12 Continuous time

Much of the theory of finance is much simpler in continuous time. All of the approximations I waved my hands over work exactly in continuous time. You have to invest just a little bit in learning how continuous time series work.

Preview:

Part I continuous time

1. Brownian motion

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

2. Differential

\[ dz_t = \lim_{\Delta \to 0} (z_{t+\Delta} - z_t) \]

3. dz and dt

\[
\begin{align*}
    dz_t &= \sqrt{dt} \\
    dz_t^2 &= dt \\
    E_t(dz_t) &= 0 \\
    var_t(dz_t) &= E_t(dz_t^2) = dz_t^2 = dt
\end{align*}
\]

4. Diffusions

\[
\begin{align*}
    dx_t &= \mu(x_t, t)dt + \sigma(x_t, t)dz_t \\
    E_t(dx_t) &= \mu(x_t, t)dt \\
    \sigma_t^2(dx_t) &= \sigma^2(x_t, t)dt
\end{align*}
\]

5. Examples

\[
\begin{align*}
    \frac{dp_t}{p_t} &= \mu dt + \sigma dz_t \\
    dx_t &= -\phi(x_t - \mu)dt + dz_t
\end{align*}
\]

6. Ito’s lemma

\[
dx_t = \mu dt + \sigma dz_t; y_t = f(x_t) \implies \\

\begin{align*}
    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
\end{align*}
\]

Part II asset pricing in continuous time

1.

\[
p_t = E_t \int_{s=0}^{\infty} e^{-\delta s} \frac{u'(c_{t+s})}{u'(c_t)} x_{t+s} ds = E_t \int_{s=0}^{\infty} \frac{\Lambda_{t+s}}{\Lambda_t} x_{t+s} ds
\]
2. What’s a “return”?

\[ dR_t = \frac{dp_t}{p_t} + \frac{x_t}{p_t} dt \]

3. What’s a “riskfree rate”?

\[ r^f_t dt \]

4. What’s an “excess return”?

\[ dR_t - r^f_t dt = \frac{dp_t}{p_t} + \frac{x_t}{p_t} dt - r^f_t dt \]

5. What’s the equivalent of \( 1 = E(mR) \)?

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

6. What’s the equivalent of \( R^f = 1/E(m) \)?

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

7. What’s the equivalent of \( E(R^e) = -R^f \text{cov}(m, R^e) \)?

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

8. What’s the equivalent of \( m_{t+1} \approx 1 - \delta - \gamma \Delta c_{t+1} \)?

\[ \frac{d\Lambda_t}{\Lambda_t} = -\delta dt - \gamma \frac{dc_t}{c_t} + \gamma (\gamma + 1) \frac{dc_t^2}{c_t^2} \]

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

\[ r^f = \delta dt + \gamma \mu_c - \gamma (\gamma + 1) \sigma_c^2 \]

where \( \frac{dc_t}{c_t} = \mu_c dt + \sigma_c dz_t \)

10. What’s the equivalent of \( E(R^e) \approx \gamma \text{cov}(R^e, \Delta c) \)?

\[ E_t (dR_t) - r^f dt = \gamma E_t \left[ \frac{dc_t dp_t}{c_t p_t} \right] = \gamma \text{cov}_t \left[ \frac{dc_t}{c_t}, \frac{dp_t}{p_t} \right] \]
12.1 Continuous time series

1. Preview. In discrete time we formed time series models such as the AR(1)

\[ x_t = \rho x_{t-1} + \sigma \varepsilon_t \]

by difference equations, building up from a i.i.d. (normal) shock \( \varepsilon_t \) with variance \( \sigma^2(\varepsilon) = 1 \). We “solve” such difference equations to represent

\[ x_t = \sum_{j=0}^{\infty} \rho^j \sigma \varepsilon_{t-j} + \rho^t (x_0) \]

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

2. Random walk-Brownian motion.

(a) This is the building block like \( \varepsilon_t \). Start with the sum \( z_t \)

\[ z_t - z_0 = \sum_{j=1}^{t} \varepsilon_t. \]

As you’ll quickly see, in continuous time it’s easier to start with the cumulative sum \( z_t \) of the \( \varepsilon_t \) shocks, rather than start thinking about what an “instantaneous shock” means. Our “shock series” will then be defined as the difference of \( z_t \).

