A Bayesian Approach to Determination of Convergence, Divergence and Oscillation of Infinite Series with Application to Riemann Hypothesis

Sucharita Roy and Sourabh Bhattacharya*

Abstract

In the classical literature on infinite series there are various tests to determine if a given infinite series converges, diverges, or oscillates. But unfortunately, for very many infinite series all the existing tests can fail to provide definitive answers. In this article we propose a novel Bayesian theory for assessment of convergence properties of any given infinite series. Remarkably, this theory attempts to provide conclusive answers to the question of convergence even where all the existing tests of convergence fail. We apply our ideas to seven different examples, obtaining very encouraging results. Importantly, we also apply our ideas to investigate the Riemann Hypothesis, and obtain results that do not completely support the conjecture.

We also extend our ideas to develop a Bayesian theory on oscillating series, where we allow even infinite number of limit points. Analysis of Riemann Hypothesis using Bayesian multiple limit points theory yielded almost identical results as the Bayesian theory of convergence assessment.

Keywords: *Bayesian theory; Dirichlet process; Infinite series; Möbius function; Riemann Hypothesis; Tests of series convergence.*

1 Introduction

Determination of convergence, divergence or oscillation of infinite series has a very rich tradition in mathematics, and a large number of tests exist for the purpose. Unfortunately, there does not seem to exist any universal test that provides conclusive answers to all infinite series; see, for example, Ilyin and Poznyak (1982), Knopp (1990), Bourchtein *et al.* (2012). Attempts to resolve the issue as much as possible using hierarchies of tests, with the successive tests in the hierarchy providing conclusive answers to successively larger ranges of infinite series, are provided by Knopp (1990), Bromwich (2005), Bourchtein *et al.* (2011) and Liflyand *et al.* (2011). These tests are based on the Kummer approach for positive series and the chain of the Ermakov tests for positive monotone series. The hierarchy of tests provided in Bourchtein *et al.* (2012) are based on Bromwich (2005) and are related to the well-known Cauchy's test (see, for example, Fichtenholz (1970), Rudin (1976), Spivak (1994)). Below we briefly discuss the approach of Bourchtein *et al.* (2012).

1.1 Hierarchical tests of convergence

The tests of Bourchtein *et al.* (2012) are based on the following theorem, which is a refinement of a result of Bromwich (2005).

Theorem 1 (Bourchtein *et al.* (2012)) *Let $\sum_{i=1}^{\infty} F'(i)$ be a divergent series where $F(x) > 0$, $F'(x) > 0$ and $F'(x)$ is decreasing. If $\sum_{i=1}^{\infty} X_i$ is a positive series, then denoting $\frac{\log\left\{\frac{F'(i)}{X_i}\right\}}{\log F(i)} = W_i$, the following hold:*

$$\text{If} \liminf_{i\to\infty} W_i > 1, \text{ then } \sum_{i=1}^{\infty} X_i \text{ converges;}$$
$$\text{If} \limsup_{i\to\infty} W_i < 1, \text{ then } \sum_{i=1}^{\infty} X_i \text{ diverges.}$$

*Sucharita Roy is an Assistant Professor in St. Xavier's College, Kolkata, pursuing PhD in Interdisciplinary Statistical Research Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700108. Sourabh Bhattacharya is an Associate Professor in Interdisciplinary Statistical Research Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700108. Corresponding e-mail: sourabh@isical.ac.in.

Letting $F(z) = z$ in the above theorem, Bourchtein *et al.* (2012) obtain their first test, which we provide below.

Theorem 2 (Test T_1 of Bourchtein *et al.* (2012)) *Consider a positive series* $\sum_{i=1}^{\infty} X_i$ *and let* $T_{1,i} = \frac{i}{\log i}\left(1 - X_i^{\frac{1}{i}}\right)$. *Then*

$$\textit{If} \liminf_{i\to\infty} T_{1,i} > 1, \textit{ then } \sum_{i=1}^{\infty} X_i \textit{ converges};$$
$$\textit{If} \limsup_{i\to\infty} T_{1,i} < 1, \textit{ then } \sum_{i=1}^{\infty} X_i \textit{ diverges}.$$

This result is the same as that of Bromwich (2005), but a proof was not supplied in that work.

Now choosing $F(z) = \log z$, Bourchtein *et al.* (2012) form their second test of the hierarchy; we provide the result below. Again, the result has been formulated by Bromwich (2005), but a proof was not given.

Theorem 3 (Test T_2 of Bourchtein *et al.* (2012)) *Consider a positive series* $\sum_{i=1}^{\infty} X_i$ *and let* $T_{2,i} = \frac{\log i}{\log\log i}(T_{1,i} - 1)$. *Then*

$$\textit{If} \liminf_{i\to\infty} T_{2,i} > 1, \textit{ then } \sum_{i=1}^{\infty} X_i \textit{ converges};$$
$$\textit{If} \limsup_{i\to\infty} T_{2,i} < 1, \textit{ then } \sum_{i=1}^{\infty} X_i \textit{ diverges}.$$

Setting $F(z) = \log\log z$, the following result has been proved by Bourchtein *et al.* (2012):

Theorem 4 (Test T_3 of Bourchtein *et al.* (2012)) *Consider a positive series* $\sum_{i=1}^{\infty} X_i$ *and let* $T_{3,i} = \frac{\log i}{\log\log i}(T_{2,i} - 1)$. *Then*

$$\textit{If} \liminf_{i\to\infty} T_{3,i} > 1, \textit{ then } \sum_{i=1}^{\infty} X_i \textit{ converges};$$
$$\textit{If} \limsup_{i\to\infty} T_{3,i} < 1, \textit{ then } \sum_{i=1}^{\infty} X_i \textit{ diverges}.$$

Successively selecting $F(z) = \log\log\log z$, $F(z) = \log\log\log\log z$, etc. successively more refined tests T_4, T_5, etc. can be constructed, with each test having wider scope compared to the preceding test with regard to obtaining conclusive decision on convergence or divergence of the underlying series.

However, if, say, at stage k, $\liminf_{i\to\infty} T_{k,i} < 1 < \limsup_{i\to\infty} T_{k,i}$ so that T_k is inconclusive, then all the subsequent tests will also fail to provide any conclusion. Thus, in spite of the above developments, conclusion regarding the series can still be elusive. For instance, an example considered in Bourchtein *et al.* (2012) is the following series:

$$S_1 = \sum_{i=3}^{\infty}\left(1 - \frac{\log i}{i} - \frac{\log\log i}{i}\left\{\cos^2\left(\frac{1}{i}\right)\right\}\left(a + (-1)^i b\right)\right)^i, \tag{1.1}$$

where $a \geq 0$ and $b \geq 0$. For $a = b = 1$, $\liminf_{i\to\infty} T_{2,i} = 0 < 1 < 2 = \limsup_{i\to\infty} T_{2,i}$. Hence, the hierarchy of tests $\{T_k; k \geq 1\}$ fails to provide definitive answer to the question of convergence of the above series.

In fact, we can generalize the series (1.1) such that the hierarchy of tests fails for the general class of series. Indeed, consider

$$S_2 = \sum_{i=3}^{\infty} \left(1 - \frac{\log i}{i} - \frac{\log\log i}{i} f(i) \left(a + (-1)^i b\right)\right)^i, \tag{1.2}$$

where $0 \leq f(i) \leq 1$ for all $i = 1, 2, 3, \ldots$, and $f(i) \rightarrow 1$ as $i \rightarrow \infty$. Such a function can be easily constructed as follows. Let $g(i)$ be positive and monotonically increase to c, where $c > 0$. Then let $f(i) = g(i)/c$, for $i = 1, 2, 3, \ldots$. A simple example of such a function g is $g(i) = c - \frac{1}{i}$; $g(i) = \cos^2\left(\frac{1}{i}\right)$ is another example, showing the generality of (1.2) compared to (1.1).

1.2 Riemann Hypothesis and series convergence

It is well-known that the famous Riemann Hypothesis is equivalent to convergence of an infinite series on a certain interval. A brief introduction to the problem, along with the necessary background, is provided in Section 6. Studying the relevant infinite series, if at all possible, is then the most challenging problem of mathematics. The existing mathematical literature, however, does not seem to be able to provide any directions in this regard. Hence, innovative theories and methods for analyzing infinite series should be particularly welcome.

In this paper, we attempt to provide an alternative method of characterization of series convergence and divergence using Bayesian theory, whch we also subsequently extend to infinite series with multiple or even infinite number of limit points. For the Bayesian purpose we must formulate our theory stochastically, that is, in terms of random infinite series, noting that the theory regarding deterministic infinite series is a special case of our Bayesian formulation.

2 The key concept

Let us consider the random infinite series

$$S = \sum_{i=1}^{\infty} X_i. \tag{2.1}$$

It is required to determine whether the series of the above form converges, diverges or oscillates. Observe that convergence or divergence of the sum S may be thought of as a mapping $f(S) = p$, where f is some appropriate transformation and p is either 0 or 1, where 0 stands for divergence and 1 is associated with convergence. Since we assume that it is not known if the underlying series S converges of diverges, the value of p is unknown, signifying that we must acknowledge uncertainty about p. Conceptually, given the the value of a partial sum of the form $\sum_{i=m}^{n} X_i$, for large m and n, one may have a subjective expectation whether or not the series S converges, which may be quantified as

$$E\left(\mathbb{I}_{\left\{\left|\sum_{i=m}^{n} X_i\right| \leq c_{m,n}\right\}}\right) = P\left(\left|\sum_{i=m}^{n} X_i\right| \leq c_{m,n}\right) = p_{m,n},$$

where, for any set A, $\mathbb{I}_A$ denotes indicator of A, and $c_{m,n}$ are non-negative quantities satisfying $c_{m,n} \rightarrow 0$ as $m, n \rightarrow \infty$. Thus, the expectation depends on how large m and n are. With this, one may expect that

$$f(S) = \lim_{m,n\rightarrow\infty} P\left(\left|\sum_{i=m}^{n} X_i\right| \leq c_{m,n}\right) = \lim_{m,n\rightarrow\infty} p_{m,n} = p.$$

To convert this key concept to a practically useful theory, one requires the Bayesian paradigm, where, for each pair (m, n), belief regarding $p_{m,n}$ needs to be quantified using prior distributions. The terms X_i need to be viewed as realizations of some random process so that the partial sums $\sum_{i=m}^{n} X_i$ provide coherent probabilistic information on p when quantified by the posterior distribution of $p_{m,n}$. As m and n are updated, the posterior of $p_{m,n}$ must also be coherently updated, utilizing the new partial sum information. In particular, as $m, n \to \infty$, it is desirable that the posterior of $p_{m,n}$ converges to either $\delta_{\{1\}}$ or $\delta_{\{0\}}$ in some appropriate sense, accordingly as S converges or diverges. Here, for any x, $\delta_{\{x\}}$ denotes point mass at x.

In Section 3 we devise a recursive Bayesian methodology that achieves the goal discussed above. It is important to remark that no restrictive assumption is necessary for the development of our ideas, not even independence of X_i. With this methodology, we then characterize convergence and divergence of infinite series in Section 4, illustrating in Section 5 our theory and methods with seven examples. In Section 6 we apply our ideas to Riemann Hypothesis, obtaining results that are not in complete favour of the conjecture. We then extend our theory and methods to infinite series with multiple or inifinite number of limit points; details are provided in Section 7. Illustrations of our Bayesian multiple limit point theory are provided in Sections 8 and 9, the latter section detailing the application to Riemann Hypothesis in order to vindicate our results obtained in Section 6. Finally, we make concluding remarks in Section 10.

3 A recursive Bayesian procedure for studying infinite series

Since we view X_i as realizations from some random process, we first formalize the notion in terms of the relevant probability space. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, where Ω is the sample space, $\mathcal{A}$ is the Borel σ-field on Ω, and μ is some probability measure. Let, for $i = 1, 2, 3, \ldots$, $X_i : \Omega \mapsto \mathbb{R}$ be real valued, random variables measurable with respect to the Borel σ-field $\mathcal{B}$ on $\mathbb{R}$. As in Schervish (1995), we can then define a σ-field of subsets of $\mathbb{R}^\infty$ with respect to which $X = (X_1, X_2, \ldots)$ is measurable. Indeed, let us define $\mathbb{B}^\infty$ to be the smallest σ-field containing sets of the form

$$B = \big\{X : X_{i_1} \leq r_1, X_{i_2} \leq r_2, \ldots, X_{i_p} \leq r_p, \text{ for some } p \geq 1,$$
$$\text{some integers } i_1, i_2, \ldots, i_p, \text{ and some real numbers } r_1, r_2, \ldots, r_p\big\}.$$

Since B is an intersection of finite number of sets of the form $\big\{X : X_{i_j} \leq r_j\big\}$; $j = 1, \ldots, p$, all of which belong to $\mathcal{A}$ (since X_{i_j} are measurable) it follows that $X^{-1}(B) \in \mathcal{A}$, so that X is measurable with respect to $(\mathbb{R}^\infty, \mathbb{B}^\infty, P)$, where P is the probability measure induced by μ.

3.1 Development of the stage-wise likelihoods

For $j = 1, 2, 3, \ldots$, let

$$S_{j,n} = \sum_{i=(j-1)n+1}^{jn} X_i, \tag{3.1}$$

where $n \geq 1$. In view of this definition, and for notational convenience we shall often denote S by $S_{1,\infty}$. Also let $\{c_j\}_{j=1}^\infty$ be a decreasing sequence and

$$Y_{j,n} = \mathbb{I}_{\{|S_{j,n}| \leq c_j\}}(S_{j,n}). \tag{3.2}$$

Let, for $j \geq 1$ and $n \geq 1$,

$$P\left(Y_{j,n} = 1\right) = p_{j,n}. \tag{3.3}$$

Hence, the likelihood of $p_{j,n}$, given $y_{j,n}$, is given by

$$L\left(p_{j,n}\right) = p_{j,n}^{y_{j,n}}\left(1 - p\right)^{1 - y_{j,n}} \tag{3.4}$$

It is important to relate $p_{j,n}$ to convergence, divergence or oscillation of the underlying series. Note that $p_{j,n}$ is the probability that $|S_{j,n}|$ falls below c_j. Thus, $p_{j,n}$ can be interpreted as the probability that the series $S_{1,\infty}$ is convergent when the data observed is $S_{j,n}$. If $S_{1,\infty}$ is convergent, then it is to be expected *a posteriori*, that

$$p_{j,n} \to 1 \quad \text{as } j \to \infty \text{ for all } n \geq 1. \tag{3.5}$$

Note that the above is expected to hold not only for large n but all $n \geq 1$. This is related to Cauchy's criterion of convergence of partial sums: for every $\epsilon > 0$ there exists a positive integer N such that for all $n \geq m \geq N$, $|\sum_{i=m}^{n} X_i| < \epsilon$. Hence, in the case of convergence, by Cauchy's criterion, for any $\epsilon > 0$ there must exist a positive integer j_0 such that for $j \geq j_0$, $|S_{j,1}| = |X_j| < \epsilon$. Indeed, as we will formally show, condition (3.5) is both necessary and sufficient for convergence of the series.

On the other hand, if the series is divergent (but not oscillatory), then there exist $j_0 \geq 1$ and $n_0 \geq 1$ such that $|S_{j,n}| > c_j$ for all $j > j_0$ and $n > n_0$. Here we expect, *a posteriori*, that

$$p_{j,n} \to 0 \quad \text{as } j \to \infty, \text{ for } n \geq n_0. \tag{3.6}$$

Again, we will prove formally that the above condition is both necessary and sufficient for divergence.

Now, if the series $S_{1,\infty}$ is oscillating with two limit points occurring with equal frequency, then one may expect that in the long run, in 50% cases $|S_{j,n}|$ will fall below the respective c_j. Thus, here one may expect, *a posteriori*, that

$$p_{j,n} \to \frac{1}{2} \quad \text{as } j \to \infty, \text{ for } n \geq n_0. \tag{3.7}$$

We will formally prove that the above is both necessary and sufficient for oscillations of the sbove type. A general theory, which encompasses finite as well as infinite number of limit points, with perhaps unequal frequencies of occurrences, is developed in Section 7.

In what follows we shall first construct a recursive Bayesian methodology that formally characterizes convergence, divergence and oscillation in terms of formal posterior convergence related to (3.5), (3.6) and (3.7), respectively.

3.2 Development of recursive Bayesian posteriors

We assume that $\{y_{j,n}; j = 1, 2, \ldots\}$ is observed successively at stages indexed by j. That is, we first observe $y_{1,n}$, and based on our prior belief regarding the first stage probability, $p_{1,n}$, compute the posterior distribution of $p_{1,n}$ given $y_{1,n}$, whch we denote by $\pi(p_{1,n}|y_{1,n})$. Based on this posterior we construct a prior for the second stage, and compute the posterior $\pi(p_{2,n}|y_{1,n}, y_{2,n})$. We continue this procedure for as many stages as we desire. Details follow.

