Modify Thompson's approach and applying ROC curve to measuring similarity in Bayesian historical borrowing

Yu (Bel) Zhang

December 8, 2023

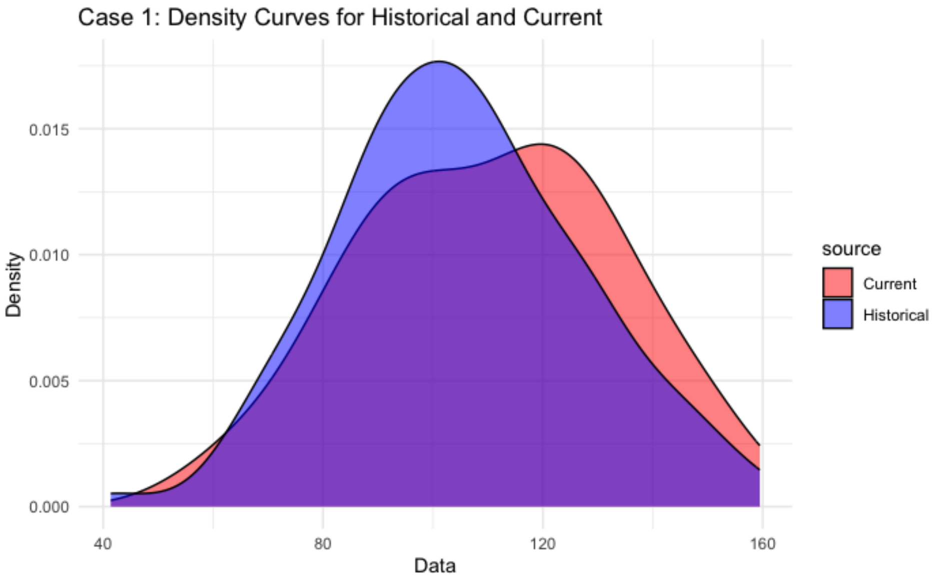
1 Thompson et al. approach.

1.1 The potential issue of Laura's method.

When using Laura's method ([Thompson et al., 2021]), her equation (6) (following equation) from the paper restricts the borrowing only occur when observed difference between the sample means, $\bar{D}_0$ and $\bar{D}_1$, falls within $(-\delta, \delta)$. θ_0 and θ are the means come from the posterior distribution given D_0 and D_1 alone, respectively. Here, D_0 is the historical data and D_1 is an interim subset from the current data,

$$\alpha_0(D_0, D_1) = \begin{cases} P(|\theta - \theta_0| < \delta | D_0, D_1), & \text{if} |\bar{D}_0 - \bar{D}_1| < \delta \\ 0 & o.w \end{cases} \tag{6}$$

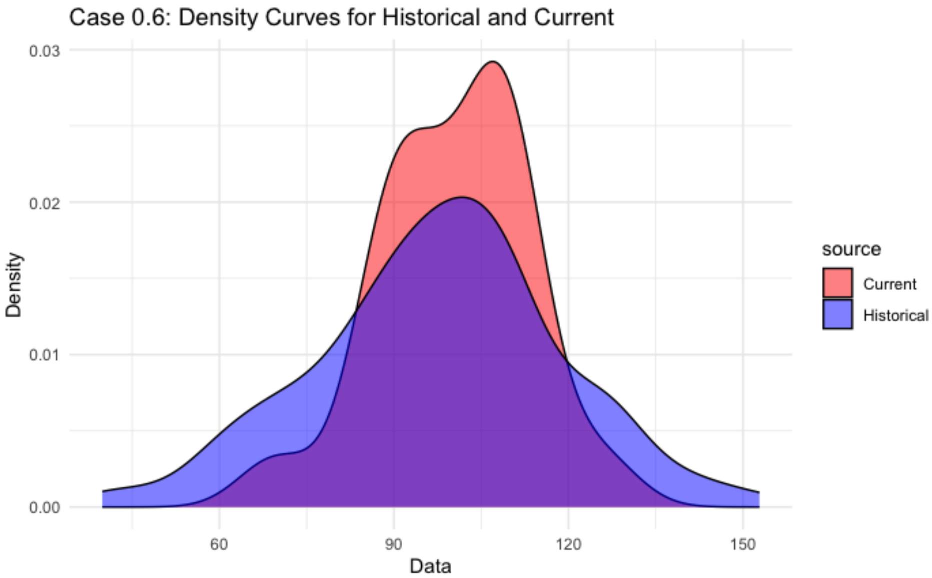
For example, if the historical data D_0 comes from $N(100, 20^2)$ with size $n_0 = 100$, current data D_1 comes from $N(100, (r_\sigma \cdot 20)^2)$ with size $n_1 = 80$. The r_σ is the ratio of current standard deviation to historical standard deviation, say σ_1/σ_0. We take 25% from D_1 as our interim subset. Assume $r_\sigma = 1$, the two samples density plots look like in 1, the observed difference between two sample means is 5.17, the pre-specified $\delta = 8$. So, in this case, we should borrow base on Laura's method. Directly applying equation (6), given $N(100, 100)$ as the mean prior and $IG(0.001, 0.001)$ as the prior for standard deviation, the power parameter here is $\alpha_0 = 0.9773$. If we use the two suggestions that Laura mentioned in her paper to cap the power parameter value, one is $\alpha_{max} = 1 - \delta/|\bar{D}_0|$, we have $\alpha_{max1} = 0.9237$. The other one is $min(1 - \delta/|\bar{D}_0|, n_1/n_0)$, we get $\alpha_{max2} = 0.8$.



