RESEARCH SUPPORT
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

DECEMBER 1996

CHARACTERISSATIONS, STOCHASTIC EQUATIONS, AND THE GIBBS SAMPLER

WORKING PAPER #9612-33

STEPHEN WALKER
IMPERIAL COLLEGE, 180 QUEEN'S GATE, LONDON

AND

PAUL DAMIEN
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

Characterisations, Stochastic Equations, and the Gibbs Sampler

Stephen Walker[1]
and
Paul Damien[2]

[1] Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ.

[2] Department of Statistics and Management Science, School of Business, University of Michigan, Ann Arbor, 48109-1234, USA.

SUMMARY

We obtain characterisations of distributions on the real line and solve stochastic equations using the Gibbs sampler. Particular stochastic equations considered are of the type $X =_d B(X + C)$ and $X =_d BX + C$ and we will be considering solutions when B and C are not necessarily independent.

1 Introduction

Dufresne (1996) considers the stochastic difference equation

$$X_{n+1} = B_{n+1}(X_n + C_n),$$

where $\{B_n, n \geq 1\}$ and $\{C_n, n \geq 0\}$ are independent iid copies of B and C having densities f_B and f_C, respectively. Vervaat (1979) provides

$$E(\log B) < 0, \quad E(\log |C|)_+ < \infty$$

as being sufficient conditions for the existence and uniqueness of the limit distribution of X_n. Under the above conditions let $X_n \to_d X$. then the limit X satisfies the stochastic equation

$$X =_d B(X + C).$$

where B. C and X are independent. Solutions to this stochastic equation are apparently hard to come by (Dufresne. 1996). In this paper we will characterise solutions via the Gibbs sampler (Smith and Roberts. 1993). A closely related stochastic equation is given by $X =_d BX + C$. We will find solutions to this stochastic equation when B and C are not independent.

However, the paper is firstly concerned with characterising continuous densities on the real line. for example. the normal. gamma. exponential and beta distributions. We do this by considering a Gibbs sampler over the joint density $f(x.u)$. As is well known the Gibbs sampler generates a sequence $\{X_n\}$ by taking U_n from $f(.|X_n)$ and then taking X_{n+1} from $f(.|U_n)$. In many instances it is possible to define $U_n = h_2(V_2.X_n)$, for some appropriate random variable V_2 and function h_2. and also to define $X_{n+1} = h_1(V_1.U_n)$. again for some appropriate random variable V_1 and function h_1. Putting these together it may be possible to construct the sequence directly via

$$X_{n+1} = h(V.X_n).$$

It is well known that under mild regularity conditions $X_n \to_d X$ and the limit variable X satisfies

$$X =_d h(V.X).$$

This is a stochastic equation with solution $f_X(x) = \int f(x,u)du$. Moreover. this solution is unique and therefore $X =_d h(V.X)$ also characterises X (provided $f(x|u)$ and $f(u|x)$ define $f_X(x)$ uniquely).

We start in Section 2 with the characterisations of some well known densities on the real line. The characterisation of the exponential density leads to a solution of the stochastic equation $X =_d B(X + C)$ when B is uniform on (0.1) and C has an exponential distribution. The characterisation of the gamma distribution leads to a solution of the stochastic equation $X =_d BX + C$ when B has a beta distribution and C an exponential distribution. In Section 3 we obtain new solutions to the stochastic equation $X =_d B(X + C)$ including a solution with B and C not independent.

2 Characterisations of densities

Let X be a random variable defined on the real line, $\mathcal{U}(0,1)$ represent the uniform distribution on the interval $(0,1)$, $\mathcal{E}(\lambda)$ represent the exponential distribution with mean $1/\lambda$ ($\lambda > 0$), and $=_d$ denote equality in distribution. Our characterisations are of the type

$$X =_d h(V, X),$$

where h is a function and V a collection of independent variables, also independent of X.

Exponential(1) density

THEOREM 1. *Let* V_1 *and* V_2 *be* $\mathcal{U}(0,1)$. *Then*

$$X =_d V_1 (X - \log V_2) \tag{1}$$

if, and only if, $X \sim \mathcal{E}(1)$.

