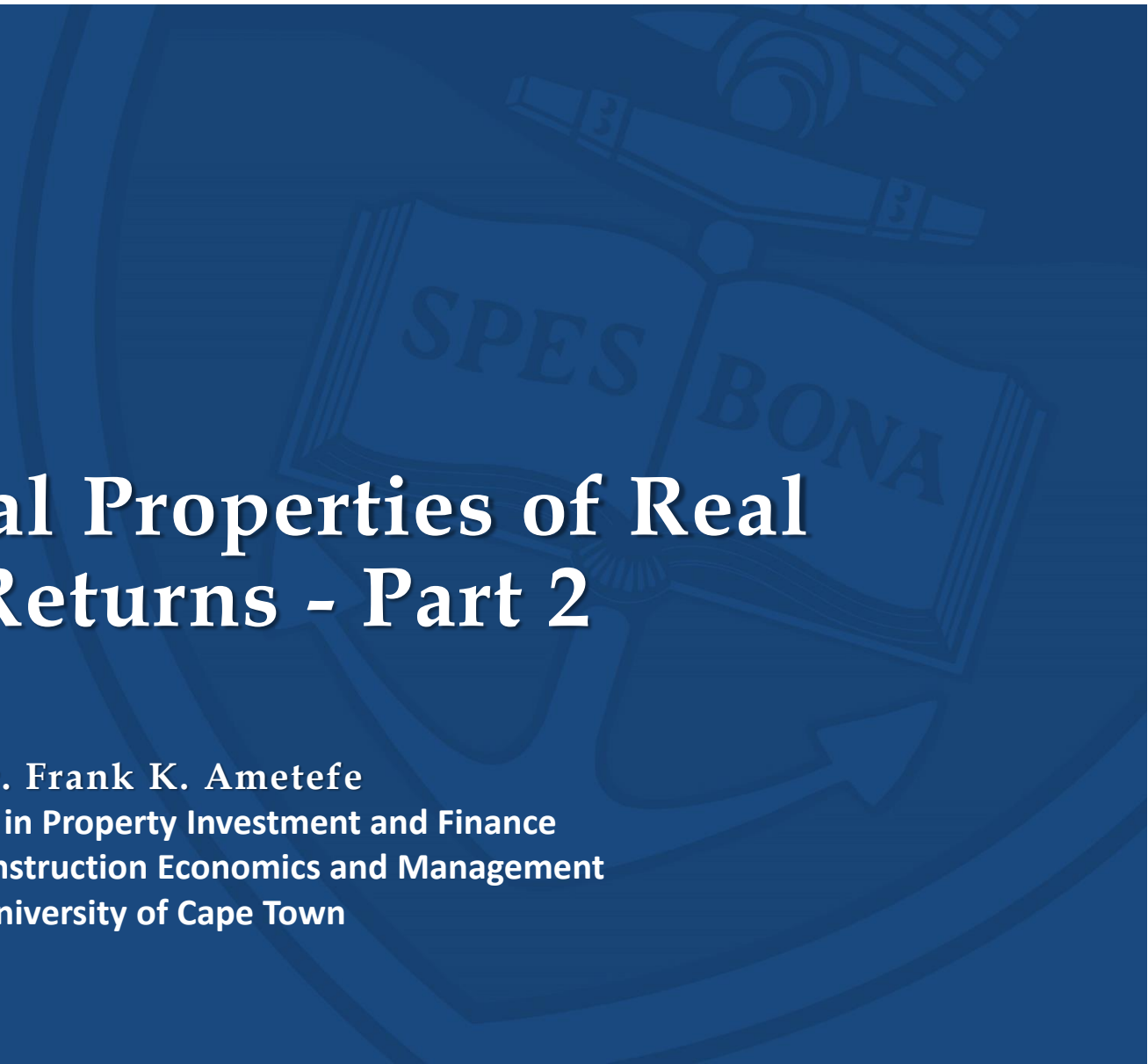


The background of the slide features a large, faint watermark of the University of Cape Town crest. The crest is circular and contains a shield with a book and a quill. The words 'SPES' and 'BONA' are visible on the book. The crest is surrounded by a banner and other heraldic elements.

