Probability statements about population based on the sample

1. **Choice of a distribution model**: Comparison of the empiric histogram with theoretical density functions.

2. **Estimation of parameters of the population.** Computation of the estimates of population parameters from the sample following computation rules for estimators.

Example:

- Corn field with 100.000 plants
- Plants height measured in cm (x)

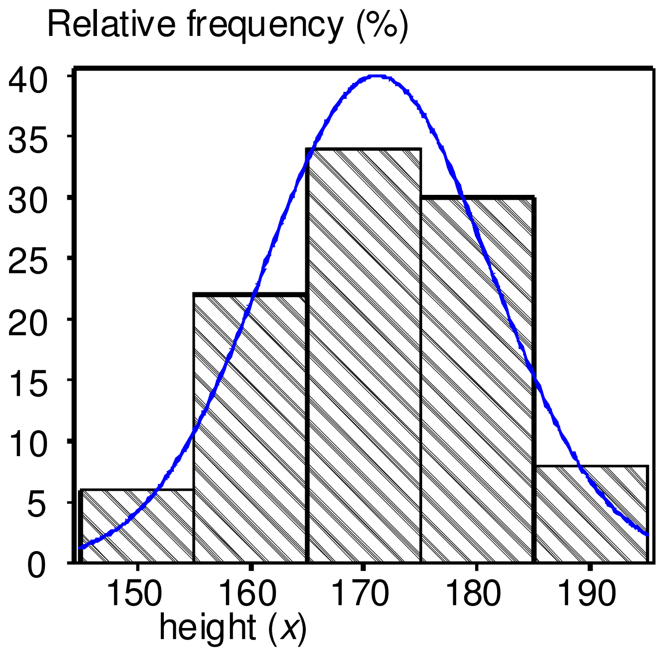
Choice of a distribution model: From comparison of the empiric histogram with theoretical density functions we assume, that the height are normally distributed.

Estimation of parameters of the population: Estimates of population parameters from the sample are as follows:

$$\hat{\mu} = 170; \ \hat{\sigma} = 10$$

Histogram of corn data with overlying density function of a normal distribution:

$\mu = 170; \ \sigma = 10$



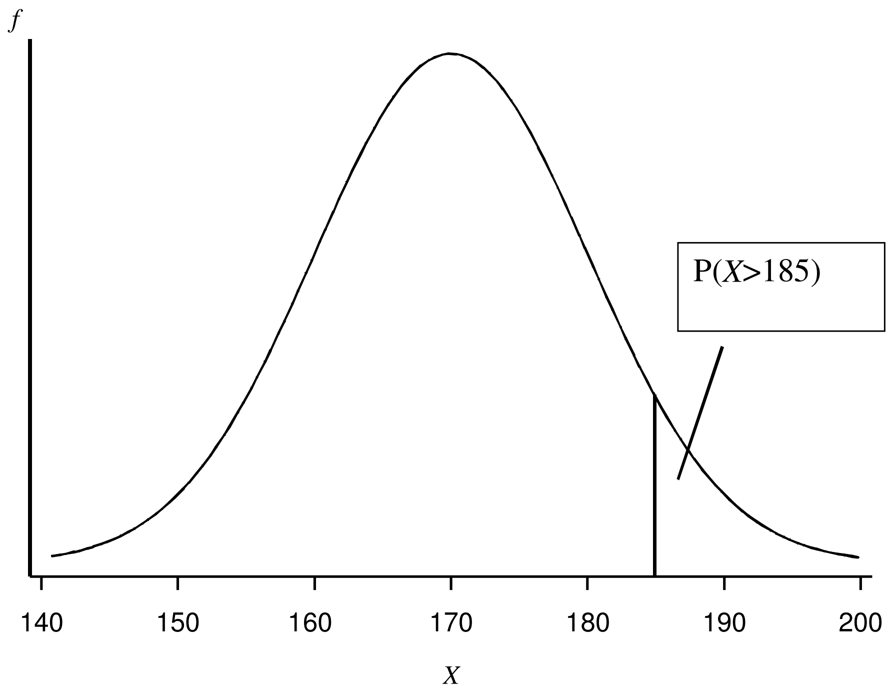
(1) about 68% of the corn plants have a height between 170 – 10 and 170 + 10 cm, so between 160 and 180 cm.

(2) about 95% of the corn plants have a height between 170 – 2*10 and 170 + 2*10 cm, so between 150 and 190 cm.

(3) about 99.7% of the corn plants have a height between 170 – 3*10 and 170 + 3*10 cm, so between 140 and 200 cm.

Calculation of probabilities. Variable: Height

Question 1: Compute a probability, that one randomly taken plant of the field has a height of $X = 185$ cm or longer



$$P(X > 185) = \int_{185}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{185}^{\infty} f(x)dx = 0.0668$$

Question: How we can compute the probability manually?

We can use the table for standard normal distribution.

The **standard normal distribution** is a distribution with $\mu = 0$ and $\sigma^2 = 1$

Z-Transformation and Standard Normal Distribution

The simplest case of the normal distribution, known as the Standard Normal Distribution, has the expected value zero and the variance one.

This is written as $\boldsymbol{N(0,1)}$.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

- The standardization of an arbitrary normal distribution is possible
- We need a table only for one distribution!

Z-Transformation:

$$Z = \frac{(X-\mu)}{\sigma}$$

Z-Transformation for height of $X = 185$: $\mu = 170$; $\sigma = 10$

$$z = \frac{(185 - 170)}{10} = 1.5$$

$$P(X > 1.5) = \int_{1.5}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \int_{1.5}^{\infty} f(z) dz = 0.0668$$

The area under the standard normal distribution to the left of $z = 1.5$ is equal to a probability, that $X > 185$.

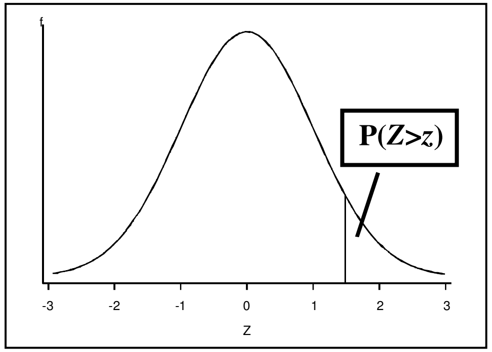
In General:

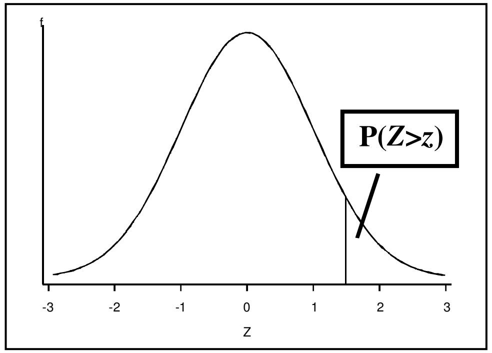
The probability $P(Z > z)$ is called **exceedance probability.**

It represents a the probability that a standard normal random variable (Z) is greater than a given value (z)

Exceedance probability $P(Z > z)$ for standard normal distribution. Example: $P(Z > 1.96) = 0.025$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014



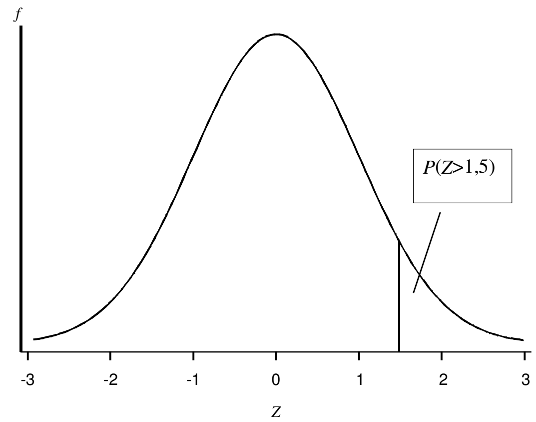


Exceedance probability $P(Z > z)$ for standard normal distribution.

Example: $P(Z > 0.33) = 0.3707$

z	0.00	0.01	0.02	0.03	0.04
0.0	0.5000	0.4960	0.4920	0.4880	0.4840
0.1	0.4602	0.4562	0.4522	0.4483	0.4443
0.2	0.4207	0.4168	0.4129	0.4090	0.4052
0.3	0.3821	0.3783	0.3745	**0.3707**	0.3669
0.4	0.3446	0.3409	0.3372	0.3336	0.3300
0.5	0.3085	0.3050	0.3015	0.2981	0.2946

Example:



$$\bar{x} = 185;\ \sigma = 10$$

Then

$$z = \frac{(x - \mu)}{\sigma} = \frac{185 - 170}{10} = 1.5$$

$$P(X > 185) = P(Z > 1.5) = 0.0668$$

In General:

If X is normal distributed with mean μ and variance σ^2

then

$$Z = \frac{X - \mu}{\sigma}$$

is standard normal distributed with mean 0 and variance 1

Question 2: Compute a probability, that the corn plants have a height between 175 and 185 cm.

$$\mu = 170; \ \sigma = 10$$

Then

$$z_1 = \frac{(x-\mu)}{\sigma} = \frac{175-170}{10} = 0.5$$

$$z_2 = \frac{(x-\mu)}{\sigma} = \frac{185-170}{10} = 1.5$$

From the table:

$P(Z > 1.5) = 0.0668$ and $P(Z > 0.5) = 0.3085$

So the area between 175 and 185 is equal to $0.3085 - 0.0668 = 0.2417$

The probability, that the corn plants have a height between 175 and 185 cm is 0.2417 or 24.17%

Question 3: Compute a probability, that the corn plants have a height smaller than 145 cm.

$$z_1 = \frac{(x - \mu)}{\sigma} = \frac{145 - 170}{10} = -2.5$$

$$P(Z < -2.5) = P(Z > 2.5) = 0.0062$$

Distribution symmetry

The probability, that the corn plants have a height smaller than 145 cm is 0.0062 or 0.62%

Question 4: Compute a probability, that the corn plants have a height between 145 and 175cm.

$$P(145 < Z < 175) - ?$$

Full area under the standard normal distribution curve

$$P(145 < Z < 175) = P(-2.5 < Z < 0.5) = 1 - 0.0062 - 0.3085 = 0.6853$$

Area to the left of -2.5

Area to the right of 0.5

The probability, that the corn plants have a height between 145 and 175 cm is 0.6853 or 6.53%

Question 4: Compute a probability, that the corn plants are higher than 170 cm

$$z = \frac{170 - 170}{10} = 0$$

$$P(Z > 0) = 0.5$$

The probability, that the corn plants have a height greater than 170 cm is 0.5 or 50%

Important:

By continuous distributions $P(X = x) = 0$!

So

$$P(X > x) = P(X \geq x)$$

and

$$P(X < x) = P(X \leq x)$$

Central Limit Theorem

The theorem states that, given certain conditions, the sum or the mean of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed.

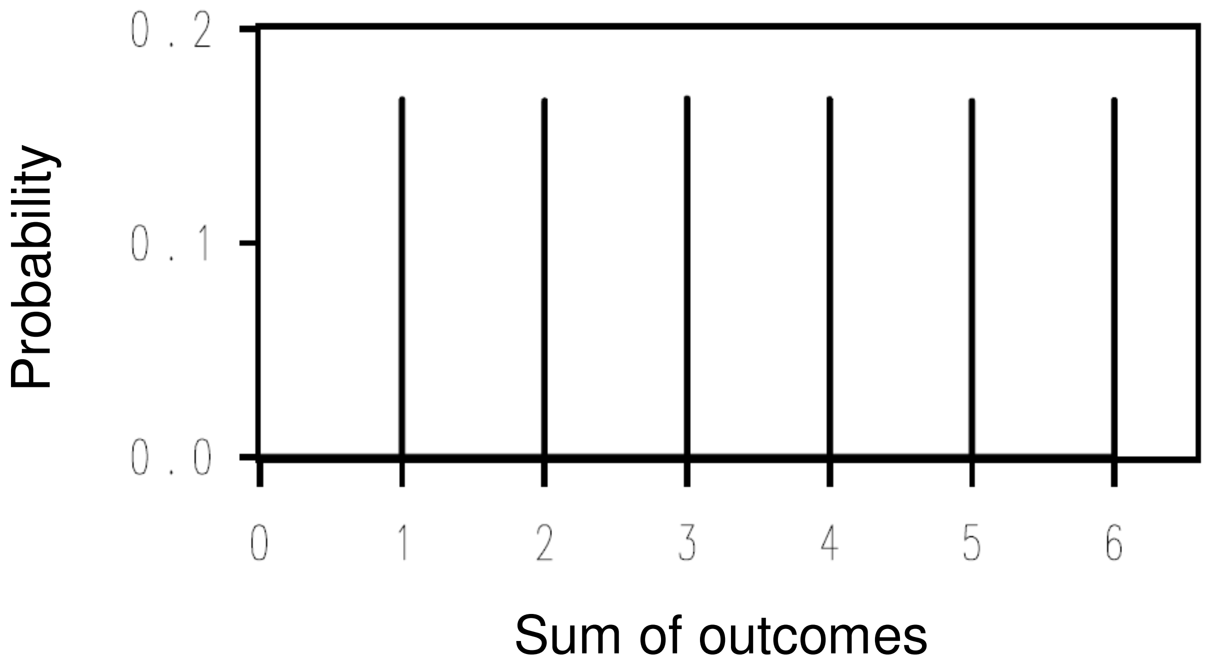
The Central Limit Theorem basically says that for non-normal data, the distribution of the sample means or sums has an approximate normal distribution, no matter what the distribution of the original data looks like, as long as the sample size is large enough (usually at least 30) and all samples have the same size.

