Azzolini Riccardo 2020-05-19

Test di ipotesi

1 Ipotesi e test

Un tipo di problema comune nell'ambito della statistica è quello in cui si ha un'ipotesi relativa al valore di un parametro di una popolazione (ad esempio la media o la varianza), e si vuole determinare, in base ai dati di un campione, se tale ipotesi sia vera o falsa.

Un'ipotesi del genere è chiamata **ipotesi statistica**, e il procedimento che consente di *rifiutarla* o *accettarla* utilizzando i dati di un campione è detto **test di ipotesi**.

1.1 Esempi

Per illustrare i principali concetti relativi alla verifica delle ipotesi, è utile considerare alcuni esempi (che, per semplicità, considerano tutti la media come parametro della popolazione su cui formulare le ipotesi):

1. Si vuole stabilire se si possa concludere che il tempo medio μ richiesto per svolgere una certa operazione è minore di 30 minuti.

   Si parte dal presupposto che il tempo medio sia invece maggiore o uguale a 30 minuti, che prende il nome di **ipotesi nulla**, H_0. Questa è, solitamente, l'ipotesi che si vuole mostrare falsa.

   L'affermazione che contraddice l'ipotesi nulla, cioè (solitamente) quella che si vuole mostrare essere vera, prende invece il nome di **ipotesi alternativa**, H_1. In questo caso, essa è l'ipotesi che il tempo medio sia minore di 30 minuti.

   Riassumendo, l'ipotesi nulla e l'ipotesi alternativa sono:

   $$\begin{aligned} \mathrm{H}_0 &: \quad \mu \geq 30\,\mathrm{min} \\ \mathrm{H}_1 &: \quad \mu < 30\,\mathrm{min} \end{aligned}$$

2. Il contenuto dichiarato dal produttore delle bottiglie d'acqua minerale di una certa marca è 920 ml. Un'associazione di consumatori sostiene che, in realtà, le bottiglie contengano in media una quantità di acqua μ inferiore.

L'ipotesi che l'associazione di consumatori sostiene, e vuole quindi verificare, è $\mu < 920\,\mathrm{ml}$. Questa è dunque H_1. Allora, come sempre, H_0 è l'ipotesi opposta:

$$\begin{aligned}\mathrm{H}_0 &: \quad \mu \geq 920\,\mathrm{ml} \\ \mathrm{H}_1 &: \quad \mu < 920\,\mathrm{ml}\end{aligned}$$

Infatti, dal punto di vista dell'associazione di consumatori, non sarebbe un problema se il contenuto medio fosse in realtà superiore a quello dichiarato: perciò, non serve verificare che sia esattamente $\mu = 920\,\mathrm{ml}$.

3. Si vuole verificare se le lattine di caffè confezionate automaticamente da una ditta contengono in media il peso dichiarato $\mu = 250\,\mathrm{g}$. A tale scopo, si estrae un campione di 30 lattine, se ne pesa il contenuto, e si calcola il peso medio (cioè la media campionaria), per stabilire se il peso medio della popolazione differisca da $250\,\mathrm{g}$.

   In questo caso, si suppone che il peso medio sia effettivamente quello dichiarato, e si prova a mostrare che invece è diverso. Le ipotesi sono quindi:

$$\begin{aligned}\mathrm{H}_0 &: \quad \mu = 250\,\mathrm{g} \\ \mathrm{H}_1 &: \quad \mu \neq 250\,\mathrm{g}\end{aligned}$$

4. Un ingegnere suggerisce alcune modifiche che si potrebbero apportare a una linea produttiva per aumentare il numero di pezzi prodotti giornalmente. Per decidere se applicare queste modifiche, occorre che i dati sperimentali indichino con forte evidenza che la macchina modificata è più produttiva di quella originaria.

   Indicando con μ_0 il numero medio di pezzi prodotti al giorno prima della modifica, e con μ quello dopo la modifica, si parte dal presupposto che le modifiche non comportino alcun miglioramento, cioè si scelgono le ipotesi:

$$\begin{aligned}\mathrm{H}_0 &: \quad \mu \leq \mu_0 \\ \mathrm{H}_1 &: \quad \mu > \mu_0\end{aligned}$$

2 Possibili conclusioni

È importante sottolineare che, con la verifica delle ipotesi (e, in generale, con l'inferenza statistica, cioè ogni volta che si usano dati campionari per trarre conclusioni sull'intera popolazione), non si arriva alla *dimostrazione* di un'ipotesi: si ha solo un'*indicazione* del fatto che l'ipotesi sia o meno sostenuta dai dati disponibili. Le possibili conclusioni per un test di ipotesi sono infatti le seguenti:

- Se l'ipotesi nulla H_0 è rifiutata, si conclude che l'ipotesi alternativa H_1 è probabilmente vera.

