Weighted Least Squares: Example free network adjustment

Using the constraint $D_1'\hat{x} = [1 \quad 1 \quad 1 \quad 1 \quad 1]\hat{x} = 0$ the free levelling network solution is found from the extended normal equations

$$\begin{bmatrix} A'PA & D_1 \\ D_1' & 0 \end{bmatrix}\begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} N & D_1 \\ D_1' & 0 \end{bmatrix}\begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} A'Py \\ 0 \end{bmatrix}$$

after two iterations
(stop criterion $\|\hat{x}\|_2 < 10^{-10}$)

$\hat{e}'P\hat{e} = 22.3\,\text{mm}^2$

$\hat{x}$ [mm]	$\hat{H}$ [m]	$\hat{e}$ [mm]	$\hat{h}$ [m]
–0.9	93.4581	2.9	14.2981
–2.8	107.7562	–2.6	9.9975
–3.4	103.4556	0.0*	7.0060
5.1	100.4641	0.0*	17.5000
2.1	110.9581	–1.6	4.3006

* For the reason that points 4 and 5 are polar points their estimated residuals turn out to be always zero !

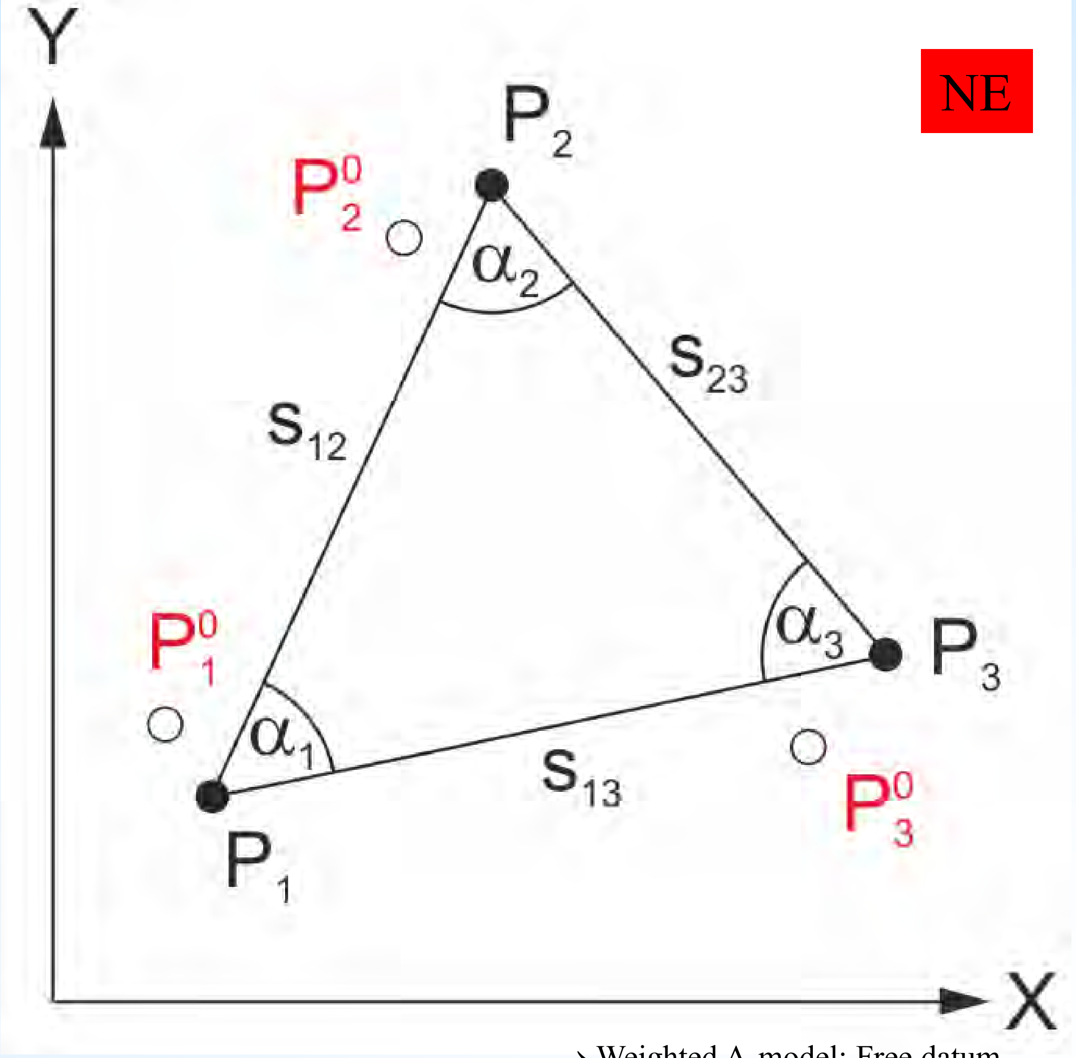
→ Weighted Least Squares: Example free network adjustment

Weighted Least Squares: Example free network adjustment

Example: Free adjustment of a triangle observed by angles and distances

	Observations	Weights
α_1	63.140 gon	p_α=10000 [1/gon^2]
α_2	51.520 gon	
α_3	85.350 gon	
s_{12}	122.400 m	p_s= 10000 [1/m^2]
s_{13}	91.000 m	
s_{23}	105.200 m	

	Approximate coordinates	
	X^0 [m]	Y^0 [m]
P_1	150.74	121.68
P_2	197.67	234.72
P_3	240.19	138.53



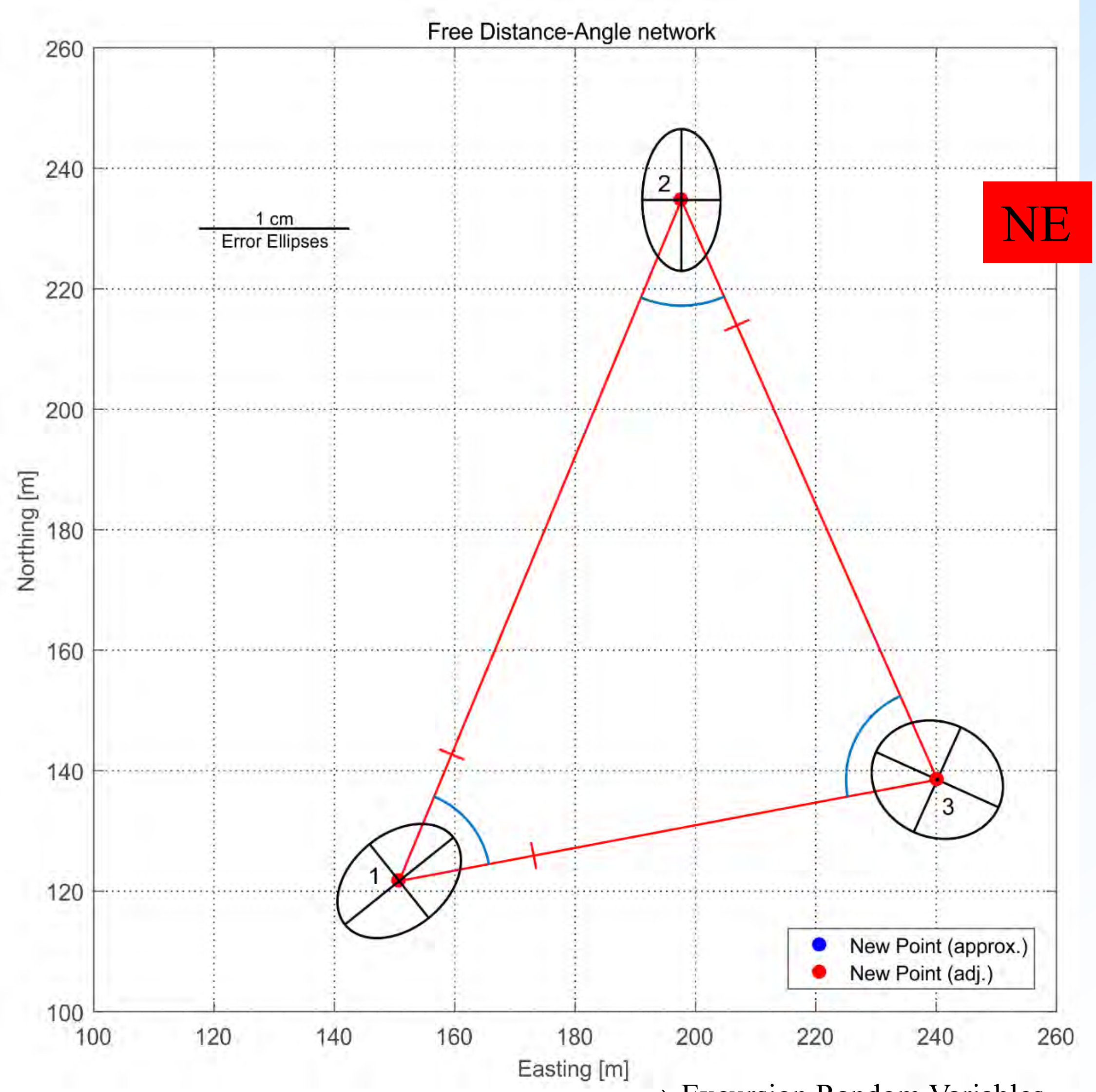
→ Weighted A-model: Free datum

Weighted Least Squares: Example free network adjustment

	Adjusted coordinates	
	$\hat{X}$ [m]	$\hat{Y}$ [m]
P_1	150.757	121.685
P_2	197.660	234.739
P_3	240.183	138.506

	Adjusted observations
$\hat{\alpha}_1$	63.1274 gon
$\hat{\alpha}_2$	51.5244 gon
$\hat{\alpha}_3$	85.3482 gon
$\hat{s}_{12}$	122.397 m
$\hat{s}_{13}$	90.994 m
$\hat{s}_{23}$	105.209 m

Sum of squares of coordinate corrections: 0.0014 m^2



NE

→ Excursion Random Variables

Excursion Random Variables

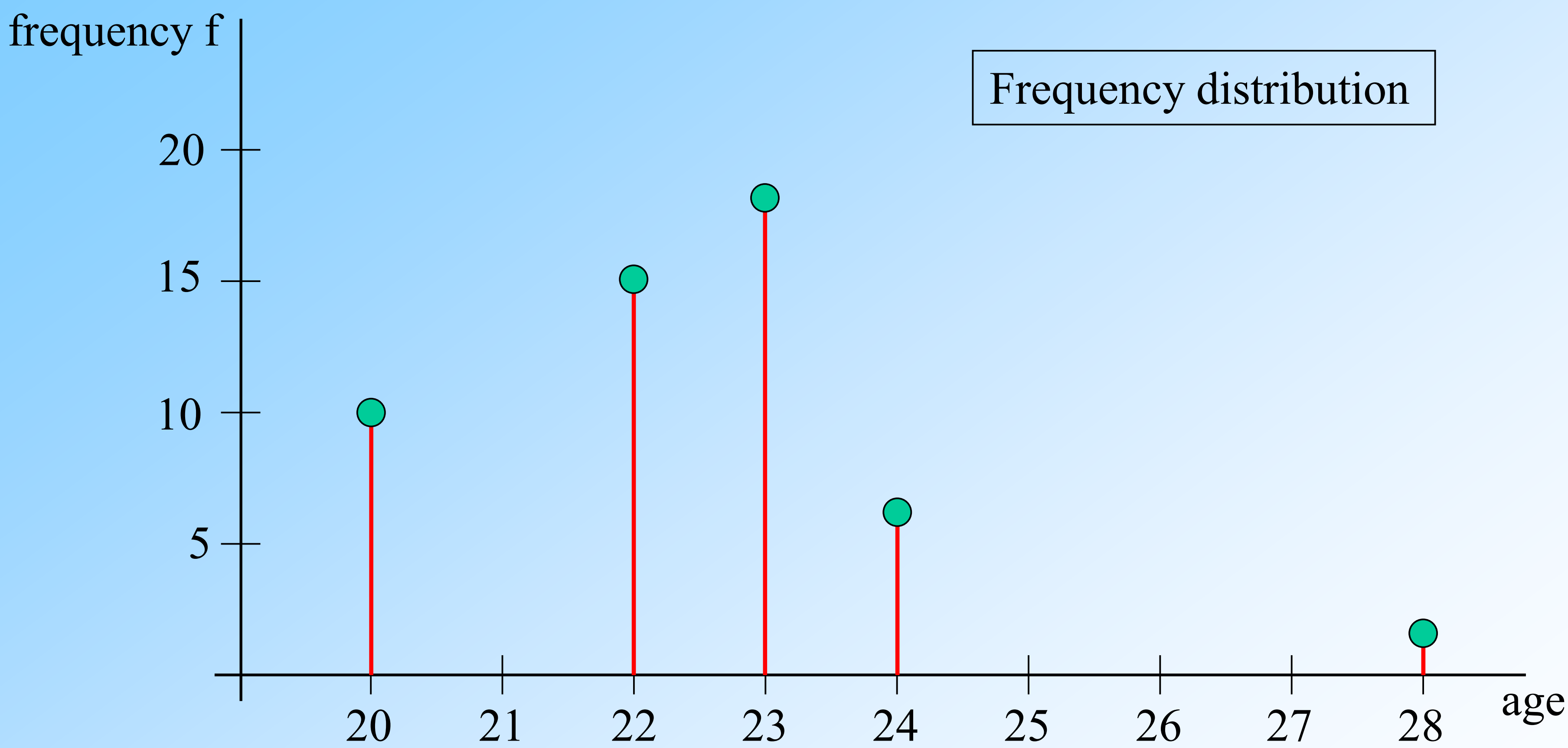
Example: We are interested in the age ($\underline{z}$) of Stuttgart University students and – as a representative for that – take a sample (z) in our class of m=50 students.

i (class)	z_i (age)	number of students = f_i (frequency)
1	20	10
2	21	0
3	22	15
4	23	18
5	24	6
6	25	0
7	26	0
8	27	0
9	28	1
$n = 9$ classes		$m = \sum_{i=1}^{n=9} f_i = 50$ students

The entire population (all students at our university) is specified by an underlined random variable $\underline{z}$, while the sample z (realizations of $\underline{z}$) is not underlined. $\underline{z}$ is a discrete random variable because we require the ages (here) to be positive <u>integers</u>.

→ Excursion Random Variables

Excursion Random Variables



$$\overline{z}\ \text{(average age)} = \frac{10\times 20+15\times 22+\ldots+1\times 28}{10+15+\ldots+1} = \frac{1116}{50} = 22{,}32$$

→ Excursion Random Variables

Excursion Random Variables

$$\bar{z} = \frac{1}{m}\sum_{i=1}^{n=9} z_i f_i = \sum_{i=1}^{n=9} z_i \frac{f_i}{m} = \sum_{i=1}^{n=9} z_i \frac{f_i}{\sum_{j=1}^{n=9} f_j}$$

relative frequencies

$$\text{as } n \to \infty: \quad \bar{z} = \int_{-\infty}^{\infty} z \frac{f(z)}{\int_{-\infty}^{\infty} f(z)dz} dz = \boxed{\int_{-\infty}^{\infty} z\, f(z)\, dz = E\{\underline{z}\}}$$

"Expectation" (1^{st} moment, mean)

probability density function ("pdf")

Note: Random variables are underlined, sample values (realizations) are not !

Generalization: higher moments (k-th moment): $E\{\underline{z}^k\} = \int_{-\infty}^{\infty} z^k\, f(z)\, dz$

Moments about the expected values (central moments): $E\{(\underline{z} - E\{\underline{z}\})^k\}$

k=2 (second central moment): $\sigma^2 = E\{(\underline{z} - E\{\underline{z}\})^2\} = \int_{-\infty}^{\infty} (z - E\{\underline{z}\})^2 f(z)\, dz$

σ^2 "Variance" $\Rightarrow \sigma$ "standard deviation"

→ Excursion Random Variables

Excursion Random Variables

Generalization: more than one random variable, e.g. $\underline{x}=[\underline{x}_1,\underline{x}_2]'$

1. moment: mean, expectation $\underset{2\times 1}{E\{\underline{x}\}} = E\left\{\begin{bmatrix}\underline{x}_1\\ \underline{x}_2\end{bmatrix}\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\begin{bmatrix}x_1\\ x_2\end{bmatrix} f(x_1,x_2)\,dx_1 dx_2$

joint pdf

2. central moment: dispersion, variance-covariance (matrix)

$$\underset{2\times 2}{\Sigma_x} = D\{\underline{x}\} = E\{[\underline{x}-E\{\underline{x}\}][\underline{x}-E\{\underline{x}\}]'\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x-E\{\underline{x}\}][x-E\{\underline{x}\}]' f(x_1,x_2)dx_1 dx_2$$

$$\Sigma_x = \begin{bmatrix}\sigma^2_{x_1} & \sigma_{x_1x_2}\\ \sigma_{x_2x_1} & \sigma^2_{x_2}\end{bmatrix},\quad \sigma^2_{x_1},\sigma^2_{x_2}\ \text{"variance"},\quad \sigma_{x_1x_2}=\sigma_{x_2x_1}\ \text{"covariance"}$$

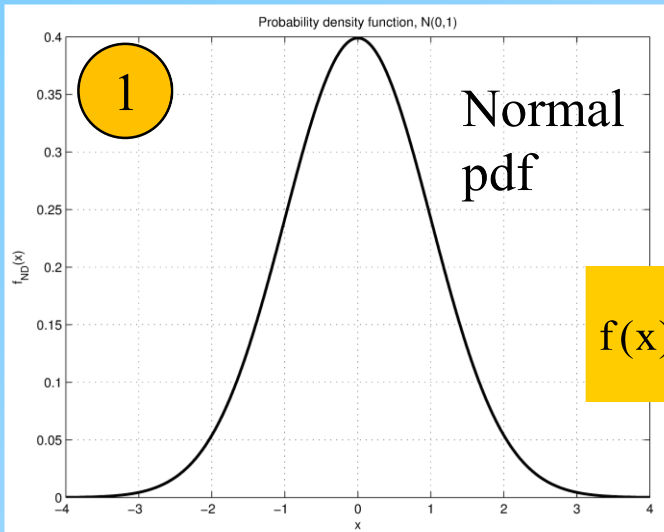
$$R_x = \begin{bmatrix}\sigma^{-1}_{x_1} & 0\\ 0 & \sigma^{-1}_{x_2}\end{bmatrix}\Sigma_x\begin{bmatrix}\sigma^{-1}_{x_1} & 0\\ 0 & \sigma^{-1}_{x_2}\end{bmatrix} = \begin{bmatrix}1 & r_{x_1x_2}\\ r_{x_2x_1} & 1\end{bmatrix}\quad \text{"correlation matrix"}$$

$$r_{x_1x_2} = \sigma_{x_1x_2}/(\sigma_{x_1}\sigma_{x_2})\quad \text{"correlation coefficient"}$$

