Stochastic Analysis for the Curious Postgraduate

Or: What some of the words mean and why you might care

Tim Sullivan
http://www.tjsullivan.org.uk/

Postgraduate Seminar
Warwick Mathematics Institute
29 October 2008

Brownian motion

- J. Ingenhousz 1785, R. Brown 1827: "large" particles suspended in water move "randomly" because they are bombarded by randomly moving water molecules.
- T.N. Thiele 1880: initial analysis of these random motions.
- L. Bachelier 1900: stochastic analysis of the stock and option markets.
- A. Einstein 1905, M. Smoluchowski 1906: these Brownian paths are random continuous paths with stationary, independent, normally distributed increments.

T. Lucretius Carus, c. 60BC, *De Rerum Natura*

"Observe what happens when sunbeams are admitted into a building and shed light on its shadowy places. You will see a multitude of tiny particles mingling in a multitude of ways. . . their dancing is an actual indication of underlying movements of matter that are hidden from our sight. . . It originates with the atoms which move of themselves. Then those small compound bodies that are least removed from the impetus of the atoms are set in motion by the impact of their invisible blows and in turn cannon against slightly larger bodies. So the movement mounts up from the atoms and gradually emerges to the level of our senses, so that those bodies are in motion that we see in sunbeams, moved by blows that remain invisible."

Questions

- How to choose a continuous path in $\mathbb{R}^n$ "at random"?
- How to do measure/probability theory on infinite-dimensional spaces?
- How to calculate integrals (expectations) like

$$\mathbb{E}[f] = \int_{C^0([0,T];\mathbb{R}^n)} f(x) \, \mathrm{d}x? \quad \text{(A. Einstein, N. Wiener...)}$$

$$\mathbb{E}[f] = \int_{\text{universes}} f(u) \, \mathrm{d}u? \quad \text{(R. Feynman, S. Hawking...)}$$

- How to make sense of a differential equation like

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = b(t, X(t)) + \text{"noise"}?$$

Definition

A measure space is a triple $(\mathscr{X}, \mathscr{F}, \mu)$ where

- $\mathscr{X}$ is a set;
- $\mathscr{F}$ is a σ-algebra: a family of subsets of $\mathscr{X}$, containing $\mathscr{X}$, and closed under countable unions, intersections, set differences...;
- $\mu \colon \mathscr{F} \to [0, +\infty]$ is a measure, satisfying

$$\mu(\emptyset) = 0 \text{ and } \mu\left(\biguplus_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mu(A_k).$$

- If $\mu(\mathscr{X}) = 1$, then μ is called a probabilty measure and $(\mathscr{X}, \mathscr{F}, \mu)$ a probability space.

Definition

Consider a Hausdorff topological space $\mathcal{X}$ and a Borel measure $\mu\colon \mathcal{B}(\mathcal{X}) \to [0, +\infty]$

- μ is strictly positive if every open set U has $\mu(U) > 0$.
- μ is locally finite if every point $x \in \mathcal{X}$ has a (open) neighbourhood N_x with $\mu(N_x) < +\infty$.
- μ is invariant under $T\colon \mathcal{X} \to \mathcal{X}$ if $T_* \mu = \mu$, i.e.

$$\text{for all Borel } B \subseteq \mathcal{X}, \mu\big(T^{-1}(B)\big) = \mu(B).$$

- μ is quasi-invariant under $T\colon \mathcal{X} \to \mathcal{X}$ if $T_* \mu \approx \mu$, i.e.

$$\text{for Borel } B \subseteq \mathcal{X}, \mu\big(T^{-1}(B)\big) = 0 \iff \mu(B) = 0.$$

Definition

Lebesgue measure on $\mathbb{R}^n$ is the "completion" or "extension" of the usual notion of n-dimensional volume for rectangular boxes in $\mathbb{R}^n$ to $\mathcal{B}(\mathbb{R}^n)$.

