Probabilità e giochi
Laboratorio PLS

F. Morandin e A. Saracco

15 dicembre 2017

1 Piano generale

Si suggerisce di seguire il concetto di laboratorio, presentando di volta in volta piccoli passi avanti, secondo un percorso storico o comunque logico.

2 Black Jack

Il Black Jack è un gioco d'azzardo di carte che vede il banco sfidato da un numero variabile di giocatori. Vincono i giocatori che realizzano un punteggio più alto del banco e non superiore a 21.

> **A:** [1]Si chiede agli studenti di cercare[2] le regole precise su internet e scriverle in modo ordinato.

Usando la matematica e in particolare la probabilità, vorremmo analizzare il gioco in modo da rispondere agli interrogativi che dovrebbe porsi un giocatore.

> **A:** Gli studenti a gruppi provano a giocare un paio di mani (non più di 5 minuti). Idealmente in ogni gruppo uno fa il banco e gli altri i giocatori. Le puntate sono virtuali, si può far finta che ogni giocatore abbia 10 fiches

> **Q:** [3]*Come giocatori, a quali domande vorreste risposta?*

Dovrebbe emergere la richiesta di sapere quando fermarsi. Anzi, meglio di sapere quando fermarsi in funzione della carta scoperta del banco.

La domanda vera è quindi di determinare la *policy* di gioco più conveniente. Per fare ciò un prerequisito indispensabile è calcolare la probabilità

[1]Con questa notazione indichiamo una attività da proporre agli studenti.

[2]Se il laboratorio si svolge in aula computer, si può fare sul momento, altrimenti si dovrà chiedere agli studenti di farlo in anticipo.

[3]Con questa notazione indichiamo una domanda da porre agli studenti.

di vittoria a seconda del punteggio a cui si ferma il giocatore, che è quello che ci proponiamo di fare, almeno in parte.

Infatti le regole del gioco sono complesse e dettagliate, e in certi aspetti ammettono diverse varianti. Per analizzare il gioco dal punto di vista matematico, è bene prima semplificarlo adeguatamente. (Si può sempre poi realizzare un modello più preciso in seguito.)

> **Q:** *Quali regole eliminiamo o semplifichiamo?*

Dopo la ricerca delle regole ufficiali su internet e la prova pratica di gioco, dovrebbe emergere una proposta vicina alla seguente.

- Estrazioni delle carte indipendenti. In realtà non lo sono, perché se ad esempio si gioca con un mazzo solo e i 4 assi sono già usciti, è impossibile fare blackjack. Se il numero di mazzi è però elevato, le estrazioni sono quasi indipendenti. (E i giocatori che contano le carte sono svantaggiati.)
- Niente *split*, *double*, *surrender*, *insurance* e *side bet*.
- Un singolo giocatore contro il banco.
- Il blackjack vale come 21 fatto in qualsiasi altro modo e non di più.

Anche proposte diverse sono accettabili, tanto in realtà dovremo inizialmente semplificare molto di più, e alla fine la complessità da scegliere e gestire sarà responsabilità degli studenti.

2.1 Un primo calcolo delle probabilità

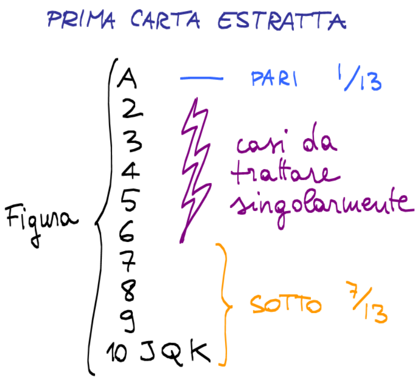
Semplificato adeguatamente il gioco, si propone agli studenti di cimentarsi in un primo calcolo.

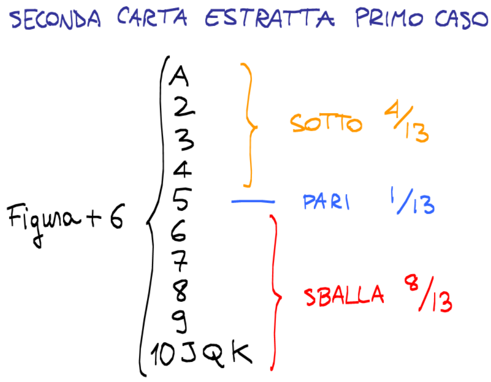
> **A:** Tentiamo di calcolare una prima probabilità in un caso molto semplice: il banco ha una figura da 10 punti, il giocatore ha girato carte fino ad ottenere 21. Tocca al banco. Qual è la probabilità che il banco pareggi, o sballi o finisca con meno di 21 punti[4]?

La domanda è apparentemente innocente, ma la soluzione è molto complessa e richiede certamente la assistenza del docente, che può guidare gli studenti.

- Alla prima carta estratta, su 13 possibilità possono succedere queste cose: se esce l'asso A il banco pareggia, se esce una carta dal 7 al re K il banco perde, se invece esce un'altra carta da 2 a 6 il banco estrae ancora.

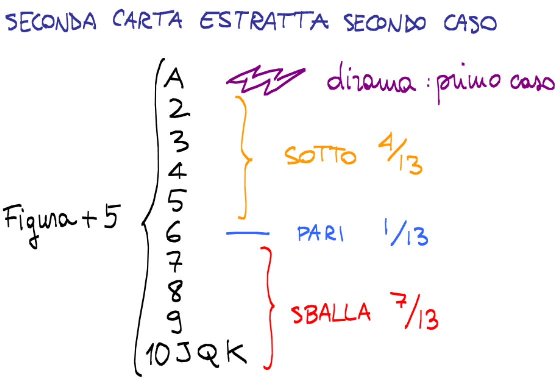
[4]Si ricorda che quando il banco arriva ad un punteggio di 17 o più deve obbligatoriamente fermarsi.

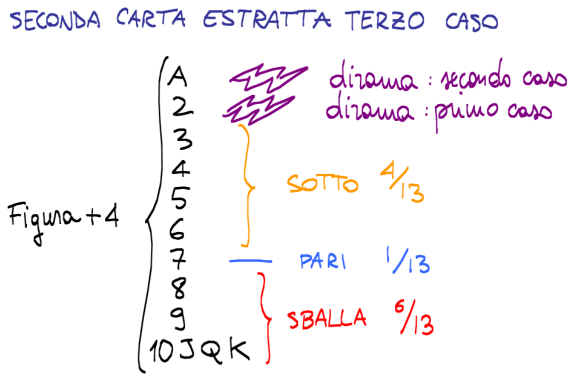




- Si aprono quindi cinque altri casi, il più semplice dei quali è che la prima estratta sia un 6. In questo caso alla seconda carta estratta, su 13 possibilità possono succedere queste cose: se esce 5 il banco pareggia, se esce una carta dall'asso A al 4 il banco perde restando sotto, se esce una carta dal 6 al re K il banco perde sballando.

