

2 Continuous Time Summary/Review

We need a statistical model that can apply to stock prices, returns, etc. The point of these notes is to quickly present the standard diffusion model that we use in asset pricing.

We use continuous time when it's simpler. In point of fact all our data are discretely sampled. Whether time is "really" continuous or discrete is an issue we'll leave to physicists and philosophers. Most of the approximations that took a lot of effort in Chapter 1 of *Asset pricing* work much more easily in continuous time.

2.1 Preview/summary:

2.1.1 Continuous time

1. Brownian motion

$$z_{t+\Delta} - z_t \sim N(0, \Delta)$$

2. Differential

$$dz_t = \lim_{\Delta \searrow 0} (z_{t+\Delta} - z_t)$$

3. dz and dt

$$\begin{aligned} dz_t &\sim O(\sqrt{dt}) \\ dz_t^2 &= dt \\ E_t(dz_t) &= 0 \\ \text{var}_t(dz_t) &= E_t(dz_t^2) = dz_t^2 = dt \end{aligned}$$

4. Diffusions

$$\begin{aligned} dx_t &= \mu(x_t, t)dt + \sigma(x_t, t)dz_t \\ E_t(dx_t) &= \mu(x_t, t)dt \\ \sigma_t^2(dx_t) &= \sigma^2(x_t, t)dt \end{aligned}$$

5. Examples

$$\begin{aligned} \frac{dp_t}{p_t} &= \mu dt + \sigma dz_t \\ dx_t &= -\phi(x_t - \mu)dt + \sigma dz_t \end{aligned}$$

¹Please report typos to john.cochrane@chicagobooth.edu. Make sure you have the latest version of the document. If you can report the context of the typo and not just the page number that will help me to find it more quickly.

6. Ito's lemma

$$dx_t = \mu dt + \sigma dz_t; y_t = f(t, x_t) \implies$$

$$\begin{aligned} dy_t &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dx_t^2 \\ dy_t &= \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \mu + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 \right] dt + \left[\frac{\partial f}{\partial x} \sigma \right] dz \end{aligned}$$

7. Stochastic calculus: “do second-order expansions, keep the dz and dt terms.”

2.1.2 Solving stochastic differential equations

1. Stochastic integral “add up the changes”

$$z_T - z_0 = \int_{t=0}^T dz_t \leftrightarrow z_T - z_0 = \sum_{t=1}^T \varepsilon_t$$

the result is just a way of saying z_T is a normally distributed random variable

$$\int_{t=0}^T dz_t = z_T - z_0 \sim N(0, T)$$

2. Example 1: diffusion

$$\begin{aligned} dx_t &= \mu dt + \sigma dz_t \\ \int_{t=0}^T dx_t &= \mu \int_{t=0}^T dt + \sigma \int_{t=0}^T dz_t \\ x_T - x_0 &= \mu T + \sigma (z_T - z_0) \end{aligned}$$

3. Example 2: lognormal diffusion

$$\begin{aligned} \frac{dx_t}{x_t} &= \mu dt + \sigma dz_t \\ d \ln x_t &= \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t \\ \ln x_T - \ln x_0 &= \left(\mu - \frac{1}{2} \sigma^2 \right) T + \sigma \int_0^T dz_t \\ x_T &= x_0 e^{(\mu - \frac{1}{2} \sigma^2) T + \sigma \int_0^T dz_s} \end{aligned}$$

4. Example 3: $AR(1)$

$$\begin{aligned} dx_t &= -\phi(x_t - \mu) dt + \sigma dz_t \\ x_T - \mu &= e^{-\phi T} (x_0 - \mu) + \sigma \int_{t=0}^T e^{-\phi(T-t)} dz_{T-t} \end{aligned}$$

5. Continuous-time MA processes

$$x_T = \int_{t=0}^T w(T-t) dz_t$$

6. With solutions, you can find moments. Examples

(a) Lognormal diffusion

$$\begin{aligned} x_T &= x_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \sigma \int_0^T dz_s} \\ E(x_T) &= x_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \frac{1}{2}\sigma^2 T} = x_0 e^{\mu T} \end{aligned}$$

(b) AR(1)

$$\begin{aligned} (x_t - \mu) &= e^{-\phi t} (x_0 - \mu) + \int_{s=0}^t e^{-\phi(t-s)} \sigma dz_s \\ E_0(x_t - \mu) &= e^{-\phi t} (x_0 - \mu) \\ \sigma^2(x_t - \mu) &= \int_{s=0}^t e^{-2\phi(t-s)} \sigma^2 ds = \frac{1 - e^{-2\phi t}}{2\phi} \sigma^2 \end{aligned}$$

(c) Continuous-time MA.

$$\begin{aligned} x_t &= \int_{s=0}^t w(t-s) dz_s, \\ E_0(x_t) &= E_0 \left(\int_{s=0}^t w(t-s) dz_s \right) = \int_{s=0}^t w(t-s) E_s(dz_s) = 0 \\ E_0(x_t^2) &= E_0 \left(\int_{s=0}^t w(t-s) dz_s \right)^2 = \int_{s=0}^t w(t-s)^2 ds. \end{aligned}$$

2.1.3 Finding Moments and Densities

1. Backward equation. What is

$$f(x_t, t) = E_t[\phi(x_T)]?$$

We can work backwards using Ito's lemma to find

$$\frac{\partial f(x, t)}{\partial t} + \frac{\partial f(x, t)}{\partial x} \mu(x, t) + \frac{1}{2} \frac{\partial^2 f(x, t)}{\partial x^2} \sigma^2(x, t) = 0$$

Then solve this partial differential equation with boundary condition

$$f(x, T) = \phi(x).$$

2. Densities and the Forward Equation.

- (a) Define $f(x, t|x_0)$ = the probability density of x_t given x_0 at time 0. It's a delta function at $t = 0$, $x = x_0$. Looking forward it solves

$$\frac{\partial f(x, t|x_0)}{\partial t} + \frac{\partial [\mu(x)f(x, t|x_0)]}{\partial x} = \frac{1}{2} \frac{\partial^2 [\sigma^2(x)f(x, t|x_0)]}{\partial x^2}$$

- (b) This fact gives a nice formula for the stationary (unconditional) density. The formula is:

$$f(x) = k \frac{e^{\int^x -2\frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(x)}$$

where k is a constant, so we integrate to one. Evaluating k ,

$$f(x) = \left[\int \frac{e^{\int^x 2\frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(s)} ds \right]^{-1} \frac{e^{\int^x 2\frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(x)}$$

