# RENYI ENTROPIES FOR BERNOULLI DISTRIBUTIONS

A. Bialas and W. Czyz

M. Smoluchowski Institute of Physics, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland
e-mail: bialas@th.if.uj.edu.pl
and
Institute of Nuclear Physics
Radzikowskiego 152, 31-342 Kraków

*(Received July 12, 2001)*

An asymptotic formula for Renyi entropies characterizing a Bernoulli distribution is derived and compared with numerical estimates. Its physical consequences are discussed.

PACS numbers: 05.10.–a, 13.85.Hd

**1.** As shown in our previous papers [1, 2], the Renyi entropies [3] may serve as an useful characteristics of the multiparticle spectra. To obtain a better insight into the meaning of such measurements, we investigate in the present note the Renyi entropies for Bernoulli distributions.

The Bernoulli distribution of $N$ particles in $M$ bins

$$P(p_1,...p_M;n_1,...n_M) = \frac{N!}{n_1!...n_M!} p_1^{n_1}...p_M^{n_M} , \tag{1}$$

where $p_j$ is the probability of one particle falling into the $j$-th bin and $N$ is the total multiplicity

$$n_1+n_2+...+n_M = N \tag{2}$$

represents the simplest model of particle production with no correlations between particles (except those induced by the fixed number of particles - $N$). Uncorrelated production is a general consequence of some important mechanisms of particle production (as, *e.g.*, in the bremsstrahlung model [4]) and therefore it is worth to be investigated in detail. To our knowledge, however, nobody as yet discussed to what an extent the distribution (1) is violated in real data[1]. Our calculation may thus provide a tool for this kind of investigation.

The coincidence probabilities are defined as

$$C_l(N,M) = \sum_{n_1+...+n_M=N} [P(p_1,...p_M;n_1,...n_M)]^l \,, \tag{3}$$

and thus, using (1) we have

$$C_l(N,M) = \sum_{n_1,...,n_M} \delta_{n_1+...+n_M,N} \left( \frac{N! p_1^{n_1}...p_M^{n_M}}{n_1!...n_M!} \right)^l \,. \tag{4}$$

**2.** We want to evaluate the sum (4) in the limit

$$N \to \infty; \quad M \quad \text{fixed} \tag{5}$$

(which seems to be the most interesting one for "practical" applications).

In this limit we can replace factorials by the Stirling formula[2]

$$n! \approx \sqrt{2\pi n + 1}\, n^n \mathrm{e}^{-n} \tag{6}$$

and obtain

$$\begin{aligned} \frac{N!}{n_1! \dots n_M!} &= \frac{\sqrt{2\pi N + 1}\, N^N \mathrm{e}^{-N}}{\sqrt{(2\pi n_1 + 1) \dots (2\pi n_M + 1)}\, n_1^{n_1} \dots n_M^{n_M} \mathrm{e}^{-n_1 - \dots - n_M}} \\ &= \left(\sqrt{2\pi N + 1}\right)^{1-M} \frac{1}{\sqrt{h(x_1) \dots h(x_M)} \left(x_1^{x_1} \dots x_M^{x_M}\right)^N} \,, \end{aligned} \tag{7}$$

where we have introduced

$$x_i = \frac{n_i}{N}\,; \quad h(x_i) = \frac{2\pi n_i + 1}{2\pi N + 1} = \frac{x_i + 1/2\pi N}{1 + 1/2\pi N} \,. \tag{8}$$

Substituting this into (4) and replacing the sum by an integral, we have

$$C_l(N,M) = N^{M-1} \left(\sqrt{2\pi N + 1}\right)^{(1-M)l} X \,, \tag{9}$$

[1] The observed correlations can often be ascribed to the fact that the total multiplicity distribution differs from the Poisson one. Here we are talking about the correlations for a *fixed* total multiplicity.

[2] We use $\sqrt{2\pi N + 1}$ instead of traditional $\sqrt{2\pi N}$. This gives the correct limit for $N \to 0$ and thus represents a much better approximation at small $N$.

where

$$X=\int dx_1 \ldots dx_M \delta(x_1+\ldots+x_M-1)\,(h(x_1)\ldots h(x_M))^{-l/2}\left(\frac{p_1^{x_1}\cdots p_M^{x_M}}{x_1^{x_1}\ldots x_M^{x_M}}\right)^{Nl}. \tag{10}$$

One should keep in mind, however, that this replacement of the sum by the integral can be justified only if all $p_i$ are finite, different from 0. In the case when one of $p_i$ vanishes, the corresponding sum contains still one term ($n_i = 0$), whereas the integral vanishes.