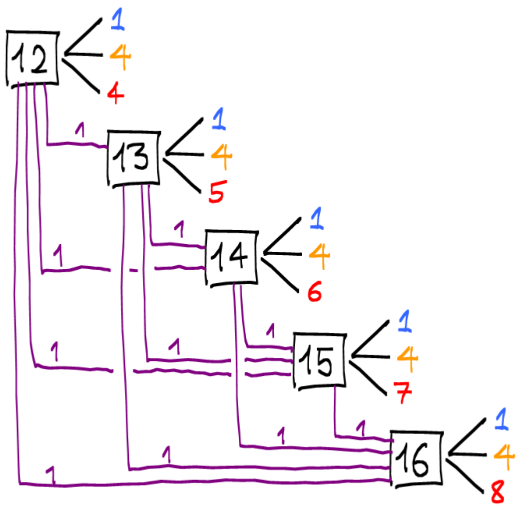
- Il secondo caso è che la prima carta estratta sia un 5. In questo caso alla seconda carta estratta, su 13 possibilità possono succedere queste cose: se esce 6 il banco pareggia, se esce una carta dal 2 al 5 il banco perde restando sotto, se esce una carta dal 7 al re K il banco perde sballando, infine, se esce l'asso A ci si ritrova nel caso precedente, visto che il banco ha 16 punti e deve girare una carta.





- Il terzo caso è che la prima carta estratta sia un 4 ed è ancora più complicato. In questo caso alla seconda carta estratta, su 13 possibilità possono succedere queste cose: se esce 7 il banco pareggia, se esce una carta dal 3 al 6 il banco perde restando sotto, se esce una carta dall'8 al re K il banco perde sballando, inoltre, se escono l'asso A o il 2 ci si ritrova nei due casi precedenti, visto che il banco ha 15 o 16 punti e deve girare una carta.

Non dettaglio nemmeno gli altri due casi, che sono dello stesso tipo ma un po' più complicati. Lo schema generale dell'albero delle possibilità è riportato qui sotto. I numeri interi che compaiono fuori dai riquadri vanno interpretati come 13-esimi di probabilità.



Mettere insieme tutto è un lavoro di pazienza e di ordine, non concettuale, inoltre potrebbe essere noioso per gli studenti, ecco perché non mi ci soffermerei. Se gli studenti non desiderano fortemente fare il calcolo, lascerei qui questa situazione e passerei alla sezione successiva.

Se invece gli studenti sono curiosi e soprattutto a condizione che sia possibile accedere ad un foglio di calcolo, il conto può essere fatto come segue.

	A	B	C	D	E	F	G
1					pari	sotto	sballa
2					11,1%	67,6%	21,2%
3					537824	3265175	1023810
4							
5	**10**	1	7	0	1,44852	8,79407	2,75742
6	**12**	1	4	4	1,34505	5,3802	6,27475
7	**13**	1	4	5	1,24898	4,9959	6,75512
8	**14**	1	4	6	1,15976	4,63905	7,20118
9	**15**	1	4	7	1,07692	4,30769	7,61538
10	**16**	1	4	8	1	4	8
11							

Le celle delle colonne B, C e D contengono gli stessi numeri della figura precedente, oltre a quelli della prima figura per la riga del **10**. La cella E5 contiene la formula

```
=B5+SOMMA(E6:E10)/13
```

A questo punto basta copiare e incollare la cella E5 nelle celle E5:G10 e i calcoli sono conclusi. Le probabilità finali sono quelle che compaiono nella riga 2, calcolate semplicemente tramite `E2=E5/13` e analoghi.

Volendo con un po' di sforzo è possibile esprimere i risultati ottenuti come frazioni: i numeratori sono i numeri interi che compaiono nella riga 3 e il denominatore è 13^6.

2.2 Il gioco del 6

L'attività precedente dovrebbe avere chiarito che anche semplificando le regole, la combinatoria coinvolta nel Black Jack è troppo complicata per essere divertente, tuttavia il calcolo è fattibile ed istruttivo. Adesso semplifichiamo ulteriormente e drasticamente il gioco, per mostrare come applicare della combinatoria più raffinata.

Il gioco del 6 si fa con un normale dado a 6 facce, invece che con le carte. Il punteggio massimo, equivalente al 21 del Black Jack è 6. Prima tira il giocatore, fermandosi quando vuole, poi tira il banco che finché il suo punteggio è strettamente inferiore a quello del giocatore non si ferma in nessun caso. Il banco vince se raggiunge o supera il giocatore senza superare il 6.

> **A:** Tiro ripetutamente il dado sommando sempre i risultati. Qual è la probabilità che ad un certo punto il totale sia esattamente 6?

Si può dividere in casi a seconda del numero di lanci necessari, da 1 a 6. Il numero di modi di ottenere 6 sommando k numeri interi positivi è per un risultato classico $\binom{6-1}{k-1}$. Se si sommano k lanci di dado il risultato è lo stesso, perciò la probabilità cercata p_6 vale

$$p_6 = \binom{5}{0}\left(\frac{1}{6}\right)^1 + \binom{5}{1}\left(\frac{1}{6}\right)^2 + \binom{5}{2}\left(\frac{1}{6}\right)^3 + \binom{5}{3}\left(\frac{1}{6}\right)^4 + \binom{5}{4}\left(\frac{1}{6}\right)^5 + \binom{5}{5}\left(\frac{1}{6}\right)^6$$

Se gli studenti la conoscono o se l'insegnante vuole introdurla, si può approfittare della formula del binomio di Newton:

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$

ottenendo

$$p_6 = \sum_{k=0}^{5} \binom{5}{k} \left(\frac{1}{6}\right)^{k+1} = \frac{1}{6}\left(\frac{1}{6}+1\right)^5 = \frac{7^5}{6^6} = \frac{16807}{46656}$$

Ora una domanda più facile.

> **Q:** *Se il giocatore ha fatto 6, e il banco deve ancora iniziare a tirare, qual è la probabiltà che pareggi? E che vinca il giocatore?*

La probabilità di pareggio coincide con p_6, quella dell'attività precedente, come è facile capire. La probabilità di vittoria del giocatore è il complementare e la denotiamo con $f_6 = 1 - p_6$. (f perché il giocatore si $f$erma.)

> **A:** Se il giocatore ha fatto 5, e il banco deve ancora iniziare a tirare, qual è la probabiltà che pareggi o superi il giocatore?

La probabilità cercata, che denoteremo con $p_{\geq 5}$ coincide con quella che tirando ripetutamente il dado ad un certo punto il totale sia esattamente 5 o che ad un certo punto il totale sia esattamente 6.

