Naive Bayes with quantile-based distributions

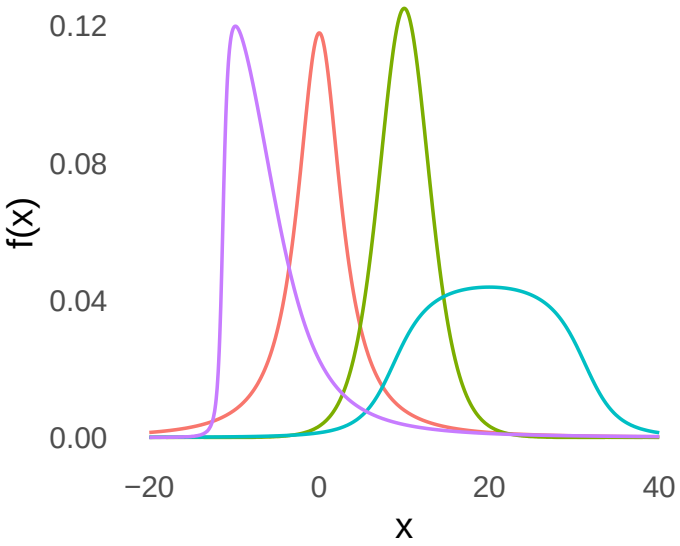
Edoardo Redivo
edoardo.redivo@unibo.it

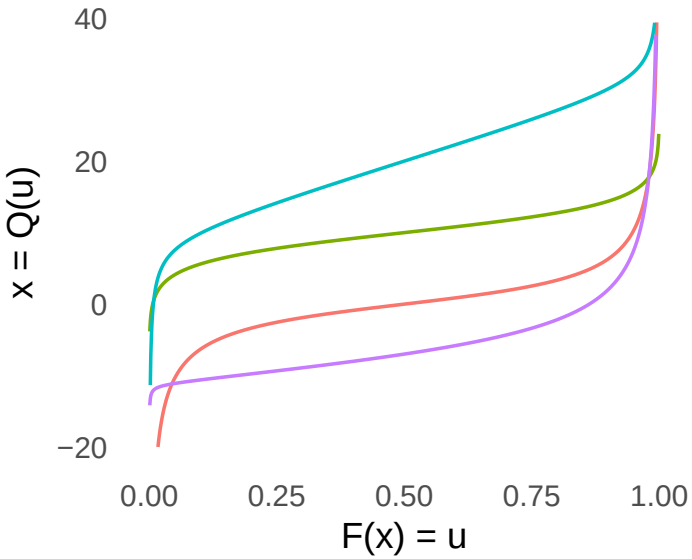
Cinzia Viroli
cinzia.viroli@unibo.it

University of Bologna

Working Group on Model-Based Clustering
Athens, 26 October 2021

Quantile-based distributions

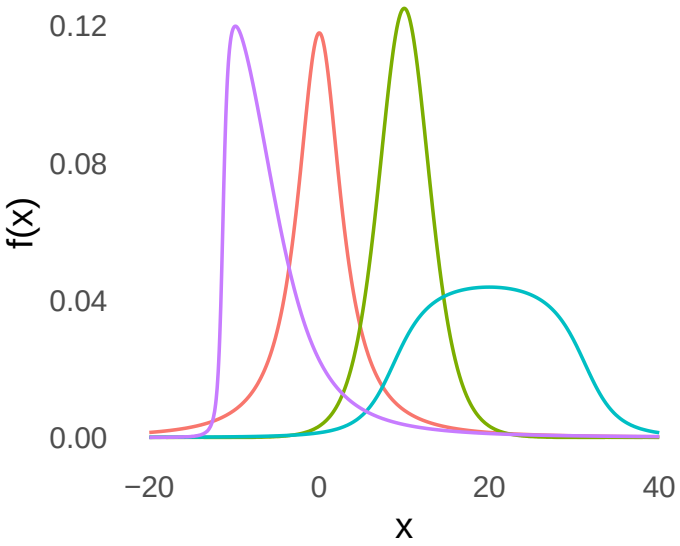


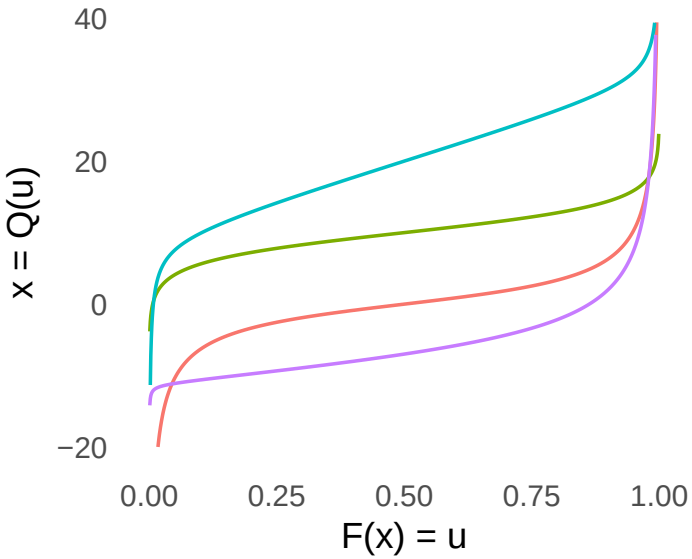


$$F(x) = u \rightarrow F^{-1}(u) = Q(u)$$

$$Q_1(u) + \cdots + Q_p(u) = Q(u)$$

Quantile-based distributions





$$F(x) = u \rightarrow F^{-1}(u) = Q(u)$$

