Econ 201 Review Handout 1: Rules for Expected Value and Variance

Edward Vytlacil, Yale University

May 11, 2023

This handout reviews expected value, variance and covariance of random variables, as well as rules for the expected value, variance, and covariance of weighted sums random variables. We will use these rules repeatedly throughout the course.

Contents

Expected Value and Variance of a Random Variable

THE EXPECTED VALUE of a random variable X, also referred to as the population mean of X, is a measure of central tendency, also called a measure of central location, and is defined as follows,

Definition 1: Expected value

Suppose X is a discrete r.v. taking values in $\{x_1, x_2, ..., x_K\}$. Then the *expected value* of X is defined by

$$\mathbb{E}(X) = \sum_{k=1}^{K} x_k \Pr\{X = x_k\}.$$

Expected value is the most traditional measure of central tendency. There are other measures that get used, such as the median of X.

The expected value of X is a weighted average of the values that X can take, weighted by the probability that X equals each of those values. While X is a random variable, $\mathbb{E}[X]$ is a constant.

Expected value plays a key role in economic theory of decision making under uncertainty, with individuals typically modeled as maximizing expected utility and firms modeled as maximizing expected profits. It plays a central role in finance, with investors desiring a higher expected return on investments. It is also plays a central role in econometrics and statistics.

Remark 1. *We will often consider the expected value of a function of a random variable, such as $Y = f(X)$ for some function f and random variable X. In that case, Y is a random variable with*

$$\mathbb{E}(Y) = \mathbb{E}(f(X)) = \sum_{k=1}^{K} f(x_k) \Pr\{X = x_k\}.$$

For example, $\mathbb{E}[X^2] = \sum_{k=1}^{K} x_k^2 \Pr\{X = x_k\}$, and for any constant c we have $\mathbb{E}[\ (X-c)^2\ ] = \sum_{k=1}^{K} (x_k - c)^2 \Pr\{X = x_k\}$.

Remark 2. *For simplicity, we have defined expected value for a discrete random variable taking a finite number of values. The definition can be extended to variables taking an infinite number of values, including for continuous random variables. One complication that arises for random* *variables X that take an infinite number of values is that the expected value of X need not exist, and when it does, it need not be finite. For example, the expected value of a Cauchy random variable does not exist. All of the following results go through for general random variables X, including those that take an infinite number of values, as long as the relevant expected values exist and are finite.*

While the expected value of X is a measure of central location, the variance of a random variable X is a measure of dispersion, how much the random variable varies. In particular, it is defined as the expected squared distance between the variable and its population mean.

While variance is the most traditional measure of dispersion, there are other measures that get used such as mean absolute distance, $E[\,|X - \mathbb{E}(X)|\,]$, and the inter-quartile range (75th quantile minus 25th quantile).

Definition 2: Variance

Suppose X is a r.v. with $\mathbb{E}[X] = \mu$. Then the *variance* of X is defined by

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mu)^2\right].$$

The square root of the variance is called the *standard deviation,* often denoted by $\mathrm{sd}(X) = \sqrt{\mathrm{Var}(X)}$.

Variance is a common measure of uncertainty in economic theory. In finance, variance of an asset return is a measure of risk for that investment, with investors assumed to dislike risk. Variance also plays a central role in econometrics and statistics, and is used in describing the uncertainty of estimates and plays a crucial role in inference based on normality.

While X is a random variable, $\mathrm{Var}(X)$ and thus $\mathrm{sd}(X)$ are constants.

Example 1 (Flipping a Weighted Coin). *Let X denote the result of flipping a weighted coin, with $X = 1$ if heads and $X = 0$ if tails. Let $p = \Pr[X = 1]$ so that $1 - p = \Pr[X = 0]$. Then*

$$\begin{aligned} \mathbb{E}[X] &= 1 \cdot p + 0 \cdot (1 - p) = p, \\ Var[X] &= (1 - p)^2 \cdot p + (0 - p)^2 \cdot (1 - p) = p \cdot (1 - p). \end{aligned}$$

For example, for a fair coin, $p = 1/2$ so that $\mathbb{E}[X] = 0.5$ and $Var(X) = 0.25$. Note that the expected value is 0.5, even though X never takes the value 0.5.

