ltcc5.tex Week 5 29 October 2019

Chapter V. ITÔ (STOCHASTIC) CALCULUS. WEAK CONVERGENCE.

§1. Quadratic Variation.

A *partition* π_n of $[0,t]$ is a finite set of points t_{ni} such that $0 = t_{n0} < t_{n1} < \ldots < t_{n,k(n)} = t$; the *mesh* of the partition is $|\pi_n| := \max_i(t_{ni} - t_{n,(i-1)})$, the maximal subinterval length. We consider *nested* sequences (π_n) of partitions (each refines its predecessors by adding further partition points), with $|\pi_n| \to 0$. Call (writing t_i for t_{ni} for simplicity)

$$\pi_n B := \sum_{t_i \in \pi_n} (B(t_{i+1}) - B(t_i))^2$$

the *quadratic variation* of B on (π_n). The following classical result is due to Lévy (in his book of 1948); the proof below is from [Pro], §I.3.

THEOREM (Lévy). $\pi_n B \to t$ $(|\pi_n| \to 0)$ in mean square.

Proof.

$$\begin{aligned}\pi_n B - t &= \sum_{t_i \in \pi_n} \{(B(t_{i+i} - B(t_i))^2 - (t_{i+1} - t_i)\} \\ &= \sum_i \{(\Delta_i B)^2 - (\Delta_i t)\} \\ &= \sum_i Y_i,\end{aligned}$$

where since $\Delta_i B \sim N(0, \Delta_i t)$, $E[(\Delta_i B)^2] = \Delta t_i$, so the Y_i have zero mean, and are independent by independent increments of B. So

$$E[(\pi_n B - t)^2] = E[(\sum_i Y_i)^2] = \sum_i E(Y_i^2),$$

as variance adds over independent summands. Now as $\Delta_i B \sim N(0, \Delta_i t)$, $(\Delta_i B)/\sqrt{\Delta_i t} \sim N(0,1)$, so $(\Delta_i B)^2/\Delta_i t \sim Z^2$, where $Z \sim N(0,1)$. So $Y_i = (\Delta_i B)^2 - \Delta_i t \sim (Z^2 - 1)\Delta_i t$,

$$E[(\pi_n B - t)^2] = \sum_i E[(Z^2 - 1)^2](\Delta_i t)^2 = c \sum_i (\Delta_i t)^2,$$

writing c for $E[(Z^2 - 1)^2]$, $Z \sim N(0,1)$, a finite constant. But since

$$\sum_i (\Delta_i t)^2 \leq \max_i \Delta_i t. \sum_i \Delta_i t = |\pi_n|.t,$$

$$E[(\pi_n B - t)^2] \leq c.t.|\pi|_n \to 0 \qquad (|\pi_n| \to 0). \qquad \bullet$$

Note. 1. From convergence in mean square, one can always extract an a.s. convergent subsequence.
2. The conclusion above extends in full generality to a.s. convergence, but an easy proof requires the reversed martingale convergence theorem, which we omit.
3. There is an easy extension to a.s. convergence under the extra restriction $\sum_n |\pi_n| < \infty$, using the Borel-Cantelli lemma and Chebychev's inequality.
4. If we consider the theorem over $[0, t + dt]$, $[0, t]$ and subtract, we can write the result formally as

$$(dB_t)^2 = dt.$$

This can be regarded either as a convenient piece of symbolism, or acronym, or as the essence of *Itô calculus*, to which we turn below.

Background on Other Integrals.

For simplicity, we fix $T < \infty$ and work throughout on time-set $[0, T]$ (in financial applications, T is the *expiry* time, when – or by when – financial derivatives such as options to buy or sell expire.

We want to define integrals of the form

$$I_t(f)(\omega) = \int_0^t f(s, \omega) dB_s,$$

with suitable stochastic processes f as integrands and BM B as integrator.

