Corso di Laurea Triennale in Matematica
Calcolo delle Probabilita' I (docenti G. Nappo, F. Spizzichino)

Seconda prova in itinere, venerdì 11 giugno 2004 (tempo a disposizione: 3 ore).
Scrivere su ogni foglio NOME e COGNOME. Fare il meglio possibile 3 esercizi (a scelta) e lasciare alla fine l'eventuale soluzione del quarto.
Le risposte devono essere giustificate sui fogli protocollo e riportate nel foglio RISPOSTE.

Esercizio 1. Indichiamo con $X_1, X_2, ...X_{200}$ i punteggi ottenuti in 200 lanci di un dado perfetto a quattro facce.

a) Trovare la distribuzione di probabilità, il valore atteso e la varianza della variabile aleatoria

$$U_1 = \min(X_1, X_2).$$

Si ponga

$$U_i = \min(X_{2i-1}, X_{2i}), \qquad i = 1, 2, ..., 100 \qquad \text{e} \qquad Y_n = \frac{\sum_{i=1}^{n} U_i}{n}.$$

Si ha che le variabili aleatorie $\{U_i;\ i = 1, 2, ..., 100\}$ sono completamente indipendenti.

b) Trovare il valore atteso e la varianza di Y_{100}.

c) Calcolare, in termini della disuguaglianza di Chebyishev, una limitazione inferiore per la probabilità dell'evento $\{\frac{7}{4} \leq Y_{100} \leq 2\}$.

d) Fornire l'approssimazione gaussiana per $P(\frac{7}{4} \leq Y_{100} \leq 2)$ in termini del teorema centrale del limite.

e) FACOLTATIVO Mostrare che U_1 ed U_2 sono indipendenti.

Esercizio 2. Una moneta perfetta viene lanciata 6 volte e indichiamo con X_i $(1 \leq i \leq 6)$ la variabile aleatoria binaria, indicatrice dell'evento $E_i \equiv \{\text{risultato testa all'i-esimo lancio}\}$.

Consideriamo ora le variabili aleatorie

$$T = \sum_{i=1}^{2} X_i, \quad U = \sum_{i=3}^{4} X_i, \quad V = \sum_{i=5}^{6} X_i$$

$$Z = T + U, \qquad W = T + V.$$

a) Z e W sono stocasticamente indipendenti?

b) Trovare la distribuzione di probabilità, il valore atteso e la varianza di Z.

c) Calcolare la covarianza fra Z e W

d) Calcolare il valore atteso di $2ZW$

e) Calcolare la covarianza fra $(2Z + 1)$ e W

f) Trovare la distribuzione di probabilità congiunta di Z e W.

Esercizio 3. Sia X una variabile aleatoria con funzione di densità di probabilità

$$f_X(x) = \begin{cases} kx & \text{per } 0 \leq x \leq 1 \\ k(2-x) & \text{per } 1 \leq x \leq 2 \\ 0 & \text{altrimenti} \end{cases}$$

a) Trovare il valore della costante k.

b) Trovare il valore atteso di X (**attenzione:** è possibile trovarlo senza fare calcoli?)

c) Trovare la funzione di distribuzione di X.

Esercizio 4. Alla fine della sessione estiva viene fatta una statistica fra 20 studenti del primo anno. Per ogni studente si conta quanti esami $(0, 1, 2,$ o $3)$ abbia superato fra quelli dei corsi del secondo semestre (ovvero AM1, ALG1, e CP1) e si pone

X_0 = numero degli studenti, nel campione intervistato, che non hanno superato nessuno di quegli esami

e analogamente, per $i = 1, 2, 3$

X_i = numero degli studenti, nel campione intervistato, che hanno superato esattamente i di quegli esami.

Si assuma per semplicità che per ciascuno studente la probabilità di superare tutti e tre gli esami sia $\frac{1}{2}$, di superarne due sia $\frac{1}{4}$ e di superarne uno solo sia $\frac{1}{8}$ (e quindi la probabilità di non superarne nessuno sia $\frac{1}{8}$), e che eventi relativi a studenti diversi siano tutti indipendenti.

a) Calcolare (scrivere l'espressione per) la probabilità che *esattamente 13 studenti superino esattamente 2 esami*

b) Calcolare (scrivere l'espressione per) la probabilità che *esattamente 18 studenti superino almeno un esame.*

c) Calcolare (scrivere l'espressione per) la probabilità che *esattamente 4 studenti superino esattamente 1 esame, 7 studenti superino esattamente 2 esami e 6 studenti superino esattamente 3 esami.*

d) **FACOLTATIVO** Sapendo che $X_0 = 2$, calcolare (scrivere l'espressione per) la probabilità che $X_2 = 13$.

