

Introduzione alla probabilità

in 4 lezioni

Fabio Fagnani

Dipartimento di Matematica

Politecnico di Torino

fabio.fagnani@polito.it

1 La nascita della probabilità

La maggior parte delle teorie scientifiche si sono sviluppate a piccoli passi partendo da nozioni molto semplici, la genesi delle quali difficilmente si può determinare con precisione. Ad esempio l'aritmetica, che forma adesso una delle più complesse e profonde teorie matematiche, molto probabilmente cominciò con il bisogno di qualche uomo primitivo di contare il suo gregge; forse quest'uomo primitivo utilizzò le dita della mano e questo ha condotto all'utilizzo della base 10 nel nostro sistema numerico. Simili considerazioni potrebbero essere fatte per la geometria, per l'astronomia come per altre discipline scientifiche.

Anche primordiali ragionamenti probabilistici sono sicuramente apparsi già nei tempi antichi forse collegati ai primi giochi d'azzardo. Si pensi che il dado ha fatto la sua apparizione in Egitto intorno al 2000 a.c e prima di lui già nel 3600 a.c. gli Assiri e i Sumeri sembra si cimentassero nel lancio dell'astragalo (l'osso sopra il ginocchio) di pecore, di cervi o di animali di simile stazza. Per la sua forma, un astragalo può ricadere in quattro modi diversi, anche se, a differenza del dado, questi quattro modi non risultano 'egualmente possibili' ed ogni astragalo ha le sue proprie caratteristiche di aleatorietà. Si può pensare che il lancio dell'astragalo sia uno dei primi giochi d'azzardo sperimentato dall'uomo.

A differenza delle altre discipline tuttavia la teoria matematica delle probabilità ha una data di origine chiaramente riconoscibile: il 1654 in Francia. Naturalmente come ogni cosa profonda, essa non nasce dal nulla ed è chiaramente riconoscibile un periodo di gestazione: un contributo particolarmente importante, dove sono riconoscibili molti elementi che si ritrovano poi nella teoria della probabilità, è il libro sui giochi *Liber de Ludo Aleae* scritto da Cardano nel 1520, ma pubblicato soltanto molto tempo dopo nel 1663.

Che cosa accadde nel 1654? A quell'epoca in Francia il gioco d'azzardo era di gran moda e le bische clandestine diffusissime. Un gioco estremamente alla moda allora era il seguente: la 'casa' scommetteva alla pari con un giocatore che quest'ultimo, lanciando per 4 volte un dado, avrebbe ottenuto almeno una volta il numero 6. Come vedremo successivamente questo gioco è leggermente favorevole alla casa che 'in media' vince

il 52% delle volte. Un distinto ed intelligente francese dell'epoca, Antoine Gombauld Cavalier de Méré, frequentatore delle case da gioco, conosceva bene quel gioco ed era incuriosito da una sua possibile variante: la casa scommetteva alla pari con un giocatore che quest'ultimo, lanciando per 24 volte una coppia di dadi, avrebbe ottenuto almeno una volta il doppio 6. Anche questo gioco, secondo il Cavalier de Méré, avrebbe dovuto essere leggermente favorevole alla casa per questo motivo: quando si lancia un dado vi sono 6 possibili risultati, quindi la probabilità che esca il 6 sarà $1/6$. Invece la probabilità che esca un doppio 6 sarà di $1/36$ (essendo 36 i risultati distinti del lancio di due dadi), quindi 6 volte più bassa; lanciando la coppia di dadi 6 volte di più ($24 = 6 \times 4$) si dovrebbe controbilanciare l'effetto di considerare un evento meno probabile di un fattore 6 e si dovrebbe avere quindi la stessa probabilità. Invece non è così, quest'ultimo gioco non è favorevole alla casa, ma al giocatore; ne era consapevole il Cavalier de Méré, non è chiaro se per averlo provato a sue spese o per qualche intuizione teorica. In ogni caso decise di parlarne con un altro brillante francese, Blaise Pascal, che si diletta di lettere, di filosofia, di teologia, ma anche di matematica. Pascal risolse il problema postogli dal de Méré provando anche che con 25 lanci il gioco sarebbe allora stato favorevole alla casa. Da quel momento Pascal cominciò ad occuparsi di altri problemi probabilistici coinvolgendo un grandissimo matematico dell'epoca Pierre de Fermat. Insieme svilupparono i primi elementi della teoria delle probabilità. Presto altri grandi pensatori dell'epoca se ne cominciarono ad interessare, Huygens, Leibnitz, Bernoulli e per la fine del XVII secolo molte idee e concetti probabilistici avevano ormai preso corpo.

Sono passati più di 300 anni da allora e ne ha fatta molta di strada la probabilità dalle bische clandestine dove era nata. E' ora una vera e propria disciplina matematica al pari dell'aritmetica, della geometria, dell'algebra o dell'analisi. Ha il suo corpo autonomo di concetti, di idee e ci sono matematici che si occupano solo di essa, probabilisti appunto, o meglio di piccole parti di essa; è in effetti così sviluppata che è ormai impossibile che una singola mente possa abbracciarla tutta. A cosa serve? Beh, sicuramente i bravi giocatori di poker ancora oggi la utilizzano e serve per capire a fondo la struttura dei vari giochi di azzardo; ma ha sicuramente applicazioni ben più importanti. La fisica moderna si fonda per molti aspetti su di essa: la descrizione di sistemi complessi come un gas, un fluido, è affidata a modelli probabilistici, mentre nella meccanica quantistica, la probabilità è il linguaggio stesso in cui si scrivono le leggi della natura. Ancora, la probabilità ha applicazioni nella biologia, nell'economia, nelle scienze finanziarie. Infine essa ha importantissime applicazioni nelle telecomunicazioni e sta alla base dei risultati scientifici che hanno permesso la rivoluzione tecnologica digitale degli ultimi anni.

La probabilità è uno strumento formidabile per indagare il mondo che ci circonda. Essa è tuttavia entrata ancora molto poco nella cultura generale e questo ha conseguenze abbastanza serie. In effetti, come vedremo, in molti problemi specifici il ragionamento probabilistico è spesso non-intuitivo; di conseguenza la non-conoscenza della probabilità non è meramente un vuoto culturale; essa porta ad avere idee sbagliate su questioni che fanno parte del mondo che ci circonda.

2 I modelli probabilistici

Il ragionamento probabilistico interviene ogni volta che effettuiamo, assistiamo ad un esperimento l'esito del quale non è completamente determinato a priori e può avere un certo numero di diversi risultati. Ad esempio, se lanciamo una moneta essa può atterrare lasciando in vista le due faccie (le vecchie testa o croce) e non siamo in grado di predeterminare in anticipo la faccia che rimarrà in vista. Similmente nel lancio del dado a 6 faccie, vi sono 6 possibili risultati distinti. Tutti i giochi d'azzardo si fondano proprio sul principio di avere più risultati possibili non prevedibili.

Naturalmente, vi sono molti altri contesti che si possono adattare ad una modellizzazione probabilistica. Ad esempio l'esperimento può consistere nel contare il numero di telefonate che arrivano ad un determinato apparecchio telefonico nell'arco di 24 ore; nel contare il numero di giornate di pioggia a Piombino nei primi 10 giorni di Novembre; nell'osservare il colore della biglia estratta casualmente da un'urna dove vi sono una biglia nera, due bianche, tre rosse e quattro verdi. Un altro esperimento ancora può consistere nell'osservare il tempo di durata di una lampadina o di qualche altro apparecchio elettrico.

In cosa consiste un modello probabilistico? Innanzitutto dobbiamo fissare un insieme che contenga come elementi i possibili esiti dell'esperimento sotto considerazione. Questo insieme verrà generalmente indicato con il simbolo Ω e chiamato *spazio degli eventi elementari*. Vediamo chi è Ω negli esempi precedenti:

- 1) Lancio della moneta: $\Omega = \{T, C\}$.
- 2) Lancio del dado: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- 3) Numero telefonate in un giorno: $\Omega = \mathbb{N} = \{0, 1, 2, \dots\}$.
- 4) Numero giornate di pioggia: $\Omega = \{0, 1, 2, \dots, 10\}$.
- 5) Colore biglia: $\Omega = \{N, B, R, V\}$.
- 6) Durata della lampadina. $\Omega = \mathbb{R}^+ =$ 'numeri reali non negativi'.

