

### 3. La probabilità alla prova... esperimenti e simulazioni

*Nel capitolo precedente ci siamo occupati di valutazioni di probabilità, ossia di come attribuire un numero al grado di fiducia che riponiamo nel realizzarsi di un evento. Però non ci siamo ancora occupati di “cosa succede” nella pratica, ossia non abbiamo ancora discusso esperimenti materiali e simulazioni al calcolatore. Ciò è avvenuto solo per ragioni di chiarezza espositiva: teoria e pratica sono approcci indissolubilmente intrecciati, che si arricchiscono ed integrano a vicenda, e che vanno dunque considerati contestualmente.*

*Certo, chi ha esperienza didattica sa che le attività laboratoriali sono molto costose in termini di tempo e di risorse investite. Pertanto, prima di passare alla loro analisi, è quanto mai opportuno fermarsi a riflettere sulle ragioni didattiche che giustificano l'effettuazione di prove di fenomeni aleatori.*

*È proprio quanto ci proponiamo di fare in questo capitolo: inizieremo con la discussione di tali motivi didattici e metteremo in luce quali attenzioni prestare affinché le attività risultino efficaci per la costruzione dei saperi; proseguiamo esaminando criticamente alcune attività di simulazione mediante il foglio elettronico o esperimenti materiali condotti con oggetti, quali dadi o monete.*

*Al docente spetterà poi scegliere quali attività proporre, in quale ordine cronologico svilupparle e in che modo integrarle con le considerazioni teoriche esposte nel capitolo precedente.*

#### 3.1 Esperimenti e simulazioni

##### Perché?

*In generale, esperimenti e simulazioni sono attività laboratoriali e, in quanto tali, permettono di coinvolgere gli studenti in processi didatticamente significativi.*

- *Favoriscono la costruzione di significati degli oggetti matematici coinvolti e l'attribuzione di un senso ai risultati teorici. In particolare, essi consentono allo studente di formarsi un'idea di “cosa succede nella pratica”, ossia di come evolvono gli esiti di un esperimento casuale. Un passo essenziale per superare vari misconcetti, così ben radicati nella conoscenza di senso comune. Infatti la sola informazione teorica non basta: “in teoria le cose dovrebbero andare così, ma nella pratica vanno diversamente ...”<sup>1</sup>. A tale proposito, provate a comunicare l'intenzione di scommettere sull'uscita dei numeri 1, 2, 3, 4, 5 al gioco del lotto. Quali commenti vi aspettate? Saranno diversi quelli degli interlocutori a cui è noto che ogni cinquina ha la stessa probabilità di uscire ... teoricamente?*
- *Consentono l'attivazione dello studente in prima persona: attore, non spettatore di un racconto. Un diktat, questo, che ispira ormai anche l'organizzazione di vari percorsi museali, alcuni dei quali sono evoluti dalla pura divulgazione ad una più costruttiva e democratica interazione con il visitatore. In effetti l'impatto sul visitatore è completamente diverso quando si passa da un approccio del tipo “ti spiego come vivono le api” a “sei un ape, come ti muoveresti?”*

*Tali considerazioni riguardano le attività laboratoriali in genere. Le simulazioni mediante il foglio elettronico presentano, a nostro avviso, un valore aggiunto.*

---

<sup>1</sup> Più generale, quando gli studenti rispondono ad una domanda senza essere del tutto convinti della correttezza della risposta, esordiscono con frasi del tipo “in teoria ...”. Questa modalità di intendere il termine “teoria” dovrebbe far riflettere.

- Permettono di effettuare “molte” prove di un dato esperimento in tempi ragionevoli. È questa una condizione indispensabile per osservare fatti significativi. Infatti, lo studente difficilmente potrà formarsi un’idea adeguata della distribuzione di probabilità nel lancio di due dadi a partire dall’esame di soli 20 esiti.
- Esigono di curare la progettazione e la formalizzazione dei procedimenti. Progettazione e formalizzazione a volte sono percepite come astratte richieste del docente. Ma in attività con il foglio elettronico si trasformano esigenze concrete del software. Vediamo di precisare meglio. Se la sintassi dei comandi non è corretta, il software non produce alcun output; al più restituisce un codice di errore. Inoltre, per implementare le formule, si deve spesso far riferimento agli indici che identificano le celle; in sostanza, si è costretti ad esplicitare in un preciso linguaggio ciò che si ha intenzione di fare. Infine, affinché la macchina sia in grado di eseguire il procedimento, si rende necessaria la scomposizione in passi elementari. Ad esempio, per calcolare le frequenze nel lancio di due dadi, si devono prima generare i numeri casuali che rappresentano le uscite dei singoli dadi, poi sommarli per ottenere i punteggi e infine contare le occorrenze di ogni possibile punteggio. L’aspetto didatticamente significativo è che ciascuno di questi passi si può rappresentare visivamente mediante una colonna del foglio di calcolo.

### Con quali attenzioni?

Gli esperimenti e le simulazioni sono attività potenzialmente feconde. Però, per non vanificarne l’efficacia, vanno attuate prestando precise attenzioni<sup>2</sup>.

- Interpretare gli esiti delle prove: la matematica non deve restare “negli occhi di chi osserva”. Ciò richiede una lettura degli esiti che, da una parte, sia corretta matematicamente e, dall’altra, sia espressa in un linguaggio adeguato alla classe. In particolare, è bene tener presente che gli esiti di un numero necessariamente finito di prove non possono costituire una dimostrazione.
- Integrare l’esame “teorico” con le attività sperimentali, visto che i due punti di vista si arricchiscono a vicenda. E, in quest’ottica, non contrapporre approccio classico e approccio frequentista nella valutazione della probabilità: lo schema frequentista non si introduce solo per modellizzare situazioni in cui sembra aver poco senso seguire l’approccio classico (a tale proposito si veda, ad esempio, quanto osservato nel paragrafo 2.6 relativamente alla modellizzazione della questione di nascere maschio o femmina).

A tali attenzioni si aggiungono quelle, necessariamente più specifiche, da riservare alle attività mediante il foglio elettronico.

- Aver chiaro che il calcolatore non genera numeri casuali. Non può farlo, visto che è vincolato ad operare in modo deterministico. Però può produrre delle sequenze di numeri che ai fini pratici esibiscono (quasi) lo stesso comportamento delle sequenze generate in modo casuale<sup>3</sup>.

<sup>2</sup> Si tratta di alcune osservazioni puntuali. Tali questioni sono studiate in letteratura nell’ambito più generale della mediazione semiotica; in tale contesto, il foglio elettronico è solo uno degli oggetti che vengono indicati come artefatti. Un riferimento è M. G. Bartolini Bussi, M. G. e M. A. Mariotti, Mediazione semiotica nella didattica della Matematica: artefatti e segni nella tradizione di Vygotskij, in L’Insegnamento della Matematica e delle Scienze Integrate, vol. 32 A-B, pp. 270-294, 2009. L’articolo si può reperire all’indirizzo <http://5.144.189.58/percontare/wp-content/uploads/2014/05/Mediazione-semiotica-nella-didattica-della-matematica-artefatti-e-segni-nella-tradizione-Vygotskij.pdf>

<sup>3</sup> Ciò giustifica l’appellativo di *pseudo casuali*, riservato a tali sequenza. Un efficace quanto elementare approfondimento della questione si trova in Grinstead, Snell, “Introduction to probability”, AMS, 1997 e in Prodi, “Metodi matematici e statistici”, McGraw-Hill, 1992.

