BREVI CENNI DI CALCOLO DELLE PROBABILITA'

SPAZI DI PROBABILITA' FINITI

S : SPAZIO DEI CAMPIONI (FINITO)
(E' L'INSIEME DI TUTTI I POSSIBILI RISULTATI DI UN ESPERIMENTO ALEATORIO)

$i \in S$: EVENTO ELEMENTARE / ESITO DELL'ESPERIMENTO

$A \subseteq S$: EVENTO

S : EVENTO CERTO

$\emptyset$: EVENTO NULLO

$A, B \subseteq S$ T.C. $A \cap B = \emptyset$: EVENTI MUTUAMENTE ESCLUSIVI

ESEMPIO

PER L'ESPERIMENTO DEL LANCIO DI DUE MONETE DISTINGUIBILI:

$$S = \{TT, TC, CT, CC\}$$

EVENTO DI OTTENERE ESATTAMENTE UNA TESTA: $\{TC, CT\}$

ASSIOMI DI PROBABILITÀ

UNA DISTRIBUZIONE DI PROBABILITÀ IN UNO SPAZIO DI CAMPIONI S E' UNA FUNZIONE REALE $Pr\{\ \}$ (PROBABILITÀ) SULL' INSIEME DEGLI EVENTI DI S CHE SODDISFA GLI ASSIOMI DI PROBABILITÀ

1. $Pr\{A\} \geq 0$, PER OGNI EVENTO $A \subseteq S$

2. $Pr\{S\} = 1$

3. $Pr\{A \cup B\} = Pr\{A\} + Pr\{B\}$, PER QUALSIASI COPPIA A, B DI EVENTI MUTUAMENTE ESCLUSIVI

$$Pr\{\bigcup_i A_i\} = \sum_i Pr\{A_i\},$$

PER QUALSIASI SEQUENZA FINITA (O NUMERABILE) DI EVENTI MUTUAMENTE ESCLUSIVI A COPPIE

ALCUNE SEMPLICI PROPRIETÀ

- $Pr\{A\} = Pr\{A \cup \emptyset\} = Pr\{A\} + Pr\{\emptyset\} \Longrightarrow Pr\{\emptyset\} = 0$

- $A \subseteq B \Longrightarrow Pr\{B\} = Pr\{A \cup (B \setminus A)\} = Pr\{A\} + Pr\{B \setminus A\}$
$\geq Pr\{A\}$

PER $A \subseteq B$, PONIAMO $\bar{A} = S \setminus A$ (EVENTO COMPLEMENTARE)

- $1 = Pr\{S\} = Pr\{A \cup \bar{A}\} = Pr\{A\} + Pr\{\bar{A}\} \Longrightarrow Pr\{\bar{A}\} = 1 - Pr\{A\}$

- A, B EVENTI QUALSIASI $\longrightarrow$
$Pr\{A \cup B\} = Pr\{A\} + Pr\{B\} - Pr\{A \cap B\}$
$\leq Pr\{A\} + Pr\{B\}$

ESEMPIO

$$S = \{TT, TC, CT, CC\}$$

SUPPONIAMO CHE $Pr\{TT\} = Pr\{TC\} = Pr\{CT\} = Pr\{CC\} = \frac{1}{4}$.

ALLORA

- probabilità di ottenere almeno una testa $= Pr\{TT, TC, CT\} = \frac{3}{4}$

- probabilità di ottenere due monete diverse $= Pr\{TC, CT\} = \frac{1}{2}$

DISTRIBUZIONI DI PROBABILITÀ DISCRETE

- SONO DISTRIBUZIONI DI PROBABILITÀ DEFINITE IN UNO SPAZIO DI CAMPIONI FINITO O INFINITO NUMERABILE

- $Pr\{A\} = \sum_{s \in A} Pr\{s\}$

- SE S E' FINITO E PER OGNI $s \in S$ SI HA $Pr\{s\} = \frac{1}{|S|}$ SI HA UNA DISTRIBUZIONE DI PROBABILITÀ UNIFORME IN S

ESEMPIO

- n LANCI DI UNA MONETA PERFETTA ($Pr\{T\} = Pr\{C\} = \frac{1}{2}$)

  DISTRIBUZIONE DI PROBABILITÀ UNIFORME IN $S = \{T, C\}^n$

- OGNI EVENTO ELEMENTARE IN S HA PROBABILITÀ $\frac{1}{2^n}$

- SIA $A = \{$ SI OTTENGONO ESATTAMENTE k TESTE $\}$

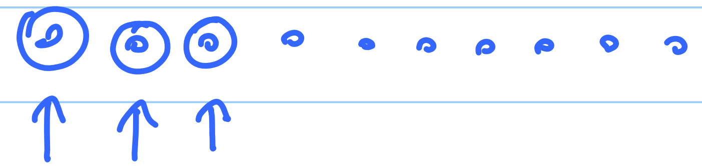
  SI HA: $|A| = \binom{n}{k} \implies Pr\{A\} = \binom{n}{k} / 2^n$

