Statistical Methods for Economists

Lecture 1

Basic Statistical Concepts and Data Characteristics

SILESIAN UNIVERSITY
SCHOOL OF BUSINESS ADMINISTRATION IN KARVINA

David Bartl
Statistical Methods for Economists
INM/BASTE

Outline of the lecture

- Reading list
- Measures of central tendency (arithmetic mean, mode, median)
- Measures of variability (range, variance, coefficient of variation)
- Measures of data concentration (skewness, kurtosis)
- Moment characteristics
- Two statistical variables
- SUPPLEMENT: The expected values of the functions of random variables

Reading list

<u>Compulsory:</u>

- TOŠENOVSKÝ, Filip: *Statistical Methods for Economists.* Karviná: SU OPF, 2014. ISBN 978-80-7510-033-7

- ANDERSON, David R., SWEENEY, Dennis J., WILLIAMS, Thomas A., FREEMAN, James, SHOESMITH, Eddie: *Statistics for Business and Economics.* 4th Edition. Cengage Learning, 2017. ISBN 978-1-4737-2656-7

- KELLER, Gerald: *Statistics for Management and Economics.* 11th Edition. Cengage Learning, 2017. ISBN 978-1-337-09345-3

Reading list

<u>Free Online Textbooks:</u>

- Many textbooks on statistics and other disciplines can be found at https://**freetextbook.org**/
- Online Statistics Education: An Interactive Multimedia Course of Study http://**onlinestatbook.com**/
- The Electronic Statistics Textbook by StatSoft, Inc. (2013) **www.statsoft.com/textbook**
- The printed version of the latter textbook:
  HILL, T. & LEWICKI, P. (2007). *STATISTICS: Methods and Applications.* StatSoft, Tulsa, OK.

Reading list

Recommended I:

- SIEGEL, Andrew: *Practical Business Statistics.* 7th Edition. Academic Press, 2016. ISBN 978-0-12-804250-2

- ÖZDEMIR, Durmuş: *Applied Statistics for Economics and Business.* 2nd Edition. Springer, 2016. ISBN 978-3-319-26495-0 (hardcover). ISBN 978-3-319-79962-9 (softcover).

- UBØE, Jan: *Introductory Statistics for Business and Economics: Theory, Exercises and Solutions.* 1st Edition. Springer, 2017. ISBN 978-3-319-70935-2 (hardcover). ISBN 978-3-319-89016-6 (softcover).

Reading list

Recommended II:

- QUIRK, Thomas: *Excel 2016 for Business Statistics: A Guide to Solving Practical Problems.* 1st Edition. Springer, 2016. ISBN 978-3-319-38958-5 (softcover).

- HERKENHOFF, Linda, FOGLI, John: *Applied Statistics for Business and Management using Microsoft Excel.* 1st Edition. Springer, 2013. ISBN 978-1-4614-8422-6 (softcover).

Reading list

Optional:

- DANIEL, W. W., TERREL, J.: *Business Statistics for Management and Economics.* Houghton Mifflin, 1995. ISBN 0-395-73717-6
- WOOLDRIDGE, J. M.: *Introductory Econometrics: A Modern Approach.* Mason, OH: Thomson/South-Western, 2006. ISBN 0-324-28978-2
- VAN MATRE, J. G., GILBREATH, G. H.: *Statistics for Business and Economics.* BPI/IRWIN, Homewood, 1997. ISBN 0-256-03719-1

Basic Statistical Concepts

- Data — Data unit — Data item — Observation — Dataset
- Population — Sample — Data item
- Population & Sample

Data — Data unit — Data item — Observation — Dataset

Data — (plural) — measurements and observations

Data unit — one entity (e.g. a person) in the *population*, under study, about which the data are collected

Data item — a characteristics (an attribute) of a data unit (e.g. the date of birth, gender, income, …), also called a **variable**

Observation — an occurrence of a specific data item recorded about a data unit, also called a **datum** (singular of “data”)

Dataset — a complete collection of all observations

Population — Sample — Data item

Population — a collection of all data units of the same specification

Sample — a selected subset of the population

Data item — a property or an attribute of a data unit of the population

Data items – **statistical variables** – are:

- **qualitative** (categorical), such as the gender, colour, taste, satisfaction
- **quantitative** (numerical), such as the revenue, price, number of customers

Population & Sample

Assume that we have a set (i.e. a "**population**") of values of some phenomenon, which we observe / measure / study / deal with. In practice, this set may be very very large (e.g. some data item, the data units being all the people living on the Earth), thus unknown to us. Another example might be the set of all results of some experiment, yet the instances which we have not done yet.

Assume however, that the set exists (in theory at least) and that the set is finite (for simplicity).

Population & Sample

Let

$$\Omega = \{1, 2, 3, \dots, N\}$$

be the underlying set of all data units. We assume for simplicity that the set Ω is finite and that N is the number of its elements.

Now, considering some variable or data item $X: \Omega \to \mathbb{R}$, we assume that the values

$$x_1, x_2, x_3, \dots, x_N$$

of the variable exist (in theory at least).

We assume that $x_1, x_2, x_3, \dots, x_N \in \mathbb{R}$, i.e. the values are real numbers.

Population & Sample

All the values

$$x_1, x_2, x_3, \dots, x_N$$

which exist (in theory at least), are called the **population.**

Notice that we may not know the whole population for various reasons (we cannot do all the measurements since the population is very large; the values include the results of future experiments, which have not been done yet).

Population & Sample

Our method of examination of the population $x_1, x_2, x_3, \ldots, x_N$ consists in **the selection of a sample**

$$x_1, x_2, \ldots, x_n$$

out of the population.

Notice that the same letter ("x") is used to denote the values of both the population and the sample. No misunderstanding occurs because it is always clear from the context whether we mean the population or the sample.