- Controllare i risultati e operare consapevolmente.  
*Certo il software presenta l'indiscutibile vantaggio di eseguire in modo automatico molti passi; d'altro canto ciò rende indispensabile l'attivazione di strategie per controllare i risultati così prodotti. Perché allora non cogliere l'occasione per sviluppare un'abilità così importante? Più in generale, è cruciale che gli studenti ripercorrano l'intero procedimento di implementazione, riflettano sui vari passi compiuti e sappiano motivare le scelte ad essi sottese.*
- *Ricordare che le simulazioni offrono delle rappresentazioni della realtà, ma non sono la realtà. Di conseguenza non sostituiscono la manipolazione degli oggetti. Ad esempio, se non si sono mai toccate le facce di un dado, non è facile comprendere come l'equiprobabilità della loro uscita segua da una nostra decisione e non corrisponda ad una presunta verità oggettiva.*

## 3.2 Simuliamo i lanci di dadi

Torniamo al problema guida relativo al lancio di due dadi, discusso nel capitolo 2. Abbiamo concluso che il punteggio (somma dei due numeri usciti su ciascun dado) più probabile è il "7" e abbiamo assegnato all'evento "il punteggio è 7" il valore di probabilità  $\frac{1}{6}$ .

A partire da tali risultati, in questa sezione vogliamo discutere la questione:

**cosa si può prevedere riguardo agli esiti dei lanci di due dadi?**

*Per comprendere a fondo cosa accade nella "pratica", ci si accorge ben presto che non bastano pochi lanci materiali, dai quali, ribadiamo, è comunque significativo partire. Infatti, l'esame di un numero esiguo di esiti<sup>4</sup>, difficilmente permette di scorgere regolarità nella sequenza ottenuta: in tale situazione non accade spesso che la distribuzione delle frequenze relative approssimi quella delle probabilità a priori. Serve dunque effettuare molte prove e, per realizzarle in tempi ragionevoli, diventa praticamente indispensabile ricorrere agli strumenti informatici.*

*Prima di procedere però, precisiamo che nella pratica didattica l'idea di elaborare i dati mediante le frequenze relative e i valori di probabilità non dovrebbe venir calata dall'alto, ma è opportuno costituisca il punto di approdo di una discussione guidata, che procede per passi successivi. In sostanza, essa dovrebbe assumere i connotati di una vera e propria conquista, almeno nella percezione degli studenti.*

*Più in generale, vorremmo articolare ciascuna attività in tre momenti: la discussione comune sulle modalità di realizzazione degli obiettivi<sup>5</sup>, il lavoro individuale al calcolatore e la discussione finale in cui si tracciano collettivamente le conclusioni. Ciascuno di essi contribuisce in maniera fondamentale alla valenza formativa del laboratorio. O riteniamo sia sufficiente che lo studente ci mostri il file realizzato al calcolatore?*

### Attività. Lanci di due dadi con Excel – base

L'attività con il foglio elettronico prevede innanzitutto la registrazione degli esiti di un numero consistente di lanci di due dadi, ad esempio 200, e la loro elaborazione mediante le frequenze relative dei vari punteggi. Si prosegue con l'esame qualitativo di tali frequenze, che si realizza anche attraverso il confronto con le probabilità a priori<sup>6</sup> di ogni punteggio.

*Precisamente, le richieste che, a nostro avviso, ha senso rivolgere agli studenti sono schematizzate in una traccia di lavoro collocata in fondo al capitolo, in appendice (A1 Attività).*

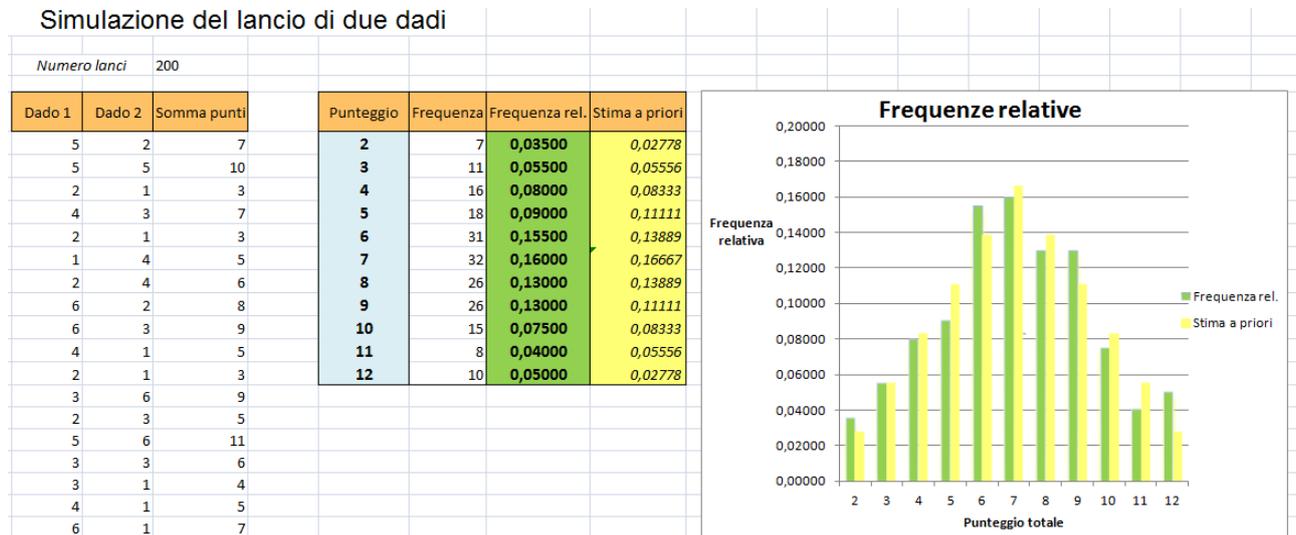
<sup>4</sup> Ottenuti, ad esempio, in 20 prove.

<sup>5</sup> Di queste fanno parte sia gli aspetti più strettamente legati all'implementazione mediante il foglio elettronico sia la strategia da seguire (oggetti matematici da considerare, principali passi algoritmici ...). Anzi, in varie situazioni è didatticamente significativo separare l'esame di tali due tipologie di aspetti, per evitare che l'indagine dello studente si fermi tristemente al nome della funzione Excel da utilizzare ad un dato passo.

<sup>6</sup> Ossia le probabilità valutate mediante lo schema classico.

*Spetta al docente adattare tale schema alle esigenze didattiche e alla situazione della classe, in modo da renderlo un vero e proprio figlio di lavoro<sup>7</sup>.*

La figura che segue mostra una possibile realizzazione dell'attività in Excel ([LancioDueDadi.xlsx](#)). Invece le conclusioni saranno discusse nella sezione successiva.



### Attività. Lanci di due dadi – approfondimento (VBA)

Per investigare più a fondo l'esperimento dei due dadi, si dovrebbero considerare più sequenze costituite da numeri diversi di lanci: da pochi lanci, fino a "molti"<sup>8</sup>. Ora, in tali condizioni, utilizzare il trascinarsi come nell'attività precedente diventa disagiata e difficilmente controllabile. Pertanto preferiamo ricorrere ad applicazioni come **VBA** (Visual Basic for Applications), che consentono di automatizzare ulteriormente la simulazione.

