Support the author!
If you like this book, buy a copy!

Appendix C

Other Financial Economics Answers

This appendix contains answers to the questions posed in chapter 3.

Answer 3.1: This is a very old problem, and a common interview question. The probability that the first head occurs on toss k is $\left(\frac{1}{2}\right)^k$; this event carries with it a payoff of $\$2^k$. The contribution of toss k to the expected payoff is thus $\left(\frac{1}{2}\right)^k \times \$2^k = \$1$. This is the same for each k. The expected payoff to the game as a whole is the summation over all k of these payoffs. This is $\$1 + \$1 + \$1 + \cdots = \$\infty$. The expected payoff to the game is infinite!

This is called the "St. Petersburg Game." The fact that the expected payoff to the game is infinite, and that no one in his or her right mind would pay more than a few hundred dollars to play, is why it is sometimes called the "St. Petersburg Paradox." There are several ways that you can think about this sensibly.

One way is to note that "value" is not the same thing as "expected payoff";[1] value equals *utility* of expected payoff. Most people cannot distinguish between very large amounts of money.[2] This means that $\$2^{50}$ is not worth twice $\$2^{49}$. However, these very large amounts are counted in exactly this way when calculating expected payoff to the game as a whole. If you think that $\$2^k = \$2^{50}$ (essentially) for all $k \geq 50$, then the expected payoff to the game is finite:

$$\$50 + \$2^{50} \times \left(\frac{1}{2^{51}} + \frac{1}{2^{52}} + \frac{1}{2^{53}} + \ldots\right) = \$51$$

A spread could be quoted around this, maybe (\$10, \$200). How much would you pay your customer to play? How much would you charge your customer?

[1]It is important to note that the Weak Law of Large Numbers fails if the expectation is not finite (Feller [1968, pp251]).

[2]Bernoulli ([1738]; [1954]) suggests that utility of payoffs should depend upon how wealthy you are. For a practitioner's view of utility, see Kritzman (1992a).

A second way to think about this is in terms of default risk.[3] We need to quote the bid (what we pay) and the ask (what the customer pays). For the bid, it is the customer's default risk we need to worry about. Let us assume a wealthy customer who defaults above one million dollars. In this case, the customer defaults after (about) 20 tosses. Assuming the investment bank is of large scale, a payoff from the customer between two dollars and one million dollars is of relatively small size. The investment bank takes such bets every day, and this one is uncorrelated with all the others. At this level, we could argue that the investment bank is risk-neutral and so the bid is exactly \$20 with no risk premium.

For the ask, it is the company that risks bankruptcy and default. Let us assume that the company files for bankruptcy after losing one billion dollars (on the order of magnitude of Barings, and Metallgesellschaft)—approximately $\$2^{30}$. The expected value of the game to the customer is thus about \$30—the bank defaults after 30 tosses. However, your career and the holdings of all the shareholders can be destroyed by this bet, so you had better add a considerable risk premium. You might want to go all the way up to \$200 and quote a bid-ask of (\$20, \$200)—it depends upon your degree of risk aversion.

Each of these two solutions uses a truncation method. Another related way to think about this is in terms of feasibility. If it does take more than 50 tosses to get a head, then the payoff is not feasible because $\$2^{50}$ is more dollars than there are atoms in the universe, and whoever sold the ticket to the game is—by the laws of physics—unable to pay. See also Feller (1968, pp251–253).

Answer 3.2: This is a frequent question. Assuming continuously compounded returns follow an arithmetic Brownian motion (see Crack [2009]), variance of returns grows linearly with the compounding period. This is because consecutive returns in a random walk are independent, and the variance of a sum of independent random variables is just the sum of the variances. This means that the four-year σ^2 equals four times the one-year σ^2. It follows that the four-year σ is two times the one-year σ. The answer is, therefore, 20%. See also Question 2.18.

Answer 3.3: This is a very common term-structure question. You should be able to do this in your head almost instantly. Think of it this way: the rate over the first five years and the rate over the second five years must average out to give the rate over the full 10 years. That is, the average of 10% and the unknown forward rate must give 15%. The unknown must be 20%. To work it out quickly, note that the unknown (20%) is as far above the average (15%) as the known (10%) is below it.

In fact, if you work it out exactly, the forward rate is

$$\left[\frac{(1.15)^{10}}{(1.10)^5}\right]^{\frac{1}{5}} - 1 = 20.227\%.$$

[3] I thank Olivier Ledoit for suggesting this solution technique. I am responsible for any errors.

You are making a "first-order" approximation when you do the simple averaging, but you end up quite close. For a practitioner's viewpoint on the term-structure of interest rates, take a look at Kritzman (1993b).

Answer 3.4: There are many different types of bond yield. The "yield" on a bond is usually the "internal rate of return" or "yield-to-maturity" or "promised yield"; it is what you earn if you hold the bond to maturity assuming a constant reinvestment rate. In practice, given that reinvestment rates can vary significantly from your initial promised yield, your actual return can be higher or lower.

The "rate of return" on a bond is the internal rate of return of the realized cash flows to the bond-holder including reinvestment. If the bond is sold before maturity, the (realized) rate of return can be positive or negative.

Suppose you buy a bond promising 5%, but interest rates rise dramatically soon after your purchase. If you then sell the bond, you record a capital loss and a negative rate of return. However, if you hold the bond to maturity, you get your promised 5% plus a bit more because of higher reinvestment rates on the coupons.

Answer 3.5: Chaos theory came out of MIT in the early 1960's. Professor Edward N. Lorenz (Professor Emeritus in the department of Earth, Atmospheric and Planetary Science from 1987 until his death in early 2008) discovered that computer-simulated nonlinear mathematical equations describing the evolution of weather patterns are very sensitive to the starting values of the variables (Lorenz [1963]).[4]

This "sensitive dependence on initial conditions" is the first of three characteristics most often associated with chaos theory. The second characteristic is that the nonlinear systems describing chaotic systems are non-random. That is, they are "deterministic," not "stochastic." However, the output of the (often very simple nonlinear) systems can appear quite random. The third characteristic is "self-similarity": the physical system looks similar at different levels of magnification. It is self-similarity that gives rise to the "fractals" that you may have seen elsewhere. Fractals are often associated with the mathematician Benoit Mandelbrot (still a Professor Emeritus of Mathematics at Yale in 2009).

There are several different definitions of "chaos" in the literature. These definitions are beyond the scope of this book. See Brock et al. (1991, pp8–17) for further details. For a low-level broad introduction to chaos, see Gleick (1987).

Can you use chaos theory in finance? This was a hot topic in the late 1980's and early 1990's. Many academic economics and finance papers were written on the subject. The few that made any sense found nothing reliable. The

[4] I had the pleasure of attending some Independent Activities Period (IAP) classes taught by Prof. Lorenz at MIT in 1994/1995. He reminded me a little of a slim Dave Thomas (you know, the Wendy's guy). Although in his late seventies he seemed younger and very good-natured.

others were written by ignorant people who jumped on the bandwagon; they should never have published their empty papers.

After reading more than 150 journal articles and half a dozen books on chaos theory and writing a 100-page Masters-level thesis on chaos theory applied to financial economics and publishing one paper in the *Journal of Finance*, I am quite pessimistic. My co-author and I hypothesized considerable discreteness-induced bias in the popular "BDS test" for chaos in equity data (Crack and Ledoit [1996]). Our hypothesis has now been confirmed (Krämer and Runde [1997]).

If you want to predict stock returns, I recommend that you use neural nets or some other nonlinear technique. In my opinion, any predictability that you can discern with chaos theory (e.g., "nearest neighbor" prediction techniques) is better investigated using the other nonlinear techniques available to you. Give it up—chaos theory is great in the physical sciences, but it is a lost cause in finance.

Answer 3.6: Look at table C.1 on page 180. The slope of the price-yield curve is $-\frac{D}{(1+r)}P$, where D is Macaulay duration, P is bond price, and r is yield. Changing slope (i.e., curvature) is driven almost entirely by changing P, because Macaulay duration, D, changes very little with changing yield, r (Crack and Nawalkha [2001]). D does, however, fall slowly with rising yield for a standard bond with coupons, and this does contribute marginally to curvature.

Note that curvature (i.e., changing slope) of the plot does *not* always imply that the Macaulay duration of a bond is changing (Crack and Nawalkha [2001])! This is a common misconception (it is easy to misconstrue this in Fabozzi and Fabozzi [1995, pp97–98]). For example, consider a zero-coupon bond with ten years to maturity. The plot of bond value versus bond yield is downward sloping with curvature. Whatever the yield, however, the bond's duration is ten years because it is a ten-year zero

A mathematical explanation of the convexity: you know that the curve slopes downward, it goes to a vertical asymptote at yield -1 and a horizontal asymptote at yield infinity. You know that the curve must be smooth because the pricing relationship is simple and well-behaved. The only way to get a well-behaved smooth curve in this situation is to have it be convex.

Answer 3.7: This question is an interesting intersection of theory and empirical reality. The question is not necessarily well-posed, but you should do your best to answer it. I give what I think is the best answer possible.

If the empirical security market line (SML) is wholly above the theoretical one, this means that stocks are under-priced relative to the CAPM. I propose two possible causes: First, maybe there is only one risk factor (the Market), but market participants require higher compensation per unit of beta-risk than suggested by the CAPM; second, maybe there is more than one risk factor, and market participants require compensation for factors not mentioned by

the CAPM. Conversely, if the empirical SML is wholly below the theoretical one, then stocks are overpriced relative to the CAPM. In this case, market participants do not require as much compensation per unit of beta-risk as theory suggests.

I think the best answer is to say that the CAPM does not account for all priced risk factors. It is likely, however, that beta is priced. It follows that stocks require a premium over and above that suggested by CAPM, and you could think of this as an empirical SML plotting above the theoretical one. For more on factor models and estimation, see Kritzman (1993a).

There have been several papers pronouncing the CAPM either dead or alive (Wallace [1980]; Fama and French [1992]; Black [1993]; Fama and French [1996]). For a friendly introduction to the CAPM, see Mullens (1992). Ferguson and Shockley (2003) show that even if the traditional single-factor CAPM holds, use of an equity-only proxy for the World Market Portfolio of Risky Assets leads to an errors-in-variables mis-estimation of CAPM betas. This error in turn means that variables related to leverage will help to explain returns because they serve as instruments for the missing beta risk. The upshot of all this is that many empirical anomalies we see are in fact consistent with the single-beta CAPM.

You should note that there are some theoretical problems with both the question and my answer. It is quite difficult (if not impossible) to get either of the empirical SML's mentioned. This is not because the CAPM is "correct," or because there is only one risk factor. Rather, it is because there is a *very* tight mathematical relationship between betas and returns (Sharpe [1964]; Roll [1977a]). You would certainly need that the market proxy is not mean-variance efficient to get the plots suggested. It is probably not sufficient to simply assume that there are risk factors not accounted for by the CAPM. Go with the answer above, but realize that there is more here than meets the eye.

Answer 3.8: This question is very similar to Question 3.3 (and is just as common). You should be able to do it in your head almost instantly. If you cannot, then go back and try Question 3.3 again before reading on.

The rate over the first year and the rate over the second year must average out to give the rate over the full two years. That is, the average of 7.15% and the unknown forward rate must give 7.60%. The unknown rate must be around 8.05% (remember, it is as far above 7.60% as 7.15% is below 7.60%).

In fact, if you work it out exactly, the forward rate is

$$\left[\frac{(1.0760)^2}{(1.0715)^1}\right] - 1 = 8.052\%.$$

You are making a "first-order" approximation when doing the simple averaging, and the answer is quite accurate.

Answer 3.9: This is introductory finance theory; it uses no-arbitrage and not much more. Assume for the sake of simplicity that interest rates are constant at r per unit time, today is time t, and the forward contract matures at time T. The forward price, $F(t,T)$, is related to the spot price, $S(t)$, as follows:

$$F(t,T) = S(t)e^{r(T-t)} \geq S(t)$$

The discount bond sells at a forward *premium* because of no-arbitrage.

The coupon bond is a different story. If you assume a continuous coupon of ρ per unit time, then the forward price, $F(t,T)$, is related to the spot price, $S(t)$, as follows:

$$F(t,T) = S(t)e^{(r-\rho)(T-t)} \leq S(t),$$

where the inequality follows because we were told that $r \leq \rho$. The coupon bond sells at a forward *discount* because of no-arbitrage.

For a practitioner's view on futures, forwards, and hedging, see Kritzman (1993c).

Answer 3.10: This is a classic question, and a very good test of your dexterity with elementary finance theory. If you have not yet figured it out, and you are peeking at the answers for a hint, I strongly recommend that you go back to the question and try again; read no further. If you are still reading, here is a hint: think about your investment horizon, and an immunization strategy. Now go back and try again.

Your investment horizon is very short. You want to profit from the change in the relationship between short- and long-term rates. However, you want to protect yourself from the level of the yield curve. That is, you want your position to be insensitive to parallel shifts in the yield curve, but positively sensitive to a steepening. This suggests that you should go short long-term debt, go long short-term debt, and match both the duration and price of the positions (i.e., use very low coupon short-term debt and very high coupon long-term debt).[5]

You may think of this as a "zero-duration" portfolio (to match your horizon). However, in just the same way that a zero net investment stock portfolio has no well-defined beta but can still be market-neutral, a zero net investment bond portfolio has no well-defined duration but can still be insensitive to parallel shifts in the yield curve.

Traders tell me that this strategy originated with the Salomon Bond arbitrage ("bond-arb") group. However, it is now so well known that profits may be slim.

For more on "Yield Curve Strategies," see the excellent papers by Jones (1991) and Litterman and Scheinkman (1991). Jones describes the statistical relationship between changes in level, slope, and curvature of the yield curve.

[5]If you cannot match durations of the positions, you can match on the product of duration × price. However, this will no longer be a zero net investment strategy.

Answer 3.11: For a standard bond, the Macaulay duration (Macaulay [1938]) is just the weighted-average term-to-maturity of the bond:

$$D \equiv \frac{\sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}}{\sum_{s=1}^{T} \frac{C_s}{(1+r)^s}} = \sum_{t=1}^{T} \omega_t \times t, \quad \text{where } \omega_t \equiv \frac{\frac{C_t}{(1+r)^t}}{\sum_{s=1}^{T} \frac{C_s}{(1+r)^s}},$$

and C_t are the cash flows (both coupon and principal). The weights ω_t are applied to the timing of the bond's cash flows. Each weight is equal to the present value of the particular cash flow as a proportion of the total value of the bond. It follows that the duration of a zero-coupon bond equals its term-to-maturity—because the weight of the final cash flow is unity (i.e., +1). Duration is measured in units of time, as is the term-to-maturity.[6]

Duration is a measure of how sensitive a bond's price is to changes in interest rates. Duration is related to, but differs from, the slope of the plot of bond price versus yield-to-maturity.

I find the following construction to be an instructive way of understanding how duration works.[7] Suppose that you have a liability due in the future and that you buy a bond now with the intention of using the bond (and its accumulated coupons) to meet the liability (the maturity of the bond is assumed to be greater than or equal to the maturity of the liability). Suppose that the present value of the bond is identical to the present value of the liability. Suppose that you open a bank account that earns the market interest rate (the yield-to-maturity of the bond). You deposit all cash in-flows from the bond in the bank account and let them compound through time (with no taxes or transaction costs). When your liability falls due, you sell your bond and close your bank account. Call the proceeds of the bond sale together with your final bank balance the "Terminal Value."

Can you meet your liability with the Terminal Value? Well, there are two risks involved. A fall in interest rates immediately after you purchase the bond pushes up the price at which you are able to sell your bond. However, a fall in interest rates also decreases your final bank balance because you earn less interest on the coupons. The opposite obtains with a rise in interest rates. That is, higher interest rates decrease the price at which you can sell the bond, but your closing bank balance is higher because you earn more interest on the coupons. These two risks are known as *price risk* and *coupon reinvestment rate risk*, respectively.

Price risk and coupon reinvestment rate risk have opposite influences on the Terminal Value. The Terminal Value differs depending upon which influence is

[6]Duration is usually measured in years, but this is not essential. If the dummy variable t in the formula counts half-years (so $T = 20$ for a 10-year bond), then a 10-year zero will have duration 20 (half-years). The only reason I mention this is that when valuing semi-annual bonds, you do sometimes count in half-years, and this can lead to confusion in the duration calculation. Obviously 20 half-years is the same as 10 years.

[7]I have never seen this construction using an artificial bank account in any literature.

strongest. It can be proved that if your liability falls due before the weighted-average term-to-maturity of your bond, the price risk has the stronger influence on Terminal Value. If your liability falls due after the weighted-average term-to-maturity of your bond, the coupon reinvestment rate risk has the stronger influence on Terminal Value. If your liability falls due precisely at the weighted-average term-to-maturity of the bond, the Terminal Value is relatively insensitive to an immediate change in interest rates.

By definition, the weighted-average term-to-maturity of the bond is just its Macaulay duration. This means that if you know when your liability falls due, you should provide for it by purchasing bonds with a duration equal to your investment horizon. You are "immunized" against a change in interest rates if the duration of your bond portfolio equals your investment horizon.[8]

I must emphasize that the bank account/Terminal Value construction is an artificial one. The fact that you must rebalance your position as time passes (in order to remain immunized) means that you cannot stick with the same bond until the liability is due. Indeed, the only bond that you can hold until the liability falls due (while remaining immunized) is a zero-coupon bond with maturity equal to the maturity of the liability; and in this case, the absence of coupons removes the need for the bank account in the construction.

Thus, the bank account/Terminal Value construction tells you how sensitive the Terminal Value is to an *immediate* parallel shift in the term-structure; for it is only in the *immediate* future that your chosen bond portfolio is immunized. This means that: if you can, open a bank account that pays the yield-to-maturity on your bond, purchase a coupon bond with duration and present value the same as those of the liability, and deposit all coupons in the bank account until the liability falls due—then, if there is one and only one parallel shock to the flat term-structure of interest rates between now and your liability falling due and if that single shift in interest rates occurs immediately, the Terminal Value will meet your liability. If anything else happens, you may be in trouble.

Other things being equal, duration increases with increasing term-to-maturity.[9] Other things being equal, duration decreases with increasing coupons (larger cash flows early on decrease the proportional importance of the repayment of principal at maturity).

Compared to duration, convexity is a higher-order measure of sensitivity of bond price to interest rates. Convexity measures the rate at which the sensitivity of bond price to interest rates changes with changing interest rates.[10] Convexity is related to, but differs from, the rate of change of slope of the plot of bond price versus yield-to-maturity. See the summary in table C.1. For a practitioner's view of Macaulay duration and convexity, see Kritzman (1992b).[11]

How do the definitions of duration and convexity arise? Suppose the price of a bond, P, is expanded in terms of yield-to-maturity, r, using a second-order Taylor series (that is, one that stops at the quadratic term):[12]

$$P(r+\Delta r) - P(r) \approx \frac{\partial P(r)}{\partial r} \times \Delta r + \frac{\frac{\partial^2 P(r)}{\partial r^2}}{2!} \times (\Delta r)^2$$

Letting $\Delta P \equiv P(r+\Delta r) - P(r)$, use $P(r) = \sum_{t=1}^{T} \frac{C_t}{(1+r)^t}$ to find that[13]

$$\Delta P \approx \frac{-\Delta r}{1+r} \sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t} + \frac{(\Delta r)^2}{2!(1+r)^2} \sum_{t=1}^{T} \frac{t \times (t+1) \times C_t}{(1+r)^t}.$$

Now divide both sides by P to get

$$\frac{\Delta P}{P} \approx \frac{-\Delta r}{1+r} D + \frac{(\Delta r)^2}{2!} \mathcal{C},$$

where

$$D \equiv \frac{1}{P} \sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}$$

is the standard Macaulay duration, and

$$\mathcal{C} \equiv \frac{1}{(1+r)^2 P} \sum_{t=1}^{T} \frac{t \times (t+1) \times C_t}{(1+r)^t}$$

is a measure of curvature, or "convexity," in the plot of bond price versus yield-to-maturity. Other things (i.e., duration and price) being equal, $\mathcal{C}$ increases

[8] Please note that an initial immunized position protects you from exactly one parallel shock to a *flat* term-structure. You are no longer immunized after a shock has hit. You get your planned future value only if no more shocks hit. You must re-balance after each shock to stay immunized. In fact, to stay immunized, you must rebalance even if no shocks hit. This is because changes in bond duration are not generally in lock-step with the passage of time. Thus, your horizon and your bond's duration decrease at different speeds, and you become non-immunized.

