

Approfondimenti sulle v.c.

Task Description

Due Date

Status

☐			
☐			
☐			
☐			
☐			
☐			
☐			
☐			
☐	Combinazione lineare di n v.c.		
☐	Media e varianza della combinazione lineare di due variabili casuali		
☐	Media e varianza della combinazione lineare di n generiche variabili casuali		

MEMO INTERPRETAZIONE

$\sigma_{xy} > 0 \leadsto$ concordanza

$\sigma_{xy} < 0 \leadsto$ discordanza

ATTENZIONE $\sigma_{xy} = 0 \not\Rightarrow$ indipendenza

assenza legame lineare

$$\rho_{xy} = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y} \rightarrow -1 \leq \rho_{xy} \leq +1$$

perfeito legame
lineare inverso

perfeito
legame lineare
diretto

V.C. MULTIVARIATE

Una n -ple di v.c.

$$(X_1, X_2, \dots, X_n)$$

costituisce una v.c. n -variata (le X_i possono essere discrete o continue)

INDIPENDENZA

$$f(x_1, x_2, \dots, x_n) = f_{x_1}(x_1) f_{x_2}(x_2) \dots f_{x_n}(x_n)$$

funzione di probabilità
o di densità congiunta

prodotto delle n funzioni di
probabilità o di densità associate
alle singole componenti (le marginali)

condizione di indipendenza

COMBINAZIONE LINEARE DI DUE VARIABILI CASUALI

siano X e Y due generiche v.c.

$$X \begin{cases} \mu_x \\ \sigma_x^2 \end{cases}$$

$$Y \begin{cases} \mu_y \\ \sigma_y^2 \end{cases}$$

Si definisce combinaz. lineare di (X, Y) una funzione W :

$$W = aX + bY$$

pesi della combinazione

Due particolari combinazioni

• $a = b = 1 \rightarrow W = X + Y$ somma delle due v.c.

• $a = b = \frac{1}{2} \rightarrow W = \frac{X + Y}{2}$ media delle due v.c.

$$E(W) = \mu_w = a\mu_x + b\mu_y$$

la media della combinazione lineare è pari alla combinazione lineare delle due medie

CASO SOMMA $\leadsto \mu_w = \mu_x + \mu_y$

CASO MEDIA $\leadsto \mu_w = \frac{\mu_x + \mu_y}{2}$

$$\text{Var}(W) = \sigma_w^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab \underbrace{\sigma_{xy}}_{=0}$$

se X e Y sono indipendenti

Consideriamo sia X che Y come v.c. discrete (per semplicità)

$$\bullet \mu_w = E(w) = \sum_w w p(w) =$$

$$= \sum_i \sum_j (a x_i + b y_j) \underbrace{p(x=x_i, y=y_j)}_{p_{ij}} =$$

$$= \sum_i \sum_j a x_i p_{ij} + \sum_i \sum_j b y_j p_{ij} =$$

$$= a \underbrace{\sum_i x_i \underbrace{\sum_j p_{ij}}_{p_{i\cdot}}}_{\mu_x} + b \underbrace{\sum_j y_j \underbrace{\sum_i p_{ij}}_{p_{\cdot j}}}_{\mu_y} =$$

$$= a \mu_x + b \mu_y$$

c.v. ol.

$$\bullet \sigma_w^2 = \text{Var}(w) = \sum_w (w - \mu_w)^2 p(w) =$$

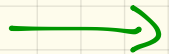
$$= \sum_i \sum_j [(\alpha x_i + b y_j) - (\alpha \mu_x + b \mu_y)]^2 \underbrace{p(X=x_i, Y=y_j)}_{p_{ij}}$$

$$= \sum_i \sum_j [(\alpha x_i - \alpha \mu_x) + (b y_j - b \mu_y)]^2 p_{ij} =$$

$$= \sum_i \sum_j [\alpha(x_i - \mu_x) + b(y_j - \mu_y)]^2 p_{ij} =$$

$$= \sum_i \sum_j \alpha^2 (x_i - \mu_x)^2 p_{ij} + \sum_i \sum_j b^2 (y_j - \mu_y)^2 p_{ij} +$$