Distributional Properties of Real Estate Returns - Part 2

Dr. Frank K. Ametefe
Senior Lecturer in Property Investment and Finance
Department of Construction Economics and Management
University of Cape Town

Detecting Non-normality in Property Data

- Nearly all inferential statistics used in property economics, finance and investment are based on the assumption of normality.
- Many observers today believe that deviations from normality of asset returns are too significant to ignore.
- What if returns are not normally distributed?
 - Statistical measures such as the standard deviation are no longer adequate
 - It is therefore important to check that data is normally distributed before proceeding with any statistical analysis.

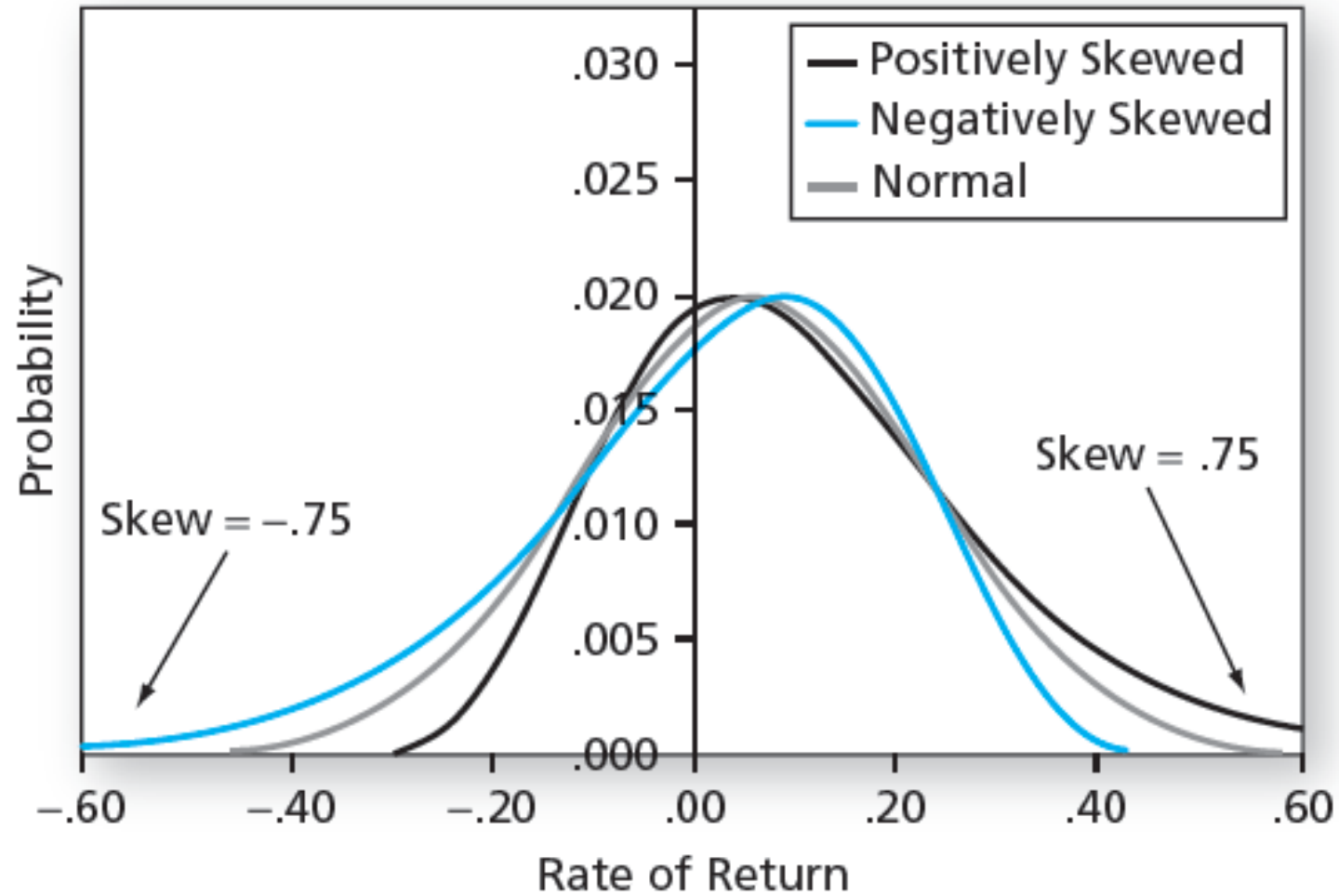
Detecting Non-normality in Property Data