$$Q_1(u) + \cdots + Q_p(u) = Q(u)$$

$$Q(u) = \theta_1 \, h_1(u) + \cdots + \theta_p \, h_p(u)$$

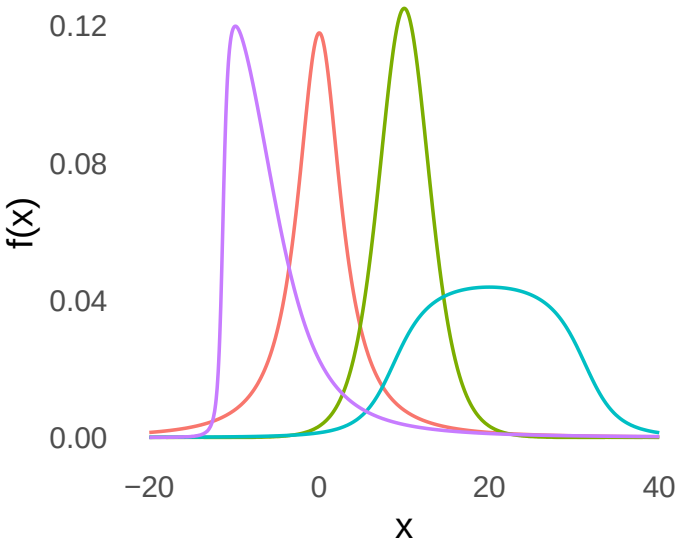
$$\mathbb{E}[X_{(i)}] = \int_0^1 Q(u) \, g(u) \, du$$

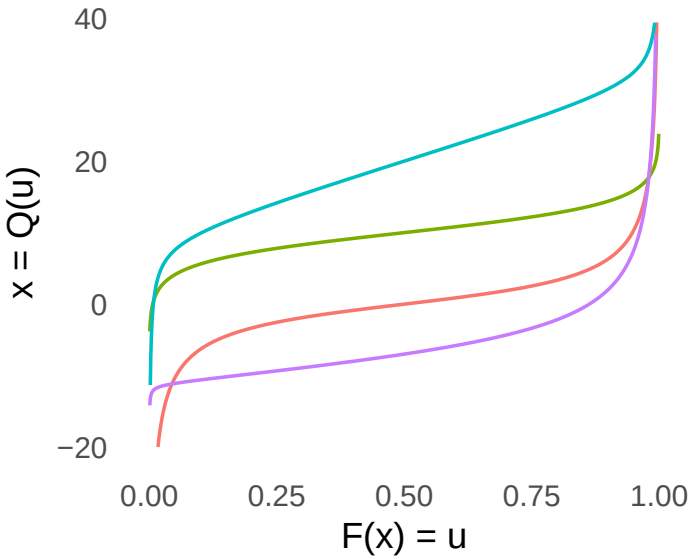
$$\Longrightarrow \mathbb{E}[X_{(i)}] = \theta_1 \, b_{1i} + \cdots + \theta_p \, b_{pi} = \boldsymbol{\theta}^\top \boldsymbol{b}_i$$

$$\sum_{i=1}^{n} (x_{(i)} - \mathbb{E}[X_{(i)}])^2 = (\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\theta})^\top (\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\theta})$$