(b) The variance of \( z \) grows linearly with horizon

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

Remember how a random walk works,

\[ \begin{align*}
    z_{t+2} - z_t &= \varepsilon_{t+1} + \varepsilon_{t+2} \\
    var(z_{t+2} - z_t) &= 2\sigma^2(\varepsilon) \\
    z_{t+k} - z_t &= \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+k} \\
    var(z_{t+k} - z_t) &= k\sigma^2(\varepsilon)
\end{align*} \]

(c) Thus, generalize this idea for any time interval. With \( \sigma = 1 \) (we can later multiply by a number \( \sigma \)) let’s define

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

for any \( \Delta \), not just integers. This is a Brownian motion.

(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 \( E(z_{t+\Delta} - z_t, z_{t+\Delta} - z_t) = 0 \), or more generally the same thing for any nonoverlapping interval.

(e) \( dz \). Define

\[ \begin{align*}
    dz_t &= \lim_{\Delta \to 0} (z_{t+\Delta} - z_t) \\
    dz_t &\leftrightarrow \varepsilon_t \end{align*} \]

This is a tiny forward-difference operator.
(f) \( dz = \sqrt{dt} \). Note

\[
\sigma^2(z_{t+\Delta} - z_t) = \Delta; \\
\sigma(z_{t+\Delta} - z_t) = \sqrt{\Delta}
\]

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 \( dz/dt \to +/ - \infty \), moving “infinitely fast” (up and down) (It’s fractal; jumpy at any time scale; it has an infinite path length). That’s why we do not write the usual integral notation

\[
\text{Yes} : ~ z_t - z_0 = \int_0^t dz_s \\
\text{No} : ~ z_t - z_0 = \int_0^t \left( \frac{dz(s)}{ds} \right) ds
\]

d\( z \), d\( z_t \), and d\( z(t) \) are equivalent notations.

(g) Moments:

\[
E_t(dz_t) = 0; \iff E_t(z_{t+\Delta} - z_t) = 0 \\
\text{var}_t(dz_t) = dt; \iff \text{var}(z_{t+\Delta} - z_t) = \Delta, \\
\text{cov}(dz_t, dz_s) = 0 \iff \text{cov}(z_{t+\Delta} - z_t, z_{s+\Delta} - z_s) = 0
\]

Watch out for notation. \( E_t(dz_t) \neq d\!z_t \). \( d \) 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 \( z_{t+\Delta} - z_t \)

(h) \( d\!z_t^2 = dt \). Variances overwhelm means, so variance and second moment are the same. \( \text{var}_t(dz_t) = E_t(dz_t^2) \). Moreover, \( d\!z_t^2 = dt \) means that

\[
\text{var}_t(dz_t) = E_t(dz_t^2) = d\!z_t^2 = dt
\]

Second moments are nonstochastic! Similarly if we have two Brownians

\[
\text{cov}_t(dz_t, dw_t) = E_t(dz_t dw_t) = d\!z_t d\!w_t.
\]

You may get a sense already how continuous time will simplify lots of things.

(i) (More on the last point, optional note. \( d\!z_t^2 \) is not a \( \chi^2 \); it’s a number. (To see why, compare \( dz_t \) and \( d\!z_t^2 \). \( z_{t+\Delta} - z_t \) is \( N(0, \Delta) \), \( (z_{t+\Delta} - z_t)/\sqrt{\Delta} \cdot N(0, 1) \) so the probability that, say, \( ||z_{t+\Delta} - z_t|| > 2\sqrt{\Delta} \) is 5%. On the other hand, \( (z_{t+\Delta} - z_t)^2/\Delta \cdot \chi^2 \). \( \chi^2 \) or “chi-squared with one degree of freedom” just means “the distribution of the square of a standard normal.”) that means that the probability that, \( (z_{t+\Delta} - z_t)^2 > 2\Delta \) is fixed as \( \Delta \) gets small. \( (z_{t+\Delta} - z_t)^2 \) is of order \( dt \) and hence nonstochastic.)