Corso di Laurea Triennale in Matematica
Calcolo delle Probabilita' I

Seconda prova in itinere, venerdì 11 giugno 2004 - FOGLIO RISPOSTE

NOME E COGNOME ....

docente G. Nappo □, F. Spizzichino □
textbfEsercizio 1.

a) Distribuzione di probabilità di U_1 svolto □ non svolto □

valore atteso di U_1=............ varianza di U_1=............

b) valore atteso di Y_{100}=........... varianza di Y_{100}=...........

c) dalla disuguaglianza di Chebyishev, si ha $P(\frac{7}{4} \leq Y_{100} \leq 2)$.................

d) $P(\frac{7}{4} \leq Y_{100} \leq 2) \simeq$.................

e) **FACOLTATIVO** Mostrare che U_1 ed U_2 sono indipendenti. svolto □ non svolto □

Esercizio 2.

a) Z e W sono stocasticamente indipendenti? si □ no □

b) la distribuzione di probabilità svolto □ non svolto □

valore atteso di Z=.......... varianza di Z=..........

c) $Cov(Z, W)$=..................

d) valore atteso di $2ZW$=....................

e) $Cov(2Z + 1, W)$=....................

f) distribuzione di probabilità congiunta di Z e W svolto □ non svolto □

Esercizio 3.

a) k=................

b) valore atteso di X=.....................

c) funzione di distribuzione di X svolto □ non svolto □

Esercizio 4.

a) probabilità che esattamente 13 studenti superino esattamente 2 esami=.......................................

b) probabilità che 18 studenti superino almeno un esame=.......................................

c) probabilità che esattamente 4 studenti superino esattamente 1 esame 7 studenti superino esattamente 2 esami e 6 studenti superino esattamente 3 esami=.......................................

d) **FACOLTATIVO** Sapendo che $X_0 = 2$, la probabilità che $X_2 = 13$ vale...

Corso di Laurea Triennale in Matematica
Calcolo delle Probabilita' I (docenti G. Nappo, F. Spizzichino)

Seconda prova in itinere, venerdì 11 giugno 2004- SOLUZIONE

Esercizio 1. Indichiamo con $X_1, X_2, ...X_{200}$ i punteggi ottenuti in 200 lanci di un dado perfetto a quattro facce.

a) **Trovare la distribuzione di probabilità, il valore atteso e la varianza della variabile aleatoria**

$$U_1 = \min(X_1, X_2).$$

- **soluzione di a)** Si ha

$$P(U_1 = k) = \frac{9-2k}{16} \qquad \text{per } k = 1, 2, 3, 4$$

ovvero

$$P(U_1 = 1) = \tfrac{7}{16} \quad P(U_1 = 2) = \tfrac{5}{16} \quad P(U_1 = 3) = \tfrac{3}{16} \quad P(U_1 = 4) = \tfrac{1}{16}$$
$$E(U_1) = \tfrac{15}{8}(= 1,875) \qquad Var(U_1) = \tfrac{55}{64}(= 0.859375 \simeq 0.859)$$

Infatti U_1 può assumere solo i valori $k = 1, 2, 3, 4$ e inoltre

$$P(U_1 > k) = P(X_1 > k, X_2 > k) = P(X_1 > k)P(X_2 > k) = \frac{(4-k)^2}{4^2}$$

da cui

$$P(U_1 = k) = P(U_1 > k-1) - P(U_1 > k) = \frac{\big((4-(k-1)\big)^2}{4^2} - \frac{(4-k)^2}{4^2}$$
$$= \frac{4^2 - 8(k-1) + (k-1)^2 - 4^2 + 8k - k^2}{4^2} = \frac{8-2k+1}{4^2}$$

$$E(U_1) = 1 \cdot \tfrac{7}{16} + 2 \cdot \tfrac{5}{16} + 3 \cdot \tfrac{3}{16} + 4 \cdot \tfrac{1}{16} = \tfrac{7+10+9+4}{16} = \tfrac{30}{16} = \tfrac{15}{8} = 1,875$$
$$E(U_1^2) = 1^2 \cdot \tfrac{7}{16} + 2^2 \cdot \tfrac{5}{16} + 3^2 \cdot \tfrac{3}{16} + 4^2 \cdot \tfrac{1}{16} = \tfrac{7+20+27+16}{16} = \tfrac{70}{16} = \tfrac{35}{8}$$

da cui

$$Var(U_1) = \tfrac{35}{8} - \left(\tfrac{15}{8}\right)^2 = \tfrac{1}{64}(35 \cdot 8 - 15^2) = \tfrac{1}{64}(280 - 225) = \tfrac{55}{64} = 0.859375 \simeq 0.859$$

Si ponga

$$U_i = \min(X_{2i-1}, X_{2i}), \qquad i = 1, 2, ..., 100 \qquad \text{e} \qquad Y_n = \frac{\sum_{i=1}^n U_i}{n}.$$

Si ha che le variabili aleatorie $\{U_i;\ i = 1, 2, ..., 100\}$ **sono completamente indipendenti.**

b) **Trovare il valore atteso e la varianza di** Y_{100}.