2.1.4 Asset pricing in continuous time

1. Present value formula

$$p_t = E_t \int_{s=0}^{\infty} e^{-\delta s} \frac{u'(c_{t+s})}{u'(c_t)} x_{t+s} ds = E_t \int_{s=0}^{\infty} \frac{\Lambda_{t+s}}{\Lambda_t} x_{t+s} ds$$

2. What's a "return"?

$$dR_t = \frac{dp_t}{p_t} + \frac{x_t}{p_t} dt$$

3. What's a "cumulative value process"

$$\frac{dV_t}{V_t} = dR_t$$

4. What's a "riskfree rate?"

$$r_t^f dt$$

5. What's an "excess return?"

$$dR_t - r_t^f dt = \frac{dp_t}{p_t} + \frac{x_t}{p_t} dt - r_t^f dt$$

6. What's the equivalent of $1 = E(mR)$?

$$E_t \left[\frac{d(\Lambda_t p_t)}{\Lambda_t p_t} \right] - \frac{x_t}{p_t} dt = 0$$

or $E_t [d(\Lambda_t V_t)] = 0$

7. What's the equivalent of $R^f = 1/E(m)$?

$$r^f dt = -E_t \left[\frac{d\Lambda_t}{\Lambda_t} \right]$$

8. What's the equivalent of $E(R^e) = -R^f cov(m, R^e)$?

$$E_t(dR_t) - r^f dt = -E_t \left[\frac{d\Lambda_t}{\Lambda_t} dR_t \right]$$

9. What's the equivalent of $m_{t+1} \approx 1 - \delta - \gamma \Delta c_{t+1}$?

$$\begin{aligned} \frac{d\Lambda_t}{\Lambda_t} &= -\delta dt - \gamma \frac{dc_t}{c_t} + \frac{1}{2} \gamma (\gamma + 1) \frac{dc_t^2}{c_t} \\ \frac{d\Lambda_t}{\Lambda_t} &= -\delta dt - \gamma d[\log c_t] + \frac{1}{2} \gamma^2 d[\log c_t]^2 \end{aligned}$$

10. What's the equivalent of $R^f \approx \delta + \gamma E(\Delta c)$?

$$\begin{aligned} r^f &= \delta + \gamma \mu_c - \frac{1}{2} \gamma^2 \sigma_c^2 \\ \text{where } d \log c_t &= \mu_c dt + \sigma_c dz_t \end{aligned}$$

11. What's the equivalent of $E(R^e) \approx \gamma cov(R^e, \Delta c)$?

$$E_t(dR_t) - r^f dt = \gamma E_t \left[\frac{dc_t}{c_t} dR_t \right] = \gamma cov_t \left[\frac{dc_t}{c_t}, dR_t \right]$$

2.2 A quick discrete time review

I'll explain continuous time by analogy with discrete time. This is a quick review of the essential concepts can't hurt.

1. A sequence of i.i.d. random variables (like a coin flip) forms the basic building block of time series.

$$\varepsilon_t \sim i.i.d.$$

i.i.d. means "independent and identically distributed." In particular, it implies that the conditional mean and variance are constant through time. Our building block series has a zero mean,

$$\begin{aligned} E_t(\varepsilon_{t+1}) &= 0 \\ \sigma_t(\varepsilon_{t+1}) &= \sigma^2 = \text{const.} \end{aligned}$$

The *concept* of conditional mean and variance is really important here! E_t means "expectation conditional on all information at time t ." We sometimes specify that ε_t are normally distributed, but they don't have to be.

2. We build more interesting time series models from this building block. The AR(1) is the first canonical example

$$x_{t+1} = \rho x_t + \varepsilon_{t+1}.$$

The sequence of conditional means follows

$$\begin{aligned} E_t x_{t+1} &= E_t (\rho x_t + \varepsilon_{t+1}) = \rho x_t \\ E_t x_{t+2} &= E_t (\rho x_{t+1} + \varepsilon_{t+2}) = \rho^2 x_t \\ E_t x_{t+k} &= \rho^k x_t. \end{aligned}$$

The AR(1) displays some *persistence* since after a shock it is expected to stay up for a while, and then decay gradually back again. A good exercise at this point is to work out the sequence of conditional variances of the AR(1). The answer is

$$\begin{aligned} \sigma_t^2(x_{t+1}) &= \sigma_\varepsilon^2 \\ \sigma_t^2(x_{t+2}) &= (1 + \rho^2)\sigma_\varepsilon^2 \\ \sigma_t^2(x_{t+k}) &= (1 + \rho^2 + \dots + \rho^{2(k-1)})\sigma_\varepsilon^2 \end{aligned}$$

3. We can also write the AR(1) as

$$x_{t+1} = \rho x_t + \sigma \varepsilon_{t+1}; \quad \sigma_\varepsilon^2 = 1$$

The earlier notation is more common in discrete time; the latter notation is more common as we move to continuous time. In either case, note $E_t \varepsilon_{t+1} = 0$ so it's uncorrelated with x_t . Thus you can estimate an AR(1) by an OLS forecasting regression.

4. The canonical example 2 is the MA(1).

$$x_{t+1} = \varepsilon_{t+1} + \theta \varepsilon_t.$$

Its sequence of conditional means follows

$$\begin{aligned} E_t(x_{t+1}) &= \theta \varepsilon_t \\ E_t(x_{t+2}) &= 0 \\ E_t(x_{t+k}) &= 0. \end{aligned}$$

It displays some persistence too but only for one period. Problem: work out its conditional variances.

5. You can transform from AR to MA representations. You can think of this operation as “solving” the AR(1)

$$\begin{aligned} x_t &= \rho x_{t-1} + \varepsilon_t \\ x_t &= \rho(\rho x_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \rho^2 x_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \\ x_t &= \rho^3 x_{t-3} + \rho^2 \varepsilon_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \\ x_t &= \rho^t x_0 + \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j} \end{aligned}$$

“Solving” the difference equation means, really, expressing x_t as a random variable, given time 0 information. Now we know x_t has

$$\begin{aligned} E_0(x_t) &= \rho^k x_0; \\ \sigma_0^2(x_t) &= (1 + \rho^2 + \dots + \rho^{2(k-1)})\sigma_\varepsilon^2. \end{aligned}$$

If the ε are normal, so is x_t .

