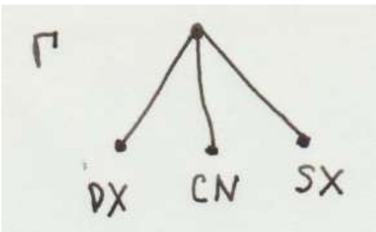
PARADOSSO DI MONTY HALL, VARIAZIONI SUL TEMA

Esaminiamo la seguente **situazione** Σ, ossia un **evento** (che deve avvenire, o è avvenuto, o è solo immaginato, in ogni caso però in un punto ben preciso dello **spazio-tempo**) + una sua descrizione attraverso un grafo che rappresenta le diverse **modalità** con le quali l'evento può verificarsi (o essersi verificato).

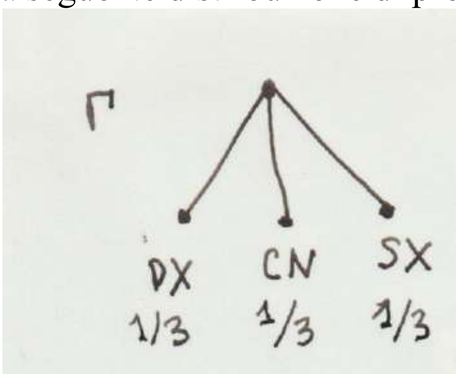
Una persona, che chiamerò il conduttore, C, invita un giocatore, G, al seguente gioco.

C sceglie da un mazzo di carte francesi diciamo l'asso di cuori + il 2 di fiori e il 2 di picche. Le mescola un poco, poi le pone davanti a sé coperte, una alla sua destra, DX, una al centro, CN, una alla sua sinistra, SX. Sottolineiamo il fatto che nemmeno C conosce la disposizione iniziale delle 3 carte.

C invita G a scommettere su dove si trova l'asso, siamo al tempo t_1. G lo fa, e diciamo che scommette che l'asso di trovi al centro. Siamo fin qui all'interno del semplice classico gioco delle 3 carte, ossia ad una situazione che è descritta dal seguente grafo Γ:



al quale viene associata la seguente distribuzione di probabilità (a priori):

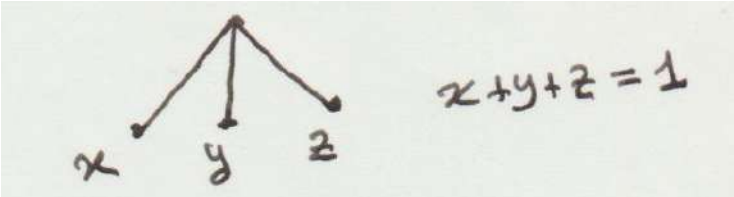


(parleremo allora di un **grafo probabilistico**; questo particolare è una tricotomia **semplice**, corrispondente cioè all'**incertezza assoluta**):

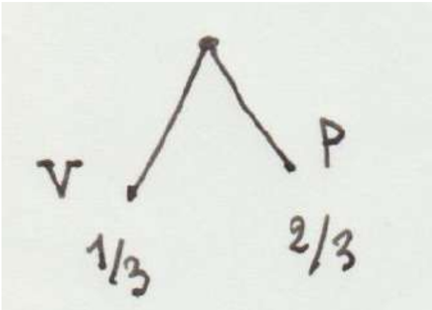
Ricordiamo che una distribuzione di probabilità su grafi di questo genere è una funzione che va da Γ all'insieme dei numeri razionali compresi tra 0 e 1. Con l'usuale simbolismo matematico:

$$p : \Gamma \to [0,1]_{\mathbb{Q}}$$
,

una funzione che rappresenteremo in forma grafica con:



Poiché in seguito dovremo forzatamente considerare anche tale circostanza, aggiungiamo che nella definizione del gioco è necessario comprendere anche l'associato grafo probabilistico **vinci/perdi** (una dicotomia nel presente caso però non semplice):



Ritornando sulla (banale) distribuzione di probabilità associata a quelli che definiamo gli **stati dell'informazione** di C e di G rispetto alla situazione Σ a questo tempo iniziale t_1 - diciamoli rispettivamente $I(C,\Sigma,t_1)$, $I(G,\Sigma,t_1)$ - potremo scrivere per esempio, in maniera pesante ma molto chiara:

$p(DX,C,I(C,\Sigma,t_1),t_1) = 1/3$ *etc.*

(probabilità della modalità "asso a destra" dal punto di vista di C all'istante t_1 dato uno stato iniziale dell'informazione che è uguale evidentemente al vuoto - fatta salva ovviamente la descrizione stessa della situazione).

