Prof. Giulio Vidotto (Università di Padova)

Lez. 13 - ***Distribuzione normale e standardizzazione delle misure***

Argomenti della lezione

- ➔ **Standardizzazione**
- ➔ **Distribuzione Normale**
- ➔ **Distribuzione Normale Standard**

Standardizzazione

Standardizzazione

La standardizzazione ha lo scopo di rendere i dati direttamente confrontabili, caratteristica che i dati grezzi in se non possiedono se vengono mantenuti nella forma originale

Punti z

Indicano la posizione dei dati in termini di distanza dalla media (numero di deviazioni standard)

$$z_i = \frac{X_i - \overline{X}}{s} = \frac{x_i}{s}$$

quindi...

$$X_i = \overline{X} + z_i \cdot s$$

1° Proprietà Formale

Somma algebrica dei Punti z

$$\sum_{i=1}^{n} z_i = 0$$

2° Proprietà Formale

Media Aritmetica dei Punti z

$\bar{z} = 0$

in quanto: $$\bar{z} = \frac{\sum_{i=1}^{n} z_i}{n} = \frac{0}{n} = 0$$

3° Proprietà Formale

Sommatoria dei Punti z

in quanto: $$\sum_{i=1}^{n} z_i^2 = n$$

4° Proprietà Formale

Deviazione dei Punti z

$$s_z = s_z^2 = 1$$

in quanto s_z^2

$$= \frac{\sum_{i=1}^{n} (z_i - \bar{z})^2}{n} = \frac{\sum_{i=1}^{n} z_i^2}{n} = \frac{n}{n} = 1$$

Un esempio

Uno Soggetto è sottoposto a due test:
Memoria: Pg=50, M=60 e DS=5
Intelligenza: Pg=80, M=100 DS=10
I risultati (Pg, Punteggio grezzo) potrebbero far protendere per una migliore riuscita al secondo Test...

In realtà, per paragonare i punteggi, e meglio calcolare i relativi punti z

Pz (memoria) = (50-60)/5 = -2

Pz (intelligenza) = (80-100)/10= -2

Se i dati originali hanno una distribuzione normale, si ha una corrispondenza fra tutti i punti z ed i ranghi percentili (Rp). Anche i ranghi percentili (Rp) sono indipendenti dall'unità di misura

Trasformazione lineare f(z)

Data una trasformazione lineare f(z)=a+bz chiamiamo Z una qualsiasi trasformazione lineare di z
I Punti Z così ricavati dalla trasformazione dei punti *z*, diventano interi e positivi per $a \geq \min(z)$ e $b > 0$

1° Proprietà Formale

La media dei Punti z

$$\bar{Z} = a$$

in quanto

$$\bar{Z} = a + b \cdot \bar{z} = a + b \cdot 0 = a$$

2° Proprietà Formale

La deviazione standard di un Punto z

$$s_Z = b$$

in quanto

$$s_Z^2 = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z})^2}{n} = \frac{\sum_{i=1}^{n} (a + b \cdot z_i - a)^2}{n} = b^2 \cdot \frac{\sum_{i=1}^{n} z_i^2}{n}$$

In Psicologia, ad esempio, alcune trasformazioni molto in uso sono

Punti T = 50 + 10z
Punti Q.I. = 100 + 10z
Punti St = 5 + 1z

Distribuzione Normale

Distribuzione Normale (DN)

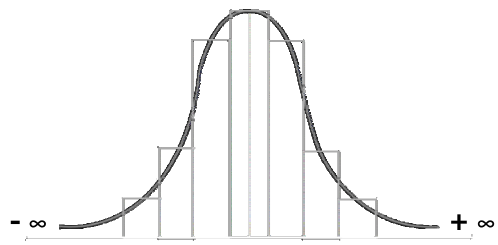
La DN è relativa a variabili casuali continue. Riveste un ruolo assai importante nella teoria della probabilità e in statistica. I dati di molte ricerche psicologiche si distribuiscono in modo molto simile a tale funzione

Prof. Giulio Vidotto (Università di Padova)

Lez. 13 - ***Distribuzione normale e standardizzazione delle misure***

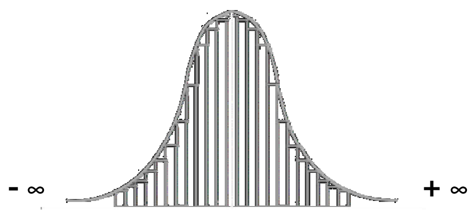
Si può descrivere la DN come una forma particolare del poligono di frequenza di dati distribuiti normalmente. Nelle ricerche psicometriche si incontra spesso che una Distribuzione di Frequenza possieda le caratteristiche di: simmetria e forma a campana...

Poligono di frequenza di dati



La Curva della DN, quindi, può essere concepita come la forma limite del poligono di frequenza di una serie di dati distribuiti normalmente...

