Concetti fondamentali di statistica descrittiva ed inferenziale

Andrea Onofri
Dipartimento di Scienze Agrarie ed Ambientali
Università degli Studi di Perugia

Indice

1 Descrizione delle osservazioni sperimentali **2**
1.1 Descrizione collettivi 2
1.2 Arrotondamenti 4
1.3 Altre statistiche descrittive 4

2 Modellizzazione dell'errore sperimentale **5**
2.1 Inferenza statistica 7
2.2 Le proprietà delle distribuzioni normali 7
2.3 Variabili standardizzate 9
2.4 La distribuzione delle medie campionarie: errore standard . . 12

3 Presentazione dei risultati **13**

4 Il test d'ipotesi **14**
4.0.1 Esempio 6 14
4.1 Ipotesi alternative semplici e complesse 16

Sommario

Lo scopo di questa lezione è quello di richiamare alcuni concetti fondamentali di statistica descrittiva erd inferenziale, che risultano fondamentali per lo studio della biometria.

1 Descrizione delle osservazioni sperimentali

1.1 Descrizione di un collettivo: analisi chimiche e altre misurazioni fondamentali

Chiunque si occupi di biometria sa che il metodo fondamentale per fronteggiare l'imprecisione degli strumenti di misura è quello di effettuare più repliche della stessa misurazione, in modo che alla fine dell'esperimento ci si ritrova con un collettivo (più o meno piccolo) di valori quantitativi.

In questa comune situazione, la descrizione dei dati sperimentali è affidata ad un indice di tendenza centrale, in genere la media, accompagnato da un indice per descrivere la variabilità dei dati intorno ad essa. **Almeno questi due indici debbono essere sempre presenti quando si riportano risultati sperimentali**.

La media aritmetica è un indicatore di tendenza centrale molto intuitivo e non necessita di particolari spiegazioni: si indica con $\bar{x}$ e si calcola banalmente.

Tuttavia, la media da sola non ci informa su come le unità sperimentali tendono a differire l'una dall'altra: ad esempio una media pari a 100 può essere ottenuta con tre individui che misurano 99, 100 e 101 rispettivamente o con tre individui che misurano 1, 100 e 199. E' evidente che in questo secondo gruppo gli individui sono molto più differenti tra loro (dispersi) che nel primo gruppo.

Pertanto, i risultati di un processo di misurazione non possono essere descritti solo con la media, ma è necessario anche calcolare un indice di variabilità. Tra essi, il più semplice è il *campo di variazione*, che è la differenza tra la misura più bassa e la misura più alta. In realtà, non si tratta di un vero e proprio indice di variabilità, in quanto dipende solo dai termini estremi della distribuzione e non necessariamente cresce al crescere della variabilità degli individui.

Esistono diversi indici di variabilità, tra cui i più diffusi sono la devianza, la varianza, la deviazione standard ed il coefficiente di variabilità.

La **devianza** (generalmente nota come SS, cioè somma dei quadrati) è data da:

$$SS = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Si tratta di un indicatore caratterizzato da significato geometrico molto preciso, collegabile alla somma dei quadrati delle distanze euclidee di ogni os-

servazione rispetto alla media. Come misura di 'distanza', ha alcune importanti proprietà (che vedremo meglio in seguito), ma essendo una somma, il valore finale dipende dal numero di scarti da sommare e quindi non è possibile operare confronti tra collettivi formati da un diverso numero di individui.

Si può quindi definire un altro indice, detto *varianza* (nei software di uso più corrente si parla di varianza campionaria, oppure di quadrato medio, cioè Mean squares = MS), e definito come segue:

$$s^2 = MS = \frac{SS}{n-1}$$

La varianza permette di confrontare la variabilità di collettivi formati da un numero diverso di individui, anche se permane il problema che questo indicatore è espresso in un'unità di misura al quadrato, rispetto a quella delle osservazioni originali: ad esempio se le osservazioni sono espresse in metri, la varianza è espressa in metri quadrati.

Per eliminare questo problema si ricorre alla radice quadrata della varianza, cioè la *deviazione standard*, che si indica con s. La deviazione standard è espressa nella stessa unità di misura dei dati originari ed è quindi molto informativa sulla banda di oscillazione dei dati rispetto alla media.

Spesso la variabilità dei dati è in qualche modo proporzionale alla media: collettivi con una media alta hanno anche una variabilità alta e viceversa. Per questo motivo viene utilizzato spesso il *coefficiente di variabilità*:

$$CV = \frac{s}{\bar{x}} \times 100$$

che è un numero puro e non dipende dall'unità di misura e dall'ampiezza del collettivo, sicché è molto adatto ad esprimere ad esempio l'errore degli strumenti di misura e delle apparecchiature di analisi.