[9] Deeply discounted coupon bonds (bonds paying coupons far below current market rates) can be an exception (Fisher and Weil [1971, table 4, p418]).

[10] Strictly speaking, this is not true. See Crack and Nawalkha (2001) for details.

[11] The standard Macaulay duration is a relatively simple concept. People on The Street expect you to know that they use more complex tools. For example, the standard Macaulay duration can be generalized to allow for immunization against parallel shifts in yield curves that are *not* flat. This generalization was originally proposed by Macaulay (1938), but was made popular by Fisher and Weil (1971). An even more sophisticated measure of duration is presented by Cox, Ingersoll, and Ross (1979). Duration measures for bonds with embedded options are also important (Mehran and Homaifar [1993]).

[12] Note that this is similar to expressing the change in the price of a call option (given a change in the level of the underlying) in terms of the "delta" and the "gamma." The delta is the rate of change of call price with respect to underlying, and the gamma measures the "convexity" of call price with respect to underlying.

[13] I used the result $\frac{\partial P(r)}{\partial r} = \sum_{t=1}^{T} \frac{\partial}{\partial r} \frac{C_t}{(1+r)^t} = \frac{-1}{1+r} \sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}$ and an analogous result for $\frac{\partial^2 P(r)}{\partial r^2}$.

with increasing coupons. Even a zero-coupon bond has positive convexity (because $C_1 = C_2 = \cdots = C_{T-1} = 0$, but $C_T = FACE > 0$).

In addition to immunization, duration and convexity enable you to estimate the impact on bond price of a change in interest rates. A "first-order" estimate uses duration; a "second-order" estimate uses duration and convexity. Higher-order approximations are more accurate.

Take a 20-year bond paying an annual coupon of 7%. Assume a face value of \$1,000. Assume that the term-structure is flat at 10%. The price of the bond is \$744.59 under these assumptions.

If the entire term-structure rises by one percentage point (i.e., 0.01), what is the new price of the bond? This can be estimated using the equation we derived previously:[14]

$$\frac{\Delta P}{P} \approx \frac{-\Delta r}{1+r} D + \frac{(\Delta r)^2}{2!} \mathcal{C}$$

The Macaulay duration of this bond is calculated as 10.0018 years, the convexity $\mathcal{C}$ can be calculated as 130.04676, $\Delta r = +0.01$, $r = 0.10$, and $P = \$744.59$, thus:

$$\begin{aligned}
\Delta P &\approx \frac{-\Delta r}{1+r} D \times P + \frac{(\Delta r)^2}{2!} \mathcal{C} \times P \\
&= \frac{-0.01}{1.10} \times 10 \times \$744.59 + \frac{(0.01)^2}{2} \times 130.04676 \times \$744.59 \\
&= -\$67.69 + \$4.84 \\
&= -\$62.85
\end{aligned}$$

Thus, $P(r + \Delta r) \approx P + \Delta P = \$744.59 - \$62.85 = \$681.74$. *Direct evaluation* gives the answer as \$681.47 (the estimate is 27 cents too high and would have been out by roughly \$5 if not for the convexity term).

If the entire term-structure falls by one percentage point (i.e., 0.01), the change in bond price is estimated as follows:

$$\begin{aligned}
\Delta P &\approx \frac{-\Delta r}{1+r} D \times P + \frac{(\Delta r)^2}{2!} \mathcal{C} \times P \\
&= \frac{+0.01}{1.10} \times 10 \times \$744.59 + \frac{(0.01)^2}{2} \times 130.04676 \times \$744.59 \\
&= +\$67.69 + \$4.84 \\
&= +\$72.53
\end{aligned}$$

Thus, $P(r + \Delta r) \approx P + \Delta P = \$744.59 + \$72.53 = \$817.12$. *Direct evaluation* gives the answer as \$817.43 (the estimate is 31 cents too low and would have been out by roughly \$5 if not for the convexity term).[15]

Note that the "27 cents too high" and the "31 cents too low" in the above examples can be reduced to pennies (at least) by using a third term in the expansion—a measure of rate of change of convexity with respect to yield. Mehran and Homaifar (1993) refer to this third term as "velocity." Thus, they represent change in bond price as a function of duration, convexity, and velocity—see Mehran and Homaifar (1993) for more details.[16]

For a standard bond, Macaulay duration $D = \frac{1}{P}\sum_{t=1}^{T} \frac{t \times C_t}{(1+r)^t}$ may be written in a closed-form formula (i.e., no summation term). With coupons $C_t = C$, a constant, for $t = 1, 2, \ldots, T-1$, and $C_T = C + F$, where F is face value, the standard Macaulay duration of the bond may be written as follows:

$$D = \frac{1+r}{r} - \frac{\left\{(1+r) + T\left[\frac{C}{F} - r\right]\right\}}{\frac{C}{F}\left[(1+r)^T - 1\right] + r}, \quad \text{where } r \neq 0$$

The proof of this result uses the standard closed-form formula for an annuity and, although not difficult, may be a little tedious—a similar type of expression exists for convexity.

Finally, let me exorcise a myth. Most of the foregoing is predicated on parallel shifts in yield curves. Other things (i.e., price and duration) being equal, the higher the convexity of a bond, the better off you are if there is a parallel shift (up or down) in a yield curve: hence the myth that you should pay for convexity.[17] In reality, these shifts are anything but parallel (Jones [1991]; Litterman and Scheinkman [1991]). Other things (i.e., price and duration) being equal, if the yield curve steepens, additional convexity will probably hurt you. Whether additional convexity helps or hurts depends upon the bonds you consider, and the "twist" in the yield curve that occurs. Crack and Nawalkha (2000) derive simple expressions that allow bond portfolio managers to capture the combined effects of term-structure height, slope and curvature shifts on duration, convexity, and higher-order bond risk measures. See Kahn and Lochoff (1990) and Lacey and Nawalkha (1993) for empirical evidence.

[14] Note that the term $\frac{D}{1+r}$ that multiplies $-\Delta r$ is often called the "modified duration," frequently denoted D^*. It follows that the first-order approximation using modified duration is $\Delta P \approx -\Delta r D^* P$.

[15] Why am I *estimating* the change in bond price when direct evaluation gives the exact answer? For purposes of demonstration, it is convenient to be able to show you exactly how the duration and convexity measures work and where the approximations break down. This simple example is a good way to do that. In a real world situation, you might know the current value of your bond portfolio and its duration and convexity. It may be easier (and much faster) to *estimate* how your portfolio changes in value with changes in interest rates—using current value, duration, and convexity—than it is to *directly evaluate* each bond individually.

[16] People on The Street tell me that duration measures accounting for embedded options are important. Mehran and Homaifar (1993) discuss duration and convexity for bonds with embedded options. Before looking at Mehran and Homaifar (1993), be sure that both your mathematics and finance are up to scratch. They have the ideas correct, but their notation is contrary to conventional symbolic mathematics.

[17] This win-win situation is not kosher. A model that allows only parallel shifts in the yield curve freely admits arbitrage opportunities: match on price and duration and go long high convexity and short low convexity (Lacey and Nawalkha [1993]).

Table C.1: Duration/Convexity Summary

The bond pays C_t for $t = 1, \ldots, T$, and has discretely compounded annual yield r.	
Bond Price	$P = \sum_{t=1}^{T} C_t(1+r)^{-t}$
Modified Duration	$D^* \equiv \frac{-\frac{\partial P}{\partial r}}{P} = \frac{\sum_{t=1}^{T} t\,C_t(1+r)^{-(t+1)}}{P} = \frac{1}{(1+r)} \sum_{t=1}^{T} t\,\omega_t$, where $\omega_t \equiv \frac{C_t(1+r)^{-t}}{P}$ & $\sum_{t=1}^{T} \omega_t = 1$.
Macaulay Duration	$D = D^*(1+r) = \sum_{t=1}^{T} t\,\omega_t$ where $\omega_t \equiv \frac{C_t(1+r)^{-t}}{P}$ & $\sum_{t=1}^{T} \omega_t = 1$.
Bond Convexity	$\mathcal{C} \equiv \frac{\frac{\partial^2 P}{\partial r^2}}{P} = \frac{\sum_{t=1}^{T} t\,(t+1)\,C_t(1+r)^{-(t+2)}}{P} = \frac{1}{(1+r)^2} \sum_{t=1}^{T} t\,(t+1)\,\omega_t$, where $\omega_t \equiv \frac{C_t(1+r)^{-t}}{P}$ & $\sum_{t=1}^{T} \omega_t = 1$.
Slope of Price-Yield Curve	$\text{SLOPE} = \frac{\partial P}{\partial r} = -D^*P = -\frac{D}{(1+r)}P$
Curvature of Price-Yield Curve	$\text{CURVATURE} = \frac{\partial^2 P}{\partial r^2} = \mathcal{C}P$
Taylor Series	$\Delta P \approx \frac{\partial P}{\partial r}(\Delta r) + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}(\Delta r)^2 = -D^*P(\Delta r) + \frac{1}{2}\mathcal{C}P(\Delta r)^2$

Note: In the table, D is Macaulay duration, D^* is modified duration, P is bond price, and $\mathcal{C}$ is convexity. Try proving that $\frac{\partial D^*}{\partial r} = (D^*)^2 - \mathcal{C}$. Many of these relationships simplify substantially when we use continuously compounded yields (e.g., D and D^* are identical using continuously compounded yields y, so $\frac{\partial D}{\partial y} = D^2 - \mathcal{C}$, which in turn equals zero if the bond is a pure discount bond (Crack and Nawalkha [2001])).

Answer 3.12: From empirical investigations, it is known that stock returns do not have constant variance through time and that periods of high (low) volatility tend to follow periods of high (low) volatility (Fama [1965]; Akgiray [1989]). The GARCH model attempts to capture this empirical fact.

Suppose you estimate a simple linear model like $r_{it} = \alpha_i + \beta_i r_{mt} + u_{it}$ (return on stock i at time t is a constant plus a constant times return on the market plus a residual). If you do not take account of changes in the variance of u_{it} through time, you can draw faulty statistical inferences about α_i and β_i. Note that the standard ordinary least squares (OLS) regression does not account

for changing variance. In remedying this problem, the GARCH estimation captures a portion of stock price behaviour that might otherwise be interpreted as non-Normality and might lead to faulty inferences.

The GARCH model is a generalization of the ARCH model first presented in Engle (1982).[18] The formal GARCH(1,1) model for the residuals of a market model of stock returns is[19]

$$
\begin{aligned}
r_{it} &= \alpha_i + \beta_i r_{mt} + u_{it} \\
u_{it}|\mathcal{F}_{i\,t-1} &\sim \mathcal{N}(0, h_{it}) \\
h_{it} &= \gamma_{0i} + \gamma_{1i} u_{i\,t-1}^2 + \gamma_{2i} h_{i\,t-1}.
\end{aligned}
$$

The residuals, u_{it}, may be assumed to be independently distributed across stocks i. The market return, r_{mt}, is assumed common to all stocks. $\mathcal{F}_{i\,t-1}$ is the information set relative to stock i available just prior to date t; $\mathcal{F}_{i\,t-1}$ contains $u_{i\,t-1}$, $h_{i\,t-1}$ and all past returns on stock i. Note that conditional Normality is not required for the GARCH model (Bollerslev [1987]).

The GARCH model estimation differs from a straightforward ordinary least squares (OLS) estimation; you do not have a nice closed-form expression for $\hat{\alpha}_i$ or $\hat{\beta}_i$. In the GARCH estimation, you typically run OLS to get an initial guess for α_i and β_i. Then you adjust guesses of the γ_{ji}'s, α_i and β_i until you obtain what seem to be the most likely parameter estimates. This is a "maximum likelihood estimation" technique.[20]

Answer 3.13: You reduce exposure (i.e., hedge) by shorting T-bond futures contracts. Each Chicago Board of Trade (CBOT) T-bond contract covers a face value of \$100,000 of T-bonds. If the duration of your bond is the same as the duration of the cheapest-to-deliver (CTD) T-bond, then you short $\frac{\$50{,}000{,}000}{\$100{,}000} = 500$ contracts.[21] If the duration of your bond is different from the duration of the CTD T-bond, then you adjust for durations: go short $\frac{D_B}{D_F} \times \frac{\$50{,}000{,}000}{\$100{,}000} = \frac{D_B}{D_F} \times 500$ contracts, where D_B is the duration of your bonds, and D_F is the duration of the CTD T-bond.

Note that you could hedge by shorting Eurodollar futures (the underlying is the interest rate on a three-month \$1 million Eurodollar deposit). However,

[18] The review paper by Bera and Higgins (1993) is the best overview of ARCH and GARCH models that I have seen. Following this, you might look at Bera et al. (1988) as an introduction to ARCH, and also as an introduction to Engle (1982). For an introduction to statistical models for financial market volatility, see Engle (1993) and his references. For a higher-level review of ARCH modelling in finance, see Bollerslev et al. (1992). For a concise overview of the broad econometric peculiarities of the ARCH(1) model, see Hendry (1986).

[19] If you remove the term $h_{i\,t-1}$ from the second moment of the GARCH(1,1) model, you get the ARCH(1) model.

[20] See Berndt et al. (1974) for details on a good maximum likelihood estimation technique. See Bollerslev (1986) and Greene (1993) for more on the GARCH model.

[21] For more details on the CTD bond, see Hull (1997, pp92–93); for details on duration-based hedging, see Hull (1997, pp102–104).

the short end of the yield curve does not move with the long end. It, therefore, makes sense to use a hedging instrument whose underlying interest has maturity as close as possible to the portfolio to be hedged.

Answer 3.14: "Brady bonds" are sovereign bonds issued by developing countries in exchange for previously rescheduled bank loans. They are either "Par" bonds or "Discount" bonds. The former were issued at the par value of the loans but carry a below-market interest rate; the latter were issued at a discount from the face value of the former loans but carry a (floating) market interest rate. About a quarter of the market value of Brady bonds is collateralized by US Treasury issues. The size of this collateralization means that Brady bonds are sensitive to changes in US interest rates. In fact, something like a quarter of the variation in price movements of Brady bonds is (statistically) explained by moves in US Treasuries (sometimes with a lag of one day). Mexico has retired all its Brady bonds, but the arguments apply more generally, so I kept the question.[22]

Let us assume that the yield on the Brady bond increases by 25 bps (i.e., one quarter of the US Treasury yield change). If we assume that the duration of the Brady bond is about 15 years, that the bond is trading at around par of \$1,000, and that the Mexican yield curve is flat at around 8%, then the price response would be (denoting yield by y)

$$\begin{aligned}\Delta P &\approx -DP\frac{\Delta y}{(1+y)}\\ &= -15\times\$1,000\times\frac{0.0025}{1.08}\\ &= -\$15,000\times\frac{0.0025}{1.08}\\ &= -\frac{\$37.50}{1.08}\approx -\$35.\end{aligned}$$

With these assumptions, my guess is that the Brady bond price goes down by about three or four percentage points.

Answer 3.15: This question is similar to Question 3.10. The zero-coupon corporate bond has the same duration as longer-term coupon-bearing treasuries. You should short the corporate bond and buy treasuries that have the same duration and value as the corporate. By matching on duration and value, you create a zero-net investment portfolio that reaps profits.

Answer 3.16: First of all, the 5/10 time span is not relevant. The same result holds for a 1/2-year time span. That is, if the one-year interest rate is 10%, and two-year interest rate is 15%, then the forward rate for the second year is close to, but strictly greater than, 20%. Second, the order of the rates is

[22]This summary benefited from an unpublished research report prepared for Merrill Lynch by a group of my former students at MIT.

not important. That is, if the two-year rate is 15%, and the forward rate for the second year is 10%, then the one-year rate is close to, but strictly greater than, 20%. Third, the result holds for effective (i.e., simple) interest rates but does not hold for continuously compounded interest rates (for which the approximation is exact).

The argument relies upon the way in which the interest on your interest accumulates. If you are offered 10% for the first year and 20% for the second year, you will not do as well as if you are offered the average (15%) for two years. Although the interest on the principal is the same in both cases (and equal to 30%), the interest on the interest is not the same (15% of 15% equals 2.25% and exceeds 20% of 10%, which is only 2%). To avoid arbitrage, the "plug" rate has to exceed 20%. That was the "plain English" approach.

The result can be proved using math. Let R_1 and R_2 be two different interest rates, then

$$\begin{aligned}\left(\frac{R_1+R_2}{2}\right)^2 - R_1R_2 &= \frac{1}{4}(R_1^2+2R_1R_2+R_2^2-4R_1R_2)\\ &= \frac{1}{4}(R_1^2-2R_1R_2+R_2^2)\\ &= \frac{1}{4}(R_1-R_2)^2 > 0.\end{aligned}$$

It follows that $\left(\frac{R_1+R_2}{2}\right)^2 > R_1R_2$. This means that the interest on the interest is better at the average rate than at the product of rates—as stated above.

The result may also be written as $\left(\frac{R_1+R_2}{2}\right) > \sqrt{R_1R_2}$. This is a special case of a more general result that an arithmetic average exceeds a geometric average. This result is true beyond the case $n=2$ and can be extended to encompass harmonic averages also. Let $\mathcal{A}$, $\mathcal{G}$, and $\mathcal{H}$ denote the arithmetic, geometric, and harmonic averages of the positive numbers $x_1, x_2, \ldots, x_n$ as follows:

$$\begin{aligned}\mathcal{A} &\equiv \frac{1}{n}\sum_{i=1}^{n}x_i = \frac{x_1+x_2+\cdots+x_n}{n},\\ \mathcal{G} &\equiv \sqrt[n]{\Pi_{i=1}^{n}x_i} = \sqrt[n]{x_1x_2\ldots x_n}, \text{ and}\\ \mathcal{H} &\equiv \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots\frac{1}{x_n}}\end{aligned}$$

Then the following result holds (Spiegel [1968]):[23]

$$\mathcal{A}\geq\mathcal{G}\geq\mathcal{H},$$

and the inequalities are equalities only in the special case where

$$x_1 = x_2 = \cdots = x_n.$$

[23]To help you remember the ranking $\mathcal{A}\geq\mathcal{G}\geq\mathcal{H}$, note that it is the same as the ranking of the letters A, G, and H in the Latin alphabet.

Answer 3.17: If the one-year rate is 12%, and the two-year rate is 18%, then the forward rate for the second year is 24% to a first-order approximation (it is exactly 24% if these are continuously compounded rates). Let us assume this is 12% per half-year in the second year. Then your discounted expected payoff to the game is approximately

$$\begin{aligned}\left(\frac{1}{2}\times -\$2\right)+\left(\frac{1}{2}\times\frac{\$7}{(1.12)(1.12)}\right) &\approx -\$1+\frac{\$3.50}{1.25}\\ &= -\$1+\frac{14}{5}\\ &= \$1.80.\end{aligned}$$

If you can play repeatedly, then you are risk-neutral, and you would pay anything up to about \$1.80 to play this game. If you can play only once, then you might argue that the amount is so small you are still risk-neutral. If you multiply everything by a factor of one million, then you'll need to add a risk premium to the discount rates, and you will not pay as much to play.[24]

> **Story:** 1. Announced she hadn't had lunch and proceeded to eat a hamburger and french fries in the interviewer's office. 2. Without saying a word, candidate stood up and walked out during the middle of the interview.
>
> Interview Horror Stories from Recruiters
> Reprinted by kind permission of *MBA Style Magazine*
> ©1996–2009 MBA Style Magazine, `www.mbastyle.com`

Answer 3.18: No one wants to trade with the informed (i.e., insider) trader because you almost always lose to someone who is better informed than you are. The identity of the informed trader has not been announced. This means that *any* trade could be a losing trade. Traders will, therefore, be reluctant to trade. This leads directly to decreased trading volume.

Here is another way to look at it. Uncertainty over the identity of the informed trader means that traders widen their bid-ask spreads to compensate (on average) for any potential losses. Wider bid-ask spreads is one component of a decrease in liquidity, and it is usually associated with a decreased volume of trade (Chordia and Subrahmanyam [1995]).

Answer 3.19: The very first thing I check (see the next question!) is whether $\rho=\frac{\min(\sigma_1,\sigma_2)}{\max(\sigma_1,\sigma_2)}$. Finding that it is not (because this ratio is 2/3 here and $\rho=0.50$), we proceed as follows.

