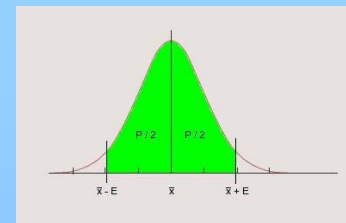
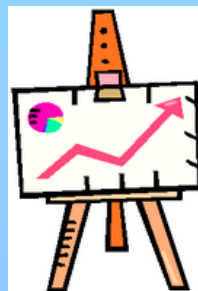
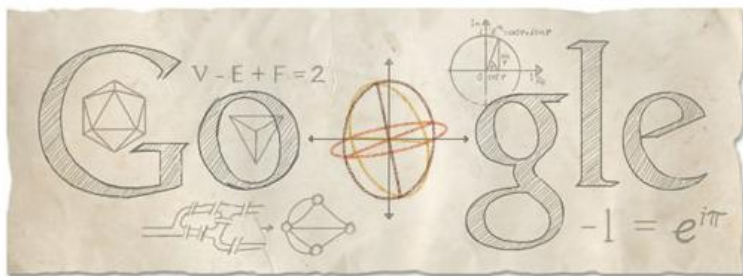




AGORÀ MATEMATICO

3 MAGGIO 2013





DATI e PREVISIONI

Statisticamente improbabile?



Giuliana Gnani
Dipartimento di Matematica

DIDATTICA DELLA STATISTICA

Problematiche

- **Formazione dei docenti**
- **Metodi didattici**
 - Uso di dati reali, contesti familiari agli studenti
 - Approccio storico

TEMI DELLA CONFERENZA

- Introduzione
- Spunti didattici sui concetti di base per un apprendimento integrato
- Materiali didattici tratti da fatti reali
- Legame tra didattica e ricerca

ASPETTI PROBLEMATICI NELL'INSEGNAMENTO DELLA STATISTICA

- Metodo diverso da quello della matematica (Dati-numeri, metodo induttivo-metodo deduttivo)
- Come inserire la statistica (descrittiva)in un programma di matematica già molto vasto

ASPETTI PROBLEMATICI NELL'INSEGNAMENTO DELLA STATISTICA

- Porre un'attenzione particolare per favorire un apprendimento di tipo **interdisciplinare**, offrendo spunti che mostrano come i concetti statistici sono collegati alle scienze sociali , sperimentali e alla matematica

DEFINIZIONI DI PROBABILITÀ A CONFRONTO

- Definizione classica -a priori
- **Limiti della definizione:** lo spazio degli eventi elementari è finito- equiprobabilità nelle ipotesi
- Definizione frequentista– legge empirica del caso- a posteriori-la frequenza come stima della probabilità
- **Limiti della definizione:** tante prove-stesse condizioni
- **Aspetti positivi:** definizioni operative

ALTRE IMPOSTAZIONI

- Approccio soggettivista (grado di fiducia nel verificarsi dell'evento - è il prezzo equo che sei disposto a pagare per ottenere 1 in caso di successo e 0 in caso di insuccesso)
- Approccio assiomatico
- Approccio logico (dalle algebre di Boole alle algebre polivalenti o di Lukasiewicz (1920))

IL CONCETTO DI DISTRIBUZIONE

- Da Maria Pannone, CIRDIS, Università di Perugia un esempio di distribuzione e sue rappresentazioni.

Tabella. La percentuale di maschi di 15 anni e più in sovrappeso in Italia alla data indicata.

	Classi di età			
	15-24	25-44	45-64	65 e più
Italia	17,8	38,0	50,6	48,6

IL CONCETTO DI DISTRIBUZIONE

- Mancata comprensione del concetto di distribuzione

=> errori nelle rappresentazioni grafiche e nella determinazione dei valori medi.

Tabella. La percentuale di maschi di 15 anni e più in sovrappeso in Italia alla data indicata.

	Classi di età			
	15-24	25-44	45-64	65 e più
Italia	17,8	38,0	50,6	48,6

La tabella rappresenta **quattro collettivi** distinti classificati rispetto al carattere “sovrappeso”.

- L'insidia sta nell'interpretare i dati come un'unica distribuzione di frequenza.

