

Variabili aleatorie

Una variabile casuale X è una funzione definita sullo spazio campionario Ω che associa ad ogni evento $E_i \subset \Omega$ uno ed un solo numero reale.

$$X : \Omega \longrightarrow \mathbb{R}$$

Attraverso X operiamo la traduzione dei risultati dell'esperimento casuale nel campo dei numeri reali.

- ▶ V.C. DISCRETE: possono assumere un insieme finito o un'infinità numerabile di valori (il supporto può essere messo in corrispondenza biunivoca con l'insieme dei numeri interi).
- ▶ V.C. CONTINUE: possono assumere un'infinità non numerabile di valori (il supporto è un intervallo di $\mathbb{R}$)

La v.c. X assume i valori $\{x_1, \dots, x_i, \dots\}$.

Esempio: lancio di un dado. Associao all'esperimento una v.c. che assume valore 0 se esce una faccia pari e 1 se esce una faccia dispari. Pertanto $x_i = 0, 1$.

Esempio: il numero dei clienti che visitano un negozio in un intervallo di tempo (variabile di flusso) costituisce una v.c. che assume valori $x_i = 0, 1, 2, \dots, \infty$.

Simili considerazioni valgono per il numero degli studenti presenti in aula (variabile di stato).

Associa ai valori x_i che la v.c. può assumere una misura di probabilità

$$P(X = x_i) = P(x_i)$$

tale che:

$$P(x_i) \geq 0, \quad \sum_i P(x_i) = 1.$$

x_i	$P(x_i)$
x_1	$P(x_1)$
$\vdots$	$\vdots$
x_n	$P(x_n)$

$P(x_i)$ è la probabilità dell'unione degli eventi elementari per i quali vale l'assegnazione $X = x_i$.

Esempio

Un'urna contiene 20 palline gialle, 30 palline rosse e 50 palline nere. Si estraggono CON reimmissione due palline.

Scrivere la probabilità della variabile aleatoria X = “numero di palline gialle estratte”

$$P(G)=2/10$$

$$P(R) = 3/10$$

$$P(N)=5/10$$

Palline		Prob
GG		$0.2 \times 0.2 = 0.04$
GR		$0.2 \times 0.3 = 0.06$
GN		$0.2 \times 0.5 = 0.1$
RG		0.06
RR		0.09
RN		0.15
NG		0.1
NR		0.15
NN		0.25

Palline	X	Prob
GG	2	$0.2 \times 0.2 = 0.04$
GR	1	$0.2 \times 0.3 = 0.06$
GN	1	$0.2 \times 0.5 = 0.1$
RG	1	0.06
RR	0	0.09
RN	0	0.15
NG	1	0.1
NR	0	0.15
NN	0	0.25

Palline	X	Prob
GG	2	$0.2 \times 0.2 = 0.04$
GR	1	$0.2 \times 0.3 = 0.06$
GN	1	$0.2 \times 0.5 = 0.1$
RG	1	0.06
RR	0	0.09
RN	0	0.15
NG	1	0.1
NR	0	0.15
NN	0	0.25

X	p(x)
0	0.64
1	0.32
2	0.04

Funzione di ripartizione

Funzione di ripartizione, $F(x_0)$, di una variabile aleatoria X esprime la probabilità che X non supero il valore x_0 ,

$$F(x_0) = P(X \leq x_0) = \sum_{x \leq x_0} p(x)$$

Proprietà:

1. Monotonicità: $F(x)$ è non-decrescente.

$$x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$$

2. Il codominio è l'intervallo $[0, 1]$,

$$\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$$

3. E' una funzione continua a destra:

$$\lim_{x \rightarrow x_i^+} F(x) = F(x_i)$$

Esempio

Lancio del dado FINO a quando si osserva il numero 6:
X "numero di tentativi prima di osservare il numero 6"

Quali possibili valori può assumere X?

Con quale probabilità $X=0$? $X=1$? $X=2$?

$$P(X = x) = \left(\frac{5}{6}\right)^x \frac{1}{6} \quad x = 0, 1, 2, \dots$$

Esempio

- Esperimento: Lancio di 3 monete
- X "numero di teste"

Ω	X
HHH	0
HHT	1
HTH	1
HTT	2
THH	1
THT	2
TTH	2
TTT	3

x	p(x)
0	1/8
1	3/8
2	3/8
3	1/8

Valore atteso

La valore atteso $E(X)$ di una variabile aleatoria X è definito come:

$$E(X) = \sum_x x p(x)$$

La notazione $\sum$ indica somma su tutti i valori di x , il valore atteso è indicato come μ_X .

