On Discrete Distributions: Uniform, Monotone, Multivariate

M.C. Jones
Department of Mathematics and Statistics, The Open University, U.K.

August 3, 2018

Abstract

In this partly expository article, I am concerned with some simple yet fundamental aspects of discrete distributions that are either uniform or have monotone probability mass functions. In the univariate case, building on work of F.W. Steutel published in 1988, I look at Khintchine's theorem for discrete monotone distributions in terms of mixtures of discrete uniform distributions, along with their 'stronger' subset of discrete α-monotone distributions. In the multivariate case, I develop a new general family of multivariate discrete distributions with uniform marginal distributions associated with continuous copulas and consider families of multivariate discrete distributions with monotone marginals associated with these.

Keywords: α-monotonicity; Khintchine's theorem; multivariate uniform distribution

1 Introduction

In this partly expository article, I am concerned with some simple yet fundamental aspects of distributions on $\mathbb{N}_0 \equiv 0, 1, \ldots,$ whose probability mass functions (p.m.f.'s) p are either uniform or more generally monotone nonincreasing, together with certain extensions of these distributions to $\mathbb{N}_0^d \equiv \mathbb{N}_0 \times \cdots \times \mathbb{N}_0$, especially $\mathbb{N}_0^2$, and subsets thereof. As a prime example of a univariate distribution with a non-uniform monotone nonincreasing p.m.f. — a 'monotone p.m.f.' for short — think of the geometric distribution.

The main points to be considered in this article, by section, are:

§2 Khintchine's theorem for monotone distributions on $\mathbb{N}_0$, re-interpreted in terms of mixtures of discrete uniform distributions, and a consequent variance inequality for univariate discrete monotone distributions;

§3 a general family of multivariate discrete distributions with uniform marginal distributions associated in an attractive yet novel way with continuous copulas;

§4 families of multivariate discrete distributions with monotone marginals associated with the multivariate uniform distributions of Section 3;

§5 univariate α-monotone distributions on $\mathbb{N}_0$ which, for $0 < \alpha \leq 1$, are a 'stronger' subset of monotone distributions. Originally introduced by Steutel (1988), I pursue further interpretation and properties.

All mathematical manipulations made in this article have the major benefit of being simple and direct.

As I go along, it will often be useful to point out analogies and connections with results for continuous data which have uniform or monotone probability density functions (p.d.f.'s) f on $\mathbb{R}^+$, and their multivariate extensions.

2 Discrete Khintchine's Theorem

Let f be a monotone p.d.f. on $\mathbb{R}^+$. Then, the renowned Khintchine's Theorem (Khintchine, 1938, Feller, 1971) says that $X \sim f$ can be written as a uniform scale mixture, either as

$X = UY$, where U and Y are independent, $U \sim \text{Uniform}(0, 1)$ and $Y \sim g$ for some p.d.f. g on $\mathbb{R}^+$, or equivalently as $X|Y = y \sim \text{Uniform}(0, y)$, $Y \sim g$. If f is differentiable, then typically $g(x) = -xf'(x)$. (The distribution of Y is slightly different if f has support $(0, b)$ say, when $b < \infty$ and $f(b) > 0$; see Section 5.)

Implicit in Steutel's (1988) paper on "discrete α-monotonicity" — of which, more in Section 5 — is a corresponding result to Khintchine's theorem in the discrete case. It is framed in terms of binomial thinning, as first proposed by Steutel and van Harn (1979). Let $N \sim p$ and $M \sim q$ on $\mathbb{N}_0$. For values of $\theta \in [0, 1]$, p is the binomially thinned version of q if

$$N \equiv \theta \circ M \equiv \sum_{j=1}^{M} B_j \tag{1}$$

where the sum is understood to be zero if $M = 0$. Here, $B_1, ..., B_M$ are Bernoulli(θ) random variables independent of each other and of M. (Note that if $\theta = 1$, $N = M$ and if $\theta = 0$, $N = 0$.) A useful equivalent way of expressing $\theta \circ M$ is as

$$N|M = m \sim \text{Binomial}(m, \theta), \qquad M \sim q, \tag{2}$$

where Binomial$(0, \theta)$ is understood to be the degenerate distribution at zero.

