IL CONCETTO DI PROBABILITÀ*

IVAN CASAGLIA

Liceo Scientifico Castelnuovo, Firenze

1. Probabilità e decisioni

In molte situazioni della vita quotidiana ci capita di dover prendere delle decisioni in condizioni d'incertezza. L'incertezza non riguarda soltanto eventi futuri, ma anche quegli eventi passati sui quali non siamo in possesso di informazioni complete. Può essere naturale, ad esempio, chiedersi se domani pioverà e intuire che si possa dare una risposta in termini di probabilità. Ma, anche quando ci chiediamo se sia piovuto in città il 15 agosto del 1974, siamo in una situazione d'incertezza. Certo, in questo caso possiamo sperare di rintracciare in una biblioteca la copia di un quotidiano del 16 agosto del 1974, o accedere ai dati rilevati da un centro di osservazione meteorologica, e dare una risposta certa. Finché però non disponiamo di queste informazioni, restiamo in una condizione d'incertezza anche se l'evento in questione appartiene al passato. Possiamo ritenere piuttosto improbabile che sia piovuto nel giorno di Ferragosto, ma non possiamo neppure escluderlo del tutto (del resto i temporali estivi sono un fenomeno tutt'altro che raro). A ben vedere l'incertezza può riguardare anche un terreno che, a torto o a ragione, è stato da sempre considerato il regno delle verità indiscutibili, quale può essere la matematica. Se in questo momento ci chiediamo, ad esempio, quale sia la decima cifra della rappresentazione decimale di π, siamo di nuovo in una situazione d'incertezza. E in tale condizione siamo costretti a restare fino a che non ricorriamo a un qualche strumento di calcolo che ci permetta di stabilire quale sia quella cifra (ma potrebbe bastare anche qualche manuale o uno dei numerosi siti internet dedicati a π).

Il calcolo delle probabilità, come teoria matematica, nasce per tentare di affrontare l'incertezza in modo ragionevole. Le sue origini non sono particolarmente nobili, dal momento che si collocano tra i tavoli dei giochi d'azzardo, di dadi o di carte, che di questa nuova teoria scientifica costituirono il primo laboratorio. Ma già in questa prima fase il legame della probabilità col problema della scelta di fronte all'incertezza è subito evidente: le questioni da cui presero le mosse Galileo, Pascal, Fermat per fondare il calcolo delle probabilità, riguardavano proprio quale fosse il comportamento più adeguato che i giocatori avrebbero dovuto tenere nei diversi giochi d'azzardo considerati. Oggi la probabilità, o per meglio dire l'insieme delle discipline statistico-probabilistiche, costituisce uno dei settori più vivaci della ricerca matematica contemporanea.

---

* Lezione tenuta a Pistoia il 3 novembre 2009, presso l'Aula Magna dell'ISPIA-Pacinotti, nell'ambito dell'edizione 2009 di *Pianeta Galileo*.

Trova applicazioni vastissime che vanno dallo studio della struttura della materia al controllo dell'affidabilità degli impianti tecnologici, dai modelli per le previsioni meteorologiche ai sondaggi sugli orientamenti elettorali, dall'analisi del comportamento del consumatore in microeconomia allo studio dell'evoluzione nel tempo del valore di un titolo finanziario, offrendo sempre, in contesti così diversi, strumenti concettuali utili ad orientarsi in condizioni di incertezza.

Parlare della probabilità non è un compito che possa esaurirsi nel tempo destinato a questa lezione-incontro. Mi limiterò dunque a svolgere alcune considerazioni sulla probabilità e sulle convinzioni più diffuse intorno a questo concetto, che spero possano costituire, per ciascuno di noi, uno stimolo a studiare e approfondire l'argomento e a riflettervi sopra con la dovuta attenzione.

Per fare questo mi servirò di due situazioni problematiche che mi sembrano molto interessanti e che sono tratte da un libro che non parla di probabilità e non è stato scritto da un matematico. Il libro è *L'illusione di sapere* e l'autore è Massimo Piattelli Palmarini [8], uno studioso di psicologia cognitiva. In questo saggio del 1993 rivolto ai non specialisti vengono presentati i risultati di un settore della ricerca psicologica[1] che ha fatto emergere la presenza di una sorta di *inconscio cognitivo* che a nostra insaputa interviene nei nostri giudizi, facendoci commettere degli errori di valutazione, talvolta clamorosi, che vengono denominati *tunnel della mente*.

È davvero sorprendente constatare quanti di questi tunnel mentali coinvolgano le nostre scelte e i nostri comportamenti di fronte all'incertezza, cioè un nostro punto di vista più o meno spontaneo sulla probabilità.

I casi che voglio proporvi permetteranno a un tempo di riflettere sui diversi e possibili significati della probabilità e di vedere come si possa utilizzare questo concetto, in concreto, per affrontare delle situazioni caratterizzate dall'incertezza. Cominciamo da un caso giudiziario.

SOFISMA DEL GIURATO

> Siete membro di una giuria popolare. Un tassista è accusato di aver investito un passante in una notte tempestosa, e di essere poi fuggito senza prestare aiuto. Il pubblico ministero, nel richiedere la condanna dell'imputato, basa <u>tutto</u> sulla testimonianza di una anziana signora che dalla sua finestra, a una certa distanza, ha visto l'incidente. La signora afferma di aver visto investire il malcapitato da un taxi blu, e di aver poi visto fuggire il taxi. L'imputato lavora in una compagnia di taxi che possiede solo macchine blu. Nel corso dell'istruttoria e del dibattito processuale è emerso quanto segue:
> 1) in quella città operano due sole compagnie di taxi, una che ha tutte le vetture verdi, e una che ha tutte le vetture blu. Di tutti i taxi circolanti quella notte circa l'85% erano verdi e circa il 15% erano blu.
> 2) La signora, testimone a carico, sulla base di ripetute prove di acutezza visiva, effettuate dal giudice istruttore in condizioni di illuminazione molto simili a quelle della notte dell'incidente, ha dimostrato di saper correttamente identificare un taxi blu, rispetto ad uno verde, 80 volte su 100.

Sulla base della testimonianza giurata della signora, e sulla base dei dati 1) e 2), qual è la probabilità che il taxi fosse veramente blu?

2. Probabilità classica e probabilità statistica

Per prima cosa interroghiamoci sul significato di probabilità. Il testo infatti esplicita questo termine solo nella domanda finale. Eppure, a ben vedere, di probabilità troviamo già nella prima parte alcune *tracce* che ci conducono in modo naturale a considerare due situazioni *tipiche* che sono all'origine delle concezioni della probabilità più diffuse e più conosciute.