  (È UN CASO PARTICOLARE DI DISTRIBUZIONE BINOMIALE)

ESERCIZIO

DA UN MAZZO DI 10 CARTE NUMERATE DALL'1 AL 10 NE VENGONO RIMOSSE 3, UNA ALLA VOLTA.
QUAL E' LA PROBABILITA' CHE LE CARTE RIMOSSE SIANO SELEZIONATE IN ORDINE CRESCENTE?

SOLUZIONE



$10 \cdot 9 \cdot 8$ terne ordinate

$\#$ terne $= 10 \cdot 9 \cdot 8$

$\#$ terne crescenti $= \dfrac{10 \cdot 9 \cdot 8}{6}$

probabilità che una terna sia ordinata $= \dfrac{\frac{10 \cdot 9 \cdot 8}{6}}{10 \cdot 9 \cdot 8} = \dfrac{1}{6}$.

PROBABILITA' CONDIZIONATA

- COME CALCOLARE LA PROBABILITA' CHE SI VERIFICHI UN CERTO EVENTO A DATO IL VERIFICARSI DI UN ALTRO EVENTO B ?

- SIA B UN EVENTO TALE CHE $Pr\{B\} \neq 0$.

  ALLORA DEFINIAMO

$$Pr\{A \mid B\} =_{def} \frac{Pr\{A \cap B\}}{Pr\{B\}}$$ (PROBABILITA' CONDIZIONATA)

ESEMPIO $S = \{TT, TC, CT, CC\}$

$A = \{$ SI OTTENGONO DUE TESTE $\}$

$B = \{$ SI OTTIENE ALMENO UNA TESTA $\}$

$$Pr\{A \mid B\} = \left(\frac{1}{4}\right) \Big/ \left(\frac{3}{4}\right) = \frac{1}{4} \cdot \frac{4}{3} = \frac{1}{3}$$

ESERCIZIO

SIA B UN EVENTO AVENTE PROBABILITÀ NON NULLA.

DIMOSTRARE CHE $\Pr\{A|B\} + \Pr\{\bar{A}|B\} = 1$.

DIM

$$\Pr\{A|B\} = \frac{\Pr\{A\cap B\}}{\Pr\{B\}} \;; \quad \Pr\{\bar{A}|B\} = \frac{\Pr\{\bar{A}\cap B\}}{\Pr\{B\}},$$

POICHÉ $A\cap B$ E $\bar{A}\cap B$ SONO MUTUAMENTE ESCLUSIVI, SI HA:

$$\Pr\{A\cap B\} + \Pr\{\bar{A}\cap B\} = \Pr\{(A\cap B)\cup(\bar{A}\cap B)\}$$
$$= \Pr\{(A\cup\bar{A})\cap B\} = \Pr\{S\cap B\} = \Pr\{B\}$$

QUINDI: $\Pr\{A|B\} + \Pr\{\bar{A}|B\} = \dfrac{\Pr\{A\cap B\}}{\Pr\{B\}} + \dfrac{\Pr\{\bar{A}\cap B\}}{\Pr\{B\}}$

$$= \frac{\Pr\{A\cap B\} + \Pr\{\bar{A}\cap B\}}{\Pr\{B\}} = \frac{\Pr\{B\}}{\Pr\{B\}} = 1,$$

$\blacksquare$

EVENTI INDIPENDENTI

- A, B SONO INDIPENDENTI SE $Pr\{A \cap B\} = Pr\{A\} \cdot Pr\{B\}$.

- SE $Pr\{B\} \neq \emptyset$, CIÒ EQUIVALE A $Pr\{A | B\} = Pr\{A\}$.