In questo corso ci occuperemo, quasi esclusivamente di situazioni nelle quali lo spazio degli eventi elementari è finito come nei casi 1), 2), 4) e 5) precedenti. Supponiamo dunque:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

Gli elementi ω_i vengono detti *eventi elementari*. Per completare la descrizione del modello probabilistico dobbiamo ora specificare la probabilità di accadimento dei singoli eventi elementari. Dobbiamo cioè dare n numeri $\Pr(\omega_1), \Pr(\omega_2), \dots, \Pr(\omega_n)$:

$$\Pr(\omega_i) = \text{probabilità che accada l'evento } \omega_i \tag{1}$$

Convenzionalmente si richiede che la somma delle probabilità degli eventi elementari faccia 1, cioè si richiede che

$$\Pr(\omega_1) + \Pr(\omega_2) + \dots + \Pr(\omega_n) = 1 \tag{2}$$

Inoltre le probabilità devono essere numeri non negativi:

$$\Pr(\omega_i) \geq 0 \quad i = 1, \dots, n \quad (3)$$

Ω con l'insieme delle probabilità $\Pr(\omega_i)$ viene detto uno *spazio di probabilità*. Rimane però da capire come si fa a decidere quanto devono valere le singole probabilità $\Pr(\omega_i)$. Questa è un punto essenziale che storicamente a dato origine a non poche controversie. Non c'è in realtà un modo unico per fissare queste probabilità ed è chiaro che soltanto la conoscenza del contesto specifico di applicazione del modello può far decidere per una scelta anziché per un'altra. Ci sono comunque due modi principali per operare questa scelta. La prima è la cosiddetta *ipotesi classica* e consiste nell'ipotizzare che tutti gli eventi elementari siano equiprobabili, cioè abbiano la stessa probabilità di accadere. Si impone quindi:

$$\Pr(\omega_1) = \Pr(\omega_2) = \dots = \Pr(\omega_n) \quad (4)$$

Usando la proprietà (2) si ottiene quindi

$$\Pr(\omega_i) = \frac{1}{n} \quad \text{per ogni } i = 1, 2, \dots, n \quad (5)$$

Questa strada è sicuramente da seguire ogni qual volta non si ha nessun tipo di informazione per poter ritenere un evento più probabile degli altri e si adatta benissimo agli esempi 1), e 2). Nel lancio della moneta si ha quindi:

$$\Omega = \{T, C\} \quad \Pr(T) = \Pr(C) = 1/2$$

mentre nel lancio del dado

$$\Omega = \{1, 2, 3, 4, 5, 6\} \quad \Pr(i) = 1/6, \quad i = 1, 2, \dots, 6$$

L'ipotesi classica non si addice tuttavia agli altri esempi visti prima. Nell'estrazione della biglia dell'esempio 4) dove

$$\Omega = \{N, B, R, V\}$$

ci aspettiamo in effetti che, per le proporzioni tra le biglie di diversi colori, si abbia

$$\Pr(B) = 2\Pr(N), \quad \Pr(R) = 3\Pr(N), \quad \Pr(V) = 4\Pr(N)$$

Poiché la somma delle probabilità deve fare 1 si ha che

$$1 = \Pr(N) + 2\Pr(N) + 3\Pr(N) + 4\Pr(N) = 10\Pr(N)$$

Quindi si ha,

$$\Pr(N) = 1/10, \quad \Pr(B) = 2/10, \quad \Pr(R) = 3/10, \quad \Pr(V) = 4/10$$

Si può mostrare come il conto appena fatto possa essere comunque ricondotto all'ipotesi classica (si pensi al perchè). Di ben diversa natura è invece l'esempio 5): non ci aspettiamo che nessuna giornata di pioggia o 2 giornate di pioggia nei primi 10 giorni di Novembre a Piombino siano due eventi da poter considerare equiprobabili. In realtà in

questo caso l'unica cosa sensata da fare è guardare le statistiche degli anni precedenti e assegnare la probabilità in base alla frequenza di accadimento. Ad esempio, supponiamo di sapere che negli ultimi 100 anni, n_i volte si sono avute i giornate di pioggia a Piombino nei primi 10 giorni di Novembre: Allora si pone

$$\Pr(i) = \frac{n_i}{100} = \text{percentuale di anni con } i \text{ giorni di pioggia nel periodo specificato}$$

Si noti che poichè $n_1 + n_2 + \dots + n_{10} = 100$ si ha che

$$\Pr(1) + \Pr(2) + \dots + \Pr(10) = 1$$

Questo modo di procedere si definisce *ipotesi frequentista* ed è utile ogni qualvolta si abbiano dati precedenti di un certo esperimento.

Assegnare invece le probabilità agli eventi elementari degli esempi 3) e 6) è una questione decisamente più complicata e per il momento la lasciamo da parte.

3 Il calcolo delle probabilità

Che cosa resta da fare una volta che abbiamo fissato uno spazio di probabilità, cioè lo spazio degli eventi elementari Ω e le probabilità dei singoli eventi elementari? Ci possiamo a questo punto porre il problema di calcolare la probabilità di eventi non elementari. Ogni sottoinsieme A di Ω è detto *evento complesso*. Diremo che si verifica l'evento A se effettuando l'esperimento accade un qualsiasi evento elementare che sta dentro A . E' logico quindi porre la probabilità di A eguale alla somma delle probabilità degli eventi elementari che compongono A . Chiariamo il concetto con un paio di esempi: consideriamo prima l'estrazione della biglia dall'urna. Si ha

$$\Omega = \{N, B, R, V\}, \quad \Pr(N) = 1/10, \quad \Pr(B) = 2/10, \quad \Pr(R) = 3/10, \quad \Pr(V) = 4/10$$

Consideriamo l'evento complesso $A = \{N, R, V\}$. A accade cioè se viene estratta una biglia di colore nero o rosso o verde, o equivalentemente, se non viene estratta una biglia bianca. Si ha che

$$\Pr(A) = \Pr(N) + \Pr(R) + \Pr(V) = \frac{1}{10} + \frac{3}{10} + \frac{4}{10} = \frac{8}{10} = \frac{4}{5}$$

Nel caso particolare in cui si stia considerando uno spazio di probabilità con l'ipotesi classica il calcolo delle probabilità degli eventi complessi assume una forma particolarmente semplice. In effetti sia $\Omega = \{\omega_1, \dots, \omega_n\}$ e sia $\Pr(\omega_i) = 1/n$ per ogni i . Sia $A \subseteq \Omega$ e supponiamo che A consista di k eventi elementari. Poichè tutti gli eventi elementari hanno la stessa probabilità si ha che

$$\Pr(A) = \underbrace{\frac{1}{n} + \dots + \frac{1}{n}}_k = \frac{k}{n} = \frac{\text{casi favorevoli}}{\text{casi possibili}}$$

Consideriamo come esempio il lancio del dado dove, ricordiamo, $\Omega = \{1, 2, 3, 4, 5, 6\}$. Consideriamo l'evento complesso $A = \{2, 4, 6\}$ Quindi A accade se esce un numero pari. Si ha che

$$\Pr(A) = \frac{\text{casi favorevoli}}{\text{casi possibili}} = \frac{3}{6} = \frac{1}{2}$$

La probabilità degli eventi complessi gode di una serie di proprietà piuttosto intuitive, ma molto importanti; siano A, B sottoinsiemi di Ω . Allora abbiamo

$$\Pr(\Omega) = 1, \quad \Pr(\emptyset) = 0$$

$$A \subseteq B \Rightarrow \Pr(A) \leq \Pr(B)$$

$$A \cap B = \emptyset \Rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$$

(6)

$$\Pr(A^c) = 1 - \Pr(A)$$

dove abbiamo indicato con A^c il complementare di A dentro Ω , cioè A^c contiene tutti e soli gli elementi di Ω che non stanno in A . Commentiamo brevemente queste proprietà senza fornire dimostrazioni tutte peraltro molto semplici. La prima dice che la probabilità che accada l'evento Ω è 1, cioè che l'evento Ω accadrà senz'altro: fatto ovvio in quanto Ω contiene tutti gli eventi elementari e uno di questi dovrà pur capitare. Invece la probabilità che accada l'evento \emptyset (l'insieme vuoto) è 0, cioè è impossibile che l'evento \emptyset accada: per lo stesso motivo di prima, un evento elementare deve comunque accadere. La seconda dice che la probabilità ha comportamento monotono: più l'evento è grande, più cresce la probabilità. La terza esprime la proprietà cruciale di additività della probabilità: se si suppone che A e B siano disgiunti e quindi che non possano verificarsi simultaneamente, allora la probabilità dell'evento $A \cup B$ si ottiene sommando le probabilità di A e B . Infine la quarta asserisce che la probabilità di un evento A e dell'evento complementare A^c sono tra loro complementari rispetto ad 1, cioè che la loro somma fa 1. Si noti che si verifica A^c se e soltanto se non si verifica A , A^c è cioè la negazione dell'evento A ; la quarta così deriva direttamente dalla terza proprietà.

Il primo obiettivo del calcolo delle probabilità è il calcolo di probabilità di eventi complessi. Guardando gli esempi precedenti si potrebbe pensare che calcolare la probabilità di eventi complessi sia una cosa sempre molto semplice in quanto consiste semplicemente nell'effettuare una somma. Tuttavia, se l'evento complesso A è costituito da molti elementi ed è definito in modo implicito attraverso qualche proprietà caratterizzante, il calcolo della sua probabilità, seppure chiaro da un punto di vista teorico, potrebbe non risultare così semplice da un punto di vista operativo. L'analisi combinatoria sarà allora uno strumento indispensabile per effettuare questi calcoli.

4 Una raccolta di problemi

Presentiamo ora una serie di problemi di tipo probabilistico che faranno da guida al resto del corso. Come vedremo, alcuni di questi problemi necessitano, per una loro esatta

formulazione matematica, l'introduzione di altri concetti probabilistici e di ipotesi specifiche. La loro risoluzione passa attraverso una esatta formalizzazione del modello probabilistico (fatto che spesso rappresenta la difficoltà maggiore che si incontra in problemi di questo tipo), e un successivo calcolo di probabilità con tecniche combinatoriche.

P1) (Problema di De Méré) Si calcoli la probabilità che lanciando per 24 volte una coppia di dadi, si ottenga almeno una volta il doppio 6.

* * * * *

P2) Si calcoli la probabilità che in un gruppo di 23 persone ve ne siano almeno due che hanno il compleanno nello stesso giorno.

* * * * *

P3) Supponendo eventi equiprobabili la nascita di un maschio e di una femmina, si calcoli la probabilità che in una famiglia con 5 figli almeno 2 siano femmine.

* * * * *

P4) Si determinino le probabilità di avere un colore servito e un poker servito in una mano di poker. Si confrontino le due probabilità.

* * * * *

P5) 5 uomini e 5 donne si siedono casualmente ad un tavolo rotondo. Quale è la probabilità che tra due uomini ci sia sempre una donna?

* * * * *

P6) Un gruppo di 20 coppie mariti e mogli vanno ad una festa danzante. Uomini e donne vengono abbinati in modo casuale per il ballo. Qual'è la probabilità che nessun uomo danzi con sua moglie?