The above is binomial thinning for fixed θ, an extension to which is to mix over a distribution for its random variable version, Θ. So, consider the distribution of $N = \Theta \circ M$ where $\Theta \sim h$ on $(0, 1)$, independently of $M \sim q$. This distribution can be expressed as

$$N|M = m \sim \text{BinMix}(m), \qquad M \sim q, \tag{3}$$

with the binomial mixture distribution 'BinMix' defined as follows: $N_m \equiv \Theta \circ m \sim$ BinMix(m) if

$$N_m|\Theta = \theta \sim \text{Binomial}(m, \theta), \qquad \Theta \sim h. \tag{4}$$

Steutel's (1988) observation is that taking $\Theta \sim \text{Uniform}(0, 1)$ is equivalent to p being a monotone p.m.f. on $\mathbb{N}_0$. I now note that in that case, where $h(\theta) = I(0 < \theta < 1)$,

$$\Theta \circ M \sim \text{Uniform}(0, 1, ..., m),$$

that is, the binomial mixture distribution reduces to the discrete uniform distribution on $0, 1, ..., m$. To see this, note that, for each $x = 0, 1, ..., m$,

$$\int_0^1 \binom{m}{x} \theta^x (1-\theta)^{m-x}\, d\theta = \binom{m}{x} B(x+1, m-x+1) = \frac{1}{m+1}$$

(here, $B(\cdot, \cdot)$ is the beta function).

The discrete analogue of Khintchine's theorem can therefore be given most simply — and not unexpectedly given its continuous analogue — as a discrete uniform mixture, as in Result 2.1.

`Result 2.1`

A p.m.f. p on $\mathbb{N}_0$ is monotone if and only if $N \sim p$ can be written as

$$N|M = m \sim \text{Uniform}(0, 1, ..., m), \qquad M \sim q,$$

where q is any p.m.f. on $\mathbb{N}_0$. In fact, p and q are related by

$$p(n) = \sum_{m=n}^{\infty} \frac{q(m)}{m+1}, \qquad q(m) = (m+1)\, \{p(m) - p(m+1)\}\,. \tag{5}$$

`Example 2.1`

(a) Let $N \sim \text{Geometric}(p)$, $0 < p < 1$, which has strictly decreasing p.m.f. In this case,

$$q(m) = (m+1)\, p^2 (1-p)^m,$$

that is, $M \sim \text{NegativeBinomial}(2, p)$, which is the distribution of the sum of two independent Geometric(p) random variables.

(b) Let $N \sim \text{Poisson}(\mu)$ with $0 < \mu \leq 1$. Then, p is monotone on $\mathbb{N}_0$, and Result 2.1 applies with

$$q(m) = (m + 1 - \mu)\, p(m).$$

One of a number of ways of interpreting q is that it is the distribution of $M_0 + B$ where $B \sim \text{Bernoulli}(\mu)$, independent of $M_0 \sim \text{Poisson}(\mu)$.

(c) Now let $M \sim \text{Poisson}(\lambda)$, $\lambda > 0$. Then, N has the strictly decreasing p.m.f.

$$p(n) = \frac{e^{-\lambda}}{\lambda} \sum_{j=n+1}^{\infty} \frac{\lambda^j}{j!} = \frac{1}{\lambda} \Gamma(\lambda; n+1)$$

where $\Gamma(\cdot;\cdot)$ is the incomplete gamma function ratio. From (6) below, $\mathbb{E}(N) = \frac{1}{2}\lambda$ and $\mathbb{V}(N) = \frac{1}{2}\lambda + \frac{1}{12}\lambda^2$, so p is overdispersed as well as decreasing.

(d) The distribution of part (c) is a special case of taking $q(m) = (m+1)\, r(m+1)/\mu_r$ where r is an arbitrary p.m.f. on $\mathbb{N}_0$ with finite mean μ_r. Then, $p(n) = \overline{R}(n)/\mu_r$ where $R(n) = P(R > n)$ and $R \sim r$, so p is clearly monotone.

(e) There is no distribution satisfying $p = q$. If there were, p must satisfy $p(m+1)/p(m) = m/(m+1)$, $m = 0, 1, ...$, and this was shown by L. Katz in the 1940s not to correspond to a valid distribution (see Johnson, Kemp and Kotz, 2005, Section 2.3.1).

Either as a consequence of more general results for mixed binomial thinning or directly, it is easy to show that

$$\mathbb{E}(N) = \mathbb{E}(M)/2, \qquad \mathbb{V}(N) = \left[4\mathbb{V}(M) + 2\mathbb{E}(M) + \{\mathbb{E}(M)\}^2\right]/12. \tag{6}$$

Since $\mathbb{V}(M) \geq 0$ and $\mathbb{E}(M) = 2\mathbb{E}(N)$, the following variance-mean inequality arises.

