Addendum to “Reply to Reflections”

by A. Ronald Gallant[1]

October 17, 2015

“Reply to Comment on Reflections” (Gallant, 2015b) constructed the dominating measure for Chris Sims’s continuous example (Sims, 2015). This addendum does the same for his discrete example, which is Table 1 of Sims (2015).

The issue addressed here is that standard Bayesian computing methods assume that likelihoods are with respect to either Lebesgue measure or counting measure whereas the density $p^*(x \mid \theta)$ constructed in “Reflections” (Gallant, 2015a) may not be. Explicit construction of the dominating measure is required to use these computing methods or, more correctly, explicit construction of the Radon-Nikodym derivative $\mathrm{adj}(x, \theta)$ of the dominating measure with respect to either Lebesgue or counting measure is required.

One can regard the adjustment $\mathrm{adj}(x, \theta)$ as a partial specification of the prior or a completion of the definition of the likelihood, depending on one’s point of view. It is a matter of how one groups terms. The grouping

$$p(\theta \mid x) \propto p^*(x \mid \theta) \left[\mathrm{adj}(x, \theta)\pi^*(\theta)\right]$$

suggests a data dependent prior. Or, more precisely, this grouping suggests that the absence of $\mathrm{adj}(x, \theta)$ implies a data dependent prior. The grouping

$$p(\theta \mid x) \propto \left[\mathrm{adj}(x, \theta)p^*(x \mid \theta)\right] \pi^*(\theta)$$

suggests a particular choice of likelihood. Above, $p^*(x \mid \theta)$ is defined by equation (17) of “Reflections” and the prior $\pi^*(\theta)$ by the second paragraph of Section 3 of “Reflections.”

Repeating a definition from “Reflections”, let

$$C^{(\theta,z)} = \{x \in \mathcal{X} \,:\, Z(x, \theta) = z\}. \tag{15R}$$

Constructing $\mathrm{adj}(x, \theta)$ for Sims’s discrete example amounts to choosing some distribution on each $C^{(\theta,z)}$.

[1]©2015 A. Ronald Gallant, POB 659, Chapel Hill NC 27514, aronaldg@gmail.com

We will give two constructions of the posterior for Sims's example. The first is an intuitive construction based on the the principle that probability for the dominating measure $\mathrm{adj}(x, \theta)$ is distributed uniformly over $C^{(\theta,z)}$. Following this construction we will repeat the construction using the notion of a representer from "Reflections," which, as we shall see, amounts to enforcing the principle that probability for the dominating measure $\mathrm{adj}(x, \theta)$ is distributed uniformly over $C^{(\theta,z)}$.

Tables 1 through 5 display the intermediate steps in the first construction of the posterior for Sims's discrete example. Table 1, Table 2, Table 3, Table 6, Table 7, and Table 5, in that order, display the intermediate steps in the second construction of the posterior for Sims's discrete example

His example does not satisfy $\mathcal{E}Z = 0$ but this does not matter for what follows. Also, in what follows, set $\pi = \pi^*$.

Table 1. Preimages and Probabilities for $Z(x, \theta)$ Defined by Table 1 of Sims (2015).

			$P(Z = z \mid \Lambda = \theta)$		
Preimage	z	$P(Z = z)$	$\theta = 1$	$\theta = 2$	$\theta = 3$
$C_1 = \{(1,1), (3,3), (4,2), (4,3)\}$	1	Ψ_1	Ψ_1	Ψ_1	Ψ_1
$C_2 = \{(1,2), (2,1), (2,3), (3,2), (4,1)\}$	2	Ψ_2	Ψ_2	Ψ_2	Ψ_2
$C_3 = \{(1,3), (2,2), (3,1)\}$	3	Ψ_3	Ψ_3	Ψ_3	Ψ_3

The sets that can occur when it is known that $\Lambda = \theta$ has occurred are those preimages C_z that contain (x, θ) for some x in the support $\mathcal{X}$ of X. Let O_θ be the union of the sets that can occur when it is known that $\Lambda = \theta$ has occurred. Conditional probability is computed as $P(Z = z \mid \Lambda = \theta) = P(C_z \cap O_\theta)/P(O_\theta)$. In this instance, O_θ is the support Θ of Λ so that $P(C_z \cap O_\theta) = \Psi_z$ and $P(O_\theta) = 1$.

