

1 – Introduzione

Problema 1. Come possiamo confrontare i dati di una rilevazione statistica?

- Il prezzo del bene A l'anno scorso era a ed oggi è α
- Il prezzo del bene A è a_1 nella città 1 ed a_2 nella città 2
- Il prezzo del bene A è a ed il prezzo del bene B è b

Soluzione: se i dati sono numerici, il confronto può avvenire con le seguenti operazioni:

- la **differenza o variazione**
- il **rapporto**

oppure attraverso una loro combinazione

Operazioni elementari

In base ai casi, possiamo essere interessati a:

- a) confrontare le frequenze corrispondenti a due diverse modalità;
- b) confrontare la frequenza relativa ad una data quantità ed il totale;
- c) confrontare due modalità di un carattere;
- d) confrontare due modalità di caratteri diversi.

Le ultime due opzioni sono possibili soltanto quando il carattere o i caratteri sono quantitativi, cioè le cui modalità sono numeri.

2 – La differenza o variazione

Sia X un carattere quantitativo (variabile) che assume valore x_t al tempo t ed x_0 al tempo 0 definiamo **differenza assoluta** o **variazione assoluta**^a del carattere X nell'intervallo temporale $[0, t]$ la quantità

$$\text{Diff}_{a[0,t]}(X) = x_t - x_0 \quad ^b \quad (1)$$

Se il risultato dell'operazione (1) risulta *negativo* allora siamo in presenza di una *diminuzione* altrimenti di un *aumento* del valore del carattere X registrato nel periodo $[0, t]$.

^a *assoluta* significa reale ovvero per come i dati x_0 e x_t sono stati acquisiti.

^b La (1) rappresenta il *range* della serie di X nel periodo $[0, t]$.

Operazioni elementari

Talvolta è necessario avere un'informazione di variazione avulsa dall'unità di misura, ovvero *in scala di rapporto*, che quantifichi la differenza in termini globali. Definiamo **differenza relativa** o **variazione relativa** del carattere X nel periodo $[0, t]$ la quantità

$$\text{Diff}_{r[0,t]}(X) = \frac{x_t - x_0}{x_0} \quad \text{a} \quad (2)$$

dove x_0 è il dato di riferimento o dato dal quale considerare la variazione. Se poi la (2) la moltiplichiamo per 100 otteniamo la **differenza percentuale** o **variazione percentuale** del carattere X nel periodo $[0, t]$ espresso dalla formula

$$\text{Diff}_{\%[0,t]}(X) = \frac{x_t - x_0}{x_0} \times 100 \quad (3)$$

^a La quantità (2) è un numero maggiore di 0 e minore di 1.

Operazioni elementari

Esempio 1. Nel mese di Ottobre 2010 il concessionario Toyota di Catanzaro ha venduto 35 auto, mentre in Gennaio 2011, il medesimo, ha venduto 12 automobili. Allora se indichiamo con X il carattere vendita auto del concessionario Toyota di Catanzaro in base alle: (1), (2) e (3) otteniamo i seguenti risultati

- $\text{Diff}_{a_{[ott'10, gen'11]}}(X) = 12 - 35 = -23;$
- $\text{Diff}_{r_{[ott'10, gen'11]}}(X) = \frac{12-35}{35} = -0.66;$
- $\text{Diff}_{\%_{[ott'10, gen'11]}}(X) = -0.66 \times 100 = -66\%;$ □

★☆☆ **Esercizio 1.** La Benzina a Dicembre 2010 aveva un costo medio di 1.465 €/litro, mentre a Febbraio 2011 ha un valore medio di 1.502 €/litro. Calcolare la variazione assoluta, la variazione relativa e la variazione percentuale nel periodo $[dic'10; feb'11]$.

