Bayesian inference for exponential distribution model using relative belief

STA4522 The Measurement of Statistical Evidence

Name: Tian Han (Eric) **Guan**

Student ID: 998978058

Department: Statistics

Program: MSc

Email: tianhan.guan@mail.utoronto.ca

Abstract

In recent years, Bayesian inference has drawn more attention in various fields of statistical analysis, however, how to measure statistical evidence precisely in Bayesian analysis still remains a challenging problem. This project discusses how to make inference for the exponential distribution model in the context of Bayesian analysis, using a particular measure of evidence – the relative belief ratio, to illustrate how it can be used to make statistical reasoning precisely. This project has the following structure. First, it introduces the specific problem along with the choice of the model and prior, including control of bias due to the choices. Then, using the observed data, a detailed model checking and prior-data conflict analysis are performed. Finally, inferences for the quantities of interest are made based on the relative belief ratios.

Introduction

Statistical reasoning has a variety of uses in almost all scientific fields, and in order to make the reasoning reliable, a complete theory of statistical inference needs to be developed. One of the important concerns of a suitable theory of statistical inference is how to measure evidence properly so that precise inferences can be made. Within the Bayesian inference context, one popular method to measure statistical evidence is to use the Bayes factors, where large values of Bayes factors correspond to strong evidence for hypothesis assessments. However, it has been shown that in some examples such as the Jeffreys-Lindley's paradox, the Bayes factors can be problematic and therefore some other measure of evidence should be used to fit into the general theory of statistical reasoning that is free from paradoxes and can always provide clear results to statistical problems.

This project discusses how to use a particular measure of evidence – the relative belief ratio in the context of Bayesian inference for an exponential distribution model in a real problem. As opposed to measures such as Bayes factors, for the relative belief principle, measure of evidence is separated from the measure of its strength, and therefore can solve many of the paradoxes. In other words, the relative belief ratio seeks to measure evidence by the amount of belief change from priori to posteriori, and uses probability to only measure the strength of evidence but not the evidence itself. Throughout the project, the evidential approach to the proposed statistical problem is emphasized, and analyses such as prior elicitation, control of biases of model and prior, model checking and prior-data conflict are performed in order to make sure the final inferences (estimation and hypothesis assessments) made are accurate and meaningful.

1. Problem and model

Decay problems are very common in real life, ranging from number of radioactive isotopes to amplitude of damped harmonic oscillators. This project focuses on a particular decay problem, the decay of beer foam (or beer froth), which has been discussed in many mathematics and physics textbooks and experiments.

1.1 Data measurements

The dataset used in this project contains the measurements of beer foam height at various time points after a bottle of Shiner Bock beer was poured into a glass under the room temperature. From a statistical modelling perspective, the data that needs to be modelled is a one-dimensional dataset with *13* data points, where the height was measured every 15 seconds in the first 90 second interval and every 30 seconds in the next 210 second interval. In addition, the height of the beer foam was measured up to an accuracy of *0.01 cm* which was considered to be accurate enough for the experiment.

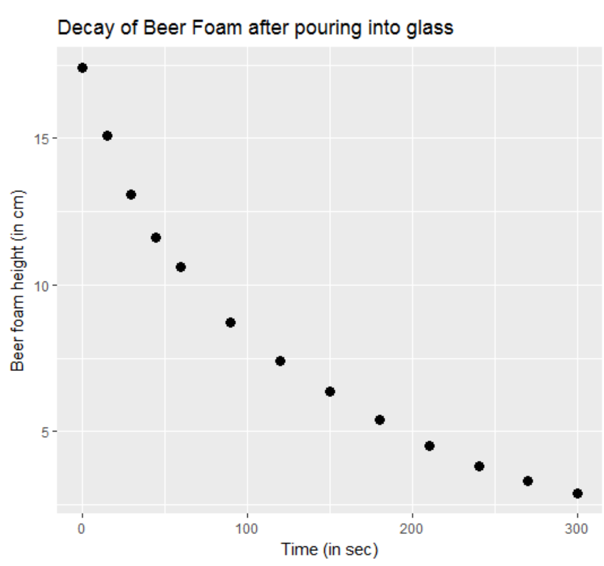

Figure 1. *Measurement of Shiner Bock beer foam height (in cm) after pouring into a glass*

1.2 Choice of the model

Similar to many other substances, various past studies have shown that the height of beer foam is subject to **exponential decay** – where the height decreases at a rate that is proportional to its current value. Mathematically, an exponential decay process can be defined as follows.

Definition 1 (Exponential decay)

Let X be the height of the beer foam and let $\lambda > 0$ be the exponential decay rate. Then, an exponential decay process of the height can be defined using the following differential equation:

$$\frac{dX}{dt} = -\lambda X$$

The solution to this first order linear differential equation is:

$$X(t) = X_0 e^{-\lambda t}, where\ t\ is\ time\ and\ X_0\ is\ the\ initial\ height$$

Given that the height of the beer foam follows the exponential decay process, it can be believed that *distribution* of the height can be modelled by an **exponential distribution** as it can be shown that exponential decay is just a scalar multiple of the exponential distribution, with the same decay rate parameter λ. Therefore, an exponential distribution model is used to model the height of the beer foam in this problem.

