Introduction to Econometrics

Review Lecture

Edward Vytlacil

Yale University, Department of Economics

June 20, 2023

Overview for Today

- **Agenda**
  1. Review Expectation, Rules for Expectations
  2. Review Variance, Rules for Variance
- **Application**
  - Asset Diversification
- **R Code**
  - Code used in this lecture (.Rmd, .html, .pdf)

Random Variables – Notation

- We denote random variables with capital letters, like X or Y.
  - We will say things like: “*Let X denote the outcome of rolling a fair die*”

Random Variables – Notation

- We denote random variables with capital letters, like X or Y.
  - We will say things like: "*Let X denote the outcome of rolling a fair die*"

- For discrete random variables, we will enumerate the set of possible realizations of X as $\{x_1, x_2, \ldots, x_K\}$.
  - For example, if X is the outcome of rolling a 6-sided die, then the set of possible realization is $\{1, 2, 3, 4, 5, 6\}$.

Random Variables – Notation

- We denote random variables with capital letters, like X or Y.
  - We will say things like: "*Let X denote the outcome of rolling a fair die*"

- For discrete random variables, we will enumerate the set of possible realizations of X as $\{x_1, x_2, \ldots, x_K\}$.
  - For example, if X is the outcome of rolling a 6-sided die, then the set of possible realization is $\{1, 2, 3, 4, 5, 6\}$.

- For ease of exposition in this course, we will typically express results mathematically for discrete random variables, though our results will generally also hold for continuous random variables.

Characterizing Central Tendency – Expected Value (Review)

- The **expected value** of a random variable is the probability-weighted average of its outcomes.

$$\mathbb{E}[X] \equiv \mu_X = \sum_{k=1}^{K} x_k \times \Pr\{X = x_k\}.$$

Characterizing Central Tendency – Expected Value (Review)

- The **expected value** of a random variable is the probability-weighted average of its outcomes.

$$\mathbb{E}[X] \equiv \mu_X = \sum_{k=1}^{K} x_k \times \Pr\{X = x_k\}.$$

  - For example, if X denotes the outcome of rolling a fair die, then

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=1}^{6} k \times \frac{1}{6} \\ &= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} \\ &= 3.5.\end{aligned}$$

Expected Value – Example: Dummy Variables

- We say that X is a *dummy variable* (also called *indicator variable*) if it takes the value 1 if some underlying event is true, and equals zero otherwise,

$$X = \begin{cases} 1 & \text{if event occurs,} \\ 0 & \text{if event does not occur.} \end{cases}$$

Expected Value – Example: Dummy Variables

- We say that X is a *dummy variable* (also called *indicator variable*) if it takes the value 1 if some underlying event is true, and equals zero otherwise,

$$X = \begin{cases} 1 & \text{if event occurs,} \\ 0 & \text{if event does not occur.} \end{cases}$$

- Since X takes values in $\{0, 1\}$,

$$\begin{aligned} \mathbb{E}[X] &= \Pr\{X = 0\} \cdot 0 + \Pr\{X = 1\} \cdot 1 \\ &= \Pr\{X = 1\} \end{aligned}$$

where $\Pr\{X = 1\} = \Pr\{\text{event occurs}\}$.

Expected Value – Example: Dummy Variables

For example, let X denote the outcome of rolling a fair die, and . . .

- Let Y denote a dummy variable for rolling an ace (a one):

$$Y = \begin{cases} 1 & \text{if } X = 1, \\ 0 & \text{if } X \neq 1. \end{cases}$$

then $\mathbb{E}[Y] = \Pr\{X = 1\} = 1/6$.

Expected Value – Example: Dummy Variables

For example, let X denote the outcome of rolling a fair die, and ...

- Let Y denote a dummy variable for rolling an ace (a one):

$$Y = \begin{cases} 1 & \text{if } X = 1, \\ 0 & \text{if } X \neq 1. \end{cases}$$

then $\mathbb{E}[Y] = \Pr\{X = 1\} = 1/6$.

- Let Z denote a dummy variable for rolling an even number:

$$Z = \begin{cases} 1 & \text{if } X = 2, 4 \text{ or } 6, \\ 0 & \text{if } X = 1, 3 \text{ or } 5. \end{cases}$$

then $\mathbb{E}[Z] = \Pr\{X \text{ even }\} = 1/2$.

A Useful Rule for Expected Values – Linearity

- For any two random variables X and Y, and constants a, b, and c,

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

A Useful Rule for Expected Values – Linearity

- For any two random variables X and Y, and constants a, b, and c,

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

- Note that this implies the following special cases:

$$\begin{aligned} \mathbb{E}[a + X] &= a + \mathbb{E}[X] \\ \mathbb{E}[b \cdot X] &= b \cdot \mathbb{E}[X] \\ \mathbb{E}[X + Y] &= \mathbb{E}[X] + \mathbb{E}[Y]. \end{aligned}$$

Central Tendency and Dispersion

- We typically care about dispersion in addition to caring about central tendency.
- Consider the following scenario:
  - You are going to retire in one month, you have $1 million in retirement savings, and need to select between two investments for those savings for the next month:

Central Tendency and Dispersion

- We typically care about dispersion in addition to caring about central tendency.
- Consider the following scenario:
  - You are going to retire in one month, you have $1 million in retirement savings, and need to select between two investments for those savings for the next month:
    1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have $1 million dollars plus a $1,500 interest payment.