PROOF. First we describe how such a characterisation arose. The density for a $\mathcal{E}(1)$ variable is given up to proportionality by $\exp(-x)I(x > 0)$. We construct a joint density for X and U, a random variable defined on $(0,1)$, given up to proportionality by $I(0 < u < \exp(-x), 0 < x)$. Clearly the marginal density for X is $\mathcal{E}(1)$. We can now construct a random sequence $\{X_n\}$ such that $X_n \to_d \mathcal{E}(1)$ using the Gibbs sampler (Smith and Roberts, 1993). This involves sampling X_{n+1} from $f(.|U_n)$ having taken U_n from $f(.|X_n)$. We can define $U_n = V_{2n} \exp(-X_n)$ where $V_{2n} \sim \mathcal{U}(0,1)$ (independent of X_n) and $X_{n+1} = -V_{1n} \log U_n$ where $V_{1n} \sim \mathcal{U}(0,1)$ (independent of V_{2n} and X_n). Joining these two together gives

$$X_{n+1} = V_{1n}(X_n - \log V_{2n}).$$

From the convergence properties of the Gibbs sampler and provided the sequence is uniquely associated with the exponential density (that is, the two conditional densities uniquely determine the joint density of X and U) then the Theorem is proved.

However, we will only use the above arguement to propose the characterisations and will rely on more traditional methods to prove the result.

Therefore let us assume that $X \sim \mathcal{E}(1)$ and consider the Laplace transform of $Y = V_1(X - \log V_2)$ noting that $-\log V_2 \sim \mathcal{E}(1)$. If $\phi_Y(\theta) = E_Y \exp(-\theta Y)$ then clearly

$$\phi_Y(\theta) = \int_0^1 \left[\frac{\phi_X(\theta u)}{1 + \theta u}\right] du. \tag{2}$$

If now $\phi_X(\theta) = 1/(1+\theta)$ then $\phi_Y(\theta) = 1/(1+\theta)$ proving the 'if' assertion.

To prove uniqueness we consider moments. Let $\mu_i = \int x^i dF_X(x)$. First, it is straightforward to show using (2) that if (1) is true then

$$\sum_{n=0}^{\infty}(-1)^n\theta^n\mu_n/n! = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-1)^{m+n}\theta^{n+m}\mu_n/[n!(m+n+1)].$$

Therefore for all i

$$(-1)^i\mu_i/i! = \sum_{m+n=i}(-1)^{m+n}\mu_n/[n!(m+n+1)].$$

leading to

$$\mu_i/i! = (i+1)^{-1}\sum_{j=0}^{i}\mu_j/j!$$

from which the result $\mu_i = i!$ follows. Since $\sum_i \mu_i t^i/i!$ is absolutely convergent for $|t| < 1$ the moments characterise the distribution of X which is identified as being $\mathcal{E}(1)$, completing the proof.

COROLLARY 1. *Let V_1 and V_2 be iid $\mathcal{U}(0,1)$. Then*

$$X =_d V_1\left(X - \lambda^{-1}\log V_2\right) \tag{3}$$

if, and only if, $X \sim \mathcal{E}(\lambda)$.

The proof to this follows the same procedure as outlined in the proof to Theorem 1. It is clear that (3) provides a solution to the stochastic equation $X =_d B(X + C)$.

The characterisation of $\mathcal{E}(\lambda)$ given in Corollary 1 is not new (Kotz and Steutel, 1988), however, that a proof is available via the Gibbs sampler is intriguing.

PROPOSITION 1. *If $B \sim \mathcal{U}(0,1)$ and $C \sim \mathcal{E}(\lambda)$ then the unique solution to the stochastic equation $X =_d B(X + C)$ is given by $X \sim \mathcal{E}(\lambda)$.*

Normal(0,1) density.

In this section let $\mathcal{N}(\nu, \sigma^2)$ represent the normal distribution with mean ν and variance σ^2. Proceeding as for the exponential density we define the joint density of X and U up to proportionality by $I(0 < u < \exp(-x^2/2))$. The Gibbs sequence is defined as follows; $U_n = V_{2n}\exp(-X_n^2/2)$ and $X_{n+1} = (2V_{1n} - 1)\sqrt{-2\log U_n}$. Putting these together gives

$$X_{n+1} = (2V_{1n} - 1)\sqrt{X_n^2 - 2\log V_{2n}}.$$

Therefore,

THEOREM 2. *Let V_1 and V_2 be iid $\mathcal{U}(0,1)$. Then*

$$X =_d (2V_1 - 1)\sqrt{X^2 - 2\log V_2} \qquad (4)$$

if, and only if, $X \sim \mathcal{N}(0,1)$.