Ricordiamo adesso che abbiamo definito la **quantità d'informazione** di una distribuzione di probabilità p(t) su una 3-tomia mediante la seguente naturale formula:

$$QI(p) = QI = \frac{3}{2}\left[\left(x-\frac{1}{3}\right)^2 + \left(y-\frac{1}{3}\right)^2 + \left(z-\frac{1}{3}\right)^2\right]$$

(qui 3/2 vale n/(n-1), che è il coefficiente previsto nel caso di una n-tomia, vedi:
http://www.cartesio-episteme.net/probabilita-finale.pdf +
http://www.cartesio-episteme.net/probabilita-appendice.pdf),
la quale implica che la quantità d'informazione inerente ad una singola modalità μ della situazione in esame è uguale a:

$QI(p|\mu) = \frac{n}{n-1}[p(\mu)-\frac{1}{n}]^2$

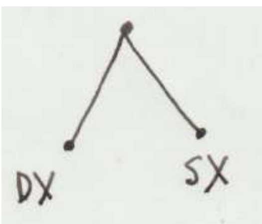
un'espressione che ammette un massimo uguale a (n-1)/n per p(μ)=1 (2/3 è quindi proprio il massimo possibile per QI(p) nel caso n=3), ed un minimo uguale a 0 per p(μ)=1/n. Per una certezza assoluta p, per esempio (1,0,...,0), si ha: QI(p)=[(n-1)/n]+(n-1)*[1/n(n-1)]= [(n-1)/n]+1/n=1, mentre per l'incertezza assoluta si ha evidentemente QI(p)=0. Si noti esplicitamente che p(μ)=0 non implica affatto QI(p|μ)=0, bensì soltanto QI(p|μ)=1/n(n-1), ossia un valore che dipende da n ed è sempre più vicino allo zero al crescere di n ma non è mai 0, mentre, come abbiamo detto, risulta QI(p|μ)=(n-1)/n per p(μ)=1, cioè un valore che tende ad 1 al crescere di n, . Tradotto, è più facile passare (ovvero, basta una piccola quantità di conoscenza) da 1/n a 0, che non da 1/n a 1.

Nota 1 - Per inciso, si tratta di una definizione della QI che è diversissima dal $\log_2(p(\mu)^{-1})$ di Shannon, la quale ha tutt'altra interpretazione semantica. La nostra ha un significato evidente e naturale, ossia il quadrato normalizzato con un certo fattore della distanza pitagorica che intercorre tra il punto (x,y,z) che rappresenta p nello spazio R^3 dal punto (1/3,1/3,1/3) che corrisponde all'incertezza assoluta, per cui appunto QI(p)=0 ses p è l'incertezza assoluta, e QI(p)=1 ses p è una delle tre certezze assolute, (1,0,0), (0,1,0), (0,0,1) - questo risultato è ovvia conseguenza della scelta del fattore di normalizzazione. Da tutto ciò deriva per esempio che la QI iniziale di C e di G rispetto alla data situazione è uguale 0.