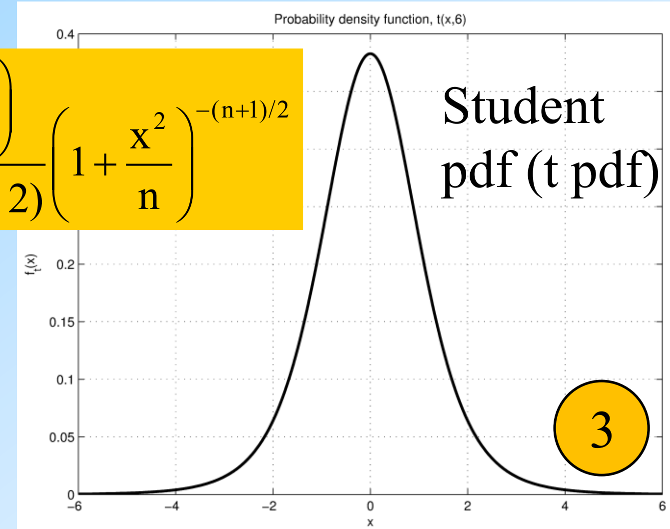
If two random variables $\underline{x}_1,\underline{x}_2$ are independent $\Rightarrow f(x_1,x_2)= f(x_1)f(x_2) \Rightarrow r_{x_1x_2}=0$

→ Excursion Random Variables

Excursion Random Variables

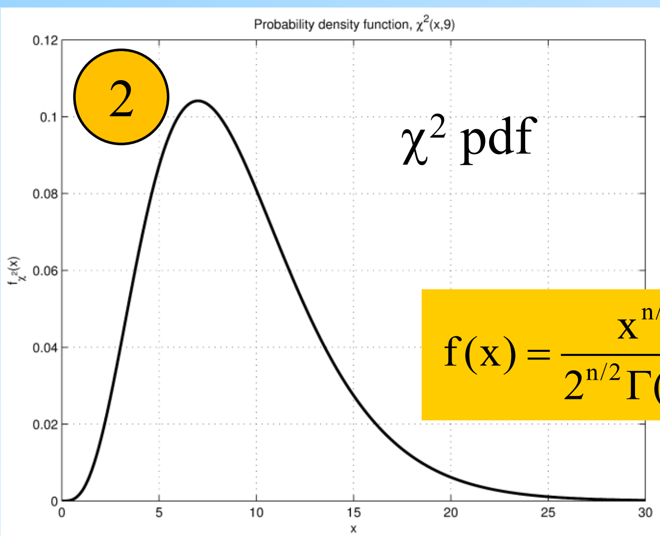


$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

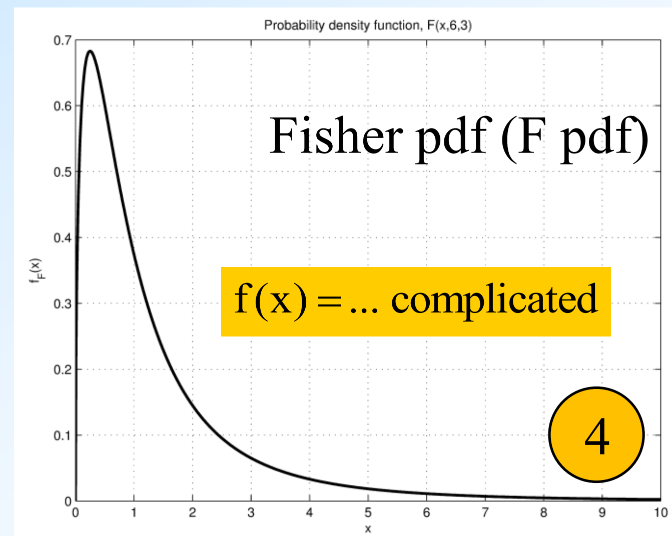


$$f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}$$

MATLAB $\rightarrow$
disttool, randtool



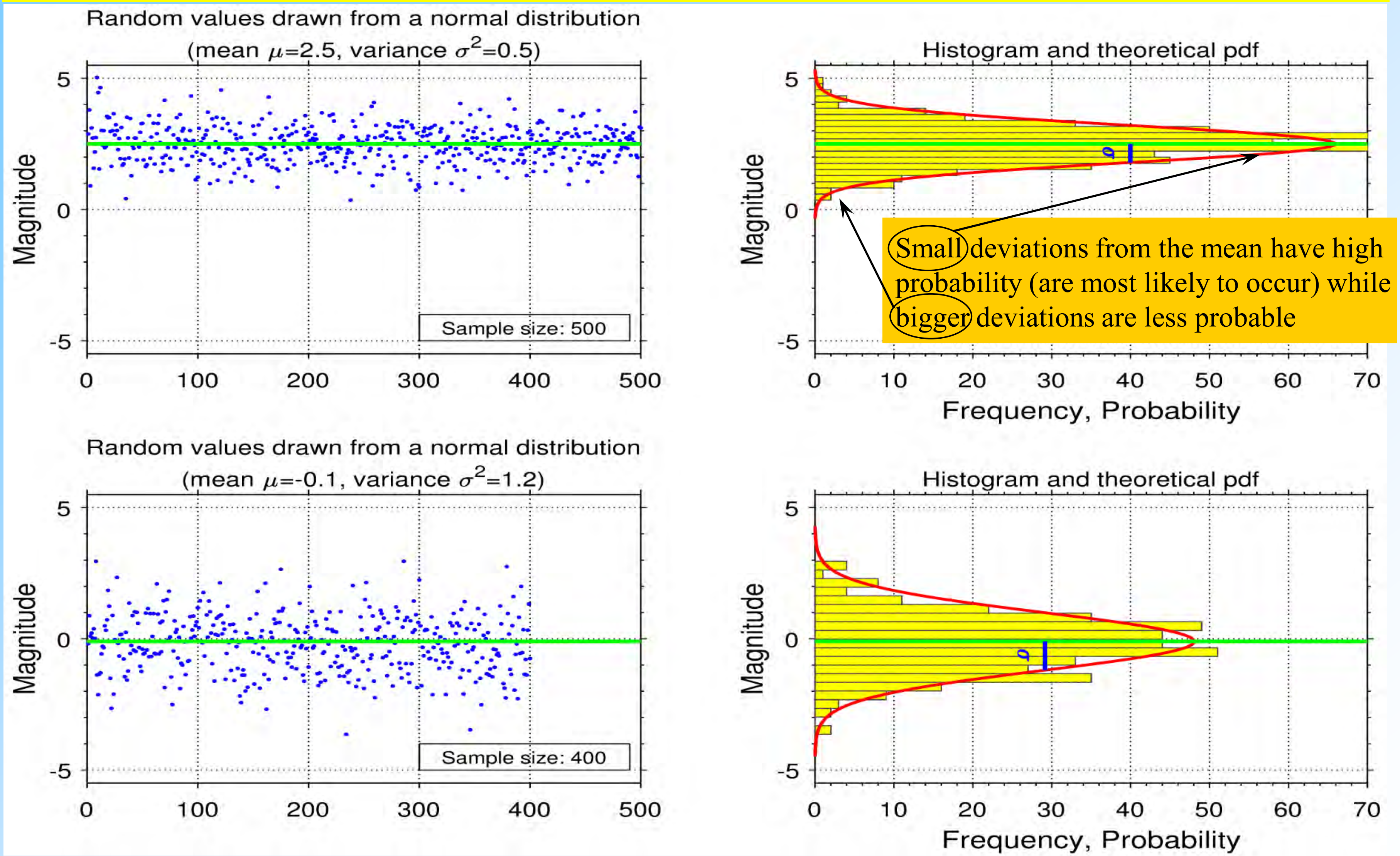
$$f(x)=\frac{x^{n/2-1}}{2^{n/2}\Gamma(n/2)}e^{-\frac{x}{2}}$$



$f(x) = \ldots$ complicated

$\rightarrow$ Excursion Random Variables

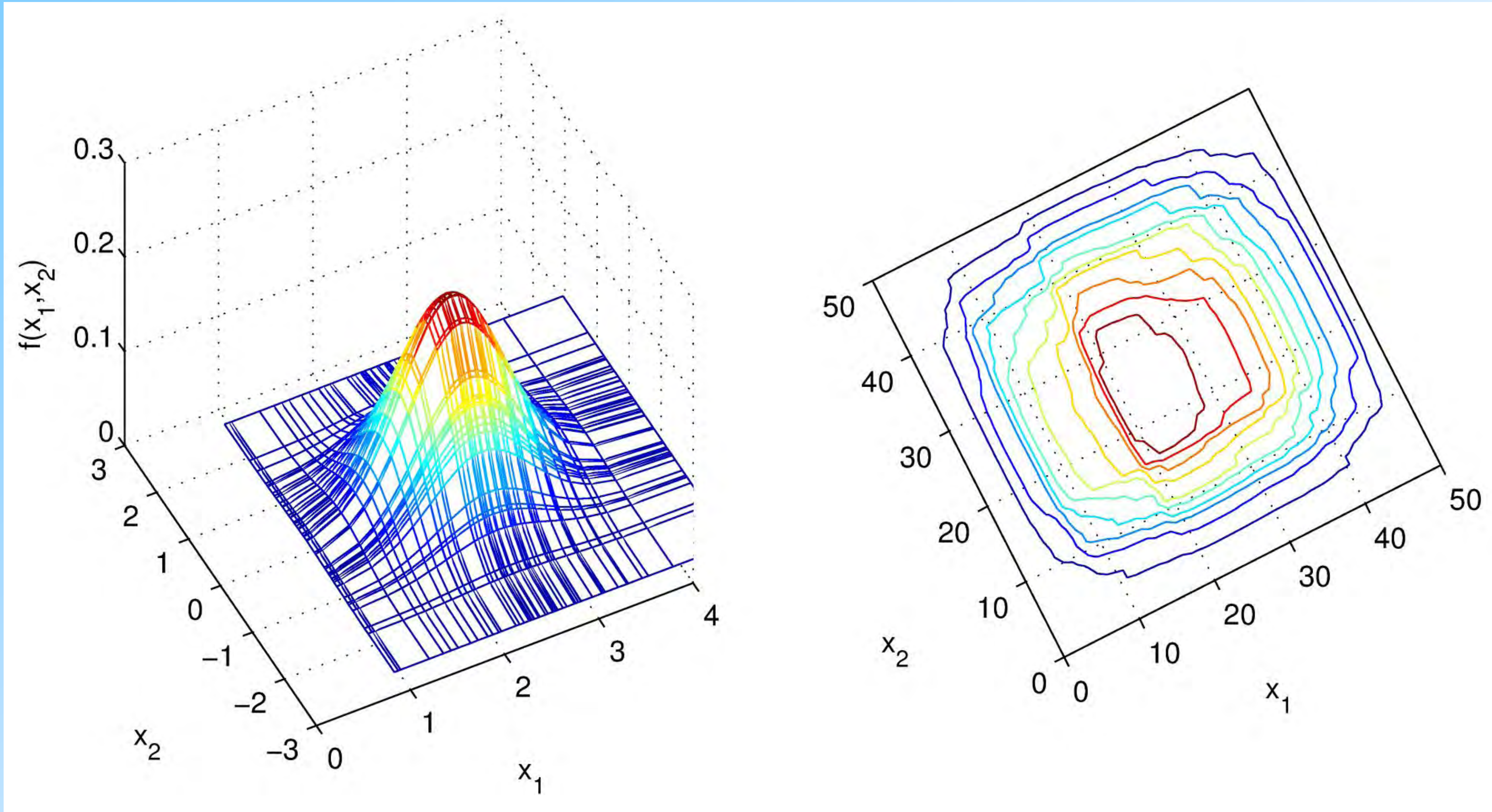
Excursion Random Variables



→ Excursion Random Variables

Excursion Random Variables

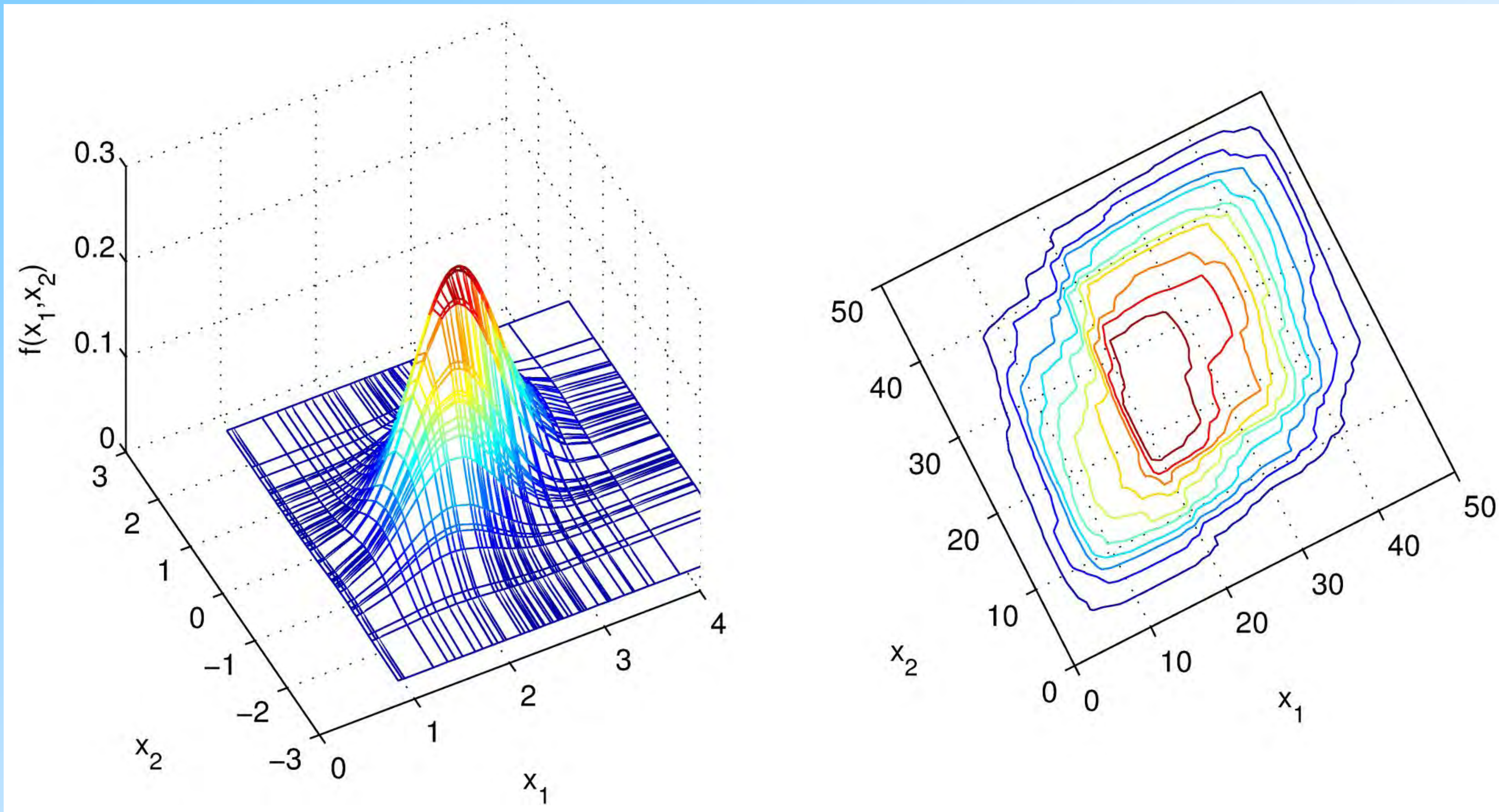
2-dimensional normal pdf: correlation coefficient r = 0



