0.1 (Not) choosing a Renyi order

The entropy of a R.V. X tells us how difficult it is to predict, which is more or less exactly what we want to quantify (the more diverse a community is, the harder it should be to predict what a random sample from said community will contain). And it relates nicely to a tangible quantity by way of the uniform distribution over an outcome set of size n, i.e.

$$H(\boldsymbol{u}^n) = -\sum^n p_i \log p_i = -\log \frac{1}{n} \sum^n \frac{1}{n} = \log n \Leftrightarrow \exp H(\boldsymbol{u}^n) = n. \quad (1)$$

This is a nice interpretation - $\exp H(\boldsymbol{x})$ is the number of *uniformly distributed* species that would be required to produce the same entropy as our specific community $\boldsymbol{x}$. Generalizing further, we get the q-parametrized equivalent species count

$$F_q(\boldsymbol{x}) := \exp\left[\frac{1}{1-q} \log \sum_i x_i^q\right], \quad (2)$$

from which we can recover the trivial species count simply by inserting $q = 0$:

$$F_0(\boldsymbol{x}) = \exp\left[\frac{1}{1} \log \sum_i x_i^0\right] = \exp \log n = n. \quad (3)$$

Different values of q gives different metrics, reflecting various different properties of the distribution. If it were true that $F_q(\boldsymbol{x}_1) > F_q(\boldsymbol{x}_2)$ for some q implies that $F_q(\boldsymbol{x}_1) > F_q(\boldsymbol{x}_2)$ for all q, then the choice of q wouldn't matter much, since the internal ordering would be preserved. But we know this isn't the case. The choice of q matters. But since we can't sensibly argue for one choice being superior to another, there's only unbiased thing we can do – include *all possible* values of q! Alas, the straight integral is divergent, while the asymptotic average

$$\lim_{q\to\infty} \frac{1}{q} \int F_q(\boldsymbol{x}) dq \quad (4)$$

is pointless for another reason - it vanishes for any $\boldsymbol{x}$. To see why, note that $\lim_{q\to\infty} \frac{1}{q} F_q(\boldsymbol{x}) \approx \frac{1}{q} \exp \frac{-1}{q} \log x_{\max}^q = \frac{1}{q} \frac{1}{x_{\max}}$.

However, by multiplying the integrand with (almost) *any* convergent function of q, we get something finite, but bounded. In fact, defining a diversity metric

$$\bar{F}(\boldsymbol{x}) := \int_0^\infty g(q) F_q(\boldsymbol{x}) dq \quad (5)$$

with *any* integrable function g that is strictly positive and monotonically decreasing for all $q \geq 0$ will produce a finite and positive diversity metric well-defined for any distribution $\boldsymbol{x}$. An example would be the Gaussian e^{-q^2} or most decaying exponentials (or indeed any function that decays more rapidly than $1/x$ asymptotically).

I don't know if it seems alien and messy at first glance, but it actually turns out to be quite elegant and should share a lot of nice qualities with the Laplace transform among other things.