6. Continuing, we can write the AR(1) as

$$x_t = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}$$

Thus, an AR(1) is the same as an MA(∞). You can similarly write an MA(1) as an AR(∞). Choose which representation is easiest for what you’re doing. An example: Let’s find the unconditional mean and variance of the AR(1). One way to do this is with the MA(∞) representation

$$\begin{aligned} E(x_t) &= \sum_{j=0}^{\infty} \rho^j E(\varepsilon_{t-j}) = 0. \\ \text{var}(x_t) &= \text{var}\left(\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}\right) = \sum_{j=0}^{\infty} \rho^{2j} \sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \end{aligned}$$

Notice: *the ε are all uncorrelated with each other.* That’s how we got rid of the covariance term and all σ^2 are the same. We can also find the same quantities from the AR(1) representation

$$\begin{aligned} E(x_t) &= \rho E(x_{t-1}) + 0 \rightarrow E(x_t) = 0 \\ \text{var}(x_t) &= \rho^2 \text{var}(x_{t-1}) + \sigma_\varepsilon^2 \rightarrow \text{var}(x_t) = \frac{\sigma_\varepsilon^2}{1 - \rho^2}. \end{aligned}$$

To come to the latter conclusions, you have to understand that $E(x_t) = E(x_{t-1})$ and $\text{var}(x_t) = \text{var}(x_{t-1})$.

7. All of the familiar tricks for linear ARMA models, including lag operator notation, have continuous time counterparts. It takes a little work to make the translation. If you’re interested, see my “Continuous-time linear models” *Foundations and Trends in Finance* 6 (2011), 165–219 DOI: 10.1561/05000000037. It’s on my webpage, http://faculty.chicagobooth.edu/john.cochrane/research/papers/continuous_time_linear_models_FT.pdf

2.3 Continuous time series

1. Preview. In discrete time, our building block is the i.i.d. (normal) shock ε_t with variance $\sigma^2(\varepsilon_t) = 1$. We build up more complex time series from this building block with difference equation models such as the AR(1),

$$x_{t+1} = \rho x_t + \sigma \varepsilon_{t+1}$$

We “solve” such difference equations to represent x_t as a function of its past shocks,

$$x_t = \sum_{j=0}^{t-1} \rho^j \sigma \varepsilon_{t-j} + \rho^t x_0.$$

Our task is to do the exact same sort of thing in continuous time.

2. Random walk-Brownian motion z_t .

(a) In discrete time, a random walk z_t is defined as the sum of independent shocks,

$$z_t - z_0 = \sum_{j=1}^t \varepsilon_j.$$

Obviously, we recover ε_t as differences of z_t .

$$\Delta z_t = (z_t - z_{t-1}) = \varepsilon_t.$$

It will turn out to be easier to think first about z_t in continuous time and then take differences than to think directly about ε_t

(b) Now, let’s examine the properties of z_t . The variance of z grows linearly with horizon.

$$\begin{aligned} z_{t+2} - z_t &= \varepsilon_{t+1} + \varepsilon_{t+2} \\ E(z_{t+2} - z_t) &= 0 \\ \text{var}(z_{t+2} - z_t) &= 2 \end{aligned}$$

$$\begin{aligned} z_{t+k} - z_t &= \varepsilon_{t+1} + \varepsilon_{t+2} + \dots + \varepsilon_{t+k} \\ E(z_{t+k} - z_t) &= 0 \\ \text{var}(z_{t+k} - z_t) &= k \end{aligned}$$

(remember, I set $\sigma_\varepsilon = 1$.)

(c) To think about continuous time, just generalize this idea for any time interval. Let’s *define the process* z_t as one with normal increments,

$$z_{t+\Delta} - z_t \sim N(0, \Delta)$$

for *any* Δ , not just integers. This is a *Brownian motion*, the continuous-time version of a *random walk*.

(d) *I.i.d. property* In discrete time $E(\varepsilon_t \varepsilon_{t+1}) = 0$, i.e. $E(z_{t+2} - z_{t+1}, z_{t+1} - z_t) = 0$. The natural generalization is $E(z_{t+\Delta} - z_t, z_{t+\Delta+\eta} - z_{t+\Delta}) = 0$, or more generally *Nonoverlapping differences of $\{z_t\}$ are uncorrelated with each other.*

3. The differential dz

- (a) Now let's take the "derivative" as we took the difference $\varepsilon_t = z_t - z_{t-1}$. Define

$$\begin{aligned} dz_t &= \lim_{\Delta \searrow 0} (z_{t+\Delta} - z_t) \\ dz_t &\leftrightarrow \varepsilon_t \text{ fundamental building block} \end{aligned}$$

d is a tiny *forward*-difference operator.

- (b) dz is of size \sqrt{dt} . What does this mean? Note

$$\begin{aligned} \sigma^2(z_{t+\Delta} - z_t) &= \Delta; \\ \sigma(z_{t+\Delta} - z_t) &= \sqrt{\Delta} \end{aligned}$$

Thus, dz_t is of order (typical size) \sqrt{dt} ! This means that sample paths of z_t are *continuous* but not *differentiable*. dz makes sense, but dz/dt does not. dz_t is of order \sqrt{dt} so $dz/dt \rightarrow +/\infty$, moving "infinitely fast" (up and down). The z_t process is fractal; it's jumpy at any time scale; it has an infinite path length.

