Statistics for EES
2. Standard error

Dirk Metzler

April 17, 2020

Contents

Contents

1 The standard error SE

The Standard Error

$$SE = \frac{\text{sd}}{\sqrt{n}}$$

describes the variability of the sample mean.

n: sample size sd: sample standard deviation

1.1 example: drought stress in sorghum

drought stress in sorghum

References

[BB05] V. Beyel and W. Brüggemann. Differential inhibition of photosynthesis during pre-flowering drought stress in Sorghum bicolor genotypes with different senescence traits. *Physiologia Plantarum*, 124:249–259, 2005.

- 14 sorghum plants were not watered for 7 days.
- in the last 3 days: transpiration was measured for each plant (mean over 3 days)
- the area of the leaves of each plant was determined

transpiration rate

=

(amount of water per day)/area of leaves

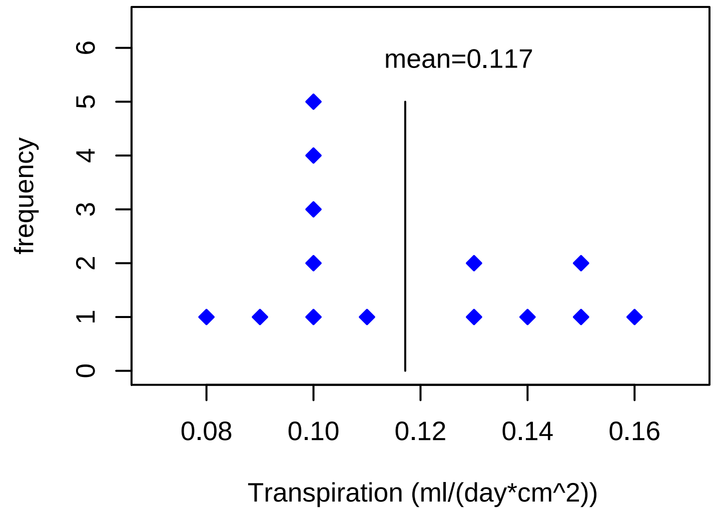
$$\left[\frac{\mathrm{ml}}{\mathrm{cm}^2 \cdot \mathrm{day}}\right]$$

Aim: Determine mean transpiration rate μ under these conditions.

If we hade many plants, we could determine μ quite precisely.

Problem: How accurate is the estimation of μ with such a small sample? ($n = 14$)

drought stressed sorghum (variety B, $n = 14$)



$$\text{transpiration data: } x_1, x_2, \ldots, x_{14}$$

$$\overline{x} = \left(x_1 + x_2 + \cdots + x_{14}\right)/14 \; = \frac{1}{14}\sum_{i=1}^{14} x_i$$

$$\overline{x} = 0.117$$

our estimation:

$$\mu \approx 0.117$$

how accurate is this estimation?

How much does $\overline{x}$ (our estimation) deviate from μ (the true mean value)?

