Sample Point	Number of heads	Number of runs
HHH	3	1
HHT	2	2
HTH	2	3
HTT	1	2
THH	2	2
THT	1	3
TTH	1	2
TTT	0	1

$X \in \{0, 1, 2, 3\}$
$Y \in \{1, 2, 3\}$

Number of heads

Number of runs

$y \backslash x$	0	1	2	3
1	$\frac{1}{8}$	0	0	$\frac{1}{8}$
2	0	$\frac{2}{8}$	$\frac{2}{8}$	0
3	0	$\frac{1}{8}$	$\frac{1}{8}$	0
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Def

Let X and Y be r.v. with joint density f_{xy} and marginal densities f_x, f_y. X and Y are independent if $f_{xy}(x,y) = f_x(x) f_y(y)$ for all x and y.

Ex

$$\left.\begin{array}{l} f_{xy}(3,3) = 0 \\ f_x(3) = \frac{1}{8} \\ f_y(3) = \frac{2}{8} \end{array}\right\} \quad f_{xy}(3,3) \neq f_x(3) f_y(3)$$

Note

If we variate the probability we can make f_{xy} indpendant.

Def

Let X and Y be continuous r.v. The ordered pair (X, Y) is called a two-dimensional continuous r.v. A function f_{xy} such that:

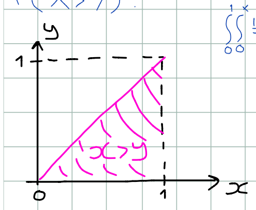
1. $f_{xy}(x,y) \geq 0$, all x, y
2. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1$
3. $P(a \leq X \leq b \text{ and } c \leq Y \leq d) = \int_c^d \int_a^b f(x,y)\,dx\,dy$

Ex

$f(x,y) = \frac{12}{7}(x^2 + xy)$, $0 \leq x, y \leq 1$

$P(X > Y)$?

$$\int_0^1 \int_0^x \frac{12}{7}(x^2 + xy)\,dy\,dx = \frac{9}{14}$$



Def

Let (X, Y) be a two-dimensional r.v with joint density f_{xy}. H is a function then

$$E[H(X,Y)] = \begin{cases} \sum_{\text{all } x} \sum_{\text{all } y} H(x,y) f(x,y) & \text{discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y) f(x,y) & \text{continuous} \end{cases}$$

Remark

1. $H(x,y) = x \cdot y$
2. If X, Y are independent, $E[X \cdot Y] = E[X] \cdot E[Y]$

Def: Covariance

Let X and Y be r.v with means $E[X]$ and $E[Y]$. The covariance between X and Y, denoted $Cov(X,Y)$ or σ_{XY}, is given by:

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

Proposition

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

Def

Let X and Y be r.v with expectations $E[X]$ and $E[Y]$, and variances $Var[X]$ and $Var[Y]$. The correlation, ρ_{XY}, between X and Y is given by:

$$-1 \leq \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}} \leq 1$$

Remark

If $X = Y$ $\rho_{XY} = \frac{Cov(X,X)}{\sqrt{Var[X]\,Var[X]}} = \frac{Var[X]}{Var[X]} = 1$

If $X = -Y$ $\rho_{XY} = \frac{Cov(X,-X)}{\sqrt{Var[X]\,Var[X]}} = \frac{-Var[X]}{Var[X]} = -1$

Note

Correlation measures only linear depandence.

Prelims

1. Systems which changes states in discrete time
2. The collection of all possible states $I = \{i_0, \dots, i_n\}$ is called state space.
3. An initial distribution $\lambda = \{\lambda_{i_0}, \dots, \lambda_{i_n}\}$ defined on I, specifies the starting state.
4. The random mechanism is described by transition matrix P.
   Entry P_{ij} gives the probability that the system will change state from i to j in a unit of time.
5. Each entry in P is non-negative and not greater than one. The sum of entries on each row equals to one.
   A matrix with these properties is called stochastic.

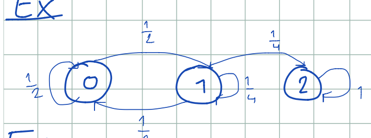
Def

A sequence of random variables X_n with values in a finite set I is a discrete-time Markov chain (DTMC) with initial distribution λ and transition matrix P if for all $i_0, \dots, i_n \in I$, the joint probability $P(X_0 = i_0, \dots, X_n = i_n) = \lambda_{i_0} \cdot P_{i_0 i_1} \cdot \dots \cdot P_{i_{n-1} i_n}$

Def

A state i of MC is called absorbing if it is impossible to leave it. ($P_{ii} = 1$)
A MC is called absorbing if it has atleast one absorbing state and it is possible to reach that state from all other states. It is not necessary to be able to go directly there.

Ex



Theorem

1. $P(X_n = j) = (\lambda P^n)_j$ ← jth element, P^n – n:th power of matrix P
2. $P_{ij}^{(n)} = P(X_{k+n} = j \mid X_k = i)$ n-step transition probability

Ex

$$P(X_{n+1} = j \mid X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = \frac{P(X_0 = i_0, \dots, X_n = i, X_{n+1} = j)}{P(X_0 = i_0, \dots, X_n = i)} = \frac{\lambda_{i_0} P_{i_0 i_1} \cdot \dots \cdot P_{ij}}{\lambda_{i_0} P_{i_0 i_1} \cdot \dots \cdot P_{i_{n-1} i}} = P_{ij}$$