In genere, la deviazione standard, per le sue caratteristiche, viene utilizzata come indicatore dell'incertezza assoluta associata ad una determinata misurazione, mentre il coefficiente di variabilità (incertezza relativa percentuale; CV), è molto adatto ad esprimere l'errore degli strumenti di misura e delle apparecchiature di analisi.

1.2 Arrotondamenti

Come si è visto nell'esempio precedente, il calcolo della media e della deviazione standard (sia a mano che con il computer) porta all'ottenimento

di un numero elevato di cifre decimali. E' quindi lecito chiedersi quante cifre riportare nel riferire i risultati della misura. L'indicazione generale, da prendere con le dovute cautele è che nel caso della media si riportano un numero di cifre decimali pari a quello effettivamente rilevato nella misura (in base alla precisione dello strumento), mentre per gli indicatori di variabilità si dovrebbe utilizzare un decimale in più.

1.3 Altre statistiche descrittive

Quando i collettivi sono sufficientemente numerosi, è possibile definire altre indicatori descrittivi importanti. Il più semplice indicatore di tendenza centrale, utilizzabile con qualunque tipo di dati è la *moda*, cioè il valore della classe che presenta la maggior frequenza. Ovviamente, se la variabile è quantitativa, si assume come moda il punto centrale della classe con maggior frequenza. L'individuazione della moda è banale e non richiede calcoli di sorta.

Nel caso di distribuzioni di frequenza per caratteri ordinabili (qualitativi e quantitativi), oltre alla moda possiamo calcolare la *mediana*, data dal valore che bipartisce la distribuzione di frequenza in modo da lasciare lo stesso numero di termini a sinistra e a destra.

Se abbiamo una serie di individui ordinati in graduatoria, la mediana è data dall'individuo che occupa il posto $(n + 1)/2$ o, se gli individui sono in numero pari, dalla media delle due osservazioni centrali.

La mediana è legata al concetto di ripartizione ed è il primo di una serie di indicatori detti *quantili*, o, se parliamo di frequenze percentuali, *percentili*. Un percentile bipartisce la popolazione normale in modo da lasciare una certa quantità di termini alla sua sinistra e la restante quantità alla sua destra. Ad esempio, il primo percentile bipartisce la popolazione in modo da lasciare a sinistra l' 1% dei termini e alla destra il restante 99%. Allo stesso modo l' ottantesimo percentile bipartisce la popolazione in modo da lasciare a sinistra l' 80% dei termini e alla destra il restante 20% (figura 1).

In relazione all'uso dei percentili, possiamo introdurre il concetto di *boxplot* (grafico Box-Whisker). Si tratta di una scatola che ha per estremi il 25esimo e il 75esimo percentile ed è tagliata da una linea centrale in corrispondenza della mediana. Dalla scatola partono due linee verticali che identificano il valore massimo e il minimo. Se il massimo (o il minimo) distano dalla mediana più di 1.5 volte la differenza tra la mediana stessa e il 75esimo (o 25esimo) percentile, allora le linee verticali si fermano ad un valore pari ad 1.5 volte il 75esimo (o il 25esimo) percentile rispettivamente. La figura 2

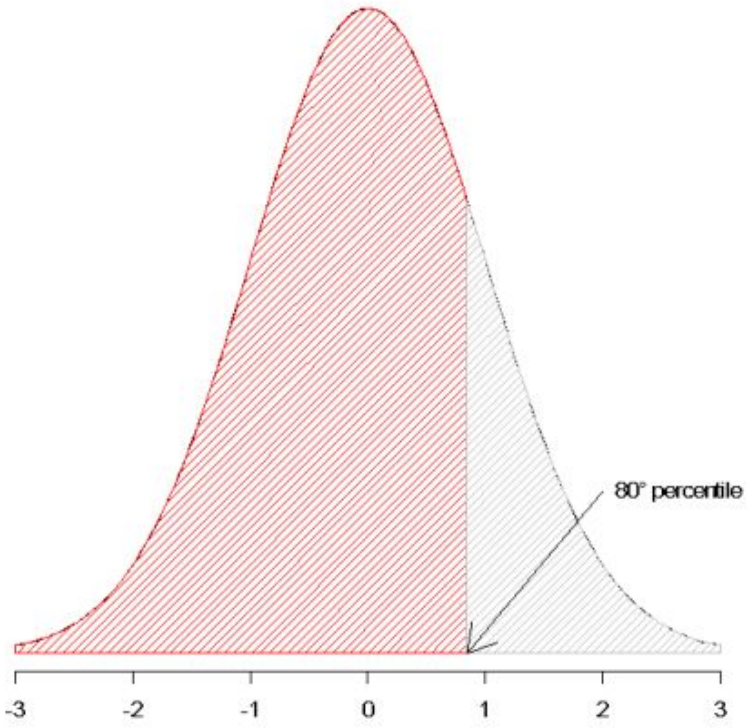

Figura 1: Raffigurazione dell'80esimo percentile

riporta un esempio di boxplot, disegnato per un campione caratterizzato da mediana pari a 167.