- (c) dz , dz_t , and $dz(t)$ are equivalent notations.
 (d) *Moments*:

$$\begin{aligned} E_t(dz_t) &= 0; \iff E_t(z_{t+\Delta} - z_t) = 0 \\ \text{var}_t(dz_t) &= dt; \iff \text{var}(z_{t+\Delta} - z_t) = \Delta, \\ \text{cov}(dz_t, dz_s) &= 0 \text{ } s \neq t \iff \text{cov}(z_{t+\Delta} - z_t, z_{s+\Delta} - z_s) = 0 \end{aligned}$$

Watch out for notation. $E_t(dz_t) \neq dz_t$! d is a forward difference operator, so $E_t(dz_t)$ means expected value of how much dz will change in the *next* instant. It really means $E_t(z_{t+\Delta} - z_t)$, which is obviously not the same thing as $z_{t+\Delta} - z_t$

- (e) $dz_t^2 = dt$. Variance and second moment are the same.

$$\text{var}_t(dz_t) = E_t(dz_t^2) = dz^2 = dt.$$

Second moments are nonstochastic! It's not just dz^2 is of order dt , but in fact $dz^2 = dt$. We often write $E(dz^2)$ to remind ourselves that for any discrete interval these are second moments of random variables. But in the limit, in fact, second moments are nonstochastic. Similarly if we have two Brownians dz and dw , they can be correlated and the cross product is

$$\text{cov}_t(dz_t, dw_t) = E_t(dz_t dw_t) = dz_t dw_t.$$

- (f) (Optional note.) The fact that squares of random variables are deterministic is hard to swallow. If dz_t is a normal with standard deviation \sqrt{dt} , why is dz_t^2 a number, not a χ^2 distributed random variables? To see why, compare dz_t and dz_t^2 . $z_{t+\Delta} - z_t$ is $N(0, \Delta)$, $(z_{t+\Delta} - z_t)/\sqrt{\Delta} \sim N(0, 1)$ so the probability that, say, $\|z_{t+\Delta} - z_t\| > 2\sqrt{\Delta}$ is 5%. The "typical size" of a movement $z_{t+\Delta} - z_t$ goes to zero at rate $\sqrt{\Delta}$.
 On the other hand, $(z_{t+\Delta} - z_t)^2/\Delta \sim \chi_1^2$. (χ_1^2 or "chi-squared with one degree of freedom" just means "the distribution of the square of a standard normal.") The fact that the distribution $(z_{t+\Delta} - z_t)^2/\Delta$ stays the same as Δ gets small means that the

probability that, say, $(z_{t+\Delta} - z_t)^2 > 2\Delta$ is fixed as Δ gets small. Thus $(z_{t+\Delta} - z_t)^2$ goes to zero at rate Δ . The “typical size” of a movement $(z_{t+\Delta} - z_t)^2$ or the probability that it exceeds any value goes to zero at rate Δ . Things whose movements go to zero at rate Δ are differentiable, and hence nonstochastic. $(z_{t+\Delta} - z_t)$ is of order dt and hence nonstochastic.

4. **Diffusion processes.** We have the building block, the analog to ε_t . Now, we build more complex process like $AR(1)$, ARMA, etc. in the same way as we do in discrete time.

(a) Random walk with drift.

$$\text{discrete: } x_{t+1} = \mu + x_t + \sigma\varepsilon_{t+1}$$

$$x_{t+1} - x_t = \mu + \sigma\varepsilon_t$$

$$\text{continuous: } dx_t = \mu dt + \sigma dz_t$$

i. Moments of the random walk with drift, which gives you some practice in dz and dt . The mean is

$$E_t(dx_t) = \mu dt + \sigma E_t(dz_t) = \mu dt$$

For the variance, notice

$$dx_t^2 = (dx_t)^2 = \mu^2 dt^2 + \sigma^2 dz^2 + 2\mu dt dz_t = \sigma^2 dt$$

The first parenthesis clears up notation. In calculus, you learned to keep terms of order dt , and ignore terms of order dt^2 and so on, because they get arbitrarily small as the time interval shrinks. The new rule is, we keep terms of order $dz = \sqrt{dt}$ and dt , but we ignore terms of order $dt^{3/2}$, dt^2 , etc. just as in regular calculus. So terms like dt^2 and $dt dz$ drop out. Only the $\sigma^2 dt$ term remains. So, the variance is

$$\text{var}_t(dx_t) = E_t(dx_t^2) = dx_t^2 = \sigma^2 dt$$

ii. The mean drops out of the variance. Loosely, diffusions vary so much that variances overwhelm means and it doesn't matter if you take the mean out when you take variance or not. You might have objected, and wanted to define

$$\text{var}_t(dx_t) = E_t \left\{ [dx_t - E_t(dx_t)]^2 \right\}$$

But watch:

$$E_t \left\{ [dx_t - E_t(dx_t)]^2 \right\} = E_t \left\{ [dx_t - \mu dt]^2 \right\} = E_t \left\{ [\sigma dz_t]^2 \right\} = \sigma^2 E_t(dz_t^2) = \sigma^2 dt$$

We get the same result whether we take the mean out or not. The dz_t term is of order \sqrt{dt} , so almost always infinitely bigger than the dt term.

(b) Geometric Brownian motion with drift. This is the standard asset price process, so worth careful attention.

$$\frac{dp_t}{p_t} = \mu dt + \sigma dz_t$$

dp_t is the change in price, so dp_t/p_t is the instantaneous percent (not annualized) change in price or instantaneous rate of return if there are no dividends. We'll see how to add dividends in a moment.

i. The moments are

$$\begin{aligned} E_t \left(\frac{dp_t}{p_t} \right) &= \mu dt \\ \text{var}_t \left(\frac{dp_t}{p_t} \right) &= \sigma^2 dt \end{aligned}$$

so μ and σ can represent the mean and standard deviation of the arithmetic percent rate of return. Since they multiply dt , they have annual units. $\mu = 0.05$ and $\sigma^2 = 0.10^2$ are numbers to use for a 5% mean return and 10% standard deviation of return.

ii. *Don't forget the dt at the end of these expressions*

$$E_t \left(\frac{dp_t}{p_t} \right) = \mu, \text{var}_t \left(\frac{dp_t}{p_t} \right) = \sigma^2$$

is tempting but incorrect. It's ok to write

$$\frac{1}{dt} E_t \left(\frac{dp_t}{p_t} \right) = \mu, \frac{1}{dt} \text{var}_t \left(\frac{dp_t}{p_t} \right) = \sigma^2$$

but it's better to write

$$E_t \left(\frac{dp_t}{p_t} \right) = \mu dt, \text{var}_t \left(\frac{dp_t}{p_t} \right) = \sigma^2 dt$$

iii. Sticklers like to write this process

$$dp_t = \mu p_t dt + \sigma p_t dz_t$$

so it is in the canonical form

$$dp_t = \mu(p_t, t, \dots) dt + \sigma(p_t, t, \dots) dz_t.$$

iv. You can also see here the reason we call it a “stochastic differential equation.” Without the dz_t term, we would “solve” this equation to say

$$p_t = p_0 e^{\mu t}.$$

i.e. value grows exponentially. We will come back and similarly “solve” the *stochastic differential equation* with the dz_t term.

