Independent Events

Two events E and F are *independent* if

$$P(EF) = P(E)\,P(F)\,.$$

In this case

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{P(E)\,P(F)}{P(F)} = P(E)\,,$$

(assuming $P(F)$ is not zero).

Thus

knowing F *occurred doesn't change the probability of* E *.*

EXAMPLE : If E and F are *independent* then so are E and F^c .

PROOF : $E = E(F \cup F^c) = EF \cup EF^c$, where

$$EF \quad \text{and} \quad EF^c \quad \text{are } \textit{disjoint} .$$

Thus

$$P(E) = P(EF) + P(EF^c) ,$$

from which

$$\begin{aligned} P(EF^c) &= P(E) - P(EF) \\ &= P(E) - P(E) \cdot P(F) \quad \text{(since } E \text{ and } F \text{ independent)} \\ &= P(E) \cdot (1 - P(F)) \\ &= P(E) \cdot P(F^c) . \end{aligned}$$

EXAMPLE : Draw *one* card from a deck of 52 playing cards.

Counting outcomes we find

$$P(\text{Face Card}) = \tfrac{12}{52} = \tfrac{3}{13} ,$$

$$P(\text{Hearts}) = \tfrac{13}{52} = \tfrac{1}{4} ,$$

$$P(\text{Face Card and Hearts}) = \tfrac{3}{52} ,$$

$$P(\text{Face Card}|\text{Hearts}) = \tfrac{3}{13} .$$

We see that

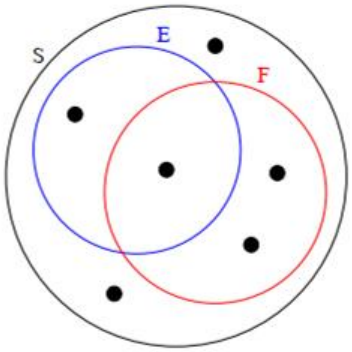
$$P(\text{Face Card and Hearts}) = P(\text{Face Card}) \cdot P(\text{Hearts}) \quad (= \frac{3}{52}) .$$

Thus the events "*Face Card*" and "*Hearts*" are *independent*.

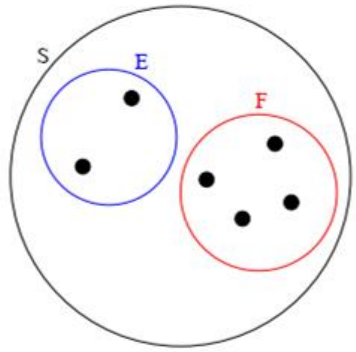
Therefore we also have

$$P(\text{Face Card}|\text{Hearts}) = P(\text{Face Card}) \quad (= \frac{3}{13}) .$$

NOTE : *Independence* and *disjointness are different things* !



Independent, but not disjoint.



Disjoint, but not independent.

(The six outcomes in S are assumed to have equal probability.)

If E and F are *independent* then $P(EF) = P(E)\ P(F)$.

If E and F are *disjoint* then $P(EF) = P(\ \emptyset\ ) = 0$.

If E and F are *independent and disjoint* then one has *zero probability* !

Three events E , F, and G are *independent* if

$$P(EFG) \ = \ P(E)\ P(F)\ P(G)\ .$$

and

$$\begin{aligned} P(EF) &= P(E)\ P(F)\ . \\ P(EG) &= P(E)\ P(G)\ . \\ P(FG) &= P(F)\ P(G)\ . \end{aligned}$$

DISCRETE RANDOM VARIABLES

DEFINITION : A *discrete random variable* is a *function* $X(s)$ from a *finite* or *countably infinite* sample space $\mathcal{S}$ to the real numbers :

$$X(\cdot) \quad : \quad \mathcal{S} \quad \rightarrow \quad \mathbb{R} \,.$$

EXAMPLE : Toss a coin 3 times in sequence. The sample space is

$$\mathcal{S} \ = \ \{HHH\,,\,HHT\,,\,HTH\,,\,HTT\,,\,THH\,,\,THT\,,\,TTH\,,\,TTT\}\,,$$

and examples of random variables are

- $X(s) \ =$ the number of Heads in the sequence ; *e.g.*, $X(HTH) \ = \ 2$,

- $Y(s) \ =$ The index of the first H ; *e.g.*, $Y(TTH) \ = \ 3$,

  0 if the sequence has no H , *i.e.*, $Y(TTT) \ = \ 0$.

