FinM 345/Stat 390 Stochastic Calculus, Autumn 2009

Floyd B. Hanson, Visiting Professor

Email: fhanson@uchicago.edu

Master of Science in Financial Mathematics Program
University of Chicago

Lecture 2 (Corrected October 9, 2009)

Diffusion & Jump Stochastic Integration

6:30-9:30 pm, 05 October 2009, Kent 120 in Chicago

7:30-10:30 pm, 05 October 2009 at UBS Stamford

7:30-10:30 am, 06 October 2009 at Spring in Singapore

Continuation of Lecture 1:
1.4. Time-Dependent (NonHomogeneous) Poisson Process:

- Financial markets are very time-dependent, so modelers need to think critically about constant coefficient models, understanding that in some cases time-dependence of coefficients may be difficult to estimate, but perhaps not much more difficult to analyze. Thus, consider $\lambda = \lambda(t)$ so the Poisson process $\mathbf{P(t)}$ will be **nonstationary**.
- Thus, the Poisson parameter differential is $d\Lambda(t) \equiv \lambda(t)dt$, while the integral parameter, assuming $\Lambda(0) = 0$ as in the constant jump rate case, is

$$\Lambda(t) = \int_0^t \lambda(s)ds.$$

- Then, the Poisson parameter increment is defined by

$$\Delta\Lambda(t) \equiv \Lambda(t+\Delta t) - \Lambda(t) = \int_t^{t+\Delta t} \lambda(s)ds.$$

Thus, $\Delta\Lambda(t) \sim \lambda(t)\Delta t$ only when $\Delta t \ll 1$, i.e., is small, but if not the integral must be used.

- The temporal Poisson distributions $\mathrm{Prob}[dP(t)=k] = p_k(\Lambda_{[1:3]}(t))$ for the three cases $\Delta P_{[1:3]}(t) = [dP(t), \Delta P(t), P(t)]$ and parameters $\Delta\Lambda_{[1:3]}(t) = [d\Lambda(t), \Delta\Lambda(t), \Lambda(t)]$, are the same

$$\Phi_{\Delta P_i(t)}(k; \Delta\Lambda_i(t)) = e^{-\Delta\Lambda_i(t)} \frac{(\Delta\Lambda_i(t))^k}{k!},$$

for **i = 1:3** and **k = 0, 1, 2, . . .** jumps, $t \geq 0$ and $\Delta t \geq 0$. (**Hint: In MATLAB, $1:n = [j]_{1\times n}$ is a row-vector.**)

- Note that all three Poisson processes are increment processes, even $\Delta P_3(t) = P(t) = P(t) - P(0)$, where $P(0) \equiv 0$. Also, $\Lambda(t)$ is continuous as integrals with $\lambda(t) > 0$ for $t > 0$.
- While the basic statistics for the set of Poisson *increment* processes are similar to the simple constant rate case, i.e., $\mathrm{E}[\Delta P_i(t)] = \Delta\Lambda_i(t) = \mathrm{Var}[\Delta P_i(t)]$. However, the exponential distribution of the interjump times are much more complicated, but see Hanson's (2007), pp. 22-23, and cited background references.
- Some theory becomes more complicated in the nonstationary case, but changing the *clock* from $\mathbf{t}$ to $\Lambda(t)$, by changing the rate to the constant one, will remove most of the difficulties (see text, p. 22).

1.5. Martingale Properties of Markov Processes — Expectations Conditioned on the Past:

- **Simple Definition 1.7:** *A **martingale M(t)** is a stochastic process that principally satisfies*

$$\mathrm{E}[M(t) \mid M(s), 0 \le s < t] = M(s),$$

*with some technical side conditions in probability space that **M(t) is absolutely integrable**, i.e.,* $\mathrm{E}[|M(t)|] < \infty$ *on **[0,T]** for some finite horizon time* $T < \infty$.
(Comment: The term Martingale comes from horse racing and abstractly symbolizes a **fair game** since

$$E[M(t) - M(s) \mid M(s)] = 0, \quad 0 \le s < t,$$

i.e., there being no net gain on the average conditioned on past data. Alternately, $E[\Delta M(t) \mid M(t)] = 0, t \ge 0$.)

- **Poisson Examples** (assuming $0 \le s < t$):

  1. Expanding in increments, $\mathrm{E}[P(t)|P(s)] = \mathrm{E}[(P(t) - P(s)) + P(s)|P(s)] = \Lambda(t;s) + P(s)$, where $\mathrm{E}[P(t) - P(s)|P(s)] = \Lambda(t;s) \equiv \Lambda(t) - \Lambda(s)$ so $P(t)$ is **not a martingale**, but the **zero-mean Poisson**, $\widehat{P}(t) \equiv P(t) - \Lambda(t)$ is a martingale, because $\mathrm{E}[\widehat{P}(t)|\widehat{P}(s)] = \widehat{P}(s)$; hence $\mathrm{E}[\Delta\widehat{P}(t)|\widehat{P}(t)] = 0$, so implies a fair game.

  2. Again expanding, $\mathrm{E}[P^2(t)|P^2(s)] = \mathrm{E}[((P(t) - P(s) - \Lambda(t;s)) + (P(s) + \Lambda(t;s)))^2|P(s)] = \mathrm{Var}[P(t)|P(s)] + 2 \cdot 0 + (P(s) + \Lambda(t;s))^2 = P^2(s) + \Lambda(t;s)(2P(s) + \Lambda(t;s) + 1)$, so $P^2(t)$ **cannot be converted into a martingale** since the cross-term $\Lambda(t;s)(2P(s) + \Lambda(t;s) + 1)$ prevents additive separability into **t** and **s** terms.

- **Wiener Examples** (assuming $0 \leq s < t$):

  1. Since the Wiener process is a zero mean process **W(t) is a martingale**, i.e., $\mathrm{E}[W(t)|W(s)] = \mathrm{E}[(W(t) - W(s)) + W(s)|W(s)] = W(s)$ and $\mathrm{E}[\Delta W(t)|W(t)] = 0$ implies a fair game. (*Comment: Zero-meanness helps, but is not sufficient in general. Note also that* $\mathrm{E}[|W(t)|] = \sqrt{2t/\pi} < \sqrt{2T/\pi} < \infty$ *by Table 1.1*.)

