Econ 5023: Statistics for Decision Making

Univariate Statistics (VII): Continuous Variables and Moments

Le Wang

Itinerary:

1. Four Moments and What and how do they characterize?
2. Applications in Finance: Moments and Risks

Again, to characterize the whole distribution, we need infinite order moments!

This is not unlikely! We focus on the first four moments in practice to obtain some useful features of the distribution.

Pearson's System: Based on the first four moments, we obtain

Mean: the central value about which the measurements scatter

Pearson's System: Based on the first four moments, we obtain

- **Mean**: the central value about which the measurements scatter
- **Variance** How far **most** of the measurements scatter about the mean. (Standard Deviation is the square root of the variance.)

Pearson's System: Based on the first four moments, we obtain

- **Mean**: the central value about which the measurements scatter
- **Variance** How far **most** of the measurements scatter about the mean. (Standard Deviation is the square root of the variance.)
- **Skewness** (Symmetry) The degree to which the measurements pile up on only one side of the mean

Pearson's System: Based on the first four moments, we obtain

- **Mean**: the central value about which the measurements scatter
- **Variance** How far **most** of the measurements scatter about the mean. (Standard Deviation is the square root of the variance.)
- **Skewness** (Symmetry) The degree to which the measurements pile up on only one side of the mean
- **Kurtosis** How far **rare measurements** scatter from the mean.

Motivation for Variance:

Suppose that I have the following portfolio options.

```
x1 <- c(-10,0,10)
x2 <- c(-100,0,100)

x1

## [1] -10   0  10

x2

## [1] -100    0  100
```

Question: Which one will you pick?

Variance of x_1: $\frac{(-10-0)^2+(0-0)^2+(10-0)^2}{3-1} = 100$

```r
x1
```

```
## [1] -10   0  10
```

```r
x2
```

```
## [1] -100    0  100
```

```r
mean(x1);mean(x2)
```

```
## [1] 0
## [1] 0
```

```r
median(x1);median(x2)
```

```
## [1] 0
## [1] 0
```

```r
var(x1);var(x2)
```

```
## [1] 100
## [1] 10000
```

Variance is indeed considered to be a good measure of risk!

In economics, what else can it be used to measure?

From Quantopian,

"Although variance and standard deviation tell us how volatile a quantity is, they do not differentiate between deviations upward and deviations downward. Often, such as in the case of returns on an asset, we are more worried about deviations downward."

Semivariance is defined as the part smaller than the mean

$$\frac{\sum_{X_i<\mu}(X_i-\mu)^2}{N<}$$

where $N<$ is the number of observations smaller than the mean.

We can program this by ourselves

$$\frac{\sum_{X_i<\mu}(X_i-\mu)^2}{N<}$$

```
x1 <- rnorm(1000)
sum(x1[x1 < mean(x1)]-mean(x1))^2/length(x1[x1<mean(x1)])
```

```
## [1] 283.1336
```

```
var(x1)
```

```
## [1] 0.9196963
```

An alternative measure of risk is

Interquartile Range (IQR) (the difference between 75^{th} percentile and 25^{th} percentile)

In the previous example, mean is equal to median. Both can be regarded as a "typical" value for forecast or to characterize the general tendencies.

Now, consider the following example

```
x<- c(-100,-10,0,10)
mean(x)

## [1] -25

median(x)

## [1] -5
```

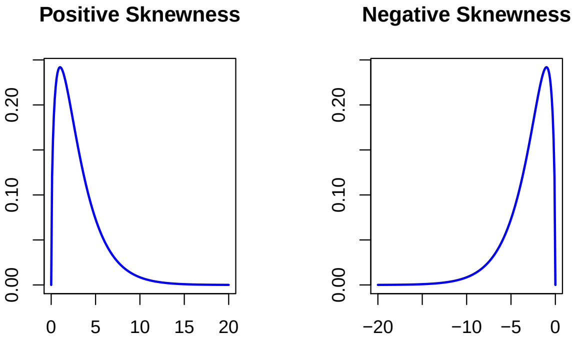
Question: Where (or which region) is the mass concentrated?