(c) AR(1), *Ornstein Uhlenbeck* process

$$\begin{aligned} \text{discrete:} \quad x_{t+1} &= (1 - \rho)\mu + \rho x_t + \sigma \varepsilon_{t+1} \\ x_{t+1} - x_t &= -(1 - \rho)(x_t - \mu) + \sigma \varepsilon_{t+1} \\ E_t(x_{t+1} - x_t) &= -(1 - \rho)(x_t - \mu) \\ \text{var}_t(x_{t+1} - x_t) &= \sigma^2 \end{aligned}$$

$$\begin{aligned} \text{continuous} & : dx_t = -\phi(x_t - \mu)dt + \sigma dz_t \\ E_t dx_t & = -\phi(x_t - \mu)dt \\ \text{var}_t dx_t & = \sigma^2 dt \end{aligned}$$

The part in front of dt is called the “drift” and the part in front of dz_t is called the “diffusion” coefficient.

- (d) In general, we build complex time series processes from the *stochastic differential equation*

$$dx_t = \mu(x_t, t, \dots)dt + \sigma(x_t, t, \dots)dz_t.$$

μ and σ may be nonlinear functions (ARMA in time series are just linear). Note we often suppress state and time dependence and just write μ, σ or $\mu(\cdot), \sigma(\cdot)$ or μ_t and σ_t .

$$\begin{aligned} E_t(dx_t) & = \mu(x_t, t, \dots)dt \\ \text{var}_t(dx_t) & = \sigma^2(x_t, t, \dots)dt \end{aligned}$$

- (e) Simulation/approximation. We often use continuous time to derive approximations for discrete time. We also will often want to simulate forward continuous time processes on a computer. The “Euler approximation” is a natural way to do this. From

$$dx_t = \mu(x_t, t)dt + \sigma(x_t, t)dz_t$$

write

$$x_{t+\Delta} - x_t = \mu(x_t, t)\Delta + \sigma(x_t, t)\sqrt{\Delta}\varepsilon_{t+\Delta}$$

where Δ is a short time interval and $\varepsilon_{t+\Delta} \sim N(0, 1)$. For example, you could use $\Delta = 1/365$ to simulate daily data.

5. **Ito’s lemma.** Given a diffusion process for x_t , we can construct a new process $y_t = f(x_t)$. How can we find a diffusion representation for y_t given the diffusion representation of x_t ? For example, what does the log price process $\log(p_t)$ look like?

- (a) Nonstochastic answer: Given

$$\frac{dx_t}{dt} = \mu$$

we use the chain rule.

$$dy_t = \frac{\partial f}{\partial x} dx_t = \frac{\partial f}{\partial x} \mu dt$$

But that assumes dx/dt is meaningful, i.e. $dx_t = \mu dt$. What if $dx_t = \mu_t dt + \sigma dz_t$?

- (b) So our question:

$$\begin{aligned} dx_t & = \mu dt + \sigma dz_t \\ y_t & = f(x_t) \\ dy_t & = ? \end{aligned}$$

Answer:

- Do a *second* order Taylor expansion. Keep terms of order dt and $dz_t = \sqrt{dt}$, ignore higher order terms. Don't forget $dz_t^2 = dt$.

The answer is called Ito's lemma:

$$dy_t = \frac{\partial f}{\partial x} dx_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dx_t^2$$

$$dy_t = \left[\frac{\partial f}{\partial x} \mu + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 \right] dt + \left[\frac{\partial f}{\partial x} \sigma \right] dz$$

Intuition: Jensen's inequality says, $E[y(x)] < y[E(x)]$ if $y(x)$ is concave. $E[u(c)] < u[E(c)]$ for example. The second derivative term takes care of this fact. You will see this in the "convexity" terms of option and bond pricing.

- (c) An example. Geometric growth

$$\frac{dp_t}{p_t} = \mu dt + \sigma dz_t$$

$$y_t = \ln p_t$$

$$dy_t = \frac{1}{p_t} dp_t - \frac{1}{2} \frac{1}{p_t^2} dp_t^2 = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t$$

- (d) Conversely,

$$dy_t = \mu dt + \sigma dz_t$$

$$p_t = e^{y_t}$$

$$dp_t = e^{y_t} dy_t + \frac{1}{2} e^{y_t} dy_t^2$$

$$\frac{dp_t}{p_t} = \left(\mu + \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t$$

- (e) Another example. We will often multiply two diffusions together. What is the product diffusion? *Use the chain rule but go out to second derivatives.*

$$d(x_t y_t) = \frac{\partial(xy)}{\partial x} dx_t + \frac{\partial(xy)}{\partial y} dy_t + \frac{1}{2} \times 2 \times \frac{\partial^2(xy)}{\partial x \partial y} dx_t dy_t$$

$$= y_t dx_t + x_t dy_t + dx_t dy_t$$

The second partials with respect to x and y $\partial^2(xy)/\partial x^2$ are zero. Notice the extra term $dx_t dy_t$ relative to the usual chain rule.

- (f) Ito's lemma more generally. If

$$y = f(x_t, t)$$

then

$$dy_t = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dx_t^2$$

$$dy_t = \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \mu + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 \right] dt + \left[\frac{\partial f}{\partial x} \sigma \right] dz$$

The version with partial derivatives starting from $y_t = f(x_{1t}, x_{2t}, t)$ is obvious enough and long enough that I won't even write it down.

2.4 Solving stochastic differential equations.

1. The problem: We want to do the analogue of “solving” difference equations. For example, when we write an AR(1),

$$x_t = \rho x_{t-1} + \varepsilon_t$$

we solve backward to

$$x_t = \rho^t x_0 + \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}.$$

How do we do this in continuous time?

2. Another motivation. If you have a differential equation

$$dx_t = \mu dt$$

you know how to solve this. Integrate both sides from 0 to T ,

$$x_T = x_0 + \mu \int_{s=0}^T dt = x_0 + \mu T.$$

So, how do you solve a stochastic differential equation,

$$dx_t = \mu dt + \sigma dz_t?$$

3. A natural idea: Why don't we just integrate both sides,

$$\int_{t=0}^T dx_t = \mu \int_{t=0}^T dt + \sigma \int_{t=0}^T dz_t.$$

What meaning can we give to the integrals? Why not just “add up all the little changes.”

$$\int_0^T dz_t = (z_\Delta - z_0) + (z_{2\Delta} - z_\Delta) + (z_{3\Delta} - z_{2\Delta}) + \dots + (z_T - z_{T-\Delta}) = z_T - z_0.$$

Then our solution reads

$$x_T - x_0 = \mu T + \sigma (z_T - z_0).$$

Now, we defined $(z_T - z_0)$ to mean “a normally distributed random variable with mean 0 and variance T .” So this solution means, “ $x_T - x_0$ is a normally distributed random variable with mean μT and variance $\sigma^2 T$.” Similarly, if dz_t are the little changes, you get z back if you add them up.

