Bayesian Semiparametric Sales Projections for the Texas Lottery

A. Majumdar and R. Eubank
Arizona State University

Abstract: State lotteries employ sales projections to determine appropriate advertised jackpot levels for some of their games. This paper focuses on prediction of sales for the Lotto Texas game of the Texas Lottery. A novel prediction method is developed in this setting that utilizes functional data analysis concepts in conjunction with a Bayesian paradigm to produce predictions and associated precision assessments.

Key words: curve registration, functional data analysis, posterior inference, predictive distribution.

1. Introduction

State lotteries throughout the U.S market portfolios of games that provide a source of revenue for their respective states. Typically, at least one lottery game is jackpot-driven in the sense that players buy tickets in hopes of winning a publicly advertized top prize. For such games, sales are responsive to the size of the top prize or the *jackpot* level which entails that higher gross revenue will result from larger jackpots. On the other hand, the advertized jackpot figure must be set prior to the actual drawing and, hence, before the sales that are needed to pay winners have been realized. As a result, lotteries find it necessary to carry out sales projections to avoid the negative revenue consequences of arbitrarily setting prize amounts too low or too high. In this paper, we describe one such sales projection technique that was developed for use by the Texas lottery with their signature Lotto Texas game.

The Texas lottery began operation in 1992 with a single game known as Lotto Texas wherein players attempted to match the 6 numbers selected by the Lottery (without replacement) from the numbers 1 to 52. Since that inception point, four other games have been added (including the multi-state Mega Millions game) and the original Lotto Texas ball configurations have been altered on three occasions. The focal data set for this article derives from a period in 2000-2002 where a "6 of 54" scenario was being used with 6 balls (i.e., numbers) being selected

without replacement from a set of 54. This also happens to be the current game configuration.

The drawings for Lotto Texas take place at 10 pm CST on Saturday and Wednesday nights. A new jackpot level needs to be set shortly after each draw upon determination of whether or not there has been a jackpot winner. If one or more player tickets match the selected ball numbers, then the jackpot for the next draw is set at a standard level which has traditionally been 4 million dollars for Texas. If there is no winner on a given Wednesday or Saturday, then the jackpot prize pool carries over to the next draw and a value for the jackpot for the subsequent Saturday or Wednesday draw must be selected. This value is then used in Lottery promotions that include billboards and radio/television spots.

In theory it is quite easy to determine an "optimal" jackpot value given some basic information about lottery sales allocation. Specifically, the ideal jackpot or, equivalently, the largest jackpot that can be supported by sales, is obtained via the formula

$$(\text{payout}) \times (\text{jackpot allocation}) \times (\text{annuity factor}) \times (\text{cumulative sales at draw time}) . \tag{1.1}$$

Here the "payout" is the proportion of sales for the game that is returned to players which was 0.5 for the data we consider. The "jackpot allocation" is the proportion of the total prize pool (i.e., 50% of total game sales) that is allocated to the top tier prize. This was fixed at 0.64 for the time period in our study. For Texas, advertised jackpots are annuitized figures that are obtained by multiplying the present value of jackpot funds by an "annuity factor." This annuity factor depends on interest rates that vary from draw to draw. It typically fell in a range between 1.6 to 1.7 for the data used in our analysis.

Formula (1.1) provides a prescription for setting fiscally responsible jackpots. However, to use it one needs a 3 or 4 day ahead prediction of sales depending on whether the draw is to occur on a Saturday or Wednesday. When this work began, the projections were accomplished using a heuristic, but overall quite effective, approach based on "nearest neighbor" principles. To predict sales, marketing personnel would determine one or more previous *roll cycles* or *runs* (i.e., a sequence of consecutive draws without a jackpot winner) that had similar jackpot levels and sales behavior to the current roll cycle and then use this information to determine a suitable sales level to employ in (1.1). It has been the Texas policy to only increase jackpots in 1 million dollar increments which provides a built in margin for error in the calculations. In addition, the personnel conducting the projections were skilled, well-experienced with their approach and well-informed about the Texas marketing environment. Thus, from a practical perspective, their pointwise predictions were more than satisfactory and unlikely to be improved in any meaningful sense. The goal of this work was rather to provide

something they could not gauge with their methodology: namely, the precision of point-wise predictions.

