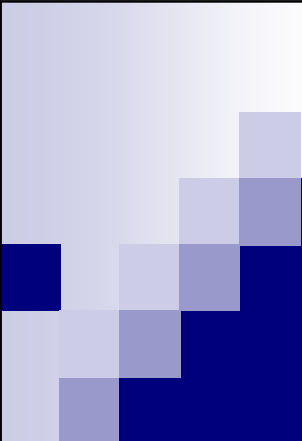




Ricerca Operativa 2° modulo

A.A. 2008/2009

2. Statistica descrittiva e calcolo
delle probabilità (pillole)



Pillole di statistica descrittiva

- **Statistica:** insieme di metodi scientifici applicati alla raccolta, classificazione, analisi e presentazione dei dati quantitativi allo scopo di giungere a delle conclusioni che non si possono ottenere in maniera intuitiva
- **Fase descrittiva:** raccolta dei dati, elaborazione, sintesi
- **Fase induttiva:** interpretazione, previsione

Fase descrittiva: passi

1. Raccolta dei dati
2. Classificazione e rappresentazione
3. Sintesi

Raccolta dei dati: definizioni

- **Campione**: insieme di esperimenti / osservazioni
- **Unità statistica**: elementi del campione
- **Variabile**: caratteristica rilevata/misurata dell'unità statistica
- **Modalità**: valori distinti assunti da una variabile (ci limitiamo a dati quantitativi)
- **Variabile statistica** Y : corrispondenza empirica tra le unità statistiche i e le modalità dei valori y_i ad esse associati

Raccolta dei dati: Matrice dei dati

VALORI

VARIABILI

UNITÀ STAT.

Unità statistica	Vendite settore 1	Vendite settore 2	Costi settore 1	...
Mese 1	10	6	8	...
Mese 2	11	12	9	...
Mese 3	15	6	5	...
...	...	...	...	...

VARIABILE STATISTICA

Ricerca Operativa (2° modulo) – Ing. Luigi De Giovanni

Lezione 2 – pag. 5

Classificazione (o seriazione)

- Se i dati sono numerosi o se la variabile è continua è conveniente definire **classi di modalità** contigue.
- Esempio: vendite in €
 - Classe 0 – 1000: vendite $\in [0, 1000)$
 - Classe 1000 – 2000: vendite $\in [1000, 2000)$
 - ...
 - Nota: 1000 – 2000: vendite $\in (1000, 2000]$
- Definizione delle classi:
 - Numero di classi adeguato
 - Disgiunte
 - Esaustive delle modalità osservate
- D'ora in avanti parleremo, in generale, di classe (indice j)

Ricerca Operativa (2° modulo) – Ing. Luigi De Giovanni

Lezione 2 – pag. 6

Rappresentazione: frequenze

- Data una variabile e una classe j
- **Frequenza assoluta** n_j : numero di unità portatrici della classe
- **Frequenza relativa** f_j : frazione di unità portatrici della modalità/classe
- **Frequenza cumulativa** F_j : frequenza relativa di j sommata a tutte le classi precedenti (classi ordinate)
- Sia N il num. di unità statistiche e J il num. di classi

$$f_j = \frac{n_j}{\sum_{j=1}^J n_j} = \frac{n_j}{N}$$

$$F_j = \sum_{i=1}^j f_i$$

Ricerca Operativa (2° modulo) – Ing. Luigi De Giovanni

Lezione 2 – pag. 7

Rappresentazione: grafici

- **Istogramma** delle frequenze assolute o relative: una barra per classe, di altezza proporzionale alla frequenza
- **Torta** delle frequenze relative
- **Poligono** (o serie) di frequenza
- **Serie temporale**: se le unità statistiche sono riferite a intervalli temporali consecutivi

Ricerca Operativa (2° modulo) – Ing. Luigi De Giovanni

Lezione 2 – pag. 8

Sintesi: indici di posizione

- **Media aritmetica** della variabile Y : $E(Y) = \mu_Y = \mu = \bar{Y}$

- Su dati grezzi:
$$E(Y) = \frac{1}{N} \sum_{i=1}^N y_i$$

- Su dati classificati (media aritmetica ponderata da n_j):

$$E(Y) = \frac{1}{N} \sum_{j=1}^J y_j n_j = \sum_{j=1}^J y_j f_j$$

- Classificazione su variabili continue (valore centrale)

$$E(Y) = \frac{1}{N} \sum_{j=1}^J y_j^c n_j = \sum_{j=1}^J y_j^c f_j \quad \text{dove} \quad y_j^c = \frac{(y_{j-1} + y_j)}{2}$$

Indici di variabilità

- **Varianza** di Y : $V(Y) = \sigma_Y^2 = \sigma^2$. Media aritmetica del quadrato della variabile di scarto $Y - E(Y)$

$$V(Y) = E((Y - E(Y))^2) = \frac{1}{N} \sum_{j=1}^J (y_j - E(Y))^2 n_j = \sum_{j=1}^J (y_j - E(Y))^2 f_j$$

- **Scarto quadratico medio**: $\sigma_Y = \sigma$

$$\sigma_Y = \sqrt{V(Y)}$$

- In caso di classi continue si usa y_j^c invece di y_j

Pillole di calcolo delle probabilità: distribuzione binomiale o di Bernoulli

- Descrive la probabilità della ri-presentazione di un evento *esclusivo* (il suo verificarsi esclude tutti gli altri) in caso di prove indipendenti ripetute
 - N: numero di prove eseguite
 - x: numero di successi voluti
 - p: probabilità di successo dell'evento in una singola prova
 - q: probabilità di insuccesso ($q=1-p$)

$$P(x) = \frac{N!}{x!(N-x)!} p^x q^{N-x} = \binom{N}{x} p^x q^{N-x}$$

Distribuzione binomiale: direzioni applicative

- Si hanno a disposizione 12 ore di produzione su un macchinario con resa di 10 pezzi/ora e probabilità del 3% di produrre un pezzo difettoso.
 - Che probabilità abbiamo di riuscire ad evadere un ordine di 110 pezzi senza rimanenze in magazzino?
 - Qual è la probabilità se ammettiamo di avere pezzi sovrabbondanti? (si utilizzino le probabilità cumulate)
- Si ha disposizione una macchina per la produzione di pezzi (un pezzo per ciclo produttivo). La probabilità che un ciclo produttivo dia un pezzo difettoso è dello 0,5%. Si vuole determinare il numero di cicli produttivi per avere una probabilità di almeno il 97% di evadere un ordine di 100 pezzi.
- Si vuole acquistare un macchinario per la produzione di 100 pezzi dei quali al massimo 10 possono risultare difettosi. Sul mercato esistono macchinari con diversi prezzi e diversi livelli di accuratezza, misurati dalla probabilità che un pezzo prodotto risulti difettoso. La scelta del macchinario deve tenere conto della probabilità di ripresentazione dell'evento "10 o meno pezzi difettosi su 100 ripetizione della produzione". Si confrontano le distribuzioni binomiali relative alle diverse opzioni (probabilità di un singolo pezzo difettoso) del mercato.
- ...

Statistica in Excel

Statistica descrittiva

- Se si hanno dati grezzi
 - `MEDIA`(intervallo)
 - `VAR.POP`(intervallo)
 - `DEV.ST.POP`(intervallo)
- Se si hanno frequenze assolute/relative
 - Calcolare le formule "a mano"

Distribuzione binomiale

- `DISTRIB.BINOM`(x,N,p,falso)