DIAGNOSTICS OF ROBUST IDENTIFICATION OF MODEL

ROBUST 2014

LEDEN 19 - 24, 2014
HOTEL BELLEVUE, JETŘICHOVICE

RESEARCH WAS SUPPORTED BY GRANT OF THE CZECH SCIENCE FOUNDATION PROJECT NO. 13-01930S

"ROBUST METHODS FOR NONSTANDARD SITUATIONS, THEIR DIAGNOSTICS AND IMPLEMENTATIONS".

JAN ÁMOS VÍŠEK

FSV UK, PRAHA

Content

Notations

The most frequent econometric (statistical) framework is:

Regression model

$$Y_i = X_i'\beta^0 + \varepsilon_i = \sum_{j=1}^{p} X_{ij}\beta_j^0 + \varepsilon_i, \quad i = 1, 2, ..., n$$

Notations

The most frequent econometric (statistical) framework is:

Regression model

$$Y_i = X_i'\beta^0 + \varepsilon_i = \sum_{j=1}^{p} X_{ij}\beta_j^0 + \varepsilon_i, \quad i = 1, 2, ..., n$$

$$Y = X\beta^0 + \varepsilon$$

Crucial task

Identification of regression model

$$\hat{\beta}^{(n)}(Y, X) \rightarrow R^p,$$

Crucial task

Identification of regression model

$$\hat{\beta}^{(n)}(Y, X) \rightarrow R^p,$$

but we need also $\hat{\sigma}^2_{(n)}(Y, X) \rightarrow R^+$

$\rightarrow$ to be able to establish significance of explanatory variables,

Crucial task

Identification of regression model

$$\hat{\beta}^{(n)}(Y, X) \rightarrow R^p,$$

but we need also $\hat{\sigma}^2_{(n)}(Y, X) \rightarrow R^+$

$\rightarrow$ to be able to establish significance of explanatory variables,

usually we have $\hat{\sigma}^2_{(n)}\left(Y, X, \hat{\beta}^{(n)}\right) \rightarrow R^+$.

Content

Returning to the roots

First of all, let's recall the classical estimators

1. the (Ordinary) Least Squares $\hat{\beta}^{(OLS,n)}$

and

Returning to the roots

First of all, let's recall the classical estimators

1. the (Ordinary) Least Squares $\hat{\beta}^{(OLS,n)}$

   and

2. the Maximum Likelihood $\hat{\beta}^{(ML,n)}$

Returning to the roots

How have we developed the estimation of regression model ?
What key steps do we teach in introductory courses ?

Returning to the roots

How have we developed the estimation of regression model ?
What key steps do we teach in introductory courses ?

1. Deriving $\hat{\beta}^{(OLS,n)}$, $\hat{\beta}^{(ML,n)}$ and $\hat{\sigma}^2_{(OLS,n)}$,

Returning to the roots

How have we developed the estimation of regression model ?
What key steps do we teach in introductory courses ?

1. Deriving $\hat{\beta}^{(OLS,n)}$, $\hat{\beta}^{(ML,n)}$ and $\hat{\sigma}^2_{(OLS,n)}$,
2. establishing the significance of $\hat{\beta}_\ell^{(OLS,n)}$,

Returning to the roots

How have we developed the estimation of regression model ?
What key steps do we teach in introductory courses ?

1. Deriving $\hat{\beta}^{(OLS,n)}$, $\hat{\beta}^{(ML,n)}$ and $\hat{\sigma}^2_{(OLS,n)}$,
2. establishing the significance of $\hat{\beta}_\ell^{(OLS,n)}$,
3. evaluating the determination of model,

Returning to the roots

How have we developed the estimation of regression model ?
What key steps do we teach in introductory courses ?

1. Deriving $\hat{\beta}^{(OLS,n)}$, $\hat{\beta}^{(ML,n)}$ and $\hat{\sigma}^2_{(OLS,n)}$,
2. establishing the significance of $\hat{\beta}_\ell^{(OLS,n)}$,
3. evaluating the determination of model,
4. stressing the importance of normality of disturbances,

Returning to the roots

How have we developed the estimation of regression model ?
What key steps do we teach in introductory courses ?

1. Deriving $\hat{\beta}^{(OLS,n)}$, $\hat{\beta}^{(ML,n)}$ and $\hat{\sigma}^2_{(OLS,n)}$,
2. establishing the significance of $\hat{\beta}_\ell^{(OLS,n)}$,
3. evaluating the determination of model,
4. stressing the importance of normality of disturbances,
5. testing the submodels, etc.

Returning to the roots

How have we built up accompanynig tools ?

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),
2. deriving an indication of collinearity (index number, Farrar-Glauber),

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),
2. deriving an indication of collinearity (index number, Farrar-Glauber),
3. establishing the correlation of disturbances (Durbin-Watson),

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),
2. deriving an indication of collinearity (index number, Farrar-Glauber),
3. establishing the correlation of disturbances (Durbin-Watson),
4. testing the homoscedasticity (White, Breuch-Pagan),

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),
2. deriving an indication of collinearity (index number, Farrar-Glauber),
3. establishing the correlation of disturbances (Durbin-Watson),
4. testing the homoscedasticity (White, Breuch-Pagan),
5. verifying the specification of model (Hausman),

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),
2. deriving an indication of collinearity (index number, Farrar-Glauber),
3. establishing the correlation of disturbances (Durbin-Watson),
4. testing the homoscedasticity (White, Breuch-Pagan),
5. verifying the specification of model (Hausman),
6. studying the stability of model (Chow),

Returning to the roots

How have we built up accompanynig tools ?