Una variante di questo difficile problema è stato per la prima volta formulato nel 1708 dal matematico Montmort.

* * * * *

P7) In una casa abita una famiglia con due figli dei quali ignoriamo il sesso. Alla porta viene ad aprire uno dei due figli che è un maschio. Supponendo eventi equiprobabili la nascita di un maschio e di una femmina, calcolare la probabilità che l'altro sia femmina.

* * * * *

P8) Vi sono 4 urne identiche ciascuna contenente due palline. In un'urna le due palline sono entrambe bianche, in una entrambe nere, e nelle restanti due urne ve ne sono, in ciascuna, una nera e una bianca. Scegliamo a caso una delle urne ed estraiamo da essa una pallina che risulta nera. Quale è la probabilità che l'altra pallina nell'urna scelta sia bianca?

Qualche analogia con il problema precedente?

* * * * *

P9) In una scuola di un paesino ci sono due classi dell'ultimo anno. Nella classe A ci sono 10 maschi e 11 femmine, nella B invece 8 maschi e 17 femmine. Nell'esame di fine anno ottengono il massimo punteggio 2 maschi e 2 femmine nella A e, 3 maschi e 6 femmine nella B. Sono più bravi i maschi o le femmine?

Un solerte giornalista del posto da la seguente risposta: nella classe A la percentuale dei maschi che hanno ottenuto il massimo è $2/10$ mentre per le femmine è $2/11$. Nella B invece le percentuali sono $3/8$ per i maschi e $6/17$ per le femmine. Nota poi il giornalista che

$$2/10 > 2/11 \quad 3/8 > 6/17$$

dal quale conclude che senza ombra di dubbio i maschi sono complessivamente migliori. Sareste in grado di smontare questa interpretazione?

* * * * *

P10) In una società un po' fuori moda ogni famiglia vuole avere uno ed un solo figlio maschio ed adotta così la strategia di fare figli fermandosi quando ne nasce uno maschio. Supponendo equiprobabile la nascita di un maschio o di una femmina, si dica se, in questa società, il numero dei maschi sarà in media superiore od inferiore al numero delle femmine.

Questo problema era già noto a Laplace che ne presenta una soluzione elementare sebbene non del tutto rigorosa. Una soluzione formale richiede un po' di matematica meno elementare.

* * * * *

P11) Se sei ingiustamente accusato di un crimine che non hai commesso, ti conviene chiedere un test alla macchina della verità? Specifichiamo meglio il problema. La macchina della verità, più precisamente il poligrafo, un apparecchio che registra variabili fisiologiche come battito del cuore, respirazione, pressione sanguigna che sono sotto il controllo del sistema vegetativo di un individuo. I sostenitori dell'uso della macchina della verità sostengono che questi parametri fisiologici, se misurati mentre una persona sta mentendo, mostrano un tracciato caratteristico che un esaminatore esperto può rivelare. Un esperimento del 1986 ha mostrato come la sensibilità media del poligrafo sia del 76%, cioè su 100 mentitori, il poligrafo ne scopre 76; la sua specificità è invece del 63%, cioè su 100 persone che dicono effettivamente la verità, per il poligrafo 63 di queste dicono la verità mentre le altre 37 mentono. In una popolazione di 1000 persone delle quali esattamente 5 sono mentitori, si sceglie una persona a caso e la si sottopone ad un test con una macchina della verità che ha le caratteristiche di sensibilità e specificità sopra esposte. Se viene positivo al test, qual'è la probabilità che sia effettivamente un mentitore?

* * * * *

P12) Uno dei modi per identificare un portatore di AIDS è tramite un esame del sangue che individua certi anticorpi che vengono prodotti in reazione al virus dell'AIDS. Tale test, per quanto buono, è soggetto a due tipi di errore: A) può risultare negativo, cioè fallire l'identificazione del virus, in un portatore; B) può risultare positivo, falsamente identificando il virus, in una persona invece sana. Supponiamo che il test dia, in media, risultato giustamente positivo sul 95% dei portatori e risultato giustamente negativo sul 99% delle persone sane. Supponiamo inoltre che la distribuzione media dei portatori tra la popolazione sia di 1 su 1000. Se ad una persona scelta a caso nella popolazione viene effettuato il test e questo da esito positivo, quale è la probabilità che questa persona sia effettivamente un portatore di AIDS?

* * * * *

P13) Attraverso una linea telefonica disturbata si vogliono trasmettere 4 numeri binari (ciascun numero è quindi 0 o 1). Il rumore sulla linea fa sì che il ricevente possa commettere un errore nell'identificare un numero inviato (cioè scambiare uno 0 con un 1 o viceversa) con probabilità 0.1 (cioè un errore in media ogni 10 numeri trasmessi). Supponendo che la linea permetta di trasmettere un numero al secondo, sareste in grado di inventarvi un modo che, utilizzando la linea per non più di 8 secondi, permetta di trasmettere i 4 numeri in modo tale che la probabilità di un errore nella ricezione sia inferiore a 0.2?

Si noti che spedendo i 4 numeri, senza alcuna precauzione, la probabilità di non avere commesso neppure un errore è $0.9^4 = 0.656$ (capite perchè?) Quindi la probabilità di aver commesso almeno un errore è $1 - 0.656 = 0.344$ ben maggiore della richiesta 0.2. Che fare?

Questa è una versione semplificata di problemi tipici che si affrontano nella teoria dei codici i cui risultati sono alla base degli sviluppi odierni della tecnologia delle comunicazioni.

5 L'impostazione del problema probabilistico

Illustreremo ora alcuni modi importanti e diffusi per la costruzione di modelli probabilistici che saranno utili per la soluzione dei vari problemi proposti.

Esperimenti ripetuti indipendenti Consideriamo un esperimento che ha come insieme dei possibili risultati

$$\Omega_o = \{\omega_1, \omega_2, \dots, \omega_n\}$$

Supponiamo ora di ripetere l'esperimento, un certo numero di volte, diciamo k e di annotare in ordine i risultati ottenuti. Avremo alla fine una sequenza ordinata di k elementi di Ω_o :

$$(a_1, a_2, \dots, a_k) \quad a_1, \dots, a_n \in \Omega_o$$

dove a_i indica l'esito dell' i -esimo esperimento. Questi sono i possibili esiti dei k esperimenti ripetuti in ordine. L'insieme di queste sequenze di lunghezza k è detto il prodotto cartesiano di Ω_o per se stesso k volte ed è indicato

$$\Omega_o^k = \{(a_1, a_2, \dots, a_k) \mid a_1, \dots, a_n \in \Omega_o\}$$

Si noti che la cardinalità di questo insieme, cioè il numero dei suoi elementi è

$$|\Omega_o^k| = n^k.$$

Che tipo di probabilità ha senso introdurre su Ω_o^k ? Questo dipende dal tipo di probabilità esistente su Ω_o che governava il singolo esperimento. Se l'esperimento di partenza soddisfaceva l'ipotesi classica, cioè se

$$\Pr(\omega_1) = \Pr(\omega_2) = \dots = \Pr(\omega_n) = \frac{1}{n}$$

è naturale continuare ad imporre l'ipotesi classica anche su Ω_o^k : non c'è motivo alcuno di ritenere una sequenza di risultati più probabile di un'altra. In effetti, si sta implicitamente anche facendo un'altra ipotesi e cioè che i vari esperimenti non si influenzino vicendevolmente, cioè che essi siano tra di loro indipendenti. Ritourneremo dopo su questo concetto. Con l'ipotesi classica si ha che

$$\Pr(a_1, a_2, \dots, a_k) = \frac{1}{n^k}$$

qualunque sia la sequenza $(a_1, a_2, \dots, a_k) \in \Omega^k$.

Risoluzione di P1) Cominciamo col considerare il primo gioco d'azzardo dei tempi di De Méré, cioè quello di ottenere almeno un 6 in 4 lanci di un dado.

In generale, se lanciamo un dado k volte, siamo esattamente nelle situazione appena descritta. Lo spazio degli eventi elementari è dato da

$$\Omega = \{1, 2, 3, 4, 5, 6\}^k.$$

Nel nostro caso specifico $k = 4$ e l'evento complesso del quale si vuole calcolare la probabilità è

$$A = \{(a_1, a_2, a_3, a_4) \in \Omega \mid \text{almeno un } a_i = 6\}$$

Poichè vale l'ipotesi classica si ha che

$$\Pr(A) = \frac{|A|}{|\Omega|}$$

Sappiamo che $|\Omega| = 6^4$. Non è tuttavia così semplice calcolare $|A|$. Proviamo a considerare il complementare

$$A^c = \{(a_1, a_2, a_3, a_4) \in \Omega \mid a_i \neq 6 \text{ per ogni } i\}$$

Chiaramente, $|A^c| = 5^4$ e quindi

$$\Pr(A) = 1 - \Pr(A^c) = 1 - \frac{5^4}{6^4} \sim 0.52$$

Consideriamo ora il problema P1). In questo caso l'esperimento consiste in 24 lanci di una coppia di dadi. Si ha dunque

$$\Omega = (\{1, 2, 3, 4, 5, 6\}^2)^{24} = \{1, 2, 3, 4, 5, 6\}^{48}$$

che contiene 6^{48} elementi! L'evento sul quale si scommette è

$$A = \{((a_1, b_1), \dots, (a_{24}, b_{24})) \mid \text{almeno una coppia } (a_i, b_i) = (6, 6)\}$$

Di nuovo ci conviene lavorare con il complementare

$$A^c = \{((a_1, b_1), \dots, (a_{24}, b_{24})) \mid (a_i, b_i) \neq (6, 6) \text{ per ogni } i\}.$$

Poichè le possibili coppie $(a_i, b_i) \in \{1, 2, 3, 4, 5, 6\}^2$ sono $6^2 = 36$ e si esclude una coppia $(6, 6)$, le coppie che possono apparire nelle sequenze di A^c sono 35. Dunque si ha che $|A^c| = 35^{24}$. Possiamo ora calcolare la probabilità voluta:

$$\Pr(A) = 1 - \Pr(A^c) = 1 - \frac{35^{24}}{36^{24}} = 0.49$$

Questo gioco è dunque leggermente sfavorevole alla casa come aveva intuito il De Méré. Se invece avessimo considerato 25 lanci, con gli stessi conti di prima avremmo ottenuto

probabilità di avere almeno una coppia di 6 uguale a circa 0.51: in questo caso il gioco sarebbe ritornato favorevole alla casa.