Central Tendency and Dispersion

- We typically care about dispersion in addition to caring about central tendency.
- Consider the following scenario:
  - You are going to retire in one month, you have $1 million in retirement savings, and need to select between two investments for those savings for the next month:
    1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have $1 million dollars plus a $1,500 interest payment.
    2. Second option: invest full $1 million in a startup. With probability $1/2$ the startup will go bankrupt and you will lose the full investment, but with probability $1/2$ the startup will succeed in which case you will have $2,005,000 at the end of the month.

Central Tendency and Dispersion

- We typically care about dispersion in addition to caring about central tendency.
- Consider the following scenario:
  - You are going to retire in one month, you have \$1 million in retirement savings, and need to select between two investments for those savings for the next month:
    1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have \$1 million dollars plus a \$1,500 interest payment.
    2. Second option: invest full \$1 million in a startup. With probability $1/2$ the startup will go bankrupt and you will lose the full investment, but with probability $1/2$ the startup will succeed in which case you will have \$2,005,000 at the end of the month.
  - **What is the expected value of each?**

Central Tendency and Dispersion

- We typically care about dispersion in addition to caring about central tendency.
- Consider the following scenario:
  - You are going to retire in one month, you have $1 million in retirement savings, and need to select between two investments for those savings for the next month:
    1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have $1 million dollars plus a $1,500 interest payment.
    2. Second option: invest full $1 million in a startup. With probability $1/2$ the startup will go bankrupt and you will lose the full investment, but with probability $1/2$ the startup will succeed in which case you will have $2,005,000 at the end of the month.
  - **What is the expected value of each?**
  - **Which option would you prefer?**

Characterizing Dispersion – Variance

- The **variance** of a random variable is the expected squared deviation from it's expected value:

$$
\begin{aligned}
\text{Var}[X] \equiv \sigma_X^2 &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\
&= \sum_{i=1}^{K} (x_i - \mathbb{E}[X])^2 Pr\{X = x_i\}.
\end{aligned}
$$

Characterizing Dispersion – Variance

- The **variance** of a random variable is the expected squared deviation from it's expected value:

$$\begin{aligned} \mathrm{Var}[X] \equiv \sigma_X^2 &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \sum_{i=1}^{K} (x_i - \mathbb{E}[X])^2 Pr\{X = x_i\}. \end{aligned}$$

- The **standard deviation** of a random variable is the square root of it's variance:

$$sd(X) \equiv \sigma_x = \sqrt{Var(X)}.$$

Central Tendency and Dispersion

- You are going to retire in one month, you have \$1 million in retirement savings, and need to select between two investments for those savings for the next month:

    1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have \$1 million dollars plus a \$1,500 interest payment.
    2. Second option: invest full \$1 million in a startup. With probability $1/2$ the startup will go bankrupt and you will lose the full investment, but with probability $1/2$ the startup will succeed in which case you will have \$2,005,000 at the end of the month.

Central Tendency and Dispersion

- You are going to retire in one month, you have \$1 million in retirement savings, and need to select between two investments for those savings for the next month:

  1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have \$1 million dollars plus a \$1,500 interest payment.

  2. Second option: invest full \$1 million in a startup. With probability $1/2$ the startup will go bankrupt and you will lose the full investment, but with probability $1/2$ the startup will succeed in which case you will have \$2,005,000 at the end of the month.

- **What is the variance of each?**

Central Tendency and Dispersion

- You are going to retire in one month, you have \$1 million in retirement savings, and need to select between two investments for those savings for the next month:

    1. First option: invest in 30 day U.S. treasury bills, in which case at the end of the month you will with certainty have \$1 million dollars plus a \$1,500 interest payment.
    2. Second option: invest full \$1 million in a startup. With probability $1/2$ the startup will go bankrupt and you will lose the full investment, but with probability $1/2$ the startup will succeed in which case you will have \$2,005,000 at the end of the month.

- **What is the variance of each?**
- **Typically we want higher expected return but lower risk (variance), typically a tradeoff.**

Application – Asset Diversification

- Let r denote the rate of return on a security, defined as:

$$r = \frac{p_1 + d - p_0}{p_0},$$

where

$$\begin{aligned} p_1 &= \text{price of security at end of time period,} \\ d &= \text{dividends (if any) paid during time period,} \\ p_0 &= \text{price of security at beginning of time period.} \end{aligned}$$

Application – Asset Diversification

- Let r denote the rate of return on a security, defined as:

$$r = \frac{p_1 + d - p_0}{p_0},$$

where

$$\begin{aligned} p_1 &= \text{price of security at end of time period,} \\ d &= \text{dividends (if any) paid during time period,} \\ p_0 &= \text{price of security at beginning of time period.} \end{aligned}$$

- Investors like high expected returns, but dislike risk.
  - Risk typically quantified by the variance of the return.
  - Risk-Return tradeoff.