Consider the sequences $\{\alpha_j\}_{j=1}^{\infty}$ and $\{\beta_j\}_{j=1}^{\infty}$, where $\alpha_j = \beta_j = 1/j^2$ for $j = 1, 2, \ldots$. At the first stage of our recursive Bayesian algorithm, that is, when $j = 1$, let us assume that the prior is given by

$$\pi(p_{1,n}) \equiv Beta(\alpha_1, \beta_1), \tag{3.8}$$

where, for $a > 0$ and $b > 0$, $Beta(a, b)$ denotes the Beta distribution with mean $a/(a + b)$ and variance $(ab)/\left\{(a + b)^2(a + b + 1)\right\}$. Combining this prior with the likelihood (3.4) (with $j = 1$), we obtain the following posterior of $p_{1,n}$ given $y_{1,n}$:

$$\pi(p_{1,n}|y_{1,n}) \equiv Beta\left(\alpha_1 + y_{1,n}, \beta_1 + 1 - y_{1,n}\right). \tag{3.9}$$

At the second stage (that is, for $j = 2$), for the prior of $p_{2,n}$ we consider the posterior of $p_{1,n}$ given $y_{1,n}$ associated with the $Beta(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$ prior. That is, our prior on $p_{2,n}$ is given by:

$$\pi(p_{2,n}) \equiv Beta\left(\alpha_1 + \alpha_2 + y_{1,n}, \beta_1 + \beta_2 + 1 - y_{1,n}\right). \tag{3.10}$$

The reason for such a prior choice is that the uncertainty regarding convergence of the series is reduced once we obtain the posterior at the first stage, so that at the second stage the uncertainty regarding the prior is expected to be lesser compared to the first stage posterior. With our choice, it is easy to see that the prior variance at the second stage, given by

$$\left\{(\alpha_1+\alpha_2+y_{1,n})(\beta_1+\beta_2+1-y_{1,n})\right\}/\left\{(\alpha_1+\alpha_2+\beta_1+\beta_2+1)^2(\alpha_1+\alpha_2+\beta_1+\beta_2+2)\right\},$$

is smaller than the first stage posterior variance, given by

$$\left\{(\alpha_1+y_{1,n})(\beta_1+1-y_{1,n})\right\}/\left\{(\alpha_1+\beta_1+1)^2(\alpha_1+\beta_1+2)\right\}.$$

The posterior of $p_{2,n}$ given $y_{2,n}$ is then obtained by combining the second stage prior (3.10) with (3.4) (with $j=2$). The form of the posterior at the second stage is thus given by

$$\pi(p_{2,n}|y_{2,n}) \equiv Beta\left(\alpha_1+\alpha_2+y_{1,n}+y_{2,n}, \beta_1+\beta_2+2-y_{1,n}-y_{2,n}\right). \tag{3.11}$$

Continuing this way, at the k-th stage, where $k>1$, we obtain the following posterior of $p_{k,n}$:

$$\pi(p_{k,n}|y_{k,n}) \equiv Beta\left(\sum_{j=1}^{k}\alpha_j+\sum_{j=1}^{k}y_{j,n}, k+\sum_{j=1}^{k}\beta_j-\sum_{j=1}^{k}y_{j,n}\right). \tag{3.12}$$

It follows from (3.12) that

$$E\left(p_{k,n}|y_{k,n}\right) = \frac{\sum_{j=1}^{k}\alpha_j+\sum_{j=1}^{k}y_{j,n}}{k+\sum_{j=1}^{k}\alpha_j+\sum_{j=1}^{k}\beta_j}; \tag{3.13}$$

$$Var\left(p_{k,n}|y_{k,n}\right) = \frac{(\sum_{j=1}^{k}\alpha_j+\sum_{j=1}^{k}y_{j,n})(k+\sum_{j=1}^{k}\beta_j-\sum_{j=1}^{k}y_{j,n})}{(k+\sum_{j=1}^{k}\alpha_j+\sum_{j=1}^{k}\beta_j)^2(1+k+\sum_{j=1}^{k}\alpha_j+\sum_{j=1}^{k}\beta_j)}. \tag{3.14}$$

Since $\sum_{j=1}^{k}\alpha_j=\sum_{j=1}^{k}\beta_j=\sum_{j=1}^{k}\frac{1}{j^2}$, (3.13) and (3.14) admit the following simplifications:

$$E\left(p_{k,n}|y_{k,n}\right) = \frac{\sum_{j=1}^{k}\frac{1}{j^2}+\sum_{j=1}^{k}y_{j,n}}{k+2\sum_{j=1}^{k}\frac{1}{j^2}}; \tag{3.15}$$

$$Var\left(p_{k,n}|y_{k,n}\right) = \frac{(\sum_{j=1}^{k}\frac{1}{j^2}+\sum_{j=1}^{k}y_{j,n})(k+\sum_{j=1}^{k}\frac{1}{j^2}-\sum_{j=1}^{k}y_{j,n})}{(k+2\sum_{j=1}^{k}\frac{1}{j^2})^2(1+k+2\sum_{j=1}^{k}\frac{1}{j^2})}. \tag{3.16}$$

4 Characterization of convergence properties of the underlying infinite series

Based on our recursive Bayesian theory we have the following theorem that characterizes convergence of $S_{1,\infty}$ in terms of the limit of the posterior probability of $p_{k,n}$, as $k\rightarrow\infty$.

Theorem 5 *$S_{1,\infty}$ is almost surely convergent if and only if*

$$\pi\left(\mathcal{N}_1|y_{k,n}\right)\rightarrow 1, \tag{4.1}$$

$k\rightarrow\infty$, almost surely for all $Y_n=\{y_{k,n}\}_{n=1}^{\infty}$, for any $n\geq 1$, where $\mathcal{N}_1$ is any neighborhood of 1 (one).

Proof. If $S_{1,\infty}$ is convergent, then there exists a finite j_0 such that for all $j>j_0$, $|S_{j,n}|\leq c_j$ for all n, so that $y_{j,n}=1$ for all $j>j_0$, for all n. Hence, in this case, $\sum_{j=1}^{k}y_{j,n}=k-k_0$, where $k_0\geq 0$. Also,

$\sum_{j=1}^{k}\frac{1}{j^2}\rightarrow\frac{\pi^2}{6}$, as $k\rightarrow\infty$. Consequently, it is easy to see that

$$\mu_k = E\left(p_{k,n}|y_{k,n}\right)\sim\frac{\frac{\pi^2}{6}+k-k_0}{k+\frac{\pi^2}{3}}\rightarrow 1,\ \text{as } k\rightarrow\infty,\ \text{and,} \tag{4.2}$$

$$\sigma^2_k = Var\left(p_{k,n}|y_{k,n}\right)\sim\frac{(\frac{\pi^2}{6}+k)(\frac{\pi^2}{6})}{(k+\frac{\pi^2}{3})^2(1+k+\frac{\pi^2}{3})}\rightarrow 0 \text{ as } k\rightarrow\infty. \tag{4.3}$$

In the above, for any two sequences $\{a_k\}_{k=1}^{\infty}$ and $\{b_k\}_{k=1}^{\infty}$, $a_k\sim b_k$ indicates $\frac{a_k}{b_k}\rightarrow 1$, as $k\rightarrow\infty$. Now let $\mathcal{N}_1$ denote any neighborhood of 1, and let $\epsilon>0$ be sufficiently small such that $\mathcal{N}_1\supseteq\{1-p_{k,n}<\epsilon\}$. Combining (4.2) and (4.3) with Chebychev's inequality ensures that (4.1) holds.

Now assume that (4.1) holds. Then for any given $\epsilon>0$, for all $n\geq 1$,

$$\pi\left(p_{k,n}>1-\epsilon|y_{k,n}\right)\rightarrow 1,\ \text{as } k\rightarrow\infty. \tag{4.4}$$

Hence,

$$E\left(p_{k,n}|y_{k,n}\right)\rightarrow 1; \tag{4.5}$$

$$Var\left(p_{k,n}|y_{k,n}\right)\rightarrow 0, \tag{4.6}$$

as $k\rightarrow\infty$, for all $n\geq 1$. If $S_{1,\infty}$ does not converge then there exist j_0 and n_0 such that for $n\geq n_0$, $|S_{j,n}|>c_j$, for $j\geq j_0$. Hence, in this situation, for $n\geq n_0$, $0\leq\sum_{j=1}^{k}y_{j,n}\leq j_0$. Substituting this in (3.15) and (3.16), it is easy to see that, for $n\geq n_0$, as $k\rightarrow\infty$,

$$E\left(p_{k,n}|y_{k,n}\right)\rightarrow 0; \tag{4.7}$$

$$Var\left(p_{k,n}|y_{k,n}\right)\rightarrow 0, \tag{4.8}$$

so that (4.5) is contradicted.

A somewhat different stuation may arise if $S_{1,\infty}$ does not converge in the sense that it is an oscillatory series. Then it need not hold that $|S_{j,n}|>c_j$, for $j\geq j_0$ for $n\geq n_0$. Instead, for $n\geq n_0$, we may have $|S_{j_0+2r,n}|>c_j$ and $|S_{j_0+2r+1,n}|\leq c_j$, for $r=0,1,2,\ldots$. Assuming without loss of generality that k is even, it follows that for $n\geq n_0$, $\sum_{j=1}^{k}y_{j,n}=\frac{k}{2}-k_0$, for $k_0\geq 0$. Substituting this in (3.15) and (3.16), we obtain, for $n\geq n_0$, that

$$E\left(p_{k,n}|y_{k,n}\right)\rightarrow\frac{1}{2}; \tag{4.9}$$

$$Var\left(p_{k,n}|y_{k,n}\right)\rightarrow 0, \tag{4.10}$$

as $k\rightarrow\infty$, so that (4.5) is again contradicted. ■

We now prove the following theorem that provides necesary and sufficient conditions for divergence of $S_{1,\infty}$ in terms of the limit of the posterior probability of $p_{k,n}$, as $k\rightarrow\infty$.

Theorem 6 *$S_{1,\infty}$ is almost surely non-oscillating divergent if and only if there exists $n_0\geq 1$ such that*

$$\pi\left(\mathcal{N}_0|y_{k,n}\right)\rightarrow 1, \tag{4.11}$$

$k\rightarrow\infty$, almost surely for all $Y_n=\{y_{k,n}\}_{n=1}^{\infty}$, for all $n\geq n_0$, where $\mathcal{N}_0$ is any neighborhood of 0 (zero).

Proof. Assume that $S_{1,\infty}$ is non-oscillating divergent. Then then there exist j_0 and n_0 such that for $n\geq n_0$, $|S_{j,n}|>c_j$, for $j\geq j_0$. From the proof of the sufficient condition of Theorem 5 it follows that (4.7) and (4.8) hold. Let $\epsilon>0$ be small enough so that $\mathcal{N}_0\supseteq\{p_{k,n}<\epsilon\}$. Then combining Chebychev's inequality with (4.7) and (4.8) it is easy to see that (4.11) holds.

Now assume that (4.11) holds. Then for any given $\epsilon > 0$, for all n,

$$\pi\left(p_{k,n} < \epsilon | y_{k,n}\right) \rightarrow 1, \text{ as } k \rightarrow \infty. \tag{4.12}$$

It follows that

$$E\left(p_{k,n} | y_{k,n}\right) \rightarrow 0; \tag{4.13}$$

$$Var\left(p_{k,n} | y_{k,n}\right) \rightarrow 0, \tag{4.14}$$

as $k \rightarrow \infty$.

If $S_{1,\infty}$ is convergent, then by Theorem 5, $\pi\left(\mathcal{N}_1 | y_{k,n}\right) \rightarrow 1$ as $k \rightarrow \infty$, so that $E\left(p_{k,n} | y_{k,n}\right) \rightarrow 1$, which is a contradiction to (4.13).

If $S_{1,\infty}$ is oscillatory, then from the proof of the sufficient condition of Theorem 5, $E\left(p_{k,n} | y_{k,n}\right) \rightarrow \frac{1}{2}$, which is again a contradiction to (4.13). ■

We now characterize the oscillatory behaviour of $S_{1,\infty}$ in terms of the limit of the posterior probability of $p_{k,n}$, as $k \rightarrow \infty$.

Theorem 7 *$S_{1,\infty}$ is almost surely oscillatory if and only if there exists $n_0 \geq 1$, such that*

$$\pi\left(\mathcal{N}_{\frac{1}{2}} | y_{k,n}\right) \rightarrow 1, \tag{4.15}$$

as $k \rightarrow \infty$, almost surely for all $Y_n = \{y_{k,n}\}_{n=1}^{\infty}$, for all $n \geq n_0$. In the above, $\mathcal{N}_{\frac{1}{2}}$ is any neighborhood of $1/2$.

Proof. Here we assume that k is even, that is, $k = 2\tilde{k}$, where $\tilde{k} \rightarrow \infty$. If $S_{1,\infty}$ is oscillatory, then by the proof of sufficiency of Theorem 5 it follows that there exists $n_0 \geq 1$, such that (4.9) and (4.10) hold for $n \geq n_0$. Now let $\mathcal{N}_{\frac{1}{2}}$ be any neighborhood of $1/2$. Let $\epsilon > 0$ be sufficiently small so that $\mathcal{N}_{\frac{1}{2}} \supseteq \left\{|p_{k,n} - \frac{1}{2}| < \epsilon\right\}$. Then by Chebychev's inequality $\pi\left(\mathcal{N}_{\frac{1}{2}} | y_{k,n}\right) \rightarrow 1$, as $k \rightarrow \infty$, for all $n \geq n_0$. Thus, (4.15) holds.

Now assume that there exists $n_0 \geq 1$ such that (4.15) holds for $n \geq n_0$. Then $\pi\left(|p_{k,n} - \frac{1}{2}| < \epsilon | y_{k,n}\right) \rightarrow 1$, as $k \rightarrow \infty$, for $n \geq n_0$. Combining this with Chebychev's inequality it follows that (4.9) and (4.10) hold for $n \geq n_0$. If $S_{1,\infty}$ is non-oscillating, then by Theorems 5 and 6 $E\left(p_{k,n} | y_{k,n}\right)$ tends to either 0 or 1, which contradict (4.9). Hence $S_{1,\infty}$ must be oscillatory. ■

5 Illustrations

We now illustrate our ideas with seven examples. These seven examples can be categorized into three categories in terms of construction of the upper bound $c_{j,n}$. With the first example we demonstrate that it may sometimes be easy to devise an appropiate upper bound. In Examples 2 – 5, we show that usually simple bounds such as that in Example 1, are not adequate in practice, but appropriate bounds may be constructed if convergence and divergence of the series in question is known for some values of the parameters; the resultant bounds can be utilized to learn about convergence or divergence of the series for the remaining values of the parameters. In Examples 6 and 7, the series in question are stand-alone in the sense they are not defined by parameters with known convergence/divergence for some of their values which might have aided our construction of $c_{j,n}$. However, we show that these series can be embedded into appropriately parameterized series, facilitating similar analysis as Examples 2 – 5.

For these examples, we construct the partial sums $S_{j,n}$ setting $n = 10^6$, and run our recursive Bayesian methodology for $K = 10^5$ stages. Since we needed to sum 10^6 terms at each step of 10^5 stages, the associated computation is extremely demanding. For the purpose of efficiency, we parallelized the computation of the sums of 10^6 terms, splitting the job on many processors, using the Message Passing Interface (MPI) protocol. In more details, we implemented our parallelized codes, written in

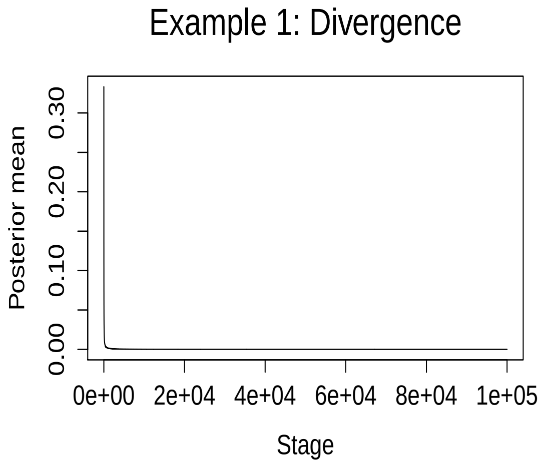


Figure 5.1: Example 1: The series (5.1) is divergent.

C, in VMware consisting of 60 double-threaded, 64-bit physical cores, each running at 2793.269 MHz. Parallel computation of our methods associated with Examples 1 to 5 take, respectively, 1 minute, 4 minutes, 7 minutes, 6 minutes, and 9 minutes. Examples 6 and 7 require about 6 minutes and 4 minutes of computational time.

5.1 Example 1

In their first example Bourchtein *et al.* (2012) study the following divergent series with their methods:

$$S = \sum_{i=2}^{\infty} \frac{1}{\log(i)}. \tag{5.1}$$

We test our Bayesian idea on this series choosing the monotonically decreasing sequence as $c_{j,n} = 1/\sqrt{nj}$, where we represent c_j as $c_{j,n}$ to reflect dependence on n. Figure 5.1, a plot of the posterior means of $\left\{p_{k,n}; k = 1, \ldots, 10^5\right\}$, clearly and correctly indicates that the series is divergent. We also constructed approximate 95% highest posterior density credible intervals at each recursive step; however, thanks to very less variances at each stage, the intervals turned out to be too small to be clearly distinguishable from the plot of the stage-wise posterior means.

5.2 Example 2

Example 2 of Bourchtein *et al.* (2012) deals with the following series:

$$S^a = \sum_{i=2}^{\infty} \left(1 - \left\{\frac{\log(i)}{i}\right\} - a\frac{\log\log(i)}{i}\right)^i, \tag{5.2}$$

where $a \in \mathbb{R}$. Bourchtein *et al.* (2012) prove that the series converges for $a > 1$ and diverges for $a \leq 1$.

5.2.1 Choice of c_j

Now, however, selecting the monotone sequence as $c_{j,n} = 1/\sqrt{nj}$ turn out to be inappropriate for this series, the behaviour of which is quite sensitive to the parameter a, particularly around $a = 1$. Hence, any appropriate sequence $\{c_{j,n}\}_{j=1}^{\infty}$ must depend on the parameter a of the series (5.2).

Denoting $c_{j,n}$ by $c_{j,n}^a$ to reflect the dependence on a and n, we first set

$$u_{j,n}^a = S_{j,n}^{a_0} + \frac{(a - 1 - 9 \times 10^{-11})}{\log(j+1)}, \tag{5.3}$$

and then let

$$c_{j,n}^{a} = \begin{cases} u_{j,n}^{a}, & \text{if } u_{j,n}^{a} > 0; \\ S_{j,n}^{a_0}, & \text{otherwise.} \end{cases} \tag{5.4}$$

where $a_0 = 1 + 10^{-10}$. The reason behind such a choice of $c_{j,n}^{a}$ is provided below.

Let, for $\epsilon > 0$,

$$\tilde{S} = \sup\{S^a : a \geq 1 + \epsilon\}. \tag{5.5}$$

Thus, $\tilde{S}$ may be interpreted as the convergent series which is closest to divergence given the convergence criterion $a \geq 1 + \epsilon$. Since S^a is decreasing in a, it easily follows that equality of (5.5) is attained at $a_0 = 1 + \epsilon$.

From the above arguments it follows that $S_{j,n}^{a_0}$ in (5.4) is decreasing in j and n due to Cauchy's convergence criterion. We assume that ϵ is chosen to be so small that convergence properties of the series for $\{a \leq 1\} \cup \{a \geq 1 + \epsilon\}$ are only desired. Indeed, since $\left(1 - \left\{\frac{\log(i)}{i}\right\} - a\frac{\log\log(i)}{i}\right)^i$ is decreasing in a for any given $i \geq 3$, our method of constructing $c_{j,n}^{a}$ need not be able to correctly identify the convergence properties of the series for $1 < a < 1 + \epsilon$.

For the purpose of illustrations we choose $\epsilon = 10^{-10}$. Note that for $a > 1$ the term $\frac{(a-1-9\times10^{-11})}{\log(j+1)}$ inflates c_j^a making $S_{j,n}^{a}$ more likely to fall below $c_{j,n}^{a}$ for increasing a, thus paving the way for diagnosing convergence. The same term also ensures that for $a \leq 1$, $c_{j,n}^{a} < S_{j,n}^{a_0}$, so that $S_{j,n}^{a}$ is likely to exceed c_j^a, thus providing an inclination towards divergence. The term -9×10^{-11} is an adjustment for the case $a = 1 + 10^{-10}$, ensuring that $c_{j,n}^{a}$ marginally exceeds $S_{j,n}^{a}$ to ensure convergence. The scaling factor $\log(j + 1)$ ensures that the part $\frac{(a-1-9\times10^{-11})}{\log(j+1)}$ of (5.4) tends to zero at a slow rate so that $c_{j,n}^{a}$ is decreasing with j and n even if $a - 1 - 9 \times 10^{-11}$ is negative.

Figure 5.2, depicting our Bayesian results for this series, is in agreement with the results of Bourchtein *et al.* (2012). In fact, we have applied our methods to many more values of $a \in A_\epsilon$ with $\epsilon = 10^{-10}$, and in every case the correct result is vindicated.

5.3 Example 3

Let us now consider the following series analysed by Bourchtein *et al.* (2012):

$$S = \sum_{i=3}^{\infty} \left(1 - \left(\frac{\log(i)}{i}\right) a^{\frac{\log\log(i)}{\log(i)}}\right)^i, \tag{5.6}$$

where $a > 0$. As is shown by Bourchtein *et al.* (2012), the series converges for $a > e$ and diverges for $a \leq e$.