Skenewness and Symmetry

Population Skewness $(= \frac{\mathbb{E}[(y-\mu)^3]}{\sigma_y^3})$ measures whether a distribution is **symmetrical**.

Sample skewness is given by $\frac{\frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{y})^3}{\widehat{\sigma}^3}$.

1. **Skewed to the right** > 0 (mean $>$ median)
2. **Skewed to the left** < 0 (mean $<$ median)
3. **Symmetric** $= 0$ (mean $=$ median)



```r
library(moments)
x5<-c(-10,0,10, 100)
x6<-c(-100,-10,0,10)

mean(x5)
```

```
## [1] 25
```

```r
median(x5)
```

```
## [1] 5
```

```r
# Calculate skewness
skewness(x5)
```

```
## [1] 1.065605
```

```r
skewness(x6)
```

```
## [1] -1.065605
```

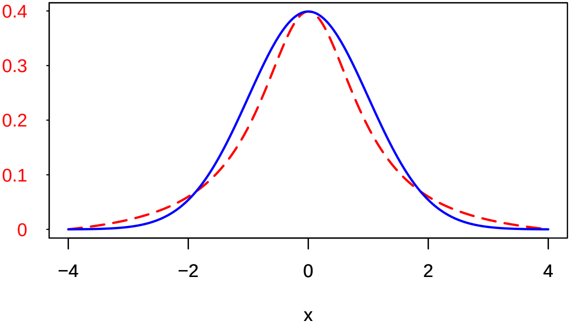
Negative skewness indicates that the mean of the data values is less than the median, and the data distribution is left-skewed.

Positive skewness would indicates that the mean of the data values is larger than the median, and the data distribution is right-skewed.

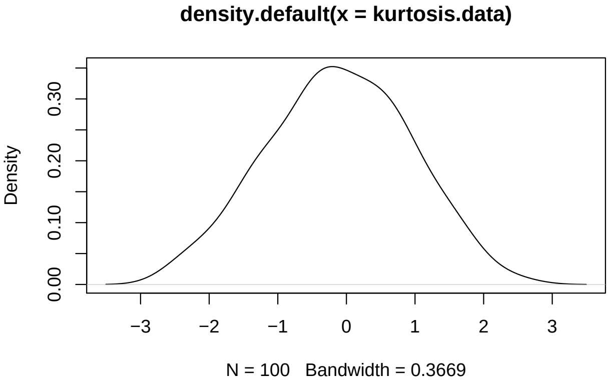
Population Kurtosis $\left(= \frac{\mathbb{E}[(y-\mu)^4]}{\sigma_y^4}\right)$ measures whether a distribution has **fat tails** (meaning more probability in the extreme tails, that is, very large positive or negative values).

Sample kurtosis is given by $\frac{\frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{y})^4}{\widehat{\sigma}^4}$. A (normal) distribution is symmetric and thus has a kurtosis of 3.

Thus, any distribution with kurtosis greater than 3 has thick tails, while any distribution with kurtosis less than 3 has thin tails.


0.4
0.3
0.2
0.1
0
−4
−2
0
2
4
x

Example: Fat tail vs Thin Tail?



Example: Fat tail vs Thin Tail

Let's calculate the kurtosis of the data.

```
kurtosis(kurtosis.data)

## [1] 2.544437
```

When a random variable displays **excess kurtosis**, **fat tail**, it is also called **leptokurtic**.

When a random variable displays **thin tail**, it is also called **platykurtic**.

Those of you who are familiar with Greek may think that these definitions are errors, since *lepto* means *slim* and *platy* means *fat*. As explained by Kendall and Stuart (1977, p.88), these terms were originally used to refer not to the tails but to the central part of the distribution; e.g., fat central part is "thin tail".

Question: How to Program it yourself?

