Study Guide on Modeling Financial Series for the Casualty Actuarial Society (CAS) Exam 7

(Based on Gary Venter's Paper, "Advances in Modeling Financial Series")

Study Guide Created by G. Stolyarov II, ARe, AIS - *Spring 2011*

Source: Venter, G.G., "Advances in Modeling of Financial Series," Society of Actuaries, 2010.

This is an open-source study guide and may be revised pursuant to suggestions.

S7-AMFS-1. Let X be a time series of the inflation rates from November 2002 to December 2011.
(a) Let Y be the *first lag* of X. What does Y stand for in real terms?
(b) Let Z be the *second lag* of X. What does Z stand for in real terms?
(c) Let L be the lag operator. Represent Y in terms of L and X.
(d) Represent Z in terms of L and X.
(e) Represent the *first difference* of series X in terms of X and Y.
(f) Represent the *first difference* of series X in terms of L and X.
(g) Represent the *second difference* of series X in terms of X, Y, and Z.
(h) Represent the *second difference* of series X in terms of L and X.
(Venter, p. 3)

Solution S7-AMFS-1.
(a) Y is the time series of inflation rates from October 2002 to November 2011.
(b) Z is the time series of inflation rates from September 2002 to October 2011.
(c) Y = LX.
(d) $Z = L^2X$.
(e) First difference is X-Y.
(f) First difference is X-LX = (1-L)X.
(g) Second difference is the first difference of the first difference: (X – Y) – (Y – Z) = X + Z – 2Y.
(h) Second difference is $(1-L)^2X$.

S7-AMFS-2.
(a) The *first autocorrelation* of a time series X is a correlation with *what*?
(b) The *second autocorrelation* of a time series X is a correlation with *what*?
(c) What is the implication for a shock if most autocorrelations in the time series are zero?
(d) What is the implication for a shock if autocorrelations for many lags in the time series are high? (Venter, p. 3)

Solution S7-AMFS-2.
(a) First autocorrelation is between **X and LX.**
(b) Second autocorrelation is between **X and L^2X.**
(c) If most autocorrelations are zero, the effect of a shock quickly goes away. This is because the next event does not depend so much on the previous events.
(d) If many autocorrelations are high, then the effect of the shock can persist for a long time.

S7-AMFS-3. Give the formula for the AR(1) time-series model (first-order autoregressive process). Define the variables used. (Venter, p. 3)

Solution S7-AMFS-3.
AR(1) formula: $\mathbf{r_{i+1} = a + br_i + s\varepsilon_{i+1}}$.
ε_{i+1} is the standard normal variate.
a, b, and s are parameters.
r_0 is the starting value in the series; r_i is the (i+1)st value.

S7-AMFS-4. In the AR(1) formula, what is the role of the parameter b? What happens as b approaches 1, and what happens as b approaches 0? (Venter, p. 3)

Solution S7-AMFS-4. The parameter b indicates the magnitude of the autocorrelation. The kth autocorrelation is b^k. As b approaches 1, the autocorrelation between r_i and r_{i+1} increases and persists longer. As b approaches 0, the autocorrelation decreases and quickly disappears.

S7-AMFS-5. The following questions pertain to the AR(1) time-series model.
(a) What is the expected value of r_t for some value of t?
(b) What is the limit of the expected value of r_t as t increases without bound?
(c) What is the variance of r_t?
(d) What is the limit of the variance of r_t as t increases without bound?
(Venter, p. 3)

Solution S7-AMFS-5.
(a) $\mathbf{E(r_t) = r_0b^t + a(1-b^t)/(1-b)}$.
(b) Since $|b| < 1$, as $t \to \infty$, $b^t \to 0$. Thus, $\mathbf{\lim_{t\to\infty}E(r_t) = a/(1-b)}$.
(c) $\mathbf{Var(r_t) = s^2(1-b^{2t+1})/(1-b^2)}$.
(d) $\mathbf{\lim_{t\to\infty}Var(r_t) = s^2/(1-b^2)}$.