ES. $A = \{\text{prima moneta } T\}$ $\quad Pr\{A\} = \frac{1}{2}$

$B = \{\text{una } T \text{ e una } C\}$ $\quad Pr\{B\} = \frac{1}{2}$

$$Pr\{A \cap B\} = Pr\{\text{prima moneta } T, \text{ seconda moneta } C\} = \frac{1}{4}$$

DUNQUE $Pr\{A \cap B\} = Pr\{A\} \cdot Pr\{B\} \Rightarrow A, B$ INDIPENDENTI

VARIABILI CASUALI DISCRETE

UNA VARIABILE CASUALE (DISCRETA) X E' UNA FUNZIONE REALE SULLO SPAZIO DI CAMPIONI S (FINITO O NUMERABILE)

$$x \in \mathbb{R}, \quad X = x \underset{def}{\equiv} \{s \in S : X(s) = x\}.$$

QUINDI: $$Pr\{X = x\} = \sum_{\substack{s \in S \\ X(s)=x}} Pr\{s\}$$

- $f(x) = Pr\{X = x\}$: FUNZIONE DI DENSITA' DI PROBABILITA' DELLA VARIABILE CASUALE X

ESEMPIO: LANCIO DI DUE DADI NORMALI A 6 FACCE

- SPAZIO DEI CAMPIONI : 36 EVENTI ELEMENTARI
- GLI EVENTI ELEMENTARI SONO EQUIPROBABILI, CIASCUNO CON PROBABILITA' $\frac{1}{36}$
- $X =_{\text{def}}$ max tra i due valori che presentano i dadi
- $Pr\{X=3\} = \frac{5}{36}$
- $Y =_{\text{def}}$ somma dei due valori che presentano i dadi
- $Pr\{Y=1\}=0$, $Pr\{Y=2\}=\frac{1}{36}$, $Pr\{Y=7\}=\frac{1}{6}$

- SIANO X, Y VARIABILI CASUALI SU UN MEDESIMO SPAZIO DI CAMPIONI S

LA FUNZIONE $f(x,y) = Pr\{X=x, Y=y\}$ E' LA FUNZIONE DENSITÀ DI PROBABILITÀ CONGIUNTA DI X E Y

- FISSATO y, SI HA:

$$Pr\{Y=y\} = \sum_x Pr\{X=x, Y=y\}$$

- ANALOGAMENTE, FISSATO x, SI HA:

$$Pr\{X=x\} = \sum_y Pr\{X=x, Y=y\}$$

- SI HA:

$$Pr\{X=x \mid Y=y\} = \frac{Pr\{X=x, Y=y\}}{Pr\{Y=y\}}$$

- LE VARIABILI CASUALI X, Y SONO INDIPENDENTI SE PER OGNI x, y GLI EVENTI $X = x$ E $Y = y$ SONO INDIPENDENTI, CIOÈ SE $\Pr\{X = x, Y = y\} = \Pr\{X = x\} \cdot \Pr\{Y = y\}$.

VALORE ATTESO DI UNA VARIABILE CASUALE (MEDIA)

$$E[X] =_{def} \sum_x x \Pr\{X = x\}$$

$$\begin{cases} X(T) = +3 \\ X(C) = -2 \end{cases}$$

Es. LANCIO DI 2 MONETE.

- SI VINCONO 3 EURO PER OGNI T
- SI PERDONO 2 EURO PER OGNI C

$$E[X] = 6 \cdot \Pr\{2T\} + 1 \cdot \Pr\{1T, 1C\} - 4 \Pr\{2C\}$$

$$= 6 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} - 4 \cdot \frac{1}{4}$$

$$= 1$$

LINEARITÀ DEL VALORE ATTESO

- $E[X+Y] = E[X] + E[Y]$
- SIANO X VARIABILE ALEATORIA, $g: \mathbb{R} \to \mathbb{R}$

  SI CONSIDERI LA VAR. ALEATORIA $g(X)$,

  ALLORA:

$$E[g(X)] = \sum_x g(x) \Pr\{X=x\}$$

$$g(x) = ax \implies E[aX] = a\,E[X]$$

$$E[aX+bY] = aE[X] + bE[Y]$$

- SE X, Y SONO INDIPENDENTI, ALLORA

$$E[XY] = \sum_x \sum_y xy \Pr\{X=x, Y=y\}$$

$$= \sum_x \sum_y xy \Pr\{X=x\} \Pr\{Y=y\}$$

$$= \left(\sum_x x \Pr\{X=x\}\right) \cdot \left(\sum_y y \Pr\{Y=y\}\right)$$

$$= E[X] \cdot E[Y]$$

- PIÙ IN GENERALE, SE $X_1, X_2, \dots, X_n$ SONO MUTUAMENTE INDIPENDENTI, SI HA:

$$E[X_1 \cdot X_2 \cdot \ldots \cdot X_n] = E[X_1] \cdot E[X_2] \cdot \ldots \cdot E[X_n]$$