Vediamo di capire a questo punto se era così sbagliato il ragionamento che a De Méré aveva fatto credere che questo gioco avrebbe avuto per la ‘casa’ la stessa probabilità di vittoria di quello classico illustrato all’inizio. In entrambi i casi si sta ripetendo un esperimento un certo numero di volte, diciamo k . Ognuna di queste volte si sta guardando l’accadimento di un certo evento ω che ha probabilità diciamo p : nel primo caso questo evento è che esca 6 e la probabilità è $p = 1/6$, nel secondo caso è invece che esca la coppia (6, 6) e la probabilità è $p = 1/36$. La probabilità che non accada mai ω nei k lanci è data da

$$\Pr(\text{non accade mai } \omega) = (1 - p)^k$$

(ritorneremo su questo fatto, accontentiamoci di notare che questo è vero nei nostri due casi). Quindi la probabilità che si verifichi almeno una volta ω nei k esperimenti è data da

$$\Pr(\text{accade almeno una volta } \omega) = 1 - (1 - p)^k$$

Consideriamo la funzione

$$f(x) = (1 - x)^k, \quad x \in [0, 1]$$

Per x molto vicini a 0 essa si può approssimare con la sua retta tangente per il punto $(0, f(0))$; calcoliamo la derivata

$$f'(x) = -k(1 - x)^{k-1}, \quad f'(0) = -k$$

Poichè $f(0) = 1$ si ha che la retta tangente è data da $y - 1 = -kx$, cioè $y = 1 - kx$. Dunque

$$(1 - x)^k \simeq 1 - kx \quad \text{per } x \text{ molto vicini a } 0$$

Ritornando alle nostre probabilità, se p è abbastanza piccolo, si ha che

$$\Pr(\text{accade almeno una volta } \omega) \simeq 1 - (1 - kp) = kp$$

Nel primo caso si ha che $k = 4$ e $p = 1/6$, mentre nel secondo caso si ha che $k = 24$ e $p = 1/36$. In entrambi i casi si ha quindi

$$kp = 2/3$$

Quindi il De Méré non aveva poi sbagliato di tanto nel ritenere che i due giochi dovessero avere la stessa probabilità di vittoria in quanto questo prodotto si manteneva uguale nei due casi. In effetti i due risultati, come abbiamo visto, non differiscono di molto: tuttavia essi stanno da parti opposte rispetto a 0.5, fatto che li rende, dal punto di vista del gioco d’azzardo, ben diversi.

Estrazioni ripetute ordinate Supponiamo di avere un’urna con n biglie numerate da 1 ad n e di estrarre una biglia a caso dall’urna. Il modello probabilistico naturale è dunque

$$\Omega_o = \{1, 2, \dots, n\}, \quad \Pr(i) = \frac{1}{n}$$

Ci sono due modi di ripetere questo esperimento. Il primo consiste nel rimettere prima la biglia estratta nell'urna, dopo aver annotato il risultato, e quindi di effettuare una seconda estrazione. Ripetendo k volte l'estrazione in questo modo ci troviamo esattamente nella situazione descritta precedentemente di esperimenti ripetuti indipendenti. Un secondo modo è quello invece di non rimettere la biglia estratta nell'urna, e di effettuare un'altra estrazione con le biglie rimaste; se si effettuano k estrazioni consecutive di questo tipo (dette *estrazioni ordinate senza riimmisione*) alla fine otterremo una sequenza di k biglie in ordine fuori dell'urna. Gli eventi elementari sono quindi ancora sequenze (a_1, a_2, \dots, a_k) dove a_i indica il numero della i -esima biglia estratta. Non tutte le sequenze di questo tipo possono tuttavia apparire, il fatto che le biglie estratte non vengano riimmesse impone che i vari a_i siano tra loro diversi. Lo spazio degli eventi giusto per questo esperimento non è più il prodotto cartesiano Ω_o^k ma il suo sottoinsieme

$$\Omega_{od}^k = \{(a_1, a_2, \dots, a_k) \in \Omega_o^k \mid a_i \neq a_j \text{ se } i \neq j\}$$

Ω_{od}^k è detto insieme delle *disposizioni ordinate di k oggetti dell'insieme Ω_o* e la sua cardinalità è data da

$$|\Omega_{od}^k| = n(n-1) \cdots (n-k+1)$$

La probabilità naturale da considerare è quella classica, così che

$$\Pr(a_1, a_2, \dots, a_k) = \frac{1}{n(n-1) \cdots (n-k+1)}$$

per ogni estrazione $(a_1, a_2, \dots, a_k) \in \Omega_{od}^k$.

Nel caso in cui si compiano tante estrazioni quante sono le biglie, si ottengono in questo caso le disposizioni ordinate di n oggetti da un insieme che ne contiene n che vengono anche dette le *permutazioni di un insieme di n oggetti*, intuitivamente i modi distinti in cui posso ordinare in fila uno dopo un altro gli n oggetti. Si noti che in questo caso

$$|\Omega_{od}^n| = n(n-1) \cdots 3 \cdot 2 \cdot 1 = n!.$$

Esempio 1: Supponiamo di avere un'urna con 3 biglie numerate 1, 2, 3 e di compiere tre estrazioni consecutive dalla biglia senza riimmisione. Vogliamo calcolare la probabilità che in ognuna delle tre estrazioni non compaia mai una biglia riportante lo stesso numero dell'estrazione in corso; in altre parole vogliamo che alla prima estrazione non compaia 1, che alla seconda non compaia 2 e che alla terza non compaia 3. Il modello probabilistico è dato da

$$\Omega = \{1, 2, 3\}_d^3$$

con la probabilità classica, mentre

$$A = \{(2, 3, 1), (3, 1, 2)\}$$

Dunque,

$$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{2}{6} = \frac{1}{3}$$

Commenti su P5): Un modo di modellare l'esperimento in questione è il seguente: si supponga di aver attaccato un'etichetta numerata da 1 a 20 sulle donne e sugli uomini in modo tale che marito e moglie abbiano entrambi lo stesso numero. Si consideri poi un'urna con 20 biglie numerate da 1 a 20 e si compiano 20 estrazioni ordinate senza riimmissione; le estrazioni determinano gli accoppiamenti per il ballo nel modo seguente: se all' i -esima estrazione esce il numero k l'uomo con il numero i ballerà con la donna avente il numero k . Dunque gli eventi elementari dell'esperimento consistono nelle permutazioni di un insieme di 20 oggetti e la probabilità da considerare è evidentemente quella classica. L'evento complesso che ci interessa è dato dall'insieme A delle permutazioni in cui alla posizione i -esima non compare mai il numero i , cioè quelle corrispondenti alle estrazioni per le quali non capita mai che all' i -esima estrazione esca il numero i . Tali permutazioni si dicono *complete*. Si noti che siamo esattamente nella situazione illustrata nell'esempio precedente con l'unica differenza che prima avevamo 3 oggetti e ora 20. Nell'esempio precedente avevamo calcolato la cardinalità di A semplicemente elencando le permutazioni che lo componevano, solo due in quel caso; in questo caso i numeri in gioco non ci permettono di farlo! In realtà si riesce a risolvere esattamente il seguente problema generale: consideriamo lo spazio di probabilità Ω_n delle permutazioni di n oggetti con la probabilità classica e sia A_n l'evento complesso delle permutazioni di Ω_n complete nel senso definito prima. Allora, si può dimostrare, anche se non è affatto semplice che vale la seguente formula

$$\Pr(A_n) = \frac{1}{2!} - \frac{1}{3!} + \cdots \pm \frac{1}{n!} \quad (7)$$

dove l'ultima frazione va presa con il segno $+$ o $-$ a seconda che n sia pari o dispari. Nel nostro caso dunque in cui $n = 20$ si ottiene

$$\Pr(A) = \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{1}{20!} \sim 0.36788$$

La cosa interessante da notare è che la somma a secondo membro della formula (7) tende a stabilizzarsi velocemente, al crescere di n , ad un numero molto particolare $1/e$ dove e è il numero di Nepero. Per $n = 8$ siamo già estremamente vicini a questo numero (quattro cifre decimali eguali). Quindi le probabilità in questione cambiano pochissimo se alla festa da ballo ci fossero state ad esempio 10 o 50 coppie anziché 20!