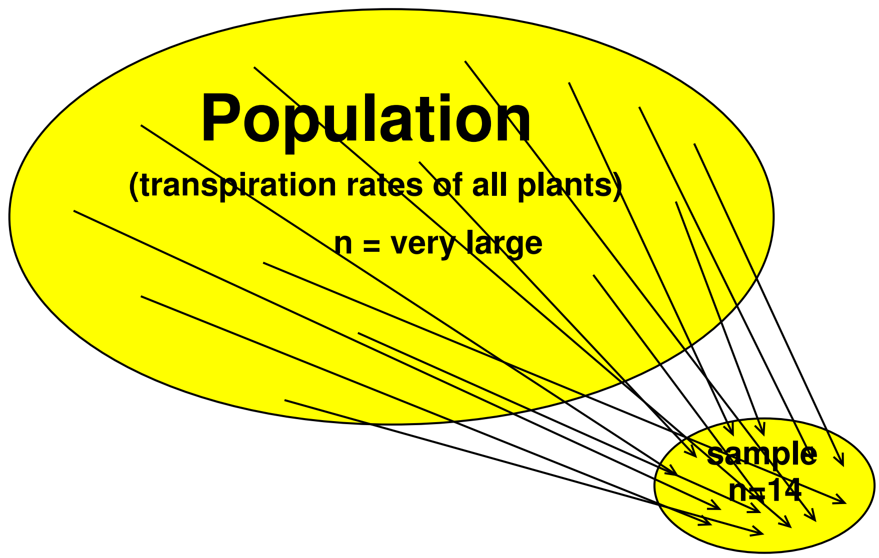
1.2 general consideration

Assume we had made the experiment not just 14 times, but repeated it 100 times, 1000 times, 1000000 times

We consider our 14 plants as

random sample

from a very large population of possible values.



We estimate

the population mean

μ

by the sample mean

$\overline{x}$.

μ is a parameter.

$\overline{x}$ is a statistic.

parameter: real or theoretical value within mathematical model

- example: μ

- non-random (in classical frequentistic stats)
- assumed usually in stats: there is a true value that is unknown

statistic: a function of the sampled data (that is, calculated from the data)

- example: $\overline{x}$
- are random variables because data is also random due to
  - randomly sampling from natural variation
  - random process
  - measurement error

estimator: statistic to estimate the value of a parameter; example: $\overline{x}$ is an estimator for μ

Another example

The **statistics**

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2} \qquad \text{and} \qquad \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

are **estimators** for the population standard deviation σ, which is a **parameter**.

Each new sample gives a new value of $\overline{x} = (x_1 + x_2 + \cdots + x_n)/n$.

$\overline{x}$ depends on randomness: it is a *random variable*

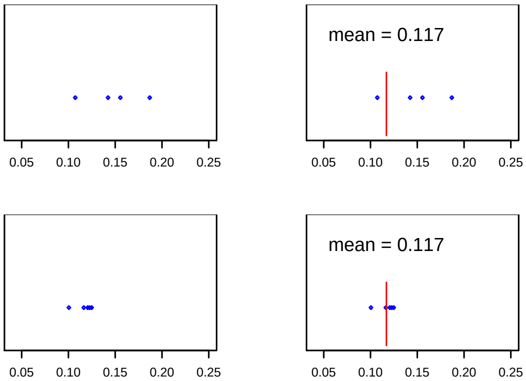
Problem: How variable is $\overline{x}$?

More precisely: What is the typical deviation of $\overline{x}$ from μ?

What does the variability of $\overline{x}$ depend on?

What does the variability of $\overline{x}$ depend on?

1. On the the variability of the single observations $x_1, x_2, \ldots, x_n$



x varies a lot
$\Rightarrow$ $\overline{x}$ *varies a lot*

x varies little
$\Rightarrow$ $\overline{x}$ *varies little*

What does the variability of $\overline{x}$ depend on?