Quando si dice, ad esempio, che quella notte era in circolazione un 85% di taxi verdi e un 15% di taxi blu si forniscono due informazioni che interpretabili come probabilità. Nel senso che se avessimo potuto scegliere a caso un taxi quella notte, come si estrae un numero dal sacchetto della tombola, avremmo avuto una probabilità dell'85% che fosse verde e un 15% che fosse blu. Questa impostazione è quella che va sotto il nome di "probabilità classica": si immagina di condurre un esperimento o un gioco in cui si hanno degli esiti semplici (nel lancio di una moneta, Testa o Croce; nel lancio di un dado, una delle facce, 1 o 2 o 3 o 4 o 5 o 6). Quando non si ha ragione di dubitare che monete o dadi o qual si voglia altro dispositivo, siano truccati, si può ipotizzare che i singoli eventi elementari abbiano la stessa probabilità e, assumendo che quest'ultima possa esprimersi come un rapporto e una percentuale, si dice, ad esempio, che la probabilità di ottenere Testa o Croce è del 50%. Analogamente, la probabilità che l'esito del lancio di un dado sia una qualsiasi delle facce è $\frac{1}{6}$. In questo schema la probabilità di un evento E viene definita come il numero:

$$\mathbf{P}(E) = \frac{\text{numero dei casi favorevoli}}{\text{numero dei casi possibili}}.$$

Tornando al nostro esempio, supponiamo per semplicità che i taxi in circolazione, la notte del delitto, fossero 20, e di questi 17 fossero verdi e 3 blu. La probabilità che un taxi scelto a caso fosse verde è:

$$\mathbf{P}(V) = \frac{17}{20} = 0{,}85 = 85\%,$$

mentre quella che fosse blu è:

$$\mathbf{P}(B) = \frac{3}{20} = 0{,}15 = 15\%.$$

Quando invece, al punto 2), si dice che la testimone ha dimostrato di saper identificare correttamente il colore del taxi all'80% siamo di fronte all'altro schema tipico, quello della probabilità *statistica* o *frequentista.* Questo schema interviene quando si esamina un fenomeno *ripetibile* e, osservato un numero *grande* di queste ripetizioni (o prove), si va a considerare il numero dei *successi*. In questo caso la probabilità di un evento E è identificata con la *frequenza relativa* dei successi, cioè col numero:

$$\mathbf{P}(E) = \frac{\text{numero dei successi}}{\text{numero delle prove}}.$$

Se, ad esempio, la testimone del sofisma è stata sottoposta per 10 volte al riconoscimento del colore di un taxi in condizioni di illuminazione simili a quelle della notte del delitto, e per 8 volte lo ha identificato correttamente, sbagliando in due casi, per l'evento:

E = La testimone riconosce correttamente il colore del taxi

avremo:

$$\mathbf{P}(E) = \frac{8}{10} = 80\%.$$

Questa seconda impostazione ha il merito di potersi applicare in molte situazioni nelle quali non si può utilizzare quella classica. Prendiamo ad esempio il lancio di una moneta. Se riteniamo che la moneta non sia truccata giudicheremo (per quelle che in Fisica si direbbero *ragioni di simmetria*) che le probabilità che escano Testa o Croce siano uguali:

$$\mathbf{P}(T) = \mathbf{P}(C) = 50\%.$$

Ma cosa accade se invece di una moneta lanciamo una puntina da disegno? Si tratta di un oggetto non simmetrico dal punto di vista fisico. Qual è la probabilità che la puntina caschi 'di piatto' piuttosto che 'di punta'? Non abbiamo la possibilità di risolvere questa questione in termini di casi possibili e di casi favorevoli e allora non ci resta che lanciare ripetutamente la puntina da disegno, registrare gli esiti dei lanci, ed esprimere la probabilità come rapporto tra il numero dei successi e il numero complessivo dei lanci.

Qual è il collegamento tra queste due impostazioni della probabilità? Quella che nel linguaggio comune viene detta *legge dei grandi numeri* ma che sarebbe più corretto indicare come *postulato empirico del caso*, secondo il quale, quando un certo evento E è caratterizzato da una probabilità p e si può osservare una sequenza di n prove relative a quell'evento, la *frequenza relativa* dei successi f_n, cioè il rapporto tra il numero dei successi e il numero n degli eventi, tende a p al crescere di n.

In simboli:

$$f_n \to p \text{ quando } n \to \infty.$$

La decisione di valutare una probabilità, che non sappiamo calcolare in modo classico, ricorrendo allo schema delle frequenze (purché si abbia a disposizione un numero *sufficientemente grande* di prove), si fonda sull'accettazione, più o meno esplicita, di questo postulato.

Intorno a queste due impostazioni che, come abbiamo già detto, sono quelle più diffuse anche fuori dalla cerchia di chi si occupa della probabilità, si sono sviluppate nel tempo numerose discussioni e polemiche, delle quali non possiamo certo dare conto qui, ma che non sono difficili da prospettare. Entrambe le definizioni presentano infatti delle difficoltà sia di ordine logico che di ordine, per così dire, pratico, applicativo.

Nell'impostazione classica della probabilità, ad esempio, si presuppone che tutti gli eventi elementari abbiano la stessa probabilità. Questo presupposto presenta una doppia difficoltà: di ordine logico perché si vuol definire un concetto, quello di probabilità, utilizzandolo nella definizione quando si parla di eventi *equiprobabili*; di ordine applicativo perché, se il dispositivo che si utilizza (moneta, dado, ecc.) non è simmetrico, questa definizione ci lascia del tutto impotenti nel fare previsioni. Che dire poi della definizione statistica, che pure riceve i maggiori favori? Cosa significa che un evento è ripetibile? In un certo senso ogni evento è di per sé irripetibile. Per quanto si cerchi di ricostruire, ad esempio, le condizioni di illuminazione della notte del delitto, come essere sicuri di esserci riusciti? Quali e quanti fattori oggettivi e soggettivi possono avere influito sul riconoscimento del colore del taxi? E ancora, quand'è che il numero delle prove fatte può essere considerato sufficientemente grande per far sì che la frequenza relativa rappresenti un'approssimazione accettabile della probabilità? Ce la sentiremmo di trarre delle conclusioni che preludono ad una sentenza sulla base di dieci prove di acutezza visiva? E anche ammesso che questo numero, o uno maggiore, sia ritenuto adeguato, come ci si può affidare alla frequenza relativa sapendo che questa dipende dalla particolare sequenza di prove che si sono effettuate? Supponiamo, per capire meglio questa domanda, che la sequenza delle dieci prove abbia dato i seguenti esiti, indicando con 1 il riconoscimento del colore (successo) e con 0 l'errore:

$$1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 1.$$

Se consideriamo la frequenza relativa sulle prime cinque prove, dovremmo concludere che la testimone è in grado di riconoscere correttamente il colore del taxi nel 100% dei casi?

Naturalmente, mostrando questi elementi critici delle due impostazioni, non voglio con ciò affermare che esse siano prive di significato o di importanza. Se così fosse questo mio intervento potrebbe chiudersi qui, senza dare una risposta al problema iniziale. Si capisce però che in qualche modo si debba tentare di uscire dalle difficoltà che abbiamo evidenziato, se non vogliamo dare l'idea che di probabilità si possa parlare solo a condizione di essere in presenza di casi particolarissimi, rinunciando cioè a una formulazione generale del concetto di probabilità.

2.1 Probabilità soggettiva

La via di uscita, che a me pare più convincente, è quella offerta dall'impostazione *soggettiva* della probabilità, che dobbiamo essenzialmente ai contributi di un grande matematico italiano del secolo scorso: Bruno De Finetti. Secondo questa definizione la probabilità di un evento E è il *grado di fiducia* (espresso con un numero compreso tra 0 e 1) che un individuo attribuisce al verificarsi di quell'evento. Questa definizione può apparire, in prima battuta, un po' troppo generica per poter essere applicata, ma così non è. Proviamo comunque a riformularla in un modo più operativo, partendo da un esperimento mentale. Supponiamo di trovarci nella città in cui si è svolto il delitto e supponiamo inoltre che esista un'unica centrale telefonica che smista le richieste di taxi. Devo chiamare un taxi per andare alla stazione e, prima di fare la telefonata o mentre aspetto che mi rispondano, scommetto con un amico sul colore del taxi che mi verrà a prelevare. Quanto sarò disposto a scommettere per guadagnare, ad esempio, 1 euro nel caso che a presentarsi alla porta sia un taxi blu (e niente se si presenta un taxi verde)? È proprio questa somma, il prezzo per partecipare alla scommessa, che nell'impostazione soggettiva rappresenta il grado di fiducia, cioè la probabilità che un individuo attribuisce al verificarsi dell'evento oggetto della scommessa. La definizione più operativa di probabilità soggettiva è dunque la seguente:

> la probabilità $\mathbf{P}(E)$ di un evento E è il prezzo che un individuo giudica equo pagare per ricevere 1 euro nel caso in cui si verifichi l'evento considerato (e niente altrimenti).