Asset Diversification – Expected Returns and Risks

- Consider two Securities, Ford and Tesla stock, with returns r_F and r_T:

$$r_F = \begin{cases} .20 & \text{with probability } .5 \\ -.10 & \text{with probability } .5. \end{cases}$$

$$r_T = \begin{cases} .60 & \text{with probability } .5 \\ -.40 & \text{with probability } .5. \end{cases}$$

Asset Diversification – Expected Returns and Risks

- Consider two Securities, Ford and Tesla stock, with returns r_F and r_T:

$$r_F = \begin{cases} .20 & \text{with probability } .5 \\ -.10 & \text{with probability } .5. \end{cases} \Rightarrow \begin{array}{lcl} \mathbb{E}[r_F] & = & .05, \\ Var[r_F] & = & .0225 \end{array}$$

$$r_T = \begin{cases} .60 & \text{with probability } .5 \\ -.40 & \text{with probability } .5. \end{cases}$$

Asset Diversification – Expected Returns and Risks

- Consider two Securities, Ford and Tesla stock, with returns r_F and r_T:

$$r_F = \begin{cases} .20 & \text{with probability } .5 \\ -.10 & \text{with probability } .5. \end{cases} \Rightarrow \begin{aligned} \mathbb{E}[r_F] &= .05, \\ Var[r_F] &= .0225 \end{aligned}$$

$$r_T = \begin{cases} .60 & \text{with probability } .5 \\ -.40 & \text{with probability } .5. \end{cases} \Rightarrow \begin{aligned} \mathbb{E}[r_T] &= .10, \\ Var[r_T] &= .25 \end{aligned}$$

Asset Diversification – Expected Returns and Risks

- Consider two Securities, Ford and Tesla stock, with returns r_F and r_T:

$$r_F = \begin{cases} .20 & \text{with probability .5} \\ -.10 & \text{with probability .5.} \end{cases} \Rightarrow \begin{array}{lcl} \mathbb{E}[r_F] & = & .05, \\ Var[r_F] & = & .0225 \end{array}$$

$$r_T = \begin{cases} .60 & \text{with probability .5} \\ -.40 & \text{with probability .5.} \end{cases} \Rightarrow \begin{array}{lcl} \mathbb{E}[r_T] & = & .10, \\ Var[r_T] & = & .25 \end{array}$$

Which has higher expected return?

Which has higher risk?

Asset Diversification – Expected Returns and Risks

- Consider two Securities, Ford and Tesla stock, with returns r_F and r_T:

$$r_F = \begin{cases} .20 & \text{with probability } .5 \\ -.10 & \text{with probability } .5. \end{cases} \Rightarrow \begin{aligned} \mathbb{E}[r_F] &= .05, \\ Var[r_F] &= .0225 \end{aligned}$$

$$r_T = \begin{cases} .60 & \text{with probability } .5 \\ -.40 & \text{with probability } .5. \end{cases} \Rightarrow \begin{aligned} \mathbb{E}[r_T] &= .10, \\ Var[r_T] &= .25 \end{aligned}$$

Which has higher expected return?

Which has higher risk?

Which asset is the better investment?

Asset Diversification – Expected Returns and Risks

- Consider two Securities, Ford and Tesla stock, with returns r_F and r_T:

$$r_F = \begin{cases} .20 & \text{with probability } .5 \\ -.10 & \text{with probability } .5. \end{cases} \Rightarrow \begin{array}{rcl} \mathbb{E}[r_F] & = & .05, \\ Var[r_F] & = & .0225 \end{array}$$

$$r_T = \begin{cases} .60 & \text{with probability } .5 \\ -.40 & \text{with probability } .5. \end{cases} \Rightarrow \begin{array}{rcl} \mathbb{E}[r_T] & = & .10, \\ Var[r_T] & = & .25 \end{array}$$

Which has higher expected return?

Which has higher risk?

Which asset is the better investment?

What if you could hold both securities in a portfolio?

Asset Diversification – Risk and Return Tradeoff

```
> library(ggplot2)
>
> Company <- c("Tesla", "Ford")
> Er      <- c(0.1,0.05)
> Var     <- c(0.25,0.0225)
>
> df <- data.frame(Company,Er,Var)
>
> ggplot(df, aes(x = Var, y = Er, color = Company))+
+     geom_point(size = 5) +
+     theme_bw() + ggtitle("Risk-Return Tradeoff") +
+     xlab("Variance") + ylab("Expected Returns")
```

Asset Diversification – Risk and Return Tradeoff

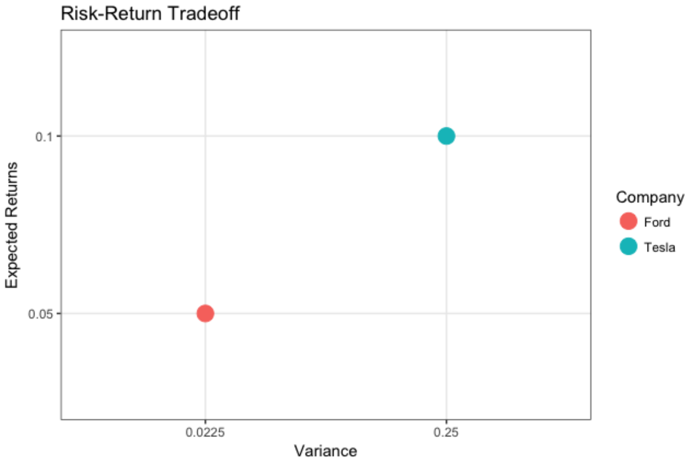


Figure 1.1: Two Risky Assets in the Variance-Expected Return Space

Asset Diversification – Expected Portfolio Return and Risk

- Let w_T and w_F denote the fraction of funds invested in Tesla and Ford.

