Numeri pseudocasuali

L'idea è di creare una successione di numeri casuali distribuiti uniformemente tra 0 e 1 (la random) per poter fare qualsiasi tipo di simulazione (metodi Monte Carlo, studi sulla fisica delle particelle, simulare un evento come un processo stocastico, per la crittografia). Si parte da un generatore di numeri casuali distribuiti tra 0 e 1. Con delle tecniche matematiche si passa a poter generare qualsiasi variabili casuali, anche se non uniformi.

La prima idea fu quella di generare numeri casuali tramite un fenomeno fisico, scrivere una tabella di questi numeri e immagazzinarli al computer. Quest'idea venne abbandonata per mancanza di memoria e si andò verso delle operazioni che generassero numeri casuali e indipendenti tra di loro. Queste richieste non furono completamente rispettate, sono dei decimali, non reali. Soprattutto non c'è indipendenza, c'è n'è sempre un po'. Le tecniche sviluppate negli hanno lo scopo di ridurre al minimo la correlazione e siccome i numeri vengono generati tramite operazioni a un certo punto si può presentare lo stesso numero iniziale e da quel numero iniziale si ripete la sequenza, quindi si ha un periodo. Si cerca di fare in modo che il periodo sia il più lungo possibile. Da qui il nome di numeri pseudocasuali.

PRIMO ALGORITMO (VON NEUMANN): METODO DEL MEDIO QUADRATO

- SEME INIZIALE ($n° \in \mathbb{N}$ "grande") s_0
- $s_0^2 \to$ si prendono le cifre centrali come s_1
- $s_1^2 \to$ " " " come s_2

Prendiamo un numero naturale, facciamo il quadrato e isoliamo le cifre centrali del quadrato, da usare come numero successivo.

esempio (n° 4 cifre)

$S_0 = 9354$	$S_0^2 = 87\,\underline{4973}\,16$	$X_0 = 9354/10^4$
$S_1 = 4973$	$S_1^2 = 24\,\underline{7307}\,29$	$X_1 = 4973/10^4$
$S_2 = 7307$	$S_2^2 = 53\,\underline{3922}\,49$	$X_2 = 7307/10^4$
$S_3 = 3922$	$S_3^2 = 15\,\underline{3820}\,84$	$X_3 = 3922/10^4$

Ovviamente, con lo stesso seme si ottiene la stessa sequenza.

PRO: ALGORITMO SEMPLICE ANCHE DA IMPLEMENTARE

CONTRO: 1) SE NELLA SEQUENZA RICOMPARE UN NUMERO S_k GIÀ USCITO SI RIPETONO ANCHE $S_{k+1}, S_{k+2} \dots$

2) C'È UNA CORRELAZIONE TRA I NUMERI VICINI ED IN PARTICOLARE SE $S_i \simeq S_k$ ANCHE $S_{i+1} \simeq S_{k+1}$

METODI CONGRUENZIALI

Si considerano 3 numeri fissati

m = modulo

a = incremento

x_0 = valore iniziale detto seme

$$x_i = (a\, x_{i-1}) \bmod m \qquad X_i = \frac{x_i}{m}$$

es. $a = 7^5$, $m = 2^{31} - 1$

GENERATORI DI FIBONACCI

$$x_i = (x_{i-p} \odot x_{i-q}) \bmod m$$

con p e $q < i$

$\odot$ rappresenta un'operazione aritmetica o binaria

N.B. Tutte queste sequenze di numeri sono comunque periodiche. Il generatore sarà "buono" se il periodo sarà molto grande

Esercizi

Q 1 10/02/2020

SCATOLA CONTENENTE

7 monete equilibrate

2 " con $P(T) = \frac{1}{3} \Rightarrow P(C) = \frac{2}{3}$

1 " " $P(T) = \frac{3}{4} \Rightarrow P(C) = \frac{1}{4}$

estrazione di una moneta dalla scatola

L_T = 'uscita di T con un lancio'

$P(L_T)$

$$P(H_1) = \frac{7}{10}$$
$$P(H_2) = \frac{2}{10}$$
$$P(H_3) = \frac{1}{10}$$

$$P(L_T) = P(L_T|H_1)P(H_1) + P(L_T|H_2)P(H_2) + P(L_T|H_3)P(H_3) =$$

$$= \frac{1}{2} \cdot \frac{7}{10} + \frac{1}{3} \cdot \frac{2}{10} + \frac{3}{4} \cdot \frac{1}{10} =$$

$$= \frac{42 + 8 + 9}{120} = \frac{59}{120}$$

Una variabile casuale per i tipi di moneta.