  2. Expanding, $\mathrm{E}[W^2(t)|W(s)] = \mathrm{E}[((W(t) - W(s)) + W(s))^2|W(s)] = (t - s) - 2W(s) \cdot 0 + W^2(s)$, so rearranging we see that $(W^2(t) - t)$ **is a martingale**, $0 \leq t < T < \infty$, since $\mathrm{E}[W^2(t) - t|W(s)] = W^2(s) - s$. (**Caution: Note no time-dependent coefficients.**)

FinM 345 Stochastic Calculus:

2. Stochastic Integration for Stochastic Differential Equations:

2.0. Somewhat General Jump-Diffusion Stochastic Differential Equation (SDE):

Our interest in stochastic calculus is the integration or numerical simulation of given jump-diffusion stochastic differential equations of a type like

$$dX(t) = f(X(t), t)dt + g(X(t), t)dW(t) + h(X(t), t)dP(t),$$

$X(0) = x_0$ initially with probability one, with stochastic jump-diffusion terms like differential Wiener process $dW(t)$ and a simple Poisson differential process $dP(t)$ here.

The coefficients $\{f(x,t), g(x,t), h(x,t)\}$ could be nonlinear in the state $X(t)$, e.g., an asset, but most likely to be linear or affine in finance, e.g., $f(x,t) = f_0(t) + f_1(t)x$. More generally and usefully, the Poisson process will be a compound Poisson process $dP(t) = dP(t;Q)$, doubly stochastic with a independent, identically distributed (IID) amplitude variable Q and will be treated later in the course.

Integration of SDEs is not the same as those for ODEs (else there would not be this course) due to the facts that $W(t)$ is **nondifferentiable** and $P(t)$ is **discontinuous** (but piecewise continuous) and because of this the above SDE is really ***symbolic*** since the integrated or simulated solution depends on the integration rule, which must be specified.

2.1. Stochastic Integration for Diffusions:

We will follow the regular calculus model, but modified for jumps and nondifferentiability in a nonabstract way. Unlike regular calculus where the integral in independent of the approximation point as in the **Riemann integral**,

$$F(t) = \int_0^t f(X(s),s)ds \simeq \sum_{i=0}^{n} f(X(t_{i+\theta}),t_{i+\theta})\Delta t_i,$$

which converges to $F(t)$ as $n \to \infty$ for sufficiently nice $f(X(t),t)$, where $t_i = t_0 + i\Delta t_i$, $t_{i+\theta} = t_i + \theta\Delta t_i$, $t_0 = 0$, $t_{n+1} = t$ and θ is the **fractional approximation point spacing**, $0 \le \theta \le 1$. In the case of **fixed** $(n+1)$ **time-steps**, $\Delta t_i = t/(n+1)$, so $\sum_{i=0}^{n} \Delta t_i = t$.

Of more interest in stochastic calculus is the **Stieltjes or Riemann-Stieltjes integral**,

$$G(t) = \int_0^t g(X(s), s)dX(s) \simeq \sum_{i=0}^{n} g(X(t_{i+\theta}), t_{i+\theta})\Delta\mathrm{X}(t_i),$$

which converges to $G(t)$ as $n \longrightarrow \infty$ for sufficiently nice $g(X(t), t)$ and continuous $X(t)$, having bounded variation, where $\Delta\mathrm{X}(t) \equiv X(t_{i+1}) - X(t_i)$. However, one can show that $W(t)$ does not have bounded variation anywhere.

The primary stochastic calculus used today is the Itô calculus developed for the Wiener process in the 1940s by **Kyosi Itô** (d. 11/10/2008) apparently in his PhD thesis research, but he also extended the calculus to Poisson processes in his 1954 memoir. Although also known for his abstract analysis on stochastic processes. For real application simulation or integration purposes, his emphasis on the **forward or Euler or left-hand integration rule**, i.e., $\theta \equiv 0$, is of interest since the forward rule is consistent with the Markov properties of forgetting the past and independent increments, as in the diffusion and jump processes considered, but the stochastic property requires at least mean square convergence. When $\theta \neq 0$, then integrals vary with θ and calculations are very complicated since they must be split into many combinations of independent increments for calculations. . . .

. . . *(See Itô appreciation Homework #4 $\smile$.)*

For integration of a derivative, the fundamental theorem of calculus holds for $W(t)$, so $\int_0^t dW(s) = W(t)$ or $\int_0^t dG(W(s)) = G(W(t)) - G(0)$. For general Wiener integration,

$$I[W dW](t) \equiv \int_0^t W(s) dW(s)$$

is the fundamental "counter-example" to regular integration, e.g.,

$$\int_0^t X(s) dX(s) = \int_0^t dX^2(s)/2 = (X^2(t) - X^2(0))/2,$$

if $X(t)$ is differentiable.

The Itô forward approximation **(IFA) when $\theta \equiv 0$** is

$$I_n[WdW](t) \equiv \sum_{i=0}^{n} W(t_i)\Delta W(t_i),$$

noting that $W(t_i) = W(t_i) - W(0)$ and $\Delta W(t_i) = W(t_{i+1}) - W(t_i)$ are independent increments, preserving a crucial Markov property. Hence, since for each increment $\mathrm{E}[\Delta W(t)] = 0$, then by independence and zero-mean properties, $\mathrm{E}[I_n[WdW](t)] = \sum_{i=0}^{n} 0 \cdot 0 = 0$.

In the following, we simplify the index notation by letting $W_i \equiv W(t_i)$ and $\Delta W_i \equiv \Delta W(t_i)$.

For approximating and resumming, we need the following from Hanson (2007),

- **Lemma 2.1.1. Critical Sum Identities:** *Let* $\{x_i | i = 0 : n+1\}$ *be any sequence of numbers, and let* $\Delta \mathrm{x}_i = x_{i+1} - x_i$ *for* $i = 0 : n$*. Then*

$$\sum_{i=0}^{n} \Delta \mathrm{x}_i = x_{n+1} - x_0,$$

$$\sum_{i=0}^{n} x_i \Delta \mathrm{x}_i = \frac{1}{2}\left(x_{n+1}^2 - x_0^2 - \sum_{i=0}^{n} (\Delta \mathrm{x}_i)^2\right).$$

Proof: *See the text, p. 35, but first part is trivial and is a generalization of the identity* $\sum_{i=0}^{n} \Delta t_i = t$ *for partitions of* $[0, t]$.

As a deterministic example, in the case of a variable mesh, Δt_i is not constant, but then a maximal mesh size must be specified, $0 < \Delta t_i \le \delta t_n \equiv \max_i[\Delta t_i]$, such that $\delta t_n \to 0^+$ to avoid degeneracy as $n \to +\infty$, along with $\sum_{i=0}^{n} \Delta t_i = t$ **fixed**. Using the forward approximation $(\theta \equiv 0)$ on elementary integrals,

$$\int_0^t ds = \sum_{i=0}^{n} \Delta t_i = t, \text{ exactly;}$$

$$\int_0^t (ds)^2 \simeq \sum_{i=0}^{n} (\Delta t_i)^2 \le \delta t_n \sum_{i=0}^{n} \Delta t_i = t \cdot \delta t_n \to 0^+;$$

$$\int_0^t (ds)^{1+\alpha} \simeq \sum_{i=0}^{n} (\Delta t_i)^{1+\alpha} \le (\delta t_n)^{\alpha} \sum_{i=0}^{n} \Delta t_i = t \cdot (\delta t_n)^{\alpha} \to 0^+,$$

where $0 < \alpha$, reaffirming the notion of **dt-precision**.

