spl7.tex
Lecture 7. 24.10.2011.

Absolute continuity.

Theorem. If $f \in L(\mu)$,

$$\int_A f d\mu \to 0 \qquad (\mu(A) \to 0).$$

Proof. Write $f_n = f$ if $|f| \leq n$, 0 otherwise. Then $|f_n| \uparrow |f|$. So $\int |f_n| d\mu \uparrow \int f d\mu$: for all $\epsilon > 0$ there exists N with

$$\int |f| d\mu < \int |f_n| d\mu + \epsilon/2 \qquad (n \geq N).$$

Then for $A \in \mathcal{A}$ with $\mu(A) < \epsilon/(2N)$,

$$|\int_A f d\mu| \leq \int_A |f| d\mu = \int_A |f_N| d\mu + \int_A (|f| - |f_N) d\mu < N.\epsilon/(2N) + \int_\Omega (|f| - |f_N|) d\mu$$

$$< \epsilon/2 + \epsilon/2 = \epsilon. \qquad //$$

For f non-negative mble, consider

$$\nu(A) := \int_A f d\mu \qquad (*)$$

as a set-function for $A \in \mathcal{A}$. If $A_n \in \mathcal{A}$ are disjoint with union A,

$$I_{\cup_1^n A_i} = \sum_1^n I_{A_i} \uparrow \sum_1^\infty I_{A_i} = I_{\cup_1^\infty A_i} = I_A.$$

So by monotone convergence,

$$\nu(\cup_1^n A_i) \uparrow \nu(A).$$

So the non-negative set-function ν is σ-additive, that is, ν is a *measure.* By the theorem above,
(i) $\nu(A) \to 0$ as $\mu(A) \to 0$.
Also
(ii)$\nu(A) = 0$ if $\mu(A) = 0$,

as the integral of anything over a null set is 0. Either property (i) or (ii) can be called *absolute continuity* of ν with respect to μ, written

$$\nu << \mu.$$

So if $\nu = \int f d\mu$ as in $(*)$, then ν is absolutely continuous (ac) w.r.t. μ.

The converse is false in general, but true for σ-finite measures, which are all we need. This is the content of the *Radon-Nikodym theorem* (Johann RADON (1887-1956) in 1913 in Euclidean space, Otto NIKODYM (1887-1974) in 1930 in the general case). We quote this result (for proof, see e.g. [S] Ch. 19); the Radon-Nikodym theorem is related to (is in a sense equivalent to) the martingale convergence theorem, which we discuss later.

When $\nu << \mu$, the function f in $(*)$ is called the *Radon-Nikodym derivative* (RN derivative) of ν wrt μ; we then write

$$f = d\nu/d\mu.$$

Then $(*)$ becomes

$$\nu(A) = \int_A \frac{d\nu}{d\mu} d\mu,$$

which we may write symbolically as

$$d\nu = \frac{d\nu}{d\mu}.d\mu,$$

reminiscent of the Chain Rule in ordinary Differential Calculus.

If $\nu << \mu$ and also $\mu << \nu$, we call μ, ν *equivalent.* Then
(i) they have the same null sets;
(ii) both $d\nu/d\mu$ and $d\mu/d\nu$ exist, and

$$d\nu/d\mu = 1/(d\mu/d\nu).$$

Again, note the similarity with ordinary Differential Calculus.
Note. 1. Passing from μ to ν is a *change of measure.* The key result for change of measure in Probability Theory is *Girsanov's theorem.* It plays a crucial role in Mathematical Finance, where the essence of the subject can be reduced to two steps:
(i) discount everything (wrt the risk-free rate of interest) – i.e. pass from prices in nominal terms to prices in real terms;

(ii) take conditional expectations (see later!) under the equivalent martingale measure, EMM (the probability measure equivalent to the original one under which discounted asset prices become martingales – see later!) See [BK], Preface and text, for details.
2. In Probability and Statistics, we typically encounter two distinct kinds of distribution:
(i) *discrete* distributions, such as the Bernoulli, binomial or Poisson distributions, and
(ii) distributions with a *density*, such as the Normal (Gaussian), exponential, chi-square (χ^2) and (Fisher) F distributions.
In case (ii), the density is the RN derivative of the probability measure wrt *Lebesgue measure* ($\phi(x) = \exp\{-x^2/2\}/\sqrt{2\pi}$ for the standard normal distribution $\Phi = N(0,1)$, e^{-x} ($x \geq 0$) for the exponential distribution, etc.) In case (i), the probability mass function ($e^{-\lambda}\lambda^k/k!$ ($k = 0, 1, 2, \ldots$ for the Poisson distribution $P(\lambda)$ with parameter λ, etc.) is the RN derivative of the probability measure wrt *counting measure*. Measure Theory gives us a framework in which we can deal with both these cases together, rather than having to do everything twice, once with sums (discrete case) and once with integrals (density case).

Differentiation and integration; Lebesgue decomposition.

Recall the Fundamental Theorem of Calculus (FTC), that says that differentiation and integration are inverse processes (there are various ways in which this can be made precise). Now that we have a new definition of integration, we need a new version of FTC, as follows.

Theorem (Lebesgue's differentiation theorem, 1904). If $f \in L_1$ and

$$F(x) := \int_{-\infty}^{x} f(u)du,$$

then F is differentiable a.e. and

$$F'(x) = f(x) \qquad a.e.$$

The set-function corresponding to F, the signed measure

$$\mu(A) := \int_A f(x)dx,$$

is absolutely continuous wrt Lebesgue measure λ, with RN-derivative f.

Let F be a probability distribution function; then F is non-decreasing, right-continuous and bounded (between 0 and 1). Such an F can have at most countably many discontinuities, each of which is a *jump* (at most 1 of jump ≥ 1; at most 2 of jump $\geq 1/2$, ..., at most n with jump $\geq 1/n$, etc.). These are the points x_n at which the LS measure μ_F corresponding to F has positive mass $m_n := F(x_n) - F(x_n-)$. Then

$$\mu_{F,j}(A) := \sum_{n: x_n \in A} m_n$$

('j for jump') is a measure with countable support (the set $\{x_n\}$ of atoms), the *jump* component of F. There may be an absolutely continuous component $F_{ac}(x) = \int_{-\infty}^{x} f(u)du$ with RN derivative f say ($f \geq 0$ as F is non-decreasing). There may be a third component F_s, which is continuous and singular (no jumps; grows only on a Lebesgue-null set). According to *Lebesgue's decomposition theorem*, F has a unique decomposition into these three components:

$$F = F_{ac} + F_j + F_s.$$

In this course, F_s will be absent, and usually only one of F_{ac}, F_j will be present – the *density case* (e.g. normal) and *discrete case* (e.g. Poisson) above.

Transformation of integrals.

If $T : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ is measurable ('T for transformation'), and $f = I_A$ for $A \in \mathcal{A}_2$,

$$\int f dT(\mu) = T(\mu)(A) = \mu(T^{-1}(A)) = \int I_{T^{-1}(A)} d\mu = \int I_A(T(x))d\mu(x)$$

$$= \int f(T(x))d\mu(x) = \int f(T)d\mu.$$

The formula

$$\int f dT(\mu) = \int f(T)d\mu$$

extends from $f = I_A$ to simple functions f by linearity. It then extends further to general integrable functions f by approximation (increasing limits of simple functions in the non-negative case, and by considering positive and negative parts separately in the general case), as in L6. For details, see [S] Ch. 14. We will use this formula in Ch. II below for expectations.