(a) Sample distribution of D_0 and D_1, $r_\sigma = 1$.

Figure 1

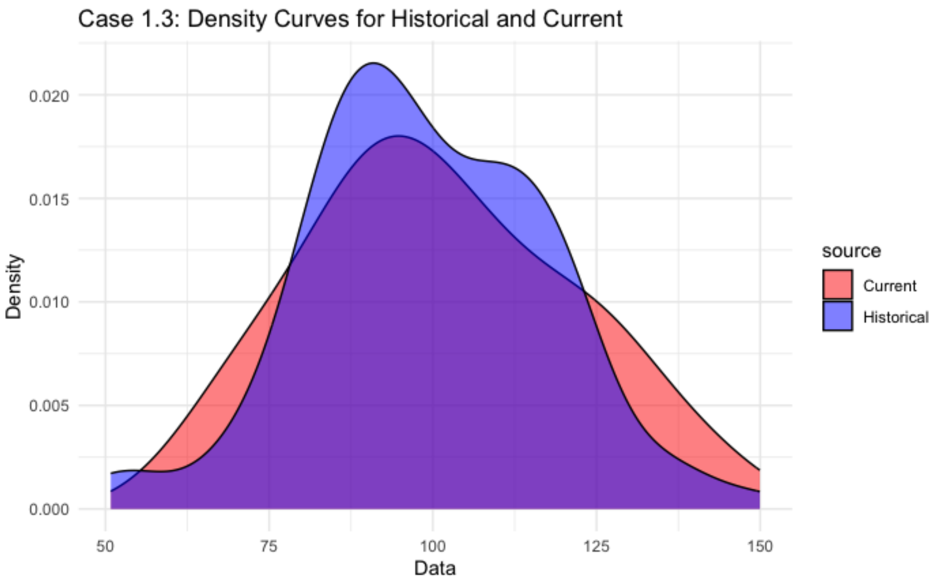
Now, suppose the current data still comes from a normal distribution with mean 100, but with a smaller standard deviation, say 12. The $r_\sigma = 0.6$ in this case. The density plots is shown in 10. $|\bar{D}_0 - \bar{D}_1| = 2.72 < 8$, borrowing happens. $\alpha_0 = 0.9917$, $\alpha_{max1} = 0.9180$ and $\alpha_{max2} = 0.8$. Note that, even the mean of current data is the same as the historical mean , with such a different standard deviation, we doubt the two data are similar. However, Laura's approach failed to capture the influence of changes in deviation on borrowing.



(a) Sample distribution of D_0 and D_1, $r_\sigma = 0.6$.

Figure 2

Moreover, if we keep other setup the same but change the $r_\sigma = 1.3$ (the plot is given in 3), the $|\bar{D}_0 - \bar{D}_1| = 1.78 < 8$. And $\alpha_0 = 0.9794$, $\alpha_{max1} = 0.9192$ and $\alpha_{max2} = 0.8$.



(a) Sample distribution of D_0 and D_1, $r_\sigma = 1.3$.

Figure 3

A summary of these three cases is given in Table 2.

Scenario	$\lvert\bar{D}_0 - \bar{D}_1\rvert$	α_0	$\alpha_{max1} = 1 - \delta/\lvert\bar{D}_0\rvert$	$\alpha_{max2} = min(1 - \delta/\lvert\bar{D}_0\rvert, n1/n0)$
$r_{sd} = 1$	5.17	0.9773	0.9237	0.8
$r_{sd} = 0.6$	2.72	0.9917	0.9180	0.8
$r_{sd} = 1.3$	1.78	0.9794	0.9192	0.8

Table 1: One-off example using Laura's method with Dr. Tubb's comment cases. All cases are set with $n_0 = 100$, $\sigma_0 = 20$, $n_1 = 80$ and $\delta = 8$. $\mu_0 = 100$, $\mu_1 = 100$, 25% percentage of current data is used to construct the α_0.

