

Tecniche di ricerca psicologica e analisi dei dati

A.A. 2022/23

PROF.SSA LISA DI BLAS

DIBLALI@UNITS.IT

Contenuti, obiettivi e bibliografia

Il corso presenta le tecniche di base di analisi dei dati, insieme ad elementi di metodologia su ruoli e relazioni tra variabili.

Gli argomenti principali sono

- dal punteggi grezzo ai punteggi standardizzati e normalizzati nel testing psicologico;
- analisi della varianza per confronti tra gruppi, per misure ripetute e modelli fattoriali misti;
- analisi della correlazione semplice e multipla
- analisi della regressione semplice e multipla (con particolare attenzione ai concetti di parzializzazione, spiegazione e previsione di una variabile, controllo di una relazione)
- tecniche di analisi di riduzione dei dati (analisi fattoriale esplorativa)
- elementi psicometrici di base nel testing psicologico (validità e affidabilità)

Contenuti, obiettivi e bibliografia

Esercitazioni pratiche, produzione e interpretazione di output di analisi dei dati (programma statistico Jamovi), **esercitatore dott. Nicolo Petruzzella**

M. Gallucci, L. Leone e M. Berlingeri (2017, seconda ed). “Modelli statistici per le scienze sociali” (capp. 2, 3, 4, 8, 9, 12, 13) Pearson Italia, MilanoTorino.

Obiettivi formativi:

- acquisizione di conoscenze di base metodologiche e delle tecniche di analisi (ANOVA, regressione lineare e riduzione dei dati,
- funzionali a comprendere e utilizzare correttamente il linguaggio tecnico psicometrico,
- comprendere i punteggi standardizzati nel testing psicologico,
- sviluppare la capacità di valutare in modo autonomo e critico gli strumenti di misurazione psicologica e le verifiche empiriche di proposte teoriche di un articolo scientifico,
- riuscire a comprendere, formulare e difendere correttamente ipotesi di indagine
- organizzare, raccogliere e analizzare dati con un'autonomia di base.

Competenze di base e propedeuticità

Il corso di Psicometria 1 è propedeutico all'esame del presente corso, vale a dire, è obbligatorio superare l'esame di Psicometria 1 (o eventuali diverse etichette) per poter iscriversi e sostenere l'esame di Tecniche di ricerca psicologica e analisi dei dati.

Competenze di base della Psicometria, con particolare riferimento a Statistica descrittiva, elementi di base della statistica inferenziale, interpretazione del p valore osservato, teoria del limite centrale, distribuzioni campionarie degli stimatori (gaussiana, t di student); elementi di base della metodologia: variabile (aleatoria) e variabilità, campionamento, disegno di ricerca pre-sperimentali, quasi-sperimentali e sperimentali; disegni a misure ripetute.

Prima di iniziare

La valutazione della didattica

Le vostre competenze acquisite

Per imparare a leggere affermazioni apparentemente semplici da comprendere con nuove lenti, distintive della vostra futura professione

1. Marco è una persona particolarmente aggressiva.
2. Mediamente, gli studenti di questa scuola sono poco aggressivi.
3. Assassini seriali che commettono questo tipo di omicidi sono generalmente uomini, di 35-45 anni, di razza caucasica, con livelli di istruzione elevati, professioni di rilievo, contesto familiare non integro.
4. Alcune persone sono più aggressive di altre.
5. A seguito del trattamento, il senso di soddisfazione corporea è maggiore nel gruppo sperimentale rispetto alla condizione di controllo.
6. L'aggressività è maggiore negli uomini rispetto alle donne.
7. A seguito del training didattico, i livelli di apprendimento linguistico sono maggiori nella scuola in cui è stato applicato il training rispetto alla scuola di controllo, in cui non è stato applicato

Per imparare a leggere affermazioni apparentemente semplici da comprendere con nuove lenti, distintive della vostra futura professione

1. L'aggressività aumenta quando aumenta la frustrazione percepita.
2. La frustrazione è una determinante causale dell'aggressività.
3. Le condizioni depressive aumentano nelle persone costrette all'isolamento sociale per tre mesi.
4. L'intelligenza risente delle condizioni socio-economiche.