Let $\sigma_1=0.20$, $\sigma_2=0.30$, and $\rho=0.50$ be the standard deviations and correlation, respectively. Let ω be the weight put into Stock 1. The portfolio

[24] The risk premium as a function of the size of the bet is discussed by Tversky and Kahneman (1981) and Kahneman and Tversky (1982). Tversky and Kahneman (1974) is an earlier article you might like to read before reading these two.

variance is just $\sigma^2=\omega^2\sigma_1^2+(1-\omega)^2\sigma_2^2+2\omega(1-\omega)\sigma_1\sigma_2\rho$. Differentiate this with respect to ω to get

$$\begin{aligned}\frac{\partial\sigma^2}{\partial\omega} &= 2\omega\sigma_1^2-2(1-\omega)\sigma_2^2+2(1-\omega)\sigma_1\sigma_2\rho-2\omega\sigma_1\sigma_2\rho\\ &= 2[\omega(\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho)-\sigma_2^2+\sigma_1\sigma_2\rho].\end{aligned}$$

This is zero when $\omega=\frac{\sigma_2(\sigma_2-\sigma_1\rho)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho}$. In our particular case this ratio is $\omega=\frac{0.06}{0.07}=0.8571...$ and this gives $\sigma=0.1964$.

It is good practice to check the second order condition: $\frac{\partial^2\sigma^2}{\partial\omega^2}=2(\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho)=0.14>0$, so it is a minimum.

Answer 3.20: In the previous question I said that the very first thing I check is whether $\rho=\frac{\min(\sigma_1,\sigma_2)}{\max(\sigma_1,\sigma_2)}$. It was not in that question, but it is in this one (i.e, $\sqrt{\frac{0.10}{0.40}}=0.50$), so the approach is different.

I have never seen it written down in a book, but it is well known that there are several cases for ρ in the two-asset portfolio (these can each be deduced from the first order condition in Answer 3.19):

- $\rho=-1$: then $\omega=\frac{\sigma_2}{\sigma_1+\sigma_2}$, and $\sigma=0$. This is the case of perfect negative correlation and a zero-risk portfolio. No shorting is required.
- $-1<\rho<\frac{\min(\sigma_1,\sigma_2)}{\max(\sigma_1,\sigma_2)}$: then $\omega=\frac{\sigma_2(\sigma_2-\sigma_1\rho)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho}$, this does not involve short selling, and variance reduction occurs below that of either asset. This is the case in Question 3.19.
- $\rho=\frac{\min(\sigma_1,\sigma_2)}{\max(\sigma_1,\sigma_2)}$: then all money goes into the lowest volatility asset to minimize volatility (so $\omega=1$ if $\sigma_1<\sigma_2$, and $\omega=0$ if $\sigma_2<\sigma_1$), and $\sigma=\min(\sigma_1,\sigma_2)$.
- $\frac{\min(\sigma_1,\sigma_2)}{\max(\sigma_1,\sigma_2)}<\rho<1$: then $\omega=\frac{\sigma_2(\sigma_2-\sigma_1\rho)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho}$, this does involve short selling, and variance reduction occurs below that of either asset.
- $\rho=1$: then $\omega=\frac{\sigma_2}{\sigma_2-\sigma_1}$, and the high volatility asset is shorted to over-invest in the low volatility asset, and the optimum is a zero-risk portfolio.

This question is for the middle case above, and we should put all our money into the low volatility asset. Unlike the previous question, you just have to remember the ratio of the standard deviations (only one ratio is feasible for the standard deviations because the reciprocal ratio would be bigger than 1), and if the correlation equals the feasible ratio then all money goes into the low volatility asset. The interviewer is trying to find out whether you know this result.

Support the author!
If you like this book, buy a copy!

Appendix D

Statistics Answers

This appendix contains answers to the questions posed in chapter 4.

Answer 4.1: This sort of question is common. Begin by calculating the expected payoff to the game. This is given by the summation over the product of potential outcomes times their probability of occurrence:

$$\left(\tfrac{1}{6} \times \$1\right) + \left(\tfrac{1}{6} \times \$2\right) + \left(\tfrac{1}{6} \times \$3\right) + \left(\tfrac{1}{6} \times \$4\right) + \left(\tfrac{1}{6} \times \$5\right) + \left(\tfrac{1}{6} \times \$6\right) = \$3.50$$

If you are selling tickets to repeated plays of this game, you are effectively risk-neutral.[1] This means you should charge the expected payoff ($3.50) plus a margin for profit. You choose how wide to make the margin—it depends on your overhead, monopoly power, greed, and so on. You cannot charge $6.00 or above, since no one will play. If the game is to be played only once, then you are risk-averse. You should charge the expected value, plus your profit margin, plus a risk premium. The risk premium depends upon how risk-averse you are.

Answer 4.2: Another die-rolling question; they are very popular. You want to get as many dollars as possible. You let me roll once and look at which number comes up. You must compare this number to the possible payoffs on the remaining two rolls. If it seems likely that you can do better by not stopping the game, then you proceed, otherwise you stop me.[2]

You must work backwards to deduce the best strategy. This is analogous to pricing an American-style option using a tree method. So, suppose that

[1] This is an application of the "Weak Law of Large Numbers." The law says, essentially, that if you independently draw repeated observations from the same random distribution, then for very many drawings, the sample mean is very close to the population mean (DeGroot [1989, p229–231]). In other words, after many repeated plays of the game, the ticket seller can be sure that his average payout per game is very close to the expected payout per game. Because all that matters is the expected payout, not the variance of payouts, the ticket seller is effectively risk-neutral. Similarly, casinos are effectively risk-neutral. With repeated plays, and odds slightly in the favor of "the house," the casino expects to be the winner for sure in the long run.

[2] I thank Bingjian Ni for suggesting the solution technique. I am responsible for any errors.

you have seen the second roll and are trying to decide whether to ask for a third. You must compare the outcome of the second roll to the distribution of possible outcomes on the third roll:

Table D.1: Distribution of Payoff to Third Roll of a Die

Maximum Payoff	\$1	\$2	\$3	\$4	\$5	\$6
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

The expected value of the distribution in table D.1 is \$3.50; the variance is \$2.92; the standard deviation is \$1.71. If you see a 4 or higher on the second roll, you might stop the game because you probably will not do better. If you get a 3 or lower, you might continue because you expect to do better.

Now, stepping backwards again, suppose that you have just seen the first roll. You must decide whether to ask for a second roll (which may lead to a third). You must compare the outcome of the first roll to the distribution of possible outcomes if you proceed to a second (and possibly third) roll.

If you ask for the second roll, there is one-half a chance that it yields a 1, 2, or 3, and one-half a chance that it yields a 4, 5, or 6. Using the argument above, in the first case (1, 2, or 3 on roll two) you proceed to a third roll; in the second case (4, 5, or 6 on roll 2) you do not proceed. There is thus one-half a chance that you proceed to a third roll (expected value \$3.50 from table D.1), and one-half a chance that you stop the game at roll two (expected value $\frac{\$4+\$5+\$6}{3}$). It follows that the expected value of asking for a second roll is

$$\left(\frac{1}{2}\times\$3.50\right)+\left[\frac{1}{2}\times\left(\frac{\$4+\$5+\$6}{3}\right)\right]=\$4.25.$$

Thus, you would ask for a second roll only if you get a 1, 2, 3, or 4 on roll one. If you have a 5 or 6 on roll one, you should stop the game.

In simple terms, then, the strategy is to stop the game at roll number one if a 5 or 6 appears (probability $\frac{1}{3}$), otherwise continue (probability $\frac{2}{3}$). If you continue, stop the game at roll number two if a 4, 5, or 6 appears, otherwise continue.

Please note that my argument involving expected payoffs assumes that you are risk-neutral; your stopping rule might use lower acceptable payoffs if you are risk-averse, or higher payoffs if you are risk-loving.[3]

The overall expected value of the game may now be calculated.

$$\text{Value}=\left(\frac{2}{3}\times\$4.25\right)+\left[\frac{1}{3}\times\left(\frac{\$5+\$6}{2}\right)\right]=\frac{\$14}{3}\approx\$4.67.$$

[3]In addition, you should question my treatment of discreteness. For example, although you cannot roll a "$3\frac{1}{2}$" or a "$4\frac{1}{4}$," I use these as cutoff points when deciding whether to proceed or not.

If you are charging entry to repeated plays of this game, you are effectively risk-neutral.[4] You charge the expected value (\$4.67) plus a commission. You add a risk premium to the ticket price if there is only one or a few plays of the game; the more plays, the lower the risk premium. You would never charge more than six dollars because the player can never earn more than six dollars.

In the amended game (where I roll the die three times and pay you the maximum number of the three rolls), you need the distribution of the maximum payoff to three rolls of a die; this distribution is given in table D.2.[5]

Table D.2: Distribution of Maximum Payoff in Three Rolls of a Die

Maximum Payoff	\$1	\$2	\$3	\$4	\$5	\$6
Probability	$\frac{1}{216}$	$\frac{7}{216}$	$\frac{19}{216}$	$\frac{37}{216}$	$\frac{61}{216}$	$\frac{91}{216}$

The mean of the distribution of the maximum payoff from three rolls of the die is $\$\frac{1071}{216}=\$4.96$; the variance is $\$\frac{61047}{46656}=\$1.31$; and the standard deviation is \$1.14 (all calculated using information in table D.2). You should, therefore, charge a ticket price of \$4.96 plus some profit margin for repeated plays. Again, you cannot charge more than six dollars because no one will play the game.

The second game is more expensive than the first game (\$4.96 > \$4.67) because it strictly dominates it. That is, the payoff to the second game is never less than, and often exceeds, the payoff to the first game. This is because the second game *guarantees* the maximum of three rolls without risk, but the first game does not.

Answer 4.3: This has been a very popular question. Assume that neither of you peek into your envelopes. Assume that you have $\$X$ in your envelope, where $\$X$ has a fifty-fifty chance of being either $\$m$ or $\$2m$. This means that your opponent's envelope has a fifty-fifty chance of containing $\$2X$ or $\$\frac{1}{2}X$. The expected value of switching is

$$\left(\frac{1}{2}\times\$2X\right)+\left(\frac{1}{2}\times\$\frac{1}{2}X\right)=\$1.25X.$$

The expected *benefit* of switching is, therefore, $\$0.25X$. On this basis, it looks as though you should switch envelopes. Of course, if your opponent does not peek, and she has $\$Y$ in her envelope, exactly the same argument shows that she has an expected benefit to switching of $\$0.25Y$. So, it looks as though she

[4]This is the "Weak Law of Large Numbers" again. See Footnote 1 (on page 187) and DeGroot (1989, p229–231).

[5]Can you use elementary statistics to prove that this probability distribution is described by $\text{Prob}(Max=m)=\frac{m^3-(m-1)^3}{216}=\frac{3m(m-1)+1}{216}$, where m is the maximum of three rolls of the die? If you cannot, you need to work on your statistics.

should switch also. This is the first part of the "Exchange Paradox": it seems that you *both* benefit from switching.

Now, suppose that neither of you peek and that you do switch envelopes once. If you still do not peek, then a repeat of exactly the same argument suggests an expected benefit of 0.25 of the contents of your envelope if you switch again. The same applies to your opponent. This is the second part of the "Exchange Paradox": it seems that you could happily switch forever (like a dog chasing its own tail). The foregoing is the naive answer.

The problem is twofold: First, you are assuming that value is expected payoff (this is so only if you are genuinely risk-neutral);[6] second, your "prior" beliefs are that you have a fifty-fifty chance of having either $\$m$ or $\$2m$. The first problem is a function of your individual risk preferences and is difficult to address. The second problem can be tackled using two approaches: the first approach is to reconsider the nature of your prior; the second approach is to "update" your prior probability assessment (this is "Bayesian" statistics as opposed to "classical" statistics).

The first approach is to reconsider the nature of your priors. Our previous (paradoxical) calculation yielded $\$1.25X$ as the expected payoff to switching. However, this assumes that for any given X, it is equally likely that your opponent has $\$2X$ or $\$\frac{1}{2}X$. If you do not peek, then you are assuming a "diffuse level prior" because you assume this equality of likelihood for *any* X. Your prior is, therefore, not a valid probability density function (pdf) because the probabilities—across X—do not sum to **1**. However, for any *particular* m, it is equally likely that you received one of $\$m$ or $\$2m$. Thus, for any particular m, your priors are a pdf and any paradoxes should disappear. The expected value of switching should be zero. This is easily demonstrated. Let $P(\$m)$ denote the probability that *you* got $\$m$ (the lower amount); let $E(V)$ denote the expected value to switching; then $E(V)$ is given by

$$\begin{aligned} E(V) &= [E(V|\$m) \times P(\$m)] + [E(V|\$2m) \times P(\$2m)] \\ &= \left(+\$m \times \frac{1}{2}\right) + \left(-\$m \times \frac{1}{2}\right) \\ &= \$0. \end{aligned}$$

The expected value is zero, and you are thus indifferent—resolving the paradox.[7] Note that $E(V|\$m) = +\$m$ because, conditional on your having been

[6] An aside is in order. In corporate finance, the present value of a projected random payout is the discounted expected cash flow. The discounting is done at a rate that incorporates risk (e.g., using the CAPM), and the expectation is a mathematical one using real world probabilities (Brealey and Myers [1991]). An alternative to the real world expected cash flow coupled with the risk-adjusted discount rate is a risk-neutral world expected cash flow coupled with a riskless discount rate. The former is popular in corporate finance; the latter is popular in option pricing (see Arnold and Crack [2004] and Arnold, Crack and Schwartz [2009, 2010]). With no discounting (e.g., the envelope question), value is expected payoff only if you are risk-neutral.

[7] I thank Andres Almazan for suggesting this type of solution technique. I am responsible for any errors.

given the envelope containing only $\$m$, you gain $\$m$ by switching.

The second approach is to update your prior. To update your prior, you need information. The most obvious source of information is to peek into your envelope. So, assume that both you and your opponent peek into your envelopes. Now it gets subjective. If you see an amount that *seems* very high, then you update your prior probabilities: the probability that you have the high-value envelope increases, and the probability that you have the low-value envelope decreases. You no longer see value in switching envelopes.[8] If you see an amount that *seems* very low, then you see value in switching. The problem now is that you must subjectively assess the amount in the envelope as being either "low" or "high." The "Bayesian Resolution of the Exchange Paradox" is covered in detail in Christensen and Utts (1992).

If you have both peeked, and you do switch, then you will not switch again. This is because one of you gained, and that person will not want to lose by switching back. A similar question (but with an upper bound on the quantities possible) appears in Dixit and Nalebuff (1991, chapter 13). The Dixit and Nalebuff book on strategic thinking is well worth a look.

Answer 4.4: There are many different possible answers; I give only two. Toss the coin twice. If you get HT, give the apple to the first child. If you get TH, give the apple to the second child. If you get HH, give the apple to the third child. If you get TT, then start again. This effectively takes TT out of the sample space.

A second solution is to toss the coin three times and assign the outcomes to the three children. Let T win. If one child beats the other two (i.e., the outcome is some permutation of $\{T, H, H\}$), then give the apple to the child who was allocated the T. Otherwise, toss three more times. This is isomorphic to a tournament where each child competes against each other child until one child beats both others.

Answer 4.5: For my first answer in Answer 4.4: If you toss the coin twice and get TT, then you have to start again. There is thus a chance that you will take more than 2 tosses to complete the strategy. We can use a recursion to find the expected number of tosses. Let N be the number of tosses required, then there is a three quarter chance that N will be 2, but a one quarter chance that N will be $2 + E(N)$ because you have to start again in the TT outcome. Thus $E(N) = \left[\frac{3}{4} \cdot 2\right] + \left[\frac{1}{4} \cdot (2 + E(N))\right] = 2 + \frac{E(N)}{4}$. Simple algebra now yields $E(N) = \frac{8}{3} = 2\frac{1}{3}$.

For my second answer in Answer 4.4: We will need at least three tosses in this case, but there are five chances in eight that we will need more than three

[8] However, you might argue that if you see an amount that seems so high that even one-half of it is more money than you can comprehend, you might switch envelopes just for the hell of it; it is worth the gamble.

tosses. Using a recursive argument similar to that above, we figure that the expected number of tosses in this case is $E(N) = 8$.

Answer 4.6: This is a very common question and has been in use since at least 1990. Well, the first thing to notice is that you are trying to replicate a \$100 bet on Team A to win the series and you are doing this via a series of small bets. At each step, there are two possible outcomes: Team A wins, or Team A loses. With replication, time steps, and binomial outcomes, the obvious thing to do is build a lattice for a replicating strategy (see figure D.1). To deduce the betting strategy in figure D.1, I first drew the lattice and identified the boundary nodes at which the game must end (marked with large dots). You start with \$100 in wealth. In the case where Team A has won four games, you must end up with an accumulated wealth of \$200 through \$100 in betting profits; In the case where Team B has won four games, you must end up with an accumulated wealth of \$0 through \$100 in betting losses. It is simple to step back from each pair of ending nodes (starting with the right-most pair) to deduce how much you must bet in each case (working back to \$31.25 on the first game) in order to end up replicating the payoffs at the boundary. If you follow this betting strategy, you are *guaranteed* to replicate the payoff to betting \$100 on Team A to win the series. The given probabilities are "red herrings" because you do not need any probabilities, physical world or risk-neutral, to solve this problem. Note, finally, that I have assumed that you earn no interest on your wealth.

Answer 4.7: This is elementary statistics, and one of the easiest questions in this book. The rules of the game have effectively removed the 1 from the sample space (i.e., the collection of possible outcomes). It follows that there are five possible outcomes (2 to 6), and each is equally likely. The expected outcome is simply

$$\frac{\sum_{i=2}^{i=6} i}{5} = \$4.$$

To do the sum in your head, remember that the dots on the opposing faces of the die add to seven. The sum must be three times seven, less one to give 20. Now divide by five to get the expected payoff of \$4.

> **Story:** Late one winter's evening at MIT (1994 I think), I was helping Nobel Prize winner Franco Modigliani operate our photocopier. We somehow got onto the topic of the Crash of 1987 and he said "Yes, that is when I made all my money." He said he had been watching the market and, thinking it overvalued, he had bought out-of-the-money index puts (presumably S&P500 index options at that time). He said he made a bundle. He had tried it several times since then without success. At my office doorway another time, he told me that when pronouncing his name I should "drop the 'g'—it's the mark of a true Italian"—and that is how he pronounced it.