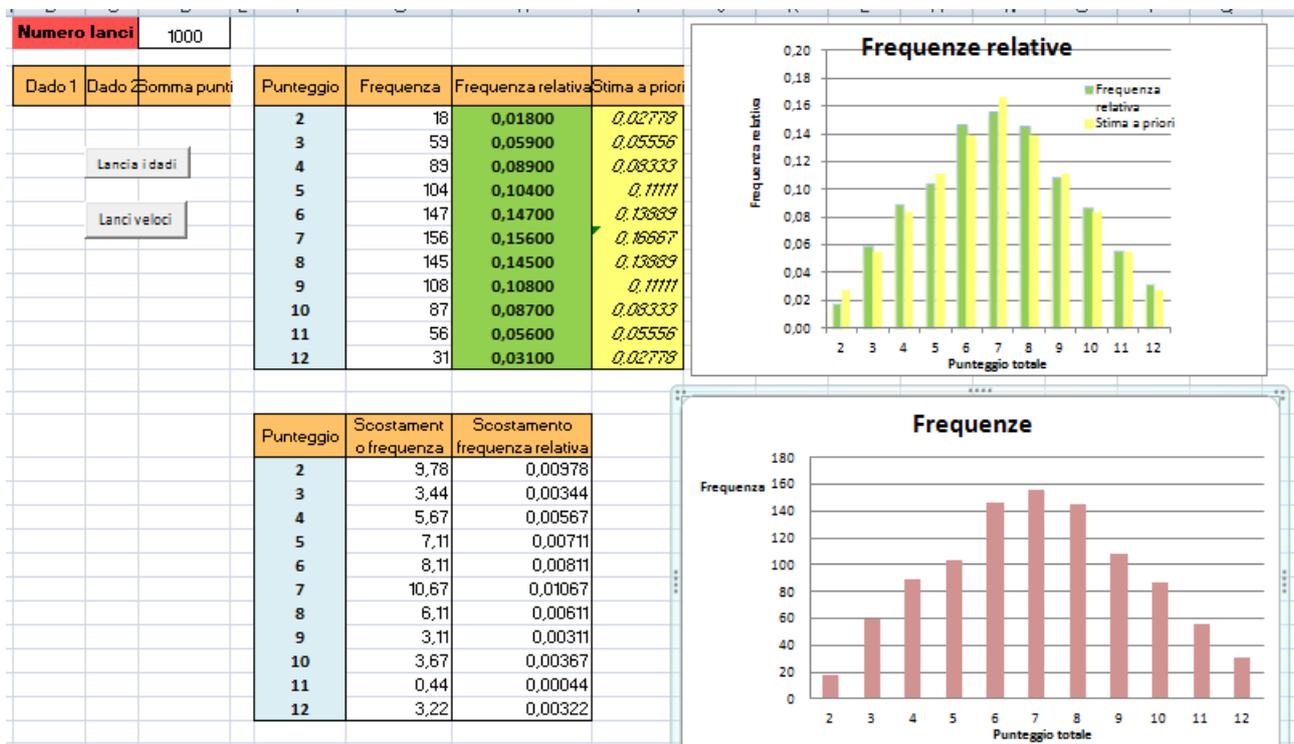
*Ciò nonostante, gli studenti dovrebbero essere in grado di operare consapevolmente, grazie all'esperienza maturata nello svolgimento dell'attività precedente. Precisamente, essi possono operare direttamente sul file **VBA** predisposto dal docente: esaminano l'andamento delle frequenze assolute e relative anche dal punto di vista quantitativo, considerando gli scostamenti<sup>9</sup> tra i valori teorici e quelli osservati. Ancora più nel dettaglio, possono seguire la traccia di lavoro predisposta in fondo al capitolo, in appendice (A2 Attività).*

<sup>7</sup> In tal senso sono da intendersi anche le altre tracce di lavoro che sono proposte in queste note.

<sup>8</sup> Ad esempio, 100, 1.000, 10.000...

<sup>9</sup> Ossia i moduli delle differenze.

La figura che segue mostra una possibile realizzazione dell'attività in Excel ([LancioDueDadi VBA.xlsm](#)); gli studenti possono lavorare sul file, in parte già predisposto, [LancioDueDadiStudenti VBA.xlsm](#).



### Attività. Lanci di un dado

Se si ritiene che l'attività del lancio di due dadi sia troppo delicata per la classe, si può pensare di proporre una del tutto analoga, ma più semplicemente rivolta all'esame dei lanci di un solo dado.

Quest'ultima, d'altra parte, può venir utilizzata come approfondimento per i ragazzi che abbiano già analizzato a fondo il lancio di due dadi; in tal caso però, si richiede loro una maggior autonomia.

Una possibile realizzazione dell'attività in Excel è [LancioDado.xlsx](#), oppure la versione avanzata è [LancioDado VBA.xlsm](#).

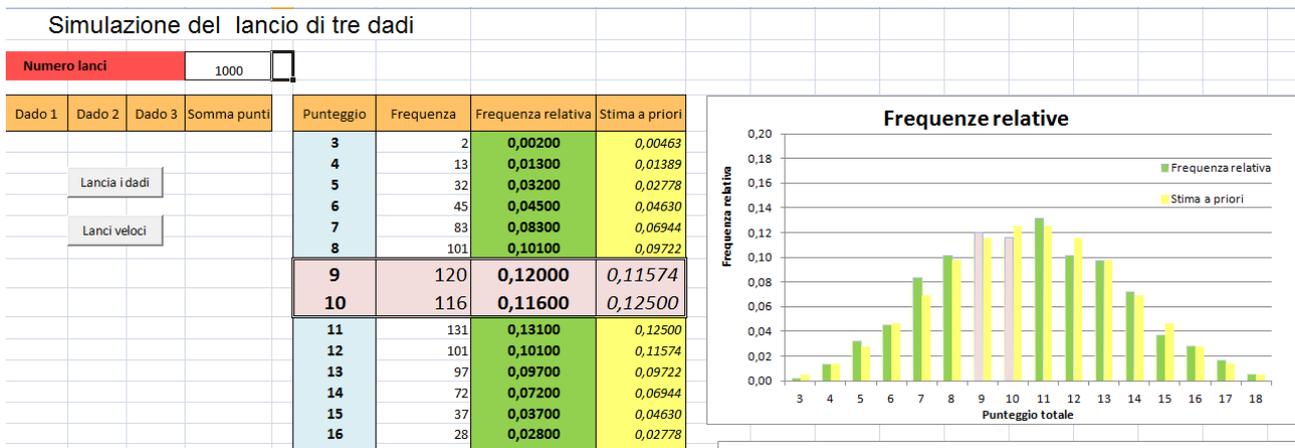
### Attività. Giochiamo alla zara (ovvero lanci di tre dadi) \*

Più articolato da esaminare ma anche più stimolante, è l'esperimento del lancio di tre dadi, che abbiamo già presentato nella veste di gioco della zara.

L'intera questione si può condensare nella richiesta: "conviene puntare sull'uscita del 9 o su quella del 10?"

Gli studenti possono utilizzare il foglio elettronico in due modalità differenti, a seconda che abbiano o meno già risolto la questione per via teorica. Si tratta di verificare la correttezza della risposta (ricordiamo, non di dimostrarla!), oppure di effettuare esperimenti numerici a partire dai quali avventurarsi in congetture sul punteggio più probabile. Naturalmente, se si opta per l'alternativa è opportuno eliminare dal file Excel la colonna delle stime a priori.

La figura seguente mostra una possibile realizzazione dell'attività in Excel ([LancioTreDadi.xlsx](#)), per osservare esiti più espressivi si può utilizzare il file [LancioTreDadi VBA.xlsm](#).



Ricordiamo che la questione è affrontata dal punto di vista teorico da M. Barra, nell'articolo già citato nel capitolo 2<sup>10</sup>.

### 3.3 Conclusioni delle attività di lancio dei dadi

Le attività di simulazione proposte prevedono che gli studenti effettuino varie prove e, sulla base di queste, si misurino con domande che precisano la questione più generale "cosa succede nella pratica?".

In questo contesto l'insegnante riveste un ruolo cruciale. Infatti, egli dovrebbe sollecitare gli studenti a tenere traccia scritta delle risposte: così, da una parte li induce a precisare il proprio pensiero e ad acquisirne consapevolezza; dall'altra, permette di portare alla luce convinzioni che altrimenti rischierebbero di rimanere inesprese. In questo modo il docente dispone degli strumenti per guidare efficacemente il confronto tra le varie congetture formulate dai ragazzi: il sapere si costruisce anche socialmente.

La discussione collettiva dovrebbe condurre a dei risultati analoghi a quelli che seguono. Come già osservato nel testo, abbiamo deciso di esporli ad un livello formale adeguato per lo studente del primo biennio, cioè in una forma che per lui abbia senso, anzi che esprima il senso che attribuisce ad essi e che riconosce come proprio; la controparte è che dobbiamo accettare di perdere la precisione e l'univocità del formalismo matematico.

#### Le differenze

Le differenze  $d$  tra frequenza relativa osservata di ciascun punteggio e la sua stima a priori<sup>11</sup>, in modulo,

- non sempre decrescono al crescere del numero di lanci  $n$
- è "molto" probabile che esse siano "piccole" se  $n$  è "grande"; un po' più precisamente, al crescere di  $n$  è sempre più<sup>12</sup> probabile che si "avvicinino" a zero<sup>13</sup>.