Estrazioni ripetute disordinate Ci sono circostanze in cui è preferibile non mantenere l'ordine in cui si sono ripetuti gli esperimenti. Se noi estraiamo k biglie senza riimmissione da un'urna che ne contiene n e vogliamo considerare l'insieme delle k biglie estratte senza tener conto dell'ordine di estrazione, gli eventi elementari da considerare non sono più le disposizioni come prima, quanto i sottoinsiemi di k elementi di un insieme Ω_o che contiene n elementi (le possibili biglie). L'insieme dei sottoinsiemi di Ω_o che contengono k elementi si indica con il simbolo

$$\mathcal{P}_k(\Omega_o)$$

Gli elementi di $\mathcal{P}_k(\Omega_o)$ sono anche detti *combinazioni di k oggetti distinti dell'insieme Ω_o* . La cardinalità di tale insieme è data dal coefficiente binomiale

$$|\mathcal{P}_k(\Omega_o)| = \binom{n}{k}$$

Risoluzione di P2) In questo caso l'esperimento consiste nell'annotare la data di compleanno, cioè un numero da 1 a 365 (si suppone per semplicità che non ci siano persone nate il 29 febbraio di anni bisestili), per tutte e 23 le persone. L'insieme degli eventi elementari è dunque

$$\Omega = \{1, 2, 3, \dots, 365\}^{23}$$

e, senza altre informazioni, si può ipotizzare che non ci siano giorni preferiti per nascere e che quindi si possa utilizzare l'ipotesi classica. Chi è A ?

$$A = \{(a_1, \dots, a_{23}) \mid \text{almeno una coppia } a_i = a_j\}$$

Consideriamo il complementare

$$A^c = \{(a_1, \dots, a_{23}) \mid a_i \neq a_j \text{ se } i \neq j\}$$

Si noti che A^c coincide con l'insieme delle disposizioni ordinate di 23 oggetti da un insieme di 365. Dunque

$$|A^c| = 365 \cdot 364 \cdots 344 \cdot 343$$

e, con qualche calcolo si vede che

$$\Pr(A) = 1 - \Pr(A^c) = 1 - \frac{365 \cdot 364 \cdots 344 \cdot 343}{365^{23}} = 0.507$$

sorprendentemente maggiore di $1/2$! Nel caso le persone siano di più, la probabilità di una coincidenza di compleanni aumenta velocemente. Riportiamo sotto alcuni valori

Numero persone	Prob. almeno due persone con compleanni coincidenti
23	0.507
25	0.569
30	0.706
50	0.970

Risoluzione di P3) Dal punto di vista probabilistico nascita di femmina o maschio equivale ad uscita di testa T o croce C nel lancio di una moneta perfettamente equilibrata. P3) può essere quindi riformulato nel modo seguente: qual'è la probabilità che in 5 lanci di una moneta (perfettamente equilibrata) si ottenga almeno due volte T ? Risolviamo un problema più generale: calcoliamo intanto la probabilità che in n lanci di una moneta perfettamente equilibrata appaiano esattamente k volte T . L'esperimento è modellato dallo spazio prodotto

$$\Omega = \{T, C\}^n$$

con la probabilità classica. L'evento complesso che vogliamo considerare è

$$A = \{(a_1, a_2, \dots, a_n) \in \{T, C\}^n \mid \text{ci sono esattamente } k \text{ } T \text{ tra gli } a_i\}$$

Dobbiamo calcolare la cardinalità di A . Si noti che una sequenza in A è univocamente determinata una volta stabilite le k posizioni in cui deve esserci T . Le sequenze in A

sono dunque tante quanti i modi in cui posso scegliere k posizioni tra le n disponibili, cioè tante quante le combinazioni di k elementi da un insieme di n . Dunque:

$$|A| = \binom{n}{k}$$

e quindi

$$\Pr(A) = \frac{\binom{n}{k}}{2^n}$$

Se consideriamo invece l'evento complesso

$$B = \{(a_1, a_2, \dots, a_n) \in \{T, C\}^n \mid \text{ci sono almeno } k \text{ } T \text{ tra gli } a_i\}$$

si avrà che, per il principio di additività di eventi disgiunti,

$$\Pr(B) = \sum_{j=k}^n \frac{\binom{n}{j}}{2^n}$$

Vediamo di effettuare questo calcolo nell'esempio particolare di almeno due teste in 5 lanci:

$$\Pr(B) = \frac{\binom{5}{2}}{2^5} + \frac{\binom{5}{3}}{2^5} + \frac{\binom{5}{4}}{2^5} + \frac{\binom{5}{5}}{2^5} = \frac{10 + 10 + 5 + 1}{32} = \frac{26}{32}$$

Certo questo caso si sarebbe più agevolmente trattato considerando il complementare B^c : 'esce al più una T in 5 lanci'. In effetti,

$$\Pr(B^c) = \frac{\binom{5}{0}}{2^5} + \frac{\binom{5}{1}}{2^5} = \frac{6}{32}$$

$$\Pr(B) = 1 - \Pr(B^c) = 1 - \frac{6}{32} = \frac{26}{32}$$

Risoluzione di P4) Un mazzo di carte da poker contiene 32 carte divise in 4 colori: cuori, quadri, fiori, picche. Ogni colore contiene 8 carte con valori ordinati come segue $A, K, Q, J, 10, 9, 8, 7$. Una mano di poker consiste in 5 carte estratte dal mazzo. Il *colore* si ha quando le 5 carte sono appunto dello stesso colore ma non in scala, cioè non sono 5 valori consecutivi secondo l'ordine sopra riportato (si escludono così le configurazioni $A, K, Q, J, 10$, $K, Q, J, 10, 9$, $Q, J, 10, 9, 8$ e $J, 10, 9, 8, 7$. Queste configurazioni, quando le carte sono dello stesso colore si dicono *scala reale*. Il *poker* si ha invece quando le 5 carte estratte contengono quattro valori eguali, ad esempio quattro assi A o quattro 7. Quale è il corretto modello probabilistico? Si tratta di 5 estrazioni senza riimmisione da un insieme di 32 elementi. Si noti tuttavia che per il tipo di problemi ai quali siamo interessati, l'ordine delle estrazioni non riveste alcuna importanza e conviene trascurarlo. Lo spazio giusto è dunque quello delle combinazioni di 5 oggetti da un insieme di 32: Sia

Ω_o l'insieme delle 32 carte da poker. Allora lo spazio degli eventi elementari delle mani di poker è dato da

$$\Omega = \mathcal{P}_5(\Omega_o)$$

e

$$|\Omega| = \binom{32}{5}.$$

Se le estrazioni sono veramente casuali, è logico ipotizzare che valga l'ipotesi classica. Consideriamo l'evento complesso COLORE. Per calcolare il numero dei suoi elementi, procediamo nel modo seguente: il colore si può avere in 4 diversi colori, per ognuno di essi esso può verificarsi in tanti modi quanti sono quelli di scegliere 5 carte da un insieme di 8 togliendo le configurazioni che mi danno scala reale, cioè $\binom{8}{5} - 4$. Dunque,

$$|\text{COLORE}| = 4 \cdot \left[\binom{8}{5} - 4 \right] = 208$$

$$\Pr(\text{COLORE}) = \frac{|\text{COLORE}|}{|\Omega|} = \frac{208}{201376}$$

Consideriamo ora l'evento complesso POKER. Per calcolare il numero dei suoi elementi, procediamo nel modo seguente: il poker si può avere in 8 diversi valori; ho quindi 8 modi distinti di scegliere le 4 carte di valore eguale, rimane la quinta carta che posso scegliere in $32 - 4 = 28$ modi distinti. Dunque,

$$|\text{POKER}| = 8 \cdot 28 = 224$$

$$\Pr(\text{POKER}) = \frac{|\text{POKER}|}{|\Omega|} = \frac{224}{201376}$$

Quindi il poker ha leggermente più probabilità di accadere rispetto al colore. Per quanto riguarda invece la scala reale il conto è molto semplice:

$$|\text{SCALA REALE}| = 4 \cdot 4 = 16$$

$$\Pr(\text{SCALA REALE}) = \frac{|\text{SCALA REALE}|}{|\Omega|} = \frac{16}{201376}.$$

Si noti che, poichè POKER e SCALA REALE sono eventi disgiunti

$$\begin{aligned} \Pr(\text{COLORE} \cup \text{SCALA REALE}) &= \Pr(\text{COLORE}) + \Pr(\text{SCALA REALE}) \\ &= \frac{208}{201376} + \frac{16}{201376} = \frac{224}{201376} \\ &= \Pr(\text{POKER}) \end{aligned}$$

Quindi abbiamo il fatto curioso che la probabilità di un colore o scala reale è uguale alla probabilità di poker.

Esercizio 1: Calcolare la probabilità che in una mano di poker si abbia un *full* (si ha il full quando si hanno tre carte dello stesso valore, cioè un tris, e le altre due carte anch'esse dello stesso valore, cioè una coppia).

Risoluzione di P5) La cosa più complicata è capire chi è in questo caso lo spazio degli eventi elementari. Si può pensare di avere fissato i 10 posti al tavolo rotondo numerati da 1 a 10 e di aver numerato le dieci persone con un'etichetta 1, 2, 3, 4, 5 le donne, 6, 7, 8, 9, 10 gli uomini. Si prende un'urna con le biglie da 1 a 10 e si compiono 10 estrazioni ordinate senza riimmissione: se alla i -esima estrazione esce il numero k la persona k si siede al posto i . Quindi Ω corrisponde all'insieme delle permutazioni di 10 oggetti. Quindi

$$|\Omega| = 10!$$

La probabilità è evidentemente quella classica. Chi è l'evento complesso A corrispondente ad avere uomini e donne alternati al tavolino. Si può scrivere facilmente come unione di due insiemi disgiunti:

$$A_1 = \{(a_1, \dots, a_{10}) \mid \text{se } i \text{ dispari } a_i \in \{1, 2, 3, 4, 5\} \text{ se } i \text{ pari } a_i \in \{6, 7, 8, 9, 10\}\}$$

$$A_2 = \{(a_1, \dots, a_{10}) \mid \text{se } i \text{ pari } a_i \in \{1, 2, 3, 4, 5\} \text{ se } i \text{ dispari } a_i \in \{6, 7, 8, 9, 10\}\}$$