Varianza e deviazione standard

Sia X una variabile aleatoria discreta. Il valore atteso del quadrato della distanza dalla media, $(X - \mu_X)^2$, si chiama varianza ed è indicato come σ^2_X

$$\sigma^2_x = E(X - \mu_x)^2 = \sum_x (x - \mu_x)^2 p(x)$$

σ_X , indica la deviazione standard,
 $\mu_X = E(X)$

Varianza (Formula Alternativa)

La varianza di una v.a. discrete può anche essere scritta come:

$$\begin{aligned}\sigma_x^2 &= E(X^2) - \mu_x^2 \\ &= \sum_x x^2 p(x) - \mu_x^2\end{aligned}$$

Valore atteso

$$E(g(X)) = \sum_i g(x_i) p(x_i)$$

$$E(E(X)) = E(X)$$

$$E(X - E(X)) = 0$$

$$E(a + bX) = a + bE(X)$$

Varianza

$$Var(X) = E(X - E(X))^2$$

$$Var(X) = E(X^2) - E(X)^2$$

$$Var(a + bX) = b^2 Var(X)$$

Esempio

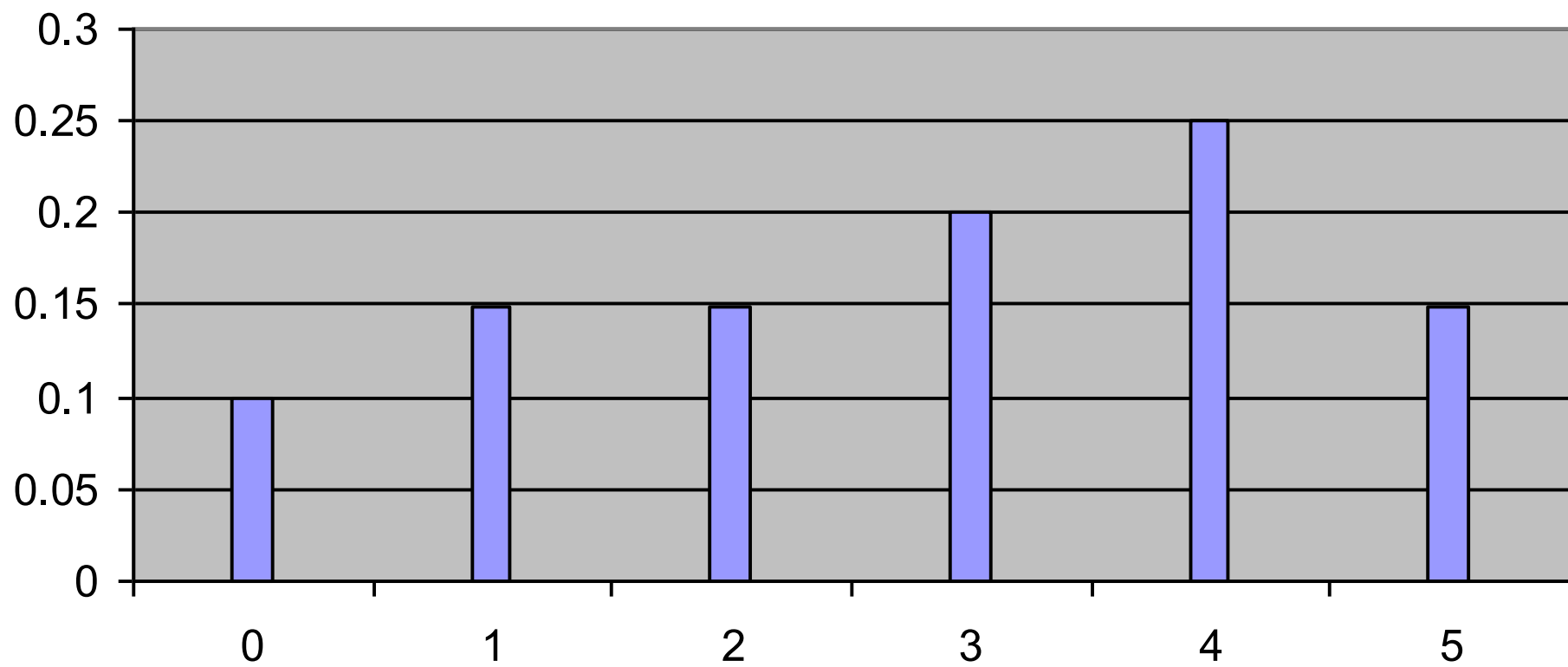
Un dattilografo scrive **100** parole al minuto, questa è la distribuzione di probabilità del numero di errori fatti ogni 100 parole

