Rényi pseudo-distances

Bc. Jan Kučera

FNSPE, CTU

Work supervisor: Ing. Václav Kůs Ph.D.

26.6.2012

Table of contents

Decomposable pseudo-distance

We say, that $\mathfrak{D} : \mathcal{P} \to \mathbb{R}$ is *pseudo-distance* if $\forall P, Q \in \mathcal{P}$

$$\mathfrak{D}(P, Q) \geq 0 \qquad \text{and} \qquad \mathfrak{D}(P, Q) = 0 \;\Leftrightarrow\; P = Q$$

and this pseudo-distance is *decomposable* if there exist functionals so that $\mathfrak{D}^0 : \mathcal{P} \to \mathbb{R}$, $\mathfrak{D}^1 : \tilde{\mathcal{P}} \to \mathbb{R}$ and measurable $\rho_\theta : \mathcal{X} \to \mathbb{R}$, $\theta \in \Theta$, so that $\forall \theta \in \Theta$ and $\forall Q \in \tilde{\mathcal{P}}$ there exists finite $\int \rho_\theta \mathrm{d}Q$ and

$$\mathfrak{D}(P_\theta, Q) = \mathfrak{D}^0(P_\theta) + \mathfrak{D}^1(Q) + \int \rho_\theta \mathrm{d}Q.$$

- we don't presume triangle inequality or symmetry

Minimal distance estimator

Functional $T_{\mathfrak{D}} : \tilde{\mathcal{P}} \to \Theta$ defines *minimal distance* estimator if $\mathfrak{D}(P_\theta, Q)$ is decomposable and $T_{\mathfrak{D}}(Q) \in \Theta$ minimizes

$$T_{\mathfrak{D}}(Q) = \arg\min_{\theta \in \Theta} \left[\mathfrak{D}^0(P_\theta) + \int \rho_\theta \mathrm{d}Q \right], \quad \forall Q \in \tilde{\mathcal{P}}$$

Rényi pseudo-distance

Let for some $\beta > 0$

$$p^{\beta}, q^{\beta}, \ln p \in \mathrm{L}_1(Q), \quad \forall P \in \mathcal{P}, Q \in \tilde{\mathcal{P}}.$$

holds. Then for some α, $0 < \alpha \leq \beta$, $P \in \mathcal{P}$, $Q \in \tilde{\mathcal{P}}$ is decomposable, that is

$$\mathcal{R}_{\alpha}(P, Q) = \mathcal{R}_{\alpha}^{0}(P) + \mathcal{R}_{\alpha}^{1}(Q) - \frac{1}{\alpha} \ln \left(\int p^{\alpha} \mathrm{d}Q \right),$$

where

$$\mathcal{R}_{\alpha}^{0}(P) = \frac{1}{1+\alpha} \ln \left(\int p^{\alpha} \mathrm{d}P \right), \quad \mathcal{R}_{\alpha}^{1}(Q) = \frac{1}{\alpha(1+\alpha)} \ln \left(\int q^{\alpha} \mathrm{d}Q \right).$$

Moreover for $\alpha \searrow 0$

$$\mathcal{R}_{0}(P, Q) = \int (\ln q - \ln p) \, \mathrm{d}Q$$

If we replace Q by the empirical distribution P_n, we get

$$\theta_{\mathfrak{R}_\alpha,n} = \begin{cases} \arg\max\limits_{\theta\in\Theta} \left(\int p_\theta^{1+\alpha}(x)\,\mathrm{d}x\right)^{-\frac{\alpha}{1+\alpha}} \frac{1}{n}\sum\limits_{i=1}^{n} p_\theta^{\alpha}(x_i), \\ \qquad \text{for } 0 < \alpha \leq \beta, \\ \arg\max\limits_{\theta\in\Theta} \frac{1}{n}\sum\limits_{i=1}^{n} \ln p_\theta(x_i), \\ \qquad \text{for } \alpha = 0. \end{cases}$$