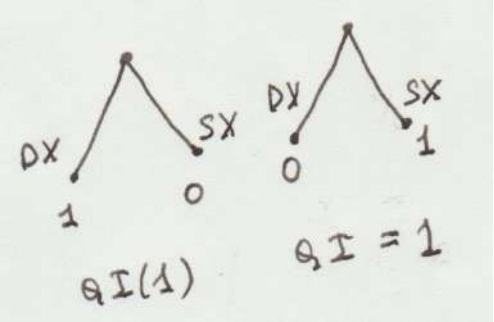
Bene, qui giunti, diciamo al tempo t_2, C potrebbe semplicemente guardare la carta scelta dal giocatore (ricordiamo, abbiamo supposto quella centrale), comunicare se G ha vinto o ha perso e il gioco finirebbe qui. Ma a questo punto invece - ed a questo punto soltanto, ribadiamo - C prende in mano qualche carta e la guarda, ma **non** guarda la carta centrale, che si limita a spostare alla propria destra, mentre guarda le due carte rimanenti, cioè quella alla sua sinistra e quella alla sua destra. La QI di C passa adesso al 100%, lui sa bene se G ha vinto o ha perso, ma tiene per sé il risultato di tale notizia, e procede in una **nuova** direzione.

Prende una delle due scartine che ha in mano e la mostra a G, diciamo il 2 di fiori, indi pone la rimanente carta che ha in mano alla propria sinistra, e chiede a G: dove sta l'asso, confermi la tua scelta iniziale (che corrisponde adesso alla carta DX) o preferisci cambiarla (scegliendo quindi la carta SX)?

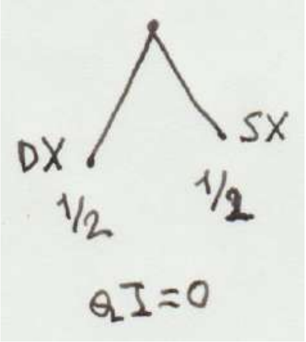
E' chiaro che si tratta di una situazione interamente **differente** da quella precedente, diciamola Σ'. Essa è rappresentata da una dicotomia, e non da una tricotomia:

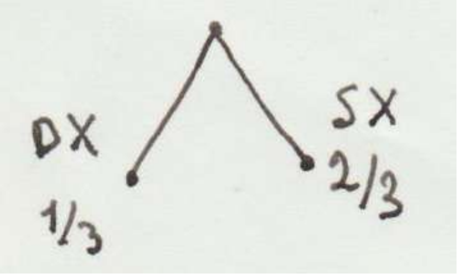


per la quale C ha una distribuzione di probabilità associata corrispondente ad uno dei due seguenti grafi probabilistici:



mentre per quanto riguarda G, ed il suo stato dell'informazione $I(C,\Sigma,t_2)$, si è discusso, e si continua a discutere, se la sua scelta più razionale sia il grafo probabilistico riportato qui sotto a sinistra oppure quello a destra:



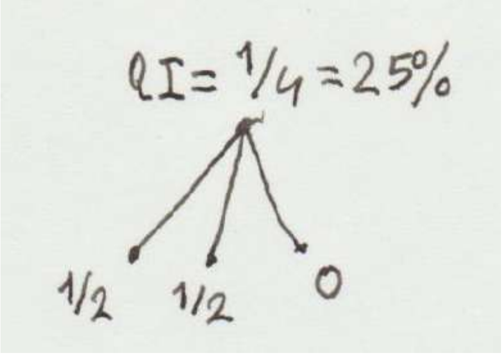


Nel primo caso, come indicato in figura, $QI(I(G,\Sigma,t_2)) = 0$, mentre nel secondo:

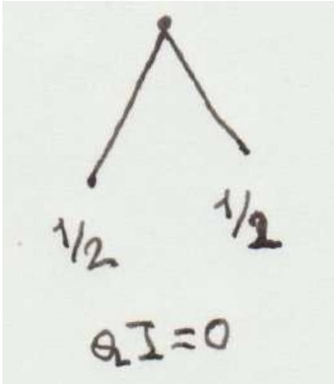
$$QI(I(G,\Sigma,t_2)) = \frac{3}{2}\left[\left(\frac{1}{3}-\frac{1}{2}\right)^2+\left(\frac{2}{3}-\frac{1}{2}\right)^2\right] = 1/9 \approx 11\%$$

(e ribadiamo che invece $QI(I(C,\Sigma,t_2)) = 1$).