x	p(x)
0	0.1
1	0.15
2	0.15
3	0.2
4	0.25
5	0.15

Trovare valore
atteso e varianza



$p(x)$



$$\begin{aligned}
 E(X) &= \sum_{i=0}^5 ip(i) = \\
 &= 0 \times p(0) + 1 \times p(1) + 2 \times p(2) + 3 \times p(3) + \\
 &\quad + 4 \times p(4) + 5 \times p(5) = \\
 &= 0 \times 0.1 + 1 \times 0.15 + 2 \times 0.15 + 3 \times 0.2 + \\
 &\quad + 4 \times 0.25 + 5 \times 0.15 = \\
 &= 2.8
 \end{aligned}$$

Ogni 100 parole scritte
ci aspettiamo **2.8 errori**

$$\begin{aligned}
 E(X^2) &= \sum_{i=0}^5 i^2 p(i) = \\
 &= 0^2 \times p(0) + 1^2 \times p(1) + 2^2 \times p(2) + 3^2 \times p(3) + \\
 &\quad + 4^2 \times p(4) + 5^2 \times p(5) = \\
 &= 0 \times 0.1 + 1 \times 0.15 + 4 \times 0.15 + 9 \times 0.2 + \\
 &\quad + 16 \times 0.25 + 25 \times 0.15 = \\
 &= 10.3
 \end{aligned}$$

$$Var(X) = E(X^2) - (E(X))^2 = 10.3 - 2.8^2 = 2.46$$

Variabile aleatoria doppia

Una v.c. multipla è una funzione definita sullo spazio campionario Ω che associa ad ogni evento $E \subset \Omega$ una n -upla di numeri reali:

$$\{X_1, \dots, X_n\} : \Omega \rightarrow \mathbb{R}^n$$

Distribuzione di probabilità di una v.c. discreta doppia

A Ω associamo una coppia di numeri reali $\{X, Y\}$ (ovvero $\{X_1, X_2\}$).

In generale, denoteremo con $P(x, y)$ (ovvero mediante $P(x_i, y_j)$) la funzione di probabilità congiunta di una v.c. doppia; essa è tale che $P(x, y) \geq 0$ e $\sum_x \sum_y P(x, y) = 1$.

Essa fornisce $P(X = x, Y = y)$.

Distribuzione di probabilità di una v.c. discreta doppia

X	Y					
	y_1	$\dots$	y_j	$\dots$	y_k	
x_1	$P(x_1, y_1)$	$\dots$	$P(x_1, y_j)$	$\dots$	$P(x_1, y_k)$	$P(x_1)$
x_2	$P(x_2, y_1)$	$\dots$	$P(x_2, y_j)$	$\dots$	$P(x_2, y_k)$	$P(x_2)$
$\vdots$	$\vdots$	$\dots$	$\vdots$	$\dots$	$\vdots$	$\vdots$
x_i	$P(x_i, y_1)$	$\dots$	$P(x_i, y_j)$	$\dots$	$P(x_i, y_k)$	$P(x_i)$
$\vdots$	$\vdots$	$\dots$	$\vdots$	$\dots$	$\vdots$	$\vdots$
x_m	$P(x_m, y_1)$	$\dots$	$P(x_m, y_j)$	$\dots$	$P(x_m, y_k)$	$P(x_m)$
	$P(y_1)$	$\dots$	$P(y_j)$	$\dots$	$P(y_k)$	1

- Distribuzioni marginali

$$P(x) = \sum_y P(x, y), \quad P(y) = \sum_x P(x, y)$$

- Distribuzioni condizionate

$$P(x|y) = \frac{P(x, y)}{P(y)}, \quad P(y|x) = \frac{P(x, y)}{P(x)}$$

Il valore atteso condizionato di Y , dato $X = x$, è definito nel caso discreto da

$$E(Y|X = x) = \sum_y yP(y|x).$$

Esercizio

Si consideri il lancio di tre monete.