Note: Based on the cases showed above, when $\mu_0 = \mu_1$, the observed mean difference $|\bar{D}_0 - \bar{D}_1|$ is smaller when two population standard deviations are apart, comparing to $\sigma_0 \approx \sigma_1$. This makes sense because one data set is more condensed around the mean than the other, regardless of whether it is current or historical. With my understanding, this is the reason why the borrowing still occurs even there exists heterogeneity between historical and current data, by using Laura's method.

2 Modified method.

2.1 Restriction on posterior predictive probability of observe mean difference.

As shown previously, the method by Thompson et al. failed to detect the heterogeneity. The motivation of this research is not to answer the question, 'Do the data come from the assumed model?' (to which the answer is almost always no), but to quantify the discrepancies between data and model, and assess whether they could have arisen by chance, under the model's own assumptions [Gelman and Carlin, 2013]. We proposed a modified method that may catch the heterogeneity better (or not). The main idea of this new method is to restrict on the probability of the posterior predictive distribution (PPD) the future observed mean difference between historical and current data. The equation (6) now will update to the following,

$$\alpha_0(D_0, D_1) = \begin{cases} P(|\theta - \theta_0| < \delta|D_0, D_1), & \text{if} P(|\bar{D}_0^* - \bar{D}_1^*| < \delta|D_0, D_1) > \phi \\ 0 & o.w \end{cases} \tag{6*}$$

Here, the $\bar{D}_0^*$ and $\bar{D}_1^*$ are the new observations from the historical and current data, based on the posterior predictive distribution, respectively.

The posterior predictive distribution is obtained by averaging the predictive distribution over all possible values of the parameters according to the posterior distribution.

$$P(\text{new data}|\text{data}) = \int P(\text{new data}|\theta) \times P(\theta|\text{data})\, d\theta$$

Since the posterior predictive checking is to help assess the fit of one statistical model to observed data, in other word, it looks for systematic discrepancies between real and simulated data, we thought restricting on PPD may better catch the heteroscedasticity. (Similar to the predictive confidence interval under frequentist paradigm, the PPD combines the uncertainty from both the parameter estimates and from the sampling errors for the prediction.) In my opinion, posterior predictive checks are helpful in assessing if your model gives you "valid" predictions about the reality, do they fit the observed data or not. (But I may understand it wrongly.)

2.1.1 Examples.

The following examples are based on the same pair of historical and current dataset, which are generated with same mean, $\mu_0 = \mu_1 = 100$, size for historical is $n_0 = 100$ and for current $n_1 = 80$. We took 25% of current data D_1 to do the analysis. The δ value is set to be 8 just for illustration purpose. We simulated 30000 iterations and the summary of each case (with different current variances) is provided below.

2.1.2 Restriction on posterior predictive observe mean difference according to (6*).

1. Case 1, $rsd = 0.6$, i.e $\sigma_0 = 20$, $\sigma_1 = 12$. By Thompson's method, the $|\bar{D}_0 - \bar{D}_1| < 8$, we should borrow from D_0.

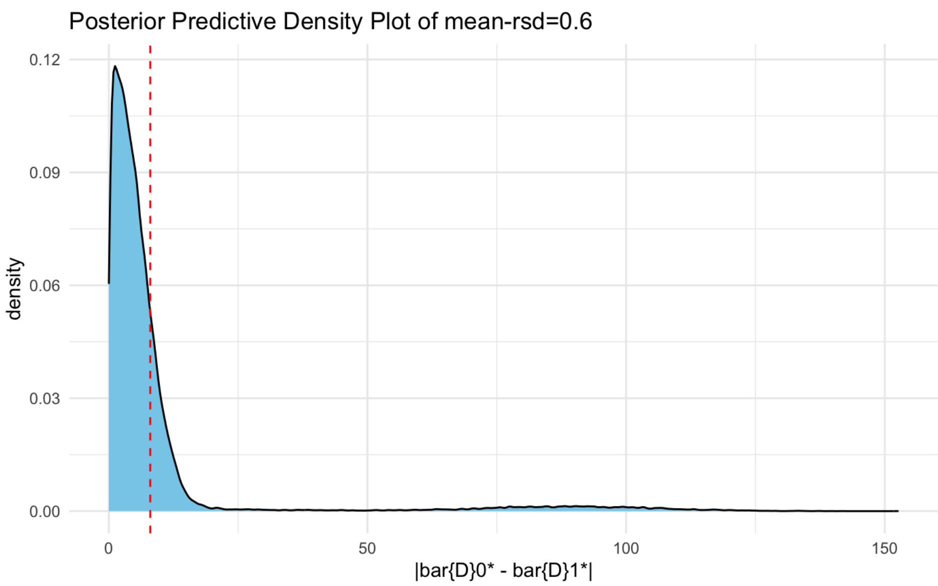


Figure 4: case1 ppd of $P(|\bar{D}_0^* - \bar{D}_1^*| < \delta | D_0, D_1)$.