Per imparare a leggere affermazioni apparentemente semplici da comprendere con nuove lenti, distintive della vostra futura professione

1. Il senso di solitudine si associa a condizioni depressive, soprattutto nelle persone con attaccamento relazionale insicuro.
2. Se le persone sono emotivamente vulnerabili, l'esposizione a Instagram favorisce la loro insoddisfazione corporea.
3. Nell'arco di un anno, l'insoddisfazione corporea aumenta nelle persone che si espongono quotidianamente a Instagram.
4. Quando preadolescenti sono quotidianamente esposti a Instagram, la loro insoddisfazione corporea aumenta sensibilmente nell'arco di un anno; l'effetto non si osserva nelle persone adulte.

Per imparare a leggere affermazioni apparentemente semplici da comprendere con nuove lenti,
distintive della vostra futura professione

L'umore di Mario fluttua notevolmente nell'arco della giornata.

Le fluttuazioni umorali a livello individuale sono maggiori nelle persone sottoposte a forte stress lavorativo.

Per imparare a leggere affermazioni apparentemente semplici da comprendere con nuove lenti,
distintive della vostra futura professione

Mario utilizza un vocabolario esteso e conosce tante cose, è una persona intelligente (colta?)

Mario prende un giocattolo senza giocarci, non ascolta quello che dici, disegna facendo solo scarabocchi, è un bambino distratto (o soffre di depressione? di qualche deficit cognitivo?)

Mario non protesta mai, accetta sempre tutto e non esprime mai quello che pensa, è una persona accomodante (sottomessa, inaffidabile, ...?)

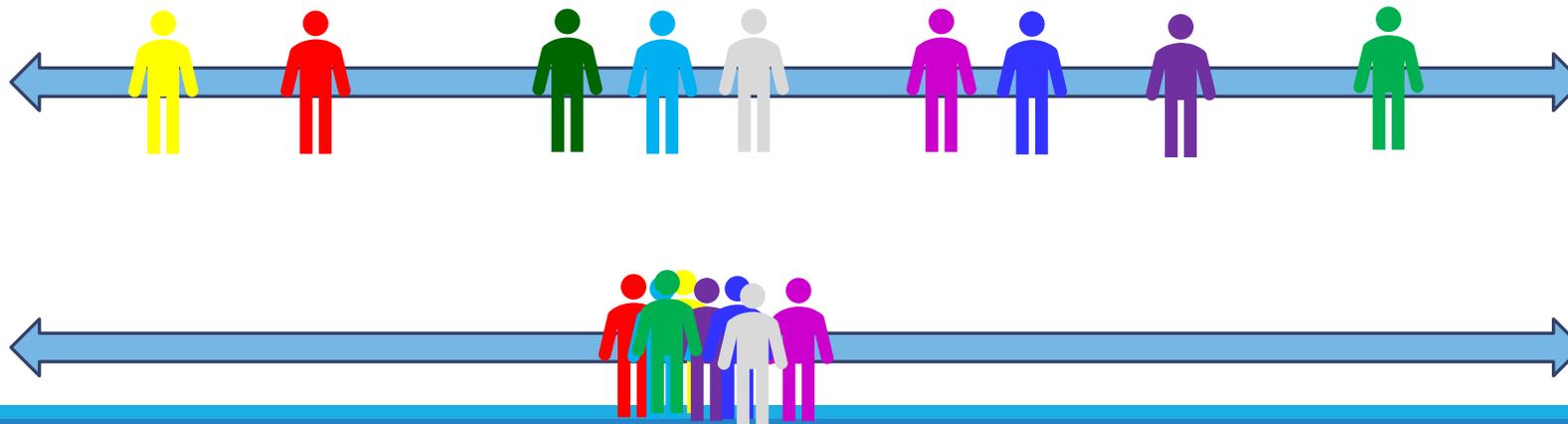
1 variabile

1 variabile e la sua variabilità

La variabilità è essenziale, noi studiamo variabili 😊

Le persone sono differenti le une dalle altre → differenze inter-personali (between people)

Condensiamo quella variabilità attraverso indicatori quantitativi: $M \pm DS$



Statistiche descrittive e punteggi individuali

La coorte di student* 2019/20 ha conseguito il titolo di laurea triennale con una media di 98/110

Quale esito ha conseguito lo studente Mario ?