Calcoliamo $|A_1|$: i modi in cui posso collocare i numeri $\{1, 2, 3, 4, 5\}$ nelle 5 posizioni dispari è esattamente uguale al numero di permutazioni di un insieme di 5 oggetti, cioè è uguale a $5!$. Analogamente i modi in cui posso collocare i numeri $\{6, 7, 8, 9, 10\}$ nelle 5 posizioni pari è uguale a $5!$. Quindi $|A_1| = (5!)^2$. Similmente si vede che $|A_2| = (5!)^2$.

$$A = A_1 \cup A_2, \quad |A| = |A_1| + |A_2| = 2 \cdot (5!)^2$$

Dunque

$$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{2 \cdot (5!)^2}{10!} = \frac{1}{2 \cdot 9 \cdot 7} \sim 0.008$$

C'è un altro modo di affrontare e risolvere questo problema, decisamente più semplice. Se non distinguiamo tra loro le 5 donne e i 5 uomini e numeriamo nuovamente i posti a tavolo da 1 a 10, la configurazione al tavolo è completamente determinata se diciamo in quali posti siedono le donne; essendo queste 5 si tratta semplicemente di determinare 5 posizioni tra le 10 possibili. Se fissiamo

$$\Omega_o = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

si ha che lo spazio degli eventi elementari secondo questa nuova interpretazione è dato dalle combinazioni di 5 oggetti dall'insieme Ω_o cioè è dato da

$$\Omega = \mathcal{P}_5(\Omega_o)$$

In questo nuovo modello l'evento del quale vogliamo calcolare la probabilità è

$$A = \{\{1, 3, 5, 7, 9\}, \{2, 4, 6, 8, 10\}\}$$

queste sono in effetti le uniche configurazioni che vedranno le donne alternarsi agli uomini. Dunque,

$$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{2}{\binom{10}{5}} = \frac{2}{\frac{10!}{5!5!}} = \frac{2(5!)^2}{10!}$$

che da lo stesso risultato (meno male!) di prima.

6 La probabilità condizionata

Consideriamo uno spazio di probabilità (Ω, \Pr) che modella un certo esperimento. Supponiamo che in qualche modo veniamo in possesso dell'informazione che si verificherà senz'altro un evento elementare che sta in un certo sottoinsieme A . A fronte di questa nuova informazione, come dovremmo modificare il nostro originale modello probabilistico?

Iniziamo con un esempio, il lancio del dado. Ricordiamo che in questo caso si ha

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \quad \Pr(i) = 1/6 \text{ per } i = 1, \dots, 6$$

Sia

$$A = \{2, 4, 6\}$$

Supponiamo di sapere che si verificherà un evento elementare in A cioè che uscirà un numero pari (non ha importanza come si sia avuta questa informazione, si può ad esempio pensare che il dado sia stato già lanciato ed il risultato sia a noi ignoto e che qualcun'altro ci fornisca questa informazione). E' logico a questo punto cambiare il modello probabilistico e costruire una nuova probabilità (che indicheremo $\tilde{\Pr}$ per distinguerla da quella originale) che concentri tutta la probabilità su A :

$$\tilde{\Pr}(1) = \tilde{\Pr}(3) = \tilde{\Pr}(5) = 0$$

Come si dovrà ripartire la probabilità sui numeri 2, 4, 6? E' logico imporre che siano rispettate le proporzioni tra le vecchie probabilità; poichè in questo caso 2, 4, 6 avevano la stessa probabilità di accadimento prima, la devono avere anche nel nuovo modello. Si pone quindi

$$\tilde{\Pr}(2) = \tilde{\Pr}(4) = \tilde{\Pr}(6) = 1/3.$$

Quella appena costruita è un esempio di probabilità condizionata che ora sotto definiamo nel caso generale.

Sia (Ω, \Pr) uno spazio di probabilità e sia $A \subseteq \Omega$ un evento complesso tale che $\Pr(A) > 0$. Si definisce *la probabilità condizionata di un evento elementare ω_i sapendo A* e si indica con il simbolo $\Pr(\omega_i|A)$, come

$$\omega_i \notin A \Rightarrow \Pr(\omega_i|A) = 0$$

$$\omega_i \in A \Rightarrow \Pr(\omega_i|A) = \frac{\Pr(\omega_i)}{\Pr(A)}$$

Quindi nel nuovo modello probabilistico, condizionato dall'informazione ricevuta, gli eventi elementari che non stanno in A hanno probabilità 0, quelli che stanno in A hanno probabilità data dalla vecchia probabilità divisa per la probabilità di A (vedremo tra un attimo il motivo di questa divisione). Si noti che in questo modo le proporzioni tra le nuove probabilità degli eventi elementari in A sono uguali alle vecchie e

$$\sum_{\omega_i \in \Omega} \Pr(\omega_i|A) = \sum_{\omega_i \in A} \Pr(\omega_i|A) = \sum_{\omega_i \in A} \frac{\Pr(\omega_i)}{\Pr(A)} = 1$$

Quindi $\Pr(\omega_i|A)$ è una vera e propria probabilità su Ω (in realtà concentrata sul sottoinsieme A). La divisione per $\Pr(A)$ serve proprio per avere la normalizzazione ad 1.

Come ogni probabilità, la probabilità condizionata si estende agli eventi complessi. Sia $B \subseteq \Omega$ un qualunque altro evento complesso. Si definisce *la probabilità condizionata di B sapendo A* come

$$\Pr(B|A) = \sum_{\omega_i \in B} \Pr(\omega_i|A) \quad (8)$$

Si noti che si può scrivere

$$\sum_{\omega_i \in B} \Pr(\omega_i|A) = \sum_{\omega_i \in B \cap A} \frac{\Pr(\omega_i)}{\Pr(A)} = \frac{\Pr(B \cap A)}{\Pr(A)}$$

Si ha quindi

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} \quad (9)$$

formula molto importante che ci sarà utile in seguito.

Risoluzione di P7) Per illustrare correttamente l'esperimento considerato in P7) dobbiamo innanzitutto considerare il sesso dei due figli che supponiamo essere distinguibili in primogenito, secondogenito (supponiamo dunque che non siano gemelli). Questo dà luogo a spazio degli eventi elementari dato da $\{M, F\}^2$ (M sta per maschio, F per femmina). Dobbiamo poi considerare l'aspetto di chi viene ad aprire la porta, può essere il primogenito od il secondogenito. Lo spazio degli eventi elementari corretto è dunque

$$\Omega = \{M, F\}^2 \times \{1, 2\}$$

I suoi elementi sono terne, i primi due elementi delle quali sono lettere F o M e stanno ad indicare il sesso rispettivamente del primogenito e del secondogenito, il terzo elemento può essere 1 o 2 e sta ad indicare chi dei due viene ad aprire la porta. Ad esempio (MF2) significa che il primogenito è maschio, il secondogenito è femmina e viene ad aprire il secondogenito. Che probabilità mettiamo su Ω ? Poichè per ipotesi nascita di maschio e femmina sono equiprobabili, se ipotizziamo che con la stessa probabilità può venire ad aprirci il primogenito od il secondogenito è naturale imporre l'ipotesi classica su Ω . Consideriamo gli eventi complessi A : 'viene ad aprirci un maschio' e B : 'chi non viene ad aprirci è femmina'. Allora vogliamo calcolare proprio $\Pr(B|A)$. Si ha che

$$A = \{MM1, MM2, MF1, FM2\}, \quad B = \{MF1, FM2, FF1, FF2\}$$

$$B \cap A = \{MF1, FM2\}$$

Quindi,

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{2/8}{4/8} = \frac{1}{2}.$$

E' importante notare come il meccanismo aleatorio di quale dei due figli viene ad aprire alla porta giuochi un ruolo fondamentale nel problema P7). Se in effetti considerassimo semplicemente le famiglie con due figli avremmo che lo spazio degli eventi elementari naturale sarebbe

$$\Omega = \{M, F\}^2$$

Consideriamo ora gli eventi complessi A : 'uno dei due figli è maschio' e B : 'uno dei due figli è femmina'. Allora

$$A = \{MM, MF, FM\}, \quad B = \{MF, FM, FF\}$$

Quindi,

$$A \cap B = \{MF, FM\}$$

e

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{2/4}{3/4} = \frac{2}{3}.$$

Cioè sapendo che uno dei due figli è maschio, la probabilità che l'altro sia femmina è $2/3$ e non $1/2$ come si poteva aver pensato e come accadeva nel problema P7)! Apparentemente l'informazione che ho è la stessa nei due casi: so che uno dei due figli è maschio. Tuttavia il modello probabilistico è diverso nei due casi.

Descriviamo ora due importanti risultati che riguardano il concetto di probabilità condizionata e che svolgono un ruolo chiave nel calcolo delle probabilità.

Theorem 1 (delle probabilità totali) *Sia (Ω, \Pr) uno spazio di probabilità e siano A_1, A_2, \dots, A_r eventi complessi di Ω tra loro disgiunti, cioè $A_i \cap A_j = \emptyset$ se $i \neq j$ e tali che*

$$A_1 \cup A_2 \cup \dots \cup A_r = \Omega \tag{10}$$

(Si dice in tal caso che gli A_i formano una partizione di Ω). Supponiamo inoltre che $\Pr(A_i) > 0$ per ogni $i = 1, \dots, r$. Allora per ogni evento complesso $B \subseteq \Omega$ si ha che

$$\Pr(B) = \Pr(B|A_1)\Pr(A_1) + \Pr(B|A_2)\Pr(A_2) + \dots + \Pr(B|A_r)\Pr(A_r) = \sum_{i=1}^r \Pr(B|A_i)\Pr(A_i)$$

Proof Segue dalla (10) che

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_r)$$

Inoltre, poichè gli A_i sono tra loro disgiunti, anche i $B \cap A_i$ sono tra loro disgiunti. Utilizzando dunque la proprietà di additività della probabilità si ha che

$$\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \dots + \Pr(B \cap A_r)$$

D'altra parte, segue dalla formula (9), che

$$\Pr(B \cap A_i) = \Pr(B|A_i)\Pr(A_i)$$

Sostituendo si ottiene esattamente la formula voluta. ■

Quando è utile il Teorema delle probabilità totali? Se conosciamo esplicitamente la probabilità di tutti gli eventi elementari, evidentemente $\Pr(B)$ si può calcolare direttamente in base alla definizione, sommando le probabilità di tutti gli eventi elementari che lo compongono. Ci sono tuttavia situazioni in cui \Pr non è conosciuta esplicitamente; si ha invece una partizione dello spazio Ω in sottoinsiemi A_i dei quali è conosciuta la probabilità $\Pr(A_i)$ e sono pure conosciute le probabilità condizionate sugli elementi della partizione $\Pr(B \cap A_i)$. Da questi, con il Teorema delle probabilità totali, si può allora calcolare la probabilità di un qualsiasi evento complesso B .