Remark. We first learn integration with x as integrator, to get $\int_0^t f(x)dx$, first as a *Riemann* integral (this is just the 'Sixth Form integral' in the 'epsilon' language of undergraduate mathematics), then as the *Lebesgue* integral (better, as more general, and easier to manipulate, thanks to the monotone and dominated convergence theorems, etc.). Later, e.g. in handling distribution functions, which may have jumps, we learn $\int_0^t f(x)dF(x)$ for F monotone. One can extend this by linearity to F a difference of two monotone functions – a function locally of *finite variation*, FV. Again, such integrals $\int_0^t f dF$ come in two kinds, Riemann-Stieltjes (R-S) and Lebesgue-Stieltjes (L-S). If we want $\int_0^t f dF$ to exist for all continuous f – as we do – then one needs F to be FV (see e.g. [Pro], Th. 52 of I.7 – though see [Mik], §2.1 for the surprising lengths to which one can push the R-S integral). Now BM has finite *quadratic* variation, so infinite *ordinary* variation. So one needs something new to handle BM as an integrator.

§2. Itô Integral.

We have our filtration $(\mathcal{F}_t)$ on Ω, where $\mathcal{F}_t$ handles randomness up to time t, and the Borel σ-field $\mathcal{B}$ (the smallest containing the intervals) to handle the time-interval $[0, T]$. Write $\mathcal{F}_t \times \mathcal{B}$ for the smallest σ-field containing all $A \times B$, where A is $\mathcal{F}_t$-measurable and B

is $\mathcal{B}$-measurable. Call $f(.,.)$ *measurable* if it is $\mathcal{F}_T \times \mathcal{B}$- measurable and *adapted* if $f(t,.)$ is $\mathcal{F}_t$-measurable for each $t \in [0,T]$. Finally, write $H^2 = H^2[0,T]$ for the class of measurable adapted f satisfying the integrability condition

$$E[\int_0^T f^2(t,\omega)dt] < \infty.$$

For $(a,b] \subset [0,T]$, $f = I_{(a,b)}$, the only plausible candidate for $I(f) := \int f dB$ is

$$I(f)(\omega) = \int_a^b dB_t = B_b - B_a.$$

Similarly for other kinds of interval, $(a,b]$, $[a,b]$, $[a,b)$, since B_t is continuous in t.

Next, integration should be *linear*, so I should extend from indicators to *simple* functions by linearity. Write H_0^2 for the class of simple square- integrable functions – those f of the form

$$f(t,\omega) = \sum_{i=0}^{n-1} a_i(\omega) I(t_i < t \leq t_{i+1})$$

with a_i $\mathcal{F}(t_i)$-measurable, $E(a_i^2) < \infty$ and $0 = t_0 < \ldots < t_n = T$. The only plausible candidate for $I_t(f)$ for $f \in H_0^2$ is

$$I_t(f)(\omega) = \sum_{i=0}^{n-1} a_i(\omega)(B(t \wedge t_{i+1}) - B(t \wedge t_i)).$$

By continuity of B_t in t, $I_t(f)$ is continuous in t. Next, for $u \geq s$, $E(B_u|\mathcal{F}_s) = B_s$ as B is a mg, while for $u \leq s$, $E(B_u|\mathcal{F}_s) = B_u$ as then B_u is known at time s. Combining,

$$E(B_u|\mathcal{F}_s) = B(\min(u,s)).$$

Thus for $s \leq t$,

$$E[I_t(f)(\omega)|\mathcal{F}_s] = \sum_{i=0}^{n-1} a_i(\omega)(B(s \wedge t \wedge t_{i+1}) - B(s \wedge t \wedge t_i),$$

which – as $s \leq t$ – is

$$\sum_{i=0}^{n-1} a_i(\omega)(B(s \wedge t_{i+1}) - B(s \wedge t_i),$$

which is $I_s(f)$. Combining, this says that $I_t(f)$ is a martingale. So, from the mg property, the products of increments over disjoint intervals have zero mean (all relevant expectations exist as we are assuming square-integrability). Consider

$$E[I_t(f)^2] = E[(\sum_{i=0}^{k-1} a_i(B(t_{i+1}) - B(t_i)) + a_k(B(t) - B(t_k))^2] \qquad (t_k \leq t < t_{k+1}).$$

Expand out the square; the cross-terms have zero expectation, leaving

$$E[\sum_{i=0}^{k-1} a_i^2(B(t_{i+1}) - B(t_i))^2 + a_k^2(B(t) - B(t_k))^2] = \sum_{i=0}^{k-1} E(a_i^2)(t_{i+1} - t_i) + E(a_k^2)(t - t_k),$$

which is $\int_0^t E(f^2(s,\omega))ds$ as f^2 is a_i^2 on $(t_i, t_{i+1}]$. Combining:

PROPOSITION. For $f \in H_0^2$ a simple function, and $(I_t f)(\omega)$ defined as above,
(i) $I_t f$ is a continuous martingale,
(ii) $E[(I_t f)(\omega)^2] = E[\int_0^t f^2(s,\omega)ds]$ (Itô isometry).