$X_1 =$ 'n° T su 100 lanci se moneta H_1'

$$X_1 \sim B\left(100, \tfrac{1}{2}\right) \stackrel{\cdot}{\sim} N\left(100 \cdot \tfrac{1}{2}, 100 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2}\right) = N(50, 25)$$

$\sim N(0,1)$

$$P(X_1 \geq 48) = P(X_1 \geq 47.5) = P\left(\frac{X_1 - 50}{5} \geq \frac{47.5 - 50}{5}\right) =$$

$$= P\left(Z \geq \frac{-2.5}{5}\right) = P\left(Z \geq -\frac{1}{2}\right) = 1 - F_Z\left(-\frac{1}{2}\right)$$

$X_2 =$ 'n° T su 100 lanci di moneta H_2'

$$X_2 \sim B\left(100, \tfrac{1}{3}\right) \underset{TLC}{\stackrel{\cdot}{\sim}} N\left(\frac{100}{3}, 100 \cdot \frac{1}{3} \cdot \frac{2}{3}\right) = N\left(\frac{100}{3}, \frac{200}{9}\right)$$

$$P(X_2 \geq 48) = P(X_2 \geq 47.5) = P\left(\frac{X_2 - \frac{100}{3}}{\frac{10}{3}\sqrt{2}} \geq \frac{47.5 - \frac{100}{3}}{\frac{10}{3}\sqrt{2}}\right) =$$

$$= P\left(Z \geq \frac{47.5 - \frac{100}{3}}{\frac{10}{3}\sqrt{2}}\right) = 1 - F_Z\left(\frac{47.5 - \frac{100}{3}}{\frac{10}{3}\sqrt{2}}\right)$$

$X_3 =$ 'n° T su 100 lanci se moneta H_3' $\quad X_3 \sim B\left(100, \frac{3}{4}\right) \simeq N\left(\overset{75}{\frac{300}{4}}, \overset{\frac{75}{4}}{\frac{300}{16}}\right)$

$$P(X_3 \geq 48) = P(X_3 \geq 47.5) = P\left(\frac{X_3 - 75}{\sqrt{\frac{75}{4}}} \geq \frac{47.5 - 75}{\sqrt{\frac{75}{4}}}\right) =$$

$$= 1 - F_Z\left(\frac{47.5 - 75}{\sqrt{\frac{75}{4}}}\right)$$

$$P(\text{"almeno 48 T su 100 lanci"}) = P(X \geq 48 | H_1) P(H_1) + P(X \geq 48 | H_2) \times P(H_2) + P(X \geq 48 | H_3) P(H_3) =$$

$$= P(X_1 \geq 48) P(H_1) + P(X_2 \geq 48) P(H_2) + P(X_3 \geq 48) P(H_3)$$

Se T esce esattamente 48 volte su 100 lanci, quale è la prob di H_1?

$$P(H_1 | X = 48) = \frac{P(X = 48 | H_1) P(H_1)}{P(X = 48)}$$

Teorema di Bayes, teorema del limite centrale.

$$P(X_K = 48) = P(47.5 \leq X_K \leq 48.5) \qquad K = 1, 2, 3$$

Q2 10/02/2020

$X =$ 'altezza individuo popolazione' $\sim N(\mu, \sigma^2)$

CAMPIONE DI N INDIVIDUI

- Usando Čebyčev stimare N

$$P(|\overline{X} - \mu| > \sigma) \leq 0.01$$

Čebyčev: Y con $E[Y] = \mu_Y$, $Var(Y) = \sigma_Y^2$

$$P(|Y - \mu| \geq r) \leq \frac{Var(Y)}{r^2}$$

$E[\overline{X}] = \mu$, $Var(\overline{X}) = \frac{\sigma^2}{N}$

$$P(|\overline{X} - E[\overline{X}]| \geq \overset{\sigma}{r}) \leq \frac{Var(\overline{X})}{r^2} = \frac{\sigma^2}{N}\frac{1}{\sigma^2} = \frac{1}{N}$$