- Return on portfolio:

$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$

Asset Diversification – Expected Portfolio Return and Risk

- Let w_T and w_F denote the fraction of funds invested in Tesla and Ford.
- Return on portfolio:

$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$

- Expected return on portfolio:

$$E[r_p] = w_F \cdot \mathbb{E}[r_F] + w_T \cdot \mathbb{E}[r_T].$$

Asset Diversification – Expected Portfolio Return and Risk

- Let w_T and w_F denote the fraction of funds invested in Tesla and Ford.
- Return on portfolio:
$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$
- Expected return on portfolio:
$$E[r_p] = w_F \cdot \mathbb{E}[r_F] + w_T \cdot \mathbb{E}[r_T].$$

What is the risk of the holding the portfolio?

How does the risk of holding the portfolio compare to holding just Ford or Tesla?

Asset Diversification – Expected Portfolio Return and Risk

- Let w_T and w_F denote the fraction of funds invested in Tesla and Ford.
- Return on portfolio:
$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$
- Expected return on portfolio:
$$E[r_p] = w_F \cdot \mathbb{E}[r_F] + w_T \cdot \mathbb{E}[r_T].$$

What is the risk of the holding the portfolio?

How does the risk of holding the portfolio compare to holding just Ford or Tesla?

Depends on dependence between r_F and r_T, how they ***covary****.*

Characterizing Dependence – Covariance

- The **covariance** of two random variables X and Y is the expected value of the product of their deviations from their individual expected values.

$$\begin{aligned}\text{Cov}(X, Y) \equiv \sigma_{XY} &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \sum_{k=1}^{K} \sum_{j=1}^{J} (x_k - \mathbb{E}[X]) \cdot (y_j - \mathbb{E}[Y]) \Pr\{X = x_k, Y = y_j\}.\end{aligned}$$

Characterizing Dependence – Covariance

- The **covariance** of two random variables X and Y is the expected value of the product of their deviations from their individual expected values.

$$\begin{aligned}\text{Cov}(X, Y) \equiv \sigma_{XY} &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \sum_{k=1}^{K}\sum_{j=1}^{J}(x_k - \mathbb{E}[X]) \cdot (y_j - \mathbb{E}[Y]) \Pr\{X = x_k, Y = y_j\}.\end{aligned}$$

- Covariance is a measure of linear dependence.

Characterizing Dependence – Covariance

- The **covariance** of two random variables X and Y is the expected value of the product of their deviations from their individual expected values.

$$\begin{aligned}\text{Cov}(X, Y) \equiv \sigma_{XY} &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \sum_{k=1}^{K}\sum_{j=1}^{J}(x_k - \mathbb{E}[X]) \cdot (y_j - \mathbb{E}[Y]) \Pr\{X = x_k, Y = y_j\}.\end{aligned}$$

- Covariance is a measure of linear dependence.
- $\text{Cov}(X, X) = \text{Var}(X)$.

Example, Covariance

- Consider returns to Morgan Stanley (MS) and Genworth Financial (GNW) shares.

```
> library(readstata13)  #need this library for read.dta13
    function, data set is in STATA format.
>
> data <- read.dta13("https://edward-vytlacil.github.io/Data
    /financeR.dta")
>
> cov(data$r_B,data$r_C)
>
> ggplot(data,aes(x=data$r_B,y=data$r_C))+
+     geom_point()+
+     xlab("Morgan Stanley Returns in Dollars")+
+     ylab("Genworth Financial Returns in Dollars")
```

Example, Covariance

- Consider returns to Morgan Stanley (MS) and Genworth Financial (GNW) shares.

```
> library(readstata13)  #need this library for read.dta13
    function, data set is in STATA format.
>
> data <- read.dta13("https://edward-vytlacil.github.io/Data
    /financeR.dta")
>
> cov(data$r_B,data$r_C)
>
> ggplot(data,aes(x=data$r_B,y=data$r_C))+
+       geom_point()+
+       xlab("Morgan Stanley Returns in Dollars")+
+       ylab("Genworth Financial Returns in Dollars")
```

- which gives:

```
> cov(data$r_B,data$r_C)
[1] 0.01325161
```

Scatter Plot, Morgan Stanley and Genworth Financial Returns in Dollars

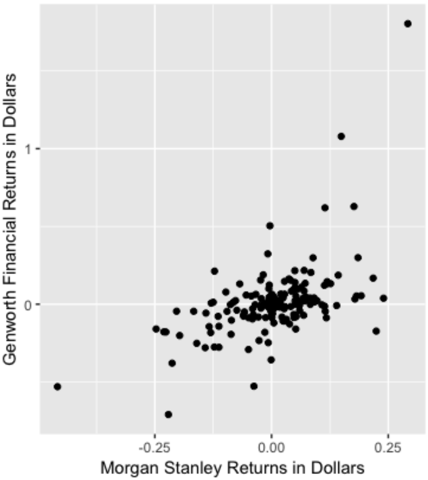


Figure 1.2: Covariance is 0.01325

Characterizing Random Variables – Correlation

Related Concept: Correlation

- The **correlation** of two random variables X and Y is a scaled version of their covariance:

$$\mathrm{Corr}(X, Y) \equiv \rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}.$$

Characterizing Random Variables – Correlation

Related Concept: Correlation

- The **correlation** of two random variables X and Y is a scaled version of their covariance:

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

```
> cor(data$r_B,data$r_C)
[1] 0.5437336
```

Characterizing Random Variables – Correlation

Related Concept: Correlation

- The **correlation** of two random variables X and Y is a scaled version of their covariance:

$$\mathrm{Corr}(X,Y) \equiv \rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}}.$$

```
> cor(data$r_B,data$r_C)
[1] 0.5437336
```

- Why should we not be surprised that the returns to Morgan Stanley (a financial firm) and Genworth Financial are strongly positively correlated?
- When would we expect the returns to two assets to be highly correlated? Negatively correlated?