5.3.1 Choice of c_j

Here we first set

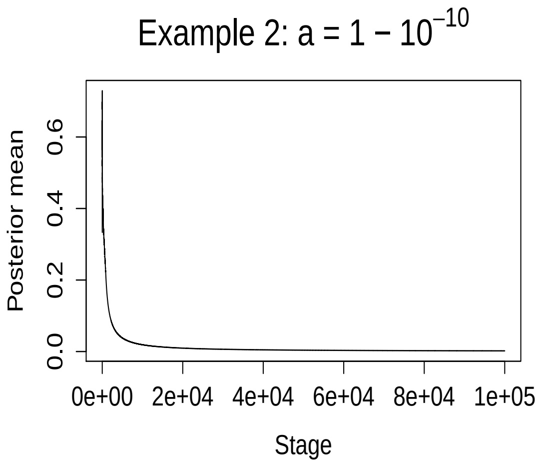
$$u_{j,n}^{a} = S_{j,n}^{a_0} + \frac{(a - e - 9 \times 10^{-11})}{\log(j + 1)}, \tag{5.7}$$

and then let $c_{j,n}^{a}$ defined by (5.4). In this example we set $a_0 = e + 10^{-10}$. The rationale behind the choice remains the same as detailed in Section 5.2.1.

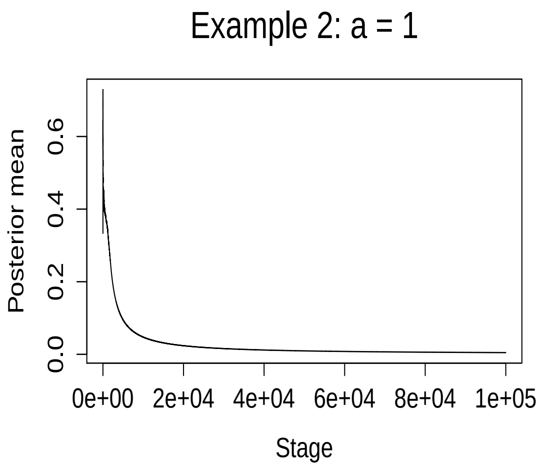
As before, the results obtained by our Bayesian theory, as displayed in Figure 5.3, are in complete agreement with the results obtained by Bourchtein *et al.* (2012).

5.4 Example 4

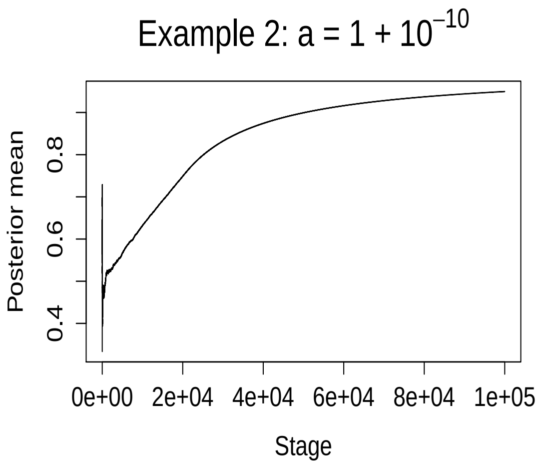
We now consider the series (1.1). It has been proved by Bourchtein *et al.* (2012) that the series is convergent for $a - b > 1$ and divergent for $a + b < 1$. As mentioned before, the hierarchy of tests of Bourchtein *et al.* (2012) are inconclusive for $a = b = 1$.



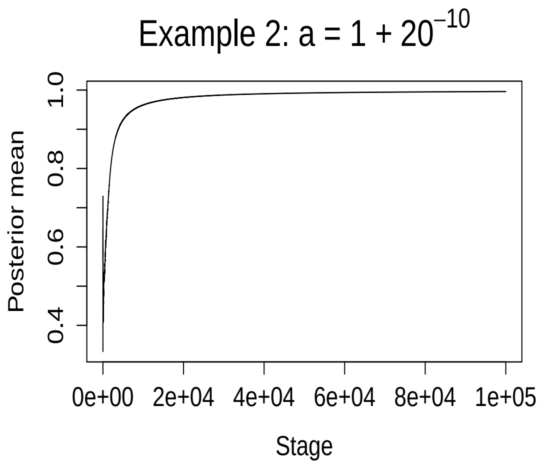
(a) Divergence: $a = 1 - 10^{-10}$.



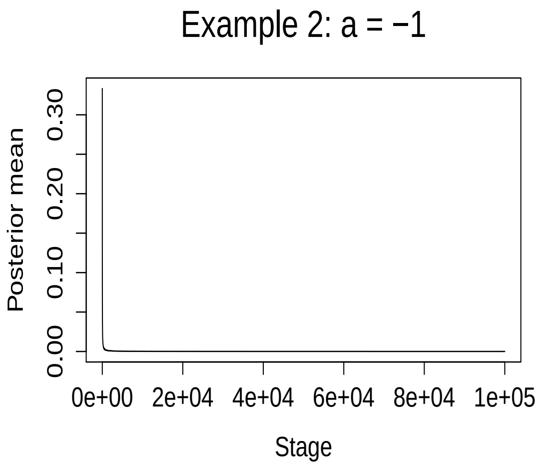
(b) Divergence: $a = 1$.



(c) Convergence: $a = 1 + 10^{-10}$.

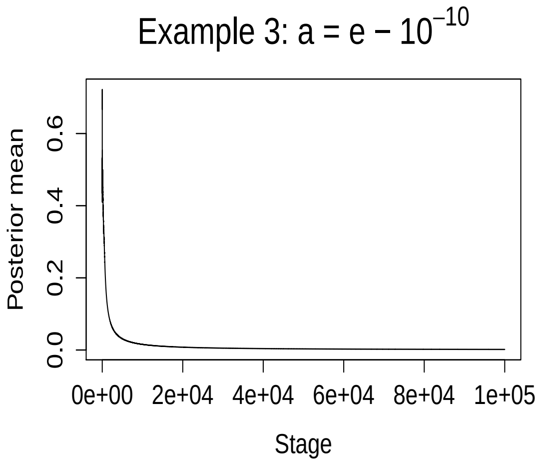


(d) Convergence: $a = 1 + 20^{-10}$.

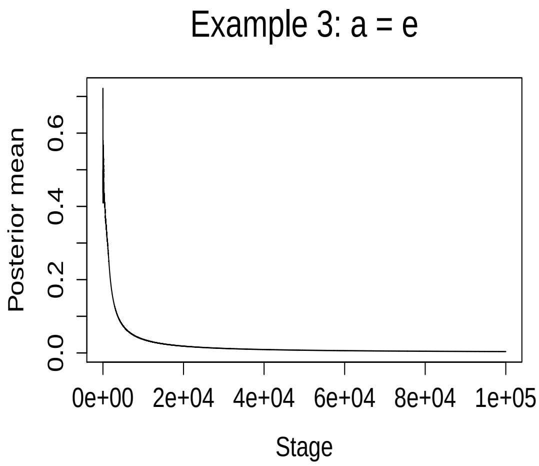


(e) Divergence: $a = -1$.

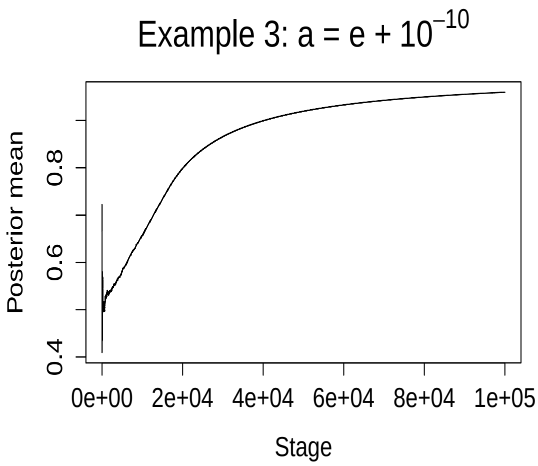
Figure 5.2: Example 2: The series (5.2) converges for $a > 1$ and diverges for $a \leq 1$.



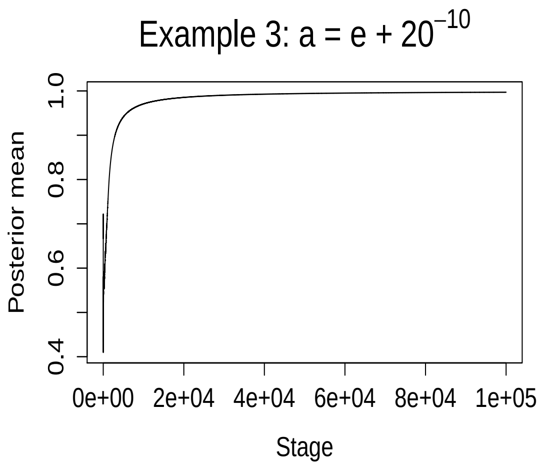
(a) Divergence: $a = e - 10^{-10}$.



(b) Divergence: $a = e$.



(c) Convergence: $a = e + 10^{-10}$.



(d) Convergence: $a = e + 20^{-10}$.

Figure 5.3: Example 3: The series (5.6) converges for $a > e$ and diverges for $a \leq e$.

In this example we denote the partial sums by $S_{j,n}^{a,b}$ and the actual series S by $S^{a,b}$ to reflect the dependence on both the parameters a and b.

$$S_{j,n}^{a,b} = \sum_{i=3+n(j-1)}^{3+nj-1} \left(1 - \frac{\log i}{i} - \frac{\log\log i}{i}\left\{\cos^2\left(\frac{1}{i}\right)\right\}\left(a+(-1)^i b\right)\right)^i, \tag{5.8}$$

We then have the following lemma, the proof of which is presented in Appendix A.

Lemma 8 *For the series (1.1), for $j \geq 1$ and n even, $S_{j,n}^{a,b}$ given by (5.8) is decreasing in a but increasing in b.*

Since $S^{a,b}$ is just summation of the partial sums, it follows that

Corollary 9 *$S^{a,b}$ is decreasing in a and increasing in b.*

We let

$$A_\epsilon = \{a : 0 \leq a \leq 1\} \cup \{a : a \geq 1+\epsilon\}, \tag{5.9}$$

and

$$\tilde{S} = \inf_{a\in A_\epsilon} \sup_{b\geq 0} \left\{S^{a,b} : a - b > 1\right\}. \tag{5.10}$$

It is easy to see in this case, due to Corollary 9 and the convergence criterion $a-b>1$, that $\tilde{S}$ is attained at $a_0 = 1+\epsilon$ and $b_0 = 0$. As before, we set $\epsilon = 10^{-10}$. Hence, arguments similar to those in Section 5.2.1 lead to the following choice of the upper bound for $S_{j,n}^{a,b}$, which we denote in this example by $c_{j,n}^{a,b}$:

$$c_{j,n}^{a,b} = \begin{cases} u_{j,n}^{a,b}, & \text{if } u_{j,n}^{a,b} > 0; \\ S_{j,n}^{a_0,b_0}, & \text{otherwise,} \end{cases} \tag{5.11}$$

where $a_0 = 1+10^{-10}$, $b_0 = 0$, and

$$u_{j,n}^{a,b} = S_{j,n}^{a_0,b_0} + \frac{(a-1-b-9\times 10^{-11})}{\log(j+1)}. \tag{5.12}$$

Note that $-b$ in (5.12) takes account of the fact that the partial sums are increasing in b, thus favouring divergence for increasing b.

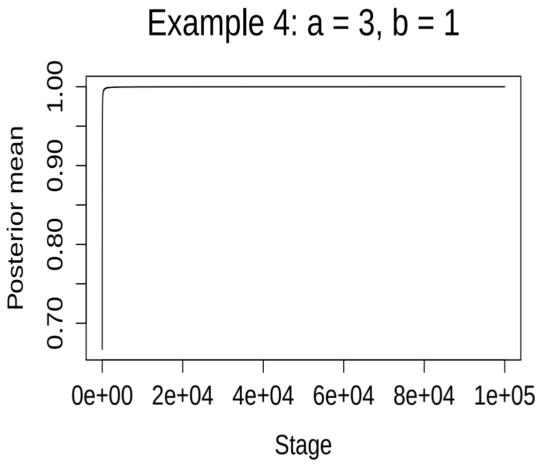
Setting aside panel (c) of Figure 5.5, observe that the remaining panels of Figures 5.4 and 5.5 are in agreement with the results of Bourchtein *et al.* (2012), but in the case $a = b = 1$, the tests of Bourchtein *et al.* (2012) turned out to be inconclusive. Panel (c) of Figure 5.5 demonstrates that the series is divergent for $a = b = 1$.

5.5 Example 5

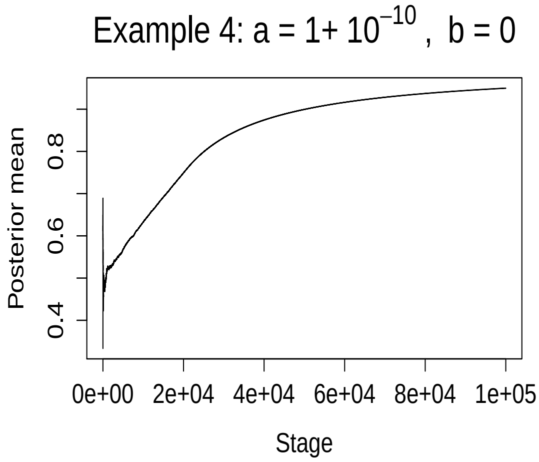
Now consider the following series presented and analysed in Bourchtein *et al.* (2012):

$$S = \sum_{i=3}^{\infty} \left(1 - \left(\frac{\log(i)}{i}\right)\left(a\left(1+\sin^2\left(\sqrt{\left(\frac{\log(\log(i))}{\log(i)}\right)}\right)\right) + b\sin\left(\frac{i\pi}{4}\right)\right)\right)^i; \; a>0, b>0. \tag{5.13}$$

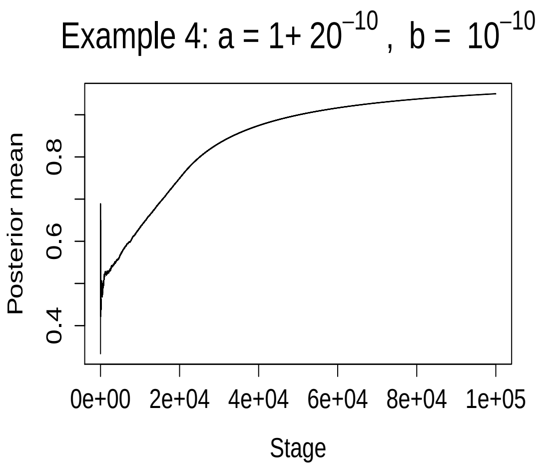
Bourchtein *et al.* (2012) show that the series converges when $a-b>1$ and diverges when $a+b<1$. Again, as in the case of Example 4, the following lemma holds in Example 5, the proof of which is provided in Appendix B. Note that for mathematical convenience we consider partial sums from the 5-th term onwards. We also assume n to be a multiple of 4.



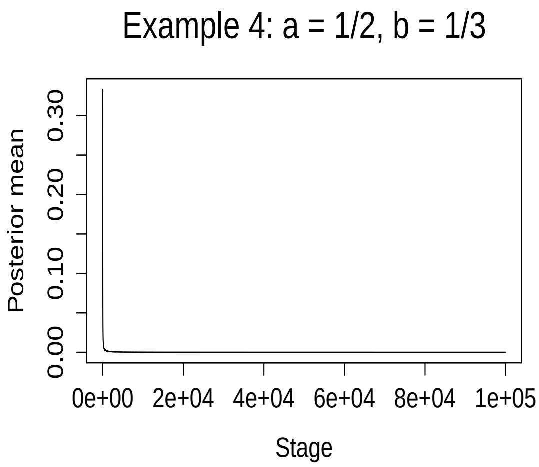
(a) Convergence: $a = 3, b = 1$.



(b) Convergence: $a = 1 + 10^{-10}, b = 0$.

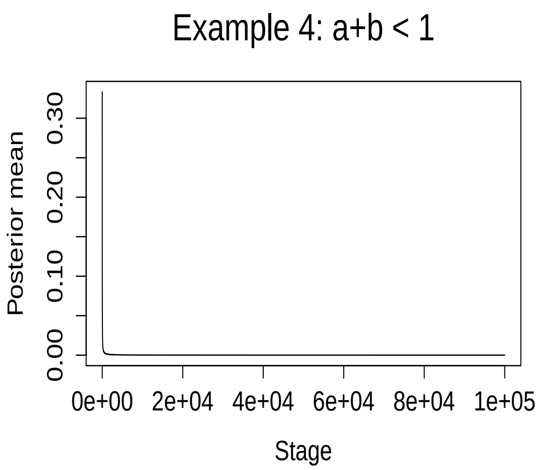


(c) Convergence: $a = 1 + 20^{-10}, b = 10^{-10}$.

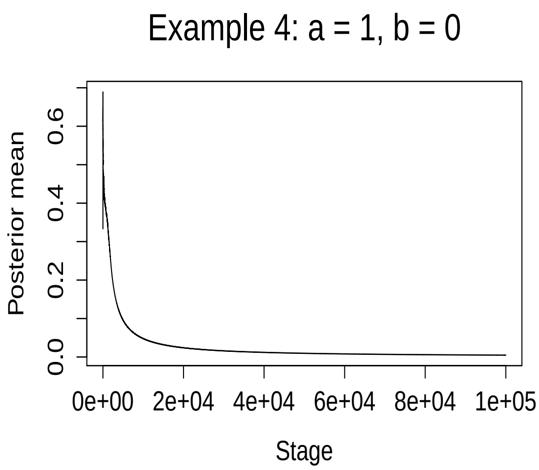


(d) Divergence: $a = 1/2, b = 1/3$.

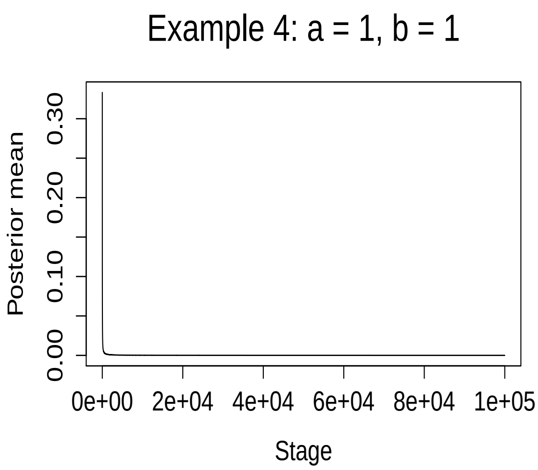
Figure 5.4: Example 4: The series (1.1) converges for $(a = 3, b = 1)$, $\left(a = 1 + 10^{-10}, b = 0\right)$, $\left(a = 1 + 20^{-10}, b = 10^{-10}\right)$ and diverges for $(a = 1/2, b = 1/3)$.



(a) Divergence: $a = \frac{1}{2}\left(1 - 10^{-11}\right), b = \frac{1}{2}\left(1 - 10^{-11}\right)$.



(b) Divergence: $a = 1, b = 0$.



(c) Divergence: $a = 1, b = 1$.

Figure 5.5: Example 4: The series (1.1) diverges for $\left(a = \frac{1}{2}\left(1 - 10^{-11}\right), b = \frac{1}{2}\left(1 - 10^{-11}\right)\right)$, $(a = 1, b = 0)$ and $(a = 1, b = 1)$.

