Lezione 13: Variabili Casuali discrete

Corso di Statistica
Facoltà di Economia
Università della Basilicata

Prof. Massimo Aria

- aria@unina.it

variabile casuale Uniforme Discreta

La prova che genera una v.c. Uniforme discreta si può assimilare all'estrazione di una pallina da un'urna che contiene *k* palline "identiche" e numerate da 1 a *k*.

L'aggettivo identico sta ad indicare la perfetta uguaglianza delle palline (in generale eventi) nella forma, peso, ecc.
Ciò fa si che ognuna di esse abbia la stessa probabilità di essere estratta.

Si pensi ad esempio alla prova "lancio di un dado", oppure "estrazione di una carta da un mazzo", ecc.

Si definisce quindi v.c. **Uniforme discreta**, la variabile X che assume valori x=1,2,…,k con probabilità costante pari a 1/k.

Essa si indica con

$$X \sim Ud(k)$$

con $P(X=x)=1/k$ per ogni x=1,2,…,k.

La v.c. Ud è simmetrica, non presenta moda e possiede media e varianza pari a:

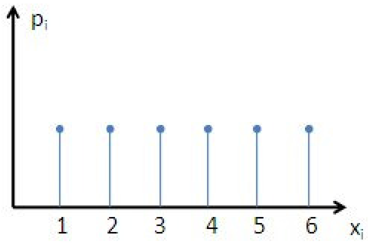
E(X)$=(k+1)/2$ **Var(X)**$=(n^2-1)/12$

Esempio: variabile casuale uniforme discreta

La prova consiste in un lancio di un dado. I possibili eventi sono k=6, equiprobabili

$$X \sim Ud(k=6)$$

Evento	X	Probabilità
	1	1/6
	2	1/6
	3	1/6
	4	1/6
	5	1/6
	6	1/6
		1



E(X)=7/2=3,5 **Var(X)**=$(6^2$-1)/12=2,917

variabile casuale di Bernoulli

La v.c. di Bernoulli trae origine da una prova nella quale interessa verificare esclusivamente se un evento E si è verificato (successo) oppure no (insuccesso).

Sono molteplici le situazioni reali in cui si incontra questo modello di probabilità.
Una qualunque prova dicotomica o dicotomizzabile può essere rappresentata da una v.c. di Bernoulli:
• *prove dicotomiche*.
Il verificarsi di "testa" nel lancio di una moneta, la scelta di una pallina bianca da un'urna che ne contiene di bianche e nere, ecc.
• *prove dicotomizzabili*.
Il verificarsi di un punteggio superiore al 5 nel lancio di un dado, l'estrazione di una pallina con numero pari da un'urna, un tasso di inflazione inferiore al 2% nel prossimo anno, ecc.

Si definisce quindi v.c. di **Bernoulli**, la variabile X che assume valore 1 (il "successo") con probabilità π (pi greco) e valore 0 (l'"insuccesso") con probabilità 1- π.

Essa si indica con

$$X \sim Ber(\pi)$$

con $P(X=x)= \pi^x(1- \pi)^{1-x}$ con x=0,1.

La v.c. di Bernoulli possiede media e varianza pari a:

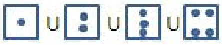
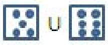
E(X)= π (la media coincide con la probabilità di successo)
Var(X)= π(1- π) (può assumere valori compresi tra 0 e 0,25, questo ultimo nel caso di massima incertezza)

Esempio: variabile casuale di Bernoulli

La prova consiste nel verificare se, in un lancio di un dado, il risultato è maggiore o uguale a 5.

I possibili eventi sono 2:
Successo (risultati 5 o 6) **Insuccesso** (risultati 1 o 2 o 3 o4)

$$X \sim Ber\left(\pi = \frac{2}{6}\right)$$

	Evento	X	Probabilità
"Insuccesso"		0	$(1-\pi)=\frac{4}{6}$
"Successo"		1	$\pi = \frac{2}{6}$
			1

E(X)=2/6=0,33 **Var(X)**=(0,33*0,67)=0,22

variabile casuale Binomiale

Se si ripete, per *n* volte e nelle medesime condizioni, lo schema *successo-insuccesso* della v.c. di Bernoulli, si genera una sequenza di *n* sottoprove indipendenti a ciascuna delle quali si può associare una v.c. di Bernoulli.
Tale modello prende il nome di v.c. Binomiale e può essere intesa come una somma di v.c. bernoulliane.