Some useful Rules for (Co)Variances

- Variance of a Sum of Random Variables:

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

proof

Some useful Rules for (Co)Variances

- Variance of a Sum of Random Variables:

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

proof

- Covariance of a linear function of X and Y:

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

Some useful Rules for (Co)Variances

- Variance of a Sum of Random Variables:

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

proof

- Covariance of a linear function of X and Y:

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

- Can show that:

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

Some useful Rules for (Co)Variances

$$\begin{aligned}
\mathrm{Var}\big(a + bX + cY\big) &= b^2 \cdot \mathrm{Var}\big(X\big) + c^2 \cdot \mathrm{Var}\big(Y\big) + 2 \cdot b \cdot c \cdot \mathrm{Cov}\big(X, Y\big), \\
\mathrm{Cov}\big(a + bX, c + dY\big) &= b \cdot d \cdot \mathrm{Cov}\big(X, Y\big), \\
\mathrm{Var}\big(X\big) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2. \\
\mathrm{Cov}\big(X, Y\big) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

- These rules are for population variance, covariance.
- However, parallel rules also hold for sample variance, covariance.

ASSET DIVERSIFICATION

Asset Diversification – Portfolio Return and Risk (cont'd)

- Return on portfolio:

$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$

Asset Diversification – Portfolio Return and Risk (cont'd)

- Return on portfolio:

$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$

- Expected return on portfolio:

$$\mathbb{E}[r_p] = w_F \cdot \mathbb{E}[r_F] + w_T \cdot \mathbb{E}[r_T].$$

Asset Diversification – Portfolio Return and Risk (cont'd)

- Return on portfolio:

$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$

- Expected return on portfolio:

$$\mathbb{E}[r_p] = w_F \cdot \mathbb{E}[r_F] + w_T \cdot \mathbb{E}[r_T].$$

- Variance of return on portfolio:

$$Var(r_p) = w_F^2 \cdot Var(r_F) + w_T^2 \cdot Var(r_T) + 2 \cdot w_F \cdot w_T \cdot Cov(r_F, r_T).$$

Asset Diversification – Portfolio Return and Risk (cont'd)

- Return on portfolio:

$$r_p = w_F \cdot r_F + w_T \cdot r_T.$$

- Expected return on portfolio:

$$\mathbb{E}[r_p] = w_F \cdot \mathbb{E}[r_F] + w_T \cdot \mathbb{E}[r_T].$$

- Variance of return on portfolio:

$$Var(r_p) = w_F^2 \cdot Var(r_F) + w_T^2 \cdot Var(r_T) + 2 \cdot w_F \cdot w_T \cdot Cov(r_F, r_T).$$

Need joint distribution of (r_F, r_T) to calculate $Cov(r_F, r_T)$!

Asset Diversification – Portfolio Return and Risk (cont'd)

- We will calculate $Cov(r_F, r_T)$ under two scenarios:

  1. Main uncertainty is total demand for cars, implying strong positive correlation in returns for Ford and Tesla.
  2. Main uncertainty is nature of demand for cars, whether electric cars largely replace conventional cars, implying strong negative correlation in returns for Ford and Tesla.

First Scenario: Positive Return Covariance

- Suppose Ford's and Tesla's performance largely driven by overall demand for cars, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	0.4	0.1
Tesla -.4	0.1	0.4

First Scenario: Positive Return Covariance

- Suppose Ford's and Tesla's performance largely driven by overall demand for cars, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	0.4	0.1
Tesla -.4	0.1	0.4

$$Cov(r_F, r_T) = \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])]$$

First Scenario: Positive Return Covariance

- Suppose Ford's and Tesla's performance largely driven by overall demand for cars, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	0.4	0.1
Tesla -.4	0.1	0.4

$$
\begin{aligned}
Cov(r_F, r_T) &= \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])] \\
&= 0.4 \cdot (.2 - .05) \cdot (.6 - .1) + 0.1 \cdot (-.1 - .05) \cdot (.6 - .1) \\
&+ 0.1 \cdot (.2 - .05) \cdot (-.4 - .1) + .4 \cdot (-.1 - .05) \cdot (-.4 - .1)
\end{aligned}
$$

First Scenario: Positive Return Covariance

- Suppose Ford's and Tesla's performance largely driven by overall demand for cars, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	0.4	0.1
Tesla -.4	0.1	0.4

$$\begin{aligned} Cov(r_F, r_T) &= \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])] \\ &= 0.4 \cdot (.2 - .05) \cdot (.6 - .1) + 0.1 \cdot (-.1 - .05) \cdot (.6 - .1) \\ &+ 0.1 \cdot (.2 - .05) \cdot (-.4 - .1) + .4 \cdot (-.1 - .05) \cdot (-.4 - .1) \\ &= 0.045. \end{aligned}$$

