CASE STUDY

A statistical learning exercise based on a modified Rock-Paper-Scissors game

Paola Bortot, University of Bologna, Bologna, Italy. Paola.Bortot@unibo.it.
Stuart Coles, Smartodds Limited, London, England. Stuart.Coles1111@gmail.com.

Abstract

The standard version of the game Rock-Paper-Scissors is interesting in terms of game theory, but less so in terms of Statistics. However, we show that with a small rule change it can be made into an interactive exercise for degree-level students of Statistics that leads to a Bayesian change-point model, for which the Gibbs sampler provides an intuitive method of inference. First, students play the game to generate the data. Second, they are encouraged to formulate a model that reflects their experience from having played the game. And third, they participate in the development of a suitable MCMC algorithm to fit the model.

Keywords: Bayesian Statistics, change-point analysis, Gibbs sampler, teaching.

1. Introduction

The processes of data collection, model building and inference are the main themes of any statistical analysis. For teachers of Statistics, however, there are few opportunities to involve students in the whole operation. This is especially true for analyses requiring advanced statistical techniques — the physical and time constraints of standard teaching environments are simply not conducive to this aspect of statistical learning. The aim of this article is to suggest a teaching exercise which illustrates the entire sequence of a statistical analysis from data collection to inference, with model and inference developed from a knowledge of the data-generating mechanism that is also part of the exercise. We have run the exercise ourselves with degree-level students following a course in Computational Statistics, though it could work equally well as a practical lesson in a Bayesian inference course. In either case, some basic knowledge of Bayesian Statistics is required, as is an understanding of the Gibbs sampler and Markov chain Monte Carlo (MCMC) in general. Our objective is to provide an interactive platform through which students can see and exploit the links between Bayesian theory, model building and simulation-based inference.

The starting point for our developments is the well-known game of Rock-Paper-Scissors. Each of two players, labeled A and B respectively, simultaneously selects by hand caricature one of the three elements Rock, Paper and Scissors, which we abbreviate to $\{R, P, S\}$. If both players select the same element the game is a draw; otherwise, using obvious notation, $R > S$, $S > P$ and $P > R$. Several rounds are usually played, and the overall winner is the player with the most rounds won. The possible outcomes of the game are labeled A, X and B, corresponding to a win for A, a draw, and a win for B, respectively. Clearly, if each player chooses between $\{R, P, S\}$ at random, the A/X/B probabilities are identically 1/3. In general we write $\theta = (\pi_A, \pi_X, \pi_B)$ to denote the vector of A/X/B probabilities, and use θ_0 to denote the special vector $(1/3,\ 1/3,\ 1/3)$.

From a game-theoretic point of view, Rock-Paper-Scissors is a simple zero-sum game whose Nash equilibrium solution corresponds to each player playing the elements of $\{R, P, S\}$, each with probability 1/3 (see van den Nouweland, 2007, for example). There have also been various statistical studies of Rock-Paper-Scissors. For example, Wang, Xu and Zhou (2014) compare player behaviour in laboratory conditions, with expected behaviour under optimal Nash Equilibrium rules. At a more general level, Walker and Walker (2004) provide a strategy handbook for players. They describe, for example, empirical evidence suggesting that players generally tend to choose S on only around 30% of occasions. Knowing this, and assuming everything else equal, opponents have a slight advantage in selecting P more often than random selection would imply. However, strategy issues like this are more interesting for their psychological or game-theoretical implications than for their statistical relevance.

2. A Modified Version of Rock-Paper-Scissors

Though the standard version of Rock-Paper-Scissors is of limited interest from a statistical point of view, a modification of the rules can lead to versions of the game that are more stimulating. Specifically, we suggest a novel variation which limits the play options for one of the players. In this way the other player has an advantage, but only once they correctly identify the limitations imposed on their opponent. The idea is then to challenge students to play this version of the game and develop a statistical model that will allow an observer to learn about the way the game has been played from the collected data.

