

Accurate Regression Parameters and Summary Statistics Estimation in Data with Censored Missing Values

Yuliya Karpievitch

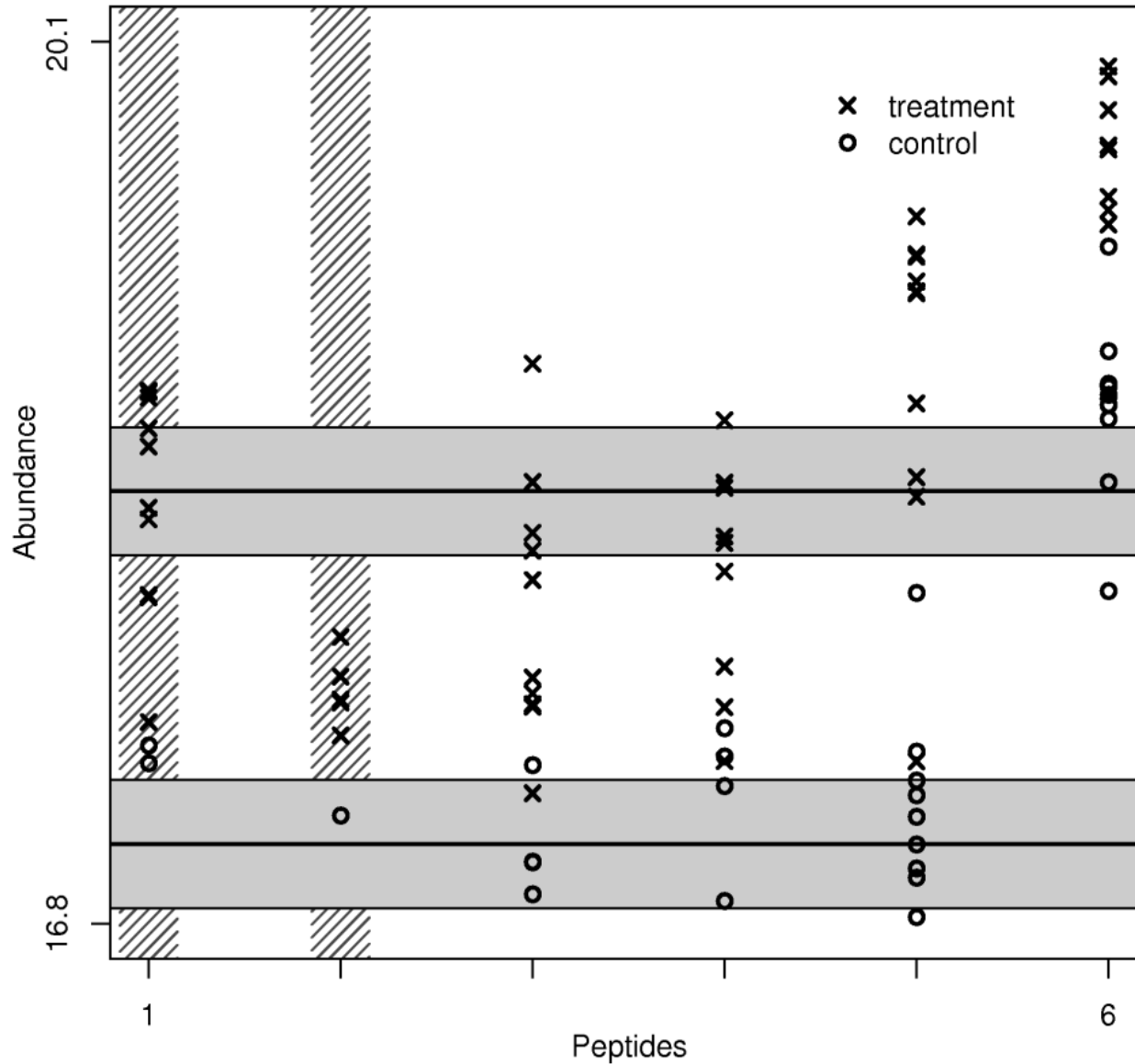
Imputation And Censored Values

Yuliya Karpievitch

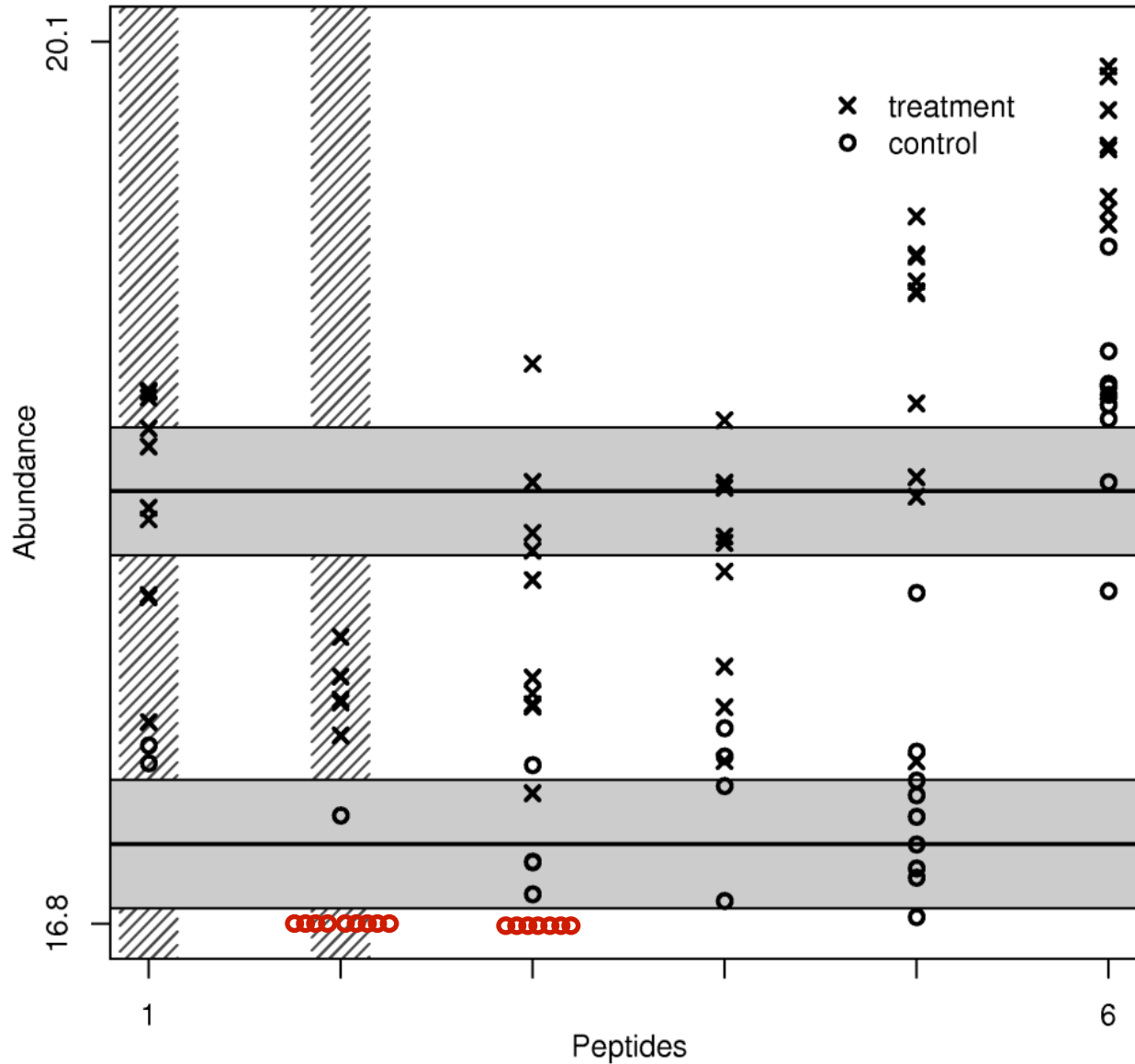
Missing Values and Imputation

- Some algorithms expect a 'complete' matrix
- Differential Expression Analysis is affected by missing values
- Censored values are lower than the observed ones
- Simple imputation methods are not good enough