1. On the the variability of the single observations $x_1, x_2, \ldots, x_n$

2. On the sample size n

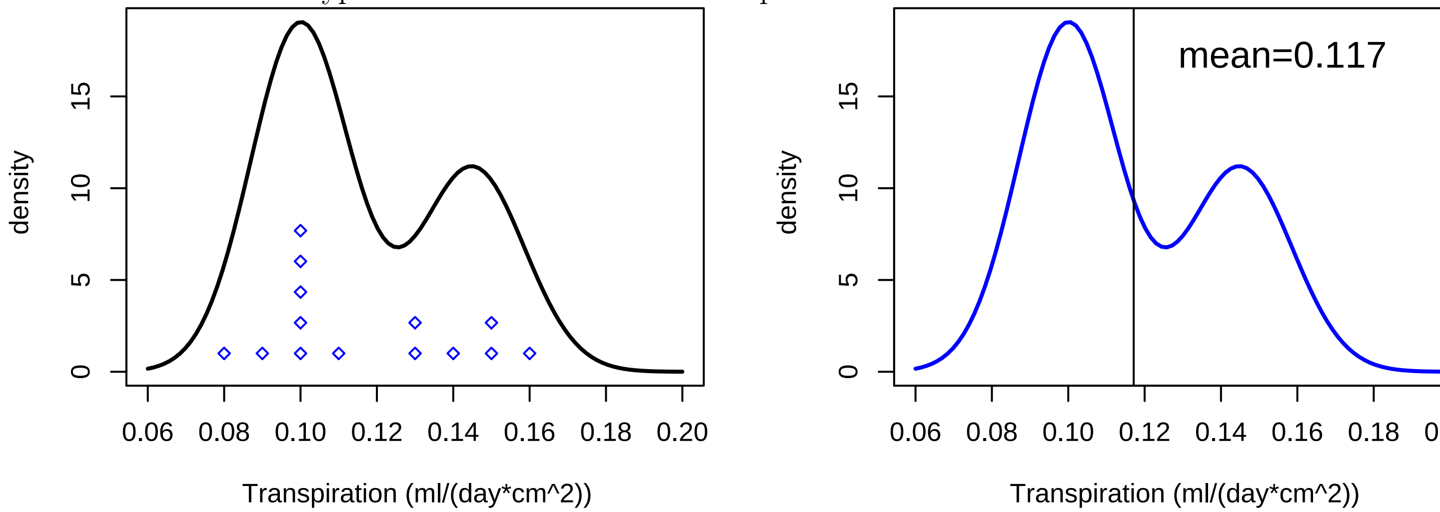
The larger n, the smaller is the variability of $\overline{x}$.

To explore this dependency we perform a (Computer-)Experiment.

Experiment: Take a population, draw samples and examine how $\overline{x}$ varies.

We assume the distribtion of possible transpriration rates looks like this:

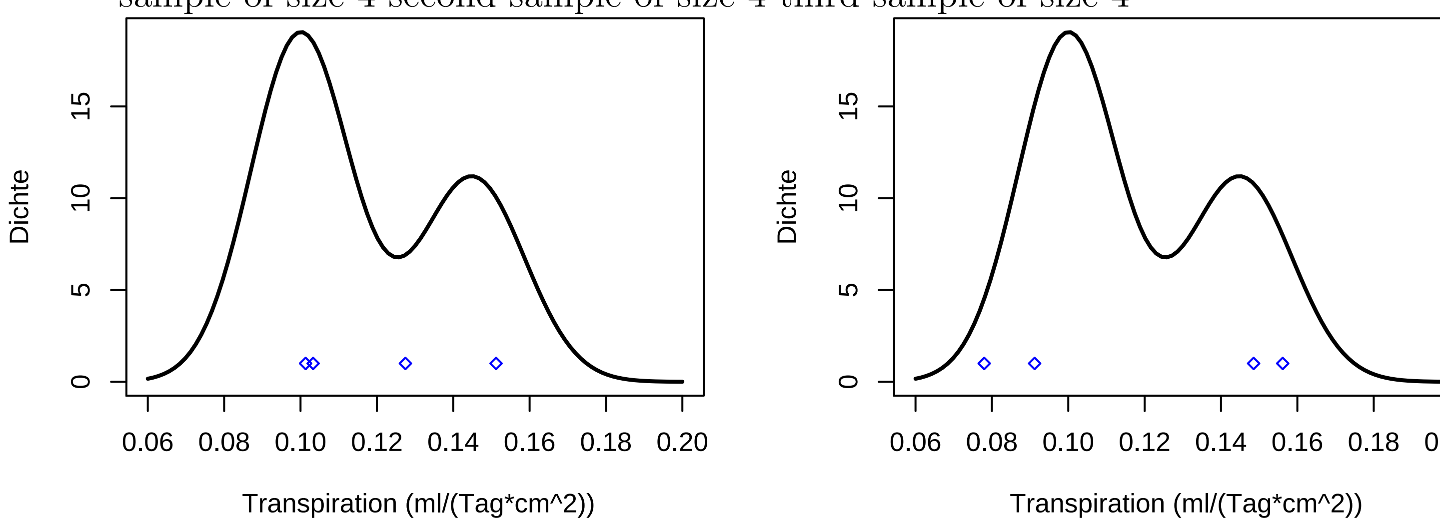
hypothetical distribution of transpiration rates