4. In sum, the building block of solutions is the opposite of our dz_t differencing operator, and completes the analogy we started when defining z_T in the first place

$$z_T - z_0 = \int_{t=0}^T dz_t \leftrightarrow z_T - z_0 = \sum_{t=1}^T \varepsilon_t$$

$\int_{t=0}^T dz_t$ is called a *Stochastic integral*. This is a fundamentally different definition of “integral” which gives mathematicians a lot to play with. For us, it just means “add up all the little changes.” Recall that z_t is not differentiable. That’s why we do not write the usual integral notation

$$\begin{aligned} \text{Yes:} \quad z_T - z_0 &= \int_0^T dz_t \\ \text{No :} \quad z_T - z_0 &= \int_0^T \left(\frac{dz(t)}{dt} \right) dt \end{aligned}$$

We do not define the integral as the “area under the curve” as you do in regular calculus. What we mean by “ $\int_0^T dz_t$ ” in the end is just a normally distributed random variable with mean zero and variance t .

$$\int_0^T dz_t = z_T - z_0 \sim N(0, T)$$

5. Now, to more complex stochastic differential equations. In discrete time, we “solve” the difference equation:

$$\begin{aligned} x_t &= (1 - \rho)\mu + \rho x_{t-1} + \varepsilon_t \\ x_t - \mu &= \rho(x_{t-1} - \mu) + \varepsilon_t \\ x_T - \mu &= \rho^T(x_0 - \mu) + \sum_{j=0}^{T-1} \rho^j \varepsilon_{T-j}. \end{aligned}$$

This means we know the conditional distribution of the random variable x_T , and really that’s what it means to have “solved” the stochastic difference equation.

$$x_T | x_0 \sim N \left[\mu + \rho^T(x_0 - \mu), \sigma_\varepsilon^2 \sum_{j=0}^{T-1} \rho^{2j} \right]$$

Let’s do the same thing in continuous time.

6. I’ll just do some examples. I won’t use anything fancier than the basic idea: plop $\int_{t=0}^T$ on both sides of an expression and interpret the results.

(a) *Random walk*. If we start with dz_t and integrate both sides,

$$\begin{aligned} z_T - z_0 &= \int_{t=0}^T dz_t \\ z_T - z_0 &\sim N(0, T) \end{aligned}$$

(b) *Random walk with drift*

$$\begin{aligned}
 dx_t &= \mu dt + \sigma dz_t \\
 \int_{t=0}^T dx_t &= \int_{t=0}^T \mu dt + \int_{t=0}^T \sigma dz_t \\
 x_T - x_0 &= \mu T + \sigma \int_{t=0}^T dz_t \\
 x_T - x_0 &= \mu T + \varepsilon; \quad \varepsilon \sim N(0, T\sigma^2) \\
 x_T - x_0 &\sim N(\mu T, T\sigma^2)
 \end{aligned}$$

As in discrete time,

$$x_T = x_0 + \mu T + \sum_{t=1}^T \varepsilon_t$$

(c) *Lognormal price process.*

$$\frac{dp_t}{p_t} = \mu dt + \sigma dz_t$$

By Ito's lemma

$$d \ln p_t = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t$$

now integrate both sides,

$$\begin{aligned}
 \int_0^T d \ln p_t &= \int_{t=0}^T \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \int_{t=0}^T \sigma dz_t \\
 \ln p_T - \ln p_0 &= \left(\mu - \frac{1}{2} \sigma^2 \right) T + \sigma \int_{t=0}^T dz_t \\
 p_T &= p_0 e^{(\mu - \frac{1}{2} \sigma^2) T + \sigma \int_{t=0}^T dz_t}
 \end{aligned}$$

i.e.,

$$p_T = p_0 e^{(\mu - \frac{1}{2} \sigma^2) T + \sigma \sqrt{T} \varepsilon} \quad \varepsilon \sim N(0, 1)$$

p_T is lognormally distributed. (Lognormal means log of p_T is normally distributed.)

- *A geometric diffusion, where the arithmetic return is instantaneously normal, means that the finite-period return is lognormally distributed.*

(d) *AR(1)*

$$dx_t = -\phi(x_t - \mu) dt + \sigma dz_t$$

The solution is,

$$x_T - \mu = e^{-\phi T} (x_0 - \mu) + \sigma \int_{s=0}^T e^{-\phi s} dz_{T-s}$$

this looks just like the discrete time case

$$x_T - \mu = \rho^T (x_0 - \mu) + \sum_{j=0}^{T-1} \rho^j \varepsilon_{t-j}.$$

i. To derive this answer, you have to be a little clever. Find

$$\begin{aligned} d\left(e^{\phi t}(x_t - \mu)\right) &= \phi e^{\phi t}(x_t - \mu) dt + e^{\phi t} dx_t \\ &= \phi e^{\phi t}(x_t - \mu) dt + e^{\phi t} [-\phi(x_t - \mu)dt + \sigma dz_t] \\ &= \sigma e^{\phi t} dz_t \end{aligned}$$

Now integrate both sides,

$$\begin{aligned} \int_{t=0}^T d\left(e^{\phi t}(x_t - \mu)\right) &= \int_{t=0}^T \sigma e^{\phi t} dz_t \\ e^{\phi T}(x_T - \mu) - (x_0 - \mu) &= \sigma \int_{t=0}^T e^{\phi t} dz_t \\ (x_T - \mu) &= e^{-\phi T}(x_0 - \mu) + \sigma \int_{t=0}^T e^{\phi(t-T)} dz_t \end{aligned}$$