Nella realtà lo schema binomiale si può ritrovare ad esempio: nel lancio di *n* monete e si è interessati al numero di testa che si presenta complessivamente; in un'indagine in cui si intervistano n persone e si è interessati al numero di favorevoli ad una certa proposta di legge; ecc.

Si definisce quindi v.c. **Binomiale**, la variabile X che rappresenta il numero di successi che si verificano in una sequenza di *n* sottoprove indipendenti nelle quali è costante la probabilità π di un successo.
La v.c. binomiale dipende da due parametri, il numero *n* delle sottoprove e la probabilità π.

Essa si indica con

$$X \sim B\,(n,\pi)$$

con $P(X=x)=$ *[coefficiente binomiale]* $* \pi^{x}(1-\pi)^{n-x}$ con x=0,1,2,…,n.

La v.c. Binomiale possiede media e varianza pari a:

E(X)= $n*\pi$ (n volte la media della bernoulliana)
Var(X)= $n*\pi(1-\pi)$ (n volte la varianza della bernoulliana)

Esempio 1: variabile casuale Binomiale

La prova consiste nel verificare , nel lancio di 5 dadi, quanti di essi diano risultato 1.

I possibili eventi in una singola sottoprova sono:

Successo (risultato 1) **Insuccesso** (risultati 2 o 3 o 4 o 5 o 6)

$$X \sim B\left(n = 5, \pi = \frac{1}{6}\right)$$

Eventi	X	Probabilità
"nessun 1 nei cinque lanci"	0	0,4019
"un solo 1 nei cinque lanci"	1	0,4019
"due 1 nei cinque lanci"	2	0,1608
"tre 1 nei cinque lanci"	3	0,0322
"quattro 1 nei cinque lanci"	4	0,0031
"tutti 1 nei cinque lanci"	5	0,0001

E(X)=5*(1/6)=0,8333

Var(X)=5*1/6*5/6=0,6944

Dove la probabilità è calcolata come $P(X = x) = \binom{n}{x} \cdot \pi^x \cdot (1-\pi)^{n-x}$

Con $\binom{n}{x} = \frac{n!}{x! \cdot (n-x)!}$ ← Coeff. Binomiale:
numero di casi in cui si presentano x successi in n prove

Esempio 2: variabile casuale Binomiale

Un'urna contiene 15 palline di cui 5 rosse e 10 verdi. Calcolare la probabilità che estraendo 6 palline, reimmettendo ogni volta la pallina estratta nell'urna, 5 siano rosse. Calcolare inoltre, il valore medio e la varianza.

L'evento E, ossia l'evento che estraendo una pallina dall'urna sia rossa, ha probabilità di verificarsi ad ogni estrazione uguale a:

$$\boxed{p(E) = 5/15 = 1/3} \leftarrow \pi$$

mentre la probabilità dell'evento contrario $\bar{E}$, ossia dell'evento che ad ogni estrazione la pallina sia verde, è:

$$\boxed{q(\bar{E}) = 1 - p(E) = 1 - 1/3 = 2/3} \leftarrow 1-\pi$$

Pertanto, la probabilità che su 6 estrazioni 5 volte venga estratta una pallina rossa, è:

$$p_{6,5} = \binom{6}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^{6-5} = \frac{6!}{5!(1!)}\frac{1}{243}\frac{2}{3} = 6\frac{1}{243}\frac{2}{3} = \frac{4}{243} = 0{,}0164$$

Il valore medio e la varianza sono:

$$\mu = M(X) = 6\frac{1}{3} = 2, \qquad Var(X) = 6\frac{1}{3}\frac{2}{3} = \frac{4}{3} = 1{,}3\bar{3}$$

Condizioni di applicabilità del modello binomiale

La v.c. binomiale trova applicazione in prove dove sono verificate le seguenti condizioni:

• la prova è composta da *n* sottoprove indipendenti
Il risultato di una sottoprova non modifica la probabilità della successiva.

• ogni sottoprova è svolta sempre nelle medesime condizioni
La probabilità π del successo (in una singola sottoprova bernoulliana) è costante in tutte le n sottoprove.

variabile casuale di Poisson

Si consideri una prova che può avere solo due possibili esiti chiamati, generalmente, Successo (S) e Insuccesso (I). Si è interessati all'evento S.
In particolare si è interessati a contare quante volte si verifica l'evento S in un certo arco temporale prefissato (oppure anche in un certo ambito spaziale: ad esempio un'area prefissata).