2 Modellizzazione dell'errore sperimentale

La descrizione dei risultati di un singolo esperimento/misurazione ha un interesse abbastanza limitato, Infatti, il metodo scientifico ha bisogno di esperimenti ripetibili e dobbiamo quindi essere in grado di prevedere, sulla base di quanto osserviamo, quali potrebbero essere i risultati delle future ripetizioni.

Per far questo assumiamo che l'errore sperimentale segua una certa distribuzione di probabilità, generalmente gaussiana (distribuzione normale).

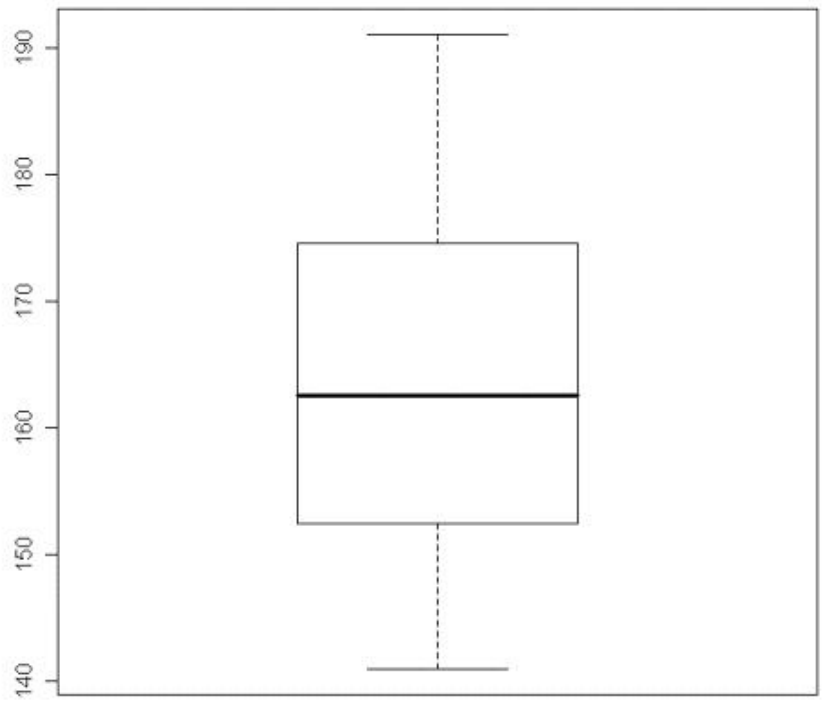

Figura 2: Esempio di boxplot Box-Whisker

In effetti, è ragionevole accettare il fatto che ripetendo la misurazione con uno strumento sufficientemente sensibile e in presenza del solo errore casuale (cioè in assenza di errore sistematico), i risultati tendono a differire tra di loro, muovendosi intorno ad un valore medio, rispetto al quale le misure superiori ed inferiori sono equiprobabili e tendono ad essere più rare, via via che ci si allontana dal valore medio.

Questo andamento 'a campana' può essere descritto con una funzione continua detta **curva di Gauss**:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

ove $P(x)$ è la densità (tralasciamo il significato di questa parola, ma precisiamo che si tratta di una quantità legata alla probabilità) che una certa

quantità x ha di verificarsi, mentre μ e σ sono rispettivamente la media e la deviazione standard della popolazione (per la dimostrazione si rimanda a testi specializzati). Le densità di frequenza che possono essere descritte con la curva di Gauss, prendono il nome di *distribuzioni normali.*

2.1 Inferenza statistica

Immaginiamo di aver fatto un'analisi chimica avendo ottenuto i seguenti risultati: 119, 120 e 121 ng g−1.

1. Descriviamo i risultati considerando che la media è pari a $\bar{x} = 120$ e la deviazione standard è pari ad $s = 1$.
2. Assumiamo che l'esperimento eseguito costituisca una realizzazione casuale, estratta di una popolazione normale, con $\mu = 120$ e σ=1
3. A questo punto siamo in grado di prevedere cosa potrebbe succedere ripetendo l'esperimento: cioè operando una nuova estrazione casuale dalla stessa popolazione.