→ Excursion Random Variables

Excursion Random Variables

2-dimensional normal pdf: correlation coefficient r = 0.8

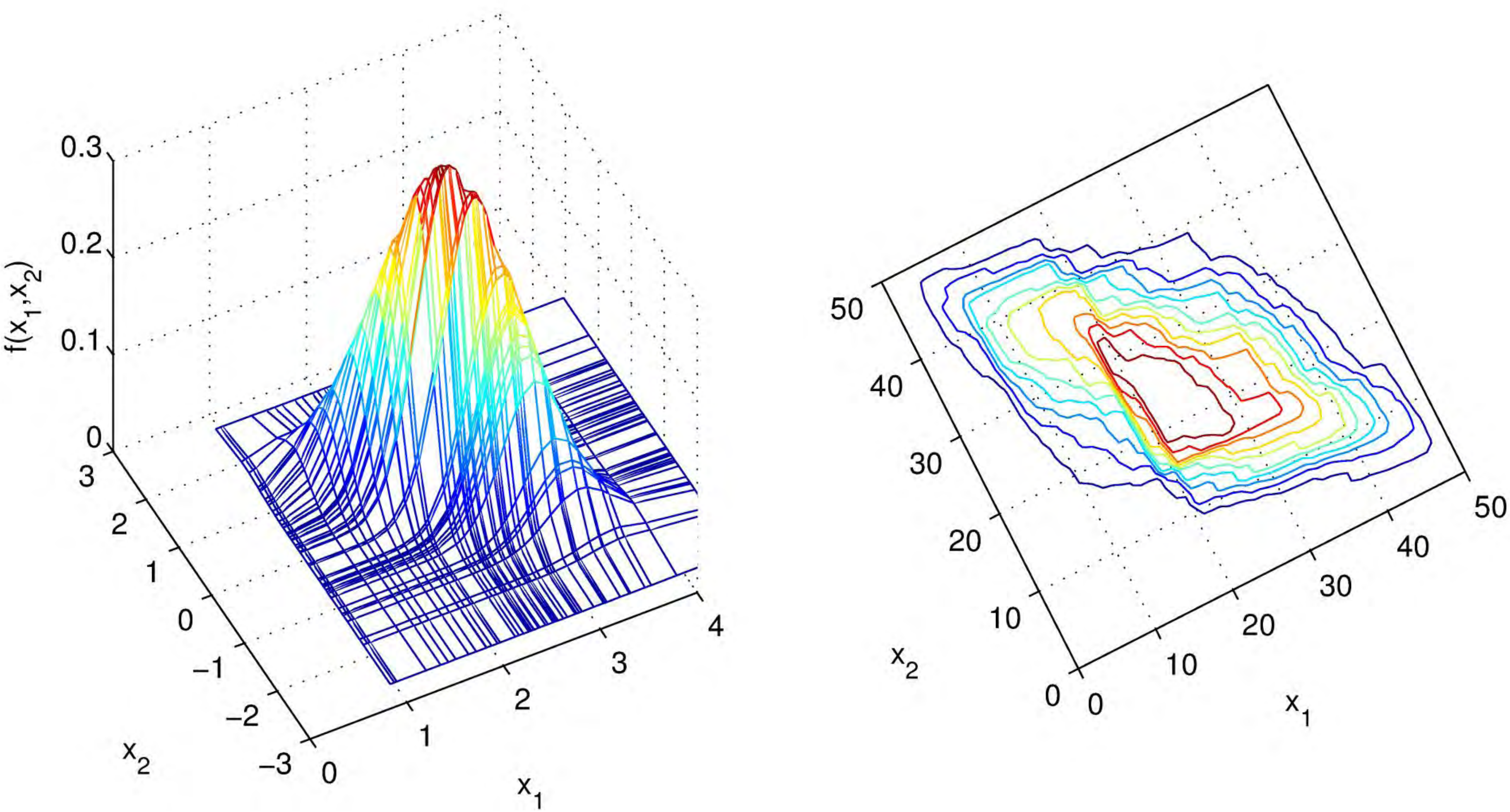


→ Excursion Random Variables

 GIS

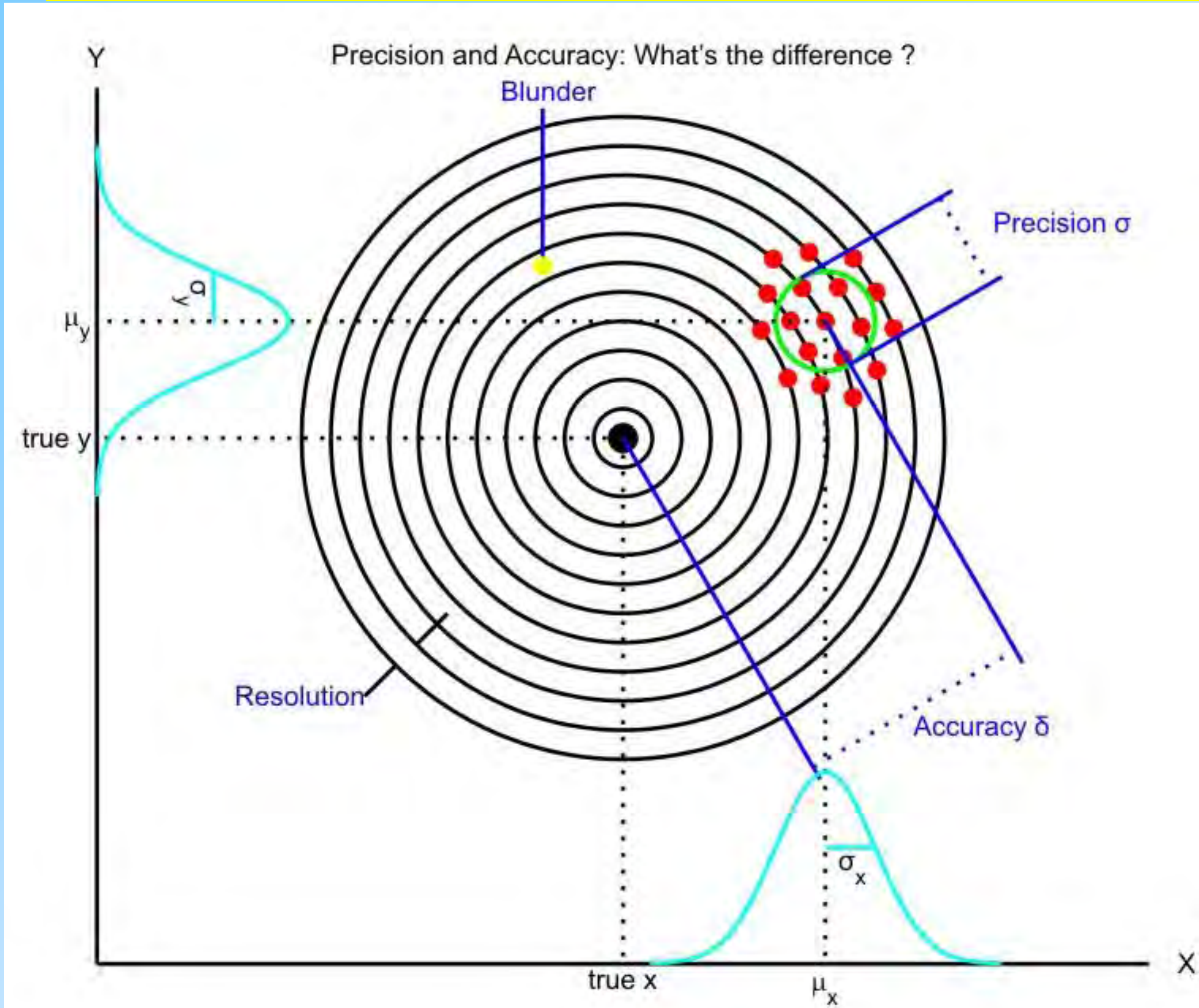
Excursion Random Variables

2-dimensional normal pdf: correlation coefficient r = –0.8



→ Excursion Random Variables

Excursion Random Variables



marks left behind by an arrow in the game of darts, in which one tries to aim at bull's eye

→ "New" adjustment model

"New" adjustment model

Inconsistencies ("measurement errors $\underline{e}$") are assumed to be random variables with the characteristics to be "mean-free", i.e. $E\{\underline{e}\}=0$, and to have a certain probability density function with given variance-covariance information $D\{\underline{e}\}$. $E\{\ldots\}$ and $D\{\ldots\}$ are expectation and dispersion operator, respectively. As from now on $\underline{e}$ is taken as a vector of random variables, observations become random, too. So to speak, stochasticity propagates through the model equation y=Ax+e, i.e. $\underline{y}=Ax+\underline{e}$. The new <u>mathematical model</u> now consists of two parts

$$\underbrace{\underset{m\times 1}{\underline{y}} = \underset{m\times n}{A}\ \underset{n\times 1}{x} + \underset{m\times 1}{\underline{e}}}_{\substack{\text{functional model,}\\ \text{observational model}}} \quad , \quad \underbrace{E\{\underline{e}\} = \underset{m\times 1}{0}\,,\ \underset{m\times m}{\Sigma_y} = D\{\underline{y}\} = D\{\underline{e}\} = \sigma^2 \underset{m\times m}{Q_y}}_{\text{stochastic model}}\ .$$

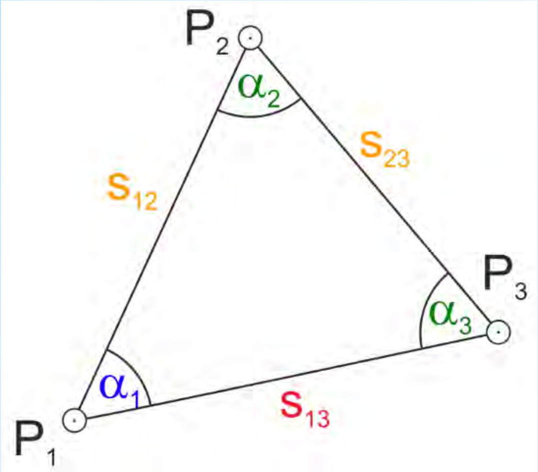
Given m×m matrix Q_y is called the matrix of cofactors (relative variances and covariances), and σ^2 is the variance factor (or variance of unit weight). It may be known or unknown, and accounts for the level of precision of the observations. If known, it is absorbed in Q_y and $D\{\underline{y}\}=Q_y$ is written. Alternatively, we may now write

$$\underbrace{E\{\underline{y}\} = Ax}_{\substack{\text{functional model,}\\ \text{observational model}}} \quad , \quad \underbrace{E\{\underline{e}\} = 0,\ D\{\underline{y}\} = D\{\underline{e}\} = \sigma^2 Q_y}_{\text{stochastic model}}$$

→ "New" adjustment model: Example

"New" adjustment model: Example

Triangle 1-2-3 has been observed with 4 different kinds of instruments given the following manufacturer specifications:



1st class theodolite ("best quality"):	$\sigma = \pm 0.001$ gon
3rd class theodolite ("acceptable quality"):	$\sigma = \pm 0.005$ gon
2nd class EDM instrument ("good quality"):	$\sigma = \pm 0.002$ m
4th class tape ("low quality"):	$\sigma = \pm 0.01$ m

Observations | **Variance-covariance matrix** (symmetric)

$$\underline{y} = \begin{bmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \\ \underline{\alpha}_3 \\ \hline \underline{s}_{12} \\ \underline{s}_{13} \\ \underline{s}_{23} \end{bmatrix}_{\text{gon},\text{m}}, \quad \Sigma_y = \left[\begin{array}{ccc|ccc} 2.5\times10^{-5} & & & & & \\ & 10^{-6} & & & 0 & \\ & & 10^{-6} & & & \\ \hline & & & 4\times10^{-6} & & \\ & 0 & & & 10^{-4} & \\ & & & & & 4\times10^{-6} \end{array}\right]_{\text{gon}^2,\text{m}^2}$$

variances | covariances | [gon × m]

covariances | variances

Often in practice, Σ_y is not known in absolute terms but only up to a (scale) factor σ^2 → relative variances and covariances (cofactors) $Q_y = \sigma^{-2}\Sigma_y$, $\Sigma_y = \sigma^2 Q_y$

→ Error propagation

Error propagation

If $\underline{z}$ is an arbitrary linear function of $\underline{y}$, i.e. $\underline{z} = L\underline{y}$ (we transform $\underline{y}$ into $\underline{z}$ using an arbitrary s×m matrix L), we are interested in $E\{\underline{z}\}$ and $D\{\underline{z}\}$.

Starting from the properties of $\underline{y}$, $E\{\underline{y}\}=\mu_y$ and $D\{\underline{y}\}=\Sigma_y$, we have $E\{\underline{z}\}=E\{L\underline{y}\}=LE\{\underline{y}\}=L\mu_y$, since the expectation operator is a linear operator and L is a non-stochastic quantity. In order to derive $D\{\underline{z}\}$ we use the definition of the dispersion operator:

$$\underset{s\times s}{\Sigma_z} = D\{\underline{z}\} = E\{[\underline{z} - E\{\underline{z}\}][\underline{z} - E\{\underline{z}\}]'\} = E\{[L\underline{y} - E\{L\underline{y}\}][L\underline{y} - E\{L\underline{y}\}]'\} =$$

$$= E\{[L\underline{y} - LE\{\underline{y}\}][L\underline{y} - LE\{\underline{y}\}]'\} = E\{L[\underline{y} - E\{\underline{y}\}][\underline{y} - E\{\underline{y}\}]'L'\} =$$

$$= LE\{[\underline{y} - E\{\underline{y}\}][\underline{y} - E\{\underline{y}\}]'\}L' =$$

$$= LD\{\underline{y}\}L' = \underset{s\times m}{L}\ \underset{m\times m}{\Sigma_y}\ \underset{m\times s}{L'}$$

(Linear) Error Propagation Law

Let us apply these rules to all quantities we know in the A-model !

→ Weighted A-model (revisited)

Weighted A-model (revisited)

1) $\hat{\underline{x}} = (A'PA)^{-1} A'P\underline{y} = L\underline{y}$, $L := (A'PA)^{-1} A'P$

2) $\hat{\underline{y}} = A\hat{\underline{x}}$ 3) $\hat{\underline{e}} = \underline{y} - \hat{\underline{y}}$

1a) $E\{\hat{\underline{x}}\} = LE\{\underline{y}\} = (A'PA)^{-1} A'PE\{\underline{y}\} \overset{E\{\underline{y}\}=Ax}{=} (A'PA)^{-1} A'PAx = x$

The expectation of the adjusted parameters $\hat{\underline{x}}$ equals the true but unknown values x even if only a finite number of observations $\underline{y}$ is available. But as the number of observations goes to infinity the estimate $\hat{\underline{x}}$ is certainly the true but not accessible x. $\hat{\underline{x}}$ is an unbiased estimate of x.

1b) $\Sigma_{\hat{x}} := D\{\hat{\underline{x}}\} = (A'PA)^{-1} A'P\Sigma_y PA(A'PA)^{-1} = L\Sigma_y L' = \sigma^2 LQ_y L'$

(Will be simplified soon)

2a) $E\{\hat{\underline{y}}\} = AE\{\hat{\underline{x}}\} \overset{E\{\hat{\underline{x}}\}=x}{=} Ax = E\{\underline{y}\}$

2b) $\Sigma_{\hat{y}} := D\{\hat{\underline{y}}\} = A\Sigma_{\hat{x}}A' = AL\Sigma_y L'A' = \sigma^2 ALQ_y L'A'$

3a) $E\{\hat{\underline{e}}\} = E\{\underline{y}\} - E\{\hat{\underline{y}}\} \overset{E\{\underline{y}\}=E\{\hat{\underline{y}}\}}{=} 0 = E\{\underline{e}\}$

3b) $\Sigma_{\hat{e}} := D\{\hat{\underline{e}}\} = \Sigma_y - \Sigma_{\hat{y}} = P_A^{\perp}\Sigma_y$

→ Weighted A-model (revisited)

Weighted A-model (revisited)

Three questions: (a) Is there any connection between weight matrix P and variance-covariance matrix Σ_y or cofactor matrix Q_y?

(b) Is $\underline{\hat{x}}$ an optimal estimate of x, i.e. is it a best estimate ?

(c) Is $\Sigma_{\hat{x}}$ the smallest variance-covariance matrix among all possible variance-covariance matrices of $\underline{\hat{x}}$?

Probabilistic approach Least-Squares Adjustment (BLUE)

B(est) - Variances in $\Sigma_{\hat{x}}$ minimal $\Leftrightarrow \operatorname{tr} D\{\underline{\hat{x}}\} = \min$

L(inear) - $\underline{\hat{x}} = L\underline{y}$ "Gauß approach"

U(nbiased) - $E\{\underline{\hat{x}}\} = x \qquad \forall x \in \mathbb{R}^n$

Estimate

$\rightarrow$ BLUE

Best Linear Unbiased Estimate (BLUE)

1) Unbiasedness: $E\{\hat{\underline{x}}\} = E\{L\underline{y}\} = LE\{\underline{y}\} = LAx \overset{!}{=} \underset{n\times 1}{x} \Leftrightarrow LA = I_n$

2) Best:

$$\begin{aligned} \operatorname{tr} D\{\hat{\underline{x}}\} &= \operatorname{tr} E\{(\hat{\underline{x}} - E\{\hat{\underline{x}}\})(\hat{\underline{x}} - E\{\hat{\underline{x}}\})'\} \\ &= \operatorname{tr} E\{(L\underline{y} - E\{L\underline{y}\})(L\underline{y} - E\{L\underline{y}\})'\} \\ &= \operatorname{tr} E\{(L\underline{y} - LE\{\underline{y}\})(L\underline{y} - LE\{\underline{y}\})'\} \\ &= \operatorname{tr} LE\{(\underline{y} - E\{\underline{y}\})(\underline{y} - E\{\underline{y}\})'L'\} \\ &= \operatorname{tr} LD\{\underline{y}\}L' \\ &= \operatorname{tr} L\Sigma_y L' \end{aligned}$$

3) Lagrange function: $\mathcal{L}(L,\Lambda) = \frac{1}{2}\operatorname{tr}\left(\underset{n\times m}{L}\ \underset{m\times m}{\Sigma_y}\ \underset{m\times n}{L'}\right) + \operatorname{tr}\left[\underset{n\times n}{\Lambda'}(\underset{n\times m}{A'}\ \underset{m\times n}{L'} - I_n)\right] = \min_{L,\Lambda}$

necessary conditions for a minimum of $\mathcal{L}$:

$$\frac{\partial\mathcal{L}}{\partial L}(\hat{L},\hat{\Lambda}) = \underset{m\times m}{\Sigma_y}\ \underset{m\times n}{\hat{L}'} + \underset{m\times n}{A}\ \underset{n\times n}{\hat{\Lambda}} = \underset{m\times n}{0} \Rightarrow \hat{L}' = -\Sigma_y^{-1}A\hat{\Lambda}$$

$$\frac{\partial\mathcal{L}}{\partial \Lambda}(\hat{L},\hat{\Lambda}) = \underset{n\times m}{A'}\ \underset{m\times n}{\hat{L}'} - I_n = \underset{n\times n}{0}$$

→ BLUE

Best Linear Unbiased Estimate (BLUE)

$$\Downarrow$$

$$-A'\Sigma_y^{-1}A\hat{\Lambda} = I_n$$

$$\Downarrow$$

$$\hat{\Lambda} = -(A'\Sigma_y^{-1}A)^{-1}$$

$$\Downarrow$$

$$\hat{L}' = \Sigma_y^{-1}A(A'\Sigma_y^{-1}A)^{-1}$$

$$\Downarrow$$

$$\hat{L} = (A'\Sigma_y^{-1}A)^{-1}A'\Sigma_y^{-1} = (A'\cancel{\sigma}^{-2}Q_y^{-1}A)^{-1}A'\cancel{\sigma}^{-2}Q_y^{-1} = (A'Q_y^{-1}A)^{-1}A'Q_y^{-1}$$

4) Solution: $\hat{L} = (A'Q_y^{-1}A)^{-1}A'Q_y^{-1}$, $\hat{\underline{x}} = (A'Q_y^{-1}A)^{-1}A'Q_y^{-1}\underline{y}$

5) Comparison with P-weighted approach: $Q_y^{-1} = P$ (high weights $\Leftrightarrow$ high precision instruments $\Leftrightarrow$ low variances $\Leftrightarrow$ low standard deviations).

→ Error propagation (revisited)

Error propagation (revisited)

Due to $Q_y^{-1} = P$ the variance-covariance matrix of the unknown parameters $\Sigma_{\hat{x}}$ becomes $\Sigma_{\hat{x}} = \sigma^2 (A' Q_y^{-1} A)^{-1} = \sigma^2 (A'PA)^{-1}$. It can be computed solely from the design matrix A and the precision of the instruments (Σ_y) prior to the measurement campaign $\Rightarrow$ Precision of the (adjusted) network parameters can be controlled by the choice of the equipment.

From error propagation we obtain

$$\underline{\hat{y}} = A\underline{\hat{x}} \quad \Rightarrow \Sigma_{\hat{y}} = A\Sigma_{\hat{x}}A' \quad = \ldots = P_A \Sigma_y \quad (P_A: \text{projector})$$

$$\underline{\hat{e}} = \underline{y} - \underline{\hat{y}} \Rightarrow \Sigma_{\hat{e}} = \Sigma_y - \Sigma_{\hat{y}} \quad = \ldots = P_A^{\perp} \Sigma_y \quad (P_A^{\perp}: \text{projector})$$

Visualization of $\Sigma_{\hat{x}}$: It is a positive definite n×n matrix consisting (in the 2D-case) of 2×2 submatrices, which represent on the diagonal variance-covariance matrices of single points (point coordinates). These are also positive definite and can be displayed as local error ellipses, delivering information on the uncertainties of the estimated coordinates.

