**Review of Statistics: Notation and Definitions**

Consider two random variables $X$ and $Y$ defined over $m$ distinct possible events. Event $i$ occurs with probability $p_i$, in which case $X$ and $Y$ take on values $x_i$ and $y_i$. Thus the probabilities of the various events occurring are $p_1$, $p_2$, ..., $p_m$, and $X$ and $Y$ take on possible values $x_1$, $x_2$, ..., $x_m$ and $y_1$, $y_2$, ..., $y_m$ respectively. If we have considered all possible events, then it must be that the sum of the $m$ probabilities is one: $\sum_{i=1}^{m} p_i = 1$.

For example, consider the following game. We flip a coin once. In the event that it lands heads-up, Beth and Bob each get \$1. In the event the coin lands tails-up, Beth gets \$2 and Bob gets nothing. Here there are two possible events (heads and tails), which occur with equal probability $p_1 = p_2 = 0.5$. The payoffs to Beth and Bob are each random variables which we could call $X$ and $Y$.

We will use the following notation

$$\begin{aligned}
\mathbf{E}(X) &= \text{The Expected Value of } X \\
\mathbf{Var}(X) &= \text{The Variance of } X \\
\mathbf{Cov}(X,Y) &= \text{The Covariance of } X \text{ and } Y \\
\mu_X &= \mathbf{E}(X) \\
\sigma_X &= \text{The Standard Deviation of } X \\
\rho_{XY} &= \text{The Correlation of } X \text{ and } Y
\end{aligned}$$

These quantities are defined as follows.

$$\begin{aligned}
\mathbf{E}(X) &= p_1 x_1 + p_2 x_2 + ... + p_m x_m \\
&= \sum_{i=1}^{m} p_i x_i \\
\mathbf{Var}(X) &= p_1(x_1 - \mu_X)^2 + p_2(x_2 - \mu_X)^2 + ... + p_m(x_m - \mu_X)^2 \\
&= \sum_{i=1}^{m} p_i (x_i - \mu_X)^2 \\
&= \mathbf{E}[(X - \mu_X)^2] \\
\sigma_X &= \sqrt{\mathbf{Var}(X)} \\
\mathbf{Cov}(X,Y) &= p_1(x_1 - \mu_X)(y_1 - \mu_Y) + p_2(x_2 - \mu_X)(y_2 - \mu_Y) + ... + p_m(x_m - \mu_X)(y_m - \mu_Y) \\
&= \sum_{i=1}^{m} p_i (x_i - \mu_X)(y_i - \mu_Y) \\
&= \mathbf{E}[(X - \mu_x)(Y - \mu_Y)] \\
\rho_{XY} &= \frac{\mathbf{Cov}(X,Y)}{\sigma_X \sigma_Y}
\end{aligned}$$

The expected value of $X$ is an average or mean value of the realizations $x$ which occur with frequencies $p$. The variance of $X$ is a measure of the average amount of variation of the realizations $x$ around their mean value. The covariance of $X$ and $Y$ is a measure of the degree to which values of $X$ which are larger than average tend to coincide with values of $Y$ which are larger or smaller than average. For example, a negative covariance indicates that when the realization of $X$ is greater than $\mu_X$, the corresponding realization of $Y$ tends, on average, to be less than $\mu_Y$. The correlation of $X$ and $Y$ is the covariance normalized so that this value falls between $-1$ and 1. We say that $X$ and $Y$ are perfectly correlated if this correlation is $-1$ or 1.

**Review of Statistics: Identities**

Let $a$, $b$, and $c$ be arbitrary constants. The following identities follow from the definitions above.

$$\begin{aligned}
\mathbf{E}(a) &= a \\
\mathbf{E}(a+X) &= a + \mathbf{E}(X) \\
\mathbf{E}(bX) &= b\mathbf{E}(X) \\
\mathbf{E}(a+bX) &= a + b\mathbf{E}(X) \\
\mathbf{E}(X+Y) &= \mathbf{E}(X) + \mathbf{E}(Y) \\
\mathbf{E}(\sum_{i=1}^{n} X_i) &= \sum_{i=1}^{n} \mathbf{E}(X_i)
\end{aligned}$$