- **soluzione di b)**

$$E(Y_{100}) = \tfrac{100E(U_1)}{100} = E(U_1) = \tfrac{15}{8} \qquad Var(Y_{100}) = \tfrac{1}{100^2}\, 100\, Var(U_1) = \tfrac{1}{100}\, \tfrac{55}{64} \simeq 0.00859$$

Infatti per calcolare il valore atteso basta applicare la proprietà di linearità del valore atteso, e osservare che le variabili aleatorie U_i hanno tutte la stessa distribuzione, e quindi lo stesso valore atteso. Per calcolare la varianza basta ricordare che $Var(aX) = a^2 Var(X)$ e che la varianza della somma è la somma delle varianze quando le variabili aleatorie sono non correlate, e le variabili aleatorie U_i sono indipendenti, e quindi non correlate.

c) **Calcolare, in termini della disuguaglianza di Chebyishev, una limitazione inferiore per la probabilità dell'evento $\{\frac{7}{4} \leq Y_{100} \leq 2\}$.**

- **soluzione di c)**

$$P(\{\tfrac{7}{4} \leq Y_{100} \leq 2\}) = P(\{\tfrac{14}{8} \leq Y_{100} \leq \tfrac{16}{8}\}) = P(-\tfrac{1}{8} \leq Y_{100} - \tfrac{15}{8} \leq \tfrac{1}{8})$$
$$\geq 1 - \frac{55}{64} \frac{1}{\left(\frac{1}{8}\right)^2 100} = 1 - \frac{55}{100} = \frac{45}{100}$$

d) **Fornire l'approssimazione gaussiana per $P(\frac{7}{4} \leq Y_{100} \leq 2)$ in termini del teorema centrale del limite.**

- **soluzione di d)**

$$P(\{\frac{7}{4} \leq Y_{100} \leq 2\}) = P(-\frac{1}{8} \leq Y_{100} - \frac{15}{8} \leq \frac{1}{8}) = P(-\frac{1}{8} \leq \frac{S_{100} - 100\frac{15}{8}}{100} \leq \frac{1}{8})$$

$$= P(-\sqrt{\frac{100}{Var(U_1)}} \cdot \frac{1}{8} \leq \sqrt{\frac{100}{Var(U_1)}} \cdot \frac{S_{100} - 100\frac{15}{8}}{100} \leq \sqrt{\frac{100}{Var(U_1)}} \cdot \frac{1}{8})$$

$$= P(-\sqrt{\frac{100}{\frac{55}{64}}} \cdot \tfrac{1}{8} \leq \frac{S_{100} - 100E(U_1)}{\sqrt{100Var(U_1)}} \leq \sqrt{\frac{100}{\frac{55}{64}}} \cdot \tfrac{1}{8})$$

$$= P(-\sqrt{\frac{100}{55}} \leq \frac{S_{100} - 100E(U_1)}{\sqrt{100Var(U_1)}} \leq \sqrt{\frac{100}{55}})$$

$$\simeq \Phi(\sqrt{\frac{100}{55}}) - \Phi(-\sqrt{\frac{100}{55}}) = 2\Phi(\sqrt{\frac{100}{55}}) - 1$$

$$\simeq 2\Phi(1,34) - 1 = 2 \cdot 0.9099 - 1 = 1.8198 - 1 = 0.8198$$

e) **FACOLTATIVO Mostrare che U_1 ed U_2 sono indipendenti.**
(SOLUZIONE IN APPENDICE)

Esercizio 2. Una moneta perfetta viene lanciata 6 volte e indichiamo con X_i $(1 \leq i \leq 6)$ la variabile aleatoria binaria, indicatrice dell'evento $E_i \equiv$ {risultato testa all'i-esimo lancio}.
Consideriamo ora le variabili aleatorie

$$T = \sum_{i=1}^{2} X_i, \quad U = \sum_{i=3}^{4} X_i, \quad V = \sum_{i=5}^{6} X_i$$

$$Z = T + U, \qquad W = T + V.$$

a) **Z e W sono stocasticamente indipendenti?**

- **soluzione di a)** Le variabili aleatorie Z e W non sono stocasticamente indipendenti. Infatti ad esempio
$$P(Z=0,W=0)=\left(\tfrac{1}{2}\right)^6$$
mentre
$$P(Z=0)\,P(W=0)=\left(\tfrac{1}{2}\right)^4\,\left(\tfrac{1}{2}\right)^4$$