1. Testing normality of disturbances (Jarque-Bera, normal plot),
2. deriving an indication of collinearity (index number, Farrar-Glauber),
3. establishing the correlation of disturbances (Durbin-Watson),
4. testing the homoscedasticity (White, Breuch-Pagan),
5. verifying the specification of model (Hausman),
6. studying the stability of model (Chow),
7. testing character of effects (Hausman-Taylor), etc.

A comparison of classical and robust regression analysis

Have we done an analogy in robust regression ?

A comparison of classical and robust regression analysis

Have we done an analogy in robust regression ?

NO !

A comparison of classical and robust regression analysis

Have we done an analogy in robust regression ?

NO !

We are in a permanent pursuit for new and new estimators, new and new principles, new and new point of view.

A comparison of classical and robust regression analysis

Have we done an analogy in robust regression ?

NO !

We are in a permanent pursuit for new and new estimators, new and new principles, new and new point of view.

What about to select one flexible robust estimator and for it to develop the same diagnostics, we are used to for $\hat{\beta}^{(OLS,n)}$ or $\hat{\beta}^{(ML,n)}$?

A comparison of classical and robust regression analysis

Have we done an analogy in robust regression ?

NO !

We are in a permanent pursuit for new and new estimators, new and new principles, new and new point of view.

What about to select one flexible robust estimator and for it to develop the same diagnostics, we are used to for $\hat{\beta}^{(OLS,n)}$ or $\hat{\beta}^{(ML,n)}$?

But which one?

Returning to the roots

Let’s recall the clasical estimators once again.

What features of estimators we are (“automatically”) used to ?

Equivariance of $\hat{\beta}^{(n)}$

$$\hat{\beta}(Y, X) : M(n, p+1) \to R^p$$

$$\textit{scale-equivariant} : \quad \forall c \in R^+ \quad \hat{\beta}(cY, X) = c\hat{\beta}(Y, X)$$

$$\textit{regression-equivariant} : \quad \forall b \in R^p \quad \hat{\beta}(Y + Xb, X) = \hat{\beta}(Y, X) + b$$

Equivariance of $\hat{\beta}^{(n)}$

$$\hat{\beta}(Y, X) : M(n, p+1) \rightarrow R^p$$

scale-equivariant : $\forall c \in R^+ \quad \hat{\beta}(cY, X) = c\hat{\beta}(Y, X)$

regression-equivariant : $\forall b \in R^p \quad \hat{\beta}(Y + Xb, X) = \hat{\beta}(Y, X) + b$

Examples : $\hat{\beta}^{(OLS,n)} = (X'X)^{-1} X'Y$, $\quad \hat{\beta}^{(L_1,n)} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^n |Y_i - X_i'\beta|$,

e. g.

we don't need recalculate estimate when we change (linearly) the units of data.

Equivariance of $\hat{\beta}^{(n)}$

$$\hat{\beta}(Y, X) : M(n, p+1) \to R^p$$

$$\textit{scale-equivariant} \ : \quad \forall c \in R^+ \quad \hat{\beta}(cY, X) = c\hat{\beta}(Y, X)$$

$$\textit{regression-equivariant} \ : \quad \forall b \in R^p \quad \hat{\beta}(Y + Xb, X) = \hat{\beta}(Y, X) + b$$

Examples : $\hat{\beta}^{(OLS,n)} = (X'X)^{-1} X'Y$, $\quad \hat{\beta}^{(L_1,n)} = \arg\min_{\beta \in R^p} \sum_{i=1}^n |Y_i - X_i'\beta|$,

e. g.

we don't need recalculate estimate when we change (linearly) the units of data.

Unfortunately, the most popular robust estimators, the *M*-estimators

$$\hat{\beta}^{(M,\rho,n)} = \arg\min_{\beta \in R^p} \sum_{i=1}^n \rho\left(Y_i - X_i'\beta\right)$$

don't possess this property.

Equivariance of $\hat{\beta}^{(n)}$

$$\hat{\beta}(Y, X) : M(n, p+1) \to R^p$$

$$\textit{scale-equivariant} \;:\; \forall c \in R^+ \quad \hat{\beta}(cY, X) = c\hat{\beta}(Y, X)$$

$$\textit{regression-equivariant} \;:\; \forall b \in R^p \quad \hat{\beta}(Y + Xb, X) = \hat{\beta}(Y, X) + b$$

Examples : $\hat{\beta}^{(OLS,n)} = (X'X)^{-1} X'Y$, $\quad \hat{\beta}^{(L_1,n)} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^n |Y_i - X_i'\beta|$,

e. g.

we don't need recalculate estimate when we change (linearly) the units of data.

Unfortunately, the most popular robust estimators, the *M*-estimators

$$\hat{\beta}^{(M,\rho,n)} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^n \rho\left(Y_i - X_i'\beta\right)$$

don't possess this property.

They require studentization of residuals!!

Bickel, P. J. (1975): One-step Huber estimates in the linear model. *J. Amer. Statist. Assoc. 70, 428–433.*

To reach *scale-* and *regression-equivariance* of an *M*-estimator, say

$$\hat{\beta}^{(M,\rho,n)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \rho\left(\frac{Y_i - X_i'\beta}{\hat{\sigma}_{(n)}}\right)$$

$\hat{\sigma}_{(n)}$ has to be *scale-equivariant* and *regression-invariant*.