Least squares estimator: $\hat{\boldsymbol{\theta}} = (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{x}$

Quantile-based distributions





$$F(x) = u \to F^{-1}(u) = Q(u)$$

$$Q_1(u) + \cdots + Q_p(u) = Q(u)$$

$$Q(u) = \theta_1\, h_1(u) + \cdots + \theta_p\, h_p(u)$$

$$\mathbb{E}[X_{(i)}] = \int_0^1 Q(u)\, g(u)\, du$$

$$\implies \mathbb{E}[X_{(i)}] = \theta_1\, b_{1i} + \cdots + \theta_p\, b_{pi} = \boldsymbol{\theta}^\top \boldsymbol{b}_i$$

$$\sum_{i=1}^{n} (x_{(i)} - \mathbb{E}[X_{(i)}])^2 = (\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\theta})^\top (\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\theta})$$

Least squares estimator: $\hat{\boldsymbol{\theta}} = (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{x}$

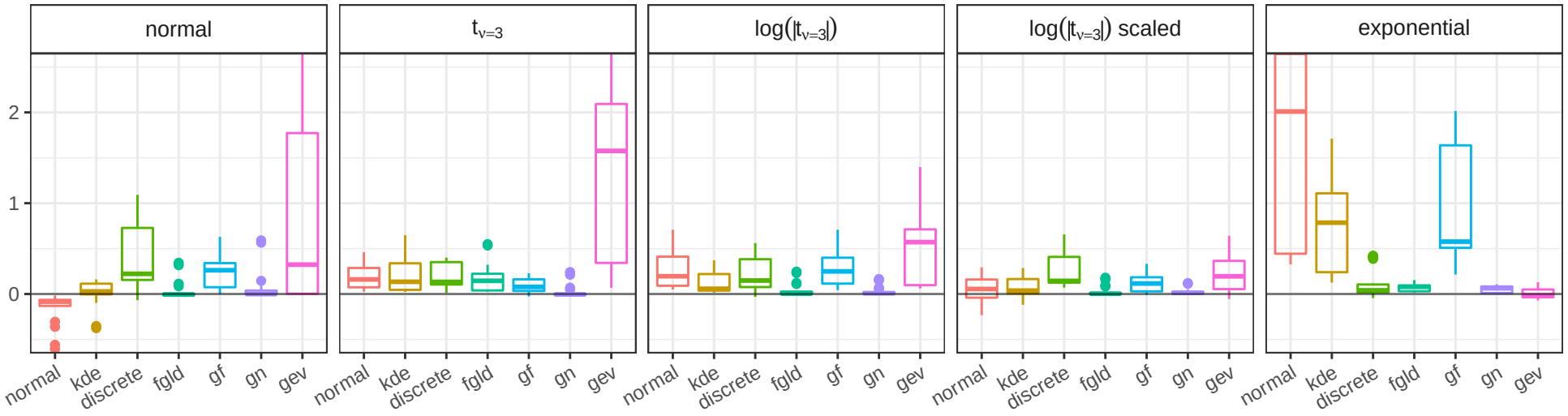
$$Q(u \mid \theta) = \theta_1 + \theta_2\, u + \theta_3\, \frac{1}{\sqrt{1-u}} - \theta_4\, \frac{1}{\sqrt{u}} + \quad \texttt{(gfgld)}$$
$$+ \theta_5\, \log u - \theta_6\, \log(1-u)$$

Other functions used are termed: `fgld`, `gn`, `gf`.