Theorem (The wonders of Lebesgue measure)

Lebesgue measure on $\mathbb{R}^n$ is locally finite, strictly positive, and invariant under all translations. Moreover, up to multiplication by a positive constant, it is the only Borel measure with these properties.

A naïve attempt to construct a "Lebesgue measure" on

$$C_0 := \left\{ x \colon [0,T] \to \mathbb{R}^n \middle| x \text{ continuous}, x(0) = 0 \right\}$$

might go something like this...

- I wish to integrate $f \colon C_0 \to \mathbb{R}$ "with respect to $x \in C_0$".
- I pick a partition $\Pi = \{0 = t_0 < t_1 < \ldots < t_k = T\}$ of $[0,T]$, let $\bar{x}^\Pi$ be a "piecewise constant version of x", and calculate

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(\bar{x}^\Pi)\, \mathrm{d}x(t_0)\mathrm{d}x(t_1)\ldots\mathrm{d}x(t_k)$$

- Now take the limit as $k \to \infty$ and the mesh of the partition Π tends to zero.

$$\int_{C_0} f(x)\, \mathcal{D}x := \overset{(!!!)}{\lim_{\mathrm{mesh}(\Pi)\to 0}} \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(\bar{x}^\Pi)\, \mathrm{d}x(t_0)\mathrm{d}x(t_1)\ldots\mathrm{d}x(t_k).$$

- This is known as a path integral.

Theorem (Bad news for physicists!)

Let $\mathcal{X}$ be an infinite-dimensional, separable Hilbert (or even just Banach) space. Then the only locally finite and translation-invariant Borel measure on $\mathcal{X}$ is the trivial (zero) measure.

Theorem (Bad news for physicists!)

Let $\mathcal{X}$ be an infinite-dimensional, separable Hilbert (or even just Banach) space. Then the only locally finite and translation-invariant Borel measure on $\mathcal{X}$ is the trivial (zero) measure.

Proof.

Suppose $\mathbb{B}_r(0)$ has finite measure. Then there exists a countably infinite family $\{\mathbb{B}_{r/4}(x_i)\}_{i \in \mathbb{N}}$, all contained in $\mathbb{B}_r(0)$, all having the same measure. For their union to have finite measure, they must each have measure zero. Since $\mathcal{X}$ is separable, it can be covered by a countable family $\{\mathbb{B}_{r/4}(y_i)\}_{i \in \mathbb{N}}$, and so has measure zero! $\square$

Theorem (Bad news for physicists!)

Let $\mathcal{X}$ be an infinite-dimensional, separable Hilbert (or even just Banach) space. Then the only locally finite and translation-invariant Borel measure on $\mathcal{X}$ is the trivial (zero) measure.

Proof.

Suppose $\mathbb{B}_r(0)$ has finite measure. Then there exists a countably infinite family $\{\mathbb{B}_{r/4}(x_i)\}_{i \in \mathbb{N}}$, all contained in $\mathbb{B}_r(0)$, all having the same measure. For their union to have finite measure, they must each have measure zero. Since $\mathcal{X}$ is separable, it can be covered by a countable family $\{\mathbb{B}_{r/4}(y_i)\}_{i \in \mathbb{N}}$, and so has measure zero! □

The moral of the tale is that infinite-dimensional spaces are big and weird. Life is no better in non-separable spaces: μ might not turn out to be trivial, but won't be strictly positive.

Definition

Gaussian measure on $\mathbb{R}$ with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$ is the Borel measure γ defined by

$$\gamma(A) := \frac{1}{\sigma\sqrt{2\pi}} \int_A \exp\left(-\frac{|x-m|^2}{2\sigma^2}\right) \mathrm{d}x.$$

Definition

Gaussian measure on $\mathbb{R}$ with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$ is the Borel measure γ defined by

$$\gamma(A) := \frac{1}{\sigma\sqrt{2\pi}} \int_A \exp\left(-\frac{|x-m|^2}{2\sigma^2}\right) \mathrm{d}x.$$