In the continuous time tradition, we usually express integrals going forward in time, as this one is. To connect with discrete time, you can also rearrange the integral to go back in time,

$$\int_{t=0}^T e^{-\phi(t-T)} dz_t = \int_{s=0}^T e^{-\phi s} dz_{T-s}.$$

ii. To check this answer, write it as

$$x_t - \mu = e^{-\phi t}(x_0 - \mu) + \sigma \int_{s=0}^t e^{-\phi(t-s)} dz_s$$

and take dx_t . The first term is just the usual time derivative. Then, we take the derivative of the stuff inside the integral as time changes, and lastly we account for the fact that the upper end of the integral changes,

$$\begin{aligned} dx_t &= -\phi e^{-\phi t}(x_0 - \mu) dt + \sigma \int_{s=0}^t (-\phi) e^{-\phi(t-s)} dz_s + \sigma dz_t \\ dx_t &= -\phi \left[e^{-\phi t}(x_0 - \mu) dt + \sigma \int_{s=0}^t e^{-\phi(t-s)} dz_s \right] + \sigma dz_t \\ dx_t &= -\phi(x_t - \mu) dt + \sigma dz_t \end{aligned}$$

(e) This example adds an important tool to our arsenal. Notice we can weight sums of dz_t terms, just as we weight sums of ε_t terms. We can produce or write results as *continuous time moving average processes*

$$x_T = \int_{t=0}^T w(T-t) dz_t$$

7. Once you have “solved an sde” you have expressed x_T as a random variable. Naturally, you will want to know moments – means and variances of x_T . Here are some examples.

(a) Example: Diffusion:

$$\begin{aligned}x_T - x_0 &= \mu T + \sigma \int_{t=0}^T dz_t \\E(x_T - x_0) &= \mu T + \sigma \int_{t=0}^T E(dz_t) = \mu T \\var(x_T - x_0) &= \sigma^2 E\left(\int_{t=0}^T dz_t\right)^2 = \sigma^2 \int_{t=0}^T E(dz_t^2) = \sigma^2 \int_{t=0}^T dt = \sigma^2 T\end{aligned}$$

I used $dz_t dz_s = 0$ for $t \neq s$ in the last line.

(b) Example: Lognormal price process

$$\frac{dp_t}{p_t} = \mu dt + \sigma dz_t.$$

The solution is

$$p_t = p_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma \int_{s=0}^t dz_s}$$

and then we can find the expected value

$$E_0(p_t) = p_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \frac{1}{2}\sigma^2 \int_{s=0}^t ds} = x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \frac{1}{2}\sigma^2 t} = x_0 e^{\mu t}$$

(In the last line I use the fact that for normal y , $E(e^y) = e^{E(y) + \frac{1}{2}\sigma^2(y)}$.)

(c) Example: AR(1). Start with the solution, then use the fact that $E(dz_t) = 0$, $E(dz_t^2) = dt$, $E(dz_t dz_s) = 0$, for $t \neq s$.

$$\begin{aligned}(x_t - \mu) &= e^{-\phi t}(x_0 - \mu) + \int_{s=0}^t e^{-\phi(t-s)} \sigma dz_s \\E_0(x_t - \mu) &= e^{-\phi t}(x_0 - \mu) \\\sigma^2(x_t - \mu) &= E\left[\left(\int_{s=0}^t e^{-\phi(t-s)} \sigma dz_s\right)^2\right] \\&= \int_{s=0}^t e^{-2\phi(t-s)} \sigma^2 ds = \frac{1 - e^{-2\phi t}}{2\phi} \sigma^2\end{aligned}$$

as in discrete time

$$\begin{aligned}E_0(x_{t+j} - \mu) &= \rho^j(x_0 - \mu) \\\var_0(x_t - \mu) &= E\left[\left(\sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}\right)^2\right] = \sum_{j=0}^{t-1} \rho^{2j} \sigma_\varepsilon^2 = \frac{1 - \rho^{2t}}{1 - \rho^2} \sigma_\varepsilon^2\end{aligned}$$

(d) Example: Continuous-time MA. If you have a solution

$$x_t = \int_0^t g(t-s) dz_s,$$

then

$$E_0(x_t) = E_0 \left(\int_0^t g(t-s) dz_s \right) = \int_0^t g(t-s) E_s(dz_s) = 0$$

and

$$E_0(x_t^2) = E_0 \left(\int_0^t g(t-s) dz_s \right)^2 = \int_0^t g(t-s)^2 ds;$$

8. In general, solving SDE's is not so easy! As with regular differential equation, you can't just integrate both sides, because you might have an x on the right side as well that you can't easily get rid of. The idea is simple, from

$$dx_t = \mu(x_t, t)dt + \sigma(x_t, t)dz_t$$

you can write

$$x_T = x_0 + \int_0^T \mu(x_t, t)dt + \int_0^T \sigma(x_t, t)dz_t$$

We end up describing a random variable $f(x_T|x_0)$. The hard part is finding closed-form expressions, when the μ and σ depend on x . The first term is just a standard differential equation. (And remember how "easy" that is!)

9. Simulation. You can easily find distributions and moments by simulating the solution to a stochastic differential equation. Use the random number generator and program

$$x_{t+\Delta} = x_t + \mu(x_t, t)\Delta + \sigma(x_t, t)\sqrt{\Delta}\varepsilon_{t+\Delta}; \varepsilon_{t+\Delta} \sim N(0, 1)$$

10. A last example. The CIR (Cox Ingersoll Ross) *square root process*. Suppose we generate x_t from z_t by

$$dx_t = -\phi(x_t - \mu)dt + \sigma\sqrt{x_t}dz_t.$$

The AR(1) process ranges from $-\infty$ to ∞ . As $x \rightarrow 0$, volatility goes down, and the drift is pulling it back, so x_t can't cross zero. That fact, and the fact that it's more volatile when the level is higher, makes it a good process for nominal interest rates. It's a nonlinear process, not in ARMA class. A strong point for continuous time is that we can handle many of these nonlinear processes in continuous time, though we really can't really do much with them in discrete time. For example, the discrete-time square root process is

$$x_t = \rho x_{t-1} + \sqrt{x_{t-1}}\varepsilon_t$$

Now what do you do? There are closed form solutions for the CIR process and many other nonlinear processes. I won't give those here.

2.5 Finding Moments more generally

(The remaining sections will not be used in *Asset Pricing* or my courses.)

