LA CURVA NORMALE

1

Nozioni preliminari

2

Il concetto di probabilità

- La probabilità può essere definita come il numero degli eventi favorevoli sul totale degli eventi possibili.
- Evento = uno dei possibili risultati di una prova

Esempio: qual è la probabilità che lanciando un dado si ottenga 5?

Gli **eventi possibili** sono **6** (corrispondenti alle 6 facce del dado), **l'evento favorevole** (faccia con 5) è **1**, perciò la **probabilità** di ottenere 5 lanciando un dado è **1/6 = 0,166666…. (arrotondato a 0,17)**.

3

- La probabilità può essere espressa come:
  - ✓ **frazione (1/6);**
  - ✓ **proporzione (0,17);**
  - ✓ **percentuale (17%).**

Esempio: qual è la probabilità che lanciando un dado si ottenga 5?

Gli **eventi possibili** sono **6** (corrispondenti alle 6 facce del dado), **l'evento favorevole** (faccia con 5) è **1**, perciò la **probabilità** di ottenere 5 lanciando un dado è **1/6** (**= 0,17** = **17%**).

4

Quando si parla di probabilità ci si riferisce sempre ad **eventi casuali**. Per esempio il risultato del lancio di un dado è casuale se il dado non è truccato.
Se l'evento non è casuale, non è possibile conoscerne la probabilità. Se, ad esempio, il dado è più pesante da un lato, la probabilità dell'uscita di un numero sarà diversa da 1/6, a seconda che il numero si trovi dalla parte più pesante o più leggera del dado. Se volessimo calcolare la probabilità dell'uscita di un certo numero su dadi truccati, dovremmo considerare tutti i possibili modi in cui può essere truccato un dado (es. più pesante di 0,0001 grammi o 0,0002 grammi o 0,0003 grammi ecc. su una delle 6 facce e così per tutte le 6 facce) e calcolare la probabilità di ogni numero in tutte queste condizioni. Poiché il peso è una variabile continua, avremmo una quantità infinita di distribuzioni di probabilità. Se invece facciamo riferimento ad un evento casuale (dado non truccato) il valore della probabilità per ogni evento è noto (nel caso del dado è uno solo: 1/6).

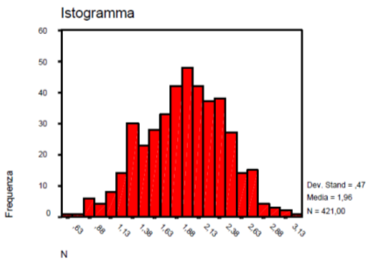
5

Probabilità e distribuzioni di frequenza

6

E' possibile calcolare la probabilità di ottenere un certo punteggio (es. dalla somministrazione di un test ad un campione casuale) osservando la distribuzione di frequenza dei punteggi all'interno di quel campione.

Per esempio in questo caso il punteggio 1,88 ha frequenza 50 (cioè 50 persone hanno ottenuto il punteggio 1,88). Poiché il campione è composto da 421 persone (N = 421), la probabilità di ottenere il punteggio 1,88 è 50/421 = 0,12 (12%). Ciò significa anche che il 12% del campione ha ottenuto il punteggio 1,88.



7

Qual valore otteniamo se sommiamo le probabilità di tutti i punteggi rappresentati nell'istogramma precedente?

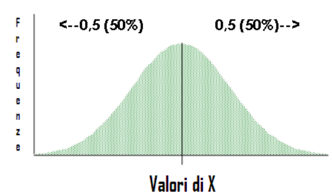
Otteniamo 1 se abbiamo calcolato le probabilità come proporzioni, otteniamo 100% se le abbiamo calcolate come percentuali. Il valore di probabilità 1 (o 100%) indica la certezza. Infatti la probabilità, nel campione precedente, di ottenere uno qualsiasi dei punteggi del test è uguale a 1, cioè la certezza.
La probabilità è quindi sempre compresa tra 0 e 1 (oppure tra 0% e 100%).