One seeks to extend I_t from simple functions $f \in H_0^2$ to general $f \in H^2$ (an extension analogous to the extension of the Lebesgue integral from simple to measurable functions, in the most basic non-random measure-theoretic set-up). It is not at all obvious, but it is true, that – with H_0^2, H^2 regarded as Hilbert spaces with the norm $\|f\| := (\int_0^T f^2(s,\omega)ds)^{1/2}$ – H_0^2 is *dense* in H^2 - that is, each $f \in H^2$ is the limit in norm of an approximating sequence $f_n \in H_0^2$. It is further true that the map I_t extends from H_0^2 to H^2 via

$$(I_t f)(\omega) := \lim_{n\to\infty}(I_t f_n)(\omega)$$

(the limit is in the Hilbert-space norm), and that the limit above does not depend on the particular choice of approximating sequence. That is:

PROPOSITION. For $f \in H^2$ and $(I_t f)(\omega)$ defined by approximation as above:
(i) $(I_t f)(\omega)$ is a continuous martingale,
(ii) The Itô isometry $E[(I_t f)(\omega)^2] = E[\int_0^t f^2(s,\omega)ds$ holds.

We must refer to a rigorous measure-theoretic treatment for details of the proof. Full accounts are in [R-Y], IV, [K-S], Ch. 3. See also [Pro], Ch. II. An accessible recent account is [Ste], Ch. 6; [Mik], Ch. 2 is non-rigorous, but useful.

We now call $I_t f$ the *Itô integral* of $f \in H^2$, and use the suggestive Leibniz integral sign (which dates from 1675):

$$(I_t f)(\omega) = \int_0^t f(s,\omega)dB(s,\omega) \quad \text{or} \quad \int_0^t f_s dB_s \qquad (0 \le t \le T).$$

An alternative notation is

$$(f.B)_t := \int_0^t f_s dB_s.$$

Predictability. We used $(t_i, t_{i+1}]$ to ensure left-continuity, so predictability. The distinction is not critical with Brownian motion as an integrator, or more generally a continuous mg integrator, but it is critical with a general semimg integrator.

The appropriate topology (or convergence concept, in metric-space situations as here) is that of uniform convergence in probability on compact time-sets (ucp). The simple predictable functions (class S) are dense in the adapted càglàd functions (class L) in the ucp topology ([Pro], II.4), and this allows the approximation and extension results we need (and assume) to go through.

An Example. We calculate $\int_0^t B_s dB_s$, using Brownian motion both as (continuous, so previsible) integrand and as integrator. Take a sequence of partitions π_n with mesh $|\pi_n| \to 0$, and write t_i for the partition points t_{ni} of π_n, as above. The approximation properties sketched above allow us to identify $\int_0^t B_s dB_s$ with the limit of

$$\sum_{t_i \in \pi_n} B(t_i)(B(t_{i+1} - B(t_i)).$$

But this is

$$\sum \frac{1}{2}(B(t_{i+1}) + B(t_i))(B(t_{i+1}) - B(t_i)) - \sum \frac{1}{2}(B(t_{i+1}) - B(t_i))(B(t_{i+1}) - B(t_i)).$$

The first sum is $\frac{1}{2}\sum(B(t_{i+1})^2 - B(t_i)^2)$, which telescopes to $\frac{1}{2}B(t)^2$ $(B(0) = 0)$. The second sum is $\frac{1}{2}\sum(\Delta_i B)^2$, which tends to$\frac{1}{2}t$ by Lévy's theorem on the quadratic variation of BM. Combining:

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

This formula of course differs dramatically from that for ordinary (Newton-Leibniz) calculus, or its Riemann-Stieltjes or Lebesgue-Stieltjes extensions. The role of the second term – the correction term or *Itô term* – illustrates both the contrast between the Itô and earlier integrals and the decisive influence of the quadratic variation on the Itô integral.