- SE $X : S \longrightarrow \mathbb{N}$, SI HA:

$$E[X] = \sum_{i=0}^{\infty} i \Pr\{X = i\}$$

$$= \sum_{i=0}^{\infty} i \left(\Pr\{X \geq i\} - \Pr\{X \geq i+1\}\right)$$

$$= \sum_{i=1}^{\infty} \Pr\{X \geq i\}$$

VARIABILI CASUALI INDICATRICI

- S : SPAZIO DEI CAMPIONI
- $A \subseteq S$: EVENTO

VARIABILE CASUALE INDICATRICE $I\{A\}$ ASSOCIATA ALL'EVENTO A :

$$I\{A\}(s) = \begin{cases} 1 & \text{SE } s \in A \\ 0 & \text{SE } s \notin A \end{cases}$$

ES. $S = \{T, C\}$, $Pr\{T\} = Pr\{C\} = \frac{1}{2}$

$X_T =_{df}$ VARIABILE CASUALE INDICATRICE ASSOCIATA ALL'EVENTO T

$$E[X_T] = E[I\{T\}] = 1 \cdot Pr\{T\} + 0 \cdot Pr\{C\}$$

$$= 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2}$$

$$= \frac{1}{2}$$

LEMMA SE S E' LO SPAZIO DEI CAMPIONI E A E' UN EVENTO IN S, PONENDO $X_A \equiv I\{A\}$, SI HA:

$$E[X_A] = \Pr\{A\}.$$

DIM. SI HA:

$$
\begin{aligned}
E(X_A) &= E[I\{A\}] \\
&= 1 \cdot \Pr\{A\} + 0 \cdot \Pr\{\overline{A}\} \qquad (\overline{A} \equiv S \setminus A) \\
&= \Pr\{A\}.
\end{aligned}
$$

$\blacksquare$

- PROVE RIPETUTE DI BERNOULLI

  ESPERIMENTI (MUTUAMENTE INDIPENDENTI) CON DUE SOLI POSSIBILI RISULTATI:

  SUCCESSO, CON PROBABILITÀ p

  INSUCCESSO, CON PROBABILITÀ $q = 1-p$

- DALLE PROVE RIPETUTE DI BERNOULLI DERIVANO DUE IMPORTANTI DISTRIBUZIONI DI PROBABILITÀ:

  - DISTRIBUZIONE GEOMETRICA
  - DISTRIBUZIONE BINOMIALE

DISTRIBUZIONE GEOMETRICA

DATA UNA SEQUENZA DI PROVE RIPETUTE DI BERNOULLI (CON PROB. SUCCESSO p), QUANTE PROVE RIPETUTE DEVONO ESSERE EFFETTUATE PRIMA DI OTTENERE UN SUCCESSO?

- SIA X IL NUMERO DI PROVE RIPETUTE RICHIESTE PER OTTENERE UN SUCCESSO

- PER $k \geq 1$, SI HA

$$\Pr\{X = k\} = q^{k-1} \cdot p$$

(DISTRIBUZIONE GEOMETRICA)
(CON $q = 1-p$)

$$E(X) = \sum_{k=1}^{\infty} k \cdot q^{k-1} \cdot p = \frac{p}{q} \sum_{k=}^{\infty} k \cdot q^{k} = \frac{p}{q} \cdot \frac{q}{(1-q)^2} = \frac{1}{p},$$

- DUNQUE IN MEDIA OCCORRONO $\frac{1}{p}$ PROVE RIPETUTE PER OTTENERE UN SUCCESSO.