Definition 2 (Exponential distribution)

Let $X_1, \ldots, X_n$ be exponential random variables with rate parameter λ that represent the height of beef foam. Then the probability density function $f(x)$ is given by:

$$\boldsymbol{f(x) = \lambda e^{-\lambda x}, x > 0, \lambda > 0}$$

1.3 Quantity of interest

Since in this problem, the exponential distribution is used to model the exponential decay process because they share the same rate parameter, the quantity of interest here is simply the rate parameter λ. It should be noted that the value of λ controls how fast the foam decays, the larger the value, the faster the foam decays, as a result, statistical inference about λ is very important in analyzing the decay behavior of the beer foam.

2. Elicitation procedure for the prior

The key difference between frequentist inference and Bayesian inference is that the later analysis involves putting a prior distribution $\pi(.)$ on the parameter of the sampling model. By making the parameter itself a random variable, this allows more uncertainties and expert knowledge to be added into the modelling process, which is usually considered very useful in case where the dataset has a very limited sample size (as in this problem). The following procedures are taken to select a good prior distribution for the model in this problem, i.e. selecting a distribution for the rate parameter λ.

Step 1. Choose a family of distribution for λ

For this problem, the prior distribution $\pi(.)$ is selected to be the **gamma distribution** for the following reasons:

- Gamma distribution has support over $(0, \infty)$ as the rate parameter λ needs to be strictly positive

- Gamma distribution can take very flexible shapes (therefore we can make the prior less informative if needed)
- The conjugate prior for the exponential distribution is the gamma distribution (unless there is a strong reason to reject the gamma family, using a conjugate prior is convenient)

It should be noted that the gamma prior in this analysis takes the following parametrization.

Definition 3 (Gamma prior)

Let the rate parameter λ be a gamma random variable with shape parameter α and

rate parameter β . Then the probability density function $\pi(\lambda)$ is given by:

$$\boldsymbol{\pi(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}, \qquad \lambda > 0, \alpha > 0, \beta > 0}$$

Step 2. Choose parameters for the prior distribution

Definition 4 (Half-life)

In a decay process, the half-life *t'* is defined as the time required for a quantity to reduce to half of its initial value. In exponential decay, *t'* is computed to be:

$$\mathrm{t}' = \frac{ln(2)}{\lambda}, where\ \lambda \text{ is the rate parameter}$$

The next step is to select values for parameter α and β. This is done using two pieces of *"prior expert knowledge"* on half-life of beer foam:

(1) Past studies based on some popular beers have shown that the beef foam half-life *t'* is around *60s*. This corresponds to a decay rate $\lambda_1 = \frac{ln(2)}{4}$ (note that the denominator 4 corresponds to 60s as the height is measured every 15s in the first interval).

(2) An extreme decay rate (for some beer foams that decay very fast, taken to be 99% percentile or 1 in 100 beers) is taken to be *30s*. This corresponds to a decay rate $\lambda_2 = \frac{ln(2)}{2}$

Using information (1) and (2), select parameters α and β such that:

- The *prior mode is equal to* λ_1, (α, β) satisfies $\frac{\alpha-1}{\beta} = \frac{ln(2)}{4}$.
- The *0.99 quantile of the prior is equal to* λ_2, (α, β) satisfies $\Pi(\lambda) = \frac{1}{\Gamma(\alpha)}\gamma(\alpha, \beta\lambda) = \frac{ln(2)}{2}$

  (where $\Pi(\lambda)$ is the gamma distribution function).

Solved numerically, the above two equations gives the prior parameters $(\alpha, \beta) \approx (11{,}59)$.