$$+ \sum_i \sum_j 2 \alpha b (x_i - \mu_x)(y_j - \mu_y) p_{ij}$$



$$\begin{aligned}
 & \overset{\sigma_x^2}{=} a^2 \underbrace{\sum_i (x_i - \mu_x)^2 \underbrace{\sum_j p_{ij}}_{p_{i \cdot}}}_{\sigma_x^2} + b^2 \underbrace{\sum_j (y_j - \mu_y)^2 \sum_i p_{ij}}_{\sigma_y^2} + \\
 & \quad + 2ab \underbrace{\sum_i \sum_j (x_i - \mu_x)(y_j - \mu_y) p_{ij}}_{\sigma_{xy}} = \\
 & = a^2 \sigma_x^2 + b^2 \sigma_y^2 + 2ab \sigma_{xy}
 \end{aligned}$$

ESEMPIO $\leadsto a=b=1$

$$\sigma_w^2 = \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}$$

scegliendo componenti X e Y inversamente correlate si ottiene una diminuzione nell'variaz. di W

← principio alla base della diversificazione del rischio

COMBINAZIONI LINEARI DI m VARIABILI CASUALI

Siano

$$X_1, X_2, \dots, X_m$$

m generiche v.c. tali che

$$\left\{ \begin{array}{l} E(X_i) = \mu_i \\ \text{Var}(X_i) = \sigma_i^2 \\ \text{Cor}(X_i, X_j) = \sigma_{ij} \end{array} \right. \quad \begin{array}{l} i = 1, m \\ j = 1, m \end{array}$$

Si definisce combinazione lineare delle m v.c. la funzione

$$W = \underbrace{a_1}_{\text{pesi}} X_1 + \underbrace{a_2}_{\text{pesi}} X_2 + \dots + \underbrace{a_m}_{\text{pesi}} X_m = \sum_{i=1}^m a_i X_i$$

$$E(W) = \mu_W = \sum_{i=1}^m a_i \mu_i$$

la media della combinazione lineare
è la combinazione lineare delle medie

$$\text{Var}(W) = \sigma_W^2 = \sum_{i=1}^m a_i^2 \sigma_i^2 + \sum_{i=1}^m \sum_{j \neq i}^m a_i a_j \sigma_{ij}$$

Task Description

Due Date

Status

☐	Media e varianza della combinazione lineare di n variabili casuali indipendenti e identic. distribuite		
☐	Due combinazioni lineari di particolare interesse: la somma e la media di n v.c. i.i.d.		
☐	Distribuzione della somma e della media di n v.c. i.i.d. come una Bernoulli (π)		
☐	Distribuzione della combinazione lineare di n v.c. i.i.d. come una $N(\mu, \sigma^2)$		
☐	Distribuzione della combinazione lineare di n v.c. i.i.d.: il teorema del limite centrale		
☐	Un'applicazione rilevante del teorema del limite centrale: approssimazione normale della v.c. binomiale		
☐			
☐			
☐			
☐			

COMBINAZIONI LINEARI DI m VARIABILI CASUALI

Siano

$$X_1, X_2, \dots, X_m$$

m generiche v.c. tali che

$$\left\{ \begin{array}{l} E(X_i) = \mu_i \\ \text{Var}(X_i) = \sigma_i^2 \\ \text{Cor}(X_i, X_j) = \sigma_{ij} \end{array} \right. \quad \begin{array}{l} i = 1, m \\ j = 1, m \end{array}$$

Si definisce combinazione lineare delle m v.c. la funzione

$$W = \underbrace{a_1}_{\text{pesi}} X_1 + \underbrace{a_2}_{\text{pesi}} X_2 + \dots + \underbrace{a_m}_{\text{pesi}} X_m = \sum_{i=1}^m a_i X_i$$

$$E(W) = \mu_W = \sum_{i=1}^m a_i \mu_i$$

la media della combinazione lineare
è la combinazione lineare delle medie

$$\text{Var}(W) = \sigma_W^2 = \sum_{i=1}^m a_i^2 \sigma_i^2 + \sum_{i=1}^m \sum_{j \neq i}^m a_i a_j \sigma_{ij}$$