NOTE : In this example $X(s)$ and $Y(s)$ are actually *integers* .

Value-ranges of a random variable correspond to *events* in $\mathcal{S}$.

EXAMPLE : For the sample space

$\mathcal{S} \ = \ \{HHH\,,\,HHT\,,\,HTH\,,\,HTT\,,\,THH\,,\,THT\,,\,TTH\,,\,TTT\}\,,$

with

$$X(s) \ = \ \text{the number of Heads}\ ,$$

the value

$X(s) \ = \ 2\ ,$ corresponds to the event $\{HHT\ ,\ HTH\ ,\ THH\}\ ,$

and the values

$1 < X(s) \leq 3\ ,$ correspond to $\{HHH\ ,\ HHT\ ,\ HTH\ ,\ THH\}\ .$

NOTATION : If it is clear what $\mathcal{S}$ is then we often just write

$$X \qquad \text{instead of} \qquad X(s)\ .$$

Value-ranges of a random variable correspond to *events* in $\mathcal{S}$,

and

events in $\mathcal{S}$ have a *probability* .

Thus

Value-ranges of a random variable have a *probability* .

EXAMPLE : For the sample space

$\mathcal{S} \; = \; \{HHH \, , \, HHT \, , \, HTH \, , \, HTT \, , \, THH \, , \, THT \, , \, TTH \, , \, TTT\} \, ,$

with $X(s) \; = \;$ the number of Heads ,

we have

$$P(0 < X \leq 2) \;\; = \;\; \frac{6}{8} \, .$$

QUESTION : What are the values of

$P(X \leq -1) \, , \, P(X \leq 0) \, , \, P(X \leq 1) \, , \, P(X \leq 2) \, , \, P(X \leq 3) \, , \, P(X \leq 4) \, ?$

NOTATION : We will also write $p_X(x)$ to denote $P(X = x)$.

EXAMPLE : For the sample space

$\mathcal{S} \; = \; \{HHH \, , \, HHT \, , \, HTH \, , \, HTT \, , \, THH \, , \, THT \, , \, TTH \, , \, TTT\} \, ,$