Si definiscano due variabili aleatorie:

X = “numero di realizzazioni dell’evento
“testa”,

Y = “numero di cambi nella sequenza”,
ovvero il numero di volte in cui si passa da
testa a croce e viceversa.

Ω	X	Y
TTT	3	0
TTC	2	1
TCT	2	2
CTT	2	1
TCC	1	1
CTC	1	2
CCT	1	1
CCC	0	0

X	Y			
	0	1	2	Totale
0				
1				
2				
3				
Totale				

X	Y			
	0	1	2	Totale
0	$1/8$	0	0	$1/8$
1	0	$2/8$	$1/8$	$3/8$
2	0	$2/8$	$1/8$	$3/8$
3	$1/8$	0	0	$1/8$
Totale	$2/8$	$4/8$	$2/8$	1

Indipendenza di Variabili Aleatorie

Siano X e Y due variabili aleatorie, X e Y sono indipendenti se

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

per ogni $x, y \in R$

Due variabili aleatorie sono indipendenti se e solo se ogni coppia di eventi $\{X=x\}$ e $\{Y=y\}$ sono indipendenti

Covarianza

- Se X and Y hanno varianza finita, la covarianza è definita come

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

- La covarianza è una misura di associazione
Covariance is a measure of association.

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Correlazione

- Se X e Y hanno varianza finita, la loro correlazione è definita come

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

Indipendenza

- Se X e Y sono indipendenti,

$$E(XY) = E(X)E(Y)$$

$$\text{Cov}(X, Y) = 0$$

Non vale il viceversa

Somma di Variabili aleatorie

- $E(X+Y)=E(X)+E(Y)$
- $Var(X+Y)=Var(X)+Var(Y)+2\times Cov(X,Y)$

Esempio

Immagina di voler creare un portafoglio che consiste in una combinazione di corporate bonds e common stocks. Le seguenti informazioni sul rendimento annuale (per \$1,000) di ogni investimento sotto le diverse condizioni sono disponibili

Esempio

Probability	State of the economy	Corporate bonds in \$	Common stock in \$
0.10	Recession	-30	-150
0.15	Stagnation	50	-20
0.35	Slow growth	90	120
0.30	Moderate growth	100	160
0.10	High growth	110	250

- Expected return for the corporate bond fund

$$E(X) = -30 \times 0.10 + 50 \times 0.15 + 90 \times 0.35 + 100 \times 0.30 + 110 \times 0.10 = 66$$

$$E(Y) = -150 \times 0.10 - 20 \times 0.15 + 100 \times 0.35 + 160 \times 0.30 + 250 \times 0.10 = 97$$

- Variance

$$E(X^2) = (-30)^2 \times 0.10 + 50^2 \times 0.15 + 90^2 \times 0.35 + 100^2 \times 0.30 + 110^2 \times 0.10 = 7510$$

$$Var(X) = E(X^2) - E(X)^2 = 1581 \quad \sigma_X = 39.76$$

$$E(Y^2) = (-150)^2 \times 0.10 + (-20)^2 \times 0.15 + 100^2 \times 0.35 + 160^2 \times 0.30 + 250^2 \times 0.10 = 21280$$

$$Var(Y) = E(Y^2) - E(Y)^2 = 11871 \quad \sigma_Y = 108.95$$

- Covariance

$$\begin{aligned} E(XY) &= (-30) \times (-150) \times 0.10 + 50 \times (-20) \times 0.15 + 90 \times 100 \times 0.35 + \\ &\quad + 100 \times 160 \times 0.30 + 110 \times 250 \times 0.10 = 11630 \end{aligned}$$

$$Cov(X, Y) = E(XY) - E(X)E(Y) = 11630 - 66 \times 97 = 4161$$

- Scegliaresti di investire nei corporate bond fund o nei common stock fund?
- Stock Y gives the investor a higher expected return than Stock X, but also has a higher variance. A risk-averse investor should invest in stock X, but an investor willing to sustain a higher risk can expect a higher return from Stock Y.
- Calcola valore atteso della portafoglio e il rischio del portafoglio per le seguenti proporzioni di investimento nei corporate bonds
- 1) 0.10 2) 0.30 3) 0.50 4) 0.70 5) 0.90
- Quale dei precedenti portafogli sceglieresti come investimento?