If all $p_i's$ are finite, the integral (10) can be evaluated by the saddle point method. To this end we first perform one integration (over $x_M$)

$$X=\int dx_1 \ldots dx_{M-1}\,(h(x_1)\ldots h(x_{M-1})h(y))^{-l/2}\left(\frac{p_1^{x_1}\cdots p_{M-1}^{x_{M-1}}p_M^y}{x_1^{x_1}\ldots x_{M-1}^{x_{M-1}}y^y}\right)^{Nl}, \tag{11}$$

where now

$$y = 1 - x_1 - \ldots - x_{M-1}\,. \tag{12}$$

We now write the integral in the form

$$X=\int dx_1 \ldots dx_{M-1}\,(h(x_1)\ldots h(x_{M-1})h(y))^{-l/2}\exp\left(\Phi(x_1,\ldots,x_{M-1})\right)\,, \tag{13}$$

and search for a maximum of $\Phi$. From (13) we have

$$\Phi(x_1,\ldots,x_{M-1}) = -Nl\left(\sum_{i=1}^{M-1} x_i\log\left(\frac{x_i}{p_i}\right)+y\log\left(\frac{y}{p_M}\right)\right)\,. \tag{14}$$

The saddle-point condition (vanishing of the first derivatives) gives

$$\frac{\partial\Phi}{\partial x_i} = -Nl\log\left(\frac{x_i}{p_i}\right)+Nl\log\left(\frac{y}{p_M}\right)=0\,. \tag{15}$$

The solution is

$$x_i = p_i. \tag{16}$$

The second derivatives are

$$\begin{aligned}\frac{\partial^2\Phi}{(\partial x_i)^2} &= -\frac{Nl}{x_i}-\frac{Nl}{y} = -\frac{Nl}{p_i}-\frac{Nl}{p_M}\,;\\ \frac{\partial^2\Phi}{\partial x_i\partial x_j} &= -\frac{Nl}{y} = -\frac{Nl}{p_M}\,.\end{aligned} \tag{17}$$

As long as all $p_i's$ are finite, we can thus approximate the integral (13) by

$$\begin{aligned} X &= (h(p_1)\dots h(p_M))^{-l/2} \int dx_1 \dots dx_{M-1} \\ &\quad \exp\left( -\frac{Nl}{2p_M} \left[ \sum_{i=1}^{M-1} \frac{p_M}{p_i}(x_i - p_i)^2 + \sum_{i,j=1}^{M-1} (x_i - p_i)(x_j - p_j) \right] \right) \\ &= (h(p_1)\dots h(p_M))^{-l/2} (2\pi p_M)^{(M-1)/2} (Nl)^{(1-M)/2} \frac{1}{\sqrt{D_M}} , \end{aligned} \quad (18)$$

where $D_M$ is the determinant of the $(M-1)$-dimensional matrix

$$D_{ij} = \frac{p_M}{p_i}\delta_{ij} + d_{ij} , \quad (19)$$

and $d_{ij}$ is the matrix with all elements equal to 1. $D_M$ can be calculated:

$$D_M = \frac{(p_M)^{M-1}}{p_1 \dots . p_M} , \quad (20)$$

so that

$$X = (h(p_1)\dots h(p_M))^{-l/2} (p_1 \dots p_M)^{1/2} (2\pi)^{(M-1)/2} (Nl)^{(1-M)/2} . \quad (21)$$

Introducing this into (9) we obtain

$$C_l(N, M) = (2\pi N + 1)^{(l-1)/2} [(2\pi\rho_1 + 1)\dots(2\pi\rho_M + 1)]^{(1-l)/2} \Omega' , \quad (22)$$

where

$$\begin{aligned} \Omega' &\equiv \left( \frac{(2\pi N + 1)(2\pi\rho_1)\dots(2\pi\rho_M)}{l^{M-1}(2\pi N)(2\pi\rho_1 + 1)\dots(2\pi\rho_M + 1)} \right)^{1/2} \\ &= \left( \frac{(2\pi N + 1)(2\pi\rho_1/l)\dots(2\pi\rho_M/l)}{(2\pi N/l)(2\pi\rho_1 + 1)\dots(2\pi\rho_M + 1)} \right)^{1/2} , \end{aligned} \quad (23)$$

and where

$$\rho_i \equiv Np_i , \quad (24)$$

are the *average* values of the multiplicity (*i.e.*, particle density) per bin.