Notice also the notation: N = the number of elements of the **population**

n = the number of elements of the **sample**

Measures of central tendency

- Arithmetic mean
- Mode
- Median
- Frequencies of occurrence
- Weighted arithmetic mean

Measures of central tendency

Assume that a variable (data item) is numerical, i.e. quantitative, discrete or continuous. We then consider several measures of central tendency of the variable:

- Arithmetic mean
- Mode
- Median

Arithmetic mean

Population arithmetic mean:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Sample arithmetic mean:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Arithmetic mean

Notice the notation:

Greek letters denote quantities relating to the population:

μ = the **population** mean (theoretical, may not be known exactly)

Latin letters denote quantities relating to the sample:

$\bar{x}$ = the **sample** mean (the result of measurements really done)

Median & Mode

Median $\tilde{x}$ is the “middle value” such that

- one half of the values is ≤ the median
- one half of the values is ≥ the median

Mode $\hat{x}$ is the “most frequent value” such that

- the probability distribution attains a local maximum at the mode
- there may be more than one mode:
  - unimodal probability distribution (one mode)
  - bimodal probability distribution (two modes)
  - etc.

The median and the mode are defined both for the population and for the sample.
The definition is the same.

Sample Mean / Median / Mode in Excel

In Excel, use the functions:

=**AVERAGEA**()	to calculate the sample arithmetic mean
=**MEDIAN**()	to find the sample median
=**MODE.SNGL**()	to find one of the sample modes
=**MODE.MULT**()	to find many of the sample modes (matrix function, press “Ctrl-Shift-Enter”)
=MODE()	to find one of the sample modes (the same as =MODE.SNGL(), deprecated)

Frequencies of occurrence

Consider the population

$$x_1, x_2, x_3, \dots, x_N$$

Let $x_1^*, x_2^*, \dots, x_K^*$ be the all the unique values in the population, i.e. values such that

$$x_1^* < x_2^* < \cdots < x_K^* \qquad \text{and} \qquad \{x_1^*, x_2^*, \dots, x_K^*\} = \{x_1, x_2, x_3, \dots, x_N\}$$

Let

f_1 be the frequency of the value x_1^* in the population,

f_2 be the frequency of the value x_2^* in the population,

…

Frequencies of occurrence

Mathematically written, we have:

$$f_1 = |\{i \in \{1,2,3,\dots,N\} : x_1^* = x_i\}|$$

$$f_2 = |\{i \in \{1,2,3,\dots,N\} : x_2^* = x_i\}|$$

$$\vdots$$

$$f_K = |\{i \in \{1,2,3,\dots,N\} : x_K^* = x_i\}|$$

Then the population mean is

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{\sum_{k=1}^{K} f_k}\sum_{k=1}^{K} f_k x_k^*$$

Arithmetic mean

Weighted population mean:

Assuming that weights $w_1, w_2, w_3, \dots, w_N$ (positive real numbers)

of the values $x_1, x_2, x_3, \dots, x_N$ are given, the weighted population mean is

$$\mu_w = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i}$$

Example: Employees (a sample of the Dataset)

ID	Gender	Age	Marital Status	Education	Position	Salary per Year	Evaluation
5060	M	65	divorced	secondary	worker	258800	4
1030	M	60	divorced	university	manager	630000	2
3049	M	60	married	primary	operator	436600	5
5047	M	60	widowed	primary+vocational	worker	240600	3
5061	M	60	widowed	primary+vocational	worker	241800	1
5087	M	60	widowed	secondary	worker	239500	—
5133	F	60	married	secondary	worker	241100	4
5177	F	60	widowed	secondary	worker	239600	4
3030	F	58	widowed	primary	operator	422600	1
3014	F	56	widowed	university	operator	303600	3
5012	F	56	widowed	primary+vocational	worker	223100	4
5056	M	56	divorced	primary	worker	225200	5
5101	M	56	unmarried	primary+vocational	worker	224600	4
5106	M	56	married	primary+vocational	worker	226100	7
5146	F	56	married	primary+vocational	worker	224900	3
5153	M	56	divorced	secondary	worker	224500	4
5189	M	56	married	primary+vocational	worker	224600	1
5196	M	56	widowed	primary+vocational	worker	222800	3
1031	M	55	married	university	manager	429000	—
5016	M	55	divorced	secondary	administrative officer	259000	5
5021	F	55	married	primary+vocational	worker	220200	—
5062	F	55	widowed	primary+vocational	worker	221400	5
5107	M	55	divorced	primary+vocational	worker	220500	4
5154	F	55	widowed	primary+vocational	worker	219200	5
5195	M	55	married	primary+vocational	worker	219400	6

sample

Example: Employees — data item “Age”

The true (population) values:

Population size:

$$N = 200$$

Population mean (Age):

$$\mu = 39.9$$

Population median:

$$\tilde{x} = 42$$

Population mode:

$$\hat{x} = 45$$

The estimated (sample) values:

Sample size:

$$n = 8$$

Sample mean (Age):

$$\bar{x} = 60.625$$

Sample median:

$$\tilde{x} = 60$$

Sample mode:

$$\hat{x} = 60$$

Measures of variability

- Range
- Variance (dispersion)
- Coefficient of variation

Measures of variability

Assume that a variable (data item) is numerical, i.e. quantitative, discrete or continuous. We then consider several measures of variability of the variable:

- Range
- Variance (dispersion)
- Coefficient of variation

Range

Population range:

$$R = \max_{i=1,\dots,N} x_i - \min_{i=1,\dots,N} x_i$$

Sample range:

$$R = \max_{i=1,\dots,n} x_i - \min_{i=1,\dots,n} x_i$$

Variance (dispersion)

Population variance:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

Sample variance:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Variance (dispersion)

Notice that once the population $x_1, x_2, x_3, \dots, x_N$ of the values is fixed, then the population mean μ and the population variance σ^2 are given, i.e. these theoretical values are fixed (though not known exactly sometimes).

If the sample $x_1, x_2, \dots, x_n$ of the values is selected from the population randomly (select an element randomly n-times; the same element may be chosen repeatedly several times), then the resulting values of the

sample mean $\bar{x}$ and sample variance s^2

are random variables too!!!