<sup>10</sup> L'articolo si trova anche in rete all'indirizzo:

<http://www.sbai.uniroma1.it/accascina Monti/ssis/linguaggiodelincertezza1/1%20TeorLimiteCentrStampa2.pdf>

<sup>11</sup> Ricordiamo che con stima a priori intendiamo il valore di probabilità dell'evento in questione valutato secondo l'approccio classico (ovvero rapporto tra numero di casi favorevoli e numero di casi possibili).

<sup>12</sup> Questo però non significa che la probabilità tenda a zero in modo monotono.

<sup>13</sup> Analogamente a quanto osservato nel capitolo 2 sezione 5, ci stiamo qui rivolgendo agli studenti del primo biennio della secondaria di secondo grado. In particolare utilizziamo non un linguaggio formale, ma il linguaggio naturale: pertanto se da una parte si guadagna in comprensibilità, dall'altra si deve accettare di perdere in univocità e

Le differenze  $D$  tra frequenza e frequenza teorica<sup>14</sup>, in modulo, al crescere di  $n$  non sempre decrescono. Anzi se  $n$  è “grande”, è “molto” probabile che esse siano “grandi”<sup>15</sup>.

*Per non generare false convinzioni, è opportuno far notare agli studenti che tali risultati non sono dimostrati dagli esiti delle simulazioni effettuate. Infatti essi riguardano specifiche sequenze di lanci, mentre la dimostrazione è relativa alla sequenza generica. Inoltre, ogni sequenza prodotta dal software è necessariamente costituita da un numero finito di lanci, e quindi non fornisce alcuna informazione sull'andamento definitivo delle frequenze.*

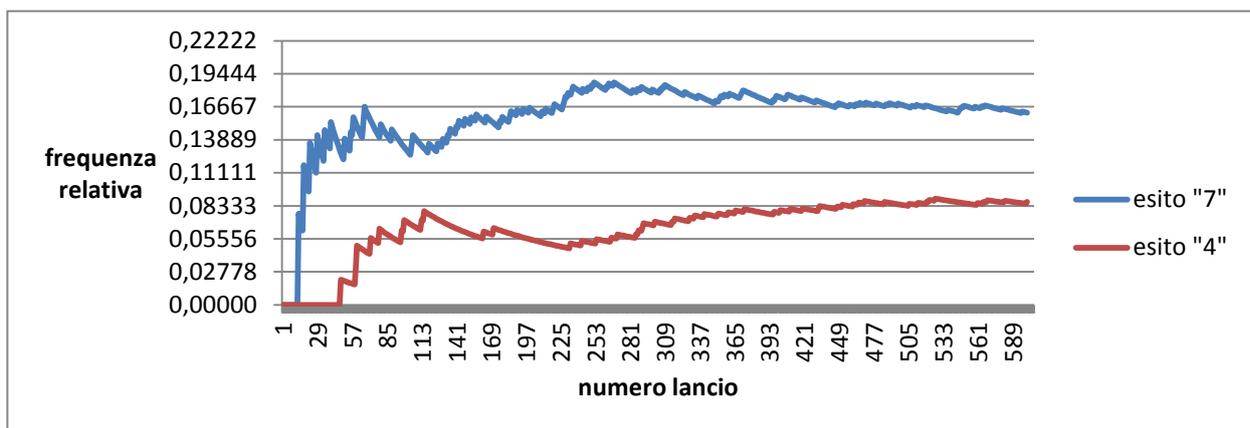
## Le frequenze relative

In altri termini, è vero che non si può prevedere quale punteggio uscirà ad un dato lancio, però si può affermare qualcosa sugli esiti di “molti” lanci:

al crescere del numero di lanci è sempre più probabile che la frequenza relativa di un dato punteggio si “avvicini” al valore di probabilità dell'evento ...

è probabile, ma non è certo che questo accada (ad esempio, è teoricamente possibile che esca il punteggio 12 ad ogni lancio<sup>16</sup>).

Ma più delle parole, un'espressiva rappresentazione dell'andamento delle frequenze relative è suggerito dalla figura seguente<sup>17</sup>. Essa mostra il grafico della frequenza relativa osservata del punteggio “7” (o del “4”), al variare del numero dei lanci, ossia della funzione che al lancio  $m$ -esimo associa la frequenza relativa del punteggio “7” (o del “4”) osservata sui primi  $m$  lanci.



L'immagine è stata ricavata mediante il file [LancioDueDadi Frequenze.xlsx](#), che permette di simulare sequenze costituite da 600 lanci e di visualizzarne graficamente l'elaborazione.

precisione. In ogni caso è importante che gli studenti siano consapevoli di tale “approssimazione”. In questo senso indichiamo tra virgolette i termini “molto”, “piccole”, “grande”, ecc. (ciò vale anche per il seguito della sezione).

<sup>14</sup> Ricordiamo che per frequenza teorica dell'evento  $E$  su un numero  $n$  di prove intendiamo il prodotto  $p(E) \cdot n$ , dove  $p(E)$  è il valore di probabilità dell'evento  $E$  secondo l'approccio classico.

<sup>15</sup> “Grandi” dell'ordine della radice di  $n$ .

<sup>16</sup> Ad esempio consideriamo l'evento “esce il punteggio 7”. Può succedere che esca infinite volte il punteggio 12, pertanto la distanza tra frequenza osservata e valore teorico di probabilità non si avvicina a 0, ma rimane costantemente uguale a  $|0 - \frac{1}{6}|$ .

<sup>17</sup> Ancora una volta, è cruciale la mediazione del docente nel precisare che il grafico non costituisce certo una dimostrazione, semmai fornisce un valido aiuto per formarsi un'idea del risultato. Con queste precisazioni, potremmo essere soddisfatti come docenti se i nostri studenti fossero in grado di abbozzare e commentare criticamente un grafico di questo tipo. Riteniamo che lo sviluppo di tale abilità sia prioritario rispetto a quello di ripetere nel linguaggio verbale le pur significative conclusioni formulate nel testo sull'andamento delle frequenze relative.

### 3.4 Facciamo il punto

Più in generale, fissiamo un esperimento e un evento a esso relativo.

Consideriamo prove ripetute dell'esperimento, indipendenti e condotte nelle "stesse" condizioni. In tali ipotesi:

al crescere del numero di prove è sempre più probabile che la frequenza relativa dell'evento si "avvicini" al suo valore di probabilità ...  
è probabile, ma non è certo che questo accada.

Tale risultato esprime la sostanza della Legge dei grandi numeri, che è una legge teorica ossia si può dimostrare in opportune ipotesi nell'ambito della teoria.

Inoltre il risultato è verificato dall'esperienza e questo fatto viene anche indicato come "legge" empirica del caso. Quest'ultima afferma che ciò che è più probabile in teoria, accade più spesso anche nella pratica.

*Le simulazioni e gli esperimenti descritti conducono in modo naturale ad un ulteriore modo di valutare la probabilità (schema frequentista).*

Si può decidere di assegnare come valore di probabilità ad un evento la frequenza relativa dell'evento, osservata su un "grande" numero di prove. Si assume che le prove avvengano nelle "stesse" condizioni.

### 3.5 E le frequenze assolute? \*

*Nella sezione precedente abbiamo esaminato l'andamento delle frequenze relative. A meno che la questione non risulti troppo articolata per la classe, è interessante esaminare anche un altro andamento: quello delle frequenze assolute. Infatti, il confronto tra l'andamento di questi due oggetti statistici, permette di comprendere più a fondo la sostanza della Legge dei grandi numeri.*

La Legge dei grandi numeri descrive l'andamento delle frequenze relative, all'aumentare del numero di prove.

Le frequenze assolute seguono un andamento analogo?