Example 2 (Rolling a Die). *Let X denote the result of rolling a fair die. Then $X = 1, 2, 3, ..., 6$, $P\{X = j\} = \frac{1}{6}$ for $j = 1, 2, ..., 6$, and*

$$\begin{aligned} \mathbb{E}[X] &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + ... + 6 \cdot \frac{1}{6} \\ &= 3\frac{1}{2}, \end{aligned}$$

and

$$\begin{aligned} Var[X] &= (1 - 3.5)^2 \cdot \frac{1}{6} + (2 - 3.5)^2 \cdot \frac{1}{6} + ... + (6 - 3.5)^2 \cdot \frac{1}{6} \\ &= \frac{35}{12}. \end{aligned}$$

Note that the expected value is 3.5, *even though the die will never equal* 3.5.

Example 3 (Asset Return). *Suppose* r_A, *and* r_B *are return on assets* A *and* B *where* $\Pr\{r_A = 0.17\} = \Pr\{r_A = 0.07\} = \frac{1}{2}$ *and* $\Pr\{r_B = 0.22\} = \Pr\{r_B = 0.02\} = \frac{1}{2}$. *Then the expected return on the assets are*

$$\begin{aligned} \mathbb{E}[r_A] &= 0.17 \cdot \frac{1}{2} \; - \; 0.07 \cdot \frac{1}{2} \; = 0.12, \\ \mathbb{E}[r_B] &= 0.22 \cdot \frac{1}{2} \; - \; 0.02 \cdot \frac{1}{2} \; = 0.12, \end{aligned}$$

and the variances of the asset returns are

$$\begin{aligned} Var(r_A) &= (0.17 - 0.12)^2 \cdot \frac{1}{2} \; + \; (0.07 - 0.12)^2 \cdot \frac{1}{2} \; = 0.0025, \\ Var(r_B) &= (0.22 - 0.12)^2 \cdot \frac{1}{2} \; + \; (0.02 - 0.12)^2 \cdot \frac{1}{2} \; = 0.01. \end{aligned}$$

Note that the expected return on assets A *and* B *are both* 0.12, *even though neither asset will have a return of* 0.12. *Assets* A *and* B *have the same expected return, though asset* B *is riskier (more spread to its distribution).*

Theorem 1: Expected Value and Variance of a Linear Function

Suppose X is a random variable. Let a and b denote constants. Then $a + b \cdot X$ is a *linear function* of X, and

$$\begin{aligned} \mathbb{E}[a + b \cdot X] &= a + b \cdot \mathbb{E}[X], \\ \mathrm{Var}(a + b \cdot X) &= b^2 \cdot \mathrm{Var}(X). \end{aligned}$$

Example 4. *Recall that* $\mathbb{E}[X]$ *is a constant, and thus, taking* $a = \mathbb{E}[X]$, $b = 0$, *we have*

$$\begin{aligned} \mathbb{E}[E(X)] &= \mathbb{E}(X) \\ Var[E(X)] &= 0, \end{aligned}$$

while taking $a = -\mathbb{E}[X]$, $b = 1$ *we have*

$$\begin{aligned} \mathbb{E}[\; X - \mathbb{E}(X) \;] &= \mathbb{E}[X] - \mathbb{E}[X] = 0, \\ Var[\; X - \mathbb{E}(X) \;] &= Var(X). \end{aligned}$$

Example 3. *(Asset returns, continued) An investor who desires higher expected return but dislikes risk (variance of returns) would prefer asset* A *to* B *as* A *has the same expected return as asset* B *but is less risky. Now additionally consider an asset C with* $\mathbb{E}[r_C] = 0.20$, *and* $Var[r_C] = 0.04$. *The* *tradeoff A and C is not clear, as A has lower expected return but lower risk than C. One common measure that balances the tradeoff between expected return and risk is the Sharpe Ratio. The Sharpe Ratio for asset j is defined by*

$$S_j = \frac{\mathbb{E}(r_j - r_f)}{\sigma_{j-f}},$$

where $\sigma_{j-f} = \sqrt{Var(r_j - r_f)}$, *and* r_f *is the risk free rate taken to be a constant in theory. An asset with a higher Sharpe Ratio is seen as a better investment. Using Theorem 2 and taking* r_f *to be a constant, we have that* $\sigma_{j-f} = \sigma_j$ *and*

$$S_j = \frac{\mathbb{E}(r_j - r_f)}{\sigma_{j-f}} = \frac{\mathbb{E}(r_j) - r_f}{\sigma_j}.$$

Taking $r_f = 0.02$, *we have*

$$S_A = \frac{0.12 - 0.02}{\sqrt{0.0225}} = 2$$
$$S_C = \frac{0.20 - 0.02}{\sqrt{0.04}} = 0.9,$$

so that in this example, according to their Sharpe ratios, asset C has a worse risk-return tradeoff despite having the higher expected return.