Figure 1 shows daily sales data corresponding to a particular roll cycle consisting of 35 days and 10 draws. The run began on a Sunday with a 4 million dollar jackpot after a winning ticket was sold prior to a Saturday night draw. The plot then tracks the sales progression over a period of 5 weeks with the jackpot finally being won on the last Saturday of the run. There are 9 Wednesday/Saturday draws shown in the plot for which there were no winning tickets. These junctures are termed *rolls* and represent instances where the jackpot prize pool accumulates and new jackpot levels must be set. The daily sales values depicted in the plot demonstrate a common pattern: sales activity is initially low but becomes more intense as the draw day approaches. This produces a sequence of peaks and valleys with the problem being one of obtaining an estimate for the height of the next peak using only the knowledge that is available at the time of the prior peak in the sequence.
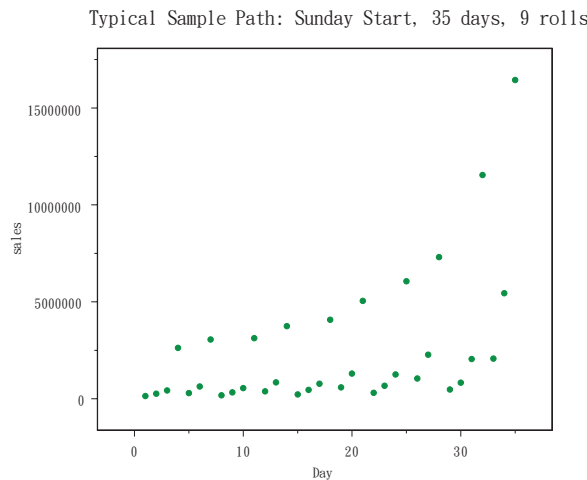

Figure 1: Example sample path

There are many strategies that could be employed for prediction purposes with data such as that in Figure 1. Perhaps the most obvious is to use information from repeated roll cycles to estimate an underlying regression function with sales projections then being obtained by standard mean estimation methods. However, this approach turns out to be too crude to be of much value in this setting because it largely ignores information from the roll cycle that is currently underway in lieu of ensemble information from previous cycles. In contrast, our approach to the problem uses information about mean behavior but incorporates it with

information about the "trajectory" of the current roll cycle to carry out the prediction process.

From our perspective, data sets such as the one in Figure 1 are examples of functional data in the sense of representing discretized readings on a continuous time stochastic process. The development of methodology for analysis of functional data is an area of current research interest due, in large part, to the foundational work of Ramsay and Silverman (2005). Although the techniques we develop here have certain similarities to others that have appeared in the functional data analysis (FDA) literature, our specific approach is, to our knowledge, novel both as a result of its origin as well as its utilization of Bayesian posterior predictions.

In the next section, we describe the methodology we have developed for prediction of Texas Lottery sales. Like most FDA problems there are registration issues that arise and we first consider how to pre-process the data in a way that makes it amenable for use in aggregate inference. We then turn to the development of an appropriate model for the (registered) process mean function and detail how this is to be used in prediction of future sample paths. The actual data analysis and sales projection results are the subject of Section 3. Concluding remarks are collected in Section 4.

2. Methodological Approach

The data in frame (a) of Figure 2 shows two roll cycles corresponding to Sunday and Thursday starts after a *hit* (i.e., a winning ticket was sold) on a Saturday and Wednesday night. While both sample paths experience the same characteristic sales run-ups on draw days, we can see that the peaks are misaligned due to the different starting days. The (b) frame in the figure shows how this effect is manifest over the collection of runs that comprise the entire data set. The gaps in the plot are typical of functional data that require realignment or registration due to the progression of sample paths along different time scales.

There are now some very sophisticated methods for registration of functional data (e.g., Gervini and Gasser 2004 and Brumback and Lindstrom 2004). However, for our case there is a relatively simple solution that can be accomplished through a uniform rescaling of time. The basic idea is to simply make time run somewhat slower from Sunday through Wednesday in a sense that is described in Table 1. This produces a new time or "Day" index that runs from 1 to 6 over any 7 day drawing period regardless of whether a run starts on a Thursday or Sunday. The end product is the satisfactorily registered sample paths that are shown in panel (c) of Figure 2.

(a)

Before Registration

(b)

Unregistered Sample Paths

(c)

Registered Sample Paths

(d)