Variance (dispersion)

Calculating

— the expected value $\mathrm{E}\bar{x}$ of the sample mean and

— the expected value $\mathrm{E}s^2$ of the sample variance,

we obtain that

$$\mathrm{E}\bar{x} = \mu \qquad \text{and} \qquad \mathrm{E}s^2 = \sigma^2$$

Variance (dispersion)

That is,

— taking a sample of randomly selected n elements of the population (¡ where one element of the population may be present several times in the sample !)

— calculating the sample mean $\bar{x}$ and the sample variance s^2,

— repeating the above process infinitely many times, and

— calculating **the average value of the sample mean and the average value of the sample variance,**

we obtain precisely

the population mean μ and the population variance σ^2

Variance (dispersion)

<u>Conclusion:</u> We often do <u>not</u> know the exact values μ and σ^2 in practice.

However, if we take a sample of n elements selected randomly with repetition (i.e. an element can be selected several times) from the population and calculate the sample mean $\bar{x}$ and the sample variance s^2, then we have

$$\bar{x} \approx \mu \qquad \text{and} \qquad s^2 \approx \sigma^2$$

i.e. the sample mean $\bar{x}$ and the sample variance s^2 are good estimates of the unknown population mean μ and population variance σ^2.

$\rightarrow$ **<u>That is why</u>** we divide by $(n-1)$ in the sample variance s^2, not by n.

Standard deviation

Population standard deviation:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Sample standard deviation:

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Variance (dispersion) & Standard deviation

Notice the notation:

Greek letters denote quantities relating to the population:

σ^2 = the **population** variance (theoretical, may not be known exactly)

σ = the **population** standard deviation

Latin letters denote quantities relating to the sample:

s^2 = the **sample** variance (the result of measurements really done)

s = the **sample** standard deviation

Sample Variance / Standard deviation

In Excel, use the functions:

=**VARA**()	to calculate the sample variance
=**STDEVA**()	to calculate the sample standard deviation
=VAR.S()	to calculate the sample variance (skipping text values)
=VAR()	to calculate the sample variance (skipping text values) (the same as =VAR.S(), deprecated)

Population Variance / Standard deviation

In Excel, use the functions:

=**VARPA**()	to calculate the population variance
=**STDEVPA**()	to calculate the population standard deviation
=VAR.P()	to calculate the population variance (skipping text values)

Coefficient of variation

Coefficient of variation:

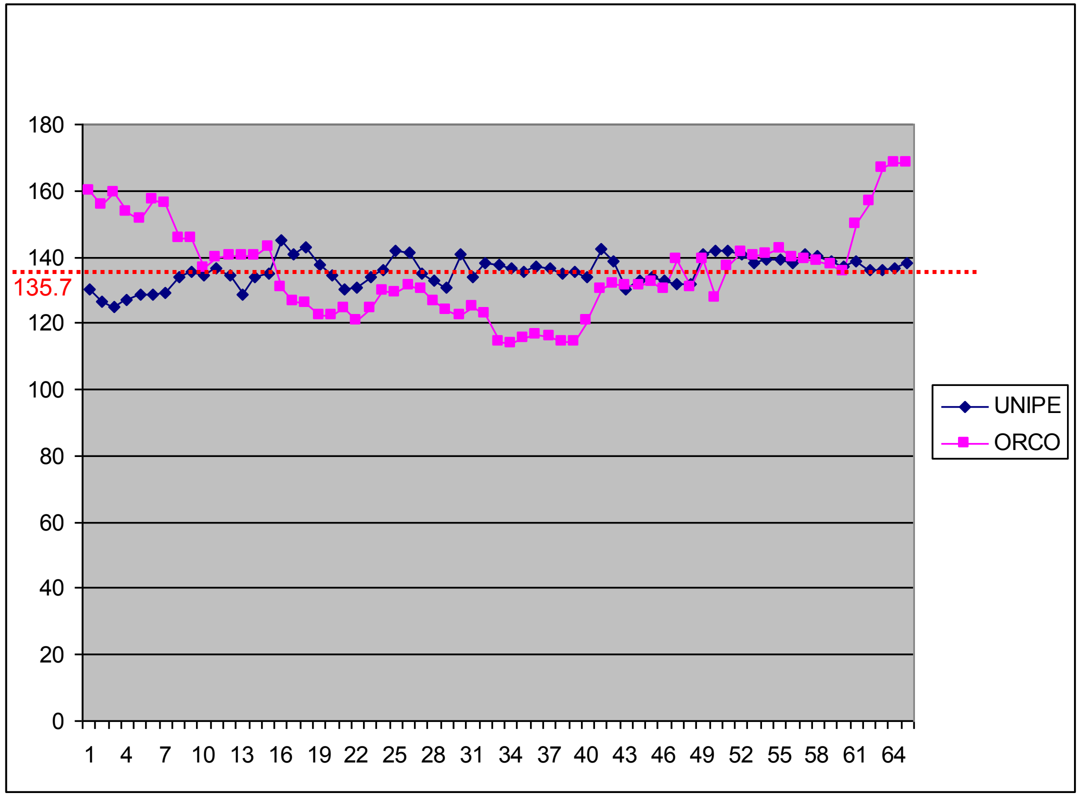
$$V = \frac{\sigma}{|\mu|}$$

<u>Sample</u> coefficient of variation:

$$v = \frac{s}{|\bar{x}|}$$

Example

The prices of two stocks (ORCO and UNIPE) during a period of time:



The average price of both stocks is the same:

$$\bar{x}_{UNIPE} = \bar{x}_{ORCO} = 135.7$$

Example

We have

$$\bar{x}_{\text{UNIPE}} = 135.7 \qquad \text{and} \qquad s_{\text{UNIPE}} = 2.09$$

hence

$$v_{\text{UNIPE}} = \frac{s_{\text{UNIPE}}}{|\bar{x}_{\text{UNIPE}}|} = \frac{2.09}{135.7} = 0.0154$$

We have

$$\bar{x}_{\text{ORCO}} = 135.7 \qquad \text{and} \qquad s_{\text{ORCO}} = 3.72$$

hence

$$v_{\text{ORCO}} = \frac{s_{\text{ORCO}}}{|\bar{x}_{\text{ORCO}}|} = \frac{3.72}{135.7} = 0.0274$$

Measures of data concentration

- Skewness
- Kurtosis

Measures of data concentration

Assume that a variable (data item) is numerical, i.e. quantitative, discrete or continuous. We then consider several measures of data concentration of the variable:

- Skewness
- Kurtosis

Skewness: Pearson's moment coefficient of skewness

<u>Population</u> skewness:

$$\mathrm{Skew}(X) = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^3}{\sigma^3}$$

<u>Sample</u> skewness:

$$\mathrm{Skew}(X) = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^3}{s^3}$$

Skewness: Properties and interpretation

Pearson's moment coefficient of skewness

$$\mathrm{Skew}(X) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^3}{\sigma^3} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^3$$

is a sum of the third powers of the fractions $\frac{x_i-\mu}{\sigma}$.

