Capitolo 6

Probabilità

6.1 Probabilità

Usiamo i dati provenienti da un campione al fine di trarre conclusioni riguardanti la popolazione da cui il campione è estratto. Per esempio, in un esperimento clinico potremmo osservare che un campione di pazienti a cui è stato somministrato un nuovo farmaco fornisce una risposta migliore rispetto ai pazienti sottoposti al vecchio trattamento. Ciò che vorremmo sapere è se sarà possibile riscontrare tale miglioramento nell'intera popolazione dei pazienti, ed in questo caso valutare l'entità del miglioramento. La teoria della probabilità ci pone nella condizione di mettere in relazione campione e popolazione, e quindi di trarre delle conclusioni in relazione alla popolazione a partire dai campioni a nostra disposizione. Inizieremo la discussione relativa alla probabilità considerando alcuni semplici strumenti che generano fenomeni aleatori, quali monete e dadi; la loro rilevanza risulterà subito evidente anche in relazione ai problemi di ambito medico.

Per prima cosa ci chiediamo cosa si intenda esattamente con il termine "probabilità". In questo libro, adotteremo l'accezione frequentista del termine: la **probabilità** che un evento accada in presenza di determinate circostanze è definita come la frequenza relativa del verificarsi dell'evento quando tende all'infinito il numero di ripetizioni delle circostanze per le quali l'evento potrebbe verificarsi. Per esempio, se lanciamo in aria una moneta, i possibili esiti sono *testa* oppure *croce*. Prima di lanciarla, non abbiamo modo di conoscere l'esito del lancio, ma sappiamo con certezza che esso sarà o testa o croce. Una volta effettuato il lancio, ovviamente, sappiamo con esattezza qual è il risultato ottenuto. Se poi andiamo avanti a lanciare la moneta, otterremo una serie di teste e di croci, esiti dei rispettivi lanci. Proseguendo a lanciare per molte volte, ci aspettiamo di ottenere come esito globale approssivamente lo stesso numero di teste e di croci. Pertanto la probabilità di ottenere "testa" quale esito del lancio verrà quantificato come $1/2$, poiché $1/2$ è la proporzione di volte in cui, sul lungo periodo, si è ottenuto l'esito desiderato ("testa" appunto). Il numero di teste che è possibile ottenere effettuando diversi lanci della moneta è una **variabile aleatoria**, ovvero una variabile che può assumere più di un valore con determinate probabilità. Allo stesso modo un dado, prima di essere lanciato, ha la possibilità di esibire una delle sue sei facce, numerate da 1 a 6, con la medesima probabilità. In questo caso possiamo essere interessati a variabili aleatorie quali "il numero di 6 usciti in un determinato numero di lanci", "il numero di lanci che occorrono fino alla prima uscita del numero 3", e così via. Esiste una definizione alternativa e più generale di probabilità che conduce ad un differente approccio alla statistica, quello che prende il nome di Scuola Bayesiana (Bland e Altman 1998), ma questo esula dagli obiettivi di questo libro.

La definizione frequentista di probabilità si applica anche a misure di tipo continuo, quali l'altezza delle persone. Per esempio, supponiamo che l'altezza mediana di una popolazione di donne sia 168 cm; dunque la metà delle donne è più bassa di 168 cm. Ora, se scegliessimo una donna dalla popolazione in maniera casuale (ovvero in modo che le caratteristiche della donna non influenzino tale scelta) e ripetessimo molte volte questa operazione, nel lungo periodo otterremo che la metà delle donne selezionate sono più basse di 168 cm. La probabilità che una donna sia più bassa di 168 cm è $1/2$. Analogamente, se $1/10$ delle donne di una popolazione ha altezza superiore a 180 cm, una donna scelta a caso all'interno di quella popolazione avrà altezza superiore a 180 cm con probabilità $1/10$. Possiamo anche calcolare la probabilità di avere un'altezza appartenente ad un determinato intervallo di valori. Comunque quando misuriamo una quantità continua, dobbiamo sempre ricordarci che siamo condizionati dai limiti imposti dai metodi con cui si effettuano le misurazioni; pertanto quando diciamo che l'altezza di una donna è 170 cm, intendiamo piuttosto che essa è compresa, per esempio, tra $169,5$ e $170,5$, a seconda dell'accuratezza dello strumento con cui stiamo effettuando la misura. Tutto ciò per dire che spesso siamo interessati alla probabilità che una variabile aleatoria assuma un valore all'interno di certi limiti più che allo specifico valore assunto.

