 4 settembre 1988	**PDF**	PC	PSA	PSS
domenica 27 agosto 1989	PSA	**PC**	PDF	PSS
domenica 3 settembre 1989	**PSA**	PC	PSS	PDF
domenica 26 agosto 1990	**PDF**	PC	PSA	PSS
domenica 2 settembre 1990	PSS	**PSA**	PDF	PC
domenica 25 agosto / domenica 8 settembre 1991	PDF	**PC**	PSA	PSS
domenica 1 settembre 1991	PC	PSS	PDF	**PSA**
domenica 30 agosto 1992	**PC**	PDF	PSS	PSA
domenica 6 settembre 1992	PC	PDF	**PSA**	PSS
domenica 29 agosto 1993	PSA	PDF	PSS	**PC**
domenica 5 settembre 1993	PDF	PC	PSS	**PSA**
domenica 28 agosto 1994	PSS	PDF	PC	**PSA**
domenica 4 settembre 1994	PDF	**PSA**	PC	PSS
domenica 25 giugno 1995	PSA	**PDF**	PSS	PC
domenica 3 settembre 1995	PSS	PDF	PSA	**PC**
domenica 16 giugno 1996	**PDF**	PSS	PC	PSA
domenica 1 settembre 1996	PSA	**PSS**	PC	PDF
domenica 22 giugno 1997	**PDF**	PSS	PSA	PC
domenica 7 / sabato 13 settembre 1997	PSS	PDF	PC	**PSA**
domenica 21 giugno 1998	PSS	PDF	PC	**PSA**
domenica 6 settembre 1998	PSA	**PDF**	PSS	PC
domenica 20 giugno 1999	PC	PDF	**PSS**	PSA
domenica 5 settembre 1999	PSS	PDF	**PSA**	PC
domenica 18 giugno 2000	PSA	PSS	**PDF**	PC
domenica 3 settembre 2000	PDF	PSA	PSS	**PC**
sabato 9 settembre 2000	PDF	PSA	PSS	**PC**
domenica 17 giugno 2001	PDF	**PC**	PSS	PSA
domenica 2 settembre 2001	PC	**PDF**	PSS	PSA
sabato 22 giugno 2002	PSS	PSA	PDF	**PC**
domenica 1 settembre 2002	PC	PSA	PDF	**PSS**
sabato 21 giugno 2003	**PSA**	PDF	PSS	PC
domenica 7 settembre 2003	PSA	PSS	PC	**PDF**
sabato 19 giugno 2004	PSS	**PSA**	PC	PDF
domenica 5 settembre 2004	PC	PSA	PDF	**PSS**
sabato 18 giugno 2005	PC	PSA	**PDF**	PSS
domenica 4 settembre 2005	**PC**	PDF	PSS	PSA
sabato 17 giugno 2006	PSS	PDF	**PSA**	PC
sabato 2 settembre 2006	PDF	PC	**PSS**	PSA
sabato 23 giugno 2007	PSS	PC	PDF	**PSA**
domenica 2 settembre 2007	PSS	PC	PSA	**PDF**
sabato 21 giugno 2008	PDF	PSS	**PC**	PSA
domenica 7 settembre 2008	PC	PDF	**PSA**	PSS
sabato 20 giugno 2009	PSA	PDF	**PC**	PSS
domenica 6 settembre 2009	PDF	**PSA**	PSS	PC
sabato 19 giugno 2010	**PSA**	PC	PSS	PDF
domenica 5 settembre 2010	PSA	PSS	**PC**	PDF
sabato 18 giugno 2011	PSA	PC	PDF	**PSS**
domenica 4 settembre 2011	**PC**	PSS	PSA	PDF
sabato 23 giugno 2012	PSA	PDF	**PSS**	PC
domenica 2 settembre 2012	**PSS**	PDF	PSA	PC
sabato 22 giugno 2013	PSA	PDF	PC	**PSS**
domenica 1 settembre 2013	PDF	**PSA**	PSS	PC
sabato 21 giugno 2014	PDF	PSS	**PSA**	PC
domenica 7 settembre 2014	PSA	PDF	PC	**PSS**
sabato 20 giugno 2015	PSS	PSA	PDF	**PC**
domenica 6 settembre 2015	**PSA**	PSS	PDF	PC

Bibliografia

[1] W. FELLER, *An Introduction to Probability Theory and Its Applications*, vol. I, 3rd edition, John Wiley & Sons, 1950.

[2] F. N. DAVID, *Dicing and Gaming (A Note on the History of Probability)*, Biometrika, 1955.

[3] G. LIBRI, *Histoire des Sciences Mathématiques en Italie, depuis la Renaissance des Lettres jusqu'à la Fin du Dix-Septieme Siécle.*

[4] C. DISSENNATI, *Le mille lance del Saracino*, Tipografia D. Badiali, 1966.

[5] C. FARDELLI, *1966-2004: Giostra del Saracino*, Arti grafiche Cianferoni, 2004.

[6] R. PARNETTI, *E vidi correr giostra: Arezzo e la Giostra del Saracino*, Gruppo Genesi Editoriale, 2006.

[7] Libro del Cancelliere della Giostra del Saracino (1992-2015)

[8] M. SCALINI, *Il Saracino e gli spettacoli cavallereschi della Toscana*, Firenze, S.P.E.S., 1987