2.2 Le proprietà delle distribuzioni normali

A questo punto diviene fondamentale conoscere le principali proprietà matematiche della curva di Gauss, che viene a rappresentare la 'replicabilità' degli esperimenti e delle misurazioni. Senza voler entrare troppo in dettaglio, il semplice esame grafico della curva di Gauss consente le seguenti osservazioni:

1. La forma della curva dipende da solo da μ e σ (figure 3 e 4). Ciò significa che, se prendiamo un gruppo di individui e partiamo dal presupposto (assunzione parametrica) che in relazione ad un determinato carattere quantitativo (es. altezza) la distribuzione di frequenza è normale (e quindi può essere descritta con una curva di Gauss), allora basta conoscere la media e la deviazione standard degli individui e immediatamente conosciamo l'intera distribuzione di frequenza;
2. la curva ha due asintoti e tende a 0 quando x tende a infinito. Questo ci dice che se assumiamo che un fenomeno è descrivibile con una curva di Gauss, allora assumiamo che tutte le misure sono possibili, anche se la loro frequenza decresce man mano che ci si allontana dalla media;
3. la probabilità che la x assuma valori compresi in un certo intervallo è data dall'integrale della curva di Gauss in quell'intervallo;
4. Se la curva di Gauss è stata costruita utilizzando le frequenze relative, l'integrale della funzione è uguale ad 1. Infatti la somma delle frequen-

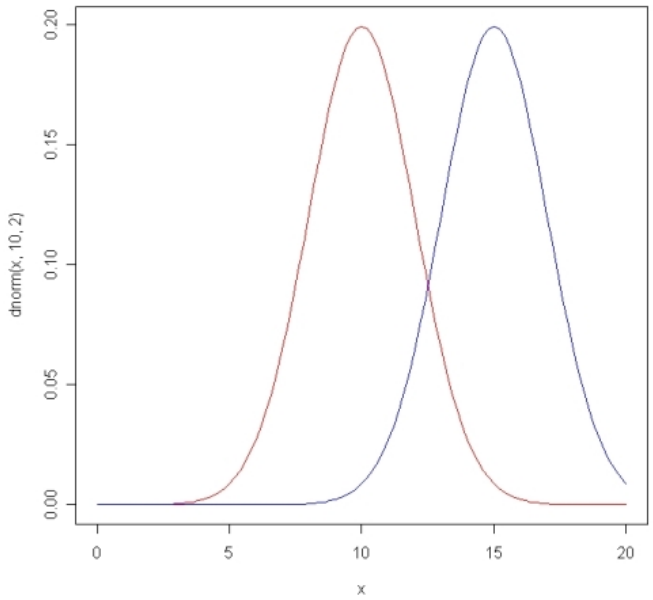

Figura 3: esempio di due distribuzioni normali, con la stessa media e diversa deviazione standard

ze relative di tutte le varianti possibili non può che essere uguale ad 1;

5. la curva è simmetrica. Questo indica che la frequenza dei valori superiori alla media è esattamente uguale alla frequenza dei valori inferiori alla media.

6. dato σ, possiamo dire che la frequenza dei valori superiori a $\mu + \sigma$ è pari al 15.87% ed è uguale alla frequenza dei valori inferiori a $\mu - \sigma$;

Insomma, il calcolo di probabilità per una distribuzione normale equivale al calcolo di un integrale, che viene eseguito numericamente, dato che la funzione di Gauss non ha primitive.

Sempre utilizzando metodi numerici è possibile calcolare i quantili per una distribuzione normale, noti che siano μ e σ.

2.3 Trasformazione e standardizzazione delle variabili

Le distribuzioni normali sono infinite (perché infiniti sono i valori possibili per μ e σ), ma con opportune trasformazioni dei dati possono tutte essere

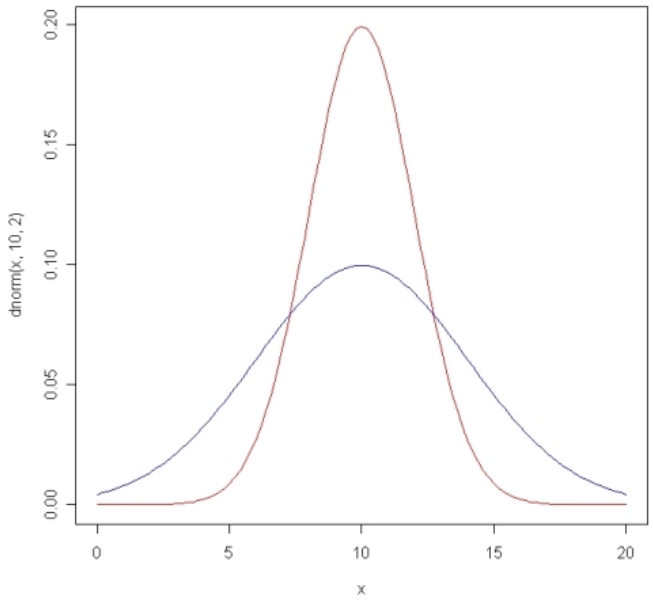

Figura 4: esempio di due distribuzioni normali, con la stessa deviazione standard e diversa media

ricondotte ad una sola distribuzione di riferimento con $\mu = 0$ e $\sigma = 1$, detta **distribuzione normale standardizzata**.