`Result 2.2`

If N follows a monotone p.m.f. on $\mathbb{N}_0$, then

$$\mathbb{V}(N) \geq \mathbb{E}(N)\,\{1 + \mathbb{E}(N)\}/3, \tag{7}$$

and any monotone distribution is overdispersed if $\mathbb{E}(N) > 2$.

This inequality and observation arose in Jones and Marchand (2018) from a different perspective. The inequality is the discrete analogue of the inequality $\mathbb{V}(X) \geq \{\mathbb{E}(X)\}^2/3$ of Johnson and Rogers (1951) in the continuous monotone case.

3 Multivariate Discrete Uniform Distributions

Write c and C for the p.d.f. and cumulative distribution function (c.d.f.) of a copula on $(0,1)^d$ (e.g. Nelsen, 2006, Joe, 1997, 2014). This section and the next can be seen as an investigation of a role for such multivariate continuous uniform distributions in providing the dependence properties of certain multivariate discrete distributions, starting in this section with multivariate discrete distributions with discrete uniform marginal distributions, or multivariate discrete uniform distributions for shorter. Note that this is quite different from the use of a (continuous) copula in conjunction with the discontinuous c.d.f.'s and quantile functions of discrete marginals, a common practice but with a number of "dangers and limitations", as discussed by Genest and Nešlehová (2007). That said, a multivariate discrete uniform distribution does *not* fulfil the same role for multivariate discrete distributions as a copula does for multivariate continuous distributions because discrete c.d.f.'s are not distributed as discrete uniforms (on the other hand, continuous c.d.f.'s are distributed as continuous uniforms).

The fact that a binomial distribution mixed over a continuous uniform distribution for its probability parameter is itself a discrete uniform distribution suggests that a multivariate discrete uniform distribution can be defined as the distribution of $N_1, ..., N_d$ on $(0, 1, ..., m_1) \times \cdots \times (0, 1, ..., m_d)$ such that

$$\begin{aligned} N_i|\Theta_i = \theta_i \quad &\sim \quad \text{Binomial}(m_i, \theta_i) \quad \text{independently for} \quad i = 1, ..., d, \\ \Theta^{(d)} \equiv \{\Theta_1, ..., \Theta_d\} \quad &\sim \quad c(\theta_1, ..., \theta_d). \end{aligned}$$

The joint p.m.f. of $N_1, ..., N_d$ is

$$\begin{aligned} p_U(n_1, ..., n_d \quad | \quad m_1, ..., m_d) = &\left\{ \prod_{i=1}^d \binom{m_i}{n_i} \right\} \\ \times \quad &\int_0^1 \cdots \int_0^1 \left\{ \prod_{i=1}^d \theta_i^{n_i} (1-\theta_i)^{m_i - n_i} \right\} c(\theta_1, ..., \theta_d)\, d\theta_1 ... d\theta_d. \end{aligned} \tag{8}$$

Its univariate marginal distributions are discrete uniform by construction because those of the copula are continuous uniform.

Moments of this construction are readily available and, in particular, correlations are determined by those of the copula as follows. Since $\text{Cov}(N_i, N_j|\Theta^{(d)} = \theta^{(d)}) = 0$, it is the case that

$$\begin{aligned}\mathrm{Cov}(N_i, N_j) &= \mathrm{Cov}\{\mathbb{E}(N_i|\Theta^{(d)} = \theta^{(d)}), \mathbb{E}(N_j|\Theta^{(d)} = \theta^{(d)})\} \\ &= m_i m_j \mathrm{Cov}(\Theta_i, \Theta_j). \qquad (9)\end{aligned}$$

Also, since $\mathbb{V}(N_i) = m_i(m_i + 2)/12$, $\mathbb{V}(N_j) = m_j(m_j + 2)/12$, it is the case that

$$\begin{aligned}\mathrm{Corr}(N_i, N_j) &= \frac{m_i m_j \mathrm{Corr}(\Theta_i, \Theta_j)/12}{\sqrt{m_i(m_i+2)m_j(m_j+2)}/12} \\ &= \sqrt{\frac{m_i}{m_i+2}}\sqrt{\frac{m_j}{m_j+2}}\,\mathrm{Corr}(\Theta_i, \Theta_j). \qquad (10)\end{aligned}$$

So, while the correlation of N_i and N_j has the same sign as that of Θ_i and Θ_j, it reduces to one-third that of the copula in the binary case, and increases, tending to a factor of one, as the marginal supports grow larger. Note that $\mathrm{Corr}(\Theta_i, \Theta_j)$ is Spearman's rho.