2. Case 2, $rsd = 1$, i.e $\sigma_0 = 20$, $\sigma_1 = 20$. By Thompson's method, the $|\bar{D}_0 - \bar{D}_1| < 8$, we should borrow from D_0.

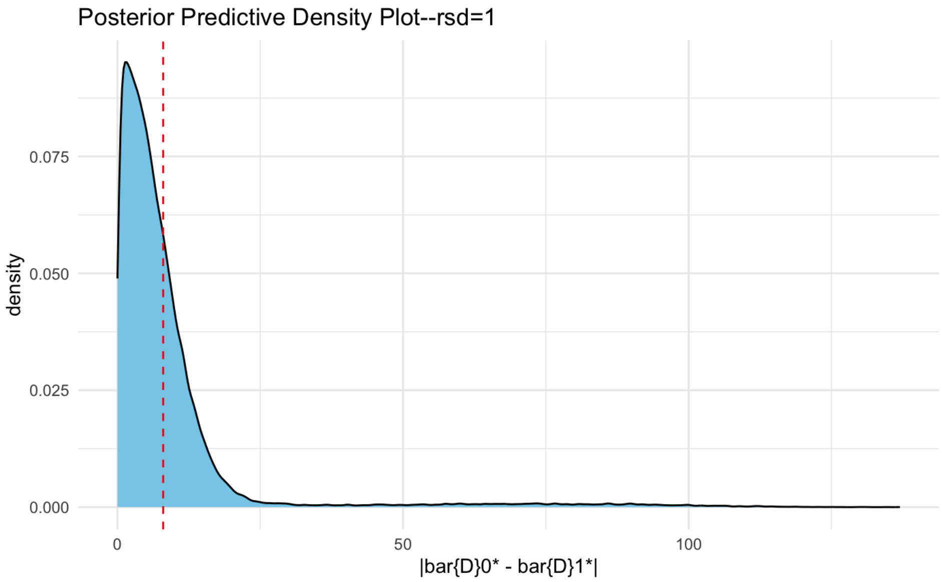


Figure 5: case1 ppd of $P(|\bar{D}_0^* - \bar{D}_1^*| < \delta | D_0, D_1)$.

3. Case 3, $rsd = 1.5$, i.e $\sigma_0 = 20$, $\sigma_1 = 30$. By Thompson's method, the $|\bar{D}_0 - \bar{D}_1| < 8$, we should borrow from D_0.

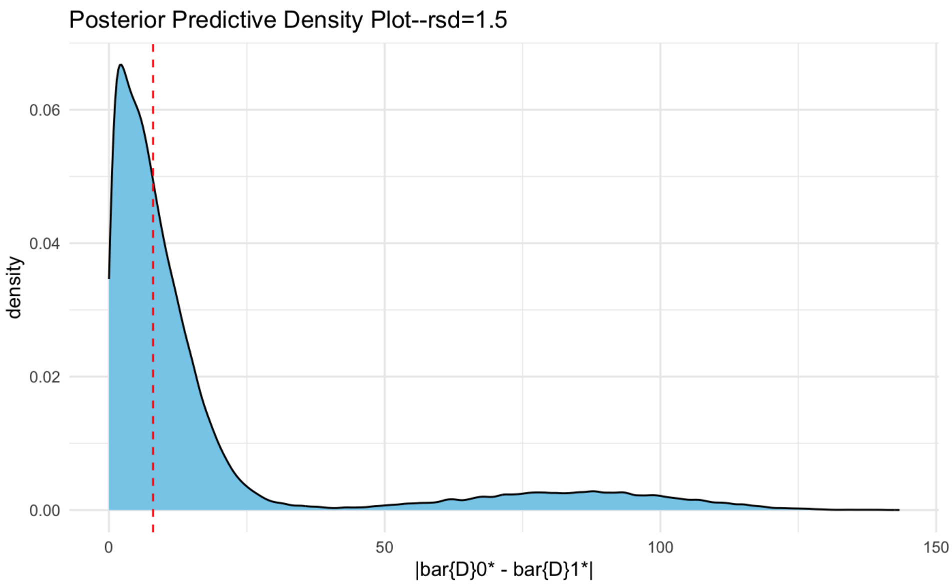


Figure 6: case1 ppd of $P(|\bar{D}_0^* - \bar{D}_1^*| < \delta | D_0, D_1)$.

