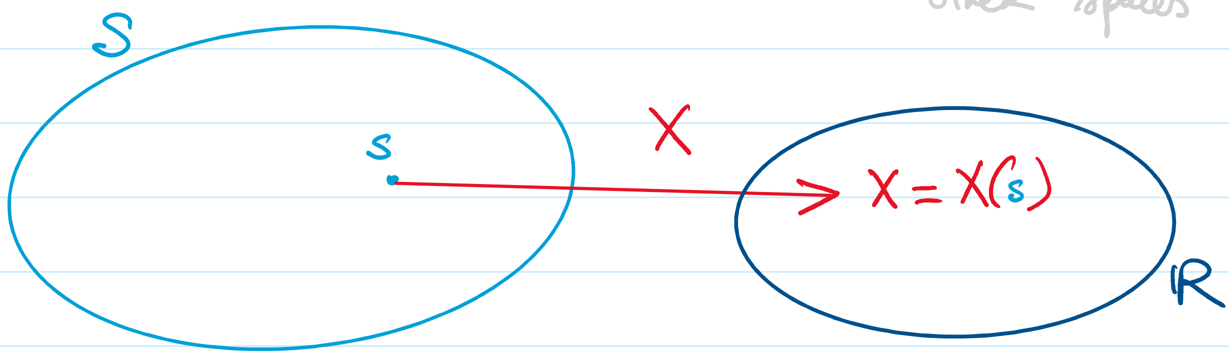
4.1 Joint and marginal distributions

Recall

Def. A random variable is a function from a sample space into the real numbers

may change to other spaces



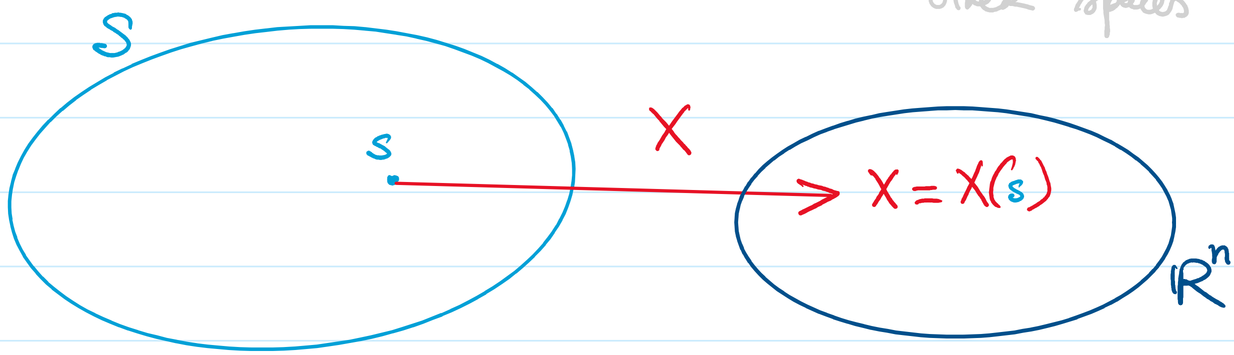
ω may be unobserved

X observed

Now, we define random variable taking values in multi-dimensional spaces

$\boxed{\text{Def.}}$ An n-dimensional Random vector is a function from a sample space into space $\mathbb{R}^n$.

may change to other spaces



ω may be unobserved

X observed

Example:

- $X = (X_1, X_2, \dots, X_n) \in \mathbb{R}^n$ associated with measurements of a (Random) person

  where X_1 = temperature

  X_2 = height

  X_3 = blood pressure

  and so on ...

- if $n = 2$, $X = (X_1, X_2) \in \mathbb{R}^2$ is called a bivariate vector.

Example: Consider bivariate R.V. $(X, Y) \in \mathbb{R}^2$ which represent the outcome of the experiment of tossing two fair dice

$$\begin{cases} X = \text{sum of the two dice} \\ Y = \text{absolute difference of the two dice} \end{cases}$$

Write sample point $s \in S$ as $s = (s_1, s_2)$

then $\begin{cases} X = s_1 + s_2 \\ Y = |s_1 - s_2| \end{cases}$ $\quad s_1, s_2 \in \{1, \dots, 6\}$

Now, we may define probability of events defined in terms of X and Y

$$P((X,Y) = (5,3)) := P(\{(s_1, s_2) \mid s_1 + s_2 = 5, |s_1 - s_2| = 3\})$$

$$= P(\{(4,1), (1,4)\})$$

$$= P(\{4,1\}) + P(\{1,4\})$$

assuming independence $\quad = P(\{4\})P(\{1\}) + P(\{1\})P(\{4\})$

fair dice $\quad = (1/6)(1/6) + (1/6)(1/6) = 1/18.$

Def. Let (X,Y) be a discrete bivariate vector. Then the function from $\mathbb{R}^2$ to $\mathbb{R}$:

$$f(x,y) := P(X=x, Y=y)$$

is called **the joint probability mass function** (pmf) of (X,Y)

Also use notations $f_{XY}(x,y)$ OR $f_{X,Y}(x,y)$.

