Shifted Dirichlet Distributions as Second-Order Probability Distributions that Factors into Marginals

David Sundgren, Love Ekenberg and Mats Danielson

University of Gävle and Royal Institute of Technology

July 15, 2009

What is a Second-Order Probability Distribution

A second-order probability distribution is meant to express belief in the possible values of an uncertain probability.
So any probability distribution that is defined on the surface defined by $\sum_{i=1}^{n} x_i = 1$ is a second-order distribution if the variables x_i are interpreted as the probabilities of the possible outcomes of an event.

How to Choose?

Which criteria are to be used when chosing a second-order probability distribution? The principle of maximum entropy is a reasonable principle, do not let the distribution express more knowledge than you have. But is this high entropy to be seen on the global level or the local? Is it the multivariate distribution or the marginal distributions that are to have maximum entropy? The distributions under consideration here do not have maximum entropy either locally or globally, the entropy is divided among on the one hand the multivariate distribution and on the other hand its marginals.

How to Choose?

Which criteria are to be used when chosing a second-order probability distribution? The principle of maximum entropy is a reasonable principle, do not let the distribution express more knowledge than you have. But is this high entropy to be seen on the global level or the local? Is it the multivariate distribution or the marginal distributions that are to have maximum entropy? The distributions under consideration here do not have maximum entropy either locally or globally, the entropy is divided among on the one hand the multivariate distribution and on the other hand its marginals.

Total Correlation

We seek second-order distributions that minimise the total correlation of the first-order probabilities, the divergence of the product of marginals from the joint distribution is to be minimal. This is achieved if the random variables are independent, but they are not in our case since the probabilities sum to one. The best we can achieve is to have

$$f(x_1, \ldots, x_n) = \frac{1}{K} \prod_{i=1}^{n} f_i(x_i),$$

where $K = \int_{\mathcal{P}} \prod_{i=1}^{n} f_i(x_i)\, d\mathbf{x}$.

Total Correlation

We seek second-order distributions that minimise the total correlation of the first-order probabilities, the divergence of the product of marginals from the joint distribution is to be minimal. This is achieved if the random variables are independent, but they are not in our case since the probabilities sum to one. The best we can achieve is to have

$$f(x_1, \ldots, x_n) = \frac{1}{K} \prod_{i=1}^{n} f_i(x_i)\,,$$

where $K = \int_{\mathcal{P}} \prod_{i=1}^{n} f_i(x_i)\, d\mathbf{x}$.

Characterisation

Theorem

A probability distribution $f(\mathbf{x})$ factors into marginals if and only if its marginal distributions are

$$f_i(x_i) = \frac{1}{(n-1)\left(1 - \sum_{j=1}^{n} a_j\right)^{\frac{1}{n-1}} (x_i - a_i)^{\frac{n-2}{n-1}}}$$

with support $[a_i, 1 - \sum_{j \neq i} a_j]$, where $\sum_{j=1}^{n} a_j < 1$.

Note that we do not have a choice for the upper bounds. Once the lower bounds are set, the distribution is set.

Characterisation

Corollary

A joint probability distribution function $f(\mathbf{x})$ *on the probability simplex* $\mathcal{P}$ *factors into marginals if and only if*

$$f(x_1, \ldots, x_{n-1}) = \frac{\left(1 - \sum_{i=1}^{n} a_i\right) \prod_{i=1}^{n} f_i(x_i)}{\Gamma^{1-n}\left(\frac{n}{n-1}\right)},$$

where $x_n = 1 - \sum_{i=1}^{n-1} x_i$, $f_i(x_i), i = 1, \ldots, n$ *are the marginal distributions of* f *and* $f_i(x_i) = 0$ *for* $x_i \geq a_i, x_i \leq 1 - \sum_{j \neq i} a_j$.