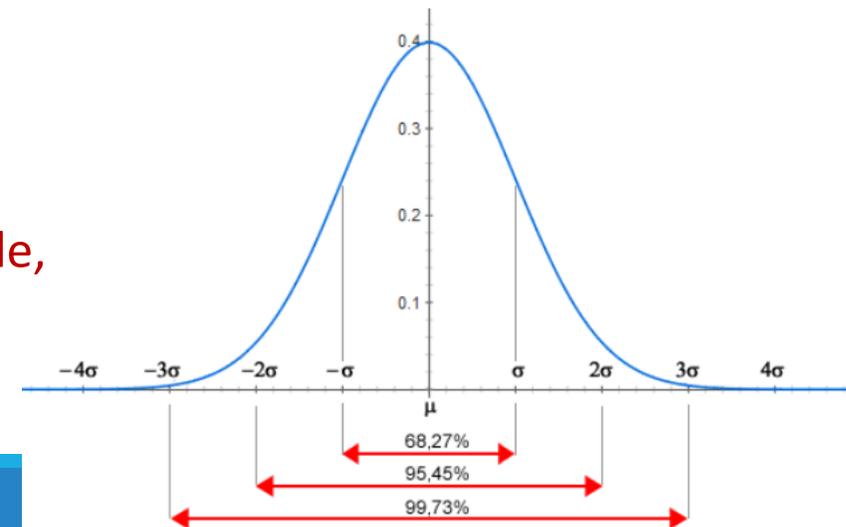


Ma di quanto ci sbagliamo nella previsione?

E se considero la variabilità intorno alla media, DS = 7, posso raffinare la mia previsione?

Quale esito ha conseguito Mario, con una probabilità di 2 su 3?

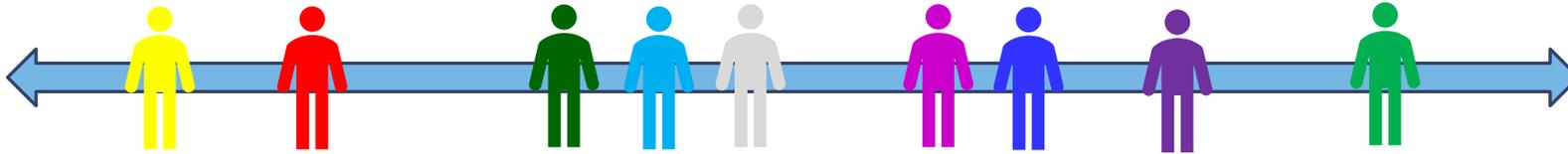
**La media mi aiuta a prevedere un esito anche a livello individuale,
con uno scarto d'errore**



Statistiche descrittive e punteggi individuali



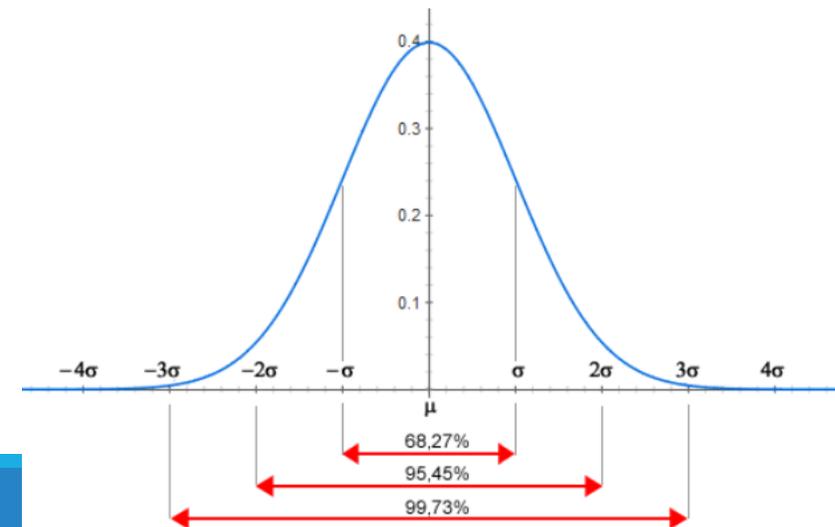
E se Mario ha conseguito un esito di 101/110, come si colloca rispetto agli altri studenti? --- meglio --- ok ma quanto meglio?



