APPENDIX I

Basic Mathematics and Statistics

Notation for Present Value Calculation

At various points, the notation $PV(\tau)$ appears. This function is the conventional present value discounting operator for a cash flow of $1 to be received in $\tau = T - t$ days. In Sec. 8.1, this is defined as the present value at time t of $1 to be received at time T. $PV(\tau)$ means the present value evaluated at $\tau = T - t$. This can be evaluated using either a continuous or discrete formulation. In order to reduce the amount of notation, $PV(\tau)$ is used to represent both continuous and discrete compounding, even though there are slight difference in the two cases. However, with correct specification of the interest rates used in the discounting process, results of the continuous and discrete compounding operations are identical.

In discrete time, the discounting operator takes the linear form, $1/(1 + rt^*)$, when t^* is less than 1, and $1/(1 + r)^{t^*}$ where t^* is an integer greater than zero. The geometric form of the discounting operator, $1/(1 + r)^{t^*}$, is also encountered where t^* can take any value. For both these cases, r is expressed an annualized interest rate. In continuous time, the discounting operator takes the form, $exp\{-rt^*\}$, where r is the continuously compounded interest rate. While the same symbol r is used as the interest rate in each of the discounting operators, equality of the present value determined by the continuous and discrete discounting operations requires the discrete and continuous interest rates to be adjusted to account for the difference between stated and effective interest.

Taylor Series Expansion

To make a comparison between the various forms of discounting operators, observe that $ln[exp\{-rt^*\}] = -rt^*$. The corresponding $ln[(1 + rt^*)^{-1}]$ for the discrete compounding case does not provide a similar simple solution but can be evaluated by taking a ***Taylor series expansion***. This useful expansion technique from mathematical analysis takes the general form, e.g., Rudin (1964):

$$f[x] = f[a] + \frac{\partial f[a]}{\partial x}(x - a) + \frac{1}{2!}\frac{\partial^2 f[a]}{\partial x^2}(x - a)^2 + \frac{1}{3!}\frac{\partial^3 f[a]}{\partial x^3}(x - a)^3 + \ldots\ldots$$

In other words, any function $f[x]$ with which has continuous derivatives over some interval $[b,c]$ can be equivalently expressed as the converging infinite sum provided by the Taylor series. This expression expands the univariate function $f[x]$ about the point fixed point a, where $b \leq a \leq c$. Each of the derivatives in the expansion are evaluated by setting $x = a$. For this expansion to be valid, the function f must have derivatives of all orders over $[b,c]$ (some of which can be zero).

A Taylor series is a specific form of a more general mathematical concept known as a power series. Other types of power series can also be used to represent an admissible function as an expansion expressed as an infinite converging sum. Examples of such expansions occur frequently, particularly in mathematical statistics. Some of the more important are the Edgeworth expansion, Cornish-Fisher expansion and the Gram-Charlier expansion. The various types of power series are distinguished primarily in the method used for determining the coefficients for each of the terms in the expansion. The Taylor series uses derivatives of the function $f[x]$ in order to determine the coefficients. An important special case of the Taylor series occurs where zero is used as the fixed point in the expansion. This form of the Taylor series is sometimes referred to as a Maclaurin series.

Applied to the function $y[r] = ln[(1 + rt^*)^{-1}]$ evaluated around the origin $(a = 0)$ gives:

$$\ln[\frac{1}{1 + rt^*}] = -\ln[1 + rt^*] = -\{\ln[1 + (0)t^*] + \frac{t^*}{1 + (0)t^*}(r - 0)$$

$$- \frac{1}{2!}\frac{(t^*)}{(1 + ((0)t^*)^2}(r - 0)^2 + \frac{1}{3!}\frac{(t^*)}{2(1 + (0)t^*)^3} - \ldots\}$$

$$= -\{0 + rt^* - r^2 t^* + r^3 t^* - \ldots\}$$

For small r, say r is .1, which covers cases involving most interest rates, the terms involving squares, cubes and so on will disappear and the expansion reduces to *ln[1/(1 + rt*)]* being approximately equal to *-rt**, which is the same as the exact value given when the log of the continuous discounting operator was evaluated. Using this approach, the difference between the continuous and discrete interest rates can be represented by the higher order terms which are being ignored in the Taylor series expansion.

Application of the Taylor series to the geometric discounting operator produces: $\ln[(1 + r)^{-t^*}] = -t^* \ln[(1 + r)]$ which is also approximately $-rt^*$. However, unlike the linear form where the approximation to the log was taken for $(1 + rt^*)$, this approximation is $-t^*$ times $\ln[1 + r]$ which will, generally, produce a different error of approximation. The solutions will be the same when t^* is an integer greater than zero.