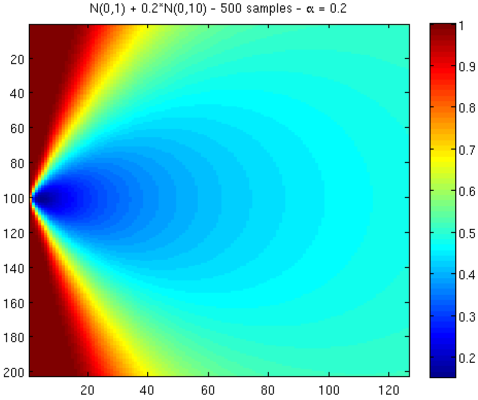
So for $\alpha = 0$ the $\theta_{\alpha,n} = \theta_{MLE}$

Normal distribution

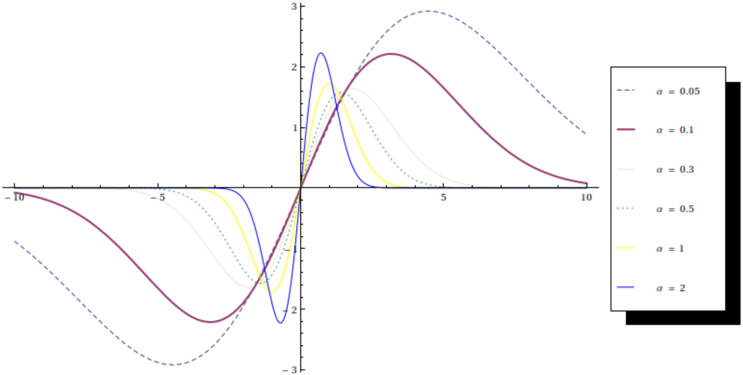
$$p_\theta = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma}\right]$$

$$\theta_{\mathfrak{R}_\alpha,n} = \arg\max_{\theta\in\Theta} \frac{1}{n\sigma^{\frac{\alpha}{1+\alpha}}} \sum_{i=1}^{n} \exp\left[-\alpha\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

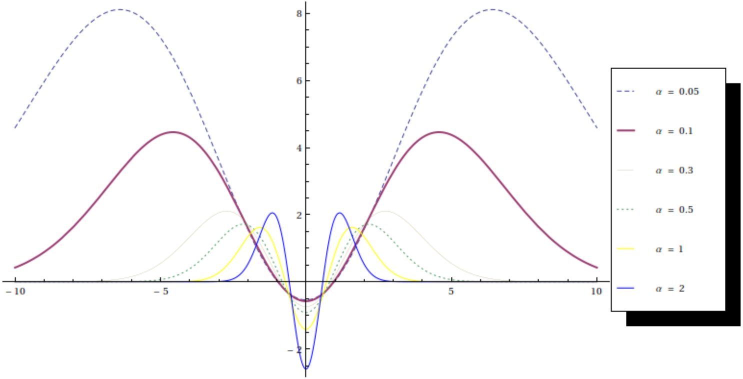
distances in the parametric space:



$$\mathrm{IF}(x; T_{\mathfrak{R}_\alpha}, \mu) = (1+\alpha)^{\frac{3}{2}}(x-\mu)e^{-\frac{\alpha}{2}(x-\mu)^2}$$



$$\mathrm{IF}(x; T_{\mathfrak{R}_\alpha}, \sigma) = \frac{(1+\alpha)^{\frac{5}{2}}\sigma}{2}\left(\left(\frac{x}{\sigma}\right)^2 - \frac{1}{1+\alpha}\right)e^{-\frac{\alpha x^2}{2\sigma^2}}$$



$\lim_{x\to\pm\infty} \mathrm{IF}(x; T_{\mathfrak{R}_\alpha}, \cdot) = 0$
$\to$ estimator is robust against outliers
$\sup_x |\mathrm{IF}(x; T_{\mathfrak{R}_\alpha}, \cdot)|$ is finite
$\to$ estimator is robust against point errors

$\alpha \backslash n$	500		
	$m(\mu)$ $m(\sigma)$	$s(\mu)$ $s(\sigma)$	$eref(\mu)$ $eref(\sigma)$
0.0	0.000 1.672	0.037 0.051	1.000 1.000
0.01	−0.000 1.652	0.037 0.050	1.032 1.031
0.05	0.001 1.572	0.033 0.046	1.248 1.244
0.1	−0.000 1.478	0.031 0.041	1.426 1.560
0.2	−0.000 1.333	0.029 0.035	1.650 2.097
0.5	−0.000 1.156	0.028 0.028	1.733 3.295