Esempio 2: Supponiamo di avere due urne: la prima U_1 contiene 3 biglie nere e 5 bianche, mentre la seconda U_2 contiene 4 biglie nere e 3 bianche, Scelgo a caso un'urna ed estraggo una biglia; quale è la probabilità che la biglia estratta sia nera? Lo spazio degli eventi elementari è un prodotto del tipo

$$\Omega = \{U_1, U_2\} \times \{N, B\}$$

Il primo fattore codifica l'urna scelta, il secondo il colore estratto. Non si può certo mettere la probabilità classica visto gli squilibri tra le biglie di diverso colore nelle due urne, quindi non è immediato capire quali siano le probabilità degli eventi elementari. Consideriamo gli eventi complessi U_1 : 'scelgo l'urna U_1 ' ed U_2 : 'scelgo l'urna U_2 '. Essi partiscono lo spazio Ω e si ha chiaramente

$$\Pr(U_1) = \Pr(U_2) = \frac{1}{2}$$

Inoltre conosco le probabilità condizionate a questi due eventi, in quanto si tratta di semplici estrazioni da un'urna fissata. Si ha in effetti, guardando le proporzioni tra colori nelle due urne

$$\Pr(N|U_1) = \frac{3}{8}, \quad \Pr(N|U_2) = \frac{4}{7}$$

dove N indica l'evento che si estragga una biglia di colore nero. Usando il Teorema delle probabilità totali si ha quindi che

$$\Pr(N) = \Pr(N|U_1)\Pr(U_1) + \Pr(N|U_2)\Pr(U_2) = \frac{3}{8} \cdot \frac{1}{2} + \frac{4}{7} \cdot \frac{1}{2} = \frac{53}{112} \sim 0.47$$

Altre applicazioni del Teorema delle probabilità totali, si vedranno successivamente congiuntamente al prossimo fondamentale risultato.

Theorem 2 (di Bayes) *Sia (Ω, \Pr) uno spazio di probabilità e siano A_1, A_2, \dots, A_r eventi complessi di Ω tra loro disgiunti, cioè $A_i \cap A_j = \emptyset$ se $i \neq j$ e tali che*

$$A_1 \cup A_2 \cup \dots \cup A_r = \Omega.$$

Supponiamo inoltre che $\Pr(A_i) > 0$ per ogni $i = 1, \dots, r$. Sia $B \subseteq \Omega$ evento complesso tale che $\Pr(B) > 0$. Si ha allora che, per ogni $k = 1, \dots, r$,

$$\Pr(A_k|B) = \frac{\Pr(B|A_k)\Pr(A_k)}{\sum_{i=1}^r \Pr(B|A_i)\Pr(A_i)}$$

Proof Sfruttando la formula (9), si ha

$$\Pr(A_k|B) = \frac{\Pr(A_k \cap B)}{\Pr(B)} = \frac{\Pr(B \cap A_k)}{\Pr(B)} = \frac{\Pr(B|A_k)\Pr(A_k)}{\Pr(B)}$$

Sostituendo al posto di $\Pr(B)$ l'espressione trovata nel Teorema delle probabilità totali si ottiene la formula cercata. ■

Come nel caso del Teorema delle probabilità totali, il Teorema di Bayes risulta utile in tutti quei casi in cui conosciamo le probabilità condizionate agli elementi della partizione $\Pr(B|A_i)$ e le probabilità dei singoli elementi della partizione $\Pr(A_i)$, e vogliamo calcolare le probabilità 'inverse' $\Pr(A_k|B)$. I prossimi esempi illustrano una varietà di applicazioni di questo importante risultato. Cominciamo illustrando un modo alternativo di risolvere il problema P7).

Risoluzione alternativa di P7) Ricordiamo che il modello è

$$\Omega = \{M, F\}^2 \times \{1, 2\}$$

con la probabilità classica. Consideriamo la partizione di Ω corrispondente alla tipologia di sesso dei due figli, cioè a seconda che siano entrambi maschi A_1 , entrambi femmine A_2 ovvero uno maschio ed uno femmina A_3 . Formalmente

$$A_1 = \{MM1, MM2\}, \quad A_2 = \{FF1, FF2\}, \quad A_3 = \{MF1, MF2, FM1, FM2\}$$

Sia poi B l'evento complesso 'viene ad aprirci un maschio'. Si noti che noi vogliamo proprio calcolare $\Pr(A_3|B)$. Usando il Teorema di Bayes si ha che

$$\Pr(A_3|B) = \frac{\Pr(B|A_3)\Pr(A_3)}{\Pr(B|A_1)\Pr(A_1) + \Pr(B|A_2)\Pr(A_2) + \Pr(B|A_3)\Pr(A_3)}$$

Si noti ora che è molto facile calcolare le probabilità $\Pr(B|A_i)$. In effetti si ha

$$\Pr(B|A_1) = 1, \quad \Pr(B|A_2) = 0, \quad \Pr(B|A_3) = 1/2$$

Inoltre si ha

$$\Pr(A_1) = \Pr(A_2) = 1/4, \quad \Pr(A_3) = 1/2$$

Sostituendo si ottiene,

$$\Pr(A_3|B) = \frac{\frac{1}{2} \frac{1}{2}}{1 \frac{1}{4} + 0 \frac{1}{4} + \frac{1}{2} \frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

Esercizio 2:) Risolvere il problema P8) utilizzando il Teorema di Bayes.

Risoluzione di P11) Il modello probabilistico che si può considerare è il seguente:

$$\Omega = \{V, F\} \times \{+, -\}$$

V ed F stanno per vero o falso a seconda che la persona scelta dica la verità oppure no. Il secondo attributo $+$ o $-$ codifica l'esito della macchina della verità. Gli eventi

elementari sono dunque le coppie $V+, V-, F+, F-$. Useremo i simboli V, F ad indicare gli eventi complessi, rispettivamente, ‘dice la verità’, ‘non dice la verità’. Similmente $+$ e $-$ indicheranno anche gli eventi complessi, rispettivamente, ‘risulta positivo al test’, ‘risulta negativo al test’. Quello che noi vogliamo calcolare è $\Pr(F|+)$. Si noti che F e V compiono una partizione dello spazio Ω e in base ai dati si ha che

$$\Pr(F) = \frac{5}{1000}, \quad \Pr(V) = \frac{995}{1000}$$

Inoltre si ha che

$$\Pr(+|F) = \frac{76}{100}, \quad \Pr(+|V) = \frac{37}{100}$$

Usando il Teorema di Bayes si ottiene così

$$\Pr(F|+) = \frac{\Pr(+|F)\Pr(F)}{\Pr(+|F)\Pr(F) + \Pr(+|V)\Pr(V)} = \frac{\frac{76}{100} \frac{5}{1000}}{\frac{76}{100} \frac{5}{1000} + \frac{37}{100} \frac{995}{1000}} = 0.01$$

Esercizio 3:) Risolvere il problema P12) utilizzando il Teorema di Bayes.

Introduciamo ora un altro concetto fondamentale del calcolo delle probabilità, quello di indipendenza. Supponiamo di avere uno spazio di probabilità (Ω, \Pr) e siano A e B due eventi complessi. Si dice che A e B sono tra loro *indipendenti* se

$$\Pr(A \cap B) = \Pr(A)\Pr(B) \tag{11}$$

Per capire il senso di questa definizione, supponiamo che $\Pr(B) > 0$. Allora si ha che, se A e B soddisfano la (11), allora

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)\Pr(B)}{\Pr(B)} = \Pr(A) \tag{12}$$

cioè, il sapere che si è verificato l’evento B non mi dà alcuna informazione probabilistica sul verificarsi dell’evento A ; in questo senso sono indipendenti. Analogamente, se $\Pr(A) > 0$, si avrà, per simmetria che

$$\Pr(B|A) = \Pr(B) \tag{13}$$

Vicendevolmente, si ha che A non dà informazioni probabilistiche su B . Avremmo potuto usare la (12) o la (13) per definire l’indipendenza; la (11) è tuttavia da preferirsi in quanto mostra la simmetria tra A e B e non necessita di ipotesi sulla positività di $\Pr(A)$ o di $\Pr(B)$.

Esempio 3: Supponiamo di considerare un certo esperimento modellato dallo spazio di probabilità (Ω, \Pr) . Se ripetiamo questo esperimento k volte, abbiamo precedentemente osservato che il corretto spazio degli eventi elementari da considerare è il prodotto cartesiano Ω^k . Avevamo anche notato che se su Ω avevamo la probabilità classica, la stessa si sarebbe dovuta avere anche sul prodotto. Che cosa accade quando non abbiamo la probabilità classica su Ω ? Quale probabilità a senso introdurre sul prodotto in questo caso?