PROOF. First we note that if X is symmetric about 0 and X^2 is chi-squared then X is normal and note that if (4) defines a random variable X on $(-\infty, +\infty)$ then it is symmetric about 0. Define $Y = (2V_1 - 1)^2(X^2 - 2\log V_2)$ and note that $(2V_1 - 1)^2$ is $\mathcal{B}(1/2, 1)$, where $\mathcal{B}$ denotes the beta distribution, and $-\log V_1$ is $\mathcal{E}(1)$. Therefore

$$\phi_Y(\theta) = B(1/2, 1)^{-1}\int_0^1 u^{-1/2}\left[\frac{\phi_{X^2}(\theta u)}{1 + 2\theta u}\right] du.$$

If now $X \sim \mathcal{N}(0,1)$ then $\phi_{X^2}(\theta) = 1/\sqrt{1+2\theta}$ leading to $\phi_Y(\theta) = 1/\sqrt{1+2\theta}$, provided $|2\theta| < 1$, proving the 'if' assertion. To show uniqueness we need to consider the integral equation

$$\phi_Y(\theta) = 1/2\int_0^1 u^{-1/2}(1 + 2\theta u)^{-1}\phi_Y(\theta u)du, \qquad (5)$$

where $Y = X^2$. Our aim is to show the unique solution is given by $\phi_Y(\theta) = 1/\sqrt{1+2\theta}$ or $f_Y(y) = (2\pi)^{-1/2}y^{-1/2}\exp(-y/2)I(y > 0)$. The moments for this density are given by $\mu_i = 2^i\Gamma(i+1/2)/\Gamma(1/2)$ and since $\sum_i \mu_i t^i/i!$ is absolutely convergent for $|t| < 1/2$ it follows that these moments characterise f_Y uniquely. It is easy to show that (5) leads to

$$\sum_{n=0}^{\infty}(-1)^n\theta^n\mu_n/n! = 1/2\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-1)^{n+m}2^m\theta^{n+m}\mu_n/[n!(n+m+1/2)].$$

Therefore

$$\mu_i/i! = (2i+1)^{-1}\sum_{j=0}^{i}2^{i-j}\mu_j/j!$$

the unique solution being given by $\mu_i = 2^i\Gamma(i+1/2)/\Gamma(1/2)$, completing the proof.

COROLLARY 2. *Let V_1 and V_2 be iid $\mathcal{U}(0,1)$. Then*

$$X =_d \left\{\nu + (2V_1 - 1)\sqrt{(X-\nu)^2 - 2\sigma^2\log V_2}\right\} \tag{6}$$

if, and only if, $X \sim \mathcal{N}(\nu,\sigma^2)$.

Gamma(a,1) density.

Let $\mathcal{G}(a,b)$ $(a,b>0)$ represent the gamma distribution with mean a/b. In the following we will additionally assume that $a > 1$. If $X \sim \mathcal{G}(a,1)$ then the density for X is given by $1/\Gamma(a)x^{a-1}\exp(-x)I(x>0)$. Here we introduce the variable U which has joint density with X given by $f(x,u) \propto I(0<u<x^{a-1}, x>0)\exp(-x)$. The Gibbs sequence is given by $U_n = V_{2n}X_n^{a-1}$ and $X_{n+1} = U_n^{1/(a-1)} - \log V_{1n}$. Combining these leads to

$$X_{n+1} = X_nV_{2n}^{1/(a-1)} - \log V_{1n}.$$

Therefore,

THEOREM 3. *Let V_1 and V_2 be iid $\mathcal{U}(0,1)$. Then*

$$X =_d XV_2^{1/(a-1)} - \log V_1 \tag{7}$$

if, and only if, $X \sim \mathcal{G}(a,1)$ $(a>1)$.

PROOF. Define $Y = XV_2^{1/(a-1)} - \log V_1$ and note that, for $a > 1$, $V_2^{1/(a-1)}$ is $\mathcal{B}(a-1, 1)$. Then we have

$$\phi_Y(\theta) = \frac{a-1}{1+\theta}\int_0^1 u^{a-2}\phi_X(\theta u)du. \tag{8}$$