IL CONCETTO DI DISTRIBUZIONE

- Il grafico mostra proprio questo sbaglio:

Popolazione italiana di maschi in sovrappeso alla data 1999-2000 suddivisi per classe di età

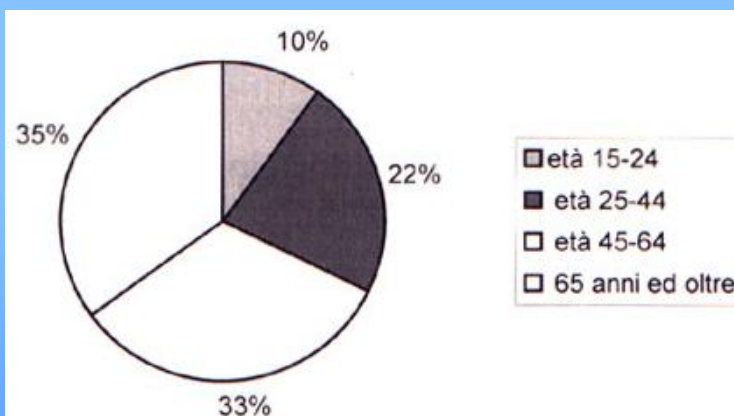


Tabella. La percentuale di maschi di 15 anni e più in sovrappeso in Italia alla data indicata.

	Classi di età			
	15-24	25-44	45-64	65 e più
Italia	17,8	38,0	50,6	48,6

L'errore si poteva evitare notando che:

- la somma dei dati della tabella è diverso da 100%,
- i dati sul grafico non corrispondono a quelli della tabella.

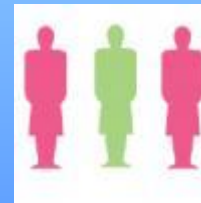
L'errore è stato facilitato dall'uso delle **procedure automatiche del software**

RAPPRESENTAZIONI GRAFICHE

Probabilità e statistica: sono ovunque, ma sono difficili da capire e può essere contro-intuitivo.

Allora, qual è il modo migliore di comunicare a un pubblico che non ha il tempo di desiderio, o di sfondo di rimanere bloccati nei numeri?

Ian Short esplora le moderne tecniche di visualizzazione e ritiene che la foto a destra davvero può valere più di mille parole



- [the Understanding Uncertainty website](http://theUnderstandingUncertaintywebsite)
- understandinguncertainty.org/

COME COSTRUIRE UN ISTOGRAMMA

Si consideri come esempio la distribuzione seguente costruita dalla prof.ssa Balestra tratto dal testo di Marco Abate- *MATEMATICA e STATISTICA-Le basi per le scienze della vita*, edizioni McGraw-Hill, Milano, 2009

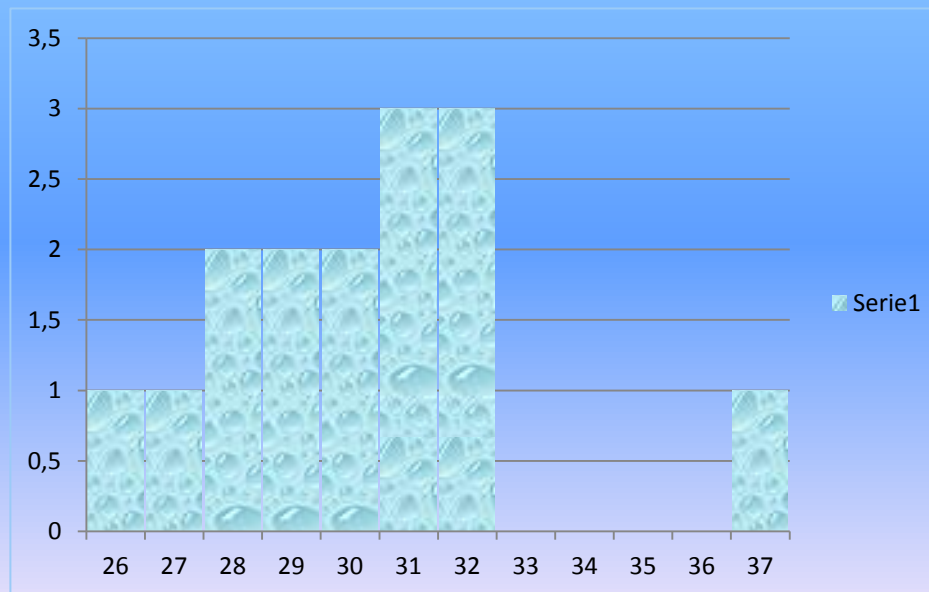
In un esperimento vengono pesate 15 cavia ottenendo i seguenti pesi

