Supplementary materials for the paper "Prior-free probabilistic prediction of future observations"

November 17, 2014

S Technical details for Section 3.3 of the paper

S.1 Existence and uniqueness of the solution (11)

Here the issue is existence and uniqueness of the solution $\theta(T, U) = (\theta_1(T, U), \theta_2(T, U))$ in the gamma problem in Section ??. The only non-trivial part is the solution θ_1 of equation $F_{\theta_1}(t_2) = u_2$, involving only (t_2, u_2). The challenge is that F_{θ_1} is a non-standard distribution. Glaser (1976), in his notation, considers the random variable

$$U^\star = \Big\{ \frac{(\prod_{i=1}^n Y_i)^{1/n}}{\frac{1}{n} \sum_{i=1}^n Y_i} \Big\}^n,$$

the n-th power of the ratio of geometric and arithmetic means of an iid $\mathsf{Gamma}(\theta_1, 1)$ sample. Then $T_2 = n^{-1} \log(U^\star)$, i.e., our T_2 is a monotone increasing function of Glaser's $U^\star$. A consequence of Glaser's Corollary 2.2 is that $U^\star$ is stochastically strictly increasing in θ_1, which implies that $F_{\theta_1}(t_2)$ is a decreasing function of θ_1 for all t_2. Therefore, if a solution exists for θ_1 in (11), it must be unique by monotonicity.

Turning to the existence of a solution for θ_1, we need to show that, for any t_2, $F_{\theta_1}(t_2)$ spans all of the interval $(0, 1)$ for u_2 as θ_1 varies. By monotonicity, it suffices to consider the limits $\theta_1 \to \{0, \infty\}$. Jensen (1986) considers the random variable $W = -1/T_2$ and shows, in his Equation (9), that θ_1/W has a limiting distribution as $\theta_1 \to \{0, \infty\}$, which implies the same for $\theta_1 T_2$. It is now clear that $F_{\theta_1}(t_2)$ converges to 1 and 0 as θ_1 converges to 0 and ∞, respectively, for all t_2. Therefore, a solution for θ_1 in (11) exists for all (t_2, u_2) pairs, as was to be shown.

S.2 Computing the solution (11)

Here we consider computing the solution $\theta(Y, U)$ in (11). The only challenging part is solving for θ_1, so we shall focus on this. Suppose t_2 and u_2 are given, and define a function $r(x) = F_x(t_2) - u_2$; the goal is to find the root for r. One can evaluate $r(x)$ by simulating $\mathsf{Gamma}(x, 1)$ variables, giving a Monte Carlo approximation of $F_x(t_2)$, and the root can then be found with any standard method, e.g., bisection. However, this can be fairly expensive computationally. A more efficient alternative approach is available based on large-sample theory. By Theorem 5.2 in Glaser (1976) and the delta theorem, if n is large, then F_x can be well approximated by a normal distribution function with mean $\psi(x) - \log(x)$ and variance $n^{-1}\{\psi'(x) - 1/x\}$, where ψ and ψ' are the digamma and trigamma functions, respectively. With this normal approximation, it is easy to evaluate $r(x)$ and find the root numerically. Though this is based on a large-sample approximation, in our experience, there is no significant loss of accuracy, even for small n.

The normal approximation discussed above is simply a tool to find the solution $\theta(Y, U)$. It also provides some intuition related to the asymptotic argument in Section 5.1. When n is large, the variance in the normal approximation is $O(n^{-1})$, so the distribution function F_x will have a steep slope in the neighborhood of the solution to the equation $t_2 = \psi(x) - \log(x)$ and, therefore, the root for $r(x)$ will be in that same neighborhood, no matter the value of u_2. The solution to the equation $t_2 = \psi(x) - \log(x)$ is the maximum likelihood estimator of θ_1 (e.g., Fraser et al. 1997), which is consistent. Therefore, when n is large, Equations (**??**) and (**??**) in the paper are essentially the same, so the approximate validity of the corresponding prediction plausibility function is clear.

One last modification that we found to be helpful was to modify that normal approximation discussed above by replacing the normal distribution function with a gamma.

That is, find solutions for the mean and variance of the normal approximation as before, but then use a gamma distribution function with mean and variance matching those obtained for the normal. See the R code available at the first author's website.

References

Fraser, D. A. S., Reid, N., and Wong, A. (1997). Simple and accurate inference for the mean of a gamma model. *Canad. J. Statist.*, 25(1):91–99.

Glaser, R. E. (1976). The ratio of the geometric mean to the arithmetic mean for a random sample from a gamma distribution. *J. Amer. Statist. Assoc.*, 71(354):480–487.

Jensen, J. L. (1986). Inference for the mean of a gamma distribution with unknown shape parameter. *Scand. J. Statist.*, 13(2):135–151.