Nota 2 - Il passaggio da una tricotomia ad una dicotomia non è ininfluente rispetto al calcolo di una QI. Vale a dire, una cosa è la QI della seguente tricotomia:

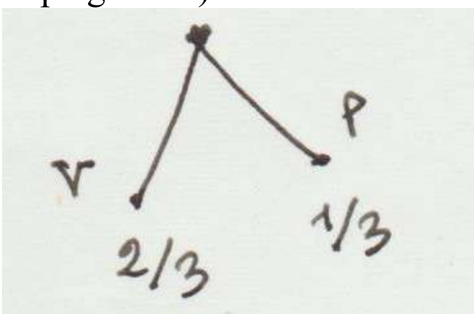


che non è 0, ed un'altra quella della "corrispondente" dicotomia:



Questa è una delle principali differenze tra l'usuale presentazione del paradosso e quella che abbiamo invece qui proposto.

Non abbiamo dubbi (almeno oggi!) che si tratti di quello a destra, e che il **nuovo** gioco proposto da C sia completato dal seguente grafo probabilistico vinci/perdi (si tratta in ultima analisi del gioco **complementare**, o **opposto**, rispetto al precedente, si vince quando prima si perdeva e viceversa, nel caso in esame si vince se si indica una scartina e non l'asso, si spera che il concetto sia chiaro senza bisogno di ulteriori spiegazioni):



Insomma, valgono le seguenti identità:

$p(DX,C,I(C,\Sigma',t_2),t_2) = p(SX,C,I(C,\Sigma',t_2),t_2) = 0$ oppure 1

$p(DX,G,I(G,\Sigma',t_2),t_2) = 1/3$

$p(SX,G,I(G,\Sigma',t_2),t_2) = 2/3$.

Su tale aspetto della discussione non intendiamo tornare, dopo quanto già detto qui (naturalmente, *mutatis mutandis*, ossia, tenuto conto del nuovo assetto del "gioco" come dianzi illustrato):

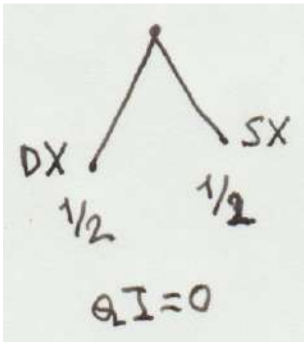
http://www.cartesio-episteme.net/monty-hall.pdf

(lo scritto, che andrebbe invero completamente rifatto, contiene anche un'appendice non banalissima concernente il concetto di grafo "inverso" di certi grafi probabilistici e la formula di Bayes).

Insomma, il famoso "paradosso" di Monty Hall consiste nella raggiunta consapevolezza che cambiare la scelta iniziale offre a G maggiori possibilità di vittoria, e si è affermato che nella **realtà**, effettuate per esempio 66 simulazioni del gioco al computer, G ne vincerebbe 44 (ovviamente circa) cambiando la scelta iniziale, e soltanto 22 mantenendola.

Bene, *so far so good*, usava sempre intercalare il mio docente di Teoria dei Numeri a Cambridge (Henry Peter Francis Swinnerton-Dyer, 1927-2018, mi piace adesso ricordarlo, e ricordare quegli anni di studio, gli ultimi tranquilli della mia complicata vicenda terrena), ma adesso può intervenire un'ulteriore interessante variazione sul tema suggeritami da uno dei mie corrispondenti, che si riconoscerà e che ringrazio.

Prima di chiudere il gioco, annunciando se G ha infine vinto o perso, C fa invece intervenire un **secondo giocatore**, diciamolo G', tenuto finora accuratamente all'oscuro di tutto, gli mostra le due carte davanti a sé, DX e SX, e gli dice di indovinare dove si trovi l'asso. Siamo adesso ad un tempo t_3, e direi che non c'è dubbio che la nuova (la terza!) situazione Σ'' sia descritta per G' dal seguente grafo probabilistico (una dicotomia semplice), che abbiamo già avuto modo di vedere in un differente contesto:



(con associato identico grafo probabilistico vinci/perdi).