Un altro modo per mostrare questo fatto potrebbe essere, ad esempio, svolgere prima il punto **c)** e dopo aver constatato che $Cov(Z,W)\neq 0$ dedurre che non possono essere indipendenti (se due variabili aleatorie sono indipendenti allora sono necessariamente scorrelate).

b) **Trovare la distribuzione di probabilità, il valore atteso e la varianza di Z.**

- **soluzione di b)**
$$Z\sim bin(4,\tfrac{1}{2})\quad \text{e quindi}\quad E(Z)=4\,\tfrac{1}{2}=2,\quad Var(Z)=4\,\tfrac{1}{2}\,\tfrac{1}{2}=1.$$
Infatti, $Z=T+U=\sum_{i=1}^{4}X_i$, con X_i variabili aleatorie binarie, indipendenti e con la stessa distribuzione: $P(X_i=1)=P(X_i=0)=\frac{1}{2}$

c) **Calcolare la covarianza fra Z e W**

- **soluzione di c)**
$$Cov(Z,W)=1$$
Infatti, per la proprietà di bilinearità della covarianza si ha
$$\begin{aligned}Cov(Z,W)&=Cov(T+U,T+Z)=Cov(T,T)+Cov(T,Z)+Cov(U,T)+Cov(U,V)\\&=Var(T),=2\tfrac{1}{2}\tfrac{1}{2}=\tfrac{1}{2}\end{aligned}$$
in quanto in generale $Cov(X,X)=Var(X)$ e $Cov(X,Y)=0$ se X ed Y sono indipendenti, e le variabili aleatorie T, U, V sono indipendenti (sia globalmente che a due a due)

d) **Calcolare il valore atteso di $2ZW$**

- **soluzione di d)**
$$E(2ZW)=9$$
Infatti
$$E(2ZW)=2E(ZW)=2\big(Cov(ZW)+E(Z)\,E(W)\big)=2\big(\tfrac{1}{2}+2\cdot 2\big)=9$$
in quanto W ha la stessa legge di Z.
Alternativamente si può calcolare direttamente (tenendo conto del fatto che $E(T^2)=Var(T)+\big(E(T)\big)^2=\frac{1}{2}+1^2=\frac{3}{2}$)
$$\begin{aligned}E(2ZW)&=2E(ZW)=2E\big((T+U)(T+V)\big)=2\big(E(T^2)+E(TV)+E(UT)+E(UV)\big)\\&=2\big(E(T^2)+E(T)E(V)+E(U)E(T)+E(U)E(V)\big)=2\big(\tfrac{3}{2}+1+1+1\big)=9\end{aligned}$$

e) **Calcolare la covarianza fra $(2Z+1)$ e W**

- **soluzione di e)**
$$Cov(2Z+1,W)=1$$
Infatti, per la proprietà di bilinearità della covarianza
$$Cov(2Z+1,W)=2Cov(Z,W)+Cov(1,W)=2\,\tfrac{1}{2}+0=1$$
($Cov(c,X)=0$ in quanto $E(c\,X)=cE(X)=E(c)E(X)$ e $Cov(c,X)=E(c\,X)-E(c)\,E(X)$)

f) Trovare la distribuzione di probabilità congiunta di Z e W.

- **soluzione di f)** La distribuzione congiunta di Z e W è individuata da
$$P(Z=i,W=j)=\sum_{h=0}^{2}P(T=h)P(U=i-h)P(V=j-h),\qquad i,j\in\{0,1,2,3,4\}$$
o meglio
$$P(Z=i,W=j)=\sum_{h=0:i-h,j-h\in\{0,1,2\}}^{2}P(T=h)P(U=i-h)P(V=j-h),\qquad i,j\in\{0,1,2,3,4\}$$
Infatti
$$P(Z=i,W=j)=\sum_{h=0}^{2}P(T=h)P(Z=i,W=j|\{T=h\})$$
e inoltre
$$P(Z=i,W=j|\{T=h\})=\frac{P(Z=i,W=j,T=h)}{P(T=h)}=\frac{P(T+U=i,T+V=j,T=h)}{P(T=h)}$$
L'evento
$$\{T+U=i,T+V=j,T=h\}=\{h+U=i,h+V=j,T=h\}=\{U=i-h,V=j-h,T=h\}$$
e quindi
$$\begin{aligned}\frac{P(T+U=i,T+V=j,T=h)}{P(T=h)}&=\frac{P(U=i-h)P(V=j-h)P(T=h)}{P(T=h)}\\&=P(U=i-h)P(W=j-h)\end{aligned}$$
Inoltre ovviamente bisogna tenere conto del fatto che
$$P(U=i-h)P(V=j-h)\neq 0\quad\text{se e solo se } i-h,j-h\in\{0,1,2\}$$
e quindi alcuni valori di $P(Z=i,W=j)$ sono nulli.
Ad esempio $P(Z=0,W=3)=P(Z=0,W=4)=P(Z=3,W=0)$ $=P(Z=4,W=0)=0$, ma anche $P(Z=1,W=4)=P(Z=4,W=1)=0$
Più precisamente si ha
$$\begin{aligned}P(Z=i,W=j)&=\sum_{h=0:i-h,j-h\in\{0,1,2\}}^{2}\binom{2}{h}\left(\tfrac{1}{2}\right)^2P(U=i-h)P(V=j-h)\\&=\sum_{h=0:i-h,j-h\in\{0,1,2\}}^{2}\binom{2}{h}\left(\tfrac{1}{2}\right)^2\binom{2}{i-h}\left(\tfrac{1}{2}\right)^2\binom{2}{j-h}\left(\tfrac{1}{2}\right)^2\\&=\sum_{h=0:i-h,j-h\in\{0,1,2\}}^{2}\binom{2}{h}\binom{2}{i-h}\binom{2}{j-h}\left(\tfrac{1}{2}\right)^6\end{aligned}$$