Occorre però precisare che cosa significa giudicare equo cioè la condizione di *equità* (o di *coerenza*):

> *le probabilità devono essere assegnate in modo che non sia possibile ottenere una vincita o una perdita certa.*

Vediamo più da vicino la questione nel nostro caso. Gli eventi che possono verificarsi sono solo due: B (taxi blu) o V (taxi verde). Supponiamo che io attribuisca, sulla base delle mie informazioni, a B la probabilità $\mathbf{P}(B) = 80\%$ e a V la probabilità $\mathbf{P}(V) = 10\%$. È facile osservare che questa valutazione non può funzionare, perché altrimenti posso combinare insieme due scommesse realizzando un guadagno certo, dando luogo a quella che viene indicata come una *scommessa olandese.* Potrei infatti scommettere, con un primo amico, sul verificarsi di B pagando 80 centesimi, e con un secondo amico, sul verificarsi di V pagando 10 centesimi. In questo modo incasserei sicuramente 1 euro, avendo pagato 90 centesimi! Si può allora concludere che è necessario assegnare le probabilità in modo che:

$$\mathbf{P}(B) + \mathbf{P}(V) = 1,$$

relazione, quest'ultima, che esprime nel caso più semplice la *condizione di coerenza*.

Ma come posso, nel senso indicato dall'impostazione soggettiva, valutare in concreto la probabilità degli eventi B e V per partecipare alla scommessa?

Se io so soltanto che i taxi in circolazione in quella città possono essere blu o verdi, non mi resta che assegnare $\mathbf{P}(B) = \mathbf{P}(V) = 50\%$: che ragioni avrei per preferire uno dei due colori?

Se invece sono un utilizzatore abituale dei taxi e ho regolarmente registrato nel tempo il colore delle vetture che mi sono venute a prendere, posso far uso della frequenza relativa calcolata a partire dalle osservazioni. Qui si capisce subito che nella valutazione della probabilità un individuo cerca di utilizzare le informazioni di cui dispone. In questa prospettiva possiamo recuperare gli schemi che sono alla base dell'impostazione classica o di quella frequentista, che possono infatti fornire le valutazioni più sensate in certe situazioni. Queste valutazioni restano comunque soggettive: la loro maggiore o minore plausibilità ed efficacia dipende dalla loro 'ragionevolezza' e non può dunque essere dedotte da nient'altro che dal grado di fiducia che esse possano rappresentare correttamente nella situazione in esame.

Ancora, supponiamo che io sappia che al momento di effettuare la scommessa, i taxi in circolazione siano 17 blu e 3 verdi. In questo caso potrò valutare:

$$\mathbf{P}(V) = 85\% \text{ e } \mathbf{P}(B) = 15\%.$$

Sempre che sia disposto a identificare la procedura di individuazione del taxi, che effettuerà il servizio, con l'estrazione di un numero della tombola da un sacchetto o da un'urna. Le cose, in effetti, avvengono in modo un po' diverso. Se, per esempio, abitassi in periferia e sapessi che i taxi della compagnia blu operano prevalentemente in centro, sarei indotto a modificare le valutazioni precedenti, incrementando $\mathbf{P}(V)$ e diminuendo $\mathbf{P}(B)$.

In seguito, non possedendo informazioni aggiuntive, partiremo dalla valutazione che abbiamo formulato sulla base delle informazioni contenute nel testo-problema e sulla base dello schema degli eventi equiprobabili, per cui $\mathbf{P}(V) = 85\%$ e $\mathbf{P}(B) = 15\%$.

3. Probabilità condizionata

Tornando ora al sofisma del giurato, provate voi a dare una prima stima della probabilità che viene richiesta. Nella mia esperienza di insegnante, quando ho sottoposto questo sofisma ai miei studenti, ho registrato reazioni simili a quelle che Piattelli Palmarini descrive nel suo libro essere state le reazioni delle persone a cui è stato sottoposto il sofisma come test. Una parte significativa colloca questa probabilità intorno all'80%, identificandola sostanzialmente con la capacità della teste di distinguere i due colori. Ma in questo modo non si tiene in alcun conto il fatto che la probabilità che un taxi capitato a caso in quella strada sia blu è davvero bassa, solo il 15%. Una parte cerca dunque di correggere questa stima indicando probabilità inferiori all'80%, ma la stragrande maggioranza opta comunque per valori superiori al 50%. Pochissimi stimano la probabilità inferiore al 50% o superiore all'80%. Come stanno davvero le cose?

Dobbiamo intanto fare un po' di chiarezza perché la probabilità che dobbiamo determinare è di un tipo leggermente diverso rispetto a quelle che abbiamo considerato finora.

Leggiamo attentamente il testo: ci viene richiesto di stimare la probabilità che il taxi visto sul luogo del delitto fosse blu *sapendo* che la testimone a carico ha dichiarato che il taxi era blu. Se ci fosse stato richiesto di indicare la probabilità che il taxi fosse blu avremmo risposto, come abbiamo visto, $\mathbf{P}(B) = 15\%$. Ma questa è una valutazione che non ha niente a che fare con le dichiarazioni della testimone in aula. Quella che invece ci viene richiesta – che nel calcolo delle probabilità viene definita *probabilità condizionata* – deve tener conto del contenuto della testimonianza. Se indichiamo con *TB* l'evento *La testimone dichiara che il taxi era blu*, vogliamo determinare la probabilità di *B* condizionata a *TB*, in simboli:

$$\mathbf{P}\left(B \middle| TB\right).$$

Facciamo ora un secondo esperimento mentale, per capire bene il significato della probabilità condizionata. Supponiamo che, come accade nelle nostre città, a ogni taxi sia associato un *codice*. Supponiamo di sapere che i codici delle compagnie della nostra città siano costituiti da una lettera seguita da un numero a due cifre, e che, in quella notte, fossero in circolazione i taxi con i seguenti codici:

COMPAGNIA TAXI VERDI			COMPAGNIA TAXI BLU
V01	V13	V32	B02
V02	V17	V39	B14
V08	V18	V42	B32
V10	V22	V51	
V11	V28	V52	
V12	V29		

Supponiamo inoltre che la notte del delitto un taxista in vacanza chiami per telefono un taxi per recarsi in aeroporto. Perché proprio un taxista? Perché possiamo supporre che conosca quali sono i taxi di turno quella notte. Supponiamo inoltre, come prima, che esista un'unica centrale telefonica che smista le chiamate. Alla richiesta telefonica di un cliente l'operatore della centrale, dopo aver cercato un taxi libero, risponde comunicando il codice del taxi che presterà il servizio. Qual è la probabilità che il taxi che andrà a prendere il nostro taxista in vacanza sia di colore blu? La domanda non ammette una sola risposta perché dipende dallo stato dell'informazione.