At first with small sample sizes:

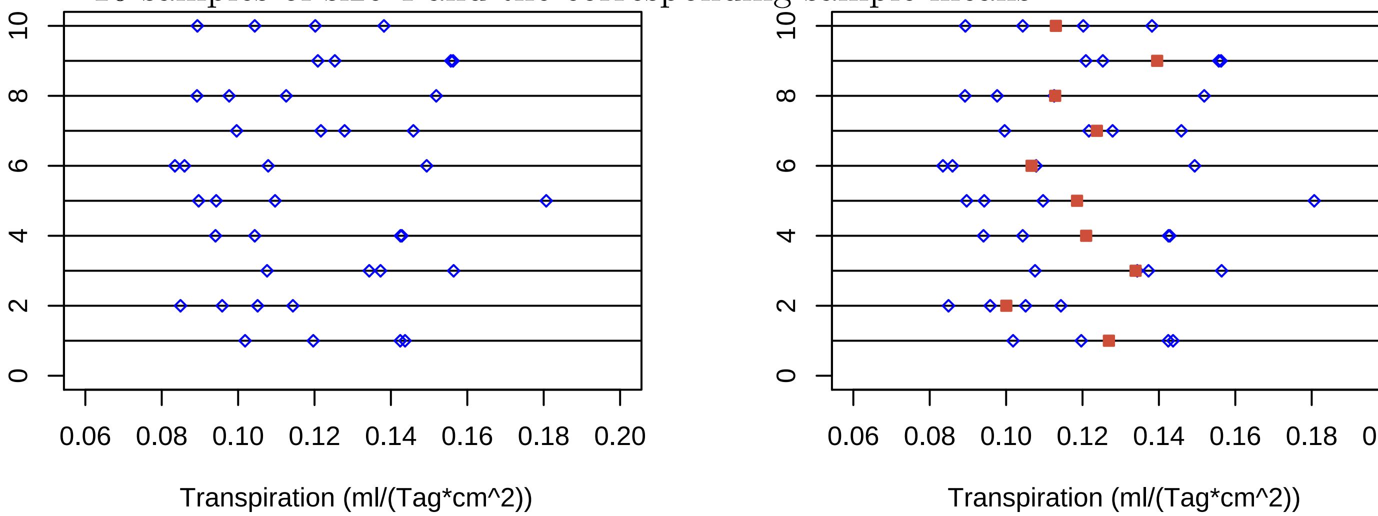
$$n = 4$$

sample of size 4 second sample of size 4 third sample of size 4

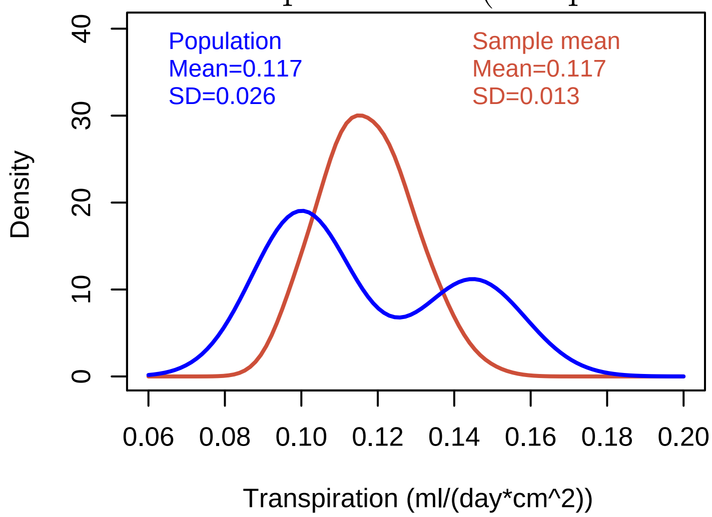


How variable are
the sample means?