Throw of six-sided dice: sum of outcomes

Sum of the two dice	Probable outcomes: 1.throw + 2.throw	Number of possible combination of dice /absolute frequency	Relative frequency
2	1+1	1	1/36≈0.038
3	1+2; 2+1	2	2/36≈0.056
4	1+3; 2+2; 3+1	3	3/36≈0.083
5	1+4; 2+3; 3+2; 4+1	4	4/36≈0.111
6	1+5; 2+4; 3+3; 4+2; 5+1	5	5/36≈0.139
7	1+6; 2+5; 3+4; 4+3; 5+2; 6+1	6	6/36≈0.167
8	2+6; 3+5; 4+4; 5+3; 6+2	5	5/36≈0.139
9	3+6; 4+5; 5+4; 6+3	4	4/36≈0.111
10	4+6; 5+5; 6+4	3	3/36≈0.083
11	5+6; 6+5	2	2/36≈0.056
12	6+6	1	1/36≈0.038
		$\sum = 36$	$\sum = 1$

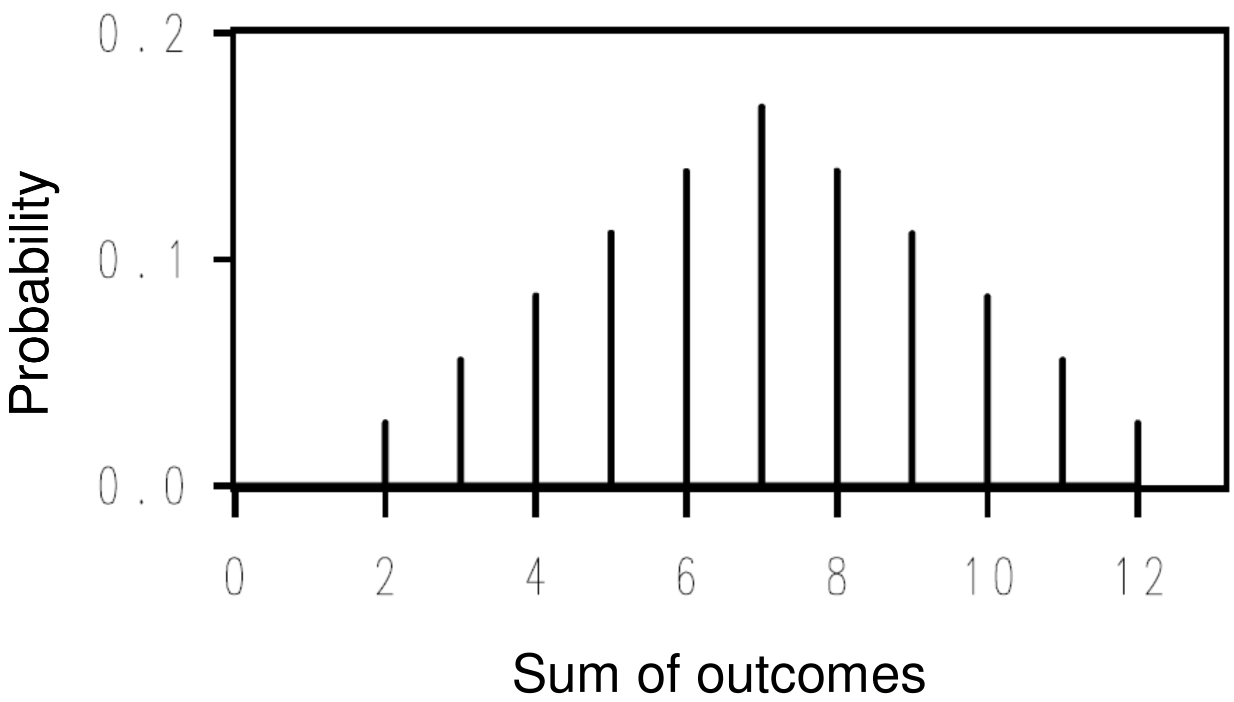
Throw of six-sided dice: sum of outcomes

1 throw. Probability of every outcome $= \frac{1}{6}$

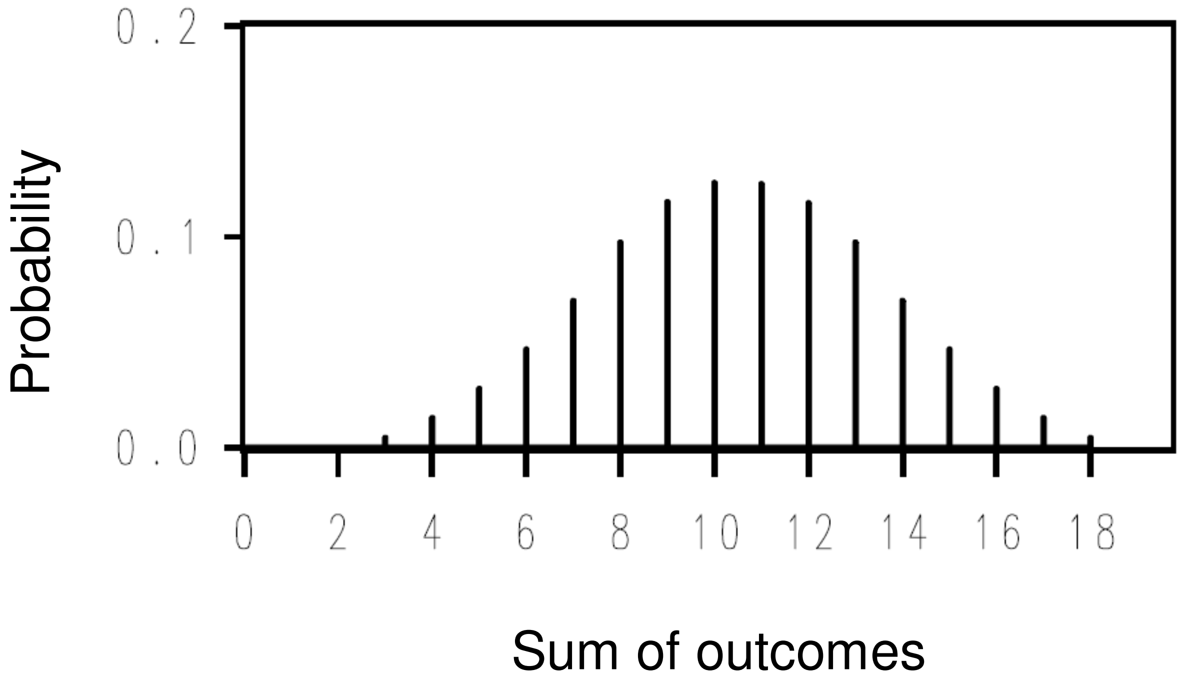


2 throws

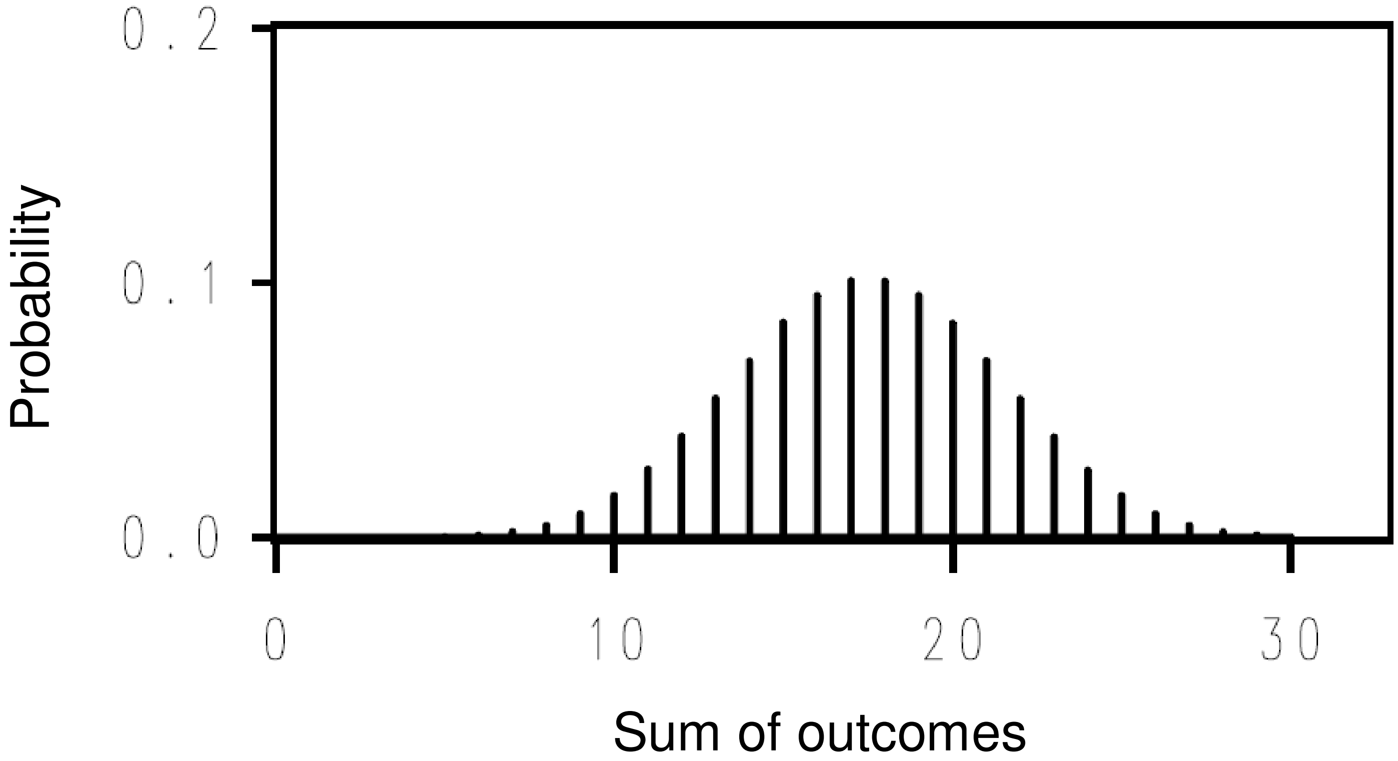
Sum of two throws	Probable outcomes: 1.throw + 2.throw	Number of possible combination of dice /absolute frequency	Relative frequency
2	1+1	1	1/36≈0.038
3	1+2; 2+1	2	2/36≈0.056
4	1+3; 2+2; 3+1	3	3/36≈0.083
5	1+4; 2+3; 3+2; 4+1	4	4/36≈0.111
6	1+5; 2+4; 3+3; 4+2; 5+1	5	5/36≈0.139
7	1+6; 2+5; 3+4; 4+3; 5+2; 6+1	6	6/36≈0.167
8	2+6; 3+5; 4+4; 5+3; 6+2	5	5/36≈0.139
9	3+6; 4+5; 5+4; 6+3	4	4/36≈0.111
10	4+6; 5+5; 6+4	3	3/36≈0.083
11	5+6; 6+5	2	2/36≈0.056
12	6+6	1	1/36≈0.038
		$\sum = 36$	$\sum = 1$