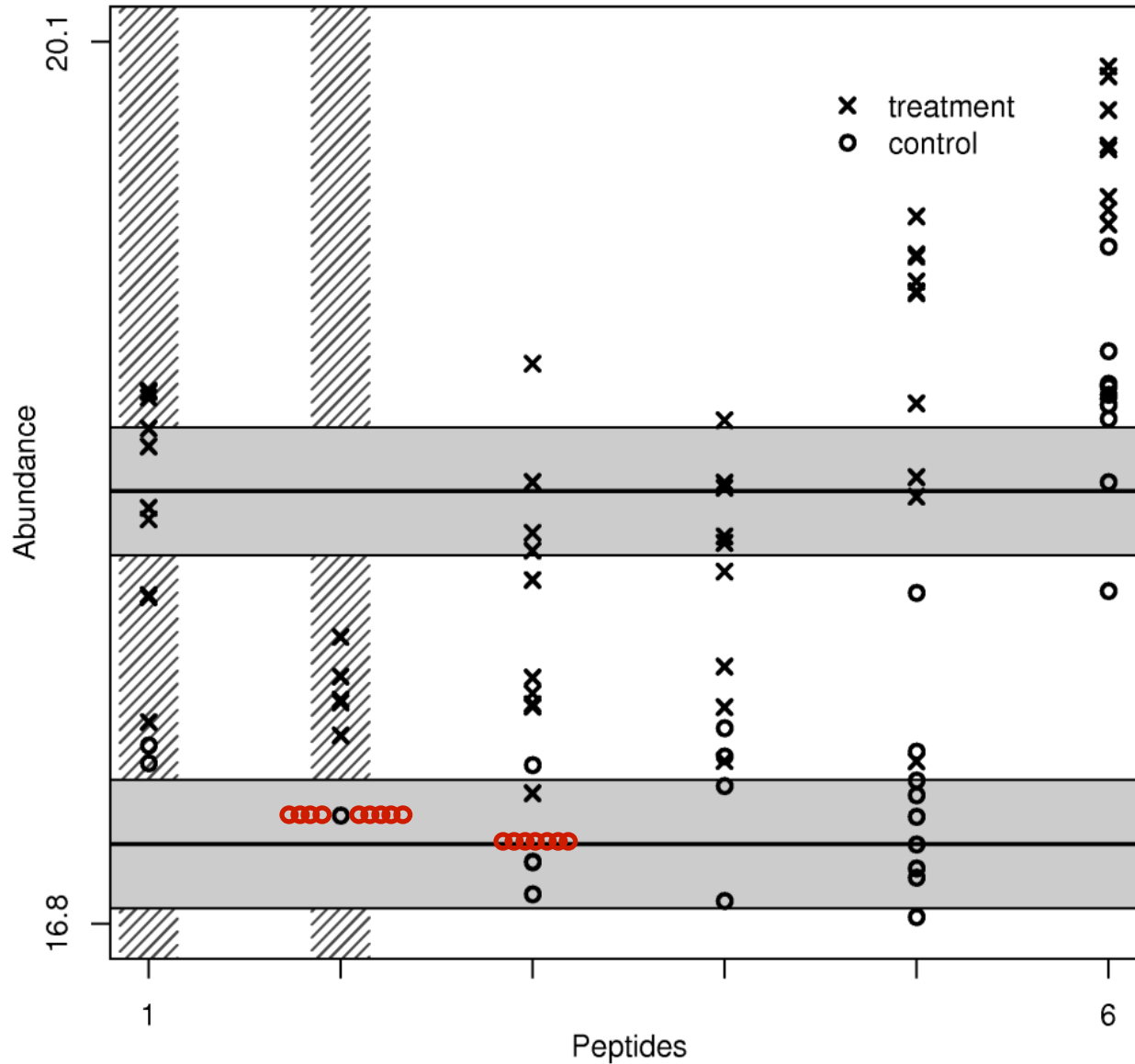
Example Protein



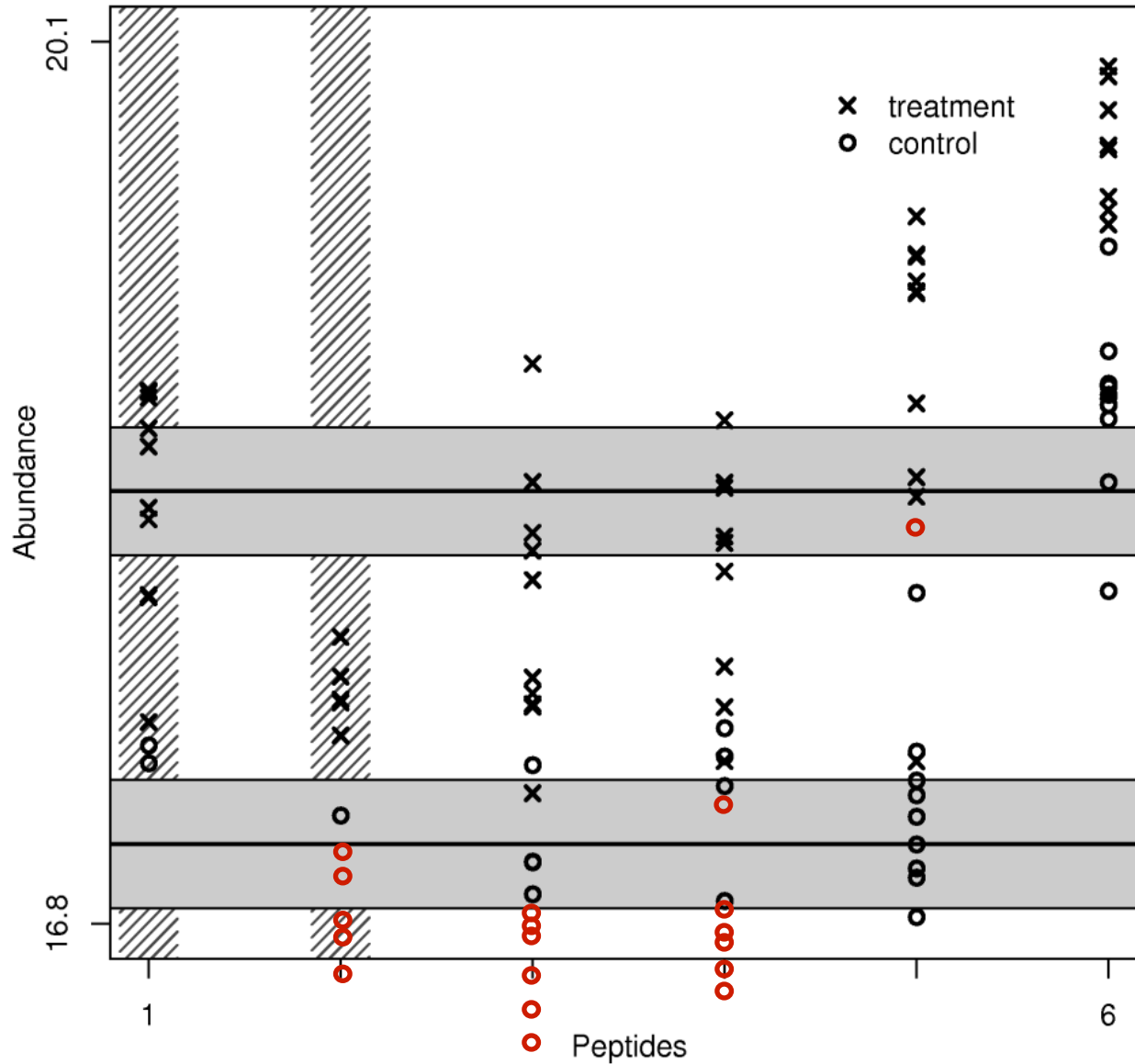
Simple Imputation – Detection Limit



