Chapter 2

Diffusion Processes

2.1 What is a Diffusion Process?

When we want to model a stochastic process in continuous time it is almost impossible to specify in some reasonable manner a consistent set of finite dimensional distributions. The one exception is the family of Gaussian processes with specified means and covariances. It is much more natural and profitable to take an evolutionary approach. For simplicity let us take the one dimensional case where we are trying to define a real valued stochastic process with continuous trajectories. The space $\Omega = C[0,T]$ is the space on which we wish to construct the measure P. We have the σ-fields $\mathcal{B}_t = \sigma\{x(s) : 0 \leq s \leq t\}$ defined for $t \leq T$. The total σ-field $\mathcal{B} = \mathcal{B}_T$. We try to specify the measure P by specifying approximately the conditional distributions $P[x(t+h)-x(t) \in A|\mathcal{B}_t]$. These distributions are nearly degenerate and and their mean and variance are specified as

$$E^P\big[x(t+h) - x(t)|\mathcal{B}_t\big] = h\, b(t,\omega)) + o(h) \tag{2.1}$$

and

$$E^P\big[(x(t+h) - x(t))^2|\mathcal{B}_t\big] = h\, a(t,\omega)) + o(h) \tag{2.2}$$

as $h \to 0$, where for each $t \geq 0$ $b(t,\omega)$ and $a(t,\omega)$ are $\mathcal{B}_t$ measurable functions. Since we insist on continuity of paths, this will force the distributions to be nearly Gaussian and no additional specification should be necessary. We will devote the next few lectures to investigate this.

Equations (2.1)and (2.2) are infinitesimal differential relations and it is best to state them in integrated forms that are precise mathematical statements.

We need some definitions.

Definition 2.1. *We say that a function $f : [0,T] \times \Omega \to R$ is progressively measurable if, for every $t \in [0,T]$ the restiction of f to $[0,t] \times \Omega$ is a measurable function of t and ω on $([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{B}_t)$ where $\mathcal{B}[0,t]$ is the Borel σ-field on $[0,t]$.*

The condition is somewhat stronger than just demanding that for each t, $f(t, \omega)$ is $\mathcal{B}_t$ is measurable. The following facts are elementary and left as exercises.

Exercise 2.1. If $f(t, x)$ is measurable function of t and x, then $f(t, x(t, \omega))$ is progressively meausrable.

Exercise 2.2. If $f(t, \omega)$ is either left continuous (or right continuous) as function of t for every ω and if in addition $f(t\, omega)$ is $\mathcal{B}_t$ measurable for every t, then f is progressively measurable.

Exercise 2.3. There is a sub σ-field $\Sigma = \Sigma_{pm} \subset \mathcal{B}[0, T] \times \mathcal{B}_T)$ such that progressive measurability is just measurability with respect to Σ_{pm}. In particular standard operations performed on progreesively measurable functions yield progressively measurable functions.

We shall always insist that the functions $b(\cdot\,,\cdot)$ and $a(\cdot\,,\cdot)$ be progressively measurable. Let us suppose in addition that they are bounded functions. The boundedness will be relaxed at a later stage.

We reformulate conditions 2.1 and 2.2 as

$$M_1(t) = x(t) - x(0) - \int_0^t b(s, \omega) ds \tag{2.3}$$

and

$$M_2(t) = [M_1(t)]^2 - \int_0^t a(s\,, \omega)) ds \tag{2.4}$$

are martingales with respect to $(\Omega), \mathcal{B}_t, P)$.

We can then define a Diffusion Process corresponding to a, b as a measure P on $(\Omega), \mathcal{B})$ such that relative to $(\Omega), \mathcal{B}_t, P)$ $M_1(t)$ and $M_2(t)$ are martingales. If in addition we are given a probability measure μ as the initial distribution, i.e.

$$\mu(A) = P[x(0) \in A]$$

then we can expect P to be determined by a, b and μ.