with

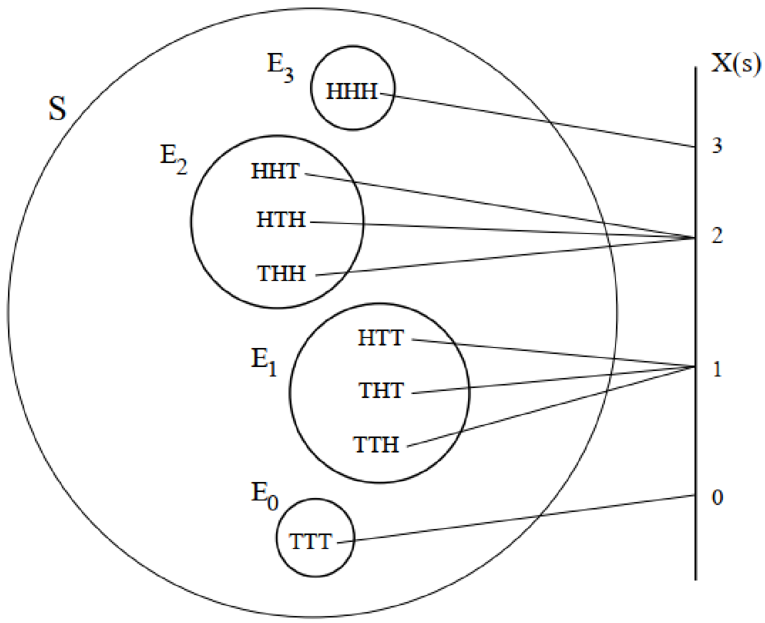
$X(s) \; = \;$ the number of Heads ,

we have

$$\begin{aligned}
p_X(0) \;&\equiv\; P(\,\{TTT\}\,) &&= \tfrac{1}{8} \\
p_X(1) \;&\equiv\; P(\,\{HTT \, , \, THT \, , \, TTH\}\,) &&= \tfrac{3}{8} \\
p_X(2) \;&\equiv\; P(\,\{HHT \, , \, HTH \, , \, THH\}\,) &&= \tfrac{3}{8} \\
p_X(3) \;&\equiv\; P(\,\{HHH\}\,) &&= \tfrac{1}{8}
\end{aligned}$$

where

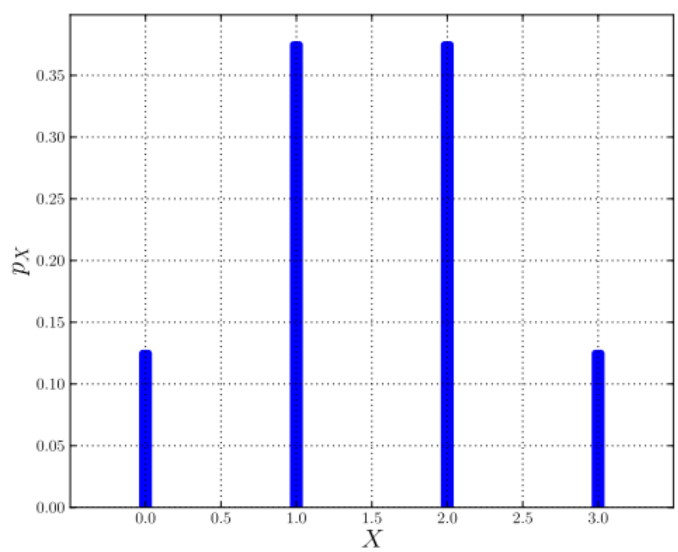
$$p_X(0) \; + \; p_X(1) \; + \; p_X(2) \; + \; p_X(3) \;\; = \;\; 1 \, . \qquad (\text{ Why ? })$$



Graphical representation of X .

The *events* E_0, E_1, E_2, E_3 are *disjoint* since $X(s)$ is a *function* !

($X : S \rightarrow \mathbb{R}$ must be defined for *all* $s \in S$ and must be *single-valued*.)



The graph of p_X .

DEFINITION :

$$p_X(x) \equiv P(X = x) ,$$

is called the *probability mass function* .

DEFINITION :

$$F_X(x) \equiv P(X \le x) ,$$

is called the (*cumulative*) *probability distribution function* .

PROPERTIES :

- $F_X(x)$ is a *non-decreasing* function of x . (Why ?)
- $F_X(-\infty) = 0$ and $F_X(\infty) = 1$. (Why ?)
- $P(a < X \le b) = F_X(b) - F_X(a)$. (Why ?)