Definition

Gaussian measure on $\mathbb{R}^n$ with mean $m \in \mathbb{R}^n$ and covariance matrix $C \in \mathbb{R}^{n\times n}$ is the Borel measure γ defined by

$$\gamma(A) := \frac{1}{\sqrt{(2\pi)^n \det C}} \int_A \exp\left(-\frac{(x-m)\cdot C^{-1}(x-m)}{2}\right) \mathrm{d}x.$$

It's easy to check that Gaussian measures on $\mathbb{R}^n$ always have Gaussian push-forward on $\mathbb{R}$ *via* any linear map $\mathbb{R}^n \to \mathbb{R}$, so we make an extended definition for more general spaces:

Definition

A Borel measure μ on a (Banach) space $\mathcal{X}$ is called a Gaussian measure if, for every $\ell \in \mathcal{X}^*$, the push-forward measure $\ell_* \mu$ on $\mathbb{R}$ is a Gaussian measure, i.e. normal distribution with some finite mean and variance.

It's easy to check that Gaussian measures on $\mathbb{R}^n$ always have Gaussian push-forward on $\mathbb{R}$ *via* any linear map $\mathbb{R}^n \to \mathbb{R}$, so we make an extended definition for more general spaces:

Definition

A Borel measure μ on a (Banach) space $\mathcal{X}$ is called a Gaussian measure if, for every $\ell \in \mathcal{X}^*$, the push-forward measure $\ell_* \mu$ on $\mathbb{R}$ is a Gaussian measure, i.e. normal distribution with some finite mean and variance.

- It's equivalent to require that $T_* \mu$ be a finite-dimensional Gaussian measure for every continuous linear map $T \colon \mathcal{X} \to \mathbb{R}^n$.
- It's also convenient to allow Dirac masses as "degenerate" Gaussian measures (with zero variance).

Definition

The mean of a (Gaussian) measure μ on $\mathcal{X}$ is the element $m \in \mathcal{X}$ such that, for every $\ell \in \mathcal{X}^*$,

$$\int_{\mathcal{X}} \ell(x - m) \, \mathrm{d}\mu(x) = 0,$$

or, if one is comfortable with vector-valued integrals,

$$m = \int_{\mathcal{X}} x \, \mathrm{d}\mu(x).$$

Definition

The covariance operator of μ is a bilinear operator

$$C_\mu \colon \mathcal{X}^* \times \mathcal{X}^* \to \mathbb{R},$$

$$C_\mu(k, \ell) := \int_{\mathcal{X}} k(x - m)\ell(x - m)\,\mathrm{d}\mu(x).$$

- By fixing one argument in $\mathcal{X}^*$, can also view C_μ as an operator

$$C_\mu \colon \mathcal{X}^* \to \mathcal{X}^{**} (\cong \mathcal{X} \text{ if } \mathcal{X} \text{ is reflexive}).$$

- Using Riesz's representation theorem that every Hilbert space $\mathcal{H}$ is isomorphic to its dual $\mathcal{H}^*$, can view C_μ for μ on a Hilbert space $\mathcal{H}$ as an operator

$$C_\mu \colon \mathcal{H} \to \mathcal{H}.$$

Theorem (Fernique)

Let μ be a (mean-zero) Gaussian measure on a separable Banach space $\mathcal{X}$. Then μ has exponentially small tails: there exists a constant $\alpha > 0$ such that

$$\int_{\mathcal{X}} \exp\left(\alpha\|x\|^2\right) \mathrm{d}\mu(x) < +\infty.$$

Hence, μ has finite mean, variance... finite moments of all orders, and all continuous linear functionals $\ell \in \mathcal{X}^*$ are integrable.