- Se l'ipotesi nulla non è rifiutata, si conclude che i dati non forniscono una sufficiente evidenza per sostenere l'ipotesi alternativa.

  In particolare, in questo caso, non si dovrebbe dire che H_0 viene accettata, ma solo che non viene rifiutata. Questo perché, tipicamente, quando un test di ipotesi non dà informazioni sufficienti a rifiutare l'ipotesi nulla, rimane ancora una probabilità abbastanza elevata che in realtà essa fosse comunque falsa.[1] Ciò nonostante, in pratica si usa spesso – impropriamente – la frase "si accetta l'ipotesi nulla".

2.1 Tipi di errori

Siccome un test di ipotesi fornisce solo un'indicazione riguardo alla validità di un'ipotesi, si corre sempre il rischio di giungere a una conclusione sbagliata. Nello specifico, si distinguono due tipi di errore:

- se H_0 è vera, ma viene erroneamente rifiutata, si commette un **errore di I specie** (o **del I tipo**);
- se invece H_0 è falsa, ma erroneamente non viene rifiutata, si commette un **errore di II specie** (o **del II tipo**).

Per convenzione, si indica con α la probabilità di commettere un errore di I specie, e con β la probabilità di un errore di II specie.

2.1.1 Problema: asciugatura della vernice

Problema: Si vuole sottoporre a test l'affermazione di un produttore di vernici, secondo cui il tempo medio μ di asciugatura di una nuova vernice è non superiore a 20 minuti. A tale scopo, si effettuano 35 prove di verniciatura con la vernice e si calcola il tempo medio di asciugatura, scegliendo di rifiutare l'affermazione del produttore se la media osservata supera il **valore critico** di 20.5 minuti, o di accettarla in caso contrario.

Assumendo che sia noto dall'esperienza che la deviazione standard del tempo di asciugatura della vernice è $\sigma = 2\,\mathrm{min}$, studiare la probabilità di commettere un errore di I specie, ossia la probabilità α che la media del campione superi 20.5 minuti quando invece la media effettiva della popolazione è minore o uguale a 20 minuti.

Soluzione: Innanzitutto, si deduce dal testo del problema che le ipotesi scelte sono:

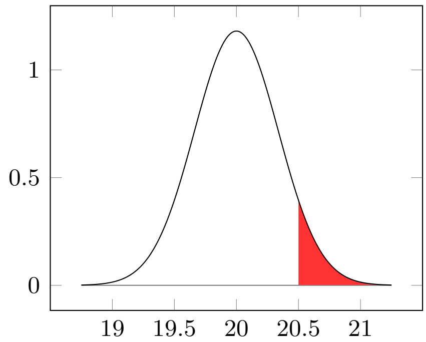
$$\begin{aligned} \mathrm{H}_0 &: \quad \mu \le 20\,\mathrm{min} \\ \mathrm{H}_1 &: \quad \mu > 20\,\mathrm{min} \end{aligned}$$

Inoltre, si sa che è stato deciso di calcolare il valore $\bar{x}$ della media campionaria $\overline{X}$ per un campione di $n = 35$ misurazioni del tempo di asciugatura, e di rifiutare H_0 se $\bar{x} > 20.5$.

[1] Viceversa, quando si rifiuta H_0, la probabilità che invece fosse vera è piuttosto bassa. Questo fatto giustifica la conclusione che sia probabilmente vera H_1.

Per determinare la probabilità di un errore di I specie, bisogna quindi calcolare $P\{\overline{X} > 20.5\} = \alpha$ supponendo che sia vera H_0, cioè che $\mu \leq 20\,\mathrm{min}$.