Se la distribuzione è gaussiana
posso calcolare un punto z
e sfruttare le proprietà della distribuzione normale

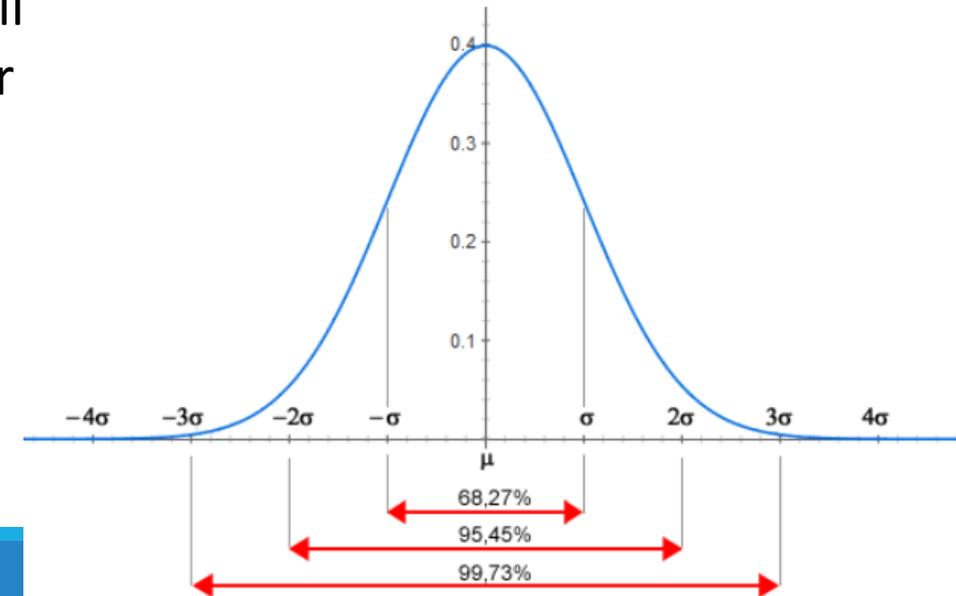
$$z = (X_i - M) / DS$$

$$z = (101 - 98) / 7 = 0.42$$



Ma cos'è una distribuzione gaussiana o normale standardizzata?

- È una distribuzione di probabilità continua
- permettere di descrivere valori che tendono a concentrarsi intorno a un valore medio
- forma a campana e simmetrica
- è caratterizzata da una **funzione di densità di probabilità**
- grazie alla quale si è in grado di **definire probabilisticamente** il numero di casi che ricadono entro un'area della campana (cfr tabelle della distribuzione gaussiana nelle slide successive)
- La media attesa è 0 e la deviazione standard è 1



2. Dal RP al punto z

Mario

TAV. II. Aree della distribuzione normale standard tra $a = 0$ e $b > 0$

<i>z</i>	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4351	.4265	.4279	.4292	.4306	.4319

TAV. II. Aree della distribuzione normale standard tra $a = 0$ e $b > 0$

z	0	1	2	3	4	5	6	7	8	9
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

STATISTICHE DESCRITTIVE E PUNTEGGI INDIVIDUALI

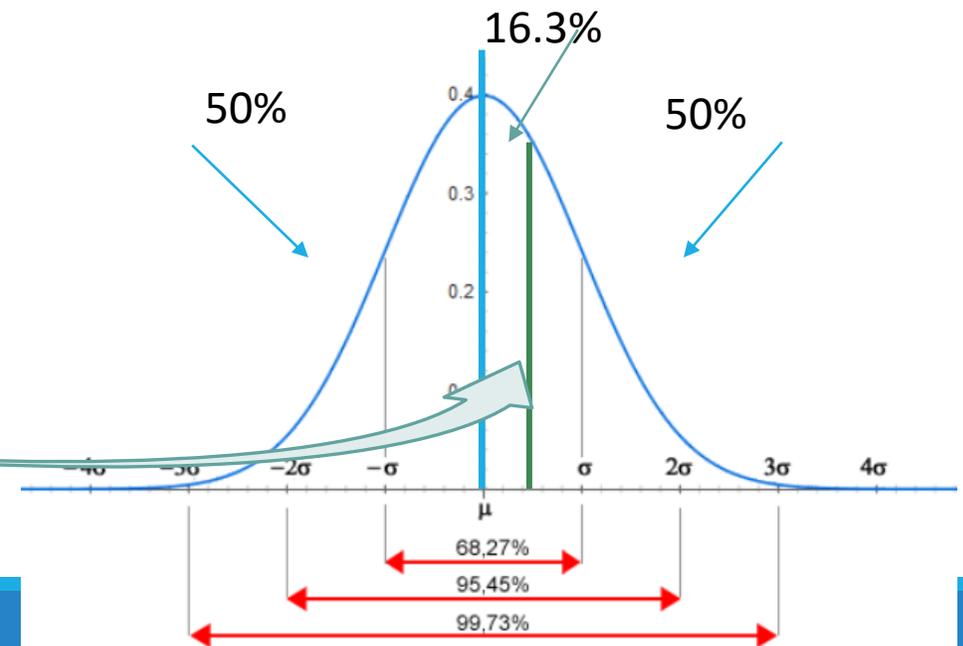
E allora quel 101/110 di Mario che corrisponde a 0.42 punti z cosa mi dice?