Now, for subset $A \subset \mathbb{R}^2$ we can find

$$\boxed{P((X,Y)\in A) = \sum_{(x,y)\in A} f(x,y).}$$

Expectation

Let $g(x,y)$ be a function from $\mathbb{R}^2 \to \mathbb{R}$ then $g(X,Y)$ is a real-valued Random variable

$$\boxed{\mathbb{E}\, g(X,Y) := \sum_{(x,y)\in\mathbb{R}^2} g(x,y)\, f_{XY}(x,y).}$$

Remark.

- for $A \subset \mathbb{R}$:

$$P(g(X,Y) \in A) = \sum_{(x,y) \in A} f_{XY}(x,y).$$

Linearity of expectation

if g_1, g_2 are real-valued functions on $\mathbb{R}^2$; $a, b \in \mathbb{R}$ then

$$\boxed{\mathbb{E}\left(a\, g_1(X,Y) + b\, g_2(X,Y)\right) = a\, \mathbb{E}\, g_1(X,Y) + b\, \mathbb{E}\, g_2(X,Y)}$$

Example: Recall bivariate vector (X,Y)

$$\begin{cases} X = \text{sum of the two dice} \\ Y = |\text{difference of the two dice}| \end{cases}$$

Let $f_{XY}(x,y)$ be the joint pmf for (X,Y)

- $P(X = Y) = ?$

$$P(X=Y) = P((X,Y) \in (\{1,1\}, \{2,2\}, \dots, \{6,6\})$$
$$= \sum_{i=1}^{6} f_{XY}(i,i) .$$

- $P(X=2) = ?$

$$P(X=2) = P((X,Y) \in \{(2,1), (2,2), \dots, (2,5), (2,6)\})$$
$$= \sum_{i=1}^{6} f_{XY}(2,i) .$$

in fact $\forall x \in \{1, \dots, 6\}$

$$P(X=x) = \sum_{y=1}^{6} f_{XY}(x,y) .$$

- Likewise, the distribution of Y is also completely determined

$$P(Y=y) = \sum_{x=1}^{6} f_{XY}(x,y).$$

Theorem Let (X,Y) be a discrete bivariate Random vector with joint pmf $f_{XY}(x,y)$

Then

X and Y are discrete Random variable with the following pmf's:

$$\begin{cases} f_X(x) = \sum_y f_{XY}(x,y) \\ f_Y(y) = \sum_x f_{XY}(x,y). \end{cases}$$

Remarks

- f_X and f_Y are called the marginal pmfs of X and Y.

- The distribution of X and distribution of Y are called marginal distributions (WRT the distribution of Random vector (X,Y)),

- The distribution of Random vector (X,Y) is also referred to as the joint distribution of X and Y.

- joint distribution / pmf completely determines its marginal distributions / pmf's

but marginal distributions / pmf's do NOT determine the joint distribution.

Example:

Let $X \sim \text{Bernoulli}(1/2)$

$Y \sim \text{Bernoulli}(1/2)$, $Y \perp\!\!\!\perp X$.

$Z = 1 - X$.

Then $X \stackrel{d}{=} Y \stackrel{d}{=} Z$

But $(X, Y) \neq (X, Z)$ in joint distribution!

Continuous bivariate Random vectors

are decsribed via joint probability density functions

Def. A function $f(x,y)$ from $\mathbb{R}^2$ to $\mathbb{R}$ is called a joint pdf of the continuous bivariate random vector (X,Y) if, for every $A \subseteq \mathbb{R}^2$

$$\boxed{P((X,Y) \in A) = \iint_{(x,y)\in A} f(x,y)\, dx\, dy.}$$

Remarks

- $f(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$
- $\iint_{\mathbb{R}^2} f(x,y)\, dx\, dy = 1$

Expectation

if $g(x,y)$ is a real-valued function then $g(X,Y)$ is a Random variable with expectation:

$$\mathbb{E}\, g(X,Y) := \iint_{\mathbb{R}^2} g(x,y)\, f(x,y)\, dx\, dy.$$

$f_{XY}(x,y)$

Marginal pdf's:

$$f_X(x) := \int_{\mathbb{R}} f(x,y)\, dy$$

$$f_Y(y) := \int_{\mathbb{R}} f(x,y)\, dx.$$

joint cdf.

$$F_{XY}(x,y) := P(X \leq x,\ Y \leq y)$$

$$= \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(s,t)\, dt\, ds$$

Remarks.

- $0 \leq F_{XY}(x,y) \leq 1 \qquad \forall (x,y) \in \mathbb{R}^2.$

- $F_{XY}(x,.) \uparrow \qquad \forall x \in \mathbb{R}$
  $F_{XY}(.,y) \uparrow \qquad \forall y \in \mathbb{R}.$

- Fundamental theorem of calculus (bivariate case) gives

$$\frac{d^2 F_{XY}(x,y)}{dx\,dy} = f_{XY}(x,y)$$

at continuous points (x,y) of function f_{XY}.