# Inequalities between the Jenson-Shannon and Jeffreys divergences.

Gavin E. Crooks
*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA*
(Dated: April 26, 2008)

Jeffreys' J-divergence and the Jensen-Shannon divergence are shown to be related by an inequality that involves a transcendental function of the Jeffreys divergence.

Jeffreys' J-divergence, a symmetrized Kullback-Leibler divergence (the average relative entropy $D(p\|q)$ between two probability distributions, $p$ and $q$) [1, 2]

$$\begin{aligned}\text{Jeffreys}(p;q) &= D(p\|q) + D(p\|q)\\ &= \sum_i p_i \log \frac{p_i}{q_i} + \sum_i q_i \log \frac{q_i}{p_i}\end{aligned}$$

and the Jensen-Shannon divergence (the mean of the relative entropy of each distribution to the mean distribution) [3–8]

$$\begin{aligned}\text{JS}(p;q) &= \tfrac{1}{2} D(p\|\tfrac{1}{2}(p+q)) + \tfrac{1}{2} D(q\|\tfrac{1}{2}(p+q))\\ &= \tfrac{1}{2}\sum_i p_i \ln \frac{p_i}{\frac{1}{2}(p_i+q_i)} + \tfrac{1}{2}\sum_i q_i \ln \frac{q_i}{\frac{1}{2}(p_i+q_i)}\end{aligned}$$

are related by the inequality

$$\text{JS}(p;q) \leq \ln \frac{2}{1+\exp\left(-\frac{1}{2}\,\text{Jeffreys}(p;q)\right)} \qquad (1)$$

This inequality is sharper than the inequality $\text{JS}(p;q) \leq \frac{1}{4}\,\text{Jeffreys}(p;q)$ described by Lin [3].

First note that the function $f(x) = \ln(1+e^{-x})$ is convex (Since the derivative in monotonically non-decreasing) and therefore that Jensen's inequality [9] $\langle f(x)\rangle \geq f(\langle x\rangle)$ implies that

$$\left\langle \ln \frac{2}{1+e^{-x}} \right\rangle \leq \ln \frac{2}{1+e^{-\langle x\rangle}}$$

Therefore,

$$\begin{aligned}\text{JS}(p;q) &= \tfrac{1}{2}\sum_i p_i \ln \frac{p_i}{\frac{1}{2}(p_i+q_i)} + \tfrac{1}{2}\sum_i q_i \ln \frac{q_i}{\frac{1}{2}(p_i+q_i)}\\ &= \tfrac{1}{2}\sum_i p_i \ln \frac{2}{1+e^{\ln(\frac{q_i}{p_i}}} + \tfrac{1}{2}\sum_i q_i \ln \frac{2}{1+e^{\ln\frac{p_i}{q_i}}}\\ &\leq \tfrac{1}{2}\ln \frac{2}{1+\exp\big(-D(p\|q)\big)} + \tfrac{1}{2}\ln \frac{2}{1+\exp\big(-D(q\|p)\big)}\\ &\leq \ln \frac{2}{1+\exp\big(-\frac{1}{2}\,\text{Jeffreys}(p;q)\big)}\end{aligned}$$

The last line follows from the previous line by a second application of the same Jensen inequality. Since the J-divergence ranges between zero and positive infinity whereas the Jensen-Shannon divergence ranges between zero and $\ln 2$ [i.e. 1 bit], inequality (1) has the correct limits for identical ($p_i = q_i$, $\text{JS}(p;q) = \text{Jeffreys}(p;q) = 0$) and orthogonal ($p_i q_i = 0$, $\text{JS}(p;q) = \ln 2$, $\text{Jeffreys}(p;q) = +\infty$) distributions.

There is no corresponding lower bound; for any value of the Jensen-Shannon divergence the Jeffreys divergence can be arbitrarily large. Consider a pair of binary distributions $(a, 1-a)$ and $(b, 1-b)$. Fix $a$ and let $b$ limit to zero. As $b$ decreases the Jensen-Shannon divergence will limit to a fixed value between 0 and $\ln 2$, but the Jeffreys divergence will increase to infinity.

Since $\ln\big(2/(1+e^{-x/2})\big) \leq x/4$ we can also deduce Lin's inequality [3].

$$\text{JS}(p;q) \leq \frac{1}{4}\,\text{Jeffreys}(p;q) \qquad (2)$$

Finally, note that we can split the Jensen-Shannon divergence into two directed parts [4], and write the equivalent of inequality (1) with respect to each part separately.

$$\tfrac{1}{2} D\left(p\|\tfrac{1}{2}(p+q)\right) \leq \tfrac{1}{2}\ln \frac{2}{1+\exp\big(-D(p\|q)\big)} \qquad (3)$$

[1] H. Jeffreys, *Theory of probability* (Oxford University Press, 1948), 2nd ed.
[2] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[3] J. Lin, IEEE Trans. Info. Theory **37**, 145 (1991).
[4] F. Topsøe, IEEE Trans. Info. Theory **46**, 1602 (2000).
[5] D. M. Endres and J. E. Schindelin, IEEE Trans. Info. Theory **49**, 1858 (2003).
[6] F. Österreicher and I. Vajda, Ann. Inst. Statist. Math **55**, 639 (2003).
[7] B. Fuglede and F. Topsøe, in *IEEE Int. Sym. on Information Theory* (2004), p. 31.
[8] A. Majtey, P. W. Lamberti, M. T. Martin, and A. Plastino, Eur. Phys. J. D **32**, 413 (2005).
[9] J. L. W. V. Jensen, Acta Mathematica **30**, 175 (1906).