3. Diffusion processes. We have the building block, the analog to \( \varepsilon_t \). Now, we build more complex process like AR(1), ARMA, etc.

(a) Random walk with drift and scaling the diffusion variance, model

\[
\begin{align*}
\text{discrete:} & \quad x_t = \mu + x_{t-1} + \varepsilon_t \\
\text{continuous:} & \quad dx_t = \mu dt + \sigma dz_t
\end{align*}
\]
(b) Moments. Practice in $dz$ and $dt$. 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. The 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. Thus,

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

(c) Geometric growth (standard asset price process.)

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

We’ll see how to add dividends in a moment. (It’s really $dp_t = \mu p_t dt + \sigma p_t dz_t$ for sticklers.) You can also see here the reason we call it a “stochastic differential equation.” Without the $dz$ term, we would “solve” this 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 $dz$ term.

(d) AR(1), Ornstein Uhlenbeck process

**discrete:**

$$ x_{t+1} = (1 - \rho)x_t + \rho x_t + \sigma \xi_{t+1} $$

$$ x_{t+1} - x_t = -(1 - \rho)(x_t - \mu) + \sigma \xi_{t+1} $$

$$ E_t(x_{t+1} - x_t) = -(1 - \rho)(x_t - \mu) $$

$$ \text{var}_t(x_{t+1} - x_t) = \sigma^2 \xi_{t+1} $$

**continuous:**

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

$$ E_t dx_t = -\phi(x_t - \mu)dt $$

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

The part in front of $dt$ is called the “drift” and the part in front of $dz_t$ is called the “diffusion” coefficient. This is the Ornstein Uhlenbeck process.

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

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

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

$$ E_t (dx_t) = \mu(x_t, t, ...) dt $$

$$ \text{var}_t(dx_t) = \sigma^2(x_t, t, ...) dt $$
4. **Ito’s lemma.** More building. Given a diffusion process for \( x_t \), how do we find the diffusion for \( y_t = f(x_t) \)?

(a) Standard 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. \( dx = \mu dt \). What if \( dx = \mu dt + \sigma dz \)?

(b) So our question:

\[ dx = \mu dt + \sigma dz \]
\[ y = f(x) \]
\[ dy = ? \]

Answer:

- Do a second order Taylor expansion. Keep terms of order \( dt \) and \( dz = \sqrt{dt} \), ignore higher order terms. Don’t forget \( dz^2 = dt \).

The answer is called Ito’s lemma:

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

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. “Convexity” in option and bond pricing.

(c) An example. Geometric growth (used all the time for asset pricing)

\[ \frac{dx_t}{x_t} = \mu dt + \sigma dz \]
\[ y_t = \ln x_t \]
\[ dy_t = \frac{1}{x_t} dx_t - \frac{1}{2} x_t^2 dt = \left( \mu - \frac{1}{2} \sigma^2 \right) \]
\[ dt + \sigma dz \]

(d) Conversely,

\[ dx_t = \mu dt + \sigma dz \]
\[ y_t = e^{x_t} \]
\[ dy_t = e^{x_t} dx_t + \frac{1}{2} e^{x_t} dx_t^2 \]
\[ \frac{dy_t}{y_t} = \left( \mu + \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t \]

(e) Another example. *Use the chain rule but go out to second derivatives.*

\[ d(x_t y_t) = y_t dx_t + x_t dy_t + dx_t dy_t \]
(f) If \( y = f(x_t, t) \)

then

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

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

12.2 Asset Pricing in Continuous Time.

I will follow exactly the same outline as we did in discrete time. Logic same, symbols different.