10 samples of size 4 and the corresponding sample means



distribution of sample means (sample size $n = 4$)

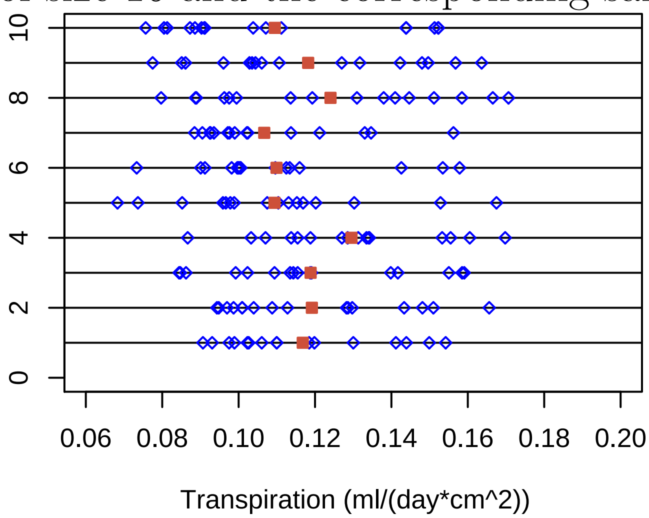


population: standard deviation $= 0.026$

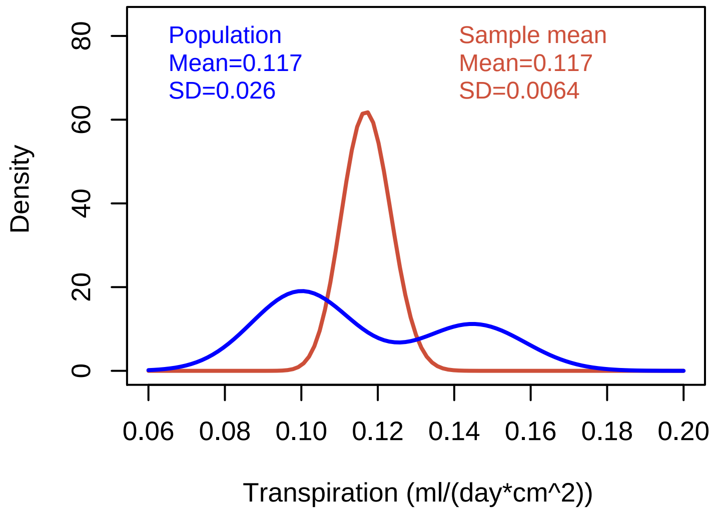
sample means ($n = 4$): $\begin{aligned} \textit{standard deviation} &= 0.013 \\ &= 0.026/\sqrt{4} \end{aligned}$

Increase the sample size from 4to 16

10 samples of size 16 and the corresponding sample means



distribution of sample means (sample size $n = 16$)



$$\text{population: standard deviation} = 0.026$$

$$\text{sample mean } (n = 16): \begin{array}{rcl} \textit{standard deviation} & = & 0.0065 \\ & = & 0.026/\sqrt{16} \end{array}$$

General Rule 1. *Let $\overline{x}$ be the mean of a sample of size n from a distribution (e.g. all values in a population) with standard deviation σ. Since $\overline{x}$ depends on the random sample, it is a random variable. Its standard deviation $\sigma_{\overline{x}}$ fulfills*

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}.$$

Problem: σ is unknown

Idea: Estimate σ by sample standard deviation s:

$$\sigma \approx s$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}} =: \text{SEM}$$

SEM stands for *Standard Error of the Mean*, or *Standard Error* for short.

