Discrete Second-order Probability Distributions that Factor into Marginals

David Sundgren
University of Gävle

Introduction

In realistic decision problems there is more often than not uncertainty in the background information. As for representation of uncertain or imprecise probability values, second-order probability, i.e. probability distributions over probabilities, offers an option. With a subjective view of probability second-order probability would seem to be impractical since it is hard for a person to construct a second-order distribution that reflects his or her beliefs. From the perspective of probability as relative frequency the task of constructing or updating a second-order probability distribution from data is somewhat easier. Here a very simple model for updating lower bounds of probabilities is employed.

But the difficulties in choosing second-order distributions may be further alleviated if structural properties are considered. Either some of the probability values are dependent in some way, e.g. that they are known to be almost equal, or they are not dependent in any other way than what follows from that the values sum to one.

In this work we present the unique family of discrete second-order probability distributions that correspond to the case where dependence is limited. These distributions are shown to have the property that the joint distributions are equal to normalised products of marginal distributions. The distribution family introduced here is a generalisation of a special case of the multivariate Pólya distribution and is shown to be conjugate prior to a compound hypergeometric distribution.

Why Discrete?

The study of second-order probability distributions has sofar mostly been restricted to continuous distributions. One family of continuous second-order distributions is a generalisation of a special case of the Dirichlet distribution with the property of the joint distribution being proportional to the product of marginal distributions. In the absence of information about dependencies among the first-order probabilities these distributions appear to offer a corresponding non-informativeness. These continuous second-order distributions have as parameters the lower bounds of the first-order probabilities.

In a continuous second-order setting, a lower bound of a probability can rarely if ever be the result of an observation. But after seeing a three-eyed dog in a kennel of ten, I know that at least one out of ten dogs in that kennel has three eyes. Outside the kennel, I cannot, based on the observation, say much more than that the probability of coming across a three-eyed dog is non-zero. Thus updating of lower bounds comes natural when there is a relative frequency interpretation of probabilities.

In this paper the discrete counterpart of the shifted Dirichlet distribution is found and an environment for updating of lower bounds is suggested. But the non-informativeness or limited dependency associated with the distributions that factor into marginals remain after updating only with a compound hypergeometric likelihood function of which there is more to learn.

Finding the Distributions

We want to find probability distributions $p(k_1,k_2,\dots,k_n)$ with k_i being integers such that $\sum_{i=1}^n k_i=N$ and

$$p(k_1,k_2,\dots,k_n)=\frac{1}{K}\prod_{i=1}^n p_i(k_i)\,,$$

where p_i is the marginal distribution corresponding to variable k_i.
Then the marginal distribution $p_i(k_i)$ equals

$$\frac{1}{K}p_i(k_i)\ast_{j\neq i}p_j\left(N-k_i\right)\,,$$

where $\ast_{j\neq i}$ is the $n-1$-fold repeated convolution $p_1\ast p_2\ast\cdots\ast p_{i-1}\ast p_{i+1}\ast\cdots\ast p_n$ and $K=\ast_{i=1}^n p_i(N)$.

$$\begin{aligned}p_2\ast p_3\ast p_4\ast\ldots\ast p_n(N-k_1)&=K\\ p_1\ast p_3\ast p_4\ast\ldots\ast p_n(N-k_2)&=K\\ &\vdots\\ p_1\ast p_2\ast p_3\ast\ldots\ast p_{n-1}(N-k_n)&=K\end{aligned}$$