$W\backslash Z$	0	1	2	3	4
0	$\frac{1}{64}$	$\frac{2}{64}$	$\frac{1}{64}$	0	0
1	$\frac{2}{64}$	$\frac{6}{64}$	$\frac{6}{64}$	$\frac{2}{64}$	0
2	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{10}{64}$	$\frac{6}{64}$	$\frac{1}{64}$
3	0	$\frac{2}{64}$	$\frac{6}{64}$	$\frac{6}{64}$	$\frac{2}{64}$
4	0	0	$\frac{1}{64}$	$\frac{2}{64}$	$\frac{1}{64}$

A titolo d'esempio calcoliamo

$$P(Z=1, W=1) = \binom{2}{0}\binom{2}{1}\binom{2}{1}\left(\tfrac{1}{2}\right)^6 + \binom{2}{1}\binom{2}{0}\binom{2}{0}\left(\tfrac{1}{2}\right)^6 = (4+2)\tfrac{1}{64}$$

Esercizio 3. Sia X una variabile aleatoria con funzione di densità di probabilità

$$f_X(x) = \begin{cases} kx & \text{per } 0 \le x \le 1 \\ k(2-x) & \text{per } 1 \le x \le 2 \\ 0 & \text{altrimenti} \end{cases}$$

a) **Trovare il valore della costante k.**

- **soluzione di a)** La costante $k = 1$.
  Infatti si tratta di trovare k in modo che $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$. Per cui,

$$\int_{-\infty}^{\infty} f_X(x)\,dx = k\left(\int_0^1 x\,dx + \int_1^2 (2-x)\,dx\right) = k\left(\tfrac{x^2}{2}\Big|_0^1 + 2 - \tfrac{x^2}{2}\Big|_1^2\right)$$
$$= k \cdot 1 = 1 \qquad \text{se e solo se } k = 1$$

b) **Trovare il valore atteso di X (attenzione: è possibile trovarlo senza fare calcoli?)**

- **soluzione di b)** Il valore atteso di X vale 1
  Infatti si tratta di calcolare

$$\int_{-\infty}^{\infty} x\,f_X(x)\,dx = \int_0^1 x^2\,dx + \int_1^2 x\,(2-x)\,dx = \tfrac{x^3}{3}\Big|_0^1 + x^2\Big|_1^2 - \tfrac{x^3}{3}\Big|_1^2$$
$$= \tfrac{1}{3} + 4 - 1 - \tfrac{8}{3} + \tfrac{1}{3} = 1$$

  Tuttavia per motivi di simmetria si ha che tale integrale deve venire 1 infatti il grafico della densità è simmetrico rispetto all'asse $x = 1$ e quindi, tenendo presente l'analogia con il baricentro, si ha immediatamente che il valore atteso deve essere 1.

c) **Trovare la funzione di distribuzione di X.**

- **soluzione di c)** La funzione di distribuzione F_X è data da

$$F_X(x) = \begin{cases} 0 & \text{per } -\infty < x < 0 \\ \frac{x^2}{2} & \text{per } 0 \le x < 1 \\ 2x - \frac{x^2}{2} - 1 & \text{per } 1 \le x < 2 \\ 1 & \text{per } 2 \le x \end{cases}$$