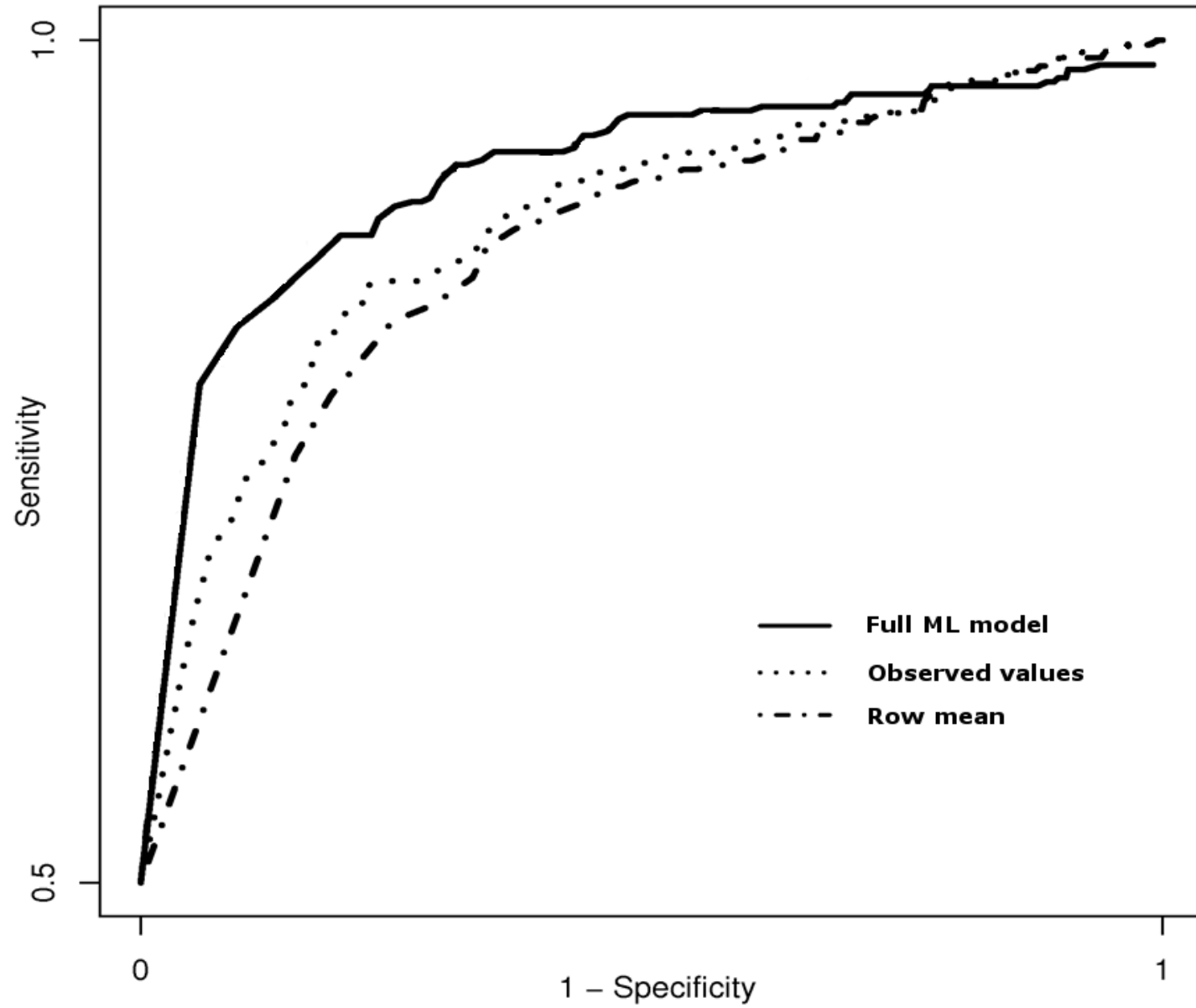
Simple Imputation - Mean



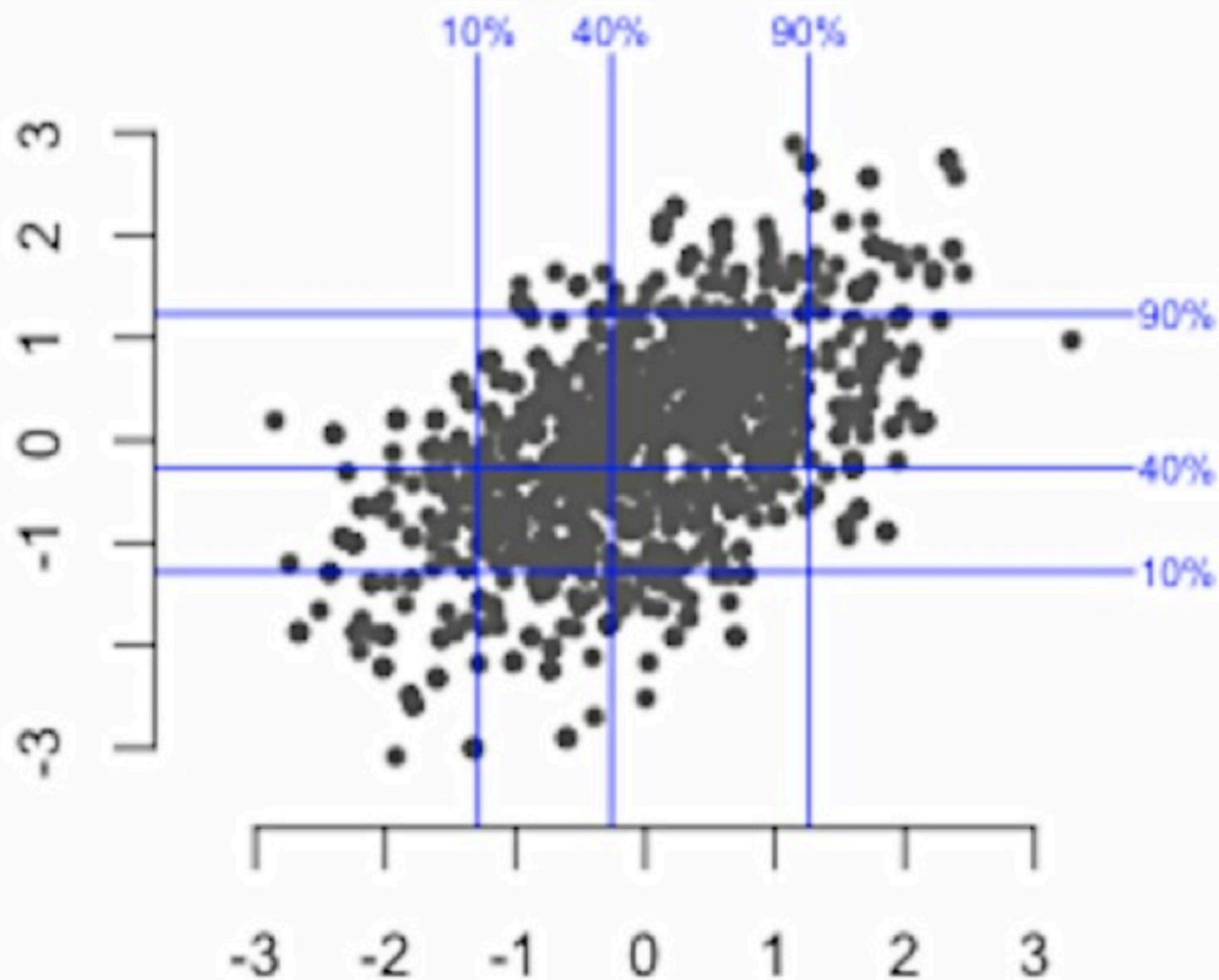
Model-based Imputation