Operare una trasformazione lineare di una popolazione (o comunque un insieme) di dati (misure) significa aggiungere ad ognuno di essi una quantità costante e/o moltiplicare ognuno di essi per una quantità costante. La trasformazione si rilette sul valore della media e della deviazione standard dei dati in modo altamente prevedibile. Trattaremo questo concetto in modo un po'empirico, mentre una trattazione un po' più dettagliata si può trovare nel mio blog, a http://onofriandreapg.blogspot.it/2014/02/the-propagation-of-measurement-errors.html.

Se aggiugiamo a tutti i dati della popolazione un valore A, la media della popolazione trasformata è pari alla media della popolazione originale + A. Lo stesso vale se tutti i dati sono moltiplicati per un numero comune B. In questo caso anche la media è uguale al prodotto della media della popolazione non trasformata per B.

Esercizio 1

Considerate i dati: (a) 12 ; 14 ; 16 ;18 ; 11. La media è pari a 14.2. Se ad ogni dato aggiungiamo il numero 2, otteniamo: (b) 14 ; 16 ; 18 ; 20 ; 13. La nuova media è 16.5. Se invece moltiplichiamo ogni dato per 2, otteniamo: (c) 24 ; 28 ; 32 ; 36 ; 22. La media è 28.4.

Se invece della media consideriamo la variabilità, le trasformazioni addittive (A + X) non hanno alcun effetto ne' sulla varianza ne' sulla deviazione standard, mentre le trasformazioni moltiplicative (B $\times$ X) fanno si che la varianza sia pari a quella originale moltiplicata per B^2, mentre la deviazione standard sia pari a quella originale moltiplicata per B.

Esercizio 2

Considerate i dati dell'esempio precedente. (a) 12 ; 14 ; 16 ;18 ; 11. La deviazione standard è pari a 2.86. Se ad ogni dato aggiungiamo il numero 2, otteniamo: (b) 14 ; 16 ; 18 ; 20 ; 13. La deviazione standard è pari ancora a 2.86 Se invece moltiplichiamo ogni dato per 2, otteniamo: (c) 24 ; 28 ; 32 ; 36 ; 22. La deviazione standard è pari a 5.72.

Ora se prendiamo un insieme di dati (x) calcoliamo la media e la deviazione standard e poi prendiamo ogni dato ci sottraiamo la media e dividiamo il risultato per la deviazione standard, secondo la funzione:

$$z = \frac{x - \mu}{\sigma}$$

otteniamo un insieme di dati trasformati la cui media è zero e la cui deviazione standard è 1.

Esercizio 3

Considerate i dati: (a) 2 ; 5 ; 8; la media è pari a 5, mentre la deviazione standard è pari a 3. Se da ogni dato sottraiamo 5 e dividiamo il risultato per 3, otteniamo la serie: (b) -1 ; 0 ; 1; che ha appunto media 0 e deviazione standard pari ad 1.

In questo modo, qualunque sia la popolazione normale di partenza, possiamo trasformarla in una popolazione normale standardizzata; ciò ci permette di risolvere il problema del calcolo di frequenza o di probabilità semplicemente ricorrendo alle tavole degli integrali della distribuzione normale standardizzata, come ad esempio quella riporta in questo link.

Esercizio 4

Qual è la probabilità che, da un pozzo con un contenuto medio di cloro pari a 1 meq l^{-1}, eseguendo l'analisi con uno strumento caratterizzato da un coefficiente di variabilità pari al 4%, si ottenga una misura pari o superiore a 1.1 meq l^{-1}? Questo problema può essere risolto immaginando che se è vero che il pozzo ha un contenuto medio di 1 meq l^{-1} i contenuti di cloro dei campioni estratti da questo pozzo dovrebbero essere distribuiti normalmente, con media pari ad 1 e deviazione standard pari a 0.04 (si ricordi la definizione di coefficiente di variabilità). Qual è la probabilità di estrarre da questa popolazione una misura pari superiore a 1.1 meq l^{-1}? Se standardizziamo, il problema è equivalente a cercare la probabilità di estrarre una misura pari a (1.1 - 1)/0.04 = 2.5 da una distribuzione normale standardizzata. Le tavole di Z ci dicono che questa probabilità è pari a 0.0062

Esercizio 5

Se ho un pozzo con un contenuto medio di cloro pari a 1 meq l^{-1}, eseguendo l'analisi con uno strumento caratterizzato da un coefficiente di variabilità pari al 4%, qual è l'intervallo (simmetrico rispetto alla media) all'interno del quale è contenuto il 95 % delle misure ottenibili? E il 99%.