Siamo interessati ad un caso particolare di combinazione lineare:

- le X_i hanno la stessa distribuzione
 $i=1, n$

identicamente
distribuite o
somigliante



$$\mu_i = \mu \quad \forall i=1, n$$

$$\sigma_i^2 = \sigma^2 \quad \forall i=1, n$$

- le X_i sono inoltre indipendenti



$$\sigma_{i,j} = 0 \quad \forall i, j$$

→ le X_i sono i. i. d.

indipendenti

identicamente
distribuite

Siano quindi $X_1, X_2, \dots, X_n$ n r.c. i.i.d. si ha che la combinazione lineare $W = \sum_{i=1}^n a_i X_i$ ha media:

$$\mu_w = \sum_i a_i \mu_i = \sum_i a_i \underbrace{\mu}_{\text{media comune a tutte le } X_i} = \mu \sum_i a_i$$

media comune
a tutte le X_i

e la sua varianza:

$$\sigma_w^2 = \sum_i a_i^2 \underbrace{\sigma_i^2}_{\text{varianza comune a tutte le } X_i} + \sum_i \sum_{j \neq i} a_i a_j \underbrace{\sigma_{ij}}_{=0 \text{ } \forall i, j} =$$

varianza
comune a
tutte le X_i

$= 0 \text{ } \forall i, j$
essendo le X_i indipendenti

$$= \sum_i a_i^2 \sigma^2 = \sigma^2 \sum_i a_i^2$$

Due particolari combinazioni lineari

$$a_i = 1 \quad \forall i=1, n \quad \rightarrow \quad W \text{ r.c. somma}$$

$$W = \sum_i X_i \quad \left\{ \begin{array}{l} \mu_W = \mu \sum_i a_i = \mu \sum_{i=1}^n 1 = n \mu \\ \sigma_W^2 = \sigma^2 \sum_i a_i^2 = \sigma^2 \sum_{i=1}^n 1^2 = n \sigma^2 \end{array} \right.$$

un esempio si ha considerando le r.c. binomiali

$$X_i \sim \text{Ber}(\pi) \quad \rightarrow \quad S_n = W = \sum_i X_i$$

$$i=1, n$$

$$S_n \sim \text{bin}(n, \pi)$$

$$E(X_i) = \pi$$

$$E(S_n) = n E(X_i) = n \pi$$

$$\text{Var}(X_i) = \pi(1-\pi)$$

$$\text{Var}(S_n) = n \text{Var}(X_i) = n \pi(1-\pi)$$

$a_i = \frac{1}{n} \quad \forall i=1, n \leadsto W$ è la media delle n v.c. i. i. d.
 $W = \sum_i X_i / n$

$$\mu_W = \mu \sum_i a_i = \mu \sum_i \frac{1}{n} = \mu \frac{1}{\cancel{n}} n = \mu$$

$$\sigma_W^2 = \sigma^2 \sum_i a_i^2 = \sigma^2 \sum_i \frac{1}{n^2} = \sigma^2 \frac{1}{\cancel{n}} n = \frac{\sigma^2}{n}$$

Se consideriamo il caso di $X_i \sim \text{Ber}(\pi)$

$$W = \frac{\sum X_i}{n} = \frac{S_n}{n} = \bar{X}_n \rightarrow \text{proporzione successi}$$

$\hookrightarrow$ media di osservazioni (0, 1)

$$E(\bar{X}_n) = \pi$$

$$\text{Var}(\bar{X}_n) = \frac{\pi(1-\pi)}{n}$$

A partire da n v.c. i.i.d. sappiamo ora calcolare
la media e la varianza della v.c. combinazione lineare

↓
per controllare completamente il processo dovremmo capire
se è possibile avere informazioni sulla distribuzione di
probabilità di W