1. Basic asset pricing equation with utility and discount factor. Let \( x_t \) be a stream of dividends,

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

2. Levels and differences.

\[
m_{t+1} = e^{-\delta} \frac{u'(c_{t+1})}{u'(c_t)} = \frac{\Lambda_{t+1}}{\Lambda_t}.
\]

As it was prettier to start with the level \( z_t \) rather than the difference \( \varepsilon_t \), it’s easier to start with the level of marginal utility \( \Lambda_t \) not the difference \( m_{t+1} \)

3. What’s a “return”? We generalize the net return,

\[
dR_t = \frac{p_{t+\Delta} - p_t + x_t \Delta}{p_t} \Rightarrow \frac{dp_t}{p_t} + \frac{x_t}{p_t} dt
\]

4. What’s a “riskfree rate”? Like a money market fund, a security with a constant price \( p_t = 1 \) so \( dp_t/p_t = 0 \), that pays interest,

\[
r^f_t dt
\]

5. What’s an “excess return”?\)

\[
dR_t - r^f_t dt = \frac{dp_t}{p_t} + \frac{x_t}{p_t} dt - r^f_t dt
\]

6. What’s the equivalent of \( 1 = E(mR)? \)

\[
E_t \left[ \frac{d(\Lambda_t p_t)}{\Lambda_t p_t} \right] - \frac{x_t}{p_t} dt = 0
\]
Not so pretty, but it will get prettier soon. Derivation:

\[ p_t = E_t \int_{s=0}^{\Delta} \frac{\Lambda_{t+s}}{\Lambda_t} x_{t+s} ds + E_t \int_{s=\Delta}^{\infty} \frac{\Lambda_{t+s}}{\Lambda_t} x_{t+s} ds \]

\[ p_t \Lambda_t \approx x_t \Lambda_t \Delta + E_t [\Lambda_{t+\Delta \delta x_{t+\Delta}}] \]

\[ 0 \approx x_t \Lambda_t \Delta + E_t [\Lambda_{t+\Delta \delta x_{t+\Delta}} - \Lambda_t \delta x_t] \]

\[ 0 = \frac{x_t}{p_t} \delta x_t + E_t \left[ \frac{d (\Lambda_t \delta x_t)}{\Lambda_t \delta x_t} \right] \]

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

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

Derivation: Apply what we just did to \( p_t = 1, x_t = r^f \),

\[ 0 = r^f dt + E_t \left[ \frac{d \Lambda_t}{\Lambda_t} \right] \]

8. What’s the equivalent of \( E(R^c) = R^f \text{cov}(m, R^c) \)?

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

Well, that’s a lot prettier! Derivation: As in discrete time we open up the second moment

\[ 0 = \frac{x_t}{p_t} \delta x_t + E_t \left[ \frac{d (\Lambda_t \delta x_t)}{\Lambda_t \delta x_t} \right] = \frac{x_t}{p_t} \delta x_t + E_t \left[ \frac{d \Lambda_t}{\Lambda_t} + \frac{dp_t}{p_t} + \frac{d \Lambda_t dp_t}{\Lambda_t p_t} \right] \]

don’t forget the last term – we keep second derivatives and \( dx^2 \) terms! Continuing,

\[ 0 = E_t \left[ \frac{dp_t}{p_t} \right] + \frac{x_t}{p_t} \delta x_t - r^f dt + E_t \left[ \frac{d \Lambda_t dp_t}{\Lambda_t p_t} \right] \]

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

9. Consumption and asset pricing in continuous time.

(a) Marginal utility and consumption. (What’s the equivalent of \( m_{t+1} \approx 1 - \delta - \gamma \Delta c_{t+1} \)?)

\[ \frac{d \Lambda_t}{\Lambda_t} = -\delta dt - \gamma \frac{dc_t}{c_t} + \gamma (\gamma + 1) \frac{dc_t^2}{c_t^2} \]

Derivation: Ito’s lemma!

\[ \Lambda_t = e^{-\delta t} u'(c_t) \]

\[ d\Lambda_t = -\delta e^{-\delta t} u'(c_t) dt + e^{-\delta t} u''(c_t) dc_t + \frac{1}{2} e^{-\delta t} u'''(c_t) dc_t^2 \]

\[ \frac{d \Lambda_t}{\Lambda_t} = -\delta dt - \left( -\frac{c_t u''(c_t)}{u'(c_t)} \right) \frac{dc_t}{c_t} + \frac{1}{2} \frac{c_t^2 u'''(c_t)}{u'(c_t)} dc_t^2 \]
For power utility

\[ u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}; \]

\[ u'(c_t) = c_t^{-\gamma}; \]

\[ u''(c_t) = -\gamma c_t^{-\gamma-1} \implies -\frac{c_t u''(c_t)}{u'(c_t)} = \gamma; \]

\[ u'''(c_t) = \gamma(\gamma + 1)c_t^{-\gamma-2} \implies \frac{c_t^2 u'''(c_t)}{u'(c_t)} = \gamma(\gamma + 1) \]

In general you can use the local derivatives of the utility function.