Operazioni elementari

Il concetto di differenza assoluta (o relativa o percentuale) può essere generalizzato per un qualsiasi intervallo temporale $[i, i + 1] \subseteq [0, t]$. La differenza assoluta (o relativa o percentuale) del carattere X nel periodo $[i, i + 1]$ è detta **tasso di variazione assoluto** (o **relativo** o **percentuale**) del carattere X nel periodo $[i, i + 1]$. In formule abbiamo

$$T_{a_i}(X) = \text{Diff}_{a_{[i, i+1]}}(X) = x_{i+1} - x_i; \quad (4)$$

$$\left(\text{opp. } T_{r_i}(X) = \frac{x_{i+1} - x_i}{x_i}; \text{ opp. } T_{\%_i}(X) = \frac{x_{i+1} - x_i}{x_i} \times 100 \right) \quad (5)$$

$$\text{Poiché } \bigcup_{i=0}^{t-1} [i, i + 1] = [0, t] \text{ allora si ha } \sum_{i=0}^{t-1} T_{a_i}(X) = \text{Diff}_{a_{[0, t]}}(X)$$

Operazioni elementari

Esempio 2. Data la serie temporale di dati del carattere X

i	0	1	2	3	4	5	6
x_i	1	1	4	2	6	2	4

allora per la (4) si ha:

$$T_{a_0}(X) = 1 - 1 = 0; \quad T_{a_1}(X) = 4 - 1 = 3; \quad T_{a_2}(X) = 2 - 4 = -2;$$

$$T_{a_3}(X) = 6 - 2 = 4; \quad T_{a_4}(X) = 2 - 6 = -4; \quad T_{a_5}(X) = 4 - 2 = 2;$$

$$\text{e quindi } \text{Diff}_{a_{[0,6]}}(X) = \sum_{i=0}^5 T_{a_i}(X) = 0 + 3 - 2 + 4 - 4 + 2 = 3 \quad \square$$

★☆☆**Esercizio 2.** Completare la serie temporale $\{6, 5, x_2, 8, x_4\}$ del carattere X sapendo che $T_2(X) = 3$ e la variazione assoluta è 4.

Operazioni elementari

Osservazione 1. Applichiamo l'approssimazione matematica

$$\ln(b) - \ln(a) = \ln\left(\frac{b}{a}\right) = \ln\left(\frac{a + b - a}{a}\right) = \ln\left(1 + \frac{b - a}{a}\right) \approx \frac{b - a}{a}$$

valida per a e b piccoli e positivi, al caso in cui

$$a = x_i \quad \text{e} \quad b = x_j \quad \text{con } i < j$$

Allora la variazione relativa del carattere X nel periodo $[i, j]$ può essere così calcolato

$$\text{Diff}_{r[i,j]}(X) \approx \ln(x_{i+1}) - \ln(x_i) \quad ^a \quad (6)$$

★☆☆ **Esercizio 3.** Calcolare la variazione relativa della benzina nel periodo $[dic'10; feb'11]$ dell'Esercizio 1 utilizzando la formula (6).

^a Se $j = i + 1$ allora la (6) può essere utilizzata per il calcolo di $T_{r_i}(X)$.

3 – Rapporti statistici

La variazione relativa e la variazione percentile del carattere X sono forme particolari di rapporti statistici. Il fine che perseguono questi indicatori descrittivi è:

- agevolare l'interpretazione del fenomeno indagato ed i confronti tra le modalità;
- ricondurre i dati ad una scala di misura standardizzata di riferimento (unitaria o percentuale) più semplice ed immediata nella lettura;
- eliminare l'influenza che altre variabili (*variabili di disturbo*) hanno sul fenomeno indagato

Operazioni elementari

I rapporti statistici possono essere classificati rispetto ai dati utilizzate che si riferiscono a:

- 1 carattere:
 - rapporti di composizione
 - rapporti di coesistenza
 - indici di eccedenza
- 2 caratteri oppure 1 carattere misurato in tempi diversi:
 - rapporti di densità
 - rapporti di durata
 - rapporti di ripetizione

3.1 – Rapporti di composizione

Il rapporto di composizione R_C è ottenuto come frazione di una quantità relativa ad una modalità del carattere X e l'ammontare complessivo.

Esempio 3. Dato il carattere X =auto possedute così rilevato,

x_i	1	2	3	≥ 4	Totale
f_{a_i}	46	32	12	5	95

il rapporto di composizione dei possessori di 1 auto è

$R_C(x_i = 1) = \frac{46}{95} \approx 0.4842$ ovvero il 48.42% mentre il rapporto di composizione di coloro che ne hanno più di 2 è

$$R_C(x_i > 2) = \frac{12 + 5}{95} \approx 0.1789 \sim 17.89\%. \quad \square$$

Operazioni elementari

Esempio 4. Il *Tasso di disoccupazione* è ottenuto come rapporto fra le persone in cerca di occupazione e le persone occupate ed in cerca di occupazione^a. Nel 2009 in Italia si è registrato nel secondo trimestre un tasso di disoccupazione pari a $\frac{1841046}{25043775} \times 100 = 7.4$.