Quadratic Variation. For continuous semimartingales X, the quadratic variation process $\langle X \rangle = (\langle X \rangle_t : t \geq 0)$ is defined by

$$\langle X \rangle = X^2 - 2\int X_- dX$$

(of course $X_- = X$ when X is continuous, as here). Alternatively, X^2 is a submg, and then

$$X^2 = \langle X \rangle + 2\int X_- dX$$

is the Doob (or Doob-Meyer) decomposition of X^2 into an increasing previsible process $\langle X \rangle$ and a mg $2\int X_- dX$.

For general (not necessarily continuous) mgs X, the quadratic variation $[X]$ involves the jumps ΔX_t of X also. Then if $[X]^c$ is the continuous part of the increasing process $[X]$,

$$[X]_t = [X]_t^c + \sum_{0 \le s \le t} (\Delta X_s)^2,$$

and if X^c is the continuous mg part of X in its semimg decomposition,

$$\langle X^c \rangle = [X^c] = [X]^c.$$

Both $\langle X \rangle$ and $[X]$ are shorthand for $\langle X, X \rangle$ and $[X, X]$. By polarization,

$$\langle X, Y \rangle := \frac{1}{4}(\langle X+Y, X+Y \rangle - \langle X-Y, X-Y \rangle)$$

(and similarly for $[X, Y]$), both quadratic forms extend to different arguments X and Y. Both $\langle X, Y \rangle$ and $[X, Y]$ are locally of BV, so are semimgs. For details, see e.g. [Pro], II.6.

Product Rule.

The quadratic covariation $[X, Y]$ is, by polarization,

$$[X, Y] = XY - \int X_- dY - \int Y_- dX.$$

Rearranging: for X, Y semimgs, so is XY, and then

$$XY = \int X_- dY + \int Y_- dX + [X, Y].$$

This is the *integration-by-parts* formula, or *product rule*. It is the principal special case of *Itô's formula* (below) – to which it is in fact equivalent.

With H, K previsible integrands and X, Y semimg integrators, we can form both the stochastic integrals $H.X = \int H dX$, $K.Y = \int K dY$. We can then form the (square-)bracket processes, for which

$$[H.X, K.Y]_t = \int H_s K_s d[X, Y]_s;$$

in particular,

$$[H.X, H.X]_t = \int H_s^2 d[X, X]_s \qquad \text{or} \qquad [H.X]_t = \int H_s^2 d[X]_s.$$

For Brownian motion, $[B]_t = \langle B \rangle_t$ (as there are no jumps) $= t$ (by Lévy's result on quadratic variation). So specialising,

$$\langle \int f dB \rangle = \int_0^t f_s^2 ds, \qquad \langle \int f_s dB_s, \int_0^t g_s dB_s \rangle = \int_0^t f_s g_s ds.$$

§3. Itô's Formula

The change-of-variable formula (or chain rule) of ordinary calculus extends to the Lebesgue-Stieltjes integral, and tells us that for smooth f $(\in C^1)$ and continuous A, of FV (on compacts),

$$f(A_t) - f(A_0) = \int_0^t f'(A_s) dA_s.$$

(This of course does *not* apply to $\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$, but there the integral is Itô, not Lebesgue-Stieltjes.)

Rather less well-known is the extension to A only *right*-continuous:

$$f(A_t) - f(A_0) = \int_{0+}^t f'(A_{s-}) dA_s + \sum_{0<s\le t} \{f(A_s(-f(A_{s-}) - f'(A_{s-})\Delta A_s\}.$$

(None of the analysis books on integration that I have to hand – by Burkill, Saks, Kestelman, McShane, Hildebrandt and Pesin – seem to contain this.) Proof is deferred, as this is a special case of the result below (or see a book on stochastics, e.g. [R-W2], IV.18).

One thus seeks a setting capable of handling both the last two displayed results together.

THEOREM (Itô's Formula). For X a semimartingale and $f \in C^2$, then $f(X)$ is also a semimartingale, and

$$\begin{aligned} f(X_t) - f(X_0) = &\int_{0+}^t f'(X_{s-}) dX_s + \frac{1}{2}\int_{0+}^t f''(X_{s-}) d[X, X]_s^c \\ &+ \sum_{0<s\le t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\}. \end{aligned}$$

First, we note the special case for X continuous.