If the fraction is "small", i.e. $\left|\frac{x_i-\mu}{\sigma}\right| < 1$, then its third power is yet smaller,

almost vanishes, $\left|\frac{x_i-\mu}{\sigma}\right|^3 < \left|\frac{x_i-\mu}{\sigma}\right| < 1$, i.e. is not counted much in the sum.

Skewness: Properties and interpretation

Pearson's moment coefficient of skewness

$$\text{Skew}(X) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^3}{\sigma^3} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^3$$

is a sum of the third powers of the fractions $\frac{x_i-\mu}{\sigma}$.

If the fraction is "large", i.e. $\left|\frac{x_i-\mu}{\sigma}\right| \geq 1$, then its third power is also large,

$\left|\frac{x_i-\mu}{\sigma}\right|^3 \geq \left|\frac{x_i-\mu}{\sigma}\right| \geq 1$, i.e. is counted in the sum properly.

Skewness: Properties and interpretation

Pearson's moment coefficient of skewness

$$\text{Skew}(X) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^3}{\sigma^3} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^3$$

can be positive or zero or negative.

- $\text{Skew}(X) < 0$ — the majority of the values is left to the mean
- $\text{Skew}(X) = 0$ — the values are distributed ≈ symmetrically around the mean
- $\text{Skew}(X) > 0$ — the majority of the values is right to the mean

Large positive or negative value — there are "outliers", i.e. values far away from the mean

Skewness in Excel

In Excel, use the functions:

=**SKEW.P**()	to calculate the population skewness
=**SKEW**()	to calculate the sample skewness

Skewness in Excel

Notice that we have defined sample skewness as the Pearson moment coefficient

$$\text{Skew}(X) = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i - \bar{x})^3}{s^3}$$

cf. the function =**SKEW.P**() in Excel.

To calculate the sample skewness, cf. the function =**SKEW**(), Excel uses the adjusted Fisher-Pearson standardized moment coefficient

$$\text{Skew}(X) = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}\frac{(x_i - \bar{x})^3}{s^3}$$

Kurtosis: Pearson's moment coefficient of kurtosis

Population kurtosis:

$$\mathrm{Kurt}(X) = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^4}{\sigma^4}$$

Sample kurtosis:

$$\mathrm{Kurt}(X) = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^4}{s^4}$$

Kurtosis: Properties and interpretation

Pearson's moment coefficient of kurtosis

$$\mathrm{Kurt}(X) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^4}{\sigma^4} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^4$$

is a sum of the fourth powers of the fractions $\frac{x_i-\mu}{\sigma}$.

If the fraction is "small", i.e. $\left|\frac{x_i-\mu}{\sigma}\right| < 1$, then its fourth power is yet smaller,

almost vanishes, $\left|\frac{x_i-\mu}{\sigma}\right|^4 < \left|\frac{x_i-\mu}{\sigma}\right| < 1$, i.e. is not counted much in the sum.

Kurtosis: Properties and interpretation

Pearson's moment coefficient of kurtosis

$$\mathrm{Kurt}(X) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^4}{\sigma^4} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^4$$

is a sum of the fourth powers of the fractions $\frac{x_i-\mu}{\sigma}$.

If the fraction is "large", i.e. $\left|\frac{x_i-\mu}{\sigma}\right| \geq 1$, then its fourth power is also large,

$\left|\frac{x_i-\mu}{\sigma}\right|^4 \geq \left|\frac{x_i-\mu}{\sigma}\right| \geq 1$, i.e. is counted in the sum properly.

Kurtosis: Properties and interpretation

Pearson's moment coefficient of kurtosis

$$\mathrm{Kurt}(X) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^4}{\sigma^4} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^4$$

can be positive or zero.

- $\mathrm{Kurt}(X) \geq 0$ is small — the values are concentrated ≈ around the mean
- $\mathrm{Kurt}(X) > 0$ is large — there are "outliers", i.e. values far away from the mean

The Skewness & Kurtosis describe the shape of the distribution of the values (i.e. the shape of the histogram).

Excess kurtosis

The kurtosis of the Gaussian normal distribution is = 3.

That is why, the number 3 is sometimes subtracted to obtain the **population excess kurtosis:**

$$\mathrm{ExKurt}(X) = \mathrm{Kurt}(X) - 3 = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - \mu)^4}{\sigma^4} - 3$$

Kurtosis in Excel

In Excel, use the function:

=**KURT**() to calculate the **sample excess kurtosis**

Kurtosis in Excel

Notice that we would define the sample excess kurtosis by using the Pearson moment coefficient

$$\mathrm{ExKurt}(X) = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^4}{s^4} - 3$$

To calculate the sample kurtosis, the function **=KURT()** in Excel uses the formula

$$\mathrm{ExKurt}(X) = \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^4}{s^4} - 3\frac{(n-1)^2}{(n-2)(n-3)}$$

Moment characteristics

- Raw moments
- Central moments
- Standardized moments

The moments

Given the population

$$x_1, x_2, \dots, x_N \in \mathbb{R}$$

we distinguish three types of moments:

- raw moment or central moment μ'_k
- central moment μ_k
- standardized moment $\tilde{\mu}_k$

for $k = 1, 2, 3, \dots$

Raw moment

The k-th raw moment:

$$\mu'_k = \mathrm{E}[X^k] = \frac{1}{N}\sum_{i=1}^{N} x_i^k$$

Notice that:

$$\mu'_1 = \mu$$

The moment is usually defined for $k = 1, 2, 3, ...,$

Central moment

The k-th central moment:

$$\mu_k = \mathrm{E}[(X-\mu)^k] = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^k$$

Notice that:

$$\mu_2 = \sigma^2$$

This moment is defined for $k = 1, 2, 3, \ldots$

Central moment and raw moments

It holds:

$$\mu_k = \mathrm{E}[(X-\mu)^k] =$$

$$= \mathrm{E}\left[\binom{k}{0}X^k\mu^0 - \binom{k}{1}X^{k-1}\mu^1 + \binom{k}{2}X^{k-2}\mu^2 + \cdots + (-1)^k\binom{k}{k}X^0\mu^k\right] =$$

$$= \binom{k}{0}\mu^0\mathrm{E}[X^k] - \binom{k}{1}\mu^1\mathrm{E}[X^{k-1}] + \binom{k}{2}\mu^2\mathrm{E}[X^{k-2}] + \cdots + (-1)^k\binom{k}{k}\mu^k\mathrm{E}[X^0] =$$

$$= \mu_k' - \binom{k}{1}\mu^1\mu_{k-1}' + \binom{k}{2}\mu^2\mu_{k-2}' + \cdots + (-1)^{k-1}\binom{k}{k-1}\mu^{k-1}\mu_1' + (-1)^k\mu^k$$

Central moment

The k-th standardized moment:

$$\tilde{\mu}_k = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^k\right] = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^k$$

Notice that:

$$\tilde{\mu}_3 = \mathrm{Skew}(X)$$

$$\tilde{\mu}_4 = \mathrm{Kurt}(X)$$

This moment is defined for $k = 1, 2, 3, \ldots$

Two statistical variables

- Two populations
- Contingency table
- Covariance
- Correlation coefficient

Two populations

Let

$$\Omega = \{1, 2, 3, \dots, N\}$$

be the underlying set of all data units. We assume for simplicity that the set Ω is finite and that N is the number of its elements.

Now, consider two statistical variable or data items

$$X: \Omega \to \mathbb{R} \qquad \text{and} \qquad Y: \Omega \to \mathbb{R}$$

Two populations

The two variables $X, Y: \Omega \to \mathbb{R}$, where $\Omega = \{1, 2, 3, \dots, N\}$, attain the values

$$x_1, x_2, x_3, \dots, x_N \qquad \text{and} \qquad y_1, y_2, y_3, \dots, y_N$$

so we have **two populations.**

Now, let $x_1^*, x_2^*, \dots, x_K^*$ and $y_1^*, y_2^*, \dots, y_L^*$ be all the unique values found in the populations, i.e. values such that

$$x_1^* < x_2^* < \dots < x_K^* \qquad \text{and} \qquad \{x_1^*, x_2^*, \dots, x_K^*\} = \{x_1, x_2, x_3, \dots, x_N\}$$

$$y_1^* < y_2^* < \dots < y_L^* \qquad \text{and} \qquad \{y_1^*, y_2^*, \dots, y_L^*\} = \{y_1, y_2, y_3, \dots, y_N\}$$

Two populations: Joint frequencies

For $i = 1, 2, \ldots, K$ and for $j = 1, 2, \ldots, L$, let

$$f_{ij} = \left|\left\{ (\omega', \omega'') \in \Omega \times \Omega : (x_{\omega'}, y_{\omega''}) = \left(x_i^*, y_j^*\right) \right\}\right|$$

be the **joint frequency** of the pair $\left(x_i^*, y_j^*\right)$ in the population of the unique pairs.

Two populations: Marginal frequencies

For $i = 1, 2, \ldots, K$, let

$$f_{i\cdot} = \sum_{j=1}^{L} f_{ij} = |\{\omega \in \Omega : x_\omega = x_i^*\}|$$

be the **marginal frequency** of the value x_i^* in the first population.

For $j = 1, 2, \ldots, L$, let

$$f_{\cdot j} = \sum_{i=1}^{K} f_{ij} = |\{\omega \in \Omega : y_\omega = y_j^*\}|$$

be the **marginal frequency** of the value y_j^* in the second population.

Contingency table — for the population

the observed frequencies of the pairs (x_i^*, y_j^*) for $i = 1, \dots, K$ and for $j = 1, \dots, L$

$X \setminus Y$	y_1^*	y_2^*	…	y_L^*	TOTAL
x_1^*	f_{11}	f_{12}	…	f_{1L}	$f_{1\cdot}$
x_2^*	f_{21}	f_{22}	…	f_{2L}	$f_{2\cdot}$
…	⋮	⋮	…	⋮	⋮
x_K^*	f_{K1}	f_{K2}	…	f_{KL}	$f_{K\cdot}$
TOTAL	$f_{\cdot 1}$	$f_{\cdot 2}$	…	$f_{\cdot L}$	N

marginal frequencies

marginal frequencies

the population size

Two samples

Let

$$\Omega' = \{\omega_1, \omega_2, \ldots, \omega_n\} \subseteq \{1, 2, 3, \ldots, N\} = \Omega$$

be a selection out of the underlying set of the data units.

(We assume $n > 1$ and $\omega_i \neq \omega_j$ if $i \neq j$.)