Esempio 5. Il *Tasso di occupazione* è ottenuto come rapporto fra le persone occupate e le persone di età compresa tra i 14 ed i 65 anni. Nel 2009 in Italia si è registrato nel secondo trimestre un tasso di occupazione pari a $\frac{22819186}{39381230} \times 100 = 57.9$.

★☆☆ **Esercizio 4.** Calcolare il *Tasso di attività giovanile* (rapporto tra le persone di età compresa tra i 15 ed i 24 anni appartenenti alle forze lavoro e la popolazione nella stessa classe di età) in Italia nel quarto trimestre del 2010.

^a *Forze lavoro* = occupati + in cerca di occupazione.

3.2 – Rapporti di coesistenza

Il rapporto di coesistenza R_C è ottenuto come frazione della frequenza (o quantità) corrispondente ad una modalità del carattere X e la la frequenza (o quantità) corrispondente ad un'altra modalità.

Esempio 6. Nel 2006 in Italia i nati vivi maschi erano 286.676 mentre i nati vivi femmine erano 269.751. Il rapporto di coesistenza

$$R_C = \frac{286.676}{269.751} = 1.06$$

detto *rapporto di mascolinità delle nascite* indicava che, se moltiplicato per 100, per ogni 100 femmine erano nati 106 maschi.

★☆☆ **Esercizio 5.** Calcolare il *rapporto di mascolinità delle morti* in Italia nel 2005 (vedi annuario del 2008 dell'Istat).

Operazioni elementari

Esempio 7. In Italia al primo gennaio del 2008 la popolazione di età uguale o superiore ai 65 anni ammontava a 11948161 unità, mentre la popolazione di età tra i 15 ed i 64 anni ammontava a 39291856.

L'indice di dipendenza degli anziani era pari a

$$\frac{11948161}{39291856} \times 100 = 30.41 \quad \square$$

Esempio 8. Nel 2006 il valore delle esportazioni dall'Italia era valutato a 332013 milioni di €, mentre il valore delle importazioni era di 352465 milioni di €. Il rapporto di coesistenza

$$\frac{332013}{352465} \times 100 = 94.2$$

detto *grado di copertura* indica che il valore della merce esportata è di 94.2 € contro un'importazione pari a 100 €. □

3.3 – Indici di eccedenza

Data una popolazione suddivisa in due classi rispetto ad un carattere, analizziamo lo squilibrio tra le classi eliminando l'influenza dell'ammontare complessivo della popolazione.

Un indice di eccedenza è il rapporto della differenza dei due valori (le frequenze delle due classi) e la somma degli stessi valori.

Esempio 9. Al primo gennaio 2008 vi erano $X_M = 28949747$ maschi e $X_F = 30669543$ femmine per un totale di 59.619.290 residenti. Lo squilibrio tra i sessi era

$$\frac{X_F - X_M}{X_F + X_M} = \frac{30669543 - 28949747}{30669543 + 28949747} = 0.029$$

Operazioni elementari

ed interpretato come $0.029 \times 100 = 2.9\%$ di eccedenza di femmine^a. \square

Esempio 10. Riprendendo i dati dell'Esempio 8 il bilancio commerciale del 2008 in Italia era

$$352465 - 332013 = 20442$$

milioni di € a favore delle importazioni per un valore totale di 684478 milioni di €. L'indice di eccedenza era allora

$$\frac{20442}{684478} = 0.03$$

ovvero del 3% a favore delle importazioni. \square

^a Avendo scelto come primo valore quello maggiore tra i due per cui la differenza al numeratore del rapporto è positiva, allora il risultato 2.9 acquista il seguente significato: su 100 residenti la metà di $100 - 2.9 = 97.1$ ovvero 48.55 erano maschi e 51.45 erano femmine.

Operazioni elementari

★☆☆**Esercizio 6.** Calcolare lo squilibrio tra la *popolazione attiva* e la *popolazione non attiva* in Italia nel 2008.

★☆☆**Esercizio 7.** Determinare su 100 bambini quanti nel 2007 in Italia sono *nati vivi* e *morti*.