28,32,37,29,31, 30,32,26,32,27,29,30,28,31,31

Si vuole rappresentare la distribuzione dei pesi delle cavia

Una prima rappresentazione naturale è un istogramma con una colonna per ogni peso ottenuto, la cui altezza indichi la frequenza di quel peso

Si suddividono i pesi possibili in intervalli di lunghezza 1 ($25,5 \leq x < 26,5$ e $26,5 \leq x < 27,5$) e si mette su ciascun intervallo una colonna di altezza (o area, visto che gli intervalli hanno tutti la stessa lunghezza) proporzionale al numero di cavia con peso ottenuto in quell'intervallo



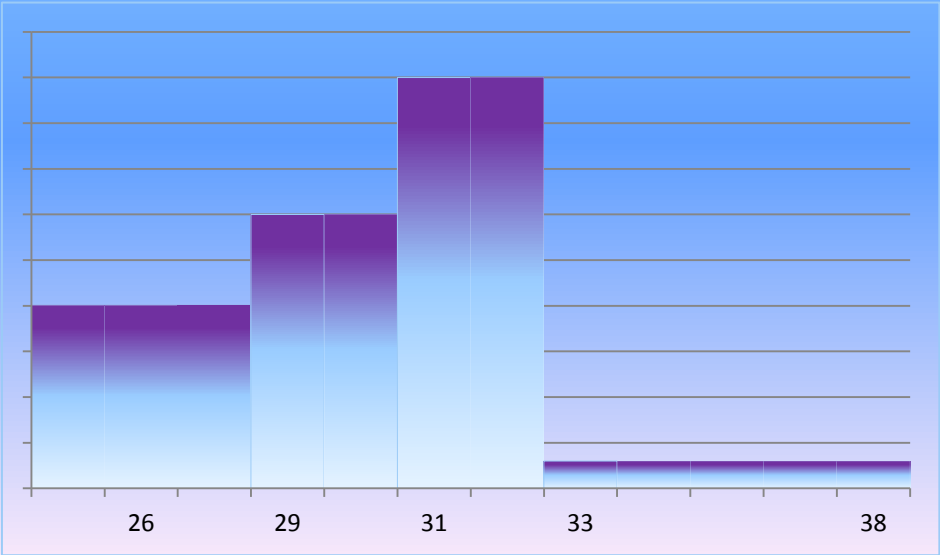
Se invece prendiamo gli intervalli

	F
$26 \leq x < 29$	4
$29 \leq x < 31$	4
$31 \leq x < 33$	6
$33 \leq x < 38$	1

E rappresentiamo la frequenza usando l'AREA

	F	$x_{i-1} - x_i$	$h = F / x_{i-1} - x_i$
$26 \leq x < 29$	4	3	1,33
$29 \leq x < 31$	4	2	2,00
$31 \leq x < 33$	6	2	3,00
$33 \leq x < 38$	1	5	0,20
	15		