We saw already that if $a \equiv 1$ and $b \equiv 0$, with $\mu = \delta_0$, we get the standard Brownian Motion. $a = a(t, x(t))$ and $b = b(t, x(t))$, we expect P to be a Markov Process, because the infinitesimal parameters depend only on the current position and not on the past history. If there is no explicit dependence on time, then the Markov Process can be expected to have stationary transition probabilities. Finally if $a(t, x) = a(t)$ is purely a function of t and $b(t, \omega)) = b_1(t) + \int_0^t c(t\,, s) x(s) ds$ is linear in ω), then one expects P to be Gaussian, if μ is so.

Because the pathe are continuous the same argument that we provided earlier can be used to establish that

$$\begin{aligned} Z_\lambda(t) &= \exp[\lambda M_1(t) - \frac{\lambda^2}{2} \int_0^t a(s\,, \omega) ds] \\ &= \exp[\lambda[x(t) - x(0) - \int_0^t b(s\,, \omega) ds] - \frac{\lambda^2}{2} \int_0^t a(s\,, \omega) ds] \end{aligned} \tag{2.5}$$

is a martingale with respect to $(\Omega), \mathcal{B}_t, P)$ for every real λ. We can also take for our definition of a Diffusion Process corresponding to a, b the condition that $Z_\lambda(t)$ be a martingale with respect to $(\Omega), \mathcal{B}_t, P)$ for every λ. If we do that we did not have to assume that the paths were almost surely continuous. $(\Omega, \mathcal{B}_t, P)$ could be any space suppporting a stochastic process $x(t\,,\omega)$ such that the martingale property holds for $Z_\lambda(t)$. If C is an upper bound for a, it is easy to check with $M_1(t)$ defined by equation (2.5)

$$E^P\left[\exp[\lambda[M_1(t) - M_1(s]\right] \leq \exp[\frac{\lambda^2 C}{2}]$$

The lemma of Garsia, Rodemich and Rumsey will guarantee that the paths can be chosen to be continuous.

Let $(\Omega, \mathcal{F}, P)$ be a Probability space. Let $\mathbf{T}$ be the interval $[0, T]$ for some finite T or the infinite interval $[0, \infty)$. Let $\mathcal{F}_T \subset \mathcal{F}$ be sub σ-fields such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s, t \in \mathbf{T}$ with $s < t$. We can assume with out loss of generality that $\mathcal{F} = \vee_{t \in \mathbf{T}} \mathcal{F}_t$. Let a stochastic process $x(t\,,\omega)$ with values in R^n be given. Assume that it is progressively measurable with respect to $(\Omega\,, \mathcal{F}_t)$. We can easily gneralize the ideas described in the previous section to diffusion processe with values in R^n. Given a positive semidefinite $n \times n$ matrix $a = a_{i,j}$ and an n-vector $b = b_j$, we define the operator

$$(\mathcal{L}_{a,b} f)(x) = \frac{1}{2} \sum_{i,j=1}^{n} a_{i,j} \partial^2 f \partial x_i \partial x_j(x) + \sum_{j=1}^{n} \partial f \partial x_j(x)$$

If $a(t\,,\omega) = a_{i,j}(t\,,\omega)$ and $b(t\,,\omega) = b_j(t\,,\omega)$ are progresssively measurable functions we define

$$(L_{t\,,\omega} f)(x) = (L_{a(t\,,\omega), b(t\,,\omega)} f)(x)$$

Theorem 2.1. *The following defintions are equivalent.* $x(t\,,\omega)$ *is a diffusion process correponding to bounded progressively measurable functions* $a(\cdot\,,\cdot), b(\cdot\,,\cdot)$ *with values in the space of symmetric positive semidefinite* $n \times n$ *matrices, and* n*-vectors if*