$$\frac{1}{N} = \frac{1}{100} \Rightarrow N = 100$$

2) $X \sim N(\mu, \sigma^2) \Rightarrow \overline{X} \sim N\left(\mu, \frac{\sigma^2}{N}\right)$

$$P(|\overline{X} - \mu| \leq \sigma) = 99\%$$

$$P(-\sigma \leq \overline{X} - \mu \leq \sigma) = 99\%$$

$$P\left(-\frac{\sigma}{\frac{\sigma}{\sqrt{N}}} \leq \underbrace{\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{N}}}}_{= Z \sim N(0,1)} \leq \frac{\sigma}{\frac{\sigma}{\sqrt{N}}}\right) = 99\%$$

$$P\left(-\sqrt{N} \leq Z \leq \sqrt{N}\right) = 99\%$$

$$F_Z(\sqrt{N}) - F_Z(-\sqrt{N}) = 99\%$$

$$F_Z(\sqrt{N}) - \left(1 - F_Z(\sqrt{N})\right) = 99\%$$

$$2F_Z(\sqrt{N}) - 1 = 0.99$$

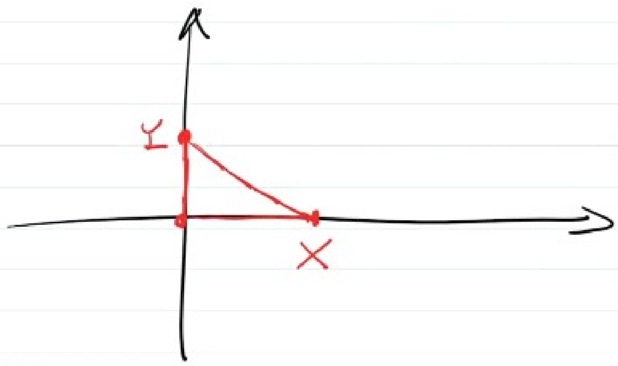
$$2F_Z(\sqrt{N}) - 1 = 0.99$$

$$F_Z(\sqrt{N}) = \frac{1.99}{2} = 0.995$$

Si cerca sulle tavole.

Q3 10/02/2020

$X \sim U(0,3)$ $\quad Y \sim U(0,3)$ INDIPENDENTI



$P(\text{Triangolo isoscele})$

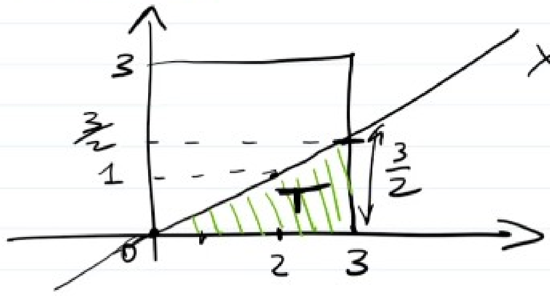
$= P(X=Y) = 0$ dato che X e Y v.c. continue

$P(\text{un cateto più lungo del doppio dell'altro cateto}) =$

$= P(X > 2Y \cup Y > 2X) = P(X > 2Y) + P(Y > 2X) =$

$= 2\,P(X > 2Y)$

$$= 2\,\underline{P(X > 2Y)} = 2\,\frac{A_{T}}{A_{RETTANGOLO}} = 2\,\frac{3 \cdot \frac{3}{2} \cdot \frac{1}{2}}{3 \cdot 3} = \frac{1}{2}$$



$X = 2Y \Rightarrow Y = \frac{X}{2}$

$$\boxed{\begin{array}{l} X \sim U(0,3) \\ E[X] = \frac{3}{2} \\ Var(X) = \frac{9}{12} = \frac{3}{4} \end{array}}$$

A = 'area del triangolo'