Siccome $n = 35 \geq 30$, per il teorema del limite centrale la distribuzione della media campionaria può essere considerata approssimabile con una normale. Ciò significa che H_0 viene rifiutata quando il valore di $\overline{X}$ ricade nella *coda* della normale a destra di 20.5:



Dato che la media $\mu = 20\,\mathrm{min}$[2] e la deviazione standard $\sigma = 2\,\mathrm{min}$ della popolazione sono note, si può calcolare la probabilità α (che corrisponde all'area della coda mostrata sopra) mediante standardizzazione:

$$\begin{aligned} P\{\overline{X} > 20.5\} &= P\left\{Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{20.5 - \mu}{\frac{\sigma}{\sqrt{n}}}\right\} \\ &= P\left\{Z = \frac{\overline{X} - 20}{\frac{2}{\sqrt{35}}} > \frac{20.5 - 20}{\frac{2}{\sqrt{35}}}\right\} \\ &= P\left\{Z = \frac{\overline{X} - 20}{0.34} > \frac{0.5}{0.34}\right\} \\ &= P\{Z > 1.47\} \\ &= 1 - P\{Z < 1.47\} \\ &\approx 1 - 0.9292 \\ &= 0.0708 = \alpha \end{aligned}$$

In conclusione, la probabilità di commettere un errore di I specie è $\alpha = 7.08\ \%$.

[2] Assumendo che H_0 sia vera, si ha che $\mu \leq 20\,\mathrm{min}$. Se diminuisce il valore di μ, sicuramente diminuiscono (almeno in media) anche i valori di $\overline{X}$, e quindi la probabilità $P\{\overline{X} > 20.5\}$. Perciò, è sufficiente considerare il "caso peggiore" $\mu = 20\,\mathrm{min}$, con il quale si ha la massima probabilità di un errore di I specie.

2.1.2 Problema: lattine di caffè

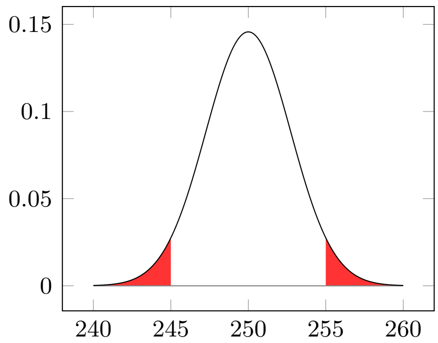
Problema: Si consideri lo scenario delle lattine di caffè presentato in precedenza (l'esempio 3 nella sezione 1.1). Assumendo che la deviazione standard della popolazione sia $\sigma = 15\,\mathrm{g}$, e scegliendo di rifiutare H_0 se la media del campione $\overline{X}$ è minore di $245\,\mathrm{g}$ o maggiore di $255\,\mathrm{g}$, studiare la probabilità α di commettere un errore di I specie, cioè che $\overline{X}$ non sia compresa tra $245\,\mathrm{g}$ e $255\,\mathrm{g}$ quando invece la media effettiva della popolazione è $\mu = 250\,\mathrm{g}$.

Soluzione: Innanzitutto, si ricorda che le ipotesi sono

$$\begin{aligned} \mathrm{H}_0 &: \quad \mu = 250\,\mathrm{g} \\ \mathrm{H}_1 &: \quad \mu \neq 250\,\mathrm{g} \end{aligned}$$

e che si estrae un campione di $n = 30$ lattine.

Come nel problema precedente, siccome $n \geq 30$, la distribuzione di $\overline{X}$ è approssimativamente gaussiana, ma, questa volta, per il rifiuto di H_0 si considerano due code *simmetriche*:



La probabilità α è data dalla somma delle aree delle due code, e, come al solito, può essere calcolata standardizzando

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{30}} \approx 2.74$$
$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} = \frac{\overline{X} - 250}{2.74}$$

e consultando le tavole (dopo aver applicato le opportune simmetrie):

$$\begin{aligned} P\big(\{\overline{X} < 245\} \cup \{\overline{X} > 255\}\big) &= P\left(\left\{Z < \frac{245-250}{2.74}\right\} \cup \left\{Z > \frac{255-250}{2.74}\right\}\right) \\ &= P\big(\{Z < -1.82\} \cup \{Z > 1.82\}\big) \\ &= P\{|Z| > 1.82\} \\ &= 2P\{Z > 1.82\} \qquad \text{(simmetria delle code)} \\ &= 2(1 - P\{Z < 1.82\}) \\ &\approx 2(1 - 0.9656) \\ &= 0.0688 = \alpha \end{aligned}$$

La probabilità di commettere un errore di I specie, ovvero di rifiutare erroneamente l'ipotesi nulla, è dunque $\alpha = 6.88$ %.

