GIS-E3010

Least-Squares Methods in Geoscience

Lecture 8/2018

Tech.Lic - M.Sc. - Eng. Fabián Barbato

Reliability

Test statistics

Quality control

Quality control

- Evaluation of the design of measuring procedure
  - Measures for precision, reliability, robustness
  - No actual measurements needed, only design matrix A and covariance matrix of observations Σ are necessary
  - Planning
- Assessment of the measurements and their influence onto the results
  - Statistical tests, outlier detection, estimated size of blunders, influence of observations to the parameters
  - Understanding the errors and influence of them onto the results

Results: adjusted parameters, adjusted observations and residuals

We have a linear model:

$E(y_{nx1}) = A_{nxu}x_{ux1}$ with $\Sigma_y = \sigma_0^2 Q_{nxn}$

$n < u,$

$rank(A) = u$

Covariance matrix of observations

Model

Least squares solution:

We can choose $P = Q^{-1}$ and σ_0^2=1

$\hat{x} = (A^T PA)^{-1}A^T Py = (A^T Q^{-1}A)^{-1}A^T Q^{-1}y = (A^T \Sigma^{-1}A)^{-1}A^T \Sigma^{-1}y$

$\hat{x} = (A^T PA)^{-1}A^T Py$

$\hat{y} = A\hat{x}$

LS-solution for parameters

Adjusted observations

Residuals: $v = A\hat{x} - y = \hat{y} - y$

(Adjusted minus observed)

Residuals

With residuals we obtain an estimate for variance factor:

$$\hat{\sigma}_0^2 = \frac{v^T Pv}{r}$$

Estimated variance factor

$r = n - u$ (in overconstraint case)

Redundancy

Quality control in planning: Precision of estimates

$$\Sigma_{\hat{x}} = \sigma_0^2 N^{-1} = \sigma_0^2 Q_x$$

$$\Sigma_{\hat{\ell}} = A\Sigma_{\hat{x}}A^T = \sigma_0^2 A Q_{\hat{x}} A^T$$

$$\Sigma_v = \Sigma_\ell - \Sigma_{\hat{\ell}} = \sigma_0^2 (\mathrm{Q}_\ell - Q_{\hat{\ell}})$$

→ Error ellipsoids

→ Relative error ellipsoids

→ Standard deviations

These matrices (or square roots of diagonal elements) are used in evaluating the influence of the random errors in observations onto the results.
It is variance propagation in LS-process

This is the **evaluation of the measuring process**, not actual measurements are needed

Influence of blunders onto the results

$Q = Q_\ell$

$\nabla_i x = Q_{\hat{x}} A^T Q^{-1} \nabla_i \ell$ External reliability, influence onto parameters

$\nabla_i \hat{l} = Q_{\hat{\ell}} Q^{-1} \nabla_i \ell = (A Q_{\hat{x}} A^T Q^{-1}) \nabla_i \ell = U \nabla_i \ell$ Contribution onto the adjusted observations

$\nabla_i v = -Q_v Q^{-1} \nabla_i \ell = (I - A Q_{\hat{x}} A^T Q^{-1}) \nabla_i \ell = -R \nabla_i \ell$ Influence onto the residuals

In planning (evaluation of the procedure):

$\nabla_i \ell$ is internal reliability measure $\nabla_0 \ell_i$

- Depends on chosen α and β

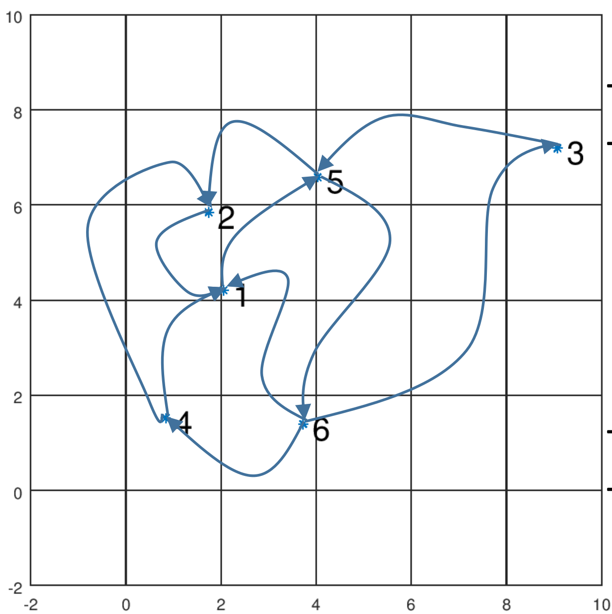
Aim: r_i large and u_i small

With actual measurements (studying the errors)