We then have **two paired samples:**

$$x_{\omega_1}, x_{\omega_2}, \ldots, x_{\omega_n} \qquad \text{and} \qquad y_{\omega_1}, y_{\omega_2}, \ldots, y_{\omega_n}$$

Two samples: Joint frequencies

Now, let $x_1^*, x_2^*, \dots, x_k^*$ and $y_1^*, y_2^*, \dots, y_l^*$ be all the unique values found in the samples, i.e. values such that

$$x_1^* < x_2^* < \cdots < x_k^* \qquad \text{and} \qquad \{x_1^*, x_2^*, \dots, x_k^*\} = \{x_{\omega_1}, x_{\omega_2}, x_{\omega_3}, \dots, x_{\omega_n}\}$$

$$y_1^* < y_2^* < \cdots < y_l^* \qquad \text{and} \qquad \{y_1^*, y_2^*, \dots, y_l^*\} = \{y_{\omega_1}, y_{\omega_2}, y_{\omega_3}, \dots, y_{\omega_n}\}$$

For $i = 1, 2, \dots, k$ and for $j = 1, 2, \dots, l$, let

$$f_{ij} = \left|\left\{ (\omega', \omega'') \in \Omega' \times \Omega' : (x_{\omega'}, y_{\omega''}) = \left(x_i^*, y_j^*\right) \right\}\right|$$

be the **joint frequency** of the pair $\left(x_i^*, y_j^*\right)$ in the population of the unique pairs.

Two samples: Marginal frequencies

For $i = 1, 2, \ldots, k,$ let

$$f_{i\cdot} = \sum_{j=1}^{l} f_{ij} = |\{\omega \in \Omega' : x_\omega = x_i^*\}|$$

be the **marginal frequency** of the value x_i^* in the first sample.

For $j = 1, 2, \ldots, l,$ let

$$f_{\cdot j} = \sum_{i=1}^{k} f_{ij} = |\{\omega \in \Omega' : y_\omega = y_j^*\}|$$

be the **marginal frequency** of the value y_j^* in the second sample.

Contingency table — for the sample

the observed frequencies of the pairs (x_i^*, y_j^*) for $i = 1, \dots, k$ and for $j = 1, \dots, l$

x \ Y	y_1^*	y_2^*	…	y_l^*	TOTAL
x_1^*	f_{11}	f_{12}	…	f_{1l}	$f_{1\cdot}$
x_2^*	f_{21}	f_{22}	…	f_{2l}	$f_{2\cdot}$
…	⋮	⋮	…	⋮	⋮
x_k^*	f_{k1}	f_{k2}	…	f_{kl}	$f_{k\cdot}$
TOTAL	$f_{\cdot 1}$	$f_{\cdot 2}$	…	$f_{\cdot l}$	n

marginal frequencies

marginal frequencies

the sample size

Arithmetic means

Population arithmetic means:

$$\mu_X = \frac{1}{N}\sum_{i=1}^{K}\sum_{j=1}^{L} f_{ij} \times x_i^* \quad \text{and} \quad \mu_Y = \frac{1}{N}\sum_{j=1}^{L}\sum_{i=1}^{K} f_{ij} \times y_j^*$$

Sample arithmetic means:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{l} f_{ij} \times x_i^* \quad \text{and} \quad \bar{y} = \frac{1}{n}\sum_{j=1}^{l}\sum_{i=1}^{k} f_{ij} \times y_j^*$$

Variances

Population variances:

$$\sigma_X^2 = \frac{1}{N}\sum_{i=1}^{K}\sum_{j=1}^{L} f_{ij} \times (x_i^* - \mu_X)^2 \quad \text{and} \quad \sigma_Y^2 = \frac{1}{N}\sum_{j=1}^{L}\sum_{i=1}^{K} f_{ij} \times \left(y_j^* - \mu_Y\right)^2$$

Sample variances:

$$s_X^2 = \frac{1}{n-1}\sum_{i=1}^{k}\sum_{j=1}^{l} f_{ij} \times (x_i^* - \bar{x})^2 \quad \text{and} \quad s_Y^2 = \frac{1}{n-1}\sum_{j=1}^{l}\sum_{i=1}^{k} f_{ij} \times \left(y_j^* - \bar{y}\right)^2$$

Covariance

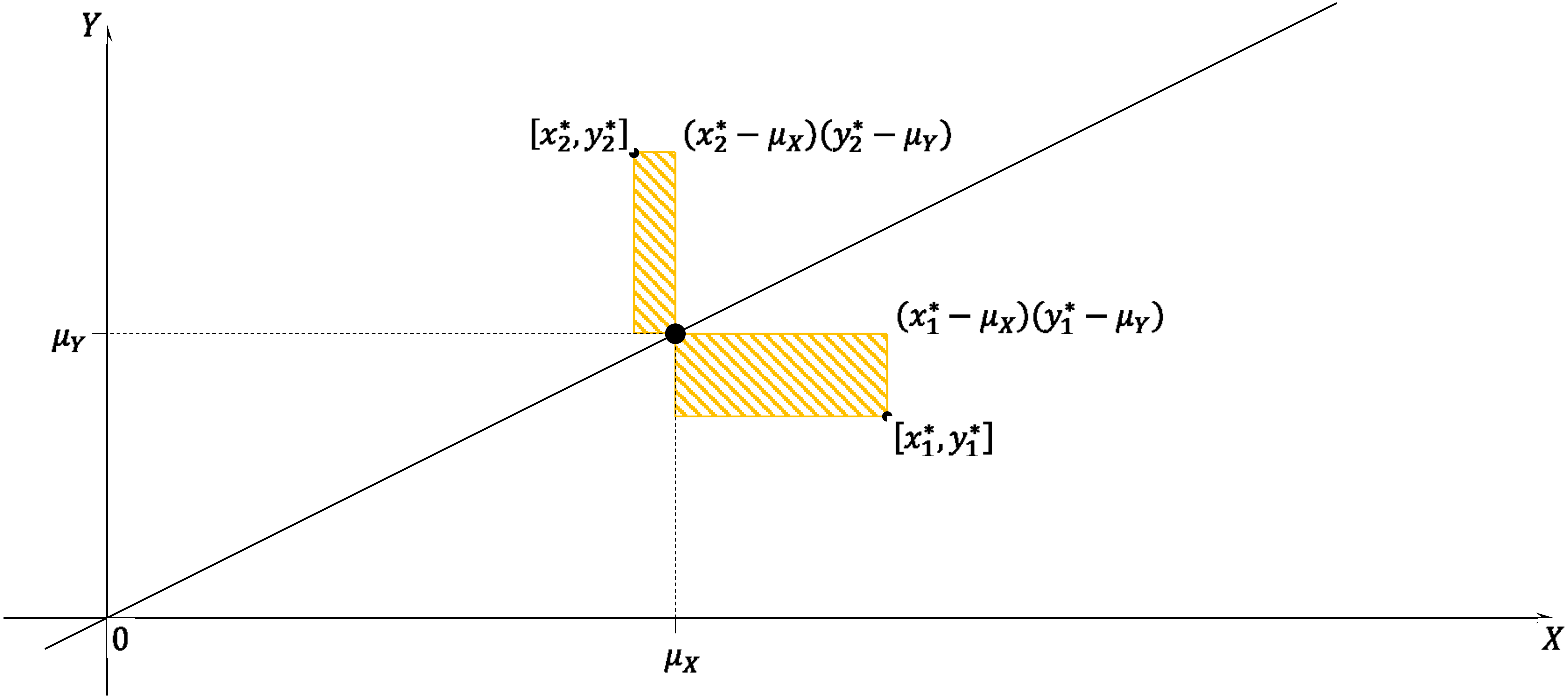
Population co-variance:

$$\operatorname{cov}(X,Y) = \frac{1}{N}\sum_{i=1}^{K}\sum_{j=1}^{L} f_{ij} \times (x_i^* - \mu_X)(y_j^* - \mu_Y)$$

Sample co-variance:

$$c_{XY} = \frac{1}{n-1}\sum_{i=1}^{k}\sum_{j=1}^{l} f_{ij} \times (x_i^* - \bar{x})(y_j^* - \bar{y})$$

Covariance



Pearson paired correlation coefficient

<u>Population</u> paired correlation coefficient:

$$\rho = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Var}(X) \times \mathrm{Var}(Y)}}$$

<u>Sample</u> paired correlation coefficient:

$$r = \frac{c_{XY}}{s_X s_Y}$$

The expected values of the functions of random variables

- The expected value of the sample mean
- Independent events
- Independent random variables
- The variance of the expected value of the sample mean
- The expected value of the sample variance

The expected value of the sample mean

Assume that the expected values $\mathrm{E}[X_1] = \mathrm{E}[X_2] = \cdots = \mathrm{E}[X_n] = \mu$.

Then

$$\mathrm{E}[\bar{X}] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[X_i] = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$$

Independent events

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

We say that events $A, B \in \mathcal{F}$ are **independent** if and only if

$$P(A \cap B) = P(A) \times P(B)$$

so that

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) \quad \text{and} \quad P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$$

$P(B) \neq 0$ $\qquad$ $P(A) \neq 0$

Independent random variables

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

We say that random variables $X, Y : \Omega \to \mathbb{R}$ are **independent** if and only if

$$P(\{\omega \in \Omega : X(\omega) \leq a\} \cap \{\omega \in \Omega : Y(\omega) \leq b\}) =$$

$$= P(\{\omega \in \Omega : X(\omega) \leq a\}) \times P(\{\omega \in \Omega : Y(\omega) \leq b\}) \qquad \text{for every} \quad a, b \in \mathbb{R}$$

in short:

$$P(\{X \leq a\} \cap \{Y \leq b\}) = P(\{X \leq a\}) \times P(\{Y \leq b\}) \qquad \text{for every} \quad a, b \in \mathbb{R}$$

Independent random variables: Theorem

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X, Y: \Omega \to \mathbb{R}$ be independent random variables such that the expected values $\mathrm{E}[|X|]$ and $\mathrm{E}[|Y|]$ are finite.

Then

$$\mathrm{E}[X \times Y] = \mathrm{E}[X] \times \mathrm{E}[Y]$$

We prove this statement in the case I, when the sample space Ω is finite ($\Omega = \{1, 2, \dots, N\}$). The proof uses limiting steps and some advanced results (Levi's Theorem) of the theory of measures and the Lebesgue integral.

Independent random variables: E(*XY*) = E(*X*) E(*Y*)

Proof (in the case I): Let

$$\{x_1, x_2, \dots, x_m\} = \{X(\omega) : \omega \in \Omega\} \quad \text{and} \quad \{y_1, y_2, \dots, y_n\} = \{Y(\omega) : \omega \in \Omega\}$$

be the ranges of the random variables X and Y, and let the ranges be finite.

(If the sample space Ω is finite [the case I], then so are the ranges.) Then

$$\mathrm{E}[X \times Y] = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \times y_j \times P(\{X = x_i\} \cap \{Y = y_j\})$$

Independent random variables: E(*XY*) = E(*X*) E(*Y*)

$$E[X \times Y] = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \times y_j \times P(\{X = x_i\} \cap \{Y = y_j\}) =$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \times y_j \times P(\{X = x_i\}) \times P(\{Y = y_j\}) =$$

$$= \sum_{i=1}^{m} x_i \times P(\{X = x_i\}) \times \sum_{j=1}^{n} y_j \times P(\{Y = y_j\}) = E[X] \times E[Y]$$

Independent random variables: Theorem II

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X', X'' : \Omega \to \mathbb{R}$ be <u>independent random variables</u> such that the expected values $\mu' = \mathrm{E}(|X'|)$ and $\mu'' = \mathrm{E}(|X''|)$ are finite. Then

$$\mathrm{E}[(X' - \mu')(X'' - \mu'')] = 0$$

Proof:

$$\begin{aligned}\mathrm{E}[(X' - \mu')(X'' - \mu'')] &= \mathrm{E}[X'X'' - X'\mu'' - \mu'X'' + \mu'\mu''] = \\ &= \mathrm{E}[X'X''] - \mathrm{E}[X'\mu''] - \mathrm{E}\,[\mu'X''] + \mathrm{E}[\mu'\mu''] = \\ &= \mathrm{E}[X']\mathrm{E}[X''] - \mathrm{E}[X']\mu'' - \mu'\mathrm{E}[X''] + \mu'\mu'' =\end{aligned}$$

The variance of the sample mean

Assume that the variances $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \cdots = \mathrm{Var}(X_n) = \sigma^2$

and that the random variables $X_1, X_2, \ldots, X_n$ are pairwise independent.