Proof Sketch

$$\int_{\mathcal{P}} \prod_{i=1}^{3} g_i(x_i)\, d\mathbf{x} = \int_0^1 \int_0^{1-x_1} g_1(x_1) g_2(x_2) g_3(1 - x_1 - x_2)\, dx_2\, dx_1 =$$

$$\int_0^1 g_1(x_1)[g_2 * g_3(1 - x_1)]\, dx_1 = g_1 * g_2 * g_3(1)\,.$$

Say that $f(\mathbf{x})$ factors into marginals then

$f_i(x_i) = \frac{1}{K} f_i(x_i) \mathop{\ast}_{j \neq i} f_j\,(1 - x_i)$,

$$f_2 * f_3 * \cdots * f_n(1 - x_1) = KH(c_1 - x_1)$$
$$f_1 * f_3 * \cdots * f_n(1 - x_2) = KH(c_2 - x_2)$$
$$\vdots$$
$$f_1 * f_2 * \cdots * f_{n-1}(1 - x_n) = KH(c_n - x_n)\,,$$

where c_i is such that $f_i(x_i) = 0$ when $x_i > c_i$.

Proof Sketch

In the Laplace domain,

$$F_2F_3\dots F_n = (Ke^{-(1-c_1)s})/s$$
$$F_1F_3\dots F_n = (Ke^{-(1-c_2)s})/s$$
$$\vdots$$
$$F_1F_2\dots F_{n-1} = (Ke^{-(1-c_n)s})/s$$

and if $f_k = g_k(x_k - a_k)H(x_k - a_k)$ where $f_k(x_k) = 0$ when $x_k < a_k$, g_k has Laplace transform $\left(\frac{K}{s}\right)^{\frac{1}{n-1}}$ leading to

$$f_k(x_k) = \frac{K^{\frac{1}{n-1}}H(x_k - a_k)}{\Gamma\left(\frac{1}{n-1}\right)(x_k - a_k)^{\frac{n-2}{n-1}}}.$$

The upper limit of the support being $1 - \sum_{j\neq i} a_j$ follows from $\mathcal{L}\left\{\mathop{\ast}_{j\neq i} f_j(1 - x_i)\right\} = \frac{Ke^{-s(\sum_{j\neq i} a_j)}}{s}$.

A Generalisation of a Special Case of the Dirichlet Distribution

When the lower bounds are zero, i.e. $a_i = 0$, the joint distribution is a Dirichlet distribution with parameters $\alpha_i = \frac{1}{n-1}$ and the marginal distributions are Beta distributions with parameters $\alpha = \frac{1}{n-1}, \beta = 1$.
Otherwise, if any $a_i > 0$, the distributions are shifted to have support only for $x_i \geq a_i$.

An Example

With $n = 3$, let us take $a_1 = 1/3, a_2 = 1/5$ and $a_3 = 1/8$. Then $1 - \sum_{i=1}^{3} a_i = \frac{120-40-24-15}{120} = \frac{41}{120}$.

$$f_1(x_1) = \frac{1}{2\sqrt{41/120(x_1 - 1/3)}},$$

$$f_2(x_2) = \frac{1}{2\sqrt{41/120(x_2 - 1/5)}}$$

and

$$f_3(x_3) = \frac{1}{2\sqrt{41/120(x_3 - 1/8)}},$$

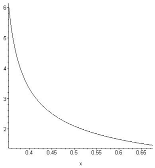
with support $[1/3, 27/40]$, $[1/5, 13/24]$ and $[1/8, 7/15]$ and mean $\frac{161}{360}$, $\frac{113}{360}$ and $\frac{43}{180}$, respectively.

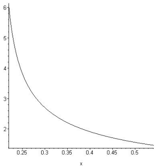
An Example

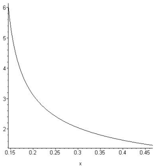
The joint distribution $f(x_1, x_2)$ is

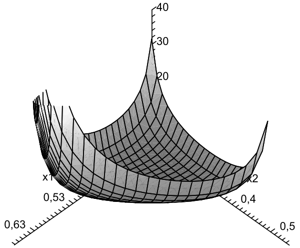
$$\frac{41 f_1(x_1) f_2(x_2) f_3(1 - x_1 - x_2)}{120\Gamma^2(3/2)} =$$

$$\frac{\sqrt{120/41}}{\Gamma^2(3/2)\sqrt{(x_1 - 1/3)(x_2 - 1/5)(7/8 - x_1 - x_2)}}.$$









Thank You!

Questions?