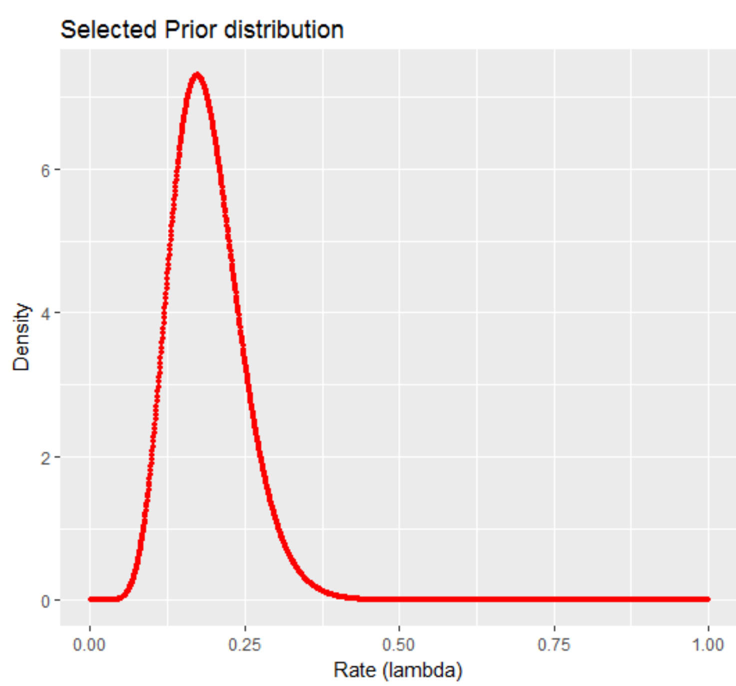

Figure 2. *Selected gamma prior using the above elicitation procedures.*

3. Control of bias due to model and prior choices

One of the challenging issues with Bayesian analysis is that the inference is always going to be affected substantially by biases induced by the prior. One famous example is the Jeffreys-Lindley's paradox, where if one makes the prior more diffuse then there will be more supports for the hypothesis based on measures such as Bayes factors. In this case, although making prior more diffuse seems to make the prior more uninformative, it is actually creating more bias in favor of the hypothesis and therefore making the inference unreliable. This section discusses how to quantify both the bias against and bias in favor for hypothesis assessments in this problem. This gives a way to see if the selected prior is inducing significant bias into the inference, and therefore to decide whether another prior should be used.

3.1 Prior predictive distribution of the model

Definition 5 (Prior predictive distribution)

Let $\pi(\lambda)$ be the prior distribution for the parameter λ, and let $f(x|\lambda)$ be the likelihood function of the observed data. Then the prior predictive distribution $m(x)$ is defined as:

$$\boldsymbol{m(x) = \int_{\lambda} f(x|\lambda)\,\pi(\lambda) d\lambda}$$

Intuitively, the prior predictive distribution predicts a new value before the samples are collected (only using the prior belief about the parameter and the sampling distribution).

Here, instead of deriving the prior predictive distribution in terms of the original random variable *X*, the *minimal sufficient statistic T(X)* is being used as we want the analysis only depends on the data through the minimal sufficient statistic. Note that for the exponential distribution, the minimal sufficient statistic is $T(X) = \sum_{i=1}^{n} X_i$ and with the gamma prior with parameters α and β, an analytical form of the prior predictive distribution is obtained as:

$$m_T(t) = \frac{\beta^{\alpha} t^{n-1} \Gamma(\alpha + n)}{\Gamma(\alpha)\Gamma(n)(t+\beta)^{\alpha+n}}, where\ t = \sum_{i=1}^{n} x_i\ , n\ is\ the\ sample\ size$$

In this problem, we are interested in the following two hypothesis assessments. For each hypothesis assessment, both the bias against and bias in favor are computed to assess whether the selected prior induces significant bias into the assessments.

1) Hypothesis assessment #1 (Primary question)

$$H_0: \lambda = \lambda_0 = \frac{ln(2)}{4}$$

(i.e. whether the beer in this dataset is a type of standard beers as used in many experiments)

2) Hypothesis assessment #2 (Secondary question)

$$H_0: \lambda = \lambda_0 = \frac{ln(2)}{2}$$

(i.e. whether the beer in this dataset is a type of “extreme” beers which has a very fast decay rate)

3.2 Bias against the hypothesis assessment

Bias against occurs when evidence in favor is not obtained when the null hypothesis is true with high prior probability. Using the principle of evidence, the bias against of $\mathrm{H}_0: \lambda = \lambda_0$ is measured by:

$M(RB_{\Lambda}(\lambda_0|x) \leq 1 \mid \lambda_0)$ *where* $M(.|\lambda_0)$ is the prior probability measure of the data given the null hypothesis is true.

In this case, using the prior predictive distribution, the bias against can also be written as:

$M(\mathrm{RB}_{\Lambda}(\lambda_0|\mathrm{x}) \leq 1 \mid \lambda_0) = \boldsymbol{M}\left(\frac{\boldsymbol{m_T(T(x)|\lambda_0)}}{\boldsymbol{m_T(T(x))}} \leq \mathbf{1} \mid \boldsymbol{\lambda_0}\right)$ which in this case is simply a frequentist probability. Using the derived prior predictive density, the bias against both hypothesis assessments are computed as follows:

1) Hypothesis assessment #1 (Primary question)

$$Bias\ against = M\left(\frac{m_T(T(x)|\lambda_0)}{m_T(T(x))} \leq 1 \mid \lambda_0\right) = \mathbf{0.354}$$

The result indicates that there is *some but not significant* bias against the null hypothesis using the selected prior.

2) Hypothesis assessment #2 (Secondary question)

$$Bias\ against = M\left(\frac{m_T(T(x)|\lambda_0)}{m_T(T(x))} \leq 1 \mid \lambda_0\right) = \mathbf{0.014}$$

The result indicates that there is *very little* bias against the null hypothesis using the selected prior.