NOTATION : When it is clear what X is then we also write

$$p(x) \text{ for } p_X(x) \quad \text{and} \quad F(x) \text{ for } F_X(x) .$$

EXAMPLE : With $X(s) =$ the number of Heads , and

$\mathcal{S} = \{HHH\,,\,HHT\,,\,HTH\,,\,HTT\,,\,THH\,,\,THT\,,\,TTH\,,\,TTT\}$,

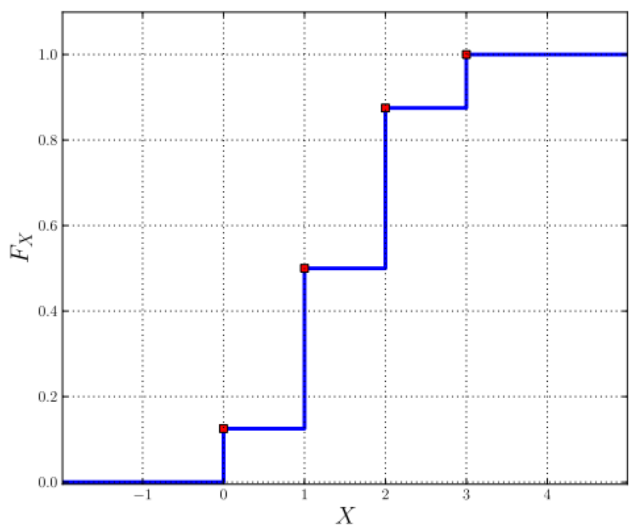
$$p(0) = \tfrac{1}{8} \quad , \quad p(1) = \tfrac{3}{8} \quad , \quad p(2) = \tfrac{3}{8} \quad , \quad p(3) = \tfrac{1}{8} \ ,$$

we have the *probability distribution function*

$$\begin{array}{lclcl} F(-1) & \equiv & P(X \le -1) & = & 0 \\ F(\ 0) & \equiv & P(X \le \ 0) & = & \tfrac{1}{8} \\ F(\ 1) & \equiv & P(X \le \ 1) & = & \tfrac{4}{8} \\ F(\ 2) & \equiv & P(X \le \ 2) & = & \tfrac{7}{8} \\ F(\ 3) & \equiv & P(X \le \ 3) & = & 1 \\ F(\ 4) & \equiv & P(X \le \ 4) & = & 1 \end{array}$$

We see, for example, that

$$\begin{aligned} P(0 < X \le 2) &= P(X = 1) + P(X = 2) \\ &= F(2) - F(0) = \tfrac{7}{8} - \tfrac{1}{8} = \tfrac{6}{8} \ . \end{aligned}$$



The graph of the *probability distribution function* F_X .

$X(s)$ is the *number of tosses* until "Heads" occurs $\cdots$

REMARK : We can also take $\mathcal{S} \equiv \mathcal{S}_n$ as *all ordered outcomes of length* n. For example, for $n = 4$,

$$\begin{aligned}
\mathcal{S}_4 \quad = \quad \{ \quad & \tilde{H}HHH \, , \quad \tilde{H}HHT \, , \quad \tilde{H}HTH \, , \quad \tilde{H}HTT \, , \\
& \tilde{H}THH \, , \quad \tilde{H}THT \, , \quad \tilde{H}TTH \, , \quad \tilde{H}TTT \, , \\
& T\tilde{H}HH \, , \quad T\tilde{H}HT \, , \quad T\tilde{H}TH \, , \quad T\tilde{H}TT \, , \\
& TT\tilde{H}H \, , \quad TT\tilde{H}T \, , \quad TTT\tilde{H} \, , \quad TTTT \quad \} \, .
\end{aligned}$$

where for each outcome the first "Heads" is marked as $\tilde{H}$.

Each outcome in $\mathcal{S}_4$ has *equal probability* 2^{-n} (here $2^{-4} = \frac{1}{16}$) , and

$$p_X(1) = \tfrac{1}{2} \quad , \quad p_X(2) = \tfrac{1}{4} \quad , \quad p_X(3) = \tfrac{1}{8} \quad , \quad p_X(4) = \tfrac{1}{16} \quad \cdots ,$$

independent of n .

Joint distributions

The *probability mass function* and the *probability distribution function* can also be functions of *more than one variable*.

EXAMPLE : Toss a coin 3 times in sequence. For the sample space

$$\mathcal{S} \ = \ \{HHH \, , \, HHT \, , \, HTH \, , \, HTT \, , \, THH \, , \, THT \, , \, TTH \, , \, TTT\} \, ,$$

we let

$X(s) \ = \ \#$ Heads , $Y(s) \ = \$ index of the first H (0 for TTT) .