Estimators of scale of disturbances which are sufficiently robust, consistent, scale-equivariant and regression-invariant:

Jurečková, J., P. K. Sen (1993): Regression rank scores scale statistics and studentization in linear models. *Proc. of the Fifth Prague Symposium on Asymptotic Statistics, Physica Verlag, 111-121.*

Croux C., P. J. Rousseeuw (1992): A class of high-breakdown scale estimators based on subranges. *Communications in Statistics - Theory and Methods 21, 1935 - 1951.*

Víšek, J. Á. (2010): Robust error-term-scale estimate. *IMS Collections. Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: Festschrift for Jana Jurečková, Vol. 7(2010), 254 - 267.*

Estimators of scale of disturbances which are sufficiently robust, consistent, scale-equivariant and regression-invariant:

Jurečková, J., P. K. Sen (1993): Regression rank scores scale statistics and studentization in linear models. *Proc. of the Fifth Prague Symposium on Asymptotic Statistics, Physica Verlag, 111-121.*

Croux C., P. J. Rousseeuw (1992):

and Time Series Analysis. Festschrift for Jana Jurečková, Vol. 7(2010), 254 - 267.

The common feature of all these estimators -

they are based on (residuals of) a preliminary estimate of regression model

which is already scale- and regression-equivariant.

Preliminary conclusion

We should prefer (robust) estimators which are *"automatically"* scale- and regression-equivarint.

Preliminary conclusion

We should prefer (robust) estimators
which are *"automatically"* scale- and regression-equivarint.

$$\hat{\beta}^{(LMS,n,h)} = \arg\min_{\beta \in R^p} r^2_{(h)}(\beta)$$

Rousseeuw, P. J. (1984): Least median of square regression.
Journal of Amer. Statist. Association 79, pp. 871-880.

Preliminary conclusion

We should prefer (robust) estimators
which are *"automatically"* scale- and regression-equivarint.

$$\hat{\beta}^{(LMS,n,h)} = \arg\min_{\beta \in R^p} \; r^2_{(h)}(\beta)$$

Rousseeuw, P. J. (1984): Least median of square regression. *Journal of Amer. Statist. Association 79, pp. 871-880.*

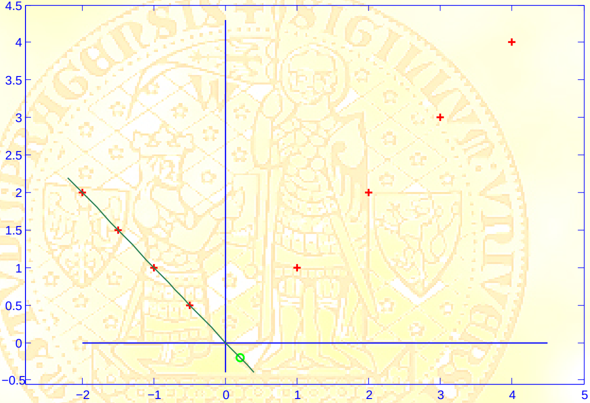
$$\hat{\beta}^{(LTS,n,h)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{h} r^2_{(i)}(\beta)$$

Hampel, F. R. et al. (1986): *Robust Statistics – The Approach Based on Influence Functions.* New York: J.Wiley & Son.

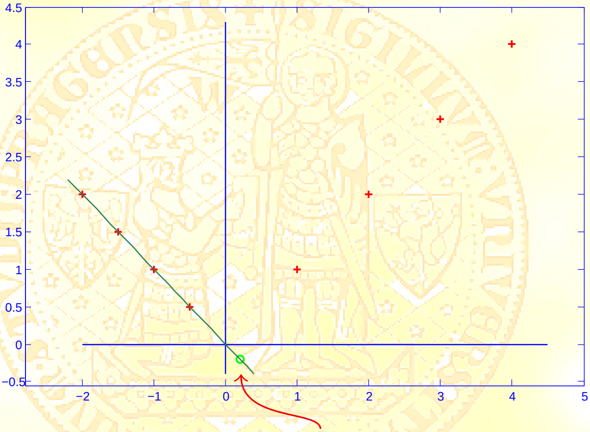
SENSITIVITY OF ANY HIGH-BREAKDOWN-POINT ESTIMATOR
TO A SMALL CHANGE OF DATA

AN (ACADEMIC) EXPLANATION BY A SHIFT OF “INLIER”

Model for the majority of data

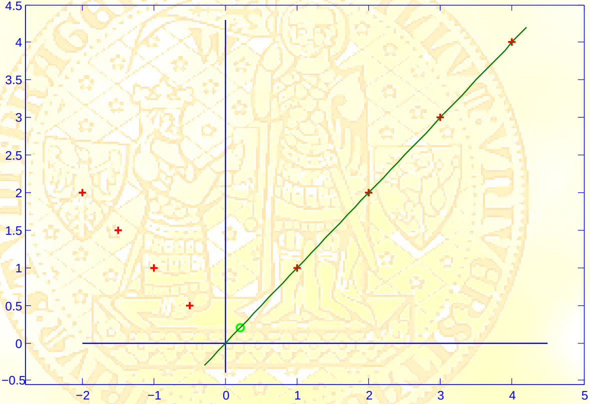


Model for the majority of data

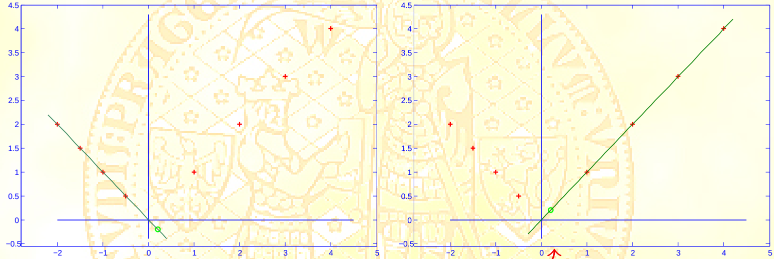


We are going to shift up this green circle " o ".

Again model for the majority of data



In both cases the model is for the majority of data



Notice: *The closer* the green circle (“ o ”) is to the *y*-axe,
the smaller shift causes the “switch” of the model.