3 Terminologia

I principali termini relativi ai test di ipotesi sono i seguenti:

- La probabilità α di commettere un errore di I specie, ossia di rifiutare un'ipotesi nulla vera, è detta **livello di significatività**.
- Il valore $(1-\alpha)\cdot 100$ % è il **grado di fiducia**, già introdotto per gli intervalli di confidenza.
- Se β è la probabilità di commettere un errore di II specie (non rifiutare un'ipotesi nulla falsa), allora la probabilità $1-\beta$ di rifiutare correttamente un'ipotesi nulla falsa prende il nome di **potenza del test**.
- Si dice **statistica test** la statistica (calcolata, come ogni statistica, dai dati del campione) in base alla quale si decide se rifiutare o meno l'ipotesi nulla. Un esempio tipico è la media campionaria.
- Si chiama **regione di rifiuto** l'insieme di valori che, se assunti dalla statistica test, conducono al rifiuto dell'ipotesi nulla. Viceversa, l'insieme di valori della statistica test che portano invece all'accettazione dell'ipotesi nulla è chiamato **regione di accettazione**. Le due regioni sono separate da uno o più valori detti **valori critici**.

  Ad esempio, nel problema precedente, la regione di accettazione è $245 < \overline{X} < 255$, la regione di rifiuto è l'unione di $\overline{X} < 245$ e $\overline{X} > 255$, e i valori critici sono 245 e 255.
- Se la regione di rifiuto è costituita da un singolo intervallo – più precisamente, da una singola coda della distribuzione – il test è **a una coda** (o **unilaterale**). L'altro caso è quello in cui la regione di rifiuto è costituita da due intervalli, ossia

da due code (tipicamente simmetriche) della distribuzione: un test di questo tipo si dice **a due code** (o **bilaterale**).

I test a una coda si usano per ipotesi del tipo

$$\begin{array}{ll} \mathrm{H}_0: & \mu \leq \mu_0 \\ \mathrm{H}_1: & \mu > \mu_0 \end{array} \quad \text{oppure} \quad \begin{array}{ll} \mathrm{H}_0: & \mu \geq \mu_0 \\ \mathrm{H}_1: & \mu < \mu_0 \end{array}$$

mentre quelli a due code si applicano quando le ipotesi sono

$$\begin{array}{ll} \mathrm{H}_0: & \mu = \mu_0 \\ \mathrm{H}_1: & \mu \neq \mu_0 \end{array}$$

4 Schema generale di un test di ipotesi

Nei problemi svolti finora, si sono fissati arbitrariamente i valori critici (ovvero la regione di rifiuto), per poi calcolare il livello di significatività α risultante. Nella pratica, è più tipico svolgere il procedimento inverso, cioè fissare α e calcolare la regione di rifiuto corrispondente. Lo schema generale per lo svolgimento di un test di ipotesi è allora il seguente:

1. Si scelgono l'ipotesi nulla e l'ipotesi alternativa.
2. Si sceglie il livello di significatività α a cui si vuole eseguire il test.
3. In funzione del tipo di test (a una o due code) e dell'α scelto, si determinano i valori critici e la regione di rifiuto.
4. Si sceglie l'ampiezza campionaria, si raccoglie il campione, si calcola dai dati campionari il valore della statistica test, e si vede se esso appartiene alla regione di rifiuto o a quella di accettazione.
5. In base al risultato del passo precedente, si decide se rifiutare o meno l'ipotesi nulla al livello di significatività stabilito.

5 Test di ipotesi sulla media a varianza nota

Si suppone di voler sottoporre a test delle ipotesi sulla media della popolazione, considerando (per ora) il caso più semplice di un test a una coda:

$$\begin{array}{ll} \mathrm{H}_0: & \mu \leq \mu_0 \\ \mathrm{H}_1: & \mu > \mu_0 \end{array}$$

Una volta fissato il livello di significatività α, è necessario trovare il valore critico x_α tale che $P\{\overline{X} > x_\alpha\} = \alpha$.

Se è nota la varianza σ^2 della popolazione, e quindi la deviazione standard σ, e se il campione è sufficientemente grande ($n = 30$), il problema di individuare il valore critico può essere riportato alla distribuzione normale standard:

$$P\{\overline{X} > x_\alpha\} = P\left\{Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > \frac{x_\alpha - \mu_0}{\frac{\sigma}{\sqrt{n}}} = z_\alpha\right\}$$

Si osserva poi che il calcolo del valore critico per la statistica standardizzata Z, cioè dello z_α tale che $P\{Z > z_\alpha\} = \alpha$, dipende esclusivamente dal livello di significatività α (e dal tipo di test – ma, per ora, si sta considerando solo il caso a una coda), e non dai dati della popolazione e del campione.