Corollary (Continuity of the covariance operator)

The covariance operator C_μ is a continuous linear operator, i.e. there exists $\|C_\mu\|_{\mathrm{op}} < +\infty$ such that

$$|C_\mu(k, \ell)| \leq \|C_\mu\|_{\mathrm{op}} \|k\|_{\mathcal{X}^*} \|\ell\|_{\mathcal{X}^*} \text{ for all } k, \ell \in \mathcal{X}^*.$$

Definition

A bounded linear operator $K : \mathcal{H} \to \mathcal{H}$ is of trace class if for some (and hence all) orthonormal bases $\{e_k\}_k$ of $\mathcal{H}$,

$$\sum_k \langle (K^*K)^{1/2} e_k, e_k \rangle_{\mathcal{H}} < \infty.$$

In this case, the trace $\operatorname{tr} K := \sum_k \langle K e_k, e_k \rangle_{\mathcal{H}}$ is absolutely convergent and is independent of the choice of the orthonormal basis.

Definition

A bounded linear operator $K: \mathcal{H} \to \mathcal{H}$ is of trace class if for some (and hence all) orthonormal bases $\{e_k\}_k$ of $\mathcal{H}$,

$$\sum_k \langle (K^* K)^{1/2} e_k, e_k \rangle_{\mathcal{H}} < \infty.$$

In this case, the trace $\operatorname{tr} K := \sum_k \langle K e_k, e_k \rangle_{\mathcal{H}}$ is absolutely convergent and is independent of the choice of the orthonormal basis.

Theorem (Classification of covariance operators)

Consider a separable Hilbert space $\mathcal{H}$ and a Gaussian measure μ on $\mathcal{H}$. Then $C_\mu : \mathcal{H} \to \mathcal{H}$ is trace class.
Conversely, if $K: \mathcal{H} \to \mathcal{H}$ is positive semi-definite, symmetric, and of trace class, then $K = C_\mu$ for some Gaussian measure μ on $\mathcal{H}$.

Definition

Two measures μ, ν on $\mathcal{X}$ are mutually singular if $\mathcal{X} = A \uplus B$ with $\mu(A) = \nu(B) = 0$

Example

- Two Dirac measures (point masses) on distinct points are mutually singular.
- Any Dirac measure and Lebesgue measure on $\mathbb{R}$ are mutually singular.

Theorem

Let μ, ν be two Gaussian measures on an infinite-dimensional Banach space $\mathcal{X}$. Then μ and ν are either equivalent or they are mutually singular.

Definition

The Cameron–Martin space for a Gaussian measure μ on a Banach space $\mathcal{X}$ is a Hilbert space $\mathcal{H}$ continuously embedded in $\mathcal{X}$ and defined equivalently by

- $\mathcal{H}$ is the completion of

  $$\left\{h \in \mathcal{X} \,\middle|\, \text{for some } h' \in \mathcal{H}^*, C_\mu(h', -) = \langle -, h \rangle \in \mathcal{H}^{**} \cong \mathcal{H}\right\}$$

  with respect to $\langle h, k \rangle_{\mathcal{H}} := C_\mu(h', k')$
- $\mathcal{H}$ is the intersection of all μ-measure-1 subspaces of $\mathcal{X}$;
- $\mathcal{H}$ is the set of all directions $v \in \mathcal{X}$ so that μ and $T_*^v \mu$ are equivalent.

Warning! If $\dim \mathcal{H} = +\infty$, then $\mathcal{H}$ has μ-measure equal to zero!

Theorem

Consider standard Gaussian measure γ *on* $\mathbb{R}^n$ *(mean zero, covariance = identity matrix), and let* T^v *denote translation by* $v \in \mathbb{R}^n$. *Then* $T^v_* \gamma$ *is equivalent to* γ *with density*

$$\frac{\mathrm{d}T^v_* \gamma}{\mathrm{d}\gamma}(x) = \exp\left(v \cdot x - \frac{1}{2}|v|^2 \right)$$

i.e., for all $f \in L^1(\mathbb{R}^n, \gamma; \mathbb{R})$,

$$\int_{\mathbb{R}^n} f(x) \, \mathrm{d}T^v_* \gamma(x) = \int_{\mathbb{R}^n} f(x) \exp\left(v \cdot x - \frac{1}{2}|v|^2 \right) \mathrm{d}\gamma(x)$$

Theorem (Cameron–Martin)