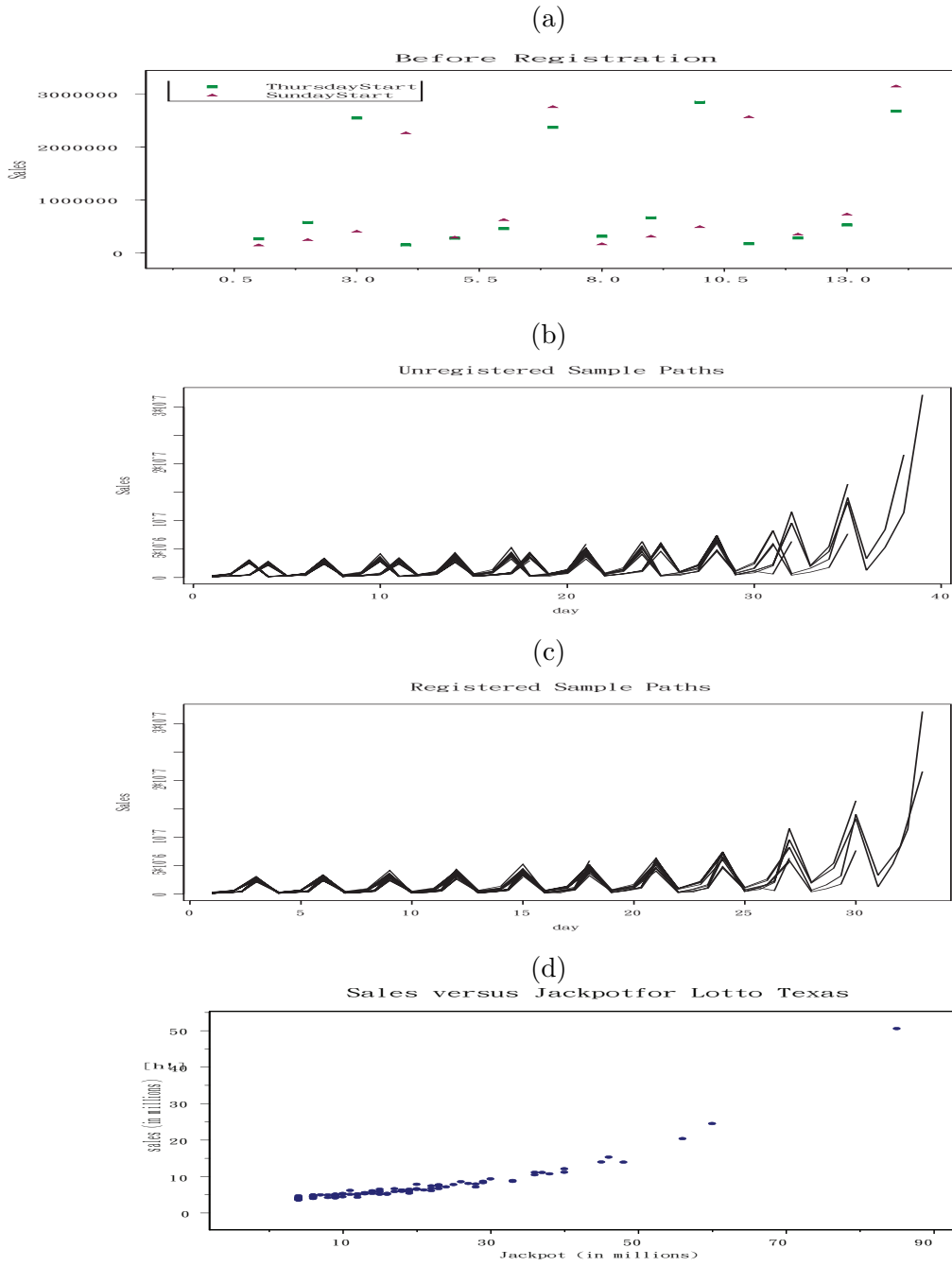**Sales versus Jackpot for Lotto Texas**

Figure 2: Lotto Texas Data: (a) example sample paths, (b) unregistered data, (c) registered data (d) Sales versus jackpot levels.

Table 1: Registered time scale

"Day" Index	1	2		3	4	4 2/3	5 1/3	6
Thurs. Start	Thurs.	Fri.		Sat.	Sun.	Mon.	Tues.	Wed.
Sun. Start	Sun.	Mon.	Tues.	Wed.	Thurs.	Fri.		Sat.
"Day" Index	1	1 2/3	2 1/3	3	4	5		6

2.1 Modeling the mean function

Once the data have been properly registered, the next step is to consider how to model any common trends in the sales sample paths. In this respect our focus will be on modeling the mean function for the (registered) process. As noted previously, the mean function is not by itself sufficient for prediction purposes. However, it nonetheless provides important information about the process behavior that is a key ingredient for our projection methodology.

Some theoretical and empirical studies have been carried out concerning the demand for lotto games in England (e.g., Farrell, *et al* 1999 and Farrell, *et al* 2000) and in Israel (e.g., Beenstock and Haitovsky 2001). One consensus is that the demand (and, hence, sales) for lotto type games are positively influenced by increasing jackpot levels. Figure 3 shows the sales and jackpot figures for the Lotto Texas data in Figure 2 aggregated across roll cycles. This clearly demonstrates that such consideration are equally applicable to the Texas scenario.

Another common factor that has been deemed important in the economic aspects of lotteries is the influence of prize pool rollovers or rolls as we have termed them here. Beenstock and Haitovsky (2001) provide arguments to suggest that the mean sales may have a form similar to $\beta H(t, z, r) + \theta r$, where β, θ are coefficients and H is a nonlinear function in time (i.e., t), jackpot level (i.e., z) and the rollover r from the previous prize pool into the current jackpot. Their empirical work then employs dummy variables to test for the presence of the θr term. When such terms are significant they refer to this as "lottomania" wherein "ticket sales increase by considerably more than implied by the unusually large jackpot." In particular, they find this effect to be present after the third rollover for their data.