In altre parole, i valori critici, e quindi le regioni di accettazione e rifiuto, sono determinati univocamente dal livello di significatività (insieme al tipo di test e alla direzione delle disuguaglianze nelle ipotesi). Allora, il test di ipotesi si riduce al confronto tra il valore assunto dalla statistica test

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

per il campione considerato e il valore critico z_α corrispondente al livello di significatività scelto, che può essere ricavato dalle tavole della distribuzione normale standard.

Le regioni di rifiuto per i livelli di significatività più comuni (0.01 e 0.05) sono riassunte nella seguente tabella:

Test	H_0	H_1	α	Valori critici	Regione di rifiuto
una coda	$\mu \le \mu_0$	$\mu > \mu_0$	0.01	2.326	$Z > 2.326$
			0.05	1.645	$Z > 1.645$
una coda	$\mu \ge \mu_0$	$\mu < \mu_0$	0.01	-2.326	$Z < -2.326$
			0.05	-1.645	$Z < -1.645$
due code	$\mu = \mu_0$	$\mu \ne \mu_0$	0.01	$-2.576, 2.576$	$Z < -2.576,\ Z > 2.576$
			0.05	$-1.96, 1.96$	$Z < -1.96,\ Z > 1.96$

5.1 Problema: durata delle lampadine

Problema: Una ditta produttrice di lampadine sostiene che la durata media delle lampadine prodotte è $\mu_0 = 1600\,\text{h}$, con una deviazione standard $\sigma = 120\,\text{h}$. Estraendo un campione di $n = 100$ lampadine, si è calcolata una durata media $\bar{x} = 1570\,\text{h}$. Stabilire se l'affermazione del produttore è corretta, usando come ipotesi alternativa che la durata media sia

1. inferiore a quella dichiarata;
2. diversa da quella dichiarata.

In entrambi i casi, usare i livelli di significatività $\alpha = 0.05$ e $\alpha = 0.01$.

Soluzione: Innanzitutto, si calcola il valore della statistica Z, che dipende solo dal campione, ed è quindi uguale per tutti i casi richiesti dal problema:

$$z = \frac{\overline{x} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} = \frac{1570 - 1600}{\dfrac{120}{\sqrt{100}}} = \frac{-30}{12} = -2.5$$

1. In questo primo caso, le ipotesi sono

$$\begin{aligned} \mathrm{H}_0 &: \quad \mu \geq 1600\,\mathrm{h} \\ \mathrm{H}_1 &: \quad \mu < 1600\,\mathrm{h} \end{aligned}$$

   dunque il test è a una coda, con la regione di rifiuto $Z < z_\alpha$.

   - Per $\alpha = 0.05$, il valore critico è $z_\alpha = -1.645$, e

   $$z = -2.5 < -1.645 = z_\alpha$$

     quindi si rifiuta l'ipotesi nulla al livello di significatività del 5 %.

   - Per $\alpha = 0.01$, il valore critico è $z_\alpha = -2.326$, e

   $$z = -2.5 < -2.326 = z_\alpha$$

     quindi H_0 è rifiutata anche al livello di significatività dell'1 %.

2. Con le ipotesi

$$\begin{aligned} \mathrm{H}_0 &: \quad \mu = 1600\,\mathrm{h} \\ \mathrm{H}_1 &: \quad \mu \neq 1600\,\mathrm{h} \end{aligned}$$

   il test è invece a due code, quindi la regione di rifiuto è composta dagli intervalli $Z < -z_{\alpha/2}$ e $Z > z_{\alpha/2}$.

   - Per $\alpha = 0.05$, si ha $z_{\alpha/2} = 1.96$, e

   $$z = -2.5 < -1.96 = -z_{\alpha/2}$$

     quindi H_0 è rifiutata al livello di significatività del 5 %.

   - Per $\alpha = 0.01$, che corrisponde a $z_{\alpha/2} = 2.576$, il valore della statistica ricade nella regione di accettazione,

   $$-z_{\alpha/2} = -2.576 < z = -2.5 < z_{\alpha/2} = 2.576$$

     quindi *non* si può rifiutare l'ipotesi nulla al livello di significatività dell'1 %.