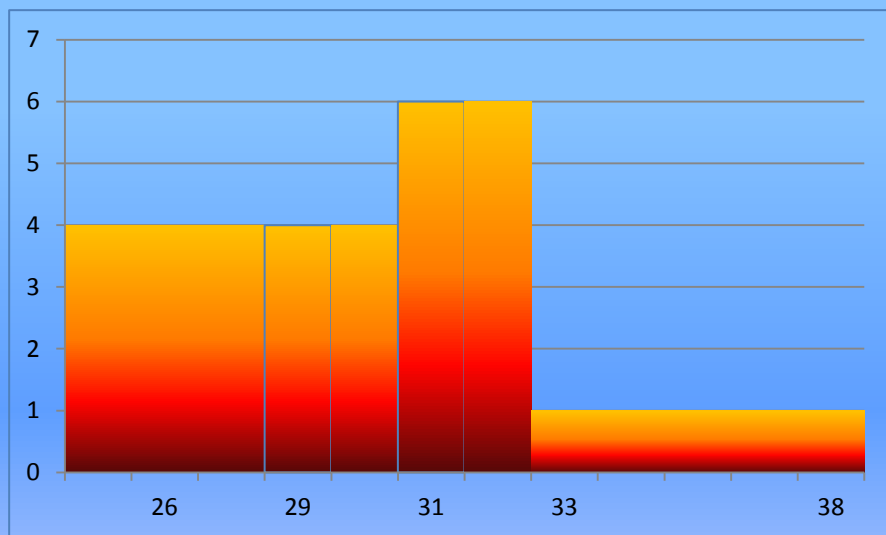
Otteniamo questo grafico



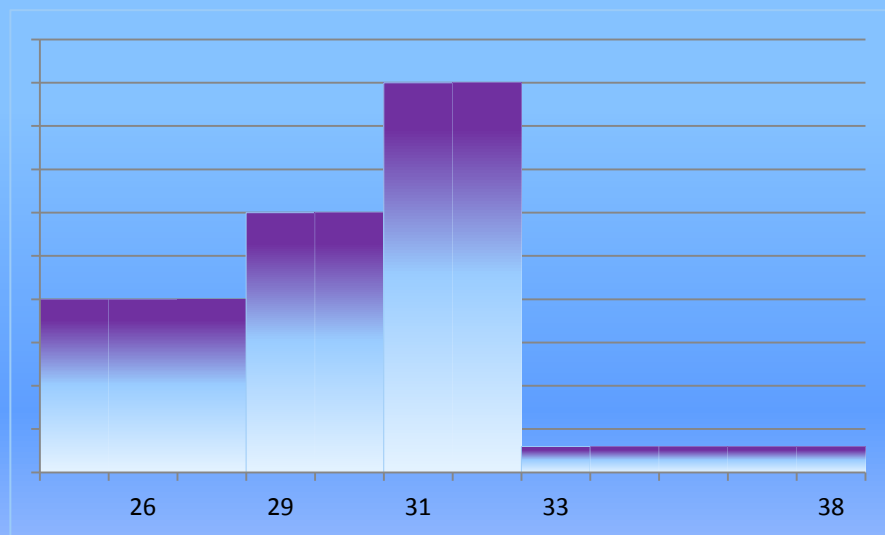
Se avessimo rappresentato la frequenza usando l'altezza avremmo ottenuto

