Refresher Lecture 1.

Note Title 7/23/2014

Material: Sections 1.1, 1.2.

1 Infinite probability spaces.

Examples of (uncountably) infinite sample spaces.

Experiments:

(i) Choose a number from a unit interval;

(ii) Toss a coin infinitely many times

(iii) Choose a continuous function on interval $[0,1]$.

Sample space Ω, outcome of an experiment $\omega \in \Omega$.

(i) $\Omega = [0,1]$;

(ii) $\Omega = \Omega_\infty$ = the set of infinite sequences of Hs and Ts, $\omega = \omega_1 \omega_2 \cdots \omega_n \ldots \in \Omega$, where $\omega_i \in \{H, T\}$, $i \in \mathbb{N}$.

(iii) $\Omega = C([0,1])$.

We want to define probabilities of subsets of Ω. Unlike in a countable case, we can not assign probabilities to all ω and then define $P(A)$ = sum of probabilities of all $\omega \in A$.

In many examples it will be true that for each $\omega \in \Omega$ $P(\{\omega\}) = 0$.

The approach we take consists of assigning probabilities to a large set of simple subsets in a consistent way and then obtain probabilities of more complicated sets according to a given set of rules. We may not be able to define probability for all subsets of Ω but only for a large class which will be sufficient for all our purposes. Such large class will be a σ-algebra.

Def. 1.1. Let $\Omega \neq \emptyset$ and $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a σ-algebra (or σ-field) if it satisfies

(i) $\emptyset \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;

(iii) $A_i \in \mathcal{F}, \forall i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

<u>Remark</u>: (ii), (iii) imply

(iii') $A_i \in \mathcal{F}, \forall i \in \mathbb{N} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$

(Indeed, by De Morgan's law

$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c \right)^c$. By (ii) $A_i^c \in \mathcal{F}, \forall i \in \mathbb{N}$

by (iii) $\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F}$, and again by (ii)

$\left(\bigcup_{i=1}^{\infty} A_i^c \right)^c \in \mathcal{F}$.)

Also (i), (ii) imply that $\Omega = \emptyset^c \in \mathcal{F}$.

Def. 1.2. Let $\Omega \neq \emptyset$ and $\mathcal{F}$ be a σ-algebra of subsets of Ω. A probability measure $\mathbb{P}$ is a function, $\mathbb{P} : \mathcal{F} \to [0,1]$ such that the following properties hold:

(i) (Normalization) $\mathbb{P}(\Omega) = 1$;

(ii) (Countable additivity) If $A_i \in \mathcal{F}$, $i \in \mathbb{N}$ are pairwise disjoint, i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

<u>Proposition</u> (Basic properties of $\mathbb{P}$)

(i) $\mathbb{P}(\emptyset) = 0$;

(ii) (Finite additivity) if $A_1, A_2, \dots, A_N \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$ $(i \neq j)$ then

$$\mathbb{P}\left(\bigcup_{i=1}^{N} A_i\right) = \sum_{i=1}^{N} \mathbb{P}(A_i);$$

(iii) $\forall A \in \mathcal{F}$ $\quad \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;

(iv) $\forall A, B \in \mathcal{F}$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B);$$

(v) If $A, B \in \mathcal{F}$ and $A \subset B$ then

$$\mathbb{P}(A) \leq \mathbb{P}(B).$$

The proof is left as an exercise (see p.3 of the text).

Example 1.3 (Uniform (Lebesgue) measure on $[0,1]$.)
Model for choosing a number <u>at random</u> from $[0,1]$.

Define $\mathbb{P}([a,b]) = b - a$, $\quad 0 \leq a \leq b \leq 1$.

Then $\mathbb{P}(\{a\}) = \mathbb{P}([a,a]) = a - a = 0$ and
$\mathbb{P}((a,b)) = \mathbb{P}((a,b]) = \mathbb{P}([a,b)) = b - a$.

This defines probabilities for many other sets, say, $\mathbb{P}\left(\left(\frac{1}{4}, \frac{1}{2}\right] \cup \left[\frac{2}{3}, \frac{5}{6}\right)\right) = \frac{1}{4} + \frac{1}{6} = \frac{1}{3}$.

Is the a way to describe all sets for which the probability is determined? Not in any simple constructive way.

It turns out that all sets for which probability is determined forms a σ-algebra, namely, the smalles σ-algebra which contains all closed intervals $[a,b]$, $0 \le a \le b \le 1$. It is called a Borel σ-algebra on $[0,1]$ and is denoted by $\mathcal{B}[0,1]$.