This discussion can also be used to clarify an approximation which is commonly used in financial calculation: $\ln[1 + x]$ is approximately equal to x when x is small. The question this operation raises is: how small can x be in order to ensure that substitution of x for $\ln[1 + x]$ does not produce sizable errors of approximation. Examining the Taylor series for $\ln[1 + x]$ reveals that the square of x must be small relative to x. For example, if x is .1 then $x^2 = .01$ which, for many applications, is sufficiently small to permit the approximation. Similarly, if $x = .05$ then $x^2 = .0025$, and the approximation is again admissible in many situations. However, for $x = .5$, $x^2 = .25$, a value which is quite large relative to x. In practice, the admissibility of the approximation will depend on the specific type of problem in which the approximation is used. It is always possible to improve the approximation by taking account of higher order terms in the expansion.

Another useful application of Taylor series is to expand the bond price function in terms of yield and time. The conventional textbook explanation for the relationship between duration and convexity, e.g., Fabozzi (1993, Chp. 4), is to treat the bond price as a univariate function of yield. Applying a Taylor series expansion to this function gives:

$$P_B(y) = P_B(y_0) + \frac{\partial P}{\partial y}(y - y_0) + \frac{1}{2}\frac{\partial^2 P}{\partial y^2}(y - y_0)^2 + \ldots\ H.O.T.$$

$$\frac{P_B(y) - P_B(y_0)}{P_B(y_0)} \cong -DUR\ (y - y_0) + \frac{1}{2}\ CON\ (y - y_0)^2$$

where *DUR* and *CON* are the duration and convexity of the bond. This result can be used to show that for two different fixed income portfolios, with equal initial yield and duration, that the higher convexity portfolio will have a more favourable percentage change in price, whether yields go up or down.

As demonstrated by Christensen and Sorensen (1994) and others, this analysis is faulty because the bond price function depends on two variables, yield and time. This expansion involves the application of a multivariate Taylor series:

$$P_B(y,t) = \sum_{t=1}^{T} \frac{C}{(1 + y)^t} + \frac{M}{(1 + y)^T}$$

$$= P_B(y_0, t_0) + \frac{\partial P}{\partial y} (y - y_0) + \frac{\partial P}{\partial t} (t - t_0)$$

$$+ \frac{1}{2} \{\frac{\partial^2 P}{\partial y^2} (y - y_0)^2 + \frac{\partial^2 P}{\partial t^2} (t - t_0)^2\} + \frac{\partial^2 P}{\partial y \, \partial t} (y - y_0)(t - t_0) + ...H.O.T.$$

Dividing through by P_B will produce the familiar expansion in terms of modified duration, convexity, theta and the cross product terms, e.g., Jordan and Chance (1995).

Some confusion about the mean-variance-skewness expected utility function originates from a misunderstanding about the approximation properties of a Taylor series. It is not possible to argue that the mean-variance-skewness function is superior to the mean-variance function because convergence of the Taylor series ensures that the mean-variance-skewness approach will provide an analytically more precise approximation to the general *EU* case. Monotonic convergence is too much to expect from Taylor's theorem, e.g., Rudin (1964), which is concerned with the limiting properties of a specific polynomial series. A Taylor series will only converge ***uniformly***. The Cauchy criterion for uniform convergence of a sequence of functions $\{f_n\}$ requires only that, for x within the interval of convergence of the sequence, there exists some integer N such that $m \geq N, n \geq N$ implies $|f_n(x) - f_m(x)| \leq \varepsilon$, for every $\varepsilon > 0$. For a uniformly convergent series, such as a Taylor series, it is possible that adding another term to the approximation may not improve the accuracy of the approximation. The essential requirement is that the series will eventually converge to the true function.

Restrictions on the Taylor Series Coefficients of the Expected Utility Function

From Sec. 2.2, once the sign restrictions have been assumed for the derivatives of the utility function, further restrictions on the coefficients can be derived by evaluating the derivatives of the Taylor series expansion which is truncated at the third derivative term. More precisely, for $U[W]' > 0$, $U[W]'' < 0$ and $U[W]''' > 0$, taking derivatives in $U[W]$ produces:

$$U[W]' = a - 2b\ (W - \Omega) + 3c\ (W - \Omega)^2 > 0$$

$$U[W]'' = -2b + 6c\ (W - \Omega) < 0$$

$$where: \quad a = U[\Omega]' \qquad b = -\frac{U[\Omega]''}{2} \qquad c = \frac{U[\Omega]'''}{3!}$$

Solving the second derivative condition produces the result: $(b/3c) = -U''/U''' > (W - \Omega)$. This condition provides a restriction on the maximum possible value for W compatible with specific values of b and c.