The existence of this simple relationship between discrete and continuous uniform correlations is a strong reason for preferring the current construction to discretisations of the copula, though the two can be very similar, as the following simple example shows.

`Example 3.1`

Consider the bivariate Farlie–Gumbel–Morgenstern (FGM) copula given by

$$C(u, v) = uv\{1 + \phi(1-u)(1-v)\}, \qquad c(u, v) = 1 + \phi(1-2u)(1-2v),$$

on $0 < u, v < 1$ with $-1 \leq \phi \leq 1$. Entering this into (8) when $d = 2$ gives

$$p_{FGM}(n_1, n_2) = \frac{1}{(m_1+1)(m_2+1)}\left\{1 + \phi\,\frac{(2n_1 - m_1)(2n_2 - m_2)}{(m_1+2)(m_2+2)}\right\};$$

its correlation, from (10) and e.g. Example 2.4 of Joe (1997), is

$$\sqrt{\frac{m_1}{m_1+2}}\sqrt{\frac{m_2}{m_2+2}}\,\frac{\phi}{3}.$$

A natural discretisation of any C in the bivariate case is

$$\begin{aligned}p'(n_1, n_2) &= C\left(\frac{n_1+1}{m_1+1}, \frac{n_2+1}{m_2+1}\right) + C\left(\frac{n_1}{m_1+1}, \frac{n_2}{m_2+1}\right) \\ &\quad - C\left(\frac{n_1+1}{m_1+1}, \frac{n_2}{m_2+1}\right) - C\left(\frac{n_1}{m_1+1}, \frac{n_2+1}{m_2+1}\right)\end{aligned}$$

which turns out in the FGM case to equate to

$$p'_{FGM}(n_1, n_2) = \frac{1}{(m_1+1)(m_2+1)} \left\{ 1 + \phi \frac{(2n_1 - m_1)(2n_2 - m_2)}{(m_1+1)(m_2+1)} \right\}; \tag{11}$$

this differs just a little from p_{FGM}. The correlation associated with this model, calculated directly from (11), is similar to that of p_{FGM}, but a little larger; it is

$$\sqrt{\frac{m_1(m_1+2)}{(m_1+1)^2}} \sqrt{\frac{m_2(m_2+2)}{(m_2+1)^2}} \frac{\phi}{3}.$$

4 Multivariate Discrete Distributions with Monotone Univariate Marginals

Combining Sections 2 and 3 further, it is now natural to develop discrete distributions on $\mathbb{N}_0^d$ with monotone univariate marginals as the distribution of $N^{(d)} \equiv \{N_1, ..., N_d\}$ where

$$\begin{aligned} N_i | M_i = m_i, \Theta_i = \theta_i &\sim \text{Binomial}(m_i, \theta_i) \quad \text{independently for} \quad i = 1, ..., d, \\ M^{(d)} \equiv \{M_1, ..., M_d\} &\sim q(m_1, ..., m_d), \\ \Theta^{(d)} \equiv \{\Theta_1, ..., \Theta_d\} &\sim c(\theta_1, ..., \theta_d), \end{aligned}$$

where q is now an arbitrary p.m.f. on $\mathbb{N}_0^d$ and $M^{(d)}$ is independent of $\Theta^{(d)}$. This is, of course, equivalent to mixing the multivariate discrete uniform distribution of Section 3 over q:

$$N | M^{(d)} = \{m_1, ..., m_d\} \sim p_U(n_1, ..., n_d | m_1, ..., m_d), \qquad M^{(d)} \sim q(m_1, ..., m_d).$$

The joint p.m.f. of $N^{(d)}$ is

$$\begin{aligned} p_D(n_1, ..., n_d) &= \sum_{m_1=n_1}^{\infty} \cdots \sum_{m_d=n_d}^{\infty} q(m_1, ..., m_d) \left\{ \prod_{i=1}^{d} \binom{m_i}{n_i} \right\} \\ &\quad \times \int_0^1 \cdots \int_0^1 \left\{ \prod_{i=1}^{d} \theta_i^{n_i} (1-\theta_i)^{m_i - n_i} \right\} c(\theta_1, ..., \theta_d) \, d\theta_1 \dots d\theta_d. \end{aligned} \tag{12}$$