Per esempio, a che punti si colloca lungo un continuum normativo, in termini di punti RP

Posso trasformare il punteggio grezzo in **punto rango percentile** (Rp)

$z = 0.42 \rightarrow$ area .163, vale a dire il 16.3%
dei casi attesi ricade tra $z = 0$ e $z = 0.42$

pertanto il numero di casi attesi
con $z = 0.42$ corrispondono a un
punto **Rp = 16.3 + 50 = 63.3**



Statistiche descrittive e punteggi individuali

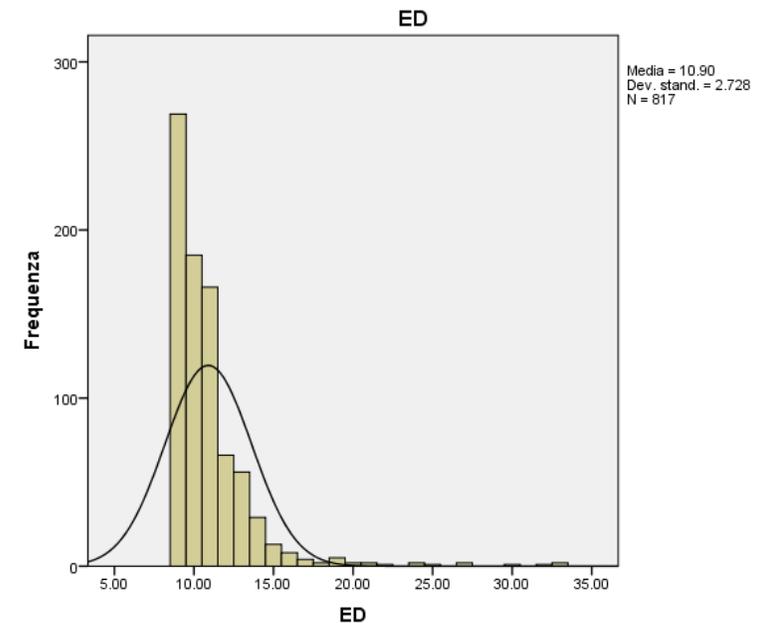
E se la distribuzione dei valori osservati di una variabile non è normale?

Non posso sfruttare le proprietà di una distribuzione gaussiana, ma posso

trasformare la distribuzione normalizzandola

I dati vengono normalizzati e standardizzati

Nel testing psicologico il punteggio grezzo viene convertito in un **punteggio normativo**



Eating Disorders

	14-20	21-30	31-40	41-50	51-60	61 and older
(N)	161	308	79	88	145	35
raw score						
11	39	37	38	38	36	38
12	45	45	45	46	43	47
13	50	50	50	51	48	53
14	54	53	54	55	53	57
15	57	56	57	58	57	61
16	58	59	61	60	60	65
17	60	62	62	63	63	66
18	61	64	63	66	66	67
19	61	65	65	70	69	69
20	63	67	68	80	70	69
21	64	68	71		73	69
22	65	70	72		80	72
23	65	71	72			80
24	66	72	72			
25	66	73	72			
26	67	75	72			
27	67	76	72			
28	67	77	72			
29	67	77	72			
30	68	77	72			
31	68	77	72			
32	69	77	72			
33	70	79	72			
34	71	80	72			
35	71		72			
36	71		72			

Tabella finale di conversione

Statistiche descrittive e punteggi individuali

Come procedere per normalizzare e standardizzare una distribuzione?

Punteggi grezzi vengono espressi in R_p

$$R_p(X) = \frac{cf_b + 0.5(f_i)}{N} \times 100$$

si calcolano per tanto per ogni punteggio osservato o osservabile lungo un continuum:

- frequenza osservata (f_i)
- frequenza cumulata (cf_i)
- rango percentile

Punti z e punti z normalizzati (servendosi delle tavole della distribuzione normale)

Useremo un file excel per le applicazioni pratiche di questa trasformazione