First Scenario: Positive Return Covariance

- Suppose Ford's and Tesla's performance largely driven by overall demand for cars, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	0.4	0.1
Tesla -.4	0.1	0.4

$$\begin{aligned} Cov(r_F, r_T) &= \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])] \\ &= 0.4 \cdot (.2 - .05) \cdot (.6 - .1) + 0.1 \cdot (-.1 - .05) \cdot (.6 - .1) \\ &+ 0.1 \cdot (.2 - .05) \cdot (-.4 - .1) + .4 \cdot (-.1 - .05) \cdot (-.4 - .1) \\ &= 0.045. \end{aligned}$$

$$\Rightarrow Corr(r_F, r_T) = .6.$$

First Scenario: Positive Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

First Scenario: Positive Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

$$\mathbb{E}[r_p] = .5 \cdot \mathbb{E}[r_F] + .5 \cdot \mathbb{E}[r_T] = .5 \cdot .05 + .5 \cdot .1 = .075.$$

- Expected return on portfolio half way between expected return on Tesla and expected return on Ford.

First Scenario: Positive Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

$$\mathbb{E}[r_p] = .5 \cdot \mathbb{E}[r_F] + .5 \cdot \mathbb{E}[r_T] = .5 \cdot .05 + .5 \cdot .1 = .075.$$

$$\begin{aligned} Var[r_p] &= .5^2 \cdot Var[r_F] + .5^2 \cdot Var[r_T] + 2 \cdot .5 \cdot .5 \cdot Cov(r_F, r_T) \\ &= .5^2 \cdot .0225 + .5^2 \cdot .25 + 2 \cdot .5^2 \cdot .045 \\ &= .09. \end{aligned}$$

- Expected return on portfolio half way between expected return on Tesla and expected return on Ford.
- Variance of portfolio return inbetween that on Ford and that on Tesla but not the average.

First Scenario: Positive Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

$$\mathbb{E}[r_p] = .5 \cdot \mathbb{E}[r_F] + .5 \cdot \mathbb{E}[r_T] = .5 \cdot .05 + .5 \cdot .1 = .075.$$

$$\begin{aligned} Var[r_p] &= .5^2 \cdot Var[r_F] + .5^2 \cdot Var[r_T] + 2 \cdot .5 \cdot .5 \cdot Cov(r_F, r_T) \\ &= .5^2 \cdot .0225 + .5^2 \cdot .25 + 2 \cdot .5^2 \cdot .045 \\ &= .09. \end{aligned}$$

- Expected return on portfolio half way between expected return on Tesla and expected return on Ford.
- Variance of portfolio return inbetween that on Ford and that on Tesla but not the average.
- How to evaluate risk-return tradeoff for portfolio vs. only Ford or Tesla?

First Scenario: Positive Return Covariance (cont'd)

```
> er_F    <- 0.05   # Expected Returns
> er_T    <- 0.1
> var_F   <- 0.0225 # Variances/Risk
> var_T   <- 0.25
> cov_FT  <- 0.045  # Covariance
>
> weights <- seq(from = 0, to = 1, length.out = 1000)
> tab <- data.frame(wF = weights, wT = 1 - weights)
> tab$er_p  <-  tab$wF * er_F + tab$wT * er_T
> tab$var_p <-  tab$wF^2*var_F +tab$wT^2*var_T +2* tab$wF *(1 -
    tab$wF)*cov_FT
>
> ggplot() + geom_point(data = tab,
+                       aes(x = var_p, y = er_p, color = wF)) +
+     ggtitle("Possible Portfolios with Positive Covariance") +
+     xlab("Volatility") +
+     ylab("Expected Returns")
```

First Scenario: Positive Return Covariance (cont'd)

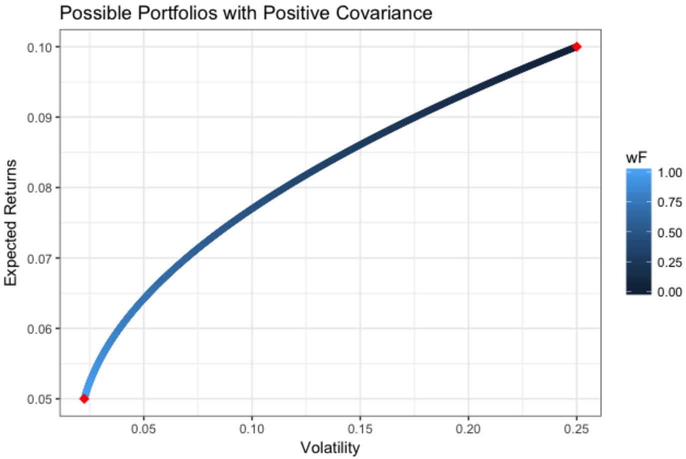


Figure 1.3: Portfolios of Two Positively Correlated Risky Assets

Second Scenario: Negative Return Covariance

- Suppose Ford's and Tesla's largely driven by whether electric cars become dominant in market place or demand for conventional cars continues to dominate, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	.1	0.4
Tesla -.4	0.4	0.1