**3.** The formula we have obtained is valid when $N$ is large and all $p_i's$ are finite, which guarantees that also all $\rho_i's$ are large. It is, however, inapplicable when one of the $p_i's$ is very small or if it vanishes. This is best seen by observing that for $l = 1$ we should have $C_1 = 1$, *i.e.*, $\Omega' = 1$, which is badly violated when some of $p_i's$ vanish. This is the consequence of the error we have made when replacing the sum (4) by the integral.

The form of the Eq. (23) suggests that the simplest way to correct for this error is to replace $\Omega'$ in (22) by

$$\Omega \equiv \left( \frac{(2\pi N + 1)(2\pi\rho_1/l + 1)\ldots(2\pi\rho_M/l + 1)}{(2\pi N/l + 1)(2\pi\rho_1 + 1)\ldots(2\pi\rho_M + 1)} \right)^{1/2} . \tag{25}$$

This prescription satisfies several natural constraints:

*(i)* It is correct in the limit $N \to \infty$ and all $p_i's$ finite, which is of course the fundamental requirement;

*(ii)* It guarantees $C_1 = 1$;

*(iii)* If some number, say $M_0$, of $p_i's$ vanishes, the formula for $C_l(N; M)$ reduces automatically to the formula for $C_l(N; M - M_0)$, as it should;

*(iv)* It satisfies the constraints $C_l = 1$ for M=1 and $C_l = 1$ for $N = 0$.

Accepting this we thus finally have the Renyi entropies:

$$H_l(N, M) = \frac{1}{1-l} \log C_l(N, M) = \frac{1}{2} \sum_{i=1}^{M} \log(2\pi\rho_i + 1) - \frac{1}{2} \log(2\pi N + 1)$$
$$- \frac{1}{2(l-1)} \left[ \log \left( \frac{2\pi N + 1}{2\pi N/l + 1} \right) + \sum_{i=1}^{M} \log \left( \frac{2\pi\rho_i/l + 1}{2\pi\rho_i + 1} \right) \right] . \tag{26}$$

In the limit $l \to 1$ we obtain the Shannon entropy:

$$\begin{aligned} S(N, M) \ &= \ \frac{1}{2} \sum_{i=1}^{M} \log(2\pi\rho_i + 1) - \frac{1}{2} \log(2\pi N + 1) \\ &\quad - \frac{\pi N}{2\pi N + 1} + \sum_{i=1}^{M} \left( \frac{\pi\rho_i}{2\pi\rho_i + 1} \right) . \end{aligned} \tag{27}$$

This completes the derivation[3].

**4.** Our final formulae (26) and (27) were derived in the limit (5) of large multiplicities. To see how this approximation works we compare in the Fig. 1, $H_2$ calculated from (26) (for the case of 4 bins) with their exact

[3] Note that now, after replacing $\Omega'$ by $\Omega$, $\rho$ in (26) and (27) can be taken arbitrarily small.

values (obtained by direct numerical evaluation). One sees that when all $p_i^\prime s$ are equal to each other, the asymptotic formula is very accurate, even down to small values of the multiplicity. When one of the $p_i^\prime s$ is much smaller than others, the formula interpolates between $M$ and $M-1$. The approximation is slightly worse but the error never exceeds a few percent. We have checked that the same is true also for the Shannon entropy.

