Alcune v.a. discrete notevoli

Variabile aleatoria Bernoulliana

Il risultato X di un esperimento aleatorio può essere classificato nel modo che segue:
"successo" oppure "insuccesso".

Indichiamo:
Successo = 1,
Insuccesso = 0.

La "probabilità di successo" è nota:

$$p(1) = P(X = 1) = p$$

Di conseguenza, sarà:

$$p(0) = P(X = 0) = q = 1 - p \text{ (evento complementare)}$$

Esperimento Bernoulliano

- Un esperimento aleatorio di questo tipo si dice esperimento bernoulliano.

Variabile aleatoria Bernoulliana

- Il valore p, probabilità dell'evento "successo", è detto parametro della v.a. di Bernoulli.
- Una v.a. è descritta attraverso i suoi parametri (uno o più).
- Si scrive $X \sim Ber(p)$

Valore atteso di una v.a. Bernoulliana

- Applichiamo la formula del v.a.:

$E[X] = \Sigma x_i * p(x_i) =$

$= 0*(1 - p) + 1*p = p$

Varianza di una v.a. Bernoulliana

- Sappiamo che $E(X) = p$
- $E[X] = \Sigma(x_i - E(X))^2 * p(x_i) =$

$= (0-p)^2 * (1-p) + (1-p)^2 * p =$

...dopo alcuni passaggi...

$= p(1-p)$

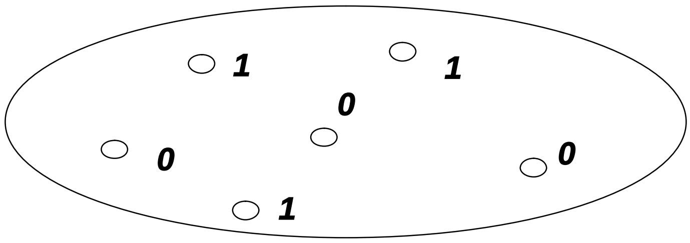
V.a. binomiale

- Il passo successivo consiste nella domanda:
- Se io ripeto n volte un esperimento bernoulliano, quante volte (x) su n ripetizioni otterrò un evento "successo"?

V.a. binomiale

- In altri termini, cosa accade se io ripeto n volte una v.a. di Bernoulli?

Lo spazio campione Ω



Le possibili n-ple

- Quante sono le possibili n-ple, non ordinate, nello spazio campionario Ω, di "1" e "0"?
- Con n = numero di prove
- x = numero di successi
- Questo numero è dato dal coefficiente binomiale:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

V.a. binomiale

- Sia X = “numero di successi su n prove bernoulliane indipendenti”; allora i valori di X sono
- x = 0, 1, ..., n   e
- X è una v.a. discreta con funzione di probabilità:

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}$$

Valore atteso di una v.a. binomiale

- Si scrive X ~ Bin(n,p)
- I parametri sono due.

- Valore atteso
- $E(X) = \Sigma E(X_i) = n*E(X_i) = np$

Varianza di una v.a. binomiale

- Poiché sappiamo che le n prove bernoulliane sono tra loro indipendenti, vale la regola:
- $Var(X) = \Sigma\, Var(X_i)$
- Pertanto:
- $Var(X) = \Sigma\, Var(X_i) = n*p*(1-p) =$

$= np(1-p)$

Esempio: 4 lanci di una moneta

Numero di "Testa"	Eventi possibili	Prob.di "Testa"
0	CCCC	0,0625
1	TCCC,CTCC, CCTC,CCCT	0,250
2	TTCC, TCTC,TCCT,CTTC CTCT,CCTT	0,3750
3	CTTT,TCTT, TTCT,TTTC	0,250
4	TTTT	0,0625

Esempio: 4 lanci di una moneta

- Stabiliamo che:
- Successo = Testa = 1
- Poiché la moneta non è truccata, si ha:
- p = ½ ; 1-p = ½
- Stabiliamo poi che n = 4.