An application: naive Bayes

Simulation study

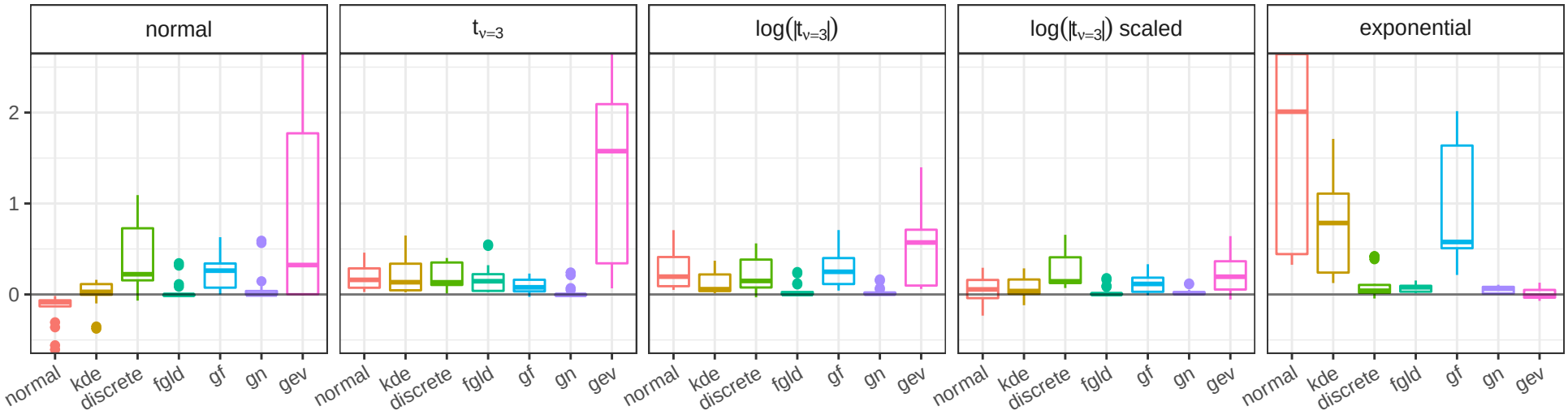
Scaled error difference from gfgld

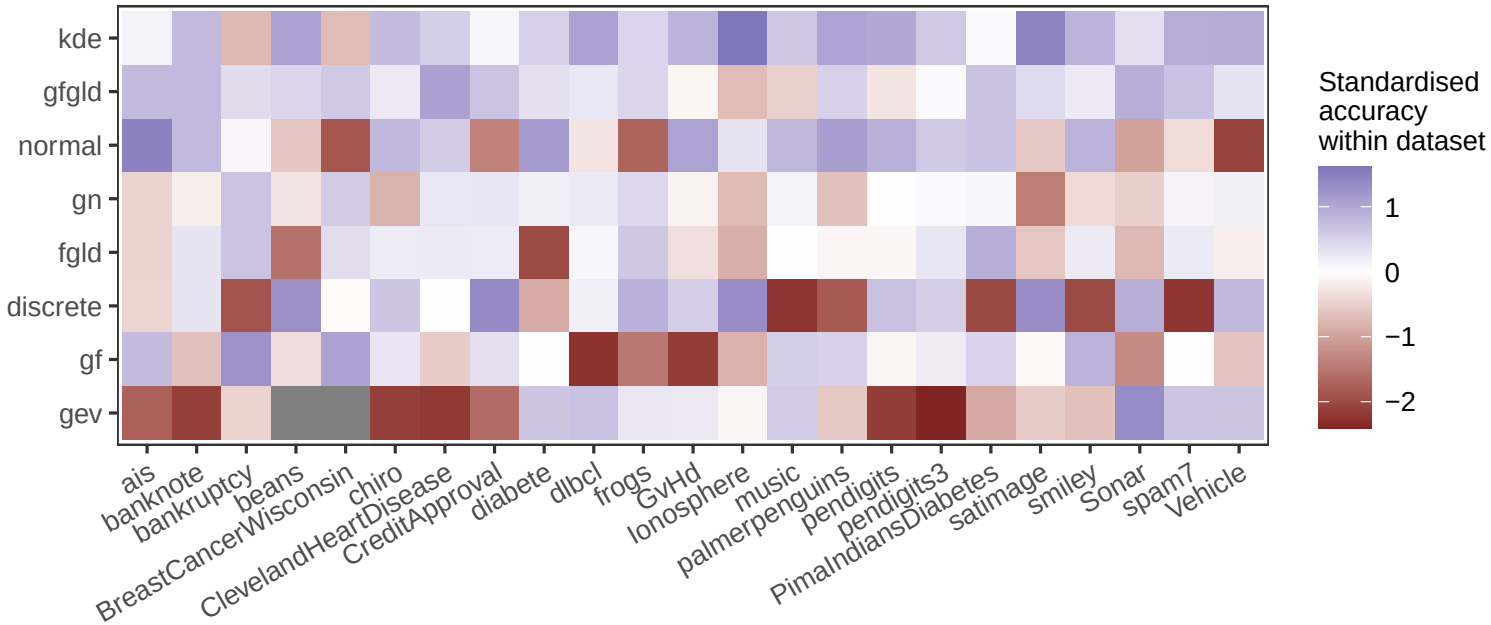


An application: naive Bayes

Simulation study

Scaled error difference from gfgld



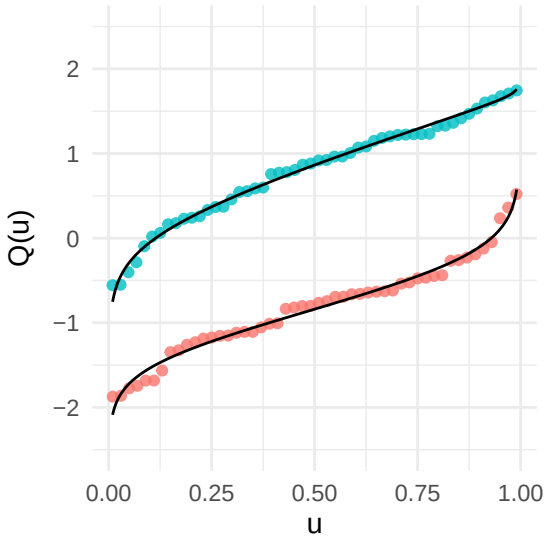


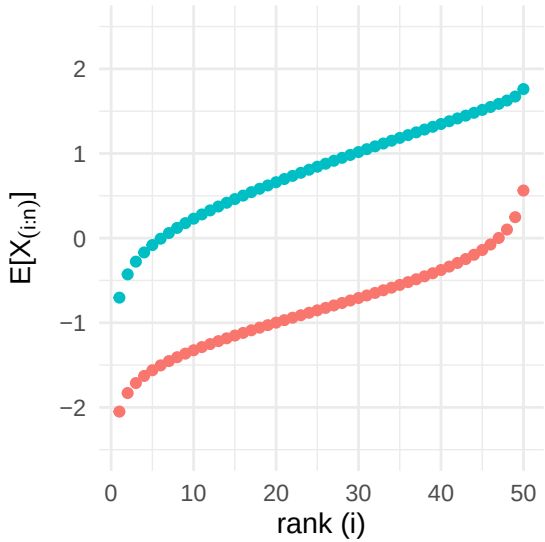
Variable importance and selection

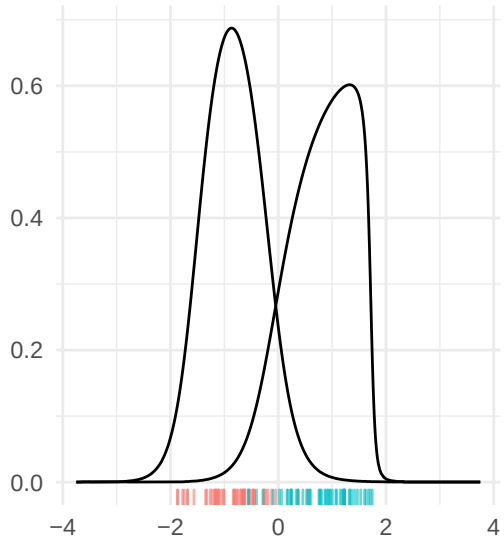
$$\hat{\mathbb{E}}[X_{(1:n)}] = \boldsymbol{B}\,\hat{\theta} \quad \rightarrow \quad ||\boldsymbol{B}\hat{\theta}_0 - \boldsymbol{B}\hat{\theta}_1||_2 = d(X \mid Y = 0,\ X \mid Y = 1)$$

Variable importance and selection

$$\hat{\mathbb{E}}[X_{(1:n)}] = \boldsymbol{B}\,\hat{\theta} \quad \rightarrow \quad ||\boldsymbol{B}\hat{\theta}_0 - \boldsymbol{B}\hat{\theta}_1||_2 = d(X \mid Y = 0,\ X \mid Y = 1)$$







Variable importance and selection

$$\hat{\mathbb{E}}[X_{(1:n)}] = \boldsymbol{B}\,\hat{\theta} \quad \rightarrow \quad ||\boldsymbol{B}\hat{\theta}_0 - \boldsymbol{B}\hat{\theta}_1||_2 = d(X \mid Y = 0,\ X \mid Y = 1)$$