(b) Risk free rate (What’s the equivalent of \( R_t \approx \delta + \gamma E(\Delta c) \)?)

\[ r^f dt = -E_t \left[ \frac{d\Lambda_t}{\Lambda_t} \right] = E_t \left[ \delta dt + \gamma \frac{dc_t}{c_t} - \gamma(\gamma + 1) \frac{dc_t^2}{c_t^2} \right] \]

\[ r^f dt = \delta dt + \gamma E_t \left( \frac{dc_t}{c_t} \right) - \gamma(\gamma + 1) \frac{dc_t^2}{c_t^2} \]

i.e. if consumption follows

\[ \frac{dc_t}{c_t} = \mu_c dt + \sigma_c dz_t \]

\[ r^f = \delta dt + \gamma \mu_c - \gamma(\gamma + 1) \sigma_c^2 \]

(c) Comment on precautionary saving. I abstracted from the variance term in discrete time. It’s “precautionary saving.” If volatility is high, and people risk averse, then they try to save, sending interest rates up. It is conventionally thought to be small. If \( \sigma_c = 0.02 \) (2% per year) and \( \gamma = 2 \), then \( \gamma(\gamma + 1)\sigma_c^2 = 2 \times 3 \times 0.02^2 = 0.0024 \) or 0.2 percentage points, not a big effect. But once we allow \( \gamma = 20 \) to explain the equity premium, then \( 20 \times 21 \times 0.02^2 = 0.168 = 16.8 \) percentage points, and precautionary saving might be very important!

(d) Risk premium and consumption. What’s the equivalent of \( E(R^e) \approx \gamma \text{cov}(R^e, \Delta c) \)?

\[ E_t (dR_t) - r^f dt = \gamma E_t \left[ \frac{dc_t}{c_t} \frac{dp_t}{p_t} \right] = \gamma \text{cov}_t \left[ \frac{dc_t}{c_t}, \frac{dp_t}{p_t} \right] \]

Derivation.

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

\[ E_t (dR_t) - r^f dt = -E_t \left\{ -\delta dt - \gamma \frac{dc_t}{c_t} + \gamma(\gamma + 1) \frac{dc_t^2}{c_t^2} \right\} \frac{dp_t}{p_t} \]

Now all the terms but \( dc_t/c_t \) drop, because they will end up being of order greater than \( dt \), so we’re done.
12.3 Solving stochastic differential equations.

1. Stochastic integral. Naturally, if $dz$ are the little changes, you get $z$ back if you add them up.

\[
  z_t - z_0 = \int_{s=0}^{t} dz_s \iff z_t - z_0 = \sum_{j=1}^{t} \varepsilon_j
\]

Thus $\int_0^t dz_s$ is just a normally distributed random variable with mean zero and variance $s$.

\[
  \int_0^t dz_s \sim N(0, t)
\]

2. 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,”

\[
  \int_0^t dz_s = (z_\Delta - z_0) + (z_{2\Delta} - z_\Delta) + (z_{3\Delta} - z_{2\Delta}) + \ldots + (z_t - z_{t-\Delta})
\]

3. Solving: the idea. In discrete time, we “solve” the difference equation:

\[
  x_t = (1 - \rho)\mu + \rho x_{t-1} + \varepsilon_t
\]

\[
  x_t - \mu = (1 - \rho)^t (x_0 - \mu) + \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}.
\]

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.

\[
  x_t | x_0 \sim N \left[ \mu + (1 - \rho)^t (x_0 - \mu), \sigma^2 \sum_{j=0}^{t-1} \rho^{2j} \right]
\]

Let’s just do the same thing in continuous time.

4. Let’s just do some examples. I won’t use anything fancier than the basic idea: plop $\int_s^t$ on both sides of an expression and interpret.