0.2
0.1
0.0
Probability
0
2
4
6
8
10
12
Sum of outcomes

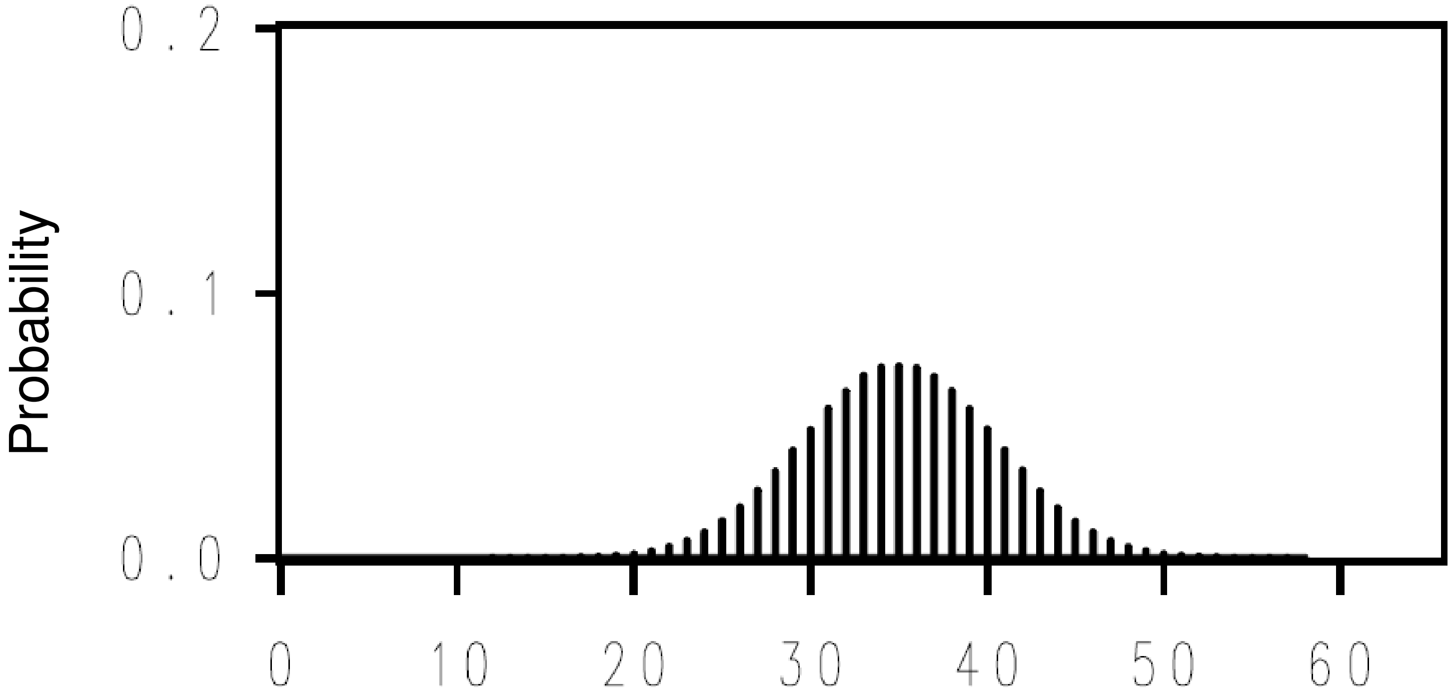
3 throws



5 throws



10 throws



Distribution of the sample mean

By the repeated sampling the sample mean $\bar{x}$ fluctuates around the same value μ as the sample values.

- The variance of the mean $\bar{x}$ is related to the variance of the sample values ${\sigma_x}^2$:

$$\sigma_{\bar{x}}{}^2 = \frac{{\sigma_x}^2}{n}$$

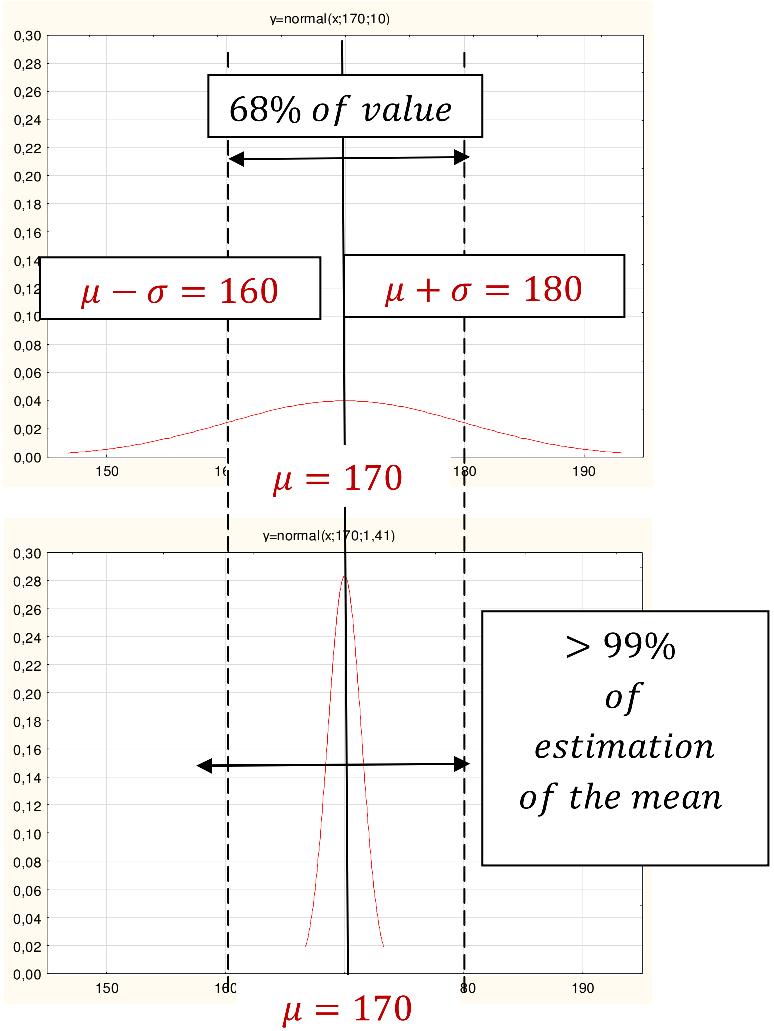
- The standard error of the mean is:

$$\sigma_{\bar{x}} = \sqrt{\frac{{\sigma_x}^2}{n}} = \frac{\sigma_x}{\sqrt{n}}$$

- Due of the Central limit theorem the sampling mean is approximately normal distributed; even so the individual values of the population are not.

- **Summary**: The mean $\bar{x}$ is approximately normal distributed with expectation μ and variance $\sigma_{\bar{x}}{}^2 = {\sigma_x}^2/n$

Example: Corn data, $n = 50,\ \mu = 170;\ \sigma = 10$



Distribution of the **individual** values of the sample:

The normal distribution with the mean $\mu = 170$

and the standard deviation $\sigma_x = 1.73$

Distribution of the **sample means** $\bar{x}_i$:

The normal distribution with the mean $\mu = 170$

and the standard deviation= standard error of the mean $\sigma_{\bar{x}} = 1.41$

⇒ The sample means lie much more closer to the mean, than the individual values

Confidence interval for population mean

A confidence interval is an interval, containing an unknown parameter with a predetermine probability.

Notation:

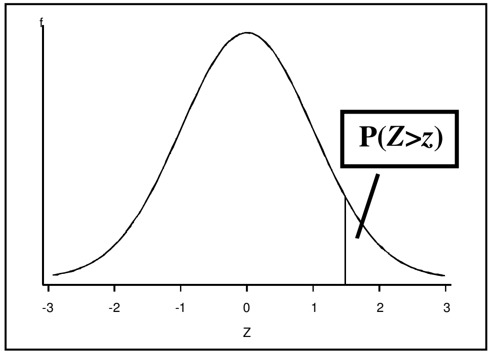
α: probability of error

$1-\alpha$: level of significance

The 95% confidence inteval overlaps the population mean μ with the probability of 95%. The probability of error equals $\alpha = 100\% - 95\% = 5\%$

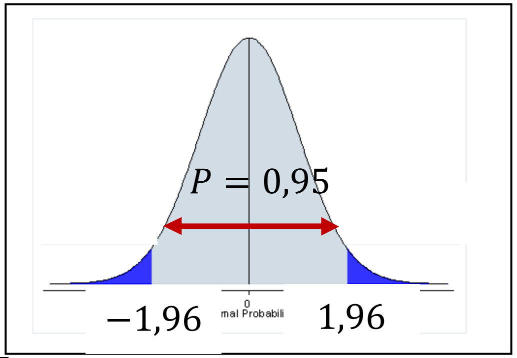
Exceedance probability $P(Z > z)$ for standard normal distribution. Example: $P(Z > 1.96) = 0.025$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014



Probability statements: Standard normal distribution

$$P(-1.96 < Z < 1.96) = 0.95$$



95% Confidence interval for the standardised random variable $\bar{X}$:

$$P\left(-1.96 < \frac{\bar{X} - \mu}{\sigma_x/\sqrt{n}} < 1.96\right) = 0.95$$

$$P\left(-1.96\frac{\sigma_x}{\sqrt{n}} < \bar{X} - \mu < 1.96\frac{\sigma_x}{\sqrt{n}}\right) = 0.95$$

$$P\left(\underbrace{\bar{X} - 1.96\frac{\sigma_x}{\sqrt{n}}}_{lower\ limit} < \mu < \underbrace{\bar{X} + 1.96\frac{\sigma_x}{\sqrt{n}}}_{upper\ limit}\right) = 0.95$$

Remember: Standardisation : $Z = \frac{X - \mu}{\sigma_x}$; $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

So $Z = \frac{\bar{X} - \mu}{\sigma_x/\sqrt{n}}$ Is the standardised estimator of the population mean

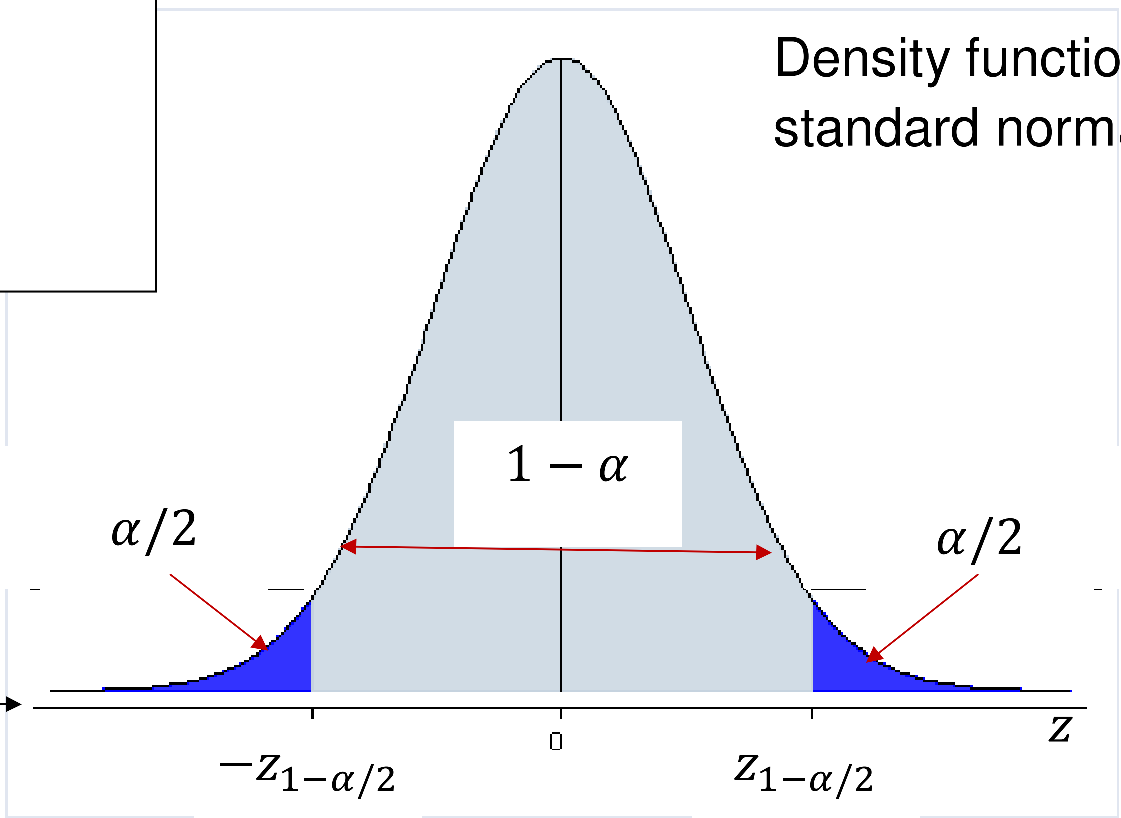
In general: $P\left(\bar{X} - z_{1-\alpha/2}\frac{\sigma_x}{\sqrt{n}} < \mu < \bar{X} + z_{1-\alpha/2}\frac{\sigma_x}{\sqrt{n}}\right) = 1-\alpha$