$A = XY$

$$E[A] = E[XY] \underset{INDIP}{=} E[X]\,E[Y]$$

$$= \left(\frac{3}{2}\right)^2 = \frac{9}{4}$$

$$\underline{E[X^2] = Var(X) + E^2[X] = \frac{3}{4} + \frac{9}{4} = \frac{12}{4} = 3}$$

$$Var(A) = E[A^2] - E^2[A] = E[X^2Y^2] - \left(\frac{9}{4}\right)^2 =$$

$$= E[X^2]\,E[Y^2] - \frac{81}{16} = 3^2 - \frac{81}{16} =$$

$$= 9\left(1 - \frac{9}{16}\right) = \frac{63}{16}$$

$$F_A(a) = P(A \le a) = P(XY \le a) = \begin{cases} 0 & \text{se } a < 0 \\ & \\ 1 & \text{se } a > 9 \end{cases}$$

Se $0 < a < 9$

$$P(XY \le a) = \iint_{XY \le a} f(x,y)\,dx\,dy = \int_0^3\left(\int_0^{\frac{a}{y}} \frac{1}{3}\cdot\frac{1}{3}\,dx\right)dy =$$

$$= \frac{1}{9}\int_0^3 \left[x\right]_0^{\frac{a}{y}} dy = \frac{1}{9}\int_0^3 \frac{a}{y}\,dy = \frac{a}{9}\left[\ln y\right]_0^3 =$$

In questo caso, non funziona.

Se $0 < a < 9$

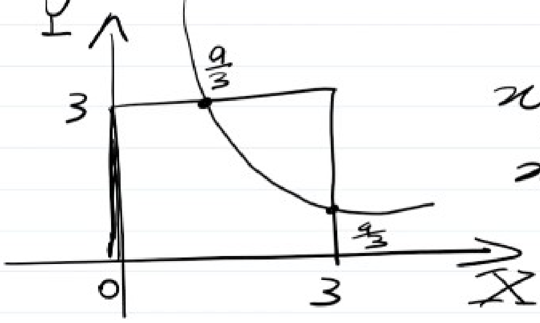
$$P(XY \le a) = \iint_{XY \le a} f(x,y)\,dx\,dy = \int_0^3\left(\int_0^{\min\left(\frac{a}{y},3\right)} \frac{1}{3}\cdot\frac{1}{3}\,dx\right)dy =$$

$$= \frac{1}{9}\int_0^3 \left[x\right]_0^{\min\left(\frac{a}{y},3\right)} dy = \frac{1}{9}\int_0^3 \frac{a}{y}\,dy = \frac{a}{9}\left[\ln y\right]_0^3 =$$

Il disegno ci aiuta.

$XY \le a$

$$= \frac{1}{9}\int_0^3 \left[x\right]_0^{\min(\frac{a}{y},3)} dy \cancel{= \frac{1}{9}\int_0^3 \frac{a}{y}\,dy} = \frac{a}{9}\left[\ln y\right]_0^3 =$$



$$xy \le a \qquad \int_0^{a/3}\left(\int_0^3 f(x,y)\,dx\right)dy +$$

$$x \le \frac{a}{y} \qquad + \int_{a/3}^3\left(\int_0^{a/y} f(x,y)\,dx\right)dy$$

Se $0 \le a \le 9$

$$P(XY \le a) = \int_0^{a/3}\left(\int_0^3 \frac{1}{9}\,dx\right)dy + \int_{a/3}^3\left(\int_0^{a/y} \frac{1}{9}\,dx\right)dy$$

$$= \frac{1}{9}\int_0^{a/3} 3\,dy + \frac{1}{9}\int_{a/3}^3 \left[x\right]_0^{a/y} dy =$$

$$= \frac{3}{9}\,\frac{a}{3} + \frac{1}{9}\int_{a/3}^3 \frac{a}{y}\,dy = \frac{a}{3} + \frac{a}{9}\left[\ln y\right]_{a/3}^3 =$$

$$= \frac{a}{3} + \frac{a}{9}\left[\ln 3 - \ln\left(\frac{a}{3}\right)\right] = \frac{a}{3} + \frac{a}{9}\ln\left(\frac{9}{a}\right)$$