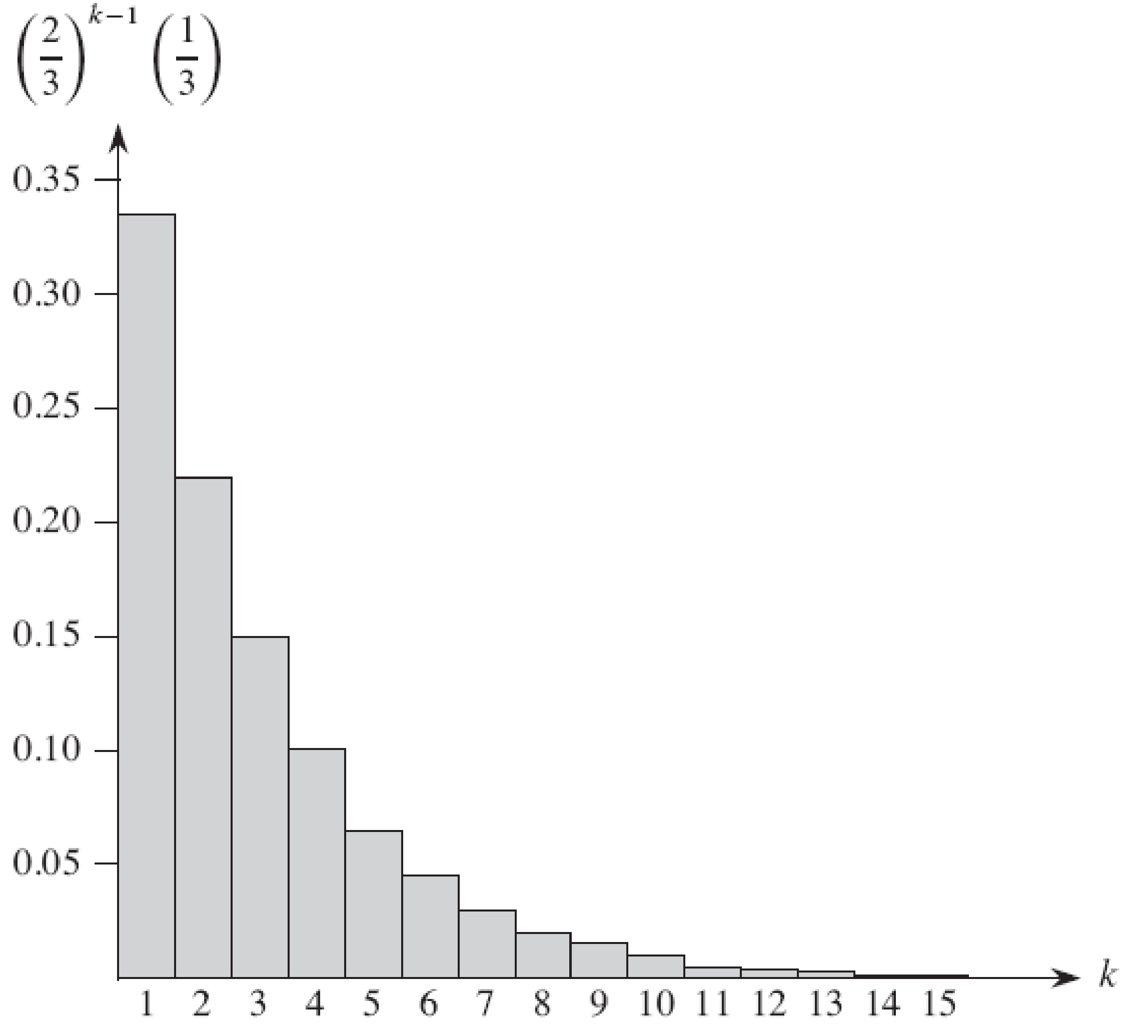


Figure C.1 A geometric distribution with probability $p = 1/3$ of success and a probability $q = 1 - p$ of failure. The expectation of the distribution is $1/p = 3$.

DISTRIBUZIONE BINOMIALE

QUANTI SUCCESSI SI VERIFICANO DURANTE n PROVE RIPETUTE DI BERNOULLI CON PROBABILITA' DI SUCCESSO p (E DI INSUCCESSO $q=1-p$) ?

- SIA X IL NUMERO DI SUCCESSI IN n PROVE RIPETUTE,
- PER $k=0,1,2,\dots,n$, SI HA:

$$\Pr\{X=k\} = \binom{n}{k} p^k q^{n-k}$$ (DISTRIBUZIONE BINOMIALE)

- PONIAMO: $$b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}$$

- SI HA: $$\sum_{k=0}^{n} b(k;n,p) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= (p+(1-p))^n = 1$$

