

INDAGINE STATISTICA

Concetto e funzioni della statistica

La statistica è la scienza che studia i fenomeni collettivi, qualitativi o quantitativi, attraverso la raccolta e l'elaborazione di dati. Obiettivo della statistica è lo studio dei fenomeni collettivi in condizioni di incertezza. La statistica produce descrizioni riassuntive di carattere quantitativo riguardanti un fenomeno collettivo e fornisce strumenti oggettivi che aiutano il processo decisionale.

Rami della statistica:

La statistica può essere:

- a) Statistica descrittiva, che raccoglie i dati di un certo fenomeno per descriverne le caratteristiche.
- b) Statistica inferenziale, che usa i dati raccolti per prevedere un determinato esito.
- a) Statistica metodologica, che concerne il metodo statistico in generale.
- b) Statistica applicata, che concerne le applicazioni del metodo statistico alle varie discipline (statistica economica, statistica sociale, giudiziaria, ecc.).

Fasi dell'indagine statistica

Un'indagine statistica consiste in una raccolta di dati che comprende le seguenti fasi:

- 1) **Formulazione**, sulla base di osservazioni casuali, **dell'ipotesi**, che si vuole sottoporre a verifica sperimentale.
- 2) **Rilevazione dei dati** scelti in modo che consentano lo studio del problema considerato (questionario);
- 3) **Spoglio dei dati** (tabella);
- 4) **Elaborazione dei dati** (scheda).

Rilevazione dei dati

La rilevazione dei dati richiede di scegliere e definire:

- a) L'**unità statistica**, che rappresenta la minima unità della quale si raccolgono i dati (individuo famiglia, impresa, ecc.).
- b) I **caratteri da rilevare** per ogni unità statistica, che possono essere qualitativi detti anche modalità (titolo di studio, professione, ecc.) o quantitativi o intensità espressi da numeri.
- c) L'**estensione nel tempo** (un istante o un intervallo di tempo) e **nello spazio** (in senso geografico o di categoria) della rilevazione.

Tipi di rilevazione

Le rilevazioni possono essere:

- 1) **Continue**, quando non presentano interruzioni nel tempo come le rilevazioni del movimento della popolazione secondo gli atti dello stato civile.
- 2) **Periodiche**, quando vengono eseguite ad intervalli regolari di tempo come i censimenti e l'indagine campionaria sulle forze di lavoro.
- 3) **Occasionali**, quando sono eseguite per conoscere un certo fenomeno e senza una periodicità assegnata come un'indagine sui pendolari nei comuni della valle Scrivia in un solo anno.
- 4) **Dirette**, quando vengono eseguite per conoscere un determinato fenomeno come il censimento generale della popolazione.
- 5) **Indirette**, quando derivano, ad esempio, da un procedimento amministrativo come il movimento delle iscrizioni e delle cancellazioni delle imprese dal Registro Imprese.

Rilevazioni complete e parziali

Particolare rilievo riveste la classificazione tra rilevazioni **complete** e **parziali**. Le prime indagano l'**universo** statistico costituito dalla totalità delle manifestazioni del fenomeno oggetto della rilevazione come la popolazione nel caso del censimento. Le rilevazioni parziali, invece, indagano solo una parte dell'universo statistico allo scopo di risparmiare tempo e denaro ma con l'intenzione di conoscere determinate caratteristiche dell'universo.

La parte dell'universo oggetto della rilevazione parziale costituisce il **campione**, che può essere **casuale**, quando è formato da unità statistiche estratte a sorte; **stratificato**, quando la popolazione oggetto della rilevazione è suddivisa in un numero determinato di strati o classi, in modo tale che le unità appartenenti ad uno strato siano omogenee rispetto al carattere in esame.

Spoglio dei dati

I dati relativi alle unità statistiche raccolti vanno ordinati e classificati secondo uno o più caratteri che sono stati rilevati, con un'operazione che prende il nome di spoglio dei dati. Le unità statistiche vengono raggruppate in classi omogenee corrispondenti alle singole modalità dei caratteri qualitativi o alle singole intensità o classi di intensità dei caratteri quantitativi. Il numero delle unità statistiche relativo alle singole classi costituisce un **dato statistico**, mentre il totale delle unità statistiche esaminate è denominato **popolazione statistica**.

Il risultato dello spoglio viene sintetizzato in appositi prospetti dette **tabelle**, che possono essere **semplici**, quando lo spoglio dei dati è stato eseguito secondo un solo carattere oppure **a doppia entrata**, quando lo spoglio dei dati è stato eseguito secondo due caratteri.