	$\mathbb{E}(r_j)$	$\mathrm{Var}(r_j)$	S_j
r_A	0.12	0.0225	2
r_B	0.12	0.0100	1
r_C	0.20	0.0400	0.9

Table 1: Returns for Example 3 with $r_f = 0.02$

Expected Value and Variance, Rules for 2 Random Variables

Often in economics we analyze more than one variable at the same time. Suppose X and Y are random variables. We will refer to (X, Y) as a *random vector.* We will often work with the expected values of weighted sums of X and Y using the following theorem.

Theorem 2: Expected Value of a Linear fn. of 2 Random Vars

Suppose (X, Y) is a random vector. Let a, b and c denote constants. Then $a + b \cdot X + c \cdot Y$ is a *linear function* of (X, Y), and

$$\mathbb{E}[a + bX + cY] = a + b \cdot \mathbb{E}[X] + c \cdot \mathbb{E}[Y].$$

Example 5. *Taking* $a = 0$, $b = c = 1$, *we have*

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

while taking $a = 0$, $b = 1$ and $c = -1$ we have

$$\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y].$$

Thus, the expected value of a sum is the sum of the expected values, and the expected value of a difference is the difference in the expected values.

Theorem 2 allows us to rewrite the variance of a random variable in a form that will often be convenient:

Theorem 3: Alternative Expression for Variance

Suppose X is a random variable. Then

$$\text{Var}(X) = \mathbb{E}[(X)^2] - (\mathbb{E}[X])^2.$$

Proof:

$$\begin{aligned}
\text{Var}(X) \equiv & \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] \\
= & \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2\right] \\
= & \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2 \\
= & \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2,
\end{aligned}$$

where the third equality is applying Theorem 2 with $Y = X^2$, $a = \mathbb{E}[X]^2$, $b = -2\mathbb{E}[X]$ and $c = 1$.

Remark 3. *Expected value of linear function is linear function of expected value. Expected value of nonlinear function is generally not the nonlinear function of expected value. For example, from Theorem 3, we have that* $\mathbb{E}[X^2] > (\mathbb{E}(X))^2$ *whenever* $Var(X) > 0$.

Now consider the variance of a linear function of (X, Y). Before we can provide a rule for the variance of $a + bX + cY$, we first need to define covariance. The *covariance* between X and Y is a measure of linear dependence between X and Y, of how much X being above its mean is related to Y being above its mean, defined as follows:

Definition 3: Covariance

Suppose (X, Y) is a random vector. Let $\mu_X \equiv \mathbb{E}(X)$ and $\mu_Y \equiv \mathbb{E}(Y)$. Then the *covariance* of X and Y is defined by

$$\text{Cov}(X, Y) = \mathbb{E}\left[\ (X - \mu_X) \cdot (Y - \mu_Y)\ \right].$$

Remark 4. *From their definitions, we have* $Cov(X, X) = Var(X)$.

Correlation is also a measure of linear dependence, which rescales covariance to be between -1 (perfect negative linear dependence) and 1 (perfect positive linear dependence).

Definition 4: Correlation

The *correlation* between X and Y is defined by

$$\text{Corr}(X, Y) \equiv \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}.$$

If $\text{Corr}(X, Y) = 0$ we say that X and Y are *uncorrelated*. One can show that $-1 \leq \text{Corr}(X, Y) \leq 1$.

Two variables are uncorrelated if and only if their covariance is zero. Two random variables being independent implies that they are uncorrelated (equivalently, have zero covariance), but two random variables being uncorrelated does not imply that they are independent.

Using the definition of covariance, we can now state a rule for the variance of a linear function of two variables.

Theorem 4: Variance of a Linear fn. of 2 Random Vars:

Suppose (X, Y) is a random vector. Let a, b and c denote constants. Then

$$\text{Var}(a + bX + cY) = b^2 \cdot \text{Var}(X) + c^2 \cdot \text{Var}(Y) + 2 \cdot b \cdot c \cdot \text{Cov}(X, Y).$$

Example 6. *If* (X, Y) *is a random vector. Then taking* $a = 0$, $b = c = 1$, *we have*

$$Var[X + Y] \; = Var[X] + Var[Y] + 2Cov(X, Y)$$

while taking $a = 0$, $b = 1$ *and* $c = -1$ *we have*

$$Var[X - Y] \; = Var[X] + Var[Y] - 2Cov(X, Y).$$

Recall that the expected value of a sum was the sum of the expected values, and the expected value of a difference was the difference in the expected values. In contrast, the variance of a sum generally does not equal the sum of the variances, and the variance of a difference does not generally equal the difference in the variances. However, if the random variables are uncorrelated, then the variance of the sum is the sum of the variances.