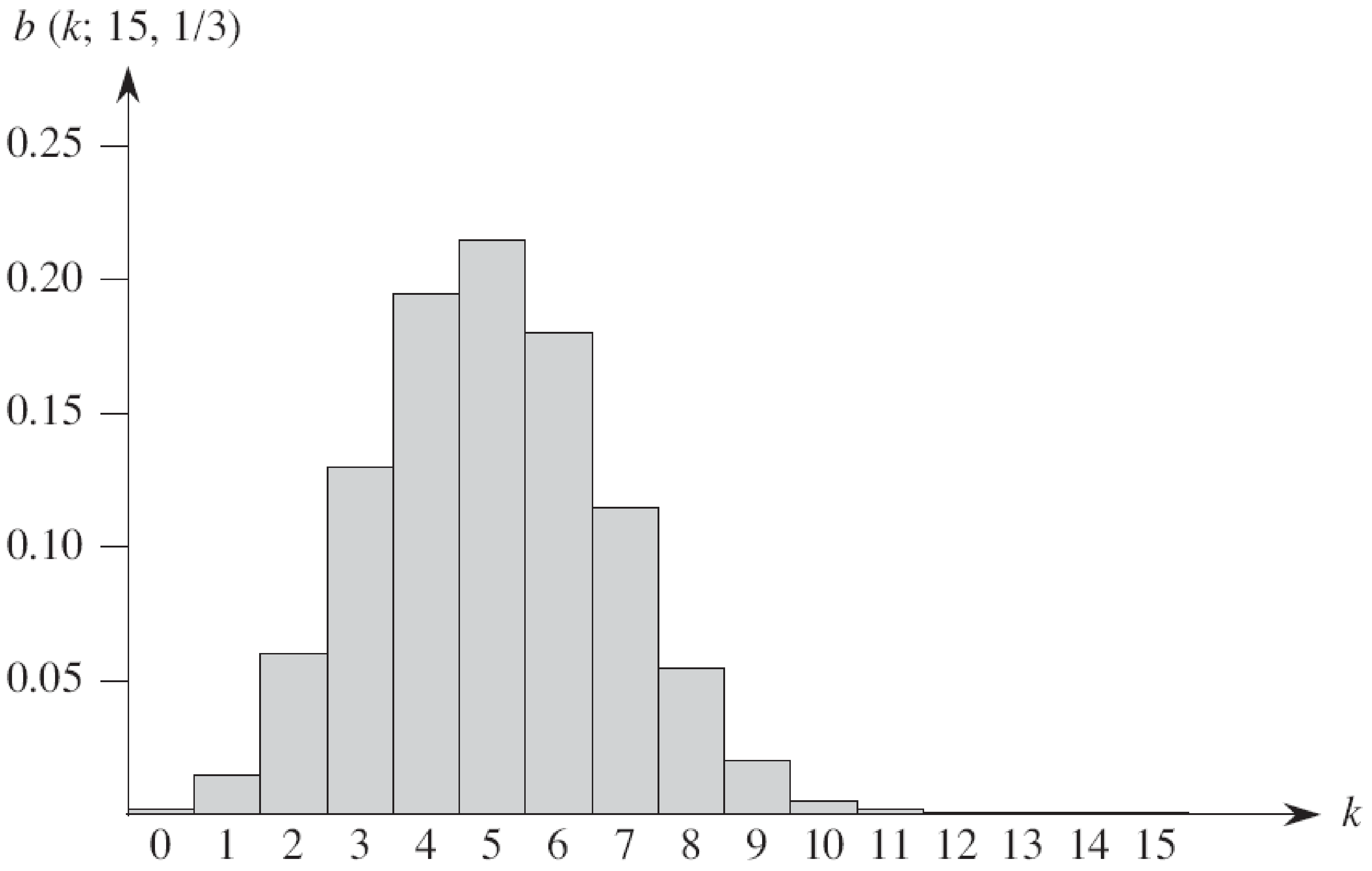


Figure C.2 The binomial distribution $b(k; 15, 1/3)$ resulting from $n = 15$ Bernoulli trials, each with probability $p = 1/3$ of success. The expectation of the distribution is $np = 5$.

$$E[X] = \sum_{k=0}^{n} k \Pr\{X=k\}$$

$$= \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \cdot \frac{n}{k} \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k}$$

$$= np \sum_{k=0}^{n-1} b(k; n-1, p)$$

$$= np$$

CALCOLO DI $E[X]$ MEDIANTE LE VARIABILI CASUALI INDICATRICI

- SIA $X_i = I\{\text{nell'i-esima prova ho successo}\}$
- DUNQUE, $X = \sum_{i=1}^{n} X_i$
- SI HA: $E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$

$$= \sum_{i=1}^{n} Pr\{\text{successo}\} = \sum_{i=1}^{n} p = n \cdot p .$$

ESERCIZIO

SI ABBIA UNA MONETA "TRUCCATA" TALE CHE

$$\Pr\{T\} = p \quad , \quad \Pr\{C\} = 1-p \quad , \text{ CON } \quad p \neq \frac{1}{2} ,$$

COME FARE A SIMULARE UNA MONETA NON TRUCCATA?

ESERCIZIO

SI ABBIA UNA MONETA "TRUCCATA" TALE CHE

$$Pr\{T\} = p \quad , \quad Pr\{C\} = 1-p \quad , \text{ CON } \quad p \neq \frac{1}{2} \, ,$$

COME FARE A SIMULARE UNA MONETA NON TRUCCATA?