↓
Se le $X_i \sim \text{Ber}(\pi)$ allora per ricavare la distribuzione di
probabilità di W possiamo sfruttare la binomiale

↳ siamo interessati in particolare a
 S_n (n , successi) e J_n (proporzioni di
successi)

Se invece le $X_i \sim N(\mu, \sigma^2)$ cosa succede alla W ?

combinazione lineare
di n v.c. normali

le proprietà riproduttive delle
v.c. normale ci assicura che W
è sempre normale (per qualunque
valore di n)

$$W \sim N\left(\underbrace{\mu \sum_1^n a_i}_{\mu_W}, \underbrace{\sigma^2 \sum_1^n a_i^2}_{\sigma_W^2}\right)$$

Se le X_i hanno invece una diversa distribuzione di probabilità $f(x_i)$ che potrei anche non conoscere

ad esempio
poisson, esponenziale, uniforme

il teorema del limite centrale ci permette di usare la v.c. normale come distribuzione approssimata (per $n \rightarrow +\infty$)

REGOLA EMPIRICA
 $n \geq 30$

la bontà dell'approssimazione dipende da due fattori

- la forma della $f(x)$
(simmetrica o meno)
- valore di n

l'approssimazione è buona indipendentemente dal modello $f(x)$ di partenza

→ l'approssimazione è migliore se $f(x)$ è "abbastanza" simmetrica

→ l'approssimazione è migliore se n è "grande"

È fatto teorema limite centrale

- distribuzione X_i

quanto più questa distribuzione è
simmetrica tanto più veloce sarà la
convergenza alla normale

- numero delle X_i che mette in combinazione

quanto maggiore è n tanto "migliore"
sarà la convergenza

REGOLA
EMPIRICA } $\rightarrow$ se $n \geq 30$ è possibile ottenere
una buona approssimazione

1) potrei avere informazioni circa la distribuzione delle X_i ma potrebbero non essere disponibili risultati teorici per derivare la distribuzione di W

2) potrei non avere informazioni circa la distribuzione delle X_i

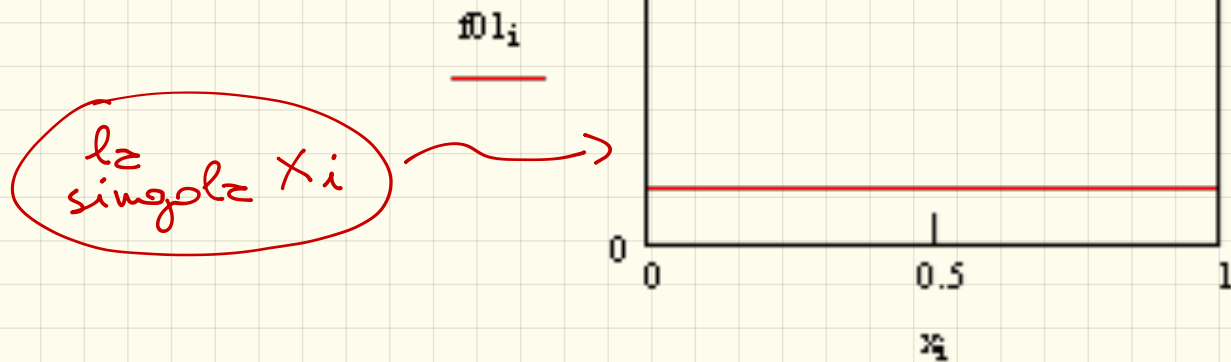
→ in entrambi i casi posso sfruttare il teorema del limite centrale (TLC) per ottenere una distribuzione approssimata per W

In termini più formali:

$$\frac{W - E(W)}{\sqrt{\text{Var}(W)}} = \frac{W - \mu_W}{\sigma_W} = \frac{W - \mu \sum_i a_i}{\sqrt{\sigma^2 \sum_i a_i^2}} \xrightarrow{(n \rightarrow \infty)} N(0, 1)$$

la combinazione lineare di n v.c. i.i.d., opportunamente standardizzata, converge ad una $N(0, 1)$