→ Error ellipses

Error ellipses

Computation of absolute and relative error ellipses from $\Sigma_{\hat{x}} = D\{\underline{\hat{x}}\}$ (or $\hat{\Sigma}_{\hat{x}} = \hat{D}\{\underline{\hat{x}}\}$)
(Variance-covariance matrix of estimated point coordinates)

$$\Sigma_{\hat{\mathbf{x}}} = \left[\begin{array}{cc|cc} \boxed{\begin{array}{cc} \sigma^2_{\hat{x}_1} & \sigma_{\hat{x}_1\hat{y}_1} \\ \sigma_{\hat{y}_1\hat{x}_1} & \sigma^2_{\hat{y}_1} \end{array}} & & \begin{array}{cc} \sigma_{\hat{x}_1\hat{x}_2} & \sigma_{\hat{x}_1\hat{y}_2} \\ \sigma_{\hat{y}_1\hat{x}_2} & \sigma_{\hat{y}_1\hat{y}_2} \end{array} & \\ \hline \begin{array}{cc} \sigma_{\hat{x}_2\hat{x}_1} & \sigma_{\hat{x}_2\hat{y}_1} \\ \sigma_{\hat{y}_2\hat{x}_1} & \sigma_{\hat{y}_2\hat{y}_1} \end{array} & & \boxed{\begin{array}{cc} \sigma^2_{\hat{x}_2} & \sigma_{\hat{x}_2\hat{y}_2} \\ \sigma_{\hat{y}_2\hat{x}_2} & \sigma^2_{\hat{y}_2} \end{array}} & \end{array}\right]$$

symmetric

not necessarily symmetric

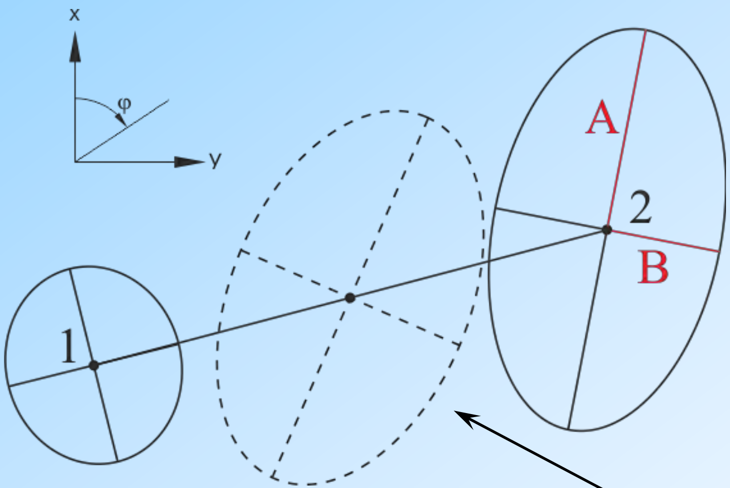
→ Error ellipses

Error ellipses

Example:

$\Sigma_{\hat{x}}$	x_1	y_1	x_2	y_2
x_1	0,01748	–0,00078	0,01552	–0,00426
y_1	–0,00078	0,01440	–0,001	0,00381
x_2	0,01552	–0,001	0,07124	0,00919
y_2	–0,00426	0,00381	0,00919	0,02568

$\Sigma_{\hat{P}_1}$ (upper-left block), $\Sigma_{\hat{P}_2}$ (lower-right block)



Absolute error ellipse (Coordinate error ellipse = Point error ellipse)

Relative error ellipse (Coordinate difference error ellipse)

NE

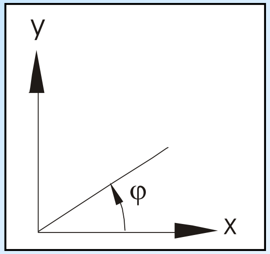
→ Error ellipses

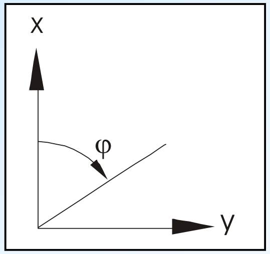
Error ellipses

Absolute error ellipse in P, bearing φ of semi major axis A

$$\tan 2\varphi = \frac{2\sigma_{\hat{x}\hat{y}}}{\sigma_{\hat{x}}^2 - \sigma_{\hat{y}}^2}$$

	$\sigma_{\hat{x}\hat{y}} > 0$	$\sigma_{\hat{x}\hat{y}} < 0$
$\sigma_{\hat{x}}^2 - \sigma_{\hat{y}}^2 > 0$	$0 < \varphi < \frac{\pi}{4}$	$\frac{3\pi}{4} < \varphi < \pi$
$\sigma_{\hat{x}}^2 - \sigma_{\hat{y}}^2 < 0$	$\frac{\pi}{4} < \varphi < \frac{\pi}{2}$	$\frac{\pi}{2} < \varphi < \frac{3\pi}{4}$





$$\sigma_{max}^2 = A^2 = \frac{1}{2}\left[\sigma_{\hat{x}}^2 + \sigma_{\hat{y}}^2 + \sqrt{\left(\sigma_{\hat{x}}^2 - \sigma_{\hat{y}}^2\right)^2 + 4\sigma_{\hat{x}\hat{y}}^2}\right] = \frac{1}{2}\left[\operatorname{tr}\Sigma_{\hat{P}} + \sqrt{\left(\operatorname{tr}\Sigma_{\hat{P}}\right)^2 - 4\det\Sigma_{\hat{P}}}\right]$$

$$\sigma_{min}^2 = B^2 = \frac{1}{2}\left[\sigma_{\hat{x}}^2 + \sigma_{\hat{y}}^2 - \sqrt{\left(\sigma_{\hat{x}}^2 - \sigma_{\hat{y}}^2\right)^2 + 4\sigma_{\hat{x}\hat{y}}^2}\right] = \frac{1}{2}\left[\operatorname{tr}\Sigma_{\hat{P}} - \sqrt{\left(\operatorname{tr}\Sigma_{\hat{P}}\right)^2 - 4\det\Sigma_{\hat{P}}}\right]$$

NE

→ Error ellipses

Error ellipses

Relative error ellipses between points j and k (Error ellipse for coordinate differences)

$$\underbrace{\Sigma_{[\widehat{x_j-x_k},\widehat{y_j-y_k}]^T}}_{2\times 2} = \underbrace{[I_2 \quad -I_2]}_{2\times 4} \underset{4\times 4}{\Sigma_{\hat{x}}} \underbrace{[I_2 \quad -I_2]^T}_{4\times 2}$$

Replace above

$$\sigma^2_{\hat{x}} \text{ by } \sigma^2_{\hat{x}_j} + \sigma^2_{\hat{x}_k} - 2\sigma_{\hat{x}_j\hat{x}_k}$$

$$\sigma^2_{\hat{y}} \text{ by } \sigma^2_{\hat{y}_j} + \sigma^2_{\hat{y}_k} - 2\sigma_{\hat{y}_j\hat{y}_k}$$

$$\sigma_{\hat{x}\hat{y}} \text{ by } \sigma_{\hat{x}_j\hat{y}_j} + \sigma_{\hat{x}_k\hat{y}_k} - \sigma_{\hat{x}_j\hat{y}_k} - \sigma_{\hat{y}_j\hat{x}_k}$$

Continuation example

φ_1=166°,5690
A_1=0,133
B_1=0,119

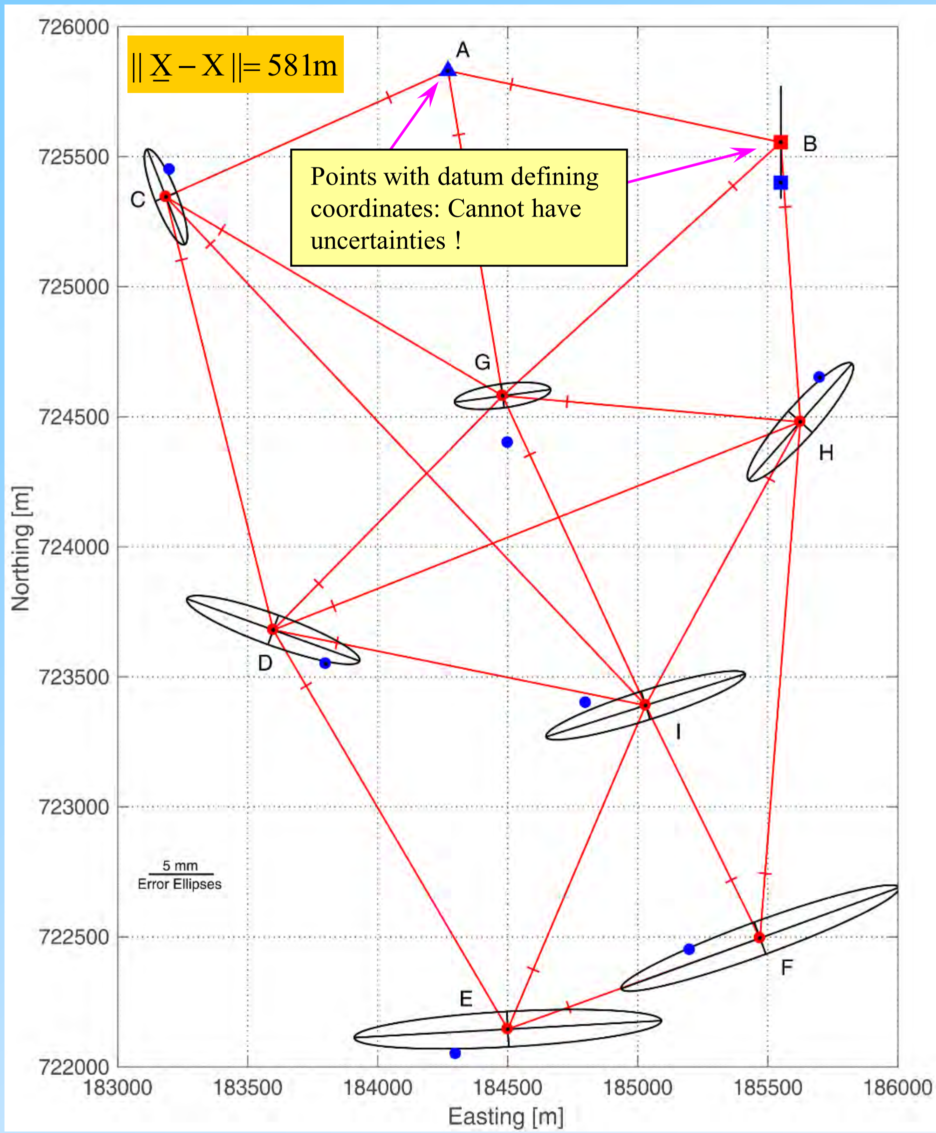
φ_2=10°,9852
A_2=0,270
B_2=0,155

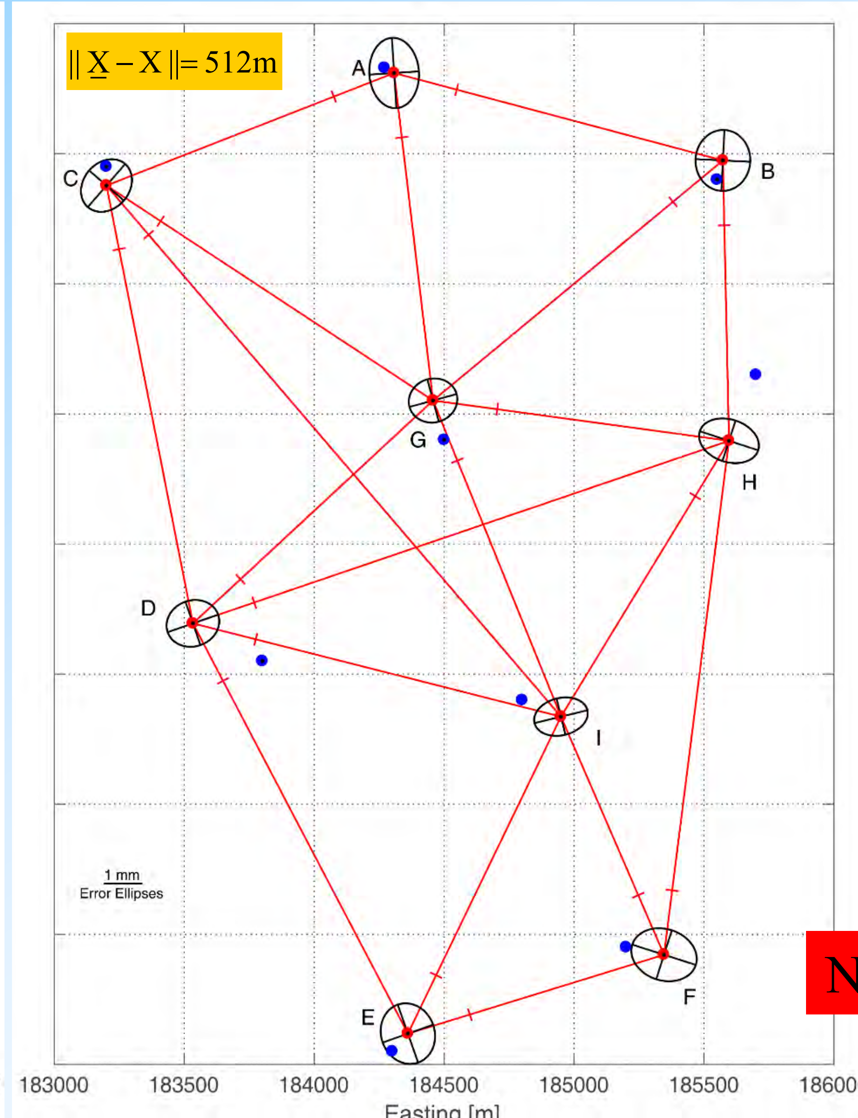
φ_{12}=23°,6549
A_{12}=0,252
B_{12}=0,1627

NE

→ Error ellipses distance network

Error ellipses distance network

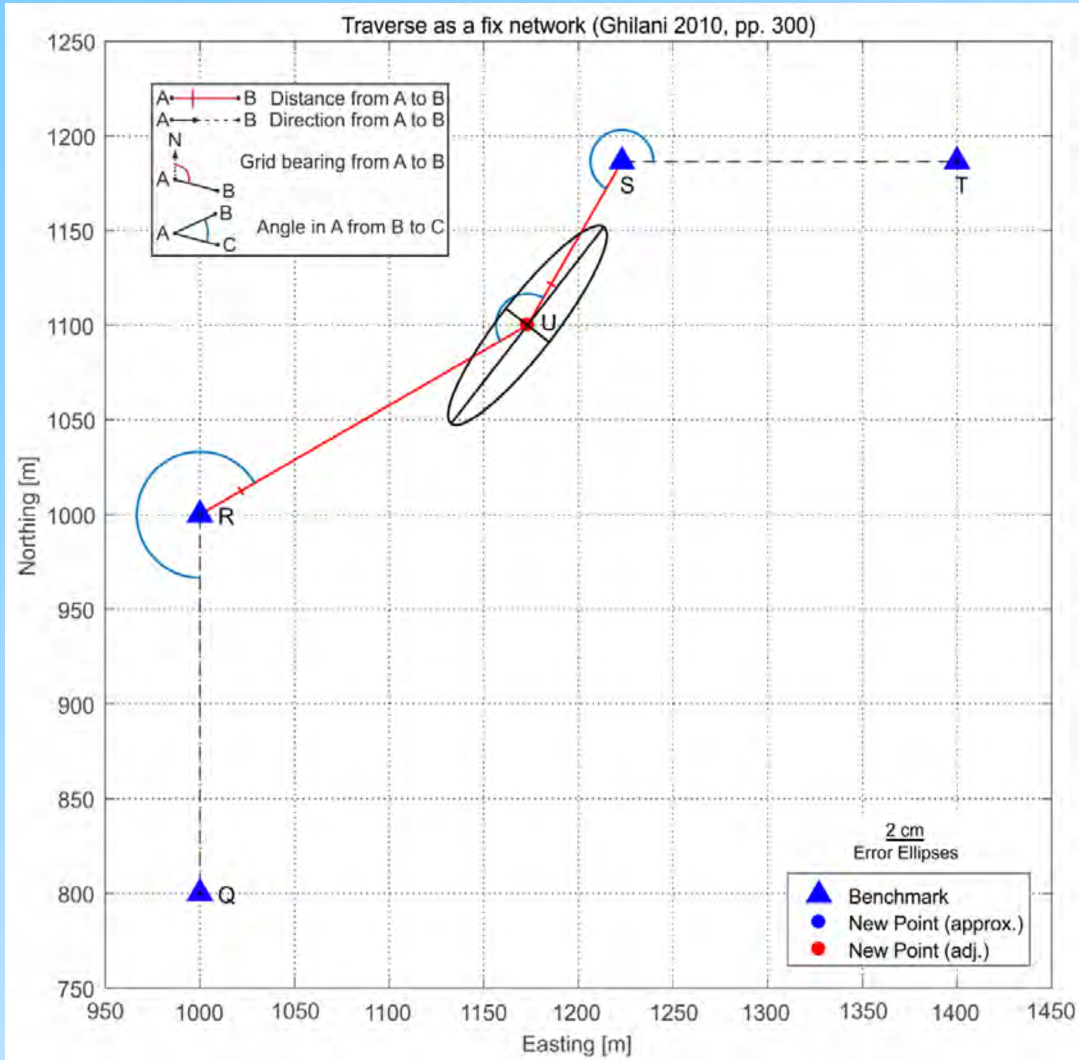




NE

→ Example Polygon Traverse

Example Polygon Traverse



	Easting x [m]	**Northing** y [m]
Q	1000.00	800.00
R	1000.00	1000.00
U	1173.20*	1100.00*
S	1223.00	1186.50
T	1400.00	1186.50

* Approximate value

	Distance [m]	σ_d [m]
R-U	200.00	0.05
U-S	100.00	0.08

in	**Angle**	σ_α
R	240°00'00"	30"
U	150°00'00"	30"
S	240°01'00"	30"

NE

→ Example Polygon Traverse

Example Polygon Traverse

$$\sigma^2 = 1 \Rightarrow P = \mathrm{diag}(\underset{[1/m]^2}{\sigma^2/\sigma^2_{d_{RU}}} \quad \underset{[1/m]^2}{\sigma^2/\sigma^2_{d_{US}}} \quad \underset{[1/°]^2}{\underbrace{\sigma^2/\sigma^2_{\alpha_R}}} \quad \underset{[1/°]^2}{\sigma^2/\sigma^2_{\alpha_U}} \quad \underset{[1/°]^2}{\sigma^2/\sigma^2_{\alpha_S}}) =$$

$$= \mathrm{diag}(\;400 \quad 156.25 \quad 14400 \quad 14400 \quad 14400\;)$$

Results ($\varepsilon = 10^{-10} \Rightarrow$ 4 iterations):

$$\widehat{\Delta x}_U = -11.1 \text{ cm} \;,\; \widehat{\Delta y}_U = 1.3 \text{ cm} \;,\; \hat{\sigma}^2 = 3.31$$

Adjusted coordinates:

$$\hat{x}_U = 1173.089 \text{ m} \pm 0.042 \text{ m} \quad , \quad \hat{y}_U = 1099.987 \text{ m} \pm 0.053 \text{ m}$$

Error ellipse elements:

Major semi axis: 6.6 cm , Minor semi axis: 1.4 cm , Grid bearing: 37°52'20"

Adjusted observations:

	Distance [m]	$\hat{e}$ [cm]	$\hat{\sigma}_{\hat{d}}$ [cm]
R-U	199.893	10.7	6.1
U-S	99.878	12.2	6.5