$\nabla_i \ell$ is estimated error $h_i \dfrac{v_i}{r_i}$

$$h_i = (0 \quad 0 \quad \cdots \quad 1 \quad 0 \quad \cdots \quad 0 \quad 0)^T$$

Example



A-matrix						From	To	dh (y-vector)	q_{ii}	σ_ℓ
0	0	0	0	-1	1	5.00000	6.00000	-0.37457	5.19911	0.0114
-1	0	0	0	0	1	1.00000	6.00000	-1.22432	3.27177	0.0090
-1	0	0	0	1	0	1.00000	5.00000	-0.84309	3.09186	0.0088
1	-1	0	0	0	0	2.00000	1.00000	0.54735	1.67002	0.0065
1	-1	0	0	0	0	2.00000	1.00000	0.53780	1.67002	0.0065
0	0	1	0	0	-1	6.00000	3.00000	-0.94636	7.88979	0.0140
0	0	0	0	-1	1	5.00000	6.00000	-0.36362	5.19911	0.0114
-1	0	0	0	1	0	1.00000	5.00000	-0.86040	3.09186	0.0088
-1	0	0	1	0	0	1.00000	4.00000	0.24709	2.95202	0.0086
0	0	-1	0	1	0	3.00000	5.00000	1.30930	5.08226	0.0113
1	0	0	-1	0	0	4.00000	1.00000	-0.25354	2.95202	0.0086
1	0	0	0	0	-1	6.00000	1.00000	1.24252	3.27177	0.0090
0	1	0	0	-1	0	5.00000	2.00000	0.33209	2.40547	0.0078
0	1	0	-1	0	0	4.00000	2.00000	-0.78142	4.42337	0.0105
0	0	0	1	0	-1	6.00000	4.00000	1.53000	2.89037	0.0085

Apriori $\sigma_0 =$ 0.005

Example: evaluation of procedure before measurements

- We know the precision of the levelling: $0.005\mathrm{m}/\sqrt{km}$
- We know the lengths of the levelling routes between points
- → We can calculate the standard deviations of observations

- We know the network design → A-matrix
- → We can evaluate the precision of the heights $\Sigma_{\hat{x}}$
- →We can evaluate the influence of possible blunders onto the results
  - Redundancy matrix and numbers R and r_i
  - Contribution matrix and numbers U and u_i
  - Lower bound for detectable error (internal reliability)
  - Theoretical sensitivity factor → influence on parameters
  - Theoretical maximum influence of undetectable error

Noncentrality
$\nabla_0 w_i = \delta_0$ **it depends on the chosen risk levels** (α, β)**.**
δ_0 **with** $\alpha = 0.001, \beta = 0.8$
delta=norminv(0.9995)+norminv(0.8)

Lower bound for detectable error (internal reliability)