w = proportion of the portfolio value assigned to asset X

Portfolio expected return:

$$E(P) = wE(X) + (1-w)E(Y)$$

Portfolio risk:

$$\sigma_P = \sqrt{w^2 \sigma_X^2 + (1-w)^2 \sigma_Y^2 + 2w(1-w)\sigma_{XY}}$$

- 1) $E(P)=95$, $\sigma_p=101.88$ 2) $E(P)=91$, $\sigma_p=87.79$
- 3) $E(P)=87$, $\sigma_p=73.78$ 4) $E(P)=83$, $\sigma_p=59.92$
- 5) $E(P)=79$, $\sigma_p=46.35$

- Which of the previous portfolio would you recommend?
- You should recommend a portfolio with 10% Stock X and 90% Stock Y if you make the decision solely on the expected return but recognize that as your expected return increases, so the portfolio risk.

Esercizio

Data la seguente distribuzione di probabilità congiunta per le variabili X ed Y,

	Y		
X	1	2	4
2	0.1	0	0.1
3	0	0.2	0
5	0.2	0.1	0.1
6	0	0	0.2

Esercizio

Si calcolino:

- Le distribuzioni di probabilità marginali della X e della Y
- La funzione di probabilità condizionate di Y dato che $X=5$
 $E(X), E(Y), E(XY)$

$$Cov(X, Y)$$

- Stabilire se X e Y sono indipendenti

	Y			
X	1	2	4	
2	0.1	0	0.1	0.2
3	0	0.2	0	0.2
5	0.2	0.1	0.1	0.4
6	0	0	0.2	0.2

X	p(x)
2	0.2
3	0.2
5	0.4
6	0.2

	Y		
X	1	2	4
2	0.1	0	0.1
3	0	0.2	0
5	0.2	0.1	0.1
6	0	0	0.2
p(x)	0.3	0.3	0.4

Y	p(x)
1	0.3
2	0.3
4	0.4

X	p(x)	$x \times p(x)$
2	0.2	0.4
3	0.2	0.6
5	0.4	2
6	0.2	1.2
	E(X)	4.2

Y	p(y)	$y \times p(y)$
1	0.3	0.3
2	0.3	0.6
4	0.4	1.6
	E(Y)	2.5

	Y		
X	1	2	4
	2	4	8
2	0.1	0	0.1
	3	6	12
3	0	0.2	0
	5	10	20
5	0.2	0.1	0.1
	6	12	24
6	0	0	0.2

$$E(XY)=11$$

$$\text{Covarianza } 11-10.5=0.5$$

	Y			
X	1	2	4	
2	0.1	0	0.1	0.2
3	0	0.2	0	0.2
5	0.2	0.1	0.1	0.4
6	0	0	0.2	0.3

Esercizio

Date due variabili aleatorie indipendenti X e Y

$$X = \begin{cases} 1 & 2 & 3 \\ 0.4 & 0.2 & 0.4 \end{cases} \quad Y = \begin{cases} 2 & 4 \\ 0.3 & 0.7 \end{cases}$$

si determini la distribuzione della variabile aleatoria $W=2X-Y$ e se ne calcoli il valore atteso.