Then we have the *joint probability mass function*

$$p_{X,Y}(x,y) \quad = \quad P(X = x \, , \, Y = y) \, .$$

For example,

$$\begin{aligned}
p_{X,Y}(2,1) \quad &= \quad P(X = 2 \, , \, Y = 1) \\
&= \quad P(\text{ 2 Heads } \, , \, 1^{\text{st}} \text{ toss is Heads}) \\
&= \quad \tfrac{2}{8} \ = \ \tfrac{1}{4} \, .
\end{aligned}$$

EXAMPLE : (continued $\cdots$) For

$\mathcal{S} = \{HHH\,,\, HHT\,,\, HTH\,,\, HTT\,,\, THH\,,\, THT\,,\, TTH\,,\, TTT\}\,,$

$X(s)$ = number of Heads, and $Y(s)$ = index of the first H ,

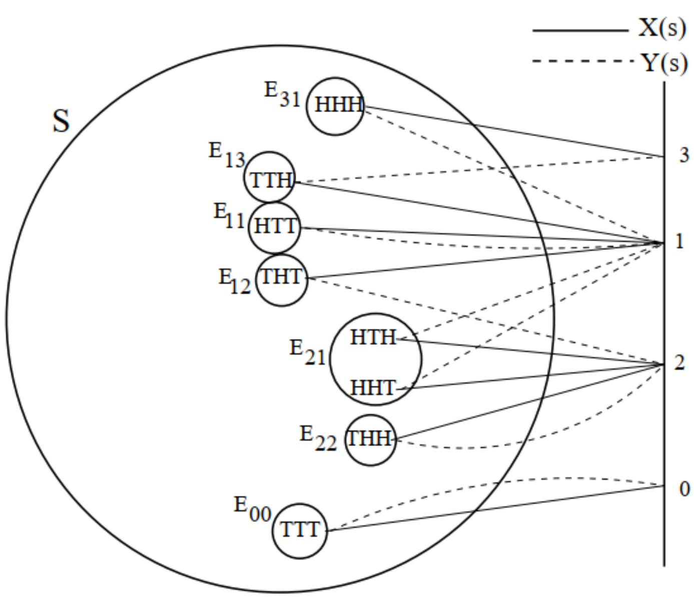
we can list the values of $p_{X,Y}(x, y)$:

Joint probability mass function $p_{X,Y}(x, y)$

	$y = 0$	$\mathbf{y = 1}$	$y = 2$	$y = 3$	$p_X(x)$
$x = 0$	$\frac{1}{8}$	0	0	0	$\frac{1}{8}$
$x = 1$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$\mathbf{x = 2}$	0	$\mathbf{\frac{2}{8}}$	$\frac{1}{8}$	0	$\frac{3}{8}$
$x = 3$	0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
$p_Y(y)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	1

NOTE :

- The *marginal probability* p_X is the probability mass function of X.
- The *marginal probability* p_Y is the probability mass function of Y.



The *events* $E_{i,j} \equiv \{\ s \in S\ :\ X(s) = i\ ,\ Y(s) = j\ \}$ are *disjoint* .

QUESTION : Are the events $X = 2$ and $Y = 1$ *independent* ?

DEFINITION :

$$p_{X,Y}(x,y) \quad \equiv \quad P(X = x \ , \ Y = y) \ ,$$

is called the *joint probability mass function* .

DEFINITION :

$$F_{X,Y}(x,y) \quad \equiv \quad P(X \leq x \ , \ Y \leq y) \ ,$$

is called the *joint* (*cumulative*) *probability distribution function* .

NOTATION : When it is clear what X and Y are then we also write

$$p(x,y) \quad \text{for} \quad p_{X,Y}(x,y) \ ,$$

and

$$F(x,y) \quad \text{for} \quad F_{X,Y}(x,y) \ .$$

EXAMPLE : *Three* tosses : $X(s) = \#$ Heads, $Y(s) =$ index 1^{st} H .