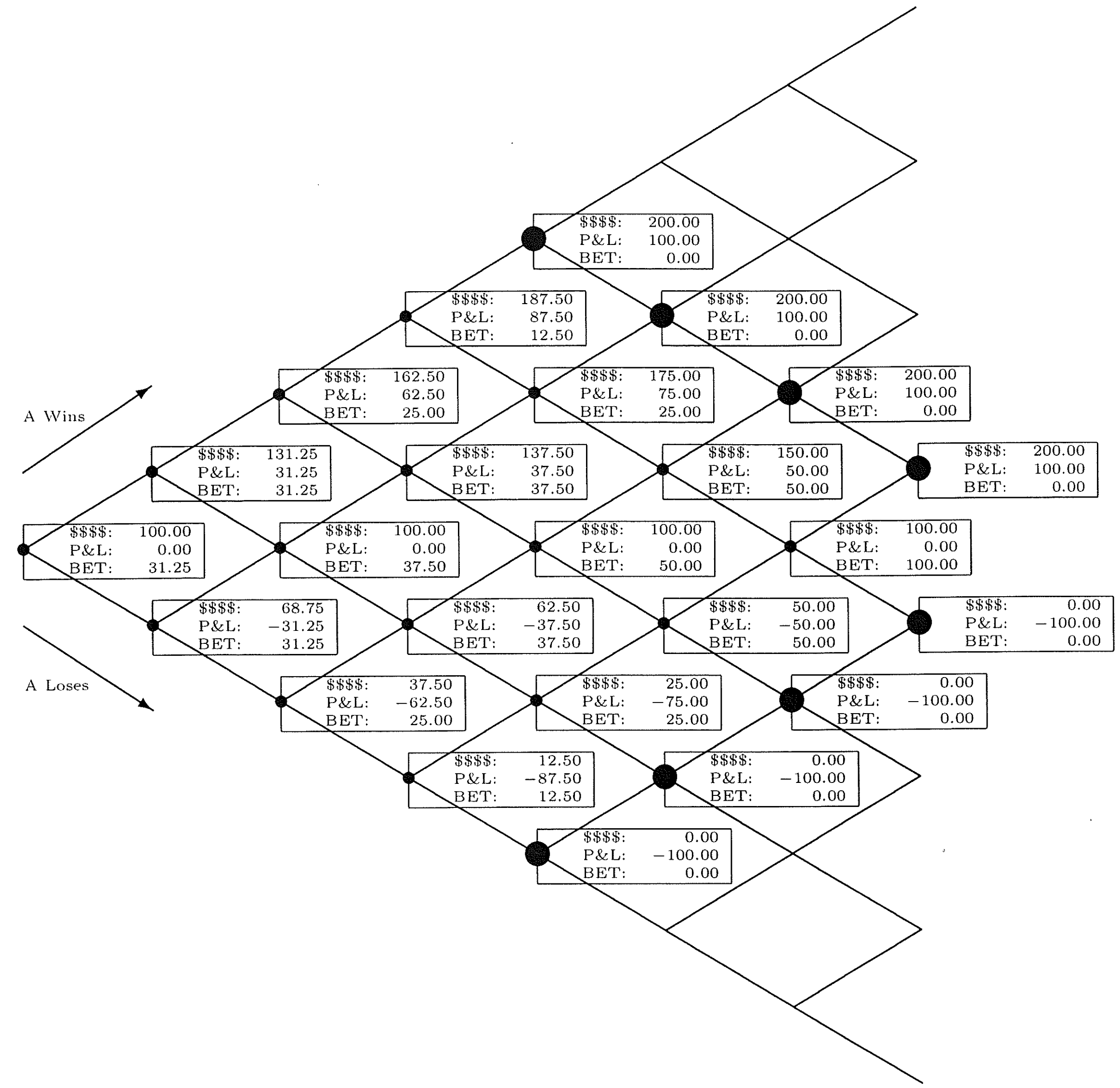

Figure D.1: World Series: Lattice of Betting Strategy

Note: See Answer 4.6 Each step has two possible outcomes for the game: Team A wins (the up step), or Team A loses (the down step). You start with \$100 in wealth. Each box shows how much wealth you have ("\$\$\$\$"), your cumulative profit or loss ("P&L") and the bet you place on Team A winning ("BET"). A large dot on a node indicates the end of the series (i.e., one team has reached four wins).

Answer 4.8: The naive answer is that the probability is just $\frac{2}{52} \approx 4\%$. This is incorrect. There are four chances that the first card dealt to you (out of a deck of 52) is a King. Conditional on the first card being a King, there are three chances that the second card dealt to you (out of the remaining deck of 51) is a King. Conditional probability says that

$$P(\text{Both are Kings}) = P(\text{Second is a King} \mid \text{First is a King}) \times P(\text{First is a King})$$

where " | " is read as "conditional upon," or "given." This is a special case of the more general conditional probability result:

$$P(A \cap B) = P(A \mid B) \times P(B)$$

Thus, $P(\text{Both are Kings}) = \frac{3}{51} \times \frac{4}{52} = \frac{1}{17} \times \frac{1}{13} = \frac{1}{221} \approx 0.5\%$. Therefore, you have roughly one chance in 200 of getting exactly two Kings dealt to you.

I wish you to avoid a common form of confusion. Please note that although you multiply probabilities to get the answer, and such multiplication is often done when dealing with independent events, the events here (King on first card, and King on second card given King on first card) are *dependent*, not independent. That is, you calculate the probability that the second card is a King given that, or *dependent upon*, the first card being a King.

I wish to emphasize that the above procedure is different from that for figuring out the probability that, for example, you get two heads in two tosses of a fair coin (this probability is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$). The outcomes of the coin tosses are genuinely *independent*, and this is why you can multiply their probabilities directly. That is, the probability of a head on the second coin toss is not influenced by the event that you get a head on the first coin toss. However, the probability that you get a King on the second card dealt *is* influenced by the event that you get a King on the first card dealt. That is why the conditional probability theory is used. Be sure you understand the distinction and how and where to apply each method. If it is not clear, go to your favorite statistics book for a review (e.g., see Feller [1968, chapter V]).

Answer 4.9: The "Let's Make a Deal" or "Monty Hall" problem is very frequently asked. Many people find it very difficult.

Assume that you choose Door 3. The host opens Door 2 and offers you the chance to switch to Door 1. Should you do it? If you have decided that it does not matter whether you switch doors or not (indifference), or that you should definitely not switch (aversion), then you should go back and think again before reading any further. Stop here and try again.

Let me begin with very simple intuition. My experience, however, is that many readers cannot accept the simple intuition, and for them I provide a formal proof using Bayes' Theorem.

SIMPLE INTUITION

Assume for a moment that you have already decided that you will switch doors. What then is the probability that you will find the prize behind the door you switch to? Well, you win the prize if you *originally* chose one of the two doors that has nothing behind it. In that case, the host shows you the other empty door, and switching yields the prize. So, the problem reduces to figuring the probability that you originally chose one of the two doors that has nothing behind it. That unconditional probability is just two thirds by construction. You thus have probability two thirds that you win by switching and one third that you lose by switching. So, you should switch![9]

FORMAL BAYES' THEOREM PROOF

If you are to play this game repeatedly, two-thirds of the time you profit by switching, and one-third of the time you lose by switching. Let B_k denote the event that the prize is behind Door number k ("B" for behind). Let H_j denote the event that you see the host open Door number j ("H" for host).

The unconditional probabilities of the location of prizes (probabilities calculated without conditioning on which door the host opens) are simply $P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$. What you need to know is the conditional probability $P(B_1|H_2)$. That is, the probability that the prize is behind Door 1 given that you see (or "conditional on") the host open Door 2. We use a straightforward application of conditional expectations and Bayes' Theorem (see Feller [1968, chapter V]), as follows:

$$P(B_1|H_2) = \frac{P(B_1 \cap H_2)}{P(H_2)} = \frac{P(H_2 \cap B_1)}{P(H_2)} = \frac{P(H_2|B_1) \times P(B_1)}{P(H_2)}$$

You know that $P(B_1) = \frac{1}{3}$, but what about $P(H_2|B_1)$ and $P(H_2)$? You know that the host is going to show you an empty door other than the door you choose (assume through all of this that it is Door 3 that you choose). The host's door must be revealed empty and cannot be the same door that you choose. Therefore, it must be that if you choose Door 3, then $P(H_2|B_1) = 1$.

Now, $P(H_2)$ is given by

$$\begin{aligned} P(H_2) &= [P(H_2|B_1) \times P(B_1)] + [P(H_2|B_2) \times P(B_2)] \\ &+ [P(H_2|B_3) \times P(B_3)], \end{aligned}$$

so some extra terms need to be calculated to get $P(H_2)$.

Well, the host's door must be shown to be empty, so it must be that $P(H_2|B_2) = 0$. The host is impartial, so it must be that $P(H_2|B_3) = \frac{1}{2}$ [and $P(H_1|B_3) = \frac{1}{2}$]. Thus, $P(H_2)$ is given by

$$\begin{aligned} P(H_2) &= [P(H_2|B_1) \times P(B_1)] + [P(H_2|B_2) \times P(B_2)] \\ &+ [P(H_2|B_3) \times P(B_3)] \\ &= \left(1 \times \frac{1}{3}\right) + \left(0 \times \frac{1}{3}\right) + \left(\frac{1}{2} \times \frac{1}{3}\right) = \frac{1}{2}. \end{aligned}$$

[9] I thank Jun Chung for this simple argument; any errors are mine.

It follows that the probability of finding the prize if you switch doors is two-thirds:

$$P(B_1|H_2) = \frac{P(H_2|B_1) \times P(B_1)}{P(H_2)} = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

The summary in table D.3 may clarify matters further. You choose Door 3. The host must choose an empty door to open. If the prize is behind Door 1, he *must* open Door 2 [$P(H_2|B_1) = 1$]. However, if the prize is behind Door 3, he can *choose* between Doors 1 and 2 [$P(H_2|B_3) = \frac{1}{2}$]. If you see Door 2, it is either because the prize is behind Door 1, and the host had no choice, or it is because the prize is behind Door 3, and the host randomly chose between Doors 1 and 2. It, therefore, follows that if you choose Door 3, and Door 2 is revealed empty by the host, the prize is twice as likely to be behind Door 1 as it is to be behind Door 3. Continuing along this line of thought, we may take

Table D.3: The Monty Hall Problem

Assume You Choose Door 3			
Prize Location B_j	Host Opens H_i	Unconditional Probability $P(H_i \cap B_j)$	Conditional Probability $P(H_i|B_j)$
1	2	$\frac{1}{3}$	1
2	1	$\frac{1}{3}$	1
3	1	$\frac{1}{6}$	$\frac{1}{2}$
	2	$\frac{1}{6}$	$\frac{1}{2}$

a frequentist approach. Suppose you play the game repeatedly and always choose Door 3. If you look at all the times the host reveals Door 2 empty, you will find that two-thirds of the time the prize lies behind Door 1, and one-third of the time it is behind Door 3. Seeing Door 2 empty is thus a stronger signal that Door 1 has the prize than it is that Door 3 has it. This argument is more general, of course. Whichever door you choose, seeing the host reveal an empty door is a signal that you should switch.

Answer 4.10: The argument is the same as the previous solution. The conclusion is the same: You should switch!

Answer 4.11: This question is solved most efficiently by trying a few possible combinations, not by some time-consuming feat of constrained linear optimization. You should begin with extreme distributions, or with symmetrical distributions. It is in the extremes or in symmetry that solutions to such problems usually lie.

The probability of selecting a white marble is maximized (at almost $\frac{3}{4}$) by placing one white marble in one jar and the remaining 99 marbles in the other. The probability of selecting a white marble is minimized (at $\frac{1}{4}$) by placing all 100 marbles in one jar (assuming you do not get a second chance if the jar you choose is empty). If zero marbles in one jar is not an acceptable answer to you, then you minimize the probability of a white marble (at just over $\frac{1}{4}$) by maximizing the probability of a black one. That is, put one black marble in one jar and the remaining 99 marbles in the other.

Answer 4.12: This is a tough "game theory" problem. Early editions of my book included a full and formal solution to this problem. It was more than five pages long and far too detailed. I have now cut my answer down to bare bones key issues only.[10]

Mr. 30 and Mr. 60 are going to shoot at each other because they do not see you as an immediate threat; you do not die first because Mr. 60 and Mr. 30 are shooting it out; you do not want to be put into a shoot-out where your opponent is a very good shot and gets to shoot first; if Mr. 30 gets to shoot before Mr. 60, it is less likely that you end up facing Mr. 60 than if Mr. 60 gets to shoot first, so you shoot in the air; if the direction of play is reversed, and Mr. 60 gets to shoot before Mr. 30, then you should help out Mr. 30 (and yourself) by shooting at Mr. 60 also, otherwise, leave it to Mr. 30; the cost of stepping in and shooting at Mr. 60 is that if you hit Mr. 60, you lose your chance to shoot first in the final shootout with Mr. 30; the benefit of stepping in and shooting at Mr. 60 is that you increase the likelihood of your facing Mr. 30 rather than Mr. 60 in the final shoot-out; there is a delicate balance between leaving it to Mr. 30 and stepping in to help him out, and it changes with the direction of play. Finally, there is a slim chance that everyone shoots at the sky, but this requires some sort of cooperation.

Answer 4.13: If you take the three-point shot, you have a 40% chance of winning. If you take the two-point shot, you have a 70% chance of a tie, and conditional on a tie you have a 50% chance of winning in overtime. Informally, the probability of winning if you take the two-point shot is thus 70% multiplied by 50%, which is 35%. This is lower than for the 40% for the three-point shot, so you should take the three-pointer.

More formally, let "W" denote winning, let "2" denote taking the two-point shot, let "T" denote sinking the two-pointer and getting a tie, and let "T^C" denote missing the two-pointer and not getting the tie (the "C" is for complement, that is, the remainder of the sample space). Then

$$\begin{aligned} P(W|2) &= P(W|T)P(T|2) + P(W|T^C)P(T^C|2) \\ &= (0.50 \times 0.70) + (0 \times 0.30) \\ &= 0.35. \end{aligned}$$

[10] I thank Olivier Ledoit for this solution technique. Any errors are mine.

Answer 4.14: This is one of the easier problems. If the cost is \$1.50 per spin, and you may play as often as you want, then yes, you should play. The expected payoff is \$1.80 per spin ($\sum_{i=1}^{i=5} \text{Payoff}_i \times \frac{1}{5} = \$1.80$). If you can play as often as you want, you are risk-neutral (in the long run, your average payoff will equal the expected payoff), and you expect to make \$0.30 per spin on average.

If you get only one spin, then whether you play or not depends upon whether the expected \$0.30 gain is sufficient to compensate you for the risk of losing \$0.50 (the \$1.50 cost less the \$1.00 worst possible payoff). With amounts this small, you would probably take the bet. It is like spending \$1.50 on a lottery ticket—it is too small to care about. If the numbers were larger, say everything multiplied by one billion, and if your job is lost if you lose, then you are significantly more risk-averse, and your boss would not want you to take the bet.

Answer 4.15: Assuming no special information on your part, each sports match presents a fifty-fifty chance of winning. Assuming each match is independent of each other, then winning is analogous to tossing a fair coin four times in a row and trying to get four heads. This probability is only $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$. The odds of winning are thus much worse than the odds offered by the bookie, and you should not play unless you are a risk-seeker. If the odds were raised to 25-to-1, this would be an attractive bet.

Answer 4.16: The standard deviation is just the square root of the expected squared deviation from the mean. Assuming equally likely probabilities, the mean of (1, 2, 3, 4, 5) is 3. The squared deviations are 4, 1, 0, 1, 4. The expected squared deviation is 2. The standard deviation is thus $\sqrt{2} \approx 1.4142$. I expect you to know $\sqrt{2}$ to four decimal places.

Answer 4.17: All you need is simple statistics. What happens if you ask the interviewer to shoot without spinning again? The first time the trigger was pulled, no bullet was found. It follows that *that* empty chamber will not be the next chamber. Also, if the first chamber was empty, then it certainly did not hold the first of the two contiguous bullets, Bullet #1, so you will not meet the second of the two bullets, Bullet #2. Thus, there are only four chambers that you might meet: three empty and one containing Bullet #1. You have one chance in four of not having to talk about your resume.

If you do ask the interviewer to spin the barrel again, then you have the same chance you had when you sat down initially. That is, there is one chance in three that you do not have to talk about your resume. It follows that you are better off not spinning.

In summary, because the first chamber did not contain a bullet, then it was not Bullet #1, so you know you will not see Bullet #2. You face only one possible bullet from the remaining four chambers. However, spinning the barrel again puts both bullets into play, and that is not a choice you want to make.

Answer 4.18: Before we look at the formal math, let's use some informal intuition. There is one chance in a thousand (unconditionally) that you plucked the two-headed coin (which would certainly explain 10 heads in a row). There is also about one chance in a thousand that a fair coin would give 10 heads in a row (because $\left(\frac{1}{2}\right)^{10} = \frac{1}{1024} \approx \frac{1}{1000}$). Looking at the event (10 heads), I'd have to say that the coin is roughly equally likely to be two-headed or fair.

Now turn to the formal math – a direct application of Bayes' Theorem. Let "TH" denote the event that your coin is the two-headed one. Let "$10H$" denote the event that you toss one of the pennies and get 10 heads. Let X^c denote the complement of an event X. Then

$$\begin{aligned} P(TH|10H) &= \frac{P(TH \cap 10H)}{P(10H)} \\ &= \frac{P(10H|TH)P(TH)}{P(10H|TH)P(TH) + P(10H|TH^c)P(TH^c)} \\ &= \frac{1 \times \frac{1}{1000}}{\left[1 \times \frac{1}{1000}\right] + \left[\left(\frac{1}{2}\right)^{10} \times \frac{999}{1000}\right]} \approx \frac{1}{2}, \end{aligned}$$

where I used the facts that $2^{10} = 1024 \approx 1000$, and $\frac{999}{1000} \approx 1$. So, given the 10 heads, you have about a half a chance that you have the two-headed coin—as per our intuition.

Answer 4.19: You win with probability 1/3. Wind is effectively absent from the sample space—it does not affect your chances of winning or losing. You lose with probability 1/3 at the first turn. You thus have only a 2/3 possibility of even getting to turn a second card. If you do get to turn the second card, there is 50% chance that it will be Fire and you lose, and a 50% chance it will not be, and you win. Thus the probability of winning is 50% of 2/3.

Answer 4.20: Assuming that the players have fifty-fifty probabilities of playing Red or Blue,[11] each player has the same expected payoff: \$1. Player B has a variance of payoffs given by

$$\left[(0-1)^2 \times \frac{1}{2}\right] + \left[(2-1)^2 \times \frac{1}{2}\right] = 1,$$

whereas player A has a variance of payoffs given by

$$\left[(1-1)^2 \times \frac{1}{4}\right] + \left[(3-1)^2 \times \frac{1}{4}\right] + \left[(0-1)^2 \times \frac{1}{2}\right] = 1.5.$$

Thus, if you are risk averse, player B's position is favored (it offers the same expected return, but less risk).

[11] A mixed strategy (Nash) equilibrium exists where B plays Red with probability $\frac{1}{4}$ and A plays Red with probability $\frac{1}{2}$. In this case, the expected payoff to playing Red equals the expected payoff to playing Blue for each player. A's expected payoff is $\frac{3}{4}$, whereas B's is 1. Thus, B is favored. I thank Alex Butler for this argument. Any errors are mine.

Answer 4.21: We seek $F_{P|H}(p) = P(P \leq p|H) = P(A|H)$, where "$A$" denotes the event that $P \leq p$, and "H" denotes the event that you get a head. Let $f(u) \equiv 1$, $0 \leq u \leq 1$ denote the unconditional pdf of P. We apply Bayes' Theorem directly for $p \in [0, 1]$ to get[12]

$$\begin{aligned} F_{P|H}(p) &= P(A|H) \\ &= \frac{P(A \cap H)}{P(H)} \\ &= \frac{\int_0^p u f(u) du}{\int_0^1 u f(u) du} \\ &= \frac{\left(\frac{p^2}{2}\right)}{\left(\frac{1}{2}\right)} = p^2. \end{aligned}$$

As $p \to 1$, $F_{P|H}(p) \to 1$, and as $p \to 0$, $F_{P|H}(p) \to 0$ (just checking). This cdf produces the pdf $f_{P|H}(p) = 2p$ that is left-skewed and has a mean of 2/3—slightly above 1/2 as you might have expected.

Let "$750H/1000$" denote the event that you flip the coin 1,000 times and get 750 heads. In this case, the (conditional) distribution function is going to look much like the step function

$$F_{(P|750H/1000)}(p) \approx \begin{cases} 0, & 0 \leq p < 0.75, \\ 1, & 0.75 \leq p \leq 1. \end{cases}$$

This conclusion relies upon a Weak Law of Large Numbers argument (see Footnote 1 [on page 187], and DeGroot [1989, p229–231]). The naive answer is to work it out mathematically, using binomial distributions and such like, but it quickly gets very messy, and the result should be essentially the same.

Answer 4.22: This is well known. More generally, if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $E(e^X) = e^{\mu + \frac{1}{2}\sigma^2}$. If X is Normal, then e^X is Lognormal. So, this is the sort of knowledge that arises in analytical Black-Scholes derivatives work, or in setting up a Monte-Carlo simulation of price paths in a Black-Scholes world. See Crack (2009, section 2.2) for detailed discussion and examples. It is straightforward to prove this from first principles because the Normal pdf has an $e^{(\cdot)}$ kernel in it, and you just have to add the exponents, complete the square, and integrate out to see what is left over.

Answer 4.23: Both games have the same expected payoff: \$3.5 million. However, the second game has much less volatility than the first. The Weak Law of Large Numbers says that your actual payoff will be much closer to the expected payoff in Game Two. As a risk-averse individual, you choose Game Two.

[12] I thank Alex Vigodner for this answer.

Answer 4.24: Start with a simple case first: What if you only need the expected number of tosses required to get *one* head? Let N be the number of coin tosses, then we want to find $E(N|1H)$ (expected number of tosses given that you seek only one head). Toss the coin once. Either you get a head (with probability p) or you get a tail. If you get a tail, then you are recursively back where you started. That is, there is probability p that $N = 1$, and probability $1-p$ that you still have $E(N|1H)$ tosses to go after the one you already tossed. In other words,

$$E(N|1H) = (p \cdot 1) + (1-p) \cdot [1 + E(N|1H)].$$

Solving for $E(N|1H)$ gives $E(N|1H) = \frac{1}{p}$. Check: the higher is p, the lower is $E(N|1H)$, and when $p = 0.50$ (a fair coin), $E(N|1H) = 2$, all of which seems reasonable.