Per rispondere, è conveniente riesaminare il comportamento delle differenze  $D$ , illustrato<sup>18</sup> nella sezione 3. Esso è coerente con il risultato generale che riportiamo di seguito:

**le frequenze assolute di un evento  $E$ , al crescere del numero  $n$  di prove, non tendono ad "avvicinarsi" alla frequenza teorica  $p(E) \cdot n$ .**

Per fissare le idee, consideriamo una situazione tipica. Lanciando più volte un dado, il numero "2" è uscito:

- 110 volte su 600 lanci:  $|frequenza - valore teorico| = |110 - 100| = 10$
- 1.030 volte su 6.000 lanci:  $|frequenza - valore teorico| = |1.030 - 1.000| = 30$
- 9.934 volte su 60.000 lanci:  $|frequenza - valore teorico| = |9.934 - 10.000| = 66$

<sup>18</sup> Ricordiamo che con  $D$  indichiamo la differenza tra frequenza e frequenza teorica in modulo.

Dunque, gli scostamenti tra frequenze assolute osservate e valori teorici, non solo non sono diminuiti ma sono addirittura aumentati, passando da 10 a 66. In particolare, non c'è stato "recupero" o "compensazione".

Possiamo utilizzare questo argomento per smontare le convinzioni errate di alcuni giocatori incalliti?

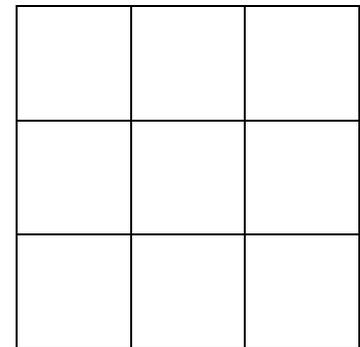
### 3.6 Ancora sull'andamento delle frequenze: un approccio grafico \*

C'è un modo davvero espressivo di visualizzare l'andamento delle frequenze assolute e relative, fondato su rappresentazioni grafiche diverse da quelle discusse nei paragrafi precedenti. L'attività non presenta particolari difficoltà nella fase di schematizzazione, ma richiede un accorto intervento del docente per orientare l'interpretazione degli esiti<sup>19</sup>.

#### La schematizzazione

L'idea è di esaminare più sequenze di lanci di una moneta<sup>20</sup>. Ad esempio, sequenze costituite rispettivamente da 9, 16, 100, 400 lanci; come si vedrà, è conveniente che il numero di lanci di ogni sequenza sia un quadrato perfetto.

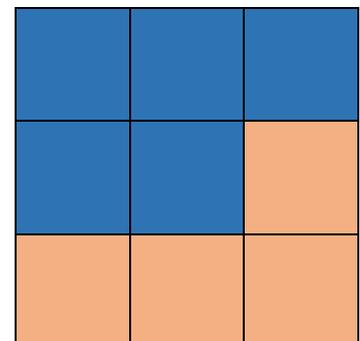
Per iniziare, si predispongono la struttura mediante la quale registrare gli esiti: si associa ad ogni sequenza di  $n$  lanci un quadrato, e lo si suddivide poi in  $n$  quadratini uguali, ciascuno dei quali rappresenta una singola prova.



Si prosegue, effettuando i lanci; ad ogni prova si colora il corrispondente quadratino con un colore diverso a seconda dell'esito. Ad esempio, si può stabilire il criterio:

- quando esce "testa" si colora un quadratino in blu
- quando esce "croce" si colora un quadratino in arancione.

Ecco dunque come si rappresenta una sequenza di 9 lanci, che hanno visto uscire "testa" per 5 volte.



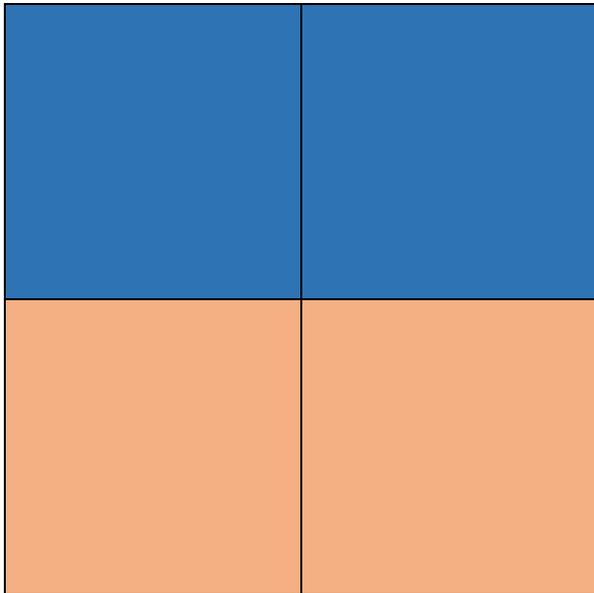
<sup>19</sup> L'attività è ispirata all'articolo di M. Barra, "Probabilità e gioco d'azzardo", in Le scienze e il loro insegnamento, nn.5-6, 2000. Essa è discussa anche sul sito dell'INDIRE, all'indirizzo:

<http://forum.indire.it/repository/working/export/249/5332.htm>

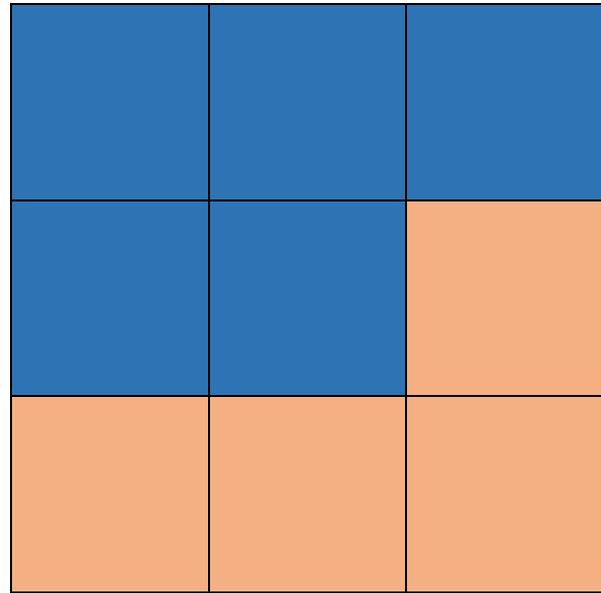
<sup>20</sup> Eventualmente simulati mediante il foglio elettronico.

## Alcune sequenze di esiti

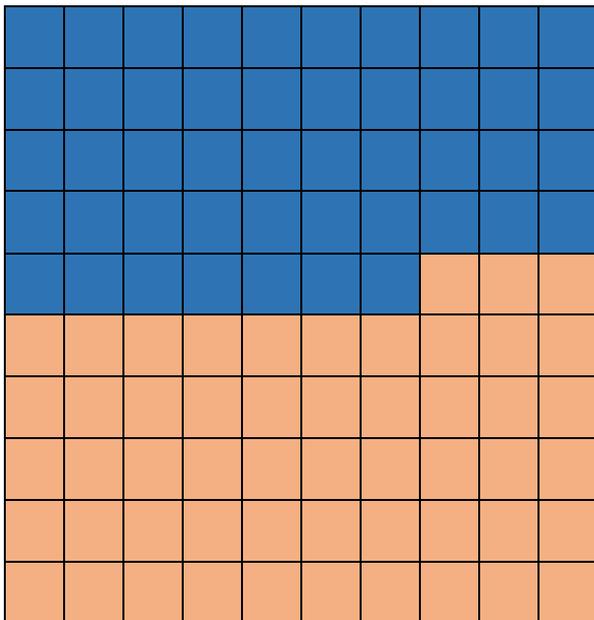
Per meglio comprendere cosa ci si può attendere dall'attività, esaminiamo gli esiti di quattro sequenze di lanci che abbiamo effettuato.