8

La curva normale

9

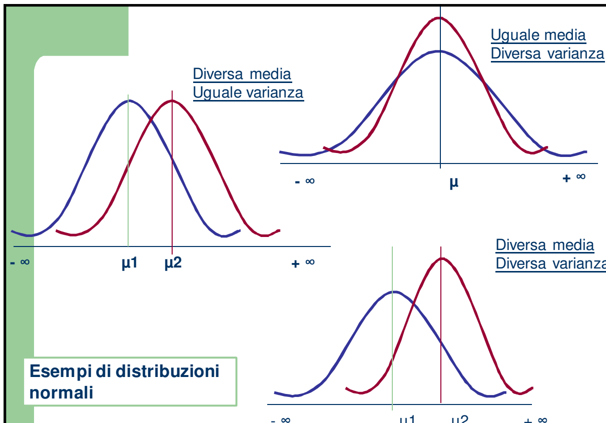
La curva normale (o curva di Gauss) può essere immaginata come un istogramma in cui ad ogni singolo valore di X sull'ascissa corrisponde un rettangolo (sottilissimo) che rappresenta la frequenza sull'ordinata. L'area sottesa all'intera curva (in verde) è quindi uguale alla somma di tutte le frequenze, cioè al 100% dei casi. La probabilità, quindi, per un valore di X di ricadere in un punto qualsiasi dell'ascissa è uguale a 1. Poiché la curva è simmetrica rispetto alla media, nella metà di destra o di sinistra ricade in totale il 50% dei casi, e la probabilità per un singolo valore di X è uguale a 0,5.



10

- La distribuzione normale, è una distribuzione teorica di probabilità, cioè non esiste nella realtà ma è calcolata con formule matematiche.
- Tale distribuzione **riguarda le variabili continue,** ciò vuol dire che essa è definita su tutto l'asse dei numeri reali, da -∞ (meno infinito) a +∞ (più infinito).
- La **proprietà che caratterizza** la curva normale è di essere definita da una particolare equazione che permette di calcolare, data un'ordinata, porzioni di area sottese alla curva stessa.
- La **curva normale** varia al variare di **μ** e di **σ (media** e **deviazione standard** della popolazione**)**. Si tratta dunque di una **famiglia di distribuzioni.**

11



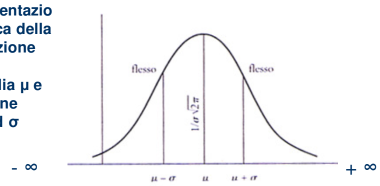
Esempi di distribuzioni normali

12

Caratteristiche della curva normale

- È **simmetrica** rispetto a **μ** e **unimodale** (moda, media e mediana coincidono), con la caratteristica **forma a campana**
- Un solo valore appare con la probabilità massima, quello centrale. La probabilità diminuisce man mano che ci si allontana dal valore centrale

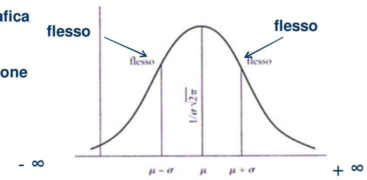
Rappresentazione grafica della distribuzione normale con media μ e deviazione standard σ



13

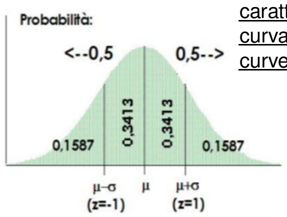
- La distribuzione normale è **asintotica all'asse delle ascisse** (tende ad infinito), cioè la curva si avvicina sempre di più all'asse delle ascisse senza mai raggiungerlo;
- La curva presenta **due flessi (punto in cui la curva passa da concava a convessa)** uno alla distanza di una deviazione standard sotto la media ed uno alla distanza di una deviazione standard sopra la media.