	F
$26 \leq x < 29$	4
$29 \leq x < 31$	4
$31 \leq x < 33$	6
$33 \leq x < 38$	1

...frequenza usando l'ALTEZZA

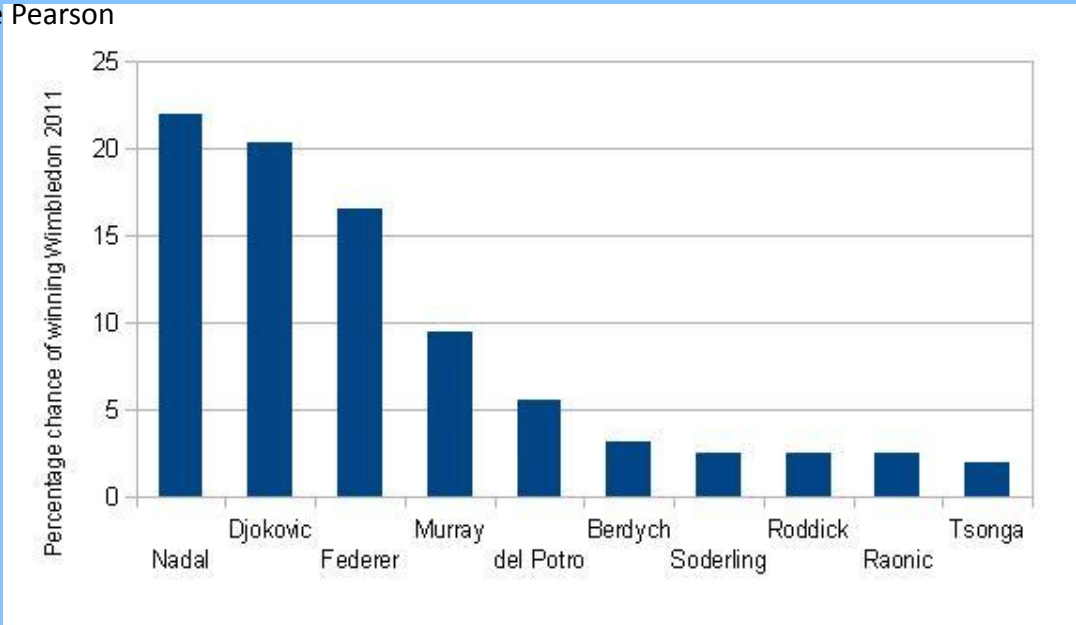


... frequenza usando l'AREA

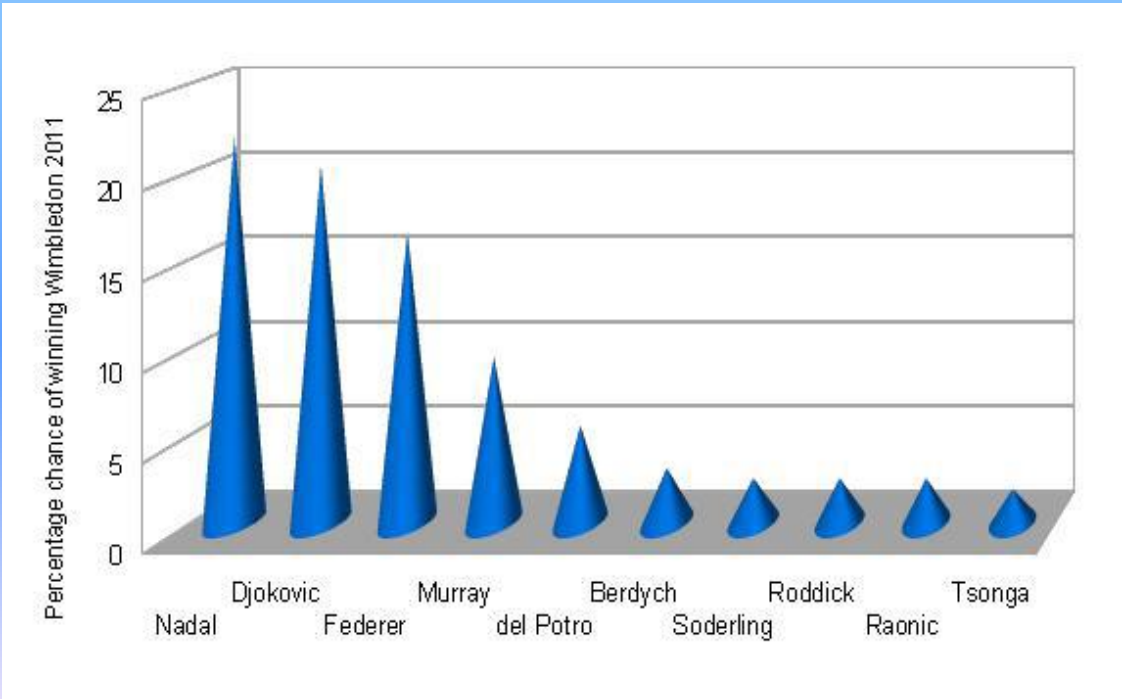


Sembra che ci siano molte più cavie nell'intervallo fra 33 e 38 di quante ce ne siano in realtà

Mike Pearson



Ian Short
Mike Pearson
William Hill



USO IMPROPRIO DEI GRAFICI STATISTICI

La regola aurea dei tre quarti

Un caso significativo è citato nel volume di R.P. Runton e A. Haber *Fondamenti di statistica* (1986) da cui si può trarre molta documentazione. Questa procedura consiste nell'eliminazione della frequenza zero dell'asse verticale ponendo l'origine in corrispondenza di un punto dell'asse delle ordinate, corrispondenti ad una frequenza maggiore di zero.