In the last identity, $X_1$, $X_2$, ... $X_n$ are $n$ random variables.

$$\begin{aligned}
\mathbf{Var}(a) &= 0 \\
\mathbf{Var}(a+X) &= \mathbf{Var}(X) \\
\mathbf{Var}(bX) &= b^2\mathbf{Var}(X) \\
\mathbf{Var}(X+Y) &= \mathbf{Var}(X) + \mathbf{Var}(Y) + 2\mathbf{Cov}(X,Y) \\
\mathbf{Var}(\sum_{i=1}^{n} X_i) &= \sum_{i=1}^{n} \mathbf{Var}(X_i) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbf{Cov}(X_i, X_j)
\end{aligned}$$

It follows from the last identity that **if** $n$ random variables are mutually uncorrelated, the variance of their sum is equal to the sum of their variances.

$$\begin{aligned}
\mathbf{Cov}(X,Y) &= \mathbf{Cov}(Y,X) \\
\mathbf{Cov}(X,X) &= \mathbf{Var}(X) \\
\mathbf{Cov}(a,X) &= 0 \\
\mathbf{Cov}(a+X, b+Y) &= \mathbf{Cov}(X,Y) \\
\mathbf{Cov}(bX,Y) &= b\mathbf{Cov}(X,Y) \\
\mathbf{Cov}(bX,cY) &= bc\mathbf{Cov}(X,Y) \\
\mathbf{Cov}(X,Y+Z) &= \mathbf{Cov}(X,Y) + \mathbf{Cov}(X,Z)
\end{aligned}$$

These identities can be proved easily by employing the definitions given above. For example, the identity $\mathbf{E}(a+X) = a + \mathbf{E}(X)$ can be proved as follows:

$$\begin{aligned}
\mathbf{E}(a+X) &= \sum_{i=1}^{m} p_i(a+x_i) \\
&= \sum_{i=1}^{m} p_i a + \sum_{i=1}^{m} p_i x_i \\
&= (a\sum_{i=1}^{m} p_i) + \mathbf{E}(X) \\
&= a + \mathbf{E}(X)
\end{aligned}$$

Similarly, the identity $\mathbf{Var}(a+X) = \mathbf{Var}(X)$ can be proved as follows:

$$\begin{aligned}\mathbf{Var}(a+X) &= \mathbf{E}[((a+X)-\mathbf{E}(a+X))^2] \\ &= \mathbf{E}[((a-\mathbf{E}(a))+(X-\mathbf{E}(X)))^2] \\ &= \mathbf{E}[(X-\mathbf{E}(X))^2] \\ &= \mathbf{Var}(X)\end{aligned}$$

### Simple Application: Reducing Risk Through Diversification

We can show that the risk of the return on a financial portfolio can be reduced by diversification among a large number of securities with risky but mutually uncorrelated returns.

Consider holding a portfolio of $N$ securities with weights $s_i$ ($s_i$ is the proportion of the dollar value of the portfolio held in security $i$) and rates of return $r_i$ ($r_i$ is the rate of return per dollar held in security $i$).

For a portfolio of dollar value $P$, the rate of return on the portfolio is

$$\begin{aligned}r_P &= \frac{s_1Pr_1 + s_2Pr_2 + ... + s_NPr_N}{P} \\ &= \sum_{i=1}^{N} s_i r_i\end{aligned}$$

Each return $r_i$ is a random variable (i.e., individual security returns are risky), and consequently so is the return on the portfolio $r_P$.