Lemma 10 *For the series (5.13), let*

$$S_{j,n}^{a,b} = \sum_{i=5+n(j-1)}^{5+nj-1} \left(1 - \left(\frac{\log(i)}{i}\right) \left(a \left(1 + \sin^2 \left(\sqrt{\left(\frac{\log(\log(i))}{\log(i)}\right)}\right)\right) + b \sin\left(\frac{i\pi}{4}\right)\right)\right)^i, \tag{5.14}$$

for $j \geq 1$ and n, a multiple of 4. Then $S_{j,n}^{a,b}$ is decreasing in a and increasing in b.

The following corollary with respect to $S^{a,b}$ again holds:

Corollary 11 *$S^{a,b}$ is decreasing in a and increasing in b.*

Thus, we follow the same method as in Example 4 to determine $c_{j,n}^{a,b}$, but we need to note that in this example $a > 0$ and $b > 0$ instead of $a \geq 0$ and $b \geq 0$ of Example 4. Consequently, here we define $b \geq \epsilon$, for $\epsilon > 0$, the set A_ϵ given by (5.9) and

$$\tilde{S} = \inf_{a \in A_\epsilon} \sup_{b \geq \epsilon} \left\{S^{a,b} : a - b > 1\right\}. \tag{5.15}$$

In this case, Corollary 11 and the convergence criterion $a - b > 1$ ensure that $\tilde{S}$ is attained at $a_0 = 1 + \epsilon$ and $b_0 = \epsilon$. As before, we set $\epsilon = 10^{-10}$. The rest of the arguments leading to the choice of $c_{j,n}^{a,b}$ remains the same as in Example 4, and hence in this example $c_{j,n}^{a,b}$ has the same form as (5.11), with $a_0 = 1 + 10^{-10}$, $b_0 = 10^{-10}$.

Figure 5.6 depicts the results of our Bayesian analysis of the series (5.13) for various values of a and b. All the results are in accordance wth those of Bourchtein *et al.* (2012).

5.6 Example 6

We now investigate whether or not the following series converges:

$$S = \sum_{i=1}^{\infty} \frac{1}{i^3 |\sin i|}. \tag{5.16}$$

This series is a special case of the generalized form of the Flint Hills series (see Pickover (2002) and Alekseyev (2011)).

For our purpose, we first embed the above series into

$$S^{a,b} = \sum_{i=1}^{\infty} \frac{i^{b-3}}{a + |\sin i|}, \tag{5.17}$$

where $b \in \mathbb{R}$ and $|a| \leq \eta$, for some $\eta > 0$, specified according to our purpose. Note that, $S = S^{0,0}$, and we set $\eta = 10^{-10}$ for our investigation of (5.16).

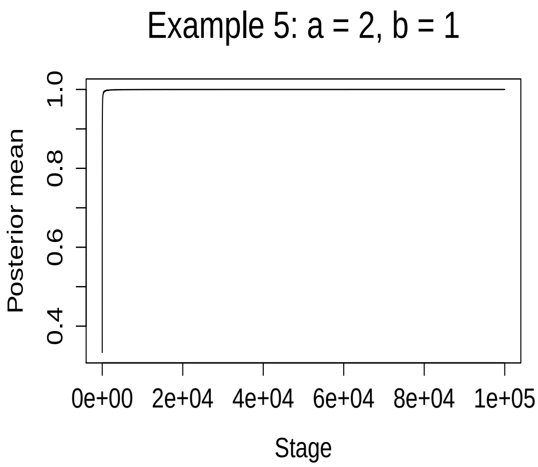
Note that for any fixed $a \neq 0$, $S^{a,b}$ converges if $b < 2$ and diverges if $b \geq 2$. Since $S^{a,b}$ increases in b it follows that the equality in

$$\tilde{S} = \sup \left\{S^{a,b} : a = \epsilon,\ b \leq 2 - \epsilon\right\} \tag{5.18}$$

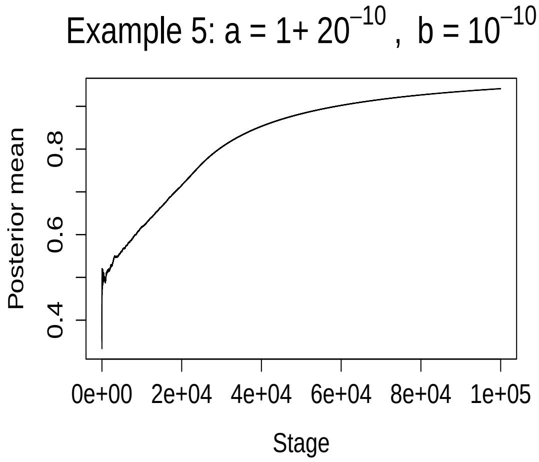
is attained at $(a_0, b_0) = (\epsilon, 2 - \epsilon)$.

Arguments in keeping with those in the previous examples lead to the following choice of the upper bound for $S_{j,n}^{a,b}$, which we again denote by $c_{j,n}^{a,b}$:

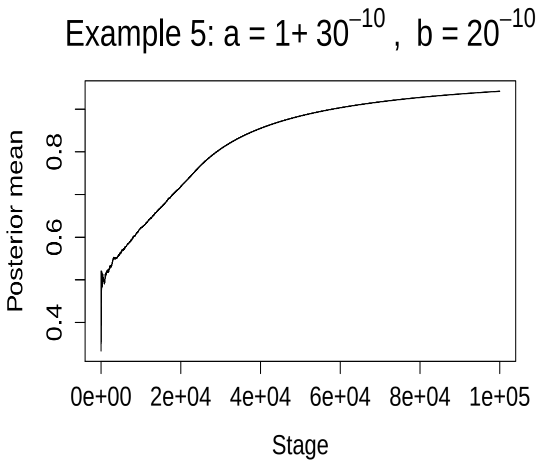
$$c_{j,n}^{a,b} = \begin{cases} u_{j,n}^{a,b}, & \text{if } b < 2; \\ v_{j,n}^{a,b}, & \text{otherwise,} \end{cases} \tag{5.19}$$



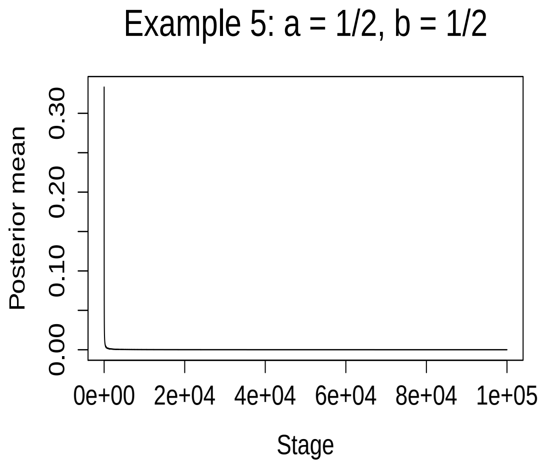
(a) Convergence: $a = 2, b = 1$.



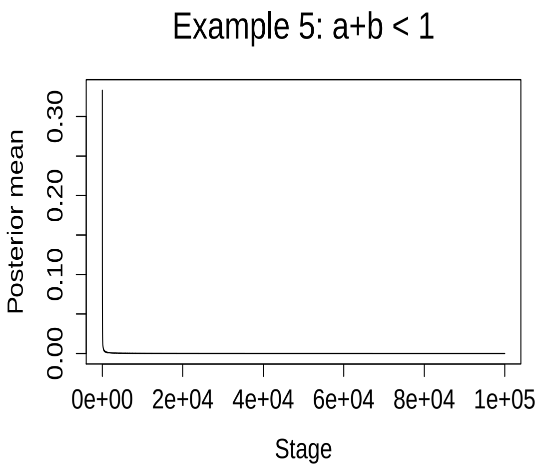
(b) Convergence: $a = 1 + 20^{-10}, b = 10^{-10}$.



(c) Convergence: $a = 1 + 30^{-10}, b = 20^{-10}$.

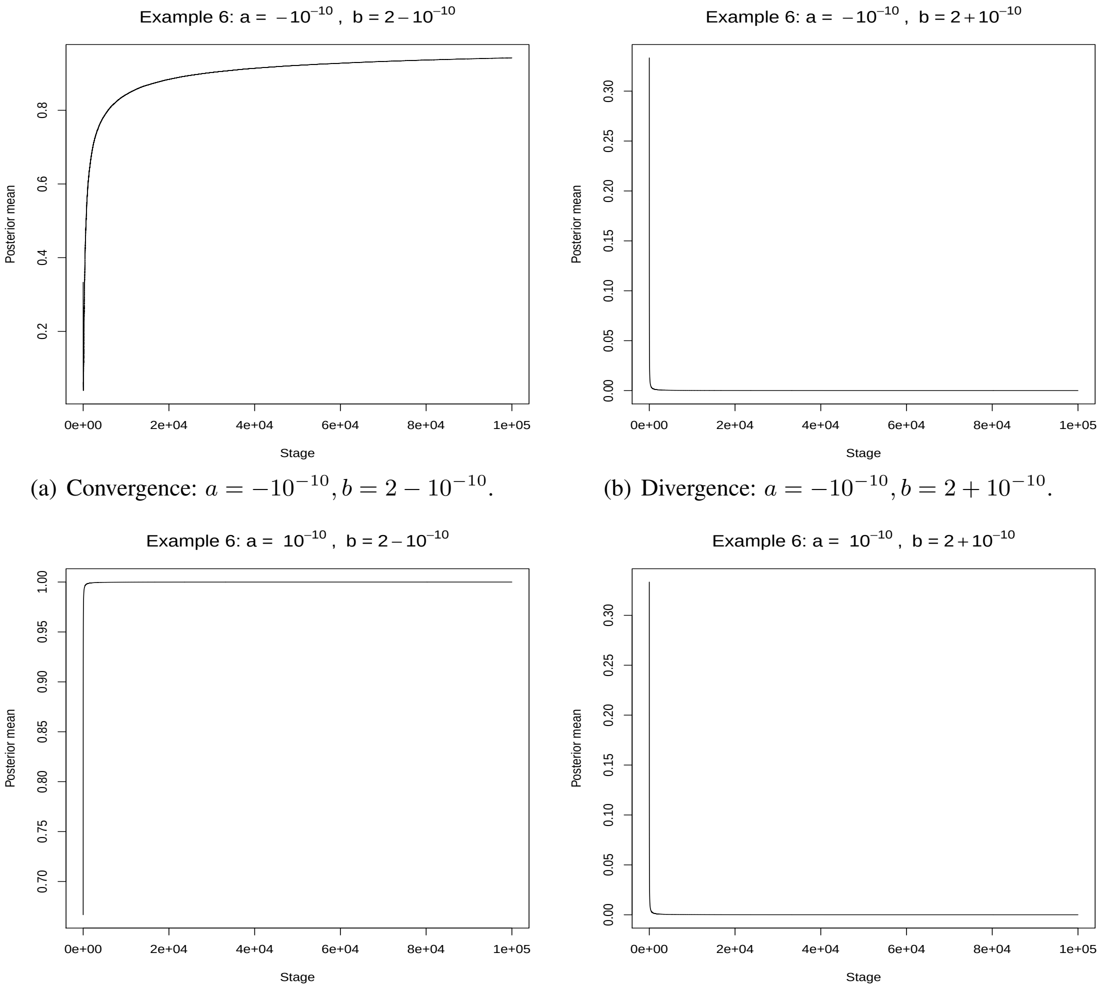


(d) Divergence: $a = 1/2, b = 1/2$.



(e) Divergence: $a = \frac{1}{2}\left(1 - 10^{-11}\right), b = \frac{1}{2}\left(1 - 10^{-11}\right)$.

Figure 5.6: Example 5: The series (5.13) converges for $(a = 2, b = 1)$, $(a = 1 + 20^{-10}, b = 10^{-10})$, $(a = 1 + 30^{-10}, b = 20^{-10})$ and diverges for $(a = 1/2, b = 1/2)$ and $\left(a = \frac{1}{2}\left(1 - 10^{-11}\right), b = \frac{1}{2}\left(1 - 10^{-11}\right)\right)$.



(a) Convergence: $a = -10^{-10}, b = 2 - 10^{-10}$.

(b) Divergence: $a = -10^{-10}, b = 2 + 10^{-10}$.

(c) Convergence: $a = 10^{-10}, b = 2 - 10^{-10}$.

(d) Divergence: $a = 10^{-10}, b = 2 + 10^{-10}$.

Figure 5.7: Example 6: The series (5.17) converges for $(a = -10^{-10}, b = 2 - 10^{-10})$, $(a = 10^{-10}, b = 2 - 10^{-10})$, and diverges for $(a = -10^{-10}, b = 2 + 10^{-10})$, $(a = 10^{-10}, b = 2 + 10^{-10})$.

where

$$u^{a,b}_{j,n} = S^{a_0,b_0}_{j,n} + \frac{(|a| - b + 2 - 2\epsilon + 10^{-5})}{\log(j+1)}; \tag{5.20}$$

$$v^{a,b}_{j,n} = S^{a_0,b_0}_{j,n} + \frac{(|a| - b + 2 - 2\epsilon - 10^{-5})}{\log(j+1)}. \tag{5.21}$$

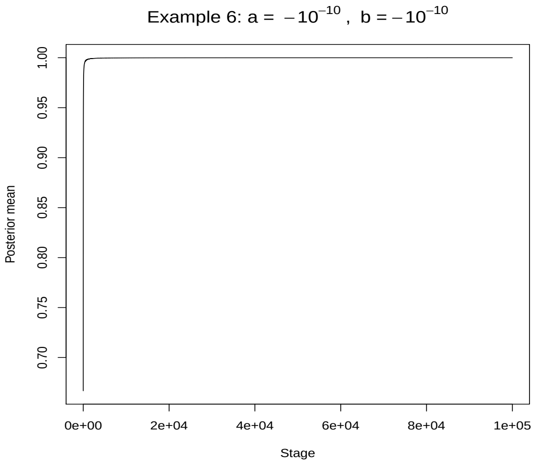
Notice that we add the term 10^{-5} when $b < 2$ so that our Bayesian method favours convergence and subtract the same when $b \geq 2$ to facilitate detection of divergence. Since convergence or divergence of $S^{a,b}$ does not depend upon $a \in [-\eta, \eta] \setminus \{0\}$, we use $|a|$ in (5.20) and (5.21).

Setting $\epsilon = 10^{-10}$, Figures 5.7 and 5.8 depict convergence and divergence of $S^{a,b}$ for various values of a and b. In particular, panel (e) of Figure 5.8 shows that our main interest, the series S, given by (5.16), converges.

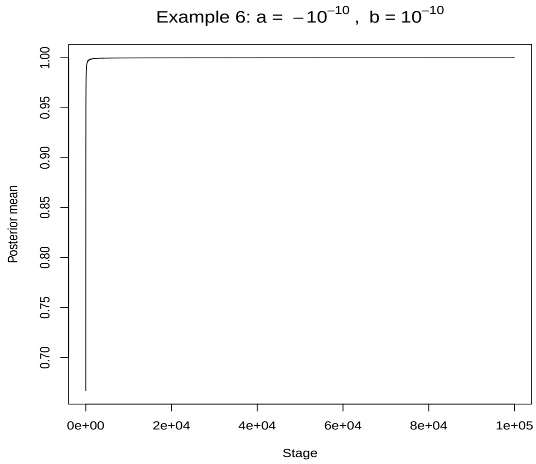
5.7 Example 7

We now consider

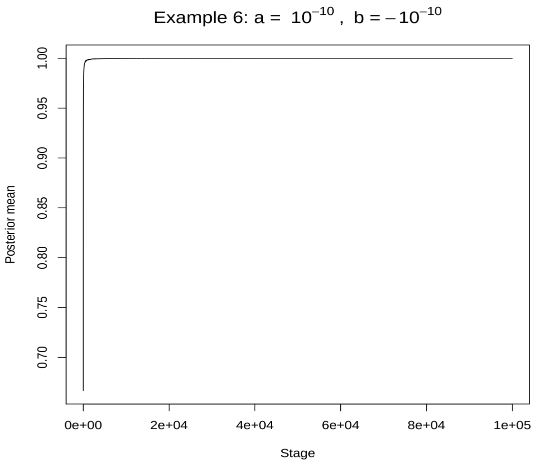
$$S = \sum_{i=1}^{\infty} \frac{|\sin\ i|^i}{i}. \tag{5.22}$$



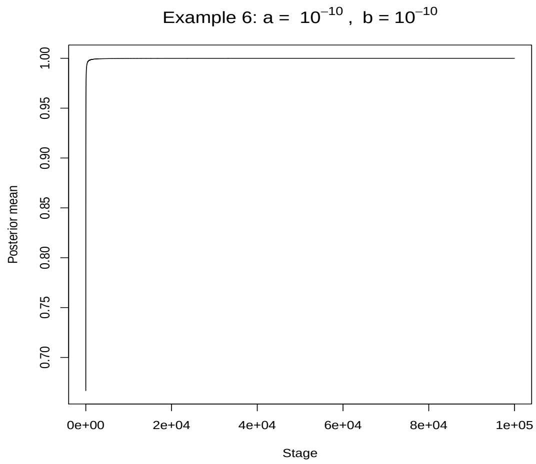
(a) Convergence: $a = -10^{-10}, b = -10^{-10}$.



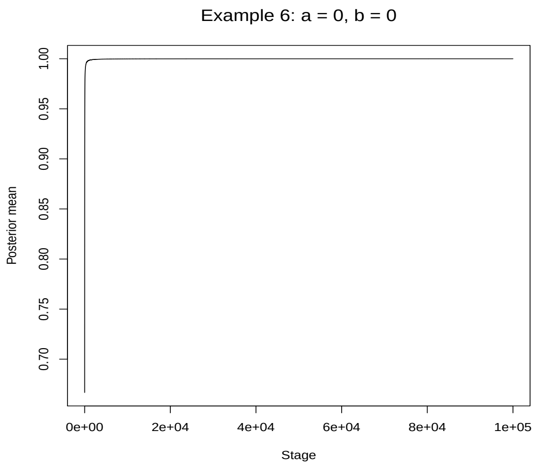
(b) Convergence: $a = -10^{-10}, b = 10^{-10}$.



(c) Convergence: $a = 10^{-10}, b = -10^{-10}$.

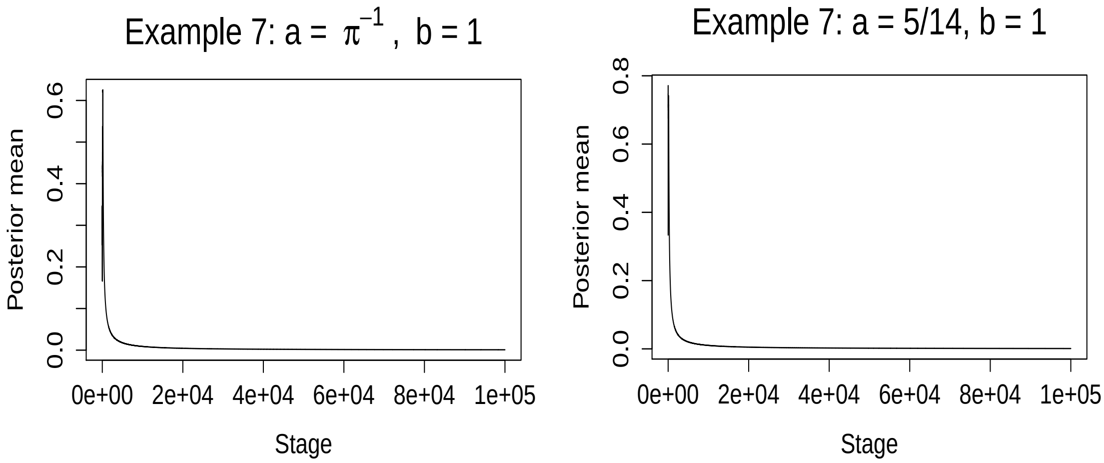


(d) Convergence: $a = 10^{-10}, b = 10^{-10}$.