S7-AMFS-6. Explain how the simple random walk model is a special case of the AR(1) model, and give the formula to show it. (Venter, p. 4)

Solution S7-AMFS-6.
The simple random walk model is the AR(1) model with a = 0 and b = 1:
$\mathbf{r_{i+1} = r_i + s\varepsilon_{i+1}}$.

S7-AMFS-7.
(a) Give the formula for the kth autocorrelation in a simple random walk model for a time series with t observations.
(b) Using a simple random walk model for a time series with 90 observations, find (i) the second autocorrelation, (ii) the 36^{th} autocorrelation, and (iii) the 78^{th} autocorrelation.
(c) Fill in the blank: In practice, it is difficult to distinguish a simple random walk from _____.
(Venter, p. 4)

Solution S7-AMFS-7.
(a) The kth autocorrelation in a simple random walk model for a time series with t observations is $(1 - k/t)^{1/2}$.
(b) (i) Second autocorrelation: $(1 - 2/90)^{1/2}$ = **0.988826469**.
(ii) 36th autocorrelation: $(1 - 36/90)^{1/2}$ = **0.7745966692**.
(iii) 78th autocorrelation: $(1 - 78/90)^{1/2}$ = **0.3651483717**.
(c) In practice, it is difficult to distinguish a simple random walk from **an AR(1) process with a high value of b (close to 1).**

S7-AMFS-8. Venter, on page 4, describes a form of persistent autocorrelation that is more difficult to model. What is it?

Solution S7-AMFS-8. The autocorrelation may be significantly less than 1 at lag 1 (e.g., the autocorrelation may be closer to ½), *but* subsequent autocorrelations decline very slowly.

S7-AMFS-9. Give the binomial expansion formula for the dth difference of a time-series lag: $(1\text{-}L)^d$. (Venter, p. 5).

Solution S7-AMFS-9. We use the following combination notation:
C(a, b) = a!/(b!*(a-b)!).
$(1\text{-}L)^d = {}_{k=0}{}^{\infty}\sum[C(d, k)*(-L)^k] = 1 - dL + d(d-1)L^2/2! - d(d-1)(d-2)L^3/3! + \ldots$.

S7-AMFS-10. Take the AR(1) formula ($r_{i+1} = a + br_i + s\varepsilon_{i+1}$) and transform it to represent a mean-reverting process, where m is the long-term mean a/(1-b), and c = 1-b. (Venter, p. 5)

Solution S7-AMFS-10.
Since m = a/(1-b) and c = 1 – b, it follows that m = a/c, and a = mc. Also, b = 1-c.
$r_{i+1} = a + br_i + s\varepsilon_{i+1} \rightarrow$
$r_{i+1} = mc + (1\text{-}c)r_i + s\varepsilon_{i+1} \rightarrow$
$\mathbf{r_{i+1} = r_i + c(m - r_i) + s\varepsilon_{i+1}}$.

S7-AMFS-11.
(a) Let m be an AR(1) process such that $m_{i+1} = m_i + h(\mu - m_i) + \sigma\eta_{i+1}$, and let r be a double mean-reverting process that incorporates m. What is the formula for process r?
(b) Describe this process in qualitative terms.
(c) How else can a double mean-reverting process be described in terms of AR(1) processes? (Venter, pp. 5-6)

Solution S7-AMFS-11.
(a) $r_{i+1} = r_i + c(m_i - r_i) + s\varepsilon_{i+1}$.
(b) In this process, r reverts toward the temporary mean m, and m reverts toward its long-term mean μ.
(c) A double mean-reverting process can be described as the *sum* of two AR(1) processes, one of which has a mean of zero.

S7-AMFS-12. Fill in the blanks (Venter, p. 6):
(a) Brownian motion is a _____ version of the random walk.
(b) In Brownian motion, the change from time q to time q+t is ____ distributed with mean _____ and variance ________. For standard Brownian motion, the parameter ____ is equal to 1.