Rappresentazione grafica della distribuzione normale con media μ e deviazione standard σ



14

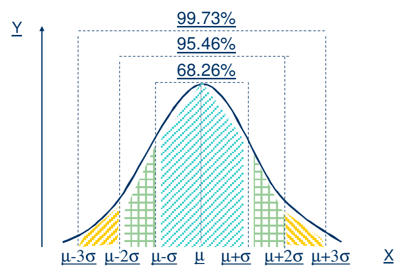
La percentuale di casi compresa tra due ordinate qualsiasi <u>è fissa ed è nota</u>, pertanto, è possibile calcolare la probabilità che un valore di X sia compreso all'interno di un determinato intervallo. Ad esempio è noto che tra la media (μ) ed una deviazione standard (σ) sopra la media ricade il 34,13% dei casi, pertanto un qualsiasi valore di X compreso tra μ e μ+σ avrà una probabilità casuale di 0,3413 (la stessa cosa vale per un valore di X compreso tra μ e μ-σ). <u>Questa particolare caratteristica distingue la curva normale da altre curve simili.</u>



15

Porzioni della distribuzione comprese tra ± 1,2,3 deviazioni standard da μ (in %)

Le aree sottese a questi intervalli **<u>sono sempre uguali</u>** in tutte le distribuzioni della famiglia delle curve normali.



16

La distribuzione normale standardizzata

Per gli usi pratici della distribuzione normale, si ricorre alla trasformazione della curva normale in **curva normale standardizzata**, che ha **μ = 0 e σ =1.** Per trasformare qualsiasi distribuzione normale in normale standardizzata si devono standardizzare i punteggi con la formula:

$$z = \frac{X - \mu}{\sigma}$$

μ = media della popolazione

σ = dev. st. della popolazione

Grazie a tale trasformazione è possibile utilizzare la **tavola della distribuzione normale standardizzata** per **individuare porzioni di area sottesa alla curva**, oppure per trovare **i valori di z (e quindi di X) che delimitano determinate aree**.

17

L'individuazione di una porzione di area sottesa alla curva ci permette di conoscere:

1. La proporzione di casi compresi in quella porzione di area
2. La percentuale di casi compresi in quella porzione di area
3. La probabilità di ottenere per caso un punteggio z (e dunque il punteggio X corrispondente) compreso in un certo intervallo.

18

Esempio

Consideriamo il punteggio ad un test. Sul manuale del test è indicato che i punteggi del campione normativo si distribuiscono normalmente con media μ = 50 e deviazione standard σ = 10.

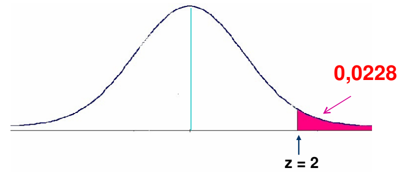
Se somministriamo il test ad un bambino scelto a caso, qual è la probabilità che il suo punteggio sia superiore a 70?

Standardizziamo il punteggio X = 70

$$z = \frac{X - \mu}{\sigma} \rightarrow \boxed{z = (70-50)/10 = 2}$$

19

L'area che corrisponde alla probabilità che ci interessa è quella evidenziata in rosa (l'area alla destra di z = 2).



Sulla tavola della curva normale standardizzata troviamo che l'area alla destra di z = 2 corrisponde ad una proporzione di 0,0228 di tutta la curva, cioè il 2,28%. Quindi la probabilità che un bambino scelto a caso ottenga un punteggio superiore a 70 è del 2,28%. La probabilità è bassa, perché abbiamo considerato un punteggio lontano dalla media (50).

20

Possiamo calcolare la probabilità che un bambino "molto bravo" in quel test ottenga un certo punteggio?

No, perché i valori di probabilità della curva normale si riferiscono solo ad eventi casuali.

21

Importanza della distribuzione normale

- Molti dei fenomeni di cui si occupano le scienze del comportamento hanno una distribuzione normale o che si approssima alla normale.
- Molti test statistici usati per la verifica delle ipotesi nella inferenza statistica hanno distribuzioni che tendono ad assumere una forma "normale" e la statistica inferenziale è basata in gran parte sulle proprietà della normale.

22