Students are placed in groups of three, comprising two players and an observer. The role of the observer is to record the sequence of match outcomes, A/X/B, though not the actual play choices of either player. Once all rounds have been played, these are the data that form the basis of the analysis.

Before the game starts the two players are randomly given one of two instructions, I1 and I2. They are additionally told that one set of instructions corresponds to complete freedom in gameplay, but that the other imposes restrictions. In practice, we write the instructions on cards and ask each player to randomly select one without replacement. The instructions are:

I1: You are free to choose from $\{R, P, S\}$ in every round;

I2: You may only choose from $\{P, S\}$ in every round;

and without loss of generality we can assume these are assigned to players A and B respectively. Therefore, Player A is playing to standard rules, while Player B is prohibited from playing R. As such, player B is at a disadvantage, but only once player A has deduced the limitation imposed on player B. The results themselves are analysed from the point of view of the observer, who sees only the sequence of A/X/B match outcomes, and is unaware of which player has which instruction, and indeed what the limitation is in instruction I2.

Though players are not constrained to play randomly from the options available to them, our experience is that they tend to do so, at least approximately. Under random selection from the available sets, the A/X/B probabilities are still $\theta = \theta_0$. However, it is likely that Player A will eventually realise after a number of rounds that their opponent never plays R, and deduce that they have an improved strategy by selecting just from $\{R, S\}$. The logic is that since player B never plays R, it is wasteful for player A to ever play P. If both players then select randomly from their reduced sets, it is easy to check that $\theta = (1/2,\ 1/4,\ 1/4)$. In real play, where players may choose not to make random plays, θ may differ slightly from this theoretical value, just as the standard version may have probabilities that differ from θ_0. In practice we have found that playing the game for 100 rounds provides a reasonable chance for player A to identify the limitations of player B and to change strategy accordingly.

Once the game is played, students are asked to develop a statistical model based only on the A/X/B data with the objectives:

1. To assess whether there has been a change of strategy during the 100 rounds, and if so to identify where it occurred;
2. To estimate $\theta = (\pi_A, \pi_X, \pi_B)$ for each round, accounting for the fact that there may have been a change of strategy for one of the players at some point.

3. Model Building

Getting students to play the game themselves serves two purposes. First, to obtain the data; second, to provide students with an experience of the data-generating process which, in turn, assists with appropriate model building. In practice what we have found is that some student pairs have made no change to strategy within the 100 rounds, and others have attempted several strategy changes. In most pairs, however, player A realises their advantage within the allocated 100 rounds and changes their play accordingly.

The discussion of these various playing strategies is an integral part of the exercise. To simplify the model development, we ask students to make three assumptions when model-building:

1. Player B maintains the same strategy at all rounds;
2. Player A also maintains a single strategy, except possibly at one point where they realise their potential advantage and change strategy for the remaining rounds;
3. In all rounds, both players make random choices from either all of, or a subset of, the options available to them.

These are reasonable assumptions from both a game-theoretic and statistical point of view, but they may be inconsistent with some players' actual strategy. This point itself can generate interesting discussion, but the bottom line is that simplifying assumptions of this type are necessary to construct a model which is both feasible and meets the stated objectives.

Step-by-step, students can be led to the natural model that these assumptions imply: a change-point model with at most one single unknown change-point corresponding to the round in which player A exploits their advantage and no longer plays R. In greater detail:

1. The game consists of n rounds, each of which is a multinomial trial:
$$Y_i \mid \theta_i \sim \text{Multinomial}(1, \theta_i), \;\; i = 1, \ldots, n,$$
where the levels of Y_i are A/X/B with probabilities given by the vector θ_i;
2. There is an unknown change-point k such that for $i \leq k$, $\theta_i = \theta^{(1)}$, while for $i > k$, $\theta_i = \theta^{(2)}$;
3. There is the possibility that $k > n$, corresponding to the situation where no change of strategy occurs within the n observed rounds;
4. In the early rounds the vector of A/X/B probabilities is likely to be close to θ_0, regardless of the strategies assigned to the players;
5. There is likely to be a change in the pattern of A/X/B results as one of the players discovers their superior strategy;
6. A priori we have no information about $\theta^{(2)}$;
7. Any change in the pattern of results is likely to occur within a reasonable number of rounds, but unlikely to occur within the first few rounds.

