della numerosità campionaria il suo valore o realizzazione tende al valore del parametro ignoto della popolazione. Ciò è possibile solo se lo stimatore T è consistente in media quadratica ovvero se tende a zero l'errore quadratico, ossia se:

$$\lim_{x\to\infty} ESQ(T_n) = \lim_{x\to\infty} (T_n - \mu)^n \tag{2.6.9}$$

Esiste inoltre la consistenza in probabilità quando è verificata la condizione al limite:

$$\lim_{x\to\infty} P[|T_n - \mu| < a] = 1 \tag{2.6.10}$$

dove $a > 0$ e T_n è lo stimatore funzione delle realizzazioni campionarie

2.6.1 Stima intervallare

Il modello inferenziale tende a superare il grosso limite della stima puntuale attraverso l'inserimento di più stime accettabili a cui viene associato un predefinito livello di affidabilità (confidenza) e, quindi, a determinare un cosiddetto intervallo di valori a cui possa appartenere, ad un certo livello di fiducia, il valore del parametro ignoto. Il concetto di stima intervallare (o per intervallo) può essere esplicitato con un esempio. Siano μ il valore medio o atteso e σ la deviazione standard della distribuzione campionaria di uno stimatore T che si distribuisce approssimativamente come una Normale (condizione che si verifica di larga massima se la numerosità del campione è maggiore di 30). A questo punto si può affermare che è più probabile che la stessa distribuzione di T sia compresa negli intervalli:

$$\begin{aligned} &\text{da } \mu - \sigma_T \text{ a } \mu + \sigma_T; \\ &\text{da } \mu - 2\sigma_T \text{ a } \mu + 2\sigma_T; \\ &\text{da } \mu - 3\sigma_T \text{ a } \mu + 3\sigma_T \end{aligned} \tag{2.6.1.1}$$

rispettivamente circa il $68,27\%$, il $95,45\%$ e il $99,73\%$ delle volte.
Questi valori rappresentano gli intervalli fiduciari rispettivamente al 68,27%, al 95,45% e al 99,73% per la stima di μ, cioè l'intervallo di valori che permettono di stimare il parametro incognito della popolazione nel caso di uno stimatore T corretto. La percentuale di fiducia è spesso detta livello di fiducia. Nella **Tabella 2.6.1.1** si riportano i valori critici della v.c. normale standardizzata z (detti anche coefficienti fiduciari) corrispondenti ai vari livelli di fiducia o di confidenza $(1-\alpha)$ e di significatività α

Tabella 2.6.1.1
Livelli di fiducia

$1-\alpha$	$0,9973$	$0,99$	$0,96$	$0,9545$	$0,95$	$0,90$	$0,6827$	$0,50$
α	$0,0027$	$0,01$	$0,04$	$0,0455$	$0,05$	$0,10$	$0,3173$	$0,50$
z critica	$3,00$	$2,58$	$2,05$	$2,00$	$1,96$	$1,645$	$1,00$	$0,6745$

Legenda: α = Livello di significatività; $1-\alpha$ = Livello di confidenza; z critica = Quantile della v.c. normale standardizzata

2.6.1.1 Stima intervallare della Media di una popolazione con varianza nota

Si vuole esprimere in termini formali la costruzione di un intervallo di stime accettabili, ad esempio per la media, e si ipotizza che la v.c. sia Normale con media incognita μ, che si vuole stimare, e varianza nota σ^2. La v.c. Z è data dalla seguente notazione:

$$Z = \frac{(\bar{X} - \mu_0) \cdot \sqrt{n}}{\sigma} \tag{2.6.1.1.1}$$

che si distribuisce come una v.c. standardizzata $X \sim N(0;1)$
La probabilità che la media sia ricompresa fra i due estremi è data dalla seguente formula:

$$P\left(\bar{X} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha \tag{2.6.1.1.2}$$

La (2.6.1.1.2) esprime la probabilità che, ad un livello di fiducia o confidenza $1-\alpha$, ovvero nel $100(1-\alpha)\%$ dei campioni, la media μ della popolazione è ricompresa negli estremi seguenti:

$$\bar{X} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \tag{2.6.1.1.3}$$

La (2.6.1.1.3) può essere denotata in modo equivalente come:

$$P(\mu_1 \leq \mu \leq \mu_2) = 1 - \alpha \tag{2.6.1.1.4}$$

dove $1 - \alpha$ è il livello di fiducia o confidenza e α è il livello di significatività, ovvero la probabilità di compiere un errore qualora si affermi che il valore del parametro della popolazione di interesse sia compreso nei limiti$(\mu_1, \ \mu_2)$. Il livello di fiducia o confidenza $1-\alpha$ è quell'intervallo di valori campionario che dovrebbe contenere il valore del parametro della popolazione di interesse ad un prefissato livello di significatività α. Tale intervallo detto anche stima intervallare rappresenta lo strumento attraverso il quale è possibile dare un giudizio di affidabilità sulla stima dei parametri della popolazione. L'intervallo di valori individuato nelle notazioni precedenti può essere definito **stimatore intervallare** e i suoi valori estremi sono rappresentati da due v.c. Esso indica se il valore del parametro incognito della popolazione da stimare è ricompreso nell'intervallo stabilito.