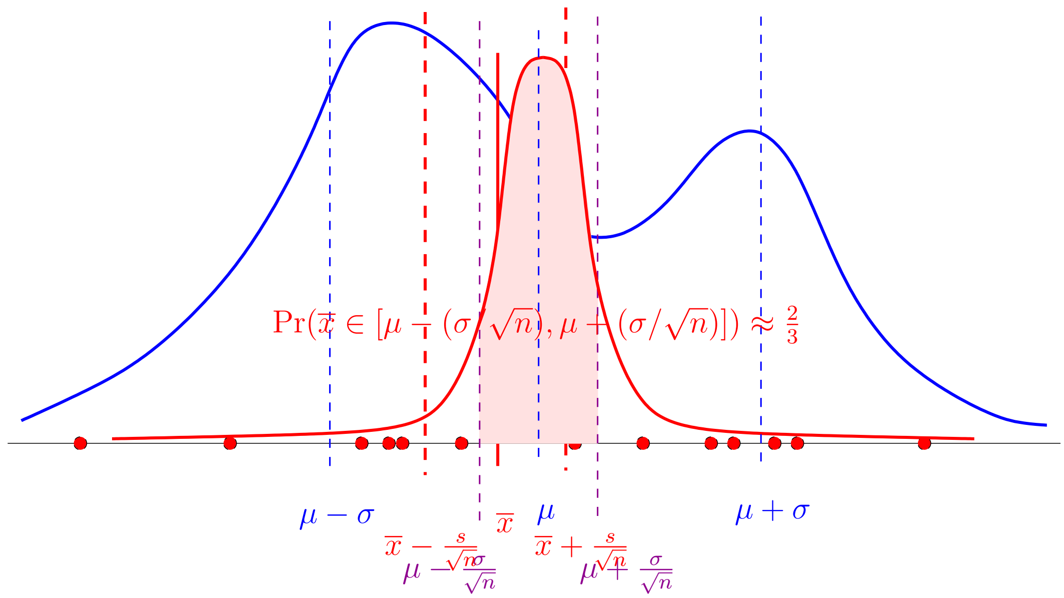
Note: The statistic SEM$= \frac{s}{\sqrt{n}}$ is an estimator

for the parameter $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$.

The distribution of $\overline{x}$

Observation

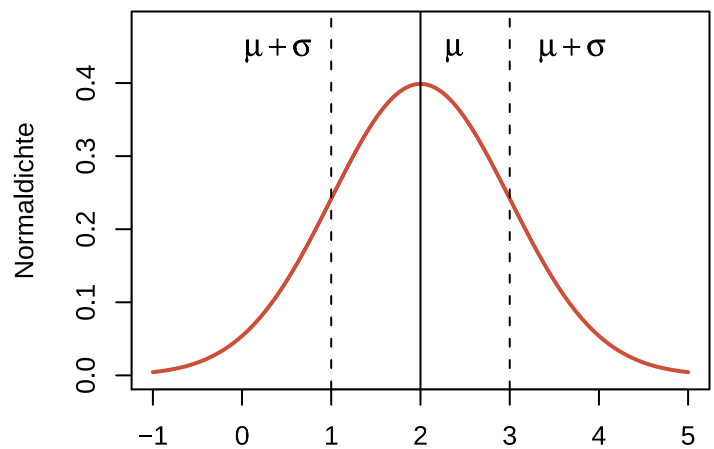
Even if the distribution of x is asymmetric and has multiple peaks, the distribution of $\overline{x}$ will be bell-shaped (at least for larger sample sizes n.)


