ENGSCI 213: MATHEMATICAL MODELLING 2SE

Oliver Maclaren
oliver.maclaren@auckland.ac.nz

NEW (& LAST) PROBABILITY TOPIC

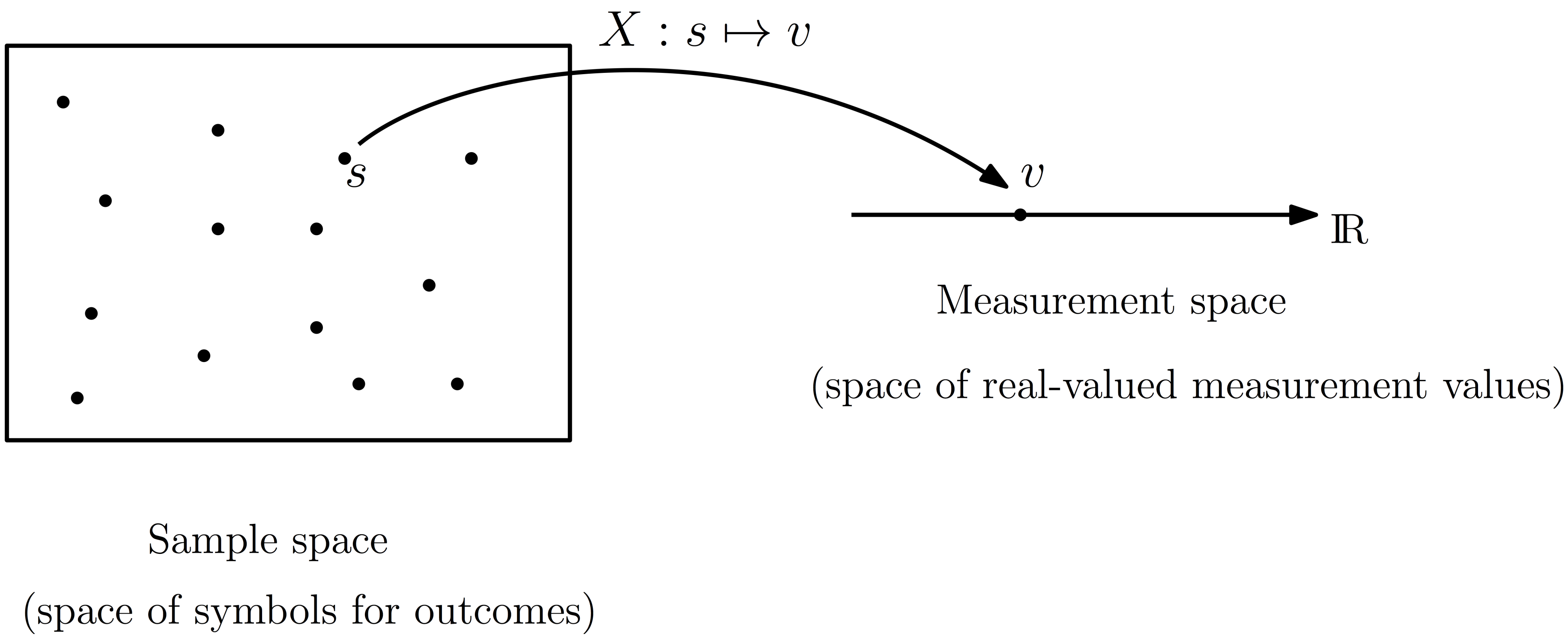
3. *Continuous probability models* [3 lectures]

   Continuous random variables and continuous probability distributions. Probability density functions and cumulative distribution functions. (Expectation and variance). Uniform, Exponential and Normal distributions.

LECTURE 7

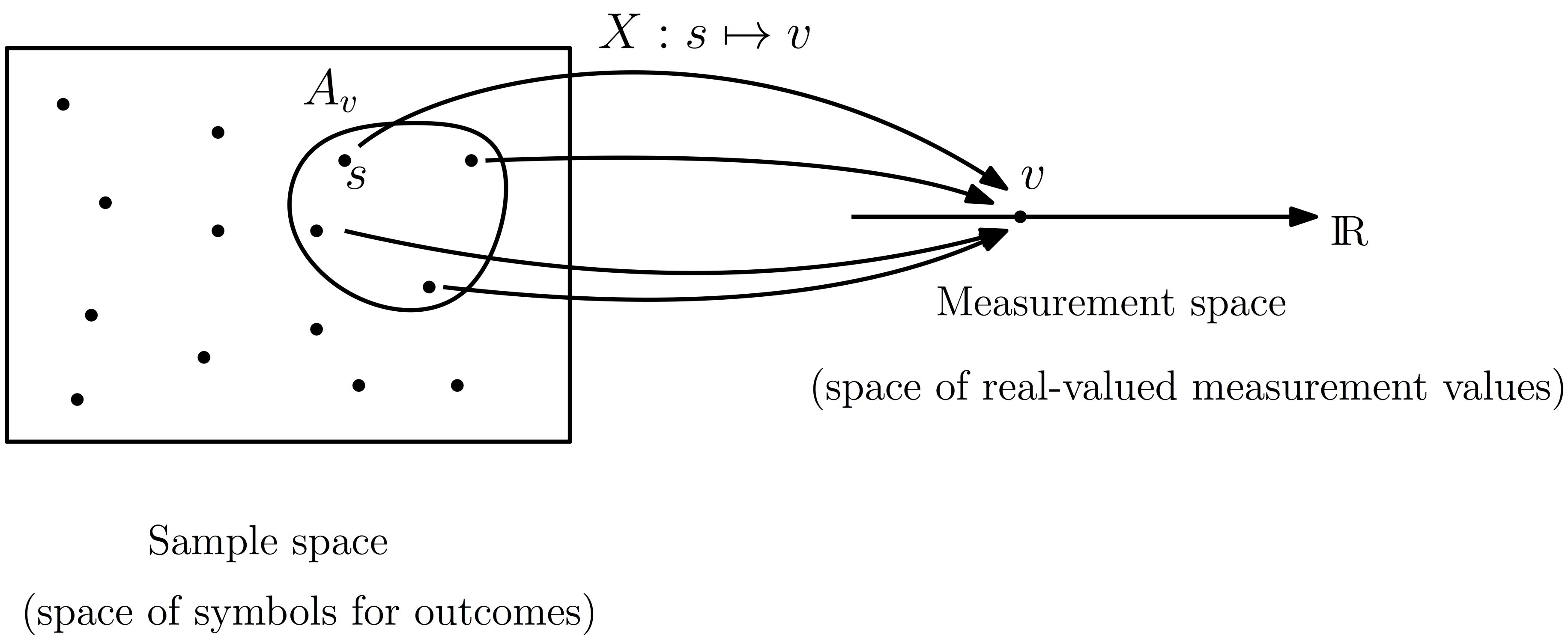
- Recap of continuous random variables and need for probability of intervals
- Cumulative distribution function
- Density function
- Expectation and variance