Consultando la tabella di Z e tenendo presente che essa rappresenta l'integrale da $-inf$ a Z, dobbiamo trovare quel valore a cui è associata una probabilità del 2.5% (0.025). Notiamo che questo valore è pari ad 1.96. Di conseguenza, nella distribuzione normale standardizzata, esiste un 2.5% di probabilità di campionare valori maggiori di 1.96 e un 2.5% di probabilità di campionare valori inferiori a -1.96 (la distribuzione è simmetrica), quindi esiste un

5% di probabilità di campionare valori fuori dall'intervallo da -1.96 a + 1.96. Se consideriamo:

$$\pm 1.96 = \frac{x - 1.0}{0.04}$$

segue che:

$$x = \pm 1.96 \times 0.4 + 1$$

Cioè l'intervallo va da 0.9216 a 1.0784 meq l^{-1}. Allo stesso modo si può trovare che l'intervallo che contiene il 99% delle misure va da 0.897 a 1.103.

2.4 La distribuzione delle medie campionarie: errore standard

In realtà, come abbiamo già avuto modo di ricordare, noi non eseguiremmo mai una singola analisi, ma ripeteremo la misura almeno due o tre volte, calcolando poi la media. Il problema allora è: esiste una variabile casuale che descrive la distribuzione delle medie di tutti gli infiniti campioni estraibili dalla popolazione anzidetta? Si può dimostrare che, data una popolazione normalmente distribuita con media μ e deviazione standard σ, le medie campionarie sono anch'esse normalmente distribuite con media μ e deviazione standard pari a:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

dove n è la dimensione del campione.

Questa quantità si dice ERRORE STANDARD.

Sempre considerando l'esempio negli esercizi 4 e 5, se dobbiamo fare una determinazione con uno strumento caratterizzato da un errore di misura del 4%, la concentrazione vera è indeterminabile, a causa dell' errore sperimentale, che trasforma i risultati delle analisi in una variabile variabile casuale normale, con media pari alla concentrazione incognita e deviazione standard pari a 0.04. Per ottenere la vera concentrazione del campione l'unico modo sarebbe ripetere infinite analisi. Tuttavia, possiamo considerare anche che, se preleviamo un campione di n individui dalla popolazione in esame (cioè se ripetiamo l'analisi n volte), abbiamo il 95% di probabilità che la media delle n determinazioni effettuate sia compresa tra 0.95 e 1.04. Questi margini

di incertezza si restringono se n aumenta e si annullano quando n diviene infinito.

Questa affermazione, così posta, è contestuale e vale solo per il mio strumento con CV = 4% e una media da determinare pari a 1 ng/l. Se volessimo fare un discorso di validità più generale, potremmo pensare alla standardizzazione delle misure, in modo da avere, qualunque sia la sostanza da analizzare e qualunque sia l'errore di misura dell'apparecchio, una distribuzione delle misure con media pari a 0 e sigma pari ad 1. Di conseguenza la distribuzione delle medie campionarie sarà normale, con media pari a 0 e deviazione standard pari all'inverso della radice del numero dei dati.

Come già detto, nella distribuzione normale standardizzata delle medie campionarie il 95% delle misure è compreso tra -1.96 e +1.96 e, quindi, dato che:

$$-1.96 < \frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} < 1.96$$

e allora:

$$\mu - 1.96 \cdot \frac{\sigma}{\sqrt{n}} < x < \mu + 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

possiamo concludere che il 95% delle misure (medie di n determinazioni) è compreso entro l'intervallo definito dalla media vera più o meno una quantità costante, pari ad un multiplo (1.96 volte) dell'errore standard.

3 Presentazione dei risultati

Dovrebbe essere chiaro che la presenza dell'errore sperimentale ci obbliga a riportare sempre per ogni misurazione un indicatore di tendenza centrale ed un indicatore di variabilità. L'assenza di quest'ultimo non è, in linea di principio, accettabile. Possiamo considerare le seguenti possibilità:

- la media associata alla deviazione standard, per descrivere la variabilità originale del fenomeno in studio;
- la media associata all'errore standard, per descrivere l'incertezza associata alla stima della media;
- la mediana, associata al 25th e 75th percentile, per descrivere dati e fenomeni che non sembrano seguire una distribuzione normale.