Prima di chiamare la centrale, o anche durante la chiamata, mentre il nostro taxista in ferie ascolta l'odiosa musichetta che lo separa dalla risposta dell'operatore, la probabilità che quel taxi sia di colore blu è ancora $\mathbf{P}(B) = 15\%$. Ma dopo che l'operatore avrà risposto, indicando il codice del taxi che sta per arrivare, la situazione è completamente modificata. Se indichiamo con *V13* l'evento «l'operatore comunica che il taxi che effettuerà il servizio è il V13», ciò a cui siamo interessati è la probabilità di *B* condizionata a V13, cioè la probabilità che si verifichi l'evento *B* sapendo che si è verificato l'evento *V13*. In questo caso $\mathbf{P}\left(B \middle| V13\right) = 0$, perché il taxi di codice V13 è di colore verde. Analogamente $\mathbf{P}\left(V \middle| V13\right) = 1$. Ma questo che abbiamo considerato è un caso limite: da un caso incerto siamo passati di colpo a un caso certo!

Analizziamo allora un caso diverso: supponiamo cha la telefonata sia fatta non dal taxista che, in ritardo, si sta affannosamente facendo il nodo alla cravatta, ma da sua moglie. Quando quest'ultima raggiunge il marito e lui le chiede qual è il numero del taxi che sta arrivando, la moglie, assonnata, risponde «Non ho capito di preciso perché la comunicazione era disturbata, salvo il fatto che il codice del taxi finisce con il 2». Indichiamo con S l'evento che rappresenta questa nuova informazione: «L'operatore comunica che il taxi che effettuerà il servizio ha un codice che termina con 2». I taxi che hanno il codice che finisce per 2 sono i seguenti.

COMPAGNIA TAXI VERDI	COMPAGNIA TAXI BLU
V32	B02
V02	
V42	B32
V22	
V52	
V12	

Con l'informazione che ha ottenuto, il taxista in vacanza può ora valutare la probabilità che il taxi che effettuerà il servizio sia blu come:

$$\mathbf{P}(B|S) = \frac{2}{8} = 25\%.$$

Una rappresentazione ci aiuterà a capire meglio come stanno le cose. Consideriamo, per cominciare, l'insieme dei taxi in circolazione la notte del delitto:

$$\Omega = \{V01, V02, V08, V10, V11, V12, V13, V17, V18, V22, V28, V29, V32, V39, V42, V51, V52, B02, B14, B32\}$$

Questo insieme, nel calcolo delle probabilità, prende il nome di *spazio delle possibilità o spazio campione* e rappresenta l'insieme di tutti gli esiti elementari dell'esperimento a cui siamo interessati (nel nostro caso la scelta a caso di un taxi fra quelli in circolazione nella fatidica notte in cui si è consumato il delitto). L'evento B = *Il taxi era blu* può essere rappresentato dall'insieme:

$$B = \{B02, B14, B32\},$$

mentre l'evento V = *Il taxi era verde*, sarà rappresentato dall'insieme:

$$V = \{V01, V02, V08, V10, V11, V12, V13, V17, V18, V22, V28, V29, V32, V39, V42, V51, V52\}$$

Tutto questo, cioè le informazioni che il taxista possiede prima di chiamare la centrale telefonica, può essere rappresentato anche con un diagramma.

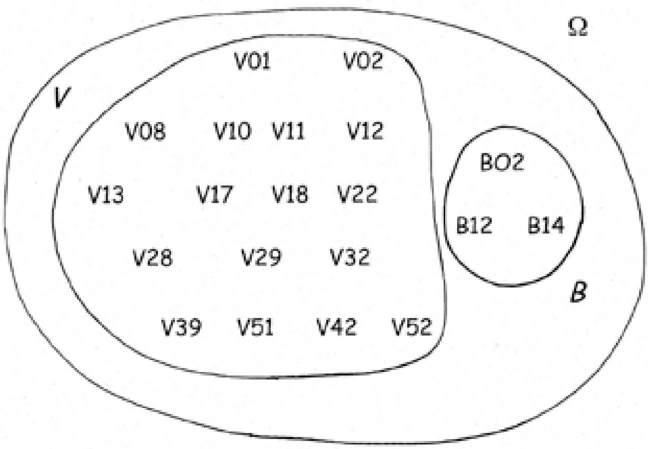

Figura 1

Evidenziamo nella figura seguente l'insieme *S* che corrisponde all'informazione che il taxista ottiene dalla moglie.

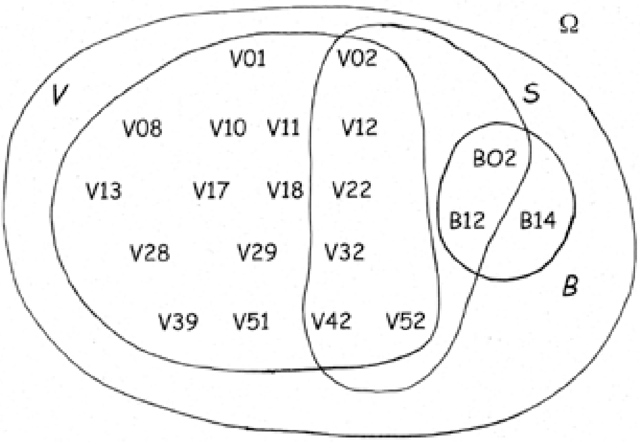

Figura 2

Quando il turista viene a sapere che il codice del taxi che verrà a prenderlo finisce con 2, questa informazione produce una *contrazione* dello spazio campione (v. Fig. 3), che a questo punto, per il prosieguo dell'esperimento è diventato ,*S* e l'evento *B*, quello a cui eravamo interessati, si è contratto in $B \cap S$.

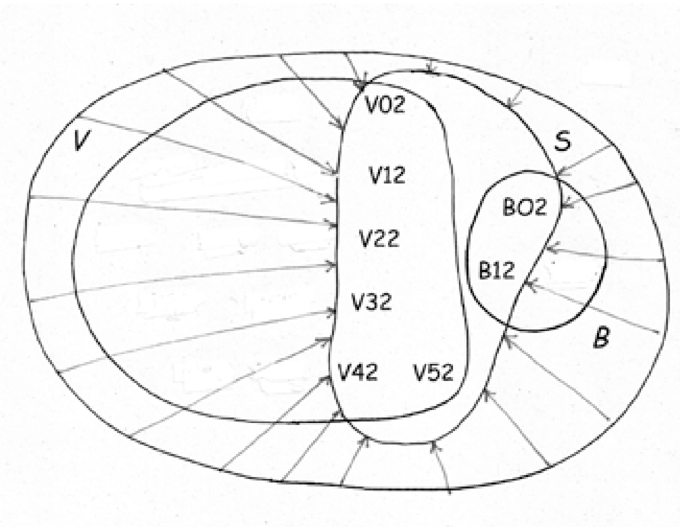

Figura 3

In questa nuova situazione la probabilità che il taxi che si presenterà alla porta sia di colore blu si è modificata. Sempre nell'ipotesi dell'equiprobabilità degli eventi elementari, avremo:

$$\mathbf{P}(B|S) = \frac{\text{numero elementi di } B \cap S}{\text{numero elementi di } S}.$$

Se vogliamo, possiamo esprimere questa relazione anche in termini di probabilità degli eventi considerati. Basta dividere numeratore e denominatore per il numero degli elementi di Ω e abbiamo:

$$\mathbf{P}(B|S) = \frac{\dfrac{\text{numero elementi di } B \cap S}{\text{numero elementi di } \Omega}}{\dfrac{\text{numero elementi di } S}{\text{numero elementi di } \Omega}} = \frac{\mathbf{P}(B \cap S)}{\mathbf{P}(S)}.$$

Questa relazione, che abbiamo ricavato nel caso particolarissimo in esame, ha però un significato del tutto generale e viene assunta come *definizione* della probabilità condizionata.