Its univariate marginal distributions have the monotone p.m.f.'s of Result 2.1 by construction. Moments remain readily available and correlations are as follows. Using (6) and (9),

$$\text{Cov}(N_i, N_j) = \{\text{Cov}(M_i, M_j) + \mathbb{E}(M_i)\mathbb{E}(M_j)\}\text{Cov}(\Theta_i, \Theta_j) + \tfrac{1}{4}\text{Cov}(M_i, M_j)$$

so that

$$\mathrm{Corr}(N_i, N_j) = \frac{\{\mathrm{Cov}(M_i, M_j) + \mathbb{E}(M_i)\mathbb{E}(M_j)\}\mathrm{Corr}(\Theta_i, \Theta_j) + 3\mathrm{Cov}(M_i, M_j)}{\sqrt{\{4\mathbb{V}(M_i) + 2\mathbb{E}(M_i) + \{\mathbb{E}(M_i)\}^2\}\{4\mathbb{V}(M_j) + 2\mathbb{E}(M_j) + \{\mathbb{E}(M_j)\}^2\}}}. \tag{13}$$

In the following two subsections, I will take a brief look at two major particular cases of this in terms of the form of distribution for M. These distributions and their properties are analogues of those given in Section 3 of Bryson and Johnson (1982) in the continuous case when $d = 2$.

4.1 When $M_1, ..., M_d$ are Independent

Let $M_i \sim q_i$, independently for $i = 1, ..., d$. This allows the dependence structure of p_U to depend only on that of C. The joint p.d.f. of $N^{(d)}$ is given by the obvious small change to (12). The correlation of N_i and N_j, given by (13), reduces to

$$\mathrm{Corr}(N_i, N_j) = \sqrt{\frac{\mathbb{E}(M_i)}{4\mathbb{D}(M_i) + \mathbb{E}(M_i) + 2}}\sqrt{\frac{\mathbb{E}(M_j)}{4\mathbb{D}(M_j) + \mathbb{E}(M_j) + 2}}\,\mathrm{Corr}(\Theta_i, \Theta_j)$$

where $\mathbb{D}(M) = \mathbb{V}(M)/\mathbb{E}(M)$ is the index of dispersion of M. Again, this has the same sign as the correlation associated with the copula and is always a reduction of the absolute value of the correlation compared with that of the copula, sometimes considerably so. This, in turn, reduces to

$$r_{ij} \equiv \mathrm{Corr}(N_i, N_j) = \frac{\mathbb{E}(M)}{4\mathbb{D}(M) + \mathbb{E}(M) + 2}\,\mathrm{Corr}(\Theta_i, \Theta_j) \tag{14}$$

when M_i and M_j have the same distribution (that of M, say).

`Example 4.1`

Following Example 2.1(a), let $q_i(m) = (m+1)\,p_i^2(1-p_i)^m$ with $\mathbb{E}(M_i) = 2(1-p_i)/p_i$ and $\mathbb{V}(M_i) = 2(1-p_i)/p_i^2$, $i = 1, ..., d$. The corresponding family of multivariate geometric distributions has joint p.m.f.

$$p_G(n_1, \ldots, n_d) = \left\{\prod_{i=1}^{d}(n_i+1)p_i^2(1-p_i)^{n_i}\right\}$$
$$\times \int_0^1 \cdots \int_0^1 \left[\prod_{i=1}^{d}\frac{\theta_i^{n_i}}{\{1-(1-p)(1-\theta_i)\}^{n_i+2}}\right] c(\theta_1, ..., \theta_d)\, d\theta_1 \ldots d\theta_d$$

and correlations

$$\mathrm{Corr}(N_i, N_j) = \frac{1}{3}\sqrt{(1-p_i)(1-p_j)}\,\mathrm{Corr}(\Theta_i, \Theta_j).$$

The correlations associated with this family of multivariate geometric distributions are therefore limited to the range $-\frac{1}{3} < \mathrm{Corr}(N_i, N_j) < \frac{1}{3}$.

4.2 When $M_1, ..., M_d$ are Equal

Let $M_1 = \cdots = M_d = M$ say, $i = 1, ..., d$, with $M \sim q_0$. This again allows the dependence structure of p_D to depend only on that of C, but with an opportunity for higher correlations. This comes at the expense, however, of having to have identical univariate marginal distributions. Let $n_{\max} = \max(n_1, ..., n_d)$. The joint p.d.f. of $N^{(d)}$ is given by

$$\begin{aligned} p_D(n_1, ..., n_d) &= \sum_{m=n_{\max}}^{\infty} q_0(m) \left\{ \prod_{i=1}^{d} \binom{m}{n_i} \right\} \\ &\quad \times \int_0^1 \cdots \int_0^1 \left\{ \prod_{i=1}^{d} \theta_i^{n_i} (1-\theta_i)^{m-n_i} \right\} c(\theta_1, ..., \theta_d)\, d\theta_1 \ldots d\theta_d. \end{aligned}$$