3.3 Bias in favor of the hypothesis assessment

Bias in favor occurs when evidence against is not obtained when the null hypothesis is false with high prior probability. Using the principle of evidence, the bias in favor of $\mathrm{H}_0: \lambda = \lambda_0$ is measured by

$M(\mathrm{RB}_{\Lambda}(\lambda_0|\mathrm{x}) \geq 1 \mid \lambda_0) - M(\mathrm{RB}_{\Lambda}(\lambda_0|\mathrm{x}) \geq 1 \mid \lambda_0)\mathit{\Pi}_{\Lambda}(\lambda_0)$

= $M(\mathrm{RB}_{\Lambda}(\lambda_0|\mathrm{x}) \geq 1 \mid \lambda_0)$, as in this problem $\mathit{\Pi}_{\Lambda}(\lambda_0) = 0$ for all values of λ_0

$= \boldsymbol{M}\left(\frac{\boldsymbol{m_T(T(x)|\lambda_0)}}{\boldsymbol{m_T(T(x))}} \geq \mathbf{1} \mid \boldsymbol{\lambda_0}\right)$ which again is simply a frequentist probability.

In addition, it should be noted that for *continuous priors*, it is not practical to consider parameter values that are very close to the hypothesized value as we cannot distinguish the difference between them. As a result, the bias in favor is often computed as:

$$sup_{\lambda:\mathrm{d}(\lambda,\lambda_0)\geq\delta}\ M\left(\frac{m_T(T(x)|\lambda_0)}{m_T(T(x))} \geq 1 \mid \lambda_0\right), where\ d(.)\ is\ some\ distance\ measure$$

Since $M\left(\frac{m_T(T(x)|\lambda_0)}{m_T(T(x))} \geq 1 \mid \lambda_0\right)$ decreases when moving away from λ_0, the bias in favor can also be computed over the supremum over $\{\boldsymbol{\lambda}: \mathbf{d}(\boldsymbol{\lambda}, \boldsymbol{\lambda_0}) = \boldsymbol{\delta}\}$ which implies

$$Bias\ in\ favour = \boldsymbol{sup_{\lambda:\mathrm{d}(\lambda,\lambda_0)=\delta}}\ \boldsymbol{M}\left(\frac{\boldsymbol{m_T(T(x)|\lambda_0)}}{\boldsymbol{m_T(T(x))}} \geq \mathbf{1} \mid \boldsymbol{\lambda_0}\right)$$

In this problem, we choose the distance function to be the *Euclidean distance and* δ *to be 0.01* (the same value used for discretizing the relative belief ratio). Therefore, the bias in favor is taken over the supremum of the set $(\boldsymbol{\lambda_0} - \mathbf{0.01}, \boldsymbol{\lambda_0} + \mathbf{0.01})$.

1) Hypothesis assessment #1 (Primary question)

$$Bias\ in\ favour = sup_{(\lambda_0-0.01,\lambda_0+0.01)}\ M\left(\frac{m_T(T(x)|\lambda_0)}{m_T(T(x))} \geq 1 \mid \lambda_0\right) = \mathbf{0.688}$$

The result indicates that there is *some* bias in favor of the null hypothesis induced by the selected prior, at least when a deviation of 0.01 from the hypothesized value is considered to be meaningful.

2) Hypothesis assessment #2 (Secondary question)

$$Bias\ in\ favour = \ sup_{(\lambda_0 - 0.01, \lambda_0 + 0.01)} \ M\left(\frac{m_T(T(x)|\lambda_0)}{m_T(T(x))} \geq \ 1 \mid \lambda_0\right) = \mathbf{0.989}$$

The result indicates that there is *significant bias* in favor of the null hypothesis induced by the selected prior, at least when a deviation of 0.01 from the hypothesized value is considered to be meaningful.

Conclusion

By analyzing both biases on the two hypothesis assessments, it is concluded that the *selected prior is a reasonable choice* as in most of the cases the bias is not significant.

Remarks

1. For a given prior, both the bias against and bias in favor will decrease toward to 0 when the sample size increases, which implies the bias induced by any prior can be controlled by the study design. This makes sense as in Bayesian analysis, the prior will have less impact on inference when the amount of available data increases.
2. For bias in favor, a smaller deviation will result in more bias in favor to be seen.

4. Model checking based on the observed data

In all statistical problems, model checking is a key step before actually making inference as all analyses are based on assumptions. In Bayesian analysis, the failure of the modelling process can be caused by two reasons: the sampling distribution and the prior. Therefore, it is essential to first perform some model checking procedures for the sampling model before discussing prior-data conflict. In this problem, the exponential distribution model is checked against the observed data using both numerical and graphical assessments.