Solution S7-AMFS-12.
(a) Brownian motion is a **continuous** version of the random walk.
(b) In Brownian motion, the change from time q to time q+t is **normally** distributed with mean **0** and variance **st^2**. For standard Brownian motion, the parameter **s** is equal to 1.

S7-AMFS-12.
(a) Give the formula for Brownian motion that incorporates a deterministic time trend.
(b) Give the formula for Brownian motion that incorporates a *mean-reverting* deterministic time trend.
(c) Give the formula for *geometric* Brownian motion. (Venter, p. 6)

Solution S7-AMFS-12. (a) $dr_t = a*dt + s*dW_t$. The deterministic component is a*dt.
(b) $dr_t = a*(m-r_t)*dt + s*dW_t$. The mean is m.
(c) $d[\ln(r_t)] = dr_t/r_t = s*dW_t$.

S7-AMFS-13.
(a) Fill in the blanks: In a compound Poisson process, the number of events in time t is Poisson-distributed in ___, and each event size is a(n) _________ draw from a single distribution.
(b) Give the formula for a compound Poisson process if N(μ) denotes the number of events with Poisson mean μ and X_k is the kth jump size. (Venter, p. 6)

Solution S7-AMFS-13.
(a) In a compound Poisson process, the number of events in time t is Poisson-distributed in **λt**, and each event size is an **independent** draw from a single distribution.
(b) $dr_t = d[{}_{i=0}{}^{N(\lambda t)}\sum(X_i)]$.

S7-AMFS-14. Describe a common way of simulating Brownian motion and compound Poisson processes in order to more realistically represent the instantaneous change dt. (Venter, p. 6)

Solution S7-AMFS-14. Instead of having an instantaneous change, short time periods (such as single days) are used in place of dt. For instance, if there are 252 trading days in a year, then t = 1/252 represents one trading day, and the variance is calculated as $st^2 = s/252^2$.

S7-AMFS-15. What are the two most common short-rate models for modeling Treasury interest rates? Give the name and the formula for each model. (Venter, p. 7)

Solution S7-AMFS-15. The two most common short-rate models for modeling Treasury interest rates are as follows:
1. **Vasicek model:** $dr_t = (b-ar_t)*dt + s*dW_t$.
2. **Cox-Ingersoll-Ross (CIR) model:** $dr_t = (b-ar_t)*dt + s*r_t^{1/2}*dW_t$.

S7-AMFS-16. Identify three advantages and one disadvantage of the Cox-Ingersoll-Ross (CIR) model, as discussed by Venter on p. 7.

Solution S7-AMFS-16.
Advantages
1. The CIR model corresponds to empirical observations that higher interest rates are associated with higher volatility.
2. The CIR model produces more realistic heavier-tailed distributions of rates.
3. The CIR model makes it impossible to produce negative interest rates, since the random term is zero when the short rate is zero.

Disadvantage
The CIR model requires simulating on short intervals and complicates the estimation of parameters, since the distribution for a short interval is approximately normal, but this is not the case for a longer interval.

S7-AMFS-17. What is the assumption of the theory of arbitrage-free yield curves regarding the price of a bond? (Venter, p. 7)

Solution S7-AMFS-17. The price of a bond is assumed to be the expected present value of the bond payments, discounted back along all possible paths of the short rate, from now to the time of maturity.

S7-AMFS-18. If the market price of risk is λ, what term could be added to the Vasicek and CIR models to get the risk-neutral short-rate process? (Venter, p. 8)

Solution S7-AMFS-18. The term to be added is $\lambda*s*r_t*dt$.

S7-AMFS-19.
(a) Let P(t, T) be the price of a bond at time t and maturity time of T. Let R(t, T) be the continuously compounded interest rate for this bond at time t. Suppose that the payment that occurs at T has magnitude 1. Give the mathematical relationship between 1 and P(t, T) and R(t, T).
(b) Now, instead of R(t, T), suppose that you have an *annually* compounded rate Y(t, T). Give the mathematical relationship between 1 and P(t, T) and Y(t, T).