Esempio: 4 lanci di una moneta

Con n = 4, le “disposizioni con ripetizione” possibili di “teste” e di “croci” (successi e insuccessi)
sono date da:

$$2^n = 2^4 = 16$$

Per una nota regola del calcolo combinatorio.

Poiché è n =4; p = ½

La f. di probabilità assume la forma:

$$P(X=x)=\binom{n}{x}\cdot p^x\cdot(1-p)^{n-x}=\binom{4}{x}\cdot(1/2)^x\cdot(1/2)^{n-x}$$

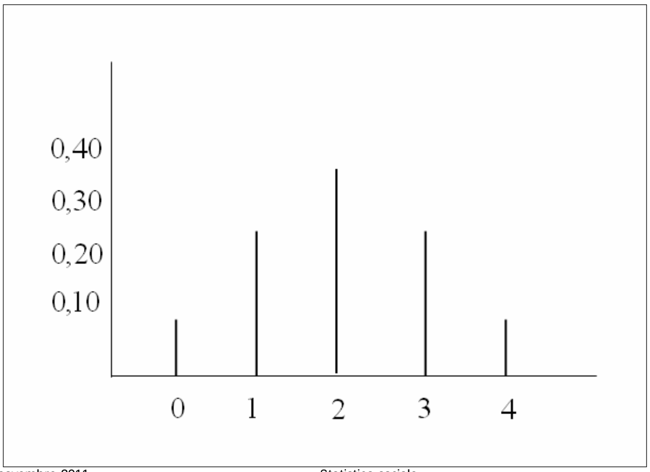
Pertanto, se vogliamo calcolare, ad esempio, p(3),
Avremo:

$$P(X=3)=\binom{4}{3}\cdot(1/2)^3\cdot(1/2)^1=\frac{4!}{3!(4-3)!}\cdot(1/2)^3\cdot(1/2)^1=$$

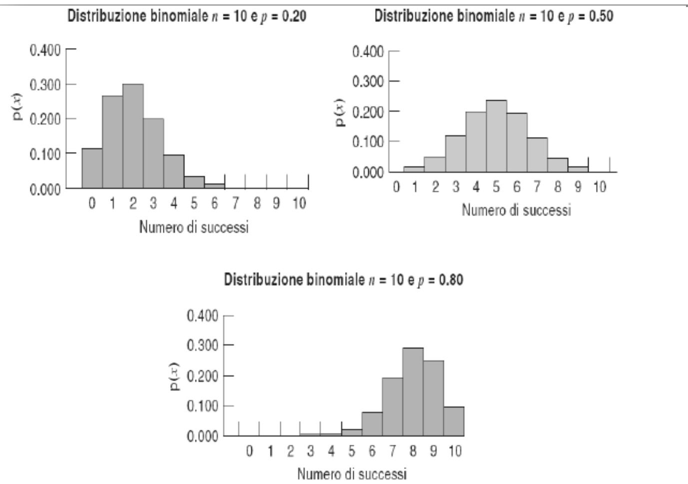
$$=\frac{4!}{3!\cdot 1!}\cdot(1/8)\cdot(1/2)=\frac{4\cdot 3\cdot 2}{3\cdot 2}=4\cdot\frac{1}{16}=\frac{1}{4}=0{,}25$$

Che era appunto il valore che avevamo già visto nella tabella.

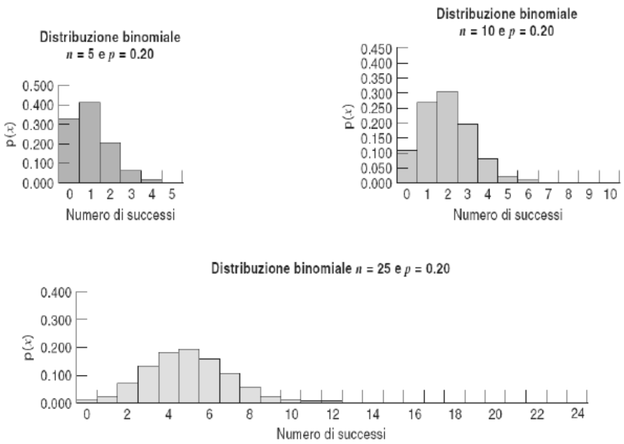
In termini grafici:



La forma della v.a. binomiale varia a seconda del parametro p;
è simmetrica se p = ½



La forma della v.a. binomiale varia anche a seconda dell'altro parametro, cioè il numero di prove n; si noti che la distribuzione tende a diventare simmetrica all'aumentare di n, anche se p è diversa da ½.