As a result of the above discussion, we chose to model the sales roll cycle mean function as

$$\mu(t, z) = \beta(t)z \qquad (2.1)$$

with z representing the jackpot level, t corresponding to the time (in the altered scale of Table 1) of the roll cycle and $\beta(\cdot)$ a smooth, unknown function. Model (2.1) is a special case of a varying coefficient model as considered in Hastie and Tibshirani (1993). The idea is that sales are, on the average, proportional to

jackpot levels but the proportionality factor is allowed to vary over the course of time thereby accounting for the nonlinearity present in Figure 3.

Nonparametric regression considerations dictate the use of a flexible form for estimation of β. In this respect, we approximate it by as piecewise quadratic function with different quadratic segments for each of the biweekly drawing periods that occur over the course of a run. That is, on each interval of the form $[3l - 2, 3l], l = 1, \ldots, L$, we will assume that

$$\beta(t) = \alpha_0^l + \alpha_1^l t + \alpha_2^l t^2. \tag{2.2}$$

Note that no continuity constraints are imposed across segments which allows for twice weekly jumps in mean sales after drawings have occurred. For the data shown in the bottom panel of Figure 2 the longest run is for 33 days which translates into fitting a coefficient function with $L = 11$ segments.

We complete the mean function modeling process by employing priors for the parameters that arise in $\beta(\cdot)$. Specifically, we will assume that $\alpha_0^l, \alpha_1^l, \alpha_2^l, l = 1, \ldots, L$, are mutually independent with common densities $\pi(\alpha_0), \pi(\alpha_1)$ and $\pi(\alpha_2)$ for the intercept, linear and quadratic coefficients, respectively, of each segment. Then, $\pi(\alpha_i)$ is modeled as being normal with mean zero and variance $\sigma_i^2, i = 0, 1, 2$. In lieu of assuming a hierarchical model for σ_0^2, σ_1^2 and σ_2^2, we will instead choose them to be large positive values thereby effectively employing diffuse priors for the segment coefficients.

2.2 Relating the sample paths and the mean function

An examination of the lower frame in Figure 2 suggests that there is a tendency for the sample paths of individual roll cycles to follow the same overall trend in the combined data apart from deviations in starting sales levels and slope or trajectory relative to the "grand mean" for the data. This suggests that we consider a model wherein the runs are sample paths from a stochastic process of the form

$$Y(t) = a + b\mu(t, z(t)) + \varepsilon(t) \tag{2.3}$$

with $Y(\cdot)$ representing sales, $z(\cdot)$ the concomitant jackpots, $\epsilon(\cdot)$ a normal white noise process (i.e., the mean of $\varepsilon(\cdot)$ is zero and the covariance between $\varepsilon(t)$ and $\varepsilon(t')$ is zero for $t \neq t'$) with variance σ_ϵ^2 and a, b are random intercepts and slopes, respectively, that are independent of $\varepsilon(\cdot)$. If we are to observe m runs then, conditional on the realized values $(a_1, b_1), \ldots, (a_n, b_n)$ of (a, b) in (2.3), the ith run will evolve from the model

$$Y_i(t) = a_i + b_i\mu(t, z(t)) + \varepsilon_i(t) \tag{2.4}$$

with $\varepsilon_i(\cdot)$ a normal white noise process. For a mature Lotto game, such as the one in Texas, the results from one roll cycle typically have little effect on subsequent

sales levels. Accordingly we will treat the a_i, b_i and $\varepsilon_i(\cdot)$ that come from different roll cycles as being uncorrelated.

The vectors $(a_j, b_j), j = 1, \ldots, m$ in (2.4) are each assumed to have bivariate normal distributions, $\pi(a_j, b_j | \Sigma)$, with common mean vector $(0,1)^T$ and common variance-covariance matrix Σ. Let

$$\text{Symm}(a, b, c) = \begin{bmatrix} a & b \\ b & c \end{bmatrix}. \tag{2.5}$$

Using this notation, $\Sigma \equiv \text{Symm}(\sigma_a^2, \sigma_{ab}, \sigma_b^2)$ is assigned an Inverse Wishart prior $\pi(\Sigma) = W^{-1}(\Psi, m')$, where $\Psi = I_{2 \times 2}$ the identitiy matrix and $m' = 4$. This entails that the prior mean is $\text{E}[\Sigma] = I_{2 \times 2}/(m' - 2 - 1)$ and the prior variances for each entry of Σ are infinite. The prior density for the error variance $\pi(\sigma_\epsilon^2)$ is chosen to be an Inverse Gamma with hyper-parameters γ_ϵ and δ_ϵ. In our analysis of the Texas Lottery data presented in the next section, we choose the shape parameter $\gamma_\epsilon = 2$ and the scale parameter $\delta_\epsilon = 1$. This choice sets the prior mean as 1 and the prior variance as infinity. These particular priors were found to perform better in terms of posterior distribution and prediction than other choices of priors from the same family with the same mean that had finite variances. All of these priors are proper and yield conjugate posterior distributions. The priors and posterior distributions as well as the posterior prediction densities under the Bayesian model are discussed in more detail in the Appendix.