Its correlations are, from (13),

$$\rho_{ij} \equiv \mathrm{Corr}(N_i, N_j) = \frac{[\mathbb{D}(M) + \mathbb{E}(M)]\mathrm{Corr}(\Theta_i, \Theta_j) + 3\mathbb{D}(M)}{4\mathbb{D}(M) + \mathbb{E}(M) + 2}. \qquad (15)$$

Comparing (14) and (15), it is found that

$$\rho_{ij} = r_{ij} + \frac{\mathbb{D}(M)\{\mathrm{Corr}(\Theta_i, \Theta_j) + 3\}}{4\mathbb{D}(M) + \mathbb{E}(M) + 2} > r_{ij}.$$

`Example 4.2`

While in Sections 3 and 4.1 the independence copula with density $c(\theta_1, ..., \theta_d) = \prod_{i=1}^{d} I(0 < \theta_i < 1)$ results in distributions with independent marginals, this is not the case here because of the commonality of M. In fact, using the independence copula, the joint p.m.f. of $N^{(d)}$ depends only on $n_{\max}$ and is given by

$$p(n_1, ..., n_d) = \sum_{m=n_{\max}}^{\infty} \frac{q_0(m)}{(m+1)^d}.$$

The corresponding correlations are all equal and positive, being determined by moments of M:

$$\mathrm{Corr}(N_i, N_j) = \frac{3\mathbb{D}(M)}{4\mathbb{D}(M) + \mathbb{E}(M) + 2}.$$

If, in an extension to Example 2.1(d), for each d, $q(m) = (m+1)^d\, r(m+1)/\mu_r^{[d]}$ where r is a p.m.f. on $\mathbb{N}_0$ with finite raw d'th moment $\mu_r^{[d]}$, then $p(n_1, ..., n_d) = \overline{R}(n_{\max})/\mu_r^{[d]}$. In this case, $\mathbb{E}(M) = (\mu_r^{[d+1]}/\mu_r^{[d]}) - 1$,

$$\mathbb{V}(M) = \frac{\mu_r^{[d+2]} - 2\mu_r^{[d+1]} + \mu_r^{[d]}}{\mu_r^{[d]}} - \left(\frac{\mu_r^{[d+1]} - \mu_r^{[d]}}{\mu_r^{[d]}}\right)^2$$

and so

$$0 < \mathrm{Corr}(N_i, N_j) = \frac{3\left\{\mu_r^{[d]}\mu_r^{[d+2]} - \left(\mu_r^{[d+1]}\right)^2\right\}}{4\mu_r^{[d]}\mu_r^{[d+2]} - 3\left(\mu_r^{[d+1]}\right)^2 - \left(\mu_r^{[d]}\right)^2} < 1.$$

`Example 4.3`

For a general copula and $M \sim$ NegativeBinomial$(2, p)$, the corresponding family of multivariate distributions with Geometric(p) marginals has correlations

$$\mathrm{Corr}(N_i, N_j) = \frac{1}{2} + \frac{(3-2p)\,\mathrm{Corr}(\Theta_i, \Theta_j)}{6}.$$

In this case too, $0 < \mathrm{Corr}(N_i, N_j) < 1$. In the case of the independence copula as in Example 4.2, $\mathrm{Corr}(N_i, N_j) = 1/2$.

5 Discrete α-Monotonicity

I now return to the univariate domain. To set the scene, I first describe the situation in the continuous case. There, α-monotonicity was introduced by Olshen and Savage (1970) (see also Dharmadhikari and Joag-Dev, 1988, and Bertin, Cuculescu and Theodorescu, 1997): the distribution of a continuous random variable X is said to be α-monotone if and only if the distribution of X^α is monotone. Then, X can be written in the form $X = A_\alpha Y$ say, where $A_\alpha \sim$ Beta$(\alpha, 1)$, independently of $Y \sim g$ on $\mathbb{R}^+$, in a similar manner to Khintchine's theorem; equivalently, $X = U^{1/\alpha}Y$ where $U \sim$ Uniform$(0, 1)$. Clearly $\alpha = 1$ corresponds to ordinary monotonicity, and α-monotone distributions with $\alpha < 1$ are also ordinary monotone.