1. $x(t\,,\omega)$ *has an almost surely continuous version and*

$$y_i(t\,,\omega) = x_i(t\,,\omega) - x_i(0\,,\omega) - \int_0^t b(s\,,\omega) ds$$

and

$$z_{i,j}(t\,,\omega) = y_i(t\,,\omega)\, y_j(t\,,\omega) - \int_0^t a_{i,j}(s\,,\omega) ds$$

are $(\Omega, \mathcal{F}_t, P)$ *martingales.*

2. *For every* $\lambda \in R^n$

$$Z_\lambda(t\,,\omega) = \exp\left[< \lambda\,, y(t\,,\omega) > - \frac{1}{2} \int_0^t < \lambda, a(s\,,\omega)\lambda > ds \right]$$

is an $(\Omega, \mathcal{F}_t, P)$ *martingale.*

3. *For every* $\lambda \in R^n$

$$X_\lambda(t\,,\omega) = \exp\left[i < \lambda\,, y(t\,,\omega) + \frac{1}{2}\int_0^t < \lambda,\, a(s\,,\omega)\lambda > ds\right]$$

is an $(\Omega, \mathcal{F}_t, P)$ *martingale.*

4. *For every smooth bounded function* f *on* R^n *with atleast two bounded continuous derivatives*

$$f(x(t\,,\omega)) - f((x(0\,,\omega)) - \int_0^t (\mathcal{L}_{s,\omega} f)(x(s\,,\omega))ds$$

is an $(\Omega, \mathcal{F}_t, P)$ *martingale.*

5. *For every smooth bounded function* f *on* $\mathbf{T} \times R^n$ *with atleast two bounded continuous* x *derivatives and one bounded continuous* t *derivative*

$$f(t\,, x(t\,,\omega)) - f(0\,, (x(0\,,\omega)) - \int_0^t (\frac{\partial f}{\partial s} + \mathcal{L}_{s,\omega} f)(s\,, x(s\,,\omega))ds$$

is an $(\Omega, \mathcal{F}_t, P)$ *martingale.*

6. *For every smooth bounded function* f *on* $\mathbf{T} \times R^n$ *with atleast two bounded continuous* x *derivatives and one bounded continuous* t *derivative*

$$\exp\Bigg[f(t\,, x(t\,,\omega)) - f(0\,, (x(0\,,\omega)) - \int_0^t (\frac{\partial f}{\partial s} + \mathcal{L}_{s,\omega} f)(s\,, x(s\,,\omega))ds \\ - \frac{1}{2}\int_0^t < (\nabla f)(s\,, x(s\,,\omega)), a(s\,,\omega)\,(\nabla f)(s\,, x(s\,,\omega)) > ds\Bigg]$$

is an $(\Omega, \mathcal{F}_t, P)$ *martingale.*

7. *Same as (6) except that* f *is replaced by* g *of the form*

$$g(t\,, x) = < \lambda,\, x > + f(t\,, x)$$

where f *is as in (6) and* $\lambda \in R^n$ *is arbitrary.*

Under any one of the above definitions, $x(t\,,\omega)$ *has an almost surely continuous version satifying*

$$P\left[\sup_{0 \le s \le t} |y(s\,,\omega) - y(0\,,\omega)| \ge \ell\right] \le 2n \exp[\frac{-\ell^2}{Ct}]$$

for some constant C *depending only on the dimension* n *and the upper bound for* a*. Here*

$$y_i(t\,,\omega) = x_i(t\,,\omega) - x_i(0\,,\omega) - \int_0^t b_i(s\,,\omega)ds$$

Proof. (1) *implies* (2). This was essentially the content of Theorem and the comments of the previous section. Also we saw that the exponential inequality is a consequence of Doob's inequality.
(2) *implies* (3). The condition that $Z_\lambda(t)$ is a martingale can be rewritten as a whole collecction of identities

$$\int_A Z_\lambda(t\,,\omega)dP = \int_A Z_\lambda(s\,,\omega)dP \tag{2.6}$$

that is valid for every $t > s$, $A \in \mathcal{F}_s$ and $\lambda \in R^n$. Both sides of eqation (2.6) are well defined when $\lambda \in R^n$ is replaced by $\lambda \in \mathbf{C}^n$, with complex components and define entire functions of the n complex variables λ. Since they agree when the values are real, by analytic continuation, they must agree for all purely imaginary values of λ as well. This is just (3).
(3) *implies* (4). This part of the proof requires a simple lemma.