$$\nabla_0 \ell_i = \delta_0 \frac{\sigma_{\ell_i}}{\sqrt{r_i}}$$

Controllability factor

$$\delta'_{0i} = \frac{\nabla_0 \ell_i}{\sigma_{\ell_i}}$$

Theoretical sensitivity factor

$$\bar{\bar{\delta}}_{0i} = \delta_0 \sqrt{\frac{u_i}{r_1}}$$

Theoretical maximum influence of undetectable error

$$\nabla_i z \leq \bar{\bar{\delta}}_{0i} \sigma_{\hat{z}}$$

In example

from	to	$\hat{\ell}$	v	v_{std}	v_{std}	$\nabla\ell_i$	r_i	u_i	$\sigma_{\hat{\ell}}$	σ_v	$\nabla_0\ell_i$	$\bar{\bar{\delta}}_i$	$\bar{\bar{\delta}}_{0i}$
5	6						0.79	0.21	0.0091	0.0178	0.044		1.75
1	6						0.74	0.26	0.0081	0.0136	0.036		2.04
1	5						0.74	0.26	0.0078	0.0133	0.035		2.01
2	1						0.64	0.36	0.0068	0.0091	0.028		2.57
2	1						0.64	0.36	0.0068	0.0091	0.028		2.57
6	3						0.56	0.44	0.0164	0.0184	0.064		3.04
5	6						0.79	0.21	0.0091	0.0178	0.044		1.75
1	5						0.74	0.26	0.0078	0.0133	0.035		2.01
1	4						0.70	0.30	0.0083	0.0126	0.035		2.26
3	5						0.36	0.64	0.0158	0.0119	0.064		4.57
4	1						0.70	0.30	0.0083	0.0126	0.035		2.26
6	1						0.74	0.26	0.0081	0.0136	0.036		2.04
5	2						0.58	0.42	0.0088	0.0104	0.035		2.91
4	2						0.72	0.28	0.0097	0.0157	0.042		2.12
6	4						0.57	0.43	0.0098	0.0113	0.039		2.97

Example: evaluation of the actual measurements, understanding the errors

- Solution: heights
- A posterior variance factor
- Residuals
- Standardized residuals
- Estimated errors
- Empirical sensitivity factor
- Influence onto the results

Adjusted parameters

no	H	σ_H	from plan
1	0.7602	0.0050	0.0028679
2	0.2228	0.0064	0.0036669
3	-1.4184	0.0133	0.0075620
4	1.0199	0.0074	0.0041857
5	-0.1030	0.0060	0.0033996
6	-0.4815	0.0063	0.0036025

Standardized residual (test statistic)

$$\delta_i = v_{i_{std}} = \frac{v_i}{\sigma_{v_i}}$$

Estimated error

$$\nabla \ell_i = \frac{v_i}{r} = v_{i_{std}} \frac{\sigma_{\ell_i}}{\sqrt{r_i}}$$

Standard deviation of estimated error

$$\sigma_{\nabla \ell_i} = \frac{\sigma_{\ell_i}}{\sqrt{r_i}}$$

Standardized size of estimated error

$$\delta_i' = \frac{\nabla \ell_i}{\sigma_{\ell_i}} = \frac{v_{i_{std}}}{\sqrt{r_i}}$$

Empirical sensitivity factor, standardized influence in the adjusted observation

$$\bar{\bar{\delta}}_i = v_{i_{std}} \sqrt{\frac{u_i}{r_1}} = \delta_i' \sqrt{u_i}$$

Actual influence of observation on the function derived from solution is not more than

$$\nabla_i z \leq \bar{\bar{\delta}}_i \sigma_{\hat{z}}$$
$$\sigma_{\hat{z}} = g^T \Sigma_{\hat{x}} g$$

Measures related to observation in design or in actual measurements

from	to	$\hat{\ell}$	v	v_{std}	v_{std}	$\nabla\ell_i$	r_i	u_i	$\sigma_{\hat{\ell}}$	σ_v	$\nabla_0\ell_i$	$\bar{\bar{\delta}}_i$	$\bar{\bar{\delta}}_{0i}$
5	6	-0.3785	-0.0039	-0.22	-0.39	-0.0050	0.79	0.21	0.0091	0.0178	0.044	-0.11	1.75
1	6	-1.2417	-0.0174	-1.28	-2.24	-0.0236	0.74	0.26	0.0081	0.0136	0.036	-0.76	2.04
1	5	-0.8632	-0.0201	-1.51	-2.66	-0.0271	0.74	0.26	0.0078	0.0133	0.035	-0.89	2.01
2	1	0.5374	-0.0099	-1.09	-1.92	-0.0155	0.64	0.36	0.0068	0.0091	0.028	-0.82	2.57
2	1	0.5374	-0.0004	-0.04	-0.07	-0.0006	0.64	0.36	0.0068	0.0091	0.028	-0.03	2.57
6	3	-0.9369	0.0095	0.51	0.90	0.0170	0.56	0.44	0.0164	0.0184	0.064	0.46	3.04
5	6	-0.3785	-0.0149	-0.83	-1.47	-0.0188	0.79	0.21	0.0091	0.0178	0.044	-0.43	1.75
1	5	-0.8632	-0.0028	-0.21	-0.37	-0.0038	0.74	0.26	0.0078	0.0133	0.035	-0.13	2.01
1	4	0.2597	0.0126	1.00	1.76	0.0181	0.70	0.30	0.0083	0.0126	0.035	0.66	2.26
3	5	1.3154	0.0061	0.51	0.90	0.0170	0.36	0.64	0.0158	0.0119	0.064	0.69	4.57
4	1	-0.2597	-0.0061	-0.49	-0.86	-0.0088	0.70	0.30	0.0083	0.0126	0.035	-0.32	2.26
6	1	1.2417	-0.0008	-0.06	-0.10	-0.0011	0.74	0.26	0.0081	0.0136	0.036	-0.03	2.04
5	2	0.3258	-0.0063	-0.61	-1.07	-0.0109	0.58	0.42	0.0088	0.0104	0.035	-0.52	2.91
4	2	-0.7971	-0.0157	-1.00	-1.76	-0.0217	0.72	0.28	0.0097	0.0157	0.042	-0.62	2.12
6	4	1.5014	-0.0286	-2.54	-4.46	-0.0502	0.57	0.43	0.0098	0.0113	0.039	-2.21	2.97