$\Pr(\overline{x} \in [\mu - (\sigma/\sqrt{n}), \mu + (\sigma/\sqrt{n})]) \approx \frac{2}{3}$
$\mu - \sigma$
$\overline{x}$
μ
$\mu + \sigma$
$\overline{x} - \frac{s}{\sqrt{n}}$
$\overline{x} + \frac{s}{\sqrt{n}}$
$\mu - \frac{\sigma}{\sqrt{n}}$
$\mu + \frac{\sigma}{\sqrt{n}}$

The distribution of $\overline{x}$ is approximately of a certain shape:

the normal distribution.

Density of the normal distribution

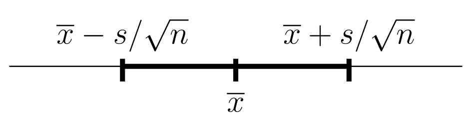


The normal distribution is also called *Gauß distribution* (after Carl Friedrich Gauß, 1777-1855)

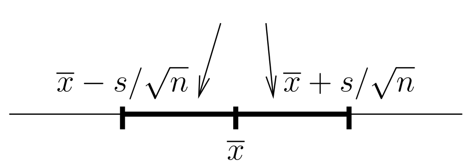
2 Taking standard errors into account

Important consequence

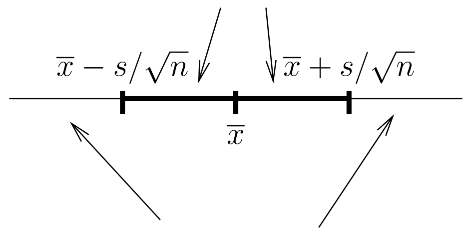
Consider the interval



This interval contains μ with probability of ca. $2/3$



This interval contains μ with probability of ca. $2/3$



probability that μ is outside of interval is ca. $1/3$

Thus:

It may happen that $\overline{x}$ deviates from μ by more than $s/\sqrt{n}$.

Application 1: Which values of μ are plausible?

$$\overline{x} = 0.12$$
$$s/\sqrt{n} = 0.007$$

Question: Could it be that $\mu = 0.115$?

Answer: Yes, not unlikely.

Deviation $\overline{x} - \mu = 0.120 - 0.115 = 0.005$.

Standard Error $s/\sqrt{n} = 0.007$

Deviations like this are not unusual.

Application 2: Comparison of mean values

Example: Galathea

Galathea: Carapace lengths in a sample

Males: $\overline{x}_1 = 3.04$ mm $s_1 = 0.9$ mm $n_1 = 25$

Females: $\overline{x}_2 = 3.23$ mm $s_2 = 0.9$ mm $n_2 = 29$

The females are apparently larger.

Is this significant?

Or could it be just *random*?

How precisely do we know the true mean value?

Males: $\overline{x}_1 = 3.04$ mm $s_1 = 0.9$ mm $n_1 = 25$

$$s_1/\sqrt{n_1} = 0.18 \text{ [mm]}$$

We have to assume uncertainty in the magnitude of ± 0.18 (mm) in $\overline{x}_1$

How precisely do we know the true mean value?

Females: $\overline{x}_2 = 3.23$ mm $s_2 = 0.9$ mm $n_2 = 29$

$$s_2/\sqrt{n_2} = 0.17 \text{ [mm]}$$

It is not unlikely that $\overline{x}_2$ deviates from the true mean by more than ± 0.17 (mm).

The difference of the means

$$3.23 - 3.04 = 0.19 \text{ [mm]}$$

is not much larger than the expected inaccuracies.

It could also be due to pure random that $\overline{x}_2 > \overline{x}_1$

MORE PRECISELY:

If the true means are actually equal $\mu_{Females} = \mu_{Males}$ it is still quite likely that the sample means $\overline{x}_2$ are $\overline{x}_1$ that different.