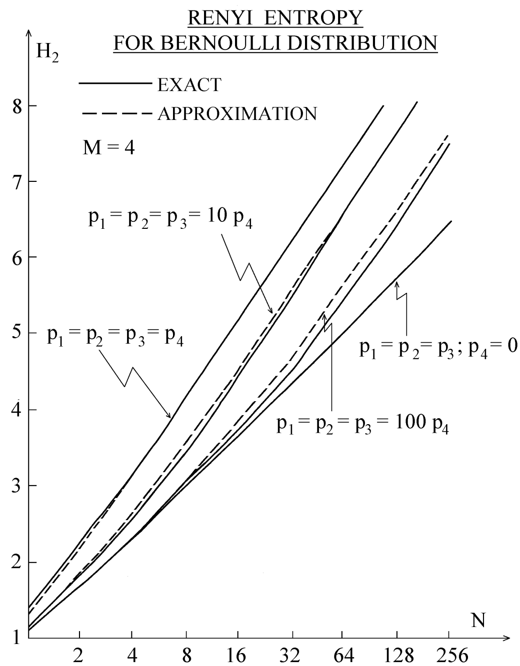


Fig. 1. Renyi entropy $H_2$ calculated for Bernoulli distribution and compared with the approximation given by Eq. (26). In the $p_1 = p_2 = p_3$; $p_4 = 0$ case, the exact and approximate curves are practically indistinguishable (as in the $p_1 = p_2 = p_3 = p_4$ case).

We conclude that the Eqs. (26) and (27) represent a good approximation to the actual values of the Renyi and Shannon entropies and may thus be used as their reliable estimates.

**5.** The first observation from the formulae (26) and (27) is that the entropy of the system is (apart from an additive constant) a sum of the contributions from individual bins. This is the reflection of the property of additivity: different bins may be considered as quasi-independent statistical systems. For a system with the same average number of particles in each bin ($\rho_i = \rho$) this implies linear dependence of entropy on number of bins and thus its linear dependence on the total number of particles. We thus find a normal situation, expected for weakly correlated systems. It should be emphasized, however, that the proportionality coefficient is not universal but depends on $\rho$.

On the other hand *for a fixed number of bins M* the dependence of the entropy on the average number of particles is rather different. In this case a change in the total number of particles implies change in the particle density $\rho$. At large $N$, and if the particle density is not too small ($2\pi\rho \gg 1$), a linear increase of the entropy with the *logarithm* of the number of particles is expected. However, in the very low density limit ($2\pi\rho \ll 1$) the entropy, $S$, becomes

$$S \approx 2\pi N - \tfrac{1}{2}\log(2\pi N + 1) - \tfrac{1}{2} + O(1/N)\,. \tag{28}$$

Thus we recover now a universal linear dependence of the leading term in $S$ on the number of particles, $N$ (for large $N$). For such a situation to occur, the number of bins, $M$, must be indeed very large, to insure that the density $\rho = N/M$ is small enough[4].

These observations show that the interpretation of the experimental measurements requires rather careful specification of their conditions. In the particular case we consider, one sees that the measurement of (Renyi) entropy at a fixed particle density (changing the number of bins) and the measurement at a fixed number and size of the bins (changing the particle density) provide entirely independent information. The first one tests the independence of the particle distribution in different bins. The second one tests to what an extent the mechanism of particle production depends on the density of the produced particles.

**6.** In conclusion, we have derived a formula which gives coincidence probabilities for an entirely random distribution of particles at a given total multiplicity. This formula predicts a rather simple dependence of Renyi entropies on the number of particles in the phase-space region considered: they are linear in $\log N$ (with the exception of the *very* low densities). It also gives a linear dependence on the number of bins taken for the analysis, reflecting the additivity of entropy for the weakly correlated systems. Since the Bernoulli distribution provides a fundamental building block for many models of particle production, it would be interesting to see how this compares with the data.

[4] It is possible — though we do not have a proof — that the linear dependence of the leading term of $S$ on $\log N$ characterizes systems at all densities except the very low ones ($\rho_i << 0.1$!!). To support this claim we may quote our previous result from [2]: The entropy of a system of Bosons condensing in the lowest (but discrete!) state acquires the leading $\log N$ dependence.

This investigation was supported in part by the Subsydium of Foundation for Polish Science 1/99.

## REFERENCES

[1] A. Bialas, W. Czyz, J. Wosiek, *Acta Phys. Pol.* **B30**, 107 (1999); A. Bialas, W. Czyz, *Phys. Rev.* **D61**, 074021 (2000); A. Bialas, W. Czyz, *Acta Phys. Pol.* **B31**, 687 (2000).

[2] A. Bialas, W. Czyz, *Acta Phys. Pol.* **B31**, 2803 (2000).

[3] A. Renyi,*On Measures of Entropy and Information*, in Proc. 4-th Berkeley Symp.. on Math. Stat. Prob. 1960, Vol.1, University of California Press, Berkeley, Los Angeles 1961, p.547.

[4] See, *e.g.*, L. Stodolski, *Phys. Rev. Lett.* **28** (1972) 60.