(e) Convergence: $a = 0, b = 0$.

Figure 5.8: Example 6: The series (5.17) converges for $(a = -10^{-10}, b = -10^{-10})$, $(a = -10^{-10}, b = 10^{-10})$, $(a = 10^{-10}, b = -10^{-10})$, $(a = 10^{-10}, b = 10^{-10})$, and $(a = 0, b = 0)$.



(a) Divergence: $a = \pi^{-1}, b = 1$. (b) Divergence: $a = 5/(2 \times 7), b = 1$.

Figure 5.9: Example 7: The series (5.23) diverges for $(a = \pi^{-1}, b = 1)$, $(a = 5/7, b = 1)$.

We embed this series into

$$S^{a,b} = \sum_{i=1}^{\infty} \frac{|\sin\ a\pi i|^i}{i^b}, \tag{5.23}$$

where $a \in \mathbb{R}$ and $b \geq 1$. The above series converges if $b > 1$, for all $a \in \mathbb{R}$. But for $b = 1$, it is easy to see that the series diverges if $a = \ell/2m$, where ℓ and m are odd integers.

Letting $a_0 = \pi^{-1}$ and $b_0 = 1 + \epsilon$, with $\epsilon = 10^{-10}$, we set the following upper bound:

$$u_{j,n}^{a,b} = S_{j,n}^{a_0,b_0} + \frac{\epsilon}{j}. \tag{5.24}$$

Thus, $u_{j,n}^{a,b}$ corresponds to a convergent series which is also sufficiently close to divergence. Addition of the term $\frac{\epsilon}{j}$ provides further protection from erroneous conclusions regarding divergence.

Panel(a) of Figure 5.9 demonstrates that the series of our interest, given by (5.22), diverges. Panel (b) confirms that for $a = 5/(2 \times 7)$ and $b = 1$, the series indeed diverges, as it should.

6 Application to Riemann Hypothesis

6.1 Brief background

Consider the Riemann zeta function given by

$$\zeta(a) = \frac{1}{1 - 2^{1-a}} \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!(n-k)!} (k+1)^{-a}, \tag{6.1}$$

where a is complex. The above function is formed by first considering Euler's function

$$Z(a) = \sum_{n=1}^{\infty} \frac{1}{n^a}, \tag{6.2}$$

then by multiplying both sides of (6.2) by $\left(1 - \frac{2}{2^a}\right)$ to obtain

$$\left(1 - \frac{2}{2^a}\right) Z(a) = \sum_{n=1}^{\infty} \frac{(-1)^{a+1}}{n^a}, \tag{6.3}$$

and then dividing the right hand side of (6.3) by $\left(1 - \frac{2}{2^a}\right)$. The advantage of the function $\zeta(a)$ in comparison with the parent function $Z(a)$ is that, $Z(a)$ is divergent if the real part of a, which we denote by $Re(a)$, is less than or equal to 1, while $\zeta(a)$ is convergent for all a with $Re(a) > 0$. Importantly, $\zeta(a) = Z(a)$ whenever $Z(a)$ is convergent.

Whenever $0 < Re(a) < 1$, $\zeta(a)$ satisfies the following identity:

$$\zeta(a) = 2^s \pi^{a-1} \sin\left(\frac{\pi a}{2}\right) \Gamma(1-a)\zeta(1-a), \tag{6.4}$$

where $\Gamma(\cdot)$ is the gamma function. This can be extended to the set of complex numbers by defining a function with non-positive real part by the right hand side of (6.4); abusing notation, we denote the new function by $\zeta(a)$. Because of the sine function, it follows that the trivial zeros of the above function occur when the values of a are negative even integers. Hence, the non-trivial zeros must satisfy $0 < Re(a) < 1$.

Riemann (1859) conjectured that all the non-trivial zeros have the real part $1/2$, which is the famous Riemann Hypothesis. For accessible account of the Riemann Hypothesis, see Borwein *et al.* (2006), Derbyshire (2004).

One equivalent condition for the Riemann Hypothesis is related to sums of of the Möbius function, given by

$$\mu(n) = \begin{cases} -1 & \text{if} \quad n \text{ is a square-free positive integer with an odd number of prime factors;} \\ 0 & \text{if} \quad n \text{ has a squared prime factor;} \\ 1 & \text{if} \quad n \text{ is a square-free positive integer with an even number of prime factors,} \end{cases} \tag{6.5}$$

where, by square-free integer we mean that the integer is not divisible by any perfect square other than 1. Specifically, the condition

$$\sum_{n=1}^{x} \mu(n) = O\left(x^{\frac{1}{2}+\epsilon}\right) \tag{6.6}$$

for any $\epsilon > 0$, is equivalent to Riemann Hypothesis. This condition implies that the Dirichlet series for the Möbius function, given by

$$M(a) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^a} = \frac{1}{\zeta(a)}, \tag{6.7}$$

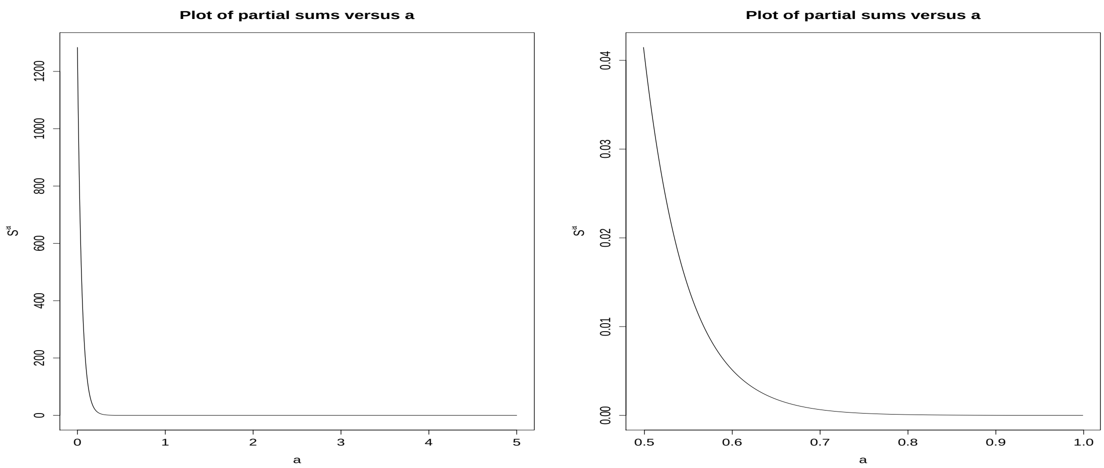
is analytic in $Re(a) > 1/2$. This again ensures that $\zeta(a)$ is meromorphic in $Re(a) > 1/2$ and that it has no zeros in this region. Using the functional equation (6.4) it follows that there are no zeros of $\zeta(a)$ in $0 < Re(a) < 1/2$ either. Hence, (6.6) implies Riemann Hypothesis. The converse is also certainly true.

The above arguments also imply that convergence of $M(a)$ in (6.7) for $Re(a) > 1/2$ is equivalent to Riemann Hypothesis, and it is this criterion that is of our interest in this paper. Now, $M(a)$ converges absolutely for $Re(a) > 1$ and $\sum_{n=1}^{N} \mu(n) = o(N)$, as $N \to \infty$, showing that $M(1) < \infty$. The latter is equivalent to the prime number theorem stating that the number of primes below x is asymptotically $x/\log(x)$, as $x \to \infty$. Thus, $M(a)$ converges for $Re(a) \geq 1$. That $M(a)$ diverges for $Re(a) \leq 1/2$ can be seen as follows. Note that if $M(a)$ converged for any a^* such that $Re(a^*) \leq 1/2$, then analytic continuation for Dirichlet series of the form $M(a)$ would guarantee convergence of $M(a)$ for all a with $Re(a) > Re(a^*)$. But $\zeta(a)$ is not analytic on $0 < Re(a) < 1$ because of its non-trivial zeros on the strip. This would contradict the analytic continuation leading to the identity $M(a) = 1/\zeta(a)$ on the entire set of complex numbers. Hence, $M(a)$ must be divergent for $Re(a) \leq 1/2$.

In this paper, we apply our ideas to particularly investigate convergence of $M(a)$ when $1/2 < a < 1$.

6.2 Choice of the upper bound and implementation details

To form an idea of the upper bound we first plot the partial sums $S^a_{j,n}$, for $j = 1000$ and $n = 10^6$, with respect to a. In this regard, panel (a) of Figure 6.1 shows the decreasing nature of the partial sums



(a) Plot of partial sums in the domain $(0, 5)$. (b) Plot of partial sums in the domain $(0.5, 1)$.

Figure 6.1: Plot of the partial sums $S^a_{1000,1000000}$ versus a. Panel (a) shows the plot in the domain $[0, 5]$ while panel (b) magnifies the same in the domain $(0.5, 1)$.

with respect to a, and panel (b) magnifies the plot in the domain $1/2 < a < 1$ that we are particularly interested in. The latter shows that the partial sums decrease sharply till about 0.7, getting appreciably close to zero around that point, after which the rate of decrease diminishes. Thus, one may expect a change point around 0.7 regarding convergence. Specifically, divergence may be expected below a point slightly larger than 0.7 and convergence above it.

Since $M(1) < \infty$, we consider this series as the basis for our upper bound, with the value of a also taken into account. Specifically, we choose the upper bound as

$$c_{j,n} = \left| S^1_{j,n} + \frac{a}{j+1} \right|. \tag{6.8}$$

Since Figure 6.1 shows that the partial sums are of monotonically decreasing nature, the above choice of upper bound facilitates detection of convergence for relatively large values of a. The part $\frac{a}{j+1}$, which tends to zero as $j \to \infty$, takes care of the fact that the series may be convergent if $a < 1$, by slightly inflating $S^1_{j,n}$.

For our purpose, we compute the first 10^9 values of the Möbius function using an efficient algorithm proposed in Lioen and van de Lune (1994), which is based on the Sieve of Eratosthenes (Horsley (1772)). We set $K = 1000$ and $n = 10^6$. A complete analysis with our VMware with our parallel implementation takes about 2 minutes.

6.3 Results of our Bayesian analysis

Panels (a)–(e) of Figure 6.2 and panels (d)–(f) of Figure 6.3 show the $M(a)$ diverges for $a = 0.1$, 0.2, 0.3, 0.4, 0.5, but converges for $a = 1 + 10^{-10}$, 2 and 3. In fact, for many other values that we experimented with, $M(a)$ converged for $a > 1$ and diverged for $a < 1/2$, demonstrating remarkable consistency with the known, existing results.

Certainly far more important are the results for $1/2 < a < 1$. Indeed, panel (f) of Figure 6.2 and panels (a)–(c) of Figure 6.3 show that $M(a)$ diverged for $a = 0.6$ and 0.7 and converged for $a = 0.8$ and 0.9. It thus appears that $M(a)$ diverges for $a < a^*$ and converges for $a \geq a^*$, for some $a^* \in (0.7, 0.8)$. Figure 6.4 displays results of our further experiments in this regard. Panels (a) and (b) of Figure 6.4 show the posterior means for the full set of iterations and the last 500 iterations, respectively, for $a = 0.71$. Note that from panel (a), convergence seems to be attained, although towards the end, the plot seems to be slightly tilted downwards. Panel (b) magnifies this, clearly showing divergence. Panels (c) and

(d) of Figure 6.4 depict similar phenomenon for $a = 0.715$, but as per panel (d), divergence seems to ensue all of a sudden, even after showing signs of convergence for the major number of iterative stages. Convergence of $M(s)$ begins at $a = 0.72$ (approximately); panels (e) and (f) of Figure 6.4 take clear note of this.

Thus, as per our methods, $M(s)$ diverges for $a < 0.72$ and converges for $a \geq 0.72$. This is remarkably in keeping with the wisdom gained from panel (b) of Figure 6.1 that convergence is expected to occur for values of a exceeding 0.7. Note that neither the upper bound (6.8), nor our methodology, is in any way biased towards $a \approx 0.7$; hence, our result is perhaps not implausible.

6.4 Implications of our result

As per our results, $M(s)$ does not converge for all $s > 1/2$, and hence does not completely support Riemann Hypothesis. However, convergence of $M(s)$ fails only for the relatively small region $0.5 < a < 0.72$, which perhaps is the reason why there exists much evidence in favour of Riemann Hypothesis.

7 Oscillatory series with multiple limit points

In this section we assume that the sequence $\{S_{1,n}\}_{n=1}^{\infty}$ has multiple limit points, including the possibility that the number of limit points is countably infinite. For the time being, let us assume that each term of the series is non-negative.

7.1 Finite number of limit points

Let us assume that there are $M\ (>1)$ limit points. We define

$$Y_j = m \ \text{ if } \ c_{j,m-1} < S_{1,j} \leq c_{m,j};\ m = 1, 2, \ldots, M, \tag{7.1}$$

where, for $j \geq 1$, $c_{0,j}, c_{1,j}, \ldots, c_{M,j}$ appropriately partition $\mathbb{R}^+$, the non-negative part of the real line. Choice of these quantities will be discussed in Section 7.2.

Note that unlike our ideas appropriate for non-oscillating series, here do not consider blocks of partial sums $S_{j,n} = \sum_{i=(j-1)n+1}^{jn} X_i$, but $S_{1j} = \sum_{i=1}^{j} X_i$. In other words, for Bayesian analysis of non-oscillating series we compute sums of n terms in each iteration, whereas for oscillating series we keep adding a single term at every iteration. Thus, computationally, the latter is a lot simpler.

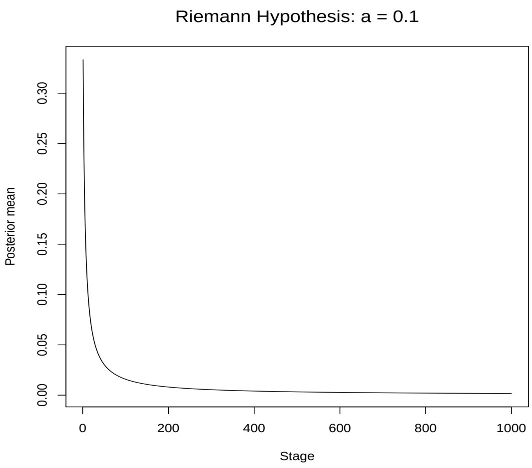
We assume that

$$(\mathbb{I}(Y_j = 1), \ldots, \mathbb{I}(Y_j = M)) \sim Multinomial\,(1, p_{1,j}, \ldots, p_{M,j})\,, \tag{7.2}$$

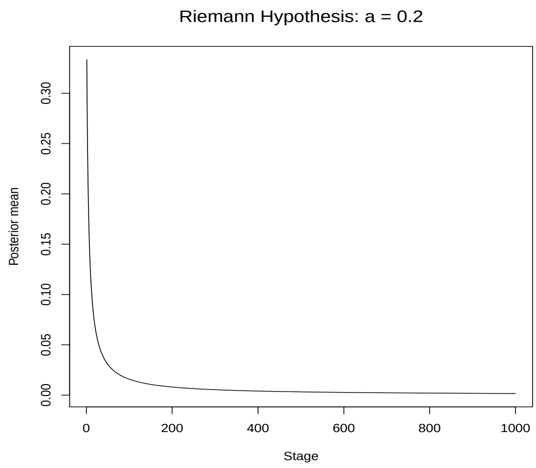
where $p_{m,j}$ can be interpreted as the probability that $S_{1,j} \in (c_{j,m-1}, c_{j,m}]$. As $j \to \infty$ it is expected that $c_{m-1,j}$ and $c_{m,j}$ will converge to appropriate constants depending upon m, and that $p_{m,j}$ will tend to the correct proprotion of the limit point indexed by m. Indeed, let $\{p_{m,0};\ m = 1, \ldots, M\}$ denote the actual proportions of the limit points indexed by $\{1, \ldots, M\}$, as $j \to \infty$.

Following the same principle discussed in Section 3, at the k-th stage we arrive at the following posterior of $\{p_{m,k} : m = 1, \ldots, M\}$:

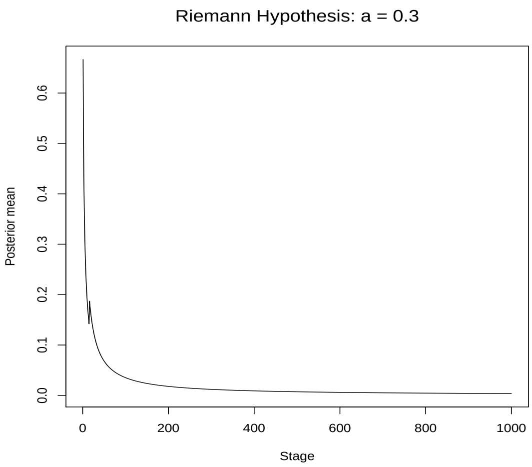
$$\pi\left(p_{1,k}, \ldots, p_{M,k}|y_k\right) \equiv Dirichlet\left(\sum_{j=1}^{k}\frac{1}{j^2} + \sum_{j=1}^{k}\mathbb{I}\left(y_j = 1\right), \ldots, \sum_{j=1}^{k}\frac{1}{j^2} + \sum_{j=1}^{k}\mathbb{I}\left(y_j = M\right)\right). \tag{7.3}$$



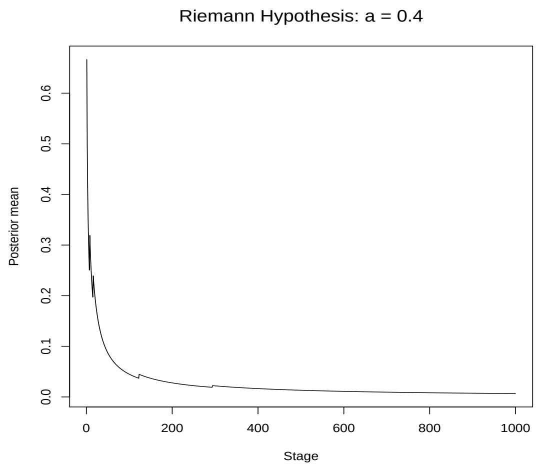
(a) Divergence: $a = 0.1$.



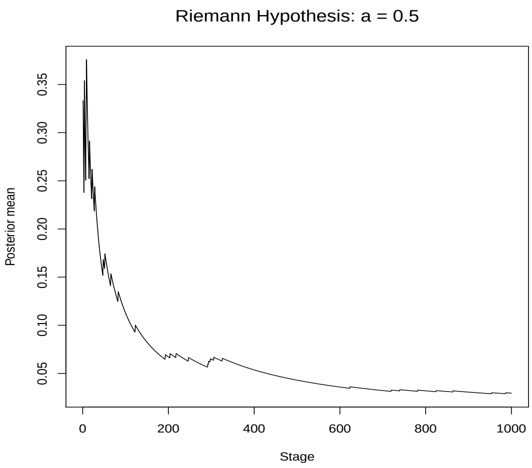
(b) Divergence: $a = 0.2$.