Usi della v.a. binomiale

- Tra gli anni ‘60 e ‘80 del ‘900 le giurie popolari negli stati del sud degli USA erano composte da non più di 4 giurati afro-americani.
- Possiamo ritenere che i giurati fossero eletti in modo “casuale”, se sappiamo che la popolazione in quegli stati e in quel periodo era per circa il 50% composta da persone afro-americane?
- Sia X la variabile casuale che conta il numero di afro-americani in una giuria di 80 persone;
- X è una variabile casuale Binomiale con parametri n = 80 e p = 0.5 (ricordiamo che la pop. è per il 50% afro-americana)

Usi della v.a. binomiale

- Si fanno 80 estrazioni e il “successo” è rappresentato dall’evento “persona afro-americana estratta”.
- Calcoliamo la probabilità che in una giuria americana vi siano meno di 4 persone afro-americane:

$$\begin{aligned} P(X \leq 4) &= P(X=0) + P(X=1) + P(X=2) + \\ &+ P(X=3) + P(X=4) \\ &= \sum_{k=0}^{4} \binom{80}{k} 0.5^k 0.5^{80-k} \\ &\simeq 0.0000000000000000014 \end{aligned}$$

Questa probabilità è più piccola della probabilità di fare 3 scale reali consecutive nel gioco del poker!

V.a. uniforme discreta

- Un esperimento può avere k possibili esiti (k>2), ciascuno dei quali ha la stessa probabilità 1/k. Solitamente i k eventi equiprobabili sono rappresentati con i numeri interi 1,2,...,k
- Con k=6 si ha il classico esempio del lancio di un dado.
- La v.a. discreta che rappresenta questo tipo di esperimento si dice uniforme discreta, ed ha la seguente funzione di p.
- $P(X = x) = 1/k, \quad x = 1, 2, ..., k$

V.a. uniforme discreta

- La funzione di ripartizione è:

$$F(x) = 0 \quad (\text{per } x < 1)$$
$$= [x]/k \quad (\text{per } 1 \leq x < k)$$
$$= 1 \quad (\text{per } x \geq k)$$

- Valore atteso e Varianza:

$E(X) = (k+1)/2$;

$Var(X) = (k^2+1)/12.$

V.a. di Poisson (eventi rari)

- Un esperimento consiste nel contare il numero X di volte in cui si verifica un determinato evento in un intervallo di tempo prefissato.
- Supponendo che il numero medio di volte in cui esso si verifica nell'intervallo definito sia pari a λ (non necessariamente intero), il numero X segue una v.a. di Poisson di parametro λ.
- X ~ P(λ)

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \; x = 0, 1, 2, \ldots$$

- Valore atteso e varianza sono entrambi uguali al parametro λ.

$E(X) = \lambda$; $E(X^2) = \lambda(\lambda + 1)$;
$Var(X) = E(X^2) - [E(X)]^2 = \lambda(\lambda+1) - \lambda^2 = \lambda$.