Remark 5. *The above rules for expected value and variance of a sum plays a central role in portfolio analysis. Suppose* (r_A, r_B) *is a random vector of* *returns on the assets A and B. Suppose the investor creates a portfolio that invests fraction ω in asset A and fraction $1-\omega$ in investment B, so that the return on the portfolio is given by*

$$r_P = \omega \cdot r_A + (1-\omega) \cdot r_B.$$

Then the expected value of the returns on the portfolio is given by

$$\mathbb{E}[r_P] = \omega \cdot \mathbb{E}[r_A] + (1-\omega) \cdot \mathbb{E}[r_B],$$

while the variance of the returns on the portfolio is given by

$$Var(r_P) = \omega^2 \cdot Var(r_A) + (1-\omega)^2 \cdot Var(r_B) + 2 \cdot \omega \cdot (1-\omega) \cdot Cov(r_A, r_B).$$

An implication is that the variance of the returns on the portfolio depends on the covariance between the assets that comprise the portfolio. The variance of a portfolio (i.e., the risk of a portfolio) can be decreased by including assets with low covariances with the previously included assets. Thus, one can hedge the risk of the returns on an asset A by diversifying the investment to also include another asset B that has a low covariance with asset A.

Example 3. *(continued). In Example 3, an investor would prefer holding asset A alone to holding asset B alone, as they have the same expected return but asset B is four times as risky. However, as illustrated by Figure 1, if the covariance in their returns is low enough, a portfolio that includes both assets can have a lower risk than holding asset A alone.*

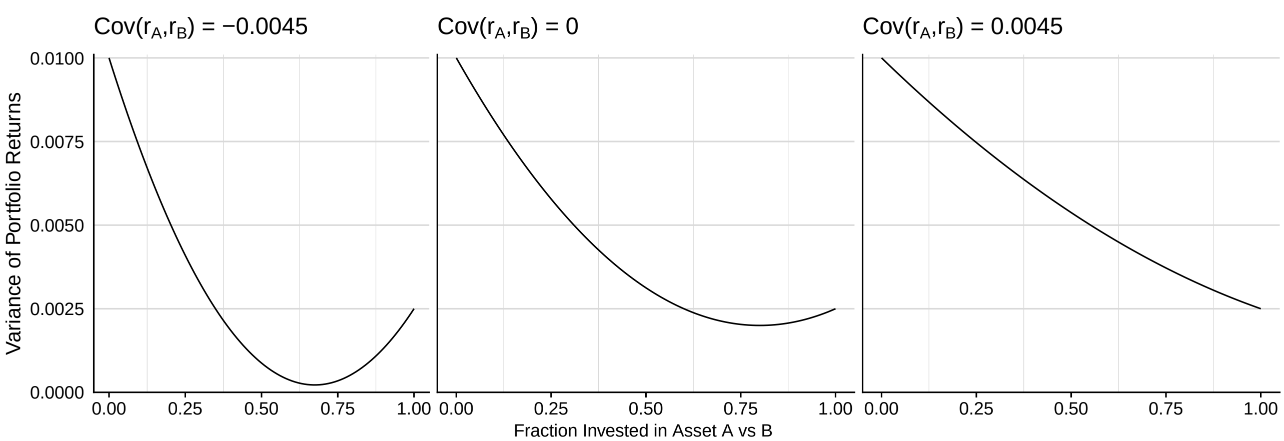

Figure 1: $\text{Var}(r_A) = 0.0025$, $\text{Var}(r_B) = 0.01$, but the variance of the portfolio including both can be below 0.0025 if $\text{Cov}(r_A, r_B)$ is sufficiently small.

We can also consider covariance of linear functions:

Theorem 5: Covariance of linear functions

Suppose (X, Y) is a random vector. Let a, b, c and d denote constants, then

$$\text{Cov}(a + bX, c + dY) = b \cdot d \cdot \text{Cov}(X, Y).$$

Remark 6. *Recall that $Cov(X, X) = Var(X)$. Thus, Theorem 5 implies that $Cov(a + bX, c + dX) = b \cdot d \cdot Var(X)$. For example, $Cov(X, -X) = -Var(X)$.*

The following theorem generalizes Theorem 5.