If $X \sim \mathcal{G}(a, 1)$ then $\phi_X(\theta) = (1+\theta)^{-a}$ and with this

$$\phi_Y(\theta) = \frac{a-1}{1+\theta}\sum_{m=0}^{\infty}(-1)^m\Gamma(a+m)\theta^m/[\Gamma(a)(m+a-1)],$$

for $|\theta| < 1$, leading to

$$\phi_Y(\theta) = (1+\theta)^{-1}\sum_{m=0}^{\infty}(-1)^m\Gamma(a-1+m)\theta^m/\Gamma(a-1),$$

and hence $\phi_Y(\theta) = (1+\theta)^{-a}$ proving the 'if' assertion. To show uniqueness we consider

$$\phi_X(\theta) = \frac{a-1}{1+\theta}\int_0^1 u^{a-2}\phi_X(\theta u)du.$$

This leads to

$$\sum_{n=0}^{\infty}(-1)^n\theta^n\mu_n/n! = (a-1)\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-1)^{n+m}\theta^{n+m}\mu_n/[n!(n+a-1)],$$

and hence

$$\mu_i/i! = (a-1)\sum_{j=0}^{i}\mu_j/[j!(j+a-1)]$$

the unique solution being given by $\mu_i = \Gamma(i+a)/\Gamma(a)$. Again, since $\sum_i \mu_i t^i/i!$ is absolutely convergent for $|t| < 1$ this implies these moments characterise uniquely the distribution for X, which is identified as being $\mathcal{G}(a, 1)$, completing the proof.

COROLLARY 3. *Let* V_1 *and* V_2 *be iid* $\mathcal{U}(0,1)$. *Then*

$$X =_d XV_2^{1/(a-1)} - b^{-1}\log V_1 \tag{9}$$

if, and only if, $X \sim \mathcal{G}(a, b)$ $(a > 1)$.

Clearly (9) provides a solution to the stochastic equation $X =_d BX + C$. This solution is already known (Vervaat, 1979, 3.8.2). The use of $V^{1/(a-1)}$ to represent the $\mathcal{B}(a-1,1)$ distribution in Vervaat (1979) takes on added significance. A characterisation of $\mathcal{G}(a,1)$ exists for all $a > 0$ by considering the joint density $f(x,u) \propto x^{a-1} I(u < \exp(-x))$. For the Gibbs sequence we obtain $U_n = V_{2n}\exp(-X_n)$ and $X_{n+1} = -V_{1n}^{1/a}\log U_n$ leading to the characterisation/stochastic equation

$$X =_d V_1^{1/a}\{X - \log V_2\}.$$

Beta(a,b) density.

We state without proof the characterisation for $\mathcal{B}(a,b)$ for $a > 1, b > 0$.

THEOREM 4. *Let V_1 and V_2 be iid $\mathcal{U}(0,1)$. Then*

$$X =_d \left\{1 - V_1^{1/b}\left(1 - XV_2^{1/(a-1)}\right)\right\} \qquad (10)$$

if, and only if, $X \sim \mathcal{B}(a,b)$ $(a > 1)$.

This provides a solution to the stochastic equation $X =_d BX + C$ when B and C are not independent, i.e. $B = V_1^{1/b}V_2^{1/(a-1)}$ and $C = 1 - V_1^{1/b}$.

Extreme value density

The extreme value density is given up to proportionality by $f(x) \propto \exp(-\exp(x))I(x > 0)$ and consider the joint density $f(x,u) \propto \exp(-u)I(u > \exp(x) > 1)$. The Gibbs sequence is given by $U_n = -\log V_{1n} + \exp(X_n)$ and $\exp(X_{n+1}) = U_n^{V_{2n}}$, leading to the characterisation/stochastic equation

$$\exp(X) =_d \{\exp(X) - \log V_1\}^{V_2}.$$

For some general theory, let f_X be a strictly monotone decreasing density on $(0,\infty)$ with inverse $g = f^{-1}$. Define the joint density of X and U by $f(x,u) = I(0 < u < f(x))$. The Gibbs sampler is then given by $U_n = V_1 f(X_n)$ and $X_{n+1} = V_2 g(U_n)$, where V_1 and V_2 are independent $\mathcal{U}(0,1)$ variables, leading to the stochastic equation $X =_d V_2 g[V_1 f(X)]$. This has the solution f_X. Therefore, PROPOSITION 2. *A random variable X defined on $(0,\infty)$ has a strictly monotone decreasing density f if, and only if, $X =_d V_2 f^{-1}[V_1 f(X)]$, where V_1 and V_2 are independent uniform(0,1) variables.*

3 $X =_d B(X+C)$ and $U =_d BU + C$

The aim in this section is to provide a solution to the above stochastic equations via the construction of a suitable Gibbs sampler. An example of this solution has already been given in Proposition 1. Solutions are based on the following Theorem,