La v.a. di Poisson approssima la v.a. Binomiale quando n è grande e p è molto piccolo

- n = 100
- p = 1/100; λ = np = 1

k	0	1	2	3	4	5
Binomiale	.366	.370	.185	.0610	.0149	.0029
Di Poisson	.368	.368	.184	.0613	.0153	.00307

Esempio

- Numero di incidenti al giorno sull'autostrada Bologna-Milano ~ Poisson (λ=1).
- Qual è la probabilità che oggi si verifichino due incidenti?

k	0	1	2	3	4	5
Binomiale	.366	.370	.185	.0610	.0149	.0029
Di Poisson	.368	.368	.184	.0613	.0153	.00307

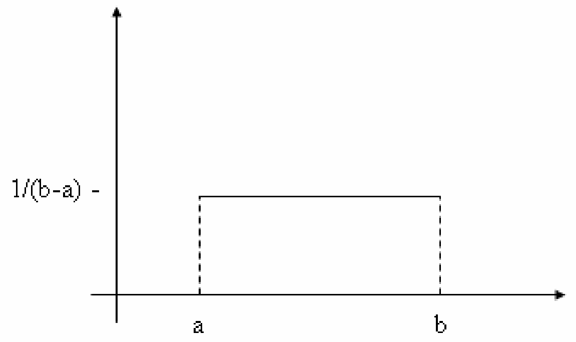
V.a. uniforme continua (o rettangolare)

- Una v.a. uniforme, definita sull'intervallo reale *[a, b]* , assegna la stessa densità a tutti i valori dell' intervallo stesso. Poiché l'area sottesa dalla funzione di densità deve essere uguale a 1, la densità di ciascun punto è il reciproco della lunghezza dell'intervallo:

- f(y) = 1/(b-a), a < y < b

- Il valore atteso e la varianza di una v.a. uniforme continua sono:
  E(Y) = (a+b)/2 ;
  Var(Y) = (b-a)^2 /12

In termini grafici (funzione di densità)



La v.a. normale

La v.a. normale o “gaussiana”

- La distribuzione normale o “gaussiana” riveste un ruolo fondamentale per l’inferenza statistica.

Carl Friedrich Gauss (1777 - 1855)

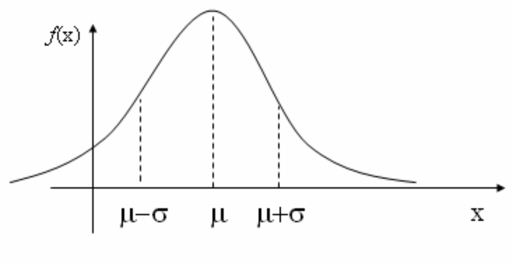
La v.a. normale ha due parametri

- Una v.a. X ha una distribuzione normale, con media μ e varianza σ^2, se:
  - ❑ la sua *funzione di densità di probabilità* è data da:

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

$$\left(-\infty < x < +\infty\right)$$

In termini grafici (funzione di densità)



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

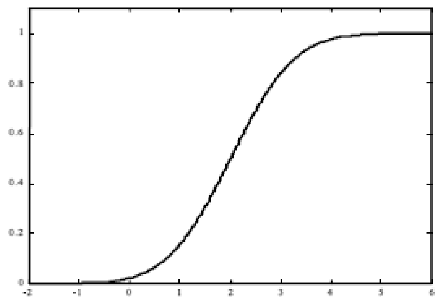
La v.a. normale

- e se la sua *funzione di ripartizione* è data da:

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

$$(-\infty < x < +\infty)$$

In termini grafici (funzione di ripartizione)



Nel grafico rappresentato nella figura,
$\mu = 2$; $\sigma^2 = 1$; cioè X ~ N(2,1)

La v.a. normale

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

- dove:

μ, σ^2 = parametri della v.a. normale;

π = la costante 3,14...;

e = numero di Nepero, base dei logaritmi naturali (2,7183...),

La v.a. normale

- La v.a. normale è continua;
- assume valori su tutto l'asse reale **R**, cioè compresi tra -∞ e +∞;
- descrive una curva di tipo *simmetrico,* di forma *campanulare*, che presenta le seguenti caratteristiche:

  - ❑ la curva è perfettamente simmetrica al livello di x in cui la funzione *f* (x) raggiunge il suo punto più alto, che è in corrispondenza di $x_i = \mu$;
  - ❑ Da ciò consegue che:
  - ❑ Media, mediana e moda coincidono.