IT WAS THE (ACADEMIC) EXAMPLE - THERE ARE REAL DATA

HETTMANSPERGER, T. P., S. J. SHEATHER (1992):
A CAUTIONARY NOTE ON THE METHOD OF LEAST MEDIAN SQUARES.
The American Statistician 46, 79–83.

Engine Knock Data ($n = 16$, $p = 4$, $h = 11$)

c	x_1	x_2	x_3	x_4	y
1	13.3	13.9	31	697	84.4
2	13.3	14.1	30	697	84.1
3	13.4	15.2	32	700	88.4
4	12.7	13.8	31	669	84.2
⋮	⋮	⋮	⋮	⋮	⋮
14	12.7	16.1	35	649	93.0
15	12.9	15.1	36	721	93.3
16	12.7	15.9	37	696	93.1

x_1 is spark timing x_2 air/fuel ratio
x_3 intake temperature x_4 exhaust temperature
y engine knock number

Engine Knock Data $(n = 16, p = 4, h = 11)$

c	x_1	x_2	x_3	x_4	y
1	13.3	13.9	31	697	84.4
2	13.3	14.1	30	697	84.1
3	13.4	15.2	32	700	88.4
4	12.7	13.8	31	669	84.2

This is the exact value of $\hat{\beta}^{(LTS,n,h)}$!

Data	Interc.	SPARK	AIR	INTK	EXHS.
Correct data ($x_{22} = 14.1$)	35.11	-0.028	2.949	0.477	-0.009
Damaged data($x_{22} = 15.1$)	-88.7	4.72	1.06	1.57	0.068

x_3 intake temperature $\quad$ x_4 exhaust temperature

y engine knock number

Content

The least weighted squares

Residuals $\quad \forall \beta \in R \quad \rightarrow \quad r_i(\beta) = Y_i - X_i'\beta$

Order statistics of squared residuals, i. e.

$$r^2_{(1)}(\beta) \le r^2_{(2)}(\beta) \le \dots \le r^2_{(n)}(\beta)$$

Definition

Let $w_i \in [0, 1]$, $\; i = 1, 2, ..., n$. Then

$$\hat{\beta}^{(LWS,n,w)} = \arg\min_{\beta \in R^p} \sum_{i=1}^n w_i r^2_{(i)}(\beta)$$

will be called *the least weighted squares* (*LWS*).

The least weighted squares

Residuals $\forall \beta \in R \quad \rightarrow \quad r_i(\beta) = Y_i - X_i'\beta$

Order statistics of squared residuals, i. e.

$$r^2_{(1)}(\beta) \leq r^2_{(2)}(\beta) \leq ... \leq r^2_{(n)}(\beta)$$

Definition

Let $w_i \in [0, 1]$, $i = 1, 2, ..., n$. Then

$$\hat{\beta}^{(LWS,n,w)} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^{n} w_i r^2_{(i)}(\beta)$$

will be called *the least weighted squares (LWS)*.

The weights are prescribed to the order statistics of squared residuals !!

The least weighted squares

Residuals $\quad \forall \beta \in R \quad \rightarrow \quad r_i(\beta) = Y_i - X_i'\beta$

Order statistics of squared residuals, i.e.

Víšek, J. Á. (2000): Regression with high breakdown point. *Robust 2000 (eds. Antoch, J. Dohnal, G.), 324 - 356.*

Let $w_i \in [0, 1]$, $\; i = 1, 2, ..., n$. Then

$$\hat{\beta}^{(LWS,n,w)} = \underset{\beta \in R^p}{\arg\min} \; \sum_{i=1}^{n} w_i r_{(i)}^2(\beta)$$

will be called <u>*the least weighted squares*</u> *(LWS)*.

The least weighted squares

Residuals $\quad \forall \beta \in R \quad \rightarrow \quad r_i(\beta) = Y_i - X_i'\beta$

Order statistics of squared residuals, i. e.

$$r^2_{(1)}(\beta) \leq r^2_{(2)}(\beta) \leq ... \leq r^2_{(n)}(\beta)$$

Definition

Let $w_i \in [0, 1]$, $\; i = 1, 2, ..., n$. Then

$$\hat{\beta}^{(LWS,n,w)} = \arg\min_{\beta \in R^p} \sum_{i=1}^n w_i r^2_{(i)}(\beta)$$

will be called *the least weighted squares* (*LWS*).

$$w_{(1)} \geq w_{(2)} \geq ... \geq w_{(n)}$$

The least median of squares $\hat{\beta}^{(LMS,h,n)}$ as well as the least trimmed squares $\hat{\beta}^{(LTS,h,n)}$ are special cases of the $\hat{\beta}^{(LWS,n,w)}$.

PROS AND CONS OF LWS

Inherited from LTS:

$\sqrt{n}$-consistency (even under heteroscedasticity)

Scale- and affine-equivariance

Pros and cons of LWS

Achieved due to the continuity of weight function:

- Breakdown point adaptable to level and character of contamination
- Quick and reliable algorithm (implemented in MATLAB)
- Modifications for nonstandard situations
  (e. g. instrumental variables, total least squares, fixed or random effects)
- Low sensitivity to the shift and deletion of observation(s)
- More diagnostic tools (D-W, White test, Hausman test, etc.)
- Applicability for panel data
- "Coping automatically" with heteroscedasticity of data

Pros and cons of LWS

Still lacking:

Diagnostics for finite sample size
- i. e. significance of explanatory variables, test of submodels, etc.

Determination of model

Content

Significance of explanatory variable - for the Least Weighted Squares (LWS)

We are going to give an idea of deriving
the significance of *individual explanatory variable* - two steps:

Significance of explanatory variable - for the Least Weighted Squares (LWS)

We are going to give an idea of deriving
the significance of *individual explanatory variable* - two steps:

The first one:

The *Least Weighted Squares* $\hat{\beta}^{(LWS,n,w)}(Y, X)$ can be
- at any point of a basic probabily space $(\Omega, \mathcal{A}, P)$ -
written as *Ordinary Least Squares* $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X})$.