Solution S7-AMFS-19.
(a) 1 = P(t, T)*exp[R(t, T)*(T-t)].
(b) 1 = P(t, T)*[1 + Y(t,T)]$^{T-t}$.

S7-AMFS-20.
(a) According to Venter (p. 9), what is the drawback of closed-form yield-curve formulas? How does this drawback adversely affect risk management?
(b) What is a possible way to remedy this problem?

Solution S7-AMFS-20.
(a) The drawback of closed-form yield-curve formulas is that the market price of risk and the short rate are the only variables that determine the entire yield curve, and this limits the kinds of shapes that can occur. For risk management, this is a problem, because some yield curves are overrepresented in the model while others are omitted – which means that all of the possible scenarios are not available, and a distorted picture of risk may emerge.
(b) A possible way to remedy this problem is to use multifactor Vasicek and CIR models – either by creating double-mean-reverting processes or expressing the interest rate as a sum of two partial interest rates.

S7-AMFS-21. Fill in the blanks (Venter, p. 9):
(a) The interest rates generated from the Vasicek model follow a ________ distribution.
(b) The interest rates generated from the CIR model are distributed as a sum of a series of _______ distributions.
(c) The two-factor CIR model requires _____________ in order to have closed-form yield curves.

Solution S7-AMFS-21.
(a) The interest rates generated from the Vasicek model follow a **normal** distribution.
(b) The interest rates generated from the CIR model are distributed as a sum of a series of **gamma** distributions.
(c) The two-factor CIR model requires **independent factors (i.e., uncorrelated factors)** in order to have closed-form yield curves.

S7-AMFS-22. Give the formulas for the Andersen and Lund model, generalizing the CIR model to fit US Treasury rates. Briefly explain the model in conceptual terms. (Venter, p. 10)

Solution S7-AMFS-22.
The Andersen and Lund model has three formulas:
(i) $\mathbf{dr_t = (b_t – a*r_t)dt + s_t*r_t^p*dW_{r\_t}}$.
(ii) $\mathbf{db_t = (m – c*b_t)dt + h*b_t^{1/2}*dW_{b\_t}}$.
(iii) $\mathbf{d[\ln(s_t^2)] = u*(v – (1/2)\ln(s_t^2))*dt + w*dW_{s\_t}}$.

Formulas (i) and (ii) represent a double-mean-reverting process. (Note how r_t is defined in terms of b_t.) Formula (iii) represents stochastic volatility.

S7-AMFS-23. Give the formulas for the Balduzzi model. Briefly explain the model in conceptual terms. (Venter, pp. 10-11)

Solution S7-AMFS-23.
The Balduzzi model has four formulas:
(i) $\mathbf{dr_t = (b_t – a*r_t)dt + s_t^{1/2}*dW_{r\_t}}$.
(ii) $\mathbf{db_t = (m – c*b_t)dt + h*dW_{b\_t}}$.
(iii) $\mathbf{ds_t = u*(v – s_t)dt + w*s_t^{1/2}*dW_{s\_t}}$.
(iv) $\mathbf{d[W_{r\_t}W_{s\_t}] = \rho*dt}$.

Formulas (i) and (ii) represent a double-mean-reverting Vasicek process. Formula (iii) represents stochastic volatility, which follows a square-root process. Formula (iv) correlates stochastic volatility with the interest rate.

S7-AMFS-24. Fill in the blanks (Venter, p. 15): Any short-rate models can be made to fit exactly to the initial term structure by making the b parameters _______ chosen to ______.

Solution S7-AMFS-24. Any short-rate models can be made to fit exactly to the initial term structure by making the b parameters **deterministic time-dependent functions** chosen to **make the fit exact**.

S7-AMFS-25.
(a) Given that P(t, T) is the price at time t of a bond that matures at time T, give the formula for the forward rate f(t, T) in terms of P(t, T).
(b) Define P(t, T) in terms of f.
(Venter, p. 15)

Solution S7-AMFS-25.
(a) $\mathbf{f(t, T) = -\partial[\ln(P(t,T))]/\partial T}$.
(b) $\mathbf{P(t, T) = \exp(-\int_t^T f(t, s)*ds)}$.