For any Gaussian measure μ *on a Banach space* $\mathcal{X}$*, and any* $v \in \mathcal{H} \subseteq \mathcal{X}$*,* $T^v_* \mu$ *is equivalent to* μ *with density*

$$\frac{\mathrm{d}T^v_* \mu}{\mathrm{d}\mu}(x) = \exp\left(\langle v, x \rangle^\sim - \frac{1}{2} \| v \|^2_{\mathcal{H}} \right)$$

i.e., for all $f \in L^1(\mathcal{X}, \mu; \mathbb{R})$*,*

$$\int_{\mathcal{X}} f(x) \, \mathrm{d}T^v_* \mu(x) = \int_{\mathcal{X}} f(x) \exp\left(\langle v, x \rangle^\sim - \frac{1}{2} \| v \|^2_{\mathcal{H}} \right) \mathrm{d}\mu(x).$$

$\langle v, x \rangle^\sim$ is, in some sense, an extension of the inner product $\langle \cdot, \cdot \rangle \colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ to something $\mathcal{H} \times \mathcal{X} \to \mathbb{R}$. It's called the Paley–Wiener integral and is our first example of a stochastic integral.

Corollary (Integration by parts)

Suppose $f : \mathcal{X} \to \mathbb{R}$ *has Fréchet derivative* $\mathrm{D}f : \mathcal{X} \to \mathcal{X}^*$. *Then integrating the Cameron–Martin formula gives*

$$\int_{\mathcal{X}} f(x+ti(h))\,\mathrm{d}\mu(x) = \int_{\mathcal{X}} f(x) \exp\left(t\langle h, x\rangle^{\sim} - \frac{1}{2}t^2\|h\|_{\mathcal{H}}^2 \right) \mathrm{d}\mu(x)$$

for any $t \in \mathbb{R}$. *Formally differentiating with respect to* t *and evaluating at* $t = 0$ *gives the integration by parts formula*

$$\int_{\mathcal{X}} \mathrm{D}f(x)(i(h))\,\mathrm{d}\mu(x) = \int_{\mathcal{X}} f(x)\langle h, x\rangle^{\sim}\,\mathrm{d}\mu(x).$$

Formulae like this give rise to the motto that "stochastic integrals are like infinite-dimensional divergence operators."

Our first kind of stochastic integral, the Paley–Wiener integral, comes right out of the abstract formulation.

Recall that the Cameron–Martin space is a Hilbert space $\mathcal{H}$ sitting inside $\mathcal{X}$; the inclusion $i \colon \mathcal{H} \hookrightarrow \mathcal{X}$ is continuous, linear and injective; for simplicity, assume that $i(\mathcal{H})$ is dense in $\mathcal{X}$.

Our first kind of stochastic integral, the Paley–Wiener integral, comes right out of the abstract formulation.

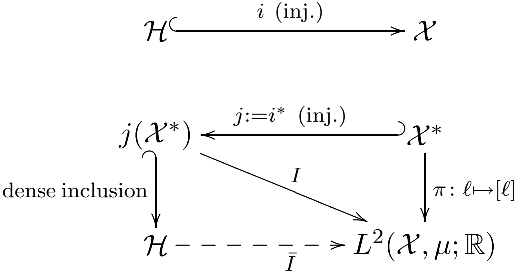
Recall that the Cameron–Martin space is a Hilbert space $\mathcal{H}$ sitting inside $\mathcal{X}$; the inclusion $i\colon \mathcal{H} \hookrightarrow \mathcal{X}$ is continuous, linear and injective; for simplicity, assume that $i(\mathcal{H})$ is dense in $\mathcal{X}$.

Lemma

Given that $i\colon \mathcal{H} \hookrightarrow \mathcal{X}$ is continuous, linear and injective with dense range, so is the adjoint $j := i^*\colon \mathcal{X}^* \to \mathcal{H}^* \cong \mathcal{H}$, where

$$i^*(\ell)(h) = \ell(i(h)) \text{ for all } \ell \in \mathcal{X}^*, h \in \mathcal{H}.$$

(All arrows are continuous linear maps.)