Thus,

$$I_n[W dW](t) = \frac{1}{2}\left(W^2(t) - \sum_{i=0}^{n}(\Delta W_i)^2\right).$$

So, using **Table 1.1** from Lecture 1 for the mean of the squared Wiener increment,

$$\mathrm{E}[I_n[W dW](t)] = \frac{1}{2}\left(t - \sum_{i=0}^{n}\Delta t\right) = (t - t)/2 = 0$$

preserving the original zero mean but suggesting that

$$I[W dW](t) \approx \frac{1}{2}\left(W^2(t) - t\right),$$

but the real target and more fundamental integral is

$$I[(dW)^2](t) = \int_0^t (dW)^2(s) \approx t,$$

but note that the expectation above is no indicator of equality.

You should have demonstrated the t limit for $I[(dW)^2](t)$ in Homework Set 1, using the Itô forward approximation,

$$I_n[(dW)^2](t) \equiv \sum_{i=0}^{n} (\Delta W_i)^2,$$

for simulation. In order to establish a weak version of equality, **Itô's mean square (IMS) limit** will be used:

- **Definition 2.1.1. Mean Square Limit or Convergence:** *The random variable* $I_n(t)$ ***converges in the mean square*** *to the random variable* $I(t)$ *if*

$$\mathrm{E}\left[(I_n(t) - I(t))^2\right] \to 0 \qquad (1)$$

as $n \to \infty$*, assuming that both random variables have bounded mean squares, i.e.,* $\mathrm{E}\left[(I_n)^2(t)\right] < \infty$ *and* $\mathrm{E}\left[I^2(t)\right] < \infty$*. If the limit* (1) *exists, then denote the* ***mean square limit*** *as*

$$I(t) = \underset{n\to\infty}{\overset{\text{ms}}{\lim}}[I_n(t)].$$

As an abbreviation, sometimes $\overset{\text{ims}}{=}$ *will be used for* $= \underset{n\to\infty}{\overset{\text{ms}}{\lim}}$*, where* $\overset{\text{ims}}{=}$ *means* ***equals in the Itô mean square****.*

(Comment: Mean square convergence is related to other types of convergence in Hanson (2007), p. 37-38.)

- **Theorem 2.1.1:** *Let*

$$I_n\left[(dW)^2\right](t) \equiv \sum_{i=0}^{n}(\Delta W_i)^2; \quad (2)$$

then

$$t = \underset{n\to\infty}{\overset{\text{ms}}{\lim}}\left[I_n\left[(dW)^2\right](t)\right] \overset{\text{ims}}{=} I\left[(dW)^2\right](t). \quad (3)$$

Proof: The mean t of $I_n[(dW)^2](t)$ is absorbed into the summation by the critical sums with $x_i = t_i$; the square of the mean square argument leads to a double sum which is separated **by the diagonalization technique** into diagonal parts ($j = i$) and off-diagonal parts ($j \neq i$), allowing the splitting of the expectations using the independent increment property, so

$$\mathrm{E}\left[\left(I_n[(dW)^2](t) - t\right)^2\right]$$
$$= \mathrm{Var}\left[I_n[(dW)^2](t)\right] = \mathrm{E}\left[\left(\sum_{i=0}^{n}(\Delta W_i)^2 - t\right)^2\right]$$
$$= \mathrm{E}\left[\left(\sum_{i=0}^{n}\left((\Delta W_i)^2 - \Delta t_i\right)\right)^2\right]$$
$$= \mathrm{E}\left[\sum_{i=0}^{n}\left((\Delta W_i)^2 - \Delta t_i\right)\sum_{j=0}^{n}\left((\Delta W_j)^2 - \Delta t_j\right)\right]$$
$$= \sum_{i=0}^{n}\mathrm{E}\left[\left((\Delta W_i)^2 - \Delta t_i\right)^2\right]$$
$$+ \sum_{i=0}^{n}\mathrm{E}\left[(\Delta W_i)^2 - \Delta t_i\right]\sum_{\substack{j=0\\ j\neq i}}^{n}\mathrm{E}\left[(\Delta W_j)^2 - \Delta t_j\right]$$
$$= \sum_{i=0}^{n}\mathrm{Var}\left[(\Delta W_i)^2\right] + \sum_{i=0}^{n} 0 \sum_{\substack{j=0\\ j\neq i}}^{n} 0$$
$$= \sum_{i=0}^{n}\left(\mathrm{E}\left[(\Delta W_i)^4\right] - E^2\left[(\Delta W_i)^2\right]\right)$$
$$= \sum_{i=0}^{n}\left(3(\Delta t_i)^2 - (\Delta t_i)^2\right) = 2\sum_{i=0}^{n}(\Delta t_i)^2.$$

(Comment: Note you need that $(\sum_i A_i)^2 = (\sum_i A_i)(\sum_j A_j)$, where j is a dummy index to avoid confusion with the original index j; try for $i = 0:1$.)

The last two of steps relying on the results of **Table 1.1** of Lecture 1. Since a variable grid must have a maximal constraint, $\Delta t_i \le \delta t_n = \max_j[\Delta t_j]$, we then have

$$\mathrm{E}\left[\left(I_n[(dW)^2](t) - t\right)^2\right] = 2\sum_{i=0}^{n}(\Delta t_i)^2 \le 2\delta t_n \sum_{i=0}^{n}\Delta t_i$$
$$= 2t\delta t_n \to 0$$

as $n \to \infty$ showing that

$$t = \underset{n\to\infty}{\overset{\text{ms}}{\lim}}\,[I_n[(dW)^2](t)].$$

Clearly both $I_n[(dW)^2](t)$ and t have bounded mean squares for bounded t.