In terms of inference there is a strong argument to be made for the use of a Bayesian rather than a classical model (Killick, 2011, for example). The arguments are two-fold, and are worth elaborating with the students. The first argument is technical: Bayesian methods are better suited than classical methods for change-point problems, since they naturally admit marginalising over the uncertainty in the change-point. The second argument is practical: we have different knowledge about the A/X/B probabilities both previous to and after any possible change-point, and this is much more naturally expressed via a Bayesian model. Previous to the change-point, for reasons discussed above, we anticipate θ_i to be close to θ_0. On the other hand, exchangeability in the players implies that θ_0 remains a reasonable 'best guess' for θ_i even after the change-point, though there is no reason to believe that the actual probabilities will be close to this value. Lack of information on $\theta^{(2)}$ implies that any element of the valid space for the A/X/B probabilities,

$$\Delta_2 = \{(z_1, z_2, z_3): z_1 \geq 0,\ z_2 \geq 0,\ z_3 \geq 0,\ z_1 + z_2 + z_3 = 1\},$$

is equally plausible for $\theta^{(2)}$ prior to observing the data.

The next point we try to emphasise to students is the interplay between model structure and inference. What we are aiming for is an understanding that while a fundamental aspect of Bayesian inference is the inclusion of prior knowledge through prior distribution specification, such knowledge is generally limited to summaries of centrality and variability. More precise details about the shape of the prior distribution can be selected on grounds of computational convenience, which usually implies exploiting conditional conjugacy. For the multinomial change-point model this means choosing Dirichlet prior distributions for $\theta^{(1)}$ and $\theta^{(2)}$, most conveniently parametrised as

$$\theta \sim \text{Dirichlet}(\phi, d),$$

where $\phi \in \Delta_2$ is the mean and $d > 0$ is a dispersion parameter. Full definitions and properties are given in the Appendix. For the change-point model we then set

$$\theta^{(1)} \sim \text{Dirichlet}(\theta_0, d_1),\ \ \theta^{(2)} \sim \text{Dirichlet}(\theta_0, d_2),$$

independently.

The extent to which students can come up with these choices themselves depends on whether they have studied conjugacy both in general, and specifically in the context of multinomial models. Nonetheless, the arguments in favour of these choices are easily understood:

1. The support of the Dirichlet distribution, Δ_2, coincides with the domain of $\theta^{(1)}$ and $\theta^{(2)}$;
2. The parameter choices ensure that θ_0 is the prior mean for both $\theta^{(1)}$ and $\theta^{(2)}$;
3. The scale parameters d_1 and d_2 enable flexibility in the prior distributions for the θ parameters. Setting $d_2 = 3$ gives a uniform prior on Δ_2 for $\theta^{(2)}$, but specifying a considerably larger value for d_1 leads to a greater concentration of the prior distribution of $\theta^{(1)}$ around θ_0.

The remaining parameter is the change-point, k, whose theoretical domain is the entire set of positive integers, as the model assumes that a change will occur, even if this might happen after the allotted n rounds. However, the Gibbs sampler is considerably simplified (see Section 4 below) if the prior distribution for k is bounded above at some pre-specified value k_{max}, which might be chosen to be much greater than n. Apart from this restriction, any choice can be made that is consistent with the prior knowledge that Player A is unlikely to learn their optimal strategy in the first few rounds, but is also unlikely to need very many rounds to learn it. To account for these aspects, we suggest a truncated Negative Binomial model for the prior probability function of the change-point:

$$h(k) \propto g(k; m, v),\ \ k = 1, \ldots, k_{max},$$

for some value of $k_{max} \geq n$, where $g(.; m, v)$ is the probability function of the Negative Binomial distribution parametrised in terms of mean m and variance v. This choice affords considerable flexibility in prior elicitation for the change-point through the specification of m, v and k_{max}. However, it is necessary that $m \leq v$ to satisfy the validity requirements of the Negative Binomial distribution. Furthermore, to avoid a monotonically decreasing prior distribution with mode at 1, the additional constraint $v \leq m^2 + m$ should also be respected.