2.6.1.2 Stima intervallare della Media di una popolazione con varianza ignota

In presenza di un campione casuale $X_1, X_2, \cdots, X_n$ estratto da una popolazione che si distribuisce secondo una Normale con media μ e varianza sconosciuta, si procede al calcolo della media e della varianza campionarie. In questo caso la v.c. scarto standardizzato si denota come t che è dato da:

$$t = \frac{(\bar{X} - \mu) \cdot \sqrt{n}}{S} \tag{2.6.1.2.1}$$

dove S è il valore della deviazione standard ovvero la radice quadrata della varianza stimata S^2.
Tale v.c. si distribuisce secondo una t di Student con $(n-1)$ gradi di libertà all'aumentare

dei quali essa tende alla distribuzione Normale. Lo stimatore intervallare per la media della popolazione al livello di confidenza $1-\alpha$ con varianza incognita è espresso dalla seguente notazione:

$$P\left(\bar{X} - t_{1-\alpha/2,n-1} \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{1-\alpha/2,n-1} \cdot \frac{S}{\sqrt{n}}\right) = 1-\alpha \qquad (2.6.1.2.2)$$

La (2.6.1.2.2) esprime la probabilità che, ad un livello di fiducia o confidenza $1-\alpha$ ovvero nel $100(1-\alpha)\%$ dei campioni, la media μ della popolazione è ricompresa negli estremi seguenti:

$$\bar{X} \pm t_{1-\alpha/2} \cdot \frac{S}{\sqrt{n}} \qquad (2.6.1.2.3)$$

2.6.1.3 Stima intervallare della Proporzione di una popolazione bernoulliana

In questo caso si vuole procedere alla costruzione di uno stimatore intervallare per una parte di unità statistiche che presentano due sole modalità (successo o insuccesso) espresse da v.c. bernoulliane. Poiché non risulta agevole tale costruzione si invoca il teorema del limite centrale secondo cui, al crescere della numerosità campionaria, la distribuzione della media campionaria si approssima alla distribuzione Normale con media p e varianza $p(1-p)$. Se si estrae un campione di numerosità n la proporzione campionaria $\hat{p} = \frac{X}{n}$ può essere considerata un buon stimatore di p. Lo scarto standardizzato Z, denotato come segue:

$$Z = \frac{(\bar{X} - p)}{\sqrt{\frac{p(1-p)}{n}}} \qquad (2.6.1.3.1)$$

si distribuisce come una v.c. standardizzata $X \sim N(0,1)$.
Lo stimatore intervallare per la proporzione di una popolazione bernoulliana p al livello di confidenza $1-\alpha$ è espresso dalla seguente notazione:

$$P\left(\hat{p} - z_{1-\alpha/2} \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \leq p \leq \hat{p} + z_{1-\alpha/2} \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}\right) = 1-\alpha \qquad (2.6.1.3.2)$$

dove $\hat{p}$ è la stima della proporzione campionaria e p è la proporzione della popolazione. La (2.6.1.3.2) esprime la probabilità che, ad un livello di fiducia o confidenza $1-\alpha$ ovvero nel $100(1-\alpha)\%$ dei campioni, la proporzione p della popolazione è ricompresa negli estremi seguenti:

$$\hat{p} \pm z_{1-\alpha/2} \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \qquad (2.6.1.3.3)$$

2.6.1.4 Stima intervallare della varianza di una popolazione Normale

Si prenda in considerazione una popolazione distribuita normalmente con valore medio e varianza entrambe incognite e con la v.c. standardizzata espressa dalla notazione $\frac{(n-1)\cdot S^2}{\sigma^2}$ che si distribuisce secondo una Chi-quadrato con *(n-1)* gradi di libertà.
Lo stimatore intervallare per la varianza è ricompreso fra due estremi ed è definito dalla notazione:

$$\frac{(n-1)\cdot S^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)\cdot S^2}{\chi^2_{1-\alpha/2,n-1}} \tag{2.6.1.4.1}$$

dove S^2 rappresenta la varianza campionaria corretta e $\chi^2_{\alpha/2,n-1}$ e $\chi^2_{1-\alpha/2,n-1}$ rappresentano i valori dei quantili critici della distribuzione Chi-quadrato con *n-1* gradi di libertà.

2.7 Test parametrici di verifica d'ipotesi

Le procedure che permettono di decidere se accettare o rifiutare una data ipotesi o di stabilire se un dato campione osservato differisce dai risultati attesi sono definite test statistici o test d'ipotesi o test di significatività dette anche regole di decisione. Si riporta nella **Tabella 2.7.1** la struttura delle ipotesi per i test unilateralidestro e sinistro e per il test bilaterale.

Tabella 2.7.1 Sistema di ipotesi

Struttura dei test	Ipotesi nulla.	Ipotesi alternativa
bilaterale	$H_0 : \mu = 0$	$H_1 : \mu \neq 0$
unilaterale sx	$H_0 : \mu = 0$	$H_1 : \mu < 0$
unilaterale sx	$H_0 : \mu = 0$	$H_1 : \mu \leq 0$
unilaterale dx	$H_0 : \mu = 0$	$H_1 : \mu > 0$
unilaterale dx	$H_0 : \mu = 0$	$H_1 : \mu \geq 0$

Legenda: H_0 = Ipotesi nulla o di interesse; H_1 = Ipotesi alternativa

Regole di decisione e tecnica del p-value.
Si ribadisce che le procedure che permettono di decidere se accettare o rifiutare una data ipotesi o di stabilire se un dato campione osservato differisce dai risultati attesi sono definite test statistici o test d'ipotesi o test di significatività dette anche regole di decisione.
Esse sono assunte secondo due procedimenti:

- il primo attraverso il metodo indiretto che prende in considerazione lo stimatore intervallare; assunta una distribuzione normale se il valore del quantile empirico z cade all'interno della regione di accettazione ricompresa fra i due estremi del quantile critico si accetta l'ipotesi nulla H_0 e si rifiuta quella alternativa H_1
- il secondo attraverso il metodo diretto che prende in considerazione la tecnica del p-value, definito anche la probabilità del valore del quantile empirico z; se il p-value