ROC Curve for Numerical Estimation



X2



X1

X2

1.0
0.5
0.0
-0.5
-1.0

-3

-2

-1

0

1

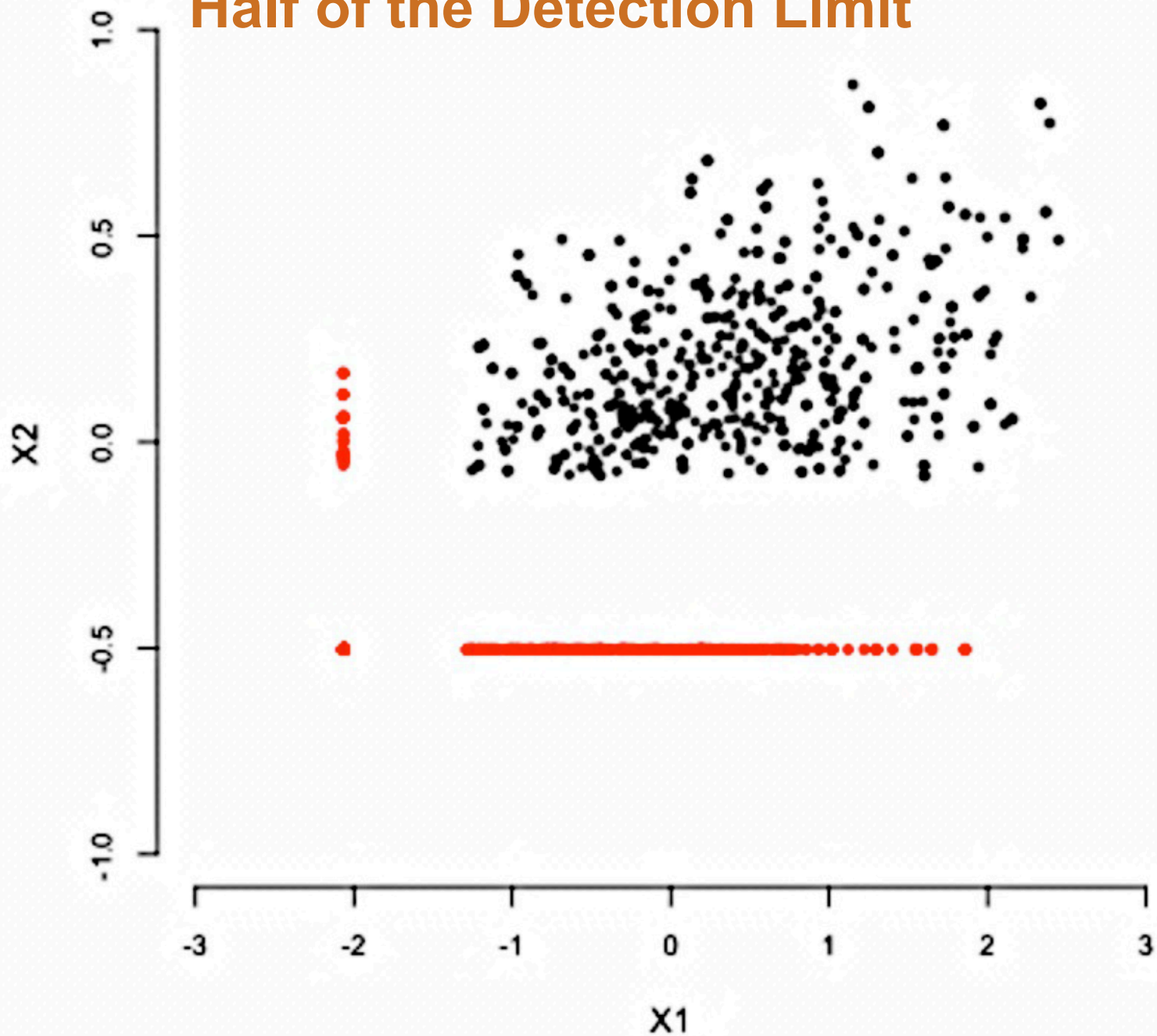