La probabilità p_5 di arrivare esattamente a 5 si calcola come nel caso precedente. Se nel caso precedente gli studenti non sono stati molto autosufficienti, questa può essere una buona occasione per vedere se hanno capito e sanno riprodurre il ragionamento. Se invece il concetto è già chiaro o se viceversa risultasse troppo astratto per gli studenti, converrà non soffermarvisi e dare direttamente il risultato senza far perdere tempo ed energie su questo aspetto e invece passando al punto successivo. Il calcolo comunque è questo,

$$\begin{aligned} p_5 &= \binom{4}{0}\left(\frac{1}{6}\right)^1 + \binom{4}{1}\left(\frac{1}{6}\right)^2 + \binom{4}{2}\left(\frac{1}{6}\right)^3 + \binom{4}{3}\left(\frac{1}{6}\right)^4 + \binom{4}{4}\left(\frac{1}{6}\right)^5 \\ &= \frac{1}{6}\left(1+\frac{1}{6}\right)^4 = \frac{7^4}{6^5} = \frac{2401}{7776} \end{aligned}$$

È invece molto più interessante fare ragionare gli studenti su come si possano mettere assieme p_5 e p_6 per ottenere $p_{\geq 5}$. Non è cosa facile, quindi una piccola guida può aiutare.

Suggerisco questa analogia. Si tirano due dadi. Si chiede la probabilità che il primo dia 3 o che il secondo dia 4. Non dovrebbe essere difficile ragionare per enumerazione di casi o con i diagrammi di Venn e capire che la probabilità dell'unione (che accada uno almeno di due eventi) è la somma delle probabilità meno la probabilità dell'intersezione.

Ciò fatto, per calcolare $p_{\geq 5}$ serve la probabilità dell'evento intersezione. Si dovrebbe chiedere agli studenti di identificare questo evento. La risposta è:

che tirando ripetutamente il dado ad un certo punto il totale sia esattamente 5 e successivamente esattamente 6. Facile a questo punto capire che

$$p_{\geq 5} = p_5 + p_6 - \frac{1}{6}p_5 = p_6 + \frac{5}{6}p_5$$

> **A:** Se il giocatore è arrivato a 5 deve decidere se tirare di nuovo con il rischio di sballare o se fermarsi, dando al banco una maggior probabilità di vincere. Quali sono le sue probabilità di vittoria nei due casi?

Ragionando brevemente gli studenti dovrebbero arrivare a capire che la probabilità di vittoria se si $f$erma è $f_5 = 1 - p_{\geq 5}$, mentre se decide di $t$irare è $t_5 = \frac{1}{6}(1 - p_6)$.

Se si sostituiscono i numeri si trova che $f_5 \approx 0.3825$ mentre $t_5 \approx 0.1066$, perciò conviene fermarsi. Denotiamo con d_5 la probabilità di vittoria se il giocatore la la $d$ecisione ottima (in questo caso si ferma), ovvero $d_5 = \max(f_5, t_5)$.

> **A:** Come si deve estendere il ragionamento precedente per decidere qual è il punteggio minimo a cui conviene fermarsi?

Discutendo ampiamente gli studenti dovrebbero concludere che occorre innanzitutto calcolare anche $p_{\geq 4}$, $p_{\geq 3}$ e $p_{\geq 2}$. (In teoria anche $p_{\geq 1}$, ma è semplicemente 1.) L'insegnante fornirà i valori di queste probabilità senza farli calcolare agli studenti.

n	p_n	$p_{\geq n}$	f_n	t_n	d_n
6	0.3602	0.3602	0.6398	0.0000	0.6398
5	0.3088	0.6175	0.3825	0.1066	0.3825
4	0.2647	0.7940	?	?	?
3	0.2269	0.9074	?	?	?
2	0.1945	0.9722	?	?	?
1	0.1667	1.0000	0.000	?	?

Poi gli studenti dovrebbero (magari aiutati) arrivare alle formule seguenti per f_n, t_n e d_n. Dovrebbero procedere per casi, partendo da $n = 4$ e scendendo ragionando con calma.

Se si è a 4, le probabilità di vittoria fermandosi e tirando sono rispettivamente $f_4 = 1 - p_{\geq 4}$ e $t_4 = \frac{1}{6}f_6 + \frac{1}{6}f_5$. Occorre vedere quale delle due è più alta e sostituirla al posto di d_4 nelle formule sotto.

Se si è a 3, le probabilità di vittoria fermandosi e tirando sono rispettivamente $f_3 = 1 - p_{\geq 3}$ e $t_3 = \frac{1}{6}f_6 + \frac{1}{6}f_5 + \frac{1}{6}d_4$. Occorre vedere quale delle due è più alta e sostituirla al posto di d_3 nelle formule sotto.

Se si è a 2, le probabilità di vittoria fermandosi e tirando sono rispettivamente $f_2 = 1 - p_{\geq 2}$ e $t_2 = \frac{1}{6}f_6 + \frac{1}{6}f_5 + \frac{1}{6}d_4 + \frac{1}{6}d_3$.

E così via. Se il tempo stringe, si può dare la tabella finale agli studenti:

n	p_n	$p_{\geq n}$	f_n	t_n	d_n
6	0.3602	0.3602	0.6398	0.0000	0.6398
5	0.3088	0.6175	0.3825	0.1066	0.3825
4	0.2647	0.7940	0.2060	0.1704	0.2060
3	0.2269	0.9074	0.0926	0.2047	0.2047
2	0.1945	0.9722	0.0278	0.2388	0.2388
1	0.1667	1.0000	0.0000	0.2786	0.2786

Q: *Guardando la tabella completa, qual è il punteggio minimo a cui conviene che il giocatore si fermi?*

Se non riescono a rispondere a questa, non hanno capito nulla. Basta infatti per ogni n confrontare la probabilità di vittoria fermandosi f_n e la probabilità di vittoria tirando ancora t_n per capire che da 1 a 3 si tira e da 4 a 6 si sta.

Se potesse essere utile, il file Excel `gioco6.xlsx` riporta tutta la tabella e la calcola in modo semplice e sintetico.

2.3 Un possibile lavoro finale

Supponendo che si sia compresa la logica per determinare la migliore *policy* nel gioco del 6, si potrebbe pensare di realizzare un progetto analogo per una versione verosimile del Black Jack, come lavoro finale da presentare al seminario di giugno.

La difficoltà di calcolo delle probabilità coinvolte (limitatamente ai valori analoghi ai p_n del gioco del 6) andrebbe gestita con il calcolatore, o facendogli fare tutti i casi, oppure facendogli svolgere la partita molte volte stimando le probabilità dei diversi esiti con le loro frequenze.

