STATISTICA DESCRITTIVA Esempi di esercizi

**1**: Ai 1000 abitanti un piccolo comune viene chiesto di esprimere un giudizio su un nuovo servizio comunale, usando una scala da 0 a 4 (0=pessimo, 4= ottimo). Le risposte ottenute sono riassunte nella tabella che segue. Fornire indici di posizione e di variabilit`a adeguati e una rappresentazione graﬁca opportuna delle risposte.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| giudizio | 0 | 1 | 2 3 | 4 |
| freq. assoluta | 251 | 260 | 80 154 | 255 |

**Soluzione:**



*Rappresentazione grafica*

*media: x* = 1*.*90

*mediana:*

*x*ˆ = 1

*moda:*

*x*˜ = 1

*varianza: s*2 = 2*.*43

*deviazione standard: s* = 1*.*56*.*

**2**: Vengono intervistati 36 torinesi, a cui viene chiesto il numero di vani presente nella propria abitazione. Le 36 risposte ottenute sono le seguenti:

1*,* 3*,* 4*,* 2*,* 2*,* 4*,* 5*,* 5*,* 1*,* 1*,* 2*,* 3*,* 4*,* 3*,* 2*,* 6*,* 6*,* 1*,*

2*,* 2*,* 3*,* 2*,* 1*,* 3*,* 4*,* 2*,* 3*,* 3*,* 3*,* 5*,* 6*,* 4*,* 2*,* 2*,* 4*,* 2*.*

a) Fornire le frequenze relative e le frequenze cumulate relative delle risposte ottenute.

b) Fornire un rappresentazione graﬁca ”a torta” delle risposte. c) Determinare media, moda, mediana e varianza delle risposte.

d) La stessa indagine `e stata svolta a Milano, e le risposte fornite dai 36 milanesi hanno dato un valor medio uguale a 2.5 ed una varianza uguale a 3.6. Confrontare la variabilit´a relativa del numero di vani delle abitazioni nei due comuni e commentare.

**Soluzione:**



*c)*

*media: x* = 3

*mediana:*

*x*ˆ = 3

*moda:*

*x*˜ = 2

*s*2 = 2*.*11

*d)*

*√ √*

*c.v.*(*T O*) = 2*.*11 = 0*.*48 *c.v.*(*M I* ) = 3*.*6 = 0*.*76

3 2*.*5

*Variabilit´a maggiore a Milano, sia in senso assoluto che relativo.*

**3**: Per 300 giorni vengono rilevati i consumi complessivi di energia elettrica presso un pic- colo comune montano. I dati ottenuti, espressi in *K W* , vengono riassunti nella tabella che segue. Rappresentare graﬁcamente i consumi osservati, e trovare in maniera approssimata media, mediana e deviazione standard dei consumi.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| classe | [0*,* 100) | [100*,* 200) | [200*,* 400) | [400*,* 600) | [600*,* 1000] |
| freq. | assoluta | 50 | 85 | 65 | 55 | 45 |

**Soluzione:**





*media: x ≃* 327*.*5

*mediana:*

*x*˜ *≃* 250

*varianza: s*2 *≃* 60870

*dev standard: s ≃* 247

**4**: Considerate le seguenti coppie (*x, y*) di dati:

(22*,* 45)*,* (17*,* 30)*,* (15*,* 28)*,* (25*,* 49)*,* (21*,* 38)*.*

a) Si discuta la correlazione tra i due caratteri *x* ed *y*;

b) si trovi la retta di regressione (minimi quadrati) delle *y* date le *x* e con essa si fornisca una stima del carattere *y* di una coppia avente *x* = 20.



**Soluzione:**

*media: x* = 20*, y* = 38*,*

*s*2 (*x*) = 12*.*8*, s*2 (*y*) = 66*.*8*,*

*cX Y* = 28*.*6*, rX Y* = 0*.*98 *(forte correlazione positiva),*

*m* =  *cX Y*

*s*2 (*x*) = 2*.*23*, q* = *x − my* = *−*6*.*7*, y* = 2*.*23*x −* 6*.*7*,*

*x*0 = 20*, y*0 = 38*.*

**5**: Si consideri il seguente elenco di coppie (*x, y*) di dati:

(2*,* 0)*,*(1*,* 0)*,* (1*,* 0)*,* (2*,* 1)*,* (0*,* 1)*,* (1*,* 1)*,* (2*,* 1)*,*

(0*,* 0)*,*(1*,* 0)*,* (0*,* 1)*,* (0*,* 1)*,* (0*,* 0)*,* (1*,* 1)*,* (1*,* 0)*.*

*a*) Si riassuma la serie di dati con una tabella a doppia entrata.

*b*) Si trovino media, mediana, moda e varianza dei singoli caratteri *x* ed *y* (presi separata- mente).

*c*) Si discuta la correlazione tra i due caratteri *x* ed *y* (sono incorrelati?).