6.2 Proprietà fondamentali

Dalla definizione di probabilità che abbiamo dato al paragrafo precedente discendono le seguenti proprietà:

1. La probabilità è un numero compreso tra 0 e 1.
   Quando un evento non accade mai gli si attribuisce probabilità 0, quando accade sempre 1.

2. **Regola della Somma**. Supponiamo che due eventi siano tra loro *mutuamente esclusivi*, ovvero che il verificarsi dell'uno escluda automaticamente la possibilità che si verifichi l'altro: in questo caso la probabilità che si verifichi l'uno o l'altro dei due eventi è pari alla somma delle probabilità dei singoli eventi. Per esempio, se gettiamo un dado, potranno uscire il numero 1 o il numero 3, ma non entrambi contemporaneamente; la probabilità di

"ottenere 1 o 3" sarà pertanto $1/6 + 1/6 = 2/6 = 1/3$.

3. **Regola del Prodotto**. Supponiamo che due eventi siano tra loro *indipendenti*, ovvero che l'eventuale verificarsi di uno dei due non dica nulla riguardo al verificarsi dell'altro: in questo caso, la probabilità che si verifichino entrambi gli eventi è pari al prodotto delle probabilità dei singoli. Per esempio, supponiamo di lanciare in aria due monete: le due monete non si influenzano vicendevolmente, dunque gli esiti dei due lanci possono essere considerati indipendenti; la probabilità di "ottenere due teste" sarà pertanto $1/2 \times 1/2 = 1/4$. Consideriamo due eventi indipendenti A e B: la proporzione di volte in cui si otterrà A nel lungo periodo corrisponde alla probabilità di A. Dal momento che A e B sono indipendenti, tra tutte le volte in cui accadrà A, ce ne saranno alcune - pari alla proporzione che corrisponde alla probabilità di B - in cui accadrà anche B. Pertanto la probabilità che A e B accadano contemporaneamente è appunto la probabilità di A moltiplicata per la probabilità di B.

6.3 Distribuzioni di probabilità e variabili aleatorie

Supponiamo di considerare un insieme di eventi tra loro mutuamente esclusivi e supponiamo che questo insieme includa tutti gli eventi possibili. La somma delle probabilità degli eventi dell'insieme è 1. L'insieme delle probabilità degli eventi dell'insieme dà luogo ad una **distribuzione di probabiltà**. Per esempio, se consideriamo il lancio di una moneta, i due possibili esiti (testa o croce) sono mutualmente esclusivi e rappresentano tutti e soli gli eventi che si possono verificare. La distribuzione di probabilità che ne deriva è:

$$\begin{aligned} \text{PROB(testa)} &= 1/2, \\ \text{PROB(croce)} &= 1/2. \end{aligned}$$

Definiamo ora una variabile, che denoteremo con il simbolo X, tale che $X = 0$ se esce "croce" e $X = 1$ se invece esce "testa". X è il numero di teste in un singolo lancio della moneta, e può assumere solo i valori 0 o 1. Prima del lancio non sappiamo che valore assumerà X, ma conosciamo con che probabilità essa assume ciascuno dei suoi possibili valori. X è una variabile aleatoria (§6.1) e la distribuzione di probabilità individuata è anche la distribuzione di X. Possiamo rappresentare tale distribuzione con un diagramma come quello in Figura 6.1(a).

Cosa succede se lanciamo due monete contemporaneamente? Sappiamo che ci sono quattro esiti possibili: "testa-testa", "testa-croce", "croce-testa", "croce-croce". Chiaramente ciascuno di questi eventi è equiprobabile e avrà pertanto probabilità 1/4. Sia ora Y il numero di teste. In questo caso, Y potrà assumere 3 diversi valori:

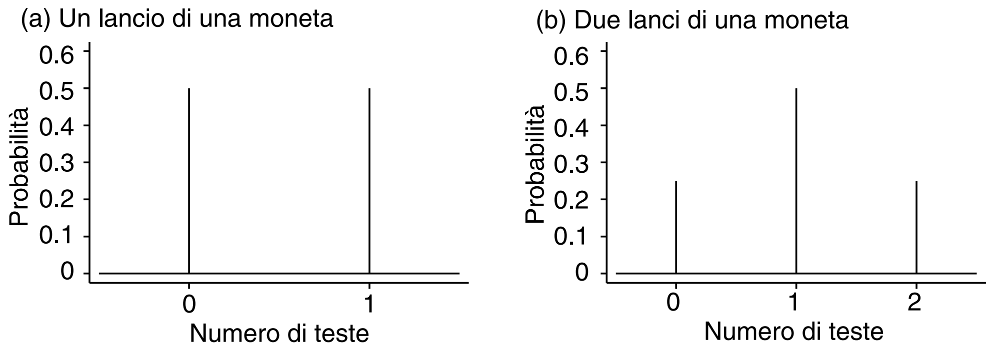

Figura 6.1: Distribuzione di Probabilità per il numero di esiti "testa" nel lancio rispettivamente di una e di due monete

0 se in nessuno dei due lanci uscirà testa (evento che ha probabilità 1/4), 2 se in entrambi lanci uscirà testa (evento che ha, come nel caso precedente, probabilità 1/4), e 1 se in almeno uno dei due lanci uscirà testa (evento che si realizza sia che si ottenga "testa-croce", sia che si ottenga "croce-testa", e che pertanto avrà probabilità $1/4 + 1/4 = 1/2$). La distribuzione di probabilità può essere dunque scritta come segue:

$$\begin{aligned} \text{PROB(Y=0)} &= 1/4, \\ \text{PROB(Y=1)} &= 1/2, \\ \text{PROB(Y=2)} &= 1/4. \end{aligned}$$

La distribuzione di probabilità di Y è mostrata in Figura 6.1(b).

6.4 La distribuzione binomiale

Nel paragrafo precedente abbiamo considerato le distribuzioni di probabilità di due variabili: X, ovvero il numero di teste ottenibili dal lancio di una moneta, che può assumere valore 0 o 1, e Y, ovvero il numero di teste ottenibili dal lancio di due monete, che può assumere invece i valori 0, 1 o 2. In realtà possiamo incrementare il numero di monete a nostro piacimento; in Figura 6.2 è mostrata la distribuzione del numero di teste che si ottiene dal lancio di 15 monete. Non necessariamente la probabilità che caratterizza l'evento di interesse, come il fatto che esca "testa", deve essere 1/2: potremmo essere interessati a contare il numero di "sei" che si ottengono dal lancio di uno o più dadi e il procedimento con cui ricaveremmo i valori della distribuzione di probabilità sarebbe del tutto analogo; in Figura 6.2 è anche raffigurata la distribuzione di probabilità del numero di "sei", esito del lancio di 10 dadi. In generale, possiamo pensare al lancio di una moneta o di un dado come ad un esperimento, il cui esito può essere un successo ("testa" nel caso della moneta, "sei" nel caso del dado) oppure un fallimento (rispettivamente "croce" e uno qualsiasi tra gli altri numeri delle restanti 5 facce). Le distribuzioni mostrate nelle Figure 6.1 e 6.2 sono tutti esempi della distribuzione Binomiale che è molto facile incontrare in ambito medico. La **distribuzione Binomiale** è quella che caratterizza il numero di successi in n prove indipendenti in cui ogni singola prova è contraddistinta da una probabilità di successo pari a p. In effetti la distribuzione Binomiale è una famiglia di distribuzioni, i cui membri sono definiti tramite i valori assunti dai **parametri** n e p.

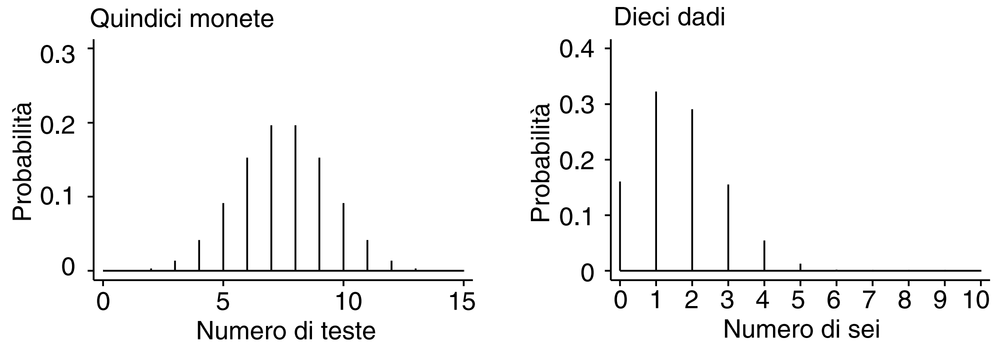

Figura 6.2: Esempi di distribuzione Binomiale: distribuzione del "numero di teste", esito del lancio di 15 monete in contemporanea, e del "numero di sei", esito del lancio di 10 dadi.