- Popular measures/approaches used to detect normality in data include:
 - Graphical: Histograms, Box Plots, QQ plots etc.
 - Skewness
 - Kurtosis
 - Jarque-Bera Test
 - Anderson-Darling Test
 - Kolmogorov-Smirnov (K-S) Test
 - Shapiro-Wilk (SW) Test

SKEWNESS

- Skewness is a measure of symmetry or rather lack of symmetry in data.
- A distribution is asymmetric when one tail is longer than the other.
- A positive skewness means the tail of the distribution extends to the right.
- When the distribution is positively skewed (the skew is greater than zero), the standard deviation overestimates risk because extreme positive deviations from expectation (which are not a concern to the investor) nevertheless increase the estimate of volatility
- If the skewness value is negative, the distribution is skewed to the left, i.e. the tail of the distribution extends to the left.
- Conversely and more importantly, when the distribution is negatively skewed, the standard deviation will underestimate the risk

Skewness



Mean = 6%, SD = 17%

Skewness

- There are a number of formulas used to calculate skewness. Most spreadsheets and statistical/econometric packages use the Fisher-Pearson standardised third central moment formula.
- The sample skewness can be calculated using the formula below:

$$\text{Sample Skewness} = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{r_i - \bar{r}}{\sigma} \right)^3$$

- n = number of observations
- r_i = observed return for each period
- $\bar{r}$ = average return
- σ = sample standard deviation

Skewness

- The population skewness can be calculated using the formula below:

$$\textit{Population Skewness}^* = \frac{1}{N} \sum_{i=1}^n \left(\frac{r_i - \bar{r}}{\sigma} \right)^3$$

- N = number of observations
- r_i = observed return for each period
- $\bar{r}$ = average return
- σ = population standard deviation.

Example 1

- The data below is for a Kavala Property Trust, a Property Fund you are interested in investing in. Determine if the returns of the fund are positively or negatively skewed by calculating: (a) The sample skewness and (b) the population skewness statistic.

Year	Return
2001	2.6
2002	1.2
2003	1.6
2004	5.6
2005	11.2
2006	2.4

Example 1a (Answer)

Year	Return	$r - \bar{r}$	$(r - \bar{r})/\sigma$	$[(r - \bar{r})/\sigma]^3$
2001	2.6	-1.50	0	-0.0612
2002	1.2	-2.90	-1	-0.4421
2003	1.6	-2.50	-1	-0.2832
2004	5.6	1.50	0	0.0612
2005	11.2	7.10	2	6.4876
2006	2.4	-1.70	0	-0.0891

Example 1a (Answer)

Average ($\bar{r}$)	4.1000
Standard Deviation (σ)	3.8068
σ^3	55.1687
n	6.0000
n-1	5.0000
n-2	4.0000
$n/(n-1)(n-2)$	0.3000
$\Sigma[(r-\bar{r})/\sigma]^3$	5.6732
Skewness	1.7020
Skewness (using excel)	1.7020

Example 2

- Using the dataset linked below, calculate the sample and population skewness of the return series:

[Return Data.xls](#)

- What conclusions can you draw?