  Infatti chiaramente per ogni $x \le 0$

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = 0,$$

per $0 \leq x \leq 1$

$$F_X(x) = F_X(0) + \int_0^x y\, dy = 0 + \tfrac{x^2}{2},$$

per $1 \leq x < 2$

$$\begin{aligned} F_X(x) &= F_X(1) + \int_1^x (2-y)\, dy = \tfrac{1}{2} + \int_1^x (2-y)\, dy = \tfrac{1}{2} + 2\,y\big|_1^x - \tfrac{y^2}{2}\Big|_1^x \\ &= \tfrac{1}{2} + 2x - 2 - \tfrac{x^2}{2} + \tfrac{1}{2} = 2x - \tfrac{x^2}{2} - 1, \end{aligned}$$

ed infine per $x \geq 2$

$$F_X(x) = F_X(2) + \int_2^x f_X(y)\, dy = 1 + 0 = 1.$$

Esercizio 4. Alla fine della sessione estiva viene fatta una statistica fra 20 studenti del primo anno. Per ogni studente si conta quanti esami ($0, 1, 2$, o 3) abbia superato fra quelli dei corsi del secondo semestre (ovvero AM1, ALG1, e CP1) e si pone

$X_0 =$ **numero degli studenti, nel campione intervistato, che non hanno superato nessuno di quegli esami**

e analogamente, per $i = 1, 2, 3$

$X_i =$ **numero degli studenti, nel campione intervistato, che hanno superato esattamente** i **di quegli esami.**

Si assuma per semplicità che per ciascuno studente la probabilità di superare tutti e tre gli esami sia $\frac{1}{2}$, di superarne due sia $\frac{1}{4}$ e di superarne uno solo sia $\frac{1}{8}$ (e quindi la probabilità di non superarne nessuno sia $\frac{1}{8}$), e che eventi relativi a studenti diversi siano tutti indipendenti.

a) **Calcolare (scrivere l'espressione per) la probabilità che** *esattamente 13 studenti superino esattamente 2 esami*

- **soluzione di a)** La probabilità cercata vale

$$\binom{20}{13} \left(\frac{1}{4}\right)^{13} \left(\frac{3}{4}\right)^{7}.$$

Infatti, per $i = 0, 1, 2, 3$ e variabili aleatorie X_i contano il numero di successi in 20 prove (prova k-sima$\equiv$ studente k-simo) indipendenti in cui successo significa che lo studente supera i esami, e quindi

$$X_i \sim bin(20, p_i), \qquad \text{dove } p_0 = \frac{1}{8}, \quad p_1 = \frac{1}{8}, \quad p_2 = \frac{1}{4}, \quad p_3 = \frac{1}{2}.$$

In particolare $X_2 \sim bin(20, \frac{1}{4})$ e l'evento

$$\{\textit{esattamente 13 studenti superino esattamente 2 esami}\} = \{X_2 = 13\}$$

da cui

$$P(X_2 = 13) = \binom{20}{13} \left(\frac{1}{4}\right)^{13} \left(\frac{3}{4}\right)^{7}.$$

b) Calcolare (scrivere l'espressione per) la probabilità che *esattamente 18 studenti superino almeno un esame.*

- **soluzione di b)** La probabilità cercata vale

$$\binom{20}{2}\left(\frac{1}{8}\right)^2\left(\frac{7}{8}\right)^{18}=\binom{20}{18}\left(\frac{7}{8}\right)^{18}\left(\frac{1}{8}\right)^2$$

L'evento

$$\{\textit{esattamente 18 studenti superino almeno un esame}\}=\{X_0=2\}$$

e quindi la probabilità cercata vale

$$P(X_0=2)=\binom{20}{2}\left(\frac{1}{8}\right)^2\left(\frac{7}{8}\right)^{18}.$$

Alternativamente si puo' osservare che la variabile aleatoria $Y=X_1+X_2+X_3$ conta il numero di successi su 20 prove (prova k-sima$\equiv$ studente k-simo) indipendenti in cui successo significa che lo studente supera almeno un esame, e quindi

$$Y\sim bin(20,p_1+p_2+p_3)=bin(20,\frac{1}{8}+\frac{1}{4}+\frac{1}{2})=bin(20,\frac{7}{8})$$

da cui l'evento

$$\{\textit{esattamente 18 studenti superino almeno un esame}\}=\{Y=18\}$$

e quindi la probabilità cercata vale

$$P(Y=18)=\binom{20}{18}\left(\frac{7}{8}\right)^{18}\left(\frac{1}{8}\right)^2.$$

c) Calcolare (scrivere l'espressione per) la probabilità che *esattamente 4 studenti superino esattamente 1 esame, 7 studenti superino esattamente 2 esami e 6 studenti superino esattamente 3 esami.*

- **soluzione di c)** La probabilità cercata vale

$$\frac{20!}{3!\,4!\,7!\,6!}\left(\frac{1}{8}\right)^3\left(\frac{1}{8}\right)^4\left(\frac{1}{4}\right)^7\left(\frac{1}{2}\right)^6$$