4.1 Numerical assessment

The *one-sample Kolmogorov-Smirnov test* is being used to check if the exponential distribution is a reasonable fit of the observed data. The one-sample Kolmogorov-Smirnov test is a nonparametric test to check the goodness-of-fit of some given continuous one-dimensional probability distribution against a data sample. The test statistic computes the distance between the empirical distribution function of the sample and the distribution function of the reference distribution, and the null hypothesis is that the samples are drawn from the reference distribution.

In this problem, the samples are tested against an exponential distribution which the rate parameter is estimated using maximum likelihood. By performing the test, a **p-value of 0.184** is obtained which shows that there is no strong evidence against the selected model. As a result, the numerical assessment indicates that the exponential distribution *is a reasonable choice* of the data.

4.2 Graphical assessment

In addition to the numerical test, a graphical assessment is also being used to see if the exponential distribution gives a reasonable fit to the sample data using a *kernel density plot*. It should be noted that even though graphical tools such as histogram is usually a popular choice, a very small sample size (as in this problem) will make it very challenging to select appropriate bandwidth for good visualization comparison. On the other hand, kernel density estimation is a nonparametric method to estimate the density function of the data and therefore can give a smooth density to be compared with the reference distribution.

From the kernel density plot below, it can be seen that the exponential distribution *is a reasonable model* for the data samples.

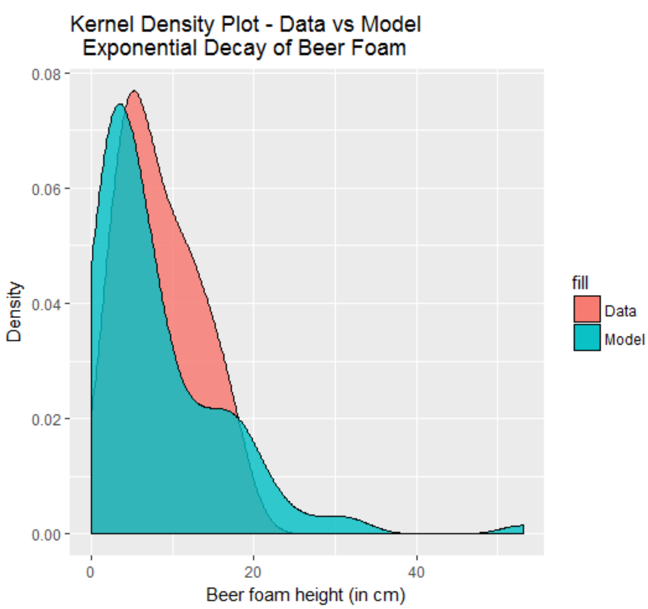

Figure 3*. Use a kernel density plot to compare the sample distribution and the sampling distribution.*

5. Checking for prior-data conflict based on the observed data

After making sure the sampling model is a reasonable choice for the observed data, now it is important to make sure the selected prior does not have significant conflict against the observed data. Prior-data conflict essentially means that the selected prior is placing most of its mass on parameter values which is surprising given the observed data.

5.1 Diagnostics

Before doing a formal analysis on prior-data conflict, the prior and posterior distributions of the rate parameter are plotted on the same graph to see how different they are. Such diagnostic is a useful tool in *low dimension and unimodal problems*, and in case where the prior and posterior distributions have very different high density regions, we can conclude that there is significant prior-data conflict.