The expected rate of return on the portfolio is

$$\begin{aligned}\mathbf{E}(r_p) &= \mathbf{E}(\sum_{i=1}^{N} s_i r_i) \\ &= \sum_{i=1}^{N} s_i \mathbf{E}(r_i)\end{aligned}$$

and if we measure the riskiness of the portfolio as the variance of its return, then this riskiness is

$$\begin{aligned}\mathbf{Var}(r_p) &= \mathbf{Var}(\sum_{i=1}^{N} s_i r_i) \\ &= [\sum_{i=1}^{N} s_i^2 \mathbf{Var}(r_i)] + [\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} s_i s_j \mathbf{Cov}(r_i, r_j)]\end{aligned}$$

Further, **if** the returns of the individual securities are mutually uncorrelated, then all the covariances above are zero, so that

$$\mathbf{Var}(r_p) = \sum_{i=1}^{N} s_i^2 \mathbf{Var}(r_i)$$

Now suppose that we hold equal shares of each security, so that $s_i = 1/N \;\; \forall i$, and for simplicity further suppose that the variance of each security's returns is identical: $\mathbf{Var}(r_i) = \sigma^2 \;\; \forall i$. Then for $N$ independent (uncorrelated) securities

$$\begin{aligned}\mathbf{Var}(r_p) &= \sum_{i=1}^{N} (\frac{1}{N})^2 \sigma^2 \\ &= N \cdot (\frac{1}{N})^2 \sigma^2 \\ &= \frac{\sigma^2}{N}\end{aligned}$$

Thus, as we increase the number of independent securities held in the portfolio while holding the dollar value of the portfolio constant (i.e., increase the degree of diversification of the portfolio), the portfolio's riskiness becomes smaller and goes to zero in the limit as $N \to \infty$.

Note that positive covariances of returns across securities would limit our ability to reduce risk through diversification, and negative covariances would increase our ability to do this with a small number of securities. Models of asset prices, such as CAPM, usually conclude that securities should only command risk premia (extranormal expected returns) for non-diversifiable risk. Thus, a security might have a highly uncertain return, but if its covariation with the rest of the market is small, it may have an expected return close to the rates on relatively riskless assets, since much of its risk can be diversified away.

**Rational Expectations**

The expectations of economic agents concerning the uncertain outcomes of future event often play important roles in determining market outcomes, both in the present and in the future. For example, the current demand and supplies of oil, bonds, and other commodities depend on the expectations of their future prices on the part of buyers and sellers in those markets. Thus, the current prices of these commodities depend on expectations of their future prices.

There are a variety of definitions of *rational expectations* in the economics literature. Intuitively, a rational expectation should be a best possible forecast based on all available information. We can operationalize this by saying that the rational expectation should be a mathematical expectation (expected value). Call $X_{t+j}^{e(t)}$ the expectation at time $t$ of the future random variable $X_{t+j}$. Then if expectations are rational,

$$X_{t+j}^{e(t)} = \mathbf{E}_t X_{t+j}$$

where $\mathbf{E}_t$ denotes the *conditional expectation* based on information available at time $t$. The specific information available at time $t$ will define different probabilities $p_i$ of events than the unconditional probabilities (i.e., long run frequencies) which correspond to the unconditional expectation $\mathbf{E}X_{t+j}$.

An example might be useful here. Consider the outcome of a sequence of two coin tosses. Value a head at \$1 and a tail at \$0. The probability of either a head or a tail in any toss is .5. Call the random variable $Z$ the cumulated return to the two tosses. The unconditional expectation of the cumulated return is $\mathbf{E}(Z) = \$0.5 + \$0.5 = \$1$. However, if in between the two tosses (call this time $t_o$), we are told that the first toss was a head, then the conditional expectation of $Z$ is now $\mathbf{E}_{t_o} Z = \$1 + \$0.5 = \$1.5$. If we had been told that the first toss had been a tail, then the conditional expectations would have been $\mathbf{E}_{t_o} Z = \$0 + \$0.5 = \$0.5$.