★★★**Esercizio 1.** Nel 2007 i diplomati del 2004 erano così distribuiti per condizione occupazionale

	Lavorano	Non Lavorano
Femmine	225780	81998
Totale	225780	140994

Stabilire lo squilibrio tra i diplomati del 2004 e gli squilibri all'interno delle classi maschi e femmine in termini occupazionali.

3.4 – Rapporti di densità

Il rapporto di densità R_d è ottenuto come frazione relativa di un'intensità di un carattere X rispetto ad una dimensione spaziale.

Le rilevazioni sono bivariate semplici in cui il secondo carattere deve indicare la superficie in cui viene osservato il primo carattere.

Esempio 11. Nel 2008 la popolazione residente in Italia era 59619290 su una superficie complessiva di 301302 km^2 . Il rapporto

$$\frac{59619290}{301336} = 197.85$$

è detto *grado di affollamento* o *densità della popolazione* ed indica che mediamente sono presenti 197.85 persone in 1 km^2 . □

Operazioni elementari

REGIONI	Classi di superficie territoriale											
	Fino a 1.000		1.001-2.000		2.001-6.000		6.001-25.000		Oltre 25.000		Totale	
	Comuni	Superficie %	Comuni	Superficie %	Comuni	Superficie %	Comuni	Superficie %	Comuni	Superficie %	Comuni	Superficie %
Piemonte	404	10,1	417	23,2	316	40,8	69	25,9	-	-	1.206	100,0
Valle d'Aosta/ Vallée d'Aoste	8	1,8	14	6,4	36	37,8	16	53,9	-	-	74	100,0
Lombardia	745	17,7	468	27,5	286	36,6	47	18,3	-	-	1.546	100,0
Trentino-Alto Adige	69	3,2	73	7,5	125	31,8	71	55,2	1	2,2	339	100,0
<i>Bolzano/Bozen</i>	9	0,6	14	2,9	49	24,9	43	67,6	1	4,1	116	100,0
<i>Trento</i>	60	6,2	59	13,1	76	40,1	28	40,5	-	-	223	100,0
Veneto	42	1,8	208	17,5	269	46,6	60	30,6	2	3,6	581	100,0
Friuli-Venezia Giulia	18	1,4	58	11,2	110	48,1	33	39,2	-	-	219	100,0
Liguria	61	7,5	83	22,5	79	48,5	12	21,5	-	-	235	100,0
Emilia-Romagna	7	0,2	19	1,4	193	33,6	117	56,2	5	8,6	341	100,0
Toscana	5	0,2	26	1,9	106	18,4	140	65,7	10	13,9	287	100,0
Umbria	1	0,1	7	1,2	38	17,4	39	51,4	7	29,9	92	100,0
Marche	20	1,6	76	11,4	102	34,6	47	49,7	1	2,8	246	100,0
Lazio	28	1,3	95	8,2	181	35,6	70	41,8	4	13,2	378	100,0
Abruzzo	24	1,6	83	11,8	157	50,7	40	31,5	1	4,3	305	100,0
Molise	3	0,5	34	12,0	85	62,8	14	24,7	-	-	136	100,0
Campania	161	6,9	141	15,3	213	53,9	36	23,9	-	-	551	100,0
Puglia	27	1,1	40	3,1	96	18,2	82	52,7	13	24,9	258	100,0
Basilicata	-	-	6	1,0	57	22,3	66	70,2	2	6,5	131	100,0
Calabria	39	1,8	92	9,3	227	53,0	50	34,1	1	1,9	409	100,0
Sicilia	53	1,4	63	3,6	142	20,2	117	54,4	15	20,4	390	100,0
Sardegna	22	0,6	56	3,6	163	25,5	132	64,3	4	6,0	377	100,0
ITALIA	1.737	3,5	2.059	10,0	2.981	34,1	1.258	44,3	66	8,0	8.101	100,0
Nord-ovest	1.218	12,5	982	23,9	717	39,6	144	23,9	-	-	3.061	100,0
Nord-est	136	1,5	358	8,7	697	38,9	281	46,2	8	4,7	1.480	100,0
Centro	54	0,7	204	5,2	427	26,0	296	53,9	22	14,2	1.003	100,0
Mezzogiorno	329	1,7	515	6,3	1.140	33,0	537	48,4	36	10,5	2.557	100,0

Fonte: Variazioni territoriali, denominazione dei comuni, calcolo delle superfici comunali (E)

(a) Le classi di superficie territoriale sono espresse in ettari. I dati della superficie dei comuni derivano dalle misurazioni dell'Agenzia del territorio al 31 dicembre 2002.