Now consider the case of two heads in a row. Well, to get two heads in a row, you first need one head "in a row," which requires the expected $\frac{1}{p}$ tosses just calculated. If you have this one head already, then there is probability p that your next toss will be a head, and probability $(1-p)$ that you are back where you started having performed $E(N|1H)$ plus one tosses already. In other words,

$$\begin{aligned} E(N|2H) &= p \cdot [E(N|1H) + 1] + (1-p) \cdot [E(N|1H) + 1 + E(N|2H)] \\ &= E(N|1H) + 1 + (1-p) \cdot E(N|2H). \end{aligned}$$

This last line implies that

$$E(N|2H) = \frac{E(N|1H) + 1}{p} = \frac{1+p}{p^2}.$$

In the case of a a fair coin, this gives $E(N|2H) = 6$. Exactly the same reasoning in the case of three heads in a row leads us to

$$E(N|3H) = \frac{E(N|2H) + 1}{p} = \frac{1+p+p^2}{p^3}.$$

In the case of a a fair coin, this gives $E(N|3H) = 14$. It should be clear that there is a pattern: in the case of J heads in a row,

$$E(N|JH) = \frac{\sum_{i=0}^{i=J-1} p^i}{p^J}.$$

This can be proved formally using numerical induction if you wish.

Answer 4.25: The short answer is that it will be roughly $\pm\frac{1}{\sqrt{1{,}000}}$ which is roughly $\pm\frac{1}{\sqrt{900}}$ which is $\pm\frac{1}{30}$ which is roughly $\pm 3\%$.

The long answer is still quite simple. Suppose you sample N people and record a 1 if they say they will vote for Candidate A and a 0 otherwise. You have N binomial trials. Let X_i denote the outcome of the ith trial, then

$$X_i = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } (1-p). \end{cases}$$

Let $Y = \sum_{i=1}^{i=N} X_i$, then Y/N is the estimator of p. It follows that

$$E(Y/N) = E(Y)/N = \sum_{i=1}^{i=N} E(X_i)/N = N \cdot p/N = p,$$

and letting "$V(\cdot)$" denote variance, we have

$$V(Y/N) = V(Y)/(N^2) \stackrel{*}{=} \sum_{i=1}^{i=N} V(X_i)/(N^2) \stackrel{**}{=} N \cdot p(1-p)/N^2 = \frac{p(1-p)}{N},$$

where "$*$" follows because the variance of the sum equals the sum of the variance when the random variables are independent, and "$**$" follows directly from the definition of X_i using the definition of variance as probability-weighted deviation about the mean: $V(X_i) = (1-p)^2 \cdot p + p^2 \cdot (1-p) = p(1-p)$. it follows that the standard error of Y/N is the square root of its variance: $\text{SE}(Y/N) = \sqrt{\frac{p(1-p)}{N}}$.

In the particular case of the question we have $\hat{p} = 0.60$, so our best guess for the standard error is $\sqrt{\frac{\hat{p}(1-\hat{p})}{N}} = \sqrt{\frac{0.60 \cdot 0.40}{1000}} = 0.01549$. So, a 95% confidence interval would be

$$\hat{p} \pm 1.96 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{N}} = 0.60 \pm 1.96 \cdot 0.01549 = 0.60 \pm 0.0304.$$

That is, the margin of error is roughly plus/minus 3%, which is what I said two-dozen lines above! So, how did I know that before working out the details? The key lies in two approximations. First, the 95% confidence interval for $\hat{p}$ is $\hat{p} \pm 1.96$ times the standard error, so let us just take that as $\pm$ 2.00 times the standard error. Now, the standard error itself is $\sqrt{\frac{p(1-p)}{N}}$, but for most cases of interest, p is roughly 0.50 (otherwise it would not be much of a horse race). So, $p(1-p)$ is roughly 0.25. In fact, $p(1-p)$ is an inverted parabola in p and it has slope zero at $p = 0.50$. So, there is a range about $p = 0.50$ where approximating $p(1-p)$ by 0.25 is good. For example, with $\hat{p} = 0.60$ we get $\hat{p}(1-\hat{p}) = 0.24$ which is very close.

So, we get a confidence interval which is

$$\hat{p} \pm (\text{roughly } 2) \times \sqrt{\frac{\text{roughly } 0.25}{N}},$$

but $\sqrt{0.25} = 0.50$, and 2 times this gives 1. So, we get a 95% confidence interval which is roughly

$$\hat{p} \pm \sqrt{\frac{1}{N}}.$$

This is a good approximation unless $\hat{p}$ is quite far from 0.50.

It is good if you can calculate approximate square roots of powers of 10 in your head: $\sqrt{1,000} \approx 30$, $\sqrt{10,000} = 100$, $\sqrt{100,000} \approx 300$, and so on. ...and you need to be able to invert them too (e.g., $1/30 \approx 0.03$).

Answer 4.26: This is an old/common interview question. It is a direct application of Bayes' Theorem. Question 4.18 was also a Bayes' Theorem question. In that case we tried some informal intuition before doing the math; let us try that here too.

Most people do not have the disease. If we were to randomly select from the disease-free population (which is 99.5% of the sample here, so that is not too far from what we are actually doing), we would get a positive test result 7% of the time. If we are randomly selecting from the *entire* population, we get a diseased person 0.5% of the time. So, overall it seems that for roughly 7.5% of the draws, we get a positive, but for only 0.5% of the draws would they have the disease. Given a positive test result, we focus only on the roughly 7.5% of the population that returns a positive. With this restriction, the likelihood of the person having the disease is the ratio of those with the disease to those with a positive: roughly 0.005/0.075. This ratio is just 2/3 with the decimal place moved, so our guess is 6.67%.

Now let us do the math. Let "+" denote a positive on the test. Let "D" denote the presence of the disease, and let "D^c" denote the complement of being diseased (i.e., no disease). Then from Bayes' Theorem

$$\begin{aligned} P(D|+) &= \frac{P(D \cap +)}{P(+)} \\ &= \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)} \\ &= \frac{(1 \times 0.005)}{(1 \times 0.005) + (0.07 \times 0.995)} \\ &\approx \frac{0.005}{0.07465} \\ &\approx 0.06698. \end{aligned}$$

So, our 6.67% estimate is not far off.

If the percentage with the disease were only 0.05%, then the answer would drop to 0.007096% (which is almost exactly the ratio of $P(D)$ to $[P(+|D^c)+P(D)]$—the dominant terms in the above calculations).

Answer 4.27: A "stars and bars" approach is the easiest.[13] Let each star denote a thousand-dollar investment and each bar denote a line of demarcation between funds. Then `*****|********||*******|` has 20 stars and four bars to indicate an allocation of ($5,000;$8,000;$0;$7,000;$0).

There are $\binom{24}{4} = \frac{24!}{4!(24-4)!} = \frac{24\cdot 23\cdot 22\cdot 21}{4\cdot 3\cdot 2\cdot 1} = 10,626$ ways to allocate the four bars to the 24 possible locations for the 24 symbols I used. The general answer is $\binom{N+k-1}{k-1}$ for N-thousand dollars invested into k funds. Please confirm that in the case $N = 20, k = 2$ this formula gives the simple answer $N + 1 = 21$.

A less elegant approach is to consider 20 independent generalized Bernoulli trials where we roll a five-sided die and record each result as a five-tuple with a 1 in the position corresponding to the die's outcome and zeroes otherwise. The five-tuple containing the cumulative total count of outcomes is distributed multinomial (DeGroot [1989, p297]). The number of possible fungible (i.e., financially indistinguishable) allocations of bills to buckets must be the same as the size of the sample space in this generalized experiment. For a multinomial with N generalized Bernoulli trials (each of k outcomes) the sample space is of size $\binom{N+k-1}{k-1}$, as above.[14]

Answer 4.28: We shall explore several solutions to the cookie dough problem because the nature of the problem allows us to practice many different skills: statistical inference, probability, combinatorics, recursion, induction, algebraic approximations, etc.

This question is a particular case of a more general question: How many chips, N, should go randomly into some cookie dough to give a probability of at least p that each of k cookies randomly cut from the batch contains at least m chips? The interviewer's particular case uses the high probability $p = 0.90$ and the low hurdle $m = 1$ chips per cookie.

Well, if you have fewer than 100 chips, then at least one cookie does not have a chip, so 100 is the lower bound on any feasible answer. You should take an initial gut instinct guess before any formal analysis. My gut instinct is that something like $N = 500$ chips is enough because I would then be fairly sure that each cookie in the batch gets a chip. Is 500 enough? We will look at two exact solutions and several approximations to find out.

EXACT SOLUTIONS

#1: Inclusion-Exclusion.[15] If p is the probability that every cookie has at least $m = 1$ chips, then $1 - p$ is the probability of the event that some cookie has no chips. This event in turn is a union of events involving individual cookies having no chips. Let A_i^0 be the event that cookie i has no chips, then

[13] I thank Chun Han for suggesting this approach to me; any errors are mine.
[14] Lyons and Hutcheson (1996).
[15] I thank Nate Coehlo for suggesting this approach; any errors are mine.

the inclusion-exclusion formula (see discussion on page 76) yields

$$\begin{aligned}
1-p &= P\left(A_1^0 \cup A_2^0 \cdots \cup A_k^0\right) \\
&= P\left(\bigcup_{i=1}^{k} A_i^0\right) \\
&= \sum_{i=1}^{k} P\left(A_i^0\right) - \sum_{1\le i<j\le k} P\left(A_i^0 \cap A_j^0\right) + \\
&\quad \sum_{1\le i<j<l\le k} P\left(A_i^0 \cap A_j^0 \cap A_l^0\right) - \cdots + (-1)^{k+1} P\left(A_1^0 \cap A_2^0 \cap \cdots \cap A_k^0\right) \\
&= \sum_{i=1}^{k} (-1)^{i+1} \sum_{\substack{\{j_1,j_2,\ldots,j_i\}\subseteq\{1,2,\ldots,k\} \\ j_1<j_2<\cdots<j_i}} P\left(\bigcap_{j\in\{j_1,j_2,\ldots,j_i\}} A_j^0\right).
\end{aligned}$$

The i^{th} term from this summation is a signed summation over probabilities of the form

$$P\left(\bigcap_{j\in\{j_1,j_2,\ldots,j_i\}} A_j^0\right) = P\left(A_{j_1}^0 \cap A_{j_2}^0 \cap \cdots \cap A_{j_i}^0\right).$$

There are $\binom{k}{i}$ such probabilities that satisfy $\{j_1, j_2, \ldots, j_i\} \subseteq \{1, 2, \ldots, k\}$ and $j_1 < j_2 < \cdots < j_i$ (i.e., i numbers chosen without regard to order from k numbers) and, by symmetry, each such probability is identical. So, we just need to figure out the probability of getting a particular sub-batch of i chip-less cookies, $\{j_1, j_2, \ldots, j_i\}$, in the batch of k cookies. For any given chip arriving in the dough, there is a probability $\frac{i}{k}$ that the chip lands in this sub-batch and a probability $\frac{k-i}{k}$ that the chip misses this sub-batch. The N chips arrive independently, so there is probability $\left(\frac{k-i}{k}\right)^N$ that this sub-batch of i cookies ends up chip-less. Plugging these results back into the above gives

$$\begin{aligned}
1-p &= \sum_{i=1}^{k} (-1)^{i+1} \sum_{\substack{\{j_1,j_2,\ldots,j_i\}\subseteq\{1,2,\ldots,k\} \\ j_1<j_2<\cdots<j_i}} P\left(\bigcap_{j\in\{j_1,j_2,\ldots,j_i\}} A_j^0\right) \\
&= \sum_{i=1}^{k} (-1)^{i+1} \binom{k}{i} \left(\frac{k-i}{k}\right)^N, \text{ for } N \ge k. \qquad \text{(D.1)}
\end{aligned}$$

Note that the final, k^{th}, term in the summation is identically zero because the cookies cannot all be chip-less. If we plug $p = 0.90$ and $k = 100$ into equation D.1 and solve for N, we need a computer to find $N = 682.52$. So, $N = 683$ chips will do the job. Do not worry about needing a computer to solve this; Your interviewer will either have stopped you before you got to this point or will have given you a computer.

I am now able to use equation D.1 to find that my gut instinct guess of $N = 500$ chips gives a probability of only $p = 51.2\%$ that I will get at least $m = 1$ chip in each of $k = 100$ cookies. At the other extreme, it would have taken $N = 916$ chips to obtain $p = 99\%$.

#2: Recursion.[16] An exact solution can also be found using a recursion approach. Let $B_{k,N}$ count the number of ways that we can put N chips into k cookies. There are k choices for the first chip, k choices for the second chip, etc. So, $B_{k,N} = k^N$. Let $C_{k,N}$ count the number of ways that we can put N chips into k cookies where every cookie gets at least one chip (call this outcome a "success"). Well, as mentioned above, $C_{k,N} = 0$ if $N < k$. We must also have $C_{1,N} = 1$ for all $N \geq 1$ (there is only one way to place N chips into one cookie). We can get a recursion going as follows: In order to get a success when placing the N^{th} chip, it must be either that the previous $N-1$ chips already distributed a chip to each cookie (success already!) and the N^{th} chip can go into any of the k cookies, or the previous $N-1$ chips distributed a chip to $k-1$ of the cookies and the N^{th} chip will go into the empty cookie; there are k ways that the latter event could take place. This produces the following recursion formula:

$$\begin{aligned} C_{k,N} &= k \cdot C_{k,N-1} + k \cdot C_{k-1,N-1} \\ &= k \cdot [C_{k,N-1} + C_{k-1,N-1}] \end{aligned}$$

If we let $P_{k,N}$ be the probability that we get a success (i.e., at least one chip per cookie) from N chips placed into k cookies, then

$$\begin{aligned} P_{k,N} &= \frac{C_{k,N}}{B_{k,N}} \\ &= \frac{k \cdot [C_{k,N-1} + C_{k-1,N-1}]}{k^N} \\ &= P_{k,N-1} + P_{k-1,N-1} \cdot \left(\frac{k-1}{k}\right)^{N-1}, \end{aligned} \tag{D.2}$$

with intitial conditions $P_{k,N} = 0$ if $N < k$ and $P_{1,1} = 1$ for all $N \geq 1$ (from above). Although the recursion does not automatically generate a closed form solution, we can compare equation D.2 with equation D.1, above, to deduce that

$$P_{k,N} = p = 1 - \sum_{i=1}^{k} (-1)^{i+1} \binom{k}{i} \left(\frac{k-i}{k}\right)^N, \text{ for } N \geq k. \tag{D.3}$$

A tedious proof by induction (omitted) easily shows that $P_{k,N}$ given in equation D.3 does indeed satisfy the recursion shown in equation D.2 and the initial condition $P_{1,1} = 1$ for all $N \geq 1$.

From an implementation standpoint, the recursion in equation D.2 is easier to execute accurately for large numbers than the closed-form solution in equation D.3—because the latter involves large factorials in the binomial coefficients.[17]

APPROXIMATE SOLUTIONS

I present two approximate solutions that can be figured using a handheld calculator. The first uses the union bound; the second relies upon near independence.

#1: Union Bound. The union bound relies upon Boole's Inequality, which uses just the first term from the inclusion-exclusion formula:

$$P\left(\bigcup_{i=1}^{k} A_i^0\right) \leq \sum_{i=1}^{k} P\left(A_i^0\right)$$

That is, the probability of at least one of the events occurring is bounded above by the sum of the probabilities of each of the events. In our case, we get

$$\begin{aligned} 1 - P_{k,N} &= P\left(\bigcup_{i=1}^{k} A_i^0\right) \\ &\leq \sum_{i=1}^{k} P\left(A_i^0\right) \\ &= (-1)^{1+1}\binom{k}{1}\left(\frac{k-1}{k}\right)^N, \text{ from the } i = 1 \text{ term in equation D.1} \\ &= k \cdot \left(\frac{k-1}{k}\right)^N. \end{aligned}$$

Thus, $P_{k,N} \geq 1 - k \cdot \left(\frac{k-1}{k}\right)^N$. In our case, we get $P_{k,N} \geq 1 - 100 \cdot \left(\frac{99}{100}\right)^N$. If we plug the known solution $N = 683$ into this we get $P_{k,N} \geq 89.6\%$, which is correct ($N = 683$ actually gives $P_{k,N} = 90.0\%$). If instead we set $P_{k,N} = 0.90$ and solve for N, we get $N \leq \frac{\log(0.001)}{\log(0.99)} = 687.3$, so, $N = 688$ is an upper bound on a sufficient number of chips.

#2: Assumed Independence. When the number of chips N is large relative to the number of cookies k, and when the minimum number of chips m required in each cookie is small, then the event that there are at least m chips in one cookie is very nearly independent of the event that there are at least m chips in another cookie. We can exploit this to get an excellent approximation to the exact solution.

Consider our case where $k = 100$ cookies and we want at least $m = 1$ chip per cookie. From the above, we know that the probability that there are no chips in the first cookie is

$$P(A_1^0) = \left(\frac{k-1}{k}\right)^N.$$

[16] I thank Torsten Schöneborn for suggesting this approach; any errors are mine.

[17] Note, however, that there are recursive approaches to calculating binomial coefficients that include large numbers: we can use $\binom{k}{i} = \frac{k}{i} \cdot \binom{k-1}{i-1}$. So, for example, $\binom{100}{6} = \frac{100}{6} \cdot \frac{99}{5} \cdot \frac{98}{4} \cdot \frac{97}{3} \cdot \frac{96}{2} \cdot \frac{95}{1}$.

It follows that the probability that there is at least one chip in the first cookie is $\left[1 - \left(\frac{k-1}{k}\right)^N\right]$. Each cookie has the same statistical properties, so, with near independence, the probability that there is at least one chip in every cookie is

$$p \approx \left[1 - \left(\frac{k-1}{k}\right)^N\right]^k.$$

Solving for N gives $N \approx \frac{\ln\left(1-p^{\frac{1}{k}}\right)}{\ln\left(1-\frac{1}{k}\right)} = \frac{\ln(1-0.9^{0.01})}{\ln(0.99)} = 682.17$. So, 683 chips will suffice, and we can figure this with a handheld calculator.

This answer relies upon the near independence between the event that there is at least one chip in one cookie and the event that there is at least one chip in another cookie.[18] For the case of at least $m = 1$ chip per cookie, I think my assumed-independence solution for N is either exactly correct, or underestimates N by only 1 for all k. I have also seen a solution that uses the Poisson distribution in the $m = 1$ case, assumes independence, and yields an answer that is algebraically very close to each of solutions presented here.

#3: Another Approximation. Finally, and this might be difficult to do in an interview, I plotted the assumed-independence solution for the $m = 1$ case, $N \approx \ln\left(1 - p^{\frac{1}{k}}\right) / \ln\left(1 - \frac{1}{k}\right)$, and the solution appeared to me to behave like $N = k \cdot \ln(b \cdot k)$ where b is a function of p only. Comparing the two functional forms revealed

$$b \approx \left[\ln\left(\frac{1}{p}\right)\right]^{-1} \cdot p^{\frac{1}{k}} \approx \left[\ln\left(\frac{1}{p}\right)\right]^{-1}.$$

This yielded the approximate solution

$$\begin{aligned} N &\approx k \cdot \ln\left\{\left[\ln\left(\frac{1}{p}\right)\right]^{-1} \cdot k\right\} \\ &= k \cdot \{\ln(k) - \ln[-\ln(p)]\}. \end{aligned}$$

This solution is almost as accurate as the assumed-independence solution. For example, when $k = 100$ cookies, and $p = 0.90$, the exact solution for the $m = 1$ case is 683, the assumed-independence solution is 683, and the above approximation yields 686 (an error of less than a half percent). Similarly, when $k = 500$ cookies, and $p = 0.90$, the exact solution for the $m = 1$ case is 4,229, the assumed-independence solution is 4,229, and the above approximation yields 4,233 (an error of less than a tenth of a percent).