What other sets are in this σ-algebra? Open intervals are there since

$$(a,b) = \bigcup_{n=1}^{\infty} \left[a+\tfrac{1}{n}, b-\tfrac{1}{n}\right] \text{ (see Def 1.1, (iii))}.$$

Essentially all sets of interest for us will be Borel sets.
$\mathcal{B}[0,1]$ can be equivalently defined as the smallest σ-algebra which contains
(i) all open intervals in $[0,1]$ ($[0,a)$, $(b,1]$ are open in $[0,1]$)
(ii) all open subsets of $[0,1]$
(iii) all closed subsets of $[0,1]$

These equivalences are based on Def 1.1 and two facts:

- the complement of every open subset is closed (and the other way around).
- every open subset is an at most countable union of open intervals.

Example 1.4 (infinite independent coin toss prob space). Let Ω_∞ be our sample space. We shall define a sequence of σ-algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_\infty$ (such sequence is called a <u>filtration</u>) and define $\mathbb{P}$ on each of them in a consistent way.

$\mathcal{F}_0 = \{\emptyset, \Omega\}$ (trivial σ-algebra) $\mathbb{P}(\emptyset)=0$
$\mathbb{P}(\Omega)=1$.

Let $A_H = \{\omega : \omega_1 = H\}$
$A_T = \{\omega : \omega_1 = T\}$. Then $A_H \cup A_T = \Omega_\infty$
$A_H \cap A_T = \emptyset$

$\mathcal{F}_1 = \{\emptyset, A_H, A_T, \Omega\}$ (information contained in the first toss: $\forall \omega \in \Omega$ and each set in $\mathcal{F}_1$ we can determine if ω is in that set looking only at the first outcome.)

$|\mathcal{F}_1| = 2^{2^1} = 4$

Define $\mathbb{P}(A_H) = p$ $(0<p<1)$ and
$\mathbb{P}(A_T) = 1-p$

Let $A_{HT} = \{\omega : \omega_1 = H, \omega_2 = T\}$
$A_{TH} = \{\omega : \omega_1 = T, \omega_2 = H\}$
$A_{TT} = \{\omega : \omega_1 = T, \omega_2 = T\}$
$A_{HH} = \{\omega : \omega_1 = H, \omega_2 = H\}$

The smallest σ-algebra which contains these 4 sets is

$$\mathcal{F}_2 = \{\emptyset, A_H, A_T, A_{HT}, A_{TH}, A_{HH}, A_{TT}, A_{HT}^c, A_{TH}^c, A_{TT}^c, A_{HH}^c, A_{HH} \cup A_{TH}, A_{TH} \cup A_{HT},$$

$$A_{TT} \cup A_{HT}, A_{HH} \cup A_{TT}, \Omega\}.$$

$|\mathcal{F}_2| = 2^{2^2} = 16$

Define $\mathbb{P}(A_{HH}) = p^2$; $\mathbb{P}(A_{HT}) = \mathbb{P}(A_{TH}) = p(1-p)$

$\mathbb{P}(A_{HH}) = (1-p)^2$.

Now we can compute the probability of any set in $\mathcal{F}_2$.

$P(A_{TH} \cup A_{HT} \cup A_{TT}) = \mathbb{P}(A_{HH}^c) = 1-p^2$.

$\mathbb{P}(A_H) = \mathbb{P}(A_{HH} \cup A_{HT}) = p^2 + (1-p)p = p$

(consistent with the previous assignment).

and so on for $\mathcal{F}_3, \mathcal{F}_4, \ldots$.

After this we can compute the probability of any set which denends on finitely many tosses. But countable additivity property allows us to compute probabilities of many set which depend on infinitely many tosses. For example, if we want $\mathbb{P}(\{HH\ldots\})$ then we consider

$\mathbb{P}(A_H) = p$

$\mathbb{P}(A_{HH}) = p^2$

$\mathbb{P}(A_{HHH}) = p^3, \ldots$ so $\mathbb{P}(\{HH\ldots\}) = \lim_{n\to\infty} p^n = 0.$

What does it have to do with the countable additivity? We have

$A_H \supset A_{HH} \supset A_{HHH} \supset \dots$ and

$$\{\text{all outcomes are } H\} = \bigcap_{n=1}^{\infty} A_{\underbrace{H\dots H}_{n}}$$

Then $\mathbb{P}(\{HH\dots\}) = \lim_{n\to\infty} \mathbb{P}(A_{\underbrace{H\dots H}_{n}}) = 0$

(this is called continuity of probability measure and it is equivalent to countable additivity, see Jacod, Protter, Probability Essentials, part (ii) of Theorem 2.3 or Appendix A, Theorem A.1.1 of the text.)