Significance of explanatory variable - for the Least Weighted Squares (LWS)

We are going to give an idea of deriving the significance of *individual explanatory variable* - two steps:

The first one:

The *Least Weighted Squares* $\hat{\beta}^{(LWS,n,w)}(Y,X)$ can be - at any point of a basic probabily space $(\Omega, \mathcal{A}, P)$ - written as *Ordinary Least Squares* $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})$.

The second one:

The classical derivation for significance of *individual explanatory variable for OLS* $\hat{\beta}^{(OLS,n)}(Y,X)$ can be generalised for $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})$.

Deriving form of $\hat{\beta}^{(LWS,n)}$

$\underline{\forall \ \ (\omega \in \Omega) \ \ \exists \ \ (\pi = \pi(\omega) = \{\pi_1(\omega), \pi_2(\omega), ..., \pi_n(\omega)\})}$ so that

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w_i \left(Y_{\pi_i} - X'_{\pi_i}\beta\right)^2$$

$$= \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \left(w_i^{\frac{1}{2}} Y_{\pi_i} - w_i^{\frac{1}{2}} X'_{\pi_i}\beta\right)^2 = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \left(\tilde{Y}_i - \tilde{X}'_i\beta\right)^2 .$$

Deriving form of $\hat{\beta}^{(LWS,n)}$

$\forall \ (\omega \in \Omega) \ \exists \ (\pi = \pi(\omega) = \{\pi_1(\omega), \pi_2(\omega), ..., \pi_n(\omega)\})$ so that

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w_i \left(Y_{\pi_i} - X'_{\pi_i}\beta\right)^2$$

$$= \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \left(w_i^{\frac{1}{2}} Y_{\pi_i} - w_i^{\frac{1}{2}} X'_{\pi_i}\beta\right)^2 = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \left(\tilde{Y}_i - \tilde{X}'_i\beta\right)^2 .$$

Put $\tilde{W} = diag\left\{w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, ..., w_n^{\frac{1}{2}}\right\}$, $\tilde{Y} = \tilde{W} Y_\pi$, $\tilde{X} = \tilde{W} X_\pi$ and $\tilde{e} = \tilde{W} e_\pi$ and consider the model

$$\tilde{Y} = \tilde{X}\beta^0 + \tilde{\varepsilon} \quad \text{with} \quad \mathcal{L}(\tilde{\varepsilon}) = \mathcal{N}\left(0, \sigma^2 \tilde{W}^2\right).$$

Deriving form of $\hat{\beta}^{(LWS,n)}$

$\underline{\forall \ (\omega \in \Omega) \ \exists \ (\pi = \pi(\omega) = \{\pi_1(\omega), \pi_2(\omega), \dots, \pi_n(\omega)\})}$ so that

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w_i \left(Y_{\pi_i} - X'_{\pi_i}\beta\right)^2$$

$$= \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \left(w_i^{\frac{1}{2}} Y_{\pi_i} - w_i^{\frac{1}{2}} X'_{\pi_i}\beta\right)^2 = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \left(\tilde{Y}_i - \tilde{X}'_i\beta\right)^2 .$$

Put $\tilde{W} = diag\left\{w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, \dots, w_n^{\frac{1}{2}}\right\}$, $\tilde{Y} = \tilde{W} Y_\pi$, $\tilde{X} = \tilde{W} X_\pi$ and $\tilde{e} = \tilde{W} e_\pi$ and consider the model

$$\tilde{Y} = \tilde{X}\beta^0 + \tilde{\varepsilon} \quad \text{with} \quad \mathcal{L}(\tilde{\varepsilon}) = \mathcal{N}\left(0, \sigma^2 \tilde{W}^2\right) .$$

Then

$$\underline{\hat{\beta}^{(LWS,n,w)}(\omega)} = \arg\min_{\beta \in R^p} \left\{\left(\tilde{Y} - \tilde{X}\beta\right)' \left(\tilde{Y} - \tilde{X}\beta\right)\right\}$$

$$= \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y} = \underline{\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X})}.$$

Deriving form of $\hat{\beta}^{(LWS,n,w)}$

Fix $\pi = \{\pi_1, \pi_2, ..., \pi_n\}$ and put

$$B(\pi) = \left\{\omega \in \Omega \ : \ \hat{\beta}^{(LWS,n,w)}(Y, X) = \hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X})\right\}$$

Deriving form of $\hat{\beta}^{(LWS,n,w)}$

Fix $\pi = \{\pi_1, \pi_2, ..., \pi_n\}$ and put

$$B(\pi) = \left\{ \omega \in \Omega \ : \ \hat{\beta}^{(LWS,n,w)}(Y, X) = \hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X}) \right\}$$

Then: $\pi^{(1)} \neq \pi^{(2)} \implies B(\pi^{(1)}) \cap B(\pi^{(2)}) = \emptyset \quad a.s.$

Deriving form of $\hat{\beta}^{(LWS,n,w)}$

Fix $\pi = \{\pi_1, \pi_2, ..., \pi_n\}$ and put

$$B(\pi) = \left\{\omega \in \Omega \; : \; \hat{\beta}^{(LWS,n,w)}(Y,X) = \hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right\}$$

Then: $$\pi^{(1)} \neq \pi^{(2)} \quad \Longrightarrow \quad B(\pi^{(1)}) \cap B(\pi^{(2)}) = \emptyset \quad a.s.$$

$$\bigcup_{over \;\; all \;\; \pi' s} B(\pi) = \Omega$$

Deriving form of $\hat{\beta}^{(LWS,n,w)}$

Fix $\pi = \{\pi_1, \pi_2, ..., \pi_n\}$ and put

$$B(\pi) = \left\{\omega \in \Omega \ : \ \hat{\beta}^{(LWS,n,w)}(Y, X) = \hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X})\right\}$$