Tabella 1: Fonte: Annuario Statistico 2008 ISTAT

Operazioni elementari

Esempio 12. Con riferimento alla Tabella 1, i comuni italiani alla data del 31 dicembre 2007 erano 8101, ripartiti in numero massimo nel Nord-ovest (3061 con più di 5000 abitanti); numero minimo nelle Isole (767 comuni in 65 km^2 e 8729 abitanti). Nel Centro (11641 abitanti) c'è un affollamento rispetto al n. di comuni presenti. \square

★☆☆ **Esercizio 8.** Nella tabella sono censiti gli istituti di credito presenti nella ex Prov. di Catanzaro nel 2008 (*Fonte: Pagine Gialle*)

Area geografica	Istituti di credito	Superficie (km^2)
Catanzaro	120	2391
Crotone	46	1717
Vibo Valentia	48	1139

Calcolare i rapporti di densità delle aree geografiche. \square

3.5 – Rapporti di durata

Il rapporto di durata interessa quella popolazione che nel periodo $[0, t]$, è soggetta ad un mutamento nel numero complessivo (cambia la *consistenza*) delle unità a causa di immissione di nuove unità (*flusso in entrata*) ed di emissione di unità presenti (*flusso in uscita*).

$$D_{[0,t]}(X) = \frac{C_0 + C_t}{E + U} = \frac{C_0 + (C_0 + E - U)}{E + U} = \frac{2C_0 + E - U}{E + U} \quad (7)$$

- $C_0 = \text{consistenza iniziale} = \text{n. di unità al tempo } 0$;
- $C_t = \text{consistenza finale} = \text{n. di unità al tempo } t$;
- $E = \text{flusso in entrata} = \text{n. di unità entrate nel periodo } [0, t]$;
- $F = \text{flusso in uscita} = \text{n. di unità uscite nel periodo } [0, t]$.

Operazioni elementari

Tale rapporto calcola allora, la *permanenza* di una generica unità statistica nel collettivo.

Esempio 13. Al 1° gennaio 2009 la popolazione italiana era di 60045068; al 31 dicembre del 2009 si registrava un *movimento naturale* di 568857 *nati vivi* e 591663 *morti* ed un *movimento migratorio* di 1850482 iscritti e di 1532416 cancellati. Il tempo di permanenza di un residente italiano riferito al periodo dal 01/01/09 al 31/12/09 è in base alla (7)

$$\frac{2 \times 60045068 + (568857 + 1850482) - (591663 + 1532416)}{(568857 + 1850482) + (591663 + 1532416)} = 26.5$$

★☆☆ **Esercizio 9.** Determinare il rapporto di durata del fenomeno residenza in Italia nel primo trimestre 2011.

Operazioni elementari

Esempio 14. La Tabella 2 illustra il fenomeno assunzione - licenziamento di un'azienda con 500 impiegati negli anni 2001-2010.

Anno	'01	'02	'03	'04	'05	'06	'07	'08	'09	'10
Assunti	13	18	27	43	34	16	37	12	26	33
Licenziati	27	20	18	33	22	10	30	38	29	37

Tabella 2: Assunzione/licenziamento di un'azienda nel periodo 2001-2010

Osservato che la consistenza iniziale è uguale alla consistenza finale ovvero 500 per tutti gli anni, l'indice di rotazione di occupazione nell'azienda, ad es. per l'anno 2005, è $\frac{2 \times 500 + 34 - 22}{34 + 22} = 18.07$.

★☆☆**Esercizio 10.** Con riferimento ai dati della Tabella 2 calcolare il tempo medio di impiego per anno.

3.6 – Rapporti di ripetizione

Il rapporto di ripetizione è l'inverso del rapporto di durata ed indica il numero di volte che il fenomeno osservato nel periodo $[0, t]$ si ripete in quel lasso temporale.

Esempio 15. Con riferimento ai dati della Tabella 2, il rapporto di ripetizione relativo all'anno 2005 è

$$\frac{1}{D_{2005}} = \frac{1}{18.07} = 0.06$$

che moltiplicato per 100 assume il seguente significato di: nell'anno 2005 si è verificata una rotazione del personale pari al 6%.

★☆☆ **Esercizio 11.** Calcolare la rotazione del personale dell'azienda per ogni anno.