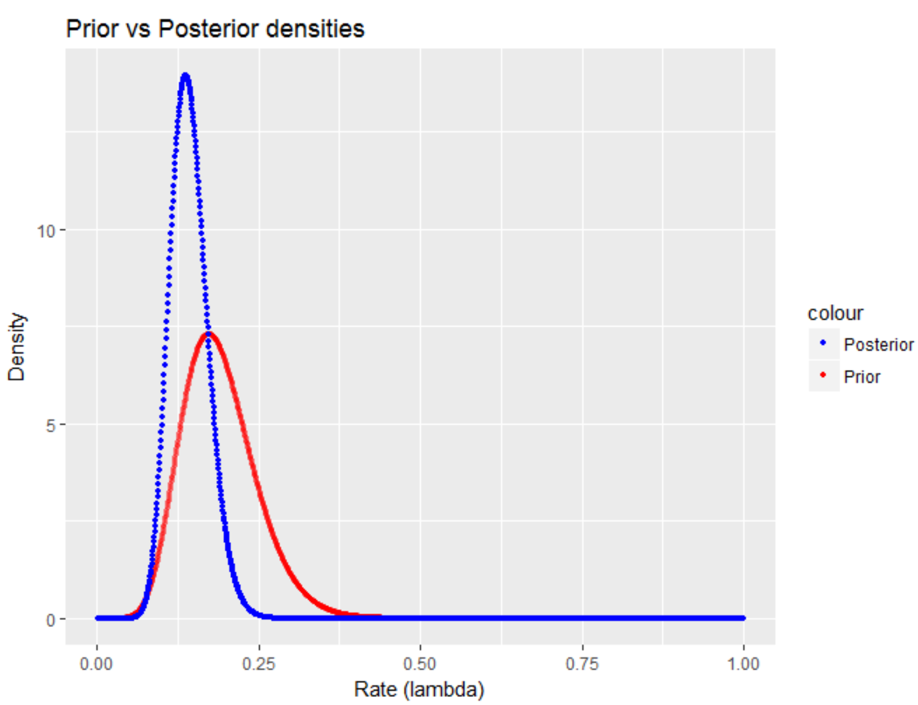

Figure 4. *Prior vs posterior density of the rate parameter lambda.*

Since the high density regions where most of the masses are not very different but have significant overlapping, the diagnostic *does not indicate significant prior-data conflict*. It can be also seen that the posterior gives a much confident (narrow) high density region, which shows that after seeing the data, we are more confident about the unknown rate parameter of the model.

5.2 Prior-data conflict using the prior predictive distribution

Even though diagnostics such as prior-posterior plot can be useful, they are mathematically informal and sometimes too conservative, and can become very challenging for high dimensional parameters. Therefore, this section checks the prior-data conflict using a formal measure which involves the prior predictive distribution with the minimal sufficient statistic.

In order to evaluate whether the observed minimal sufficient statistic *t(x)* gives surprising result using the selected gamma prior, a tail probability of $\boldsymbol{M_T(m_T(t) \leq m_T(t(x))}$, is computed numerically. Intuitively, this tail probability tells what is the prior probability of seeing a minimal sufficient statistic with probability less than that of the observed minimal sufficient statistic, which implies how extreme or surprising the prior distribution behaves when compared to the observed data. Figure 5 below shows the prior predictive distribution and where the observed minimal sufficient statistic lies in the distribution. The tail probability is computed to be $M_T(m_T(t) \leq m_T\big(t(x)\big) \approx \mathbf{0.15}$. Both the plot and the tail probability *do not indicate significant prior-data conflict*. Note that in other cases where a significant conflict is seen, a less informative prior should be chosen.

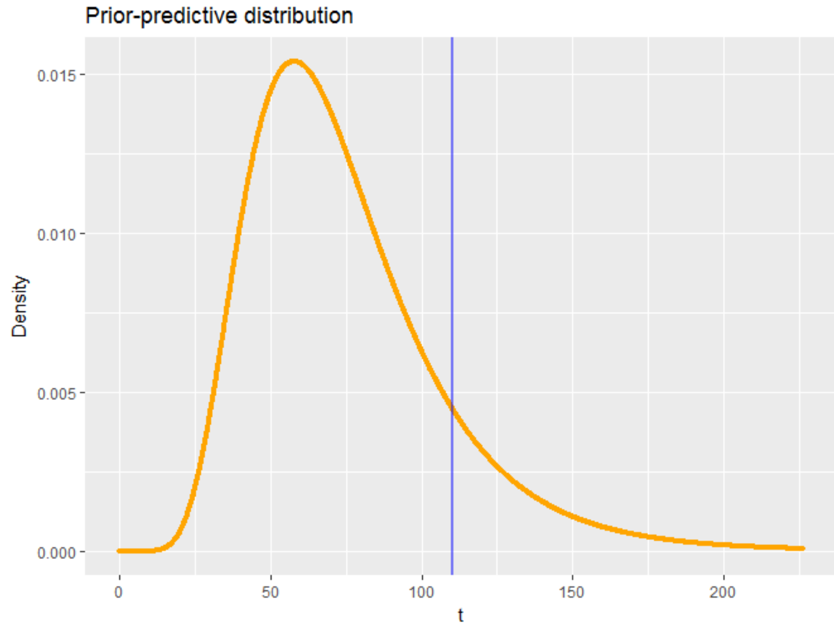

***Figure** 5. Plot of the prior predictive distribution of the minimal sufficient statistic. The blue line shows where the observed minimal sufficient statistics t(x) lies in the distribution.*

6. Inferences for the rate parameter based on relative belief

Although the Bayes factor is commonly used as the measure of evidence in most Bayesian inference problems, here, the relative belief ratio is being used to measure the statistical evidence for both estimation and hypothesis testing.

Definition 5 (Relative belief ratio)

Here, the quantity of interest is the rate parameter λ. The relative belief ratio is defined as

$$\mathrm{RB}_{\Lambda}(\lambda|\mathrm{x}) = \frac{\pi_{\Lambda}(\lambda|\mathrm{x})}{\pi_{\Lambda}(\lambda)}$$

where $\pi_{\Lambda}(.|\mathrm{x})$ *and* π_{Λ} *are the posterior and prior densities of* λ

Note that in this problem, the *posterior distribution* has a closed form:

$$\Lambda|\mathbf{x} \sim \mathbf{Gamma}(\boldsymbol{\alpha} + \boldsymbol{n}, \boldsymbol{\beta} + \boldsymbol{t}), where\ t = \sum_{i=1}^{n} x_i$$