Now this is an easy example, since we can argue (as long as we are willing to avoid deep questions about the nature of randomness) that there are objective probabilities attached to a flip of a coin.[1] However, the macro economy is much more complicated than a controlled game like a coin toss, and this has lead many economists to question the relevance of this approach. For example, John Maynard Keynes (1935) argued that the expected returns on investment projects could not be calculated as mathematical expectations because there was no way of establishing good probabilities for the success of new products and processes or for the myriad of events that might take place in the macro economy in the future. As further support for this claim, he noted that macroeconomic events which we might try to forecast (for example the likelihood of a recession) depend themselves on market participants' expectations of them; thus my expectations must depend on my beliefs about others' expectations. Thus, Keynes believed that what is now called rational expectations (the term was coined much later, in the 1960s) is not an appropriate theory of expectations formation in the macro economy.[2] If it is not clear to agents in the economy how their current information

[1] In reality, the flip of the coin is deterministic. However, since we typically don't have the computational ability to forecast the result, we pragmatically assume that the outcome is truly random. Similarly, computer random number generators are deterministic (there is an algorithm that produces the numbers, but if you don't know the algorithm it's pretty hard to predict the next number). It is now standard to model quantum events as being intrinsically random, but again this is a mater of some debate (does God play dice?).

[2] Keynes believed that mathematical expectations *were* useful in describing expectations formation in certain situations,

translates into probabilities of various outcomes occurring (what is the likelihood that inflation will be at 8% in five years?), then expectations may be indeterminate or "irrational" (the meaning of which is unclear).

At a pragmatic level, the assumption of rational expectations is often operationalized by asserting that if agents are making the best forecasts that they can given the available information, then certain conditions should hold. Particularly, consider the expectational error $\epsilon_{t+j}$ defined as

$$\epsilon_{t+j} = X_{t+j} - X_{t+j}^{e(t)}$$

This is just how far off your expectation turns out to be. If expectations are rational (i.e., people are making the best forecasts that they can), then the following conditions should hold.

1. The expectational error aught to be zero on average. I.e., the unconditional expectation of the forecast error aught to be zero:
$$\mathbf{E}\epsilon_{t+j} = 0$$
This says that expectations are not *systematically* incorrect. We do not, for example, *continuously* under- or over- estimate future rates of inflation or returns to a security over long periods of time.
2. The expectational error aught to be uncorrelated with all variables who's values are known at time $t$:
$$\mathbf{Cov}(\epsilon_{t+j}, Y) = 0$$
where $Y$ is any variable who's value is known at time $t$. This says that all known correlations between other variables and the variable that we are trying to forecast have been exploited in our forecast. If this were not true, then we could improve our forecasts by taking these correlations into consideration. For example, if current raw materials prices are correlated with the general price level three months in the future, then information on this "leading indicator" will be exploited in the formation of inflationary expectations. Similarly, if stock prices systematically behave differently in January than in other months, this typical behavior will be anticipated by market participants.

These conditions are very powerful and are the foundation of rational expectation econometrics. Some simple examples follow.

**Example 1: Are Stock Prices Excessively Volatile?**

Rational expectations is one of the foundations of efficient market theory (EMT) in financial market theory.

According to the EMT, stock prices are based on market fundamentals rather than speculative behavior not related to market fundamentals. Consider a stock with uncertain (random) future dividends $d_{t+i}$. Suppose for simplicity that there is a constant interest rate $R$ on other securities in this security's risk class. Then the EMT predicts that the price of this stock at time $t$ should be the rational expectation of the present value of future dividends paid on the stock:

$$\begin{aligned} P_t &= \mathbf{E}_t PV_t \\ &= \mathbf{E}_t \sum_{i=1}^{\infty} (\frac{1}{1+R})^i d_{t+i} \\ &= \sum_{i=1}^{\infty} (\frac{1}{1+R})^i \mathbf{E}_t d_{t+i} \end{aligned}$$

Call this the *fundamental value* of the stock at time $t$.

---

for example in insurance markets in which there are relatively stable long term frequencies of events (an 18 year old driving a loaded sports car will on average cost the insurance company more than a 35 year old economics professor driving a family wagon) and good data collection.

We will return to the derivation of this later, to ask whether *rational* speculative bubbles might (in theory) make the price deviate from its fundamental value.

For now, however, consider Shiller's (1981) empirical test for whether stock prices deviate from their fundamental values in practice. He tests to see if actual stock prices are more volatile than should be the case if they follow expected future dividends as above.