**Soluzione:**



*media: x* = 1 *y* = 0*.*429

*s*2 (*x*) = 0*.*428 *s*2 (*y*) = 0*.*245

*cX Y* = 0*.*071 *rX Y* = 0*.*22 *(incorrelazione, o debolissima correlazine positiva)*

**6**: Date le seguenti frequenze assolute (relative al numero di vani di una popolazione di

180 unit`a abitative), determinare le frequenze relative e relative cumulate, rappresentarle graﬁcamente, e determinare media, mediana, moda e deviazione standard dei dati.

modalit`a 2 3 4 5 6 freq. assoluta 30 55 45 30 20



**Soluzione:**

*media: x* = 3*.*75

*moda:*

*x*˜ = 3

*mediana:*

*x*ˆ = 4

*s*2 = 11*.*83

*s* = 3*.*44

**7**: Data la serie di dati quantitativi discreti

*{*3*,* 5*,* 1*,* 1*,* 0*,* 2*,* 8*,* 2*,* 3*, x},*

con *x ∈* IR, trovare un valore da assegnare ad *x* aﬃnch´e risulti

*a*) *x*¯ = 3;

*b*) *x*˜ = 2;

*c*) *x*ˆ = 2*.*5;

*d*) *x*ˆ *< x*¯.

Posto poi *x* = 1,

*e*) trovare la varianza e il coeﬃciente di variazione della serie di dati;

*f*) scrivere le frequenze relative e le frequenze relative cumulate;

*g*) rappresentare graﬁcamente la distribuzione dei dati.

**Soluzione:**

*a) x* = 5 *b) x* = 2 *c) x* = 3 *d) x* = 1

*e) x*¯ = 2*.*67*, s*2 = 5*.*56*, c.v.* = 0*.*88

**8**: Ai 1500 abitanti un piccolo comune viene chiesto di esprimere un giudizio su un nuovo servizio comunale, usando una scala da 1 a 5 (1=pessimo, 5= ottimo). Le risposte ottenute

sono riassunte nella tabella che segue. Fornire indici di posizione e di variabilit`a adeguati e una rappresentazione graﬁca opportuna delle risposte.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| giudizio | 1 | 2 | 3 | 4 | 5 |
| freq. assoluta | 450 | 125 | 110 | 230 | 575 |

**Soluzione:**



*media: x* = 3*.*22

*mediana:*

*x*ˆ = 4

*moda:*

*x*˜ = 5

*s*2 = 2*.*96

*s* = 1*.*72

**9**: Considerate le seguenti coppie (*x, y*) di dati:

(18*,* 30)*,* (22*,* 45)*,* (14*,* 28)*,* (25*,* 49)*,* (21*,* 38)*.*

a) Si discuta la correlazione tra i due caratteri *x* ed *y*;

b) si trovi la retta di regressione (minimi quadrati) delle *y* date le *x* e con essa si fornisca una stima del carattere *y* di una coppia avente *x* = 20.

**Soluzione:**



*media: x* = 20 *y* = 38

*s*2 (*x*) = 14 *s*2 (*y*) = 66*.*8

*cX Y* = 29*, rX Y* = 0*.*948 *(forte correlazione positiva)*

*m* =  *cX Y*

*s*2 (*x*) = 2*.*07*, q* = *x − my* = *−*3*.*4*, y* = 2*.*07*x −* 3*.*4*,*

*x*0 = 20*, y*0 = 38

**10**: Per 200 giorni vengono rilevati i consumi di acqua potabile presso un piccolo comune. I dati ottenuti, espressi in m3 , vengono riassunti nella tabella che segue. Rappresentare graﬁcamente i consumi osservati, e trovare in maniera approssimata media, mediana e deviazione standard dei consumi.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| classe | [0*,* 200) | [200*,* 300) | [300*,* 400) | [400*,* 600) | [600*,* 1000] |
| freq. | assoluta | 30 | 55 | 45 | 40 | 30 |

**Soluzione:**





*media: x ≃* 382*.*5

*moda:*

*x*ˆ *≃* 340

*s*2 *≃* 45950*, s ≃* 214*.*

**11**: Sono state intervistate 30 persone, a cui `e stato chiesto quante volte si sono recate al cinema nell’ ultimo mese. Le 30 risposte ottenute sono le seguenti:

1*,* 0*,* 4*,* 2*,* 2*,* 4*,* 5*,* 0*,* 1*,* 1*,* 2*,* 3*,* 4*,* 3*,* 2*,*

2*,* 2*,* 3*,* 0*,* 1*,* 3*,* 4*,* 0*,* 0*,* 3*,* 3*,* 5*,* 6*,* 4*,* 2*.*

a) Fornire le frequenze relative e le frequenze cumulate relative delle risposte ottenute. b) Fornire un rappresentazione graﬁca ”a torta” delle risposte.

c) Determinare media, moda, mediana e varianza delle risposte.

d) La stessa indagine `e stata svolta d’estate, e le risposte fornite dalle 30 persone hanno dato un valor medio uguale a 2.5 ed una varianza uguale a 3. Confrontare con i dati riportati sopra e commentare.