Summarizing, we have

- **Theorem 2.1.2:**

$$\int_0^t W(s)dW(s) \stackrel{\text{ims}}{=} \frac{1}{2}\left(W^2(t)-t\right) = \underset{n\to\infty}{\overset{\text{ms}}{\lim}}[I_n[WdW](t)], \qquad (4)$$

where $t < \infty$ *and*

$$I_n[WdW](t) = \sum_{i=0}^{n} W_i \Delta W_i.$$

(Comment: See Hanson (2007), pp. 39-42, for more proof details. ***Such discrete analysis is usually only needed at the start and, as in the regular calculus for the most basic results, once we get to the chain rule, here for stochastic processes, we will not use this kind of analysis again, until there is something with complexity beyond the chain rule.***)

- **Definition 2.1.2. General Itô Mean Square (IMS) Limit:**

Let

$$I[g(W,t)dW](t) = \int_{t_0}^{t} g(W(s),s)dW(s), \qquad (5)$$

where $0 \le t_0 \le t$ *and the integrand process* $g(W(t),t)$ *has a bounded mean integral of its square, i.e.,*

$$\mathrm{E}\left[\int_{t_0}^{t} g^2(W(s),s)ds\right] < \infty.$$

Further, let the ***forward integration approximation*** *be, with mesh* $\{t_{i+1}=t_i+\Delta t_i, t_{n+1}=t, \delta t_n=\max_i[\Delta t_i]\ll 1, n\gg 1\}$,

$$I_n[g(W,t)dW](t) \stackrel{\text{ifa}}{=} \sum_{i=0}^{n} g(W(t_i),t_i)\Delta W(t_i)),$$

then the ***IMS limit*** *of (5) is*

$$I[g(W,t)dW](t) \stackrel{\text{ims}}{=} \underset{n\to\infty}{\overset{\text{ms}}{\lim}} [I_n[g(W,t)dW](t)]. \qquad (6)$$

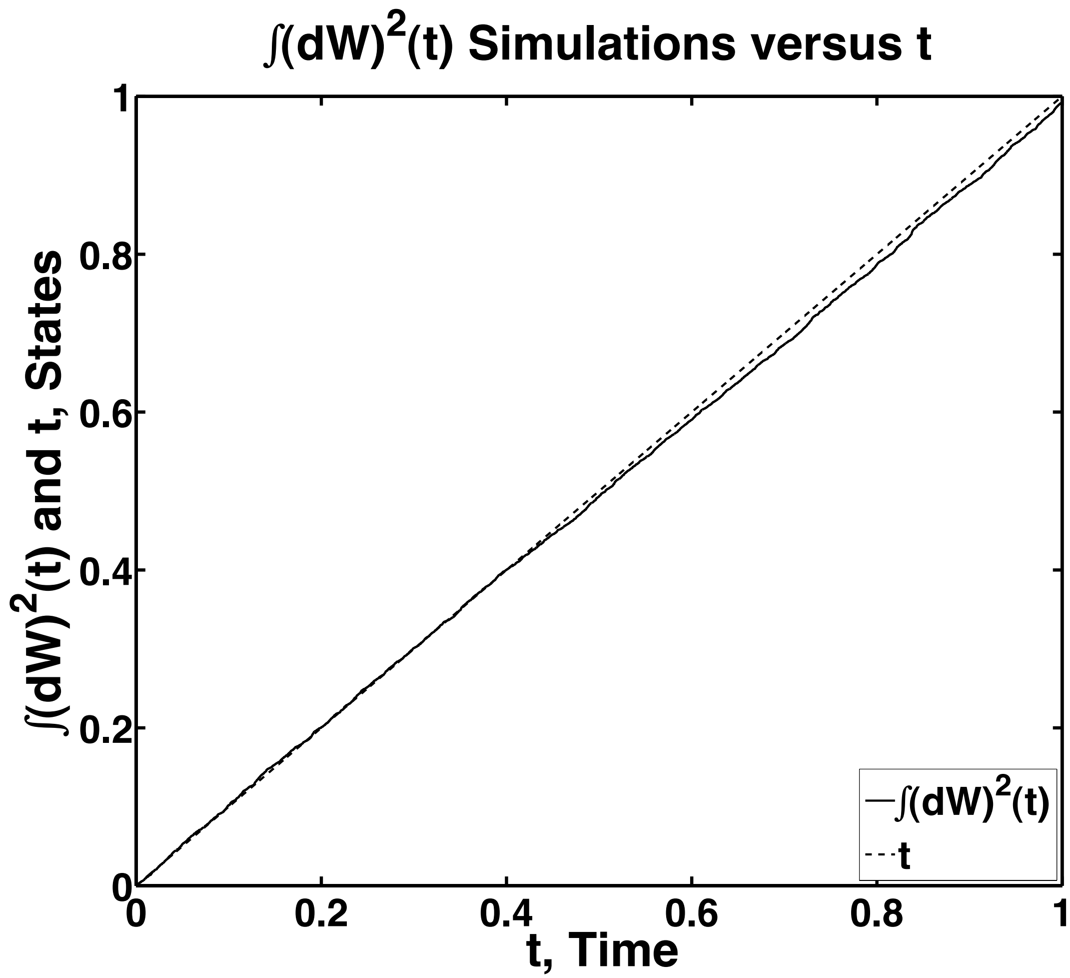

Figure 1: Simulated sample path for the Itô forward integration approximating sum of $\int(dW)^2(t) \stackrel{\text{ims}}{=} t \simeq \sum_i (\Delta W_i)^2$ for $n = 10^4$ MATLAB `randn` sample size. ***Note, the IMS limit t is deterministic: a paradox?***

- **Fundamental Stochastic Calculus Motivation 2.1.1.**

$$\int_0^t (dW)^2(t) \approx \lim_{n\to\infty} \sum_{i=0}^{n} (\Delta W_i)^2 = t:$$

Consider $\Delta W_i \stackrel{\text{dist}}{=} \mathcal{N}(w; 0, \Delta t)$, *but the short derivation on L1-p25 implies there is a* $Z_i = Z_i(t_i) \stackrel{\text{dist}}{=} \mathcal{N}(w; 0, 1)$ *such that* $\Delta W_i = \sqrt{\Delta t} \cdot Z_i$ *for all* i. *The unbiased sample variance of the* Z_i *is*

$\mathrm{Var}_{n+1}[\{Z_i\}_{i=0}^n] = \sum_{i=0}^n Z_i^2/n \to \sigma_Z^2$ *as* $n \to \infty$ *where the population variance is* $\sigma_Z^2 = 1$ *according to the statistical limit theorems in* ***FINM 331***. *Hence,*

$\sum_{i=0}^n (\Delta W_i)^2 = \Delta t \sum_{i=0}^n Z_i^2 = n \cdot t \cdot \mathrm{Var}_{n+1}[\{Z_i\}_{i=0}^n]/(n+1) \to t$ *as* $n \to \infty$, *i.e., the desired result due to the unit-variance of the underlying standard normal* Z & $\Delta t = t/(n+1)$.