[18] Let N_i be the number of chips in the i^{th} cookie. Then, for $N = 683$ and $k = 100$ the events $\mathcal{A} = (N_i \geq m)$ and $\mathcal{B} = (N_j \geq m)$ for $i \neq j$ are effectively numerically indistinguishable from independent for small m. I simulated $\frac{P(\mathcal{A})\cdot P(\mathcal{B})}{P(\mathcal{A}\cap\mathcal{B})}$ for various m and found the ratio was 1 ± 0.000001 for $m = 1$, 1 ± 0.0001 for $m = 2$, 1 ± 0.0002 for $m = 3$, 1 ± 0.0003 for $m = 4$, 1 ± 0.001 for $m = 5$, 1 ± 0.02 for $m =$ 6–10, 1 ± 0.10 for $m =$ 11–16, and rare enough that it was difficult to simulate for $m \geq 17$.

Answer 4.29: I present two solutions: the first is a full "hammer-and-tongs" solution; the second uses a recursive argument similar to Answer 4.24. Questions requiring recursive proofs have become popular in interviews; look up "recursive argument" in the index to find other examples in this book.

FIRST SOLUTION

When the game stops (i.e., you rolled a 4, 5, or 6), you have a $\frac{2}{3}$ chance of seeing the accumulated score (i.e., you got a 4 or 5 out of a possible 4, 5, or 6 to stop the game). The accumulated score is 0 with probability $\frac{1}{2}$, 1 with probability $\left(\frac{1}{2}\right)^2$, 2 with probability $\left(\frac{1}{2}\right)^3$, and so on. So the expected payoff is

$$\begin{aligned} \frac{2}{3} \cdot \sum_{i=0}^{\infty} j \cdot \left(\frac{1}{2}\right)^{j+1} &= \frac{1}{3} \cdot \sum_{i=0}^{\infty} \frac{j}{2^j} \\ &= \frac{1}{3} \cdot 2 = \frac{2}{3}, \end{aligned}$$

using the result that $\sum_{i=0}^{\infty} \frac{j}{2^j} = 2$ (see Footnote 4 on page 59).

SECOND SOLUTION[19]

Let N be the random number of times you roll a 1, 2, or 3, before a 4, 5, or 6 appears. If you now roll the die once, there is a half a chance that $N = 0$ and the game stops, and there is a half a chance that you get a 1, 2, or 3, and you are otherwise recursively back where you started (expecting to roll an additional $E(N)$ 1's, 2's, or 3's before the game ends). If we write this algebraically, we get

$$E(N) = \left(0 \cdot \frac{1}{2}\right) + \left[(1 + E(N)) \cdot \frac{1}{2}\right].$$

We can solve this directly to get $E(N) = 1$. When the game does end, we get paid only if we roll a 4 or 5. In other words, conditional upon the game ending, we have a two thirds chance of being paid N. Our expected payoff is thus $\frac{2}{3} \cdot E(N) = \frac{2}{3}$, as before.

Answer 4.30: I first heard about the "broken stick problem" appearing in MBA interviews back in the early 1990's. This old question deserves justice here. The most elegant proof I have seen for obtaining the probability that a triangle can be formed from three bits is a special case of a general proof in higher dimensions.

Bull (1948) considers the case where a stick is broken randomly into n pieces. He says that a necessary and sufficient condition for a polygon of n or fewer angles to be made out of the n pieces is that no one piece can be of a length

[19] I thank Simon West for suggesting this technique; any errors are mine.

that exceeds the sum of the lengths of the others. Equivalently, each piece must be not greater in length than half the length of the stick.[20]

Following Bull... ...let x, y, and $1-x-y$ denote the lengths of the three separate pieces. If x and y are taken as axes of coordinates in two dimensions, then all ways of breaking the stick (regardless of whether a triangle can be formed or not) are represented by points inside and on the triangle formed by the coordinate axes and the line $x+y=1$ (i.e., the boundary and interior of the triangle O–A_1–A_2 in figure D.2). If this statement is not immediately obvious, then think of it as follows. There are three pieces of lengths x, y, and $1-x-y$. It must be the case that each piece has non-negative length: $x \geq 0$, $y \geq 0$, and $1-x-y \geq 0$. After rearranging the latter inequality, these three conditions say $x \geq 0$, $y \geq 0$, and $x+y \leq 1$. The intersection of these three conditions is the boundary and interior of the triangle O–A_1–A_2 in figure D.2, as stated.

The points in figure D.2 representing ways of breaking the stick that allow a triangle to be formed are those points inside and on the area bounded by the lines $x=\frac{1}{2}$, $y=\frac{1}{2}$, and $x+y=\frac{1}{2}$ (i.e., the shaded area in figure D.2). Again, if this is not immediately obvious, think of the conditions required: each piece must not be greater in length than half the length of the stick: $x \leq \frac{1}{2}$, $y \leq \frac{1}{2}$, and $1-x-y \leq \frac{1}{2}$ (or, $x+y \geq \frac{1}{2}$). The intersection of these three conditions is the boundary and interior of the shaded triangle B_{12}–A_1'–A_2' in figure D.2, as stated.

Random breaking of the stick means that the possible triplets of pieces $(x, y, 1-x-y)$ obtained are represented by coordinate pairs (x,y) distributed uniformly over the triangle O–A_1–A_2 in figure D.2. The shaded triangle B_{12}–A_1'–A_2' corresponds to pieces that can be used to build a triangle. The probability you can form a triangle is thus

$$P(3) = \frac{\text{AREA(triangle } B_{12}\text{–}A_1'\text{–}A_2')}{\text{AREA(triangle } O\text{–}A_1\text{–}A_2)} = \frac{\frac{1}{2} - 3\cdot\left(\frac{1}{4\cdot 2}\right)}{\frac{1}{2}} = 1-(3\cdot 1/4) = \frac{1}{4}.$$

Visual inspection of figure D.2 yields the same solution. We have now solved the question posed.

Bull (1948) presents a geometric proof for the case of four pieces. His proof is the three-dimensional analogy of my figure D.2. He has three axes, x, y, z, and a tetrahedron representing all ways of forming a polygon of four or fewer angles from the pieces. At the origin, and along each axis are tetrahedrons that must be excluded if a polygon is to be formed. So, instead of my two axes and three triangles to exclude, Bull has three axes and four tetrahedrons to exclude. I won't draw the picture, but here are the calculations: The volume of a tetrahedron is $\frac{1}{3} \cdot A \cdot h$, where A is base area and h is height. The large

[20] Bull implicitly allows for two degenerate cases: triangles with zero area when one piece has length $\frac{1}{2}$; and, one or two pieces having zero length. For randomly placed breaks, these cases happen with probability zero; my allowing them does not change the final answer.

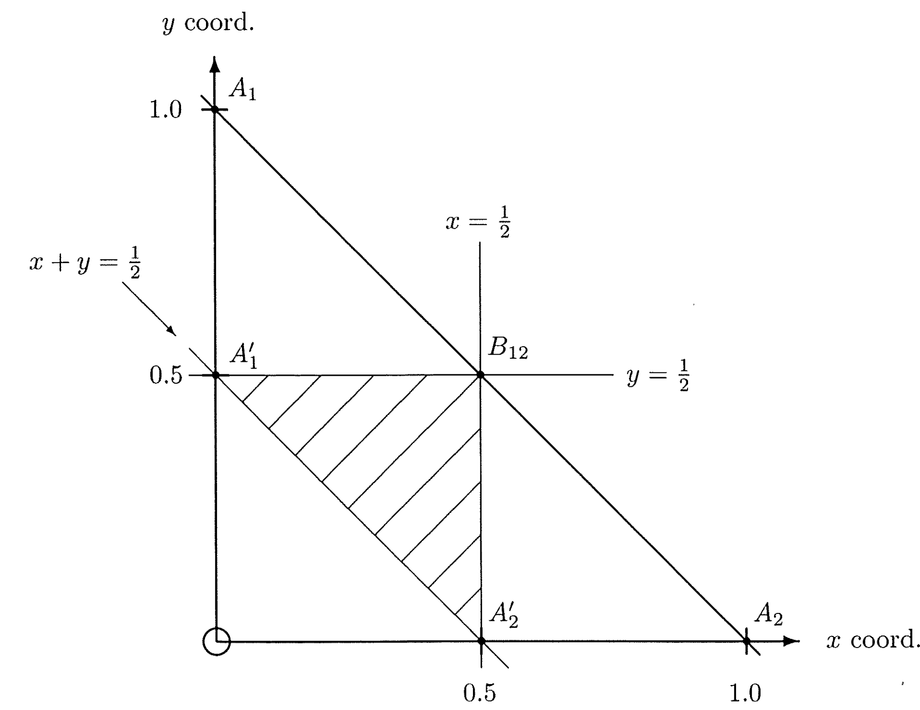

Figure D.2: Broken Stick Problem: Form a Triangle

Note: Any particular x and y coordinate pair within the triangle O–A_1–A_2 corresponds to a broken stick with three pieces of lengths x, y, and $1-x-y$, respectively. If the x and y coordinate pair fall within the shaded triangle B_{12}–A_1'–A_2' then the three pieces can be used to construct a triangle—for in this case no one piece can be of a length that exceeds the sum of the lengths of the others, or, equivalently, each piece must be not greater in length than half the length of the stick.

tetrahedron has $A = \frac{1}{2}$ and $h = 1$; the four small ones have $A = \frac{1}{8}$ and $h = \frac{1}{2}$. So, the answer for $n = 4$ is $P(4) = \frac{(\frac{1}{3}\cdot\frac{1}{2}\cdot 1) - 4\cdot(\frac{1}{3}\cdot\frac{1}{8}\cdot\frac{1}{2})}{(\frac{1}{3}\cdot\frac{1}{2}\cdot 1)} = 1 - (4 \cdot 6/48) = \frac{1}{2}$.

In the general n-piece proof, Bull has an $(n-1)$-dimensional polytope with n polytopes to be excluded.[21] The same relative probability argument yields $P(n) = \frac{\frac{1}{(n-1)!} - n\cdot\left(\frac{1}{2^{n-1}(n-1)!}\right)}{\frac{1}{(n-1)!}} = 1 - \frac{n}{2^{n-1}}$. As n gets large $P(n)$ goes to 1. This seems intuitive; If I break a stick randomly into 100 pieces, how likely is it that I can lay the pieces out to form a polygon of 100 or fewer angles? ...the answer should be very close to 1.

Answer 4.31: Unlike Answer 4.30, a triangle need not be formed here. To derive the pdf for the longest piece we will derive the cdf first and then differentiate it. We shall work with the properties of figure D.2.

Let L denote the random longest piece, and l denote a particular value of L. We want to find the cdf $F_L(l) = P(L \leq l)$. We know that $L \equiv \max(x, y, 1 - x - y)$, so the event $L \leq l$ happens if and only if $x \leq l$, $y \leq l$ and $1 - x - y \leq l$. The latter may be rearranged as $x + y \geq 1 - l$. Following the arguments in Answer 1.30, we need only find the relative area contained within the region bounded by the lines $x = l$, $y = l$ and $x + y = 1 - l$ (but also within the original triangle O–A_1–A_2) in figure D.2. There are two cases: $0.5 \leq l \leq 1$ (see figure D.3), and $\frac{1}{3} \leq l \leq 0.5$ (see figure D.4). The figure captions contain the algebra. We find that the cdf of the longest piece is

$$F_L(l) = \begin{cases} [1 - 3(1-l)^2], & 0.5 \leq l \leq 1 \\ (3l-1)^2, & \frac{1}{3} \leq l \leq 0.5, \end{cases}$$

and the pdf of the longest piece is[22]

$$f_L(l) = \begin{cases} 6(1-l), & 0.5 \leq l \leq 1 \\ 6(3l-1), & \frac{1}{3} \leq l \leq 0.5. \end{cases}$$

So, the expected length of the longest piece of broken stick is

$$\begin{aligned} E(L) &= \int_{\frac{1}{3}}^{1} l \cdot f_L(l)dl \\ &= \int_{\frac{1}{3}}^{0.5} l \cdot 6(3l-1)dl + \int_{0.5}^{1} l \cdot 6(1-l)dl \\ &= \left(6l^3 - 3l^2\right)\Big|_{\frac{1}{3}}^{0.5} + \left(3l^2 - 2l^3\right)\Big|_{0.5}^{1} \\ &= \left[\left(\frac{6}{8} - \frac{3}{4}\right) - \left(\frac{6}{27} - \frac{3}{9}\right)\right] + \left[(3-2) - \left(\frac{3}{4} - \frac{2}{8}\right)\right] = \frac{11}{18}. \end{aligned}$$

[21] A polytope is "a finite region of n-dimensional space enclosed by a finite number of hyperplanes (`mathworld.wolfram.com`)."

[22] Can you sketch the pdf to confirm that it is right-skewed and triangular?

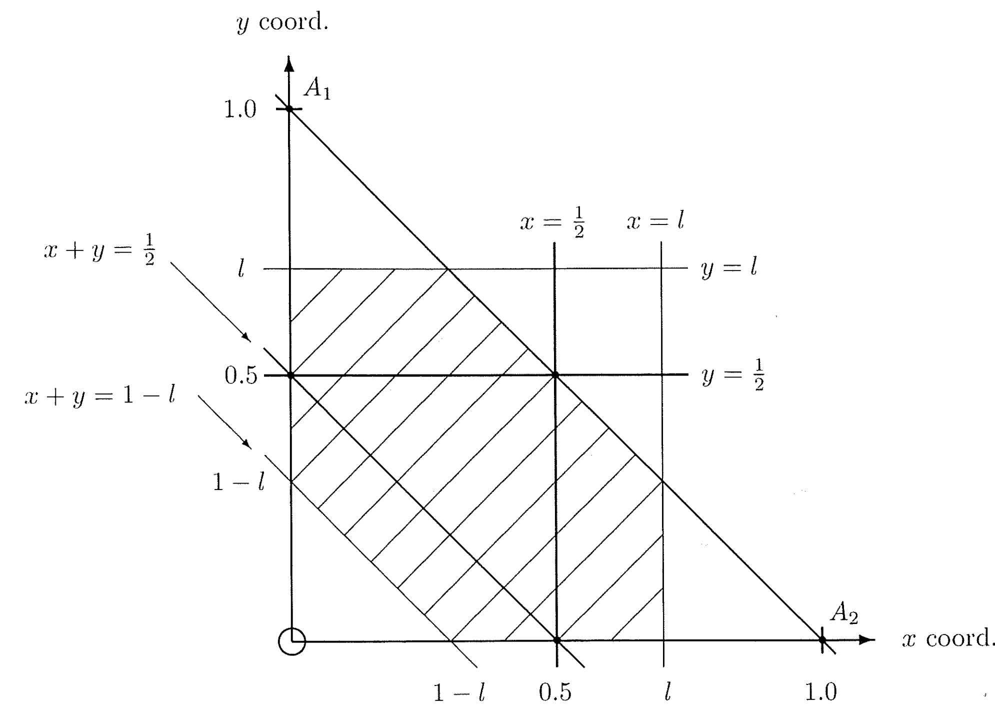

Figure D.3: Broken Stick Problem, $F_L(l) = P(L \leq l)$, Case: $0.5 \leq l \leq 1$

Note: The cdf of the length of the longest piece is given by the shaded relative probability area bounded by the three conditions $x \leq l$, $y \leq l$ and $x + y \geq 1 - l$ (but also within the original triangle O–A_1–A_2). In the case $l \geq 0.5$ this is simply the relative area obtained by excluding the three equal-sized un-shaded triangles. Each of the three un-shaded triangles has area $\frac{1}{2}(1-l)^2$, so the cdf value is simply $F_L(l) = P(L \leq l) = \frac{\frac{1}{2} - 3\cdot\left[\frac{1}{2}(1-l)^2\right]}{\frac{1}{2}} = [1 - 3(1-l)^2]$. This yields pdf $f_L(l) = 6(1-l)$ in the case $0.5 \leq l \leq 1$.

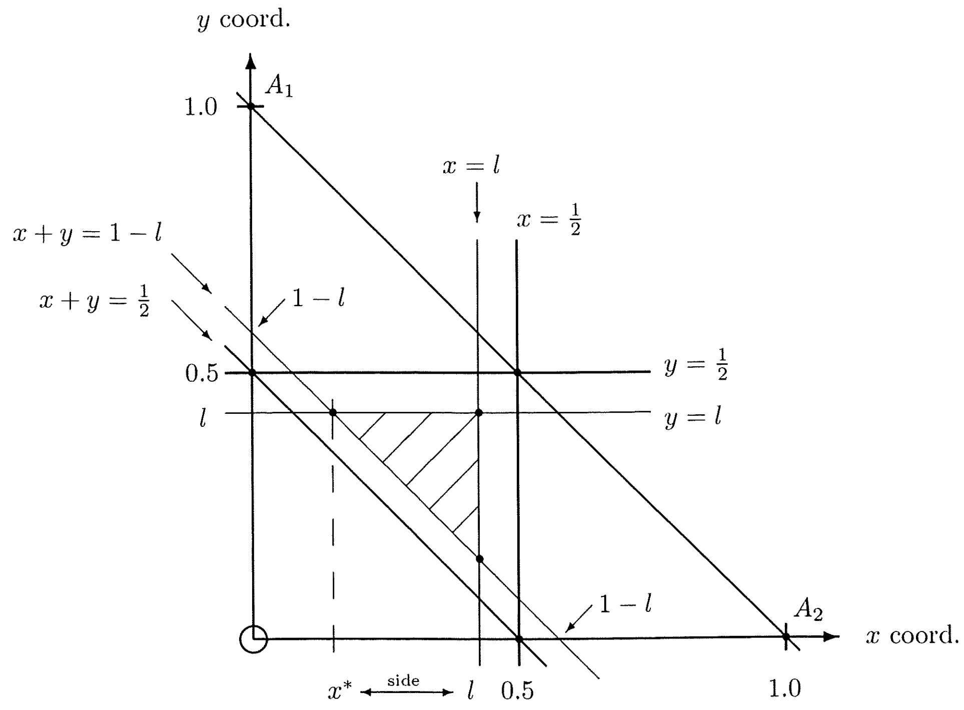

Figure D.4: Broken Stick Problem, $F_L(l) = P(L \leq l)$, Case: $\frac{1}{3} \leq l \leq 0.5$

Note: As in figure D.3, the cdf of the length of the longest piece is given by the shaded relative probability area bounded by the three conditions $x \leq l$, $y \leq l$ and $x+y \geq 1-l$ (but also within the original triangle O–A_1–A_2). In the case $l \leq 0.5$ this is simply the relative area of the shaded triangle. Close inspection shows that as l drops, the area of the triangle reduces until $l = \frac{1}{3}$ at which point the area is zero. This corresponds to the simple intuition that the longest piece of stick cannot be smaller than $\frac{1}{3}$ in length (else it is not the longest). One way to get the area of the shaded triangle is to deduce the value of x^* so that we can deduce the side length $l - x^*$. To find x^* note that it is the value of x at which the lines $x + y = 1 - l$ and $y = l$ intersect. If we plug the latter into the former, we can solve for x^*: $x^* + l = 1 - l$ implies $x^* = 1 - 2l$. The side length is thus $l - x^* = l - (1 - 2l) = 3l - 1$. The shaded triangle thus has area $\frac{1}{2}(3l-1)^2$. So, the cdf value is simply $F_L(l) = P(L \leq l) = \frac{\frac{1}{2}(3l-1)^2}{\frac{1}{2}} = (3l-1)^2$. This yields pdf $f_L(l) = 6(3l - 1)$ in the case $\frac{1}{3} \leq l \leq 0.5$.

I leave you to use the same technique to confirm that the expected length S of the shortest piece[23] is $E(S) = \frac{2}{18} = \frac{1}{9}$.

Answer 4.32: You know only that I have two children and that one is a girl. My family has (implicitly) conducted two Bernoulli trials. Without the information that one child is a girl, there are four possible equally likely outcomes: GG, GB, BG, BB. With the information that one child is a girl, the last outcome is excluded, and the sample space describing the randomness you are confronted with is three equally likely outcomes: GG, GB, BG. The probability of GG is thus $\frac{1}{3}$. See Answer 4.33 for more details.

Answer 4.33: You know only that I have two children and that one is a girl you are facing at my front door. Now the argument changes from the previous question. The only randomness is the gender of the child you cannot see. The sample space describing this randomness is just G and B. The probability that I have two girls is thus $\frac{1}{2}$.

Like in Answer 4.32, without seeing my child, there are four possible equally likely outcomes: GG, GB, BG, BB. Having pinpointed the gender of Child #1 (we might as well call her that), however, the latter two outcomes are excluded, an the sample space describing the randomness you are confronted with is two equally likely outcomes: GG, GB. The probability of GG is thus $\frac{1}{2}$.