2

3

X1



Half of the Detection Limit



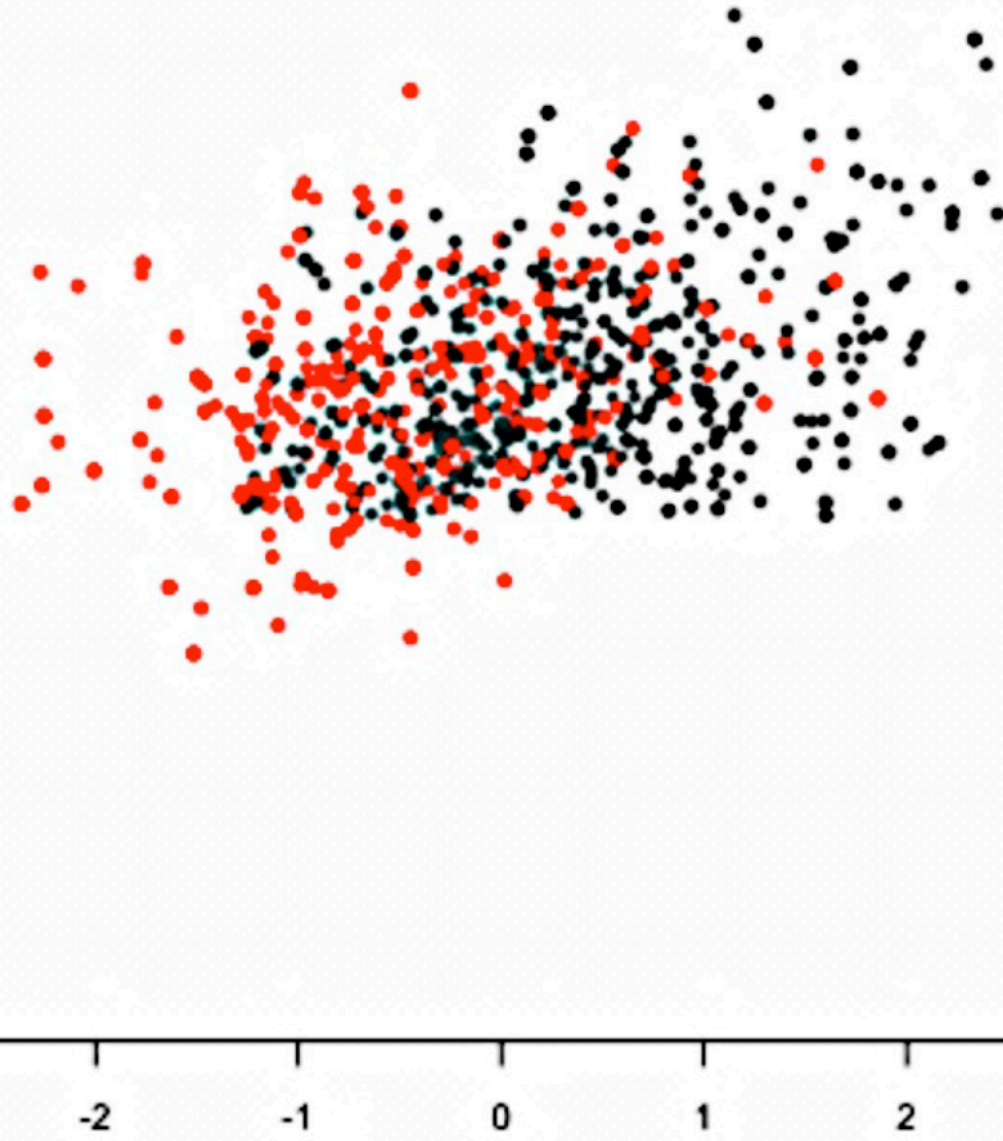
NORM

X2

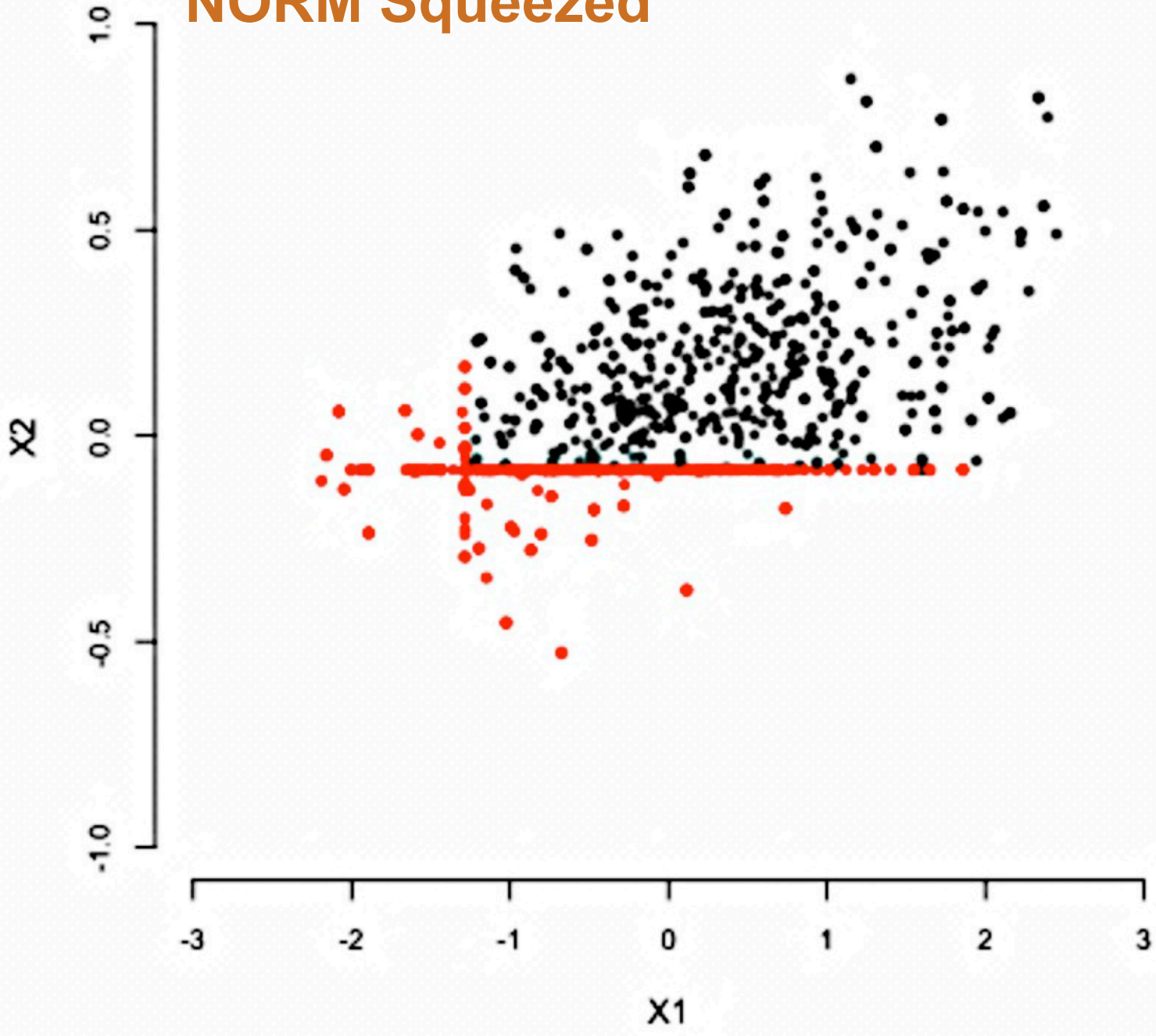