Probability

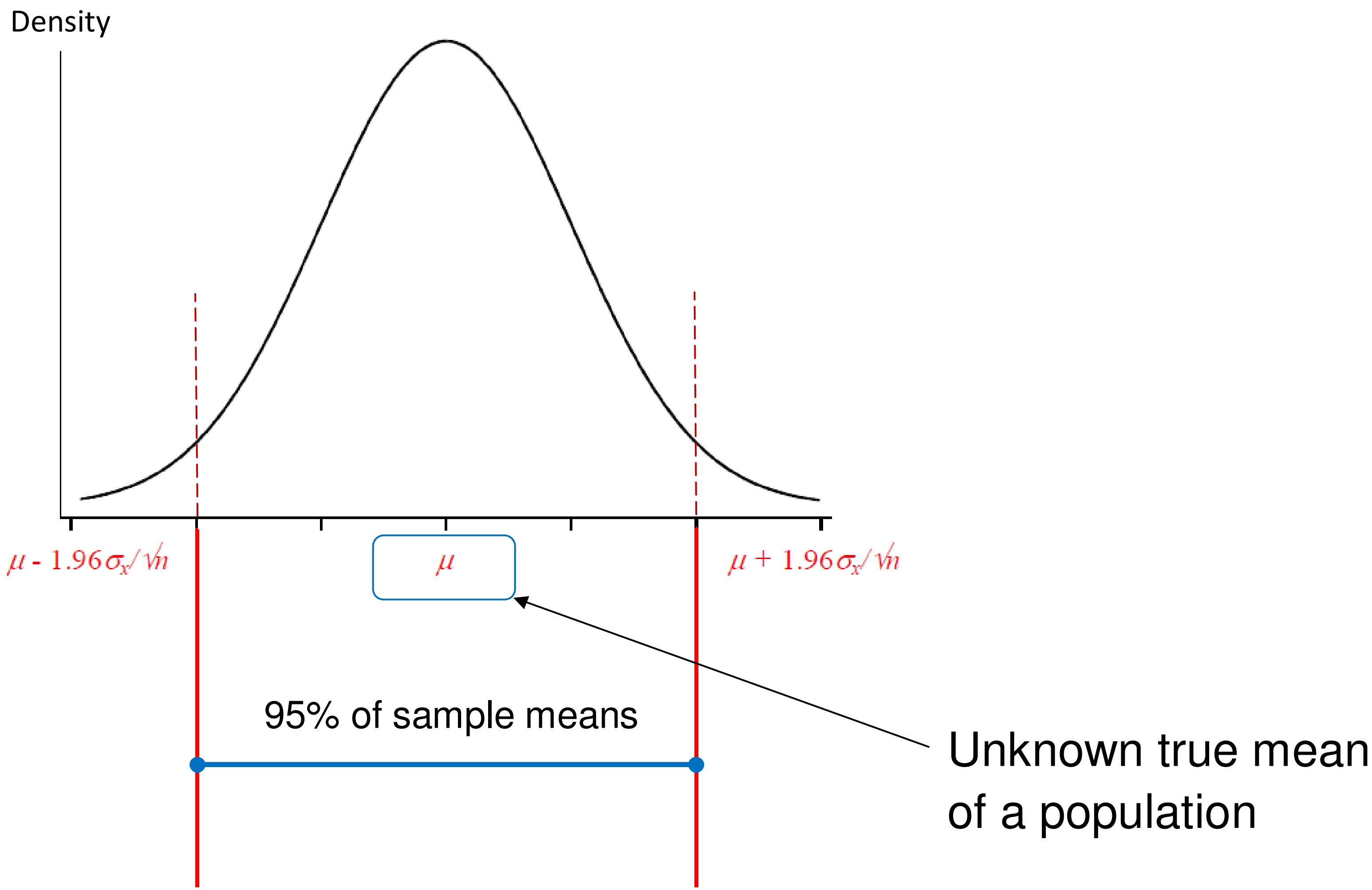
Area under the curve

A concrete value of the standard distributed random variable Z;

x-coordinate

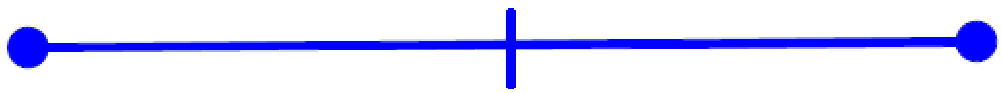


Notation: α: probability of error; $1-\alpha$: level of significance



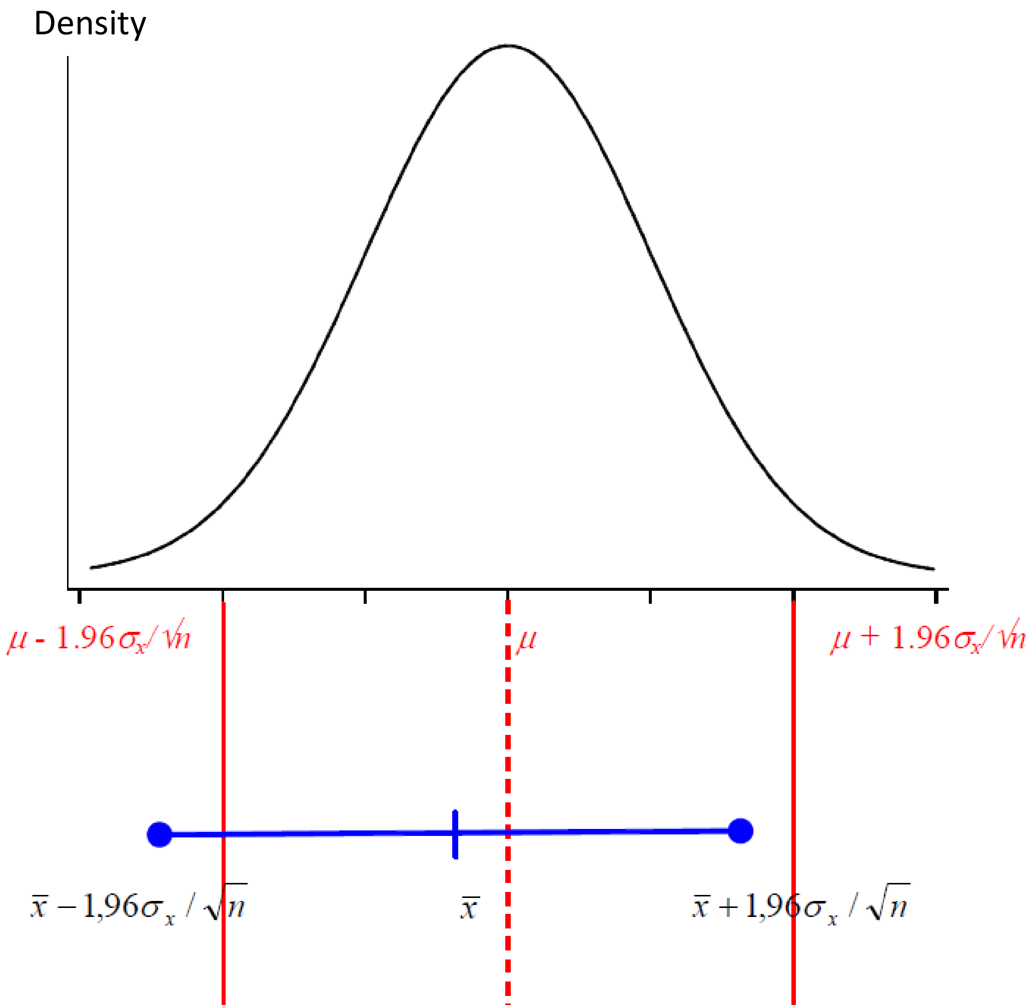
We use the length of the interval to create the confidence interval

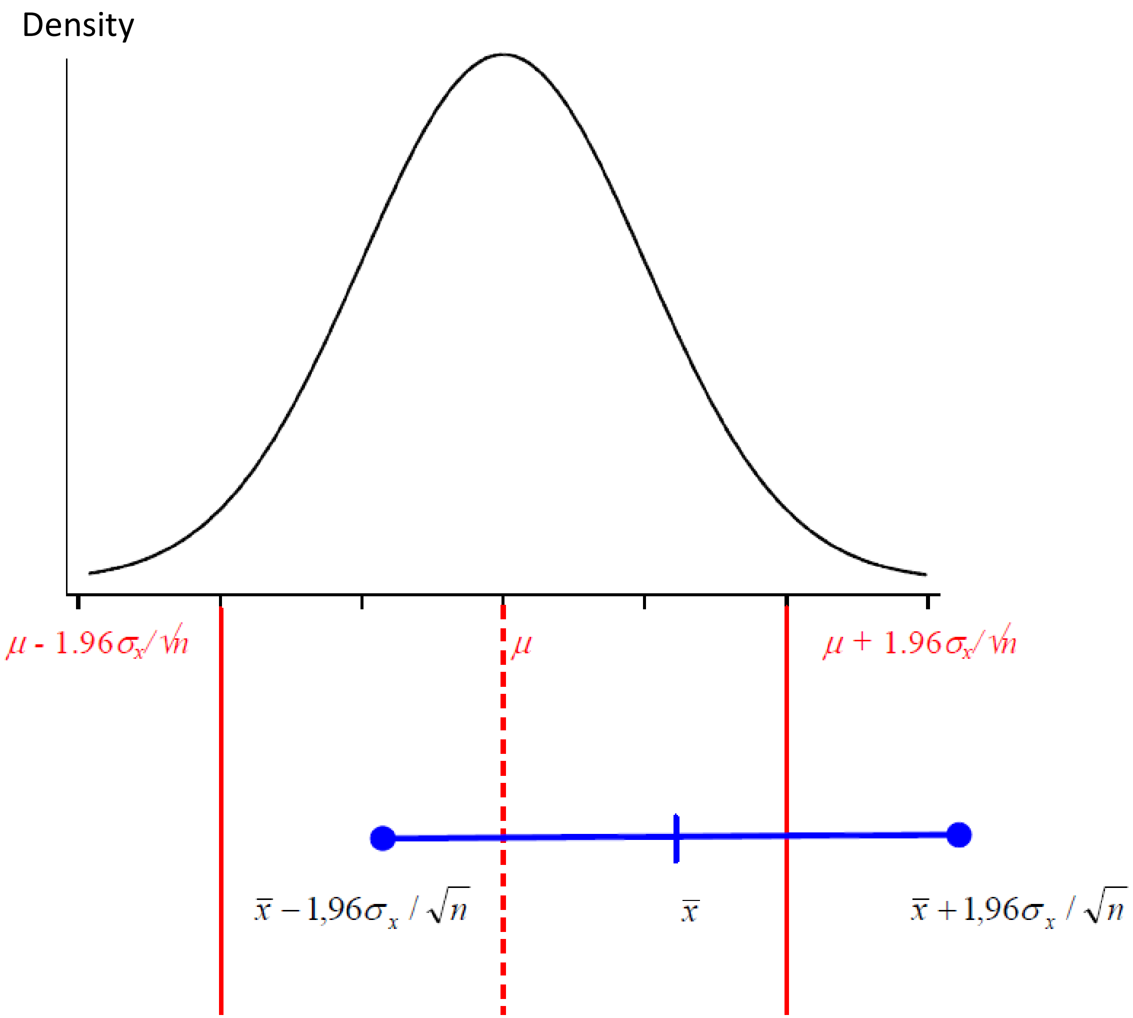
The confidence interval is a random variable!