Lemma 2.2. *Let $M(t\,,\omega)$ be a martingale relative to $(\Omega\mathcal{F}_t, P)$ which has almost surely continuous trajectories and $A(t\,,\omega)$ be a progressively measurable process that is for almost all ω a continuous function of bounded variation in t. Assume that for every t the random variable $\xi(t\,,\omega) = \sup_{0\le s\le t} |M(t)|\,Var_{[0,t]}\,A(t\,,\omega)$ has a finite expectation. Then*

$$\eta(t) = M(t)A(t) - M(0)A(0) - \int_0^T M(s)dA(s)$$

is again a martingale relative to $(\Omega, \mathcal{F}_t, P)$.

Proof. (of lemma.) We need to prove that for every $s < t$,

$$E^P\left[M(t)A(t) - M(s)A(s) - \int_s^t M(u)dA(u)\big|\mathcal{F}_s\right] = 0 \qquad \text{a.e.}$$

We can subdivide the interval $[s,t]$ into subintervals with end points $s = t_0 < t_1 < \cdots < t_N = t$, and approximate $\int_s^t M(u)dA(u)$ by $\sum_{j=1}^N M(t_j)[A(t_j) - A(t_{j-1})]$. The fact that A is continuous and $\xi(t)$ is integrable makes the approximation work in $L_1(P)$ so that

$$\begin{aligned}
E^P\left[\int_s^t M(u)dA(u)\big|\mathcal{F}_s\right] &= \lim_{N\to\infty} E^P\left[\sum_{j=1}^N M(t_j)[A(t_j) - A(t_{j-1})]\big|\mathcal{F}_s\right] \\
&= \lim_{N\to\infty} E^P\left[\sum_{j=1}^N [M(t_j)A(t_j) - M(t_j)A(t_{j-1})]\big|\mathcal{F}_s\right] \\
&= \lim_{N\to\infty} E^P\left[\sum_{j=1}^N [M(t_j)A(t_j) - M(t_{j-1})A(t_{j-1})]\big|\mathcal{F}_s\right] \\
&= E^P\left[M(t)A(t) - M(s)A(s)\right]
\end{aligned}$$

and we are done. We used the martingale property in going from the second line to the third when we replaced $M(t_j)A(t_{j-1})$ by $M(t_{j-1})A(t_{j-1})$ $\square$

Now we return to the proof of the theorem. Let us apply the above lemma with $M_\lambda(t) = X_\lambda(t)$ and

$$A_\lambda(t) = \exp[i\int_0^t < \lambda\,, b(s) > ds - \frac{1}{2}\int_0^t < \lambda\,, a(s)\lambda > ds].$$

Then a simple computation yields

$$\begin{aligned} M_\lambda(t)A_\lambda(t) - M_\lambda(0)A_\lambda(0) - \int_0^t M_\lambda(s)dA_\lambda(s) \\ = e_\lambda(x(t) - x(0)) - 1 - \int_0^t (\mathcal{L}_{s,\omega}e_\lambda)((x(s) - x(0))ds \end{aligned}$$

where $e_\lambda(x) = \exp[i < \lambda\,, x >]$. Multiplying by $\exp[i < \lambda\,, x(0) >]$, which is essentially a constant, we conclude that

$$e_\lambda(x(t)) - e_\lambda(x(0)) - \int_0^t (\mathcal{L}_{s,\omega}e_\lambda)((x(s))ds$$

is a martingale. The above expression is just what we had to prove, except that our f is special namely, the exponentials $e_\lambda(x)$. But by linear combinations and limits we can easily pass from exponentials to arbitray smooth bounded functions with two bounded derivatives. We first take care of infinitely diffrentiable functions with compact support by Fourier integrals and then approximate twice differentiable functions with those.