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

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

(b) Random walk with drift

\[
  dx_t = \mu dt + \sigma dz_t
\]

\[
  \int_{t=0}^{T} dx_t = \int_{t=0}^{T} \mu dt + \int_{t=0}^{T} \sigma dz_t
\]

\[
  x_T - x_0 = \mu T + \sigma \int_{t=0}^{T} dz_t
\]

\[
  x_T - x_0 = \mu T + \varepsilon; \varepsilon \sim N(0, T\sigma^2)
\]

\[
  x_T - x_0 \sim N(\mu T, T\sigma^2)
\]
As in discrete time,
\[ x_t = x_0 + \mu t + \sum_{j=1}^{t} \varepsilon_j \]

(c) Lognormal price process
\[ \frac{dx_t}{x_t} = \mu dt + \sigma dz_t \]

By Ito’s lemma
\[
d\ln x_t = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t
\]
\[
\int_0^T d\ln x_t = \int_0^T \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \int_0^T \sigma dz_t
\]
\[
\ln x_T - \ln x_0 = \left( \mu - \frac{1}{2} \sigma^2 \right) T + \sigma \int_0^T dz_t
\]
\[
x_T = x_0 e^{(\mu - \frac{1}{2} \sigma^2)T + \sigma \int_0^T dz_t}
\]

i.e.,
\[ x_T = x_0 e^{(\mu - \frac{1}{2} \sigma^2)T + \sigma \sqrt{T} \varepsilon} \]
\[ \varepsilon \sim N(0, 1) \]
\[ x_T \] is lognormally distributed.

(d) AR(1)
\[ dx_t = -\phi(x_t - \mu)dt + \sigma dz_t \]

The solution is,
\[ x_T - \mu = e^{-\phi T}(x_0 - \mu) + \sigma \int_{t=0}^{T} e^{-\phi(T-t)}dz_{T-t} \]

this looks just like the discrete time case
\[ x_T - \mu = \rho^T(x_0 - \mu) + \sum_{j=0}^{T-1} \rho^j \varepsilon_{t-j}. \]

It adds an important tool to our arsenal. Notice we can weight sums of \( dz \) terms, just as we weight sums of \( \varepsilon \) terms. We produce all sorts of interesting processes with \( \int_{t=0}^{T} w(t)dz_t \), just as we produce all sorts of interesting functions of time (deterministic processes) with weights \( \int_{t=0}^{T} w(t)dt \).

To derive this answer, you have to be a little clever. Find
\[
d\left( e^{\phi t} (x_t - \mu) \right) = \phi e^{\phi t} (x_t - \mu) dt + e^{\phi t} dx_t
\]
\[
= \phi e^{\phi t} (x_t - \mu) dt + e^{\phi t} [-\phi(x_t - \mu)dt + \sigma dz_t]
\]
\[
= \sigma e^{\phi t} dz_t
\]

Now integrate both sides,
\[
\int_{t=0}^{T} d\left( e^{\phi t} (x_t - \mu) \right) = \int_{t=0}^{T} \sigma e^{\phi t} dz_t
\]
\[
e^{\phi T} (x_T - \mu) - (x_0 - \mu) = \sigma \int_{t=0}^{T} e^{\phi t} dz_t
\]
\[
(x_T - \mu) = e^{-\phi T}(x_0 - \mu) + \sigma \int_{t=0}^{T} e^{\phi(T-t)}dz_t
\]
and rearrange the last integral to go back in time, if you wish. To check this answer, write it as

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

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,

$$dx_t = -\phi e^{-\phi t} (x_0 - \mu) dt + \sigma \int_{s=0}^{t} (-\phi) e^{-\phi(t-s)} dz_{t-s} + \sigma dz_t$$

I expressed the solutions with integrals going back in time, as that is closer to the discrete time tradition. It’s more common in continuous time to write them with the integrals going forward in time.

(e) In general, it’s not so easy! 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. The first term is just a standard differential equation. (And remember how “easy” that is!)

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

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

(g) CIR 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 $\infty$. As $x \to 0$, volatility goes down here, so $x$ 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. Note it’s a nonlinear process, not in ARMA class. 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

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

Now what do you do?? The fact that there are closed form solutions for the CIR process is a great advantage.
5. **Finding Moments.** Often you don’t really need the whole solution, you only want moments, i.e. $E_t(x_T)$, or $E_t[\phi(x_T)]$ (option pricing) or $E_0[\int e^{-\rho t} m_t x_t dt]$ (asset pricing).