Joint probability mass function $p_{X,Y}(x,y)$

	$y=0$	$y=1$	$y=2$	$y=3$	$p_X(x)$
$x=0$	$\frac{1}{8}$	0	0	0	$\frac{1}{8}$
$x=1$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$x=2$	0	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{3}{8}$
$x=3$	0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
$p_Y(y)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	1

Joint distribution function $F_{X,Y}(x,y) \equiv P(X \leq x, Y \leq y)$

	$y=0$	$y=1$	$y=2$	$y=3$	$F_X(\cdot)$
$x=0$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$x=1$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{3}{8}$	$\frac{4}{8}$	$\frac{4}{8}$
$x=2$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{6}{8}$	$\frac{7}{8}$	$\frac{7}{8}$
$x=3$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{7}{8}$	1	1
$F_Y(\cdot)$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{7}{8}$	1	1

Note that the distribution function F_X is a *copy* of the 4th column, and the distribution function F_Y is a *copy* of the 4th row. (Why ?)

In the preceding example :

Joint probability mass function $p_{X,Y}(x,y)$

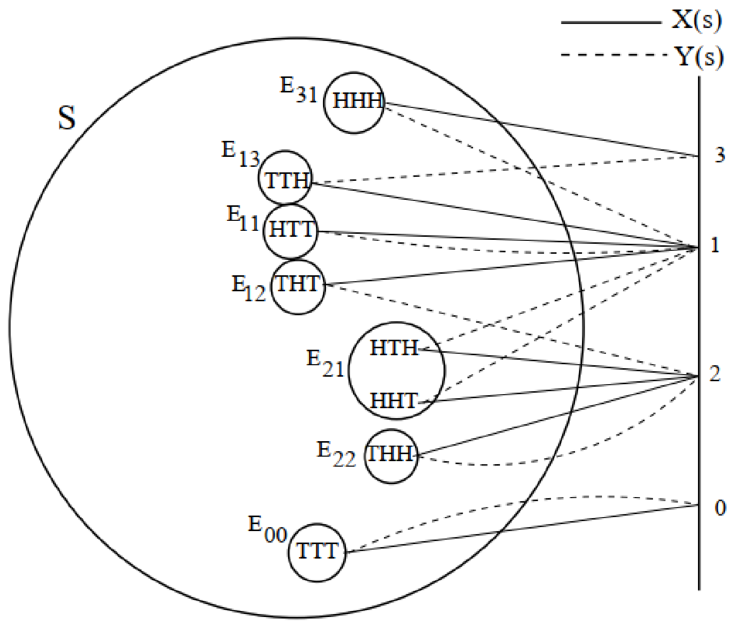
	$y=0$	$y=1$	$y=2$	$y=3$	$p_X(x)$
$x=0$	$\frac{1}{8}$	0	0	0	$\frac{1}{8}$
$x=1$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$x=2$	0	$\frac{2}{8}$	$\mathbf{\frac{1}{8}}$	$\mathbf{0}$	$\frac{3}{8}$
$x=3$	0	$\frac{1}{8}$	$\mathbf{0}$	$\mathbf{0}$	$\frac{1}{8}$
$p_Y(y)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	1

Joint distribution function $F_{X,Y}(x,y) \equiv P(X \leq x, Y \leq y)$

	$y=0$	$y=1$	$y=2$	$y=3$	$F_X(\cdot)$
$x=0$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$x=1$	$\frac{1}{8}$	$\mathbf{\frac{2}{8}}$	$\frac{3}{8}$	$\mathbf{\frac{4}{8}}$	$\frac{4}{8}$
$x=2$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{6}{8}$	$\frac{7}{8}$	$\frac{7}{8}$
$x=3$	$\frac{1}{8}$	$\mathbf{\frac{5}{8}}$	$\frac{7}{8}$	$\mathbf{1}$	1
$F_Y(\cdot)$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{7}{8}$	1	1

QUESTION : Why is

$$P(1 < X \leq 3\,, 1 < Y \leq 3) \;=\; F(3,3) - F(1,3) - F(3,1) + F(1,1)\ ?$$



Three tosses : $X(s) = \#$ Heads, $Y(s) =$ index 1^{st} H .

- What are the values of $p_X(2)$, $p_Y(1)$, $p_{X,Y}(2,1)$?
- Are X and Y *independent* ?