Then: $$\pi^{(1)} \neq \pi^{(2)} \implies B(\pi^{(1)}) \cap B(\pi^{(2)}) = \emptyset \quad a.s.$$

$$\bigcup_{over \ all \ \pi's} B(\pi) = \Omega$$

$$P(B(\pi)) = \frac{1}{n!}$$

Deriving form of $\hat{\beta}^{(LWS,n,w)}$

Fix $\pi = \{\pi_1, \pi_2, ..., \pi_n\}$ and put

$$B(\pi) = \left\{\omega \in \Omega \;:\; \hat{\beta}^{(LWS,n,w)}(Y,X) = \hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right\}$$

Then: $$\pi^{(1)} \neq \pi^{(2)} \quad \Longrightarrow \quad B(\pi^{(1)}) \cap B(\pi^{(2)}) = \emptyset \quad a.\,s.$$

$$\bigcup_{over\;\; all\;\; \pi' s} B(\pi) = \Omega$$

$$P(B(\pi)) = \frac{1}{n!}$$

- Conditional p-value

Deriving form of $\hat{\beta}^{(LWS,n,w)}$

Fix $\pi = \{\pi_1, \pi_2, ..., \pi_n\}$ and put

$$B(\pi) = \left\{\omega \in \Omega \; : \; \hat{\beta}^{(LWS,n,w)}(Y,X) = \hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right\}$$

Then: $$\pi^{(1)} \neq \pi^{(2)} \implies B(\pi^{(1)}) \cap B(\pi^{(2)}) = \emptyset \quad a.s.$$

$$\bigcup_{over\ \ all\ \ \pi' s} B(\pi) = \Omega$$

$$P(B(\pi)) = \frac{1}{n!}$$

- Conditional p-value
- Unconditional p-value

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

$$Y_i = X_i'\beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \qquad \text{or} \qquad Y = X\beta^0 + \varepsilon$$

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

$$Y_i = X_i'\beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X\beta^0 + \varepsilon$$

Conditions :

$\{(X_i', \varepsilon_i)'\}_{i=1}^{\infty}$ i.i.d., $F_{X,\varepsilon}(x, v) = F_X(x) \cdot F_\varepsilon(v)$, $F_\varepsilon(v) = \mathcal{N}(0, \sigma^2)$,

$Q = E[X_1 \cdot X_1']$ is regular.

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

$$Y_i = X_i'\beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X\beta^0 + \varepsilon$$

Conditions :

$$\{(X_i', \varepsilon_i)'\}_{i=1}^{\infty} \ \underline{\text{i.i.d.}}, \ F_{X,\varepsilon}(x, v) = F_X(x) \cdot F_\varepsilon(v), \ \underline{F_\varepsilon(v) = \mathcal{N}(0, \sigma^2)},$$
$$Q = E[X_1 \cdot X_1'] \text{ is regular.}$$

Significance of ℓ-th explanatory variable $X_{i\ell}$ $\Leftrightarrow$ $H_0: \ \hat{\beta}_\ell^{(OLS,n)} = 0$

Denote $c_{\ell,\ell}^2 = \left[(X'X)^{-1}\right]_{\ell,\ell}$ and $s_n^2 = \frac{1}{n-p}\sum_{i=1}^n \left(Y_i - X_i'\hat{\beta}^{(OLS,n)}\right)^2$.

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

$$Y_i = X_i'\beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X\beta^0 + \varepsilon$$

Conditions :

$$\{(X_i', \varepsilon_i)'\}_{i=1}^{\infty} \text{ i.i.d., } F_{X,\varepsilon}(x, v) = F_X(x) \cdot F_\varepsilon(v), \ F_\varepsilon(v) = \mathcal{N}\left(0, \sigma^2\right),$$

$$Q = E\left[X_1 \cdot X_1'\right] \text{ is regular.}$$

Significance of ℓ-th explanatory variable $X_{i\ell}$ $\Leftrightarrow$ $H_0 : \ \hat{\beta}_\ell^{(OLS,n)} = 0$

Denote $\quad c_{\ell,\ell}^2 = \left[(X'X)^{-1}\right]_{\ell,\ell} \quad$ and $\quad s_n^2 = \frac{1}{n-p}\sum_{i=1}^n \left(Y_i - X_i'\hat{\beta}^{(OLS,n)}\right)^2.$

Test is based on

$$\mathcal{L}\left(\frac{\hat{\beta}_\ell^{(OLS,n)} - \beta_\ell^0}{s_n \cdot c_{\ell,\ell}}\right) = \mathcal{L}\left(t_\ell\right) = t_{n-p}$$

(Fisher-Cochran theorem)

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Let's recall $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0 = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y}$

$$\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\widehat{\tilde{Y}} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\varepsilon}, \qquad (1)$$

- firstly, $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0$ is function of $\widehat{\tilde{Y}} = \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y}$,

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Let's recall $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X}) - \beta^0 = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y}$

$$\left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{X} \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y} = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\widehat{\tilde{Y}} = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{\varepsilon}, \qquad (1)$$

- firstly, $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X}) - \beta^0$ is function of $\widehat{\tilde{Y}} = \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y}$,
- secondly,due to normality of disturbances $\tilde{\varepsilon}$,
  also $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X}) - \beta^0$ is normally distributed.

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Let's recall $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0 = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y}$

$$\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\hat{\tilde{Y}} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\varepsilon}, \qquad (1)$$

- firstly, $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0$ is function of $\hat{\tilde{Y}} = \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y}$,
- secondly,due to normality of disturbances $\tilde{\varepsilon}$,
  also $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0$ is normally distributed.