Shiller notes that, if expectations are rational, as EMT assumes, then we can write the relationship above

$$\begin{aligned} P_t &= \mathbf{E}_t PV_t \\ &= PV_t - \epsilon_t \end{aligned}$$

where $PV_t$ is the actual present value (i.e., the present value of the actual future realizations of the dividends) and $\epsilon_t$ is the error $(PV_t - \mathbf{E}_t PV_t)$ in forecasting this present value. Rearranging, we have

$$PV_t = P_t + \epsilon_t$$

But then, taking the variance of both sides

$$\mathbf{Var}(PV_t) = \mathbf{Var}(P_t) + \mathbf{Var}(\epsilon_t) + 2\mathbf{Cov}(P_t, \epsilon_t)$$

The covariance is zero under rational expectations; since today's price is observed today, the forecast error should not be correlated with it. Thus,

$$\mathbf{Var}(PV_t) = \mathbf{Var}(P_t) + \mathbf{Var}(\epsilon_t) \geq \mathbf{Var}(P_t)$$

(variances can not be negative). Thus, we have the prediction, from the EMT, that the variance of stock prices should *not* be greater than the variance of the present value of actual future dividends on the stock. Intuitively, the price of a stock should be based on the expected values of future dividends, which are less volatile (over time for example) than the actual realizations of future dividends.

This is a testable hypothesis, since we have historical data on stock prices and subsequent dividends. Since $(\frac{1}{1+R})^i$ becomes small as $i$ becomes large, we can get a decent measure of $PV_t$ with a finite number of subsequent periods. Shiller constructs time series for stock prices and (partial) present values of subsequent dividends for a variety of stocks, and finds that the sample variances of price tends to be *greater than* the sample variances of dividends, violating the prediction of the EMT.

**Example 2: Do Consumers Follow Life-Cycle Consumption Behavior?**

The life-cycle model of consumption assumes that consumers desire to smooth consumption over their lifetimes. Therefore, each household tries to forecast its lifetime income and spread this income evenly across consumption in each year by borrowing and saving it.

If you try to keep consumption constant, and future income is uncertain, then even though you may be surprised next year by new information on your lifetime income (you are told that your company is planning to move to Tennessee, and is not planning to take you with it), the expected value of this surprise based on your current information is zero. Thus, your best guess today about your consumption tomorrow is that it will be the same as it is today:

$$\mathbf{E}_t C_{t+1} = C_t$$

We can rewrite the expected value of consumption next year as the actual ex post value minus a forecast error, so that:

$$C_{t+1} = C_t + \epsilon_{t+1}$$

This says that consumption "follows a random walk." $\epsilon_{t+1}$ can be thought of as a shock term that is zero on average. Consumption stays at its current level until it is shocked. Once shocked then consumption will jump (e.g., if the realization of $\epsilon_{t+1}$ is positive, then consumption jumps up at time $t+1$) and stay at it's new level until it is shocked again.

Therefore, a way of testing empirically whether consumption follows the life-cycle hypothesis is to test the prediction of that theory that consumption follows a random walk. To do this, we can estimate the following regression equation

$$C_{t+1} = a + bC_t + e_{t+1}$$

The joint hypotheses of life-cycle behavior and rational expectations predict that $a = 0$, $b = 1$, and $e_{t+1}$ is uncorrelated with variables known at time $t$. A typical finding (e.g., Hall(1978)) is that the residual from this regression is correlated with disposable income at time $t$ and consumption at time $t-1$: i.e., these variables have predictive power for $C_{t+1}$ not captured by $C_t$. This is evidence that either consumers are not life-cyclers or do not have rational expectations. One possible explanation is that some consumers would like to be lifecyclers but are liquidity constrained and so can not borrow against future expected income. Thus, their consumption follows current disposable income rather than expected lifetime income. Alternative explanations are that they are simply not life-cyclers or have systematically incorrect expectations. The average propensity to consume rose substantially in the US in the 1980s. One possible explanation for this fall in saving behavior is that the growth path of income fell below its previous trend and so was systematically lower than expected over the course of the decade.

## References


Hall, Robert, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, **86**, December 1978, 971–988.

Keynes, John Maynard, *The General Theory of Employment, Interest, and Money*, 1935.

Shiller, Robert, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, **71**, 1981, 421–436.