I conti a partire da quelle probabilità sono invece molto ragionevoli e non dissimili da quelli del gioco del 6.

Se il gruppo desidera cimentarsi in un progetto simile, certamente servirà un incontro per precisare come si mettono in pratica queste strategie.

3 Lotterie, ritarti e legge dei grandi numeri

Questa unità è pensata per essere svolta in un singolo incontro di circa due ore. Le lotterie in genere non destano un grande interesse tra i giovanissimi, quindi sarebbe utile stimolarli il più possibile con domande, curiosità, esperienze personali.

Lo scopo principale è conciliare la legge dei grandi numeri (LGN) con l'indipendenza delle estrazioni, chiarendo l'inutilità di studiare i ritardi. Il tutto può essere affrontato *sperimentalmente* in laboratorio informatico costruendo un opportuno file Excel.

Q: *Qui con "lotterie" intendiamo quei giochi a premi che prevedono estrazioni periodiche e a cui si può concorrere giocando uno o piùù numeri a propria scelta, non predeterminati in un biglietto. Quali giochi di questo tipo conoscete?*

Tra le risposte prevedibili ci sono Lotto, Superenalotto e Winforlife. Per la nostra analisi ci concentreremo sulle prime due, che sono quelle a cui meglio si applicano i ritardi. Perciò se escono altre proposte, occorre "archiviarle" con poche parole e verificare che gli studenti conoscano i meccanismi base di Lotto e Superenalotto. (Basta chiarire che ogni concorso c'è una estrazione con rimessa di 5 o 6 numeri da 1 a 90.)

Q: *Se vi dico che al Superenalotto il numero 74 ritarda da 132 estrazioni, cosa ne pensate?*

Qui è utile spendere alcuni minuti in una discussione di gruppo. Gli elementi (tra loro apparentemente contraddittori) che potrebbero emergere sono ad esempio:

- Cosa significa che un numero *ritarda*.
- Siccome ci sono 90 numeri e ad ogni estrazione ne vengono scelti 6 distinti, mediamente ogni $\frac{90}{6} = 15$ estrazioni i numeri sono estratti una volta ciascuno e quindi il ritardo *tipico* che ci aspettiamo per ciascun numero (anche per il 74) è di 15 estrazioni. Quindi un ritardo di 132 estrazioni sembra eccezionale.
- L'urna da cui vengono estratte le palline con i numeri (anche se forse oggi si fa in modo diverso) non possono sapere il ritardo dei diversi numeri, quindi la probabilità che esca il 74 non dovrebbe essere diversa da quella che esca un qualsiasi altro numero, anche se appena uscito. (Si dice che *il Caso non ha memoria*.)
- Un numero non può ritardare per sepre: prima o poi viene estratto per forza. (Talvolta questa affermazione viene giustificata impropriamente invocando la LGN o altri risultati generali.)
- Molti giocatori si basno sui ritardi, e chi organizza questi giochi li pubblicizza molto.

Sarà bene fare in modo che i primi tre punti almeno siano citati a qualche livello. Gli ultimi due non importa, ma se viene fuori l'ultimo è opportuno smontarlo subito, pur senza elementi convincenti che emergeranno durante il lavoro: i giocatori che si basno sui ritardi non hanno speranze di vittoria maggiori degli altri e chi pubblicizza le informazioni sui ritardi lo fa con lo scopo di confondere, ingannare e portare a giocare in modo compulsivo[5] più persone possibili.

Q: *Qualcuno sa cos'è la legge dei grandi numeri?*

Se nessuno risponde, si lascia la cosa nel mistero e si procede. Se qualcuno risponde (anche in maniera scorretta) si lasciano discutere gli studenti tra di loro e si vede cosa concludono, ma anche in questo caso si rinvia ad un momento successivo la risposta ufficiale e si prosegue.

3.1 La Lotteria del Dado

La Lotteria del Dado è un nuovissimo gioco a premi che prevede ogni settimana un singolo lancio di un normale dado a 6 facce. Giocare costa 1 euro e chi indovina ne vince 2.

A: Generiamo 3000 estrazioni della Lotteria del Dado

Su uno spreadsheet come MS Excel si crea un foglio nuovo e si riempiono le caselle da 2 a 3001 della colonna A con numeri casuali uniformi tra 1 e 6. La casella A1 si lascia vuota.

Su MS Excel conviene usare il comando

```
=CASUALE.TRA(1;6)
```

e fare copia-incolla su tutte le celle richieste. Premendo F9 i valori casuali vengono generati nuovamente e questo sarà utile in seguito. Se 3000 lanci casuali di dado sono troppi rispetto alla potenza del calcolatore usato, si scenda pure a 1000.

Una alternativa *con le mani* al comando di sopra è `=INT(6*CASUALE())+1`.

A: Verificare che ciascun numero sia uscito circa 500 volte.

Nelle caselle da C2 a C7 si inseriscono i numeri da 1 a 6. Nella casella D2 si inserisce il comando

```
=CONTA.SE(A$2:A$3001;C2)
```

che si copia-incolla nelle celle da D3 a D7. (Attenzione ai dollari!)

In questo modo si può verificare che ogni numero è uscito circa 500 volte, ma (premendo F9 ripetutamente) che sono frequenti deviazioni anche di 30 punti. Si nota inoltre che queste deviazioni non sembrano essere dovute ad una mancanza di casualità delle estrazioni del calcolatore, infatti non sono sempre sugli stessi numeri e non sono sempre nella stessa direzione[6].

[5] Si pensi infatti ad un giocatore incallito che si basa sui ritardi: sarà molto difficile che salti anche solo un'estrazione, perché il numero ritardadario potrebbe uscire in qualunque momento. Il giocatore ritardista è spesso schiavo del gioco a cui partecipa.

[6] Viceversa, se il conteggio dei 5 risultasse tipicamente maggiore di 500, vorrebbe dire che il generatore casuale del PC favorisce questo numero. Così non dovrebbe essere

> **A:** Calcolare per ogni estrazione il ritardo di ciascuno dei 6 numeri.

Si può usare il comando `CONFRONTA`, che permette di trovare la prima occorrenza di un numero in un array. L'unica cosa un po' complicata è come mettere giù le cose sul foglio di calcolo e l'uso raffinato dei dollari.

Per prima cosa è necessario supporre che le righe più in alto rappresentino estrazioni più recenti, quindi A2 è l'ultima estrazione, A3 la penultima, ... A3000 è la seconda e A3001 la prima.