Second Scenario: Negative Return Covariance

- Suppose Ford's and Tesla's largely driven by whether electric cars become dominant in market place or demand for conventional cars continues to dominate, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	.1	0.4
Tesla -.4	0.4	0.1

$$Cov(r_F, r_T) = \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])]$$

Second Scenario: Negative Return Covariance

- Suppose Ford's and Tesla's largely driven by whether electric cars become dominant in market place or demand for conventional cars continues to dominate, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	.1	0.4
Tesla -.4	0.4	0.1

$$
\begin{aligned}
Cov(r_F, r_T) &= \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])] \\
&= 0.1 \cdot (.2 - .05) \cdot (.6 - .1) + 0.4 \cdot (-.1 - .05) \cdot (.6 - .1) \\
&+ 0.4 \cdot (.2 - .05) \cdot (-.4 - .1) + .1 \cdot (-.1 - .05) \cdot (-.4 - .1)
\end{aligned}
$$

Second Scenario: Negative Return Covariance

- Suppose Ford's and Tesla's largely driven by whether electric cars become dominant in market place or demand for conventional cars continues to dominate, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	.1	0.4
Tesla -.4	0.4	0.1

$$
\begin{aligned}
Cov(r_F, r_T) &= \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])] \\
&= 0.1 \cdot (.2 - .05) \cdot (.6 - .1) + 0.4 \cdot (-.1 - .05) \cdot (.6 - .1) \\
&+ 0.4 \cdot (.2 - .05) \cdot (-.4 - .1) + .1 \cdot (-.1 - .05) \cdot (-.4 - .1) \\
&= -0.045.
\end{aligned}
$$

Second Scenario: Negative Return Covariance

- Suppose Ford's and Tesla's largely driven by whether electric cars become dominant in market place or demand for conventional cars continues to dominate, with the probabilities of the four possible outcomes given by:

Returns	Ford .2	Ford -.1
Tesla .6	.1	0.4
Tesla -.4	0.4	0.1

$$
\begin{aligned}
Cov(r_F, r_T) &= \mathbb{E}[(r_F - \mathbb{E}[r_F])(r_T - \mathbb{E}[r_T])] \\
&= 0.1 \cdot (.2 - .05) \cdot (.6 - .1) + 0.4 \cdot (-.1 - .05) \cdot (.6 - .1) \\
&+ 0.4 \cdot (.2 - .05) \cdot (-.4 - .1) + .1 \cdot (-.1 - .05) \cdot (-.4 - .1) \\
&= -0.045.
\end{aligned}
$$

$$
Corr(r_F, r_T) = -.6.
$$

Second Scenario: Negative Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

Second Scenario: Negative Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

$$\mathbb{E}[r_p] = .5 \cdot \mathbb{E}[r_F] + .5 \cdot \mathbb{E}[r_T] = .5 \cdot .05 + .5 \cdot .1 = .075.$$

- Expected return on portfolio half way between expected return on Tesla and expected return on Ford.

Second Scenario: Negative Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

$$\mathbb{E}[r_p] = .5 \cdot \mathbb{E}[r_F] + .5 \cdot \mathbb{E}[r_T] = .5 \cdot .05 + .5 \cdot .1 = .075.$$

$$\begin{aligned} Var[r_p] &= .5^2 \cdot Var[r_F] + .5^2 \cdot Var[r_T] + 2 \cdot .5 \cdot .5 \cdot Cov(r_F, r_T) \\ &= .5^2 \cdot .0225 + .5^2 \cdot .25 + 2 \cdot .5^2 \cdot (-.045) \\ &= .046. \end{aligned}$$

- Expected return on portfolio half way between expected return on Tesla and expected return on Ford.
- Recall variance on portfolio was .09 in first scenario, versus .046 here in second scenario. Explanation?

Second Scenario: Negative Return Covariance (cont'd)

- Consider portfolio with equal investment in Ford and Tesla, $w_F = w_T = 1/2$.

$$\mathbb{E}[r_p] = .5 \cdot \mathbb{E}[r_F] + .5 \cdot \mathbb{E}[r_T] = .5 \cdot .05 + .5 \cdot .1 = .075.$$

$$\begin{aligned} Var[r_p] &= .5^2 \cdot Var[r_F] + .5^2 \cdot Var[r_T] + 2 \cdot .5 \cdot .5 \cdot Cov(r_F, r_T) \\ &= .5^2 \cdot .0225 + .5^2 \cdot .25 + 2 \cdot .5^2 \cdot (-.045) \\ &= .046. \end{aligned}$$

- Expected return on portfolio half way between expected return on Tesla and expected return on Ford.
- Recall variance on portfolio was .09 in first scenario, versus .046 here in second scenario. Explanation?
- How to evaluate risk-return tradeoff for portfolio vs. only Ford or Tesla?

Second Scenario: Negative Return Covariance (cont'd)

- Consider portfolio investing mostly in Ford but partially in Tesla, $w_F = .9, w_T = .1$.

$$\mathbb{E}[r_p] = .9 \cdot \mathbb{E}[r_F] + .1 \cdot \mathbb{E}[r_T] = .9 \cdot .05 + .1 \cdot .1 = .055.$$

$$\begin{aligned} Var[r_p] &= .9^2 \cdot Var[r_F] + .1^2 \cdot Var[r_T] + 2 \cdot .9 \cdot .1 \cdot Cov(r_F, r_T) \\ &= .9^2 \cdot .0225 + .1^2 \cdot .25 + 2 \cdot .9 \cdot .1 \cdot (-.045) \\ &= .013. \end{aligned}$$

- In this scenario, would an investor ever wish to invest 100% in Ford?