3. Results from Posterior Inference

In this section we summarize the results of Bayesian posterior inference obtained through application of the results in the Appendix to model (2.4) using the Lotto Texas sales data from Figure 2. We begin with a discussion of the parameter simulations before turning to issues related to sales prediction. We have used sales in millions of dollars as our unit of measurement.

Realizations were generated from the posterior distibution of the parameters in the models (and for predictions) using a Monte Carlo Markov chain (MCMC) approach. For this purpose 5000 burn-in samples were employed after which 100,000 samples were obtained. These later samples were subsequently thinned out by 10 iterations.

Table 2 gives the quantiles obtained by sampling from the posterior distributions of Σ and σ_ε. In particular, we see from this that σ_{ab}, the covariance between the intercept a_j and b_j for $j = 1, \ldots, m$ is non-significant since its 95% posterior credible interval contains 0.

Table 2: Posterior quantiles

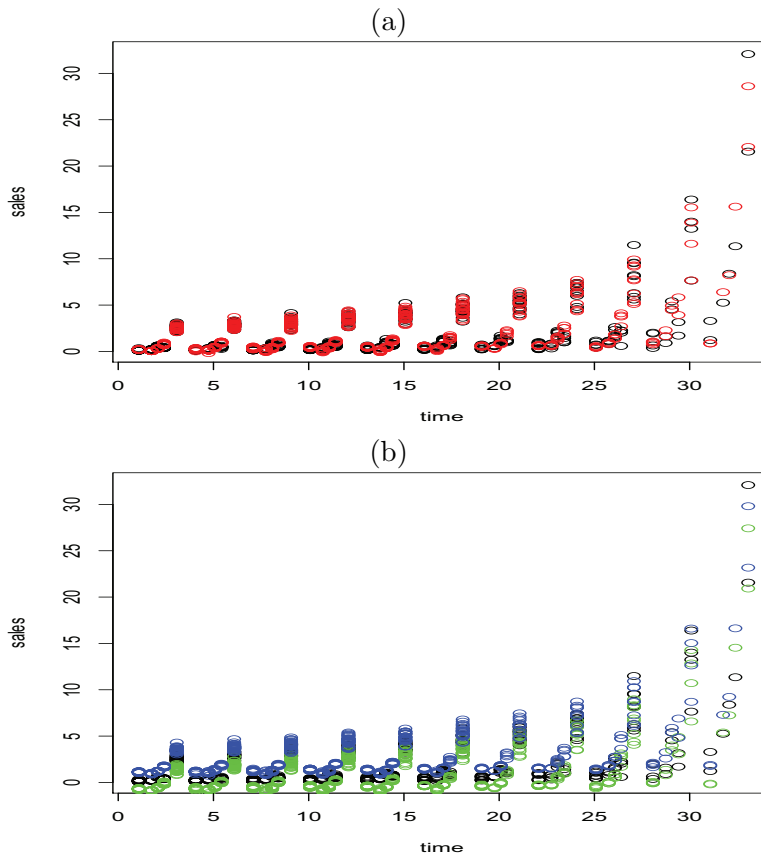| Parameters	2.5%	50.0%	97.5 %
$\sigma_a{}^2$	0.229	0.471	2.043
$\sigma_b{}^2$	1.390	2.450	4.276
σ_{ab}	-1.589	0.143	2.096
$\sigma_\epsilon{}^2$	1.012	1.108	1.220

(a)

(b)

Figure 3: (a) Real (black circles) vs. median predicted(red circles) sales. (b) Real (black circles) sales vs. 95% posterior predictive intervals – lower bounds are green circles and upper bounds blue circles.

Figure 3 summarizes the results of "predictions" for the observed Lotto Texas sales. Here we use the posterior parameter simulations described in the Appendix to produce 95% prediction intervals for the observed sales. The results are overall quite good with 96.8% of the intervals containing the actual sales. As was true for estimation of the coefficients for β, lack of data likely contributed to the problems