La rappresentazione con i diagrammi di Eulero-Venn consente una visualizzazione efficace ed un'interpretazione qualitativa della probabilità condizionata. Possiamo fare un passo in avanti se, considerato che la probabilità è un rapporto, perfezioniamo la rappresentazione in modo da valutare la probabilità in termini di rapporto tra aree delle figure che rappresentano gli eventi. Se infatti l'insieme B è costituito dal 15% degli elementi di Ω, possiamo rappresentare B come una regione con un'area pari al 15% dell'area della regione che rappresenta Ω. In questo contesto la contrazione delle possibilità determinata dall'informazione ottenuta, produce un nuovo spazio delle possibilità, l'evento S, nel quale la probabilità $\mathbf{P}(B|S)$ può essere correttamente interpretata come rapporto tra l'area di $B \cap S$ e l'area di S.

4. La legge di Bayes

Ora possiamo finalmente tornare al SOFISMA DEL GIURATO e determinare la probabilità a cui siamo interessati.

Per prima cosa rappresentiamo lo spazio campione, cioè l'insieme dei taxi in circolazione quella notte, come un quadrato di lato 1, e al suo interno visualizziamo gli eventi B (= *Il taxi era blu*) e V (= *Il taxi era verde*) con due rettangoli di basi, rispettivamente, 0,15 e 0,85. Poiché l'area della figura che rappresenta Ω è uguale a 1, le aree dei rettangoli possono essere lette direttamente come le probabilità degli eventi corrispondenti.

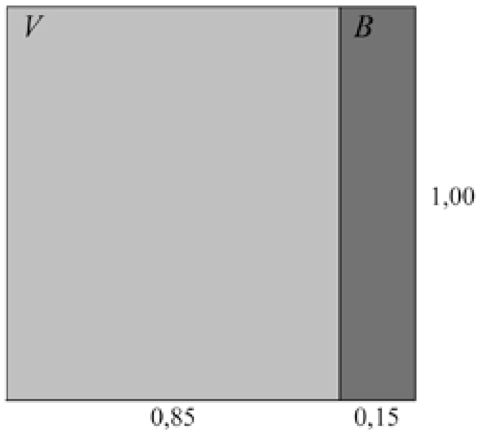

Figura 4

Allo stesso modo possiamo rappresentare nel quadrato Ω gli eventi che corrispondono al riconoscimento del colore del taxi da parte della teste con due rettangoli, che per comodità rappresenteremo in orizzontale, entrambi di base 1 e altezze 0,80 e 0,20 rispettivamente.

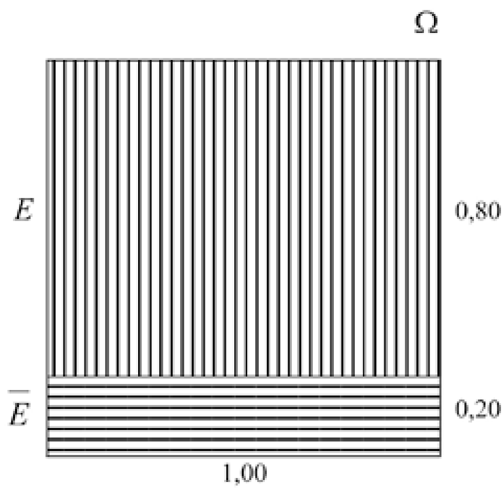

Figura 5

Nella figura, ovviamente, E rappresenta l'evento *La teste riconosce correttamente il colore* *del taxi* e $\overline{E}$ il suo contrario. Se ora combiniamo le due figure, sovrapponendole, otteniamo una nuova rappresentazione che permetterà di tenere conto di tutte le informazioni in nostro possesso.

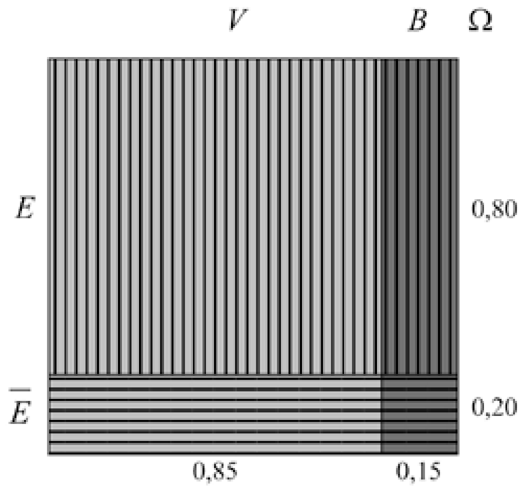

Figura 6

Che cosa rappresentano i rettangoli ottenuti sovrapponendo gli eventi B, V, E ed $\overline{E}$?

Per cominciare, consideriamo l'evento $B \cap E$ che, nella figura, è rappresentato dal rettangolo in alto a destra. Quando si verificano entrambi gli eventi B ed E, il taxi nella strada era blu e la teste ne ha riconosciuto correttamente il colore e dunque testimonierà che il taxi era blu. D'altra parte anche nel caso dell'evento $V \cap \overline{E}$, cioè nel caso in cui il taxi era verde e la testimone non riconosce il colore, essa testimonierà che il taxi era blu.

Ora, se indichiamo con *TB* l'evento *La testimone dichiara che il taxi era blu*, si ha che questo evento è ottenuto 'incollando' i due eventi $B \cap E$ e $V \cap \overline{E}$ e dunque può essere rappresentato come nella figura che segue.

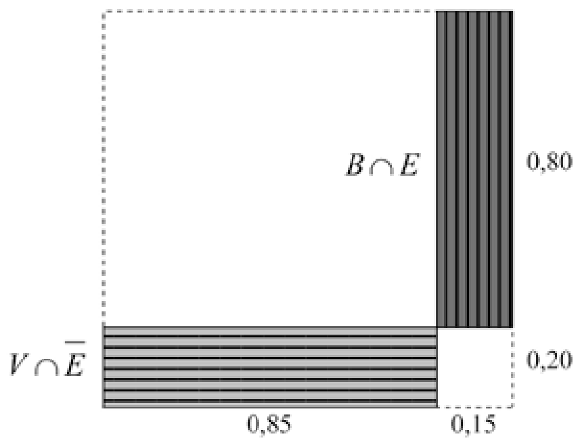

Figura 7

In modo analogo, se consideriamo l'evento TV = *La testimone dichiara che il taxi era verde*, potremo rappresentarlo così:

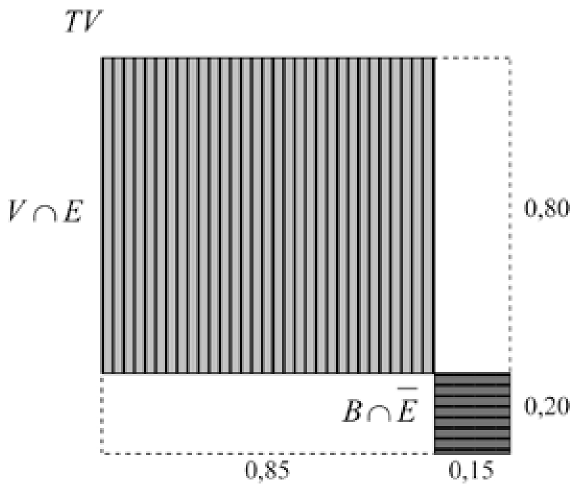

Figura 8

Siamo finalmente in grado di risolvere il problema iniziale: conoscere qual è la probabilità che il taxi che abbandonò il luogo del delitto fosse blu, sapendo che la testimone ha dichiarato che il colore del taxi era blu. Ora, posto che si sia verificato TB, dobbiamo considerare che nella contrazione dello spazio campione da Ω a TB, l'evento B si restringe a $B \cap TB$ che poi è $B \cap E$.