Providing an alternative view of an equivalent formulation of Abouammoh (1987/1988), Steutel (1988) first put forward discrete α-monotonicity in the following manner: for $0 < \alpha \le 1$, $N \sim p$ is discrete α-monotone if $N = A_\alpha \circ M_\alpha = U^{1/\alpha} \circ M_\alpha$, where $A_\alpha \sim \text{Beta}(\alpha, 1)$, $U \sim \text{Uniform}(0,1)$ and either of these is independent of $M_\alpha \sim q_\alpha$ on $\mathbb{N}_0$. The distribution of N can now be recognized, from Section 2, as being that of

$$N|M_\alpha = m_\alpha \sim \text{BetaBinomial}(m_\alpha, \alpha, 1), \qquad M_\alpha \sim q_\alpha, \tag{16}$$

where the BetaBinomial$(m_\alpha, \alpha, 1)$ distribution has p.m.f.

$$p_{BB1}(x) = \frac{\alpha\, m_\alpha!\, \Gamma(x+\alpha)}{x!\, \Gamma(m_\alpha + \alpha + 1)} \tag{17}$$

on $x = 0, 1, ..., m_\alpha$. This is because now $h(\theta) = \alpha\theta^{\alpha-1} I(0 < \theta < 1)$ in (4) so that the binomial mixture distribution becomes

$$\alpha \int_0^1 \binom{m_\alpha}{x} \theta^{x+\alpha-1}(1-\theta)^{m_\alpha - x}\, d\theta = \alpha \binom{m_\alpha}{x} B(x+\alpha, m_\alpha - x + 1) = p_{BB1}(x).$$

(16) and (17) lead directly to confirmation of Steutel's (1988) formula

$$p(n) = \alpha \frac{\Gamma(n+\alpha)}{n!} \sum_{m=n}^{\infty} \frac{m!\, q_\alpha(m)}{\Gamma(m+\alpha+1)}.$$

Steutel then observes that

$$(n+\alpha)p(n) - (n+1)p(n+1) = \alpha q_\alpha(n) \tag{18}$$

from which it can be concluded that discrete α-monotonicity corresponds to p having the simple property that

$$(n+\alpha)p(n) \ge (n+1)p(n+1).$$

I will now add that (18) can also be written

$$q(n) = (1-\alpha)p(n) + \alpha\, q_\alpha(n) \tag{19}$$

where $q = q_1$ is as at (5) in Result 2.1. To corroborate and interpret (19), an alternative way of expressing α-monotonicity arises from writing $A_\alpha = UV$ where $U \sim \text{Uniform}(0,1)$

independently of some appropriate V. Now, Beta$(\alpha, 1)$ is a distribution on a finite interval with non-zero density at its upper endpoint. As signposted at the start of Section 2, the density of V is not $-xf'(x)$ if f has support $(0,b)$ and $f(b) > 0$; in fact,

$$V \sim \begin{cases} Y & \text{with probability } 1-\alpha, \\ b & \text{with probability } \alpha, \end{cases}$$

where $Y \sim -xf'(x)/\{1-f(b)\}$ on $(0,b)$. When $b=1$ and $h(x) = \alpha x^{\alpha-1}$ so that $h(1) = \alpha$, it turns out that $-xh'(x)/\{1-h(1)\} = h(x)$. In the case of discrete α-monotonicity, it follows that $N = A_\alpha \circ M = (UV) \circ M = U \circ (V \circ M)$ so that $N = U \circ N_0$ where $U \sim$ Uniform$(0,1)$ and

$$N_0 \sim \begin{cases} N & \text{with probability } 1-\alpha, \\ M & \text{with probability } \alpha, \end{cases} \tag{20}$$

which is immediately seen to be equivalent to (19).

By any of a number of routes, it can be shown that, for α-monotone distributions,

$$\mathbb{E}(N) = \frac{\alpha \mathbb{E}(M_\alpha)}{\alpha+1}, \quad \mathbb{V}(N) = \frac{\alpha\left[(\alpha+1)^2\mathbb{V}(M_\alpha) + (\alpha+1)\mathbb{E}(M_\alpha) + \{\mathbb{E}(M_\alpha)\}^2\right]}{(\alpha+1)^2(\alpha+2)}. \tag{21}$$

Since $\mathbb{V}(M_\alpha) \geq 0$ and $\mathbb{E}(M_\alpha) = (\alpha+1)\mathbb{E}(N)/\alpha$, the following variance-mean inequality ensues.