Code follows later . . . intdwdw.m in Online Appendix C

- $\sum_i(\Delta W_i)^2$ ***and t book MATLAB code example***:

```
function intdwdw
% Fig. 2.1 Example MATLAB code for integral of (dW)^2:
clc % clear variables;
t0 = 0.0; tf = 1.0;
n = 1.0e+4; nf = n + 1; % set time grid: (n+1) steps;
dt = (tf-t0)/nf;  %        and (n+2) points;
sqrtdt = sqrt(dt); % dW(i) noise time scale so E[dW]=0;
kstate = 1; randn('state',kstate); % Fix randn state;
dW = sqrtdt*randn(nf,1); % simulate (n+1)-dW(i) sample;
t = t0:dt:tf; % get time vector t;
W = zeros(1,nf+1); % set initial diffusion condition;
sumdw2 = zeros(1,nf+1); % set initial integral sum;
for i = 1:nf % simulate integral sample path;
    W(i+1) = W(i) + dW(i); % sum diffusion noise;
    sumdw2(i+1) = sumdw2(i) + (dW(i))^2; % sum;
end % Better to use cumsum here for vector code;
fprintf('\n\nFigure 1: int[(dW)^2](t) versus t\n');
```

```
nfig = 1;
figure(nfig);
scrsize = get(0,'ScreenSize'); % fig. spacing for screen;
ss = [5.0,4.0,3.5]; % figure spacing factors
plot(t,sumdw2,'k-',t,t,'k--','LineWidth',2); % 2 plots;
title('\int(dW)^2(t) Simulations versus t'...
    ,'FontWeight','Bold','Fontsize',44);
ylabel('\int(dW)^2(t) and t, States'...
    ,'FontWeight','Bold','Fontsize',44);
xlabel('t, Time'...
    ,'FontWeight','Bold','Fontsize',44);
hlegend=legend('\int(dW)^2(t)','t'...
    ,'Location','Southeast');
set(hlegend,'Fontsize',36,'FontWeight','Bold');
set(gca,'Fontsize',36,'FontWeight','Bold','linewidth',3);
set(gcf,'Color','White','Position' ...
,[scrsize(3)/ss(nfig) 60 scrsize(3)*0.6 scrsize(4)*0.8]);
% End intdwdw Code
```

• Theorem 2.1.3. Fundamental Theorem of Itô Stochastic Diffusion Calculus.

Let $g(w)$ *be continuous and* $G(w)$ *be continuously differentiable, for* $0 \le t$*, then*

(a) $d\left[\int_0^t g(W(s))dW(s)\right] \overset{\text{ims}}{=} g(W(t))dW(t)$;

(b) $\int_0^t dG(W(s)) \overset{\text{ims}}{=} G(W(t)) - G(0)$.

(Comment: The proofs rely on straight-forward applications of the IFA and IMS principles. See Hanson (2007), pp. 45-46. Of course, the lower limit at $s = 0$ can be replaced by some t_0 such that $0 \le t_0 \le t$. ***The main point is that the Fundamental Theorem of Regular Calculus is preserved with IMS limits.***)

- **Symbolic Notation — Itô Mean Square in dt−Precision:**

Taking the mean square source as understood, we will use the previous symbol $\stackrel{\rm dt}{=}$ instead of the current symbol $\stackrel{\rm ims}{=}$, so using the first part of the Fundamental Theorem 2.1.3, to write our current results as stochastic differentials,

$$(dW)^2(t) \stackrel{\rm dt}{=} dt,$$

$$W(t)dW(t) \stackrel{\rm dt}{=} \frac{1}{2}(d(W^2)(t) - dt)$$

or

$$d(W^2)(t) \stackrel{\rm dt}{=} 2W(t)dW(t) + dt.$$

- **Quick dt–Precision Calculations by Increment Expansion:** Note that the differential increment $dW(t) \equiv W(t+dt) - W(t)$ can also be written $W(t+dt) = W(t) + dW(t)$ for any function, not just $W(t)$, leading to an alternate derivation for $d(W^2)(t)$,

$$\begin{aligned} d(W^2)(t) &\equiv W^2(t+dt) - W^2(t) \\ &= (W + dW)^2(t) - W^2(t) \\ &= (W^2 + 2W\,dW + (dW)^2 - W^2)(t) \\ &\stackrel{\text{dt}}{=} 2W\,dW(t) + dt, \end{aligned}$$

using some easy algebra and the fundamental differential $(dW)^2(t) \stackrel{\text{dt}}{=} dt$ of Wiener processes or Brownian motion.

- **Itô Piecewise-Constant Approximations (i-PWCA) for general Problems:** Associated with Wiener process integrand and the use of the forward integration approximation is the assumption that the integrand $g(W(t),t)$ can be approximated by a **right-continuous set of step-function**, i.e.,

$$g(W(s),s) \simeq Z_n(s) = \{z_i : t_i \le s < t_{i+1}, i = 1:n+1\},$$

where any $t_0 \ge 0$ and $t_{n+1} = t$. (Note that last point would contribute zero area to the integral.) The approximate values z_i depend on a sequence of overlapping, past history processes $\mathcal{W}_i = \{W(s) : t_0 \le s < t_i\}$, i.e., they are non-anticipatory or adapted to $\mathcal{W}_i$, in the sense of abstract probability analysis. The **i-PWCA** $Z_n(s)$ must converge in the mean square to $g(W(s),s)$ as $n \to \infty$. We use $z_i = g(W(t_i),t_i)$, **Itô's forward approximation (IFA)**, denoted in the limit $n \to \infty$ by $\stackrel{\text{ifa}}{=}$. The general **i-PWCA** formulation is advanced background for future work elsewhere, but we mostly rely on **IFA** here.

- **Theorem 2.1.4. Mean of Itô Stochastic Integral:**

$$\mathbf{E}\left[\int_{t_0}^{t} g(W(s),s)dW(s)\right] \overset{\text{ifa}}{=} 0,$$

for $0 \le t_0 \le t$ *, assuming for* g *the mean square integrability condition and the IFA (symbol* $\overset{\text{ifa}}{=}$*) forward approximation limit assumption.*

(Comment: The main idea of the proof is that the expectation can be passed into the integral since formally the domain $[t_0, t)$ of the integral and sum are deterministic, by interchangeability $\mathbf{E}[\sum_i g_i \Delta W_i] = \sum_i \mathbf{E}[g_i \Delta W_i]$, and since g_i and ΔW_i have **"independent increments"**, then $\mathbf{E}[g_i \Delta W_i] = \mathbf{E}[g_i] \cdot \mathbf{E}[\Delta W_i] = 0$, assuming only that $\mathbf{E}[g]$ is bounded, and reassembling by IFA, $\mathbf{E}[\int g \cdot dW] \overset{\text{ifa}}{=} 0$. The theorem is NOT true for approximating rules for $\theta > 0$.)