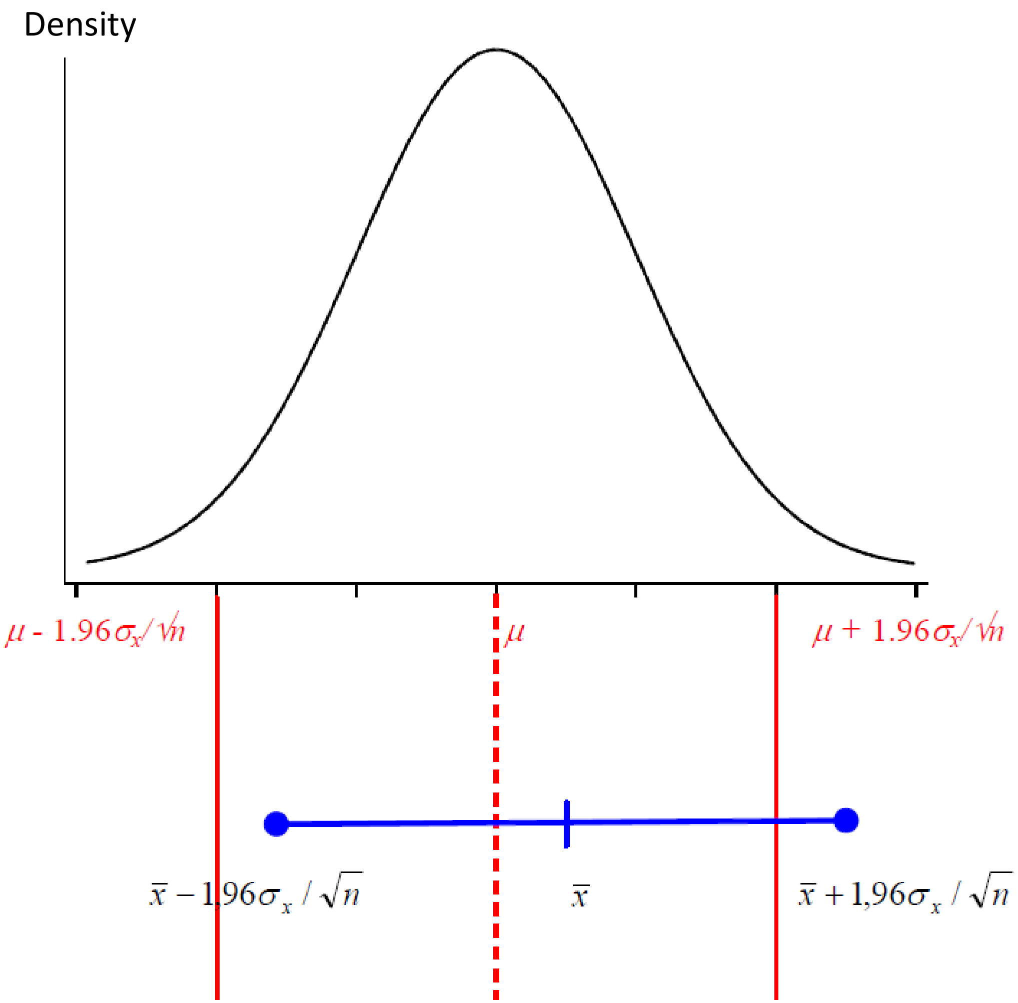


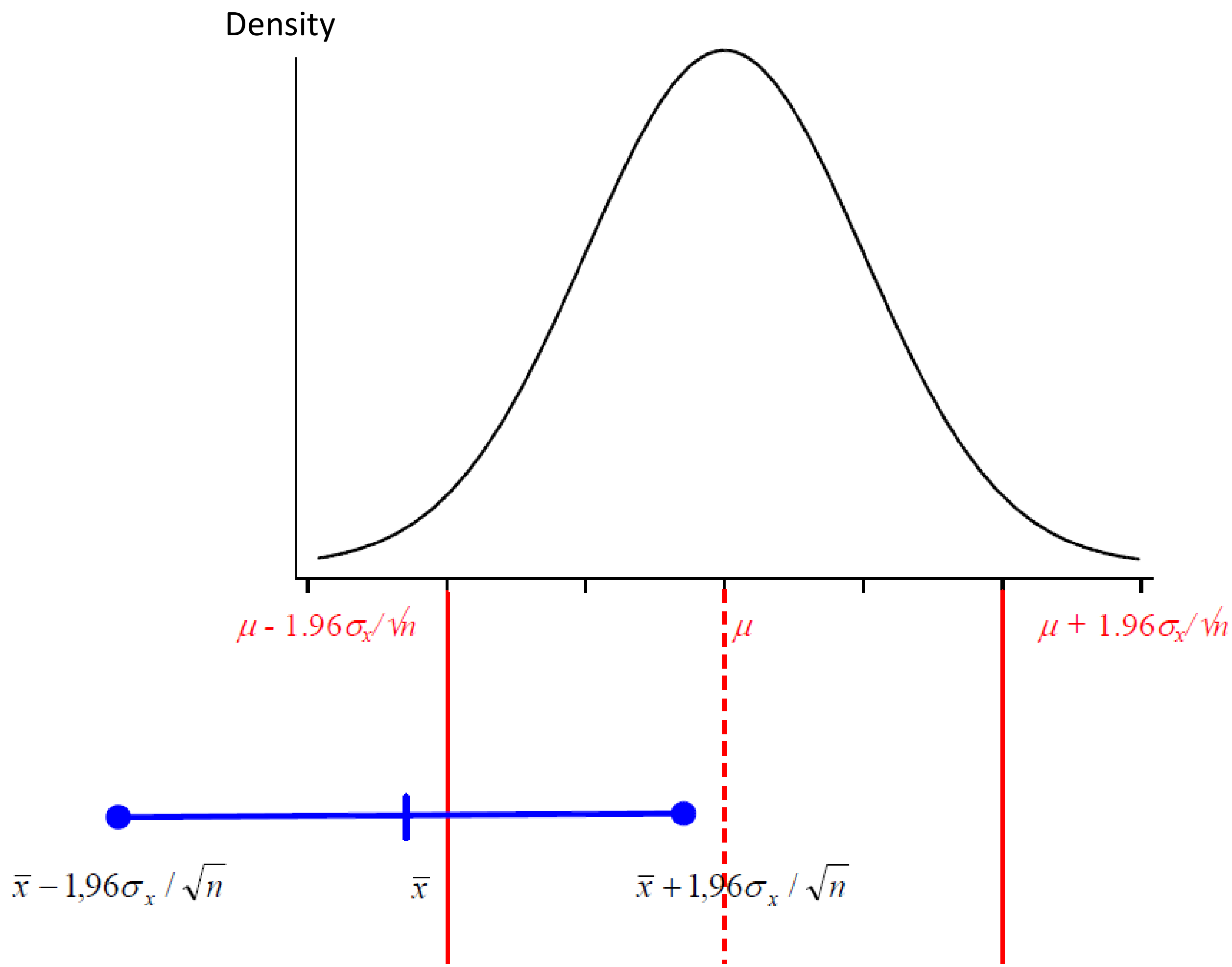
$\bar{x} - 1{,}96\sigma_x / \sqrt{n}$ $\quad$ $\bar{x}$ $\quad$ $\bar{x} + 1{,}96\sigma_x / \sqrt{n}$

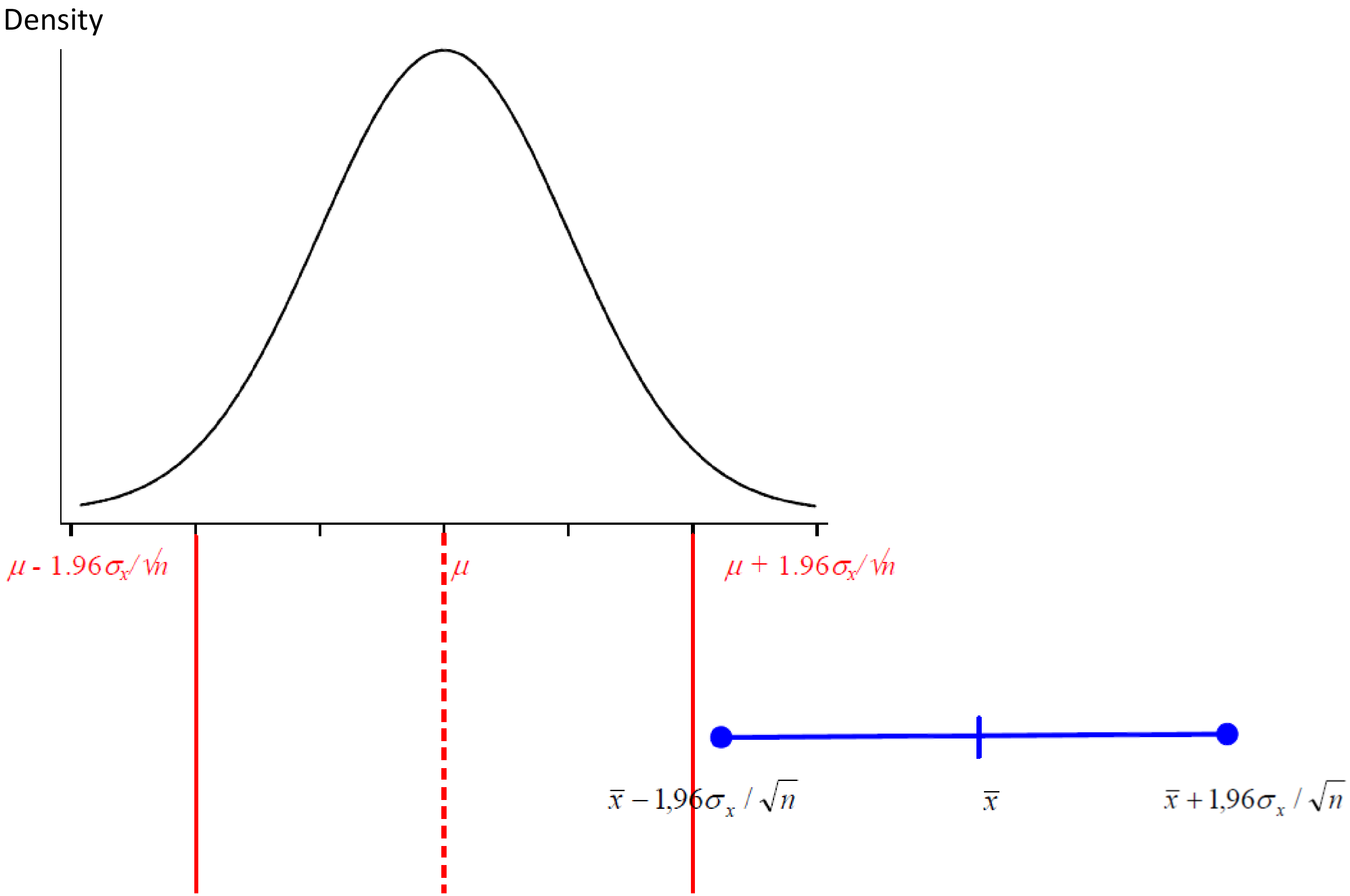
Possible constellations between the true mean μ and sample mean $\bar{x}$.




Density
μ - 1.96$\sigma_x/\sqrt{n}$
μ
μ + 1.96$\sigma_x/\sqrt{n}$
$\bar{x} - 1{,}96\sigma_x / \sqrt{n}$
$\bar{x}$
$\bar{x} + 1{,}96\sigma_x / \sqrt{n}$


Density
μ - 1.96σ_x/ $\sqrt{n}$
μ
μ + 1.96σ_x/ $\sqrt{n}$
$\bar{x} - 1{,}96\sigma_x / \sqrt{n}$
$\bar{x}$
$\bar{x} + 1{,}96\sigma_x / \sqrt{n}$


Density
μ - 1.96σ_x/√n
μ
μ + 1.96σ_x/√n
$\bar{x} - 1{,}96\sigma_x / \sqrt{n}$
$\bar{x}$
$\bar{x} + 1{,}96\sigma_x / \sqrt{n}$


Density
μ - 1.96σx/√n
μ
μ + 1.96σx/√n
x̄ − 1,96σx / √n
x̄
x̄ + 1,96σx / √n

Example: Corn plants, $n = 50$; $\bar{x} = 170\ s_x = 10$;

Boundaries of the 95% confidence inteval:

$$lower\ limit: 170 - \frac{1.96 \cdot 10}{\sqrt{50}} = 167.2$$

$$upper\ limit: 170 + \frac{1.96 \cdot 10}{\sqrt{50}} = 172.8$$

With the probability of 95% the population mean lies between 167.2 and 172.8 cm.