The influence of the observation (6- 4) onto the parameters

$$\nabla_i x = (A^T P A)^{-1} A^T P \nabla_i \ell$$

$$\nabla_{0i} x = (A^T P A)^{-1} A^T P \nabla_{0i} \ell$$

$$\nabla_{15}\ell = (0 \quad 0 \quad \cdots \quad 0 \quad 0 \quad \cdots \quad 0 \quad \nabla\ell_{15})^T = (0 \quad 0 \quad \cdots \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0.0502)^T$$

$$[(A^T P A)^{-1} A^T P]_{6x15} \cdot \nabla_{15}\ell = \begin{matrix} -0.0012717 \\ -0.0019616 \\ 0.0046664 \\ -0.0123920 \\ 0.0017232 \\ 0.0092357 \end{matrix}$$

Actual influence on parameters

Empirical sensitivity factor

$$\bar{\bar{\delta}}_i \cdot 0.0074 = -2.21 \cdot 0.0074 = -0.0163$$

Limit for influence

Theoretical sensitivity factor

Decision: remove the observation!

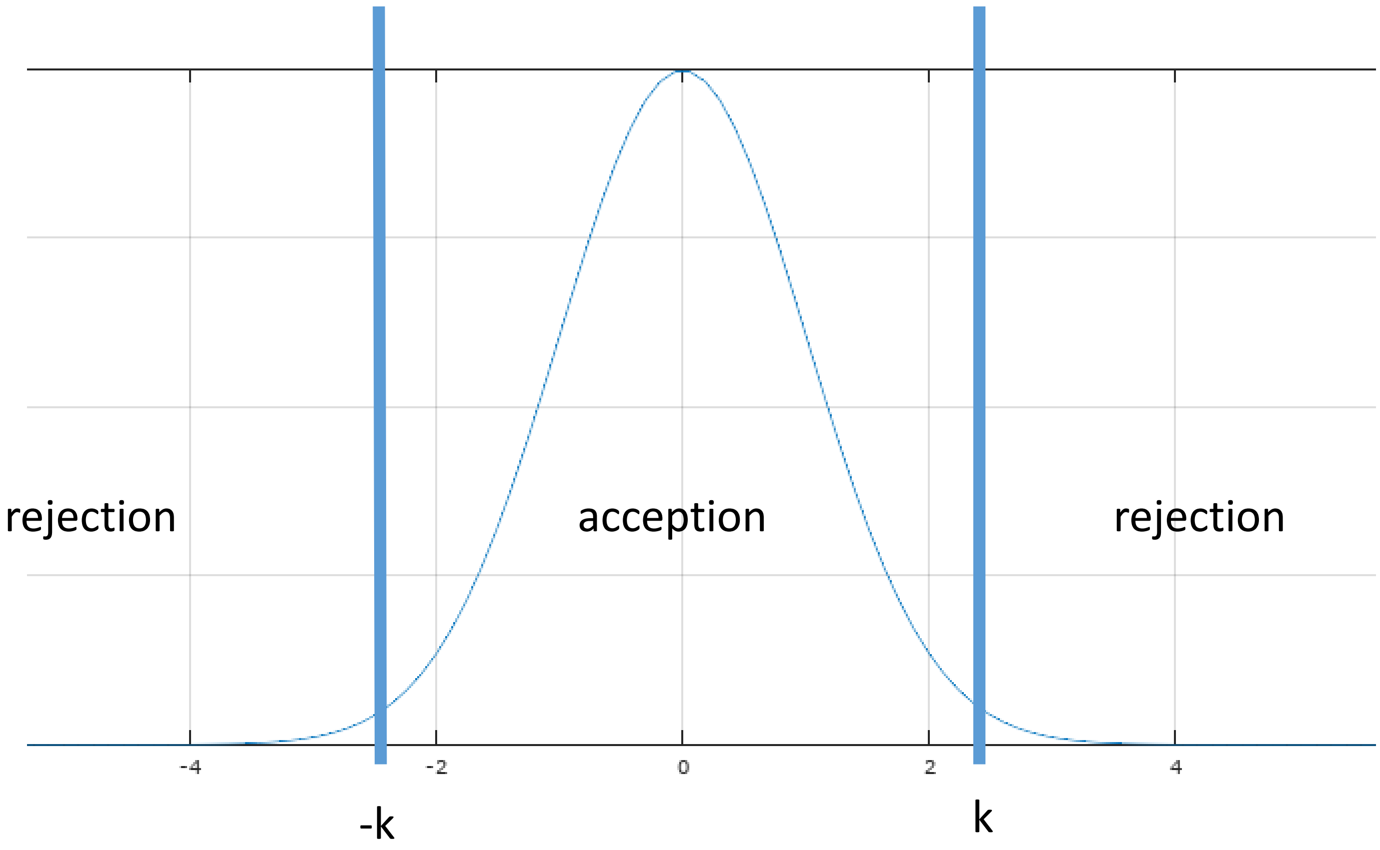
$$\bar{\bar{\delta}}_{0i} \sigma_{\hat{z}} = 2.97 \cdot 0.0074 = 0.021978$$