- **Theorem 2.1.5. Covariance of Itô Stochastic Integral:**

$$\mathrm{E}\left[\int_{t_0}^{t} f(W(s),s)dW(s)\int_{t_0}^{t} g(W(r),r)dW(r)\right] \stackrel{\text{ifa}}{=} \int_{t_0}^{t}\mathrm{E}\left[f(W(s),s)g(W(s),s)\right]ds$$

for $0 \le t_0 \le t$*, assuming that* $f(W(t),t)$ *and* $g(W(t),t)$ *satisfy the mean square integrability condition and the IFA limits assumption.*

(Comment: Extending the formal interchangeability and independent increment ideas from the prior theorem, i.e., $\mathrm{E}[\sum_i f_i \Delta W_i \sum_j g_j \Delta W_j] = \sum_i \sum_j \mathrm{E}[f_i g_j \Delta W_i \Delta W_j]$, and since f_i, g_j,ΔW_i and ΔW_j have independent increments except when $i=j$ giving $\mathrm{E}[\Delta W_i \Delta W_j] = \Delta t \delta_{i,j}$, so $\mathrm{E}[\sum_i f_i \Delta W_i \sum_j g_j \Delta W_j] = \sum_i \mathrm{E}[f_i g_i] \cdot \Delta t$, but reassembling by IFA in reverse yields the above hypothesis.)

An immediate corollory from when $f = g$, called **Itô isometry** or **martingale isometry**, follows:

- **Corollory 2.1.5. Variance of Itô Stochastic Integral:**

$$\mathrm{E}\left[\left(\int_{t_0}^{t} g(W(s),s)dW(s)\right)^2\right] \stackrel{\text{ifa}}{=} \int_{t_0}^{t} \mathrm{E}\left[g^2(W(s),s)\right] ds,$$

under prior assumptions.

(Comment: Since $\int gdW$ has zero-mean by Th. 2.4, then 2nd Moment is the variance, $\mathrm{E}[(\int gdW)^2] = \mathrm{Var}[\int gdW]$. Also, *isometry* is a distance preserving map between metric spaces, according to Wikipedia.)

- **Table 2.1.1. Summary of Itô Stochastic Diffusion Differentials with an Accuracy in dt-precision:**

Differential Diffusion Form	Itô Mean Square Limit
$dW(t)$	$dW(t)$
dt	dt
$dt\, dW(t)$	0
$(dW)^2(t)$	dt
$(dW)^m(t)$	$0, m \geq 3$
$(dt)^\alpha(dW)^m(t)$	$0, \alpha > 0,\; m \geq 1$

Some of these table entries have been the focus of homework simulation problems, but proofs can be difficult. An example is the $(dW)^m(t) \stackrel{\mathrm{dt}}{=} 0$ when $m \geq 3$ when the IFA mean and IMS approximation need be found first, with $\int_0^t (dW)^m(s) \stackrel{\mathrm{ifa}}{=} \sum_{i=0}^n (\Delta W_i)^m$.

When the power is odd, $m = 2k-1,\ k \geq 2$, the **IFA mean test** is $\mathrm{E}[\sum_{i=0}^n (\Delta W_i)^{2k-1}] = \sum_{i=0}^n \mathrm{E}[(\Delta W_i)^{2k-1}] = 0$, exactly, since $(\Delta W_i)^{2k-1}$ is an odd function. The **IMS test** uses the square of the sum that is decomposed as a diagonal and nondiagonal sums ($\sum_j = \sum_{j=i} + \sum_{j \neq i}$),

$$\mathrm{E}[(\sum_{i=0}^n (\Delta W_i)^{2k-1})^2] = \sum_{i=0}^n \sum_{j=0}^n \mathrm{E}[(\Delta W_i)^{2k-1} (\Delta W_j)^{2k-1}]$$
$$= \sum_{i=0}^n \mathrm{E}[\sum_{i=0}^n (\Delta W_i)^{2(2k-1)} + \sum_{j \neq i} (\Delta W_i)^{2k-1} (\Delta W_j)^{2k-1}]$$
$$= (4k-3)!! \sum_{i=0}^n (\Delta t_i)^{2k-1} \leq (4k-3)!! \delta t_n^{2k-2} t \to 0,$$

as $n \to \infty$ and $\sum_i \Delta t = t$ where $(4k-3)!!$ is from Table 1.1.

When the power is even, $m = 2k,\ k \geq 2$, the **IFA mean test** is different due to even function, $\mathrm{E}[\sum_{i=0}^{n}(\Delta W_i)^{2k}]$ $=\sum_{i=0}^{n}\mathrm{E}[(\Delta W_i)^{2k}]=(4k-3)!!\sum_{i=0}^{n}(\Delta t_i)^{k}\leq\delta t_n^{k-1}t\to 0$, approximately. The **IMS test** is

$$\mathrm{E}[(\textstyle\sum_{i=0}^{n}(\Delta W_i)^{2k})^2] = \sum_{i=0}^{n}\sum_{j=0}^{n}\mathrm{E}[(\Delta W_i)^{2k}(\Delta W_j)^{2k}]$$
$$= \textstyle\sum_{i=0}^{n}\mathrm{E}[\sum_{i=0}^{n}(\Delta W_i)^{4k} + \sum_{j\neq i}(\Delta W_i)^{2k}(\Delta W_j)^{2k}]$$
$$= (4k-1)!!\textstyle\sum_{i=0}^{n}(\Delta t_i)^{2k}+((2k-1)!!)^2\sum_{i=0}^{n}(\Delta t_i)^{k}\sum_{j=0}^{n}(\Delta t_j)^{k}$$
$$\leq (4k-1)!!\delta t_n^{2k-1}t+((2k-1)!!)^2\delta t_n^{2k-2}t^2\to 0$$

as $n \to \infty$ after much more algebra and analysis **(see text)**.

Hence,

$$\int_0^t (dW)^m(s) \stackrel{\text{ims}}{=} 0 \text{ and } (dW)^m(t) \stackrel{\text{dt}}{=} 0$$

for $m \geq 3$, the latter using the fundamental theorem of differential stochastic calculus on the former integral.

2.2. Stochastic Integration for Jumps:

Usually a **Itô process**, also called a **Gaussian process** is a process linear in the **Wiener process** with a mean and volatility. However, in K. Itô's early published Memoir of the American Mathematics Society on stochastic differential equations, he also treated the case of **Poisson jumps**, so perhaps those who named the diffusion process for him did not read that far into his work.