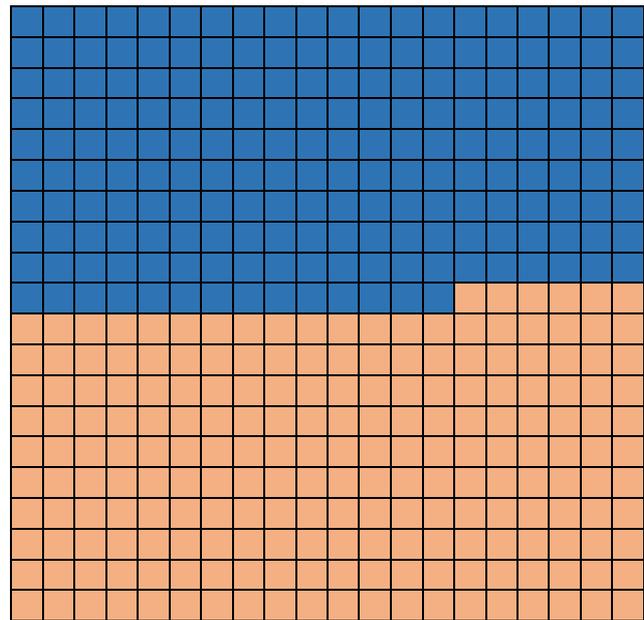
4 lanci - 3 teste



9 lanci - 5 teste



100 lanci - 47 teste



400 lanci - 194 teste

## L'interpretazione degli esiti

Tali rappresentazioni, per quanto suggestive, vanno interpretate con attenzione. Infatti, in primo luogo, esse sono relative a sequenze specifiche, pertanto non costituiscono certo una dimostrazione. Inoltre, offrendo un panorama ristretto ad un numero finito di lanci, non consentono di inferire l'andamento definitivo della sequenza.

Di questi aspetti si dovrà tenere conto nella discussione collettiva. Essa potrà essere articolata in passi successivi, che prevedono le seguenti conclusioni parziali.

- a. Al crescere del numero di lanci ...
- l' **area** degli insiemi blu si "**avvicina**" a metà dell'area del quadrato grande
  - il **numero di quadratini** blu si "**allontana**" dalla metà del numero di quadratini in cui è stato suddiviso il quadrato grande
- b. Per ogni sequenza, l'area dell'insieme blu rispetto al quadrato grande che lo contiene<sup>21</sup>, rappresenta la **frequenza relativa** dell'evento "esce testa" nella sequenza in esame.  
Il **numero di quadratini** blu rappresenta la **frequenza assoluta** dell'evento "esce testa" nella sequenza in esame.
- c. Dunque, nella situazione in esame, cosa possiamo affermare sull'andamento delle frequenze relative? E su quello delle frequenze assolute?

Le risposte a tali quesiti sono le conclusioni illustrate nelle sezioni 3 e 5 del capitolo.

Ossia, al crescere del numero delle prove è sempre più probabile che la frequenza relativa di un evento si "avvicini" alla stima a priori della sua probabilità ... è probabile, ma non è certo che questo accada.

Invece le frequenze assolute di un evento E, al crescere del numero n di prove, non tendono ad "avvicinarsi" alla frequenza teorica  $p(E) \cdot n$ .

### 3.7 Non dimentichiamo le prove con oggetti

*Le simulazioni mediante il foglio elettronico costituiscono un formidabile strumento didattico, ma, come discusso all'inizio del capitolo, non sostituiscono le prove materiali e la manipolazione di oggetti.*

*Come le altre attività operative – sperimentali, esse risultano efficaci quando sono integrate con l'esame teorico delle questioni: teoria e "pratica" sono indissolubilmente legate nella costruzione dei significati degli oggetti matematici.*

#### Attività. Lanci di un dado materiale

Per non impiegare troppo tempo nell'effettuare i lanci, conviene suddividere la classe in gruppi e raccogliere in una tabella le frequenze osservate.

gruppo	1	2	3	4	5	6
I	15	21	13	17	16	14
II	15	15	18	12	22	12
III	14	19	15	15	15	18
IV	18	15	15	19	9	20
V	13	18	20	12	17	16
VI	22	17	13	13	11	19
VII	17	17	18	16	12	16

La tabella riporta gli esiti delle prove effettuate in una scuola secondaria di La Spezia. Come si vede, ogni gruppo ha effettuato 96 lanci, tranne uno che ne ha effettuati 93; così, in poco tempo si sono avuti a disposizione  $16 \cdot 6 \cdot 7 - 3 = 669$  esiti. Vale la pena notare esplicitamente, o meglio decidere assieme agli studenti, che ciò è equivalente a ripetere l'esperimento 669 volte.

<sup>21</sup> Ossia l'area dell'insieme blu nell'unità di misura costituita dal quadrato grande.

Il passo successivo è quello di elaborare i dati raccolti:

frequenza	1	2	3	4	5	6
assoluta	114	122	112	104	102	115
relativa	0,170	0,182	0,167	0,155	0,152	0,172

Alla luce di quanto discusso nelle sezioni precedenti, il docente dovrebbe aver ormai chiaro quale tipo di analisi si può condurre a partire dai valori così calcolati.

### Attività. Lanci di monete

In modo analogo si può esaminare cosa succede lanciando più volte una moneta. Per osservare l'avvicinamento delle frequenze relative al valore teorico atteso, possono bastare 300 lanci complessivi ... probabilmente<sup>22</sup>.

Se si riesce a reperire una moneta truccata, si può proporre una variante dell'esperimento, ancora più ricca e coinvolgente<sup>23</sup>. Si consegnano agli studenti due monete, una delle quali è appunto truccata, e si chiede loro di identificare la moneta truccata, mediante una scrupolosa investigazione degli esiti.

Attenzione però a non esultare troppo presto: per quanti lanci si effettuino, si potrà mai essere certi della correttezza della risposta? La questione è delicata ed è sottesa anche alle attività che seguono.

### Attività. Test di ipotesi

*Le prove ripetute di un esperimento si rivelano efficaci anche per sondare la validità di una congettura. A tale proposito, esaminiamo ora due questioni particolarmente significative, anche per la rilevanza storica che hanno avuto.*

- L'errore di D'Alembert<sup>24</sup>.

Nel lancio di due monete oneste consideriamo i tre esiti possibili: "escono due teste", "escono una testa e una croce", "escono due croci". Essi sono equiprobabili?

Per rispondere, proviamo a lanciare molte volte due monete, proseguiamo registrando gli esiti e elaborandoli come discusso nell'attività dei lanci del dado. Al crescere dei lanci dovrebbe affiorare (probabilmente) qualche dubbio sull'equiprobabilità dei tre eventi.

Naturalmente i risultati sperimentali non costituiscono una dimostrazione, ma diventano un stimolo per investigare le ragioni di tale andamento degli esiti. L'analisi dovrebbe condurre ad approfondire il significato dell'evento "esce una testa e una croce": nel linguaggio naturale in cui tale evento è descritto, non si precisa l'ordine in cui gli esiti "testa", "croce" devono susseguirsi ...

- Il gioco della zara<sup>25</sup>.

Conviene puntare sul punteggio 9 o sul punteggio 10?

Il fatto che gli esiti di prove ripetute possano servire a formulare delle congetture, doveva essere noto già a G. Galilei, secondo il quale "... la lunga osservazione ha fatto dai giocatori stimarsi più vantaggiosi il

<sup>22</sup> Come si sostiene nel testo di Villani e altri "Non solo calcoli", ed. Springer.

<sup>23</sup> L'attività è suggerita nel testo di Villani e altri "Non solo calcoli", ed. Springer.

<sup>24</sup> L'autore dell'Encyclopédie riteneva che nel lancio di due monete, i tre esiti "due teste", "due croci", "una testa e una croce" sono equiprobabili.

<sup>25</sup> Esaminato nel capitolo 2 e anche nelle sezioni precedenti di questo capitolo.

10 che il 9 ...” Un risultato che stupisce anche noi moderni, visto che la differenza tra i due valori di probabilità è minore dell’1%! Per l’esattezza, la differenza è  $\frac{2}{216}$ .

Noi però, a differenza di Galileo, non dobbiamo per forza vestire i panni di giocatori incalliti per effettuare molte prove: possiamo utilizzare il foglio elettronico, come discusso nel paragrafo 3.2.