Q4 10/02/2020

$W \sim N(200, (50)^2)$

AMMISSIONE : $W \geq 230$

$$P(W \geq 230) = P\left(\frac{W-200}{50} \geq \frac{230-200}{50}\right) = P\left(Z \geq \frac{3}{5}\right) =$$

$$= 1 - F_Z\left(\frac{3}{5}\right)$$

$$\overline{W} = \sum_{k=1}^{N} \frac{W_k}{N} \sim N\left(200, \frac{(50)^2}{N}\right)$$

$$P(\overline{W} > 280) = 0.01$$

$$P\left(\underbrace{\frac{\overline{W}-200}{50/\sqrt{N}}}_{= Z \sim N(0,1)} > \frac{280-200}{50/\sqrt{N}}\right) = 0.01$$

$$P\left(Z > \frac{8}{5}\sqrt{N}\right) = 0.01$$

$$1 - F_Z\left(\frac{8}{5}\sqrt{N}\right) = 0.01 \Rightarrow F_Z\left(\frac{8}{5}\sqrt{N}\right) = 0.99$$

Q5 10/02/2020

$$f(x) = \begin{cases} Kx & \text{se } x \in [0,1] \\ K & \text{se } x \in]1,3] \\ 0 & \text{altrove} \end{cases} \qquad K, F, E[X]$$

$f(x)$ è funzione di densità di prob se

- $f(x) \geq 0 \quad \forall x \implies K \geq 0$
- $\int_{-\infty}^{+\infty} f(x)\,dx = 1 \implies 1 = \cancel{\int_{-\infty}^{0} 0\,dx} + \int_{0}^{1} Kx\,dx + \int_{1}^{3} K\,dx + \cancel{\int_{3}^{+\infty} 0\,dx}$

$$= \left[K\frac{x^2}{2}\right]_0^1 + \left[Kx\right]_1^3 = \frac{K}{2} + 3K - K =$$

$$= \frac{5}{2}K$$

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx = \int_0^1 x\,Kx\,dx + \int_1^3 x\,K\,dx =$$

$$= K\left[\frac{x^3}{3}\right]_0^1 + K\left[\frac{x^2}{2}\right]_1^3 = \frac{K}{3} + \frac{K}{2}(9-1) =$$

$$= \frac{K}{3} + 4K = \frac{13}{3}K = \frac{26}{15}$$

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx =$$

$$= \begin{cases} 0 & \text{se } a < 0 \\ \int_0^a \frac{2}{5}x\,dx & 0 \leq a \leq 1 \\ \int_0^1 \frac{2}{5}x\,dx + \int_1^a \frac{2}{5}\,dx & 1 < a \leq 3 \\ 1 & \text{se } a > 3 \end{cases}$$

$Y = \sqrt{X}$

$$E[Y] = \int_{-\infty}^{+\infty} \sqrt{x}\, f(x)\, dx = \int_0^1 \sqrt{x}\, \frac{2}{5} x\, dx + \int_1^3 \sqrt{x}\, \frac{2}{5}\, dx =$$

$$= \frac{2}{5}\left[\frac{2}{5} x^{5/2}\right]_0^1 + \frac{2}{5}\left[\frac{2}{3} x^{3/2}\right]_1^3 =$$

$$= \frac{4}{25} + \frac{4}{15}\left(3^{3/2} - 1\right)$$

$$F_Y(y) = P(Y \le y) = P(\sqrt{X} \le y) = P(X \le y^2) =$$

$$= \begin{cases} 0 & \text{se } y < 0 \\ \int_0^{y^2} \frac{2}{5} x\, dx & \text{se } 0 \le y \le 1 \\ \int_0^1 \frac{2}{5} x\, dx + \int_1^{y^2} \frac{2}{5}\, dx & \text{se } 1 < y \le \sqrt{3} \\ 1 & \text{se } y > \sqrt{3} \end{cases}$$