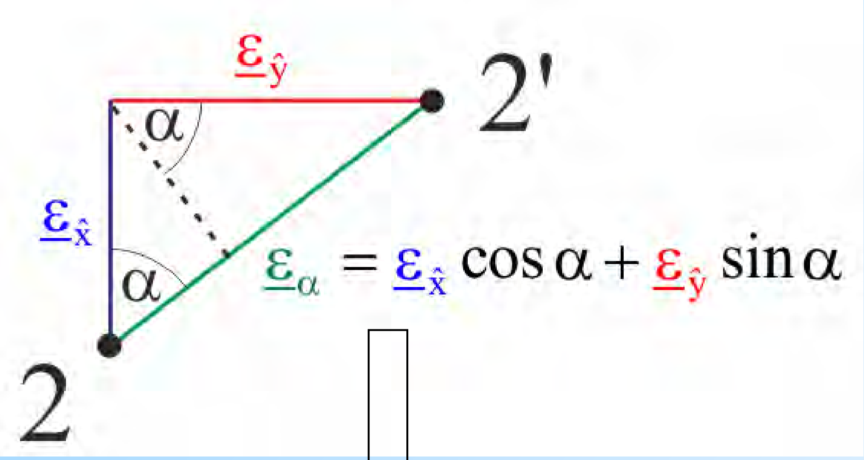
in	Angle	$\hat{e}$ ["]	$\hat{\sigma}_{\hat{\alpha}}$ ["]
R	239°59'11"	48.7	29.0
U	149°59'43"	17.2	44.1
S	240°01'06"	–5.8	35.0

NE

→ Error Curve

Error curve (pedal curve, support function)

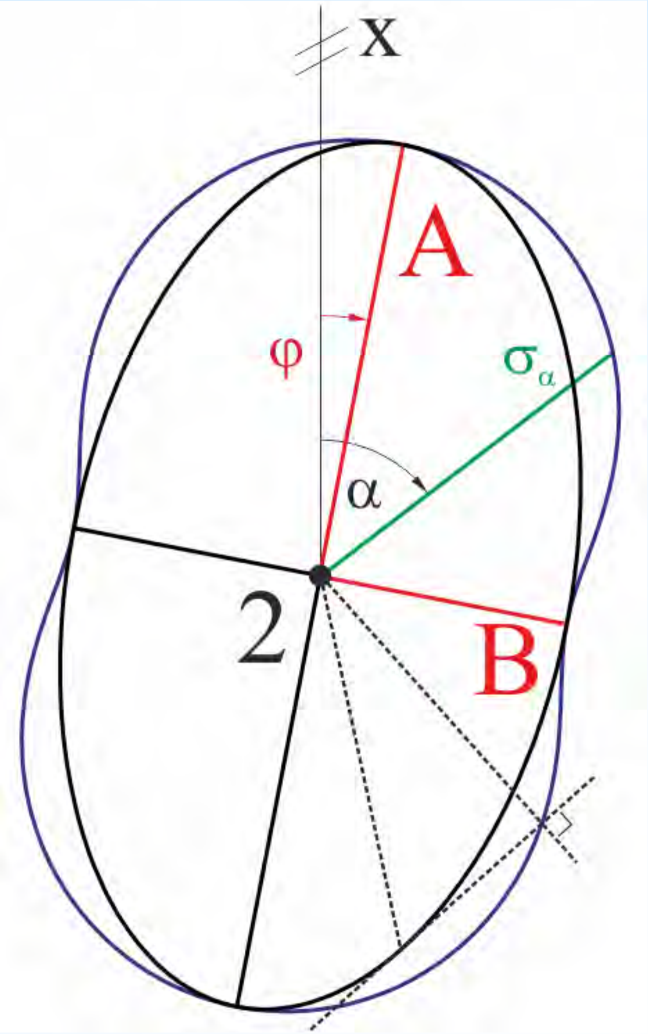
Error ellipses display minimum and maximum precision using $\sigma^2_{max} = A^2, \sigma^2_{min} = B^2$.
The error curve gives the standard deviation along any direction of azimuth α. Starting from random errors $\underline{\varepsilon}_{\hat{x}}, \underline{\varepsilon}_{\hat{y}}$ in direction of the coordinate axes the error $\underline{\varepsilon}_{\alpha}$ in direction of α is derived.



$$\underline{\varepsilon}_{\alpha} = \underline{\varepsilon}_{\hat{x}} \cos\alpha + \underline{\varepsilon}_{\hat{y}} \sin\alpha$$

Law of Error propagation

$$\sigma_\alpha = \sqrt{\sigma^2_{\hat{x}} \cos^2\alpha + \sigma^2_{\hat{y}} \sin^2\alpha + 2\sigma_{\hat{x}\hat{y}} \sin\alpha\cos\alpha}$$
$$= \sqrt{\sigma^2_{max} \cos^2(\alpha-\varphi) + \sigma^2_{min} \sin^2(\alpha-\varphi)}$$



→ Error propagation (revisited)

Error propagation (revisited)

If $\underline{z}$ is an arbitrary non-linear function of $\underline{y}$, i.e. $\underline{z} = f(\underline{y})$ (we transform $\underline{y}$ into $\underline{z}$ using an arbitrary non-linear function f), we are interested in $D\{\underline{z}\}$ $\Rightarrow$ Linearization necessary!

Example 1: Distance between two points A and B the coordinates of which are stochastic quantities, i.e.

$$\underline{s}_{AB} = \sqrt{(\underline{x}_A - \underline{x}_B)^2 + (\underline{y}_A - \underline{y}_B)^2} \Rightarrow$$

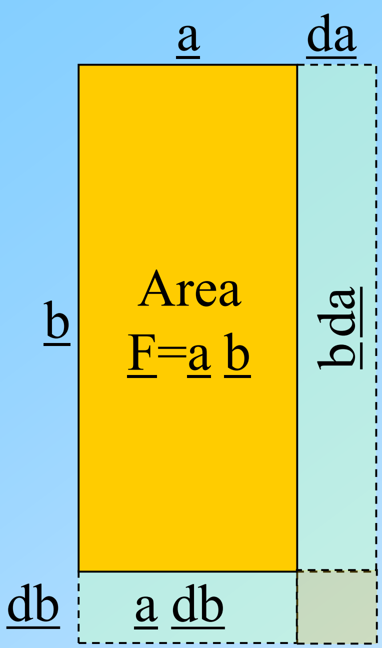
$$d\underline{s}_{AB} = \underbrace{\begin{bmatrix} \frac{x_A - x_B}{s_{AB}} & \frac{y_A - y_B}{s_{AB}} & -\frac{x_A - x_B}{s_{AB}} & -\frac{y_A - y_B}{s_{AB}} \end{bmatrix}}_{A} \begin{bmatrix} d\underline{x}_A \\ d\underline{y}_A \\ d\underline{x}_B \\ d\underline{y}_B \end{bmatrix} \Rightarrow$$

$$\sigma^2_{s_{AB}} = A \begin{bmatrix} \sigma^2_{x_A} & \sigma_{x_A y_A} & \sigma_{x_A x_B} & \sigma_{x_A y_B} \\ & \sigma^2_{y_A} & \sigma_{y_A x_B} & \sigma_{y_A y_B} \\ & \text{symmetric} & \sigma^2_{x_B} & \sigma_{x_B y_B} \\ & & & \sigma^2_{y_A} \end{bmatrix} A' \cancel{+ \ldots} = A\Sigma_x A' \cancel{+ \ldots}$$

→ Error propagation (revisited)

Error propagation (revisited)

Example 2: Variance of rectangular area $\underline{F}=\underline{a}\,\underline{b}$, measured $\underline{a}$ and $\underline{b}$



$$\underline{F} = \underline{a}\;\underline{b} \Rightarrow \underline{dF} = \frac{\partial F}{\partial a}\underline{da} + \frac{\partial F}{\partial b}\underline{db} \Rightarrow b\,\underline{da} + a\,\underline{db} = \underbrace{\begin{bmatrix} b & a \end{bmatrix}}_{A} \begin{bmatrix} \underline{da} \\ \underline{db} \end{bmatrix}$$

$$\sigma_F^2 = A \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix} A' = \begin{bmatrix} b & a \end{bmatrix} \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = b^2\sigma_a^2 + 2ab\sigma_{ab} + a^2\sigma_b^2$$

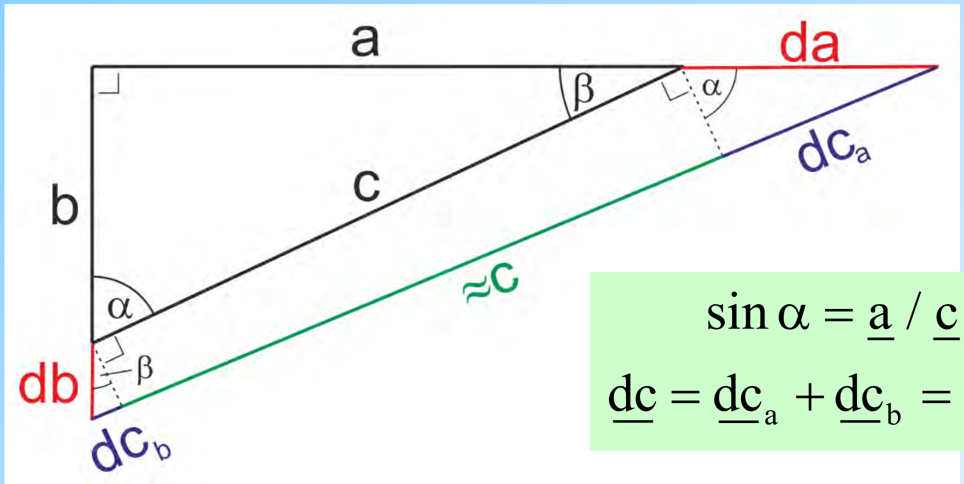
The short side should be measured with higher precision !

→ Error propagation (revisited)

Error propagation (revisited)

Example 3 (Pythagoras): Variance of hypotenuse $\underline{c} = \sqrt{\underline{a}^2 + \underline{b}^2}$, $\underline{a}$ and $\underline{b}$ observed

Total differential: ("how does $\underline{c}$ vary by variation of all other stochastic variables ?")



$$\underline{dc} = \frac{\partial c}{\partial a}\underline{da} + \frac{\partial c}{\partial b}\underline{db} = \underbrace{\frac{1}{c}\begin{bmatrix} a & b \end{bmatrix}}_{A}\begin{bmatrix} \underline{da} \\ \underline{db} \end{bmatrix}$$

$$\sin\alpha = \underline{a} / \underline{c} = \underline{dc}_a / \underline{da} \quad , \quad \sin\beta = \underline{b} / \underline{c} = \underline{dc}_b / \underline{db}$$

$$\underline{dc} = \underline{dc}_a + \underline{dc}_b = \underline{da}\sin\alpha + \underline{db}\sin\beta = \underline{a} / \underline{c}\;\underline{da} + \underline{b} / \underline{c}\;\underline{db}$$

Transition to variances

$$\sigma_c^2 = A\begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix}A' = \frac{a^2\sigma_a^2 + 2ab\sigma_{ab} + b^2\sigma_b^2}{c^2}$$
$$= \sigma_a^2 \sin^2\alpha + 2\sigma_{ab}\sin\alpha\sin\beta + \sigma_b^2\sin^2\beta$$

→ Stochastic model (revisited)

Stochastic model (revisited)

So far, the stochastic model was simply formulated in the form

$$\Sigma_y = D\{\underline{y}\} = D\{\underline{e}\} = \sigma^2 Q_y = \sigma^2 P^{-1}$$

where the equivalence of the cofactor matrix Q and the inverse weight matrix P resulted from comparing the weighted least-squares approach with the best linear unbiased estimation BLUE.

Now, there exist two possibilities, the first stating that the precision of the observation equipment – expressed through Σ_y – is exactly known, the other assuming that measurement precision is known only up to a (scale) factor, σ^2: $\Sigma_y = \sigma^2 Q_y$. This means that the relative observation weights are given, but the absolute precision level is unknown. This more general case leads to the so-called Gauß-Markoff model

$$\underbrace{\underline{y} = Ax + \underline{e}}_{\substack{\text{functional model,}\\ \text{observational model}}} \quad , \quad \underbrace{E\{\underline{e}\} = 0,\ D\{\underline{y}\} = \sigma^2 Q_y = \sigma^2 P^{-1}}_{\text{stochastic model}}.$$

→ Stochastic model (revisited)

Stochastic model (revisited)

Given matrix Q is called the "matrix of weight coefficients" or "cofactor matrix", unknown factor σ^2 the "variance of unit weight" or the "variance component". The impact of unknown σ^2 on the least-squares results will be discussed soon. In the frequent case of a diagonal variance-covariance matrix Σ_y the link between weights, cofactors and variances/covariances becomes

$$\underset{m\times m}{\Sigma_y} = \sigma^2 Q_y = \sigma^2 \begin{bmatrix} \sigma_1^2/\sigma^2 & & & 0 \\ & \sigma_2^2/\sigma^2 & & \\ & & \ddots & \\ 0 & & & \sigma_m^2/\sigma^2 \end{bmatrix}$$

$$\underset{m\times m}{P} = Q_y^{-1} = \begin{bmatrix} \sigma^2/\sigma_1^2 & & & 0 \\ & \sigma^2/\sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma^2/\sigma_m^2 \end{bmatrix} = \begin{bmatrix} p_{11} & & & 0 \\ & p_{22} & & \\ & & \ddots & \\ 0 & & & p_{mm} \end{bmatrix}$$

Although the physical units of variances/covariances are predefined by the type of observations, physical units of σ^2 and Q_y (or P) can be freely chosen: If σ^2 is chosen to be a dimensionless quantity then Q (or P) must have the dimensions of Σ_y (or Σ_y^{-1}) and vice versa.

→ Stochastic model (revisited)

Stochastic model (revisited)

How is the impact of a known or unknown σ^2 in the A-model, i.e. how do the estimated quantities are influenced by it ?

1) $\underline{\hat{x}} = \left(A'\sigma^{-2}PA\right)^{-1} A'\sigma^{-2}P\underline{y} = \left(A'PA\right)^{-1} A'P\underline{y}$
2) $\underline{\hat{y}} = A\underline{\hat{x}}$
3) $\underline{\hat{e}} = \underline{y} - \underline{\hat{y}}$

Do not depend on σ^2 and can always be computed.

4) $Q_{\hat{x}} = \left(A'PA\right)^{-1} \quad , \quad \Sigma_{\hat{x}} = \sigma^2 Q_{\hat{x}}$
5) $Q_{\hat{y}} = AQ_{\hat{x}}A' \quad , \quad \Sigma_{\hat{y}} = \sigma^2 Q_{\hat{y}}$
6) $Q_{\hat{e}} = Q_y - Q_{\hat{y}} \quad , \quad \Sigma_{\hat{e}} = \sigma^2 Q_{\hat{e}}$

Strongly depend on σ^2 and cannot be computed numerically for unknown σ^2

→ Stochastic model (revisited)

Stochastic model (revisited)

However, in case of unknown σ^2 an estimate $\underline{\hat{\sigma}}^2 = \underline{\hat{e}}' P \underline{\hat{e}} / (m - n)$ is available with the wonderful property of being unbiased, $E\{\underline{\hat{\sigma}}^2\} = \sigma^2$. So to speak, we are able to derive from a sample (of observations) an estimate $\underline{\hat{\sigma}}^2$ which equals the true but unknown σ^2 of the population and reflects the actually attained precision of the observations. This enables us to computes estimates 4')-6') for the "precision quantities" 4)-6).

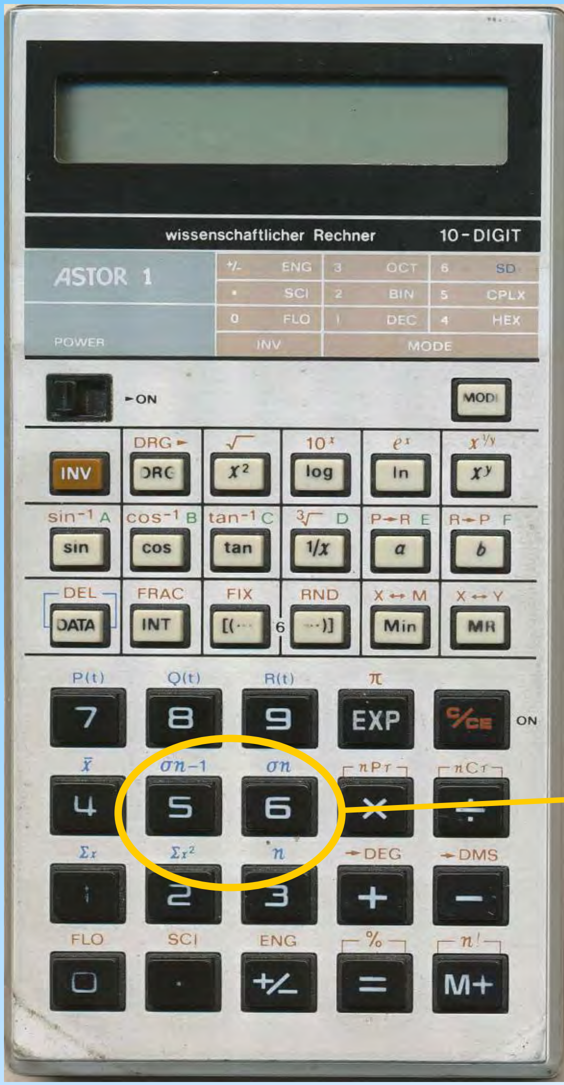
$$\left.\begin{array}{llll} 4') & Q_{\hat{x}} = (A'PA)^{-1} & , & \hat{\Sigma}_{\hat{x}} = \underline{\hat{\sigma}}^2 Q_{\hat{x}} \\ 5') & Q_{\hat{y}} = A Q_{\hat{x}} A' & , & \hat{\Sigma}_{\hat{y}} = \underline{\hat{\sigma}}^2 Q_{\hat{y}} \\ 6') & Q_{\hat{e}} = Q_y - Q_{\hat{y}} & , & \hat{\Sigma}_{\hat{e}} = \underline{\hat{\sigma}}^2 Q_{\hat{e}} \end{array}\right\} \quad \hat{\Sigma}_{\dots} \text{ independent of the chosen } \sigma^2$$

The point error ellipses in the trilateration network were computed this way.