Testing Statistical Hypothesis

What is Hypothesis Testing?

A **statistical hypothesis** is an assumption about a population parameter. This assumption may or may not be true. **Hypothesis testing** refers to the formal procedures used by statisticians to accept or reject statistical hypotheses.

The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected.

There are two types of statistical hypotheses.

- **Null hypothesis**. The null hypothesis, denoted by $\boldsymbol{H_0}$, is usually the hypothesis that sample observations result purely from chance.

- **Alternative hypothesis**. The alternative hypothesis, denoted by $\boldsymbol{H_1}$ or $\boldsymbol{H_A}$, is the hypothesis that sample observations are influenced by some non-random cause.

Hypothesis Tests

Statisticians follow a formal process to determine whether to reject a null hypothesis, based on sample data. This process, called **hypothesis testing**, consists of the following steps.

- State the hypotheses. This involves stating the null and alternative hypotheses.
- Choose an acceptable upper limit for the error probability and the according rejection region if the null hypothesis is not true.
- Choosing of a **test statistic**.
- Computing the **value of the test statistic for the given sample data.**
- Interpret results. Apply the **decision rule**: If the value of the test statistic is unlikely, based on the null hypothesis, reject the null hypothesis.

Statement of the null hypothesis H_0 and alternative hypothesis H_1

The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false.

Example:

Null hypothesis $\boldsymbol{H_0}$**:** The mean value μ of the height of trees in a population, which the sample was taken from, is equal to the theoretical value μ_0

Alternative hypothesis $\boldsymbol{H_1}$**:** The mean value μ of the height of trees in a population, which the sample was taken from, is **not** equal to the theoretical value μ_0

Notation:

$\boldsymbol{H_0}$**:** $\mu = \mu_0$

$\boldsymbol{H_1}$**:** $\mu \neq \mu_0$

One-Tailed and Two-Tailed Hypothesis

A one tailed hypothesis specifies a directional relationship between testing parameter. In this case we not only state that there will be differences between the parameters but we specify in which direction the differences will exist.

Anytime we expect a relationship to be directional (i.e. to go one specific way) we are using a one-tailed hypothesis.

This is the opposite of a two-tailed hypothesis. A two tailed hypothesis would predict that there was a difference between groups, but, would make no reference to the direction of the effect.

Examples: Two-Tailed Hypothesis

Null hypotheses $\boldsymbol{H_0}$**:** The mean value μ of the height of trees in the population, which the sample was taken from, is equal to the theoretical value μ_0

Alternative hypotheses $\boldsymbol{H_1}$**:** The mean value μ of the height of trees in the population, which the sample was taken from, is **not** equal to the theoretical value μ_0

Notation:

$\boldsymbol{H_0}$**:** $\mu = \mu_0$

$\boldsymbol{H_1}$**:** $\mu \neq \mu_0$

Examples: One-Tailed Hypothesis

Null hypotheses $\boldsymbol{H_0}$**:** The mean value μ of the height of trees in the population, which the sample was taken from, is **not smaller** than the theoretical value μ_0

Alternative hypotheses $\boldsymbol{H_1}$**:** The mean value μ of the height of trees in the population, which the sample was taken from, is **smaller** than the theoretical value μ_0

Notation:

$$\boldsymbol{H_0}\textbf{:}\ \mu \geq \mu_0$$

$$\boldsymbol{H_1}\textbf{:}\ \mu < \mu_0$$

Rules for Hypothesis stating:

- The statement is always made about parameters of population
- What is to be proved belongs to the alternative hypotheses **(H_1)**
- The equality sign belongs to the null hypotheses **(H_0)**

The asymmetry between H_0 and H_1

- If you can reject a $\boldsymbol{H_0}$, then you have a strong evidence, that $\boldsymbol{H_1}$ is true.

- If you cannot reject a $\boldsymbol{H_0}$, it doesn’t mean, that the null hypothesis is true. It means, that your data sample doesn’t let you make a definitely conclusion

Summary:

You can **reject** the null hypothesis or you can **fail to reject** the null hypothesis.

Why the distinction between "acceptance" and "failure to reject"? Acceptance implies that the null hypothesis is true. Failure to reject implies that the data are not sufficiently persuasive for us to prefer the alternative hypothesis over the null hypothesis.

Decision Errors

Two types of errors can result from a hypothesis test.

- **Type I error**. A Type I error occurs when the researcher rejects a null hypothesis when it is true. The probability of committing a Type I error is called the **significance level**. This probability is also called **alpha**, and is often denoted by α.

- **Type II error**. A Type II error occurs when the researcher fails to reject a null hypothesis that is false. The probability of committing a Type II error is called **Beta**, and is often denoted by β. The probability of **not** committing a Type II error is called the **Power** of the test.

Examples:

Type I error: „false positive“

- Medicine: Test of some disease is positive, but a person is healthy
- Alarm by airport control, but no forbidden object
- Court procedure: an innocent person is convicted

Type II error: „false negative“

- Medicine: Test of some disease is negative, but a person is ill
- No alarm by airport control, but a forbidden object
- Court procedure: a guilty person is acquitted

Type I error and Type II error

		Reality	
		H_0 true	H_0 false
Tests **Decision**	H_0 rejected	Type I error	correct decision
	H_0 not rejected	correct decision	Type II error

Test statistic

- The evaluation often focuses around a single test statistic.
- A test statistic is a some parameter, which is calculated based on a sample
- The value of a test statistic gives an information about differences of results based on a sample and assumption of hypothesis
- The test statistic under $\boldsymbol{H_0}$ have an exactly or approximately known probability distribution
- We can compute, how probable under $\boldsymbol{H_0}$ is the value of the test statistic, which was calculated based on a sample.
- If the value of the test statistic under $\boldsymbol{H_0}$ is too improbable, then the $\boldsymbol{H_0}$ will be rejected

Null hypotheses $\boldsymbol{H_0}$**:** The mean value μ of the height of trees in a population, which the sample was taken from, is equal to the theoretical value μ_0

Alternative hypotheses $\boldsymbol{H_1}$**:** The mean value μ of the height of trees in a population, which the sample was taken from, is **not** equal to the theoretical value μ_0

H_0: $\mu = \mu_0$

H_1: $\mu \neq \mu_0$

Under $\boldsymbol{H_0}$ has the difference $\delta = \mu - \mu_0$ a Normal Distribution with the mean 0 and a unknown variance σ^2/n.

Computing the value of the test statistic z for the sample

The statistic: The standardized difference

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

The sample value of the statistic:

$$z_{test} = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}$$

where: $\bar{x}$ – sample mean

$s_{\bar{x}} = s_x/\sqrt{n}$ – standard error of the mean

s_x – standard deviation of the sample

Decision Rules

In statistics, a result is called significant if it is **unlikely** to have occurred by chance alone, according to a predetermined threshold probability, the significance level.

The analysis plan includes decision rules for rejecting the null hypothesis. In practice, statisticians describe these decision rules in two ways

- with reference to a region of acceptance

or

- with reference to a p-value.

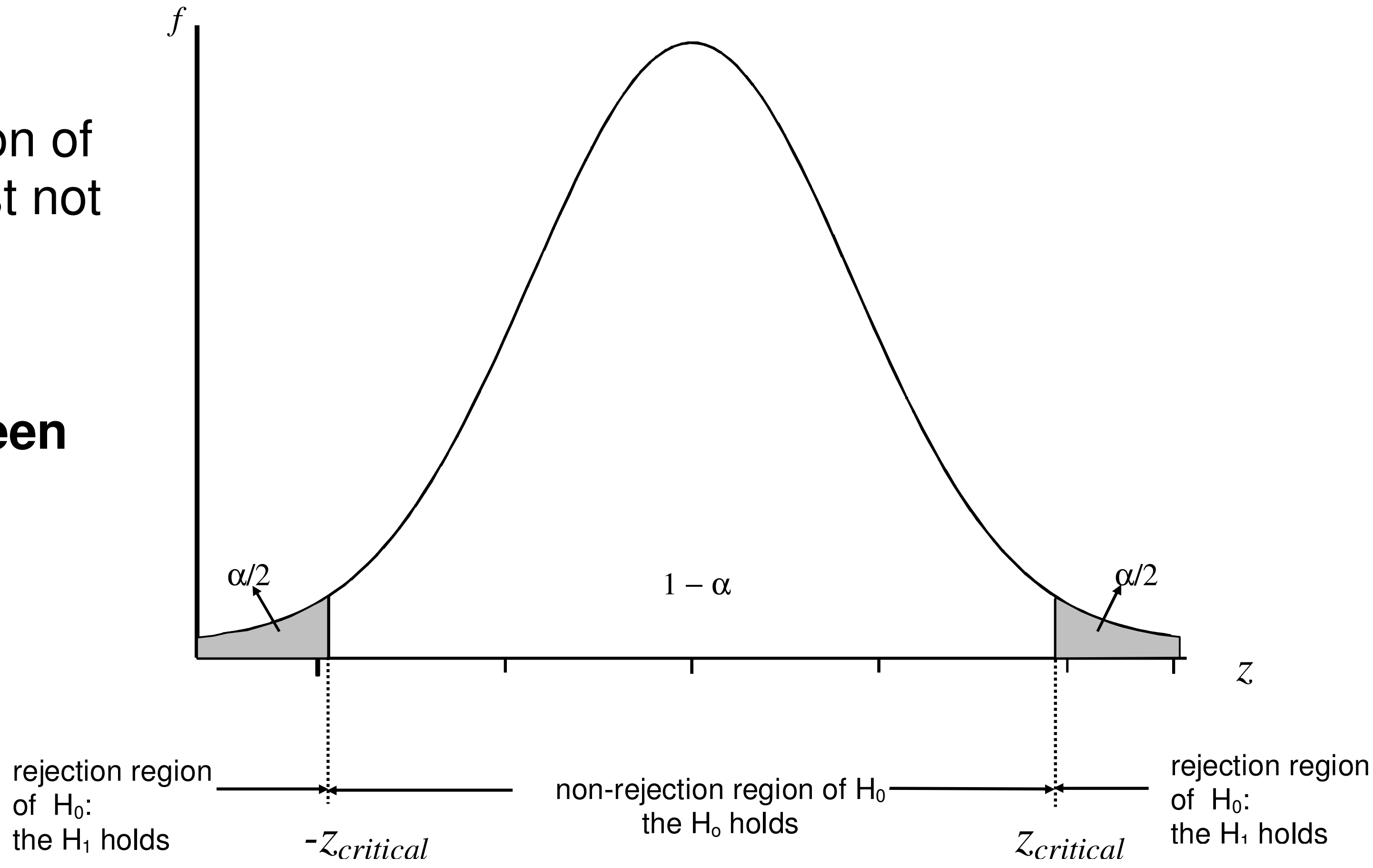
The **region of acceptance** or of **non-rejection** is a range of values. If the test statistic falls within the region of acceptance, the null hypothesis is not rejected. The region of acceptance is defined so that the chance of making a Type I error is equal to the significance level.

The set of values outside the region of acceptance is called the **region of rejection**. If the test statistic falls within the region of rejection, the null hypothesis is rejected. In such cases, we say that the hypothesis has been rejected at the α level of significance.

Construction of a two-tailed region of nonrejection/rejection of null hypothesis (for Standard Normal Distribution)

The regions of non-rejection and rejection of Null Hypothesis must not overlap. It does hold always H_0 or H_1.