<u>(4) *implies* (3)</u>. The steps can be retraced. We start with the martingales defined by (4) in the special case of f being e_λ and choose

$$A_\lambda(t) = \exp[-i\int_0^t < \lambda\,, b(s) > ds + \frac{1}{2}\int_0^t < \lambda\,, a(s)\lambda > ds]$$

and do the computations to get back to the martingales of type (3).

<u>(4) *implies* (5)</u>. This is basically a computation. If $f(t\,, x)$ can be approximated by smooth function and so we may assume with out loss of generality more

derivatives.

$$
\begin{aligned}
&E^P[f(t\,,x(t)) - f(s\,,x(s))|\mathcal{F}_s]\\
&\quad= E^P[f(t\,,x(t)) - f(t\,,x(s))|\mathcal{F}_s] + E^P[f(t\,,x(s)) - f(s\,,x(s))|\mathcal{F}_s]\\
&\quad= E^P[\int_s^t (\mathcal{L}_{u,\omega} f(t\,,\cdot))(x(u))du|\mathcal{F}_s] + E^P[\int_s^t \frac{\partial f}{\partial u}(u\,,x(s))du|\mathcal{F}_s]\\
&\quad= E^P[\int_s^t (\mathcal{L}_{u,\omega} f(u\,,\cdot))(x(u))du|\mathcal{F}_s]\\
&\qquad+ E^P[\int_s^t (\mathcal{L}_{u,\omega}[f(t\,,\cdot) - f(u\,,\cdot)])(x(u))]du|\mathcal{F}_s]\\
&\qquad\quad+ E^P[\int_s^t \frac{\partial f}{\partial u}(u\,,x(u))du|\mathcal{F}_s]\\
&\qquad\qquad+ E^P[\int_s^t [\frac{\partial f}{\partial u}(u\,,x(s)) - \frac{\partial f}{\partial u}(u\,,x(u))]du|\mathcal{F}_s]\\
&\quad= E^P[\int_s^t [\frac{\partial f}{\partial u} + (\mathcal{L}_{u,\omega} f)](u\,,x(u))du|\mathcal{F}_s] + J
\end{aligned}
$$

where

$$
\begin{aligned}
J &= E^P[\int_s^t (\mathcal{L}_{u,\omega}[f(t\,,\cdot) - f(u\,,\cdot)])(x(u))du|\mathcal{F}_s]\\
&\quad+ E^P[\int_s^t [\frac{\partial f}{\partial u}(u\,,x(s)) - \frac{\partial f}{\partial u}(u\,,x(u))]du|\mathcal{F}_s]\\
&= E^P[\int_s^t \int_u^t (\frac{\partial f}{\partial v}\mathcal{L}_{u,\omega} f)(v\,,x(u))du\,dv|\mathcal{F}_s]\\
&\quad- E^P[\int_s^t \int_s^u (\mathcal{L}_{v,\omega}\frac{\partial f}{\partial u})(u\,,(x(v))du\,dv|\mathcal{F}_s]\\
&= E^P\Big[\int\int_{s\le u\le v\le t} (\mathcal{L}_{u,\omega}\frac{\partial f}{\partial v})(v\,,(x(u))du\,dv\\
&\quad- \int\int_{s\le v\le u\le t} (\mathcal{L}_{v,\omega}\frac{\partial f}{\partial u})(u\,,(x(v))du\,dv\Big]\\
&= 0.
\end{aligned}
$$