the prediction intervals encountered for a few of the very large jackpot and we
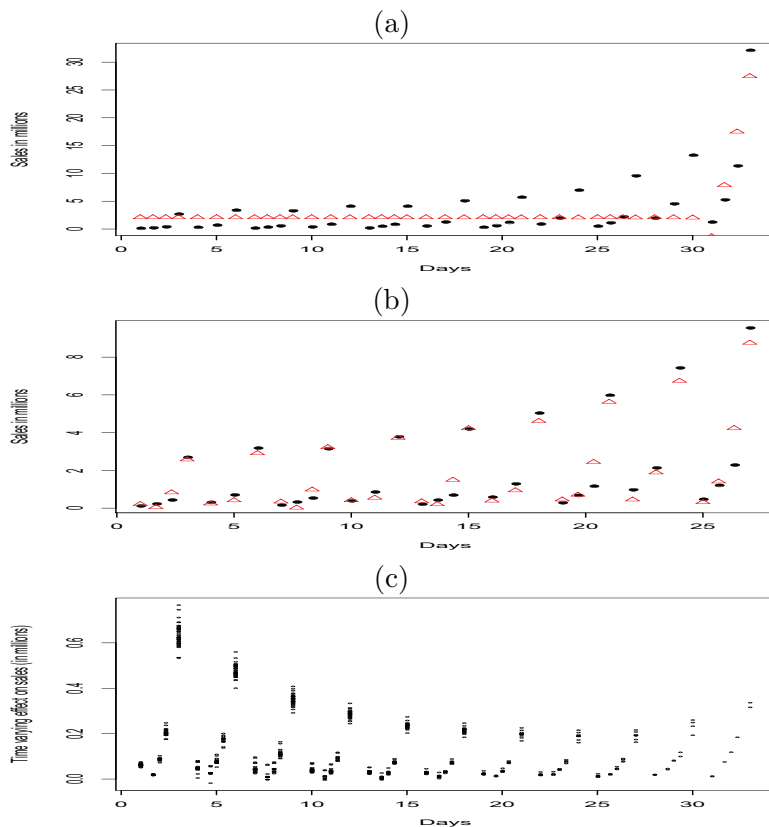will now examine this issue in more detail.

Figure 4: (Real(black dots) and median predicted sales(red triangles) for the
(a) 9th and (b) 5th runs (c) Plot of time versus median posterior values
$\mathrm{Med}(\mathrm{E}[Y_i(t_{ij})]/z_{ij})$.

We notice that of the 58 runs, 3 runs have very different values for their
intercepts a_j and slopes b_j in Figure 4. These are the 9th, 7th and 53rd run
which also happen to be the only runs that have observations in the last time
interval [31, 33]. In panel (a) of Figure 2 we plot the real sales (represented
by black dots) versus the median of the posterior means (represented by the
red triangles) for one of these (i.e., the 9th) runs that typifies what happens for
all three cases. To compare predictions, we also show one other run (i.e., the
5th) which contain no observations from the last time-interval in panel (b) of
the figure. This demonstrates that the poor predictions in the last time interval
represent isolated cases rather than an inherent problem with the method. From
a practical perspective, the last time interval corresponds to unusual situations

where unexpectedly long runs have produced unprecedently large jackpot levels. Conservative sales projections enacted by senior lottery personnel would generally be used for such instances in any event.

4. Discussion

In this paper we have reported on what we consider to be a successful blend of FDA and Bayesian methods for solution of a real world problem. Although the particular application concerns sales projection for lotteries, the problem of prediction from functional data is quite general and this basic approach might be expected to be adaptable to a much more general context. We intend to pursue this in future work.

The landscape of the Texas Lottery games changed significantly with the entrance of the muli-state Mega Millions game into their game portfolio in December of 2003. Mega Millions follows a "power ball" format with large odds that produce very large jackpot. Like Lotto Texas, sales for Mega Millions are jackpot driven and, as a result, there has been significant sales canibalization for Lotto Texas in the period since Mega Millions became available. The main consequence this has for our analysis here is that the present projection algorithm is no longer appropriate. At the very least it become necessary to include the advertized jackpot for Mega Millions into the process as a covariate. This is another topic that is currently under investigation.

Finally, let us return to the "lottomania" phenomenon that was discussed in Beenstock and Haitovsky (2001). A question arises as to what empirical symptoms would be realized in our analysis that could be indicative of the on-set of "lottomania." The answer may lie in in panel (c) ou Figure 4 that shows the pattern of sales wherein the effect of jackpots on prediction has been eliminated. Here we have plotted median predicted sales divided by the jackpot values across time for all runs. This suggest that the sales slopes tend to initially decline as jackpots increase until they flatten and begin to increase for very long runs that also have very large jackpots. The implication is that for typical, shorter runs the proportional effect on sales of increasing jackpots is diminishing. However, eventually this trend is reversed and we postulate that the associated change point may represent the beginning of "lottomania."

Appendix

Here we provide a detailed derivation of the posterior distributions that were used for the developments in Section 3. We initially focus on the parameters in model (2.4) and then describe how posterior simulations obtained for the parameters can be used for prediction purposes.

Let $\mathbf{y}$ denote the vector of sales values arranged in sub-vectors corresponding to each roll cycle and ordered by time within each sub-vector: i.e.,

$$\mathbf{y}^T = (\mathbf{y}_1^T, \dots, \mathbf{y}_m^T)$$

with

$$\mathbf{y}_j = (y_{1j}, \dots, y_{r_j j})^T$$

the results of the jth run. The vector of parameters for data from model (2.4)

$$\boldsymbol{\theta}^T = (a_1, b_1, \dots, a_m, b_m, \alpha_0^1, \alpha_1^1, \alpha_2^1, \dots, \alpha_0^L, \alpha_1^L, \alpha_2^L, \sigma_\epsilon^2, \Sigma) \qquad (A.1)$$