Example 3

- A spreadsheet containing the IPD SA All Property total returns for the period 2013 to 2022 is linked below:

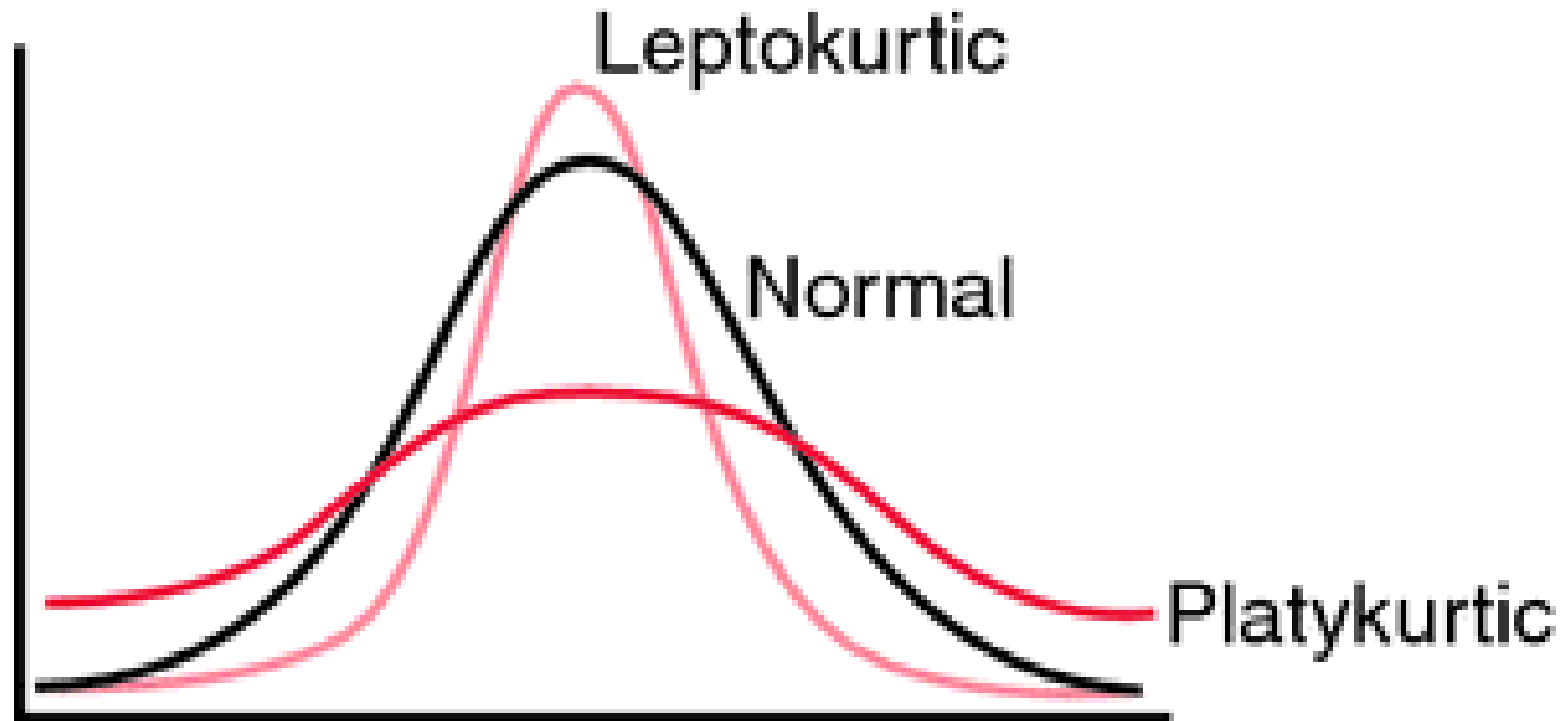
Return Data.xls

- Calculate the sample and population skewness of the total return series.
- What conclusions can you draw?