Semplici esperimenti aleatori quali quelli che si possono fare con dadi e monete sono interessanti di per sé, ma la loro rilevanza in campo medico non è di immediata evidenza. In ogni caso, supponiamo di eseguire un sondaggio casuale al fine di stimare la prevalenza p (ignota) di una determinata malattia. Dal momento che gli individui del campione sono scelti in maniera casuale dalla popolazione e sono tra loro indipendenti, per ognuno di loro la probabilità di essere malato sarà pari a p. Abbiamo dunque una serie di prove indipendenti (ciascuna con probabilità di successo p), in cui il numero di successi - ovvero il numero di individui appartenenti al campione che presentano la malattia - segue una distribuzione Binomiale. Come vedremo in seguito, le proprietà della distribuzione Binomiale ci consentono di stabilire anche quanto accurata sia la stima della prevalenza ottenuta a partire dal campione (§8.4).

Saremmo in grado di calcolare le probabilità di una distribuzione Binomiale ottenuta, per esempio, dal lancio di 15 monete, elencando tutti gli esiti possibili dell'esperimento. Tuttavia, essendoci $2^{15} = 32768$ combinazioni per gli esiti generati dal lancio delle 15 monete, seguire questa strada non risulterebbe molto conveniente. Esiste invece una formula per calcolare le probabilità a partire dal numero di lanci che si effettuano e dalla probabilità di ottenere, in ognuno di questi lanci, una testa. Questa formula ha validità generale, qualsiasi sia l'esperimento ripetuto n volte in modo indipendente e qualsiasi sia la probabilità p di successo ad ogni ripetizione. La probabilità di ottenere r successi è:

$$\text{PROB(r successi)} = \frac{n!}{r!(n-r)!} p^{\,r} (1-p)^{n-r}$$

dove $n!$, detto "n fattoriale", equivale a $n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$. Questa formula, all'apparenza piuttosto complicata, si spiega invece abbastanza facilmente così: per ognuna delle possibili serie di r successi (ciascuno avvenuto con probabilità p) e di $n-r$ insuccessi (ciascuno avvenuto con probabilità $1-p$), la probabilità che tale serie si verifichi è $p^{\,r}(1-p)^{n-r}$, in quanto ogni prova è indipendente dalle altre e pertanto è possibile applicare la regola del prodotto. Il numero di modi che si hanno per scegliere r oggetti tra n possibili è fornito infine dal coefficiente binomiale $n!/r!(n-r)!$ (Appendice 6A). Solo una combinazione alla volta può accadere; pertanto abbiamo $n!/r!(n-r)!$ modi tra loro mutualmente esclusivi di ottenere r successi in n prove, ciascuno plausibile con probabilità $p^{\,r}(1-p)^{n-r}$. Dunque la probabilità di avere r successi è la somma di queste $n!/r!(n-r)!$ probabilità, generando così la formula sopra riportata. Coloro i quali ricordassero la formula di sviluppo delle potenze di un binomio riconoscerebbero in questa formula uno dei suoi termini, da qui il nome di distribuzione Binomiale.