5.2 Problema: precisione di una macchina

Problema: La precisione di una macchina che produce componenti di dimensioni specificate viene controllata con periodiche verifiche a campione. La dimensione media richiesta è $\mu_0 = 3.5\,\mathrm{mm}$, e si ha una varianza $\sigma^2 = 0.22\,\mathrm{mm}^2$.

1. Valutare se il processo sia da ritenersi sotto controllo oppure no, quando la media riscontrata su un campione di 60 pezzi è $\bar{x} = 3.42\,\mathrm{mm}$.
2. Ripetere la valutazione nel caso che il campione sia di 150 pezzi, con media $\bar{x} = 3.41\,\mathrm{mm}$.

Soluzione: Si sceglie l'ipotesi nulla che il processo sia sotto controllo, mentre l'ipotesi alternativa è che il processi sia fuori controllo, e necessiti quindi di un qualche intervento correttivo:

$$\begin{aligned} \mathrm{H}_0: &\quad \mu = 3.5\,\mathrm{mm} \\ \mathrm{H}_1: &\quad \mu \neq 3.5\,\mathrm{mm} \end{aligned}$$

Date queste ipotesi, il test effettuato sarà a due code.

Il testo del problema non specifica il livello di significatività da considerare. In una situazione reale, α dovrebbe essere scelto valutando gli errori che si rischia di commettere:

- un errore di I specie, che ha probabilità α, significherebbe considerare il processo fuori controllo quando in realtà non lo è, ovvero comporterebbe un intervento correttivo non necessario;
- un errore di II specie, la cui probabilità è β, corrisponderebbe a considerare sotto controllo un processo che invece produce componenti fuori specifica, a scapito degli acquirenti di tali componenti.

A parità di dimensione del campione, una riduzione di α porta a un aumento di β (e viceversa), quindi si può ridurre la probabilità di uno solo dei due tipi di errori:

- se un intervento correttivo fosse molto costoso, bisognerebbe essere ben sicuri che sia effettivamente necessario, e allora si sceglierebbe un valore di α piccolo;
- ponendosi dal punto di vista degli acquirenti, se si volesse essere sicuri di cogliere uno spostamento anche piccolo dalla media specificata, si dovrebbe rendere basso β, mediante la scelta di un valore più elevato di α.

Qui, a scopo illustrativo, si scelgono $\alpha = 0.05$ e $\alpha = 0.01$.

1. Nel primo caso si hanno $n = 60$ e $\bar{x} = 3.42\,\mathrm{mm}$, da cui segue che il valore della statistica Z è:

$$z = \frac{\bar{x} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} = \frac{3.42 - 3.5}{\sqrt{\dfrac{0.22}{60}}} \approx -1.32$$

- Per $\alpha = 0.05$, la regione di rifiuto è $Z < -1.96$ e $Z > 1.96$, ma

$$-1.96 < z = -1.32 < 1.96$$

quindi l'ipotesi nulla non viene rifiutata al livello di significatività del 5 %, e il processo si ritiene dunque sotto controllo.

- Per $\alpha = 0.01$, si considera la regione di rifiuto $Z < -2.576$ e $Z > 2.576$:

$$-2.576 < z = -1.32 < 2.576$$

quindi H_0 non viene rifiutata al livello di significatività dell'1 %, ovvero il processo si ritiene ancora sotto controllo.

2. Prendendo invece il campione con $n = 150$ e $\bar{x} = 3.41\,\text{mm}$, la statistica Z assume il valore:

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{3.41 - 3.5}{\sqrt{\frac{0.22}{150}}} \approx -2.35$$

- Per $\alpha = 0.05$, con la regione di rifiuto $Z < -1.96$ e $Z > 1.96$, avviene che

$$z = -2.35 < -1.96$$

cioè si rifiuta l'ipotesi nulla al livello di significatività del 5 %, ritenendo che il processo sia fuori controllo.

- Per $\alpha = 0.01$, corrispondente alla regione di rifiuto $Z < -2.576$ e $Z > 2.576$,

$$-2.576 < z = -2.35 < 2.576$$

dunque l'ipotesi nulla non viene rifiutata al livello di significatività dell'1 %: il processo si ritiene sotto controllo.