(a) Of course you *can* solve the whole thing first and then take moments (if you can solve the whole thing).

i. Example: lognormal price process

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

$$x_t = x_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma \int_0^t dz_s}$$

$$E_0(x_t) = x_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \frac{1}{2} \sigma^2 s} = x_0 e^{\mu t}$$

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

ii. Example 2: AR(1). Start with the solution, then use the fact that $E(dz_t) = 0, E(dz_t^2) = 1, E(dz_t dz_s) = 0$,

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

$$E_0(x_t - \mu) = e^{-\phi t} (x_0 - \mu)$$

$$\sigma^2(x_t - \mu) = E \left[ \left( \int_{s=0}^{t} e^{-\phi(t-s)} \sigma dz_s \right)^2 \right]$$

$$= \int_{s=0}^{t} e^{-2\phi(t-s)} \sigma^2 dt = \frac{1 - e^{-2\phi t}}{2\phi} \sigma^2$$

as in discrete time

$$E_0(x_{t+j} - \mu) = \rho^j (x_0 - \mu)$$

$$\text{var}_0(x_t - \mu) = E \left[ \left( \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j} \right)^2 \right] = \sum_{j=0}^{t-1} \rho^{2j} \sigma^2 = \frac{1 - \rho^{2t}}{1 - \rho^2} \sigma^2$$

iii. What we just did for the AR(1) is more generally true and useful. Since $dz_s$ is unpredictable like $\varepsilon_{t+1}$, it means $g(s)$ and $dz_s$ are independent. Thus,

$$E_0 \left( \int_{0}^{t} g(s) dz_s \right) = E_0 \left( \int_{0}^{t} g(s) E_s (dz_s) \right) = 0$$

and

$$E_0 \left( \int_{0}^{t} g(s) dz_s \right)^2 = \int_{0}^{t} E_0 [g(s)^2] ds$$

(b) 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.*

i. Example.

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

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.
ii. Example: lognormal pricing

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

The mean follows

\[
\begin{align*}
\frac{dE_0(x_t)}{E_0(x_t)} &= \mu dt \\
dE_0(x_t) &= \mu E_0(x_t) dt \\
E_0(x_t) &= x_0 e^{\mu t}
\end{align*}
\]

This is the same solution we had before.

iii. Similarly, to find moments \( \phi(x) \), find the diffusion representation for \( \phi(x) \) by Ito’s lemma.

iv. Alas, this technique is limited. If \( \mu(x_t) \) then \( E_0(dx_t) = E_0(\mu(x_t))dt \).

(c) The “Backward equation.” 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 \( f \),

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

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

\[ f(x, T) = \phi(x) \]

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.
6. 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 $t + \Delta$.

$$f(x_{t+\Delta}|x_0) = \int f(x_t|x_0)f(x_{t+\Delta}|x_t)dx_t$$

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

...Ito’s lemma and lots of algebra...

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

Note the $\mu$ and $\sigma^2$ 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:

$$f(x) = k e^{\int_0^x \frac{\mu(s)}{\sigma^2(s)} ds}$$

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

$$f(x) = \left[ \int e^{\int_0^x \frac{2\mu(s)}{\sigma^2(s)} ds} \frac{1}{\sigma^2(s)} ds \right]^{-1} e^{\int_0^x \frac{2\mu(s)}{\sigma^2(s)} ds}$$

This is often pretty easy to use.

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

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

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

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

$$\sigma^2(x)f(x) = ke^{\int_0^x \frac{2\mu(s)}{\sigma^2(s)} ds}$$

$$f(x) = k e^{\int_0^x \frac{2\mu(s)}{\sigma^2(s)} ds}$$

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

$$f(x) = \left[ \int e^{\int_0^x \frac{2\mu(s)}{\sigma^2(s)} ds} \frac{1}{\sigma^2(s)} ds \right]^{-1} e^{\int_0^x \frac{2\mu(s)}{\sigma^2(s)} ds}$$