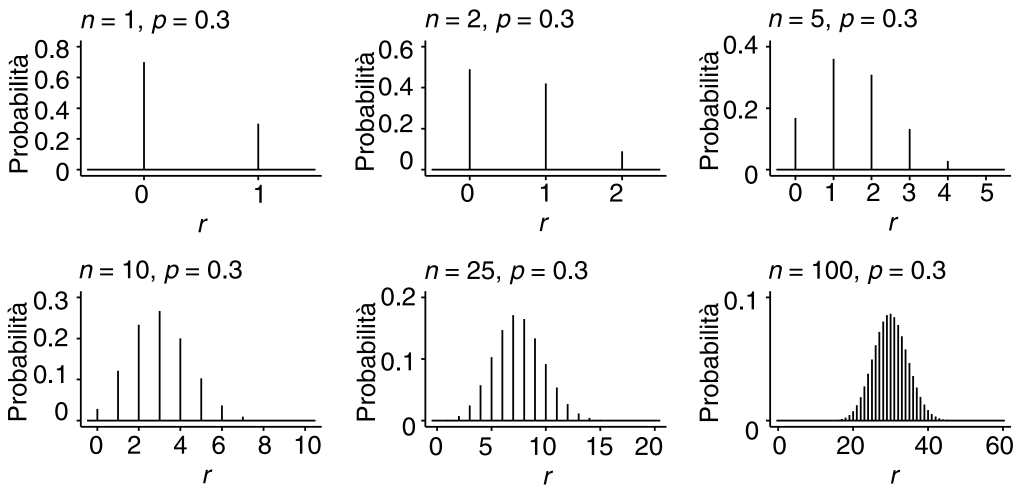

Figura 6.3: Distribuzioni Binomiali al variare di n, con $p = 0.3$

Ora che ne abbiamo compreso il significato, possiamo applicare questa formula al numero di teste che possono essere ottenute nel lancio di due monete. La variabile "numero di teste" seguirà una distribuzione Binomiale con $p = 0.5$ ed $n = 2$, pertanto la probabilità di ottenere $r = 2$ teste sarà:

$$\begin{aligned}
\text{PROB(r = 2)} &= \frac{n!}{r!(n-r)!}p^{\,r}(1-p)^{n-r} \\
&= \frac{2!}{2!0!}0.5^2 \times (0.5)^0 \\
&= \frac{2}{2 \times 1} \times 0.25 \times 1 = 0.25
\end{aligned}$$

Ricordiamo che $0! = 1$ (Appendice 6A), e che qualsiasi numero elevato alla potenza 0 risulta pari a 1.
Analogamente, per $r = 1$ ed $r = 0$ si ha:

$$\text{PROB(r = 1)} = \frac{2!}{1!1!}0.5^1 \times (0.5)^1 = \frac{2}{1 \times 1} \times 0.5 \times (0.5) = 0.5$$

$$\text{PROB}(\text{r} = 0) \;=\; \frac{2!}{0!2!}0.5^0 \times (0.5)^2 = \frac{2}{1 \times 2} \times 1 \times (0.25) = 0.25.$$

che è esattamente quanto avevamo già calcolato per la distribuzione relativa al lancio di due monete nel paragrafo 6.3. Possiamo usare questa distribuzione ogniqualvolta abbiamo a che fare con una serie di prove con due soli esiti possibili. Per esempio, se stiamo trattando un gruppo di pazienti, il numero di quanti guariscono segue una distribuzione Binomiale. Se misuriamo la pressione del sangue in un gruppo di persone scelte a caso il numero di ipertesi segue una distribuzione Binomiale.

La Figura 6.3 mostra il grafico di una distribuzione Binomiale per $p = 0.3$ e valori di n crescenti: essa converge ad una distribuzione Normale, come vedremo nel prossimo capitolo, entrando maggiormente nel dettaglio.

6.5 Media e varianza

Il numero di differenti valori di probabilità in una distribuzione Binomiale può essere molto alto e molto poco pratico da utilizzare. Quando n è grande, si presenta la necessità di riassumere in qualche modo tutte queste probabilità. Come la media e la varianza sono in grado di riassumere le principali caratteristiche di una distribuzione di frequenze, così anche le distribuzioni di probabilità e le variabili aleatorie a loro associate sono descritte da queste quantità.

La **media** è il valor medio che la variabile aleatoria assume nel lungo periodo. È anche detta **valore atteso** e si indica generalmente con il simbolo $\mathrm{E}(X)$. Consideriamo di nuovo il numero di teste ottenibili con il lancio di due monete: sul lungo periodo, otteniamo 0 in 1/4 dei lanci (ovvero con probabilità 1/4), otteniamo 1 nella metà dei lanci (ovvero con probabilità 1/2), infine 2 di nuovo in 1/4 dei lanci. Il valor medio a lungo termine si trova moltiplicando ciascuno dei valori che è possibile ottenere per la rispettiva proporzione di volte in cui ci aspettiamo che quel valore compaia e sommando il tutto:

$$0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 0 + \frac{1}{2} + \frac{1}{2} = 1.$$

Se dunque lanciassimo le due monete ripetutamente il numero medio di teste per ogni lancio sarebbe esattamente 1. Per qualsiasi variabile aleatoria che assuma valori discreti, la media (o valore atteso) si ottiene sommando tutti i possibili valori che la variabile aleatoria in questione può assumere moltiplicati per la rispettiva probabilità.

