5.6 Torniamo ai problemi iniziali

A questo punto gli studenti dovrebbero disporre degli strumenti matematici per rispondere alle questioni proposte all'inizio del percorso nonché per esaminarle criticamente. La risoluzione è accompagnata da commenti, video, esplorazioni con il software... per favorire una comprensione più profonda delle situazioni esaminate e permettere così ai ragazzi di prendere decisioni consapevoli.

a) Regolarità

Lanciamo 10 volte una moneta "onesta".
Su quale tra le due sequenze di esiti scommettete?

$$T\ T\ T\ T\ T\ T\ T\ T\ T\ T \qquad T\ C\ T\ C\ C\ T\ C\ T\ T\ C$$

Ricorrendo alla legge della moltiplicazione si ottiene che le probabilità dei due eventi sono uguali:

$$p(T\ T\ T\ T\ T\ T\ T\ T\ T\ T) = p(T)\cdot p(T)\cdot \ldots \cdot p(T) = \left(\mathbf{\frac{1}{2}}\right)^{\mathbf{10}}$$

$$p(T\ C\ T\ C\ C\ T\ C\ T\ T\ C) = p(T)\cdot p(C)\cdot \ldots \cdot p(C) = \left(\mathbf{\frac{1}{2}}\right)^{\mathbf{10}}$$

Ma, al di là della valutazione numerica, le probabilità dei due eventi sono uguali essenzialmente perché *i lanci sono* ***indipendenti.***

Più in generale si può osservare che ogni sequenza di 10 lanci della moneta ha probabilità $\left(\frac{1}{2}\right)^{\mathbf{10}}$

b) Compensazione (riformulata)

Lanciamo 10 volte una moneta "onesta".
L'esito dei primi 9 lanci è

$$T\ T\ T\ T\ T\ T\ T\ T\ T$$

È vero che al decimo lancio è più probabile ottenere C?

Al decimo lancio, come al primo, vale $p(T) = p(C) = \frac{1}{2}$ in quanto i lanci sono indipendenti[13] ovvero "la moneta **non ha memoria**".

Osservazione
Si è risposto alla questione iniziale, ma è interessante investigare insieme agli studenti l'origine del misconcetto.

- L'evento "i primi nove lanci hanno tutti esito testa" è **poco probabile**, infatti:

$$p(T\ T\ T\ T\ T\ T\ T\ T\ T) = \left(\frac{1}{2}\right)^9 < \frac{1}{500}$$

Questa è anche la probabilità di una qualsiasi sequenza fissata costituita da 9 simboli ciascuno scelto tra T, C.

[13] Ossia, più formalmente, gli eventi "esito del lancio i-esimo", "esito del lancio i+1-esimo" sono indipendenti.

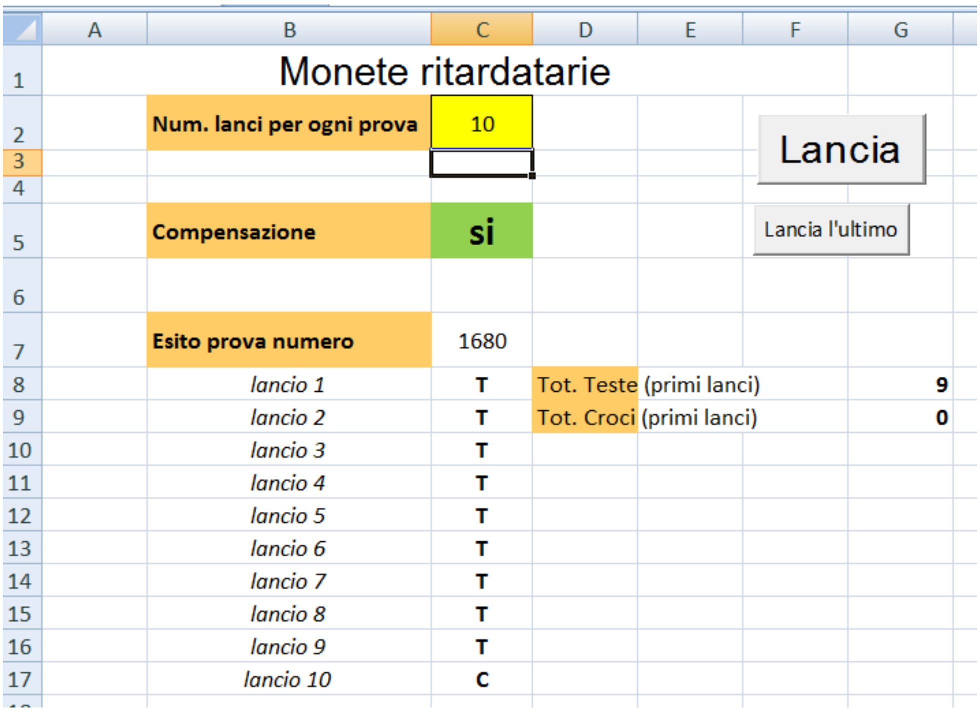
Però, secondo il testo del problema, questo evento ormai **è accaduto**, è un evento certo, è storia. Solo ciascuno dei due esiti del decimo lancio (ossia "esce T", "esce C" è un evento aleatorio, su cui ha senso scommettere.