Then

$$\mathrm{Var}[\bar{X}] = \mathrm{E}[(\bar{X} - \mathrm{E}[\bar{X}])^2] = \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right)^2\right] = \mathrm{E}\left[\frac{(\sum_{i=1}^{n}(X_i - \mu))^2}{n^2}\right]$$

$$= \frac{1}{n^2}\mathrm{E}\left[\sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i \neq j}}^{n}(X_i - \mu)(X_j - \mu)\right] =$$

The variance of the sample mean

$$\mathrm{Var}[\bar{X}] = \frac{1}{n^2}\mathrm{E}\left[\sum_{i=1}^{n}(X_i-\mu)^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}(X_i-\mu)(X_j-\mu)\right] =$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathrm{E}[(X_i-\mu)^2] + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\mathrm{E}\big[(X_i-\mu)(X_j-\mu)\big]\right) =$$

The variance of the sample mean

If X_i and X_j are independent, then $\mathrm{E}\left[(X_i - \mu)(X_j - \mu)\right] = 0$

$$\mathrm{Var}[\bar{X}] = \frac{1}{n^2}\left(\sum_{i=1}^{n} \mathrm{E}[(X_i - \mu)^2] + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} \mathrm{E}\left[(X_i - \mu)(X_j - \mu)\right]\right) =$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \mathrm{E}[(X_i - \mu)^2]\right) =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

The expected value of the sample variance

Assume that the expected values $\mathrm{E}[X_1] = \mathrm{E}[X_2] = \cdots = \mathrm{E}[X_n] = \mu$,
that the variances $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \cdots = \mathrm{Var}(X_n) = \sigma^2$,
and that the random variables $X_1, X_2, \ldots, X_n$ are pairwise independent.
Then

$$\mathrm{E}[S^2] = \mathrm{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \mathrm{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right] =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{E}\left[\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right] =$$

The expected value of the sample variance

$$\mathrm{E}[S^2] = \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{E}\left[\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right] =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{E}\left[X_i^2 - \frac{2}{n}X_i\sum_{j=1}^{n}X_j + \frac{1}{n^2}\sum_{j=1}^{n}X_j\sum_{k=1}^{n}X_k\right] =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{E}\left[X_i^2 - \frac{2}{n}X_i^2 - \frac{2}{n}\sum_{\substack{j=1\\ i\neq j}}^{n}X_iX_j + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1\\ j\neq k}}^{n}X_jX_k + \frac{1}{n^2}\sum_{j=1}^{n}X_j^2\right] =$$

The expected value of the sample variance

$$\mathrm{E}[S^2] = \frac{1}{n-1}\sum_{i=1}^{n} \mathrm{E}\left[X_i^2 - \frac{2}{n}X_i^2 - \frac{2}{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} X_i X_j + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1 \\ j\neq k}}^{n} X_j X_k + \frac{1}{n^2}\sum_{j=1}^{n} X_j^2\right] =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{n-2}{n}\mathrm{E}[X_i^2] - \frac{2}{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} \mathrm{E}[X_i X_j] + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1 \\ j\neq k}}^{n} \mathrm{E}[X_j X_k] + \frac{1}{n^2}\sum_{j=1}^{n} \mathrm{E}[X_j^2]\right) =$$

The expected value of the sample variance

Recall that $\mathrm{E}[X_1] = \cdots = \mathrm{E}[X_n] = \mu$ and $\mathrm{Var}(X_1) = \cdots = \mathrm{Var}(X_n) = \sigma^2$, and $\sigma^2 = \mathrm{Var}(X_i) = \mathrm{E}[X_i^2] - (\mathrm{E}[X_i])^2 = \mathrm{E}[X_i^2] - \mu^2$ in general. Hence $\mathrm{E}[X_i^2] = \mu^2 + \sigma^2$ for every $i = 1, \dots, n$. Since X_i and X_j are independent, we have $\mathrm{E}[X_i X_j] = \mathrm{E}[X_i]\mathrm{E}[X_j] = \mu^2$ for every $i, j = 1, \dots, n$ when $i \neq j$. Therefore

$$\mathrm{E}[S^2] = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{n-2}{n}\mathrm{E}[X_i^2] - \frac{2}{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\mathrm{E}[X_i X_j] + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{\substack{k=1\\ j\neq k}}^{n}\mathrm{E}[X_j X_k] + \frac{1}{n^2}\sum_{j=1}^{n}\mathrm{E}[X_j^2]\right) =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{n-2}{n}(\mu^2+\sigma^2) - 2\frac{n-1}{n}\mu^2 + \frac{n(n-1)}{n^2}\mu^2 + \frac{n}{n^2}(\mu^2+\sigma^2)\right) =$$

The expected value of the sample variance

$$\mathrm{E}[S^2] = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{n-2}{n}(\mu^2+\sigma^2) - 2\frac{n-1}{n}\mu^2 + \frac{n(n-1)}{n^2}\mu^2 + \frac{n}{n^2}(\mu^2+\sigma^2)\right) =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{n-1}{n}(\mu^2+\sigma^2) - \frac{n-1}{n}\mu^2\right) =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{n-1}{n}\sigma^2\right) = \sigma^2$$

Alternative formula for sample variance

We have noticed that the **sample variance** satisfies the next equation:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2$$

To see the equation, note that:

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2 = \sum_{i=1}^{n}\left(X_i^2 - \frac{2}{n}X_i\sum_{j=1}^{n}X_j + \frac{1}{n^2}\sum_{j=1}^{n}X_j\sum_{k=1}^{n}X_k\right) =$$

$$= \sum_{i=1}^{n}X_i^2 - \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_j + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}X_jX_k =$$

Alternative formula for sample variance

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} X_j X_k =$$

$$= \sum_{i=1}^{n} X_i^2 - \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j + \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n} X_j X_k =$$

$$= \sum_{i=1}^{n} X_i^2 - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j =$$

$$= \frac{1}{2n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} X_i^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j + \sum_{i=1}^{n}\sum_{j=1}^{n} X_j^2\right) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(X_i - X_j\right)^2$$