then has the associated likelihood function

$$L(\mathbf{y}|\boldsymbol{\theta}) \propto \sigma_\epsilon^{-n} \exp\left(-(2\sigma_\epsilon^2)^{-1} \sum_{j=1}^m \sum_{i=1}^{r_j} (y_{ij} - a_i - b_i z_{ij} \beta(t_{ij}))^2\right) \qquad (A.2)$$

with $(t_{ij}, z_{ij}), i = 1, \dots, r_j$, being the day indices and jackpots associated with sales for the jth run. The developments in Sections 2.1–2.2 now entail that the joint distribution of $\mathbf{y}$ and $\boldsymbol{\theta}$ is the product of $L(\mathbf{y}|\boldsymbol{\theta})$ with the prior density

$$\pi(\boldsymbol{\theta}) = \prod_{j=1}^m \pi(a_j, b_j | \Sigma) \ \pi(\Sigma) \prod_{l=1}^L \pi(\alpha_0^l) \pi(\alpha_1^l) \pi(\alpha_2^l) \ \pi(\sigma_\varepsilon^2).$$

The next step is to derive the posterior densities for each of the components of (A.1) that will allow us to carry out Bayesian inference via the multi-stage Gibbs sampler (e.g., Robert and Casella 2004, Chapter 10). For this development we will employ the symbol "\" to indicate removal of a particular parameter or parameters from $\boldsymbol{\theta}$.

We begin the marginalization process by noting that the posterior density for σ_ε^2 is

$$\pi(\sigma_\varepsilon^2 | \mathbf{y}, \boldsymbol{\theta} \backslash \sigma_\varepsilon^2) \quad \propto \quad \sigma_\varepsilon^{-n} \exp\left(-\sum_{j=1}^m \sum_{i=1}^{r_j} (y_{ij} - a_j - b_j \ \mu(t_{ij}, z_{ij}))^2 / 2\sigma_\varepsilon^2\right)$$
$$\times \sigma_\varepsilon^{-2(\gamma_\epsilon + 1)} \exp\left(-\delta_\epsilon / \sigma_\epsilon^2\right).$$

From this it follows that $\sigma_\varepsilon^2 | \mathbf{y}, \boldsymbol{\theta} \backslash \sigma_\epsilon^2$ has an inverse gamma distribution with parameters $2^{-1} n + \gamma_\epsilon$ and $2^{-1} \sum_{j=1}^m \sum_{i=1}^{r_j} (y_{ij} - a_j - b_j \mu(t_{ij}, z_{ij}))^2 + \delta_\epsilon)$.

For the joint posterior distribution of (a_j, b_j), $j = 1, \dots, m$, we obtain

$$\pi(a_j, b_j | \mathbf{y}, \boldsymbol{\theta} \backslash (a_j, b_j)) \quad \propto \quad \exp\left(-(2\sigma_\epsilon^2)^{-1} \sum_{i=1}^{r_j} (y_{ij} - a_j - b_j \mu(t_{ij}, z_{ij}))^2\right)$$
$$\times \exp\left(-2^{-1} (a_j, b_j) \Sigma^{-1} (a_j, b_j)^T\right).$$

This leads to $(a_j, b_j)^T | \mathbf{y}, \boldsymbol{\theta} \setminus (a_j, b_j)$ having a bivariate normal distribution with posterior mean

$$\tilde{\boldsymbol{\mu}} = \tilde{\Sigma} \left(\sigma_\varepsilon^{-2} \sum_{i=1}^{r_j} y_{ij}, \sum_{i=1}^{r_j} y_{ij} z_{ij} \right)^T$$

and posterior variance-covariance matrix

$$\tilde{\Sigma} = \left[\Sigma + \sigma_\epsilon^{-2} \text{Symm} \left(r_j, \sum_{i=1}^{r_j} z_{ij}, \sum_{i=1}^{r_j} z_{ij}^2 \right) \right]^{-1}.$$

The posterior distribution of α_k^l, $k = 0, 1, 2$, $l = 1, \dots, L$, can be derived as follows. First, observe that

$$\pi(\alpha_k^{\ l} | \mathbf{y}, \boldsymbol{\theta} \setminus \alpha_k^l) \quad \propto \quad \exp \left(-(2\sigma_\epsilon^2)^{-1} \sum_{j=1}^m \sum_{i=1}^{r_j} (y_{ij} - a_j - b_j \mu(t_{ij}, z_{ij}))^2 \right)$$
$$\times \exp \left(- \left(\alpha_k^l \right)^2 / \left(2\sigma_k^2 \right) \right).$$

From this we then deduce that $\alpha_k^l | \mathbf{y}, \boldsymbol{\theta} \setminus \alpha_k^l$ is normal with mean

$$\tilde{\mu} = \frac{\tilde{\sigma}^2}{\sigma_\epsilon^{\ 2}} \sum_{j=1}^m \sum_{i=1}^{r_j} \sum_{t_{ij} \in [3l-2, 3l]} \left((y_{ij} - a_j) \, b_j z_{ij} t_{ij}^k \right)$$

and variance

$$\tilde{\sigma}^2 = \left(\sigma_k^{-2} + \sigma_\epsilon^{-2} \sum_{j=1}^m \sum_{i=1}^{r_j} \sum_{t_{ij} \in [3l-2, 3l]} \left(b_j z_{ij} t_{ij}^k \right)^2 \right)^{-1}.$$