The decision between $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ is unambiguous

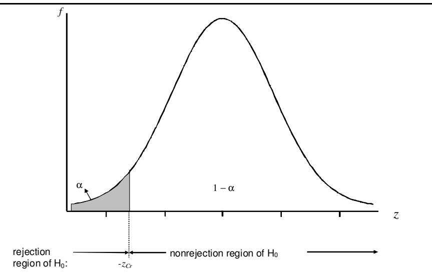
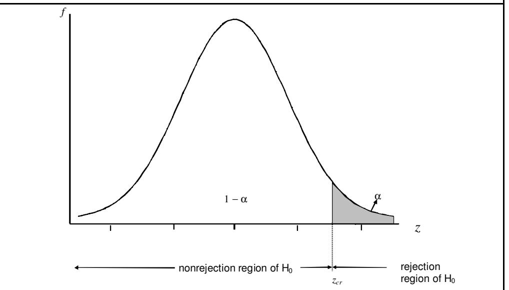
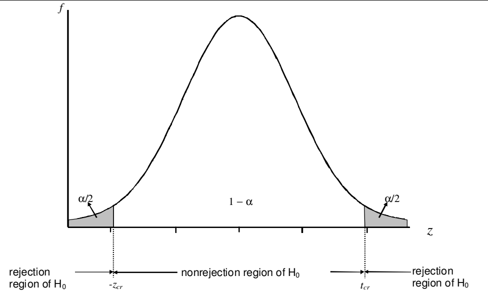


One-Tailed and Two-Tailed Tests

A test of a statistical hypothesis, where the region of rejection is only on one side of the sampling distribution, is called a **one-tailed test**.

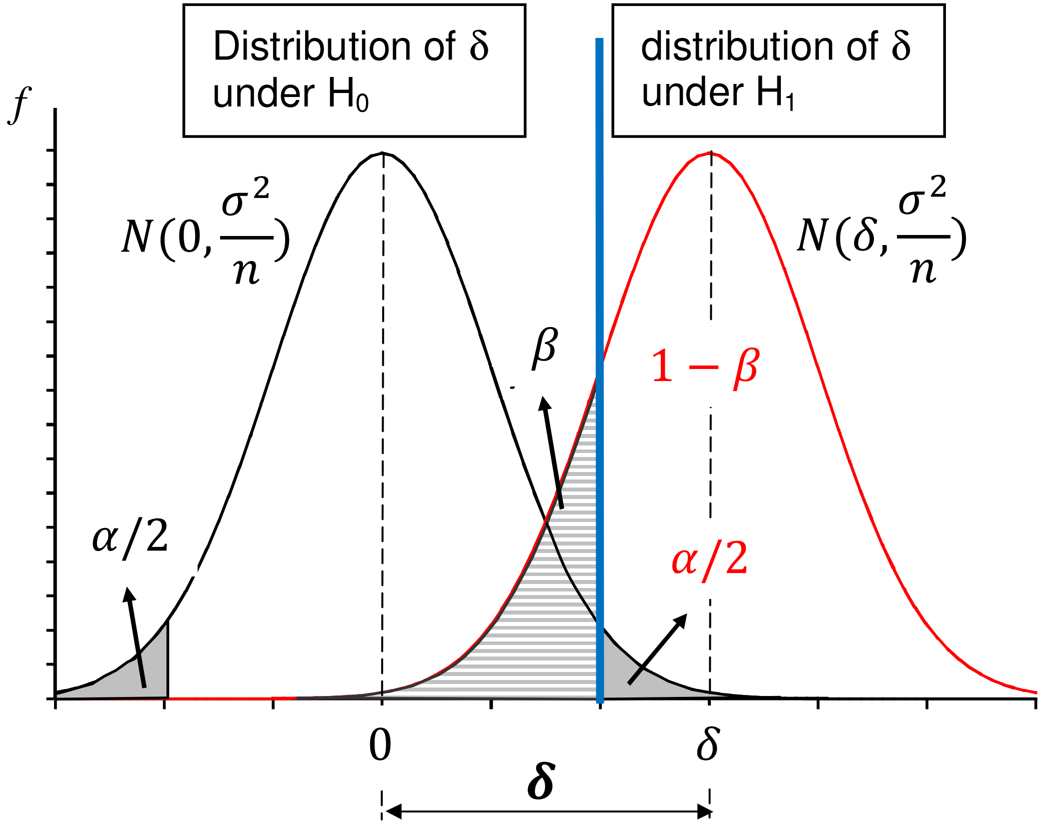
A test of a statistical hypothesis, where the region of rejection is on both sides of the sampling distribution, is called a **two-tailed test**.

Null and alternative hypothesis by testing of mean

Question	Null hypothesis H_0	Alternative hypothesis H_1	Graph
Is the mean μ of a population, which the sample was taken from, significantly **smaller**, than the theoretical value μ_0 ?	$\mu \geq \mu_0$ or $\delta = \mu - \mu_0 \geq 0$	$\mu < \mu_0$ or $\delta = \mu - \mu_0 < 0$	
Is the mean μ of a population, which the sample was taken from, significantly **greater**, than the theoretical value μ_0 ?	$\mu \leq \mu_0$ or $\delta = \mu - \mu_0 \leq 0$	$\mu > \mu_0$ or $\delta = \mu - \mu_0 > 0$	
Does the mean μ of a population, which the sample was taken from, significantly **differ** from the theoretical value μ_0 ?	$\mu = \mu_0$ or $\delta = \mu - \mu_0 = 0$	$\mu \neq \mu_0$ or $\delta = \mu - \mu_0 \neq 0$	

Type I error cannot be infinitely small

Example: Distribution of difference $\bar{X} - \mu_0$ under H_0 and H_1



The smaller the type I error (α), the bigger the chance of the type II error and the smaller **the power of the test $(\mathbf{1}-\boldsymbol{\beta})$.**

$1-\boldsymbol{\beta}$: The power of the test

The power of the test shows the probability for rejection of H_0, if the alternative hypothesis is true

The power of the test

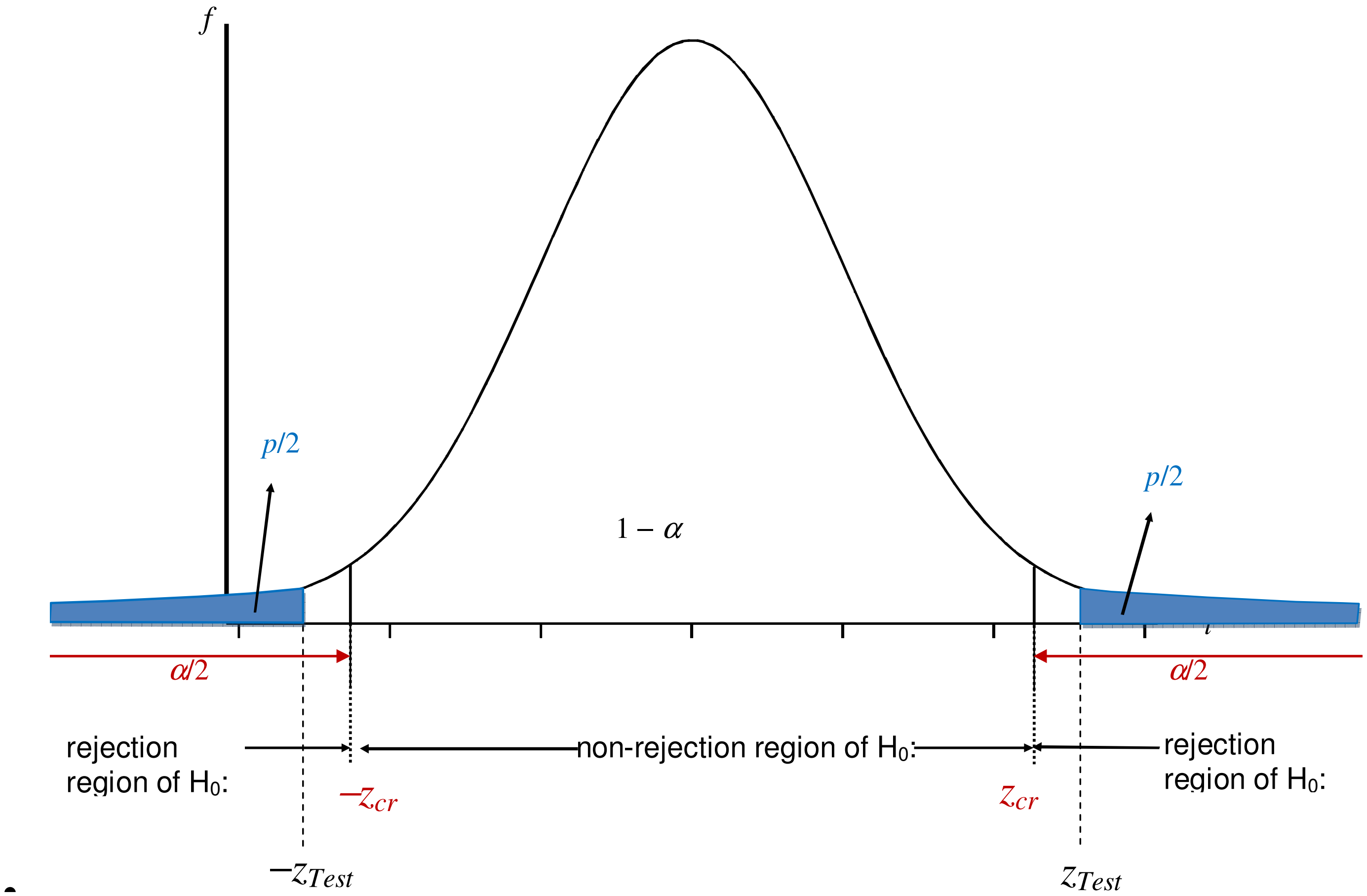
- Can be given: For computing a sample size.
- Can be computing only afterwards **(post hoc)**: After the trial is carried out

p-value

The strength of evidence in support of a null hypothesis is measured by the **p-value**. The p-value is the probability of observing a test statistic as extreme as the observed, assuming the null hypothesis is true. If the p-value is less than the significance level, we reject the null hypothesis.

- The smaller the **p-value**, the more the result speaks against the null hypothesis.
- If the null hypothesis is rejected, the result of a test is considered as significant.

Interpretation of the Computer Output – p-value
(Example: Standard normal distribution)



p-value for the two-tailed t-test

The decision rules with reference to a region of acceptance or with reference to a p-value are equivalent. Some statistics texts use the p-value approach; others use the region of acceptance approach.

p-Value $< \alpha \Rightarrow H_0$ is rejected p-Value $\geq \alpha \Rightarrow H_0$ is not rejected

Compare:

$\lvert z_{Test} \rvert \leq z_{1-\alpha/2} \Rightarrow H_0$ is not rejected $\lvert z_{Test} \rvert > z_{1-\alpha/2} \Rightarrow H_0$ is rejected

The bigger $\lvert z_{Test} \rvert$, the smaller the $\boldsymbol{p}$ $-$Value and contrariwise!

Two-tailed Gaussian test

The Gaussian test compares the sample mean with some theoretical value

Question: Does the mean μ of the variable Height of a population, which the sample was taken from, significantly differ from the theoretical value 6 ($\alpha = 5\%$)? The sample mean $\bar{x} = 7.884$; the standard deviation of the sample $s_x = 1.734$; the sample size $n = 50$.

Assumption: Normal distribution of a population; Sample size is not very small ($n > 30$)

H_0: $\delta = \mu - \mu_0 = 0$

H_1: $\delta = \mu - \mu_0 \neq 0$ (two-tailed)

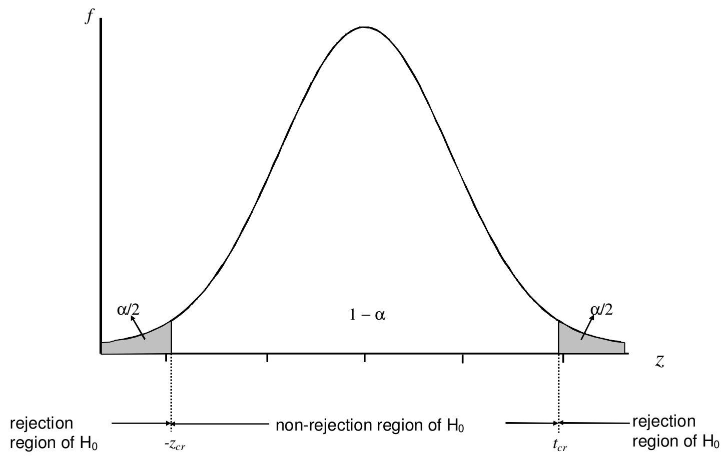
$\alpha = 5\%$

two-tailed

Test statistic: $z_{test} = \frac{\bar{x}-\mu_0}{s_x/\sqrt{n}}$

Decision rule: $|z_{Test}| \leq z_{1-\alpha/2}$ $\Rightarrow H_0$ (no significant differences)

$|z_{Test}| > z_{1-\alpha/2}$ $\Rightarrow H_1$ (significant differences)



Example

$$z_{Test} = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}} = \frac{7.884 - 6}{\frac{1.734}{\sqrt{50}}} = \frac{1.884}{0.245} = 7.68$$

$7.68 > z_{Table} = 1.96 \ \Rightarrow H_1$

The mean of the variable Height μ of a population, which the sample was taken from, differs significantly from the theoretical value 6 by $\alpha = 5\%$.