Se reputiamo i vari esperimenti tra loro indipendenti è logico imporre che la probabilità su Ω^k sia tale che se si considerano eventi relativi ad esperimenti distinti, essi risultino indipendenti nel senso della definizione data prima. Consideriamo un evento elementare $(a_1, a_2, \dots, a_k) \in \Omega^k$ e gli eventi complessi A_i : 'accade a_i nell' i -esimo esperimento'. Allora si ha

$$\{(a_1, a_2, \dots, a_k)\} = A_1 \cap A_2 \cap \dots \cap A_k$$

e da una ripetuta applicazione della (11) si ottiene che

$$\begin{aligned} \Pr(a_1, a_2, \dots, a_k) &= \Pr(A_1 \cap A_2 \cap \dots \cap A_k) \\ &= \Pr(A_1)\Pr(A_2) \dots \Pr(A_k) \\ &= \Pr(a_1)\Pr(a_2) \dots \Pr(a_k) \end{aligned}$$

Dunque, imponendo l'indipendenza, la probabilità di una sequenza di esiti è semplicemente il prodotto delle probabilità dei singoli esiti.

Esempio 4: Si noti che gli eventi A : 'viene ad aprire un maschio' e B : 'chi non viene ad aprire è femmina', introdotti nella risoluzione di P7) sono tra loro indipendenti. Si ha in questo caso

$$\Pr(A) = \Pr(B) = \frac{1}{2}, \quad \Pr(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$$

Invece i due eventi A : 'uno dei due figli è maschio' e B : 'uno dei due figli è femmina', introdotti nella variante discussa dopo la risoluzione di P7) sono tra loro dipendenti: in effetti si ha in questo caso

$$\Pr(A) = \Pr(B) = \frac{3}{4}, \quad \Pr(A \cap B) = \frac{1}{2} \neq \frac{3}{4} \cdot \frac{3}{4}$$

Commenti su P13): Ogni volta che spediamo un bit sulla rete può verificarsi un errore con probabilità $p = 0, 1$. Lo spazio degli eventi per questo esperimento è quindi

$$\Omega_o = \{c, e\}$$

dove c sta per 'corretto' ed e per 'errato'; le probabilità sono

$$\Pr(c) = 0.9 \quad \Pr(e) = 0.1$$

Se spediamo n bit, lo spazio da considerare è dunque

$$\Omega = \{c, e\}^n$$

e, considerando gli errori sulle varie spedizioni, eventi indipendenti tra loro, la probabilità da considerare sarà quella data dal prodotto come illustrata nell'Esempio 3. Quindi se $(a_1, a_2, \dots, a_n) \in \Omega = \{c, e\}^n$ è tale che ci sono k e ed $n - k$ c , si avrà che

$$\Pr(a_1, a_2, \dots, a_n) = (0.1)^k (0.9)^{n-k}$$

Sia A_k l'evento complesso 'ci sono esattamente k errori e '. A_k è costituito da sequenze che hanno tutte probabilità $(0,1)^k(0,9)^{n-k}$. Quante sono queste sequenze? Tante quanti i modi di scegliere k posizioni tra le n disponibili, cioè $\binom{n}{k}$. Si notino le analogie con le considerazioni fatte nella risoluzione di **P3**. Dunque si ha

$$\Pr(A_k) = \binom{n}{k} (0.1)^k (0.9)^{n-k} \quad (14)$$

Fatte queste premesse, arriviamo al nostro problema. Se spediamo i nostri 4 bit sulla rete direttamente la probabilità di non commettere alcun errore sarà data da

$$\Pr(A_0) = \binom{4}{0} (0.1)^0 (0.9)^4 = (0.9)^4 = 0.656$$

quindi la probabilità di aver commesso almeno un errore sarà di 0.344 ben superiore alla soglia 0.2 che non vorremmo superare. Visto che ci viene concesso di usare la linea per 8 volte a fronte dei soli 4 bit che dobbiamo spedire, potremmo pensare di sfruttare la ridondanza concessa spedendo due volte ciascun bit. Vediamo che cosa otteniamo in questo modo. Ogni bit si è trasformato in una coppia di bit identici che vengono trasmessi sulla linea. Se al ricevitore arriva effettivamente una coppia di bit identici, (0,0) ad esempio, ipotizzerà che sia stato inviato effettivamente lo 0. Cometterà errore solo nel caso fosse stata spedita la coppia (1,1) e ci siano stati due errori nella trasmissione; questo avviene con probabilità $(0.1)^2 = 0.01$. Se al ricevitore arriva invece una coppia del tipo (0,1) saprà certamente che c'è stato un errore, ma non avrà alcun modo per sapere quale bit era stato effettivamente inviato 0 o 1, può scegliere uno dei due a caso e cometterà errore con probabilità 0.5. Si può utilizzare il Teorema delle probabilità totali per calcolare la probabilità di non commettere errore con questa nuova strategia

$$\Pr(\text{ricevuto } (0,0) \text{ o } (1,1)) = \Pr(0 \text{ o } 2 \text{ errori}) = (0.9)^2 + (0.1)^2 = 0.82$$

$$\Pr(\text{ricevuto } (0,1) \text{ o } (1,0)) = \Pr(1 \text{ errore}) = 2(0.9)(0.1) = 0.18$$

$$\begin{aligned} \Pr(\text{errore}) &= \Pr(\text{errore}|\text{ricevuto } (0,0) \text{ o } (1,1))\Pr(\text{ricevuto } (0,0) \text{ o } (1,1)) \\ &+ \Pr(\text{errore}|\text{ricevuto } (0,1) \text{ o } (1,0))\Pr(\text{ricevuto } (0,1) \text{ o } (1,0)) \\ &= 0.01 \cdot 0.82 + 0.5 \cdot 0.18 = 0.0982 \end{aligned}$$

leggermente più bassa di prima. La probabilità di non commettere alcun errore in ricezione sarà data ora da

$$\Pr(A_0) = \binom{4}{0} (0.0982)^0 (0.9018)^4 = (0.9018)^4 = 0.661$$

quindi la probabilità di aver commesso almeno un errore sarà di 0.339 ancora ben superiore alla soglia 0,2. Che fare?

Si noti che i messaggi possibili che si possono spedire utilizzando 4 bit sono $2^4 = 16$. Usando 8 bit per la trasmissione si hanno a disposizione ben $2^8 = 256$ messaggi diversi.

L'idea è allora di associare ad ognuno dei 16 messaggi da spedire $a_1, a_2, \dots, a_{16} \in \{0, 1\}^4$, un messaggio tra i 256 disponibili nella trasmissione. Questa si chiama *codifica*: ogni messaggio iniziale $a_i \in \{0, 1\}^4$ viene codificato in un nuovo messaggio $m_i \in \{0, 1\}^8$ che viene poi trasmesso. Introduciamo un concetto di distanza tra gli elementi di $\{0, 1\}^8$: la distanza tra due elementi è il numero di posizioni sui quali differiscono. Ad esempio,

$$(0, 1, 1, 1, 0, 1, 0, 0), \quad (1, 1, 0, 1, 0, 1, 0, 1)$$

hanno distanza 3 poichè differiscono sul primo, sul terzo e sull'ottavo bit. Il ricevitore è a conoscenza della codifica effettuata in fase di trasmissione ed adotta la seguente strategia di decodifica: se gli arriva un m_i lo decodifica nel corrispondente a_i . Se invece gli arriva un messaggio $f \in \{0, 1\}^8$ che non è alcuno degli m_i , va a scegliere l' m_i più vicino di tutti al suo f e decodifica quindi in a_i . E' chiaro che una codifica sarà tanto più buona quanto più i 16 messaggi m_1, m_2, \dots, m_{16} saranno lontani tra loro di modo che errori di trasmissione non trasformino un m_i in un altro. Nella codifica che avevamo proposto all'inizio in cui un bit semplicemente veniva fatto raddoppiare, avevamo messaggi che avevano distanza 2 tra loro. Supponiamo ora che esista una codifica per la quale la distanza tra due qualunque m_i e m_j sia almeno 3. Per il modo in cui avviene la decodifica, si ha che se nella trasmissione degli 8 bit non si verificano errori o se ne verifica soltanto 1, il ricevitore decodifica nel corretto messaggio inviato. In effetti, ipotizzando che sia stato spedito il messaggio m_j , se non ci sono errori, al ricevitore arriva proprio m_j che lui decodificherà correttamente; se invece vi è un errore, al ricevitore arriva un messaggio f che dista 1 da m_j e (ricordando che gli m_i distano tra loro almeno 3) almeno 2 da tutti gli altri m_i . Il ricevitore dunque decodificherà di nuovo nel corretto messaggio m_j . Quindi errori si commettono solo se c'è più di un errore. D'altra parte, utilizzando la (14), si ha che

$$\Pr(0 \text{ o } 1 \text{ errore in } 8 \text{ trasmissioni}) = (0.9)^8 + 9(0.9)^7(0.1) = 0.86$$

$$\Pr(\text{più di un errore in } 8 \text{ trasmissioni}) = 0.14$$

Quindi

$$\Pr(\text{almeno un errore nella trasmissione del messaggio di 4 bit}) \leq 0.14 < 0.2!$$

Il problema è dunque se si riescono a trovare 16 sequenze binarie in $\{0, 1\}^8$ che distino tra loro almeno 3. Questo si può fare: provate a pensarci...

Il problema che abbiamo appena affrontato è un esempio di problemi che si incontrano in Teoria dei codici dove ci si pone appunto l'obbiettivo di determinare strategie di codifica e decodifica che permettano di ridurre il più possibile la probabilità di commettere errori nella trasmissione su un canale disturbato. Questi codici hanno applicazioni enormi e li trovate ad esempio al lavoro in qualsiasi modem o telefono cellulare.