INDIPENDENZA STOCASTICA

INDIPENDENZA LOGICA

a confronto

Definizione 1. Due eventi A e B si dicono **indipendenti** se $P(A \wedge B) = P(A) * P(B)$

Definizione 2. Due eventi si dicono **indipendenti** se la conoscenza del verificarsi dell'uno non modifica la probabilità del verificarsi dell'altro

Definizione 3. Due eventi si dicono **indipendenti** se il verificarsi di A non implica né il verificarsi di B né il verificarsi di B' e simmetricamente il verificarsi di B non implica né il verificarsi di A né il verificarsi di A'

SECONADARIE DI 1^ GRADO SONO UN ARGOMENTO NUOVO?

In geometria

1. Dato il rettangolo di dimensioni x e y , determinare il quadrato equivalente

Si tratta di determinare la lunghezza l del lato del quadrato: il quadrato ha area

$$l^2 = xy, \text{ da cui } l = \sqrt{xy}$$

2. *Dato il rettangolo di dimensioni x e y , determinare il quadrato isoperimetrico*

Si tratta di determinare la lunghezza l del lato del quadrato: il quadrato ha semiperimetro

$$2l = x + y \quad \text{da cui} \quad l = (x + y) / 2$$

3. *Dato il rettangolo di dimensioni x e y , determinare il rapporto tra area e semiperimetro del quadrato*

Si considera $xy / [(x + y) / 2]$, da cui $2 / (1/x + 1/y)$

DALLE PROPORZIONI

1. *Dati i numeri x, y , determinare il numero z medio proporzionale tra x e y ,*

cioè $x:z=z:y$, da cui $z=\sqrt{xy}$

2. *Dati i numeri x, y , determinare il numero z tale che $x-z=z-y$, da cui $z=(x+y)/2$*

3. *Dati i numeri x, y , determinare il numero z tale $(z-x):x=(y-z):y$, da cui $z=2xy/(x+y)$*

MEDIA GEOMETRICA, MEDIA ARITMETICA, MEDIA ARMONICA

- Altri esempi dalle scienze sperimentali?
 1. Tasso medio di rendimento ,la crescita media
 2. Passaggio a nuova unità di misura
 3. Velocità media

ESPERIENZE NELL'INSEGNAMENTO DELLA STATISTICA

- Il progetto mat@abel
- Dalla sezione dedicata alla **didattica della Statistica** nel sito della SIS è possibile scaricare **materiali e strumenti didattici** e raggiungere altri siti interessanti

<http://www.sis-statistica.it/>

- UMI – SIS La Matematica per il cittadino
- ISTAT <http://www.istat.it>



<http://www.nuffieldfoundation.org/uncertainty-1>



Contiene un percorso destinato alla formazione degli insegnanti su matematica e incertezza, in un approfondito contesto teorico.

Piaget e Inhelder(1975) segnalano come la domanda intellettuale di costruzione di conoscenza probabilistica includa non solo la comprensione di casualità e di proporzione, ma anche calcolo combinatorio.

Fischbein(1975) studia il contributo della intuizione nella costruzione dei concetti chiave della Probabilità

Prat e Noss(2002)approfondiscono lo studio della evoluzione delle conoscenze

ESPERIENZE IN AMBITO INTERNAZIONALE

- Ampio elenco sul sito della SIS e dell'ISTAT
- Da segnalare
 - IASE e International Statistical Literacy Project (materiale per gli insegnanti da usare in classe o per migliorare la conoscenza della Statistica , per valutare le conoscenze acquisite dagli studenti e molto altro!)
 - Statistics Canada materiale di supporto per insegnanti e studenti <http://www.statcan.gc.ca/edu/index-eng.htm>
 - Royal Statistical Society due riviste Teaching Statistics e Significance e varie risorse <http://www.rss.org.uk>



- **CONSORTIUM FOR THE ADVANCEMENT OF UNDERGRADUATE STATISTICS EDUCATION**

Cause è una organizzazione USA che ha lo scopo di migliorare l'insegnamento della Statistica, promuovendo lo sviluppo professionale dei docenti e il rapporto tra didattica e ricerca

- <http://www.causeweb.org/>
contiene materiale didattico



Swivel

- miniera di dati
- nato in California nel 2006
- permette di esplorare i dati , condividere con altri le proprie conclusioni.