In the language of statistics:

The difference of the mean values is (statistically) *not significant.*

not significant = can be just random

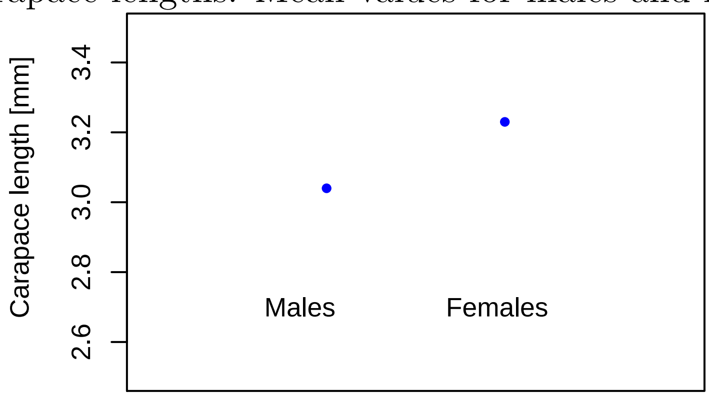
Application 3:

If the mean values are represented graphically, you should als show their precision

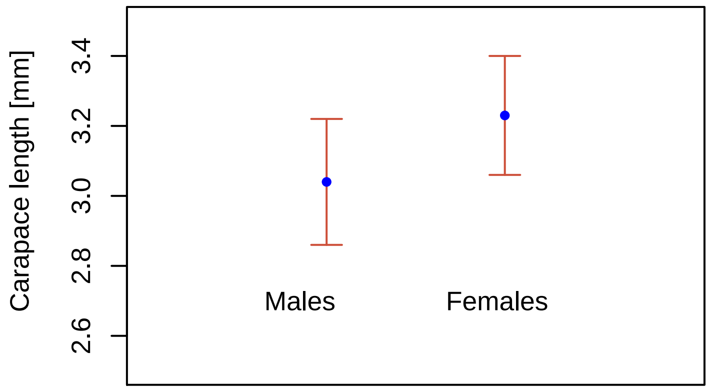
$$(\pm s/\sqrt{n})$$

Not Like this: Like that:

Carapace lengths: Mean values for males and females



Carapace lengths: Mean values ± standard errors for males and females



Application 4:
Planning an experiment: How many observations do I need? (How large should n be?)

If you know which precision you need for (the standard error $s/\sqrt{n}$ of) $\overline{x}$

and if you already have an idea of s
then you can estimate the value of n that is necessary: $s/\sqrt{n} = g$
(g = desired standard error)

Example: Stressed transpiration values in another sorghum subspecies: $\overline{x} = 0.18$ $s = 0.06$ $n = 13$

How often do we have to repeat the experiment to get a standard error of ≈ 0.01?

Which n do we need?

$$\text{Solution: desired: } s/\sqrt{n} \approx 0.01$$
$$\text{From the previous experiment we know: } s \approx 0.06, \text{ so: } \sqrt{n} \approx 6$$
$$n \approx 36$$

Summary

- Assume a population has mean value μ and standard deviation σ.
- We draw a sample of size n from this population with sample mean $\overline{x}$.
- $\overline{x}$ is a random variable with mean value μ and standard deviation $\sigma/\sqrt{n}$.
- Estimate the standard deviation of $\overline{x}$ by $s/\sqrt{n}$, where s is the standard deviation computed from the sample with the formula with $n-1$.
- $s/\sqrt{n}$ is the *Standard Error (of the Mean)*.
- Deviations of $\overline{x}$ of the magnitude of $s/\sqrt{n}$ are usual. They are *not significant:* they can be random.

Some of what you should be able to explain

- Concepts: parameters, statistics, estimators
- Why is the sample mean $\overline{x}$ a random variable?
- distribution properties of $\overline{x}$
- What is the standard error and how is it different from ...
  - ... sd?
  - ... the standard deviation $\sigma_{\overline{x}}$ of the mean?
- When calculating the standard error from data why must I once divide by n (or $\sqrt{n}$) and another time by $n-1$ (or $(\sqrt{n-1})$?
- Applications of the standard error in descriptive data analysis and experimental design.