Lemma

For all $f \in \mathcal{X}^*$, $\|j(\ell)\|_{\mathcal{H}} = \|\ell\|_{L^2(\mathcal{X},\mu;\mathbb{R})}$.

The extended map $\bar{I} \colon \mathcal{H} \to L^2(\mathcal{X}, \mu; \mathbb{R})$ is the Paley–Wiener integral; $I(h)(x)$ is what was called $\langle h, x \rangle^{\sim}$ in statement of the Cameron–Martin theorem.

Definition

The classical Wiener space is the above set-up with

- $\mathcal{X} := C_0$, continuous paths in $\mathbb{R}^n$ starting at 0, with

$$\|x\|_{\mathcal{X}} := \|x\|_\infty = \sup_{t \in [0,T]} |x(t)|;$$

- $\mathcal{H} := L_0^{2,1}$, paths in $\mathbb{R}^n$ starting at 0 with time derivative in L^2, with

$$\langle h, k \rangle_{\mathcal{H}} := \int_0^T \dot{h}(t) \cdot \dot{k}(t) \, \mathrm{d}t;$$

- and the measure is the "standard" Gaussian measure on C_0 — "standard" in the sense that on n-dimensional subspaces of $\mathcal{H}$, it's standard n-dimensional Gaussian measure. We call it Wiener measure, $\mathbb{W}$.

- This gives us a probability space $(\Omega, \mathscr{F}, \mathbb{P}) = (C_0, \mathcal{B}(C_0), \mathbb{W})$.
- The Paley–Wiener integral becomes

$$\langle h, x \rangle^{\sim} = \int_0^T \dot{h}(t) \cdot \mathrm{d}x(t) = \int_0^T \dot{h}(t) \cdot \dot{x}(t) \, \mathrm{d}t,$$

but it makes sense even if x is just continuous!

- This gives us a probability space $(\Omega, \mathscr{F}, \mathbb{P}) = (C_0, \mathcal{B}(C_0), \mathbb{W})$.
- The Paley–Wiener integral becomes

$$\langle h, x \rangle^{\sim} = \int_0^T \dot{h}(t) \cdot \mathrm{d}x(t) = \int_0^T \dot{h}(t) \cdot \dot{x}(t) \, \mathrm{d}t,$$

  but it makes sense even if x is just continuous!
- We can consider random variables $X : \Omega \to \mathbb{R}, \mathbb{R}^n, \mathcal{Y} \dots$
- An obvious one to consider is the identity map, known in this case as the canonical process, standard Wiener process or standard Brownian motion

$$W : \Omega \to C_0$$
$$\omega \mapsto \omega.$$

- What are the properties of this path-valued random variable?

Theorem

The standard Wiener process W satisfies

- $W(0) = 0$;
- $\mathbb{W}$-*almost surely,* $t \mapsto W(t)$ *is continuous;*
- $\mathbb{W}$-*almost surely,* $t \mapsto W(t)$ *is nowhere differentiable;*
- *for* $0 \leq s < t \leq T$, *given* $W(s)$, *the increments are normally distributed:*
$$W(t) - W(s) \sim \mathcal{N}(0, t - s);$$
- *for* $0 \leq s < t \leq u < v \leq T$, *the increments* $W(v) - W(u)$ *and* $W(t) - W(s)$ *are independent.*

...which is exactly what Einstein, Wiener *et al.* wanted!

- Consider the Riemann–Stieltjes integral of $f \colon [0, T] \to \mathbb{R}$ with respect to $g \colon [0, T] \to \mathbb{R}$ of class C^1 (or actually of bounded variation):

$$\int_0^T f(t) \, \mathrm{d}g(t) = \int_0^T f(t) \dot{g}(t) \, \mathrm{d}t$$
$$= \lim_{\mathrm{mesh}(\Pi) \to 0} \sum_{i=1}^{k} f(t_i) \big(g(t_{i+1}) - g(t_i) \big)$$

  where the limit is taken as the mesh of the partition $\Pi = \{0 = t_0 < t_1 < \ldots < t_k = T\}$ tends to 0.
- What would this integral be if g were merely continuous?
- This question doesn't really make sense for any particular continuous g, but can be answered in a probabilistic way using standard Gaussian Wiener measure $\mathbb{W}$ on C_0.