The summary table is given in Table 2

Scenario	$\lvert\bar{D}_0 - \bar{D}_1\rvert$	$< \delta$?/Borrow or not? (Thompson)	$P(\lvert\bar{D}_0^* - \bar{D}_1^*\rvert < \delta \mid D_0, D_1)$
$r_{sd} = 0.6$	3.45	Yes	0.77
$r_{sd} = 1$	3.52	Yes	0.667
$r_{sd} = 1.5$	3.60	Yes	0.498

Table 2: Result summary of posterior predictive probability of future observed mean difference between historical and current sample means. All cases are set with $n_0 = 100$, $\sigma_0 = 20$, $n_1 = 80$ and $\delta = 8$. $\mu_0 = 100$, $\mu_1 = 100$, 25% percentage of current data is used to construct the α_0.

According to the results presented in Table 2, as the value of r_{sd} increases, the p-value for $|\bar{D}_0^* - \bar{D}_1^*| < \delta$ decreases. It appears that the modified restriction in (6^*) has not resolved the issue observed in Thompson's approach. We expected to see the larger probability occurs with r_{sd} closer to 1. (Note that this result is just from a one-off example, not a simulation, the result may not be robust. However, the decreasing trend is consistent with the found we mentioned at the end of the section 1, that is when the $r_s d << 1$, the current data located closer to the μ_1, that may be the reason caused the $P(|\bar{D}_0^* - \bar{D}_1^*| < \delta | D_0, D_1)$ is larger than other cases with larger value of r_{sd}.

2.2 Restriction on PPD of observe mean difference and observe standard deviation.

In order to catch the heteroscedasticity, we may add a restriction on the probability of posterior predictive distribution of observe sample standard deviation. Here, the observe standard deviation ratio is using the current's to the historical's.

$$\alpha_0(D_0, D_1) = \begin{cases} P(|\theta - \theta_0| < \delta | D_0, D_1), & \text{if } P(|\bar{D}_0^* - \bar{D}_1^*| < \delta | D_0, D_1) > \phi \\ & \text{and } P(|s_1^*/s_0^*| < \xi | D_0, D_1) > \eta \\ \\ 0 & o.w \end{cases} \tag{6**}$$