Si noti che il valore atteso di una variabile aleatoria non deve necessariamente essere uno dei valori che la variabile può assumere. Per esempio, considerando il numero di teste nel lancio di una sola moneta, abbiamo che la variabile aleatoria X può assumere i valori 1 e 0 rispettivamente in caso di successo o insuccesso, ma il suo valore atteso è dato da $0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$: sebbene il numero di teste debba essere necessariamente 0 o 1, il valore atteso non rientra in nessuna di queste due alternative, ma è la media che otterremmo sul lungo periodo.

La **varianza** di una variabile aleatoria è invece il valore atteso del quadrato degli scarti dalla media. Nel caso del numero di teste nel lancio di due monete, 0 e 2 distano dalla media ($\mathrm{E}[Y] = 1$) un'unità e accade con probabilità 1/4, mentre 1 dista dalla media 0 unità e accade con probabilità 1/2. Moltiplicando ciascuno degli scarti elevato al quadrato per la probabilità del rispettivo evento e sommando il tutto si ottiene:

$$\begin{aligned}
\text{Varianza} &= (0-1)^2 \times \frac{1}{4} + (1-1)^2 \times \frac{1}{2} + (2-1)^2 \times \frac{1}{4} \\
&= (-1)^2 \times \frac{1}{4} + 0^2 \times \frac{1}{2} + 1^2 \times \frac{1}{4} \\
&= \frac{1}{2}.
\end{aligned}$$

La varianza di una qualsiasi variabile aleatoria X si denota con $VAR(X)$. In termini matematici

$$\mathrm{VAR(X)} = \mathrm{E}[(X - \mathrm{E}(X))^2]$$

La radice quadrata della varianza è la **deviazione standard** della variabile aleatoria o della distribuzione. Per indicare media e deviazione standard di una distribuzione di probabilità, spesso vengono usate le lettere greche μ ("mu") e σ ("sigma"). Di conseguenza la varianza sarà σ^2. Media e varianza per la distribuzione di variabili continue, di cui parleremo più diffusamente nel capitolo 7, sono definite in modo analogo. Ulteriori dettagli matematici sono in questo caso necessari per poterle esprimere in termini di integrali, ma questo va oltre i nostri attuali obiettivi. In breve, ciò che succede nel caso continuo è sostanzialmente analogo al caso discreto, con la differenza che la scala dei valori della variabile aleatoria continua è suddivisa in intervalli piccolissimi e i valori che la variabile aleatoria assume in questi intervalli sono moltiplicati per la probabilità di ottenere un valore in tali intervalli e quindi sommati. L'integrale nel caso continuo non è che il corrispettivo della somma nel caso discreto.

6.6 Proprietà di media e varianza

Quando usiamo media e varianza delle distribuzioni di probabilità nei calcoli statistici, non sono tanto i dettagli delle loro formule che è necessario sapere, quanto le loro proprietà. La maggior parte delle formule usate nei calcoli statistici, infatti, si ricavano da queste ultime. Risulta abbastanza facile leggere in modo non matematico tali proprietà.

Somma di una costante. Se ad una variabile aleatoria aggiungiamo una costante, la nuova variabile avrà media pari alla precedente più la costante che abbiamo aggiunto, mentre la sua varianza si manterrà inalterata. Supponiamo che la variabile aleatoria in questione sia l'altezza umana: possiamo aggiungere una costante a tale altezza misurando le persone in piedi su di un gradino. L'altezza media delle persone sul gradino risulterà quindi pari alla somma dell'altezza media delle persone più l'altezza del gradino. L'avere aggiunto il gradino, però, non influenza la variabilità dell'altezza: la differenza tra l'individuo più alto e quello più basso, per esempio, rimarrà inalterata. Allo stesso modo, anziché porre le persone su un gradino potremmo misurare la loro altezza dal suolo facendole stare in piedi nel mezzo di una buca. Questo equivarrebbe a sottrarre una quantità costante dalla misura dell'altezza: l'altezza media risulterebbe ridotta esattamente di quella costante, ma di nuovo la variabilità non ne risentirebbe. (Il mio programma Clinstat, liberamente scaricabile come già accennato nel Paragrafo 1.3, possiede un semplice tool grafico in grado di illustrare tutti questi esempi).