1.0
0.5
0.0
-0.5
-1.0

-3 -2 -1 0 1 2 3

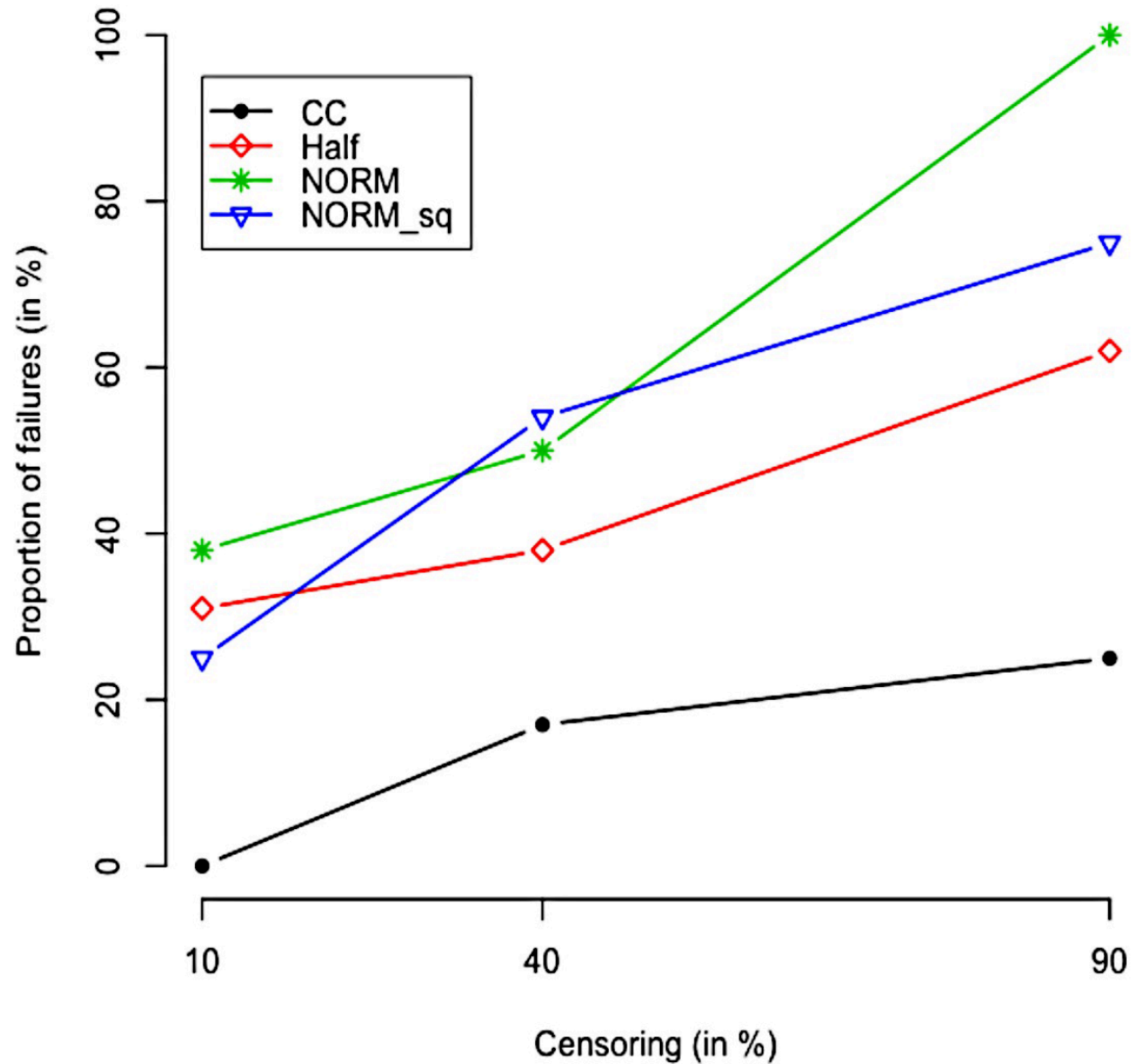
X1



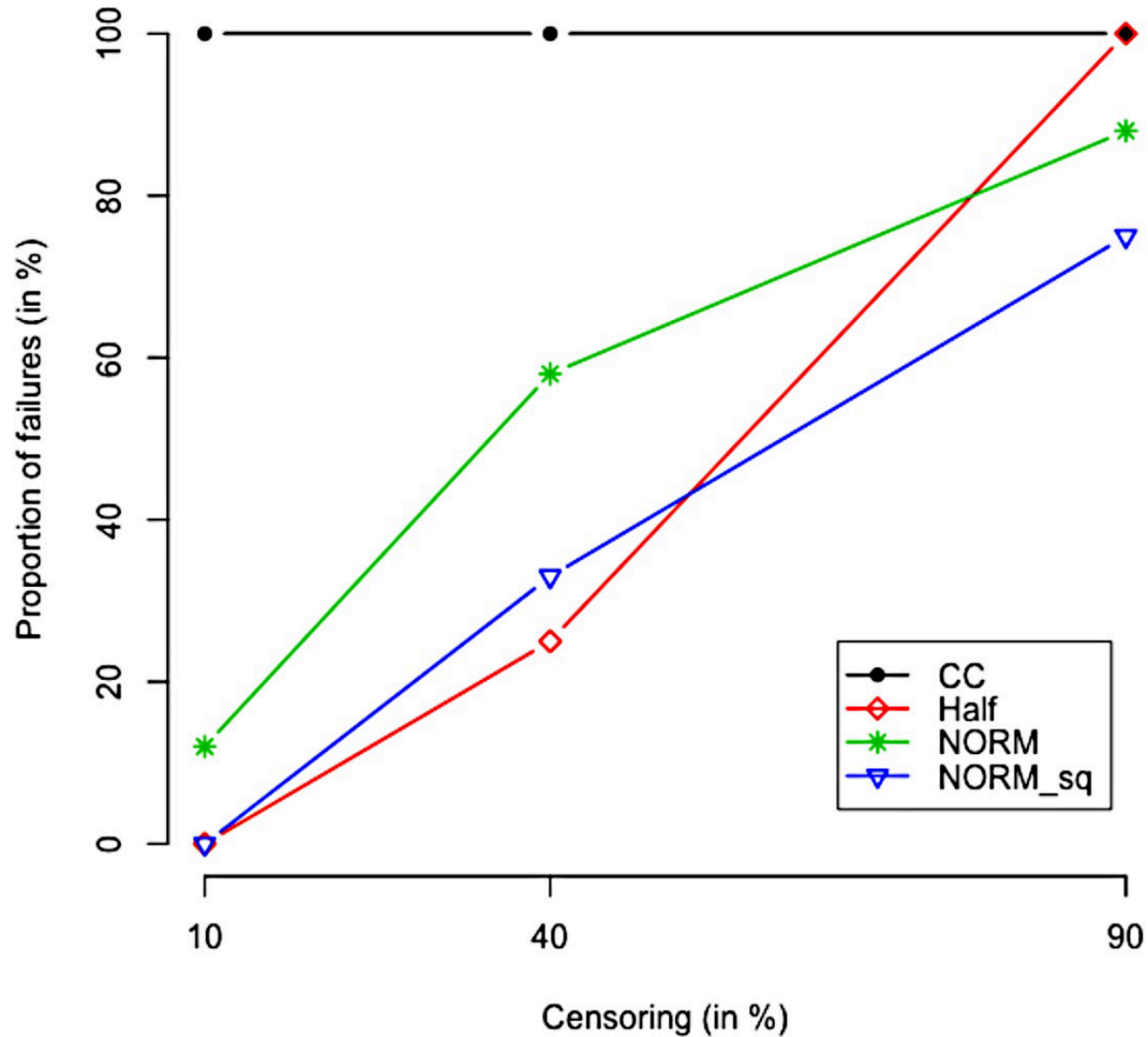
NORM Squeezed



Estimating Regression Coefficients



Estimating the Mean



Collaborators:

Firouzeh Noghrehchi
John McKinlay