La v.a. normale

- La sua funzione di densità *f* (x) è asintotica per **x=0**, con x che tende e -∞ e a +∞ (la curva si avvicina all'asse delle ascisse senza mai arrivare a toccarla);
- è crescente per valori della x che vanno da -∞ a *μ,* mentre è decrescente per valori della x che vanno da *μ* a +∞;
- è completamente caratterizzata dai due parametri *μ* e σ^2 e dalle tre costanti 2, π, *e.*

Valori caratteristici della v.a. normale

- I due parametri della v.a. normale, cioè *μ* e σ^2, corrispondono, rispettivamente, alla media E(X) e alla varianza Var(X) della distribuzione.
- Si può dimostra infatti che si ha:

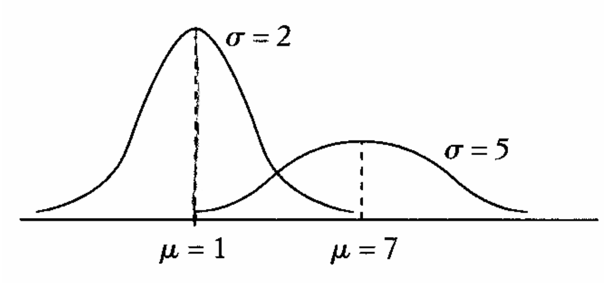
$$E(X) = \mu = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

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

Valori caratteristici della v.a. normale

- Dalle precedenti relazioni consegue che ogni v.a. normale è univocamente definita dalla propria media e dalla propria varianza.
- Varie distribuzioni normali, pertanto, possono differire per la media e la varianza, nonostante mantengano costanti le altre loro caratteristiche.

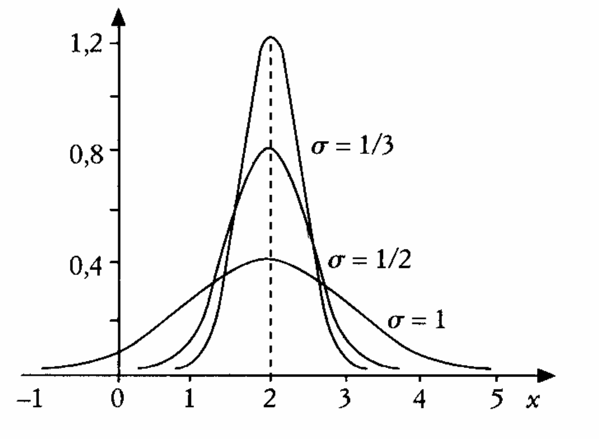
Valori caratteristici della v.a. normale: diversi valori di μ e di σ



Valori caratteristici della v.a. normale: diversi valori di μ e di σ

- Osserviamo come al variare della media e della varianza la curva subisca sia una traslazione sull'asse delle ascisse, sia un appiattimento.

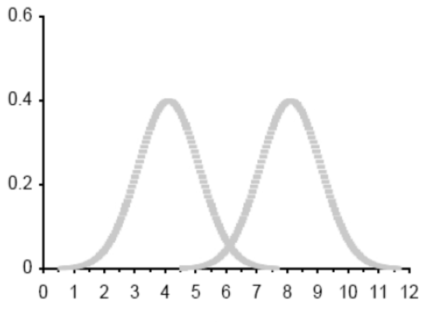
Valori caratteristici della v.a. normale: diversi valori di σ



Valori caratteristici della v.a. normale: diversi valori di σ

- Se, invece, facciamo variare soltanto la varianza, tenendo costante la media, la curva “si appiattisce” quando la varianza aumenta e diventa meno piatta quando la varianza diminuisce.
- La media rimane sempre la stessa.

Valori caratteristici della v.a. normale: diversi valori di μ



Si ha una semplice traslazione della curva.

La distribuzione normale standardizzata

- La distribuzione normale è difficilmente trattabile dal punto di vista calcolatorio, a causa dei suoi due parametri, μ e σ^2 .

- Il ricorso alla **"distribuzione normale standardizzata"** permette invece di individuare facilmente le probabilità relative agli intervalli di valori, utilizzando opportune *tavole statistiche*.