S7-AMFS-26. Fill in the blank: The forward rate is a representation of all the ________. (Venter, p. 15)

Solution S7-AMFS-26. The forward rate is a representation of all the **possible risk-neutral interest rates being discounted over at time T.**

S7-AMFS-27. You are given the extended formula for the Vasicek model, adjusted for the market price of risk: $dr_t = (b(t) – k*r_t)*dt + s*dW_t$, give expressions for the following:
(a) b(t) in terms of f(0, t) and the variables in the formula above.
(b) B(t, T), if the bond price is $P(t, T) = A(t, T)/\exp[B(t, T)*r_t]$.
(c) A(t, T), if the bond price is $P(t, T) = A(t, T)/\exp[B(t, T)*r_t]$.
(Venter, pp. 15-16)

Solution S7-AMFS-27.
(a) $b(t) = \partial[f(0, t)]/\partial[dt] + k*f(0, t) + s^2*(1-e^{-2kt})/(2k)$.
(b) $B(t, T) = (1-e^{k(t-T)})/k$.
(c) $A(t, T) = [P(0, T)/P(0, t)]*\exp[B(t,T)*f(0,t) – s^2*(1-e^{-2kt})*B(t,T)^2/(4k)]$

S7-AMFS-28. Fill in the blank: If the yield curve can match the current curve exactly then the ________ and the _________ would be calibrated to ______. (Venter, p. 16)

Solution S7-AMFS-28. If the yield curve can match the current curve exactly then the **market prices of risk** and the **partial interest rates** would be calibrated to **option prices**.

S7-AMFS-29.
(a) What way of improving the fit of multifactor short-rate models have Dai and Singleton proposed?
(b) Fill in the blanks: Dai and Singleton found a general framework where the yield curve can be calculated in closed form once _________ (what is done?). This is called a(n) ________. (Venter, p. 16)
(c) What does the Dai and Singleton approach allow for with regard to factors in the CIR model?
(d) Which model have Dai and Singleton found to be the best-fitting to empirical data, and why?

Solution S7-AMFS-29.
(a) Dai and Singleton proposed to add terms for interaction and correlations among the factors.
(b) Dai and Singleton found a general framework where the yield curve can be calculated in closed form once **two ordinary differential equations are solved numerically**. This is called an **almost-closed-form solution**.
(c) The Dai and Singleton approach allows for **positive correlation among the CIR factors** to be included.
(d) Dai and Singleton found the **Balduzzi** model to be the best-fitting, because it allows for both positive and negative correlations among the factors, which is more realistic.

S7-AMFS-30.
(a) What tradeoff does Venter (p. 16) mention with regard to the basic interest-rate models with closed-form solutions?
(b) What issue does Venter (p. 16) state is often ignored with regard to the market price of risk?

Solution S7-AMFS-30.
(a) The tradeoff is between the model being analytically tractable and being able to handle stochastic volatility.
(b) The issue often ignored is that the market price of risk probably changes with time t. The models assume that the market price of risk is constant for every T at time t.

S7-AMFS-31.
(a) With regard to analyzing US inflation rates, the double-mean-reverting AR(1) process fits relatively well for the period from 1983 onward. However, Venter (p. 19) states that this is probably not sufficient for risk-management purposes. Explain why it is not.
(b) Which process has been shown to fit the inflation data for 1947-2009 relatively well? (Venter, p. 17)

Solution S7-AMFS-31.
(a) The 1983-onward model would not suffice for risk-management purposes, because inflation rates were lower during that period than they had been previously. A model that relies solely on information from this period might understate the magnitudes of future inflation rates.
(b) A model consisting of three independent AR(1) processes for partial inflation rates has been able to model the 1947-2009 data relatively well.

S7-AMFS-32. According to Venter (p. 20), where is the association between interest rates and inflation particularly clear, and how might it be quantified?