Remark: Rather often the (a priori) assumption σ^2=1 (sometimes together with P=I) is made because no other reasonable information is available. Then, after the adjustment, a comparison between σ^2=1 and the (a posteriori) estimate $\underline{\hat{\sigma}}^2$ is made with the intention to find out if the assumption σ^2=1 was justified and to make conclusions about the observations and the model → statistical hypothesis testing

→ Stochastic model (revisited)

Stochastic model (revisited)



Data $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$

$\Downarrow$

Mean (Least squares estimate)	Mean given
$\hat{\underline{x}} = \frac{1}{n}\sum_{i}^{n} \underline{x}_i$	$\bar{x}$

Residuals

$$\underline{\bar{e}}_i = \underline{x}_i - \bar{x} \quad , \quad \underset{n\times 1}{\underline{\bar{e}}} = \underset{n\times 1}{\underline{x}} - \bar{x}$$

$$\underline{\hat{e}}_i = \underline{x}_i - \hat{\underline{x}} \quad , \quad \underset{n\times 1}{\underline{\hat{e}}} = \underset{n\times 1}{\underline{x}} - \hat{\underline{x}}$$

Square sum of residuals

$$\underline{\bar{e}}'\underline{\bar{e}}$$

$$\underline{\hat{e}}'\underline{\hat{e}}$$

$$\sigma_{n-1} = \sigma_n \sqrt{\frac{n}{n-1}}$$

$\hat{\underline{x}}$ estimated: σ_{n-1}	$\bar{x}$ given: σ_n
$\underline{\hat{\sigma}} = \sqrt{\frac{\underline{\hat{e}}'\underline{\hat{e}}}{n-1}}$	$\underline{\bar{\sigma}} = \sqrt{\frac{\underline{\bar{e}}'\underline{\bar{e}}}{n}}$
$E\{\underline{\hat{\sigma}}^2\} = \sigma^2$	$E\{\underline{\bar{\sigma}}^2\} \neq \sigma^2$

→ Stochastic model (revisited)

Stochastic model (revisited)

Balancing the physical units for an arbitrary choice of weight units using the example for distance and direction observations

$$\underline{y} = Ax + \underline{e} \;\; , \;\; \underline{\hat{x}} = (A'PA)^{-1} A'P\underline{y} = N^{-1} A'P\underline{y}$$

$$\begin{array}{r} \text{Distances:} \\ \text{Directions:} \end{array} \begin{bmatrix} \underline{y}_{s\,[m]} \\ \underline{y}_{r\,[gon]} \end{bmatrix} = \begin{bmatrix} A_{s\,[-]} \\ A_{r\,[gon/\,m]} \end{bmatrix} x_{[m]} + \begin{bmatrix} \underline{e}_{s\,[m]} \\ \underline{e}_{r\,[gon]} \end{bmatrix}$$

Choice of a priori standard deviation σ [m]

$$P = \begin{bmatrix} \underset{[-]}{P_s} & 0 \\ 0 & \underset{[m^2/gon^2]}{P_r} \end{bmatrix} \Rightarrow \underset{[-]}{N} = A'PA = \underset{[-]}{A_s'}\underset{[-]}{P_s}\underset{[-]}{A_s} + \underset{[gon/m]}{A_r'}\underset{[m^2/gon^2]}{P_r}\underset{[gon/m]}{A_r}$$

$$\underset{[m]}{A'P\underline{y}} = \underset{[-]}{A_s'}\underset{[-]}{P_s}\underset{[m]}{\underline{y}_s} + \underset{[gon/m]}{A_r'}\underset{[m^2/gon^2]}{P_r}\underset{[gon]}{\underline{y}_r}$$

$$\Rightarrow \underset{[m]}{\underline{\hat{x}}} = N^{-1}A'P\underline{y} \;\checkmark$$

$$\underset{[m^2]}{\underline{\hat{e}}'P\underline{\hat{e}}} = \underset{[m^2]}{\underline{\hat{e}}_s'P_s\underline{\hat{e}}_s} + \underset{[m^2]}{\underline{\hat{e}}_r'P_r\underline{\hat{e}}_r} \Rightarrow \underset{[m]}{\underline{\hat{\sigma}}} = \sqrt{\frac{\underline{\hat{e}}'P\underline{\hat{e}}}{m-n}}$$

Unit of a posteriori-value matches the unit of a priori-value

$$\underset{[m^2]}{\hat{\Sigma}_{\hat{x}}} = \underline{\hat{\sigma}}^2 Q_{\hat{x}} = \underset{[m^2]}{\underline{\hat{\sigma}}^2}\; \underset{[-]}{N^{-1}}$$

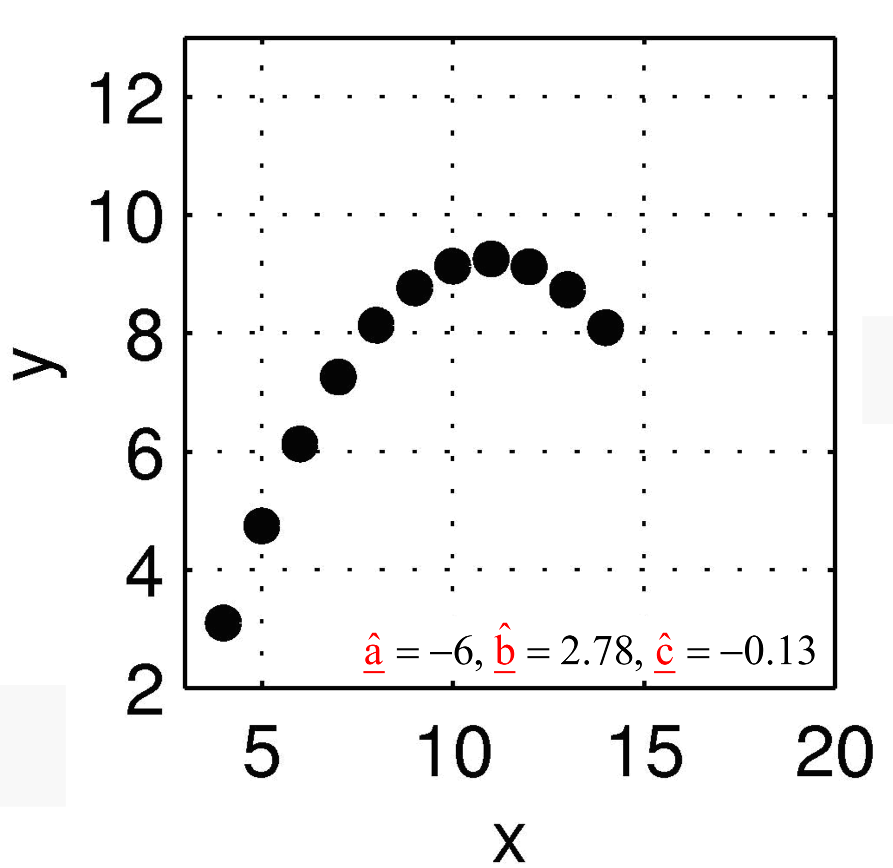
Variances of estimated parameters: correct physical units !

→ Stochastic model (revisited)

Stochastic model (revisited)

Anscombe dataset 2, quadratic regression model
(http://www.gis.uni-stuttgart.de/lehre/campus-docs/AnscombeQuartet.txt)

$$\underline{y}_i = a + bx_i + cx_i^2 + \underline{e}_i \;,\; E\{\underline{e}\} = 0 \;,\; \Sigma_y = D\{\underline{y}\} = Q_y = P^{-1} = I_m \quad (\sigma^2 = 1)$$



$$\underline{y} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \underline{e} \quad , \quad \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 4 \\ \vdots \\ 14 \end{bmatrix}$$

$$\underline{y} = A\xi + \underline{e} \quad , \quad \xi := \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$N = A'PA = \begin{bmatrix} 11 & 99 & 1001 \\ 99 & 1001 & 10989 \\ 1001 & 10989 & 127589 \end{bmatrix}$$

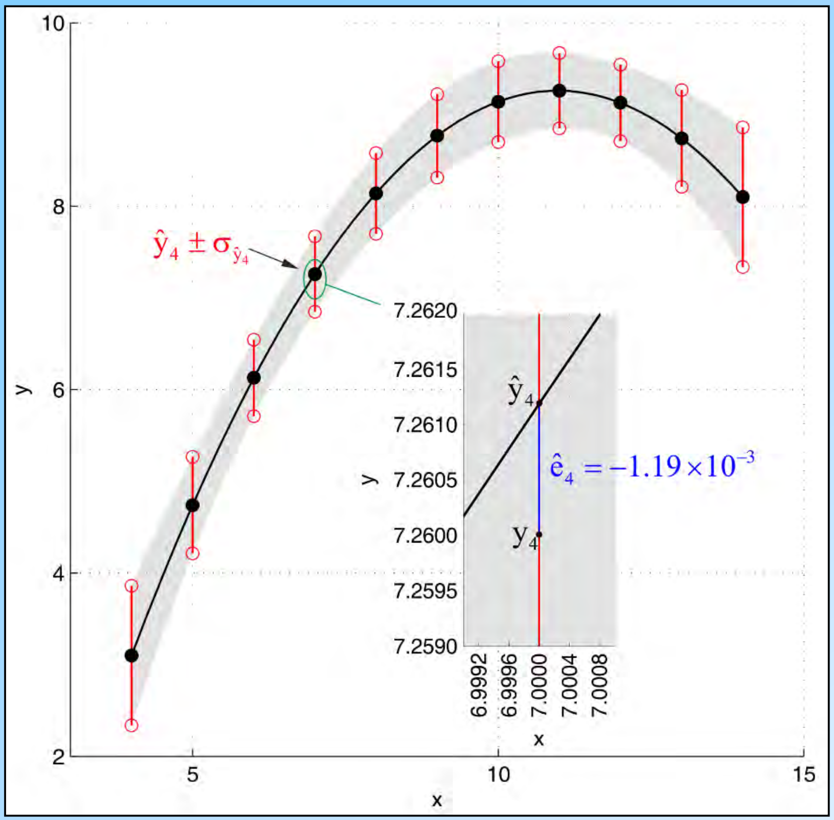
NE

How are the standard deviations of $\underline{\hat{a}}, \underline{\hat{b}}, \underline{\hat{c}}$?

→ Stochastic model (revisited)

Stochastic model (revisited)

$$\underline{\hat{e}} = 10^{-3}\left[-0.21 \;\; -0.63 \;\; 2.38 \;\; -1.19 \;\; -1.33 \;\; 1.96 \;\; -1.33 \;\; -1.19 \;\; 2.38 \;\; -0.63 \;\; -0.21\right]'$$



$$\Sigma_{[\hat{a},\hat{b},\hat{c}]} = N^{-1} \approx \begin{bmatrix} 6.70 & -1.57 & 0.08 \\ -1.57 & 0.39 & -0.02 \\ 0.08 & -0.02 & 0.001 \end{bmatrix}$$

$$\Downarrow$$

$$\sigma_{\hat{a}} \approx \pm\sqrt{6.70} \approx \pm 2.6$$

$$\sigma_{\hat{b}} \approx \pm\sqrt{0.39} \approx \pm 0.6$$

$$\sigma_{\hat{c}} \approx \pm\sqrt{0.001} \approx \pm 0.03$$

Known σ^2=1:

$$\Sigma_{\hat{y}} = A\Sigma_{[\hat{a},\hat{b},\hat{c}]}A'$$

$$\Downarrow$$

$$\sigma_{\hat{y}} \approx \pm[0.76, 0.53, 0.42, 0.41, 0.44, 0.46, \ldots$$

$$\ldots, 0.44, 0.41, 0.42, 0.53, 0.76]'$$

NE

Is this result reasonable ? Data and model seem to perfectly fit to each other ! Why parameters should be that unsafe ? Why should adjusted observations should have standard deviations like this ?

→ Stochastic model (revisited)

Stochastic model (revisited)

Unknown σ^2:

$$\underline{y}_i = a + bx_i + cx_i^2 + \underline{e}_i \, , \, E\{\underline{e}\} = 0 \, , \, \Sigma_y = D\{\underline{y}\} = \sigma^2 Q_y = \sigma^2 P^{-1} = \sigma^2 I_m$$

σ^2 ... unknown scale factor - variance component

$Q_y = I_m$ ... given matrix of cofactors

$$\Sigma_{[\hat{a},\hat{b},\hat{c}]} = \sigma^2 Q_{[\hat{a},\hat{b},\hat{c}]} = \sigma^2 N^{-1} \approx \sigma^2 \begin{bmatrix} 6.70 & -1.57 & 0.08 \\ -1.57 & 0.39 & -0.02 \\ 0.08 & -0.02 & 0.001 \end{bmatrix}$$

$$\underline{\hat{\sigma}}^2 := \frac{\underline{\hat{e}}' P \underline{\hat{e}}}{m-n} = 0.000\,002\,8$$ unbiased estimate of σ^2, i.e. $E\{\underline{\hat{\sigma}}^2\} = \sigma^2$

$$\hat{\Sigma}_{[\hat{a},\hat{b},\hat{c}]} = \underline{\hat{\sigma}}^2 N^{-1} \quad \Rightarrow \quad \hat{\sigma}_{\hat{a}} \approx \pm 0.004 \, , \, \hat{\sigma}_{\hat{b}} \approx \pm 0.001 \, , \, \hat{\sigma}_{\hat{c}} \approx \pm 0.000\,06$$

NE

$$\hat{\Sigma}_{\hat{y}} = A \hat{\Sigma}_{[\hat{a},\hat{b},\hat{c}]} A' \Rightarrow \hat{\sigma}_{\hat{y}} \approx \pm 10^{-3} \times [1.3, 0.9, 0.7, 0.7, 0.7, 0.8, 0.7, 0.7, 0.7, 0.9, 1.3]'$$

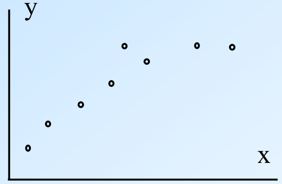
→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

The aim of hypothesis testing is to check input data for mistakes and (gross) errors and/or to inspect the significance of output from the adjustment.

Example 1: The three angles in a triangle have been measured ⇒ sum of (imperfect) measurements will not close to π (inconsistencies, measurement errors). Misclosure can tell something about the mathematical model (is the triangle really planar ?), or may indicate the existence of gross errors. In addition, a wrong stochastic model will lead to less precise least-squares estimates.

Example 2: Hypothesis testing enables to test the significance of certain parameters in the functional model: Are all of the estimated coefficients in a regression model significant, or can the model be simplified ? For the data distribution as shown in the graph, can we assume a linear relationship or would it be better to set up a parabolic model ? In other words: Is the coefficient c in $y=a+bx+cx^2$ significantly different from zero ?



→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

Example 3: Are there any significant between-epoch movements/deformations of a construction as can be detected from repeated measurements (deformation analysis) ?

Example 4: Does the additive constant of an EDM equipment comply with the manufacturer specification or is it necessary to re-calibrate the tool ?

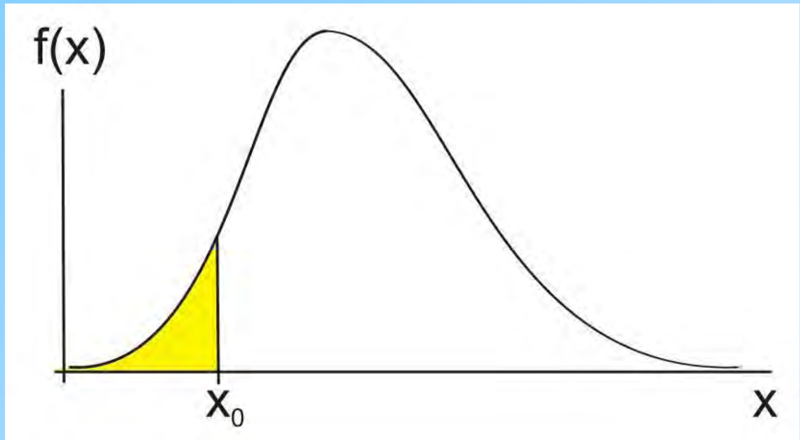
Example 5: Does a favorable/disadvantageous location of a shop affect the pricing ?

- From now on: Exclusive use of random variables
- Hypothesis testing without stochastic model not possible
- Following the central limit theorem it is assumed that observations are normally distributed.

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

Let $\underline{x}$ be a stochastic variable – an observable quantity –, x its corresponding realization – a sample – and f(x) its probability density function (pdf).



Then, the probability that $\underline{x}$ can take a value smaller or equal to x_0 is

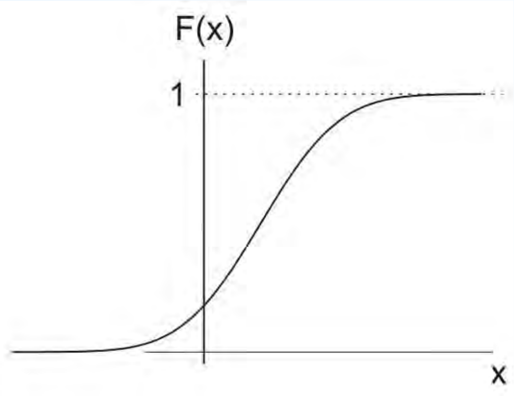
$$P(\underline{x} \le x_0) = \int_{-\infty}^{x_0} f(x)dx \,,$$

i.e. the (yellow) area under f(x) from the leftmost tail to x_0.

Clearly, $P(\underline{x} \le \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$, i.e. the probability that x can take any value between $-\infty$ and $+\infty$ is just one. It is the sure event.

Cumulative distribution function (cdf)

$$P(\underline{x} \le x) = F(x) = \int_{-\infty}^{x} f(x)dx \,,$$

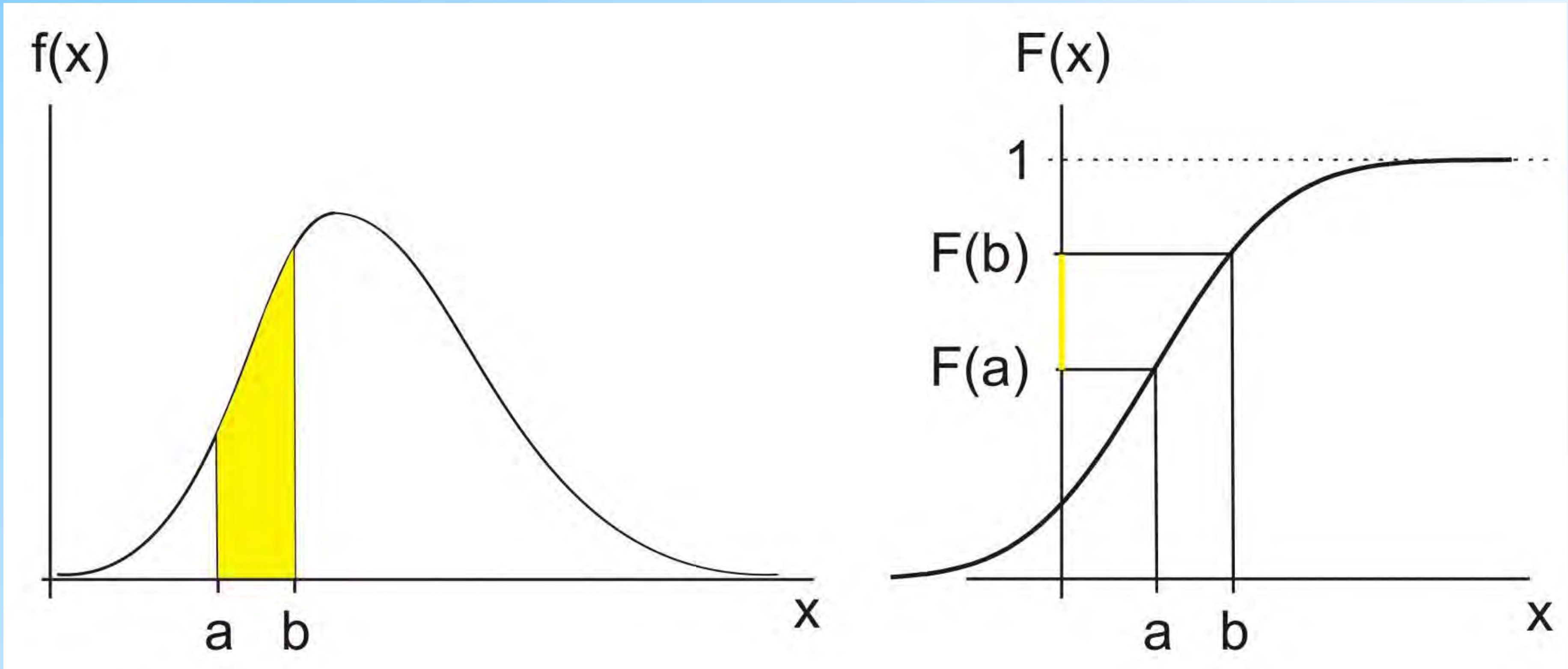


→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

The probability that $\underline{x}$ can take a value from the interval a,b is

$$P(a \le \underline{x} \le b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx = F(b) - F(a)$$



→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

A hypothesis H is the assertion that a random variable $\underline{x}$ has a specific pdf f(x) with specific parameters which describe the pdf. It is written in the form "H: $\underline{x} \sim f(x)$". The type of probability density function is not under investigation or even questionable because it is unique for the chosen random variable $\underline{x}$, and $\underline{x}$ cannot have different variants of probability density functions. Only the descriptive parameters may change.