In the transform domain,

$$\begin{aligned}\mathcal{Z}\{p_2\}\mathcal{Z}\{p_3\}\ldots\mathcal{Z}\{p_n\}&=\mathcal{Z}\{KH(c_1-k_1)\}\\ \mathcal{Z}\{p_1\}\mathcal{Z}\{p_3\}\ldots\mathcal{Z}\{p_n\}&=\mathcal{Z}\{KH(c_2-k_2)\}\\ &\vdots\\ \mathcal{Z}\{p_1\}\mathcal{Z}\{p_2\}\ldots\mathcal{Z}\{p_{n-1}\}&=\mathcal{Z}\{KH(c_n-k_n)\}\end{aligned}$$

where H is the Heaviside function and the support of p_i ends at $k_i=c_i$.
If $p_i(k_i)$ can be written as a shifted function $q_i(k_i-a_i)$,
$\mathcal{Z}\{p_i(k_i)\}=\mathcal{Z}\{q_i(k_i-a_i)\}=\mathcal{Z}\{q_i(k_i)\}z^{-a_i}$
and $\prod_{j\neq i}\mathcal{Z}\{p_j\}=\frac{\prod_{j\neq i}\mathcal{Z}\{q_j\}}{z^{\sum_{j\neq i}a_j}}=\frac{Kz}{z-1}\frac{1}{z^{\sum_{j\neq i}a_j}}$
Thus $\mathcal{Z}\{q_i\}=\left(\frac{Kz}{z-1}\right)^{\frac{1}{n-1}}=\frac{K^{\frac{1}{n-1}}\Gamma\left(k_i+\frac{1}{n-1}\right)}{k_i!\Gamma\left(\frac{1}{n-1}\right)}$
Putting back the shift k_i-a_i

$$p_i(k_i)=\frac{K^{\frac{1}{n-1}}\Gamma\left(k_i-a_i+\frac{1}{n-1}\right)}{(k_i-a_i)!\Gamma\left(\frac{1}{n-1}\right)}$$

But

$$K^{n-1}=\frac{\left(N-\sum_{j=1}^n a_j\right)!\Gamma\left(k_i-a_i+\frac{1}{n-1}\right)}{(n-1)\Gamma\left(N+1-\sum_{j=1}^n a_j+\frac{1}{n-1}\right)(k_i-a_i)!}$$

So

$$p_i(k_i)=\frac{(N-\sum_{j=1}^n a_j)!\Gamma\left(k_i-a_i+\frac{1}{n-1}\right)}{(n-1)\Gamma\left(N+1-\sum_{j=1}^n a_j+\frac{1}{n-1}\right)(k_i-a_i)!}\,,$$

and

$$p(k_1,\dots,k_n)=\frac{\left(N-\sum_{i=1}^n a_i\right)!\prod_{i=1}^n\frac{\Gamma\left(k_i-a_i+\frac{1}{n-1}\right)}{(k_i-a_i)!}}{(n-1)\Gamma\left(\frac{1}{n-1}\right)^{n-1}\Gamma\left(N+1-\sum_{i=1}^n a_i+\frac{1}{n-1}\right)}$$

Updating of Lower Bounds

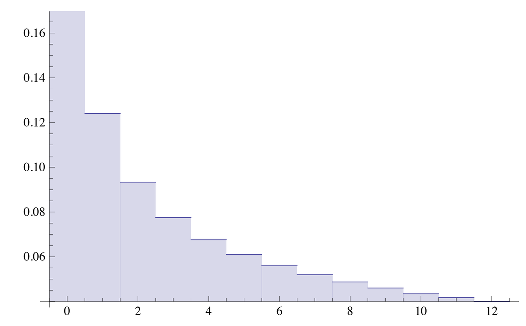
Let there be N balls of n different colours in an urn. By picking up a_i balls of type i we can conclude that there are at least a_i balls of type i among the N.
The observed number of balls gives us the lower bounds a_i of the distribution. Since the distribution is concerned with how many balls there are in total of each colour, even those picked from the urn count, as well as those still inside, that is why we leave the observed balls visible on the side.
The unknown balls are in the urn and the known balls on the plate.