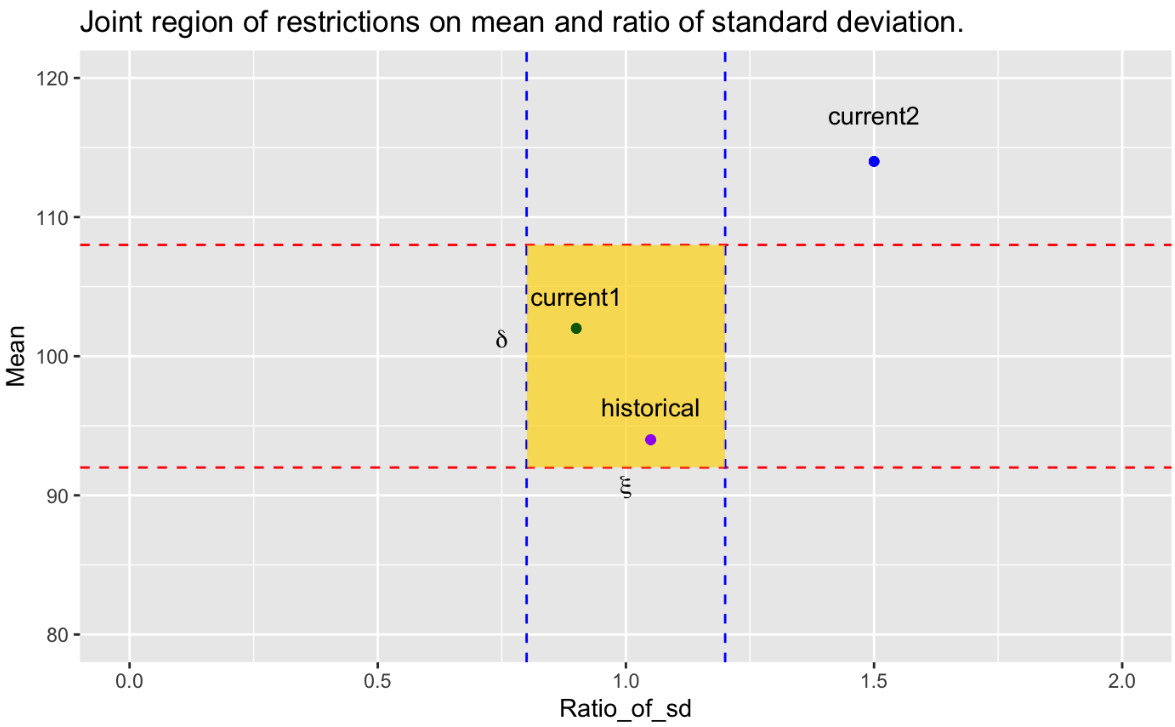
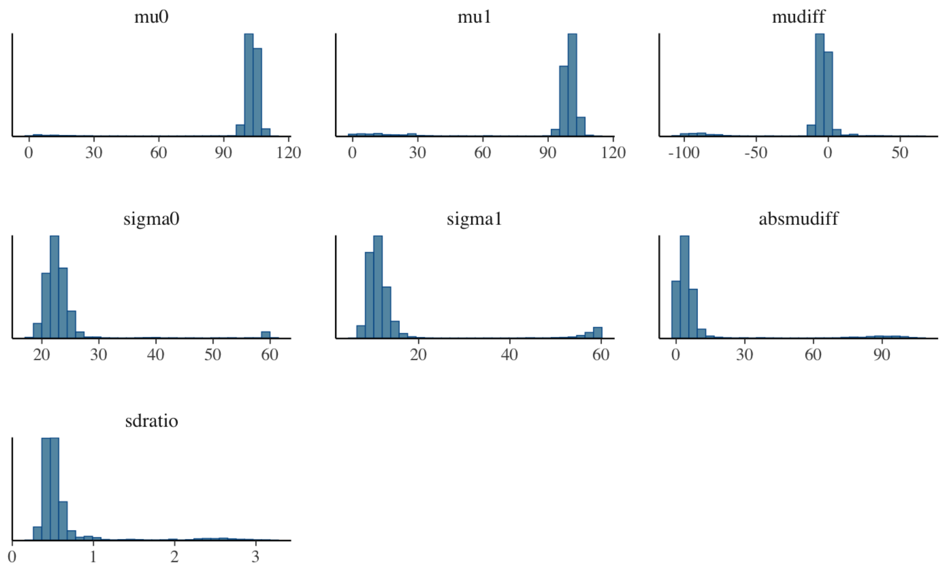


Figure 7: Illustration graph of equation (6**).

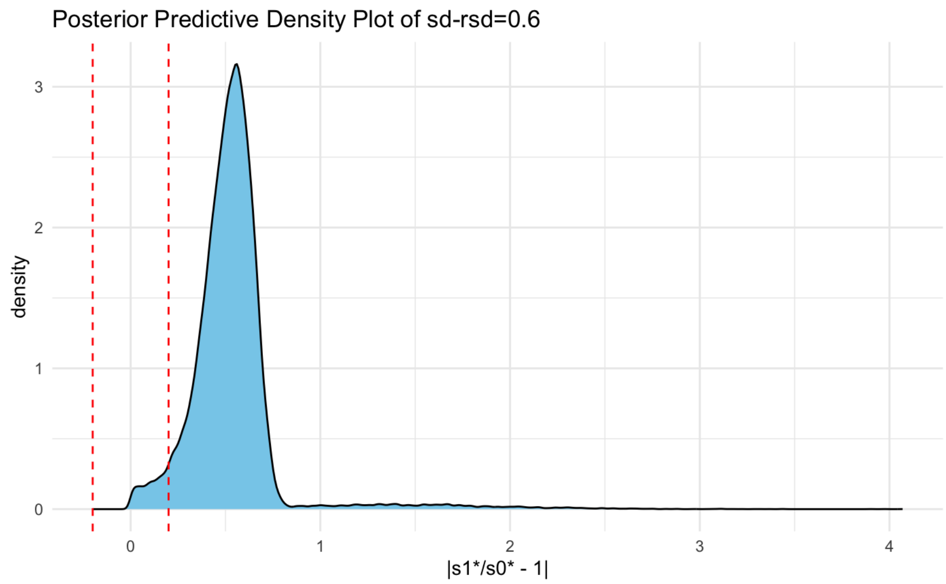
As shown in Figure 7, once the δ and ξ values are specified, if the current data and historical data fall within the yellow region, we claim the patient-level exchangeability has been obtained. Otherwise, there should not use borrowing because of the violated of exchangeability assumption.

Now, we will use the same setting as previous three cases, but add one more step to monitor the behavior of the PPD of $P|s_1^*/s_0^* - 1| < \xi$.