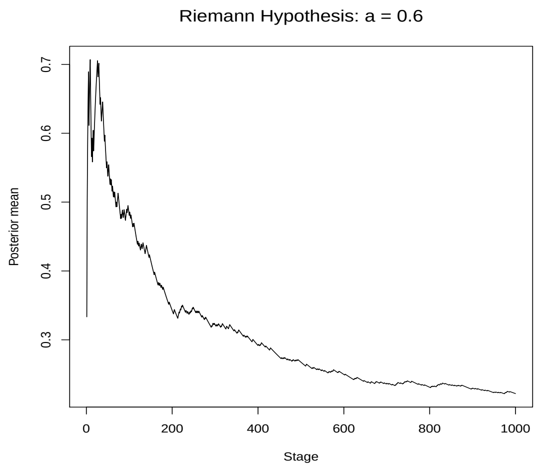
(c) Divergence: $a = 0.3$.



(d) Divergence: $a = 0.4$.

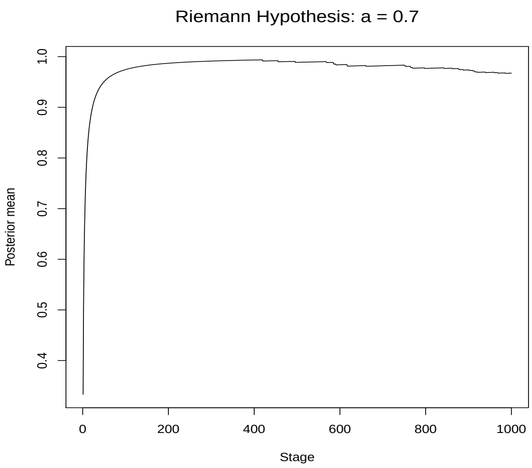


(e) Divergence: $a = 0.5$.

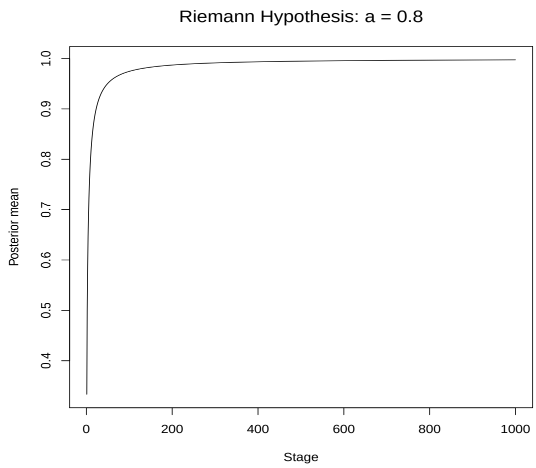


(f) Divergence: $a = 0.6$.

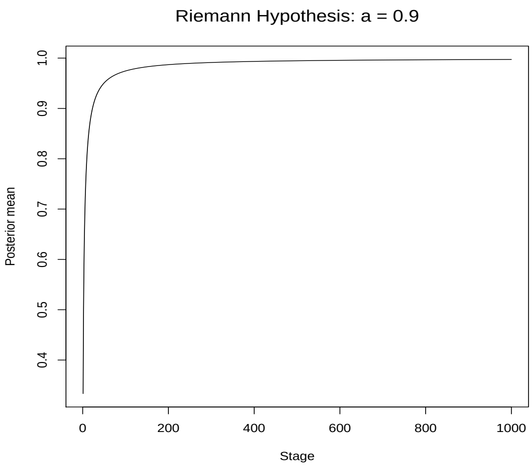
Figure 6.2: Riemann Hypothesis: The mobius function based series diverges for $a = 0.1$, 0.2, 0.3, 0.4, 0.5, 0.6.



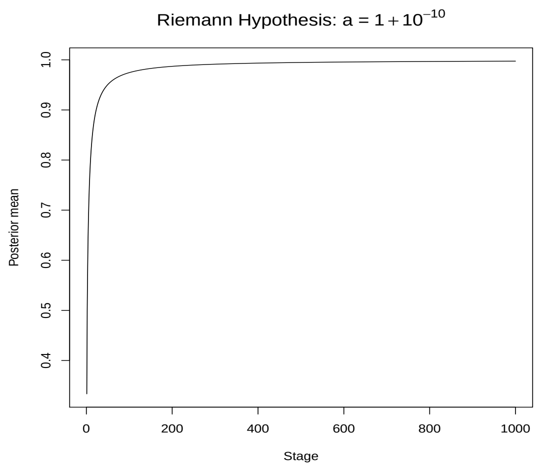
(a) Divergence: $a = 0.7$.



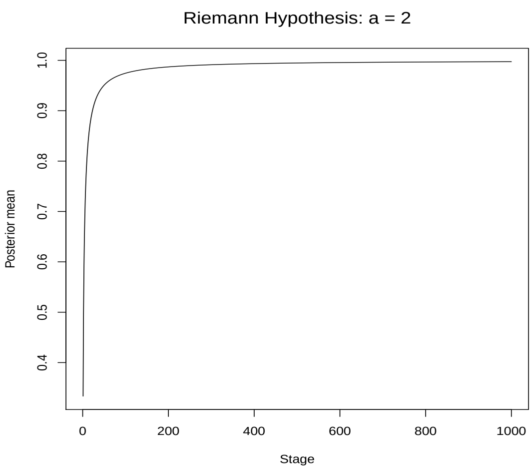
(b) Convergence: $a = 0.8$.



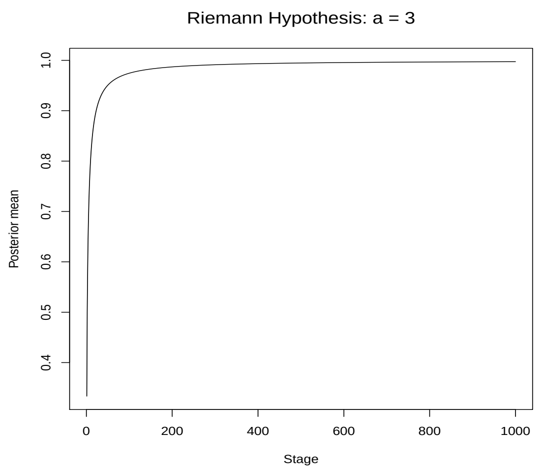
(c) Convergence: $a = 0.9$.



(d) Convergence: $a = 1 + 10^{-10}$.

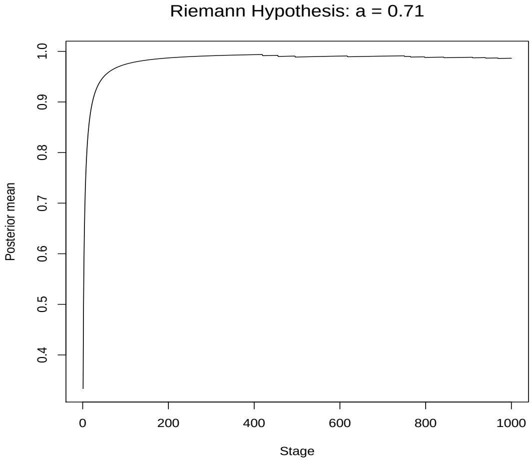


(e) Convergence: $a = 2$.

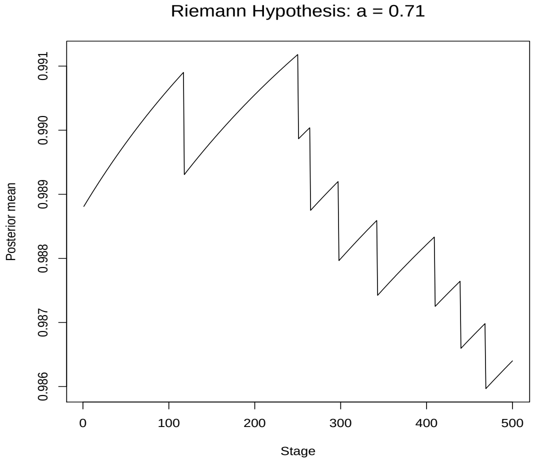


(f) Convergence: $a = 3$.

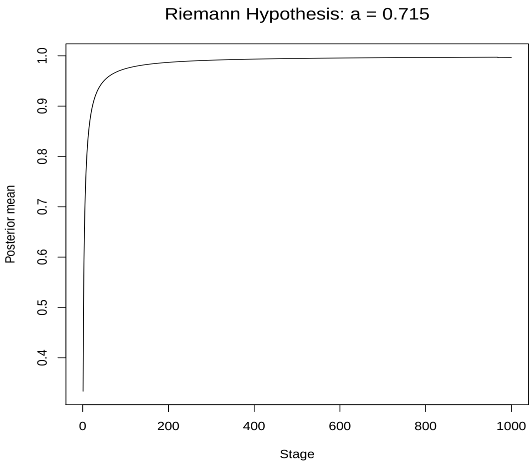
Figure 6.3: Riemann Hypothesis: The mobius function based series diverges for $a = 0.7$ but converges for $a = 0.8$, 0.9, $1 + 10^{-10}$, 2, 3.



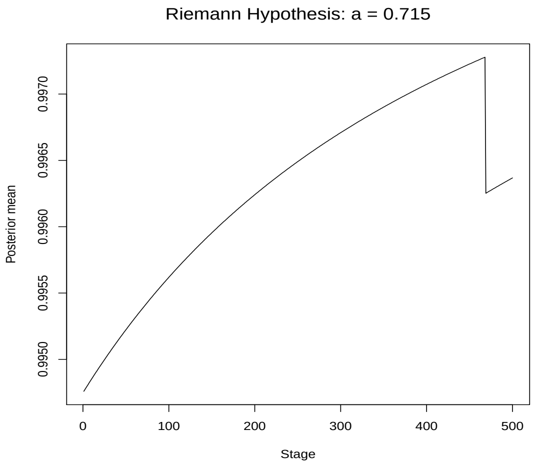
(a) Divergence: $a = 0.71$.



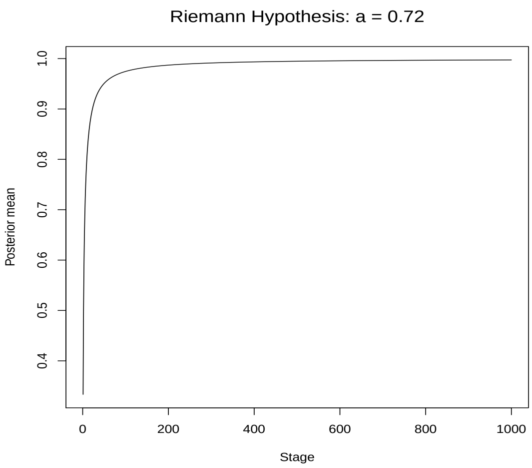
(b) Divergence: $a = 0.71$.



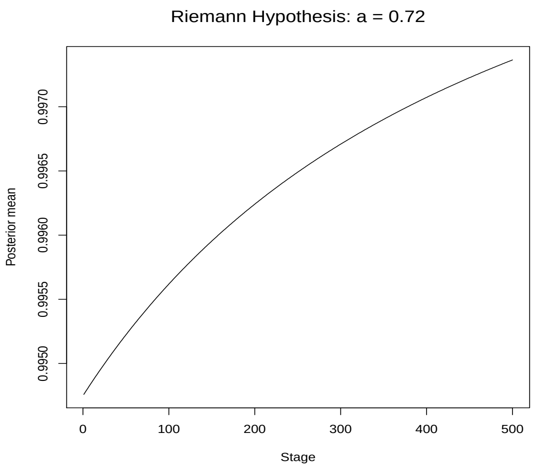
(c) Divergence: $a = 0.715$.



(d) Divergence: $a = 0.715$.



(e) Convergence: $a = 0.72$.



(f) Convergence: $a = 0.72$.

Figure 6.4: Riemann Hypothesis: The left panels show the posterior means for the full set of iterations, while the right panels depict the posterior means for the last 500 iterations, for $a = 0.71$, 0.715 and 0.72. It is evident that the mobius function based series diverges for $a = 0.71$ and 0.715 but converges for $a = 0.72$.

The posterior mean and posterior variance of $p_{m,k}$, for $m = 1, \ldots, M$, are given by:

$$E\left(p_{m,k}|y_k\right) = \frac{\sum_{j=1}^{k}\frac{1}{j^2} + \sum_{j=1}^{k}\mathbb{I}\left(y_j = m\right)}{M\sum_{j=1}^{k}\frac{1}{j^2} + k}; \tag{7.4}$$

$$Var\left(p_{m,k}|y_k\right) = \frac{\left(\sum_{j=1}^{k}\frac{1}{j^2} + \sum_{j=1}^{k}\mathbb{I}\left(y_j = m\right)\right)\left((M-1)\sum_{j=1}^{k}\frac{1}{j^2} + k - \sum_{j=1}^{k}\mathbb{I}\left(y_j = m\right)\right)}{\left(M\sum_{j=1}^{k}\frac{1}{j^2} + k\right)^2\left(M\sum_{j=1}^{k}\frac{1}{j^2} + k + 1\right)}. \tag{7.5}$$

Let $k = M\tilde{k}$, where $\tilde{k} \to \infty$. Then, from (7.4) and (7.5) it is easily seen, using $\frac{\sum_{j=1}^{k}\mathbb{I}(y_j=m)}{k} \to p_{m,0}$ almost surely as $k \to \infty$, that almost surely,

$$E\left(p_{m,k}|y_k\right) \to p_{m,0}, \text{ and} \tag{7.6}$$

$$Var\left(p_{m,k}|y_k\right) = O\left(\frac{1}{k}\right) \to 0, \tag{7.7}$$

as $k \to \infty$.

We can now characterize the m limit points of $S_{1,\infty}$ in terms of the limits of the marginal posterior probabilities of $p_{m,k}$, denoted by $\pi_m\left(\cdot|y_k\right)$, as $k \to \infty$.

Theorem 12 *$\{S_{1,n}\}_{n=1}^{\infty}$ has M (>1) limit points if and only if for $m = 1, \ldots, M$,*

$$\pi_m\left(\mathcal{N}_{p_{m,0}}|y_k\right) \to 1, \tag{7.8}$$

as $k \to \infty$, almost surely for all $Y = \{y_k\}_{k=1}^{\infty}$. In the above, $\mathcal{N}_{p_{m,0}}$ is any neighborhood of $p_{m,0}$.

Proof. Follows using the same ideas as the proof of Theorem 7. ■

7.2 Choice of $c_{j,0}, \ldots, c_{j,M}$

Let us define, for $j = 1, 2, \ldots, k$,

$$\tilde{p}_{\ell,j} = \begin{cases} 0 & \text{if} \quad \ell = 0; \\ E\left(p_{\ell,j}|y_j\right) & \text{if} \quad \ell = 1, 2, \ldots, M. \end{cases} \tag{7.9}$$

We then set $c_{j,0} \equiv 0$ for all $j = 1, 2, \ldots, k$, and, for $m \geq 1$, define

$$c_{j,m} = \frac{\sum_{\ell=1}^{m}\tilde{p}_{\ell,k}}{1 - \sum_{\ell=1}^{m}\tilde{p}_{\ell,k}}, \tag{7.10}$$

for $j = 1, 2, \ldots, k$. Thus, the inequality $c_{j,m-1} < S_{1,j} \leq c_{m,j}$ in (7.1) is equivalent to

$$\sum_{\ell=1}^{m-1}\tilde{p}_{\ell,k} < \left(\frac{S_{1,j}}{1 + S_{1,j}}\right)^{\rho(\theta)} \leq \sum_{\ell=1}^{m}\tilde{p}_{\ell,k}, \tag{7.11}$$

where $\rho(\theta)$ is some relevant power depending upon the set of parameters θ of the series, responsible for appropriately inflating or contracting the quantity $\frac{S_{1,j}}{1+S_{1,j}}$ for properly diagnosing the limit points. If $\left(\frac{S_{1,j}}{1+S_{1,j}}\right)^{\rho(\theta)} \geq 1$, we set $Y_j = M$. By (7.1) and (7.2) it then holds that the probability of the event (7.11) converges to $p_{m,0}$, as $k \to \infty$.

7.3 Infinite number of limit points

We now assume that the number of limits points of $\{S_{1,n}\}_{n=1}^{\infty}$ is countably infinite, and that $\{p_{m,0}; m = 1, 2, 3, \ldots\}$, where $0 \leq p_{m,0} \leq 1$ and $\sum_{m=1}^{\infty} p_{m,0} = 1$, are the true proportions of the limit points.

Now we define

$$Y_j = m \ \text{ if } \ c_{j,m-1} < S_{1,j} \leq c_{m,j}; \ m = 1, 2, \ldots, \infty, \tag{7.12}$$

where, for $j \geq 1$, $c_{0,j}, c_{1,j}, \ldots$, are appropriately chosen constants that partition $\mathbb{R}^+$.

Let $\mathcal{X} = \{1, 2, \ldots\}$ and let $\mathcal{B}(\mathcal{X})$ denote the Borel σ-field on $\mathcal{X}$ (assuming every singleton of $\mathcal{X}$ is an open set). Let $\mathcal{P}$ denote the set of probability measures on $\mathcal{X}$. Then, at the j-th stage,

$$[Y_j|P_j] \sim P_j, \tag{7.13}$$

where $P_j \in \mathcal{P}$. We assume that P_j is the following Drichlet process:

$$P_j \sim DP\left(\frac{1}{j^2}G\right), \tag{7.14}$$

where, the probability measure G is such that, for every $j \geq 1$,

$$G(Y_j = m) = \frac{1}{2^m}. \tag{7.15}$$

It then follows using the same previous principles that, at the k-th stage, the posterior of P_k is again a Dirichlet process, given by

$$[P_k|y_k] \sim DP\left(\sum_{j=1}^{k} \frac{1}{j^2}G + \sum_{j=1}^{k} \delta_{y_j}\right), \tag{7.16}$$

where δ_{y_j} denotes point mass at y_j. It follows from (7.16) that

$$E(p_{m,k}|y_k) = \frac{\frac{1}{2^m}\sum_{j=1}^{k}\frac{1}{j^2} + \sum_{j=1}^{k}\mathbb{I}(y_j = m)}{\sum_{j=1}^{k}\frac{1}{j^2} + k}; \tag{7.17}$$

$$Var(p_{m,k}|y_k) = \frac{\left(\sum_{j=1}^{k}\frac{1}{j^2} + \sum_{j=1}^{k}\mathbb{I}(y_j = m)\right)\left((1 - \frac{1}{2^m})\sum_{j=1}^{k}\frac{1}{j^2} + k - \sum_{j=1}^{k}\mathbb{I}(y_j = m)\right)}{\left(\sum_{j=1}^{k}\frac{1}{j^2} + k\right)^2\left(\sum_{j=1}^{k}\frac{1}{j^2} + k + 1\right)}. \tag{7.18}$$

As before, it easily follows from (7.17) and (7.18) that for $m = 1, 2, 3, \ldots$,

$$E(p_{m,k}|y_k) \to p_{m,0}, \ \text{ and} \tag{7.19}$$

$$Var(p_{m,k}|y_k) = O\left(\frac{1}{k}\right) \to 0, \tag{7.20}$$

almost surely, as $k \to \infty$.