SOLUZIONE

```
Moneta_Onesta( )
   while true
   do    x = Lancio_Moneta_Truccata()
         y = Lancio_Moneta_Truccata()
         if x ≠ y
         then  return x
   end while;
```

$$\Pr\{TC\} = \Pr\{CT\} = p(1-p)$$

$$\Pr\{CC\} = (1-p)^2$$

$$\Pr\{TC \vee CT\} = 2p(1-p)$$

$$\Pr\{TC \mid TC \vee CT\} = \frac{\Pr\{TC\}}{\Pr\{TC \vee CT\}} = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2}$$

$$\Pr\{CT \mid TC \vee CT\} = \frac{\Pr\{CT\}}{\Pr\{TC \vee CT\}} = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2}$$

IN MEDIA IL CICLO WHILE SARÀ RIPETUTO

$$\frac{1}{2p(1-p)}$$

VOLTE. ■

ESERCIZIO

- SIA DATA UNA FUNZIONE "RANDOM" RANDOM(0,1) TALE CHE

$$\Pr\{\text{RANDOM}(0,1) = 0\} = \Pr\{\text{RANDOM}(0,1) = 1\} = \frac{1}{2}$$

- SI DESCRIVA COME IMPLEMENTARE LA FUNZIONE PIÙ GENERALE RANDOM(a, b), CON $0 \le a < b$, $a, b \in \mathbb{N}$, TALE CHE

$$\Pr\{\text{RANDOM}(a,b) = i\} = \frac{1}{b-a+1}, \text{ PER } a \le i \le b$$

UTILIZZANDO LA FUNZIONE RANDOM(0,1).

SOLUZIONE

- E' SUFFICIENTE GENERARE NUMERI DA 0 A $m = b - a$ CON PROBABILITÀ UNIFORME
- SERVONO $k = \lfloor \log m \rfloor + 1$ BIT PER SCRIVERE m IN BINARIO
- CHIAMANDO k VOLTE RANDOM(0, 1), SI GENERERÀ UN NUMERO RANDOM x (IN BINARIO) TALE CHE

$$0 \leq x \leq 2^k - 1$$

CON PROBABILITÀ $\Pr\{x\} = \frac{1}{2^k}$

- SE $a + x \leq b$, $a + x$ È L'OUTPUT DELLA NOSTRA PROCEDURA, ALTRIMENTI x VIENE SCARTATO E SI RIPETE IL PASSO PRECEDENTE.

- LA PROBABILITA' DI SUCCESSO E'

$$\frac{b-a+1}{2^k} = \frac{b-a+1}{2^{\lfloor \lg(b-a) \rfloor + 1}}$$

- QUINDI, IN MEDIA SARANNO SCARTATI

$$\frac{2^{\lfloor \lg(b-a) \rfloor + 1}}{b-a+1} - 1$$

NUMERI PRIMA DI GENERARE UN OUTPUT VALIDO. $\blacksquare$

IL PROBLEMA DELLE ASSUNZIONI

- LISTA DI n CANDIDATI AVENTI RANGO DIVERSO A DUE A DUE

```
HIRE-ASSISTANT(n)
best = 0          // candidate 0 is a least-qualified dummy candidate
for i = 1 to n
    interview candidate i
    if candidate i is better than candidate best
        best = i
        hire candidate i
```

- COSTO COLLOQUIO = c_{COL}
- COSTO ASSUNZIONE = c_{ASS}
- SUPPONIAMO CHE $c_{COL} \ll c_{ASS}$

- SE VENGONO EFFETTUATE m ASSUNZIONI SI INCORRE NEL COSTO $n \cdot c_{COL} + m \cdot c_{ASS}$

- NEL CASO PESSIMO $m = n$

$\Longrightarrow$ COSTO TOTALE $= O(n \cdot c_{COL} + n \cdot c_{ASS})$

- IN MEDIA, QUANTE ASSUNZIONI SI EFFETTUERANNO?

- <u>IPOTESI</u>: I CANDIDATI ARRIVANO IN ORDINE CASUALE

- SE NON E' POSSIBILE FARE IPOTESI SULLA DISTRIBUZIONE DELL'INPUT, E' POSSIBILE FORZARE UNA DISTRIBUZIONE UNIFORME MEDIANTE LA TECNICA DELLA RANDOMIZZAZIONE

- IN GENERALE, UN ALGORITMO E' RANDOMIZZATO SE IL SUO COMPORTAMENTO E' DETERMINATO NON SOLTANTO DAL SUO INPUT, MA ANCHE DAI VALORI PRODOTTI DA UN GENERATORE DI NUMERI CASUALI (ES. RANDOM(a,b))

```
RANDOMIZED-HIRE-ASSISTANT(n)
randomly permute the list of candidates
best = 0            // candidate 0 is a least-qualified dummy candidate
for i = 1 to n
    interview candidate i
    if candidate i is better than candidate best
        best = i
        hire candidate i
```

ES. DATA LA LISTA $A_3 = [5, 2, 1, 8, 4, 7, 10, 9, 3, 6]$, HIRE-ASSISTANT EFFETTUERA' ESATTAMENTE 3 ASSUNZIONI, CIO' NON VALE PER RANDOMIZED-HIRE-ASSISTANT.