Infatti l'evento {*esattamente 4 studenti superino esattamente 1 esame, 7 studenti superino esattamente 2 esami e 6 studenti superino esattamente 3 esami*} coincide con l'evento

$$\{X_1=4\,X_2=7,\;X_3=6\}=\{X_0=3,\;X_1=4\,X_2=7,\;X_3=6\}$$

e quindi la probabilità cercata vale

$$P(\{X_0=3,\;X_1=4\,X_2=7,\;X_3=6\})=\frac{20!}{3!\,4!\,7!\,6!}\left(\frac{1}{8}\right)^3\left(\frac{1}{8}\right)^4\left(\frac{1}{4}\right)^7\left(\frac{1}{2}\right)^6.$$

d) FACOLTATIVO Sapendo che $X_0 = 2$, calcolare la probabilità che $X_2 = 13$.

- **soluzione di d)**

$$
\begin{aligned}
P(X_2 = 13 | X_0 = 2) &= \binom{18}{13} \left(\frac{p_2}{1-p_0}\right)^{13} \left(1 - \frac{p_2}{1-p_0}\right)^{5} \\
&= \binom{18}{13} \left(\frac{\frac{1}{4}}{1-\frac{1}{8}}\right)^{13} \left(1 - \frac{\frac{1}{4}}{1-\frac{1}{8}}\right)^{5} \\
&= \binom{18}{13} \left(\frac{2}{7}\right)^{13} \left(\frac{5}{7}\right)^{5}
\end{aligned}
$$

Infatti **se sappiamo che 2 studenti non hanno superato nessun esame**, allora i diciotto studenti rimasti potranno aver superato 1 esame oppure 2 oppure 3, con probabilità

$$\alpha_i = \frac{p_i}{p_1 + p_2 + p_3} = \frac{p_i}{1 - p_0}.$$

Quindi le variabili aleatorie X_i, **condizionatamente a sapere che** $\{X_0 = 2\}$, hanno distribuzione $bin(18, \alpha_i)$.

APPENDICE- soluzione del punto e) dell'ESERCIZIO 1

[**e) FACOLTATIVO**] **Mostrare che U_1 ed U_2 sono indipendenti.**

- **soluzione di e)**

Il motivo per cui sono indipendenti è il seguente: per cominciare basta determinare se le partizioni di eventi

$$\big\{\{U_1 = h\};\ h = 1,2,3,4\big\} \quad \text{e} \quad \big\{\{U_2 = k\};\ k = 1,2,3,4\big\}$$

sono indipendenti, ovvero se

$$P(U_1 = h,\ U_2 = k) = P(U_1 = h)P(U_2 = k) \qquad \text{per ogni } h,\ k = 1,2,3,4.$$

Infatti

$$\{U_1 = h\} = \{\min(X_1, X_2) = h\} \qquad \{U_2 = k\} = \{\min(X_3, X_4) = k\}$$

quindi

$$\{U_1 = h\} \in \mathcal{G}(\mathcal{H}_{1,2}), \quad \text{dove } \mathcal{H}_{1,2} \text{ è la partizione } \big\{\{X_1 = i, X_2 = j\}, i,j = 1,2,3,4\big\}$$

e

$$\mathcal{G}(\mathcal{H}_{1,2}) \text{ è l'algebra generata dalla partizione } \mathcal{H}_{1,2}$$

e

$$\{U_2 = k\} \in \mathcal{G}(\mathcal{H}_{3,4}), \quad \text{dove } \mathcal{H}_{3,4} \text{ è la partizione } \big\{\{X_3 = i, X_4 = j\}, i,j = 1,2,3,4\big\}$$

e

$$\mathcal{G}(\mathcal{H}_{3,4}) \text{ è l'algebra generata dalla partizione } \mathcal{H}_{3,4}$$

Le partizioni $\mathcal{H}_{1,2}$ e $\mathcal{H}_{3,4}$ sono indipendenti: infatti

$$
\begin{aligned}
P(\{X_1 = i,\ X_2 = j\} \bigcap \{X_3 = i',\ X_4 = j'\}) &= P(X_1 = i)P(X_2 = j)P(X_3 = i')P(X_4 = j') \\
&= P(\{X_1 = i,\ X_2 = j\})P(\{X_3 = i',\ X_4 = j'\})
\end{aligned}
$$

La precedente dimostrazione è un caso particolare del seguente: se $Z = \phi(\boldsymbol{X})$ e $W = \psi(\boldsymbol{Y})$, con $\boldsymbol{X}$ e $\boldsymbol{Y}$ variabili aleatorie (anche multidimensionali) indipendenti, allora Z e W sono indipendenti (si veda il libro di Baldi per una dimostrazione generale)