Example: In the linear model $\underline{y} = Ax + \underline{e}$ or $E\{\underline{y}\} = Ax$ we start from the assumption that $\underline{y} \sim N(E\{\underline{y}\} = Ax, D\{\underline{y}\} = \Sigma_y)$, i.e. $\underline{y}$ is normally distributed with mean value Ax (or $\underline{e}$ has mean zero: $E\{\underline{e}\}=0$) and variance-covariance matrix Σ_y. Then we know that

$$\underline{\hat{e}} = \underline{y} - \underline{\hat{y}} = [I - A(A'Q_y^{-1}A)^{-1}A'Q_y^{-1}]\underline{y} = P_A^{\perp}\underline{y}$$

$$\Sigma_{\hat{e}} = \Sigma_y - \Sigma_{\hat{y}} = P_A^{\perp}\Sigma_y$$

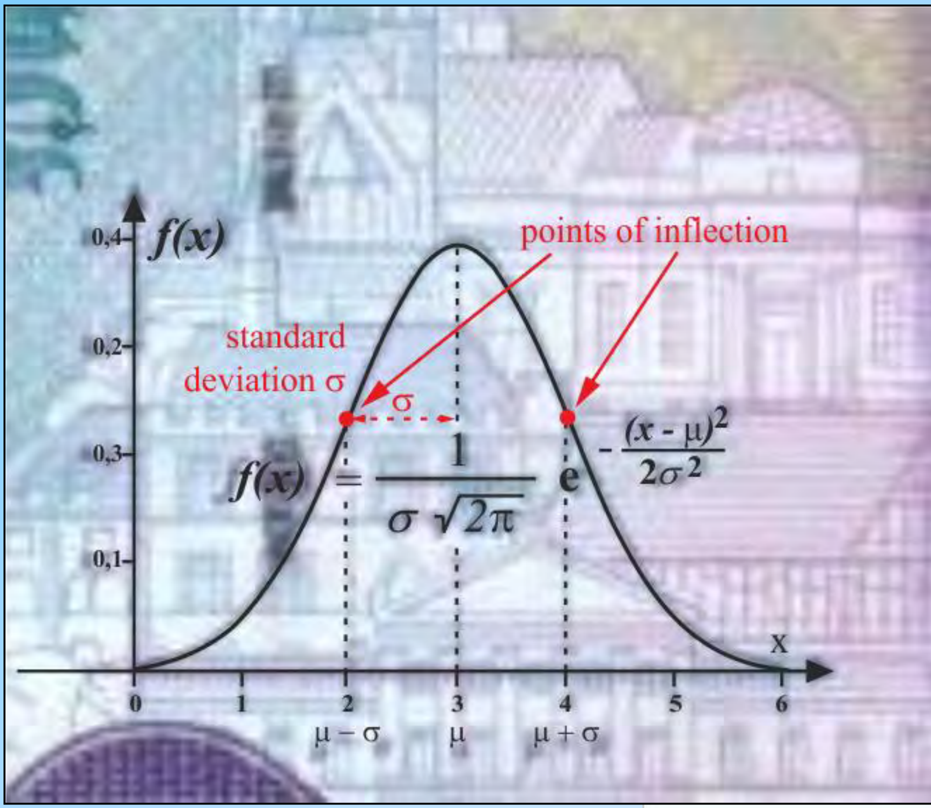
and we can derive $E\{\underline{\hat{e}}\} = 0$, $D\{\underline{\hat{e}}\} = \Sigma_{\hat{e}}$.

The hypothesis is now "H: $\underline{\hat{e}} \sim N(E\{\underline{\hat{e}}\}, D\{\underline{\hat{e}}\}) = N(0, \Sigma_{\hat{e}})$",

i.e. the estimate of e, $\underline{\hat{e}}$, obeys a normal distribution (Gauß pdf) with mean value 0 and covariance matrix $\Sigma_{\hat{e}}$. The decision on whether this is true or false is made by hypothesis testing.

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing



Gauß-pdf for a scalar random variable $\underline{x}$

$\mu := E\{\underline{x}\}$ "mean value"

$\sigma^2 = D\{\underline{x}\}$ "variance"

$\sigma = \pm\sqrt{\sigma^2}$ "standard deviation"

Church Sankt-Jacobi / Göttingen

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

Gauß memorial stone on
"Wilseder Berg" (169m ASL)

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

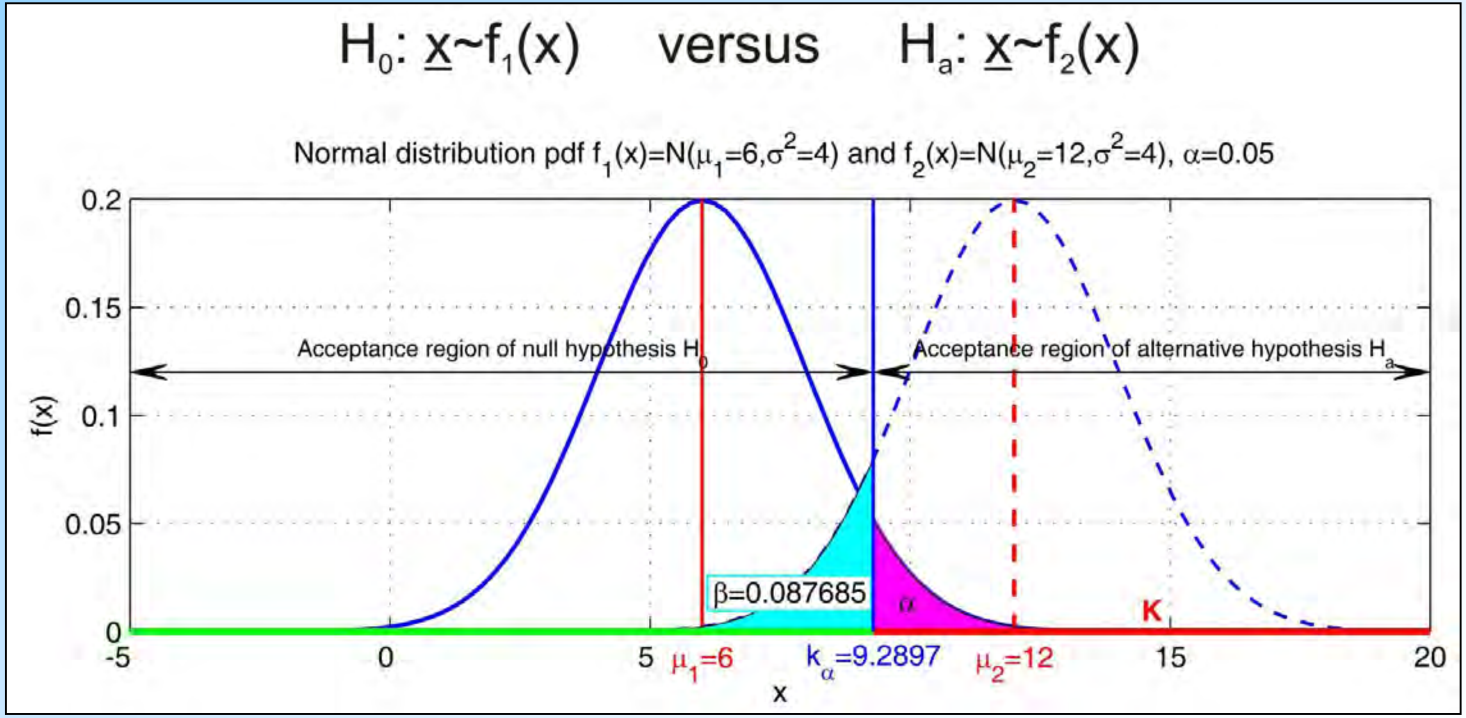
Example:
Assume that we want to know the length $\underline{x}$ of a big (6m long) table. For this reason, we measure its length with an instrument m times (→ sample y). We assume that the unavoidable measurement errors (inconsistencies) are random variables being normally distributed with mean value $E\{\underline{e}\}=0$ and variance-covariance matrix $D\{\underline{e}\}=\Sigma_e$. Then, as we introduce the linear observational model $\underline{y}=Ax+\underline{e}$ we can expect (!) that $\mu_1 \equiv E\{\underline{y}\}=Ax=6m$ (rules for the expectation operator), i.e. the mean value $\mu_1 \equiv E\{\underline{y}\}$ of the observations is 6m. The observations will be distributed around their mean $\mu_1 \equiv E\{\underline{y}\}=6m$, meaning that most of them will be close to and only a few will be far away from the mean μ_1.
In the example below the standard deviation of $\underline{e}$ (and therefore of $\underline{y}$) is assumed to be $\sigma \pm 2m$.
Unfortunately, we have used a damaged instrument with a bias of 6m. Then the assumption $\mu_1 \equiv E\{\underline{y}\}=Ax=6m$ is certainly wrong and we will have the situation $\mu_2 \equiv E\{\underline{y}\}=Ax=12m$, instead.
Now, the corrupted observations will be distributed around their mean $\mu_2 \equiv E\{\underline{y}\}=12m$. Most of the measurements will be close to and only a few far away from the mean μ_2.

Here, the mean $E\{\underline{y}\}$ is under investigation using a hypothesis test "on the mean". The density function itself is not questionable.

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

Hypothesis testing means to test whether the hypothesis H: $\underline{x} \sim f(x)$ – with f(x) carrying certain characteristic parameters with itself – is true or false. The test is performed using a (measured or given) sample value x. If the test proves H_0: $\underline{x} \sim f_1(x)$ to be false then H_a: $\underline{x} \sim f_2(x)$ must hold, instead, because $\underline{x}$ must be distributed either with $f_1(x)$ or with $f_2(x)$. One of both **must** be true ! H_0 is called the "null hypothesis", H_a carries the name "alternative hypothesis".

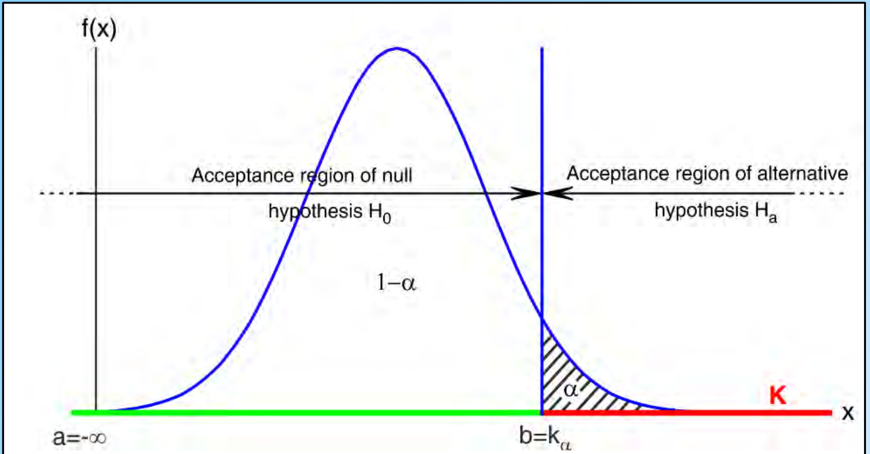


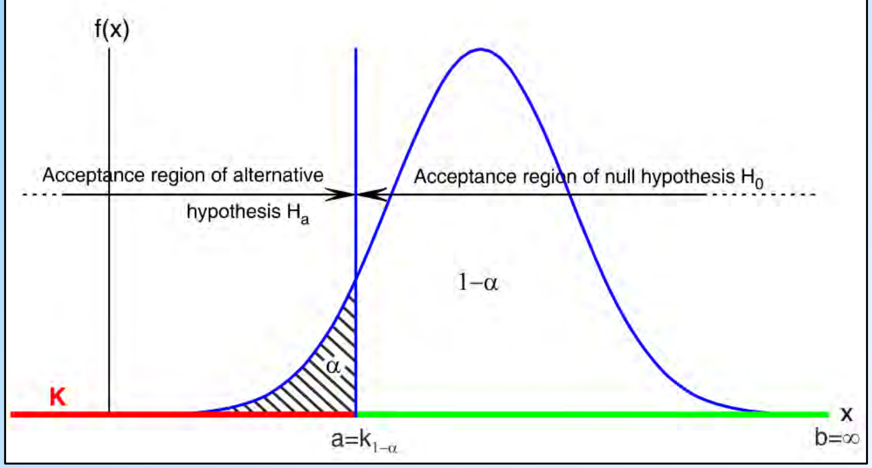
→ Probability theory and hypothesis testing

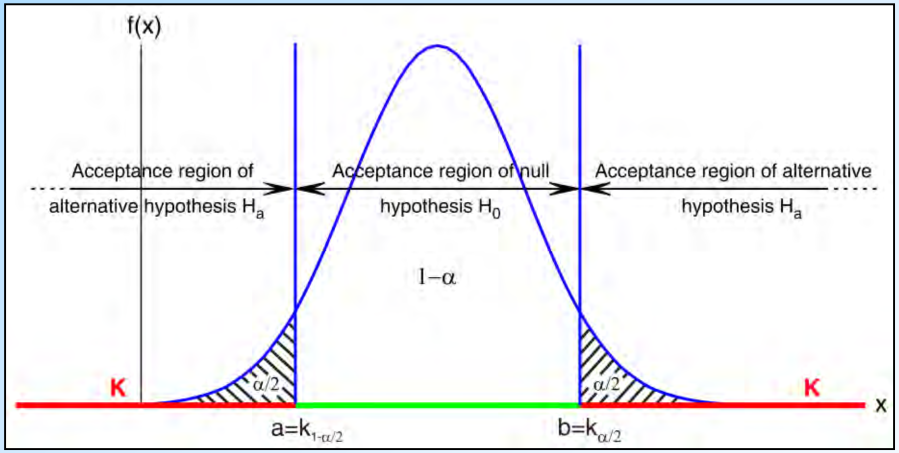
Probability theory and hypothesis testing

How to decide on the composite hypothesis H_0: $\underline{x} \sim f_1(x)$ versus H_a: $\underline{x} \sim f_2(x)$?

If – for a given α – the sample value x falls into the range [a,b] (a and b follow from specifying α) then it is said that H_0: $\underline{x} \sim f_1(x)$ is accepted with probability $1-\alpha$,







otherwise rejected. But since $\underline{x}$ **must have** some pdf, then in case of rejection of H_0, H_a: $\underline{x} \sim f_2(x)$ must hold, instead. Depending on the type of question one has to distinguish between one-sided and two-sided hypotheses, see also Appendix B of the lecture notes.

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

The numbers a,b in $P(a < \underline{x} \le b)$ or $P(a \le \underline{x} < b = \infty)$ or $P(a = -\infty < \underline{x} \le b)$ are the left and right boundaries of the acceptance or confidence region while the remaining part of the interval [a,b] is called the critical or rejection region K. The values a,b are also called critical values k. Depending on a one-sided or two-sided test k gets different indices, see figure.

Abbreviations:

$$P(a < \underline{x} \le b) = 1 - \alpha$$

$P(\underline{x} \notin [a, b]) = \alpha$ "size of the test","level of significance","error probability"

Example using the standard normal pdf $f(x) = 1/(\sigma\sqrt{2\pi})\exp[-(x-\mu)^2/(2\sigma^2)]$

a) Specify $a,b \triangleq k$ in order to find α and $1-\alpha$

$$\begin{aligned}
&k = \pm 1\sigma : P(\mu - k < \underline{x} \le \mu + k) && = 68.3\% \triangleq 1-\alpha && \Rightarrow \alpha = 0.317\\
&k = \pm 2\sigma : && = 95.5\% && \Rightarrow \alpha = 0.045\\
&k = \pm 3\sigma : && = 99.7\% && \Rightarrow \alpha = 0.003
\end{aligned}$$

b) Find k from (frequently used values of) α

$$\alpha = 0.05 \Rightarrow k = \pm 1.96 :\ P(\mu - 1.96 < \underline{x} \le \mu + 1.96) = 95\%$$

$$\alpha = 0.01 \Rightarrow k = \pm 2.58 :\ P(\mu - 2.58 < \underline{x} \le \mu + 2.58) = 99\%$$

NE

$\rightarrow$ Probability theory and hypothesis testing

GIS

Probability theory and hypothesis testing

K: critical region		Reality	
		H_0 is true/H_a is false	H_a is true/H_0 is false
Decision	Accept H_0 because $x \notin K$	Correct decision P(Correct decision)=$1-\alpha$	Wrong decision: reject H_a although it is true ("Failure to give alarm") Type-II-error $P(\text{Type-II-error})=P(x \notin K \mid H_a)=\beta$ $P(x \in K \mid H_a)=1-\beta$
	Reject H_0 because $x \in K$	Wrong decision: reject H_0 although it is true ("Wrong alarm") Type-I-error $P(\text{Type-I-error})=P(x \in K \mid H_0)=\alpha$ $P(x \notin K \mid H_0)=1-\alpha$	Correct decision P(Correct decision)=$1-\beta$ ($1-\beta$="Power of the Test")

$P(\ldots \mid \underbrace{H\ldots}) \rightarrow$ indicates "H… is true"

Note: Sometimes $1-\beta$ and β are interchanged in literature

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

Discussion:

α is the probability of a type-I-error and should be small in order to obtain a good protection against it.
β is the probability of a type-II-error and should also be small in order to have a good protection against it.

But: as α gets smaller, β gets larger and vice versa.

