Annotated selected works

Frank Nielsen

We highlight the main result of each selected work as follows:

- Nielsen, F. and Okamura, K. (2023). On f-Divergences between Cauchy distributions. *IEEE Transactions on Information Theory*, 69(5):3150–3171 : The main result is that all f-divergences $I_f(p:q) = \int p(x) f\left(\frac{q(x)}{p(x)}\right) \mathrm{d}x$ between univariate Cauchy distributions $p_{l_1,s_1}(x)$ and $p_{l_2,s_2}(x)$ are symmetric by showing that the χ^2-divergence is a *maximal invariant* Eaton (1989) for the linear fractional transform action of $\mathrm{SL}(2,\mathbb{R})$ (real fractional linear group) when Cauchy distributions $p_{l,s}$ are parametrized by a complex number $\theta = l + is$. That is $a.x \mapsto \frac{ax+b}{cx+d}$ and $A.X \sim \mathrm{Cauchy}(A.\theta)$ when $X \sim \mathrm{Cauchy}(\theta)$ for $A = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$. Since all f-divergences are invariant under this group action, they can be expressed as a scalar function h_f of the maximal invariant $\chi(l_1,s_1;l_2,s_2) = I_{\chi^2}(p_{\theta_1}:p_{\theta_2}) = \frac{(l_1-l_2)^2}{2s_1s_2}$ divergence:

$$I_f(p_{l_1,s_1}:p_{l_2,s_2}) = h_f(\chi(l_1,s_1;l_2,s_2)) = I_f(p_{l_2,s_2}:p_{l_1,s_1}).$$

- Nielsen, F. (2022). Statistical divergences between densities of truncated exponential families with nested supports: Duo Bregman and duo Jensen divergences. *Entropy*, 24(3):421 : Consider two truncated densities $p_{\theta_1}^{R_1}$ and $p_{\theta_2}^{R_2}$ of an exponential family $\{p_\theta(x) = \frac{\mathrm{d}P_\theta}{\mathrm{d}\mu}(x) = 1_{\mathcal{X}}(x)\exp(\langle\theta, t(x)\rangle - F(\theta) + k(x))\}$ where R_1 and R_2 are the supports of $p_{\theta_1}^{R_1}$ and $p_{\theta_2}^{R_2}$, respectively. A density p_θ^R of a truncated exponential family belongs to another exponential family with log-normalizer $F_R(\theta) = F(\theta) + \log Z_R(\theta)$ where $Z_R(\theta) = \int_R p_\theta(x)\mathrm{d}\mu(x)$. When $R_1 \subset R_2$ (nested support), we show that

$$D_{\mathrm{KL}}[p_{\theta_1}^{R_1} : p_{\theta_2}^{R_2}] = \int_{R_1} p_{\theta_1}^{R_1}(x) \log \frac{p_{\theta_1}^{R_1}(x)}{p_{\theta_2}^{R_2}(x)} \mathrm{d}\mu(x) = B_{F_{R_2},F_{R_1}}(\theta_2:\theta_1),$$

  where B_{F_1,F_2} is a duo Bregman pseudo-divergence:

$$B_{F_1,F_2}(\theta:\theta') = F_1(\theta) - F_2(\theta') - \langle\theta - \theta', \nabla F_2(\theta')\rangle \geq 0.$$

  This is a pseudo-divergence because when $R_1 \neq R_2$, $B_{F_{R_1},F_{R_2}} > 0$. As an example, we report the formula for the Kullback-Leibler divergence between truncated normal distributions.

References

Eaton, M. L. (1989). *Group invariance applications in statistics.*

Nielsen, F. (2022). Statistical divergences between densities of truncated exponential families with nested supports: Duo Bregman and duo Jensen divergences. *Entropy*, 24(3):421.

Nielsen, F. and Okamura, K. (2023). On f-Divergences between Cauchy distributions. *IEEE Transactions on Information Theory*, 69(5):3150–3171.