The theorem below characterizes countable number of limit points of $S_{1,\infty}$ in terms of the limit of the marginal posterior probabilities of $p_{m,k}$, as $k \to \infty$.

Theorem 13 *$\{S_{1,n}\}_{n=1}^{\infty}$ has countable limit points if and only if for $m = 1, 2, \ldots$,*

$$\pi_m\left(\mathcal{N}_{p_{m,0}}|y_k\right) \to 1, \tag{7.21}$$

as $k \to \infty$, almost surely for all $Y = \{y_k\}_{k=1}^{\infty}$.

Proof. Follows using the same ideas as the proof of Theorem 7. ■

As regards the choice of the quantities $c_{j,m}$, we simply extend the construction detailed in Section 7.2 by only letting $M \to \infty$, and with obvious replacement of the posterior means with those associated with the posterior Dirichlet process.

It is useful to remark that our theory with countably infinite number of limit points is readily applicable to situations where the number of limit points is finite but unknown. In such cases, only a finite number of the probabilities $\{p_{m,j};\ m = 1, 2, 3 \ldots\}$ will have posterior probabilities around positive quantities, while the rest will concentrate around zero. For known finite number of limit points, it is only required to specify G such that it gives positive mass to only a specific finite set.

7.4 Characterization of convergence and divergence with our approach on limit points

Note that for convergent series, $\pi_m(\mathcal{N}_1|y_k) \to 1$ as $k \to \infty$ for smaller values of m, while for divergent series, $\pi_m(\mathcal{N}_1|y_k) \to 1$ as $k \to \infty$ for much larger values of m. We formalize these statements below as the following theorems.

Theorem 14 *Let there be M number of possible limit points, where M may be infinite. Then $S_{1,\infty} = \infty$ if and only if*

$$\pi_{m,k}(\mathcal{N}_1|y_k) \to 1, \tag{7.22}$$

as $k \to \infty$ and $m \to M$.

Proof. Let $S_{1,\infty} = \infty$. Then there exists $k_0 \geq 1$ such that given any $\lambda > 0$, $S_{1,k} > \lambda$ for $k \geq k_0$. Then as $k \to \infty$,

$$\left(\frac{S_{1,k}}{1 + S_{1,k}}\right)^{\rho(\theta)} \to 1. \tag{7.23}$$

In other words, for any fixed $M\ (> 1)$, $y_k \to M$, almost surely, as $k \to \infty$. Hence, as $k \to \infty$ and $m \to M$, it easily follows using the same techniques as before, that (7.22) holds. Consequently, for infinite number of limit points (7.22) holds as $m \to \infty$.

Now assume that (7.22) holds. It then follows from the formula of the posterior mean that $y_k \to M$, almost surely, as $k \to \infty$, for fixed M. Hence, (7.23) holds, from which it follows that $S_{1,\infty} = \infty$. ■

Theorem 15 *$S_{1,\infty} < \infty$ if and only if for some finite $m_0 \geq 1$,*

$$\pi_{m_0,k}(\mathcal{N}_1|y_k) \to 1, \tag{7.24}$$

as $k \to \infty$.

Proof. Let $S_{1,\infty} < \infty$. Then as $k \to \infty$,

$$\left(\frac{S_{1,k}}{1 + S_{1,k}}\right)^{\rho(\theta)} \to c, \tag{7.25}$$

for some constant $0 \leq c < 1$. Hence, there exists some finite $m_0 \geq 1$ such that $y_k \to m_0$, almost surely, as $k \to \infty$. Using the same techniques as before, it is seen that that (7.24) holds.

Now assume that (7.24) holds. It then follows from the formula of the posterior mean, that $y_k \to m_0$, almost surely, as $k \to \infty$. Hence, (7.25) holds, from which it follows that $S_{1,\infty} < \infty$. ■

7.5 A rule of thumb for diagnosis of convergence, divergence and oscillations

Based on the above theorems we propose the following rule of thumb for detecting convergence and divergence when M is finite: if $\frac{m}{M} > 0.9$ such that $\pi_{m,k}(\mathcal{N}_1|y_k) \to 1$ as $m \to M$ and $k \to \infty$, then declare the series as divergent. If, on the other hand, $\frac{m}{M} \leq 0.9$ such that $\pi_{m,k}(\mathcal{N}_1|y_k) \to 1$, then declare the series as convergent.

If, instead, there exist m_ℓ; $\ell = 1, \ldots, L$ $(L > 1)$ such that $\pi_{m_\ell,k}\left(\mathcal{N}_{p_{m_\ell,0}}|y_k\right) \rightarrow 1$ as $k \rightarrow \infty$, where $0 < p_{m_\ell,0} < 1$ for $\ell = 1, \ldots, L$ and $\sum_{\ell=1}^{L} p_{m_\ell,0} = 1$, then say that the sequence $\{S_{1,n}\}_{n=1}^{\infty}$ has L limit points. Note that the value of $\frac{m}{M}$ is not important in this situation.

Next, we illustrate our theory on limit points with Example 5, arguably the most complex series in our set of examples (other than Riemann Hypothesis) and in Section 9, validate our result on Riemann Hypothesis with our Bayesian limit point theory.

8 Illustration of the Bayesian limit point theory with Example 5

Since there is at most one limit point in the cases that we investigated, application of our ideas to these cases must be able to re-confirm this. We consider the theory based on Dirichlet process developed in Section 7.3, assuming for the sake of illustrations that G is concentrated on M values, with $G\left(Y_j = m\right) = \frac{1}{M}$; $m = 1, 2, \ldots, M$. We set $M = 10$ for our experiments. Thus, by our rule of thumb, divergence is to be declared only if $\pi_{m=10,k}\left(\mathcal{N}_1|y_k\right) \rightarrow 1$, as $k \rightarrow \infty$.

As regards implementation, notice that here there is no scope for parallelization since at the j-th step only y_j is added to the existing $S_{1,j-1}$ to form $S_{1,j} = S_{1,j-1} + y_j$. As such, on our VMware, using a single processor, only about two seconds are required for 10^5 iterations associated with the series (5.13), for various values of a (> 0) and b (> 0).

8.1 Choice of $\rho(\theta)$ in $\left(\frac{S_{1,k}}{1+S_{1,k}}\right)^{\rho(\theta)}$

In our example, $\theta = (a, b)$. We choose, for $j \geq 1$,

$$\tilde{\rho}(\theta) = a - b + \epsilon, \tag{8.1}$$

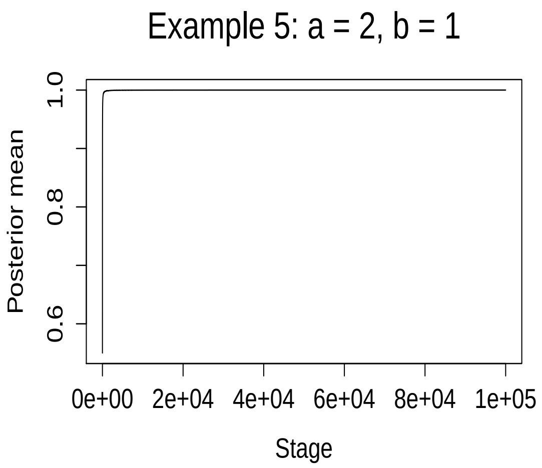
and set

$$\left(\frac{S_{1,j}}{1 + S_{1,j}}\right)^{\rho(\theta)} = \min\left\{1, \left(\frac{S_{1,j}}{1 + S_{1,j}}\right)^{\tilde{\rho}(\theta)}\right\} \tag{8.2}$$

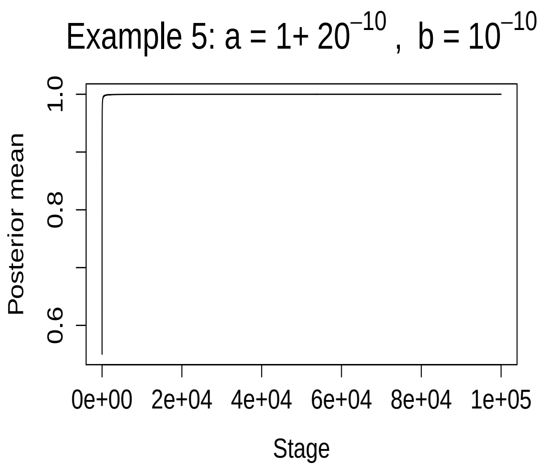
Recall that the series (5.13), defined for $a > 0$ and $b > 0$, converges for $a - b > 1$ and diverges for $a + b < 1$. In keeping with this result, (8.2) decreases as $(a - b)$ increases, so that the chance of correctly diagnosing convergence increases. Moreover, if both a and b are between 0 and 1 such that $a + b < 1$, then (8.2) tends to be inflated, thereby increasing the chance of correctly detecting divergence. The term ϵ in (8.2) prevents the power from becoming zero when $a = b$. It is important to note here that for $a + b = 1$ convergence or divergence is not guaranteed, but if $\epsilon = 0$ in (8.2), then $a = b$ would trivially indicate divergence, even if the series is actually convergent. A positive value of ϵ provides protection from such erroneous decision. Note that if $a < b - \epsilon$, the convergence criterion $a - b > 1$ is not met but the divergence criterion $a + b < 1$ may still be satisfied. Thus, for such instances, greater weight in favour of divergence is indicated. In our illustration, we set $\epsilon = 10^{-10}$.

8.2 Results

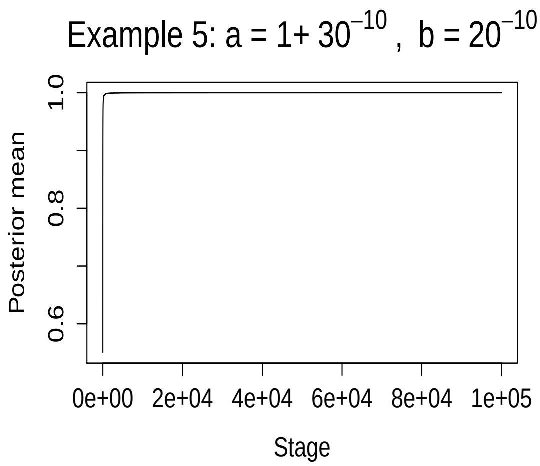
Figure 8.1 shows the results of our Bayesian analysis of the series (5.13) based on our Dirichlet process model. Based on the rule of thumb proposed in Section 7.5 all the results are in agreement with the results based on Figure 5.6.



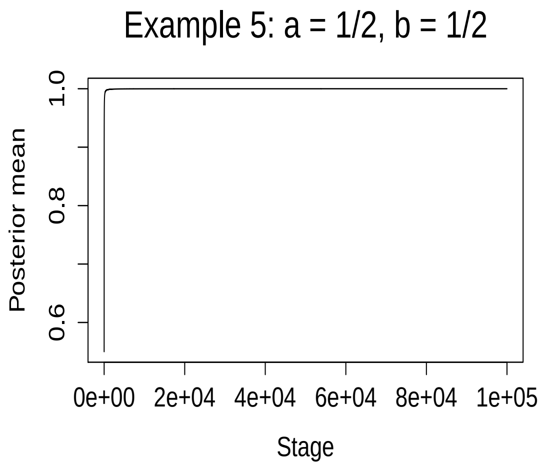
(a) Convergence: $a = 2, b = 1$. The posterior of $p_{1,k}$ converges to 1 as $k \to \infty$



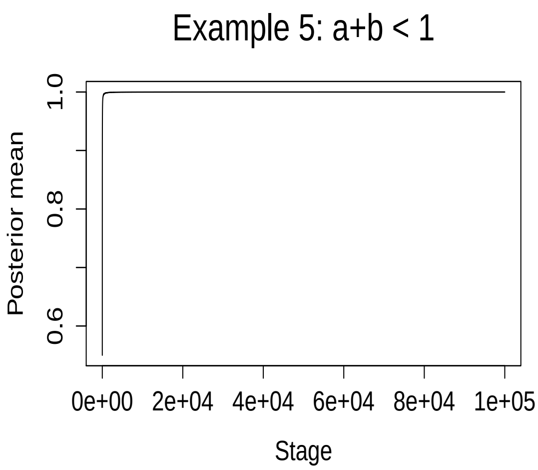
(b) Convergence: $a = 1 + 20^{-10}, b = 10^{-10}$. The posterior of $p_{2,k}$ converges to 1 as $k \to \infty$.



(c) Convergence: $a = 1 + 30^{-10}, b = 20^{-10}$. The posterior of $p_{2,k}$ converges to 1 as $k \to \infty$.



(d) Divergence: $a = 1/2, b = 1/2$. The posterior of $p_{10,k}$ converges to 1 as $k \to \infty$.



(e) Divergence: $a = \frac{1}{2}\left(1 - 10^{-11}\right), b = \frac{1}{2}\left(1 - 10^{-11}\right)$. The posterior of $p_{10,k}$ converges to 1 as $k \to \infty$.

Figure 8.1: Illustration of the Dirichlet process based theory with Example 5: For $(a = 2, b = 1)$ in the series (5.13), $\frac{m}{M} = \frac{1}{10} < 0.9$, indicating convergence, for $(a = 1 + 20^{-10}, b = 10^{-10})$, $\frac{m}{M} = \frac{2}{10} < 0.9$, indicating convergence, for $(a = 1 + 30^{-10}, b = 20^{-10})$, $\frac{m}{M} = \frac{2}{10} < 0.9$, indicating convergence, for $(a = 1/2, b = 1/2)$, $\frac{m}{M} = \frac{10}{10} > 0.9$, indicating divergence, and for $\left(a = \frac{1}{2}\left(1 - 10^{-11}\right), b = \frac{1}{2}\left(1 - 10^{-11}\right)\right)$, $\frac{m}{M} = \frac{10}{10} > 0.9$, indicating divergence.

9 Application of the Bayesian multiple limit points theory to Riemann Hypothesis

To strengthen our result on Riemann Hypothesis presented in Section 6 we consider application of our Bayesian multiple limit points theory to Riemann Hypothesis.

9.1 Choice of $\rho(\theta)$ in $\left(\frac{|S_{1,k}|}{1+|S_{1,k}|}\right)^{\rho(\theta)}$

For Riemann Hypothesis, $\theta = a$; we choose, for $j \geq 1$,

$$\tilde{\rho}(\theta) = a^6. \tag{9.1}$$

The reason for such choice with a relatively large power is to allow discrimination between $\left(\frac{|S_{1,k}|}{1+|S_{1,k}|}\right)^{\rho(\theta)}$ for close values of a. However, substantially large powers of a are not appropriate because that would make the aforementioned term too small to enable detection of divergence. In fact, we have chosen the power after much experimentation. Implementation of our methods takes about 25 seconds on our VMWare, with 10^6 iterations.

9.2 Results

The results of application of our ideas on multiple limit points are depicted in Figures 9.1, 9.2 and 9.3. The values of m/M and the thumb rule proposed in Section 7.5 show that all the results are consistent with those obtained in Section 6. There seems to be a slight discrepancy only regarding the location of the change point of convergence. In this case, unlike $a = 0.72$ as obtained in Section 6, we obtained $a = 0.74$ as the change point (see panel (b) of Figure 9.2). In fact, it turned out that $\frac{m}{M} = 1$ for all the values of $a \in (0.7, 0.74)$ that we experimented with.

This (perhaps) negligible difference notwithstanding, both of our methods are remarkably in agreement with each other, emphasizing our point that Riemann Hypothesis can not be completely supported.

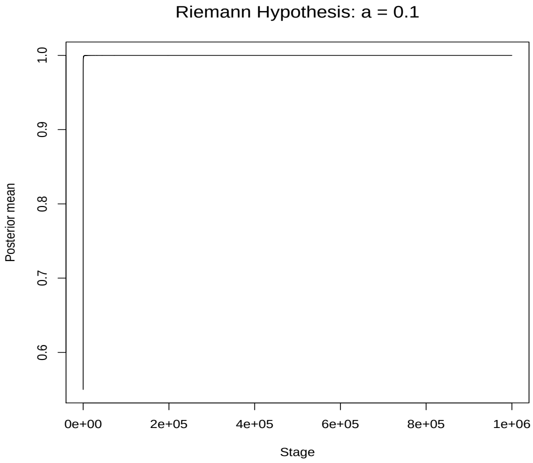
10 Summary and conclusion

In this paper, we proposed and developed a novel Bayesian methodology for assessment of convergence of infinite series; we further extended the theory to infinite series with possibilities of multiple or even infinite number of limit points. Our developments do not require any restrictive assumption, not even independence of the elements X_i of the infinite series.

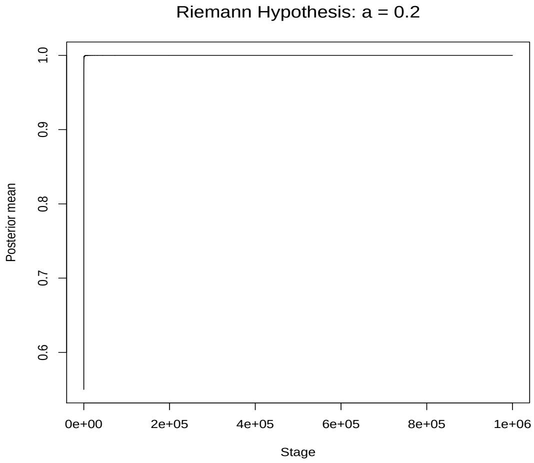
We demonstrated the reliability and efficiency of our methods with varieties of examples, the most important one being associated with Riemann Hypothesis.

Both methods proposed in this paper, namely the convergence assessment method and the multiple limit points method are almost completely in agreement that the Riemann Hypothesis can not be completely supported. Indeed, both the methods agree that there exists some a^* in the neighborhood of 0.7 such that the infinite series based on the Möbius function diverges for $a < a^*$ and converges for $a \geq a^*$. The results that we obtained by our Bayesian analyses are also supported by informal plots of the partial sums depicted in Figure 6.1.

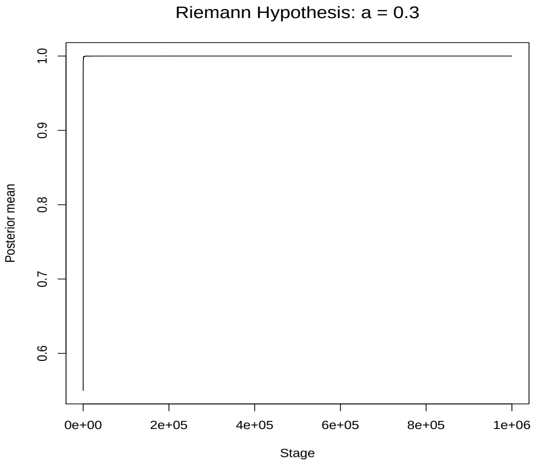
The theory that we developed readily applies to random series; we shall carry out a detailed investigation including comparisons with existing theories on random infinite series. We then intend to extend these works to complex infinite series, both deterministic and random.