The two integrals are identical, just the roles of u and v have been interchanged.
<u>(5) *implies* (4)</u>. This is trivial because after all in (5) we are allowed to take f <u>to be purely a</u> function of x.
<u>(5) *implies* (6)</u>. This is again the lemma on multiplying a martingale by a function of bounded variation. We start with a function of the form $\exp[f(t\,,x)]$ and the martingale

$$
\exp[f(t\,,x(t))] - \exp[f(0\,,x(0))] - \int_0^t (\frac{\partial e^f}{\partial s} + \mathcal{L}_{s,\omega} e^f)(s\,,x(s))ds
$$

and use

$$A(t) = \exp\big[-\int_0^t (\frac{\partial f}{\partial s} + \mathcal{L}_{s,\omega} f)(s\,,x(s))ds$$
$$-\frac{1}{2}\int_0^t < (\nabla f)(s\,,x(s))\,,a(s)(\nabla f)(s\,,x(s)) > ds\big]$$

<u>(6) *implies* (5)</u>. This just again reversing the steps.
<u>(6) *implies* (7)</u>. The problem here is that the function $<\lambda\,,x>$ are unbounded. If we pick a function $h(x)$ of one variable to equal x in the interval $[-1.1]$ and then levels off smoothly we get easily a smooth bounded function with bounded derivatives that agrees with x in $[-1,1]$. Then the sequence $h_(x) = kh(\frac{x}{k})$ clearly converges to x, $|h_k(x)| \le |x|$ and more over $|h'_k(x)|$ is uniformly bounded in x and k and $|h''_k(x)|$ goes to 0 uniformly in k. We approximate $<\lambda\,,x>$ by $\sum_j \lambda_j h_k(x_j)$ and consider the martingales

$$\exp\left[\sum_j \lambda_j h_k(x_j(t)) - \sum_j \lambda_j h_k(x_j(0)) - \int_0^t \psi_k^\lambda(s)ds\right]$$

where

$$\psi_k^\lambda(s) = \int_0^t \sum_j \lambda_j b_j(s\,,\omega)h'_k(x_j(s))ds + \frac{1}{2}\int_0^t \sum_j a_{j,j}(s\,,\omega)h''_k(x_j(s))ds$$
$$+\frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\lambda_i\lambda_j h'_i(x_i(s)h'_j(x_j(s)ds$$

and converges to

$$\psi^\lambda(s) = \int_0^t \sum_j \lambda_j b_j(s\,,\omega)ds + \frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\lambda_i\lambda_j ds$$

as $k \to \infty$. By Fatous's lemma the limit of nonnegative martingales is always a supermartingale and therefore in the limit

$$\exp\left[<\lambda\,,x(t)-x(0)> - \int_0^t \psi^\lambda(s)ds\right]$$

is a supermartingale. In particular

$$E^P\left[\exp[<\lambda\,,x(t)-x(0)> - \int_0^t \psi^\lambda(s)ds]\right] \le 1$$

If we now use the bound on ψ it is easy to obtain the estimate

$$E^P[\exp[<\lambda\,,x(t)-x(0)>] \le C_\lambda$$

This provides the necessary uniform integrability to conclude that in the limt we have a martingale. Once we have the estimate, it is easy to see that we can approximate $f(t\,,x)+<\lambda\,,x>$ by $f(t\,,x)+\sum_j \lambda_j h_k(x_j)$ and pass to the limit, thus obtaining (7) from (6). Of course (7) implies both (2) and (6). Also all the exponential estimates follow at this point. Once we have the estimates there is no difficulty in obtainig (1) from (3). We need only take $f(x)=x_i$ and $x_i x_j$ that can be justified by the estimates. Some minor manipulation is needed to obtain the results in the form presented. $\square$