$\alpha=0.01$ means that in 1 out of 100 cases H_0 is rejected although it is true. In order to reduce this possibility of a type-I-error the acceptance region is broadened, i.e. the critical value k is shifted to the right. Thus the probability of a sample value to fall into the acceptance region increases.

However, at the same time the probability of a type-II-error increases in equal measure.

Procedures to optimize this crucial situation exist as well as methods to find a good sample size!

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

For the following examples we assume a random variable $\underline{x}$ with known variance $\sigma^2 = D\{\underline{x}\}$ and unknown mean value $\mu = E\{x\}$. So, the test will be on the mean μ given a certain value μ_0. The significance level α is specified from the beginning. Then, from sample values $x_1, x_2, \ldots, x_m$ the sample mean $\underline{\hat{\mu}} = (\underline{x}_1 + x_2 + \ldots + x_m) / m$ is estimated (A-model !!), and the test quantity $\underline{T} = (\underline{\hat{\mu}} - \mu_0)\sqrt{m} / \sigma$ is known to be standard normally distributed: $\underline{T} \sim N(0,1)$

Example 1: One-sided test

$$H_0 : \mu \le \mu_0 \quad \leftrightarrow \quad H_a : \mu = \mu_1 > \mu_0$$

If the test quantity $\underline{T} = |\underline{\hat{\mu}} - \mu_0| \sqrt{m} / \sigma > k_\alpha$, it gives reason to reject H_0 and to say: $\mu = E\{x\} > \mu_0$.

Example 2: One-sided test

$$H_0 : \mu \ge \mu_0 \quad \leftrightarrow \quad H_a : \mu = \mu_1 < \mu_0$$

If the test quantity $\underline{T} = |\underline{\hat{\mu}} - \mu_0| \sqrt{m} / \sigma < k_{1-\alpha}$, it gives reason to reject H_0 and to say: $\mu = E\{x\} < \mu_0$.

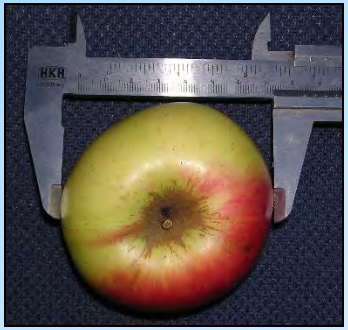
→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

Example 3: Two-sided test $H_0: \mu = \mu_0 \quad \leftrightarrow \quad H_a: \mu = \mu_1 \neq \mu_0$

If the test quantity $\underline{T} = |\underline{\hat{\mu}} - \mu_0|\sqrt{m}/\sigma > k_{\alpha/2}$ or $\underline{T} = |\underline{\hat{\mu}} - \mu_0|\sqrt{m}/\sigma < k_{1-\alpha/2}$ it would not be false to reject H_0 and to say: $\mu = E\{x\} \neq \mu_0$.

Numerical example: A farmer has picked a huge amount of apples from his trees and wants to sell them to a fruit shop. However, the shop owner wants to buy the apples only if they fit into his trays, which require the apple's diameter not to exceed μ_0=10 cm. Because it is impossible to measure the diameters of all



apples, a sample of m=50 apples is taken, the diameter of which is measured with a caliper gauge, and the instrument's precision is assumed to be σ=±0.01cm. The mean diameter as being computed from the least-squares adjustment using the A-model results to $\underline{\hat{\mu}} = 10.085$ cm and the question is ...

NE

→ Probability theory and hypothesis testing

Probability theory and hypothesis testing

… whether or not the estimated deviation $\underline{\hat{\mu}} - \mu_0 = 0.085 \text{ cm}$ is significant. A one-sided hypothesis test with α=0.05

$$H_0 : \mu \le \mu_0 = 10 \text{ cm} \quad \leftrightarrow \quad H_a : \mu = \mu_1 > \mu_0 = 10 \text{ cm}$$

is applied in order to give an answer on the basis of the standard normally distributed test quantity $\underline{T} = |\underline{\hat{\mu}} - \mu_0|\sqrt{m}/\sigma$.

For the reason that $\underline{T} = |10{,}085 - 10|\sqrt{50}/0.01 = 60.1 > k_\alpha = 1.645 = \text{norminv}(1-\alpha)$ the null hypothesis is rejected. The sample originates from a population with a mean μ significantly larger than μ_0=10 cm. (⇒ The fruit shop owner will not buy the apples from the farmer, because they do not satisfy his specification).

It is important to note that here σ was a quantity given in advance. It was not estimated from the sample data before ! In case that σ is not given and has to be replaced by an estimate, $\hat{\sigma}^2$, $\underline{T}$ is not anymore standard normally distributed.

NE

→ Test distributions

Test distributions

As we deal with normally distributed N(E{•},D{•}) random variables and also compute functions of them, we need to know distributions of the functions.

1) Central χ^2-distribution

Let the m×1 random vector $\underline{X} = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_m]'$ be standard normally distributed according to $\underline{X} \sim N(\underset{m\times 1}{0}, I_m)$, then the I_m-weighted sum of squares $\underline{y} = \underline{x}_1^2 + \underline{x}_2^2 + \ldots + \underline{x}_m^2 = \sum_{i=1}^{m} \underline{x}_i^2 = \underline{X}'\underline{X}$ is said to have the χ^2-distribution with m degrees of freedom, $\underline{y} = \underline{X}'\underline{X} \sim \chi_m^2 = \chi^2(m,0)$. $E\{\underline{y}\} = m, D\{\underline{y}\} = 2m$.

2) Central χ^2-distribution

Let the m×1 random vector $\underline{X} = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_m]'$ be non-standard normally distributed according to $\underline{X} \sim N(\underset{m\times 1}{0}, \underset{m\times m}{\Sigma})$, then the Σ^{-1}-weighted sum of squares $\underline{y} = \underline{X}'\Sigma^{-1}\underline{X}$ is said to have the χ^2-distribution with m degrees of freedom, $\underline{y} = \underline{X}'\Sigma^{-1}\underline{X} \sim \chi_m^2 = \chi^2(m,0)$. $E\{\underline{y}\} = m, D\{\underline{y}\} = 2m$.

→ Test distributions

Test distributions

Example (Gauß-Markoff model, A-model):

If $\underline{\hat{e}} = \underline{y} - \underline{\hat{y}} = \underline{y} - A\underline{\hat{x}} = [I - A(A'Q_y^{-1}A)^{-1}A'Q_y^{-1}]\underline{y}$ is the vector of estimated residuals/inconsistencies within a problem with m observations and n unknowns, then $\underline{z} := \underline{\hat{e}}'\Sigma_y^{-1}\underline{\hat{e}} \sim \chi^2_{m-n} = \chi^2(m-n, 0)$.

Remark 1: According to the definition, the rule would be to use $\underline{z} := \underline{\hat{e}}'\Sigma_{\hat{e}}^{-1}\underline{\hat{e}}$. However, the inverse of $\Sigma_{\hat{e}}$ does not exist, because it is rank deficient, and Σ_y is being used, instead.

Remark 2: The random variable $\underline{\hat{e}}'\Sigma_y^{-1}\underline{\hat{e}}$ can also be expressed as $(m-n)\hat{\sigma}^2 / \sigma^2$.

Proof: (1) $\Sigma_y = \sigma^2 P^{-1} \Rightarrow \Sigma_y^{-1} = \sigma^{-2} P \Rightarrow P = \sigma^2 \Sigma_y^{-1}$

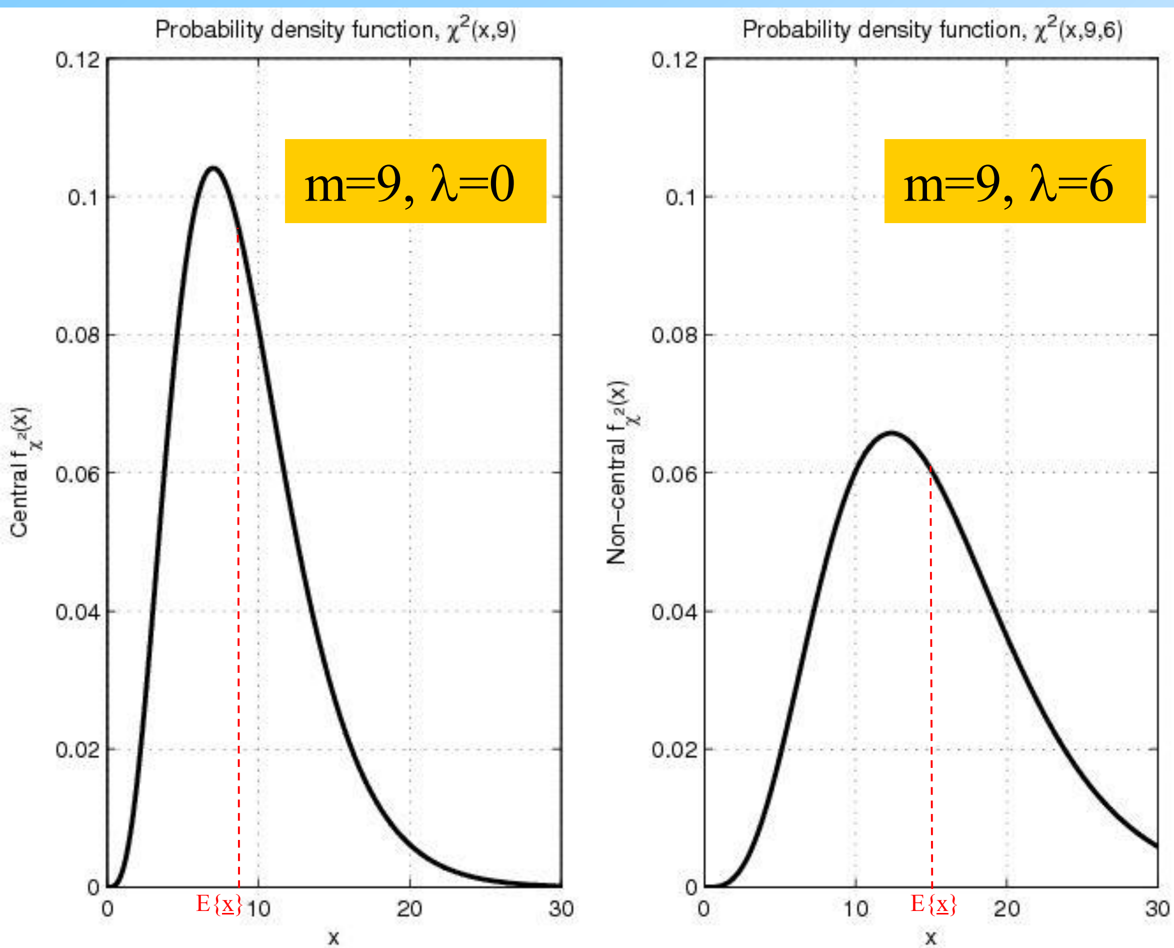
(2) $\underline{\hat{\sigma}}^2 = \dfrac{\underline{\hat{e}}'P\underline{\hat{e}}}{m-n} = \dfrac{\sigma^2\underline{\hat{e}}'\Sigma_y^{-1}\underline{\hat{e}}}{m-n} \Rightarrow \underline{\hat{e}}'\Sigma_y^{-1}\underline{\hat{e}} = (m-n)\underline{\hat{\sigma}}^2 / \sigma^2$.

→ Test distributions

Test distributions

3) Non-central χ^2-distribution

Let the m×1 random vector $\underline{Y} = [\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_m]'$ be non-standard normally



distributed according to

$\underline{Y} \sim N(\underset{m\times 1}{\mu}, \underset{m\times m}{\Sigma})$, then the sum

of squares $\underline{x} = \underline{Y}'\Sigma^{-1}\underline{Y}$ is said

to have the non-central

χ^2-distribution with m

degrees of freedom,

$$\underline{x} = \underline{Y}'\Sigma^{-1}\underline{Y} \sim \chi^2_{m,\lambda} = \chi^2(m,\lambda),$$

and non-centrality parameter

$$\lambda = \mu'\Sigma^{-1}\mu\,.$$

$$E\{\underline{x}\} = m + \lambda$$

→ Test distributions

Test distributions

4) t-distribution

Let the random variables $\underline{x}$ and $\underline{y}$ be independently distributed according to $\underline{x} \sim N(0,1)$ and $\underline{y} \sim \chi^2_m$. Then the random variable $\underline{z} = \underline{x} / \sqrt{\underline{y} / m}$ is said to have the t-distribution with m degrees of freedom, $\underline{z} = \underline{x} / \sqrt{\underline{y} / m} \sim t_m$ ("Student's distribution").

Example (Gauß-Markoff model, A-model):

If $\hat{\underline{x}}$ is the estimate of a certain unknown scalar parameter x in a problem with m observations and n unknowns, and $\hat{\underline{\sigma}}^2_{\hat{x}} = \hat{\underline{\sigma}}^2 q_{\hat{x}\hat{x}}$ its estimated variance ($\hat{\underline{\sigma}}^2$ is the estimate of the unknown variance factor σ^2 and $q_{\hat{x}\hat{x}}$ the diagonal element of $Q_{\hat{x}} = N^{-1} = (A'PA)^{-1}$ belonging to the particular $\hat{\underline{x}}$), then the random variable $\underline{z} := | \hat{\underline{x}} - c | / \hat{\underline{\sigma}}_{\hat{x}} \sim t_{m-n}$. c is a constant.

→ Test distributions

Test distributions

4) Central F-distribution (Fisher-distribution)

Let the random variables $\underline{x}_1$ and $\underline{x}_2$ be independently distributed according to

$\underline{x}_1 \sim \chi^2(m_1)$, $\underline{x}_2 \sim \chi^2(m_2)$. Then the random variable $\underline{y} = (\underline{x}_1 / m_1)/(\underline{x}_2 / m_2)$

is said to have the F-distribution with m_1 and m_2 degrees of freedom,

$$\underline{y} \sim F(m_1, m_2),\ E\{\underline{y}\} = \frac{m_2}{m_2 - 2},\ D\{\underline{y}\} = \frac{2(m_1 + m_2 - 2)m_2^2}{m_1(m_2 - 2)^2(m_2 - 4)},\ m_2 > 2.$$

Example 1 (Gauß-Markoff model, A-model):

The random variable $\underline{z} := (\hat{\underline{x}} - c)^2 / \hat{\underline{\sigma}}_{\hat{x}}^2 \sim F_{1,m-n}$. c is a constant.

Example 2:

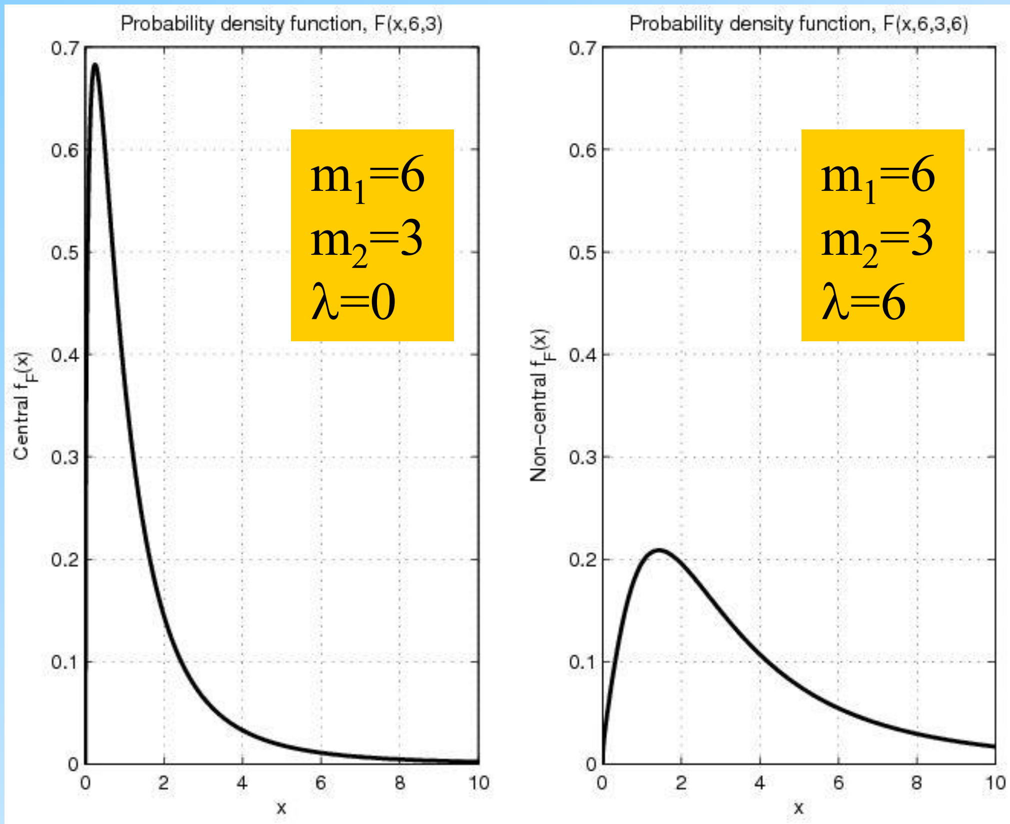
The random variable $\dfrac{\hat{\underline{\sigma}}^2}{\sigma^2} = \dfrac{\hat{\underline{e}}' P \hat{\underline{e}}}{\sigma^2(m-n)} = \dfrac{\hat{\underline{e}}' Q_y^{-1} \hat{\underline{e}}}{\sigma^2(m-n)} = \dfrac{\hat{\underline{e}}' \Sigma_y^{-1} \hat{\underline{e}}}{m-n} \sim F_{m-n,\infty} = \dfrac{\chi^2_{m-n}}{m-n}.$

→ Test distributions

Test distributions

5) Non-central F-distribution

Let the random variables $\underline{x}_1$ and $\underline{x}_2$ be independently distributed according to $\underline{x}_1 \sim \chi^2(m_1, \lambda)$, $\underline{x}_2 \sim \chi^2(m_2)$. Then the random variable $\underline{y} = (\underline{x}_1 / m_1)/(\underline{x}_2 / m_2)$



is said to have the non-central F-distribution with m_1, m_2 degrees of freedom, and non-centrality parameter λ, $\underline{y} \sim F(m_1, m_2, \lambda)$.

Test distributions

Connections between distributions

a) χ^2-distribution $\quad \chi^2_{1-\alpha;r} = rF_{1-\alpha;r,\infty}$

b) Standard normal distribtion $\quad z_{1-\alpha/2} = \sqrt{F_{1-\alpha;1,\infty}} = \sqrt{\chi^2_{1-\alpha;1}}$

c) t-distribution $\quad t_{1-\alpha/2;r} = \sqrt{F_{1-\alpha;1,r}}$

MATLAB

a) $\text{chi2inv}(1-\text{alpha},r) - r * \text{finv}(1-\text{alpha},r,10^6) \approx 0$

b) $\text{norminv}(1-\text{alpha}/2) - \text{sqrt}(\text{finv}(1-\text{alpha},1,10^6)) \approx 0$

c) $\text{tinv}(1-\text{alpha}/2,r) - \text{sqrt}(\text{finv}(1-\text{alpha},1,r)) \approx 0$

This is the end !