Finally, the posterior for Σ can be deduced from

$$\pi(\Sigma | \mathbf{y}, \boldsymbol{\theta} \setminus \Sigma)$$
$$\propto \exp \left(-\frac{1}{2} \sum_{j=1}^m (a_j, b_j)^T \Sigma^{-1} (a_j, b_j) \right) \det(\Sigma)^{-m/2} \times \pi(\Sigma)$$
$$\propto \exp \left(-\frac{1}{2} \text{trace}(\text{Symm}(\sum_{j=1}^m a_j^{\ 2}, \sum_{j=1}^m a_j b_j, \sum_{j=1}^m b_j^{\ 2}) \Sigma^{-1}) \right) det(\Sigma)^{-m/2}$$
$$\times \exp \left(-\frac{1}{2} \text{trace}(\Psi^{-1} \Sigma^{-1})) \right) \det(\Sigma)^{-(m'+2+1)/2}$$
$$= \exp \left(-\frac{1}{2} \text{trace} \left(\left\{ \text{Symm}(\sum_{j=1}^m a_j^{\ 2}, \sum_{j=1}^m a_j b_j, \sum_{j=1}^m b_j^{\ 2}) + \Psi^{-1} \right\} \Sigma^{-1} \right) \right)$$
$$\times \det(\Sigma)^{-(m+m'+2+1)/2}.$$

Thus, the posterior distribution for Σ is inverse Wishart and given by $\pi(\Sigma|\mathbf{y},\boldsymbol{\theta}\backslash\Sigma)$ $= W^{-1}(\tilde{\Psi},\tilde{m})$, where $\tilde{\Psi} = \left\{\text{Symm}(\sum_{j=1}^{m}a_j{}^2, \sum_{j=1}^{m}a_jb_j, \sum_{j=1}^{m}b_j{}^2) + \Psi^{-1}\right\}^{-1}$ and $\tilde{m} = m + m'$.

In order to find the posterior distribution of $\text{E}(Y_i(t)|\boldsymbol{\theta})/z_i(t)$, from (2.4) and (2.1) we note that $\text{E}(Y_i(t)|\boldsymbol{\theta}) = a_i + b_i\mu(t,z_i(t)) = a_i + b_i\beta(t)z_i(t)$. Thus, from the S subsamples $\boldsymbol{\theta}^s, s = 1, \ldots, S$, generated in the MCMC step we can further generate subsamples of the posterior distribution of $\text{E}(Y_i(t)|\boldsymbol{\theta})/z_i(t)$. That is, we can compute $[a_i^s + b_i^s\beta^s(t)z_i(t)]/z_i(t), s = 1, \ldots, S$ for any particular choice of t and jackpot value $z_i(t)$. We then use these S posterior subsamples to approximate the posterior median in panel (c) of Figure3 4. In order to find the posterior predictive distribution of $Y_i(t)$, we use a similar algorithm to get S subsamples $a_i^s + b_i^s\beta^s(t)z_i(t) + e_i^s(t), s = 1, \ldots, S$ where $e_i^s(t)$ is simulated from a zero mean, normal distribution with variance equal to the σ_ϵ^2 component of $\boldsymbol{\theta}^s$. We can then find the median and the 95% posterior predictive intervals by evaluating the corresponding 2.5th and 97.5th percentiles of the the posterior predictive distribution.

Acknowledgement

References

Beenstock, M. and Haitovsky, Y. (2001). Lottomania and other anomalies in the market for lotto. *J. Econ. Psych.* **22**, 721-744.

Brumback, L. and Lindstrom, M. (2004). Self modeling and flexible, random time transformations. *Biometrics* **60**, 351-363.

Farrell, L., Hartley, R., Lanot, G. and Walker, I. (2000). The demand for lotto: the role of conscious selection. *J. Bus. Econ. Statist.* **18**, 228-241.

Farrell, L., Morgenroth, E. and Walker, I. (1999). A time series analysis of U.K. lottery sales: long and short run price elasticities. *Oxford Bull. Econ. Statist.* **61**, 513-526.

Gervini, D. and Gasser, T. (2004). Self modeling warping functions. *J. Roy. Statist. Soc. B* **66**, 959-971.

Hastie, T. and Tibshirani, R. (1993). Varying coefficient models. *J. Roy. Statist. Soc. B* **55**, 757-796.

Ramsay, J. and Silverman, B. (1997). *Functional Data Analysis*. Springer.

Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods, Second Edition*. Springer.

A. Majumdar
Arizona State University
Tempe, AZ 85287-1804, USA
ananda@math.asu.edu

R. Eubank
Arizona State University
empe, AZ 85287-1804, USA
ananda@math.asu.edu