2.2 Random walks and Brownian Motion

Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables with mean 0 and variance 1. The partial sums S_k are defined by $S_0 = 0$ and for $k \geq 1$

$$S_k = X_1 + X_2 + \cdots + X_k$$

We rescale and interpolate to define stochastic processes $X_n(t) : 0 \leq t \leq 1$ by

$$X_n(\frac{k}{n}) = \frac{S_k}{\sqrt{n}}$$

for $0 \leq k \leq n$ and for $1 \leq k \leq n$ and $t \in [\frac{k-1}{n}, \frac{k}{n}]$

$$X_n(t) = (nt - k + 1)X_n(\frac{k}{n}) + (k - nt)X_n(\frac{k-1}{n})$$

Let P_n denote the distribution of the process $X_n(\cdot)$ on $\mathbf{X} = C[0\,,1]$ and P the distribution of Brownian Motion, or the Wiener measure as it is often called. We want to explore the sense in which

$$\lim_{n\to\infty} P_n = P$$

Lemma 2.3. *For any finite collection* $0 \leq t_1 < t_2 < \cdots < t_m \leq 1$ *of* m *time points the joint distribution of* $(x(t_1), \cdots, x(t_m))$ *under* P_n *converges, as* $n \to \infty$*, to the corresponding distribution under* P.

Proof. We are dealing here basically with the central limit theorem for sums independent random variables. Let us define $k_n^i = [nt_i]$ and the increments

$$\xi_n^i = \frac{S_{k_n^i} - S_{k_n^{i-1}}}{\sqrt{n}}$$

for $i = 1, 2, \cdots, m$ with the convention $k_n^0 = 0$. For each n, ξ_n^i are m mutually independent random variables and their distributions converge as $n \to \infty$ to Gaussians with 0 means and variances $t_i - t_{i-1}$ respectively. We take $t_0 = 0$. This is of course the same distribution for these increments under Brownian Motion. The interpolation is of no consequence, because the difference between the end points is exactly some $\frac{X_i}{\sqrt{n}}$. So it does not really matter if in the definition of $X_n(t)$ if we take $k_n = [nt]$ or $k_n = [nt]+1$ or take the interpolated value. We can state this convergensce in the form

$$\lim_{n\to\infty} E^{P_n}\big[f(x(t_1), x(t_2), \cdots, x(t_m))\big] = E^P\big[f(x(t_1), x(t_2), \cdots, x(t_m))\big]$$

for every m, any m time points $(t_1, t_2, \cdots, t_m)$ and any bounded continuous function f on $\mathbf{R}^m$. □

These measures P_n are on the space $\mathbf{X}$ of bounded continuous functions on $[0\,,1]$. The space $\mathbf{X}$ is a metric space with $d(f,g) = \sup_{0\le t\le 1} |f(t)-g(t)|$ as the distance between two continuous functions. The main theorem is

Theorem 2.4. *If $F(\cdot)$ is a bounded continuous function on $\mathbf{X}$ then*

$$\lim_{n\to\infty} \int_{\mathbf{X}} F(\omega)dP_n = \int_{\mathbf{X}} F(\omega)dP$$

Proof. The main difference is that functions depending on a finite number of coordinates have been replaced by functions that are bounded and continuous, but otherwise arbitrary. The proof proceeds by approximation. Let us assume Lemma 2.5 which asserts that for any $\epsilon > 0$, there is a compact set K_ϵ such that $\sup_n P_n[\mathbf{X} - K_\epsilon] \le \epsilon$ and $P[\mathbf{X} - K_\epsilon] \le \epsilon$. From standard approximation theory (i.e. Stone-Weierstrass Theorem) the continuous function F, which we can assume to be bounded by 1, can be approximated by a function f depending on a finite number of coordinates such that $\sup_{\omega\in K_\epsilon} |F(\omega)-f(\omega)| \le \epsilon$. Moreover we can assume without loss of generality that f is also bounded by 1. We can estimate

$$|\int_{\mathbf{X}} F(\omega)dP_n - \int_{\mathbf{X}} f(\omega)dP_n| \le \int_{K_\epsilon} |F(\omega) - f(\omega)|dP_n + 2P_n[K_\epsilon^c] \le 3\epsilon$$

as well as

$$|\int_{\mathbf{X}} F(\omega)dP - \int_{\mathbf{X}} f(\omega)dP| \le \int_{K_\epsilon} |F(\omega) - f(\omega)|dP + 2P[K_\epsilon^c] \le 3\epsilon$$