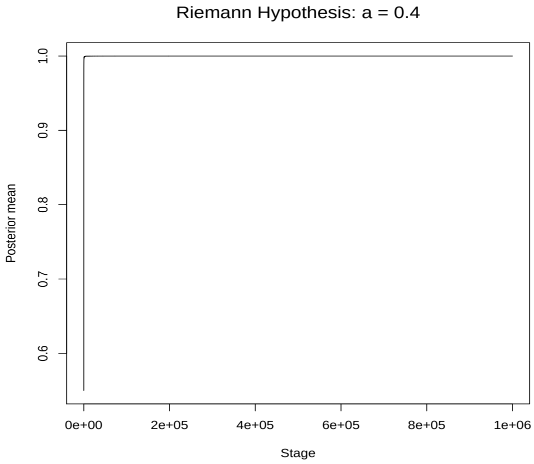
(a) Divergence: $a = 0.1$, $\frac{m}{M} = \frac{10}{10}$.



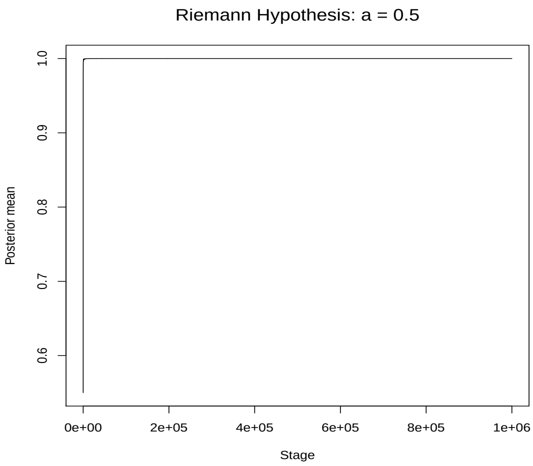
(b) Divergence: $a = 0.2$, $\frac{m}{M} = \frac{10}{10}$.



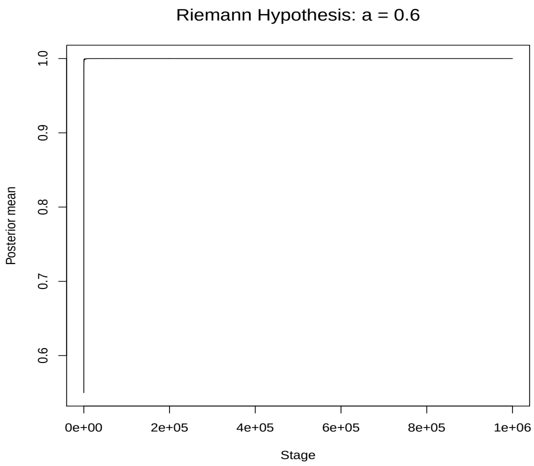
(c) Divergence: $a = 0.3$, $\frac{m}{M} = \frac{10}{10}$.



(d) Divergence: $a = 0.4$, $\frac{m}{M} = \frac{10}{10}$.

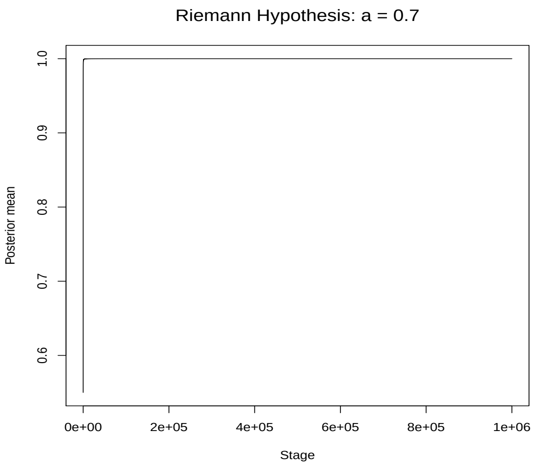


(e) Divergence: $a = 0.5$, $\frac{m}{M} = \frac{10}{10}$.

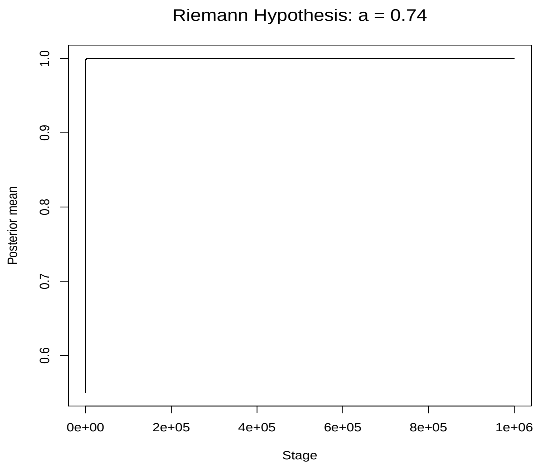


(f) Divergence: $a = 0.6$, $\frac{m}{M} = \frac{10}{10}$.

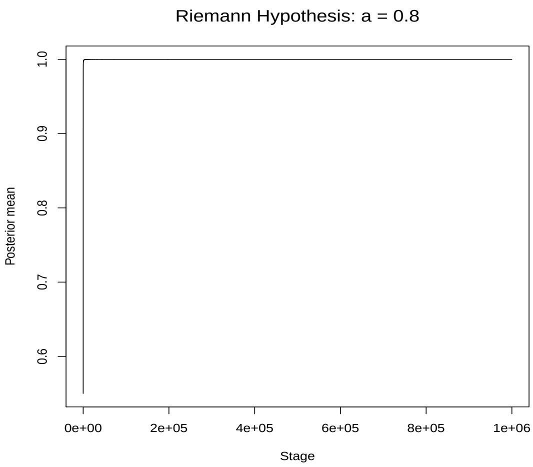
Figure 9.1: Riemann Hypothesis based on Bayesian multiple limit points theory: Divergence for $a = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$.



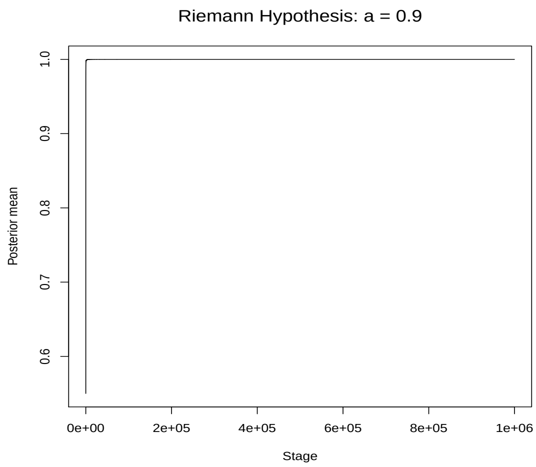
(a) Divergence: $a = 0.7$, $\frac{m}{M} = \frac{10}{10}$.



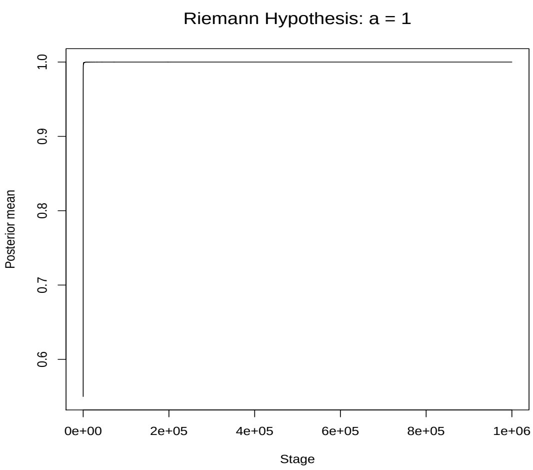
(b) Convergence: $a = 0.74$, $\frac{m}{M} = \frac{9}{10}$.



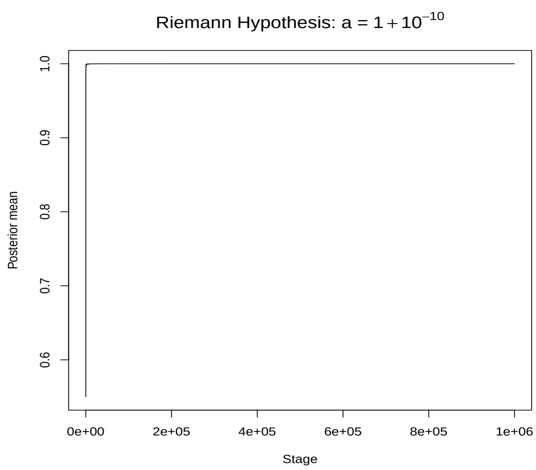
(c) Convergence: $a = 0.8$, $\frac{m}{M} = \frac{9}{10}$.



(d) Convergence: $a = 0.9$, $\frac{m}{M} = \frac{7}{10}$.

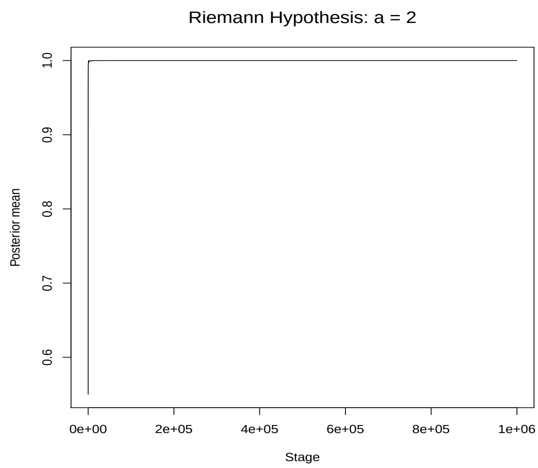


(e) Convergence: $a = 1.0$, $\frac{m}{M} = \frac{5}{10}$.

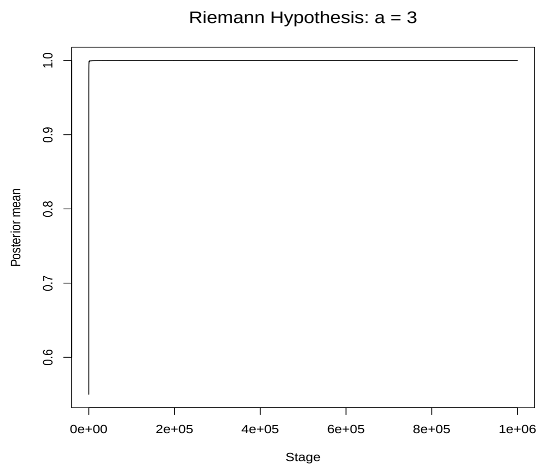


(f) Convergence: $a = 1 + 10^{-10}$, $\frac{m}{M} = \frac{5}{10}$.

Figure 9.2: Riemann Hypothesis based on Bayesian multiple limit points theory: Divergence for $a = 0.7$ but convergence for $a = 0.74, 0.8, 0.9, 1, 1 + 10^{-10}$.



(a) Convergence: $a = 2$, $\frac{m}{M} = \frac{1}{10}$.



(b) Convergence: $a = 3$, $\frac{m}{M} = \frac{1}{10}$.

Figure 9.3: Riemann Hypothesis based on Bayesian multiple limit points theory: Convergence for $a = 2$, 3.

Appendix

A Proof of Lemma 8

Since each term of the series (1.1) is decreasing in a, it is clear that $S^{a,b}_{j,n}$ is decreasing in a. We need to show that $S^{a,b}_{j,n}$ is decreasing in b.

Let, for $i \geq 3$,

$$g(i) = \left(1 - \frac{\log i}{i} - \frac{\log\log i}{i}\left\{\cos^2\left(\frac{1}{i}\right)\right\}\left(a + (-1)^i b\right)\right)^i. \tag{A.1}$$

Since n is even, observe that all our partial sums of the form $S^{a,b}_{j,n}$ for $j \geq 3$ admit the form

$$S^{a,b}_{j,n} = \sum_{i=r}^{r+n-1} g(i), \tag{A.2}$$

where $r = 3 + n(j-1)$, which is clearly odd. Now,

$$\sum_{i=r}^{r+n-1} g(i) = \{g(r) + g(r+1)\} + \{g(r+2) + g(r+3)\} + \cdots + \{g(r+n-2) + g(r+n-1)\}, \tag{A.3}$$

where the sums of the consecutive terms within the parentheses have the form

$$\begin{aligned}
&g(r+\ell) + g(r+\ell+1)\\
&= \left(1 - \frac{\log(r+\ell)}{r+\ell} - \frac{\log\log(r+\ell)}{r+\ell}\left\{\cos^2\left(\frac{1}{r+\ell}\right)\right\}\left(a + (-1)^{(r+\ell)}b\right)\right)^{(r+\ell)}\\
&\quad + \left(1 - \frac{\log(r+\ell+1)}{r+\ell+1} - \frac{\log\log(r+\ell+1)}{r+\ell+1}\left\{\cos^2\left(\frac{1}{r+\ell+1}\right)\right\}\left(a + (-1)^{(r+\ell+1)}b\right)\right)^{(r+\ell+1)}.
\end{aligned} \tag{A.4}$$

Since r is odd, and since the terms are represented pairwise in (A.3) it follows that in (A.4), $r+\ell$ is odd and $r+\ell+1$ is even. That is, in (A.4), $a + (-1)^{(r+\ell)}b = a - b$ and $a + (-1)^{(r+\ell+1)}b = a + b$.

Since $\cos^2(\theta)$ is decreasing on $\left[0, \frac{\pi}{2}\right]$, and since $\frac{1}{i} \leq \frac{\pi}{2}$ for $i \geq 3$, it follows that $\cos^2\left(\frac{1}{i}\right)$ is increasing in i. Moreover, $\frac{\log\log i}{i}$ decreases in i at a rate faster than $\cos^2\left(\frac{1}{i}\right)$ increases, so that $\frac{\log\log i}{i} \times \cos^2\left(\frac{1}{i}\right)$ decreases in i. It follows that

$$\frac{\log\log(r+\ell)}{r+\ell}\cos^2\left(\frac{1}{r+\ell}\right) > \frac{\log\log(r+\ell+1)}{r+\ell+1}\cos^2\left(\frac{1}{r+\ell+1}\right). \tag{A.5}$$

Note that in $g(r+\ell)+g(r+\ell+1)$, $\frac{\log\log(r+\ell)}{r+\ell}\cos^2\left(\frac{1}{r+\ell}\right)$ is associated with b while $\frac{\log\log(r+\ell+1)}{r+\ell+1}\cos^2\left(\frac{1}{r+\ell+1}\right)$ involves $-b$. Hence, increasing b increases $g(r+\ell)$ but decreases $g(r+\ell)$, and because of (A.5), $g(r+\ell)+g(r+\ell+1)$ increases in b. This ensures $\sum_{i=r}^{r+n-1} g(i)$ given by (A.3), is increasing in b. In other words, partial sums of the form (A.2) are increasing in b, proving Lemma 8 when n is even.

B Proof of Lemma 10

That $S_{j,n}^{a,b}$ is decreasing in a follows trivially since each term of (5.13) is decreasing in a. We need to show that $S_{j,n}^{a,b}$ is increasing in b.

Let, for $i \geq 5$,

$$g(i) = \left(1 - \left(\frac{\log(i)}{i}\right)\left(a\left(1+\sin^2\left(\sqrt{\left(\frac{\log(\log(i))}{\log(i)}\right)}\right)\right) + b\sin\left(\frac{i\pi}{4}\right)\right)\right)^i. \tag{B.1}$$

Now note that, with $r = 5 + n(j-1)$,

$$\begin{aligned}\sum_{i=r}^{r+n-1} g(i) &= \sum_{m=1}^{\frac{n}{4}} Z_{r,m} \\ &= \{Z_{r,1} + Z_{r,2}\} + \{Z_{r,3} + Z_{r,4}\} + \cdots + \left\{Z_{r,\frac{n}{4}-1} + Z_{r,\frac{n}{4}}\right\},\end{aligned} \tag{B.2}$$

where

$$Z_{r,m} = \sum_{\ell=5+4(m-1)}^{5+4(m-1)+3} g(r+\ell). \tag{B.3}$$

Now, for any $\ell \geq 1$, observe that in $\{Z_{r,\ell} + Z_{r,\ell+1}\}$, the term $Z_{r,\ell}$ consists of only negative signs of the sine-values, while in $Z_{r,\ell+1}$ the corresponding signs are positive although the magnitudes are the same. Since $\log(i)/i$ is decreasing in i, it follows that $\{Z_{r,\ell} + Z_{r,\ell+1}\}$ is increasing in b for $\ell \geq 1$. Hence, it follows that (B.2), and $S_{j,n}^{a,b}$, defined by (5.14), are increasing in b for $j \geq 1$ and n, a multiple of 4, proving Lemma 10.

References

Alekseyev, M. A. (2011). On Convergence of the Flint Hills Series. Available at "http://arxiv.org/pdf/1104.5100v1.pdf".

Borwein, P., Choi, S., Rooney, B., and Weirathmueller, A. (2006). *The Riemann Hypothesis: For the Aficionado and Virtuoso Alike*. Springer, New York.

Bourchtein, L., Bourchtein, A., Nornberg, G., and Venzke, C. (2011). A Hierarchy of the Convergence Tests for Numerical Series Based on Kummer's Theorem. *Bulletin of the Paranaense Society of Mathematics*, **29**, 83–107.

Bourchtein, L., Bourchtein, A., Nornberg, G., and Venzke, C. (2012). A Hierarchy of the Convergence Tests Related to Cauchy's Test. *International Journal of Mathematical Analysis*, **6**, 1847–1869.

Bromwich, T. J. I. (2005). An introduction to the theory of infinite series. AMS, Providence.

Derbyshire, J. (2004). *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*. Penguin, New York.

Fichtenholz, G. M. (1970). *Infinite Series: Rudiments*. Gordon and Breach Publishing, New York.

Horsley, S. (1772). ΚΟΣ ΚΙΝΟΝ ΕΡΑΤΟΣ Θ ΕΝΟΣ. or, The Sieve of Eratosthenes. Being an Account of his Method of Finding all the Prime Numbers by the Rev. Samuel Horsley, F. R. S. *Philosophical Transactions (1683–1775)*, **62**, 327–347.

Ilyin, V. A. and Poznyak, E. G. (1982). *Fundamentals of Mathematical Analysis, Vol.1*. Mir Publishers, Moscow.

Knopp, K. (1990). *Theory and Application of Infinite Series*. Dover Publishers, New York.

Liflyand, E., Tikhonov, S., and Zeltser, M. (2011). Extending Tests for Convergence of Number Series. *Journal of Mathematical Analysis and Applications*, **377**, 194–206.

Lioen, W. M. and van de Lune, J. (1994). Systematic Computations on Mertens' Conjecture and Dirichlet's Divisor Problem by Vectorized Sieving. In K. Apt, L. Schrijver, and N. Temme, editors, *From Universal Morphisms to Megabytes: a Baayen Space Odyssey*, pages 421–432, CWI, Amsterdam.

Pickover, C. A. (2002). *The Mathematics of Oz: Mental Gymnastics from Beyond the Edge*. Cambridge University Press, U. K.

Riemann, B. (1859). Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse. Monatsberichte der Berliner Akademie. In Gesammelte Werke, Teubner, Leipzig (1892), Reprinted by Dover, New York (1953). Original manuscript (with English translation). Reprinted in (Borwein et al. 2008) and (Edwards 1974).

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, New York.

Schervish, M. J. (1995). *Theory of Statistics*. Springer-Verlag, New York.

Spivak, M. (1994). Calculus, Publish or Perish.