1. Case1.



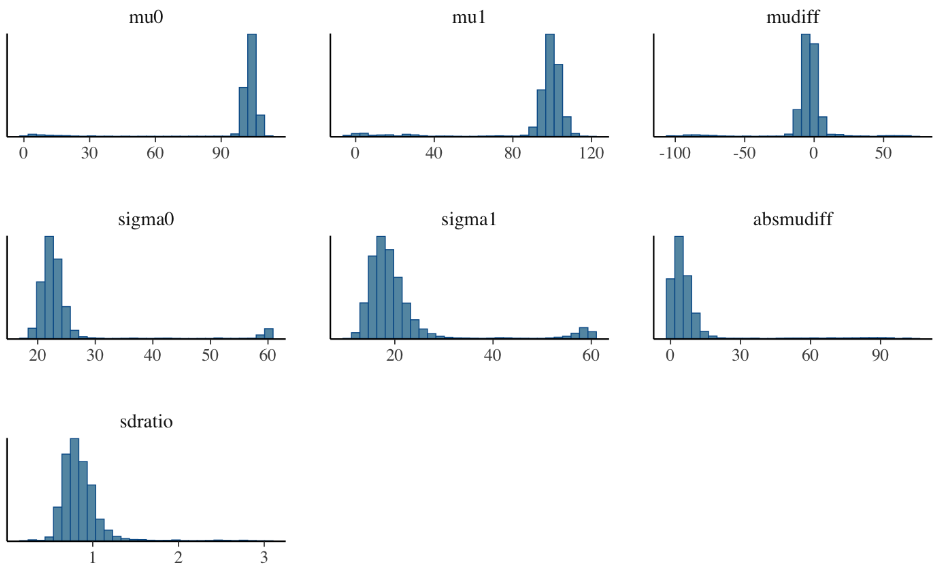
(a) case1 posterior plots.



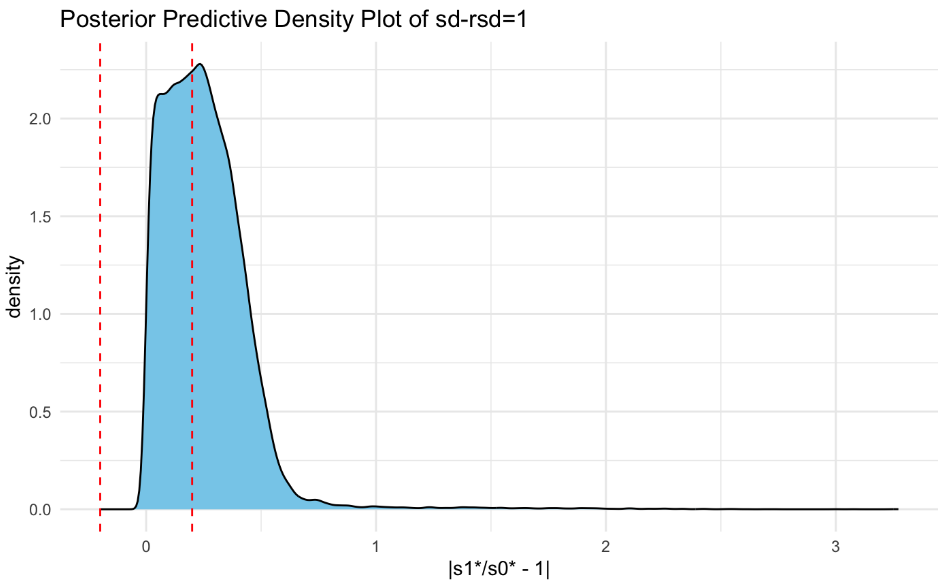
(b) case1 ppd of $P(|s_1^*/s_0^*| < \xi | D_0, D_1)$.

Figure 8

2. case2.



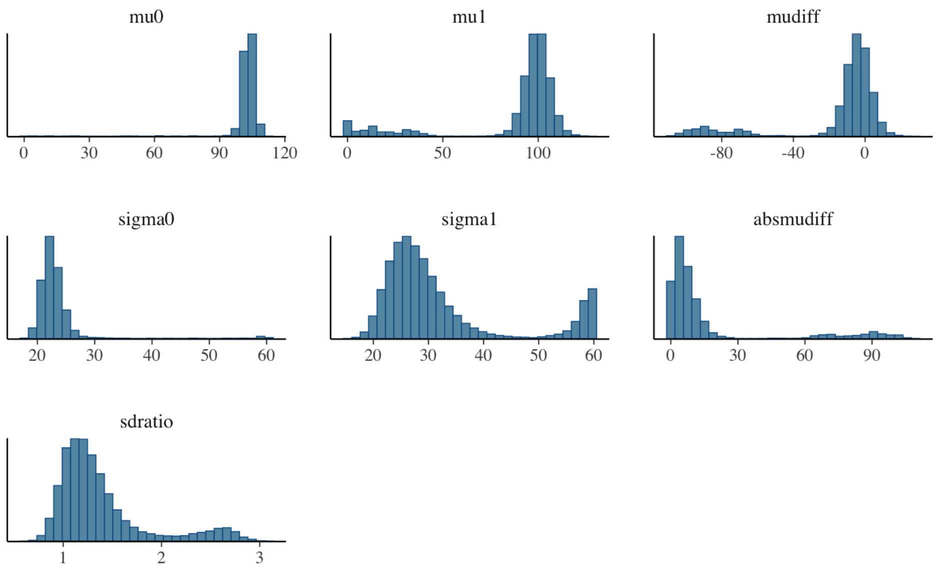
(a) case2 posterior plots.