- Standard Wiener probability space $(\Omega = C_0, \mathcal{B}(C_0), \mathbb{W})$.
- Standard Wiener process/Brownian motion $W = \mathrm{id} \colon \Omega \to C_0$.
- Another path-valued random variable (stochastic process) $X : \Omega \to C^0([0, T]; \mathbb{R}^n)$.
- A "decent" function $f \colon [0, T] \times \mathbb{R}^n \to \mathbb{R}$.
- The Itō integral

$$\int_0^T f(t, X(t)) \, \mathrm{d}W(t) \colon \Omega \to \mathbb{R}$$

  is also a random variable . . .
- defined by taking the Riemann–Stieltjes integral

$$\int_0^T f(t, X(t)) \, \mathrm{d}W(t) = \lim_{\mathrm{mesh}(\Pi) \to 0} \sum_{i=1}^{k} f(t_i, X(t_i)) \big(W(t_{i+1}) - W(t_i) \big)$$

  but with two important caveats.

$$\int_0^T f(t, X(t))\,\mathrm{d}W(t) = \lim_{\mathrm{mesh}(\Pi)\to 0} \sum_{i=1}^{k} f(t_i, X(t_i))\big(W(t_{i+1}) - W(t_i)\big)$$

- This limit cannot be taken "pointwise" in Ω; it has to be taken as a limit in $L^2(\Omega, \mathbb{W}; \mathbb{R})$. In particular, it is only determined up to sets of measure zero.

$$\int_0^T f(t, X(t))\,\mathrm{d}W(t) = \lim_{\mathrm{mesh}(\Pi)\to 0} \sum_{i=1}^{k} f(t_i, X(t_i))\big(W(t_{i+1}) - W(t_i)\big)$$

- This limit cannot be taken "pointwise" in Ω; it has to be taken as a limit in $L^2(\Omega, \mathbb{W}; \mathbb{R})$. In particular, it is only determined up to sets of measure zero.
- If you replace $f(t_i, X(t_i))$, i.e. evaluation at the left of each subinterval, with

$$f\left(\frac{t_i + t_{i+1}}{2}, X\left(\frac{t_i + t_{i+1}}{2}\right)\right),$$

i.e. evaluation at the mid-point of each subinterval, you get a different limit! This is the Stratonovich integral, often denoted

$$\int_0^T f(t, X(t)) \circ \mathrm{d}W(t).$$

This allows us to write down stochastic differential equations, i.e. "differential equations with randomness" in a rigorous way as stochastic *integral* equations. For example, the Itō SDE

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = b(t, X(t)) + \sigma(t, X(t)) \frac{\mathrm{d}W(t)}{\mathrm{d}t}^{(!!!)}$$

This allows us to write down stochastic differential equations, i.e. “differential equations with randomness” in a rigorous way as stochastic *integral* equations. For example, the Itō SDE

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = b(t, X(t)) + \sigma(t, X(t)) \frac{\mathrm{d}W(t)}{\mathrm{d}t}^{(!!!)}$$

means

$$X(t) = X(0) + \int_0^t b(s, X(s))\,\mathrm{d}s + \int_0^t \sigma(s, X(s))\,\mathrm{d}W(s).$$

Happily, the same “Lipschitz coefficients $\implies$ existence and uniqueness of solutions” arguments apply. We can also study the push-forward measure $X_*\mathbb{W}$ on $C^0([0,T];\mathbb{R}^n)$; in nice cases (Girsanov), it's a new Gaussian measure. One could also ask about the smoothness properties of the distribution of $X(t)$ — this leads to Malliavin calculus.