Si inseriscono i numeri da 1 a 6 nelle celle F1, G1, ... K1. Nella cella F2 si inserisce la formula

```
=CONFRONTA(F$1;$A2:$A$3001;0)
```

Attenzione ai dollari: precedono tutti e soli i riferimenti alla riga 1, alla riga 3001 e alla colonna A. Si fa copia-incolla di questa cella su tutto il rettangolo di celle F2:K3001.

Ora la riga F2:K2 riporta i ritardi dei 6 numeri dopo l'ultima estrazione. La riga F3:K3 dopo la penultima estrazione e così via. Si verifichi il corretto funzionamento premendo F9 un po' di volte e controllando che i ritardi riportati siano confermati guardando la colonna A. È normale che nelle ultime righe dal rettangolo la formula dia parecchi errori. (Perché?)

> **Q:** *Si può notare che in ogni momento i 6 numeri hanno ritardi che sono sei numeri distinti: perché? Il minimo è sempre 1: perché?*

Le risposte a queste domande sono semplici: gli studenti devono arrivarci senz'altro e capirne bene il significato.

> **A:** Il comportamento del ritardo massimo ad ogni estrazione è meno semplice da comprendere. Si vuole studiarne la distribuzione: quanto grande può essere?

Si inserisce la formula `=MAX(F2:K2)` nella cella M2 e si copia-incolla nelle celle sottostanti, fino alla cella M3001.

Se si vuole trovare il massimo ritardo storico delle 3000 estrazioni occore un po' di accortezza, perché siccome le ultime righe in basso danno errore, il comando `=MAX(M2:M3001)` dà certamente errore. Si può risolvere in fretta limitando il numero di righe con `=MAX(M2:M2900).` Si trova che sovente questo numero supera le 50 estrazioni, nonostante ciascun numero debba uscire in media una volta ogni 6 estrazioni.

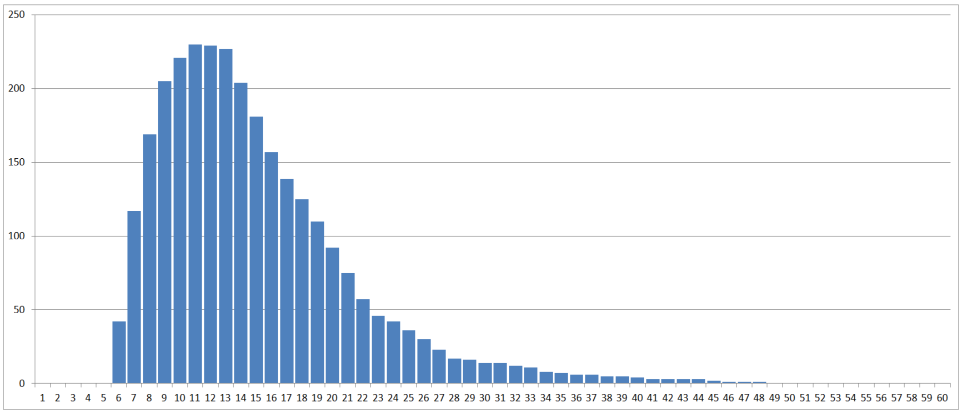
È possibile anche tracciare un istogramma dei ritardi massimi, anche se è un po' laborioso. Nelle caselle da O2 ad O61 si inseriscono i numeri da 1 a 60[7]. Si selezionano le celle da P2 a P61, lasciandole selezionate e senza cliccare da nessuna parte si scrive sulla tastiera (ma senza concluderla con Invio) la formula

```
=FREQUENZA(M2:M2900;O2:O61)
```

Invece di inserirla come al solito premendo Invio alla fine, si inserisce come formula matriciale, premendo Ctrl-Shift-Invio[8]. Si può notare che la formula che compare nella barra delle formule è apparentemente

```
{=FREQUENZA(M2:M2900;O2:O61)}
```

e le celle si riempiono con i conteggi delle frequenze dei massimi ritardi. A questo punto si possono selezionare le celle P2:P61 e inserire un grafico a colonne. La distribuzione esce abbastanza regolare e con un massimo tra 10 e 12. I ritardi tipici del numero più ritardatario sono di questo ordine di grandezza, ma il grafico mostra che hanno una probabilità non trascurabile anche ritardi di più di 40 estrazioni.



> **A:** Si calcoli la media dei ritardi di tutti i numeri per tutte le estrazioni (tranne le cento più vecchie).

La formula è semplicemente

```
=MEDIA(F2:K2900)
```

[7]Un modo veloce è inserire i primi due numeri, selezionarli entrambi con il mouse, quindi cliccare sul quadratino nero in evidenza nell'angolo in basso a destra del riquadro di selezione e trascinare verso il basso. Excel riempe le celle continuando la progressione aritmetica dei due numeri selezionati.

[8]Si tengono premuti i tasti Ctrl e lo Shift (quello per fare le maiuscole) e di preme Invio.

Dovrebbe uscire un numero molto vicino a 6, in accordo con la buona distribuzione dei sei numeri. Si osservi come ciò significhi che anche la media dei 6 ritardi su ciascuna riga sia circa 6 (anche se con più variabilità). Ovviamente, se 6 numeri interi tutti distinti e con il minimo sempre esattamente uguale a 1 hanno media circa 6, il massimo avrà tipicamente un valore intorno a 11, in accordo con le osservazioni.

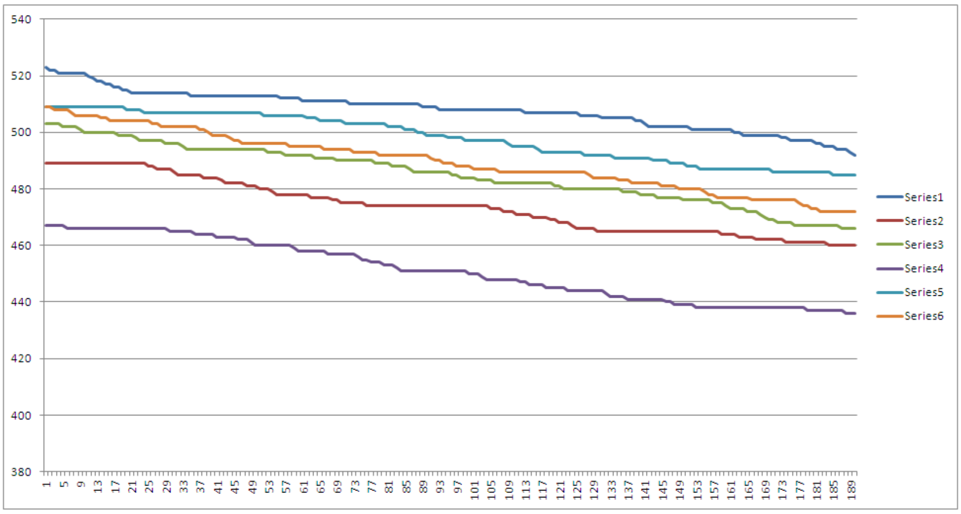
3.2 Verso la LGN

Rinominiamo il foglio usato finora Ritardi e iniziamo a lavorare su un secondo foglio. Si inseriscono i numeri da 1 a 6 nelle celle da A1 a F1. Nella cella A2 si inserisce la formula

```
=CONTA.SE(Ritardi!$A2:$A$3001;A$1)
```