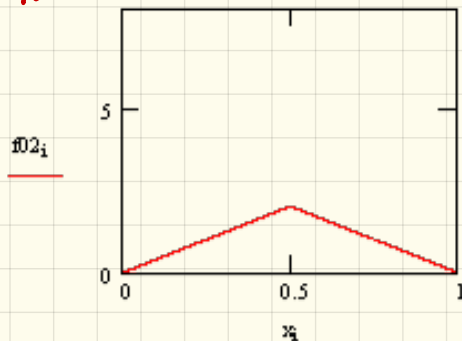
Si consideri, ad esempio,
una distribuzione uniforme
continua:



e per queste si consideri la combinazione lineare media di
 n v.c., con $n = 2, 3, 4, 8, 16, 32$

Comportamento della media di n v.c. i.i.d. come una v.c. uniforme

$n=2$



$n=3$



$n=4$



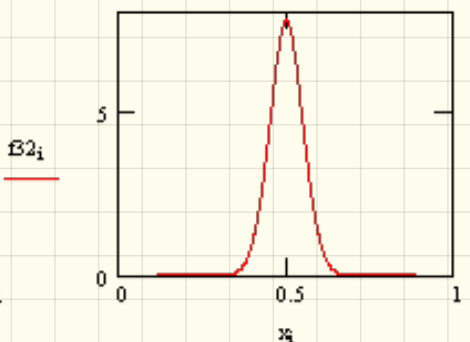
$n=8$



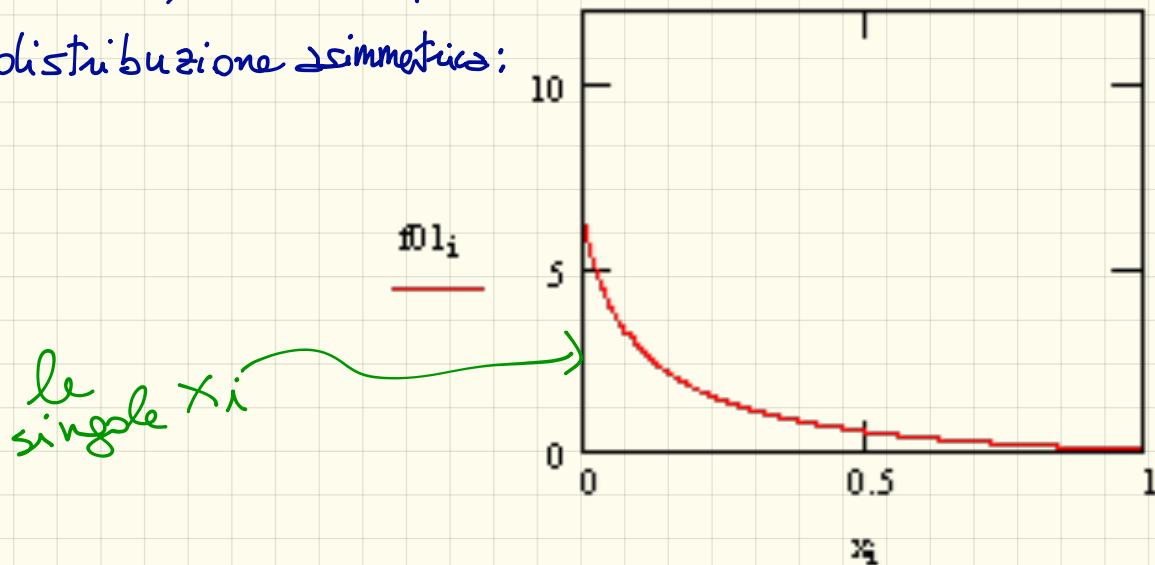
$n=16$



$n=32$



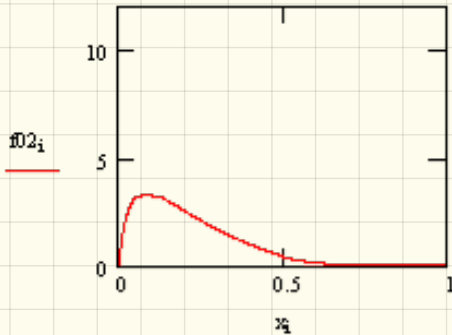
Si consideri, ad esempio,
una distribuzione asimmetrica:



e per questo si consideri la combinazione lineare media di
 n r.c., con $n = 2, 3, 4, 8, 16, 32$