Kurtosis

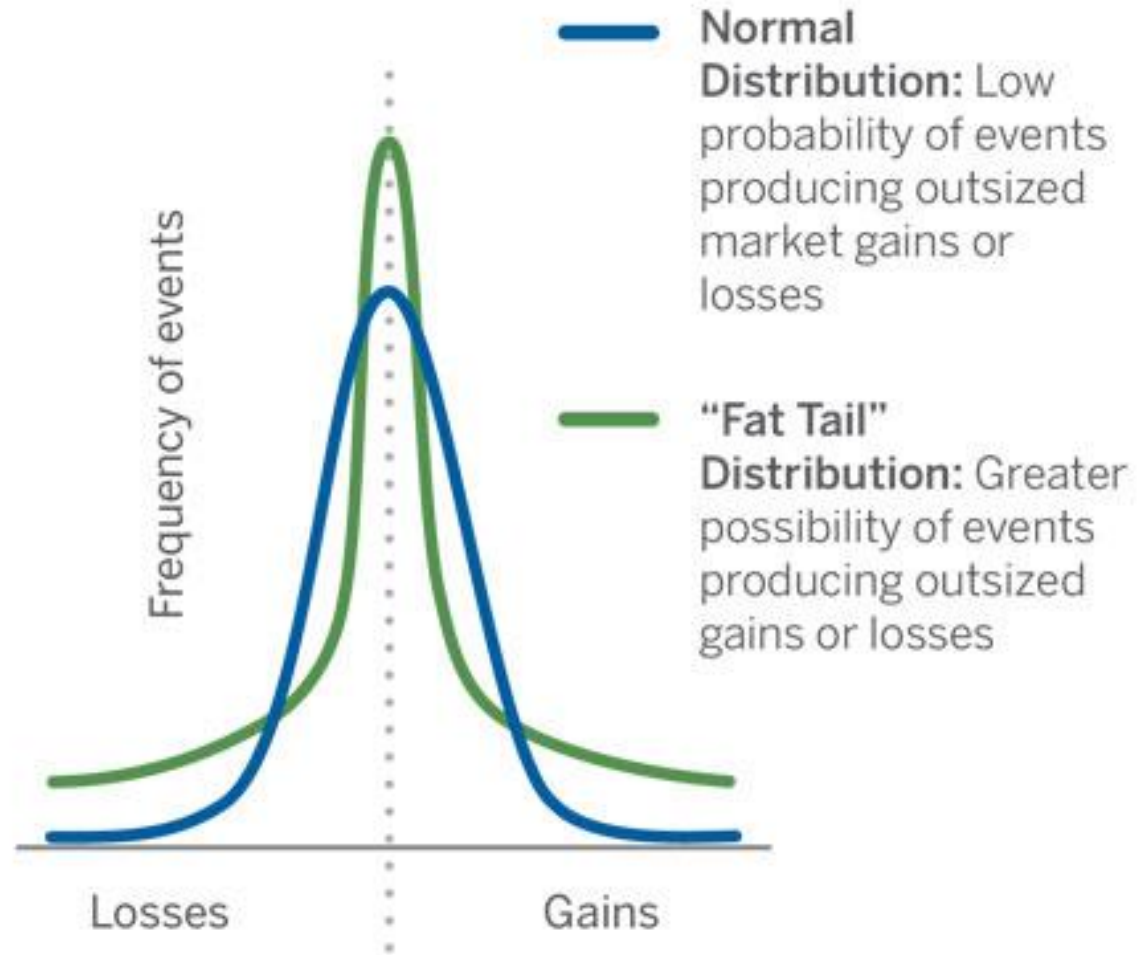
- Another important deviation from normality concerns the likelihood of extreme values on either side of the mean (fat tails).
- Kurtosis provides information on the tails (the extremes or outliers) of a distribution.
- Graphically speaking, when the tails of a distribution are 'fat' there is more probability mass in the tails of the distribution than predicted by the normal distribution at the expense of 'slender shoulders'
- Although asymmetry is still preserved, the standard deviation will underestimate the likelihood of extreme events: large losses as well as large gains

