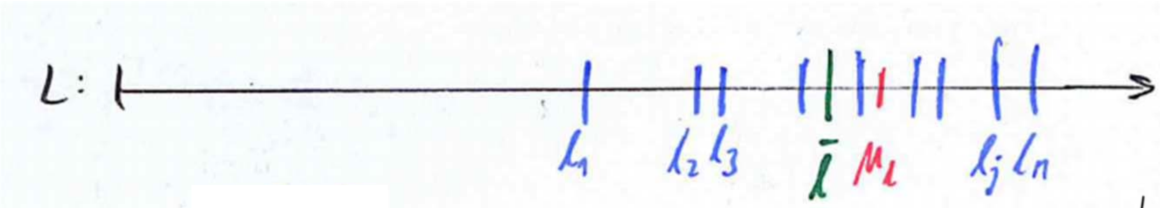
2.4 Variance and standard deviation

2.4.1 Definition of dispersion measures

<u>Given</u>: Random variable L with its realizations $l_j\ ,\ j = 1, 2, 3, \dots, n$

sample → „observation vector"



for when we integrate

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}$$ n observations for the same random variable L

All observations originate from the same population ($n \to \infty$)
$\Rightarrow$ Expectation $E(L) = \mu_L$

<ins>Wanted</ins>: Dispersion measure

Epsilon   list of n   Mean

$$\boxed{\varepsilon_j = l_j - \mu_L \quad \forall j}$$

Random Deviations, „random errors“

„actual – nominal“

$$\boxed{\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} l_1 - \mu_L \\ l_2 - \mu_L \\ \vdots \\ l_n - \mu_L \end{bmatrix} = \mathbf{l} - \mathbf{e} \cdot \mu_L}$$

Vector of deviations “error vector”

Values ε_j contain information about dispersion of single observation

Variance, if probability density is known (theoretical variance)

Definition of variance as measure of dispersion of a random variable (accuracy measure):

sigma = theoretical variance

$$\sigma_l^2 = E\left(\varepsilon^2\right) = E\{(l-\mu_L)^2\} = \boxed{E\left(l^2\right) - \mu_L^2}$$ - endpoint?

expectation of random error

$$\boxed{\sigma_l^2 = E\left(\varepsilon^2\right) = \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{j=1}^{n}\varepsilon_j^2\right\} = \lim_{n\to\infty}\left\{\frac{1}{n}\cdot\boldsymbol{\varepsilon}^{\mathrm{T}}\cdot\boldsymbol{\varepsilon}\right\}}$$ Theoretical Variance

Variance σ^2 is the mean of the squared ε_j

$$\boxed{\sigma_l = +\sqrt{\sigma_l^2}}$$ Theoretical Standard deviation

Standard deviation σ is the (positive) square root of the variance σ^2

2.4.1 Definition of dispersion measures

If we know the probability density function $f(x)$, we can compute the variance of a random variable directly (without observations):

$$\sigma_x^2 = E\left(\varepsilon^2\right) = E\{(x - \mu_X)^2\}$$

$$\boxed{\sigma_x^2 = \int_{+\infty}^{-\infty} (x - \mu_X)^2 f(x)\, dx}$$ Theoretical Variance

with $f(x) =$ density of x

2.4.2 Empirical dispersion measures

If we consider n observations (measurements, empirical values), we can compute the empirical standard deviation s as estimation for the theoretical standard deviation σ.

Attention: We have to distinguish between two cases

Case A	Case B
Expectation μ_L known	Expectation μ_L unknown

practical

theoretical

2.4.2.1 Expectation μ_L of the random variable is known (CASE A)

Given:

- Random variable L with its realizations $l_j\,,\; j = 1, 2, 3, \dots, n$
- Observation vector $\mathbf{l}^{\mathrm{T}} = (l_1 l_2 \dots l_n)$ with $n \ll \infty$
- Known expectation $E(L) = \mu_L$

$\Rightarrow$ Vector of random deviations $\boxed{\boldsymbol{\varepsilon} = \mathbf{L} - \mathbf{e} \cdot \mu_L}$

2.4.2.1 Expectation μ_L of the random variable is known (CASE A)

Wanted: Estimation s_l^2 for the theoretical variance σ_l^2

$$\boxed{s_l^2 = \frac{\varepsilon_1^2+\varepsilon_2^2+\ldots+\varepsilon_n^2}{n} = \frac{1}{n}\sum_{j=1}^{n}\varepsilon_j^2 = \frac{1}{n}\boldsymbol{\varepsilon}^{\mathrm{T}}\cdot\boldsymbol{\varepsilon}}$$ Empirical Variance of a single observation

Very important! KNOW!

$$\boxed{s_l = +\sqrt{s_l^2}}$$ Empirical Standard Deviation of a single observation

— also dispersion value ↳ mechanic

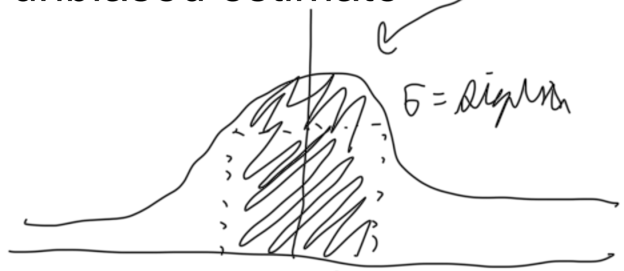
It holds: $E\left(s_l^2\right) = \sigma_l^2$ Empirical variance s_l^2 is an unbiased estimate of the theoretical variance σ_l^2.