Esercizio per casa

$P(C) = 0.9$

$P(T) = 0.06$

$P(B) = 0.04$

$P(\text{NON VEDERE GLI OCCHIALI DOVE SONO}) = 0.1$

N_C = 'occhiali non trovati nel cassetto'

$$P(T \mid N_C) = \frac{P(N_C \mid T)\, P(T)}{P(N_C)}$$

$$P(N_c) = P(N_c|C)P(C) + P(N_c|T)P(T) + P(N_c|B)P(B) =$$

$$= \frac{1}{10}\cdot\frac{9}{10} + 1\cdot\frac{6}{100} + 1\cdot\frac{4}{100} = \frac{9+6+4}{100} = \frac{19}{100}$$

$$P(T|N_c) = \frac{1\cdot\frac{6}{100}}{\frac{19}{100}} = \frac{6}{19}$$

$$P(C|N_c) = \frac{P(N_c|C)P(C)}{P(N_c)} = \frac{\frac{1}{10}\cdot\frac{9}{10}}{\frac{19}{100}} = \frac{9}{19}$$

6.

$X \sim B(2,\frac{1}{2})$, $Y \sim B(2,\frac{1}{2})$

$P(X=1, Y=1) = 1/3$,

$P(X=0, Y=1) = P(X=0, Y=2) = P(X=1, Y=0) = P(X=1, Y=2) =$

$= P(X=2, Y=2) = p$

$P(x,y)$? $Cov(X,Y)$

$Y \backslash X$	0	1	2	
0				
1				
2				

$$P(X=0)=\binom{2}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^2=\frac{1}{4}; \quad P(X=1)=\binom{2}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^1=\frac{2}{4}$$

$$P(X=2)=\binom{2}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^0=\frac{1}{4}$$

$Y \backslash X$	0	1	2	
0	p	p	p	$\frac{1}{4}$
1	p	$\frac{1}{3}$	p	$\frac{2}{4}$
2	p	p	p	$\frac{1}{4}$
	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

$$3p=\frac{1}{4}$$

$$2p+\frac{1}{3}=\frac{2}{4}$$

$$3p=\frac{1}{4}\Rightarrow p=\frac{1}{12}$$

$$2p+\frac{1}{3}=\frac{2}{4}$$

$$\rightarrow 2p=\frac{2}{4}-\frac{1}{3}=\frac{2}{12}$$

$$8p+\frac{1}{3}=1$$

$Y \backslash X$	0	1	2	
0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
1	$\frac{1}{12}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{4}$
2	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

$$Cov(X,Y) = E[XY] - E[X]E[Y] = 1 - 1 = 0$$

$$E[X] = E[Y] = np = 2 \cdot \frac{1}{2} = 1$$

$$E[XY] = 1 \cdot 1 \cdot \frac{1}{3} + 1 \cdot 2 \cdot \frac{1}{12} + 2 \cdot 1 \cdot \frac{1}{12} + 2 \cdot 2 \cdot \frac{1}{12} =$$

$$= \frac{1}{3} + \frac{4}{12} + \frac{4}{12} = 1$$

$$p(0,0) = \frac{1}{12}$$

$$\neq$$

$$p_X(0) \cdot p_X(0) = \frac{1}{4} \cdot \frac{1}{4}$$

$$\begin{array}{c|c} 1 & \frac{1}{12} \\ 2 & \frac{1}{12} \\ \hline & 1 \end{array}$$

$\Rightarrow X$ e Y SONO DIPENDENTI

Per $Z = X + Y$ $\quad E[Z]$? $\quad Var(Z)$?

$$E[Z] = E[X+Y] = E[X] + E[Y] = 2$$

$$Var(Z) = Var(X) + Var(Y) + 2\,Cov(X,Y)$$

$$Var(Z) = Var(X) + Var(Y) + 2\,Cov(X,Y) =$$

$$= Var(X) + Var(Y) = \frac{1}{2} + \frac{1}{2} = 1$$

$$Var(X) = Var(Y) = npq = 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

p_Z? $\quad Z = X + Y \in \{0, 1, 2, 3, 4\}$

$$p_Z(0) = p(0,0) = \frac{1}{12}$$

$$p_Z(1) = p(0,1) + p(1,0) = \frac{2}{12}$$

$$p_Z(2) = p(0,2) + p(2,0) + p(1,1) = \frac{1}{12} + \frac{1}{12} + \frac{1}{3}$$

$$p_Z(3) = p(1,2) + p(2,1) = \frac{2}{12}$$

$$p_Z(4) = p(2,2) = \frac{1}{12}$$