4 Il test d'ipotesi

La statistica inferenziale non si pone solo l'obiettivo di comprendere le caratteristiche di una popolazione a partire dai dati raccolti su un campione rappresentativo, ma si pone anche l'obiettivo di verificare ipotesi fatte *a priori* su alcuni aspetti di interesse biologico. Immaginiamo ad esempio di avere una popolazione nota N(μ, σ). Immaginiamo di ipotizzare che su questa popolazione abbia agito un certo trattamento sperimentale che ha spostato la media della popolazione sul valor $\nu > \mu$. Per verificare questa ipotesi, immaginiamo di prendere un campione di n individui e misurare il valore della media, trovandolo pari a $X_m \pm s, maggioredi\mu$. Ci chiediamo: il valore X_m è effettivamente superiore a μ perchè il trattamento ha avuto l'effetto ipotizzato, oppure si tratta di una normale oscillazione legata all'errore di campionamento?

Il test d'ipotesi aiuta a rispondere a domande di questo tipo, permettendo di raggiungere conclusioni su base probabilistica. Il procedimento è il seguente:

1 - Si formula l'ipotesi nulla (corrispondente ad affermare che non vi è effetto) e l'ipotesi alternativa;

2 - Si calcola la probabilità che l'ipotesi nulla sia vera;

3 - Se il livello di probabilità è inferiore ad una certa soglia α prefissata (generalmente 0.05), si rifiuta l'ipotesi nulla e si accetta l'ipotesi alternativa.

Daremo un esempio molto semplice.

4.0.1 Esempio 6

Poniamo di monitorare con uno strumento di analisi la concentrazione di una determinata sostanza in un pozzo. Le misure finora effettuate (molto numerose) mostrano che la concentrazione media è pari a 250 μ g l $^{-1}$, con una deviazione standard pari a 180 μ g l $^{-1}$. Improvvisamente, estraiamo un campione d'acqua e facendo quattro analisi replicate otteniamo un valore medio pari a 350 μ g l $^{-1}$. Possiamo sospettare che il pozzo si è inquinato o si tratta di una normale oscillazione legata all'errore sperimentale?

L'ipotesi è che la concentrazione di nitrati nel pozzo sia descrivibile con una distribuzione di frequenza normale, con media pari a 250 e deviazione standard pari a 180. Di conseguenza le medie campionarie (per il caso di quattro

repliche) sono descrivibili con una distribuzione normale con media pari a 250 e deviazione standard pari a 180/2 = 90.

La media osservata nel campione analizzato è pari a $\bar{X} = 350$.

Poniamo l'ipotesi nulla:

$H_0 : \mu = 250$

che significa che non vi è stato inquinamento e la maggior concentrazione del campione rispetto alla popolazione è imputabile solo al caso (errore di campionamento o di analisi). L'ipotesi alternativa è:

$H_1 : \mu > 250$

cioè il pozzo è inquinato e di conseguenza la sua concentrazione è cambiata rispetto a prima.

Possiamo ora calcolare la probabilità di estrarre da una popolazione di medie campionarie N(250, 90) una media pari o superiore a 350. Se vogliamo fare riferimento alla distribuzione normale standardizzata (che semplifica i calcoli) dobbiamo tener presente che estrarre 350 da una popolazione N(250, 90) equivale ad estrarre (350-250)/90 = 1.111 da una distribuzione normale standardizzata. La probabilità tabellata è pari a 0.1335.

Si tratta quindi di un caso non troppo raro, la cui probabilità è al disopra del livello prefissato, pari ad α=0.05. Per questo motivo possiamo accettare l'ipotesi nulla e concludere che non vi sono elementi per ritenere che il pozzo è inquinato. Si tratta quindi di una sorta di 'assoluzione per insufficienza di prove'.

Se invece avessimo riscontrato una concentrazione pari a 550 μ g l $^{-1}$, la probabilità sarebbe stata (considerare che (550 - 250)/90 = 3.33) pari a 0.0005, cioè trascurabile. In questo caso avremmo rifiutato l'ipotesi nulla, senza però essere totalmente sicuri che la nostra decisione corrisponde alla verità vera. Infatti, anche se abbiamo rifiutato l'ipotesi nulla, l'evento osservato (campionare una media pari a 550) non è impossibile, ma solo molto improbabile (ha lo 0.05% di probabilità di verificarsi). Questo tipo di errore (rifiutare erroneamente l'ipotesi nulla) si dice **errore di**

prima specie e la sua probabilità è pari allo 0.05%, cioè alla probabilità che ha l'ipotesi nulla di essere vera.

Oltre all'errore di prima specie (il più importatante da un punto di vista metodologico), esiste anche il cosiddetto **errore di seconda specie**, che consiste nell'accettare erroneamente l'ipotesi nulla falsa. I due tipi di errore sono esemplificati graficamente nella figura 5).

L'errore di seconda specie è detto β e il suo complemento 1 - β è detto potenza del test, in quanto è la probabilità di mettere in luce le differenze effettivamente esistenti, cioè di rifiutare correttamente l'ipotesi nulla. Un'elevata potenza si ottiene con un basso errore standard, cioè con esperimenti molto precisi (basso σ e/o con molte repliche).