Theorem 6: Covariance of a linear function

Suppose (X, Y, W, Z) is a random vector. Let a_0, b_0, c_0 and a_1, b_1, c_1 denote constants, then

$$\begin{aligned}
&\text{Cov}(a_0 + b_0 \cdot X + c_0 \cdot Y,\ a_1 + b_1 \cdot W + c_1 \cdot Z) \\
&\quad = b_0 \cdot b_1 \cdot \text{Cov}(X, W) + b_0 \cdot c_1 \cdot \text{Cov}(X, Z) \\
&\qquad + c_0 \cdot b_1 \cdot \text{Cov}(Y, W) + c_0 \cdot c_1 \cdot \text{Cov}(Y, Z).
\end{aligned}$$

Example 3. *(Asset return continued). In the portfolio example, suppose that the investor currently holds the portfolio $r_P = \frac{2}{3} \cdot r_A + \frac{1}{3} \cdot r_B$. The investor may be interested in the covariance between the returns on this portfolio and the market return, how strongly does the portfolio return tend to move with the market return? Let r_M denote the market return. Applying Theorem 6, we have*

$$Cov(r_M, r_P) = \frac{2}{3}Cov(r_M, r_A) + \frac{1}{3}Cov(r_C, r_B).$$

Expected Value and Variance, Rules for K Random Variables

The following theorem generalizes Theorem 2 to the expected value of the weighted sum of K random variables.

Theorem 7: Expected Value of a Linear fn of K Random Vars

Suppose $(X_1, X_2, ..., X_K)$ is a random vector. Let $a_o, a_1, a_2, ..., a_K$ denote constants. Then $a_0 + \sum_{k=1}^{K} a_k X_k$ is a *linear function* of $(X_1, X_2, ..., X_K)$, and

$$E\left[a_0 + \sum_{k=1}^{K} a_k X_k\right] = a_0 + \sum_{k=1}^{K} a_k \mathbb{E}[X_k].$$

The generalization of Theorem 4 to the variance of the sum of K random variables is slightly more complicated, but has a simple form in the case of uncorrelated random variables.

Theorem 8: Var of a Linear fn of K Uncorrelated Random Vars

Suppose $(X_1, X_2, ..., X_K)$ is a random vector. Suppose $Cov(X_i, X_j) = 0$ for all $i \neq j$. Let $a_0, a_1, a_2, ..., a_K$ denote constants. Then

$$\text{Var}\left[a_0 + \sum_{k=1}^{K} a_k X_k\right] = \sum_{k=1}^{K} a_k^2 \text{Var}[X_k].$$

Theorems 7 and 8 can be used, for example, to derive the expected value and variance of a sample mean of i.i.d. random variables, a key result in statistics.

Theorem 9: Mean and Variance of Sample Mean

Suppose $X_1, ..., X_N$ denotes an i.i.d. sequence of random variables. Let $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \text{Var}(X_i)$. Let $\bar{X}_N$ denote the sample mean, $\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$. Then $\mathbb{E}[\bar{X}_N] = \mu$ and $\text{Var}[\bar{X}_N] = \frac{\sigma^2}{N}$.

Proof:

$$\mathbb{E}[\bar{X}_N] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left[X_i\right] = \frac{1}{N}\sum_{i=1}^{N} \mu = \frac{N}{N}\mu = \mu;$$

and

$$\begin{aligned}\text{Var}[\bar{X}_N] = \text{Var}\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] &= \frac{1}{N^2}\sum_{i=1}^{N} \text{Var}\left[X_i\right] \\ &= \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2 = \frac{N}{N^2}\sigma^2 = \frac{\sigma^2}{N}.\end{aligned}$$

Theorem 9 implies that, for an i.i.d. sequence, the sample mean is centered at the population mean and the variance of the sample mean is decreasing in sample size. Thus, the larger the sample size, the less the variability of the sample mean around the population mean.

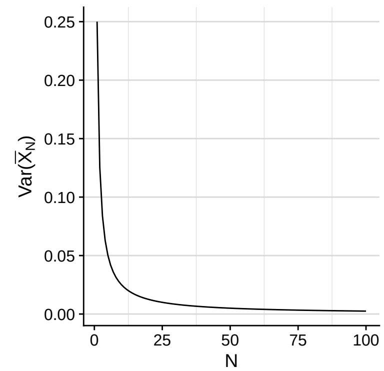

Figure 2: $\mathrm{Var}(\bar{X}_N)$ for N independent flips of a fair coin.