4. Model Inference

Assuming students have some background knowledge of Gibbs sampling, they can formulate, at least in outline, the following steps, all of which are implicitly conditional on the observed data:

1. Choose arbitrary initial values for k, $\theta^{(1)}$ and $\theta^{(2)}$. Then iterate over the following steps:
2. Given k, simulate from
$$\theta^{(1)} \sim \text{Dirichlet}(\theta_0 + c_1, d_1),$$
where $c_1 = (c_1^{(A)}, c_1^{(X)}, c_1^{(B)})$ is the vector counts of the outcomes A, X and B, respectively, among $y_1, \ldots, y_k$.
3. Similarly, again given k, simulate from
$$\theta^{(2)} \sim \text{Dirichlet}(\theta_0 + c_2, d_2),$$
where $c_2 = (c_2^{(A)}, c_2^{(X)}, c_2^{(B)})$ is the vector counts of the outcomes A, X and B, respectively, among $y_{k+1}, \ldots, y_n$, with the convention that $c_2 = (0,0,0)$ if this set is empty;
4. Since k is discrete and bounded, given $\theta^{(1)}$ and $\theta^{(2)}$, simulate from the full conditional probability function which, up to proportionality, is given by
$$f(k) \propto h(k) \prod_{i=1}^{k} f(y_i \mid \theta^{(1)}) \prod_{i=k+1}^{k_{max}} f(y_i \mid \theta^{(2)}).$$

In this expression h is the prior change-point probability function and $f(y_i \mid \theta^{(1)})$ and $f(y_i \mid \theta^{(2)})$ are, respectively, the multinomial probability functions before and after round *k*, subject to the convention that $f(y_i \mid \theta^{(2)}) = 1$ whenever $i > k_{max}$. In summary, the multinomial-Dirichlet conjugacy has been exploited to enable simple updates of the multinomial probability vectors given the current value of the change-point, while the change-point itself is updated via an enumeration of the full conditional probabilities, which is feasible because its support is discrete and bounded.

We assume that students are familiar with the Gibbs sampler and issues about mixing and convergence of MCMC series. The neat structure of the above model leads to a Gibbs sampler that behaves well in both these respects. For protocol we assume a small burn-in period, discarding the first few simulations of the simulated Markov chain, but even this is not strictly necessary.