4.1 Ipotesi alternative semplici e complesse

Nel caso precedente abbiamo valutato un'ipotesi alternativa complessa H_1: $\mu > 250$. Se *a priori* non abbiamo elementi per sostenere che $\nu > \mu$, possiamo testare ipotesi alternative semplici, del tipo $\mu \neq 250$. In questo caso, la probabilità di errore α deve essere considerata a destra e a sinistra della media, 2.5% per parte (test a due code), in quanto μ potrebbe essere sia maggiore che minore di 250. Di questo deve essere tenuto conto nel calcolo di probabilità.

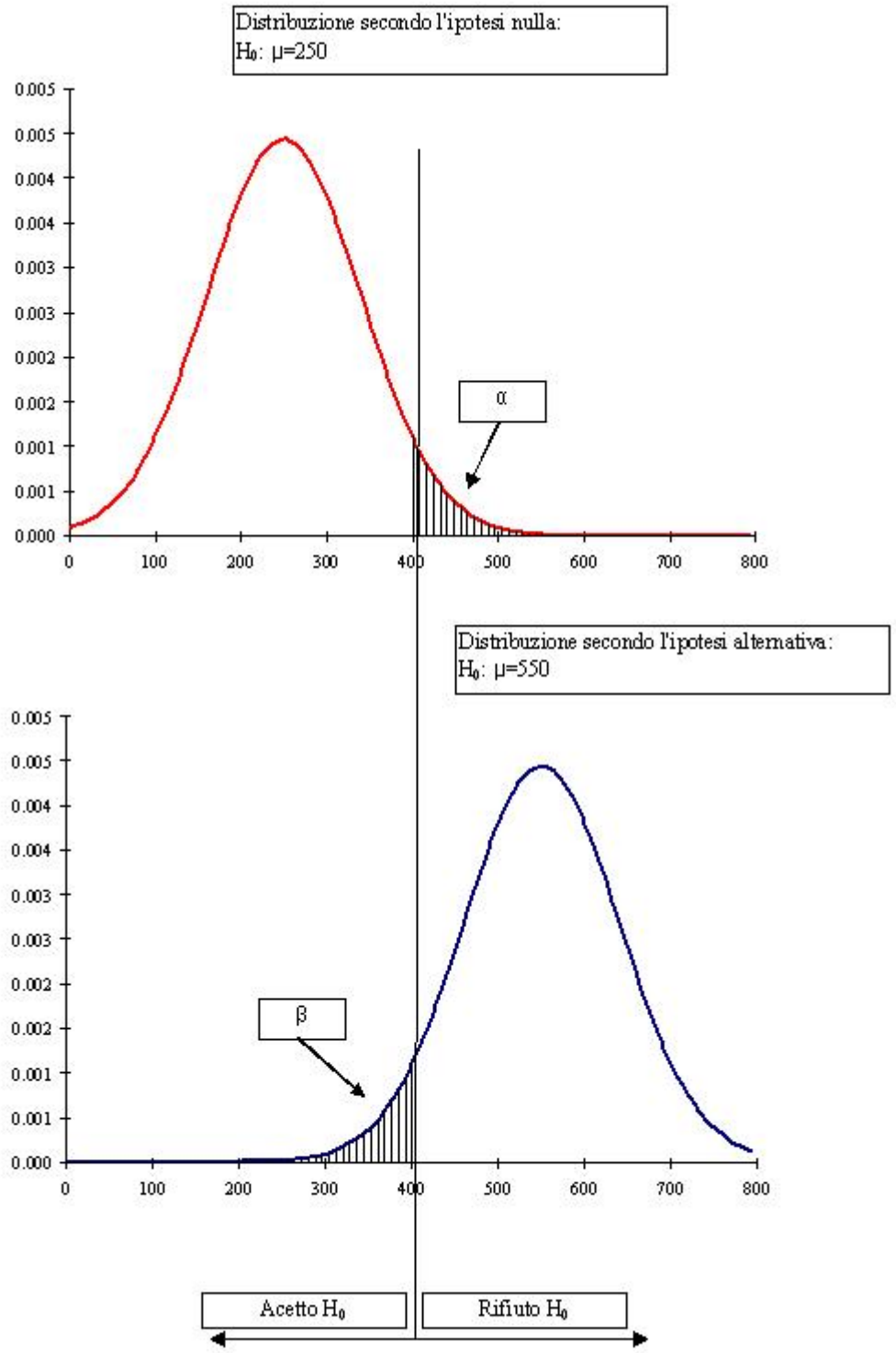

Figura 5: Esemplificazione grafica dell'errore di I e II specie