From this it is now possible to derive an essential feature of the third order Taylor series approximation. Taking expectations of the derivatives gives:

$$E[U'[W]] = a + 3c\ \sigma^2 > U'[\Omega] = a > 0$$

$$E[U''[W]] = U''[\Omega] = -2b \qquad E[U'''[W]] = U'''[\Omega] = 6c$$

The restriction of the first derivative, that the average slope of the $U[W]$ is greater than the slope at the average, is an implication of Jensen's inequality applied to the first derivative. This result can be contrasted with the restrictions provided by the Taylor series derivatives for a second order approximation which is relevant to the mean-variance case:

$$U'[W] = a - 2b\ (W - \Omega) > 0 \quad \rightarrow \quad E[U'[W]] = a = U'[\Omega]$$

$$U''[W] = -2b = U''[\Omega] < 0$$

For the mean-variance expected utility function, the average slope equals the slope at the average, a result which differs from the mean-variance-skewness case. In addition, the restriction on admissible values of W is different: $(a/2b) > (W - \Omega)$.

The Prakash et al. (1996) Claim

Taking $W_0 = \Omega$, the Prakash et al. claim is that a risk averse manager with sufficient preference for positive skewness will undertake an unfair gamble. For the weaker case of a fair gamble, this implies that the risky outcome $EU[W_1]$ can be preferred to the sure outcome of not gambling: $EU[W_1] > EU[W_0] = U[\Omega]$. Using the third order Taylor series expansion and taking expectations it follows that:

$$U[\Omega] < U[\Omega] + a\ (W - \Omega) - b\ (W - \Omega)^2 + c\ (W - \Omega)^3 \quad \rightarrow$$

$$U[\Omega] < U[\Omega] - b\ var[W] + c\ skew[W] \quad \rightarrow \quad 3(W - \Omega) < \frac{b}{c} = \frac{3\ U''}{U'''} < \frac{skew[W]}{var[W]}$$

The last inequality involving $(W - \Omega)$ is developed above.

The novelty of this claim can be demonstrated by considering the weaker case where the gamble is fair (expected return on the gamble is equal to zero). Taking initial wealth to be equal to expected wealth next period, $W_0 = \Omega = E[W_1]$, it follows that: $U[\Omega] = E[U[W_0]] < EU[W_1] \leq U[E[W_1] = U[\Omega]$, where the first inequality is the Prakash et al. proposition and the second (weak) inequality follows from Jensen's Inequality. Prakash et al. (1996) develop their seemingly impossible claim by manipulating the Taylor series expansion. Examining the coefficient values for the derivatives of $U[W]$ demonstrates that the Prakash et al. claim corresponds to the condition: $b/3c < skew[W_1]/var[W_1]$. When this result holds, the Prakash et al. claim is correct. One apparently obvious case where this condition could apply would be $b = 0$, $c > 0$ and $skew[W] > 0$. In general, if the distribution of W is negatively skewed or symmetric the Prakash et al. condition cannot apply.

To see this, observe that the Prakash et al. condition produces the third order differential equation:

$$U'''[\Omega] - \{\frac{3\ var[W]}{skew[W]}\}\ U''[\Omega] < 0$$

Taking the inequality to be an equality, this can be solved to produce an exponential utility function ($U = -exp\{-(3\ var[W]/skew[W])W\}$) which is consistent with the derivative restrictions. Horowitz then makes two observations: this solution is inconsistent with the assumption that U'''' and higher derivatives are zero; and, the parameter $3\sigma^2/\sigma^3$ is "incompatible" with "the existence of a risk preference function that validates the violations", e.g., assuming $b = 0$ is incompatible. The last observation is the same as saying that there is no utility function with a convergent Taylor series which is compatible with the type of gamble distribution implied by the Prakash et al. condition.[1] For present purposes, this debate is avoided by considering only the case of negatively skewed distributions.

Relationship of Continuous Time and Discrete Interest Rates

Use the notation that: $r(t,T) = r\ ((T\text{-}t)/365) = rt^*$, where r is the interest rate expressed as an annualized rate. In words, rt^* is the actual interest rate which is earned over the holding period. For example, if $T - t$ = 91 days then rt^* is the interest rate earned over a three month period, approximately r/4. When t=0, the formula simplifies to $r(0,T) = r\ (T/365)$. The future value of an investment of $V(t)$ at $r(t,T)$ can now be expressed as: $V(T) = V(t)(1 + r(t,T))$ and $\Delta V = V(T) - V(t) = r(t,T)\ V(t) = rt^*\ V(t)$. Recognizing that $V(t)$ and r are fixed and that $t^* = (T\text{-}t)/365$ is a time difference $\Delta t/365 = \Delta t'$ gives: $V(T) = V(t)\ r\ \Delta t'$. This can be compared with the same calculations done using continuous interest rates:

$$V(t) = V(0)\ e^{rt} \quad \Rightarrow \quad \frac{dV(t)}{dt} = r\{V(0)\ e^{rt}\} \quad \Rightarrow \quad dV(t) = V(t)\ r\ dt$$

This result is used in Chap. 9 to describe the return on the Black-Scholes riskless hedge portfolio, where $V(t) = S(t) - \beta\ C(t)$ and in Sec. 9.1 in the discussion of Fischer (1975).