Prodotto per una costante. Se moltiplichiamo una variabile aleatoria per una costante positiva, la media e la deviazione standard subiranno un'amplificazione della stessa entità, la varianza invece risulterà moltiplicata per il quadrato della costante. Per esempio, se decidiamo di cambiare l'unità di misura di un campione di osservazioni passando da pollici a centimetri, dobbiamo moltiplicare tutte le misure per 2.54. Questo avrà l'effetto di moltiplicare la media per una costante (2.54 appunto) e analogamente la deviazione standard, dato che ha la stessa unità di misura delle osservazioni. La varianza risulterà invece moltiplicata per il quadrato della costante in questione, coerentemente col fatto che la sua unità di misura è pari al quadrato di quella dei dati. L'operazione di divisione per una costante risulta del tutto analoga. Se la costante in questione è negativa, quando se ne fa il prodotto con la media, si avrà che quest'ultima cambia di segno; al contrario il segno della varianza rimane lo stesso in quanto, essendo moltiplicata per il quadrato della costante, questo sarà comunque un numero positivo. Anche la deviazione standard, essendo la radice quadrata della varianza, è sempre positiva. Essa viene moltiplicata per il valore assoluto della costante, ovvero la costante privata del segno "meno".

Somma di due variabili aleatorie. Se sommiamo tra di loro due variabili aleatorie, la media di tale somma coincide con la somma delle medie; inoltre, *se le variabili in questione sono indipendenti*, la varianza della somma coincide con la somma delle varianze. Possiamo verificare questo fatto misurando l'altezza delle persone poste su gradini di altezza casuale. La media della nuova variabile "altezza degli individui sui gradini" sarà pari all'altezza media degli individui stessi più l'altezza media dei gradini. Anche la varianza in questo caso risulterà maggiore, a causa del fatto che qualche persona di bassa statura si sarà trovata su gradini bassi, mentre qualche persona di alta statura su gradini alti. Se le due variabili non fossero state indipendenti, invece, le cose sarebbero andate diversamente. La media della somma sarebbe comunque rimasta pari alla somma delle medie, ma la varianza della somma non sarebbe coincisa con la somma delle varianze. Supponiamo per esempio che gli individui facenti parte del campione avessero deciso di stare sopra ai gradini non per il capriccio dello statistico che ha progettato l'esperimento, ma per uno scopo preciso: avvitare una lampadina. Questa operazione richiede di essere effettuata ad un'altezza precisa, pertanto le persone devono adattarsi a tale esigenza: le persone più basse dovranno usare gradini più alti per fare ciò che le persone più alte sono in grado di fare anche con gradini bassi. Il risultato della somma di queste due variabili ("altezza degli individui" + "altezza dei gradini"), non più indipendenti in quanto un ben preciso criterio ha guidato ciascun individuo nella scelta del proprio gradino, si rifletterà in una diminuzione di varianza (l'altezza raggiunta sarà pressochè uniforme per tutti, altrimenti la lampadina non potrà essere avvitata) anzichè in un aumento come nel caso precedente. Un'ulteriore possibilità è quella in cui le due variabili non siano indipendenti nel senso che vi siano precise disposizioni dello statistico per cui l'individuo più alto deve prendere il gradino più alto e così via fino al più basso. In questo caso l'esito della non indipendenza si tradurrebbe in un aumento della variabilità. Da queste considerazioni si comprende bene quanto importante sia l'assunzione di indipendenza delle variabili.

Differenza di due variabili aleatorie. Se sottraiamo una variabile aleatoria da un'altra, la media di tale differenza è pari alla differenza delle medie, e, se le due variabili sono indipendenti, la varianza della differenza è pari alla *somma* delle varianze. Supponiamo di misurare l'altezza dal suolo di individui che stanno in piedi nel mezzo di buche di profondità variabile: l'altezza media dal suolo degli individui nelle buche sarà pari all'altezza media degli individui meno la profondità media delle buche, mentre la varianza risulterà maggiore rispetto alla varianza di partenza relativa all'altezza degli individui. Questo perché alcuni individui di bassa statura possono essere capitati in buche molto profonde e al contrario individui di alta statura in buche poco profonde. Se le variabili non fossero indipendenti, cadrebbe l'additività della varianza esattamente come nel caso della somma di due variabili aleatorie non indipendenti: per esempio, se l'obiettivo di ciascun individuo fosse trovare una buca in grado di nasconderlo, riscontreremmo di nuovo una diminuzione di varianza.