Esempio di tabella semplice
Individui per sesso

Sesso	Numero
Maschi	10
Femmine	13
Totale	23

La prima colonna contiene le varie modalità del carattere qualitativo ovvero le varie intensità del carattere quantitativo. La seconda colonna contiene i numeri (**frequenze**) che indicano quante volte le unità statistiche della popolazione esaminata appartengono a ciascuna classe. La tabella fornisce le frequenze per ciascuna classe e la somma delle frequenze dà la popolazione. La tabella che contiene le frequenze di un carattere qualitativo costituisce una **serie statistica**, mentre la tabella che contiene le frequenze di un carattere quantitativo costituisce una **seriazione**. In entrambi i casi, sono distribuzioni perché la tabella dice come si distribuisce la popolazione fra le varie classi.

Esempio di tabella a doppia entrata
Individui per sesso e titolo di studio

Sesso	Titolo di studio				
	Licenza elementare	Licenza media	Diploma	Laurea	Totale
Maschi	3	5	1	1	10
Femmine	4	6	2	1	13
Totale	7	11	3	2	23

Le frequenze possono essere **relative**, quando le singole frequenze sono divise per il totale di esse e moltiplicate per 100 o per 1.000; **cumulate** (solo nelle seriazioni), quando le singole frequenze (assolute o relative) sono sommate progressivamente partendo dalla prima o dall'ultima.

I dati contenuti in una tabella possono essere rappresentati graficamente, almeno nei casi più comuni, con un diagramma a coordinate cartesiane. I grafici possono essere istogrammi, cartogrammi, grafici a barre, grafici a linee, grafici a torta, ecc.

L'elaborazione dei dati consente di esprimere i risultati dell'indagine in modo sintetico che ne facilita l'interpretazione.

Le principali forme di elaborazione dei dati sono:

- 1) I **rapporti statistici** (di composizione, di coesistenza, di derivazione, di frequenza, di durata, numeri indici, ecc.).
- 2) Le **medie** (aritmetica semplice e ponderata, quadratica, armonica, geometrica, moda, mediana, ecc.).
- 3) Gli **indici di variabilità e concentrazione** (campo di variazione, scostamento o scarto, differenza media assoluta, indici relativi di variabilità, coefficiente di Gini, ecc.).
- 4) L'**interpolazione** e la **perequazione** (interpolazione per punti e fra punti, perequazione a medie mobili, ecc.).
- 5) Gli **indici di connessione e di correlazione** (indice di contingenza media assoluta, regressione, coefficiente di correlazione lineare di Bravais, indice di cograduazione di Gin).

Individui per sesso

Sesso	Frequenze	
	Assolute	Relative
Maschi	10	43,48
Femmine	13	56,52
Totale	23	100,00

Esempio di indagine statistica

L'indagine consiste nel verificare l'andamento dell'industria in una provincia che per semplicità è composta di sole tre classi di industrie e da un campione formato da un ristretto numero di imprese come risulta dalla tabella che segue.

Provincia di X

(Andamento industria manifatturiera nel trimestre gennaio – marzo 2025)

Classi di attività	Totale imprese	Campione (2%)	Addetti campione	% addetti campione
Alimentari	150	3	15	25,0
Tessili	200	4	32	53,3
Meccaniche	50	1	13	21,7
Totale	400	8	60	100,0

Il questionario prevede le domande relative agli indicatori presenti nella tabella che segue.

Ex ante					
Produzione	Fatturato	Costi	Export	Ordinativi	Occupazione
Ex post					
Produzione	Fatturato	Costi	Export	Ordinativi	Occupazione

Risposte fornite dalle imprese alla domanda relativa all'andamento della produzione:le

Alimentari	+5	-3	+2	
Tessili	+4	-1	+2	-3
Meccaniche	+5			

L'elaborazione a livello di imprese è la seguente:

Alimentari (25 %)	+5*25/100	-3*25/100	+2*25/100	+1,00%	+1,00
Tessili (53,3%)	+4*53,3/100	-1*53,3/100	+2*53,3/100	-3*53,3/100	+1,07
Meccaniche (21,7%)	+5*21,7/100				+1,08

L'elaborazione a livello di classi di attività dà come risultato: produzione +1,06%:

Classi di attività	Addetti	Peso	Produzione		Δ industria
Alimentari (25 %)	450	23,08	+1*23,08/100	0,23	
Tessili (53,3%)	1000	51,28	+1,07*51,28/100	0,55	
Meccaniche (21,7%)	500	25,64	+1,08*25,64/100	0,28	
					1,06