Table 2. Conditional Probabilities Implied by Table 1

	$P(X = x \mid \Lambda = \theta)$		
x	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	Ψ_1	Ψ_2	Ψ_3
2	Ψ_2	Ψ_3	Ψ_2
3	Ψ_3	Ψ_2	Ψ_1
4	Ψ_2	Ψ_1	Ψ_1

$P(X = x \mid \Lambda = \theta)$ is the probability of the preimage in Table 1 that contains (x, θ), which is C_1 for $P(X = 1 \mid \Lambda = 1)$, divided by the probability of the union of all sets that contain a point of the form $(\cdot, \theta)$, which is $O_1 = C_1 \cup C_2 \cup C_3$ for $P(X = 1 \mid \Lambda = 1)$. Therefore $P(X = 1 \mid \Lambda = 1) = P(C_1)/P(O_1) = \Psi_1/1$. The column probabilities do not sum to one, which is an issue that the adjustment $\text{adj}(x, \theta)$ resolves.

Table 3. The Sets $C^{(\theta,z)}$ Implied by Table 1

z	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	$\{1\}$	$\{4\}$	$\{3,4\}$
2	$\{2,4\}$	$\{1,3\}$	$\{2\}$
3	$\{3\}$	$\{2\}$	$\{1\}$

$C^{(\theta,z)}$ is defined by equation (15R).

Table 4. Dominating Measure for Table 2

	$\mathrm{adj}(x,\theta)$		
x	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	1	$\frac{1}{2}$	1
2	$\frac{1}{2}$	1	1
3	1	$\frac{1}{2}$	$\frac{1}{2}$
4	$\frac{1}{2}$	1	$\frac{1}{2}$

When $C^{(\theta,z)}$ given in Table 4 has more than one element, probability is split evenly among the points. E.g., $C^{(1,2)} = \{2,4\}$; therefore, $\mathrm{adj}(2,1) = \mathrm{adj}(4,1) = \frac{1}{2}$ When $C^{(\theta,z)}$ is a singleton set, $\mathrm{adj}(\theta,z) = 1$; therefore $\mathrm{adj}(1,1) = 1$.

Table 5. Normalizing Factors

x	Normalizing Factors
1	$D_1 = \Psi_1\pi_1 + \frac{1}{2}\Psi_2\pi_2 + \Psi_3\pi_3$
2	$D_2 = \frac{1}{2}\Psi_2\pi_1 + \Psi_3\pi_2 + \Psi_2\pi_3$
3	$D_3 = \Psi_3\pi_1 + \frac{1}{2}\Psi_2\pi_2 + \frac{1}{2}\Psi_1\pi_3$
4	$D_4 = \frac{1}{2}\Psi_2\pi_1 + \Psi_1\pi_2 + \frac{1}{2}\Psi_1\pi_3$

For prior $\pi_\theta = P(\Lambda = \theta)$, where $\theta = 1, 2, 3$, the table shows the normalizing factors for the posterior displayed in equations (1), (2), and (3). The rows of the table are constructed from the rows of Tables 2 and 4 and the prior.