THEOREM (Itô's Formula). For X a continuous semimartingale and $f \in C^2$, $f(X)$ is also a continuous semimartingale, and

$$f(X_t) - f(X_0) = \int_0^t f'(X_s) dX_s + \frac{1}{2}\int_0^t f''(X_s) d[X, X]_s.$$

COROLLARY. If $X = X_0 + M + A$ is the decomposition of the continuous semimartingale X, that of $f(X)$ is

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dM_s + \{\int_0^t f'(X_s) dA_s + \frac{1}{2}\int_0^t f''(X_s) d[M]_s\}.$$

Proof. Let $\mathcal{A}$ be the class of C^2 functions f for which the desired result holds. Then $\mathcal{A}$ is a vector space. But $\mathcal{A}$ is also closed under multiplication, so is an *algebra*. Indeed, if $f,\ g \in \mathcal{A}$, write F_t, G_t for the semimgs $f(X_t)$ and $g(X_t)$ and use the product rule. If $X = X_0 + M + A$ is the decomposition of X into a (continuous) local mg M and a BV process A, the continuous local mg part of $F = f(X)$ is $\int f'(X_s)dM_s$. So

$$[F,G]_t = [F^{cm}, G^{cm}]_t = [\int_0^{\cdot} f'(X_s)dM_s, \int_0^{\cdot} g'(X_s)dM_s]_t,$$

which by above is

$$\int_0^t f'(X_s)g'(X_s)d[M,M]_s.$$

The product rule now says that

$$F_tG_t - F_0G_0 = \int_0^t F_s dG_s + \int_0^t G_s dF_s + \int_0^t (f'g')(X_s)d[M]_s.$$

As Itô's formula holds for f (by assumption),

$$dF_t = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d[M]_t,$$

and similarly for g. Substituting,

$$F_tG_t - F_0G_0 = \int_0^t \{F_s g'(X_s) + f'(X_s)G_s\}dX_s +$$
$$\frac{1}{2}\int_0^t \{F_s g''(X_s) + 2f'(X_s)g'(X_s) + f''(X_s)G_s\}d[M]_s.$$

This says that Itô's formula also holds for fg. So $\mathcal{A}$ is an algebra, and a vector space. Since $\mathcal{A}$ contains $f(x) = x$, $\mathcal{A}$ contains all polynomials.

One can now reduce from the local-mg to the mg case by a localization argument, and extend from $\mathcal{A}$ to C^2 by an approximation argument. For details, we refer to [R-W2], IV.32 (and for the extension to discontinuous X, VI.39). A different approach, using Taylor's formula for f, is in [Pro], II.7. •

Note. 1. *Higher dimensions.* For vector functions $f = (F_1, \dots, f_n)$ and vector processes $X = (X^1, \dots, X_n)$, we use the Einstein summation convention and $D_i f$ for $\partial f/\partial_i$. Then the Itô formula extends (with much the same proof) as

$$f(X_t) - f(X_0) = \int_0^t D_i f(X_s)dX_s^i + \frac{1}{2}\int_0^t D_{ij} f(X_s)d[X^i, X^j]_s.$$

2. With $f(x,y) = xy$, we then recover the product rule (integration-by-parts formula). As we used the second in the proof above, in this sense Itô's formula and the product rule are equivalent.
3. In particular, the class of semimgs is closed under C^2 functions. This gives a powerful – and highly non-linear – closure property of the class of semi-martingales.
4. *Differential Notation.* In the one-dimensional case, we may re-write Itô's formula in shorthand form using differential notation instead of integral notation as

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d[X]_t.$$

The left is called the *stochastic differential* of $f(X)$. So the above gives the stochastic differential of $f(X)$ in terms of the stochastic differential dX of X and the quadratic variation $[X,X]$ or $[X]$ of X.
5. *Formalism.* For the continuous case, we can obtain Itô's formula from Taylor's formula in differential form, if we adopt the following rules for differentials:

$$dX^i dX^j = d[X^i, X^j],$$

$$dX^i dX^j dX^k = 0,$$

$$dXdV = 0$$

whenever V has FV. In particular, for $X = B$ BM in one dimension,

$$dB_t^i dB_t^j = 0 \qquad (i \neq j), \qquad (dB_t^i)^2 = dt,$$

$$dBdV = 0 \qquad V \text{ of finite variation on compacts.}$$

The interpretation here is that for $i \neq j$, the Brownian increments dB^i and dB^j are independent with zero mean, so $E(dB_t^i dB_t^j) = E(dB_t^i).E(dB_t^j) = 0.0 = 0$, while for $i = j$ the earlier symbolism $(dB_t)^2 = dt$ becomes $(dB_t^i)^2 = dt$. The formalism above works well, and is a flexible and reliable tool in practice.