Therefore, the relative belief ratio can be derived as:

$$\mathbf{RB}_{\Lambda}(\boldsymbol{\lambda}|\mathbf{t}) = \frac{(\boldsymbol{\beta} + \boldsymbol{t})^{\alpha+n}\Gamma(\boldsymbol{\alpha})}{\boldsymbol{\beta}^{\alpha}\Gamma(\boldsymbol{n})}\boldsymbol{\lambda}^{n}\boldsymbol{e}^{-\lambda \mathbf{t}}$$

Figure 6 below shows the plot of the relative belief ratio. The key idea about the relative belief ratio is that it measures evidence by the amount of beliefs changing from our prior belief to the posteriori, and therefore follows the key concept of statistical reasoning: evidence is what causes our beliefs to change.

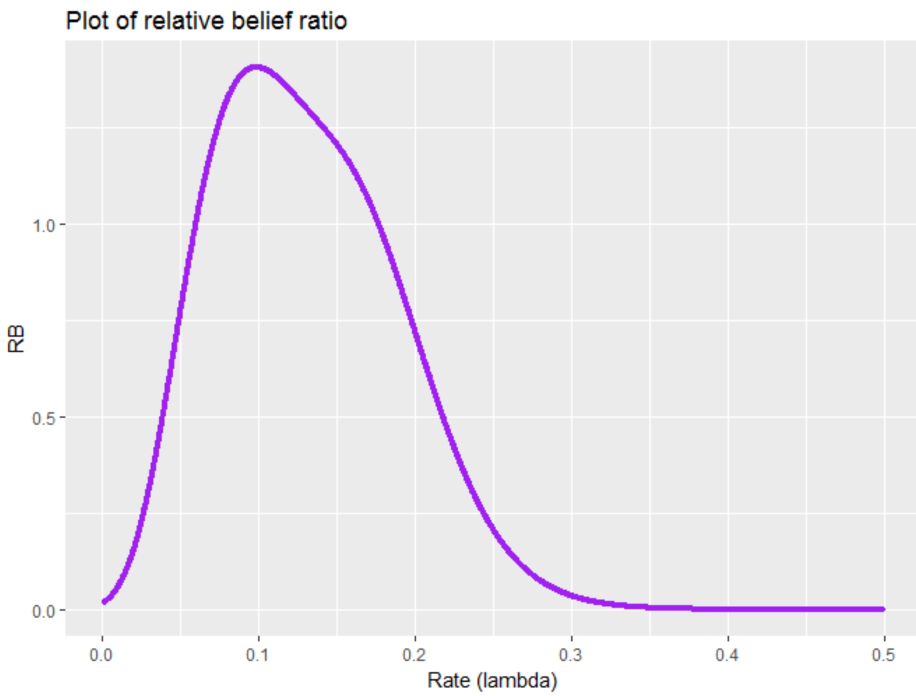

Figure 6. *Plot of the relative belief ratio (discretized version) in terms of the rate parameter lambda.*

6.1 Estimation

6.1.1. Relative belief estimate of the rate parameter

The relative belief principle says that the best estimate of the parameter λ is the value which gives the greatest amount of evidence, i.e.

$$\boldsymbol{\lambda}(\mathbf{x}) = \mathbf{argsup}\ \mathbf{RB}_{\Lambda}(\boldsymbol{\lambda}|\mathbf{x})$$

In this case, the relative belief estimate is $\boldsymbol{\lambda} \approx \mathbf{0.124}$ (and note that this is the same as the maximum likelihood estimate).

6.1.2. Accuracy of the estimate

To get an assessment of how accurate the relative belief estimate is, a *plausible region* is constructed as:

$$\mathbf{PL}_{\Lambda}(\mathbf{x}) = \{\boldsymbol{\lambda} \mid \mathbf{RB}_{\Lambda}(\boldsymbol{\lambda}|\mathbf{x}) > \mathbf{1}\}$$

Intuitively, the plausible region contains all the values of the parameter that will give the evidence in favor. In this case, the plausible region is found to be $\mathbf{PL}_{\Lambda}(\mathbf{x}) = (\mathbf{0.084}, \mathbf{0.169})$.

Remarks

Here, the plausible region has a length of *0.085* which indicates that we have a *reasonable degree of uncertainty* about the true parameter value of λ using the relative belief estimate. It is also noted that the plausible region construction only depends on the principle of evidence and not the principle of relative belief, so that all valid estimators can have the same assessment of accuracy.