Y/X	1	2	3	
				<i>0.3</i>
2	0.12	0.06	0.12	
4	0.28	0.14	0.28	<i>0.7</i>
	<i>0.4</i>	<i>0.2</i>	<i>0.4</i>	

Y/X	1	2	3	
2				0.3
4				0.7
	0.4	0.2	0.4	

Y/X	1	2	3	
2	0.12 ⁰	0.06 ²	0.12 ⁴	0.3
4	0.28 ⁻²	0.14 ⁰	0.28 ²	0.7
	0.4	0.2	0.4	

$$W = \begin{cases} -2 & 0 & 2 & 4 \\ 0.28 & 0.26 & 0.34 & 0.12 \end{cases} \quad E(W)=0.6$$

Bernoulli

Sia $X=1$ se lanciando il dado esce il numero 6

$X=0$ se esce qualsiasi altro numero al lancio del dado

X si distribuisce come una Bernoulli

$$X = \begin{cases} 0 & 1-p \\ 1 & p \end{cases}$$

$$p(x) = p^x (1-p)^{1-x}$$

$$E(X)=p \qquad \text{Var}(X)=p(1-p)$$

Lanciamo il dado n volte, Siano $X_1, X_2, \dots, X_n$ n variabili aleatorie indipendenti e identicamente distribuite come X :
cosa succede a $Y = X_1 + X_2 + \dots + X_n$?

Esperimento Binomiale

1. L'esperimento consiste di n prove identici
2. L'esito di ogni prova può essere solo o "successo" o "insuccesso".
3. $P(\text{successo}) = p$
4. Le n prove sono indipendenti l'una dall'altra
 p constant
5. La variabile aleatoria, X , rappresenta il numero di successi su n prove.

X è binomiale?

Probabilità che uno studente fuma è 0.35.

Si intervistano studenti FINO A QUANDO si trova uno che fuma

X = numero di studenti intervistati

No!

X è binomiale?

Su 20 lampadine fabbricate ce ne sono 8 difettose.

Si selezionano casualmente 5 lampadine
SENZA REIMMISSIONE”.

X = numero di lampadine difettose

No!

X è binomiale?

La probabilità che un italiano estratto a caso abbia avuto seri problemi a causa della crisi economica è 0.29.

Intervistiamo 10 Italiani a caso

Sia X = numero di Italiani che hanno avuto seri problemi a causa della crisi economica tra i 10 intervistati

Yes!

X è binomiale?

3 Studenti dicono a caso un numero da 0 a 9

Let X = numero di studenti dicono il numero 7

Yes!

Distribuzione di probabilità binomiale

$$p(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$p(x)$ = Probabilità of x 'Successo'

n = numero di esperimenti

p = Probabilità di 'Successo'

x = Numero di 'Successi' nel campione
($x = 0, 1, 2, \dots, n$)

Media e Varianza di una distribuzione Binomiale

$$\mu_X = E(X) = np$$

$$\sigma_X^2 = E[(X - \mu_X)^2] = np(1 - p)$$

Distribuzione Binomiale

IL 40% degli Americani guarda la televisione il venerdì sera. Immaginate di telefonare a 7 persone prese a caso,

Qual è la probabilità che

- nessuno stia guardando la televisione
- esattamente tre sono davanti alla televisione
- almeno 3 sono davanti alla televisione
- tutti e 7 guardano la televisione

$$P(X = 0) = \binom{7}{0}(0.4)^0(0.6)^7 = \frac{7!}{0!7!}(0.6)^7 = 0.028$$

$$P(X = 3) = \binom{7}{3}(0.4)^3(0.6)^4 = \frac{7!}{3!4!}0.064 \times 0.1296 = 0.293$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2) = \\ &= 1 - 0.028 - \binom{7}{1}(0.4)^1(0.6)^6 - \binom{7}{2}(0.4)^2(0.6)^5 = \\ &= 1 - 0.028 - \frac{7!}{6!}0.4 \times 0.0467 - \frac{7!}{2!5!}0.16 \times 0.0778 = \\ &= 1 - 0.028 - 0.1306 - 0.2614 = 0.5800 \end{aligned}$$

Esercizio

- Un recente studio demografico, mostra che il 30% dei giovani maschi italiani rimane in famiglia oltre i 35 anni. Intervistati a caso 6 giovani maschi italiani di oltre 35 anni
- qual è il numero atteso di giovani ultra 35enni che vivono in famiglia?
- Qual è la probabilità che esattamente 2 di essi vive ancora in famiglia?
- Qual è la probabilità che non più di 2 di essi viva ancora in famiglia?