Kurtosis



Kurtosis

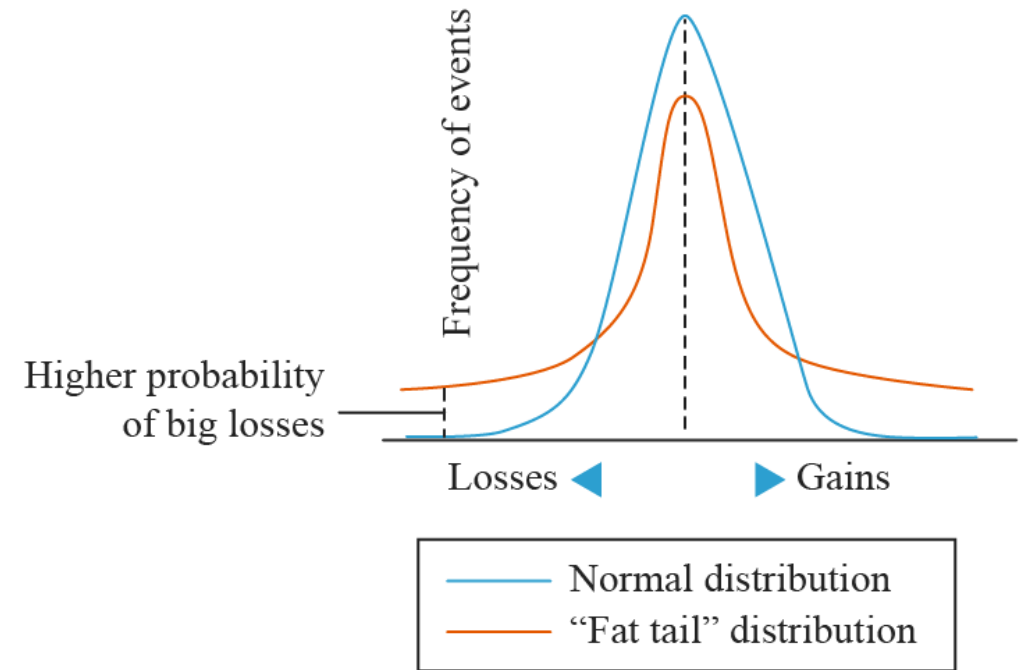
EXHIBIT 1: ARE TAILS GETTING “FATTER?”



[Source: Brown Advisory \(2017\)](#)

Kurtosis

- The kurtosis for a normal distribution is 3.
- Any Kurtosis above 3 is a sign of fatter tails than would be observed in a normal distribution.
- Similarly, a positive **excess kurtosis** (kurtosis - 3) has more extreme values (outliers) than a normal distribution.
- The excess kurtosis for a normal distribution (Kurtosis - 3) should therefore be zero.



Sample Kurtosis

- The adjusted Fisher–Pearson standardized central moment coefficient is often used as an estimate of the sample kurtosis :

$$\textbf{\textit{Kurtosis}} = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{r_i - \bar{r}}{\sigma} \right)^4 \right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

- n = number of observations (often sample size)
- r_i = observed return for each period
- $\bar{r}$ = average return
- σ = standard deviation (often sample standard deviation)

Population Kurtosis

- The Population Kurtosis is the fourth standardized moment, given as:

$$\text{Population Kurtosis} = \frac{1}{N} \sum_{i=1}^n \left(\frac{r_i - \bar{r}}{\sigma} \right)^4$$

- N = Number of observations (often population size)
 - r_i = observed return for each period
 - $\bar{r}$ = average return
 - σ = population standard deviation
-
- The formula presented here is the default Kurtosis value that Econometric packages like EViews and STATA calculate.
 - Unlike the sample Kurtosis used by Excel which calculates excess Kurtosis (Kurtosis – 3), the population Kurtosis formula here is the actual kurtosis and should be interpreted relative to 3.