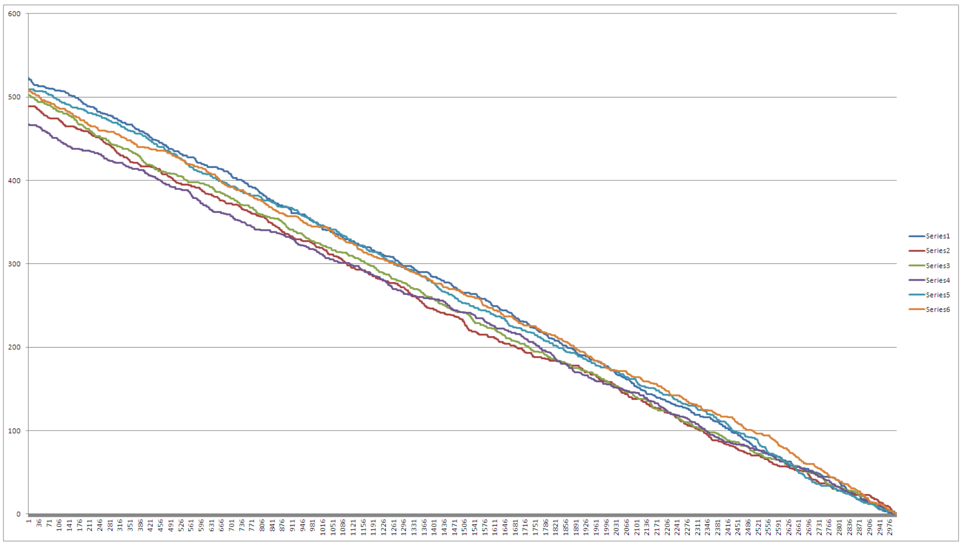
(Di nuovo, attenzione ai dollari!) Facendo copia-incolla di questa cella nelle celle del rettangolo A2:F3001 si ottiene un conteggio progressivo delle frequenze assolute dei diversi numeri.

A questo punto si può fare un grafico (a linee) delle prime cento o duecento righe, per ottenere un'impressione visiva di come le frequenze cumulate dei sei numeri crescano di pari passo in modo abbastanza regolare. (L'asse temporale è invertito, quindi all'ascissa 1 corrisponde il conteggio finale dei 3000 lanci.)

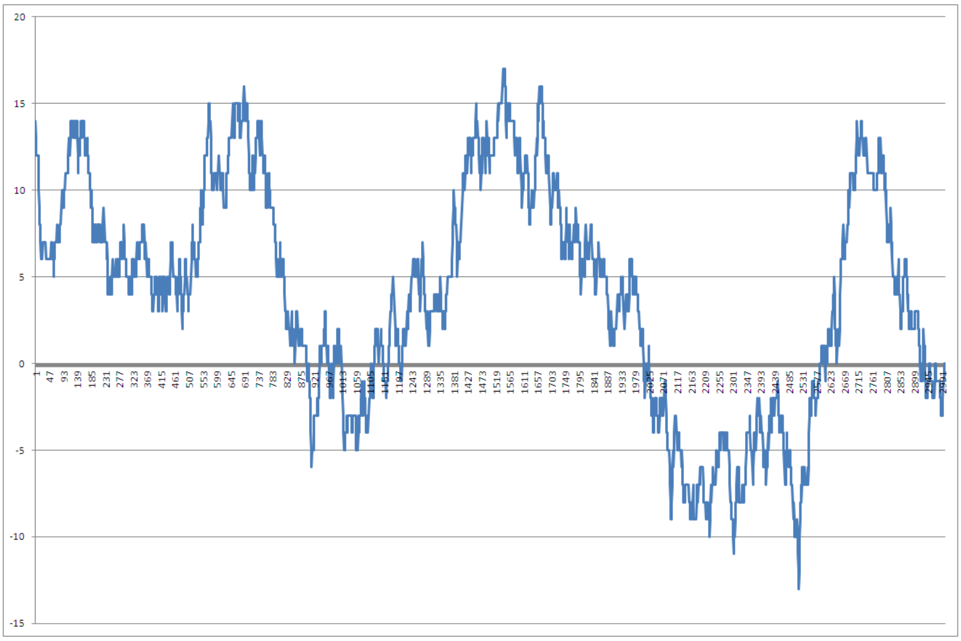


Si può notare che se un numero è più frequente in un determinato momento, tende a restarlo nelle estrazioni successive. Ad esempio in questo grafico il numero 1 è costantemente il più frequente e il numero 4 è costantemente il meno frequente. Il caso non ha memoria e non cerca in nessun modo di regolare i conti, riequilibrando le frequenze dei sei numeri.

Lo stesso grafico fatto su tutte le 3000 linee delle estrazioni fa un effetto diverso.



Si vede che le linee di frequenza si intersecano diverse volte e che la loro distanza aumenta progressivamente. Perciò non è vero che le frequenze assolute dei sei numeri tendano a pareggiarsi, anzi la loro distanza tipica aumenta lentamente. Però nulla vieta che si sorpassino a vicenda, anche molte volte. Il prossimo grafico rappresenta la differenza tra la frequenza degli 1 e quella dei 5 nello stesso esperimento dei grafici precedenti.



Giunti a questo punto, l'insegnante legge o proietta la frase seguente:

Una versione della Legge dei Grandi Numeri dice la cosa seguente. Supponiamo di essere in presenza di esperimenti identici, ripetuti in modo indipendente per n volte. Allora al crescere di n, le frequenze relative dei diversi esiti tendono alle relative probabilità.

Q: *Cosa significa questa affermazione? Come può ciò essere compatibile con gli esiti degli esperimenti osservati fino a qui?*

Il punto chiave è la differenza tra frequenze *assolute* e *relative*. Se le prime tendessero ad essere uguali, avrebbero ragione i ritardisti, ma le seconde possono tendere allo stesso valore ($\frac{1}{6}$ in questo caso) senza che ci sia bisogno di immaginare un meccanismo di bilanciamento delle estrazioni.

Per chiarire il concetto sarà utile per le stesse 3000 estrazioni del nostro gioco calcolare le frequenze relative cumulate.

A: Riprodurre il grafico delle frequenze cumulate per le frequenze relative dei sei numeri.