One-tailed Gaussian tests

Question: Is the mean of the variable Height μ of a population, which the sample was taken from, significantly **greater** than the theoretical value $\mu_0 = 6$ $(\alpha = 5\%)$?
The sample mean $\bar{x} = 7{,}884$; the standard deviation of the sample $s_x = 1{,}734$; the sample size $n = 50$.

Assumption: Normal distribution of a population; Sample size is not very small $(n > 30)$

H_0: $\delta = \mu - \mu_0 \leq 0$

H_1: $\delta = \mu - \mu_0 > 0$ (one-tailed)

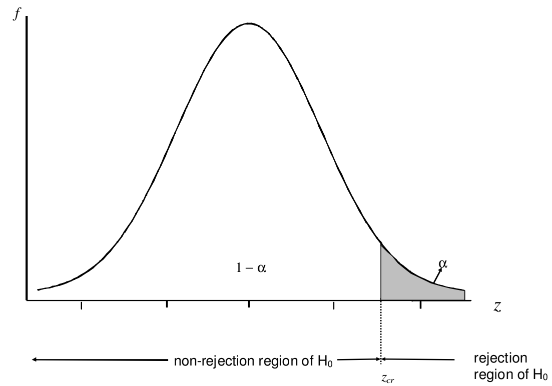
one-tailed

$\alpha = 5\%$

Test statistic: $z_{Test} = \frac{\bar{x} - \mu_0}{s_x/\sqrt{n}}$

Decision rule: $z_{Test} \leq z_{1-\alpha} \Rightarrow H_0$ (not significantly greater)

$z_{Test} > z_{1-\alpha} \Rightarrow H_1$ (significantly greater)



Example

$$z_{Test} = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}} = \frac{7{,}884 - 6}{\frac{1{,}734}{\sqrt{50}}} = \frac{1{,}884}{0{,}245} = 7{,}68$$

$7{,}68 > z_{Table} = 1{,}64 \ \Rightarrow H_1$

The mean of the variable Height μ of a population, which the sample was taken from, is significantly **greater** than the theoretical value 6 by $\alpha = 5\%$.

Question: Is the mean of the variable Height μ of a population, which the sample was taken from, significantly **smaller** than the theoretical value $\mu_0 = 9 (\alpha = 5\%)$?
The sample mean $\bar{x} = 7{,}884$; the standard deviation of the sample $s_x = 1{,}734$; the sample size $n = 50$.

Assumption: Normal distribution of a population; Sample size is not very small $(n > 30)$

H$_0$: $\delta = \mu - \mu_0 \geq 0$

H$_1$: $\delta = \mu - \mu_0 < 0$ (one-tailed)

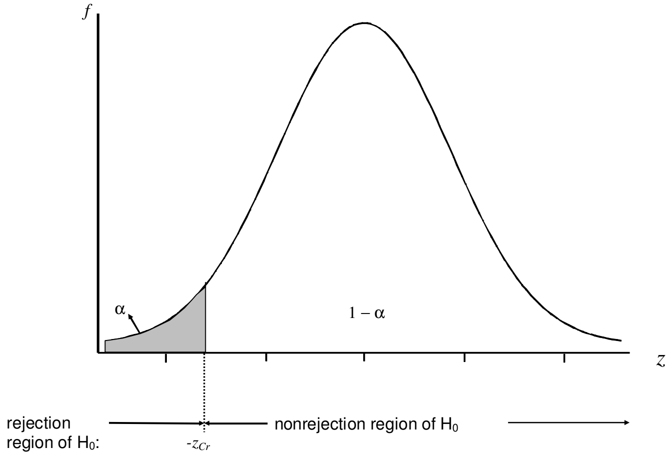
one-tailed

$\alpha = 5\%$

Test statistic: $z_{Test} = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}}$

Decision rule: $z_{Test} \geq z_{\alpha} \Rightarrow H_0$ (not significantly smaller)

$z_{Test} < z_{\alpha} \Rightarrow H_1$ (significantly smaller)



Example

$$z_{Test} = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}} = \frac{7.884 - 9}{\frac{1.734}{\sqrt{50}}} = \frac{-1.116}{0.245} = -4.56$$

$-4.56 < z_{Table} = -1.64 \;\Rightarrow H_1$

The mean of the variable Height μ of a population, which the sample was taken from, is significantly **smaller** than the theoretical value 9 by $\alpha = 5\%$.

Example: Two-tailed Gaussian test

The Gaussian test compares the sample mean with some theoretical value

Question: Does the mean of the variable Height μ of a population, which the sample was taken from, significantly differ from the theoretical value 6 ($\alpha = 5\%$)? The sample mean $\bar{x} = 7.884$; the standard deviation of the sample $s_x = 1.734$; the sample size $n = 50$.

Assumption: Normal distribution of a population; Sample size is not very small ($n > 30$)

H$_0$: $\delta = \mu - \mu_0 = 0$ (two-tailed)

H$_1$: $\delta = \mu - \mu_0 \neq 0$

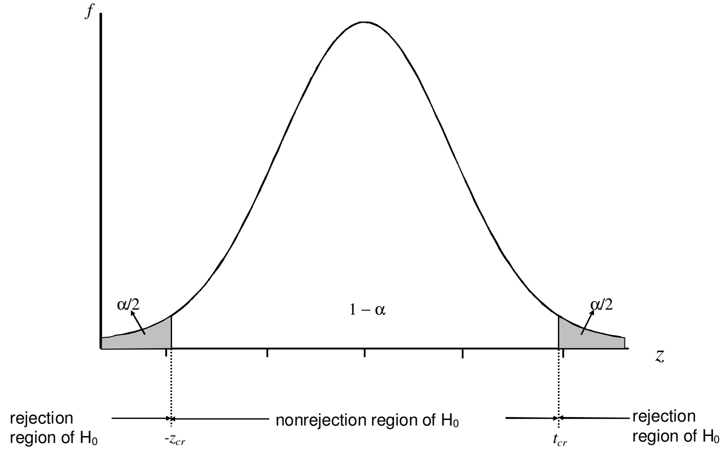
$\alpha = 5\%$

two-tailed

Test statistic: $z_{test} = \frac{\bar{x} - \mu_0}{s_x/\sqrt{n}}$

Decision rule: $|z_{Test}| \leq z_{1-\alpha/2} \Rightarrow H_0$ (no significant differences)

$|z_{Test}| > z_{1-\alpha/2} \Rightarrow H_1$ (significant differences)



Example

$$z_{Test} = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}} = \frac{7.884 - 6}{\frac{1.734}{\sqrt{50}}} = \frac{1.884}{0.245} = 7.68$$

$7.68 > z_{Table} = 1.96 \ \Rightarrow H_1$

The mean of the variable Height μ of a population, which the sample was taken from, differs significantly from the theoretical value 6 by $\alpha = 5\%$.

Confidence Interval for the difference $\bar{d} = \bar{x} - \mu_0$

$$\bar{x} = 7.884;\ s_x = 1.734;\ n = 50.\ \mu_0 = 6;\ \bar{d} = \bar{x} - \mu_0 = 1.884$$

Limits of the 95% confidence inteval for the difference:

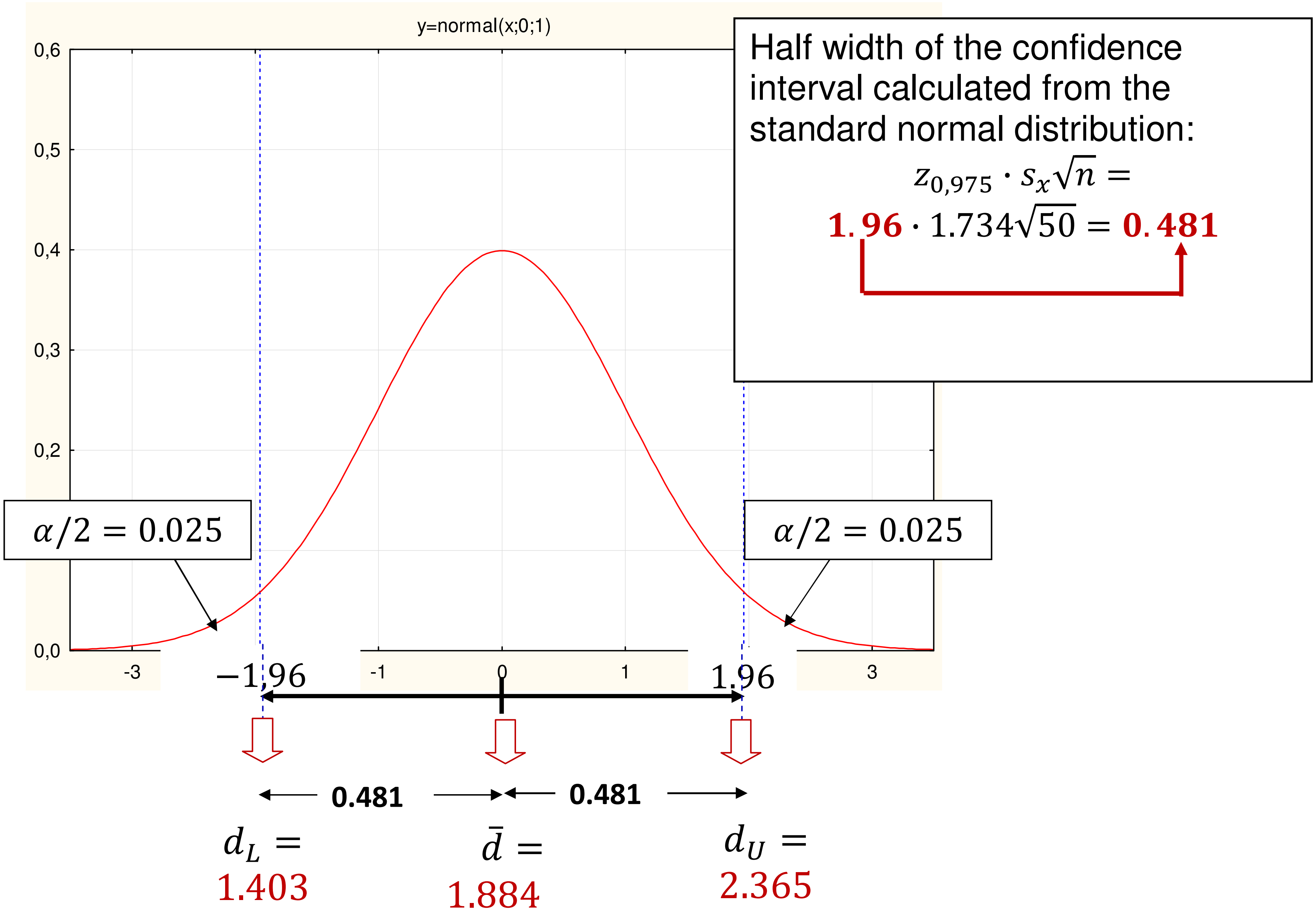
$$lower\ limit: (\bar{x} - \mu_0) - \frac{z_{1-\alpha/2} \cdot s_x}{\sqrt{n}} = 1.884 - \frac{1.96 \cdot 1.734}{\sqrt{50}} = 1.884 - 0.481 = 1.403$$

$$upper\ limit: (\bar{x} - \mu_0) + \frac{z_{1-\alpha/2} \cdot s_x}{\sqrt{n}} = 1.884 + \frac{1.96 \cdot 1.734}{\sqrt{50}} = 1.884 + 0.481 = 2.365$$

$\Longrightarrow$ The diference $\bar{x} - \mu_0$ differs signifikantly from 0 bei $\alpha = 5\%$ because the 95% confidence interval doas not include 0.

The results of the significance test and confidence intervall are identical.

Confidence interval for difference



Two-tailed Gaussian test

$$\bar{d}_{st} = z_{Test} = \frac{\bar{d}}{s_x\sqrt{n}} =$$

$$\frac{1.884}{1.734\sqrt{50}} = \mathbf{7.68}$$