Example 1

- The data below is for a Property Fund you are interested in investing in. Estimate the sample kurtosis of the return distribution:

Year	Return
2001	2.6
2002	1.2
2003	1.6
2004	5.6
2005	11.2
2006	2.4

Example 1 (Answer)

KURTOSIS:	
Average ($\bar{r}$)	4.1000
Sample Standard Deviation (σ)	3.8068
n	6.0000
n+1	7.0000
n(n+1)	42.000
n-1	5.0000
n-2	4.0000
n-3	3.0000
(n-1)(n-2)(n-3)	60.000

Example 1 (Answer)

KURTOSIS:	
$n(n+1)/(n-1)(n-2)(n-3)$	0.7000
$\sum[(r-\bar{r})/\sigma]^4$	12.711
$n(n+1)/(n-1)(n-2)(n-3) * \sum[(r-\bar{r})/\sigma]^4$	8.8974
$3(n-1)^2$	75.000
$(n-2)(n-3)$	12.000
$3(n-1)^2/(n-2)(n-3)$	6.2500
Kurtosis	2.6474

Example 2

- Using the dataset linked below, calculate the sample kurtosis of the return series:

[Return Data.xls](#)

- What conclusions can you draw?

JARQUE-BERA TEST

- The Jarque-Bera test for normality was derived and proposed in 1980 by Carlos Jarque and Anil Bera.
- The test is a simple and efficient procedure that uses the estimated skewness and kurtosis to arrive at a test statistic that can be used to draw a conclusion about the normality of a given data series.

$$Jarque\ Bera = n \left(\frac{S^2}{6} + \frac{(K - 3)^2}{24} \right) \sim \chi^2(2)$$

S = Skewness

$K - 3$ = Excess Kurtosis

n = Number of observations

Note: MS Excel calculates Excess Kurtosis directly, so there's no need to subtract 3 to arrive at Excess Kurtosis.

JARQUE-BERA TEST

- The Jarque-Bera tests the following hypothesis:
 - H_0 : *The data is normally distributed*
 - H_1 : *The data is not normally distributed*
- Once the JB test statistic is computed, it can be compared to the critical values of the χ^2 distribution with a degree of freedom of 2.
- A JB test statistic greater than the critical value of the corresponding χ^2 test statistic means we can reject the null hypothesis that the data is normally distributed.

JARQUE-BERA TEST – METHOD 1

INTERPRETATION –METHOD 1

- One of the ways to interpret the JB test statistic is to compare the computed JB test statistic to the chi-square (χ^2) critical value of a test with 2 degrees of freedom and a significance level of α , usually 5% or 1%.
- The chi-square critical value can be obtained from the a chi-square distribution table or using Excel Formula “=CHISQ.INV.RT(α , df)”
- The decision rule is to reject the null hypothesis that the data is normally distributed if the computed JB-statistic is greater than the χ^2 critical value. Rejecting the null hypothesis implies the data is not normally distributed.
- If the JB-statistic is less than the χ^2 critical value, we cannot reject the null hypothesis. This implies the data is normally distributed.
- The higher the JB Statistics, the more likely we are to reject the null hypothesis.
 - In other words, a large JB Statistic implies the data is not normally distributed.
 - Typically, a JB statistic of more than 4 means the data is very likely not normally distributed.

JARQUE-BERA TEST

INTERPRETATION –METHOD 2

- Microsoft Excel and most statistical programmes are able to calculate a precise p-value, given the JB statistic.
- The MS Excel formula is “=CHISQ.DIST.RT(JB Statistic, 2)”
- The decision rule is to reject the null hypothesis that the data is normally distributed if the p-value is less than 1% or 5% (i.e., desired significance level). Rejecting the null hypothesis implies the data is not normally distributed.
- A non-significant p-value (i.e., greater than 1% or 5%) implies we are unable to reject the null that the data is normally distributed. This implies the data is normally distributed.