Table: Renyi: $p = \mathrm{N}(0,1)$, data: $(1-\varepsilon)\mathrm{N}(0,1) + \varepsilon\mathrm{N}(0,10)$, $\varepsilon = 0.2$, $K = 10000$

$\alpha \backslash n$	500		
	$m(\mu)$ $m(\sigma)$	$s(\mu)$ $s(\sigma)$	$eref(\mu)$ $eref(\sigma)$
0.0	-0.000 0.949	0.021 0.016	1.000 1.000
0.01	-0.000 0.949	0.021 0.016	1.002 0.969
0.05	0.001 0.945	0.021 0.016	0.991 0.967
0.1	0.000 0.942	0.021 0.017	0.990 0.889
0.2	-0.001 0.931	0.021 0.017	1.032 0.865
0.5	-0.000 0.905	0.021 0.020	0.998 0.641

Table: Renyi: $p = \mathrm{N}(0,1)$, data: $0.9\mathrm{N}(0,1) + 0.1\mathrm{N}_{0.1x}(0,1)$, $K = 1000$

$\alpha \backslash n$	500		
	$m(\mu)$ $m(\sigma)$	$s(\mu)$ $s(\sigma)$	$eref(\mu)$ $eref(\sigma)$
0.0	−0.002 3.294	0.075 0.231	1.000 1.000
0.01	−0.000 3.104	0.065 0.220	1.327 1.103
0.05	0.001 2.118	0.040 0.189	3.489 1.493
0.1	0.001 1.277	0.026 0.053	8.356 18.986
0.2	0.000 1.087	0.025 0.024	8.796 91.365
0.5	−0.000 1.029	0.026 0.021	8.390 123.291

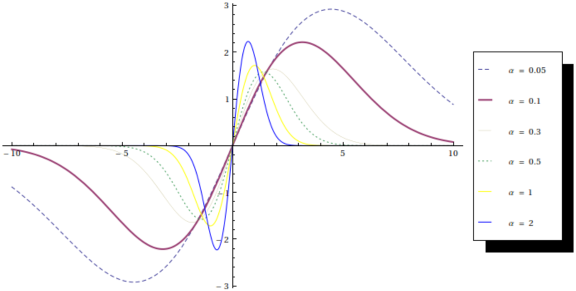
Table: Renyi: $p = \mathrm{N}(0,1)$, data: $0.9\mathrm{N}(0,1) + 0.1\mathrm{N}_{10.0x}(0,1)$, $K = 1000$

Laplace distribution

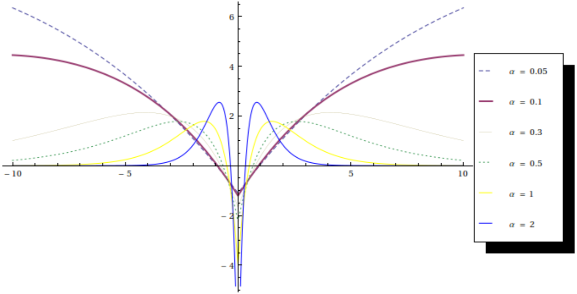
$$p_\theta = \frac{1}{2\lambda} e^{-\frac{|x-\mu|}{\lambda}}$$

$$\theta_{\mathfrak{R}_\alpha,n} = \arg\max_{\theta\in\Theta} (2\lambda)^{-\frac{\alpha}{1+\alpha}} \frac{1}{n} \sum_{i=1}^{n} \exp\left[-\alpha \frac{|x_i - \mu|}{\lambda}\right]$$

$$\mathrm{IF}(x;\, T_{\mathfrak{R}_\alpha}, \mu) = (1+\alpha)^{\frac{3}{2}}(x-\mu)e^{-\frac{\alpha}{2}(x-\mu)^2}$$



$$\mathrm{IF}(x;\, T_{\mathfrak{R}_\alpha}, \lambda) = (1+\alpha)^2\left(-\lambda + (1+\alpha)|x|\right) e^{-\frac{\alpha|x|}{\lambda}}$$