Let's also recall

$$\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right) = \left(I - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right)\tilde{\varepsilon}, \qquad (2)$$

again due to normality of disturbances,

$\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$ is normally distributed.

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Let's recall $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0 = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y}$

$$\left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y} = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\hat{\tilde{Y}} = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{\varepsilon}, \qquad (1)$$

- firstly, $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0$ is function of $\hat{\tilde{Y}} = \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{Y}$,
- secondly,due to normality of disturbances $\tilde{\varepsilon}$, also $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0$ is normally distributed.

Let's also recall

$$\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right) = \left(I - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\right)\tilde{\varepsilon}, \qquad (2)$$

again due to normality of disturbances, $\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$ is normally distributed.

Notice the orthogonality of the projection matrices in (1) and (2).

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

The orthogonality of the projection matrices in (1) and (2)

$$\Rightarrow \quad \widehat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y}, \tilde{X})\right)$$

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

The orthogonality of the projection matrices in (1) and (2)

$$\Rightarrow \quad \widehat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$$

normality $\Rightarrow$ independence of $\widehat{\tilde{Y}}$ and $\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

The orthogonality of the projection matrices in (1) and (2)

$$\Rightarrow \quad \widehat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$$

normality $\Rightarrow$ independence of $\widehat{\tilde{Y}}$ and $\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$

$\Rightarrow$ independence of $\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0$ and $\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)$

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Recalling that we consider the model

$$\tilde{Y} = \tilde{X}\beta^0 + \tilde{\varepsilon} \;\text{ with }\; \mathcal{L}(\tilde{\varepsilon}) = \mathcal{N}\left(0, \sigma^2 \tilde{W}^2\right),$$

we have from

$$\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0 = \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\varepsilon},$$

$$E\left\{\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0\right\} = 0$$

and

$$\operatorname{cov}\left\{-\beta^0\right\} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{W}^2\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}$$

$$= (X'WX)^{-1}\cdot\sum_{i=1}^{n} w_i^2\cdot\tilde{X}_i\cdot\tilde{X}_i'\,(X'WX)^{-1}.$$

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Denote

$$\left[\operatorname{cov}\left\{\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0\right\}\right]_{\ell\ell}$$

$$= \left[(X'WX)^{-1} \cdot \sum_{i=1}^{n} w_i^2 \cdot X_i \cdot X_i' \, (X'WX)^{-1}\right]_{\ell\ell} \underbrace{=}_{(denote)} d_{n,\ell}(w,X).$$

Recalling the classical regression for $\tilde{Y}, \tilde{X}$

Denote

$$\left[\operatorname{cov}\left\{\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta^0\right\}\right]_{\ell\ell}$$

$$= \left[(X'WX)^{-1} \cdot \sum_{i=1}^{n} w_i^2 \cdot X_i \cdot X_i' \, (X'WX)^{-1}\right]_{\ell\ell} \underbrace{=}_{(denote)} d_{n,\ell}(w,X).$$

Then

$$\mathcal{L}\left(\frac{\hat{\beta}_\ell^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X}) - \beta_\ell^0}{\sigma d_{n,\ell}(w,X)}\right) = \mathcal{N}(0,1).$$

Establishing the result

We can show (similarly as in the OLS-regression),

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)\cdot\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)\right)$$

$$=\mathcal{L}\left(\sigma^{-2}\cdot RSS\right)=\chi^2_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized}\left(n-p\right)$ is distribution of the sum of squares of $n-p$ independent r. v.'s normally distributed with zero mean but variance not equal one, but $\sigma^{-2}\cdot w_i$.

Establishing the result

We can show (similarly as in the OLS-regression),

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)\cdot\tilde{r}\left(\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})\right)\right)$$

$$= \mathcal{L}\left(\sigma^{-2}\cdot RSS\right) = \chi^2_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized}(n-p)$ is distribution of the sum of squares of $n-p$ independent r. v.'s normally distributed with zero mean but variance not equal one, but $\sigma^{-2}\cdot w_i$.

We conclude

THEOREM

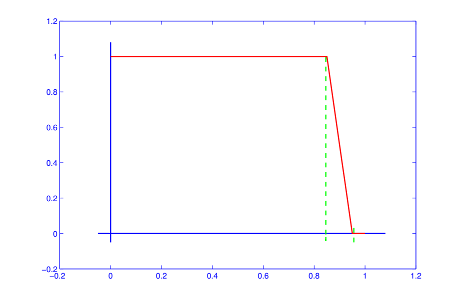
$$\mathcal{L}\left(\frac{\hat{\beta}^{(OLS,n,W,\pi)}(\tilde{Y},\tilde{X})-\beta^0}{d_{n,\ell}(w,X)}\cdot\left[\frac{\sum_{i=1}^n w_i(1-d_{ii})}{RSS}\right]^{\frac{1}{2}}\right) = t_{generalized}\left(n-p\right).$$

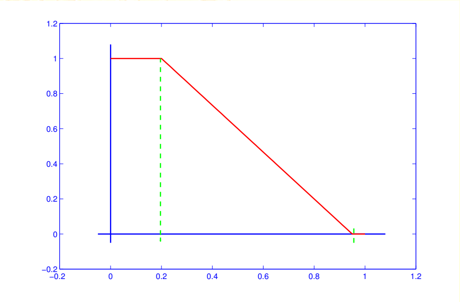
Content

Function generating the weights

Under low contamination, the intuitively optimal (left)
and really optimal (right) weight function
(in the *sense of mean square error* of the estimates of regression coefficients).