It falls in

but: $E(s_l) \neq \sigma_l$ Empirical standard deviation s_l is not an unbiased estimate of the theoretical standard deviation σ_l !

If you take

Usually $E(s_l) < \sigma_l$



for $s_l = 2.24$ cm

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

Given: - Random variable L with its realizations l_j , $j = 1, 2, 3, \dots, n$

- Observation vector $\mathbf{l}^{\mathrm{T}} = (l_1 l_2 \dots l_n)$ with $n \ll \infty$

Not known: Expectation $E(L) = \mu_L$

Wanted: Estimation s_l^2 for the theoretical variance σ_l^2

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

Mean!

Solution: Replace expectation μ_L by the mean value $\bar{l}$

$$\boxed{\bar{l} = \frac{1}{n}\sum_{j=1}^{n} l_j = \frac{1}{n} \cdot \mathbf{e}^{\mathrm{T}} \cdot \mathbf{l}}$$

Empirical mean or arithmetic mean

with $\mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

Mean

observations

$$\boxed{v_j = \bar{l} - l_j \quad \forall j}$$

Residuals

Must be 0

Check: $\sum_{j=1}^{n} v_j \overset{!}{=} 0$

$$\boxed{\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \bar{v} - v_1 \\ \bar{v} - v_2 \\ \vdots \\ \bar{v} - v_n \end{bmatrix} = \mathbf{e} \cdot \bar{l} - \mathbf{l}}$$

Vector of residuals – Slide 42

← computational error

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

$$\boxed{s_l^2 = \frac{v_1^2+v_2^2+\ldots+v_n^2}{\underbrace{(n-1)}_{f}} = \frac{1}{(n-1)}\sum_{i=1}^{n} v_i^2 = \frac{1}{n-1}\mathbf{v}^{\mathrm{T}}\cdot\mathbf{v}}$$

Empirical Variance of a single observation

f: degree of freedom – how much info do you have too much

- "redundancy"

– you measure 1Vx = 1° of freedom

$$\boxed{s_l = +\sqrt{s_l^2}}$$

Empirical Standard Deviation of a single observation

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

It holds: $E(s_l^2) = \sigma_l^2$ Empirical variance s_l^2 is an unbiased estimate of the theoretical variance σ_l^2.

but: $E(s_l) \neq \sigma_l$ Empirical standard deviation s_l is <u>not</u> an unbiased estimate of the theoretical standard deviation σ_l .

Usually $E(s_l) < \sigma_l$.

One “information” (one degree of freedom f) ”number of redundant observations” is used for the estimation of the mean $\bar{l} \Rightarrow$ we have to divide by $f = n - 1$.

2.4.2.3 Standard Deviation of an arithmetic mean

Given: - Standard deviation of a single observation s_l from random deviations (CASE A from 2.4.2.1) or from residuals (CASE B form 2.4.2.2)

- Arithmetic mean $\bar{l}$ from n observations

Definitions: Variance of an arithmetic mean $\bar{l}$ from n observations

Residual vector

scale, rotation, shear

$y \to \uparrow x = \beta$ = vector

$$\boxed{s_{\bar{l}}^2 = \frac{s_l^2}{n}}$$

Standard deviation of an arithmetic mean $\bar{l}$ from n observations

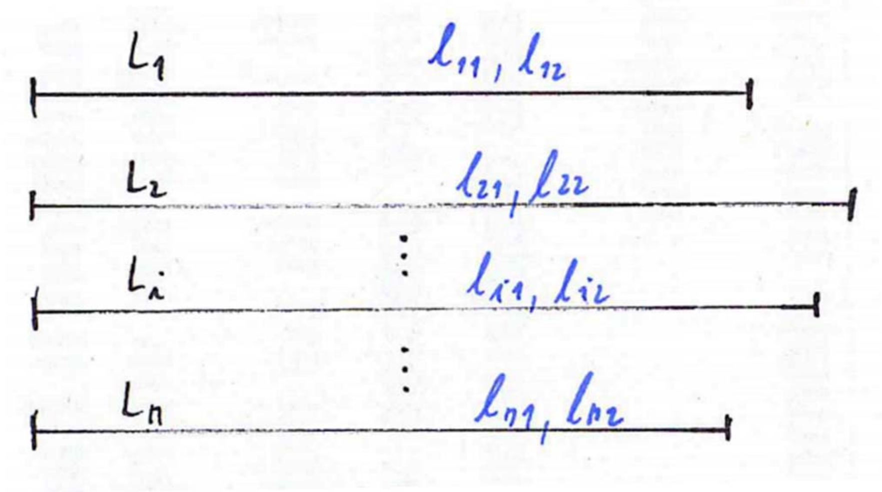
Slide 39

$$\boxed{s_{\bar{l}} = \frac{s_l}{\sqrt{n}}}$$

Standard deviation of a single observation s_l can be computed from random deviations or from residuals.