Recall that normally in calculus we simply omit all differentials of order higher than one. We do have some prior experience of retaining second-order differentials, e.g. in differential geometry (first and second quadratic forms, curvature, geodesics etc.). So retaining second-order differentials, and manipulating them by the above simple rules, is not a complete culture-shock. (Differential geometry and stochastic calculus in fact combine, as stochastic differential geometry - see e.g. [R-W2], V §5.)

One can also do stochastic calculus for *Lévy processes* (stochastic processes with stationary independent increments); see Applebaum [A]. Apart from Brownian motion, the prime example is the *Poisson process*; see Kingman [K].

§4. Weak convergence

Convergence in Distribution. If c_n, c are constants, and $c_n \to c$, then regarding c_n, c as random variables one should have $c_n \to c$ in any sense of convergence of random variables – in particular, for convergence in distribution. If F_{c_n}, F_c are their distribution functions, one has $F_{c_n}(x) \to F_c(x)$ as $n \to \infty$ for all $x \neq c$ – the point where the limit function has a jump. For a set A, write $\bar{A}$ for its *closure*, A^o for its *interior*, $\partial A := \bar{A} \setminus A^o$ for its *boundary*. With P_n, P the corresponding probability measures, $P_n((-\infty, x]) \to P((-\infty, x])$ for all $x \notin \partial(-\infty, c]$. It turns out that this is the right definition of convergence in distribution, as it generalizes – to d dimensions (random vectors), and infinitely many dimensions (stochastic processes). We confine ourselves here to processes with *continuous paths* (e.g. Brownian motion). We take the (time) parameter set as $[0, 1]$ for simplicity. If X is such a process, its paths lie in $C[0, 1]$, the space of continuous functions on $[0, 1]$. This is a metric space under the sup-norm metric $d(f, g) := \sup\{|f(t) - g(t)| : t \in [0, 1]\}$. The *Borel σ-field* $\mathcal{B} = \mathcal{B}(C([0, 1])$ of $C[0, 1]$ is the σ-field generated by the open sets of $C[0, 1]$ w.r.t. this metric. The *distribution*, or *law*, of a process X on $C[0, 1]$ is given by

$$P(B) := P(X(.) \in B), \qquad B \in \mathcal{B}.$$

If X_n, X are such processes, with laws P_n, P, one says $X_n \to X$ *in distribution*, or *in law*, or *weakly*, or $P_n \to P$ *weakly*, if

$$P_n(B) \to P(B) \qquad (n \to \infty) \qquad \forall B \in \mathcal{B} \text{ with } P(\partial B) = 0.$$

It turns out that this is equivalent to

$$\int f dP_n \to \int f dP \qquad (n \to \infty)$$

for all f bounded and continuous on $[0, 1]$ (as $[0, 1]$ is compact, f continuous on $[0, 1]$ implies f bounded also, but in general we have to require f bounded and continuous). It turns out also (Prohorov's Theorem) that weak convergence is equivalent to:
(i) convergence of finite-dimensional distributions (clearly a minimal requirement), and
(ii) *tightness*: for all $\epsilon > 0$ there exists a compact set K such that $P_n(K) > 1 - \epsilon$ for all n.
For proof, see e.g. Billingsley [B].

Statistical Applications. These include the *Kolmogorov-Smirnov* (K-S) tests for equality of two distributions F, G, in terms of the *Kolmogorov-Smirnov statistic* $D_n := \sup\{|F_n(.) - G_n(.)\}$ of the distance between their empirical distribution functions. In fact D_n has the same law as $\sup\{B_0(t) : t \in [0, 1]\}$, where B is Brownian motion and B_0 is the *Brownian bridge*: $B_0(t) := B(t) - t$. Similarly for many other functionals – *Donsker's Invariance Principle.* This is the dynamic form of the Central Limit Theorem (CLT).