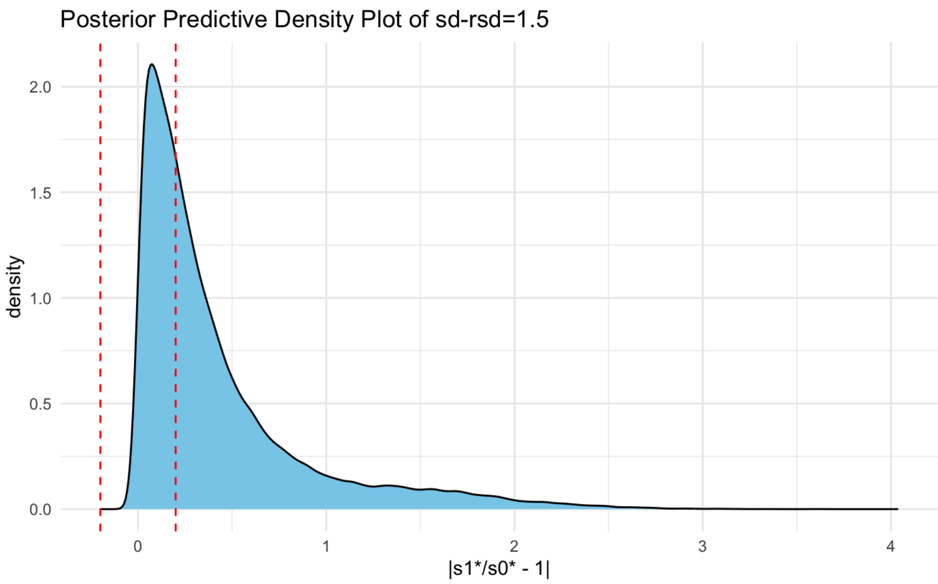
(b) case2 ppd of $P(|s_1^*/s_0^*| < \xi | D_0, D_1)$.

Figure 9

3. case3.



(a) case3 posterior plots.



(b) case3 ppd of $P(|s_1^*/s_0^*| < \xi|D_0, D_1)$.

Figure 10

Scenario	$\lvert\bar{D}_0 - \bar{D}_1\rvert$	$< \delta$?(Thompson)	$P(\lvert\bar{D}_0^* - \bar{D}_1^*\rvert < \delta \mid D_0, D_1)$	$P(\lvert\bar{s}_1^*/\bar{s}_1^* - 1\rvert < \xi \mid D_0, D_1)$
$r_{sd} = 0.6$	3.45	Yes	0.77	0.04
$r_{sd} = 1$	3.52	Yes	0.667	0.43
$r_{sd} = 1.5$	3.60	Yes	0.498	0.41
$r_{sd} = 1.8$	3.66	Yes	0.45	0.22

Table 3: Result summary of posterior predictive probability of future observe mean difference and sd ratio between historical and current sample means. All cases are set with $n_0 = 100$, $\sigma_0 = 20$, $n_1 = 80$, $\delta = 8$ and $\xi = 0.2$. $\mu_0 = 100$, $\mu_1 = 100$, 25% percentage of current data is used to construct the α_0.

Again, the summary of the result from (6**) is given in Table 3. Based on this result, we could say the (6**) may be a better way to solve the issue from Thompson than (6*). But a simulation is needed to robust this result and to calibrate a good/reasonable values of δ and ξ.

References

[Gelman and Carlin, 2013] Gelman, A. and Carlin, J.B.and Stern, H. D. D. V. A. R. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC, 3rd edition.

[Thompson et al., 2021] Thompson, L., Chu, J., Xu, J., Li, X., Nair, R., and Tiwari, R. (2021). Dynamic borrowing from a single prior data source using the conditional power prior. *Journal of Biopharmaceutical Statistics*.