Quali sono dunque i fraintendimenti sottesi?

- Il primo, come appena visto, risiede nel considerare globalmente i 10 esiti e non il solo decimo esito.
- Il secondo è nell'interpretare in modo distorto la **Legge dei grandi numeri**. Più precisamente, tale legge si applica a sequenze di esiti e non al singolo esito: dunque non si può applicare al problema in esame. In ogni caso proviamo a considerare comunque l'insieme dei 10 lanci: la legge essa afferma, sostanzialmente, che su "molti" lanci è "grande" la probabilità che la frequenza relativa delle croci sia $\frac{1}{2}$; ciò è coerente con il fatto che, sui 10 lanci, è maggiore la probabilità di ottenere una C rispetto a quella di non ottenerne nessuna (vi sono infatti 10 sequenze che contengono una sola C, mentre solo una sequenza è costituita interamente da T); ma la legge non precisa a quale lancio è più probabile esca la C, mentre nel problema in esame interessa l'uscita di C ad un lancio ben preciso: il decimo.

*Tutto questo in teoria, ma nella **pratica** cosa succede? Ecco una simulazione mediante il foglio elettronico per investigare la questione.*

L'idea dell'attività e una traccia di lavoro si trovano in fondo al capitolo nell'appendice (A3 Attività). Entrambe fanno riferimento al file Excel MoneteRitardatarieConsecutive.xlsm, predisposto per studenti.



Gli studenti possono ora rispondere in modo autonomo al quesito iniziale "Marta e i bambini".

c) Numeri ritardatari

Quanto appena discusso a proposito dei lanci di moneta si può applicare anche alle estrazioni del gioco del Lotto. Pertanto, in questo contesto potrebbe aver senso far lavorare gli studenti in modo più autonomo e piuttosto soffermarsi a discutere collettivamente le conclusioni individuali.

Il “53” non è uscito per 182 estrazioni consecutive sulla ruota di Venezia.
Qual è la probabilità che esca su tale ruota alla 183-esima estrazione?

La probabilità che esca il “53” ad **una** data estrazione su tale ruota è

$$p = \frac{5}{90} = \frac{1}{18}$$

La probabilità di uscita alla **183-esima** estrazione è ancora p: le estrazioni sono indipendenti (per il meccanismo fisico di estrazione).

<u>Osservazioni</u>

- **Ritardo del 53**

Qual è la probabilità che il “53” non esca per 182 estrazioni consecutive?

La probabilità che, ad una (qualunque) estrazione fissata, non esca il “53” è 1-p. Pertanto, per la legge della moltiplicazione la probabilità che il “53” non esca per tutte le 182 estrazioni consecutive è:

$$(1-p)^{182} \approx 0{,}000030$$

L’evento è dunque “poco probabile”, ma ormai è storia.
Comunque, per la cronaca, … il “53” è uscito su Venezia il 9 febbraio 2005, alla 183 – esima estrazione…

- **Numeri spia**

Ecco un'attività che gli studenti dovrebbero trovare stimolante...
Sul sito della Lottomatica (https://www.lottomaticaitalia.it/it/prodotti/lotto/statistiche/numero-spia) vengono elencati dei numeri speciali e sono presentati con queste parole:
“La tradizione vuole che l’estrazione di certi numeri “preannunci” l’uscita di altri. La tabella indica, per ogni numero estratto (o “**numero spia**”), 5 numeri che avrebbero la maggior probabilità di uscire.”

Venezia	
27	42 58
	65
	86
	90

Ad esempio:

L’8 gennaio 2018 è uscito il 27 sulla ruota di Venezia. Allora è vero che ciò aumenta la probabilità dell’uscita del 42 all’estrazione successiva?

d) Test clinici

Prima di affrontare i problemi motivanti proposti all'inizio del percorso relativi ai test clinici, è importante che gli studenti prendano confidenza con la situazione affrontando dei problemi più semplici, come i seguenti.