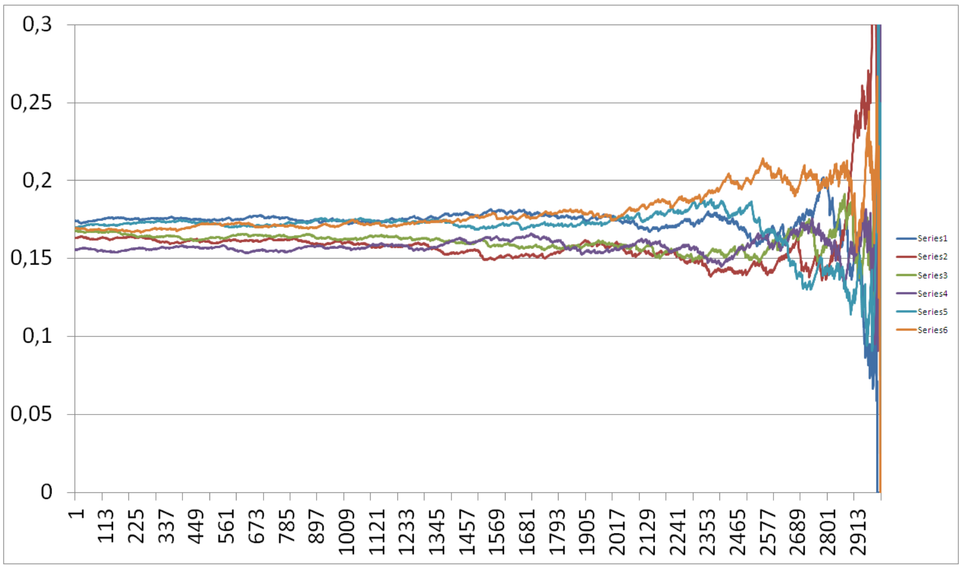
Non è assolutamente complicato. Uno dei tanti modi è questo. Si inizia un nuovo foglio, si inseriscono i numeri da 1 a 6 nelle celle da A1 a A6, si inserisce la formula

```
=CONTA(Ritardi!$A2:$A$3001)
```

nella cella H2 e si copia-incolla nelle celle H2:H3001. Si inserisce la formula

```
=CONTA.SE(Ritardi!$A2:$A$3001;A$1)/$H2
```

nella cella A2 e si copia-incolla nelle celle del rettangolo A2:F3001. Un esempio di grafico è riprodotto qui sotto. (Sempre con asse temporale invertito.)



3.3 Un possibile lavoro finale

Questo tipo di analisi può essere riprodotta per una vera lotteria, come il Superenalotto o altra di gradimento degli studenti. Si possono aggiungere dettagli di colore sul mondo dei ritardisti.

Si può anche esplorare l'argomento del cosiddetto gioco al raddoppio, che illustriamo brevemente per la Lotteria del Dado. Un giocatore scommette 1 euro (ovvero gioca una volta) sul 5 (o su qualunque altro numero). Se vince, ottiene 2 euro, con un guadagno netto di 1, ed è felice. Se perde, la settimana dopo scommette di nuovo (magari ancora sul 5) 2 euro (ovvero gioca due volte). Se vince, ottiene 4 euro, in totale ne ha investiti 3 euro, con un guadagno netto di 1, ed è anche in questo caso felice. Se perde, la settimana dopo scommette di nuovo, questa volta 4 euro, e così via. L'idea è ad ogni passaggio di scommettere abbastanza da rifarsi di tutti i soldi persi in precedenza ed avere in caso di vincita un guadagno netto di 1. Indubbiamente vi è il 100% di probabilità che prima o poi il giocatore azzecchi la scommessa e vinca. Dove sta la frequatura?

(Alcuni giocatori del mondo reale si cimentano davvero con versioni diverse di questo sistema e questo è uno dei modi più pericolosi di giocare, poiché spesso finiscono completamente rovinati.)

4 Roulette e Speranza Matematica

Questa unità andrebbe affrontata in due ore, e serve ad introdurre la speranza matematica, che sarà nuovamente utilizzata nel trattare il Poker.

Anche in questo caso, come nell'unità precedente, sarebbe utile fare l'attività in laboratorio, sfruttando Excel. I comandi che dovranno usare i ragazzi sono già stati introdotti nell'unità precedente.

Alle prime tre domande dedicate 15-20 minuti in totale.

Q: *Conoscete la roulette? Come si gioca? Quali sono le possibili puntate?*

Ci sono due versioni di roulette, la francese (36 numeri, metà rossi e metà neri, e lo 0, verde) e l'americana (36 numeri, metà rossi e metà neri, e 0 e 00, verdi). Si può puntare sul singolo numero, su una coppia di numeri vicini, su un quartetto di numeri vicini, su tre o sei numeri vicini, sui primi, gli ultimi o i 12 numeri di mezzo; su pari o dispari, rosso o nero, sui primi o gli ultimi 18 numeri. Può essere utile proiettare una foto di un tavolo della roulette. Sarebbe interessante capire quale delle due versioni conoscono i ragazzi, e adottare quella versione per l'attività.

Q: *Quanto si vince alla roulette? Cos'è la vincita?*

Il concetto di vincita è ambiguo: la vincita sono i soldi che mi vengono dati dal croupier quando vinco o quelli che mi vengono dati al netto della spesa? Solitamente per la roulette si usa la seconda convenzione, ma per altri giochi (gratta e vinci, lotterie) si usa la prima.

Per non confonderci, definiamo i seguenti termini:

- Spesa S: quanto punto in una singola giocata (posta).
- Incasso I: i soldi che si ricevono in caso di scommessa vincente.
- Bilancio B: $B = I - S$ se la scommessa è stata vincente, $B = -S$ se la scommessa è stata perdente; ovvero quanto ho guadagnato (o perso se negativo) dalla puntata.

Gli incassi nella roulette sono in seguenti (uguali per entrambe le versioni):

- numero: 36 volte la posta;
- due numeri: 18 volte la posta;
- tre numeri: 12 volte la posta;
- quattro numeri: 9 volte la posta;
- sei numeri: 6 volte la posta;
- 12 numeri: 3 volte la posta;
- pari/dispari, rosso/nero, 18 numeri: 2 volte la posta.

Q: notate uno schema negli incassi relativi alle varie giocate? secondo voi quale/i giocata/e è/sono più conveniente/i?

Stimolate la discussione fra i ragazzi: si spera che si accorgano che il prodotto tra numeri giocati e incasso è costante, sempre 36. Da questo riescono ad intuire che tutte le giocate sono (s)convenienti allo stesso modo? Come lo spiegano? Se se ne accorgono bene, se no, fate notare loro come il prodotto sia costante, e andate avanti senza scendere in dettagli.