- **Definition 2.2.1. Poisson Jump Stochastic Integration.**

$$\int_0^t h(X(s),s)dP(s) \stackrel{\text{ims}}{=} \underset{n\to\infty}{\overset{\text{ms}}{\lim}} \left[\sum_{i=0}^{n} h(X(t_i),t_i)\Delta P(t_i)\right],$$

where $X(t)$ *in the integrand function* h *depends on the jump process* $P(t)$*, but also can depend on the diffusion process* $W(t)$*. The integrand process* $h(X(t),t)$ *is also assumed to have a* ***bounded mean integral of squares****,*

$$\mathrm{E}\left[\int_0^t h^2(X(s),s)ds\right] < \infty,$$

and $h(X(t),t)$ *to satisfy the Itô piecewise constant approximations (i-PWCA) the mean square limits assumption in the sense of the* ***Itô's forward approximation (IFA)****, with a proper variable grid (*Δt_i*) partition specifications on* $[0,t]$.

(Comment: In this section, we will little time on mean square convergence, since a sufficient amount of time was spent in the last section for a course in financial applications.)

- **Theorem 2.2.1. Fundamental Theorem of Poisson Jump Calculus (FTPJC):**

Let $h(p)$ *be continuous and* $\mathcal{H}(p)$ *be continuously differentiable. Then*

(a) $d\left(\int_0^t h(P(s))dP(s)\right) \stackrel{\text{ims}}{=} h(P(t))dP(t),\ 0 \le t$,

(b) $\int_0^t d\mathcal{H}(P(s)) \stackrel{\text{ims}}{=} \mathcal{H}(P(t)) - \mathcal{H}(0),\ 0 \le t$.

The proof is very similar to that of the Wiener process in the prior section, except that bounded variation is not needed for Poisson jumps.

- **Theorem 2.2.2. Jump Integral of $\int PdP$:**

$$I[P](t) = \int_0^t P(s)dP(s) \stackrel{\text{ims}}{=} I^{(\text{ims})}[P](t) \equiv \frac{1}{2}(P(P-1))(t)$$

is the ***mean square limit integral****, with*

$$I^{(\text{ims})}[P](t) \stackrel{\text{ims}}{=} \underset{n\to\infty}{\overset{\text{ms}}{\lim}} [I_n[PdP](t)],$$

where the Itô forward integration approximation (IFA) is $I_n[PdP](t) \equiv \sum_{i=0}^{n} P(t_i)\Delta P(t_i)$.

Proof, Formally by Increments: Recalling that $d(x^2) = 2xdx$ in regular (smooth) calculus, consider the Poisson squared increment expansion

$$\Delta(P^2)(t) \equiv P^2(t+\Delta t) - P^2(t) = ((P+\Delta P)^2 - P^2)(t)$$
$$= (2P\Delta P + (\Delta P)^2)(t).$$

Taking the limit $\Delta t \to 0^+$ with $\Delta t \to dt$, replacing ΔP by dP and $(dP)^2 \stackrel{\text{dt}}{=} dP$ using the **zero-one jump law**, neglecting smaller order terms leads to $d(P^2)(t) \stackrel{\text{dt}}{=} (2PdP + dP)(t)$, in probability. Solving for the integrand-differential while forming an exact differential yields in probability

$$(PdP)(t) \stackrel{\text{dt}}{=} \frac{1}{2} d\left(P^2 - P\right)(t).$$

Therefore, integration by the fundamental theorem of stochastic jump integration leads formally to the primary result,

$$\int_0^t (PdP)(s) = \frac{1}{2}\int_0^t (d\left(P^2 - P\right)(t))(s) \stackrel{\text{ims}}{=} \frac{1}{2}\left(P^2 - P\right)(t),$$

where the initial Poisson condition $P(0) = 0$ **w.p.o** has been used to eliminate the initial value of the integral.

That takes care of the first part of the proof, but the technique is general enough for other powers.

The details of the proof of mean square convergence is given in Hanson (2007) and is too much to repeat here, even though it is only valid when the jump rate λ is constant.

(Comment: Recall that a Poisson process is a counting process and the Pythagorean counting theorem says that the sum of integers is a **triangular number**, i.e., $S_n^{(1)} = \sum_{k=0}^{n} k = n(n+1)/2$, so the primary result $P(P-1)/2$ is the $n = P(t)-1$th triangular number, such that the number is short of $P(t)$, but the sum is zero when $P(t) = 1$ due to the forward nature of IFA, using only the left-hand endpoints of each subinterval.)

- **Table 2.2.1. Some stochastic jump integrals of powers with an accuracy with error $o(dt)$ as $dt \to 0^+$:**

m	**precision-dt**: $\int_0^t (P^m dP)(s)$
0	$P(t)$
1	$(P(P-1))(t)/2$
2	$(P(P-1)(2P-1))(t)/6$
3	$(P^2(P-1)^2)(t)/4$

(Comment: The cases $m=2$ & 3 will be proved formally for homework. These cases are all cases of **supertriangular numbers** (Hanson, 2007), $S_n^{(m)} = \sum_{k=0}^{n} k^m$, where $m \leq 0$ & $n \leq 0$, again with $n = P(t) - 1$.)

If the sum form of Poisson integral holds for monomials, it should hold for homogeneous functions of $P(t)$, only.

- **Theorem 2.2.3. Pure Poisson Integral as Sum Form:**

Let $h(p)$ be a continuous function, except that $h(p)\equiv 0$ if $p<0$, and let the process $h(P(t))$ a bounded mean square. Then,

$$\int_0^t h(P(s))dP(s) \stackrel{\text{ims}}{=} \sum_{k=0}^{P(t)-1} h(k).$$

Proof: By FTPJC, $d\int_0^t h(P)dP \stackrel{\text{dt}}{=} (h(P)dP)(t)$ and $d\sum_{k=0}^{P-1} h(k) = \sum_{k=0}^{P+dP-1} h(k) - \sum_{k=0}^{P-1} h(k) \stackrel{\text{dt}}{=} (h(P)dP)(t)$ **by increment algebra**, so they have the same differential, but also the same zero initial condition at $t=0$, so are the same, i.e., the same starts and the same changes yield the same result.

- **Theorem 2.2.4. Poisson Stochastic Integrals Given Jump-Times T_k:**

Let $h(x,t)$ be a continuous function, let the process $h(X(t),t)$ have a bounded mean square, and satisfy the i-PWCA mean square limits Assumption. Then,

$$\int_0^t h(X(s),s)dP(s) \stackrel{\text{ims}}{=} \sum_{k=1}^{P(t)} h(X(T_k^-),T_k^-),$$

where T_k is the kth jump-time of Poisson process $P(t)$ and T_k^- is the prejump-time. The usual no-jump, no-amplitude convention is assumed, i.e., $\sum_{k=1}^{0} h(X(T_k^-),T_k^-) \equiv 0$.