Therefore

$$|\int_{\mathbf{X}} F(\omega)\,dP_n - \int_{\mathbf{X}} F(\omega)\,dP| \le 6\epsilon + |\int_{\mathbf{X}} f(\omega)\,dP_n - \int_{\mathbf{X}} f(\omega)\,dP|$$

and we are done. □

Remark 2.1. We shall prove Lemma 2.5 under the additional assuption that the underlying random variables X_i have a finite 4-th moment. See the exercise at the end to remove this condition.

Lemma 2.5. *Let P_n, P be as before. Assume that the random variables X_i have a finite moment of order four. Then for any $\epsilon > 0$ there exists a compact set $K_\epsilon \subset \mathbf{X}$ such that*

$$P_n[K_\epsilon] \ge 1 - \epsilon$$

for all n and

$$P[K_\epsilon] \geq 1 - \epsilon$$

as well.

Proof. The set

$$K_{B,\alpha} = \{f : f(0) = 0,\ |f(t) - f(s)| \leq B|t - s|^\alpha\}$$

is a compact subset of $\mathbf{X}$ for each fixed B and α. Theorem 1.3 can be used to give us a uniform estimate on $P_n[K^c_{B,\alpha}]$ which can be made small by taking B large enough. We need only to check that the condition (1.2) holds for P_n with some constants β, α and C that do not depend on n. Such an estimate clearly holds for the Brownian motion P.

If $\{X_i\}$ are independent identically distributed random variables with zero mean, an elementary calculation yields

$$E[(X_1 + X_2 + \cdots + X_k)^4] = kE[X_1^4] + 3k(k-1)\big[E[X_1^2]\big]^2 \leq C_1 k + C_2 k^2 \tag{2.7}$$

Let us try to estimate $E[(X_n(t) - X_n(s))^4]$. If $|t - s| \leq \frac{2}{n}$ we can estiamte

$$|X_n(t) - X_n(s)| \leq M|t - s|$$

where M is the maximum slope. There are atmost three intervals involved and

$$E[M^4] \leq n^2\, E\big[[\max|X_i|, |X_2|, |X_3|]^4\big] \leq C\, n^2$$

which implies that

$$E^{P_n}\big[|x(t) - x(s)|^4\big] \leq |t - s|^4\, E[M^4] \leq C|t - s|^2 \tag{2.8}$$

If $|t - s| > \frac{2}{n}$ we can find t', s' such that ns', nt' are integers, $|t - t'| \leq \frac{1}{n}$ and $|s - s'| \leq \frac{1}{n}$. Applying the estimate (2.8) for the end pieces that are increments over incomplete intervals and the estimate (2.7) for the piece $|x(t') - x(s')|$, we get

$$E^{P_n}[|x(t) - x(s)|^4] \leq C\, n^{-2} + \frac{C}{n}|t' - s'| + C|t' - s'|^2$$

Since both $|t - s|$ and $|t' - s'|$ are atleast $\frac{1}{n}$ we obtain (1.2). □

Exercise 2.4. To extend the result to the case where only the second moment exists, we do truncation and write $X_i = Y_i + Z_i$. The pairs $\{(Y_i, Z_i) : 1 \leq i \leq n\}$ are mutually independent identically distributed random vectors. We can asume that both Y_i and Z_i have mean 0. We can fix it so that Y_i has variance 1 and a finite fourth moment. Z_i can be forced to have an arbitrarily small variance σ^2. We have $X_n(t) = Y_n(t) + Z_n(t)$ and by Kolmogorov's inequality

$$P\big[\sup_{0 \leq t \leq 1} |Z_n(t)| \geq \delta\big] \leq \delta^{-2} E\big[[Z_n(1)]^2\big] = \delta^{-2}\sigma^2$$

which can be made small uniformly in n if σ^2 is small enough. Complete the proof.