Prodotto e divisione di due variabili aleatorie. L'effetto del prodotto tra due variabili aleatorie e della divisione di una variabile per un'altra è qualcosa di molto complesso. Fortunatamente è anche raro incontrare casi in cui sia necessario eseguire tali operazioni.

Le precedenti considerazioni ci mettono in grado di calcolare media e varianza della generica distribuzione Binomiale di parametri n e p. Consideriamo innanzitutto il caso di $n = 1$, per cui si ha:

valore	probabilità
0	1-p
1	p

La media di tale distribuzione sarà pertanto $0 \times (1-p) + 1 \times p = p$, mentre la varianza

$$(0-p)^2 \times (1-p) + (1-p)^2 \times p \quad = \quad p^2(1-p) + p(1-p)^2$$

$$
\begin{aligned}
&= p(1-p)(p+1-p) \\
&= p(1-p)
\end{aligned}
$$

Ora, una variabile aleatoria Binomiale di parametri n e p è la somma di n variabili aleatorie Binomiali indipendenti di parametri 1 e p. Dunque la sua media è la somma di n medie tutte uguali a p, così come la sua varianza la somma di n varianze tutte pari a $p(1-p)$. Pertanto una distribuzione Binomiale di parametri n e p ha media np e varianza $np(1-p)$. Per problemi che coinvolgono campioni molto ampi, queste informazioni sono molto più utili della formula che fornisce le probabilità della distribuzione Binomiale.

Le proprietà della media e della varianza di variabili aleatorie, infine, ci consentono di trovare una soluzione formale ai problemi sorti relativamente ai gradi di libertà della varianza campionaria e discussi nel Capitolo 4. Lo stimatore della varianza desiderato, infatti, deve avere valore atteso pari all'effettiva varianza della popolazione. Si può dimostrare (Appendice 6B) che il valore atteso di $\sum(x_i-\bar{x})^2$ è esattamente $(n-1)\mathrm{VAR}(x)$, pertanto divideremo per $(n-1)$, non per n, al fine di ottenere la stima della varianza.

6.7 La distribuzione di Poisson

La distribuzione Binomiale è solo una tra le moltissime distribuzioni usate in statistica. È una distribuzione discreta, che può assumere un insieme finito di valori, e in particolare è la distribuzione discreta che più di frequente si incontra nelle applicazioni mediche. Tuttavia vale la pena soffermarsi anche su un'altra distribuzione discreta, la distribuzione di Poisson. Sebbene anche la distribuzione di Poisson, così come quella Binomiale, tragga origine da un semplice modello matematico, la sua formalizzazione matematica è più complicata e verrà pertanto omessa.

Supponiamo di essere in presenza di eventi casuali indipendenti nel tempo che si verificano a tasso costante. La distribuzione del numero di eventi che si verificano in un determinato intervallo di tempo è detta **distribuzione di Poisson**. Se il tasso con cui si verificano gli eventi è pari a μ, la probabilità che si verifichino r eventi in ogni unità di tempo sarà pari a

$$\frac{e^{-\mu}\mu^r}{r!}$$

dove $e = 2.718\ldots$. Se l'evento si verifica in maniera casuale ed indipendente rispetto allo spazio anziché rispetto al tempo, la distribuzione di Poisson fornisce la probabilità per il numero di eventi nell'unità di volume o di area.

Raramente si ha l'esigenza di usare le singole probabilità di questa distribuzione, poiché nella maggior parte dei casi conoscerne media e varianza è sufficiente agli scopi che ci prefiggiamo. La media di una Poisson per il numero di eventi nell'unità di tempo è semplicemente il tasso μ, esattamente come la sua varianza. Anche la distribuzione di Poisson, così come la Binomiale, è in realtà una famiglia di distribuzioni, caratterizzate però da un solo parametro (μ appunto). Questa