Forma limite del poligono di frequenza di dati



Caratteristiche della Distribuzione Normale

Una v.c. x ha una distribuzione normale, con media μ e varianza σ^2, se la sua densità di proprietà è data dalla funzione (con $-\infty < x < +\infty$)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

...e se la sua funzione di ripartizione è data da (con $-\infty < x < +\infty$)

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2}\left(\frac{x'-\mu}{\sigma}\right)^2} dx'$$

con le seguenti caratteristiche...

- **è perfettamente simmetrica all'ordinata massima**
- **la funzione di distribuzione f(x) è asintotica di x verso -∞ e +∞**
- **è crescente per valori di X che vanno da -∞ a μ e decrescente per valori che vanno da μ a +∞**

- **è caratterizzata dai due parametri μ e σ^2 e dalle tre costanti 2, π, e**
- **presenta due punti di flesso in corrispondenza a μ + σ e μ - σ**

Per un qualsiasi valore della v.c. x (x = a) la probabilità P(a) con -∞ < x < a corrisponde all'integrale

$$P(a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Valori attesi della distribuzione normale

$$E(x) = \int_{-\infty}^{+\infty} x\, \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$Var(x-\mu) = \int_{-\infty}^{+\infty} (x-\mu)^2\, \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Distribuzione Normale Standard

Distribuzione Normale Standardizzata (DNS)

Il ricorso alla DNS consente di individuare le probabilità relative ai diversi intervalli di valori, a differenza della DN dove si utilizzava l'integrale di P(a), mediante le tavole di probabilità

Proprietà della Curva

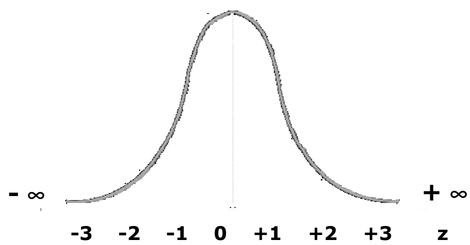
La DNS si ottiene con la trasformazione lineare dei punti grezzi in punti z. La Funzione di Densità di Probabilità della DNS f(z), da molti indicata con il simbolo f(u), diventa quindi…

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2} z^2}$$

con $-\infty < z < +\infty$

La Distribuzione Normale Standardizzata presenta le stesse caratteristiche della Distribuzione Normale non Standardizzata: forma a campana, simmetria, flessi a ± 1 deviazioni standard dalla media….

Distribuzione Normale Standardizzata



La Distribuzione Normale Standardizzata appartiene alla famiglia delle Distribuzioni Normali i cui parametri sono $\mu \in$ Re e $\sigma \in$ Re+

La Distribuzione Normale Standardizzata in particolare ha

$\mu = 0$, $\sigma = 1$

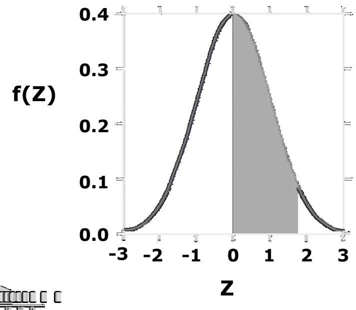
Tavole della Distribuzione

➔ **I valori delle aree della DNS sono tabulati in tavole utilizzate per**

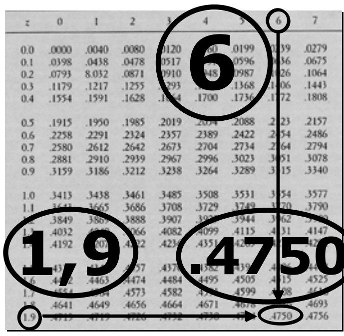
✓ **calcolare l'area compresa tra due determinati valori della variabile**

✓ **determinare la quantità di punteggi compresi tra due valori di una variabile casuale**

Ad Esempio…l'area compresa tra z = 0 e z = 1,96 è rappresentata da



Area tra z=0 e z=1,96 è 0,475 => 47,5%



Un Esempio applicato alla Psicologia

Test sulla Socievolezza

N° Soggetti = 500 con M=100 DS=15
Il Manuale del Test definisce i Soggetti con punteggi compresi tra 88 e 130 come "propensi a buoni contatti sociali", quanti soggetti ci si attende nell'intervallo compreso tra 88 e 130?

Per risolvere tale compito il ricercatore dovrà calcolare i punti zeta corrispondenti a 88 e 130, che saranno:

$$z(88) = \frac{88 - 100}{15} = -0,8$$

$$z(130) = \frac{130 - 100}{15} = +2$$

dalle tavole si ricava che:

- **area compresa tra z = -0,8 e z = 0 è 0,2881**
- **area compresa tra z = 0 e z = +2 è 0,4772**
- **area complessiva tra z = -0,8 e z = +2 sarà 0,2881 + 0,4772 = 0,7653**

0,7653 può essere letto sia come proporzione dei casi compresi tra i valori 88 e 130, sia come la probabilità che il punteggio di un soggetto cada all'interno di tale intervallo

Il numero di punteggi che ci si attende nell'intervallo compreso tra 88 e 130 si calcola:

$$0{,}7653 \cdot 500 = 382{,}65 \cong 383$$