Often you don't really need the whole solution, you only want moments. If x_T is a cash payoff at time T , you might want to know its expected value, $E_t(x_T)$. You might want the expected value

of a function of x_T , i.e. $E_t[\phi(x_T)]$. This happens in option pricing, where ϕ is the call option payoff function. The most basic asset pricing valuation formula is a moment $E_0[\int e^{-\rho t} m_t x_t dt]$.

We saw above how to find moments after you have a full solution. but that's hard. But you can often find moments without first finding the whole "solution" or distribution of the random variable, i.e. x_T or $\phi(x_T)$ itself.

1. But we can also find moments directly because *Moments follow **nonstochastic** differential equations. You don't have to solve stochastic differential equations to find the moments.*

(a) Example.

$$dx_t = \mu dt + \sigma(\cdot) dz$$

Let's suppose we want to find the mean. Now,

$$E_0(x_{t+\Delta}) - E_0(x_t) = d[E_0(x_t)] = E_0(dx_t) = \mu dt$$

Watch the birdie here. By $d[E_0(x_t)]$ I mean, how does the expectation $E_0(x_t)$ move forward in time. But now the point: Since the dz term is missing, you can find the mean of x without solving the whole equation.

(b) Example: lognormal pricing

$$\frac{dx}{x} = \mu dt + \sigma dz$$

The mean follows

$$\begin{aligned} \frac{dE_0(x_t)}{E_0(x_t)} &= \mu dt \\ dE_0(x_t) &= \mu E_0(x_t) dt \\ E_0(x_t) &= x_0 e^{\mu t} \end{aligned}$$

This is the same solution we had before. Look how much easier that was!

- (c) Similarly, to find moments $E_0\phi(x)$, find the diffusion representation for $E_0\phi(x)$ by Ito's lemma.
- (d) Alas, this technique is limited. If $\mu(x_t)$ then $E_0(dx_t) = E_0(\mu(x_t))dt$. You need to know the distribution of x_t to get anywhere.

2. The "*Backward equation.*" What is $E_t[\phi(x_T)]$? We know that $E_T[\phi(x_T)] = \phi(x_T)$. We can work backwards using Ito's lemma to find $E_t[\phi(x_T)]$. We have a process

$$dx = \mu(\cdot)dt + \sigma(\cdot)dz$$

and we want a moment, say

$$E_t[\phi(x_T)].$$

In particular, $E_t(x_T)$ is $\phi(x_t) = x_t$, $E_t(x_T^2)$ is $\phi(x_t) = x_t^2$. Define

$$f(x_t, t) = E_t[\phi(x_T)]$$

Since it's a conditional expectation, $E_t [E_{t+\Delta}(\cdot)] = E_t(\cdot)$, so

$$E_t(df_t) = 0.$$

Applying Ito's lemma to f ,

$$\begin{aligned} df &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dx^2 \\ \frac{E_t(df)}{dt} &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \mu(\cdot) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2(\cdot) = 0 \end{aligned}$$

Thus, the conditional expectation solves the "backward equation"

$$\frac{\partial f(x,t)}{\partial t} + \frac{\partial f(x,t)}{\partial x} \mu(x,t) + \frac{1}{2} \frac{\partial^2 f(x,t)}{\partial x^2} \sigma^2(x,t) = 0$$

starting from the boundary condition

$$f(x, T) = \phi(x)$$

- (a) This is a partial differential equation (ugh). The usual way to solve it is to guess and check – you guess a functional form with some free parameters, and then you see what the free parameters have to be to make it work. It is also the kind of equation you can easily solve numerically. Use the spatial derivatives at time t to find the time derivative, and hence find the function at time $t - \Delta$. Loop. That may be easier numerically than solving for $E_t \phi(x_T)$ by Monte-Carlo simulation of x_T forward. Or, the Monte-Carlo simulation may turn out to be an easy way to solve partial differential equations.

3. Finding densities. You can also find the density by an Ito's lemma differential equation rather than solving the whole thing as we did above. This is called the "Forward equation."

- (a) The density at $t + \Delta$ must be the density at t times the transition density to go from t to $t + \Delta$.

$$\begin{aligned} f(x_{t+\Delta}|x_0) &= \int f(x_t|x_0) f(x_{t+\Delta}|x_t) dx_t \\ &= \int f(x_t|x_0) N \left[\frac{(x_{t+\Delta} - x_t - \mu(x_t)\Delta)}{\sigma(x_t)\sqrt{\Delta}} \right] dx_t \end{aligned}$$

...Ito's lemma and lots of algebra, and allowing explicit $f(x, t)$, ...

$$\frac{\partial f(x, t|x_0)}{\partial t} + \frac{\partial [\mu(x)f(x, t|x_0)]}{\partial x} = \frac{1}{2} \frac{\partial^2 [\sigma^2(x)f(x, t|x_0)]}{\partial x^2}$$

Note the μ and σ^2 are inside the derivative. This is because its $\mu(x_t)\Delta$ not $\mu(x_{t+\Delta})\Delta$.

- (b) This fact gives a nice formula for the stationary (unconditional) density. The formula is:

$$f(x) = k \frac{e^{\int^x -2 \frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(x)}$$

where k is a constant, so we integrate to one. Evaluating k ,

$$f(x) = \left[\int \frac{e^{\int^x 2 \frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(s)} ds \right]^{-1} \frac{e^{\int^x 2 \frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(x)}$$

This is often pretty easy to use.

- i. Derivation: The unconditional density satisfies $\partial f / \partial t = 0$ so from the forward equation,

$$\frac{d[\mu(x)f(x)]}{dx} = \frac{1}{2} \frac{d^2[\sigma^2(x)f(x)]}{dx^2}$$

$$\mu(x)f(x) = \frac{1}{2} \frac{d[\sigma^2(x)f(x)]}{dx}$$

$$2 \frac{\mu(x)}{\sigma^2(x)} [\sigma^2(x)f(x)] = \frac{d[\sigma^2(x)f(x)]}{dx}$$

$$\sigma^2(x)f(x) = k e^{\int^x 2 \frac{\mu(s)}{\sigma^2(s)} ds}$$

$$f(x) = k \frac{e^{\int^x 2 \frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(x)}$$

Normalizing so $\int f(x)dx = 1$,

$$f(x) = \left[\int \frac{e^{\int^x 2 \frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(s)} ds \right]^{-1} \frac{e^{\int^x 2 \frac{\mu(s)}{\sigma^2(s)} ds}}{\sigma^2(x)}$$