### **Attività. Un problema “inverso”: ricostruiamo la composizione dell’urna**

Il docente mette in un sacchetto delle biglie uguali al tatto ma di colore diverso, ad esempio alcune blu altre nere, e comunica agli studenti solo il numero totale delle biglie. Lancia così la sfida:

*quante biglie blu prevedete ci siano nel sacchetto?*

Agli studenti è consentito unicamente di estrarre una biglia alla volta, registrarne il colore e reinserirla nel sacchetto.

L’idea vincente per ricostruire la composizione dell’urna è di ricorrere alle frequenze relative osservate nelle varie estrazioni ... Attenzione però: anche in questa situazione è delicato interpretare gli esiti in modo adeguato.

#### **Osservazione**

Per la legge dei grandi numeri, la previsione richiesta potrà essere attendibile (con probabilità “grande”) se si dispone di “molte” osservazioni. Pertanto può essere utile suddividere gli studenti in gruppi, assegnare ad essi sacchetti che hanno la stessa composizione (uno per gruppo) e poi unire gli esiti di tutte le prove effettuate. L’attività si può ulteriormente velocizzare, simulando le estrazioni al calcolatore.

*Per una discussione più accurata della questione si rimanda all’appendice A3, in fondo al capitolo.*

## **3.8 Non solo dadi e monete**

In un certo senso anche le rilevazioni statistiche si possono considerare come sequenze di prove ripetute, una per ciascuna unità statistica. Per chiarire meglio di cosa stiamo parlando rimandiamo ad alcuni esercizi riportati nell’appendice A4 in fondo al capitolo.

*In questo capitolo abbiamo discusso attività di simulazione al calcolatore relative essenzialmente all’andamento delle frequenze assolute e relative in prove ripetute. Torneremo ad utilizzare questo approccio didattico nel capitolo 4 “La probabilità in gioco” quando affronteremo le passeggiate casuali.*

## APPENDICE

### A1 Attività. Simulazione del lancio di due dadi.

Mediante un foglio di calcolo simula il lancio di due dadi:

- registra gli *esiti* dei lanci (ad esempio, 200)
- calcola la **frequenza** osservata e la **frequenza relativa** osservata di ogni punteggio da 2 a 12; inserisci la **probabilità** di ogni punteggio, valutata mediante lo **schema classico** (*stima a priori*)
- realizza una rappresentazione *grafica* delle frequenze relative e delle stime a priori, per *confrontarle* inoltre costruisci un *grafico* che rappresenti le frequenze osservate dei vari punteggi

Una possibile realizzazione dell'attività in Excel è il file [LancioDueDadi.xlsx](#).

Alcune funzioni Excel

- Funzione **CASUALE.TRA(num1;num2)**

restituisce un *numero casuale intero*, maggiore o uguale a num1 e minore o uguale a num2, con distribuzione uniforme<sup>26</sup>.

Un **nuovo numero casuale** viene restituito ogni volta che il foglio di lavoro viene calcolato (ad esempio mediante il **comando F9**).

**Osservazione**

Per generare numeri casuali si può anche utilizzare la funzione **CASUALE()**.

Restituisce un *numero casuale* (distribuito in maniera uniforme) maggiore o uguale a 0 e minore di 1.

In tal caso, per ottenere casuali *interi* basta utilizzare in aggiunta la funzione **INT(num)**, che restituisce la *parte intera* del numero *num* (ossia il numero intero più vicino a num e minore o uguale di num).

- Funzione **CONTA.SE(intervallo;criterio)**

Conta il numero di celle dell'intervallo specificato che soddisfano il criterio indicato.

*Esempi*

**CONTA.SE(D6:D205;2)** si considerano le celle che si trovano nella colonna D dalla posizione 6 alla 205; viene restituito il numero di celle, tra esse, che hanno valore "2".

**CONTA.SE(D6:D205;F6)** si considerano le celle che si trovano nella colonna D dalla posizione 6 alla 205; viene restituito il numero di celle, tra esse, che hanno valore uguale al contenuto della cella F6.

- Funzione **CONTA.NUMERI(intervallo)**

Conta il numero di celle dell'intervallo indicato che contengono numeri.

<sup>26</sup> Ossia ogni numeri intero appartenente all'intervallo indicato ha la stessa probabilità di venir generato.

## A2 Attività. Simulazione del lancio di due dadi - approfondimento

Utilizza il file [LancioDueDadiStudenti\\_VBA.xlsm](#)<sup>27</sup>

### 1) Cosa succede alle frequenze relative dei punteggi se aumentiamo il numero di lanci?

Osserva il **diagramma** delle frequenze relative per  $n = 10, 100, 1.000, 10.000, 100.000 \dots$  lanci. Ripeti più volte i lanci per ogni numero  $n$  considerato.

Confronta le **frequenze relative** osservate con le **stime a priori**: ti sembra che ci sia un *legame* tra loro? Eventualmente prova a descriverlo.

### 2) Proviamo a precisare l'andamento delle frequenze

Effettua  $n = 36, 360, 3.600, 36.000 \dots$  lanci<sup>28</sup>. Ripeti più volte i lanci per ogni numero  $n$  considerato. Determina, relativamente ad ogni punteggio<sup>29</sup>

- la **differenza** tra la **frequenza relativa** osservata e la **stima a priori** della probabilità; sia  $d$  il suo valore assoluto
- la **differenza** tra la **frequenza** osservata e il **prodotto di  $n$  per la stima a priori** della probabilità<sup>30</sup>; indica con  $D$  il suo valore assoluto

a) Considera i diversi valori  $d$  e  $D$  ottenuti nelle *due* prove.

- Per uno stesso  $n$  le differenze  $d$  misurate sono sempre *uguali* in ogni prova? E le differenze  $D$ ?
- Al crescere di  $n$  le differenze  $d$  **decrescono**? Precisa. [Attenzione: per ogni  $n$  vi sono più valori per  $d$ ]
- Al crescere di  $n$  le differenze  $D$  **decrescono**? Precisa.

b) Secondo te, cosa si può **prevedere** riguardo alle frequenze relative nel lancio di due dadi? E riguardo alle frequenze assolute?

Confronta gli esiti e le risposte con quelle dei compagni ... **e con quelle dell'insegnante**.

<sup>27</sup> Gli studenti possono lavorare direttamente su tale file, predisposto dall'insegnante. *Autonomamente* costruiscono la tabella delle differenze.

<sup>28</sup> Tale scelta per il numero di lanci è dovuta al fatto che per tali valori di  $n$  le differenze  $D$  assumono sicuramente valori interi. Infatti i valori di probabilità che intervengono nel calcolo di  $D$  sono frazioni che hanno denominatore 36.

<sup>29</sup> O per uno specifico punteggio, ad esempio il 7.

<sup>30</sup> Il prodotto di  $n$  per la stima a priori della probabilità si può interpretare come *frequenza teorica*. Il significato preciso di tale valore emergerà dalle conclusioni che seguono l'attività.

## A3 Ricostruzione di un'urna

Il docente mette in un sacchetto delle biglie uguali, alcune blu altre nere. Comunica solo il numero totale delle biglie.

Gli studenti estraggono una biglia alla volta, ne registrano il colore e la reinseriscono nel sacchetto.

### Quante biglie blu prevediamo ci siano nel sacchetto?

*Gli studenti fanno delle estrazioni dalle urne e poi registrano gli esiti. Per prima cosa guardano la frequenza relativa del singolo gruppo. Poi mettendo insieme tutti i risultati guardano la frequenza relativa dell'intera classe.*

*In questo modo dovrebbero osservare come la frequenza tende a stabilizzarsi intorno ad un numero. Da questo dovrebbero ricavare la possibile composizione dell'urna.*

*Eventualmente si possono anche effettuare estrazioni al calcolatore, già predisposte dal docente.*

*È delicato **interpretare** gli esiti.*

### **Osservazione**

Quando si valutano casi in cui è possibile decomporre gli eventi nella somma di eventi elementari equiprobabili e incompatibili, la definizione classica di probabilità può incorrere in difficoltà che la rendono inutilizzabile.

