Solutions

1. **Blow-up when $r + \alpha s + \beta t \neq 1$.**

   Proof. Let $c \in \mathbb{R}$ and take $f(x) \equiv c$. Then

   $$V_{\alpha,\beta,r,s,t}(P\|Q) \;\geq\; \mathbb{E}_P[c] - r\,\mathbb{E}_Q[c] \;-\; s\log\big(\mathbb{E}_Q[e^{\alpha c}]\big) \;-\; t\log\big(\mathbb{E}_Q[e^{\beta c}]\big).$$

   Since $\mathbb{E}_P[c] = c$ and $\mathbb{E}_Q[e^{\alpha c}] = e^{\alpha c}$, etc., this becomes

   $$c\,(1 - r - \alpha s - \beta t).$$

   If $1 - r - \alpha s - \beta t \neq 0$, by letting $c \to \pm\infty$ we see the supremum is $+\infty$. □

2. **Nonnegativity under $r + \alpha s + \beta t = 1$.**

   Proof. Setting $f \equiv 0$ in the variational expression gives

   $$V_{\alpha,\beta,r,s,t}(P\|Q) \;\geq\; 0 \;-\; r \cdot 0 \;-\; s\log 1 \;-\; t\log 1 \;=\; 0.$$

   Hence $V_{\alpha,\beta,r,s,t}(P\|Q) \geq 0$. □

3. **Zero at $P = Q$; converse fails.**

   Proof. If $P = Q$, then by Jensen's inequality,

   $$\mathbb{E}_P[f] \;-\; r\,\mathbb{E}_P[f] \;-\; s\log\mathbb{E}_P[e^{\alpha f}] \;-\; t\log\mathbb{E}_P[e^{\beta f}] \;\leq\; (1 - r - \alpha s - \beta t)\,\mathbb{E}_P[f] \;=\; 0,$$

   so the supremum is ≤ 0, hence zero. But one can show there exist $P \neq Q$ for which the optimal f is identically zero, so $V = 0$ yet $P \neq Q$. □

4. **Marginal-vs-joint:**

   $$V_{\alpha,\beta,r,s,t}(P_X\|Q_X) \;\leq\; V_{\alpha,\beta,r,s,t}(P_{XY}\|Q_{XY}).$$

   Proof. Every test-function $f\colon \mathcal{X} \to \mathbb{R}$ can be lifted to $f(x,y) = f(x)$, yielding the desired inequality by taking suprema. □

5. **Invariance under same channel $W_{Y|X}$.**

   Proof. Fix a channel $W(y|x)$. For any $g\colon \mathcal{Y} \to \mathbb{R}$, define $\tilde{f}(x) = \log\big(\mathbb{E}_{W(\cdot|x)}[\exp g(Y)]\big)$. Plugging $\tilde{f}$ into the X-only functional yields exactly the Y-functional for g. Optimizing over g shows equality of the two suprema. □

Data-processing and convexity.

- *Data-processing:* If $X \to Y \to Z$ is Markov, apply invariance twice to see $V(P_X \| Q_X) \geq V(P_Z \| Q_Z)$.
- *Convexity:* For fixed f, the map $(P, Q) \mapsto \mathbb{E}_P[f] - r\mathbb{E}_Q[f] - s \log \mathbb{E}_Q[e^{\alpha f}] - t \log \mathbb{E}_Q[e^{\beta f}]$ is affine in (P, Q). Supremum of affines is convex.

6. **Superadditivity over independent marginals:**

$$V_{\alpha,\beta,r,s,t}(P_{XY} \,\|\, Q_X Q_Y) \;\geq\; V_{\alpha,\beta,r,s,t}(P_X \| Q_X) + \; V_{\alpha,\beta,r,s,t}(P_Y \| Q_Y).$$

Proof. Given optimal test-functions $f(x)$ and $g(y)$ for the marginals, consider $h(x, y) = f(x) + g(y)$. Independence of $Q_X Q_Y$ yields $\mathbb{E}_{Q_X Q_Y}[\exp(\alpha h)] = \mathbb{E}_{Q_X}[e^{\alpha f}]\, \mathbb{E}_{Q_Y}[e^{\alpha g}]$, and similarly for β. Consequently the variational objective splits and the supremum is at least the sum. □

7. **The one–parameter family W_α.**

$$W_\alpha(P \| Q) \;=\; V_{\alpha, 0,\, 1-1/\alpha,\, 1/\alpha^2,\, 0}(P \| Q).$$

Limit $\alpha \to 0$. Expand the log-MGF to second order: $\log \mathbb{E}_Q[e^{\alpha f}] = \alpha \mathbb{E}_Q[f] + \frac{\alpha^2}{2} \mathrm{Var}_Q(f) + o(\alpha^2)$. Plug into the definition, rescale $f \mapsto f/\alpha$, and let $\alpha \to 0$. One obtains $\sup_f \{\mathbb{E}_P[f] - \mathbb{E}_Q[f] - \frac{1}{2}\mathrm{Var}_Q(f)\}$, whose solution is the ^{2}-divergence: $\frac{1}{2}\mathbb{E}_Q[(\frac{dP}{dQ} - 1)^2]$. □

$\boxed{\chi^2\text{-divergence superadditive}}$

Since $D_{\chi^2}(P \| Q) = 2 \lim_{\alpha \to 0} W_\alpha(P \| Q)$ and each W_α is superadditive, so is D_{χ^2}.

Closed-form for W_α, $\alpha \in (0, 1)$:

One finds that the optimal f^* satisfies

$$\frac{dP}{dQ} \;=\; \frac{\exp(\alpha f^*)}{\mathbb{E}_Q[\exp(\alpha f^*)]}$$

and hence

$$W_\alpha(P \| Q) \;=\; \frac{1}{\alpha} D_{\text{Rényi},\alpha}(P || Q) \;-\; \frac{1}{\alpha} - 1,$$

where $D_{\text{Rényi},\alpha}$ is the usual Rényi divergence of order α.

On an f–representation:

In general the five-parameter family cannot be written as a single-parameter f–divergence (except on special parameter subfamilies that reduce to KL or Rényi).