Comportamento della media di n v.c. i.i.d. come una v.c. asimmetrica

$n=2$



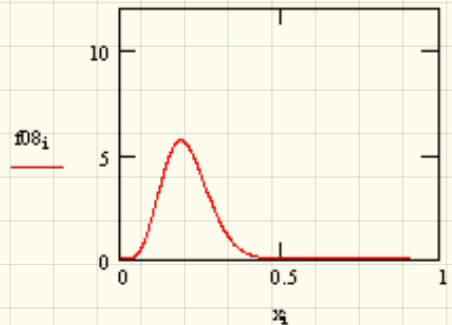
$n=3$



$n=4$



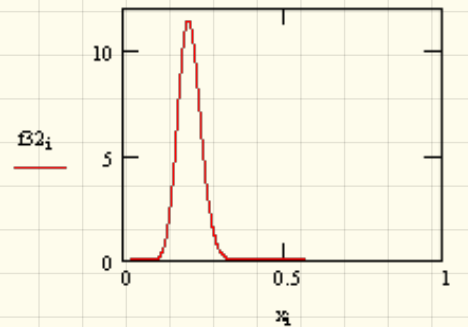
$n=8$



$n=16$



$n=32$



Ulteriori esempi sono disponibili sul sito:

http://www.statisticalengineering.com/central_limit_theorem.html

UN'APPLICAZIONE RILEVANTE DEL TLC

↳ convergenza della v.c. binomiale alla normale



la prima versione del TLC (noto come teorema di De Moivre - Laplace) riguardava la convergenza di

una v.c. binomiale ad una v.c. normale

↓
v.c. discreta

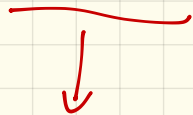


↓
v.c. continua

questo richiederà qualche accorgimento nel calcolo delle probabilità associate ai singoli valori perché $P(X = x_i)$ è calcolabile nel caso delle v.c. discrete mentre è sempre $= 0$ nel caso di una v.c. continua

$$X \sim \text{bin}(n=100, \pi=0.2)$$

$$P(X \geq 70) = \sum_{i=70}^{100} \binom{100}{i} \pi^i (1-\pi)^{n-i}$$



posso approssimare questo risultato sfruttando il TLC

$$P(X \geq 70) = P\left(\frac{X - \mu}{\sigma} \geq \frac{70 - 100 \times 0.2}{\sqrt{100 \times 0.2 \times 0.8}}\right)$$

(2)
 $E(X)$
 $\sqrt{\frac{n \pi}{1}}$

$\sqrt{\frac{n \pi (1-\pi)}{1}}$

$\text{Var}(X)$

usando le tabelle della normale ottenete un'approssimazione ragionevole (tanto più "buona" quanto maggiore è n)

$$P(X = 70) = \binom{100}{70} 0.2^{70} 0.8^{30}$$

2 volte questo potrebbe dare errori
sulla calcolatrice (nel caso di n. troppo
grandi)

se uso l'approssimaz. normale

$$P\left(\frac{X - \mu}{\sigma} = \frac{70 - 100 \times 0.2}{\sqrt{100 \times 0.2 \times 0.8}}\right) = 0 \quad \left(\begin{array}{l} \text{modello} \\ \text{continuo} \end{array}\right)$$

si approssima
usando un intervallo

→ si prende un Δ piccolo
(di solito $\Delta = 0.5$)] $P(70 - \Delta \leq X \leq 70 + \Delta)$ piccolo