Remark 7. *The standardized sample mean is defined by* $T_N = \frac{\bar{X}_N - \mu}{\sigma/\sqrt{N}} = \sqrt{N}\frac{\bar{X}_N - \mu}{\sigma}$. *By Theorem 4 and 9, for an i.i.d. sequence,* $\mathbb{E}[T_N] = 0$ *and* $Var[T_N] = 1$. *These expressions play a key role in the central limit theorem and in hypothesis testing based on asymptotic normality.*

Summary

Table 2 provides optional reading for this handout.[1] [2]

Source	Chapters
Hogg et. al (2019)	2.2-2.3, 4.2
Woolridge (2020)	B-3, B-4a-B-4d

Table 2: Optional reading

[1]

[2]

Important Definitions

Def 1: $\mathbb{E}(X) = \sum_{k=1}^{K} x_k \Pr\{X = x_k\}$;

Def 2: $\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$;

Def: 3: $\mathrm{Cov}(X, Y) = \mathbb{E}\left[\ (X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])\ \right]$;

Def: 4: $\mathrm{Corr}(X, Y) \equiv \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$.

Important Rules for Expectation and Variance

Thm 1: $\mathbb{E}[a + b \cdot X] = a + b \cdot \mathbb{E}[X]$,

$\mathrm{Var}(a + b \cdot X) = b^2 \cdot \mathrm{Var}(X)$;

Thm 2: $\mathbb{E}[a + bX + cY] = a + b \cdot \mathbb{E}[X] + c \cdot \mathbb{E}[Y]$;

Thm 3: $\mathrm{Var}(X)$ can be rewritten as

$\mathrm{Var}(X) = \mathbb{E}[(X)^2] - (\mathbb{E}[X])^2$;

Thm 4: $\mathrm{Var}(a + bX + cY)$

$= b^2 \cdot \mathrm{Var}(X) + c^2 \cdot \mathrm{Var}(Y) + 2 \cdot b \cdot c \cdot \mathrm{Cov}(X, Y)$;

Thm 7: $\mathbb{E}\left[a_0 + \sum_{k=1}^{K} a_k X_k\right] = a_0 + \sum_{k=1}^{K} a_k \mathbb{E}[X_k]$;

Thm 8: If $Cov(X_i, X_j) = 0$ for all $i \neq j$, then

$\mathrm{Var}\left[a_0 + \sum_{k=1}^{K} a_k X_k\right] = \sum_{k=1}^{K} a_k^2 \mathrm{Var}[X_k]$.

Analogous rules hold for sample means and sample variances. For example, if $Z_i = a + bX_i$, then

- $\bar{Z}_N = a + b\bar{X}_N$.
- $\hat{\sigma}^2_{Z,N} = b^2 \hat{\sigma}_{X,N}$,

where $\bar{X}_N$, $\bar{Z}_N$ and $\hat{\sigma}^2_{Z,N}$, $\hat{\sigma}_{X,N}$ are the sample means and variances of Z_i and X_i, respectively.

Important Result for Sample Mean of i.i.d. Sequence

Thm 9: If $\bar{X}_N$ is the sample mean of an i.i.d. sequence with
$\mu \equiv \mathbb{E}[X_i]$ and $\sigma^2 \equiv \text{Var}(X_i)$, then
$\mathbb{E}[\bar{X}_N] = \mu$ and $\text{Var}[\bar{X}_N] = \frac{\sigma^2}{N}$.

Self-Study Questions

1. Suppose r_C is the return on an asset with $\Pr[r_C = 0] = \Pr[r_C = .4] = \frac{1}{2}$. Show that $\mathbb{E}[r_C] = 0.2$ and $\text{Var}(r_C) = 0.04$.
2. Suppose $r_f = 0.02$. What is the Sharpe Ratio for r_C?
3. Suppose asset C is a mutual fund with a 0.2% management fee, so that the net return on the asset (net of the management fee) is $r_C - 0.002$. What is the expected value and variance of the net return?
4. Following the proof of Theorem 3, show that $\text{Cov}(X, Y)$ can be rewritten as $\text{Cov}(X, Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot E[Y]$.
5. Prove the assertion in Remark 7, that, for $\bar{X}_N$ the sample mean of an i.i.d. sequence and $T_N = \frac{\bar{X}_N - \mu}{\sigma/\sqrt{N}}$, we have $\mathbb{E}[T_N] = 0$ and $\text{Var}[T_N] = 1$.