ANALISI PROBABILISTICA DEL PROBLEMA DELLE ASSUNZIONI

X = # DI ASSUNZIONI

$$E[X] = \sum_{x=1}^{n} x \Pr\{X = x\}$$

- UTILIZZIAMO LA TECNICA DELLE VARIABILI CASUALI INDICATRICI PER CALCOLARE $E[X]$
- PONIAMO $X_i = I\{\text{IL CANDIDATO } i \text{ E' ASSUNTO}\}$,

  CIOE' $X_i = \begin{cases} 1 & \text{SE IL CANDIDATO } i \text{ E' ASSUNTO} \\ 0 & \text{SE IL CANDIDATO } i \text{ NON E' ASSUNTO} \end{cases}$

- SI HA:

$$X = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i$$

- QUINDI:

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

- PER UN LEMMA PRECEDENTE, SAPPIAMO CHE

$$E[X_i] = \Pr\{\text{IL CANDIDATO } i \text{ E' ASSUNTO}\}$$

- QUANTO VALE $\Pr\{\text{IL CANDIDATO } i \text{ E' ASSUNTO}\}$?

- PER $1 \leq j_1, j_2 \leq i$, LA PROBABILITA' CHE IL MIGLIORE CANDIDATO TRA I PRIMI i SIA IL j_1-ESIMO E' UGUALE A QUELLA CHE SIA IL j_2-ESIMO

- DUNQUE, $\Pr\{\text{IL CANDIDATO } i \text{ E' ASSUNTO}\} = \dfrac{1}{i}$

- PERTANTO :

$$E[X] = \sum_{i=1}^{n} E[X_i]$$

$$= \sum_{i=1}^{n} Pr\{\text{IL CANDIDATO } i \text{ E' ASSUNTO}\}$$

$$= \sum_{i=1}^{n} \frac{1}{i} = \ln n + O(1)$$

- IN CONCLUSIONE, IL COSTO MEDIO DI RANDOMIZED-HIRE-ASSISTANT E' $O(n \cdot c_{COLL} + c_{ASS} \cdot \ln n)$

ESERCIZIO

QUAL E' IL NUMERO ATTESO DI INVERSIONI IN UN ARRAY DI n NUMERI DISTINTI?

ES. $A = [5, 2, 6, 3, 4, 1, 8]$

1 INVERSIONE

1 INVERSIONE

3 INVERSIONI

1 INVERSIONE

4 INVERSIONI

TOTALE: 10 INVERSIONI

- SIA $X = \#$ DI INVERSIONI IN A

- PER VALUTARE $E[X]$, UTILIZZEREMO LE SEGUENTI VARIABILI INDICATRICI:

$$X_{ij} = I\{A[i] > A[j]\}, \quad \text{PER} \quad i < j$$

- SI HA: $$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$$

$$E[X] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Pr\{A[i] > A[j]\}$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{2} = \frac{1}{2} \cdot \frac{(n-1) \cdot n}{2} = \frac{1}{4}(n-1)n.$$

RANDOMIZZAZIONE DEGLI ARRAY

```
PERMUTE-BY-SORTING(A)
n = A.length
let P[1..n] be a new array
for i = 1 to n
        P[i] = RANDOM(1, n^3)
sort A, using P as sort keys
```

COMPLESSITA': $O(n \lg n)$

NOTA: 1) LA PROBABILITA' CHE LE PRIORITA' $P[i]$, PER $i = 1, \ldots, n$, SIANO A DUE A DUE DISTINTE E' ALMENO $1 - \frac{1}{n}$

2) PERMUTE-BY-SORTING USA MEMORIA ESTERNA

UNA SOLUZIONE PIÙ EFFICIENTE

RANDOMIZE-IN-PLACE(A)

```
n = A.length
for i = 1 to n
        swap A[i] with A[RANDOM(i, n)]
```

COMPLESSITÀ: $\Theta(n)$

NOTA: 1) SIA PERMUTE-BY-SORTING CHE RANDOMIZE-IN-PLACE GENERANO UNA PERMUTAZIONE CASUALE UNIFORME

2) RANDOMIZE-IN-PLACE PERMUTA L'ARRAY SUL POSTO