$\alpha \backslash n$	500		
	$m(\mu)$ $m(\lambda)$	$s(\mu)$ $s(\lambda)$	$eref(\mu)$ $eref(\lambda)$
0.0	0.000 3.723	0.065 0.355	1.000 1.000
0.01	0.003 3.593	0.063 0.338	1.090 1.102
0.05	0.002 3.239	0.063 0.306	1.063 1.348
0.1	−0.002 2.796	0.062 0.285	1.118 1.556
0.2	0.000 2.059	0.062 0.205	1.105 3.005
0.3	−0.003 1.626	0.060 0.154	1.198 5.326
0.5	0.000 1.300	0.058 0.108	1.267 10.756
1.0	−0.001 1.126	0.068 0.110	0.932 10.386

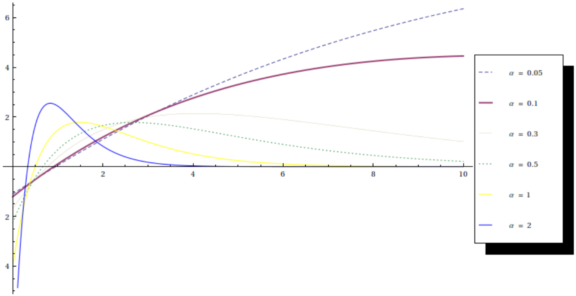
Table: Renyi: $p = \mathrm{L}(0,1)$, data: $(1-\varepsilon)\mathrm{L}(0,1) + \varepsilon\mathrm{L}(0,10)$, $\varepsilon = 0.3$, $K = 1000$

Exponential distribution

$$p_\theta = \frac{1}{\lambda} e^{-\frac{x-\mu}{\lambda}}$$

$$\theta_{\mathfrak{R}_\alpha,n} = \arg\max_{\theta\in\Theta} \lambda^{-\frac{\alpha}{1+\alpha}} \frac{1}{n} \sum_{i=1}^{n} \exp\left[-\alpha \frac{x_i-\mu}{\lambda}\right]$$

$$\mathrm{IF}(x; T_{\mathfrak{R}_\alpha}, \lambda) = (1+\alpha)^2(-\lambda + (1+\alpha)x)e^{-\frac{\alpha x}{\lambda}}$$



$\alpha \backslash n$	500		
	$m(\lambda)$	$s(\lambda)$	$eref(\lambda)$
0.0	3.901	0.432	1.000
0.01	3.614	0.356	1.475
0.05	3.241	0.317	1.862
0.1	2.796	0.274	2.495
0.15	2.389	0.239	3.258
0.2	2.049	0.209	4.281
0.3	1.622	0.153	7.999
0.5	1.310	0.110	15.330
1.0	1.139	0.109	15.806

Table: Renyi: $p = \mathrm{E}(0,1)$, data: $(1-\varepsilon)\mathrm{E}(0,1) + \varepsilon\mathrm{E}(0,10)$, $\varepsilon = 0.3$, $K = 1000$

Weibull distribution

$$p_\theta = \frac{k}{\lambda}\left(\frac{x-\mu}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x-\mu}{\lambda}\right)^{k}\right]$$

$$\begin{aligned}\theta_{\mathfrak{R}_\alpha,n} \quad = \quad & \arg\max_{\theta\in\Theta}\left(\frac{k}{\lambda}\right)^{\frac{\alpha}{1+\alpha}} (1+\alpha)^{\frac{\alpha}{1+\alpha}\frac{1+\alpha+k}{k}}\Gamma\left(\frac{1+\alpha+k}{k}\right)^{-\frac{\alpha}{1+\alpha}} \\ & \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\lambda}\right)^{\alpha(k-1)} \exp\left[-\alpha\left(\frac{x_i-\mu}{\lambda}\right)^{k}\right]\end{aligned}$$

$\alpha\backslash n$	500		
	$m(\lambda)$	$s(\lambda)$	$eref(\lambda)$
0.0	4.713	0.387	1.000
0.01	4.568	0.364	1.130
0.05	4.149	0.373	1.077
0.1	3.511	0.383	1.020
0.15	2.574	0.428	0.815
0.2	1.664	0.261	2.201
0.3	1.226	0.076	25.766
0.5	1.096	0.057	46.725
1.0	1.043	0.055	49.962

Table: Renyi: $p = \mathrm{W}(1.5, 1)$, data: $(1-\varepsilon)\mathrm{W}(1.5, 1) + \varepsilon\mathrm{W}(1.5, 10)$, $\varepsilon = 0.3$, $K = 1000$

Thank you for your attention.