6 Test di ipotesi sulla media a varianza incognita

Se la varianza è incognita, e il campione non è abbastanza grande da poterla approssimare con la varianza campionaria, ma la popolazione da cui proviene il campione ha una distribuzione normale, nei calcoli per i test di ipotesi sulla media si usa la distribuzione t di Student al posto della normale (esattamente come per la stima degli intervalli di confidenza).

In particolare, la statistica test diventa

$$T = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

(dove S è la deviazione standard campionaria), che è una variabile aleatoria avente la distribuzione t con grado di libertà $\nu = n-1$. Di conseguenza, i valori critici t_α o $t_{\alpha/2}$ non dipendono più solo dal livello di significatività, ma anche dall'ampiezza del campione, ed è quindi necessario leggerli ogni volta dalle tavole.

Come per il caso a varianza nota, si può realizzare una tabella riassuntiva, ma qui non è pratico riportare i valori critici in forma numerica (dato che essi dipendono appunto da n):

Test	H_0	H_1	α	Valori critici	Regione di rifiuto
una coda	$\mu \le \mu_0$	$\mu > \mu_0$	0.01	$t_{0.01}$	$Z > t_{0.01}$
			0.05	$t_{0.05}$	$Z > t_{0.05}$
una coda	$\mu \ge \mu_0$	$\mu < \mu_0$	0.01	$-t_{0.01}$	$Z < -t_{0.01}$
			0.05	$-t_{0.05}$	$Z < -t_{0.05}$
due code	$\mu = \mu_0$	$\mu \ne \mu_0$	0.01	$-t_{0.005}, t_{0.005}$	$Z < -t_{0.005}$, $Z > t_{0.005}$
			0.05	$-t_{0.025}, t_{0.025}$	$Z < -t_{0.025}$, $Z > t_{0.025}$

6.1 Problema: carico di rottura

Problema: Una prova del carico di rottura di 6 cavi d'acciaio costruiti da una ditta ha mostrato un carico di rottura medio $\bar{x} = 7750\,\mathrm{kg}$ e uno scarto quadratico medio $s = 145\,\mathrm{kg}$, mentre il costruttore afferma che il carico di rottura medio è di $8000\,\mathrm{kg}$. È possibile sostenere che l'affermazione del costruttore non è corretta, e che il carico di rottura medio è in realtà inferiore, ai livelli di significatività del 5 % e dell'1 %?

Soluzione: Le ipotesi sono

$$\begin{aligned} H_0&: \quad \mu \ge 8000\,\mathrm{kg} \\ H_1&: \quad \mu < 8000\,\mathrm{kg} \end{aligned}$$

perciò il test è a una coda. Il valore della statistica T è

$$t = \frac{\bar{x} - \mu_0}{\dfrac{s}{\sqrt{n}}} = \frac{7750 - 8000}{\dfrac{145}{\sqrt{6}}} \approx -4.22$$

e il grado di libertà è $\nu = n - 1 = 5$.

- Per il livello di significatività $\alpha = 0.05$, il valore critico è $t_\alpha = -t_{0.05} = -2.015$ (dalle tavole, per $\nu = 5$), ovvero la regione di rifiuto è $T < -2.015$, e

  $$t = -4.22 < -2.015$$

  quindi si rifiuta H_0 al livello di significatività del 5 %.
- Invece, per $\alpha = 0.01$, si ha il valore critico $t_\alpha = -t_{0.01} = -3.365$, che definisce la regione di rifiuto $T < -3.365$, e ancora

  $$t = -4.22 < -3.365$$

  quindi l'ipotesi nulla è rifiutata anche al livello di significatività dell'1 %.

7 Schema di applicazione delle statistiche test

Il seguente schema riassume quale sia la statistica da considerare per un test di ipotesi sulla media, a seconda delle caratteristiche del campione e della popolazione:

- Se $n \geq 30$, e la varianza σ^2 della popolazione è *nota*, si usa la statistica test

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

  che segue (approssimativamente) la distribuzione normale standard.

- Se $n \geq 30$, ma la varianza σ^2 della popolazione è *incognita*, si usa la statistica test

$$Z = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

  che segue (approssimativamente) la distribuzione normale standard.

- Se $n < 30$, ma la popolazione è *normale* e ha varianza σ^2 *nota*, si usa la statistica test

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

  che segue la distribuzione normale standard.

- Se $n < 30$, e la popolazione è *normale* con varianza σ^2 *incognita*, si usa la statistica test

$$T = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

  che segue la distribuzione t di Student con grado di libertà $\nu = n - 1$.