Second Scenario: Negative Return Covariance (cont'd)

```
library(ggplot2)
> er_F   <- 0.05   # Expected Returns
> er_T   <- 0.1
> var_F  <- 0.0225 # Variances/Risk
> var_T  <- 0.25
> cov_FT <- - 0.045  # Covariance
>
> weights <- seq(from = 0, to = 1, length.out = 1000)
> tab <- data.frame(wF = weights, wT = 1 - weights)
> tab$er_p  <-  tab$wF * er_F + tab$wT * er_T
> tab$var_p <-  tab$wF^2*var_F +tab$wT^2*var_T +2* tab$wF *(1 -
    tab$wF)*cov_FT
>
> ggplot() + geom_point(data = tab,
+                       aes(x = var_p, y = er_p, color = wF)) +
+     theme_bw() +
+     ggtitle("Possible Portfolios with Negative Covariance") +
+     xlab("Volatility") +
+     ylab("Expected Returns")
```

Second Scenario: Negative Return Covariance (cont'd)

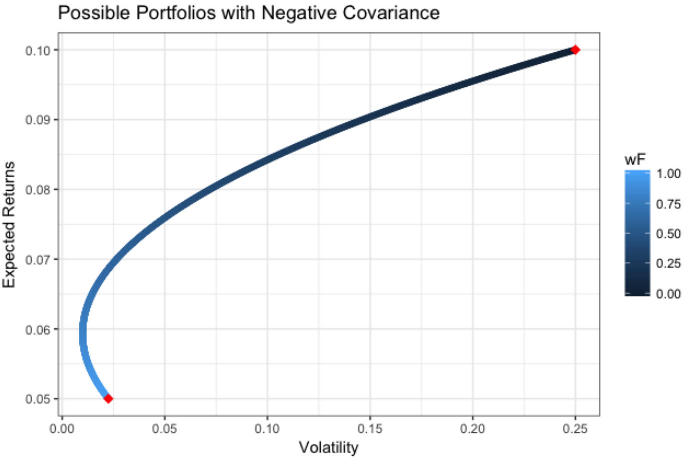
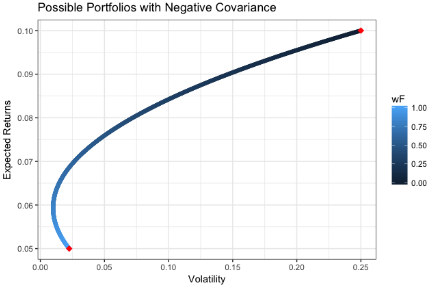
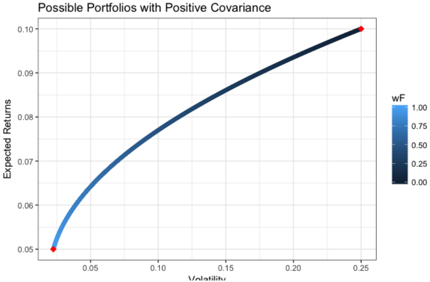


Figure 1.4: Portfolios of Two Negatively Correlated Risky Assets

Comparing the portfolios with pos. vs neg. covariance





Diversification, Insurance

- For an investor initially only holding Ford, adding Telsa to portfolio provides insurance, with the lower the correlation between their returns the better the insurance.
- For an investor initially only holding Tesla, adding Ford to portfolio provides insurance, with the lower the correlation between their returns the better the insurance.

Diversification, Insuranace

- Diversification central issue in finance, institutional investing.
- Also important for individual investing, e.g.,
  - How much should I invest in stock vs. bonds?
  - How do the equity options my employer gives me affect my investment decision?
- Central issue in private insurance and for public provision of insurance.
- Related: Risk sharing in village economies.

Pricing of Assets

- Above discussion has implication for pricing of assets.
- Value of an asset to an investor depends not just on expected return and volatility of return, but on how return covaries with returns of other assets (market return).
- Investors willing to have lower expected return for asset that moves less with market returns.
- Investors require higher expected return to invest in asset whose return is strongly correlated with market return.
- One formal model: Capital Asset Pricing Model.

Summary

- We can characterize distributions of random variables using measures of central tendency such as the expected value, dispersion such as the variance, and linear dependence such as the covariance.
- Dependencies between random variables are of central importance in many economic questions and applications.

Proof: Variance of a Sum of Random Variables

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y).$$

Proof:

$$\begin{aligned}
&Var(aX + bY) \\
&\quad = \mathbb{E}[(aX + bY)^2] - (\mathbb{E}[aX + bY])^2 \\
&\quad = \mathbb{E}[a^2X^2 + b^2Y^2 + 2aXbY] - (\mathbb{E}[aX]^2 + \mathbb{E}[bY]^2 + 2\mathbb{E}[aX]\mathbb{E}[bY]) \\
&\quad = \mathbb{E}[a^2X^2] - \mathbb{E}[aX]^2 + \mathbb{E}[b^2Y^2] - \mathbb{E}[bY]^2 + 2\left(\mathbb{E}[aXbY] - \mathbb{E}[aX]\mathbb{E}[bY]\right) \\
&\quad = a^2\left(\mathbb{E}[X^2] - \mathbb{E}[X]^2\right) + b^2\left(\mathbb{E}[Y^2] - \mathbb{E}[Y]^2\right) + 2ab\left(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\right) \\
&\quad = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y). \qquad \square
\end{aligned}$$

Back