Similarly, $\mathbb{P}(\{\omega\}) = 0$ for every $\omega \in \Omega_\infty$.

$$\mathbb{P}(\omega_1 \omega_2 \dots \omega_n) \le (\min\{p, (1-p)\})^n \to 0 \text{ as } n \to \infty$$

for any choice of $\omega_i \in \{H, T\}$.

Finally we define σ-algebra $\mathcal{F}_\infty$ as the smallest σ-algebra with contains all $\mathcal{F}_n$, $n \geq 0$. It has many sets whose probability is difficult to compute. For example,

$$A = \left\{\omega : \lim_{n\to\infty} \frac{\#H \text{ in } n \text{ tosses}}{n} = \frac{1}{2}\right\}$$

Why $A \in \mathcal{F}_\infty$? Define

$$A_{n,m} = \left\{\omega : \left| \frac{\#H \text{ in } n \text{ tosses}}{n} - \frac{1}{2} \right| \le \frac{1}{m}\right\}$$

$A_{n,m} \in \mathcal{F}_n$ for all $m \in \mathbb{N}$.

By the definition of the limit we have
$\forall m \in \mathbb{N} \; \exists N : \; \forall n \geq N$

$$\left| \frac{\# H \text{ in } n \text{ tosses}}{n} - \frac{1}{2} \right| \leq \frac{1}{m}$$

Thus, $A = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} A_{n,m} \in \mathcal{F}_{\infty}$

Recall that the strong law of large numbers says that

$$\mathbb{P}(A) = \begin{cases} 1 & \text{if } p = \frac{1}{2} \\ 0 & \text{if } p \neq \frac{1}{2} \end{cases}$$

Def 1.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If $A \in \mathcal{F}$ satisfies $\mathbb{P}(A) = 1$, we say that the event A occurs almost surely (a.s.).

2 Random variables and distributions.

Def 2.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random variable is a real-valued function $X : \Omega \to \mathbb{R}$ such that for every Borel set $B \in \mathcal{B}(\mathbb{R})$ the subset of Ω given by

$$X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\}$$

is in $\mathcal{F}$.
(The last property is called "measurability of X" with respect to $\mathcal{F}$).

Example. Let $\Omega = \Omega_\infty$ and

H_3 = # of H in the first 3 tosses

Then H_3 is $\mathcal{F}_3$-measurable ($\mathcal{F}_4, \mathcal{F}_5, \dots$ too, as $\mathcal{F}_3 \subset \mathcal{F}_4 \subset \mathcal{F}_5 \dots$) but not $\mathcal{F}_2$-measurable.

Indeed, let $B = \{3\}$. Then

$$H_3^{-1}(B) = \{\omega : H_3(\omega) = 3\} = \{A_{HHH}\}$$

But A_{HHH} is not in $\mathcal{F}_2$!

Def. 2.3 Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. The distribution <u>measure</u> of X is the probability measure μ_X on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that assigns to each $B \in \mathcal{B}(\mathbb{R})$ the mass $\mu_X(B) = \mathbb{P}(X \in B)$.

Example 2.4. Let $\mathbb{P}$ be the uniform measure as in Example 1.3. Define $X(\omega) = \omega$; $Y(\omega) = 1 - \omega$, $\omega \in [0, 1]$

It is easy to check (do it!) that X and Y are random variables.

$$\mu_X[a, b] = \mathbb{P}(\omega : a \le X(\omega) \le b) = \mathbb{P}(\omega : a \le \omega \le b)$$
$$= b - a, \quad 0 \le a \le b \le 1.$$

Therefore, the distribution measure of X is uniform. Y is different from X (in fact, $Y = 1 - X$) but its distribution measure is also uniform:

$$\mu_Y[a,b] = \mathbb{P}(\omega : a \le Y(\omega) \le b)$$
$$= \mathbb{P}(\omega : a \le 1-\omega \le b) = \mathbb{P}(\omega : 1-b \le \omega \le 1-a)$$
$$= (1-a)-(1-b) = b-a, \quad 0 \le a \le b \le 1.$$

Now, suppose we define a new probability measure on $([0,1], \mathcal{B}[0,1])$ by

$$\widetilde{\mathbb{P}}[a,b] = \int_a^b 2\omega\, d\omega = b^2 - a^2, \quad 0 \le a \le b \le 1$$

In the same way as with $\mathbb{P}$ this equation determines probabilities of all Borel sets $\mathcal{B}[0,1]$. Under $\widetilde{\mathbb{P}}$ the random variable X is no longer uniform:

$$\widetilde{\mu}_X[a,b] = \widetilde{\mathbb{P}}(\omega : X(\omega) \in [a,b]) =$$
$$= \widetilde{\mathbb{P}}(\omega : \omega \in [a,b]) = \int_a^b 2\omega\, d\omega = b^2 - a^2.$$

Moreover, the distribution measures of X and Y under $\widetilde{\mathbb{P}}$ are not the same anymore:

$$\widetilde{\mu}_Y[a,b] = \widetilde{\mathbb{P}}(\omega : a \le 1-\omega \le b) = \widetilde{\mathbb{P}}(\omega :$$
$$1-a \le \omega \le 1-b) = \int_{1-a}^{1-b} 2\omega\, d\omega = (1-b)^2 - (1-a)^2.$$

There are other ways to record the distribution of a random variable X rather than specifying μ_X.