Pensiamo ad esempio al problema di stimare la probabilità che domani nevichi. Ragionando secondo la definizione classica di casi favorevoli e casi possibili, mi pongo la questione su quali possano essere i casi elementari equiprobabili e se sia possibile elencarli. Potrei pensare che domani “nevica” o “non nevica”, quindi potrei essere portato a pensare che la probabilità sia 0,5. A questo punto posso introdurre altri possibili scenari meteorologici (soleggiato, nuvoloso, variabile) e dire che la stima della probabilità questa volta è 0,25.

Il problema principale in questo caso risiede nella possibilità o meno di isolare i casi equiprobabili. In conclusione, quando non ho la capacità di valutare quali siano gli eventi elementari equiprobabili, in modo da poter contare casi possibili e casi favorevoli, non posso utilizzare la definizione classica di probabilità. Come la valuto?

In alcune situazioni si può usare l'approccio statistico o frequentista. Esso, come discusso nel capitolo, si fonda sul fatto che all'aumentare delle prove, la frequenza relativa con cui si presenta un evento casuale “si avvicina” ad un valore che si può assumere come probabilità di quell'evento.

## A4 Esercizi

### Problema nascite<sup>31</sup>

- Nel 1989 sono nati in Italia 558.992 bambini, di cui 287.507 maschi e 271.485 femmine. Ne sono nati vivi soltanto 555.686, di cui 285.822 maschi e 269.864 femmine. A partire da questi dati risolvere i seguenti quesiti:**
  - Calcolare la probabilità che un bambino nato nel 1989, scelto a caso, sia un maschio vivo;**
  - Qual è la probabilità che un bambino maschio nato nel 1989, scelto a caso, sia nato vivo?**

a) La probabilità richiesta è

$$p = \frac{\text{Numero maschi nati vivi}}{\text{Numero totale nati}} = \frac{285.822}{558.992} \cong 0,511$$

b) La probabilità richiesta è

$$p = \frac{\text{Numero maschi nati vivi}}{\text{Numero totale maschi}} = \frac{285.822}{287.507} \cong 0,994$$

### Problema decessi<sup>32</sup>

- Esaminare la tabella B, in cui sono riportati i dati relativi al numero dei morti in Italia nel 1988, suddivisi per causa e per età; rispondere ai seguenti quesiti:**
  - Per ogni classe di età valutare la probabilità che la causa di morte sia stata tumore.**
  - Per ogni classe di età valutare la probabilità che la causa di morte sia stata incidente stradale.**
  - Per ogni classe di età valutare la probabilità che aveva un individuo (scelto a caso) di morire per tumore nel 1988. E la probabilità che aveva di morire per incidente stradale.**

#### Tabella B

#### Morti nel 1988<sup>33</sup>

Cause	Classi di età		
	15 - 24	25 - 44	45 - 64
Tumori	663	4.695	43.228
Malattie infettive	38	140	515
Disturbi psichici	492	900	1.867
Altre malattie	1.103	5.846	41.811
Incidenti stradali	2.193	2.113	2.204

<sup>31</sup> Esercizio tratto da Castelnuovo - Gori Giorgi - Valenti, *Matematica oggi 2*, ed. La Nuova Italia, 1992, pag. 820 n. 210. Il testo che proponiamo è stato modificato in parte. Il testo originale è il seguente: "Dei 558.992 bambini, nati in Italia nel 1989, ne sono nati vivi soltanto 555.686, di cui 287.507 maschi e 271.485 femmine. A partire da questi dati risolvere i seguenti quesiti:

- calcolare la probabilità  $p$  di nascita di un maschio vivo
- riprendere dall'esercizio precedente la probabilità  $q$  di nascita di un maschio e calcolare la probabilità  $r$  che un bambino nato maschio sia vivo;
- ..."

<sup>32</sup> Esercizio tratto da Castelnuovo - Gori Giorgi - Valenti, *Matematica oggi 2*, ed. La Nuova Italia, 1992, pag. 820 n. 211. La richiesta c. è stata aggiunta da noi.

<sup>33</sup> Fonte: *Annuario statistico italiano*, ISTAT, 1989 e 1990.

Altri incidenti	1.114	2.600	3.408
Totale morti	5.603	16.294	93.033
<b>Popolazione</b>	<b>9.370.886</b>	<b>16.161.657</b>	<b>13.935.159</b>

a. b. Per rispondere possiamo realizzare una tabella, come la seguente.

Classe di età	Probabilità che la causa sia tumore	Probabilità che la causa sia incidente stradale
15 - 24	$\frac{663}{5.603} \cong 0,118$	$\frac{2.193}{5.603} \cong 0,391$
25 - 44	$\frac{4.695}{16.294} \cong 0,288$	$\frac{2.113}{16.294} \cong 0,130$
45 - 64	$\frac{43.228}{93.033} \cong 0,465$	$\frac{2.204}{93.033} \cong 0,024$

### Osservazione

In questo contesto è interessante porre la questione:

*“è vero che i tumori sono la prima causa di morte?”*

A tale proposito cominciamo con l’osservare che la probabilità che la causa della morte sia un tumore è poco più dell’10% nei più giovani e arriva quasi al 50% negli over 45, mentre la probabilità che la causa sia un incidente stradale è bassa negli over 45 ed è circa il 40% nei più giovani.

Quindi la questione in esame **non ammette una risposta secca** del tipo “si/no”. Va distinta per fasce di età.

c.

Classe di età	Probabilità di morire per tumore	Probabilità di morire per incidente stradale
15 - 24	$\frac{663}{9.370.886} \cong 7,0 \cdot 10^{-5}$	$\frac{2.193}{9.370.886} \cong 2,34 \cdot 10^{-4}$
25 - 44	$\frac{4.695}{16.161.657} \cong 2,9 \cdot 10^{-4}$	$\frac{2.113}{16.161.657} \cong 1,31 \cdot 10^{-4}$
45 - 64	$\frac{43.228}{13.935.159} \cong 3 \cdot 10^{-3}$	$\frac{2.204}{13.935.159} \cong 1,58 \cdot 10^{-4}$

### Osservazioni

- La probabilità di morire per tumore differisce di un ordine di grandezza per le tre fasce di età. Mentre la probabilità di morire per un incidente stradale rimane circa costante.
- E’ importante porre **attenzione nel leggere il testo del quesito**. Infatti i testi delle domande *b* e *c* sono simili, ma non uguali. E rappresentano richieste diverse. Infatti nelle domande *a*, *b* si valuta la *causa* della morte; ci si restringe dunque all’insieme degli individui deceduti nell’anno. Mentre nella domanda *c* si valuta la probabilità di morire; dunque si considerano tutti gli individui in vita in quell’anno.

---

## Errori sul calcolo delle probabilità<sup>34</sup>

### 3. Un giocatore alla roulette dice:

***“Punto sul 10 che non è uscito in questi ultimi quattro lanci, dato che ora la probabilità che esca 10 è aumentata”.***

**Perché il giocatore sbaglia?**

Si tratta di un gioco in cui fisicamente ogni lancio è indipendente dal precedente (a meno che la roulette non sia guasta o truccata). Quindi la probabilità che esca 10 al quinto lancio è uguale a quella che esca 10 a qualsiasi lancio, ad esempio al primo. La roulette non ha memoria.

L'errore è un fraintendimento relativo alla legge dei grandi numeri. E' vero che per essa è “poco” probabile che si verifichi, ad esempio, l'evento: “ il 10 non esce per 9.999 lanci consecutivi”. Ma una volta che esso si è verificato fa parte del passato. In particolare, per quanto prima osservato, non influenza in alcun modo l'esito del 10.000<sup>esimo</sup> lancio.

---

<sup>34</sup> Esempi e problemi tratti da Castelnuovo - Gori Giorgi – Valenti, Matematica oggi 2, ed. La Nuova Italia, 1992, pag. 804 n. 83.