If they press you for the formal derivation so you can point your finger at the difference between the Answers 4.32 and 4.33, you can use Bayes' Theorem. In Answer 4.32 you are seeking $P(G \cap G|\text{told } G)$:

$$\begin{aligned}
P(G \cap G|\text{told } G) &= \frac{P(G \cap G \cap \text{told } G)}{P(\text{told } G)} = \frac{P(\text{told } G \cap (G \cap G))}{P(\text{told } G)} \\
&= \frac{P(\text{told } G|G \cap G) \cdot P(G \cap G)}{P(\text{told } G|G \cap G) \cdot P(G \cap G) + \mathbf{P(\text{told } G|B \cap G)} \cdot P(B \cap G) + P(\text{told } G|B \cap B) \cdot P(B \cap B)} \\
&= \frac{1 \cdot \frac{1}{4}}{1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{4}} = \frac{1}{3}
\end{aligned}$$

In Answer 4.33 you are seeking $P(G \cap G|\text{see } G)$:

$$\begin{aligned}
P(G \cap G|\text{see } G) &= \frac{P(G \cap G \cap \text{see } G)}{P(\text{see } G)} = \frac{P(\text{see } G \cap (G \cap G))}{P(\text{see } G)} \\
&= \frac{P(\text{see } G|G \cap G) \cdot P(G \cap G)}{P(\text{see } G|G \cap G) \cdot P(G \cap G) + \mathbf{P(\text{see } G|B \cap G)} \cdot P(B \cap G) + P(\text{see } G|B \cap B) \cdot P(B \cap B)} \\
&= \frac{1 \cdot \frac{1}{4}}{1 \cdot \frac{1}{4} + \mathbf{\frac{1}{2}} \cdot \frac{1}{2} + 0 \cdot \frac{1}{4}} = \frac{1}{2}
\end{aligned}$$

I used bold font to show where the two derivations differ: $P(\text{told } G|B \cap G) = 1$, but $P(\text{see } G|B \cap G) = \frac{1}{2}$. Otherwise everything is the same.

[23]There is only one case. I get $F_S(s) = [1 - (1 - 3s)^2]$, and $f_S(s) = 6(1 - 3s)$, for $0 \leq s \leq \frac{1}{3}$. I deduce that $E(M) = \frac{5}{18}$ for the middle-length piece because $E(S + M + L) = 1$.

Support the author!
If you like this book, buy a copy!

Appendix E

Non-Quantitative Answers (Selected)

This appendix contains answers to selected questions in chapter 5.

Answer 5.2.10: A "tombstone" is of course an advertisement that lists (like the names on a tombstone) the underwriters associated with a public issue of a security. The particular placement of the underwriters' names on the tombstone carries with it implications for the perceived status of the underwriters on the deal.

A student came to see me. He told me that he was flying to Chicago the next day for a job interview with an investment bank. I did not recognize the name of the bank. He asked me what sort of non-quantitative questions he might face, so I pulled out my book and tried several on him. When I got to the tombstone question, I stopped and asked him if he knew the definition of a tombstone. I pulled out that day's *Wall Street Journal* (WSJ) to see if there was a tombstone in the third section. The page at which I opened the WSJ contained a tombstone from the bank he was going to interview with the next day! I clipped it out and gave it to him, and he talked about it in his interview. It is worth keeping your eye on the tombstones in the third section of the WSJ in the weeks leading up to your interviews.

Answer 5.2.24: This is not necessarily a rejection. It may just be a test to see how you defend yourself. It is just the opposite way of asking you why you fit the job. I would answer "No, Sir. I fit because ..."

Answer 5.3.6: This is basic macroeconomics, and you should be fully familiar with it. The two forms of macroeconomic policy are monetary policy and fiscal policy. Monetary policy tries to achieve the broad objectives of economic policy through control of the monetary system and by operating on the supply of money, the level and structure of interest rates, and other conditions affecting the supply of credit (Pearce [1984, p291]). With monetary policy, the Federal Reserve Bank ("the Fed") sell bonds and reduces the money supply—an "open market operation." This increases interest rates (the cost of money) and makes capital expenditures more costly. This in turn slows down growth in the economy and should fight the inflationary threat.

In addition to open market operations, the Fed implements monetary policy by managing the discount rate (the rate the Fed charges banks for loans), adjusting the Fed funds rate (the rate banks charge each other for loans of federal funds), managing reserve requirements for banks (the proportion of a bank's assets required to be held in Treasury securities), and operations in the government repo (i.e., repurchase) market.[1]

Fiscal policy refers to the use of taxation and government expenditure to regulate the aggregate level of economic activity (Pearce [1984, p160]). Increasing taxes and decreasing government spending should slow down growth in the economy and fight inflationary fears. Go to any standard macroeconomic text if you want more details on fiscal or monetary policy.

Answer 5.3.12: The "Dow Jones Dogs strategy" involves buying the "Dow Jones Dogs" at the start of the year. These are the 10 Dow Jones Industrial Average (DJIA) stocks with the highest dividend yield. They are dogs because you get a relatively high dividend yield by having a low price relative to dividends. You are supposed to rebalance the portfolio every year. Historically this has been a profitable strategy. The CBOE introduced options on a Dow Jones Dogs index (ticker symbol "MUT") in 1999, but I do not see any volume in them in 2009. Deutsche Bank offers an exchange-traded note (ETN) based on the return to the Dow Jones Dogs strategy (ticker symbol "DOD"). MUT and DOD move very closely together.

Answer 5.4.8: The answer given by the interviewer was that if you are Avis or Hertz, cars are inventory. The same applies to an automobile manufacturer (or any of their distributors).

Answer 5.4.16: No. FCF does not include interest payments or repayment of principal because FCF is the cash flows generated by net assets and available to the owners of the company (both debt and equity holders). The tax benefit of interest payments is recognized in a lower after tax cost of debt in the WACC. Finally, financing costs are not cash outflows. They do not reduce cash available to owners. To the contrary, they *are* cash payments to owners and, therefore, have no net effect on cash flows available to owners (i.e., FCF).

Answer 5.5.2: The "how many somethings are there somewhere" questions are common. There is no precise algebraic solution routine. You make several rough assumptions and hope the errors cancel. For example, the US population is about 300,000,000 (mid 2009). The population of Bloomington, Indiana, is about 100,000 (when the students are there). There are about six McDonald's in Bloomington. I calculate $\frac{300,000,000}{100,000} \times 6 = 3,000 \times 6 = 18,000$. However, Bloomington is a college town, and students eat more junk food than the general populace, so I adjust my answer downwards to about 10,000 to 15,000 McDonald's outlets in the US.

[1] A "repo" is a repurchase agreement. It is an agreement to repurchase a security in the future. You give up the security now in exchange for cash, agreeing to repurchase the security at a later date for a larger amount of cash. A repo is thus a collateralized loan. A reverse repo is the other side of the deal—you purchase securities now with an agreement to sell them later. Repos range in maturity from overnight ("O/N") to as long as five years; shorter-term repos are the most popular.

Looking at online profiles of the company, I estimate that of 32,000 McDonalds restaurants (mid 2009), roughly 14,000 of them are located in the US.

In general, you grab something you know, scale it up or down, and adjust for any biases. Let us try it again a different way. I cannot believe anyone over 30 or under five would eat in a McDonald's. If lifespan is uniformly distributed between zero and 75 years, then only one-third of the population (100,000,000) is eligible for eating at McDonald's. Half of these are health nuts. That leaves 50,000,000 customers. Suppose they eat four meals per week. That works out to about 30,000,000 meals served per day. If one outlet sells a burger every 30 seconds, that is 120 an hour and about 3,000 per day. 30,000,000 meals served per day at 3,000 per outlet implies about 10,000 outlets. This is still in the ballpark.

Whether it is ping-pong balls in a 747, barbers in Chicago, or elevators in the US, find something you know and scale it up or down. Be sure to know the population of the Earth, the US, the city you live in, and the city you interview in.

Answer 5.5.4: The answer given by the interviewer was that you should threaten to kill yourself by hitting your head against the wall. The administrative nightmare that would follow would ruin the guard's upcoming weekend. He would have to give you a cigarette.

Answer 5.5.7: You have to figure that the coin is not fair. The probability of another head is essentially one. See Huff's book, "How to Lie with Statistics," for related arguments (Huff [1982, chapter 3]).

Answer 5.5.10: I have had several comments from readers on how to weigh a jet plane: Land it on an aircraft carrier and measure the displacement of the ship; land it on the ice (e.g., Arctic) and then crash into it with something big and see how far it moves; look at the size of the tire footprints and deduce it from the tire pressure. How about just looking in the manual?

Answer 5.5.16: I spent some time at NASA's web site. NASA says there are four forces on an aeroplane: thrust, drag, weight and lift. These forces move the aeroplane forward, backward, downward, and upward, respectively, and simultaneously. For example, level forward flight would require that lift balance weight, and that thrust more than compensate for drag.

NASA indicates that there is considerable debate/confusion over how lift is generated. They describe lift as a mechanical force that is generated when a solid object moves through a fluid and "turns" the fluid flow. The shape and "angle of attack" of an aeroplane's wing turn the fluid flow downwards, and, by Newton's Third Law of Motion (i.e., for every action there is an equal and opposite reaction) this provides upward lift.

NASA comments that wings are often shaped so that the wing, taken in cross section, has more surface area on the top surface than the bottom surface. The air flows more quickly over the top of the wing than the bottom. The variation in velocity of the fluid creates a pressure differential that produces lift. They state, however, that this wing shape is not necessary to create lift. Rather, it only contributes to it.

Answer 5.5.20: There are several possible responses that make sense. An obvious reason is safety: a round cover cannot fall down a round hole. Whereas if both hole and cover are either square or rectangular or oval, the cover can easily fall down the hole if lifted vertically and turned diagonally and dropped. Incidentally, I noticed while working in New Zealand that some of their manholes have rectangular covers. However, in this case, the covers are hinged and attached to a frame that is immovable—thus preventing the cover from falling.

Another reason for being round is that the (very heavy) covers may be rolled easily. Similarly, a (very heavy) round cover need not be manipulated before being returned to its hole—it may be replaced in any orientation. Finally, and with some sarcasm, manhole covers are round because the holes that they cover are round. It is easier to drill a round hole in the street than a square one. Have you ever tried drilling a *square* hole in anything?

References

Akgiray, Vedat, 1989, "Conditional Heteroscedasticity in the Time Series of Stock Returns: Evidence and Forecasts," *The Journal of Business*, Vol 62 No 1, (January), pp55–80.

Arnold, Tom, and Timothy Falcon Crack, 2004, "Using the WACC to Value Real Options," *Financial Analysts Journal*, Vol 60 No 6, (November/December), pp78–82.

Arnold, Tom, Timothy Falcon Crack, and Adam Schwartz, 2009, "Inferring Risk-Averse Probability Distributions from Options Prices Using Implied Binomial Trees: Additional Theory, Empirics, and Extensions." Working paper. Available at SSRN: http://ssrn.com/abstract=749904.

Arnold, Tom, Timothy Falcon Crack, and Adam Schwartz, 2010 "Inferring Risk-Averse Probability Distributions from Options Prices Using Implied Binomial Trees." A chapter in: G.N. Gregoriou and R. Pascalau (Eds.), Financial Econometrics Handbook. Chapman-Hall-CRC/Taylor and Francis: London, UK.

Anton, Howard, 1988, *Calculus with Analytical Geometry*, Third Edition, John Wiley and Sons: New York, NY.

Bachelier, Louis, 1900, "Théorie de la Spéculation," *Annales de l'Ecole Normale Supéieure*, Series 3, XVII, pp21–86, Gauthier-Villars: Paris. Note that an English translation by A. James Boness appears in Cootner (1964).

Ball, Clifford A. and Walter N. Torous, 1985, "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing," *The Journal of Finance*, Vol 40 No 1, (March), pp155–173.

Barone-Adesi, G. and Robert Whaley, 1987, "Efficient Analytic Approximation of American Option Values," *The Journal of Finance*, Vol 42 No 2, (June), pp301–320.

Bates, David S., 1995, "Testing Option Pricing Models," Working Paper, The Wharton School, University of Pennsylvania.

Baxter, Martin and Andrew Rennie, 1998, *Financial Calculus*, Cambridge University Press: Cambridge, England.

Bera, Anil, Edward Bubnys, and Hun Park, 1988, "Conditional Heteroscedasticity in the Market Model and Efficient Estimates of Betas," *The Financial Review*, Vol 23 No 2, (May), pp201–214.

Bera, Anil and Matthew L. Higgins, 1993, "ARCH models: Properties, Estimation, and Testing," *Journal of Economic Surveys*, Vol 7 No 4, pp305–366.

Berndt, Ernst K., Bronwyn Hall, Robert Hall, and Jerry A. Hausman, 1974, "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, Vol 3 No 4, (October), pp653–665.

Bernoulli, Daniel, 1738, "Speciman Theoriae Novae de Mensura Sortis," *Papers of the Imperial Academy of Sciences in Petersburg*, Vol V, pp175–192. Note that an English Translation appears in Bernoulli (1954).

Bernoulli, Daniel, 1954, "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, Vol 22 No 1, (January), pp23–36 (Translated from the Latin by Louise Sommer).

Biger, Nahum and John Hull, 1983, "The Valuation of Currency Options," *Financial Management*, Vol 12 No 1, (Spring), pp24–28.

Black, Fischer and Myron Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol 81 No 3, (May/June), pp637–659.

Black, Fischer, 1975, "Fact and Fantasy in the Use of Options," *The Financial Analysts Journal*, Vol 31 No 4, (July/August), pp36–41, 61–72.

Black, Fischer, 1976, "Studies of Stock Price Volatility Changes," *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section of the American Statistical Association*, pp177–181.

Black, Fischer, 1989, "How We Came Up With the Option Formula," *The Journal of Portfolio Management*, Vol 15 No 2, (Winter), pp4–8.

Black, Fischer, 1993, "Beta and Return," *The Journal of Portfolio Management*, Vol 20 No 1, (Fall), pp8–18.

Black, Fischer and Myron S. Scholes, 1972, "The Valuation of Option Contracts and a Test of Market Efficiency," *The Journal of Finance*, Vol 27 No 2, (May), pp399–417.

Black, Fischer and Myron Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol 81 No 3, (May/June), pp637–659.

Bollerslev, Tim, 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, Vol 31 No 3, (April), pp307–327.

Bollerslev, Tim, 1987, "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, Vol 69 No 3, (August), pp542–547.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, 1992, "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, Vol 52 No 1/2, (April/ May), pp5–59.

Boyce, William E. and Richard C. DiPrima, 1997, *Elementary Differential Equations and Boundary Value Problems*, Sixth Edition, John Wiley and Sons: New York, NY.

Boyle, Phelim P., 1977, "Options: A Monte Carlo Approach," *The Journal of Financial Economics*, Vol 4 No 3, (May), pp323–338.

Brealey, Richard A. and Stewart C. Myers, 1991, *Principles of Corporate Finance*, Fourth Edition, McGraw-Hill: New York, NY.

Brenner, Menachem and Marti G. Subrahmanyam, 1988, "A Simple Formula to Compute the Implied Standard Deviation," *The Financial Analysts Journal*, Vol 44 No 5, (September/October), pp80–83.

Brock, William A., David A. Hsieh, and Blake LeBaron, 1991, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, MIT Press: Cambridge, MA.

Brown, Robert, 1828, "A Brief Account of Microscopical Observations Made in the Months of June, July, and August, 1827, on the Particles Contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies," *The London and Edinburgh Philosophical Magazine and Annals of Philosophy*, Vol 4 No 21, pp161–173.

Bull, G.A., 1948, "A Broken Stick," *The Mathematical Gazette*, Vol 32 No 299, (May), pp87–88.

Chance, Don M., 1994, "Translating the Greek: The Real Meaning of Call Option Derivatives," *The Financial Analysts Journal*, Vol 50 No 4, (July/August), pp43–49.

Chesney, Marc and Louis Scott, 1989, "Pricing European Currency Options: A Comparison of the Modified Black-Scholes Model and a Random Variance Model," *The Journal of Financial and Quantitative Analysis*, Vol 24 No 3, (September), pp267–284.

Chordia, Tarun and Avanidhar Subrahmanyam, 1995, "Market Making, the Tick Size, and Payment-for-Order-Flow: Theory and Evidence," *Journal of Business*, Vol 68 No 4, (October), pp543–575.

Christensen, Peter Ove, and Bjarne Sørensen. 1994. "Duration, Convexity, and Time Value." *The Journal of Portfolio Management*, Vol 20 No 2, (Winter), pp51–60.

Christensen, Ronald and Jessica Utts, 1992, "Bayesian Resolution of the 'Exchange Paradox'," *The American Statistician*, Vol 46 No 4, (November), pp274–276.

Conze, Antoine and Viswanathan, 1991, "Path Dependent Options: The Case of Lookback Options," *The Journal of Finance*, Vol 46 No 5, (December), pp1893–1907.

Cootner, Paul H., (Ed.), 1964, *The Random Character of Stock Market Prices*, MIT Press: Cambridge, MA.

Cox, J. C., J. Ingersoll, and S. Ross, 1979, "Duration and the Measurement of Basis Risk," *The Journal of Business*, Vol 52 No 1, (January), pp51–61.

Cox, J. C. and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *The Journal of Financial Economics*, Vol 3 No 1/2, (January/March), pp145–166.

Cox, J. C., S. Ross, and M. Rubinstein, 1979, "Option Pricing: A Simplified Approach," *The Journal of Financial Economics*, Vol 7 No 3, (September), pp229–263.

Cox, J. C. and Mark Rubinstein, 1985, *Options Markets*, Prentice-Hall: Englewood Cliffs, NJ.

Crack, Timothy Falcon, 1997, *Derivatives Securities Pricing*, MBA Course Notes, Indiana University, Kelley School of Business, (Spring #II), 108pp.

Crack, Timothy Falcon, 1998, "Barrier Options: Density of Return and First Passage Time for a GBM," Notes, Indiana University, Kelley School of Business, (April 1), 20pp.

Crack, Timothy Falcon, 2004, *Basic Black-Scholes: Option Pricing and Trading*. See `www.BasicBlackScholes.com`, and the advertisement on the last page of this book, for details.

Crack, Timothy Falcon, 2009, *Basic Black-Scholes: Option Pricing and Trading*. Revised Second Edition. See `www.BasicBlackScholes.com`, and the advertisement on the last page of this book, for details.

Crack, Timothy Falcon and Olivier Ledoit, 1996, "Robust Structure Without Predictability: The 'Compass Rose' Pattern of the Stock Market," *The Journal of Finance*, Vol 51 No 2, (June), pp751–762.

Crack, Timothy Falcon and Olivier Ledoit, 2008, "Why Do Doctoral Students Choose the Wrong Central Limit Theorem When Data are Dependent?" Working Paper. Available at SSRN: http://ssrn.com/abstract=587562

Crack, Timothy Falcon and Sanjay K. Nawalkha, 2000, "Interest Rate Sensitivities of Bond Risk Measures," *The Financial Analysts Journal*, Vol. 56 No. 1, (January/February), pp34–43.

Crack, Timothy Falcon and Sanjay K. Nawalkha, 2001, "Common Misunderstandings Concerning Duration and Convexity," *Journal of Applied Finance*, Vol. 1, (October), pp82–92.

DeGroot, Morris H., 1989, *Probability and Statistics*, Addison-Wesley: Reading, MA.

Derman, Emanuel, 2004, *My Life as a Quant: Reflections on Physics and Finance*, Wiley: Hoboken, NJ.

Derman, Emanuel and Iraj Kani, 1993, "The Ins and Outs of Barrier Options," *Goldman Sachs Quantitative Strategies Research Notes*, Goldman, Sachs, (June).

Derman, Emanuel and Iraj Kani, 1994, "The Volatility Smile and its Implied Tree," *Goldman Sachs Quantitative Strategies Research Notes*, Goldman, Sachs, January.

Dixit, Avinash K. and Barry J. Nalebuff, 1991, *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*, Norton: New York, NY.

Edwards, Franklin R. and Cindy W. Ma, 1992, *Futures and Options*, McGraw-Hill: New York, NY.

Einstein, A., 1905, "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen" (On the Molecular Kinetic Theory of the Heat-Generated Motion of Particles Suspended in Fluid), *Annalen der Physik*, Series 4, Vol 17, pp549–560.

Engle, Robert F., 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, Vol 50 No 4, (July), pp987–1007.