The Urn and the Plate

Let $N=12$ and $n=3$. At first the plate is empty, that is $a_1=a_2=a_3=0$ and the prior distribution is

$$\frac{12!\Gamma(k_1+1/2)\Gamma(k_2+1/2)\Gamma(k_3+1/2)}{2\Gamma(1/2)^2\Gamma(13+1/2)k_1!k_2!k_3!}$$

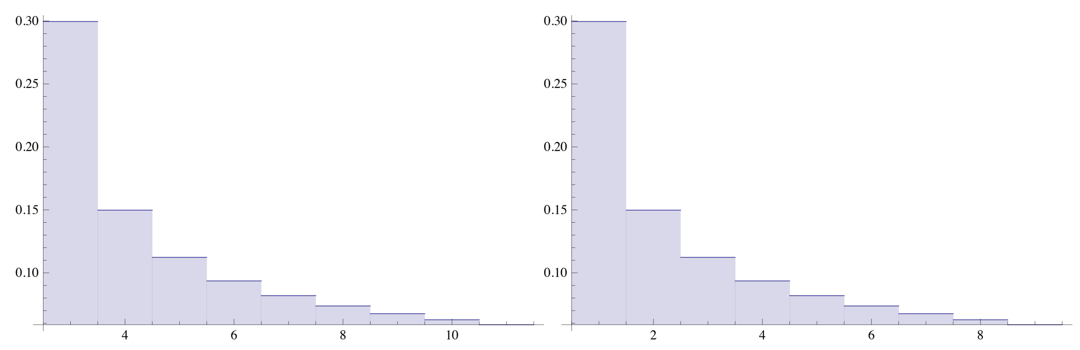
For all three marginal prior distributions



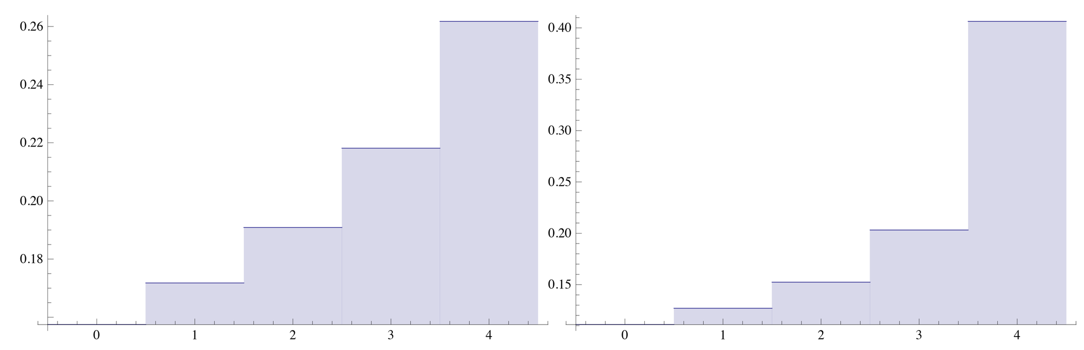
Then four balls are picked from the urn, three of which are of type 1 and one of type 2, so the posterior is

$$\frac{8!\Gamma(k_1-5/2)\Gamma(k_2-1/2)\Gamma(k_3+1/2)}{2\Gamma(1/2)^2\Gamma(9+1/2)(k_1-3)!(k_2-1)!k_3!}$$

We see the marginal distributions of k_1 and k_2 below, observe that $p_1(k)=p_2(k-2)$.



If we assume conjugacy, that the balls are distributed according to the same type of shifted Pólya as the prior even after updating, the likelihood function $p(a_i|k_i)$ also factors into marginals. We have the same expressions k_i-a_i but since a_i is the variable and k_i constant the graph is mirrored.
Given $k_1=6,k_2=4,k_3=2,\sum_{i=1}^3 a_i=4$
The likelihood function would then have marginal probabilities such that
$\Pr(a_1=0)=0.157,\Pr(a_1=1)=0.172,$
$\Pr(a_1=2)=0.191,\Pr(a_1=3)=0.218$
and $\Pr(a_1=4)=0.262$
For a_2 we have
$\Pr(a_2=0)=0.111,\Pr(a_2=1)=0.127,$
$\Pr(a_2=2)=0.152,\Pr(a_2=3)=0.202$
and $\Pr(a_2=4)=0.406$



Department of Electronics, Mathematics and Natural Sciences, University of Gävle, Sweden Mail: dsn@hig.se WWW: http://www.hig.se/dsn