- Test di gravidanza:

si veda in fondo al capitolo il testo e la risoluzione all'appendice A4

- Un problema relativo alla diagnosi di una malattia:

Una popolazione di 10.000 individui è stata sottoposta a un test per diagnosticare una certa malattia.
Sono risultate positive al test 1.726 persone e si assume che il test sia risultato positivo per il 99,0% dei malati.
Inoltre si assume che il 2,0% della popolazione avesse la malattia.
Qual è la probabilità che il test abbia fornito **indicazioni errate** su un individuo scelto a caso in tale popolazione?

Il video di questo problema e della risoluzione realizzato dal Laboratorio DicoMat Lab dell'Università degli studi di Trento si trova all'indirizzo http://youtu.be/N_sdkLtECps.

Ecco allora uno dei problemi posti ad inizio percorso relativamente a questo contesto.

Qual è la probabilità che il test "Elisa" fornisca indicazioni errate?

Il quesito, così formulato, non ammette risposta univoca: era stato volutamente presentato in forma contratta per fare emergere le convinzioni di senso comune. Pertanto, prima di procedere è opportuno precisare alcune ipotesi sul test.

Una popolazione è sottoposta al test "Elisa" per la diagnosi dell'HIV.
La probabilità che il test sia positivo sull'individuo che ha il virus è del 99,9% (sensibilità del test).
Assumiamo che la probabilità di essere negativo al test per l'individuo "sano" sia del 99,8% (specificità).
Inoltre assumiamo che lo 0,3% della popolazione abbia la malattia (prevalenza).
Qual è la probabilità che il test fornisca **indicazioni errate** su un individuo scelto a caso in tale popolazione?

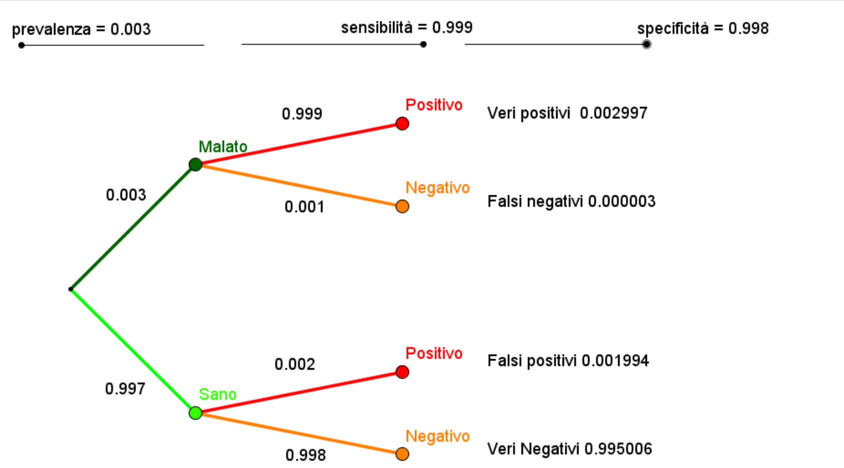
Questo problema si risolve analogamente ai due problemi precedenti. Così, modellizzando la situazione mediante un grafo ad albero, si arriva a concludere che:

$$p(\text{“}esito\ errato\text{”}) = 0{,}997 \cdot 0{,}002 + 0{,}003 \cdot 0{,}001 \approx \mathbf{0{,}002}$$

Più in generale, ha senso chiedersi come varia la probabilità richiesta al variare dei valori in ipotesi. Immergiamoci dunque in un'attività esplorativa, con il supporto del file Geogebra Test_Malattia.ggb[14]. È importante però che la risposta ottenuta empiricamente sia accompagnata da una giustificazione algebrica.

[14] Il materiale è stato ideato dalla collega Francesca Arrigoni.

Ad esempio, si può iniziare indicando con *a* la prevalenza della malattia ed esprimere la probabilità richiesta, cioè dell'evento "esito errato", in funzione di *a*; si ottiene $p = 0{,}002 - a \cdot 0{,}001$; dunque si conclude che al crescere della prevalenza *a*, la probabilità p di ottenere un esito errato diminuisce e volendo si può precisare che diminuisce linearmente.



	A	B	C	D
1		Positivi	Negativi	
2	Malati	0.002997	0.000003	0.003
3	Sani	0.001994	0.995006	0.997
4		0.004991	0.995009	1
5				
6	**Errati**	**0.001997**		
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				