Esempi possono essere:

• il numero di clienti che arrivano ad uno sportello bancario nell'arco di un'ora (S=arrivo di un cliente).

• il numero di incidenti che si verificano ad un incrocio stradale nell'arco di un anno (S=si verifica un incidente).

• il numero di chiamate che arrivano ad un centralino nel corso di 5 minuti (S=arrivo di una chiamata).

• il numero di animali presenti in una certa zona (S=l'animale è presente).

Definizione della variabile casuale di Poisson

Sia X la variabile aleatoria che conta il numero di successi in un intervallo temporale prefissato (o in un'area prefissata).

Essa segue un modello di **Poisson** se è possibile suddividere tale intervallo in tanti sottointervalli (oppure suddividere un'area in tante sottoaree) in modo tale che:

• la probabilità di osservare esattamente un successo in un sottointervallo (o in una sottoarea) è costante;

• la probabilità di osservare più di un successo in un sottointervallo (o in una sottoarea) è trascurabile (di fatto possiamo porla pari a zero);

• il verificarsi di un successo in un sottointervallo (o in una sottoarea) è stocasticamente indipendente dal verificarsi di un successo in un altro sottointervallo (o in un'altra sottoarea).

Essa si indica con

$$X \sim Po\ (\lambda)$$

con λ pari al numero medio di successi in un sottointervallo temporale o spaziale

e $P(X=x)= exp(- \lambda)*(\lambda^x/x!)$ con x=0,1,2,…

La v.c. Poisson possiede media e varianza pari a λ:

E(X)= λ **Var(X)**= λ (la media e la varianza sono pari al numero medio di occorrenze in un sottointervallo)

Esempio: variabile casuale di Poisson

Ad uno sportello bancario si presentano in media 30 clienti ogni ora.
a) Qual'é la probabilità che in un quarto d'ora si presentino esattamente 5 clienti.
b) Quanti clienti si presentano, in media, ogni 5 minuti.

a) La probabilità è pari a:

$$\lambda_h = 30 \left(\text{in un'ora}\right) \longrightarrow \lambda_q = \frac{30}{4} = 7{,}5 \left(\text{ogni 15 minuti}\right)$$

$$P(X=5) = e^{-\lambda_q} \cdot \frac{\lambda_q^x}{x!} = e^{-7,5} \cdot \frac{7{,}5^5}{5!} = 0{,}1094$$

b) La media è pari a:

$$\lambda_c = \frac{\lambda_h}{num.\ \text{sottointervalli}} = \frac{30}{12} = 2{,}5 \left(\text{ogni 5 minuti}\right)$$

Confronto tra variabili casuali discrete notevoli

V.c. di Bernoulli $X \sim Ber(\pi)$

- *Numero di successi in una sola prova* → Supporto [x = 0,1]
- *La prova può generare solo due eventi:* Successo (x=1) e Insuccesso (x=0)
- *Il numero di prove è fisso, pari ad una sola*

V.c. Binomiale $X \sim B(n, \pi)$

- *Numero di successi in n sottoprove* → Supporto [x = 0,1,2,…,n]
- *Ciascuna delle sottoprove può generare solo due eventi:* Successo (x=1) e Insuccesso (x=0)
- *Ogni sottoprova è indipendente dalle altre*
- *La probabilità dell'evento E è costante, in tutte le sottoprove*
- *Il numero di sottoprove è fisso, pari a n.*

V.c. di Poisson $X \sim Po(\lambda)$

- *Numero di successi che occorrono in un intervallo definito (tempo o spazio)*
  → Supporto [x = 0,1,2,…]
- *Il tasso medio di occorrenza dell'evento E è costante per ciascun ambito di riferimento.*
- *L'occorrenza degli eventi è indipendente in ambiti che non si sovrappongono.*
- *La probabilità di due o più occorrenze in ambiti che si sovrappongono è praticamente nulla*

Nella prossima lezione

Nella prossima lezione si affronteranno i seguenti argomenti:

- alcune variabili casuali continue notevoli
- v.c. normale
- v.c. normale standardizzata
- altre v.c. derivate dalla normale