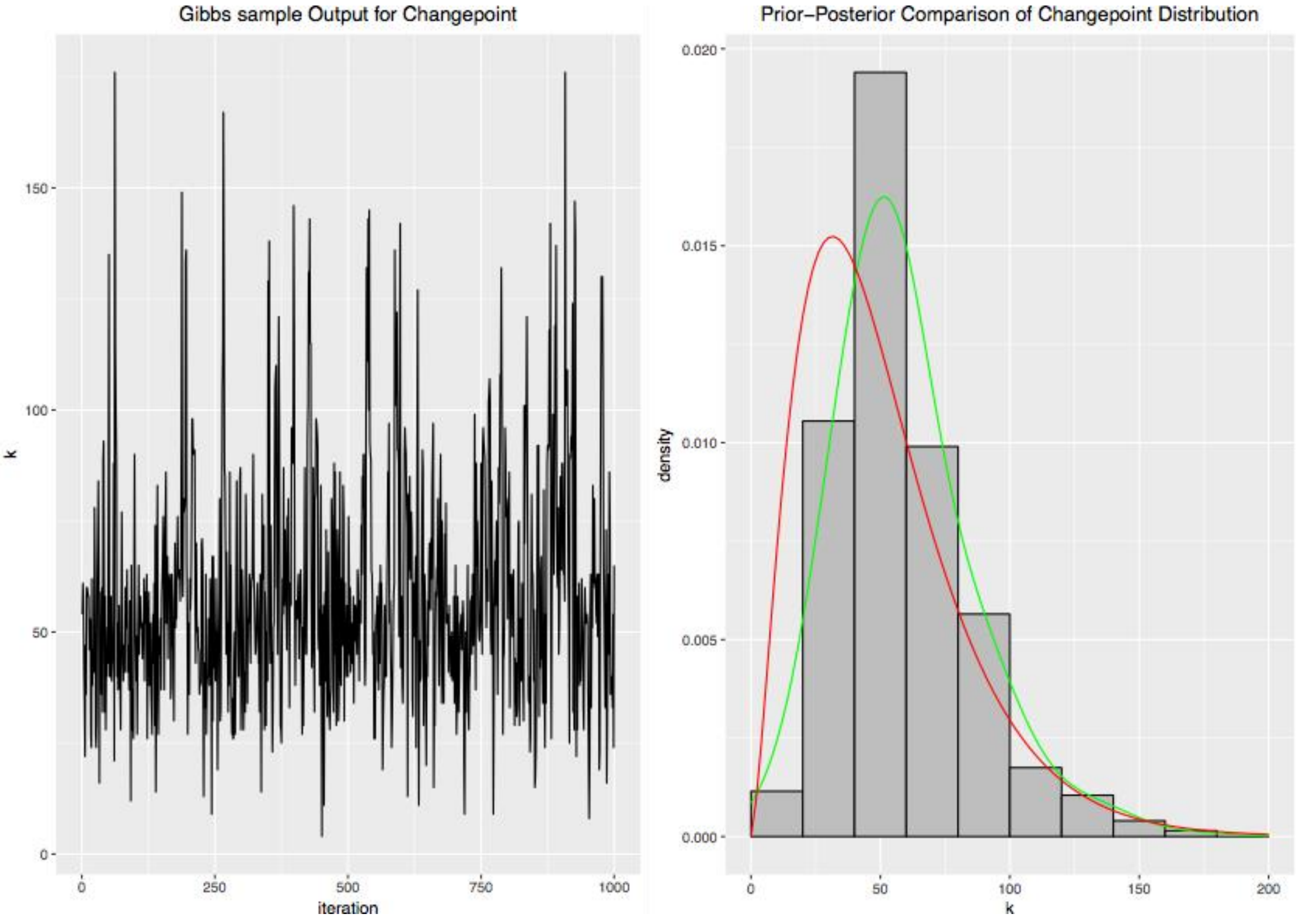

Figure 1. *Left*: Gibbs sample output of change-point parameter. *Right*: Comparison of prior (red) and posterior (green or histogram) distributions of change-point parameter.

5. Analysis of Data

Though students have generated their own data, it is easier for practical purposes to demonstrate the inference on simulated data. Using the statistical language R (R Core Team, 2017), functions for both simulating the data and fitting the model via the Gibbs sampler described above – either to genuine or simulated data – are available as supplementary material at the journal website. The only package required outside of the base R language is ggplot2 (Wickham, 2016), which we use to enable improved graphics.

As an illustration, to simulate $n = 100$ rounds with the default settings as described above, with Player A switching to the reduced set of plays $\{R, S\}$ after the 50th round, we write

```
> rps_data<-rps_tournament_changepoint(n_games=100, changepoint=50)
```

The output comprises the sequence of A/B/X results:

```
> head(rps_data$simulated_data)
```

[1] “B” “B” “X” “X” “X” “B”

The Gibbs sampler is then run on the simulated object as follows:

```
> rps_gs_out<-rps_gs(rps_data$simulated_data, d1=100, d2=3, m=50, v =1000)
```

For this example we have set $d_1 = 100$, giving a prior choice on $\theta^{(1)}$ that is strongly concentrated on θ_0, and $d_2 = 3$, giving a uniform prior for $\theta^{(2)}$ on Δ_2. The arguments for these choices have been discussed above. The prior for *k* has mean close to 50 but with a large variance. This choice of mean is slightly unfair, since in practice the true value would be unknown, but the consequences are mitigated by also having a large variance *v*. In any case, the Gibbs sampler is quick to run, so the sensitivity of results to these choices can be examined in real-time during class if that seems appropriate. Obviously, the simulated object can be replaced with actual data once collected to effect an analysis on the students' own data.