Interpretando la probabilità in termini di aree dei rettangoli, dalla figura 7, ricaviamo finalmente:

$$\mathbf{P}(B|TB) = \frac{\mathbf{P}(B \cap TB)}{\mathbf{P}(TB)} = \frac{15\% \cdot 80\%}{15\% \cdot 80\% + 85\% \cdot 20\%} \cong 41\%.$$

Naturalmente, se sostituiamo B con V, ricaviamo l'altra probabilità finale, che del resto possiamo anche ricavare tornando a considerare la figura 8, come:

$$\mathbf{P}(V|TB) = \frac{\mathbf{P}(V \cap TB)}{\mathbf{P}(TB)} = \frac{85\% \cdot 20\%}{15\% \cdot 80\% + 85\% \cdot 20\%} \cong 59\%.$$

Non ci meravigliamo certo del fatto che $\mathbf{P}(B|TB) + \mathbf{P}(V|TB) = 100\%$, si tratta pur sempre delle due sole ipotesi alternative!

Sorpresi del risultato? La cosa più sorprendente, a dire il vero, è che siamo riusciti ad affrontare il sofisma utilizzando soltanto le aree di alcuni rettangoli! Anche perché il risultato che abbiamo ottenuto va ben al di là della soluzione del sofisma. La relazione che abbiamo trovato è infatti quella che Massimo Piattelli Palmarini definisce nel suo libro, una delle «più grandi scoperte dell'intelligenza umana», un modo per «ricavare il noto dall'ignoto». Si tratta della *Legge o Teorema di Bayes,* che non ha niente di miracoloso, ma permette di utilizzare al meglio le informazioni di cui siamo in possesso, riuscendo addirittura a rovesciarne in qualche modo il significato.

Nel sofisma del giurato ci troviamo di fronte a una situazione tipica, che è caratteristica di molti contesti sperimentali. Possiamo infatti guardare al caso esaminato come a un *esperimento* svolto per confrontare due *ipotesi* alternative: che il taxi fosse di colore blu (B) o che il taxi fosse di colore verde (V). A queste ipotesi alternative sono assegnate due probabilità, valutate sulla base delle informazioni disponibili (nel nostro caso la suddivisione dei taxi in circolazione quella notte tra le due compagnie):

$$\mathbf{P}(V) = 85\% \text{ e } \mathbf{P}(B) = 15\%,$$

che sono dette *probabilità iniziali*.

L'esperimento che è stato effettuato è quello di chiedere alla testimone il colore del taxi presente sul luogo del delitto, sapendo qual era la sua capacità di riconoscere il colore giusto in termini di probabilità:

$$\mathbf{P}(TB|B) = \mathbf{P}(TV|V) = 80\% \text{ e } \mathbf{P}(TB|V) = \mathbf{P}(TV|B) = 20\%.$$

Queste probabilità sono dette *verosimiglianze*.

Il problema che si pone in un esperimento è quello di aggiornare[2] le probabilità delle ipotesi alternative (che possono essere anche più di due) alla luce dell'esito che ha avuto l'esperimento. Se torniamo alla relazione che abbiamo ricavato, possiamo interpretarla nella forma:

$$\mathbf{P}(B|TB) = \frac{\mathbf{P}(B \cap TB)}{\mathbf{P}(TB)} = \frac{\mathbf{P}(B) \cdot \mathbf{P}(TB|B)}{\mathbf{P}(B) \cdot \mathbf{P}(TB|B) + \mathbf{P}(V) \cdot \mathbf{P}(TB|V)},$$

che esprime appunto la Legge, o Teorema, di Bayes e che permette di determinare le *probabilità finali* (non in senso assoluto, ovviamente, poiché possono a loro volta rappresentare le probabilità iniziali di un esperimento successivo).

L'esame di questo semplice esempio ci permette di capire perché la Legge di Bayes sia stata interpretata come lo schema di riferimento del ragionamento *induttivo*: essa è infatti in grado di spiegare, almeno nelle situazioni più semplici, come la nostra conoscenza progredisca con l'esperienza. Con la Legge di Bayes, infatti, possiamo aggiornare progressivamente le probabilità iniziali, tenendo conto degli esiti degli esperimenti successivamente condotti.

Ma torniamo ai risultati ottenuti. È ragionevole pensare che la sorpresa, che ha accompagnato l'esito del calcolo della probabilità indicata dal problema, dipenda anche dal fatto che ci troviamo di fronte a un caso giudiziario. In modo più o meno consapevole dobbiamo esserci chiesti, immedesimandoci nella situazione che ci è stata proposta: «potremmo condannare il nostro imputato solo sulla base di una probabilità del 41%?». A parte il fatto che, anche se la probabilità fosse stata molto più alta, risponderemmo forse negativamente (si spera che ci saranno altri elementi, non come nel nostro caso, che appare come una situazione-limite, utile solo per ragionare). Però, se guardiamo all'esito dell'esperimento, abbiamo che la probabilità dell'ipotesi che il taxi fosse blu (*B*), che era inizialmente solo del 15%, sale, con la deposizione della teste, al 41%, mentre l'ipotesi che il taxi fosse verde (*V*) scende dall'80% al 59%. In un certo senso potremmo dire che l'esperimento ha *confermato* più l'ipotesi *B* che non l'ipotesi *V*, anche se la seconda resta la più probabile.

5. Il paradosso di Monty Hall

Prima di chiudere questo incontro vorrei proporvi un altro caso che Piattelli Palmarini presenta nel suo libro, definendolo addirittura come un supertunnel, per indicare quante persone abbia tratto in errore[3]. È uno dei più celebri paradossi probabilistici, il *Dilemma di Monty Hall*, che deve il suo nome a quello del conduttore di un celebre gioco a premi televisivo degli Stati Uniti, *Let's Make a Deal* (facciamo un affare), nel quale si presentava la situazione di seguito illustrata.

> DILEMMA DI MONTY HALL
>
> Nella fase finale di un concorso televisivo, il concorrente viene posto di fronte a tre porte chiuse. Dietro una di queste porte è collocata un'automobile nuova, mentre ciascuna delle altre due nasconde una capretta. Se il concorrente indovina dietro quale porta è collocata l'automobile se la porta a casa, altrimenti si porta a casa una capretta. Il conduttore del gioco, naturalmente, conosce in anticipo quale porta nasconde l'automobile. La fase finale si svolge in due tempi:
> nel primo, il concorrente indica la parta scelta. Dopo che il concorrente ha indicato la porta, il conduttore apre un'altra porta, una delle due rimanenti, dietro la quale si nasconde una capretta. Il concorrente si trova dunque due porte chiuse, una delle quali nasconde sicuramente l'automobile.
> A questo punto il conduttore da la possibilità al concorrente di confermare la prima scelta oppure di modificarla spostandosi sull'altra porta chiusa.
> Il problema che si pone è questo: è più conveniente confermare la prima scelta o modificarla?