Soluzione

X ="numero giovani maschi che rimangono in casa oltre i 35 anni"

Distribuzione Binomiale con $p=0.3$, $n=6$

$$1) E(X) = n \times p = 6 \times 0.3 = 1.8$$

$$2) P(X = 2) = \binom{6}{2} (0.3)^2 (0.7)^4 = \frac{6!}{2!4!} \times 0.09 \times 0.2401 = 0.3241$$

$$\begin{aligned} 3) P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) = \binom{6}{0} (0.3)^0 (0.7)^6 + \binom{6}{1} (0.3)^1 (0.7)^5 + 0.3241 \\ &= \frac{6!}{0!6!} \times 0.1176 + \frac{6!}{1!5!} \times 0.3 \times 0.1681 + 0.3241 = 0.1176 + 0.3026 + 0.3241 = 0.7443 \end{aligned}$$

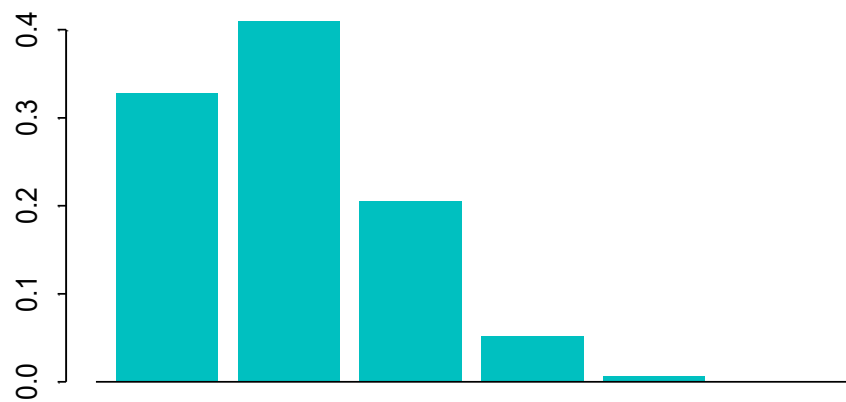
Shape of a Binomial Probability Distribution

Symmetric if $p = .50$

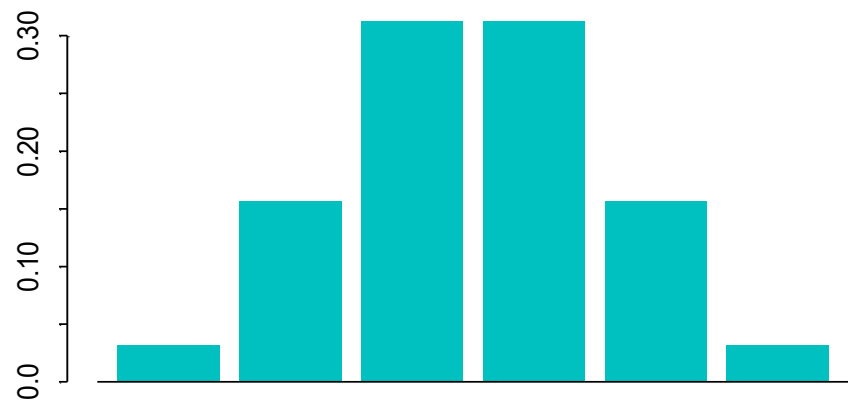
Skewed right if $p < .50$

Skewed left if $p > .50$

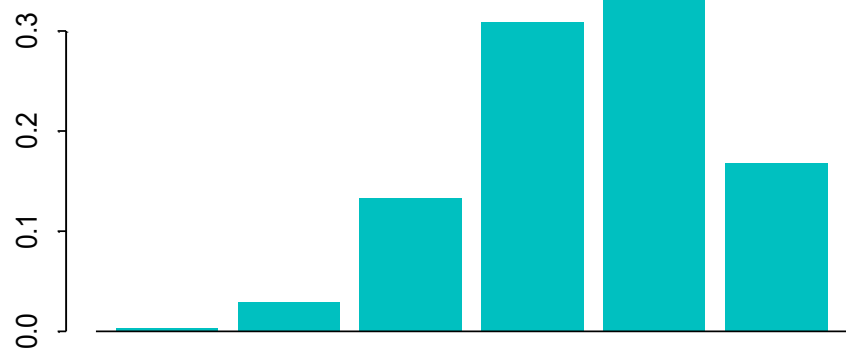
$p=.2$



$p=.5$



$p=.7$



$p=.8$