Background on Statistical Concepts

The discussion in this book uses a number of statistical concepts which may be unfamiliar, such as the conditional variance and expected value of a profit function. The use of parameters from conditional distributions is consistent with the type of valuation problems encountered in financial economics, where optimal decisions involve evaluating random variables such as future prices. Conditional expectations are required because the information set upon which the expectation is based increases over time, until the decision horizon (T) is reached. Except in restrictive cases, the parameters of the conditional distribution will differ from the parameters of the unconditional distribution which is derived from the complete information set, available at T. As a consequence of being determined from a fixed information set, unconditional statistical parameters are constant. This will typically not be the case with the parameters of conditional distributions.

A conditional expectation depends on the information set available on the decision date, t. For $t < T$ the information set will increase over time, changing the information used in forming the expectation. When the conditional parameters are constant over time, then the conditional and unconditional parameters are identical. However, conditional statistical parameters will only be constant under restrictive conditions, such as identically, independently distributed normal random variables. Given these caveats, if the conditioning information set is ignored, the mathematical operations required to derive the mean and variance of the profit function are identical for both the unconditional and conditional cases. Bearing in mind that the parameters being determined are from conditional distributions, it is conventional to drop notation involving the conditioning information set.

Given this, from Sec. 2.1 the mathematical ***conditional expectation*** of the mean of the ***speculative profit*** function conditional on information available at the decision date $t = 0$ can be stated:

$$E[\pi(1)] = Q\ \{E[F(1,T)] - F(0,T)\}$$

This result uses two properties of conditional expectations: linearity of the expectation; and, $E[F(0,T) \mid I(0)] = F(0,T)$. This follows because the information available on the decision date contains F(0,T). The associated ***conditional variance*** for the ***speculative profit*** function follows appropriately:

$$var[\pi(1)] = E[(\pi(1) - E[\pi(1)])^2] = E[(Q\{F(1,T) - F(0,T)\} - Q\{E[F(1,T)] - F(0,T)\})^2]$$

$$= E[(Q\{F(1,T) - E[F(1,T)]\})^2] = Q^2\ E[(F(1,T) - E[F(1,T)])^2] = Q^2\ var[F(1,T)] = Q^2\ \sigma_f^2$$

The ***conditional expectation*** of the stylized ***hedger profit*** function from Sec.2.3 follows appropriately:

$$E[\pi(1)] = E[Q_s\ \{S(1) - S(0)\} + Q_H\ \{F(0,T) - F(1,T)\}]$$

$$= Q_s\{E[S(1)] - S(0)]\} + Q_H\{F(0,T) - E[F(1,T)]\}$$

The associated ***conditional variance*** for the ***hedger profit*** function follows:

$$var[\pi(1)] = E[(\pi(1) - E[\pi(1)])^2]$$

$$= E[(Q_s\{S(1) - S(0)\} + Q_H\{F(0,T) - F(1,T)\} - Q_s\{E[S(1)] - S(0)\} - Q_H\{F(0,T) - E[F(1,T)]\})^2]$$

$$= E[(Q_s\{S(1) - E[S(1)]\} - Q_H\{F(1,T) - E[F(1,T)]\})^2]$$

$$= Q_s^2 E[(S(1) - E[S(1)])^2] - 2Q_sQ_H E[(S(1) - E[S(1)])(F(1,T) - E[F(1,T)])]$$

$$+ Q_H^2 E[(F(1,T) - E[F(1,T)])^2]$$

$$= Q_s^2 \sigma_s^2 - 2Q_sQ_H \sigma_{sf} + Q_H^2 \sigma_f^2$$

The mean and variance for other profit functions used in various parts of the book can also be derived using this approach.

1. Horowitz first observation would be applicable to almost any type of Taylor series expansion of a more general utility function. The second observation makes the important point that var[W] and skew[W] are functions of the optimal values of the risky investment, and that var and skew are functionally related. As a consequence, it is not possible to determine optimal values which satisfy the needed restrictions on b and c for the Prakash et al. condition to apply.