2.4.2.4 Standard Deviation for double measurements of same precision

Given: Several random variables have each been measured 2 times



$\mathbf{L}^{\mathrm{T}} = [L_1\ L_2\ \cdots\ L_n]$ Random Vector

$\mathbf{l}_1^{\mathrm{T}} = [l_{11}\ l_{21}\ \cdots\ l_{n1}]$ first series of observations

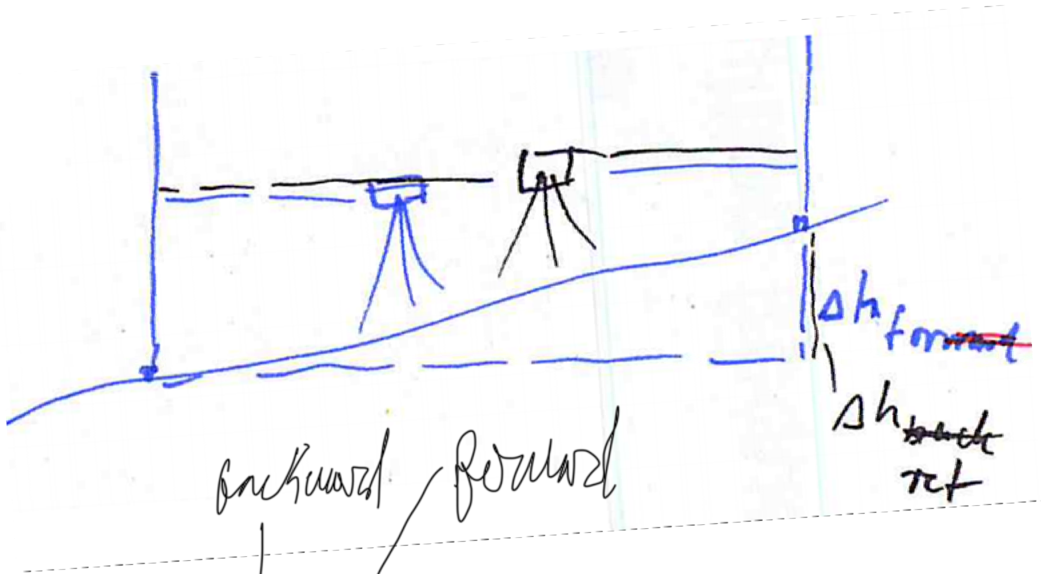
$\mathbf{l}_2^{\mathrm{T}} = [l_{12}\ l_{22}\ \cdots\ l_{n2}]$ second series of observations

Wanted: Estimation s_l^2 for the theoretical variance σ_l^2

Solution: Computation of differences from double measurements

2.4.2.4 Standard Deviation for double measurements of same precision

<u>Example</u>: Difference in the height difference from forward survey and return survey ($d = \Delta h_{for} - \Delta h_{ret}$) in differential levelling



Observation differences d_j for each double measurement

$$\boxed{d_j = l_{j2} - l_{j1}} \quad \text{for } j = 1, 2, \dots, n$$

$$\boxed{\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = \mathbf{l}_2 - \mathbf{l}_1 = \begin{bmatrix} l_{12} - l_{11} \\ l_{22} - l_{21} \\ \vdots \\ l_{n2} - l_{n1} \end{bmatrix}}$$

2.4.2.4 Standard Deviation for double measurements of same precision

Standard deviation of a single observation l_{j1} or l_{j2}

$$\boxed{s_l = \sqrt{\frac{1}{2n}\mathbf{d}^{\mathrm{T}} \cdot \mathbf{d}} = \sqrt{\frac{d_1^2 + d_2^2 + \ldots + d_n^2}{2n}} = \sqrt{\frac{\sum d^2}{2n}}}$$

Empirical Standard Deviation of a single observation

Standard Deviation $s_{\bar{l}}$ of the arithmetic mean from both observations

$$\boxed{s_{\bar{l}} = \frac{s_l}{\sqrt{2}} = \frac{1}{2}\sqrt{\frac{\sum d^2}{n}}}$$

Empirical Standard Deviation of the arithmetic mean

Adjustment Theory I

Chapter 2: Random Variables

Prof. Dr.-Ing. Frank Neitzel