As with any Gibbs sampler, the output from these functions can be studied graphically to assess the performance of the sampler and to obtain summary inferences. For example, figure 1 shows the Gibbs sample output and a comparison between the prior and posterior distributions of the change-point after 100 rounds. The Gibbs sampler tracer provides visual evidence of the satisfactory mixing and convergence of the chain. The comparison of prior and posterior distributions of the change-point show the extent to which the data have transformed prior beliefs: having observed the data, the change-point is more likely to have occurred after 50 rounds than the prior assumed, and has a greater concentration than the prior. Nonetheless, the prior and posterior are reasonably similar, but this is hardly surprising given the limited amount of data with which the inference is being made. Note that both prior and posterior distributions are discrete having support on the positive integers, but the histogram and smoothed curves are shown on a continuous scale for ease of interpretation.

Similar graphical analysis can be made on the multinomial probabilities $\theta^{(1)}$ and $\theta^{(2)}$, but more interesting in practice is the posterior for θ_i in the original model specification

$$Y_i \mid \theta_i \sim \text{Multinomial}(1, \theta_i), \; i = 1, \dots, n.$$

In other words, having observed the data, what can be said about the A/B/X probabilities for each round?

Since the change-point is unknown, the appropriate choice for θ_i between $\theta^{(1)}$ and $\theta^{(2)}$ is also unknown, but this is exactly the sort of situation in which the Gibbs sampler can be fully exploited, as the output for $(\theta^{(1)}, \theta^{(2)}, k)$ can be transformed to give a Gibbs sample for θ_i, through

$$\theta_i = \theta^{(1)} \text{ if } i \leq k$$

$$\theta_i = \theta^{(2)} \text{ if } i > k.$$

Applying this mapping to the original Gibbs sampler output generates a Gibbs sample of θ_i. Figure 2 summarises the result of this procedure for $i = 1, \dots, 100$. The three panels represent respectively the three components of θ_i, namely the A/X/B probabilities respectively. In each case, for each round i, the summary is a box plot of the Gibbs sample of the relevant component of θ_i, and therefore a graphical approximation to its posterior distribution. The central black curve is therefore a trace of the median of the distribution as a function of i; the red region is the inter-quartile range; the black stems extend to somewhere around 1.5 times the quartiles; and the yellow points indicate outlining points. Note that the posterior distributions here are conditional on data observed from all rounds, not just those up to round i; that's to say, the distributions in figure 2 provide a smooth of the data, not a filter.