4.1 Simulazione della Roulette

Facciamo ora simulare una lunga successione di giocate alla roulette (francese o americana a seconda di come i ragazzi hanno risposto alla prima domanda).

A: *Generiamo 3000 lanci di pallina alla roulette.*

Su uno spreadsheet come MS Excel si crea un foglio nuovo e si riempiono le caselle da 2 a 3001 della colonna A con numeri casuali uniformi tra 0 e 36 (tra 0 e 37 se si usa la roulette americana, intendendo con 37 il doppio0). La casella $A1$ si lascia vuota. Su MS Excel conviene usare il comando `=CASUALE.TRA(0;36)` e fare copia-incolla su tutte le celle richieste. Premendo *F*9 i valori casuali vengono generati nuovamente. Se 3000 lanci casuali della pallina sono troppi rispetto alla potenza del calcolatore usato, si scenda pure a 1000.

A: *Contiamo quante volte sono usciti i numeri da 0 a 36 e alcune combinazioni di giocate (es: i numeri 4-5; i numeri 1-2-3; i numeri 32-33-35-36; i numeri 16-17-18-19-20-21; pari; ...)*

Per contare le uscite dei singoli numeri 0,...,36 si può fare come fatto per il caso della lotteria del dado: nelle caselle da C10 a C46 si inseriscono i numeri da 0 a 36. Nella casella D10 si inserisce il comando `=CONTA.SE(A$2:A$3001;C10)` che si copia-incolla nelle celle da D11 a D46. (Attenzione ai dollari!)

Nelle caselle da C2 a C9 si possono inserire altre giocate che ci interessano, e nella cella sottostante si contano quante volte sono uscite: ad esempio per contare quante volte è uscito il carré (4 numeri) 32-33-35-36 si inserisce il comando `=D42+D43+D45+D46`

Premendo F9 si possono generare nuovi lanci casuali di pallina e osservare come cambiano i valori dei lanci.

A: *Analizziamo ora quanto ha vinto o perso un giocatore che ha puntato per 3000 volte sullo stesso evento, puntando 1 euro per ogni giocata.*

Per fare ciò, nelle caselle da E2 a E46 scriviamo quanto si vince se si realizza l'evento indicato in quella colonna: 36 nelle caselle da E2 a E36 (singolo numero), 9 nel caso del carré, 2 nel caso di pari/dispari, ecc.

La cifra vinta puntando sempre sull'evento C2 è il numero di volte che si è verificato l'evento (ovvero D2) per l'ammontare vinto ogni volta (ovvero E2); per ottenere il bilancio complessivo del giocatore, togliamo i 3000 euro puntati nelle giocate. Nella cella F2 inseriamo la formula `=D2*E2-3000` (e si copia-incolla nelle celle da F3 a F46).

I ragazzi dovrebbero osservare che avendo ripetuto 3000 volte la stessa scommessa, il giocatore è (quasi certamente) in negativo. Dividendo la cifra per le 3000 giocate effettuate nelle caselle G2-G46:`=F2/3000` (e copia-incollato nelle celle da G3 a G46) si dovrebbe osservare che i bilanci medi sono tutti molto vicini a -0,027, ovvero ad una perdita netta di 2,7 centesimi di euro a giocata.

Si può premere F9 per generare nuove successioni casuali di lanci, e vedere cosa succede.

Q: *Dai dati che abbiamo visto, sembra che alcune giocate alla roulette siano meglio di altre?*

Ovviamente ci attendiamo che i ragazzi osservino che qualunque tipo di giocata porta approssimativamente allo stesso bilancio medio negativo ...

[questa parte è da fare solo se c'è tempo, altrimenti consiglio di saltarla:

A: *Calcolare quanto è il bilancio medio dopo* 500/1000/1500/2000/2500 *giocate; o dopo* 200/400/600/800 *giocate, se abbiamo deciso di fare solo 1000 lanci*

Basta ripetere quanto fatto nelle attività precedenti, avendo cura di contare solo le occorrenze nelle prime 500/1000 ecc. celle

Q: *I bilanci medi delle varie giocate sono sempre molto simili? Variano di più o di meno all'aumentare delle giocate? Perché?*

La speranza è che i ragazzi tirino fuori la legge dei grandi numeri vista la volta prima e ragionino sulle frequenze...]

4.2 La Speranza Matematica

Q: *Perché è successo questo? Che cosa rappresenta il bilancio medio di* $-0,027$*? (*Siete liberi di sostituire questo numero con quello che sembrano suggerire i dati) Potevamo attendercelo prima di fare la simulazione?*

Stimoliamo le ipotesi dei ragazzi, facendoli discutere un po'. Si spera che -dopo aver ragionato sulle frequenze- qualcuno si accorga del fatto che la probabilità di vincere con una certa giocata (ovvero la frequenza con cui si tende a vincere, se si effettua un numero molto grande di lanci) per quanto si vince è costante, e minore di 1... E' precisamente in tutti i casi 36/37 ovvero circa 0,973. E guarda caso $1 - 0,973 = 0,027$ ...

Possiamo quindi dare un nome a questo prodotto: la *speranza matematica di vincita*.

Siccome tutte le scommesse hanno la stessa speranza matematica, e siccome il numero di tentativi effettuati è grande, la legge dei grandi numeri ci assicura che la frequenza relativa delle vincite sarà vicina alla probabilità a priori, e pertanto -per la legge dei grandi numeri- il bilancio medio sarà vicino a $-0,027$ euro a giocata.

Q: *Possiamo calcolare la speranza matematica di vincita per altri giochi? Qual è la speranza matematica di vincita per la lotteria del dado della scorsa volta? Come possiamo calcolare la speranza matematica di vincita in altri giochi?*

<u>LOTTERIA DEL DADO</u>:

$$\text{Incasso} * \text{Probabilità di Vittoria} = 2€ \times 1/6 = 1/3;$$

$$1/3 - 1 = -2/3.$$

Allo stesso modo, se scommetto 1 € su ciascuno dei 6 numeri, gioco 6 €. Vinco certamente e ottengo 2 €; $\Longrightarrow$ ho perso 4 €, cioè ho perso per ciascuna giocata $4/6 = 2/3$.

Si spera che i ragazzi abbiano interiorizzato il meccanismo e sappiano rispondere alla domanda. Cercate di guidarli verso la risposta, se hanno difficoltà.

Q: *E per il Black Jack visto la prima volta, cosa si può dire?*

Qui le cose si fanno più complesse, dato che c'è della strategia nel gioco. Sarebbe bello se nella discussione si accorgessero che le diverse possibili strategie corrispondo a diverse speranze matematiche di vincita e che scegliere la strategia migliore vuol dire massimizzare la speranza matematica di vincita.