Engle, Robert F., 1993, "Statistical Models for Financial Volatility," *The Financial Analysts Journal*, Vol 49 No 1, (January/ February), pp72–78.

Evans, Merran, Nicholas Hastings, and Brian Peacock, 1993, *Statistical Distributions*, Second Edition, John Wiley and Sons: New York, NY.

Fabozzi, Frank J. and T. Dessa Fabozzi, 1995, *The Handbook of Fixed Income Securities*, Irwin: New York, NY.

Fama, Eugene, F., 1965, "The Behavior of Stock Market Prices," *The Journal of Business*, Vol 38 No 1, (January), pp34–105.

Fama, Eugene F. and Kenneth R. French, 1988, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, Vol 96 No 2, (April), pp246–273.

Fama, Eugene F. and Kenneth R. French, 1992, "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, Vol 47 No 2, (June), pp427–465.

Fama, Eugene F. and Kenneth R. French, 1996, "The CAPM is Wanted, Dead or Alive," *The Journal of Finance*, Vol 51 No 5, (December), pp1947–1958.

Farlow, Stanley J., 1993, *Partial Differential Equations for Scientists and Engineers*, Dover: New York, NY.

Feller, William, 1968, *An Introduction to Probability Theory and its Applications*, Volume I, Third Edition, John Wiley and Sons: New York, NY.

Ferguson, Michael F. and Richard L. Shockley, 2003, "Equilibrium Anomalies," *The Journal of Finance*, Vol 58 No 6, (December), pp2549–2580.

Fisher, L. and R. Weil, 1971, "Coping with the Risk of Interest Rate Fluctuations: Returns to Bondholders from Naive and Optimal Strategies," *The Journal of Business*, Vol 44 No 4, (October), pp408–431.

Fleming, Jeff, 1998, "The Quality of Market Volatility Forecasts Implied by S&P100 Index Option Prices," *The Journal of Empirical Finance*, Vol 5 No 4, (October), pp317–345.

Fraser, Michael K., 1993, "What It Takes to Excel in Exotics," *Global Finance*, Vol 7 No 3, (March), pp44–49.

Fry, Ron, 2009, *101 Great Answers to the Toughest Interview Questions*, Sixth Edition, Thomson Course Technology PTR: Clifton Park, NY.

Garman, Mark B. and Steven W. Kohlhagen, 1983, "Foreign Currency Option Values," *Journal of International Money and Finance*, Vol 2 No 3, (December), 231–237.

Geske, Robert, 1979, "A Note on an Analytic Valuation Formula for Unprotected American Call Options on Stocks with Known Dividends," *Journal of Financial Economics*, Vol 7 No 4, (December) pp375–380.

Girsanov, I.V., 1960, "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution Measures," *Theory of Probability and its Applications*, Vol 5, pp285–301.

Gleick, James, 1987, *Chaos: Making a New Science*, Penguin: New York, NY.

Gleick, James, 1993, *Genius: The Life and Science of Richard Feynman*, Vintage Books: New York, NY.

Goldman, B.M., H.B. Sosin, and M.A. Gatto, 1979, "Path Dependent Options: 'Buy at the Low, Sell at the High'," *The Journal of Finance*, Vol 34 No 5, (December), pp1111–1127.

Goldman, B.M., H.B. Sosin, and L.A. Shepp, 1979, "On Contingent Claims That Insure Ex-Post Optimal Stock Market Timing," *The Journal of Finance*, Vol 34 No 2, (May), pp401–414.

Grabbe, J. Orlin, 1983, "The Pricing of Put and Call Options on Foreign Exchange," *Journal of International Money and Finance*, Vol 2 No 3, (December), 239–253.

Greene, William H., 1993, *Econometric Analysis*, Second Edition, MacMillan: New York, NY.

Hammer, Jerry A., 1989, "On Biases Reported in Studies of the Black-Scholes Option Pricing Model," *Journal of Economics and Business*, Vol 41 No 2, (May), pp153–169.

Harrison, J. Michael, 1985, *Brownian Motion and Stochastic Flow Systems*, John Wiley and Sons: New York, NY.

Harrison, J.M. and S.R. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, Vol 11, pp215–260.

Haug, Espen Gaarder, 2001, "The Options Genius," *Wilmott Magazine*, (July), pp1–4.

Haug, Espen Gaarder, 2007, *The Complete Guide to Option Pricing Formulas*, Second Edition, McGraw-Hill: New York, NY.

Haung, Chi-fu, 1992, *Theory of Financial Markets*, Unpublished book manuscript (incomplete), Department of Finance, Sloan School of Management, MIT, Cambridge, MA 02142.

Hendry, David F., 1986, "An Excursion into Conditional Varianceland," *Econometric Reviews*, Vol 5 No 1, pp63–69.

Holland, A.S.B., 1973, *Introduction to the Theory of Entire Functions*, Academic Press: New York, NY.

Huff, Darrell, 1982, *How to Lie with Statistics*, Norton: New York, NY.

Hull, John C., 1997, *Options, Futures, and Other Derivatives*, Third Edition, Prentice-Hall: Englewood Cliffs, NJ.

Hull, John C., 1998, *Introduction to Futures and Options Markets*, Third Edition, Prentice-Hall: Englewood Cliffs, NJ.

Hull, John and Alan White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, Vol 42 No 2, (June), pp281–300.

Hull, John and Alan White, 1993, "Efficient Procedures for Valuing European and American Path-Dependent Options," *The Journal of Derivatives*, Vol 1, (Fall), pp21–31.

Hunter, William C. and David W. Stowe, 1992, "Path-Dependent Options: Valuation and Applications," *Economic Review (Federal Reserve Bank of Atlanta)*, Vol 77 No 4, (July/August), pp30–43.

Jarrow, Robert and Andrew Rudd, 1983, "Approximate Option Valuation for Arbitrary Stochastic Processes," *Journal of Financial Economics*, Vol 10 No 3, (November), pp347–369.

Jarrow, Robert and Stuart Turnbull, 1996, *Derivative Securities*, South-Western College Publishing: Cincinnati, OH.

Johannes, Michael and Suresh Sundaresan, 2003, "The Impact of Collateralization on Swap Rates," working Paper, Columbia Business School (July).

Jones, Frank J., 1991, "Yield Curve Strategies," *The Journal of Fixed Income*, Vol 1 No 2, (September), pp43–51.

Kahn, Ronald N. and Roland Lochoff, 1990, "Convexity and Exceptional Return," *The Journal of Portfolio Management*, Vol 16 No 2, (Winter), pp43–47

Kahneman, David and Amos Tversky, 1982, "The Psychology of Preferences," *Scientific American*, Vol 246, pp160–173.

Kotz, Samuel and Norman L. Johnson (editors-in-chief), and Campbell B. Read (associate editor), 1982, *Encyclopedia of Statistical Sciences*, Vol 6, John Wiley and Sons: New York, NY.

Krämer, Walter and Ralf Runde, 1997, "Chaos and the Compass Rose," *Economics Letters*, Vol 54 No 2, (February), pp113–118.

Krause, Robert (editor), 1998, *Global Equity and Derivative Market Risk*, Morgan Stanley Dean Witter Quantitative Strategies Group, Morgan Stanley and Co.: New York, NY.

Kritzman, Mark, 1992a, "What Practitioners Need to Know About Utility," *The Financial Analysts Journal*, Vol 48 No 3, (May/June), pp17–20.

Kritzman, Mark, 1992b, "What Practitioners Need to Know About Duration and Convexity," *The Financial Analysts Journal*, Vol 48 No 6, (November/December), pp17–20.

Kritzman, Mark, 1993a, "What Practitioners Need to Know About Factor Methods," *The Financial Analysts Journal*, Vol 49 No 1, (January/February), pp12–15.

Kritzman, Mark, 1993b, "What Practitioners Need to Know About the Term Structure of Interest Rates," *The Financial Analysts Journal*, Vol 49 No 4, (July/August), pp14–18.

Kritzman, Mark, 1993c, "What Practitioners Need to Know About Hedging," *The Financial Analysts Journal*, Vol 49 No 5, (September/October), pp22–26.

Kunitomo, Naoto and Masayuki Ikeda, 1992, "Pricing Options with Curved Boundaries," *Mathematical Finance*, Vol 2 No 4, (October), pp275–298.

Lacey, Nelson J. and Sanjay K. Nawalkha, 1993, "Convexity, Risk, and Returns," *The Journal of Fixed Income*, Vol 3 No 3, (December), pp72–79.

Latané, Henry A. and Richard J. Rendleman, Jr., 1976, "Standard Deviations of Stock Price Ratios Implied in Option Prices," *The Journal of Finance*, Vol 31 No 2, (May), pp369–381.

Lewis, William Dodge, Henry Seidel Canby, and Thomas Kite Brown (editors), 1942, *The Winston Dictionary*, The John C. Winston Company: Philadelphia. PA.

Lewis, Michael M., 1990, *Liar's Poker: Rising Through the Wreckage of Wall Street*, Penguin Books: New York, NY.

Litterman, Robert, and José Scheinkman, "Common Factors Affecting Bond Returns," *The Journal of Fixed Income*, Vol 1 No 1, (June), pp54–61.

Lo, Andrew W. and A. Craig MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *The Review of Financial Studies*, Vol 1 No 1, (Spring), pp41–66.

Lo, Andrew W. and Jiang Wang, 1995, "Implementing Option Pricing Models When Asset Returns are Predictable," *The Journal of Finance*, Vol 50 No 1, (March), pp87–129.

Lorenz, Edward N., 1963, "Deterministic Nonperiodic Flow," *Journal of the Atmospheric Sciences*, Vol 20 No 2, (March), pp130–140.

Lyons, N.I., and K. Hutcheson, 1996, "Generation of Ordered Multinomial Frequencies," *Applied Statistics*, Vol 45 No 3, pp387–393.

Macaulay, Frederick Robertson, 1938, *Some Theoretical Problems Suggested by the Movements of Interest Rates and Stock Prices in the United States Since 1856*, National Bureau of Economic Research: New York, NY.

MacMillan, L.W., 1986, "Analytic Approximation for the American Put Option," *Advances in Futures and Options Research*, Vol 1 Part A, pp119–139.

Mehran, Jamshid and Ghassem Homaifar, 1993, "Duration and Convexity for Bonds with Embedded Options: The Case of Convertibles," *The Journal of Business Finance and Accounting*, Vol 20 No 1, (January), pp107–113.

Melino, Angelo and Stuart Turnbull, 1990, "Pricing Foreign Currency Options with Stochastic Volatility," *The Journal of Econometrics*, Vol 45 No 1/2, (July/August), pp239–265.

Merton, Robert C., 1973, "Rational Theory of Option Pricing," *Bell Journal of Economics and Management Science*, Vol 4 No 1, (Spring), pp141–183. Note that this appears as Chapter 8 in Merton (1992).

Merton, Robert C., 1976, "Option Pricing When Underlying Stock Returns Are Discontinuous," *The Journal of Financial Economics*, Vol 3 No 1, (January/March), pp125–144.

Merton, Robert C., 1992, *Continuous-Time Finance*, Blackwell: Cambridge, MA.

Minton, Bernadette A., 1997, "An Empirical Examination of Valuation Models for Plain Vanilla U.S. Interest Rate Swaps," *The Journal of Financial Economics*, Vol 44 No 2, (May), pp251–277.

Mongan, John, Noah Suojanen, and Eric Giguère, 2007, *Programming Interviews Exposed: Secrets to Landing Your Next Job*, Second Edition, Wrox/John Wiley and Sons: New York, NY.

Mullens, David W., 1982, "Does the Capital Asset Pricing Model Work?," *Harvard Business Review*, Vol 60 No 1, (January/February), pp105–114.

Murphy, Gareth, 1994, "When Options Price Theory Meets the Volatility Smile," *Euromoney*, No 299, (March), pp66–74.

Musiela, Marek and Marek Rutkowski, 1997, *Martingale Methods in Financial Modelling*, Springer-Verlag: Berlin.

Naik, Vasanttilak and Moon Lee, 1990, "General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns," *The Review of Financial Studies*, Vol 3 No 4, pp493–521.

Natenberg, Sheldon, 1994, *Option Volatility and Pricing: Advanced Trading Strategies and Techniques*, Irwin: Chicago, IL.

Options Clearing Corporation, 1993, *Understanding Stock Options*, (September), The Options Clearing Corporation, 440 S. LaSalle St., Suite 2400, Chicago, IL 60605.

Parkinson, Michael, 1977, "Option Pricing: The American Put," *The Journal of Business*, Vol 50 No 1, (January), pp21–39.

Pearce, David, W., 1984, *The Dictionary of Modern Economics*, The MIT Press: Cambridge, MA.

Peterson, Richard L., Christopher K. Ma, and Robert J. Ritchey, 1992, "Dependence in Commodity Prices," *The Journal of Futures Markets*, Vol 12 No 4, (August), pp429–446.

Poterba, James and Lawrence Summers, 1988, "Mean Reversion in Stock Returns: Evidence and Implications," *Journal of Financial Economics*, Vol 22 No 1, (October), pp27–60.

Protter, 2007, "The Work of Kyoshi Itô," *Notices of the American Mathematical Society*, Vol 54 No 6, (June/July 2007), pp744–745.

Rendleman, Richard J., Jr. and Brit J. Bartter, 1979, "Two-State Option Pricing," *The Journal of Finance*, Vol 34 No 5, (December), pp1093–1110.

Richardson, Matthew, 1993, "Temporary Components in Stock Prices: A Skeptic's View," *Journal of Business and Economic Statistics*, Vol 11 No 2, (April), pp199–207.

Ritchken, Peter, L. Sankarasubramanian, and Anand M. Vijh, 1993, "The Valuation of Path Dependent Contracts on the Average," *Management Science*, Vol 39 No 10, (October), pp1202–213.

Roll, R., 1977a, "A Critique of the Asset Pricing Theory's Tests: Part I: On Past and Potential Testability of the Theory," *The Journal of Financial Economics*, Vol 4 No 2, (March), pp129–176.

Roll, R., 1977b, "An Analytical Formula for Unprotected American Call Options on Stocks with Known Dividends," *The Journal of Financial Economics*, Vol 5 No 2, (November), pp251–258.

Samuelson, Paul A., 1965, "Rational Theory of Warrant Pricing," *Industrial Management Review*, Vol 6 No 2, (Spring), pp13–31.

Samuelson, Paul A., 1973, "Mathematics of Speculative Price," *SIAM Review*, Vol 15 No 1, (January), pp1–42.

Scott, Louis O., 1987, "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application," *The Journal of Financial and Quantitative Analysis*, Vol 22 No 4, (December), pp419–438.

Sharpe, W.F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, Vol 19 No 3, (September), pp425–442.

Sharpe, William F., 1978, *Investments*, Prentice-Hall: Englewood Cliffs, NJ.

Shreve, Steven E., 2004, *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer: New York, NY.

Smith, Clifford W., Jr., 1976, "Option Pricing: A Review," *The Journal of Financial Economics*, Vol 3 No 1/2, (January/March), pp3–51.

Snyder, Gerard L., 1969, "Alternative Forms of Options," *The Financial Analysts Journal*, Vol 25 No 5, (September/October), pp93–101.

Spiegel, Murray R., 1956, *College Algebra*, McGraw-Hill: New York, NY.

Spiegel, Murray R., 1968, *Mathematical Handbook*, McGraw-Hill: New York, NY.

Spiegel, Murray R., 1981, *Advanced Calculus*, McGraw-Hill: New York, NY.

Sprenkle, Case M., 1961, "Warrant Prices as Indicators of Expectations and Preferences," *Yale Economic Essays*, Vol 1 No 2, (Fall), pp178–231. Note that this paper appears in Cootner (1964).

Sullivan, Edward J. and Timothy M. Weithers, 1991, "Louis Bachelier: The Father of Modern Option Pricing Theory," *Journal of Economic Education*, Vol 22 No 2, (Spring), pp165–171.

Sullivan, Sara, 1993, "Risk reversals," *Euromoney Treasury Manager*, (December 3), p15.

Sundaresan, Suresh, 1997, *Fixed Income Markets and Their Derivatives*, South-Western College Publishing: Cincinnati, OH.

Thorp, Edward O., 1973, "Extensions of the Black-Scholes Option Model," *Proceedings of the 39th Session of the International Statistical Institute, Vienna*, appearing in *Bulletin of the International Statistical Institute*, Vol 45 Book 2, pp522–529.

Tian, Yisong, 1993, "A Modified Lattice Approach to Option Pricing," *The Journal of Futures Markets*, Vol 13 No 5, (August), pp563–577.

Trippi, Robert R., Edward A. Brill, and Richard B. Harriff, 1992, "Pricing Options on an Asset with Bernoulli Jump-Diffusion Returns," *The Financial Review*, Vol 27 No 1, (February), pp59–79.

Tversky, Amos and Daniel Kahneman, 1974, "Judgement under Uncertainty: Heuristics and Biases," *Science*, Vol 185, (September 27), pp1124–1131.

Tversky, Amos and Daniel Kahneman, 1981, "The Framing of Decisions and the Psychology of Choice," *Science*, Vol 211, (January 30), pp453–458.

Wallace, Anise, 1980, "Is Beta Dead?," *Institutional Investor*, Vol 14 No 7, (July), pp23–30.

Whaley, Robert, 1981, "On the Valuation of American Call Options on Stocks with Known Dividends," *The Journal of Financial Economics*, Vol 9 No 2, (June), pp207–211.

Wiggins, James B., 1987, "Option Values under Stochastic Volatility," *The Journal of Financial Economics*, Vol 19 No 2, (December), pp351–372.

Wilmott, Paul, 1998, *Derivatives: The Theory and Practice of Financial Engineering*, John Wiley and Sons: Chichester, England.

Wilmott, Paul, Jeff Dewynne, and Sam Howison, 1993, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press: Oxford, England.

Wilmott, Paul, Sam Howison, and Jeff Dewynne, 1997, *The Mathematics of Financial Derivatives: A Student Introduction*, Cambridge University Press: Cambridge, England.

Xu, Xinzhong and Stephen J. Taylor, 1995, "Conditional Volatility and the Informational Efficiency of the PHLX Currency Options Market," *The Journal of Banking and Finance*, Vol 19 No 5, (August), pp803–821.

Zimmerman, Heinz and Wolfgang Hafner, 2007, "Amazing Discovery: Vincenz Bronzin's Option Pricing Models," The Journal of Banking and Finance, Vol 31, pp532–546.

Alphabets and Numerical Equivalences

Greek[a]				NATO Phonetic		Roman (Latin)[a]	
α	A	Alpha	1	A	Alpha	A	50; 500
β	B	Beta	2	B	Bravo	B	300
γ	Γ	Gamma	3	C	Charlie	C	100
δ	Δ	Delta	4	D	Delta	D	500
ϵ	E	Epsilon	5	E	Echo	E	250
ζ	Z	Zeta	7	F	Foxtrot	F	40
η	H	Eta	8	G	Golf	G	400
θ	Θ	Theta	0	H	Hotel	H	200
ι	I	Iota	10	I	India	I	1
κ	K	Kappa	20	J	Juliett	J	–[b]
λ	Λ	Lambda	30	K	Kilo	K	250
μ	M	Mu	40	L	Lima	L	50
ν	N	Nu	50	M	Mike	M	1,000
ξ	Ξ	Xi	60	N	November	N	90
o	O	Omicron	70	O	Oscar	O	11
π	Π	Pi	80	P	Papa	P	400
ρ	R	Rho	100	Q	Quebec	Q	90; 500
σ	Σ	Sigma	200	R	Romeo	R	80
τ	T	Tau	300	S	Sierra	S	7;70
υ	Υ	Upsilon	400	T	Tango	T	160
ϕ	Φ[c]	Phi	500	U	Uniform	U	–[d]
χ	X[c]	Chi	700	V	Victor	V	5
ψ	Ψ[c]	Psi	700	W	Whiskey	W	–[e]
ω	Ω	Omega	800	X	X-Ray	X	10
				Y	Yankee	Y	150
				Z	Zulu	Z	2,000

[a]Some information from Lewis et al. (1942, p1161). The book is out of print and the publisher defunct.
[b]Originally the same as I.
[c]The Greek letters Φ, X, and Ψ were not needed in the medieval Latin alphabet. However, the Romans used them as numerical symbols, writing D (or M), X, and L, respectively.
[d]Originally the same as V.
[e]Not used in medieval Latin.

Index

INDEX