Una dimostrazione più diretta potrebbe anche procedere nel seguente modo: prima si calcola la distribuzione congiunta di U_1 e U_2 e poi si controlla che la densità congiunta è il prodotto delle densità marginali:

$$\begin{aligned}
P(U_1 = i, U_2 = j) = P\Big(&(\{X_1 = i, X_2 > i\} \cup \{X_1 > i, X_2 = i\} \cup \{X_1 = i, X_2 = i\}) \\
&\cap (\{X_3 = j, X_4 > j\} \cup \{X_3 > j, X_4 = j\} \cup \{X_3 = j, X_4 = j\}) \Big) \\
= P\Big(&(\{X_1 = i, X_2 > i\} \cap \{X_3 = j, X_4 > j\}) \\
&\cup (\{X_1 = i, X_2 > i\} \cap \{X_3 > j, X_4 = j\}) \\
&\cup (\{X_1 = i, X_2 > i\} \cap \{X_3 = j, X_4 = j\}) \\
&\cup (\{X_1 > i, X_2 = i\} \cap \{X_3 = j, X_4 > j\}) \\
&\cup (\{X_1 > i, X_2 = i\} \cap \{X_3 > j, X_4 = j\}) \\
&\cup (\{X_1 > i, X_2 = i\} \cap \{X_3 = j, X_4 = j\}) \\
&\cup (\{X_1 = i, X_2 = i\} \cap \{X_3 = j, X_4 > j\}) \\
&\cup (\{X_1 = i, X_2 = i\} \cap \{X_3 > j, X_4 = j\}) \\
&\cup (\{X_1 = i, X_2 = i\} \cap \{X_3 = j, X_4 = j\}) \Big)
\end{aligned}$$

e quindi, poiché gli eventi di cui si fa l'unione sono disgiunti a due a due, si ha

$$\begin{aligned}
&= P(\{X_1 = i, X_2 > i\} \cap \{X_3 = j, X_4 > j\}) \\
&+ P(\{X_1 = i, X_2 > i\} \cap \{X_3 > j, X_4 = j\}) \\
&+ P(\{X_1 = i, X_2 > i\} \cap \{X_3 = j, X_4 = j\}) \\
&+ P(\{X_1 > i, X_2 = i\} \cap \{X_3 = j, X_4 > j\}) \\
&+ P(\{X_1 > i, X_2 = i\} \cap \{X_3 > j, X_4 = j\}) \\
&+ P(\{X_1 > i, X_2 = i\} \cap \{X_3 = j, X_4 = j\}) \\
&+ P(\{X_1 = i, X_2 = i\} \cap \{X_3 = j, X_4 > j\}) \\
&+ P(\{X_1 = i, X_2 = i\} \cap \{X_3 > j, X_4 = j\}) \\
&+ P(\{X_1 = i, X_2 = i\} \cap \{X_3 = j, X_4 = j\})
\end{aligned}$$

e quindi, tenendo conto dell'indipendenza delle variabili X_k, si ha

$$
\begin{aligned}
&= P\left(\{X_1 = i, X_2 > i\}\right) P\left(\{X_3 = j, X_4 > j\}\right) \\
&+ P\left(\{X_1 = i, X_2 > i\}\right) P\left(\{X_3 > j, X_4 = j\}\right) \\
&+ P\left(\{X_1 = i, X_2 > i\}\right) P\left(\{X_3 = j, X_4 = j\}\right) \\
\\
&+ P\left(\{X_1 > i, X_2 = i\}\right) P\left(\{X_3 = j, X_4 > j\}\right) \\
&+ P\left(\{X_1 > i, X_2 = i\}\right) P\left(\{X_3 > j, X_4 = j\}\right) \\
&+ P\left(\{X_1 > i, X_2 = i\}\right) P\left(\{X_3 = j, X_4 = j\}\right) \\
\\
&+ P\left(\{X_1 = i, X_2 = i\}\right) P\left(\{X_3 = j, X_4 > j\}\right) \\
&+ P\left(\{X_1 = i, X_2 = i\}\right) P\left(\{X_3 > j, X_4 = j\}\right) \\
&+ P\left(\{X_1 = i, X_2 = i\}\right) P\left(\{X_3 = j, X_4 = j\}\right) \\
\\
&= P\left(\{X_1 = i, X_2 > i\}\right) P\left(\{U_2 = j\}\right) \\
&+ P\left(\{X_1 > i, X_2 = i\}\right) P\left(\{U_2 = j\}\right) \\
&+ P\left(\{X_1 = i, X_2 = i\}\right) P\left(\{U_2 = j\}\right) \\
\\
&= P\left(\{U_1 = i\}\right) P\left(\{U_2 = j\}\right)
\end{aligned}
$$