From Tables 2, 4, and 5 we obtain the posterior:

$$
\begin{aligned}
P(\Lambda = 1 \,|\, X = 1) &= \frac{\Psi_1 \pi_1}{D_1} \\
P(\Lambda = 1 \,|\, X = 2) &= \frac{\frac{1}{2}\Psi_2 \pi_1}{D_2} \\
P(\Lambda = 1 \,|\, X = 3) &= \frac{\Psi_3 \pi_1}{D_3} \\
P(\Lambda = 1 \,|\, X = 4) &= \frac{\frac{1}{2}\Psi_2 \pi_1}{D_4}
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
P(\Lambda = 2 \,|\, X = 1) &= \frac{\frac{1}{2}\Psi_2 \pi_2}{D_1} \\
P(\Lambda = 2 \,|\, X = 2) &= \frac{\Psi_3 \pi_2}{D_2} \\
P(\Lambda = 2 \,|\, X = 3) &= \frac{\frac{1}{2}\Psi_2 \pi_2}{D_3} \\
P(\Lambda = 2 \,|\, X = 4) &= \frac{\Psi_1 \pi_2}{D_4}
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
P(\Lambda = 3 \,|\, X = 1) &= \frac{\Psi_3 \pi_3}{D_1} \\
P(\Lambda = 3 \,|\, X = 2) &= \frac{\Psi_2 \pi_3}{D_2} \\
P(\Lambda = 3 \,|\, X = 3) &= \frac{\frac{1}{2}\Psi_1 \pi_3}{D_3} \\
P(\Lambda = 3 \,|\, X = 4) &= \frac{\frac{1}{2}\Psi_1 \pi_3}{D_4}
\end{aligned}
\tag{3}
$$

We shall next repeat the derivation of equations (1) through (3) using the notion of a representer. Recall the definition of representer given in "Reflections:" Assume that $C^{(\theta,z)}$ is not empty for any (θ, z) pair. Then for each $\theta \in \Theta$ and $z \in \mathcal{Z}$ we may choose a point $x^* \in \mathcal{X}$ for which

$$Z(x^*, \theta) = z.$$

The point x^* is the representer of $C^{(\theta,z)}$. Define

$$\Upsilon(z, \theta) = x^*. \tag{16R}$$

Table 6. A Representer for Table 1

	$x^* = \Upsilon(z, \theta)$		
z	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	1	4	3
2	2	1	2
3	3	2	1

Constructed from Table 3 by taking the first element of every set.

Table 7. The Dominating Measure for Tables 2 and 6

	$\text{adj}^*(x^*, \theta)$		
x^*	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	1	$\frac{1}{2}$	1
2	$\frac{1}{2}$	1	1
3	1		$\frac{1}{2}$
4		1	

Probabilities are given by $\frac{1}{\#C^{(\theta,z)}}$, where $\#C^{(\theta,z)}$ is the number of points in $C^{(\theta,z)}$. Entries are only required for points (x^*, θ) that satisfy $x^* = \Upsilon\big[Z(x,\theta),\theta\big]$ for some x, $x = 1, 2, 3, 4$. Note that $\text{adj}(x,\theta) = \text{adj}^*\big(\Upsilon\big[Z(x,\theta),\theta\big],\theta\big)$, where $\text{adj}(x,\theta)$ is defined by Table 4.

The formula for a posterior is

$$P(\Lambda = \theta \,|\, X = x) \propto \operatorname{adj}^*\big(\Upsilon\big[Z(x,\theta),\theta\big],\theta\big)\; p^*\big(\Upsilon\big[Z(x,\theta),\theta\big] \,\big|\, \theta\big)\; \pi(\theta). \tag{4}$$

This will give the same posterior as equations (1), (2), and (3) because the probabilities in Table 4 were chosen to satisfy

$$\operatorname{adj}(x,\theta) = \operatorname{adj}^*\big(\Upsilon\big[Z(x,\theta),\theta\big],\theta\big).$$

References

Gallant, A. Ronald (2015a), "Reflections on the Probability Space Induced by Moment Conditions with Implications for Bayesian Inference," *Journal of Financial Econometrics*, forthcoming.

Gallant, A. Ronald (2015b), "Reply to Comment on Reflections," *Journal of Financial Econometrics*, forthcoming.

Sims, Chris (2015), "Comment on Reflections," *Journal of Financial Econometrics*, forthcoming.