(Comment: This is more of a definition of the Poisson process with coefficient $h(X(t),t)$, but an elaborate constructive proof along with consistency with the results of prior theorem are given in Hanson (2007).)

- **Definition 2.2.2. Jump Function $[X](t)$:**

The jump value of the state X at the kth jump-time T_k **is really the "zeroth derivative" but** is defined as

$$[X](T_k) \equiv X(T_k^+) - X(T_k^-),$$

when the kth prejump is time T_k^-. For finite discontinuities, the jump function includes all the change of the function, the **zeroth change or discrete derivative** of the state $X(t)$. Assuming **right-continuity**, then $T_k^- < T_k^+ = T_k$.

- **Lemma 2.2.1. Mixed Differential Products** $dt\cdot dP(t)$ **and** $dP(t)\cdot dW(t)$**:**

$$\int_0^t ds\cdot dP(s) \overset{\text{ims}}{=} 0 \quad \text{or} \quad dt\cdot dP(t) \overset{\text{dt}}{=} 0,$$

and

$$\int_0^t dP(s)\cdot dW(s) \overset{\text{ims}}{=} 0 \quad \text{or} \quad dP(t)\cdot dW(t) \overset{\text{dt}}{=} 0,$$

where, recall, $W(t)$ *and* $P(t)$ *are independent random variables.*

(Comment: The proofs are similar to the proof for $dt\cdot dW(t)$ and are easy to motivate with IFA, but see the text (2007, pp. 72-73) for more information on the IMS proofs. Crudely, the first integrand is $\mathrm{O}((ds)^2)$ and the second $\mathrm{O}((ds)^{3/2})$, exceeding $\mathrm{O}(ds)$, i.e., dt-precision.)

- **Theorem 2.2.5. Mean Square Limit Form of Zero-One Law:**

Let m *be a nonnegative integer and* $\mathrm{E}[dP(t)] = \lambda(t)dt$ *with bounded maximum,* $\lambda^* = \max_t[\lambda(t)]$. *Then*

$$\int_0^t (dP)^m(s) \stackrel{\text{ims}}{=} P(t) \quad \text{or} \quad (dP)^m(t) \stackrel{\text{dt}}{=} dP(t).$$

(Comment: The mean square limit proof is left for the reader as an exercise. **Also, note that if $x > 0$ and m is an integer, then $x^m = x$ means that $x > 0$ satisfies a zero-one algebraic law: $x = 0$ or $x = 1$.**)

- **Table 2.2.2. Some Itô stochastic jump differentials with an accuracy with error $\mathrm{o}(dt)$ as $dt \to 0^+$.**

Differential Jump Form	**Itô Mean Square Limit**
$dP(t)$	$dP(t)$
dt	dt
$dt\ dP(t)$	0
$(dP)^m(t)$	$dP(t), m \geq 1$
$dP(t)\ dW(t)$	0
$(dt)^k(dP)^m(t)$	$0, k \geq 1,\ m \geq 1$
$(dt)^k(dP)^m(t)(dW)^n(t)$	$0,\ k \geq 1,\ m \geq 1,\ n \geq 1$

Remarks on Table 2.2.2:

- In **Table 2.2.2**, differential entries are just symbols of the underlying integral basis and care should be taken when applying them to find the mean square representation of differentials, especially when they appear in multiplicative combinations.

- The mean square limit justification of the differential power rule $(dP)^m(t) \stackrel{\mathrm{dt}}{=} dP(t)$ is left as an Exercise, along with the Exercise previously mentioned for $dP(t)dW(t)$, but recall that $dW(t)$ behaves like $\sqrt{dt}Z(t)$, where $Z(t)$ is a standard (zero-mean, unit-variance) process and $\mathrm{E}[dP(t)] = \lambda(t)dt$.

- **Theorem 2.2.5. Some Mean Stochastic Jump Integrals:**

Let $h(X(t), t)$ *satisfy the mean square integrability condition on* $0 \leq t_0 \leq t$ *and let* $X(t)$ *be a Markov process,*

$$\mathrm{E}\left[\int_{t_0}^{t} h^2(X(s), s)ds\right] < \infty$$

and the ***Itô Forward Approximation (IFA) Expansion Assumption*** *for* $h(X(t), t)$*, where* $\mathrm{E}[dP(t)] = \lambda(t)dt$*. Then*

1. *The* ***Itô expectation*** *of the standard jump integral is*

$$\mathrm{E}\left[\int_{t_0}^{t} h(X(s), s)dP(s)\right] \overset{\text{ifa}}{=} \int_{t_0}^{t} \mathrm{E}[h(X(s), s)]\lambda(s)ds,$$

2. *Letting* $d\widehat{P}(t) \equiv dP(t) - \lambda(t)dt$ *be the simple* ***mean-zero Poisson process****,*

$$\mathrm{E}\left[\int_{t_0}^{t} h(X(s),s)d\widehat{P}(s)\right] \stackrel{\text{ifa}}{=} 0.$$

3. *An estimate inequality in the IFA sense,*

$$\mathrm{E}\left[\left|\int_{t_0}^{t} h(X(s),s)dP(s)\right|\right] \le \int_{t_0}^{t} \mathrm{E}\left[|h(X(s),s)|\right]\lambda(s)ds.$$

4. *Let* h_1 *and* h_2 *satisfy the same mean square integrability condition as* h*; then the* ***Itô covariance*** *for jump integrals is*

$$\mathrm{E}\left[\int_{t_0}^{t} h_1(X(s),s)d\widehat{P}(s)\int_{t_0}^{t} h_2(X(r),r)d\widehat{P}(r)\right]$$
$$\stackrel{\text{ifa}}{=} \int_{t_0}^{t} \mathrm{E}\left[h_1(X(s),s)h_2(X(s),s)\right]\lambda(s)ds.$$

5. *The* ***Itô variance*** *for jump stochastic integrals is given by*

$$\mathrm{E}\left[\left(\int_{t_0}^{t} h(X(s),s)d\widehat{P}(s)\right)^2\right] \overset{\text{ifa}}{=} \int_{t_0}^{t} \mathrm{E}[h^2(X(s),s)]\lambda(s)ds.$$

(Comment: This theorem is specified for IFA for simplicity and fast calculation for financial application as given earlier with similar diffusion $W(t)$ results, but a more rigorous treatment should use Itô Mean Square Limit. **Note that h is a jump-amplitude function, and usually the variance of the compound Poisson process leads to $\mathrm{E}[h^2]$ rather than $\mathrm{Var}[h]$ as will be seen later.**)