Solution S7-AMFS-32. The association between interest rates and inflation is particularly clear in the tails: both inflation and interest rates tend to be very high and very low at the same time. A way to quantify this is through the tail association functions for the left and right tails: $L(z) = \Pr(F_X(X) < z \mid F_Y(Y) < z)$ and $R(z) = \Pr(F_X(X) > z \mid F_Y(Y) > z)$.

S7-AMFS-33.
(a) According to Venter (p. 25), what is the "starting point" for modeling equity prices?
(b) What is the corresponding formula for the process for S, the price of a stock?
(c) What problem with this process has been empirically demonstrated? (Venter, p. 25)

Solution S7-AMFS-33.
(a) The "starting point" is **geometric Brownian motion.**
(b) The corresponding formula is $d[S(t)]/S(t) = \mu * dt + \sigma * d(Z(t))$.
(c) The problem is one of insufficiently heavy tails, particularly for downward price movements.

S7-AMFS-34.
(a) Give the formula for the double-exponential jump process (DEJD), the process for S which incorporates the modeling of large upward and downward jumps via Poisson processes, given that v and w are the Poisson parameters for the downward and upward jumps, respectively.
(b) Give the corresponding distributions for the downward jump size (F(x)) and the upward jump size (F(y)).
(c) List two reasons for the popularity of the DEJD model.
(d) What is a reasonable starting assumption regarding the association between equity prices and interest rates/inflation?
(Venter, pp. 25-27)

Solution S7-AMFS-34.
(a) DEJD formula: $\mathbf{d[S(t)]/S(t) = \mu * dt + \sigma * d[Z(t)] + d({}_{i=0}{}^{N(vt)}\sum(X_i - 1) + {}_{i=0}{}^{N(wt)}\sum(Y_i - 1))}$.
(b) $\mathbf{F(x) = x^{\alpha}}$ **for** $\mathbf{0 < x < 1}$**.** $\mathbf{F(t) = 1 - y^{-\beta}}$ **for** $\mathbf{y > 1}$**.**
(c) (i) The DEJD model has high goodness of fit, and (ii) many option prices have been worked out for it.
(d) A reasonable starting assumption is that equity prices are **independent** of interest rates and inflation. Historically, the correlation coefficients differ considerably (in both magnitude and direction) depending on what timeframe is examined.

S7-AMFS-35.
(a) Give four reasons why currency exchange rates are difficult to model and have little predictive power.
(b) Given these difficulties, which models are most often proposed for currency exchange rates? (Venter, pp. 27-28)

Solution S7-AMFS-35.
(a) 1. Currency exchange rates are highly volatile, and the random element is likely to predominate in any model.
2. Currency exchange rates have complex correlation patterns.
3. Currency exchange rates have persistent autocorrelations.
4. For various currencies, exchange rates correlate strongly in the tails.
(b) Random-walk models or continuous driftless Brownian motion are most often proposed for currency exchange rates.

S7-AMFS-36.
(a) What does the theory of *uncovered* interest parity hold? Have empirical studies supported or rejected the theory?
(b) How does covered interest parity differ from uncovered interest parity?

Solution S7-AMFS-36.
(a) The theory of *uncovered* interest parity holds that the difference in interest rates in two countries' economies should be the deterministic trend in the exchange-rate process. This implies that a deposit made in either currency would have the same expected return in each currency. Empirical studies have consistently rejected this theory.
(b) Covered interest parity incorporates holding a position in forward rates

S7-AMFS-37. What sorts of risks are present in the "risky bonds" that Venter discusses on pp. 29-30? What are said "risky bonds"?

Solution S7-AMFS-37. The "risky bonds" in question are corporate and municipal bonds. They involve risks pertaining to default, liquidity, and taxation.

S7-AMFS-38. Among the following, which are the most complex to model and why, according to Venter (p. 31)? (i) Equities, (ii) Bonds, (iii) Inflation, (iv) Foreign exchange rates.

Solution S7-AMFS-38. (ii) Bonds are the most complex to model, because of interactions among various bonds and the difficulty of keeping a closed-form yield curve while maintaining a reasonable model.