Swivel contiene dati sulla politica, l'economia, il tempo, gli sport, gli affari e altri ancora

<http://www.swivel.com/>

ELENCO DI ALCUNE DIFFICOLTÀ (IN BASE ALLA LETTERATURA)

- 1) difficoltà a distinguere il concetto di **carattere** da quello di **frequenza**
- 2) difficoltà a raggruppare opportunamente i dati in classi
- 3) difficoltà riscontrabili nel **passaggio dal discreto al continuo** relative alla non consapevolezza dell'importanza della numerosità del campione, (in quanto solo con popolazioni ampie vale la legge dei grandi numeri e la convergenza della distribuzione discreta verso quella continua)
- 4) difficoltà tipica del pensiero probabilistico, ovvero l'idea che una successione "regolare" di uscite sia più improbabile di una uscita meno regolare
- 5) difficoltà di comprensione della differenza fra **fenomeno statistico** ed **evento aleatorio**
- 6) In generale, difficoltà legate ad un atteggiamento di pensiero, che potrebbero condizionare la vita sociale molto di più rispetto ai tradizionali concetti matematici

STATISTICAMENTE IMPROBABILE?

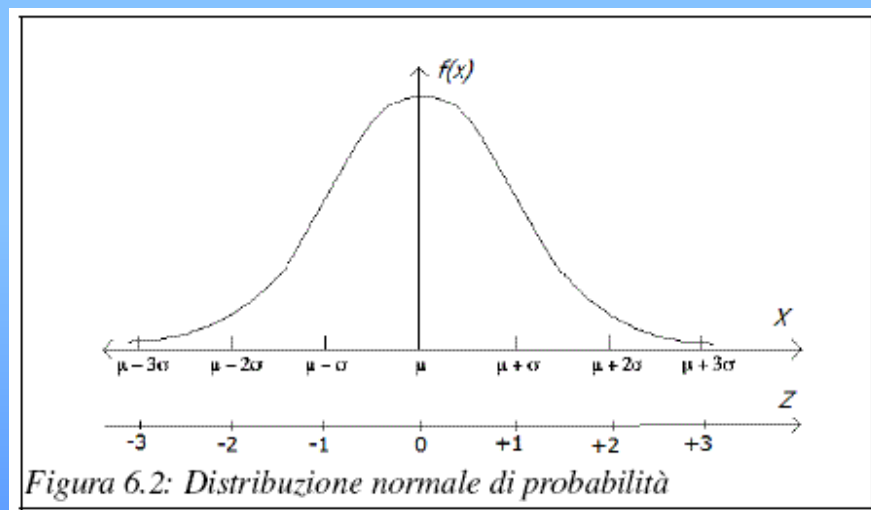
Statistica inferenziale e Calcolo delle probabilità
sono gli strumenti idonei per prendere
decisioni in situazioni di incertezza

- -uso di modelli e di test
- - studio dei campionamenti

ALCUNE DOMANDE

- Perché si considera in ambito sperimentale la media aritmetica (semplice o ponderata)
- Perché è importante associare alla media aritmetica lo scarto quadratico medio
- Importanza della distribuzione normale

Distribuzione normale



E PER FINIRE

- Significato della legge dei grandi numeri
- Confronto con la legge empirica del caso

PER VEDERE LA MATEMATICA IN MODO PROBLEMatico

- Vinicio Villani, Claudio Bernardi, Sergio Zoccante, Roberto Porcaro, *Non solo calcoli. Domande e risposte sui perché della matematica*, Springer Italia, Milano 2012,

Collana Convergenze (promossa dalla CIIM-Commissione Italiana per l'Insegnamento della Matematica, organo permanente dell'UMI-Unione Matematica Italiana)

Il libro è diviso in tre parti:

Parte I Teoria degli insiemi e Logica matematica (Claudio Bernardi)

Parte II Analisi matematica (Sergio Zoccante)

Parte III Probabilità e Statistica (Roberto Porcaro)

Un ringraziamento a tutti i partecipanti

In particolare ringrazio la prof.ssa Balestra per gli spunti e le riflessioni e ringrazio i corsisti del TFA che hanno lavorato nella preparazione di u.d. sui temi proposti nella conferenza :
ho cercato di *vedere* attraverso le difficoltà da loro incontrate