$w_\ell = w\left(\frac{\ell-1}{n}\right)$ Contamination : 4% outliers

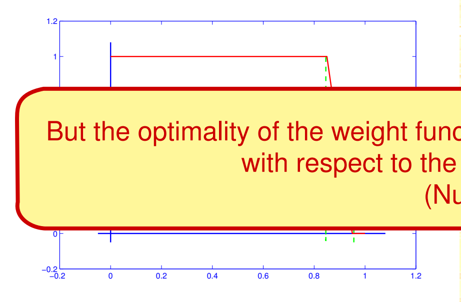


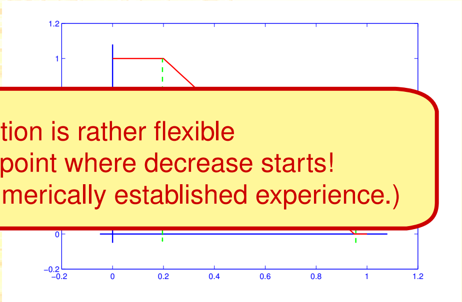


Function generating the weights

Under low contamination, the intuitively optimal (left)
and really optimal (right) weight function
(in the *sense of mean square error* of the estimates of regression coefficients).

$w_\ell = w\left(\frac{\ell - 1}{n}\right)$ Contamination : 4% outliers





But the optimality of the weight function is rather flexible
with respect to the point where decrease starts!
(Numerically established experience.)

Content

Framework of simulations

1. For fixed *n* we generated 5000 η's.

Framework of simulations

1. For fixed n we generated 5000 η's.
2. The 4875 and 4975 order statistics among thesee 5000 values were found.

Framework of simulations

1. For fixed n we generated 5000 η's.
2. The 4875 and 4975 order statistics among thesee 5000 values were found.
3. We have repeated it 100 times
   $\rightarrow$ **empirical means** and the **roots of mean square errors** over these 100 repetitions.

Framework of simulations

1. For fixed n we generated 5000 η's.
2. The 4875 and 4975 order statistics among thesee 5000 values were found.
3. We have repeated it 100 times
   $\rightarrow$ **empirical means** and the **roots of mean square errors** over these 100 repetitions.
4. We have done it for $n = 20, 30, ..., 190$.

TABLE 1

The simulated quantiles for 5%.

n	20	30	40	50	60	70
$\hat{t}^{LWS}_{0.975}(n)$	2.148 $_{(0.047)}$	2.087 $_{(0.040)}$	2.056 $_{(0.046)}$	2.027 $_{(0.045)}$	2.017 $_{(0.046)}$	2.012 $_{(0.045)}$
$t_{0.975}(n)$	2.085	2.043	2.022	2.009	2.000	1.995

n	80	90	100	110	120	130
$\hat{t}^{LWS}_{0.975}(n)$	2.008 $_{(0.040)}$	1.999 $_{(0.041)}$	1.992 $_{(0.040)}$	1.991 $_{(0.041)}$	1.990 $_{(0.040)}$	1.988 $_{(0.040)}$
$t_{0.975}(n)$	1.990	1.987	1.984	1.982	1.980	1.978

n	140	150	160	170	180	190
$\hat{t}^{LWS}_{0.975}(n)$	1.986 $_{(0.043)}$	1.989 $_{(0.041)}$	1.975 $_{(0.035)}$	1.974 $_{(0.035)}$	1.973 $_{(0.035)}$	1.973 $_{(0.035)}$
$t_{0.975}(n)$	1.977	1.976	1.975	1.974	1.974	1.973

TABLE 2

The simulated quantiles for 1%.

n	20	30	40	50	60	70
$\hat{t}^{LWS}_{0.995}(n)$	2.999 (0.100)	2.825 (0.082)	2.766 (0.080)	2.702 (0.085)	2.688 (0.077)	2.678 (0.079)
$t_{0.995}(n)$	2.845	2.748	2.705	2.678	2.661	2.651

n	80	90	100	110	120	130
$\hat{t}^{LWS}_{0.995}(n)$	2.659 (0.067)	2.644 (0.075)	2.633 (0.077)	2.627 (0.063)	2.629 (0.070)	2.626 (0.071)
$t_{0.995}(n)$	2.640	2.632	2.625	2.619	2.614	2.612

n	140	150	160	170	180	190
$\hat{t}^{LWS}_{0.995}(n)$	2.619 (0.072)	2.621 (0.073)	2.609 (0.079)	2.609 (0.070)	2.620 (0.078)	2.602 (0.078)
$t_{0.995}(n)$	2.611	2.610	2.609	2.608	2.606	2.605

TABLE 2

The simulated quantiles for 1%.

n	20	30	40	50	60	70
$\hat{t}^{LWS}_{0.995}(n)$	2.999 (0.100)	2.825 (0.082)	2.766 (0.080)	2.702 (0.085)	2.688 (0.077)	2.678 (0.079)
$t_{0.995}(n)$	2.845	2.748	2.705	2.678	2.661	2.651

n	80	90	100	110	120	130
$\hat{t}^{LWS}_{0.995}(n)$	2.659 (0.067)	2.644 (0.075)	2.633 (0.077)	2.627 (0.063)	2.629 (0.070)	2.626 (0.071)
$t_{0.995}(n)$	2.640	2.632	2.625	2.619	2.614	2.612

n	140	150	160	170	180	190
$\hat{t}^{LWS}_{0.995}(n)$	2.619 (0.072)	2.621 (0.073)	2.609 (0.079)	2.609 (0.070)	2.620 (0.078)	2.602 (0.078)
$t_{0.995}(n)$	2.611	2.610	2.609	2.608	2.606	2.605

By the way,

0.995-upper quantile of the standard normal distribution is equal to 2.575.

THANKS FOR ATTENTION