Per riassumere (e supponendo che G sia stato viceversa tenuto al corrente di tutto), ci troviamo in definitiva cioè al tempo t_3) con le seguenti identità relative alle probabilità concernenti una "medesima" modalità di quello che erroneamente viene considerato un unico evento (o un'unica situazione), pertanto solo apparentemente contraddittorie se non si usa un simbolismo adeguato (se si usa invece un simbolismo adeguato, non c'è nessun dubbio al riguardo):

$p(DX,C,I(C,\Sigma'',t_3),t_3) = 0$ oppure 1

$p(DX,G,I(G,\Sigma'',t_3),t_3) = 1/3$

$p(DX,G',I(G',\Sigma'',t_3),t_3) = 1/2$

+ identità analoghe per SX.

Potremmo ritenere di avere così concluso la nostra analisi delle varie situazioni, ma ancora il detto corrispondente solleva una naturale perplessità. Vero che si tratta di 3 diverse probabilità tra loro comunque compatibili, ma se si passa come abbiamo accennato poc'anzi **dalla teoria alla realtà**, cercando così in qualche modo di **oggettivizzare** la definizione delle probabilità in ballo, come è possibile che G vinca 2 volte su 3, e G' soltanto 1 volta su 2?

Bene, si risponde facilmente all'obiezione notando che, sulle 66 simulazioni di cui sopra, G ne vince 44 cambiando la scelta iniziale, mentre G' ne vince 33 scegliendo a caso tra DX e SX, precisamente vince la metà delle 22 che G perde, ossia 11, e la metà delle 44 che G vince, ossia 22, e 11+22=33, come volevasi dimostrare.

Ma è proprio tutto concluso, oppure c'è ancora qualcosa interessante che si puo' approfondire?

Analizziamo infatti meglio il detto passaggio dalla teoria alla realtà, supponendo di collocarci ad un tempo t_n delle 66 simulazioni che vengono effettuate.

E' chiaro che si potrebbe pensare, con ovvio simbolismo, che sussistano sempre le seguenti identità:

$p(DX,C,I(C,\Sigma^n,t_n),t_n) = 0$ oppure 1

$p(DX,G,I(G, \Sigma^n,t_n),t_n) = 1/3$

$p(DX,G',I(G', \Sigma^n,t_n),t_n) = 1/2$

+ identità analoghe per SX,

o almeno questa è stata la nostra supposizione, **ma è proprio corrispondente al vero**? Non si può invece pensare che G', scegliendo a caso come abbiamo supposto tra DX e SX, si accorga **ad un certo punto** che vince con maggiore frequenza quando sceglie SX, rispetto a quando sceglie DX? In altre parole, non può invece ritenersi che - ad un certo istante t_n, ribadiamo - G' sia razionalmente legittimato a passare dalla precedente:

$p(DX,G',I(G', \Sigma^n,t_n),t_n) = 1/2$
(+ analoga per SX)

alla:

$p(DX,G',I(G', \Sigma^n,t_n),t_n) = 1/3$

+ corrispondente per SX:

$p(SX,G',I(G', \Sigma^n,t_n),t_n) = 2/3$?

(val forse la pena di notare che, trovandoci qui in una situazione "equivalente" al lancio di una moneta, G' potrebbe accorgersi prima o poi che si tratta di una moneta "truccata").

Detto che debbo codesta osservazione a mio figlio G., viene certo subito in mente che l'obiezione in parola potrebbe essere superata ipotizzando che C ponga una volta la carta vincente a DX e una volta a SX, ossia "a caso" (probabilità 1/2). Comunque, cosa ne pensano i miei gentili lettori? Siamo davvero di fronte ad un caso - almeno secondo certe modalità organizzative del "gioco" - in cui i risultati **passati** influenzano quelli **futuri**??

Sperando di non aver commesso errori, né di scrittura né di concetto, e che questa ultima fatica della mia vita scientifica possa essere utile almeno per mostrare anche a non professionisti matematici quello che rimane l'unico modo con cui deve essere precisamente discusso ogni siffatto argomento (si tratta di una matematica che definire liceale è forse persino troppo), ringrazio per l'attenzione che sarà prestata al mio scritto, e saluto tutti cordialmente (dato il periodo, anche con tanti AUGURI pasquali),

UB - Pg, 27.III.2024