**Soluzione:**



*media: x* = 2*.*4

*mediana:*

*x*ˆ = 2

*moda:*

*x*˜ = 2

*s*2 = 2*.*64*, s* = 1*.*62*.*

*c.v.*(*inv*) = 0*.*67 *c.v.*(*est*) = 0*.*69 *(variabilit´a relativa praticamente identica)*

**12**: Si consideri il seguente elenco di coppie (*x, y*) di dati:

*{*(1*,* 0)*,* (1*,* 2)*,* (2*,* 1)*,* (2*,* 2)*,* (2*,* 2)*,* (1*,* 2)*,* (1*,* 2)*,*

(1*,* 1)*,* (2*,* 0)*,* (1*,* 1)*,* (2*,* 2)*,* (2*,* 1)*,* (1*,* 0)*,* (1*,* 2)*,* (2*,* 0)*}*

a) Si riassuma la serie di dati con una tabella a doppia entrata;

b) Si discuta la correlazione tra i due caratteri *x* ed *y* (sono incorrelati?).

c) Si fornisca una rappresentazione numerica ed una rappresentazione graﬁca del singolo carattere *y*.



**Soluzione:**

|  |  |  |
| --- | --- | --- |
| *media: x* = 1*.*47 | *y* = 1*.*2 |  |
| *s*2 (*x*) = 0*.*24*cX Y* = *−*0*.*003 | *s*2 (*y*) = 0*.*69*rX Y* = *−*0*.*07 | *(incorrelazione)* |

**13**: Cento dati di tipo quantitativo continuo vengono raggruppati in classi come segue:

classe [0*,* 1) [1*,* 2) [2*,* 4) [4*,* 8) [8*,* 15]

frequenza 10 13 28 22 27

a) rappresentare graﬁcamente tramite istogramma ed ogiva;

b) fornire una approssimazione di media, mediana e varianza;

**Soluzione:**



*media: x ≃* 5*.*51

*mediana:*

*x*ˆ *≃* 3*.*9

*s*2 *≃* 16*.*1*, s ≃* 4*.*

**14**: Si consideri una coppia (*X, Y* ) di caratteri che per dieci dati individui assume valori

*{*(0*,* 0)*,* (1*,* 0)*,* (1*,* 0)*,* (1*,* 1)*,* (0*,* 0)*,* (0*,* 1)*,* (1*,* 1)*,* (1*,* 0)*,* (1*,* 0)*,* (0*,* 0)*}.*

Relativamente alla serie data,

a) la si riassuma graﬁcamente con una tabella a doppia entrata;

b) si determinino medie e deviazioni standard di entrambi i caratteri *X* ed *Y* ;



c) si discuta la correlazione tra *X* ed *Y* .

**Soluzione:**

*media: x* = 0*.*6 *y* = 0*.*3

*s*2 (*x*) = 0*.*24 *s*2 (*y*) = 0*.*21

*cX Y* = 0*.*02 *rX Y* = 0*.*089 *(incorrelazione)*

**15**: Si considerino le seguenti coppie (*x, y*) di dati:

(6*,* 50)*,* (2*,* 70)*,* (3*,* 80)*,* (2*,* 70)*,* (2*,* 130)

a) Si discuta la correlazione tra i due caratteri *x* ed *y*;

b) si trovi la retta di regressione (minimi quadrati) e con essa si fornisca una stima del carattere *y* di una coppia avente *x* = 4.

**Soluzione:**



*media: x* = 3 *y* = 80 *s*2 (*x*) = 2*.*4 *s*2 (*y*) = 720 *cX Y* = *−*24 *rX Y* = *−*0*.*58

*m* =  *cX Y*

*s*2 (*x*) = *−*10*, q* = *x − my* = 110*, y* = *−*10*x* + 110*,*

*x*0 = 4*, y*0 = 70

**16:** Nella seg. tabella è riportata la distrib. delle famiglie per n° di componenti.

N° Componenti Famiglie

|  |  |
| --- | --- |
| 1 | 150 |
| 2 | 212 |
| 3 | 312 |
| 4 | 500 |
| 5 | 342 |
| 6 | 122 |
| 7 | 72 |
| 8 | 32 |

Determinare la media aritmetica, geometrica, armonica, quadratica nonché il valore centrale, la mediana, la moda ed il 1° e 3° quartile. Inoltre determinare il campo di variazione, la differenza interquartilica, lo scarto semplice medio, la devianza, la varianza e la deviazione standard.

**17:** Sia dato il seguente insieme di 20 dati, che rappresentano il peso al la nascita (in g) di 20 bambini nati in una settimana in una clinica

|  |  |  |  |
| --- | --- | --- | --- |
| 3280 | 3320 | 2500 | 2760 |
| 3260 | 3650 | 2840 | 3250 |
| 3240 | 3200 | 3600 | 3320 |
| 3480 | 3020 | 2840 | 3200 |
| 4160 | 2580 | 3540 | 3780 |

1.

Calcolare media e mediana del campione

2.

Determinare il diagramma a scatola del campione

3.

Ripetere il punto 1 considerando anche l’eventualità di un bambino nato prematuro di 500 g. Quale delle due misure è più robusta alla presenza di valori estremi?

