A ENTROPY OF STUDENT'S T-DISTRIBUTION

While the entropy of the Student t-distribution is well known, we derive it for completeness. Student's probability distribution defined in terms of location γ, scale factor σ_{st}^2 and ν_{st} degrees of freedom is

$$p(y;\gamma,\sigma_{st}^2,\nu_{st}) = \mathrm{St}(y;\gamma,\sigma_{st}^2,\nu_{st}) = \frac{\Gamma(\frac{\nu_{st}+1}{2})}{\sqrt{\nu_{st}\pi\sigma_{st}^2}\,\Gamma(\frac{\nu_{st}}{2})}\left(1+\frac{1}{\nu_{st}}\frac{(y-\gamma)^2}{\sigma_{st}^2}\right)^{-\frac{\nu_{st}+1}{2}}, \tag{13}$$

where Γ i a gamma function. Student's t-distribution can be written in terms of beta function $\mathrm{B}=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ if we take advantage of the fact that $\Gamma(\frac{1}{2})=\sqrt{\pi}$

$$\mathrm{p}(y;\sigma_{st}^2,\nu_{st}) = \mathrm{St}(y;\sigma_{st},\nu_{st}) = \frac{1}{\sqrt{\nu_{st}\,\sigma_{st}^2}\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})}\left(1+\frac{1}{\nu_{st}}\frac{(y-\gamma)^2}{\sigma_{st}^2}\right)^{-\frac{\nu_{st}+1}{2}}. \tag{14}$$

If we introduce a new variable $t=\frac{y-\gamma}{\sigma_{st}}$, Student's t-distribution converts into the standard form with probability density

$$\mathrm{p}(t;\nu_{st}) = \mathrm{St}(y;\nu_{st}) = \frac{1}{\sqrt{\nu_{st}}\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})}\left(1+\frac{t^2}{\nu_{st}}\right)^{-\frac{\nu_{st}+1}{2}}. \tag{15}$$

A.1 PROPOSITION

Proposition: Entropy of the generalized and standard Student's t-distributions are related via the formula

$$\mathcal{H}(y;\sigma_{st}^2,\nu_{st}) = \mathcal{H}(t;\nu_{st}) + \frac{1}{2}\log\sigma_{st}^2\,. \tag{16}$$

Proof: The transformation $t=\mathrm{g}(y)=\frac{y-\gamma}{\sigma_{st}}$ is bijective and invertible with the inverse transformation $y=g^{-1}(t)=\sigma_{st}\,t+\gamma$. The Jacobin of the transformation g^{-1} is $\mathrm{J}=\sigma_{st}$. According to the change of variable probability density formula

$$\mathrm{p}_y\,(y;\sigma_{st}^2,\nu_{st}) = \mathrm{p}_t\,(g(y);\nu_{st})\,|\,\frac{\mathrm{d}}{\mathrm{d}y}\,g(y)\,|. \tag{17}$$

Entropy transformation 16 follows directly from the definition of the entropy.

To find the generalized entropy, we just need to calculate the Shannon entropy of the standard Student's t-distribution

$$\mathcal{H}(t;\nu_{st}) = -\int_{-\infty}^{+\infty}\mathrm{p}(t;\nu_{st})\log\mathrm{p}(t;\nu_{st})\,\mathrm{d}t = \tag{18}$$

$$\log\left(\sqrt{\nu_{st}}\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})\right)\int_{-\infty}^{+\infty}\mathrm{p}(t;\nu_{st})\,\mathrm{d}t\,+ \tag{19}$$

$$\frac{\nu_{st}+1}{2}\int_{-\infty}^{+\infty}\log(1+\frac{t^2}{\nu_{st}})\,\mathrm{p}(t;\nu_{st})\,\mathrm{d}t = \log\left(\sqrt{\nu_{st}}\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})\right)+ \tag{20}$$

$$\frac{\nu_{st}+1}{2}\int_{-\infty}^{+\infty}\log(1+\frac{t^2}{\nu_{st}})\,\mathrm{p}(t;\nu_{st})\,\mathrm{d}t\,. \tag{21}$$

To find the second integral we make a substitution $x=\frac{t^2}{\nu_{st}}$ and obtain

$$\frac{\nu_{st}+1}{2}\int_{-\infty}^{+\infty}\log(1+\frac{t^2}{\nu_{st}})\,\mathrm{p}(t;\nu_{st})\,\mathrm{d}t = \frac{\nu_{st}+1}{4\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})}\int_0^{+\infty}\log(1+x)\,(1+x)^{-\frac{\nu_{st}+1}{2}}\frac{\mathrm{d}x}{\sqrt{x}} = \tag{22}$$

$$-\frac{\nu_{st}+1}{2\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})}\frac{\partial}{\partial\nu_{st}}\int_0^{+\infty}(1+x)^{-\frac{\nu_{st}+1}{2}}\frac{\mathrm{d}x}{\sqrt{x}} = -\frac{\nu_{st}+1}{2\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})}\frac{\partial}{\partial\nu_{st}}\int_0^1 x^{\frac{\nu_{st}}{2}-1}(1-x)^{\frac{1}{2}-1}\,\mathrm{d}x = \tag{23}$$

$$-\frac{\nu_{st}+1}{2\,\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2})}\frac{\partial}{\partial\nu_{st}}\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2}) = -\frac{\nu_{st}+1}{2}\frac{\partial}{\partial\nu_{st}}\log\mathrm{B}(\frac{1}{2},\frac{\nu_{st}}{2}) = \tag{24}$$

$$-\frac{\nu_{st}+1}{2}\frac{\partial}{\partial\nu_{st}}\left(\log\Gamma(\frac{\nu_{st}}{2})-\log\Gamma(\frac{\nu_{st}+1}{2})\right) = \frac{\nu_{st}+1}{2}\left(\Psi(\frac{\nu_{st}+1}{2})-\Psi(\frac{\nu_{st}}{2})\right) \tag{25}$$

where digamma function is defined as $\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$. Putting all the terms together, the entropy of the Student's t-distribution becomes

$$\mathcal{H}(t; \nu_{st}) = \frac{\nu_{st} + 1}{2} \left(\Psi(\frac{\nu_{st} + 1}{2}) - \Psi(\frac{\nu_{st}}{2}) \right) + \log \left(\sqrt{\nu_{st}}\, \mathrm{B}(\frac{1}{2}, \frac{\nu_{st}}{2}) \right) \tag{26}$$

The Shannon entropy of the random variable t is

$$\mathcal{H}(t; \nu_{st}) = \frac{\nu_{st} + 1}{2} \left(\Psi(\frac{\nu_{st} + 1}{2}) - \Psi(\frac{\nu_{st}}{2}) \right) + \log \left(\sqrt{\nu_{st}}\, \mathrm{B}(\frac{1}{2}, \frac{\nu_{st}}{2}) \right) \tag{27}$$

and the Shannon entropy of the labels y is given by our final formula

$$\mathcal{H}(y; \sigma_{st}^2, \nu_{st}) = \frac{\nu_{st} + 1}{2} \left(\Psi(\frac{\nu_{st} + 1}{2}) - \Psi(\frac{\nu_{st}}{2}) \right) + \log \left(\sqrt{\nu_{st}}\, \mathrm{B}(\frac{1}{2}, \frac{\nu_{st}}{2}) \right) + \frac{1}{2} \log \sigma_{st}^2 \,. \tag{28}$$