JARQUE-BERA TEST

Table 3. Jarque-Bera Chi-Square Test for Price Sample Normality (Jan. 2000 – June 2001)

Month	Sample Size	Computed JB-statistic	Critical Value	Decision
Jan-00	81	61.022	5.991	Reject
Feb-00	81	82.553	5.991	Reject
Mar-00	81	62.74	5.991	Reject
Apr-00	78	42.553	5.991	Reject
May-00	81	18.535	5.991	Reject
Jun-00	81	30.839	5.991	Reject
Jul-00	81	15.759	5.991	Reject
Aug-00	81	19.538	5.991	Reject
Sep-00	81	53.113	5.991	Reject
Oct-00	81	198.66	5.991	Reject
Nov-00	81	289.65	5.991	Reject
Dec-00	81	280.525	5.991	Reject
Jan-01	74	640.92	5.991	Reject
Feb-01	79	4.538	5.991	Fail to Reject
Mar-01	86	5.058	5.991	Fail to Reject
Apr-01	86	3.159	5.991	Fail to Reject
May-01	86	182.531	5.991	Reject
Jun-01	86	148.417	5.991	Reject

Source: [Fullerton and Torres \(2005\)](#)

- The table shows the normality test results as presented by Fullerton and Torres (2005)
- The authors compare the computed JB statistic with the critical value to determine whether to reject the null hypothesis that the Price data is normally distributed.
- Based on the result presented, only three of the series could be said to be normally distributed.

JARQUE-BERA TEST

Table 2. Skewness and kurtosis of the series and Jarque-Bera statistic and probability to test for normality.

	Cash Returns	Deposit Returns	Valuation Interest Liabilities	Debt Market Returns	Stock Market Returns	Inflation	Mortgage Interest
Skewness	0.108	0.151	-0.063	-0.161	-0.212	0.12	0.032
Kurtosis	1.553	1.547	1.509	2.889	2.544	1.884	1.431
Jarque-Bera	13.37 ^(***)	13.75 ^(***)	13.97 ^(***)	0.732	1.801	7.87	15.41 ^(***)
Probability	0.0012	0.0011	0.00009	0.693	0.406	0.19	0.00045

(*)(**)(***) Denote significance at the 10%/5%/ 1% level

Source: [Princep et al. \(2014\)](#)

- The table shows the normality test results as presented by Princep et al (2005)
- The authors use the probability values to determine whether to reject the null hypothesis that the different series are normally distributed.
- Only three of the series (Debt Market, Stock Market and Inflation) could be said to be normally distributed, based on the result presented.

Example 1

- The data below is for a Property Fund you are interested in investing in. Using the Jarque Bera Test, determine if the returns are normally distributed.

Year	Return
2001	2.6
2002	1.2
2003	1.6
2004	5.6
2005	11.2
2006	2.4

Example 2

- Using the dataset linked below, calculate the sample and population skewness of the return series:

[Return Data.xls](#)

- What conclusions can you draw?

Example 3

- A spreadsheet containing the IPD SA All Property total returns for the period 2013 to 2022 is linked below:

Return Data.xls

- Calculate the sample and population skewness of the total return series.
- What conclusions can you draw?

Example 4

- The average daily returns for a property company over 200 days was recorded as 1.24% with skewness of 0.232 and kurtosis of 3.319. Test the hypothesis that the distribution is normally distributed.

Example 4 (Answer)

$$\bullet \textit{Jacque Bera} = n \left(\frac{s^2}{6} + \frac{(K-3)^2}{24} \right) \sim \chi^2(2)$$

$$\bullet S = 0.232 \qquad K = 3.319 \qquad n = 200$$

$$\bullet \textit{Jacque Bera} = 200 \left(\frac{0.232^2}{6} + \frac{(3.319-3)^2}{24} \right)$$

$$\bullet \textit{Jacque Bera} = 2.6421.$$

Example 4 (Answer)

- Interpretation – Method 1
- JB-statistic = 2.6421.
- Critical value of chi-square = 5.991
- Since JB statistic is less than 5.991, we cannot reject the null hypothesis. This implies that the data is normally distributed.

Example 4 (Answer)

- Interpretation – Method 2
- JB-statistic = 2.6421.
- P-value = 0.266849
- Since the p-value is greater than 5% i.e. 0.05, we cannot reject the null hypothesis. This implies the data is normally distributed.