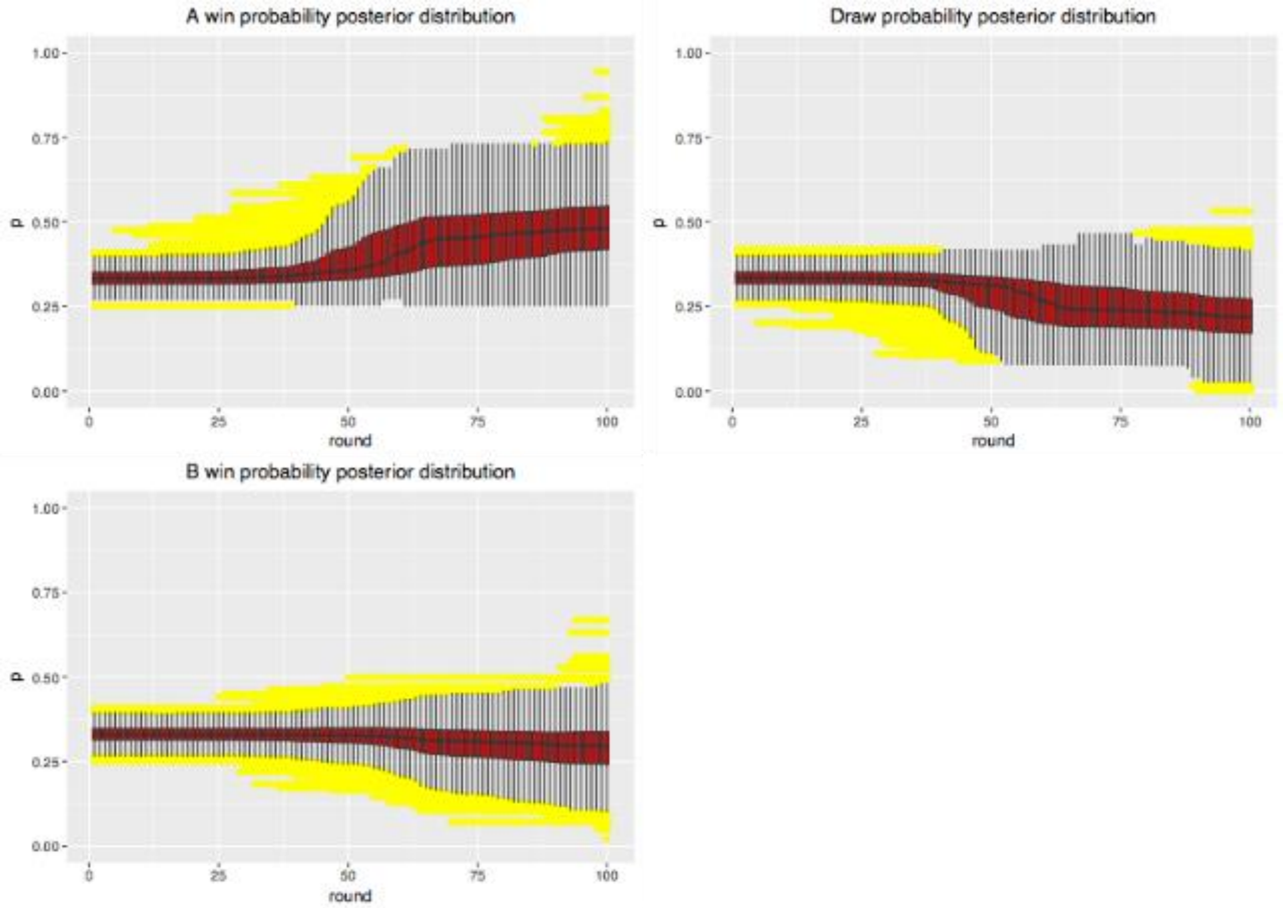

Figure 2. Gibbs sampler distributions of marginal posterior for game outcome probabilities as a function of round. Distributions are represented in standard boxplot form.

For the early rounds, the posterior mean for θ_i is approximately θ_0, with a small posterior variance, in accordance with the strong information provided by the prior. After about 35 rounds, the possibility of a change-point starts to affect the posterior distributions on the θ_i, to an increasing extent as more data become available. By the time we have the full set of data, the posterior mean has shifted much closer to the true value of $\theta = (1/2, 1/4, 1/4)$, albeit with larger posterior variances. Given what students have learnt about the process from playing the game, all aspects of the inference are entirely convincing.

6. Discussion

The exact way this exercise can be used will depend on the type and level of class in which it is introduced. The whole exercise, including data generation, model building and inference can usually be completed within two or three hours. Our own experience was with final-year undergraduate students, who had previously covered the basics of Bayesian Statistics, and were following a course on general computational methods, which included Bayesian techniques such as the Gibbs sampler. In that setting we used the exercise towards the end of the course as a way of reinforcing the links between inference and computation, emphasising the role of model construction for both aspects. In other teaching programmes which include separate modules based on case studies, this exercise could be used as one of the cases. The obvious limitation is that a knowledge of Bayesian inference and computational techniques is required to a level that is typically not studied until second-year undergraduate programmes.