THEOREM 5. *If $f(x,u)$ is a joint density with conditional densites $f(x|u)$ and $f(u|x)$ such that $f(x|u) = 1/u f_B(x/u)$ and $f(u|x) = f_C(u-x)$ where f_B and f_C are the densities for independent B and C, respectively, then $f_X(x)$ is the unique solution to the stochastic equation $X =_d B(X+C)$, $f_U(u)$ is the unique solution to the stochastic equation $U =_d BU + C$, and $X =_d BU$ and $U =_d X + C$.*

EXAMPLE. Let $f(x,u) \propto x^{a-1}(u-x)^{b-1}\exp(-cu)I(0 < x < u)$ for $a, b, c > 0$. Then

$$f(x|u) \propto x^{a-1}(u-x)^{b-1} I(x/u < 1)$$

and

$$f(u|x) \propto (u-x)^{b-1}\exp(-cu)I(u-x > 0).$$

Therefore $B \sim \mathcal{B}(a,b)$ and $C \sim \mathcal{G}(b,c)$. Therefore $X \sim \mathcal{G}(a,c)$ solves $X =_d B(X+C)$ and $U \sim \mathcal{G}(a+b,c)$ solves $U =_d BU + C$. Moreover, we deduce $\mathcal{G}(a,c) =_d \mathcal{B}(a,b) \times \mathcal{G}(a+b,c)$.

In fact we can show that the solution in the example is the only one available via the Gibbs sampler for which $C > 0$ a.s. and $0 < B < 1$ a.s. To find solutions we need to identify independent random variables B and C such that the conditionals $f(x|u) = 1/u f_B(x/u)$ and $f(u|x) = f_C(u-x)$ uniquely define the joint density $f(x,u)$. The result is contained in the following Theorem.

THEOREM 6. $f(x|u) = 1/u f_B(x/u)$ *$(0 < B < 1)$ and* $f(u|x) = f_C(u-x)$

$(C > 0)$ define $f(x,u)$ uniquely if, and only if, $C \sim \mathcal{G}(b,c)$ and $B \sim \mathcal{B}(a,b)$.

PROOF. Clearly the 'if' assertion is trivial. Now suppose that

$$f(x,u) = 1/u f_B(x/u) f_U(u) = f_C(u-x) f_X(x),$$

from which it is immediate that U and X/U are independent as are X and $U - X$ (note that $U > X > 0$). If we let $V = U - X$ and $W = X$ we see that V/W and $V + W$ are independent. Since V and W are independent Lukacs' Theorem (Lukacs, 1955) states that V and W are both gamma variables with the same scale parameter, say c, and let $U - X \sim \mathcal{G}(b,c)$ and $X \sim \mathcal{G}(a,c)$. Therefore $C \sim \mathcal{G}(b,c)$.

$$f(x,u) \propto (u-x)^{b-1} \exp(-c(u-x))\, x^{a-1} \exp(-cx) I(0 < x < u),$$

so $f(x|u) \propto x^{a-1}(u-x)^{b-1} I(0 < x < u)$ which implies that $B \sim \mathcal{B}(a,b)$.

We can obtain a further solution to the stochastic equation $X =_d BX + C$ when B and C are not independent (see also Theorem 4) by considering the Pareto distribution with density $f(x) \propto x^{-a-1} I(x > \alpha)$ and parameters $a, \alpha > 0$.

THEOREM 7. *If $B = V_1 V_2^{-1/(a+1)}$ and $C = \alpha(1 - V_1)$ where V_1 and V_2 are iid $\mathcal{U}(0,1)$ then the unique solution to the stochastic equation $X =_d BX + C$ is given by the Pareto distribution with parameters (a, α).*

References

Dufresne,D. (1996). On the stochastic equation $\mathcal{L}(X) = \mathcal{L}[B(X + C)]$ and a property of gamma distributions. *Bernoulli* **2**(3), 287-291.

Kotz,S. and Steutel,F.W. (1988). A note on the characterisation of exponential distributions. *Statist. Probab. Letters* 6, 201-203.

Lukacs,E. (1955). A characterisation of the gamma distribution. *Ann. Math. Statist.* **26**, 319-324.

Smith.A.F.M and Roberts.G.O. (1993). Bayesian computations via the Gibbs sampler and related Markov chain Monte Carlo methods. *J. R. Statist. Soc. B* **55**. 3-23.

Vervaat.W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probab.* **11**. 750-783.