The cumulative distribution function (cdf) provides an equivalent way to record the distribution of X.

$$F_X(x) := \mathbb{P}(X \le x) = \mu_X(-\infty, x] \quad , \; x \in \mathbb{R}$$

Moreover, F_X uniquely determines μ_X.
From cdf we find

$$\mu_X(x,y] = \mathbb{P}(X \in (x,y]) = F(y) - F(x)$$
for $x < y$. For $a \le b$ $\quad [a,b] = \bigcap_{n=1}^{\infty} (a - \frac{1}{n}, b]$

and $\mu_X[a,b] = \lim\limits_{n\to\infty} \mu_X(a-\frac{1}{n}, b] = F(b) - \lim\limits_{n\to\infty} F(a - \frac{1}{n}) = F(b) - F(a-)$.

From this $\mu_X[a,b]$ is determined on all $\mathcal{B}(\mathbb{R})$
In two special cases we have a more convenient way to record (and work with) distributions.

Case 1. (X has a density f) $\mu_X[a,b] = \mathbb{P}(a \le X \le b)$

$$= \int_a^b f(x)\,dx \quad , \; 0 < a \le b < \infty$$

$$\left(f(x) \ge 0 ; \quad \int_{-\infty}^{\infty} f(x)\,dx = 1 \right)$$

Case 2 (X has a probability mass function p_X)
There is a finite or infinite sequence of

$x_1, x_2, \dots, x_N, \dots$ such that $P_X(x_i) = p_i$, $\sum_i p_i = 1$

$$\mu_X(B) = \sum_{x_i \in B} P_X(x_i), \quad B \in \mathcal{B}(\mathbb{R}).$$

There are random variables whose distribution is a mixture of the 2 cases. There are also random variables whose distribution has no jumps and no density (see Appendix A, A.3).

Example 2.5 (Another construction of random variable uniformly distributed on $[0,1]$.)
Consider $(\Omega_\infty, \mathcal{F}_\infty, \mathbb{P})$, where $p = \frac{1}{2}$.

Define $Y_n(\omega) = \begin{cases} 1 & \text{if } \omega_n = H \\ 0 & \text{if } \omega_n = T \end{cases}$, $n \in \mathbb{N}$.

Set $X = \sum_{n=1}^{\infty} \frac{Y_n}{2^n}$. Then X is uniform on $[0,1]$. If $Y_1 = 1$ then $X \in [\frac{1}{2}, 1]$ (Prob. $1/2$)
if $Y_1 = 0$ then $X \in [0, \frac{1}{2}]$ (Prob $1/2$)
If $\underset{\text{(probability } 1/4)}{Y_1 = 1 \text{ and } Y_2 = 1}$ then $X \in [3/4, 1]$, and so on.

$$\mu_X\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] = \frac{1}{2^n} \quad , \; 0 \le k \le 2^n - 1$$

$$\mu_X\left[\frac{k}{2^n}, \frac{m}{2^n}\right] = \frac{m}{2^n} - \frac{k}{2^n}.$$ This leads to

$\mu_X[a, b] = b - a$, so X is uniform.

Example 2.6. Let $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

and $N(x) = \int_{-\infty}^{x} \varphi(y)\, dy, \quad x \in \mathbb{R}$.

$N(x)$ is strictly increasing ; $N : \mathbb{R} \to (0,1)$.
It has a strictly increasing inverse $N^{-1}(y)$.
Let Y be uniformly distributed random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Set $X = N^{-1}(Y)$
Then X has density φ (i.e. X is standard normal).

$$
\begin{aligned}
\mu_X [a,b] &= \mathbb{P}(\omega : a \le X(\omega) \le b) \\
&= \mathbb{P}(\omega : a \le N^{-1}(Y(\omega)) \le b) \\
&= \mathbb{P}(\omega : N(a) \le Y(\omega) \le N(b)) \\
&= N(b) - N(a) = \int_a^b \varphi(x)\, dx .
\end{aligned}
$$

HW 1: Exercises 1.1 and 1.4.
Due Thursday, July 24, at 6 p.m..