`Result 5.1`

If N follows an α-monotone p.m.f. on $\mathbb{N}_0$, then

$$\mathbb{V}(N) \geq \frac{\mathbb{E}(N)\{\alpha + \mathbb{E}(N)\}}{\alpha(\alpha+2)}.$$

This is essentially Theorem 3.1 of Abouammoh, Ali and Mashhour (1994) with $a = 0$ and Corollary 5.3.21 of Bertin et al. (1997). An α-monotone distribution is thereby guaranteed to be overdispersed if $\mathbb{E}(N) > \alpha(\alpha+1)$. Of course, Result 5.1 reduces to Result 2.2 when $\alpha = 1$.

`Example 5.1`

(a) $N \sim$ Geometric(p) is α-monotone if $p \geq 1 - \alpha$. Using (18), the corresponding p.m.f. of M_α is

$$q_\alpha(m) = \{(m+1)p - (1-\alpha)\}p(1-p)^m/\alpha.$$

The dispersion inequality for α-monotone distributions confirms the overdispersion of the geometric distribution — which actually holds for all $0 < p < 1$ — when $p < 1/(\alpha^2 + \alpha + 1)$, which tends to 1 as $\alpha \to 0$.

(b) Let $N \sim$ Poisson(μ) with $0 < \mu \leq \alpha$. Then, p is α-monotone on $\mathbb{N}_0$, and formula (18) applies to give

$$q_\alpha(m) = (m + \alpha - \mu)\, p(m)/\alpha.$$

Now, q_α is the distribution of $M_0 + B$ where $B \sim$ Bernoulli(μ/α), independent of $M_0 \sim$ Poisson(μ).

(c) Both of the above examples and other monotone binomial and negative binomial distributions are covered by the Katz family, for which

$$(1+n)\, p(n+1) = (a + bn)\, p(n)$$

(Johnson et al., 2005, Section 2.3.1). In general, $a > 0$ and $b < 1$, but monotonicity restricts the parameter ranges to $0 < a \leq 1$ and $-1 < b < 1$; α-monotonicity restricts the parameters further to $0 < a \leq \alpha$ as well as $-1 < b < 1$.

6 References

Abouammoh, A.M. (1987/1988), "On Discrete α-Unimodality," *Statistica Neerlandica*, 41, 239–244. Correction, 42, 141.

Abouammoh, A.M., Ali, A.M., and Mashhour, A.F. (1994), "On Characterizations and Variance Bounds of Discrete α-Unimodality," *Statistical Papers*, 35, 151–161.

Bertin, E.M.J., Cuculescu, I., and Theodorescu, R. (1997), *Unimodality of Probability Measures*, Dordrecht: Kluwer.

Bryson, M.C., and Johnson, M.E. (1982), "Constructing and Simulating Multivariate Distributions Using Khintchine's Theorem," *Journal of Statistical Computation and Simulation*, 16, 129–137.

Dharmadhikari, S., and Joag-Dev, K. (1988), *Unimodality, Convexity, and Applications*, Boston: Academic Press.

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications*, Volume 2, New York: Wiley.

Genest, C., and Nešlehová, J. (2007), "A Primer on Copulas for Count Data," *Astin Bulletin,* 37, 475–515.

Joe, H. (1997), *Multivariate Models and Dependence Concepts*, London: Chapman & Hall.

Joe, H. (2014), *Dependence Modeling with Copulas*, Boca Raton, FL: Chapman & Hall/CRC Press.

Johnson, N.L., Kemp, A.W. & Kotz, S. (2005), *Univariate Discrete Distributions* (3rd ed.), Hoboken, NJ: Wiley.

Johnson, N.L., and Rogers, C.A. (1951), "The Moment Problem for Unimodal Distributions," *Annals of Mathematical Statistics*, 22, 433–439.

Jones, M.C., and Marchand, É. (2018), "Multivariate Discrete Distributions Via Sums and Shares," under consideration.

Khintchine, A.Y. (1938), "On Unimodal Distributions," *Izvestiya Nauchno-Issledovatel'skogo Instituta Matematiki i Mekhaniki*, 2, 1–7.

Nelsen, R.B. (2006), *An Introduction to Copulas* (2nd ed.), New York: Springer.

Olshen, R.A., and Savage, L.J. (1970), "A Generalized Unimodality," *Journal of Applied Probability*, 7, 21–34.

Steutel, F.W. (1988), "Note on Discrete α-Unimodality," *Statistica Neerlandica*, 42, 137–140.

Steutel, F.W., and van Harn, K. (1979), "Discrete Analogues of Self-Decomposability and Stability," *Annals of Probability*, 7, 893–899.