The feedback we received from students, both informally and through student feedback questionnaires, was overwhelmingly positive. In our first attempt we were more vague about the instructions given to students, which led to inconsistencies for many pairs between the way they actually played the game and the single change-point model we had anticipated they would develop. By being more precise with the instructions, and also giving a stronger steer towards our intended model structure, we found the exercise to work much better, and students' satisfaction to be greater.

Possible tasks and extensions that can be suggested to students for further study include:

1. A study of the sensitivity of results to prior choices.
2. A more detailed analysis of the mixing and convergence properties of the Gibbs sampler.
3. A change of protocol so that both players are given the same instructions.
4. Running the Gibbs sampler on subsets of the data, using results from just the first n^* rounds, for $n^* = 10,20,\ldots,100$. How does inference on k, $\theta^{(1)}$ and $\theta^{(2)}$ change as n^* increases?

Finally, although application of a change-point model to the modified Rock-Paper-Scissors game is just an educational exercise, students can also be made aware that change-point models of the type developed here have many real-world applications, including the identification of irregularities in DNA sequences.

7. R code

A zipped version of the R Studio project is available alongside this article from the MSOR Connections journal website https://journals.gre.ac.uk/index.php/msor/. Unzipping the file and opening in Studio gives immediate access to the functions and a script we used to produce the figures.

8. Acknowledgements

We would like to thank Luigi Colombo and Paul Wikramaratna for giving us the initial idea of using the Rock-Paper-Scissors game as a teaching exercise and Philip Giles for pointing out the similarities of the Rock-Paper-Scissors change-point model with models used for DNA sequencing. We are also indebted to Professor Nicola Sartori and the students of the Statistica Computazionale 2 course in the Department of Statistics, University of Padova, for allowing us to carry out this teaching exercise. Finally, we are extremely grateful to the editor and two referees for insightful and helpful comments on a previous version of the paper.

9. Appendix: The Dirichlet distribution

The 3-dimensional Dirichlet distribution $D(\phi, d)$ where $\phi = (\phi_1, \phi_2, \phi_3) \in \Delta_2$ and $d > 0$ has probability density function

$$f(\theta) \propto \theta_1^{d\phi_1 - 1} \theta_2^{d\phi_2 - 1} \theta_3^{d\phi_3 - 1},$$

for $\theta \in \Delta_2$. Its expectation is ϕ and its variance decreases as d increases, with a limiting variance of zero as $d \to \infty$. The choice $d = 3$ and $\phi = (1/3, 1/3, 1/3)$ results in a uniform distribution on Δ_2. It is a convenient prior distribution for a random variable that comprises a probability vector both because it has the correct support and because it provides a conjugate family for the Multinomial distribution. Specifically, if x has the Multinomial distribution

$$x \mid \theta \sim \text{Multinomial}(n, \theta),$$

where $x_1 + x_2 + x_3 = n$, and

$$\theta \sim D(\phi, d)$$

then

$$\theta \mid x \sim D(\phi + x/d, d).$$

10. References

Eckley, I.A., Fearnhead, P. and Killick, R., 2011. Analysis of changepoint models. In: D. Barber, A.T. Cemgil and S. Chiappa, eds. *Bayesian Time Series Models*, Cambridge: Cambridge University Press. pp.205-224.

R Core Team, 2017. *R: A language and environment for statistical computing*. Available at: https://www.r-project.org/ [Accessed 3 September 2019].

van den Nouweland, A., 2007. Rock-Paper-Scissors; A New and Elegant Proof. *Economics Bulletin*, 3(43), pp.1-6.

Walker, D. and Walker, G., 2004. *The Official Rock Paper Scissors Strategy Guide*. Touchstone.

Wang, Z., Xu, B. and Zhou, H.-J., 2014. Social cycling and conditional responses in the Rock-Paper-Scissors game. *Scientific Reports*, 4(5830). https://doi.org/10.1038/srep05830

Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. 2nd ed. New York: Springer-Verlag.