BASIC PICTURE: RANDOM VARIABLES AS FUNCTIONS FOR LABELLING *OUTCOMES*



BASIC PICTURE: RANDOM VARIABLES AS FUNCTIONS FOR LABELLING *EVENTS*

$$\boxed{A_v = \{s | X(s) = v\} = X^{-1}(\{v\})}$$



DEFINITIONS

Random variables: functions from the sample space to real numbers. Values assigned can be any real number, regardless of whether the random variable is called 'discrete' or 'continuous'.

DEFINITIONS

Discrete random variables: the sample space is a discrete set

Continuous random variables: the sample space is a continuous set

PROBABILITY FOR RANDOM VARIABLES

The probability of a random variable taking on a value is just the *probability of the associated event*

PROBABILITY FOR RANDOM VARIABLES: CONTINUOUS CASE

For *continuous* random variables the *'elementary events'* defined by $X(s) = x$ have *'zero' (or 'infinitesimal') probability attached*.

Instead we define the *probability of* the random variable *taking values in finite intervals* e.g. $X(s) \in [a, b]$ OR define a *density* to integrate over to get a finite probability.

PROBABILITY FOR RANDOM VARIABLES: CONTINUOUS CASE

So for *continuous* RVs we use

$$P(a \leq X \leq b) = P(A_{[a,b]}) = P(\{s \mid X(s) \in [a,b]\})$$

$$= P(X^{-1}[a,b]) = \int_a^b f_X(x)dx$$

where $f_X(x)$ is called the *probability 'density' function*. We'll come back to this in a second.

THE CUMULATIVE DISTRIBUTION FUNCTION

Because the *probability mass function at a point is ill-defined for continuous distributions*, and because the density function is introduced somewhat indirectly, it can make sense to take the *cumulative distribution function* $F_X(x)$ as the *basic starting point*.

This is because we will want to talk about probabilities 'added up' over *intervals*.

THE CUMULATIVE DISTRIBUTION FUNCTION

First we *define the cumulative distribution function* $F_X(x)$ for continuous variables by

$$\boxed{F_X(x) := P(X \le x) = P(-\infty \le X \le x)}$$

PROPERTIES OF THE CUMULATIVE DISTRIBUTION FUNCTION

Key properties of the cumulative distribution function include

- $F_X(x) \in [0, 1]$
- $F_X(-\infty) = 0, F_X(\infty) = 1$
- If $a < b$ then $F_X(a) \leq F_X(b)$
- $P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a)$

THE PROBABILITY DENSITY FUNCTION

We previously mentioned the *probability density function*. It can be characterised via its *relationship with the cumulative distribution function*:

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f_X(x)dx$$

$$\frac{d}{dx}F_X(x) = f_X(x)$$

Note $f_X(x)$ is *not the same as the mass function* from the discrete case! In fact...

PROBABILITY DENSITY VS PROBABILITY MASS

...really, $\int_a^b f_X(x)dx = F(b) - F(a)$ is the *closest to the mass function*.

Or, for $b = a + \Delta a$,

$$P(a < X < a + \Delta a) \approx f_X(x)\Delta a$$

This makes sense - density×vol = (mass/vol)×vol = mass!

This is also consistent with our definition in the discrete case (recall filling in missing entries in the probability table in the tutorial)

EXPECTATION FOR CONTINUOUS RANDOM VARIABLES

We get essentially the same result for *expectation* for continuous RVs as for discrete RVs:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

...and...

VARIANCE FOR CONTINUOUS RANDOM VARIABLES

similarly for *variance* we get

$$Var(X) = \int_{-\infty}^{\infty} [x - E(X)]^2 f_X(x) dx$$

RELATIONSHIP BETWEEN VARIANCE AND MEAN

The *key relation between the variance and mean*

$$Var(X) = E(X^2) - [E(X)]^2$$

also holds again as for our discrete case.

EXAMPLES

Some examples.