6.2 Hypothesis assessments

For each hypothesis, both the direction of evidence (i.e. against or in favor) and the strength of evidence are computed. In addition, since in this problem the gamma prior has a continuous probability distribution, to avoid the strength of evidence to be meaningless when the null hypothesis is true (i.e. $\mathrm{Str}_\Lambda(\lambda_0|x) \to Unif(0,1)\ when\ H_0\ is\ true$), the relative belief ratio is computed using the *discretized versions of both the prior and posterior distributions*, namely,

$$\mathbf{RB}_\Lambda(\boldsymbol{\lambda_0}|\mathbf{x}) = \frac{\boldsymbol{\pi}_\Lambda\left([\boldsymbol{\lambda_0} - \frac{\boldsymbol{\delta}}{\mathbf{2}}, \boldsymbol{\lambda_0} + \frac{\boldsymbol{\delta}}{\mathbf{2}}]|\mathbf{x}\right)}{\boldsymbol{\pi}_\Lambda\left([\boldsymbol{\lambda_0} - \frac{\boldsymbol{\delta}}{\mathbf{2}}, \boldsymbol{\lambda_0} + \frac{\boldsymbol{\delta}}{\mathbf{2}}]\right)}$$

Choice of $\boldsymbol{\delta}$

In this discretized version of relative belief, it is important to choose a value of δ based on the specification of data collection aspects for the measurement such that if $|\lambda - \lambda_{\text{true}}| < \frac{\delta}{2}$ then the difference has no practical importance. In this problem, we know that heights are measured up to the *0.01 cm accuracy level*, therefore any measurement less than 0.01 cm is considered to be no practical importance. In addition, given the parametrization of the exponential distribution, the expected value is $\mathrm{E}[\mathrm{X}] = \frac{1}{\boldsymbol{\lambda}}$. Therefore, we want to make sure that the mean height is measured with an accuracy of 0.01 cm, i.e. $\left|\frac{1}{\lambda} - \frac{1}{\lambda_{\text{true}}}\right| < 0.01$. Given the range of the rate parameters it is noticed that $|\lambda - \lambda_{\text{true}}| < \left|\frac{1}{\lambda} - \frac{1}{\lambda_{\text{true}}}\right| < 0.01$, hence it is reasonable to select $\frac{\boldsymbol{\delta}}{\mathbf{2}} = \mathbf{0.01}$ for the discretization. In general, we can only select bigger δ but not smaller δ within the context, as data is always measured within some degree of accuracy, and the procedure of discretization is never an approximation to the reality (as in real life, none of the measurements are actually continuous).

Using the discretized version of relative belief ratio, the following results are obtained:

1) Hypothesis assessment #1 (Primary question)

$$\mathrm{H}_0: \lambda = \lambda_0 = \frac{ln(2)}{4}$$

(i.e. whether the beer in this dataset is a type of standard beer as used in many experiments)

- $\mathrm{RB}_\Lambda(\lambda_0|\mathrm{x})$ = **0.873**
- $\mathrm{Str}_\Lambda(\lambda_0) = \Pi(\mathrm{RB}_\Lambda(\lambda|\mathrm{x}) \leq \mathrm{RB}_\Lambda(\lambda_0|\mathrm{x}) \mid t) = \mathbf{0.273}$

Since the relative belief ratio is less than 1, we can conclude that there is *evidence against* the null hypothesis as the data leads to a decrease in belief that the hypothesized value is the true value. The strength of evidence tells us that there is *relatively strong evidence* against the null hypothesis (as a strength of 0 implies very strong evidence against and a strength of 1 implies very weak evidence against the null hypothesis), while there is still a posterior probability of 0.273 that the true parameter value has a smaller relative belief ratio. The analysis shows that we should have a reasonable amount of confidence on the reliability of this inference as the strength of evidence is relatively small. This shows that the beer in this dataset does not belong to the category of standard beers that have been used in many experiments and studies.

2) Hypothesis assessment #2 (Secondary question)

$$\mathrm{H}_0: \lambda = \lambda_0 = \frac{ln(2)}{2}$$

(i.e. whether the beer in this dataset is a type of extreme beer which has very fast decay rate)

- $\mathrm{RB}_\Lambda(\lambda_0|\mathrm{x})$ = **0.001**
- $\mathrm{Str}_\Lambda(\lambda_0) = P(\mathrm{RB}_\Lambda(\lambda|\mathrm{x}) \leq \mathrm{RB}_\Lambda(\lambda_0|\mathrm{x}) \mid t) \approx \mathbf{0}$

Since the relative belief ratio is less than 1 and the strength of evidence is close to 0, we can conclude that there is *very strong evidence against* the null hypothesis and we should have a great amount of confidence on the reliability of this inference. This shows that the beer in this dataset does not belong to the category of beer that has foam decays at a very fast speed.

Conclusions

This project explores how to apply the evidential approach to a real statistical problem, using the relative belief ratio as the measure of statistical evidence. Although this problem is in a one-dimension setting, in general the relative belief ratio approach can also applied to more complex problems in high dimensions, and methods such as control of bias can also be applied to estimation problems. In summary, the relative belief measure provides an efficient way of measuring evidence, and therefore makes procedures such as control of bias and prior-data conflict much more reliable. This will give us more confidence on the inference results that we make.