Global test

- **Residuals** in next slide
- Estimated variance factor: $\hat{\sigma}_0^2 = \frac{v^T P v}{n-u} = 0.00878\ ^2$
- **Global test:**
  - Test statistics: $\frac{0.00878^2}{0.005^2} = 3.0848$
  - Critical value: $\chi^2_{0.01,r} = 2.558$
  - $\boldsymbol{H_0}$ **is rejected**
  - => outliers or optimistic apriori variances
  - **An alternative hypothesis**: There is an outlier in i:th observation

Testing

- Comparing the test statistic w_i with the critical value k
  - Acceptance area [-k,k]
  - Is the test statistic in acceptance or in rejection area?



Test statistics

In these alternative hypotheses only the expectation is assumed to be influenced, not the variance

Blunder (gross error) in i:s observation:

$$\nabla_i \ell = (0 \quad 0 \quad \cdots \quad \nabla \ell_i \quad 0 \quad \cdots \quad 0 \quad 0)^T$$

Null hypothesis:

$$H_0: E(y|H_0) = Ax$$

Alternative hypotheses:

$$H_{ai}: E(y|H_{ai}) = E(y|H_0) + \nabla_i \ell = Ax + h_i \nabla \ell_i$$

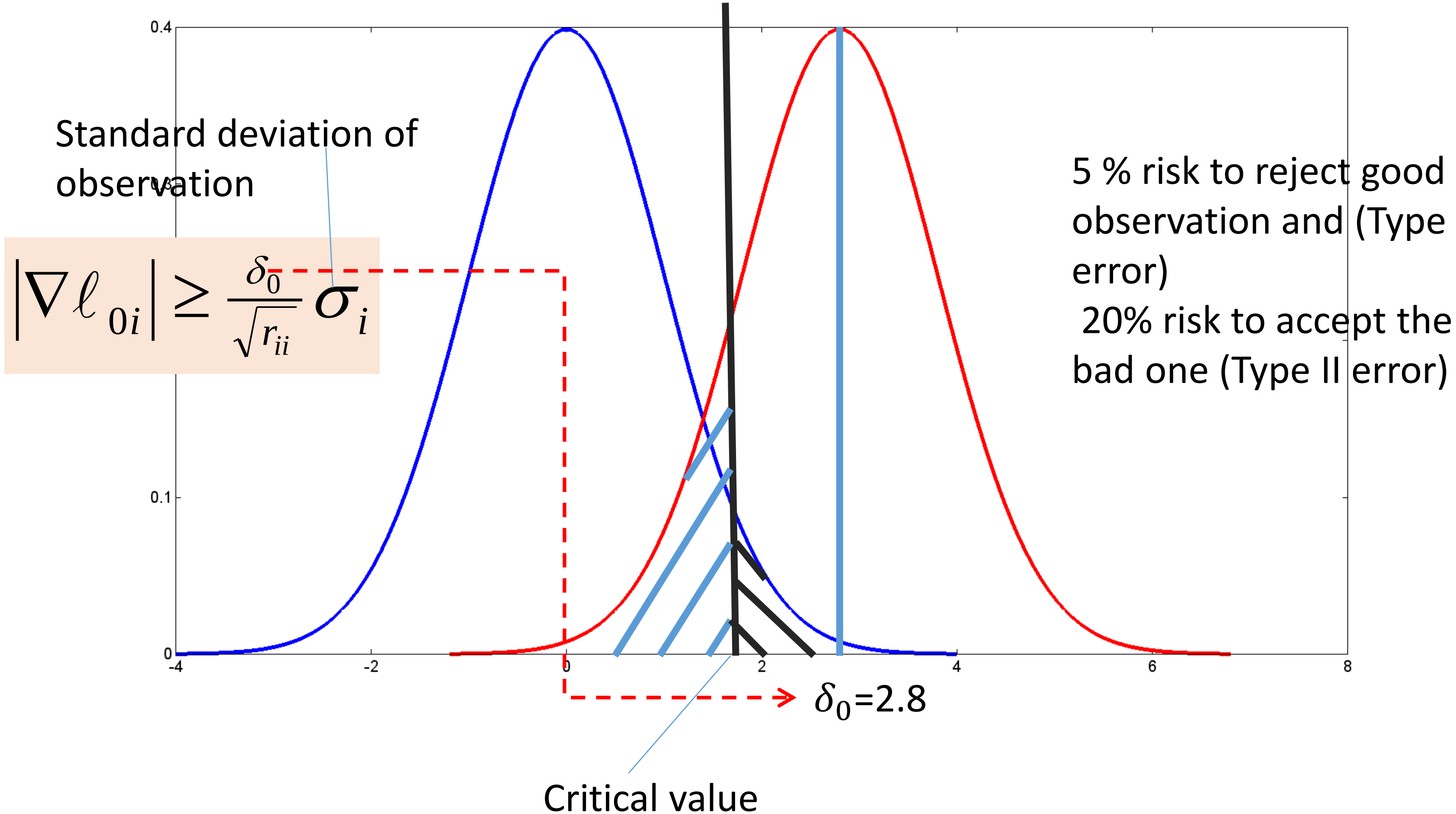
$$h_i = (0 \quad 0 \quad \cdots \quad 1 \quad 0 \quad \cdots \quad 0 \quad 0)^T$$

Test statistics for correlated observations

$$w_i = \frac{-h_i^T \Sigma_\ell^{-1} v}{\sqrt{h_i^T \Sigma_\ell^{-1} \Sigma_v \Sigma_\ell^{-1} h_i}}$$

Test statistics for uncorrelated observations

$$w_i = \frac{-v_i}{\sigma_{v_i}} = \frac{\nabla \ell_i}{\sigma_{\nabla \ell_i}} = \frac{-v_i}{\sigma_{\ell_i}\sqrt{r_i}}$$



- **Noncentrality parameter**
  - $\nabla_0 w_i = \delta_0$ it depends on the chosen risk levels (α, β).
  - $\delta_0 = 3.4175$ with $\alpha = 0.01, \beta = 0.8$

Influence onto residuals

$\nabla v_i = -R\nabla_i \ell$ An error in one observation influences all residuals and one residual is influenced by all observational errors

$\nabla_i v_j = -r_{ji}\nabla \ell_i$ Influence of i:th observation to the j:th residual

$\nabla_i v_i = -r_i \nabla \ell_i$ Influence of i:th observation to the i:th residual

r_i is a diagonal element of R