Prima di ogni altra considerazione vorrei sottolineare che ci troviamo di fronte, di nuovo, a un problema di scelta: quale strategia è la migliore?

La risposta più immediata è quella che ha formulato la stragrande maggioranza delle persone a cui è stata sottoposta questa domanda: è indifferente, la probabilità che l'auto si trovi dietro una delle due porte è del 50%, e quindi non ha senso chiedersi cosa sia più conveniente.

Per prima cosa vorrei tranquillizzarvi; se questa è stata anche la vostra reazione, siete in ottima compagnia perché numerose persone, anche tra le menti più brillanti del secolo scorso, hanno pensato nello stesso modo ed hanno faticato, anche di fronte all'evidenza, a correggere la risposta iniziale[4].

La soluzione del dilemma non è proprio immediata ma penso che, con la schematizzazione che abbiamo visto, sia possibile concludere in modo naturale.

Indichiamo le tre porte con A, B e C e indichiamo con *A*, *B*, *C* , rispettivamente, gli eventi che corrispondono al fatto che l'automobile è dietro una di queste porte. Prima di iniziare il gioco, la situazione si presenta così:

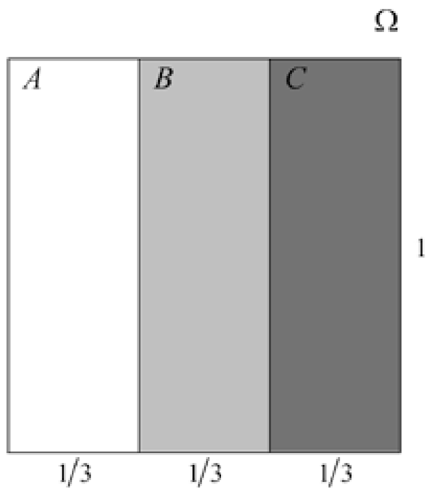

Figura 9

Anche riguardo a quale porta verrà aperta dopo la prima scelta abbiamo una situazione di indifferenza. Se consideriamo gli eventi *OA = Il conduttore apre la porta* A, ecc., inizialmente avremo la situazione rappresentata nella seguente figura.

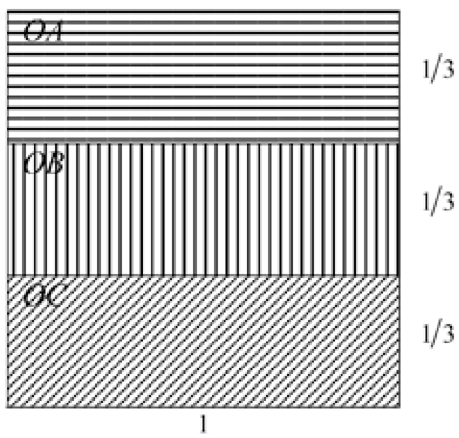

Figura 10

Ma ora immaginiamo che il concorrente scelga, ad esempio, la porta B. Che cosa succede? In termini di probabilità degli eventi *A*, *B*, *C* ancora nulla, nel senso che le probabilità che il premio si trovi dietro una delle tre porte, non vengono modificate dalla scelta del concorrente.

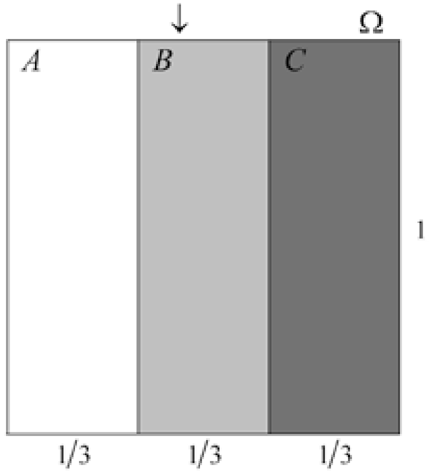

Figura 11

Sappiamo che il conduttore potrà a questo punto aprire solo una delle porte A o C. Se l'automobile si trova dietro la porta A il conduttore è costretto ad aprire la porta C e dunque si verifica l'evento *OC*. Analogamente se l'auto è dietro la porta C, si verifica l'evento *OA*. Ma se l'auto invece si trova proprio dietro la porta B, quella scelta dal concorrente, il conduttore potrà decidere di aprire o A o C. Come deciderà? Ammettiamo per un momento che scelga a caso, magari lanciando una moneta (non truccata!). In questo caso la probabilità che apra una delle due porte è del 50%. Su questa ipotesi torneremo dopo. Ora la situazione è quella rappresentata nella figura 12.

Dall'esame di questa figura si osserva che l'evento *A* è incluso in *OC,* perché, come abbiamo appena osservato, nel caso in cui il premio si trovi dietro la porta A, il conduttore è costretto ad aprire la porta C. Analogamente l'evento *C* è incluso in *OA*, mentre l'evento *B* è diviso tra *OA* e *OC,* coerentemente con l'ipotesi che il conduttore scelga a caso la porta da aprire.

Se, a questo punto, il conduttore apre la porta C, cioè se si verifica l'evento *OC*, abbiamo la restrizione dello spazio campione e guardando la figura 13, la sorpresa è grande, perché capiamo subito che non è affatto indifferente confermare la scelta iniziale o cambiarla!

Come si vede, infatti, gli unici due esiti possibili sono ora *A* e *B* ma, nella contrazione, mentre *A* è rimasto inalterato, *B* si è dimezzato e quindi:

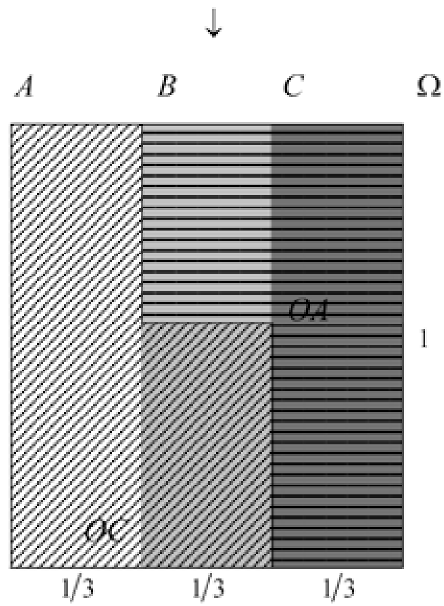

Figura 12

$$\mathbf{P}(A|OC) = \frac{\mathbf{P}(A \cap OC)}{\mathbf{P}(OC)} = \frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{6}} = \frac{2}{3},$$

mentre:

$$\mathbf{P}(B|OC) = \frac{\mathbf{P}(B \cap OC)}{\mathbf{P}(OC)} = \frac{\frac{1}{6}}{\frac{1}{3}+\frac{1}{6}} = \frac{1}{3}.$$

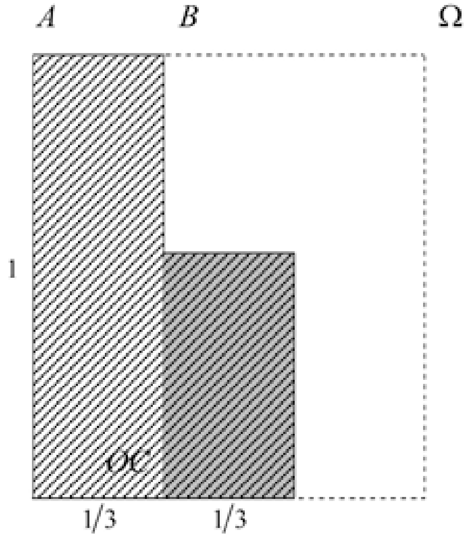

Figura 13

Questo ragionamento dimostra quindi che è più conveniente modificare la nostra scelta iniziale che non confermarla, contrariamente a quanto avevamo pensato inizialmente.