Some useful tips:

1. To view a function in a non-base R package, just type the following:

```
moments:::skewness
```

Pearson's System: Based on the first four moments, we obtain

Mean: First moment

Pearson's System: Based on the first four moments, we obtain

Mean: First moment

Variance Second moment

Pearson's System: Based on the first four moments, we obtain

Mean: First moment

Variance Second moment

Skewness (Standardized) Third moment

Pearson's System: Based on the first four moments, we obtain

- **Mean**: First moment
- **Variance** Second moment
- **Skewness** (Standardized) Third moment
- **Kurtosis** (Standardized) Fourth moment

Application I:

Financial Data for Investment Opportunities

Table 1 - Summary statistics for stock index returns

Index	Mean	Median	Min	Max	StdDev	Skewness	Kurtosis
S&P 500	0.0137	0.0687	-9.4695	10.9571	1.3501	-0.2040	9.7826
FTSE	0.0077	0.0413	-9.2645	9.3842	1.2918	-0.1203	8.0672
NIKKEI	-0.0220	0.0037	-12.1110	13.2345	1.6048	-0.2861	8.5632
IBOVESPA	0.0570	0.1379	-17.2082	28.8324	2.2520	0.3184	15.3954
BSE	0.0365	0.1062	-11.8091	15.9899	1.7193	-0.0899	8.1902
IPC	0.0642	0.1073	-14.3144	12.1536	1.5955	0.0131	9.4692

Figure: Basic Statistics of Stock Returns from Various Markets

Stylized facts about stock markets:

Usually negative skewness! (Except Mexico and Brazil)

What does it mean?

Franses and van Dijk (2000): the kurtosis of the stock returns is much larger than the kurtosis of the exchange rate returns.

Central banks can intervene in the foreign exchange market, while there are virtually no such opportunities in stock markets (with already a large amount of risk).

Application II: Government Intervention

Question: Where did all the risk go?

Consider some financial market with the following returns

```
##  [1]  0.83373317 -0.27604777 -0.35500184  0.08748742  2.25225573
##  [6]  0.83446013  1.31241551  2.50264541  1.16823174 -0.42616558
## [11] -0.99612975 -1.11394990 -0.05573154  1.17443240  1.05321861
## [16]  0.05760597 -0.73504289  0.93052842  1.66821097  0.55968789
## [21] -0.75397477  1.25655419  0.03849255  0.18953983  0.46259495
## [26] -0.42736305  0.01658600  0.70487910  0.97184932 -0.62049160
## [31] -0.85586700  0.06955833 -1.04619827 -2.74886838 -1.12985961
## [36] -0.86168477  1.56007385  1.01508837  1.04399443 -1.11590519
## [41] -1.07130450  0.96782094  0.17103255 -0.89625029  0.15828925
## [46] -0.50194847 -0.96592262 -0.11373414  1.08594981 -1.21164826
```

```r
var(return)
```

```
## [1] 1.084689
```

Now let's say that the government intervenes by adding many investment opportunites with zero or low returns

```
##   [1]  0.83373317 -0.27604777 -0.35500184  0.08748742  2.25225573
##   [6]  0.83446013  1.31241551  2.50264541  1.16823174 -0.42616558
##  [11] -0.99612975 -1.11394990 -0.05573154  1.17443240  1.05321861
##  [16]  0.05760597 -0.73504289  0.93052842  1.66821097  0.55968789
##  [21] -0.75397477  1.25655419  0.03849255  0.18953983  0.46259495
##  [26] -0.42736305  0.01658600  0.70487910  0.97184932 -0.62049160
##  [31] -0.85586700  0.06955833 -1.04619827 -2.74886838 -1.12985961
##  [36] -0.86168477  1.56007385  1.01508837  1.04399443 -1.11590519
##  [41] -1.07130450  0.96782094  0.17103255 -0.89625029  0.15828925
##  [46] -0.50194847 -0.96592262 -0.11373414  1.08594981 -1.21164826
##  [51]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [56]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [61]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [66]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [71]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [76]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [81]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [86]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [91]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
##  [96]  0.00000000  0.00000000  0.00000000  0.00000000  0.00000000
```

Application III:

How can We use Economic Theories to think about these concepts?

Systematic Way of Thinking about Various Risks

Exploring (Ignored) Higher-Order Risk Effects:

Economic theories predict: **Prudence** plays an important role in the tradeoff between risk and (negative) skewness for economic decisions made under uncertainty, as shown by Chiu (2005).

A lesser known trait affecting behavior towards risk is **temperance**, a term also coined by Kimball (1992) a temperate individual generally dislikes kurtosis.