Ma torniamo a esaminare più da vicino la soluzione che abbiamo appena trovato. Essa si fonda sull'ipotesi che il conduttore, nello scegliere la porta da aprire tra la A e la C, nel caso che l'automobile si trovi dietro la porta B indicata dal concorrente, scelga a caso, lanciando, per così dire, una moneta. Si tratta di un'ipotesi ragionevole ma non deducibile da altro che da considerazioni di simmetria: che motivo avrebbe il conduttore per preferire una porta rispetto all'altra?

Se però immaginiamo che per qualche motivo il conduttore abbia dei motivi per privilegiare una porta rispetto all'altra, le cose possono cambiare. Se ad esempio il conduttore si è affezionato a una delle due caprette aprirà senz'altro la porta dietro la quale si trova la sua preferita, in modo da impedire che il concorrente possa portarsela a casa.

Più in generale supponiamo che nello scegliere quale porta aprire tra la A e la C, il conduttore si comporti in modo che la probabilità di aprire A sia p, e la probabilità di aprire C sia q, dove p e q sono numeri compresi tra 0 e 1 (estremi compresi), tali che $p + q = 1$, e non necessariamente uguali dunque. Provate da soli a ragionare su questo caso più generale.

Come dobbiamo modificare la figura 13 per tener conto di questa nuova, più generale, ipotesi? Se riflettiamo bene le probabilità p e q che abbiamo indicato sono, in effetti due probabilità condizionate:

$$\mathbf{P}(OA|B) = p \text{ e } \mathbf{P}(OC|B) = q\,.$$

Dunque le modifiche che dobbiamo introdurre riguardano il rettangolo che rappresenta B: in esso la parte di OA dovrà essere un rettangolo di altezza p, e la parte di OC, il rettangolo complementare di altezza q.

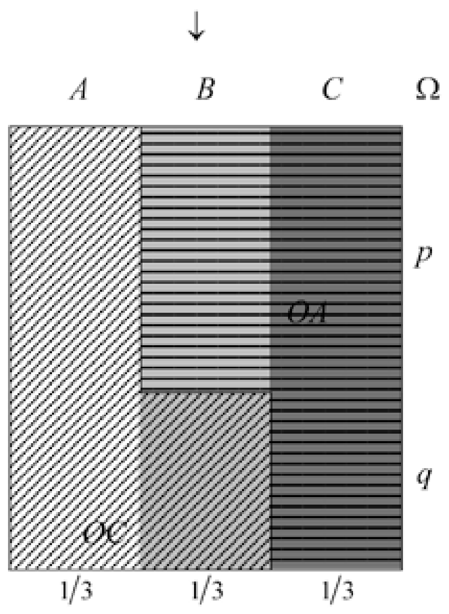

Figura 14

Se quindi calcolate le probabilità, in termini di rapporti di aree, troverete:

$$\mathbf{P}(A|OC) = \frac{\mathbf{P}(A \cap OC)}{\mathbf{P}(OC)} = \frac{\frac{1}{3} \cdot q}{\frac{1}{3} \cdot q + \frac{1}{3} \cdot 1} = \frac{q}{q+1},$$

$$\mathbf{P}(B|OC) = \frac{\mathbf{P}(B \cap OC)}{\mathbf{P}(OC)} = \frac{1}{q+1}.$$

Queste due probabilità non sono quindi, in generale, le stesse del caso iniziale, quando valeva $p = q = \frac{1}{2}$.

Per concludere, chiediamoci allora cosa succederebbe se, ad esempio, il nostro conduttore volesse impedire al concorrente di portarsi via la capretta che sta dietro la porta C. In questo caso egli aprirà certamente questa porta, che è come dire che $\mathbf{P}(OC|B) = q = 1$ e quindi che $\mathbf{P}(OA|B) = p = 0$. Se sostituiamo questi valori nelle espressioni precedenti ricaviamo:

$$\mathbf{P}(A|OC) = \frac{\mathbf{P}(A \cap OC)}{\mathbf{P}(OC)} = \frac{1}{1+1} = \frac{1}{2}$$

$$\mathbf{P}(B|OC) = \frac{\mathbf{P}(B \cap OC)}{\mathbf{P}(OC)} = \frac{1}{1+1} = \frac{1}{2}.$$

Dunque, l'unico caso in cui è davvero indifferente confermare la scelta iniziale o modificarla è proprio quello di essere di fronte a un conduttore che invece di comportarsi in modo aleatorio, scelga in modo deterministico quale porta aprire. Ma questo, generalmente, il concorrente non può saperlo.

NOTE

[1] Tra gli studiosi che hanno dato contributi decisivi a questi studi compare anche lo psicologo israeliano Daniel Kahneman che nel 2002 ha ricevuto il premio Nobel per l'economia proprio per le sue ricerche sulla teoria delle decisioni in condizioni d'incertezza.

[2] *Aggiornare* e non *correggere,* come si legge in alcuni manuali, poiché le probabilità iniziali non erano 'errate', bensì adeguate allo stato di informazione iniziale.

[3] Nel suo libro, Massimo Piattelli Palmarini racconta di aver sottoposto il problema, nell'aprile del 1991 a Trieste, ad un consesso di «illustrissimi fisici», e che tutti siano cascati nel «tranello».

[4] Paul Erdös, uno dei più geniali matematici del novecento, alla soluzione del dilemma illustratagli da un amico, reagì malissimo. Dopo aver dichiarato «No, è impossibile. Non può fare differenza», rimase solo per un'ora e al termine di questo tempo tornò dall'amico «irritatissimo», gridando «Non mi dici *perché* dovrei cambiare porta. Che cosa significa questo modo di fare?». Erdös si convinse della soluzione solo dopo che una simulazione col calcolatore gli confermò che la strategia del cambiare risulta vincente 2 contro 1. Vedi Hoffman, P., *Problemi di capre*, in [7].

BIBLIOGRAFIA

[1] Baclawski, K., Cerasoli, M., Rota, G. C., *Introduzione alla probabilità*, U.M.I., Bologna 1984.

[2] Castelnuovo, E., *Pentole, ombre, formiche*, La Nuova Italia, Firenze 1993.

[3] Costantini, D., Monari, P. (a cura di), *Probabilità e giochi d'azzardo*, Muzzio, Padova 1996.

[4] Costantini, D., *I fondamenti storico-filosofici delle discipline statistico-probabilistiche*, Bollati Boringhieri, Torino 2004.

[5] De Finett, B., *Il "saper vedere" in matematica*, Loescher, Torino 1967.

[6] De Finetti, B., *Teoria delle probabilità*, Einaudi, Torino 1970.

[7] Hoffmann, P., *L'uomo che amava solo i numeri*, Mondadori, Milano 1999.

[8] Piattelli Palmarini, M., *L'illusione di sapere*, Mondadori, Milano 1993.

[9] Piattelli Palmarini, M., *Psicologia ed economia delle scelte*, Codice, Torino 2005.

[10] Ruelle, D., *Caso e caos*, Bollati Boringhieri, Torino 1992.