Posterior distributions for likelihood ratios in forensic science

Ardo van den Hout

Department of Statistical Science, University College London

Gower Street, London WC1E 6BT, United Kingdom

E-mail: ardo.vandenhout@ucl.ac.uk

Ivo Alberink

Netherlands Forensic Institute,

Laan van Ypenburg 6, The Hague, The Netherlands

Abstract

Evaluation of evidence in forensic science is discussed using posterior distributions for likelihood ratios. Instead of eliminating the uncertainty by integrating (Bayes factor) or by conditioning on parameter values, uncertainty in the likelihood ratio is retained by parameter uncertainty derived from posterior distributions. A posterior distribution for a likelihood ratio can be summarised by the median and credible intervals. Using the posterior mean of the distribution is not recommended. An analysis of forensic data for body height estimation is undertaken. The posterior likelihood approach has been criticised both theoretically and with respect to applicability. This paper addresses the latter and illustrates an interesting application area.

Key Words: Likelihood ratio, Bayes factor, Bayesian inference, body height estimation, posterior likelihood ratios, precision, reliability.

1 Introduction and terminology

Evaluation of evidence in forensic science can be undertaken using the likelihood ratio framework. For a continuous random variable, the *likelihood ratio* (LR) is the ratio of two values of the probability function $p(x|\theta)$, given two values of model parameter θ, and data x. For values θ_1 and θ_2, we have $LR = p(x|\theta_1)/p(x|\theta_2)$, where function $p(\cdot)$ is a generic notation for a probability density function or a probability mass function.

Given two hypotheses H_1 and H_2 for assumptions for models M_1 and M_2, respectively, the *Bayes factor* (BF) in favour of H_1 is given by

$$BF = \frac{p(x|H_1)}{p(x|H_2)} = \frac{\int p(x|\phi, H_1)p(\phi|H_1)d\phi}{\int p(x|\psi, H_2)p(\psi|H_2)d\psi}. \tag{1}$$

The BF is also called a *marginal likelihood ratio* as it is the ratio of two marginal likelihoods. It is not necessarily the case that $p(x|\phi, H_1)$ is the same function as $p(x|\psi, H_2)$. These probability functions are defined by M_1 and M_2, respectively. The same holds for $p(\phi|H_1)$ and $p(\psi|H_2)$. It is because of this that the BF can be used to compare non-nested models.

If, however, M_1 and M_2 are nested, i.e., one can be derived from the other by restricting a subset of the parameters, then the BF is still different from the LR, as the latter is defined for specific parameter values and the former is defined by integrating out the parameters. It is only in the specific case where the priors given by $p(\phi|H_1)$ and $p(\psi|H_2)$ identify parameter values with probability 1 (have a point mass 1 at those values), that the BF reduces to a LR.

The following example of a Bayes factor in forensic practice is taken from Lucy [1](Section 12.5). An eyewitness height description of the male perpetrator is modelled as a normal distribution with mean 1.816 metres and standard deviation 0.054. The prosecution's hypothesis is H_p: perpetrator = suspect. The defence's hypothesis is H_d: perpetrator $\neq$ suspect. The assumed population distribution

of men is normal with mean 1.775 and standard deviation 0.098. The evidence is the height $E = 1.855$ of the suspect.

The Bayes factor is in this case equal to the probability density of E under H_p divided by the probability density of E under H_d. That is, $BF = f(E|\mu_p, \sigma_p)/f(E|\mu_d, \sigma_d)$, where f is the density of a normal distribution with mean μ and standard deviation σ [1]. For $\mu_p = 1.816, \sigma_p = 0.054, \mu_d = 1.775, \sigma_d = 0.098$ this leads to a Bayes Factor of 1.951.

We would like to add the following explanation in terms of the BF. The BF in this case is defined as

$$BF = \frac{p(E|H_p)}{p(E|H_d)} = \frac{\int p(E|\boldsymbol{\theta}, H_p)p(\boldsymbol{\theta}|H_p)d\boldsymbol{\theta}}{\int p(E|\boldsymbol{\eta}, H_d)p(\boldsymbol{\eta}|H_d)d\boldsymbol{\eta}}. \tag{2}$$

There are no background data, i.e., there are no sample data from the relevant population. The models under both hypotheses are completely specified normal distributions. This means that $p(\boldsymbol{\theta}|H_p)$ specifies $\boldsymbol{\theta} = (\mu_p, \sigma_p)$ with probability one. Likewise $p(\boldsymbol{\eta}|H_b)$ specifies $\boldsymbol{\eta} = (\mu_d, \sigma_d)$ with probability one. As a result both integrals disappear in (2) and we end up with $p(E|\boldsymbol{\theta}, H_p) = f(E|\mu_p, \sigma_p)$ and $p(E|\boldsymbol{\eta}, H_d) = f(E|\mu_d, \sigma_d)$.

Note that there is no uncertainty associated with the BF. Consider the case where background data are used for the estimation of μ_d and σ_d. In that case, the denominator of (2) would have been

$$p(E|H_d, B) = \int p(E|\boldsymbol{\eta}, H_d, B)p(\boldsymbol{\eta}|H_d, B)d\boldsymbol{\eta} \tag{3}$$

$$= \int p(E|\boldsymbol{\eta}, H_d, B)\frac{p(B|\boldsymbol{\eta}, H_d)p(\boldsymbol{\eta}|H_d)}{p(B|H_d)}d\boldsymbol{\eta}, \tag{4}$$

where $p(B|\boldsymbol{\eta}, H_d)$ is the likelihood and $p(\boldsymbol{\eta}|H_d)$ is the prior density. Because the BF is in this case defined conditional on background data B, there is still no uncertainty associated with the BF. The uncertainty with respect to $\boldsymbol{\eta}$ is integrated out. Nevertheless, if a new data set B were sampled, another BF would

be the result. By conditioning on B, this sample uncertainty is not accounted for.

In Section 2, the posterior distribution of the likelihood ratio is explained within the context of forensic science. Section 3 presents an evaluation of evidence where the posterior distribution of the likelihood ratio is used for the measurement of body height. Background data in this case consist of measurements on test persons. A comparison is made with the Bayes factor approach. For the posterior sampling we use WinBUGS (Lunn *et al.* [2]). Section 4 concludes the paper.

2   Posterior likelihood ratio

As an alternative method for simple null hypothesis testing, Aitkin [3] advocates using a Bayesian framework and working with the posterior distribution of the LR. Instead of eliminating the uncertainty by maximising (LR test) or by integrating (BF), Aitkin proposes to retain uncertainty in the LR via parameter uncertainty derived from the posterior distributions.

Bayesian inference focusses on the posterior density of parameters. If θ is the parameter and x are the data, then the posterior is given by $p(\theta|x) = p(x|\theta)p(\theta)/p(x)$, where $p(x|\theta)$ is the likelihood of the data and $p(\theta)$ is the prior density of θ. Thus the posterior is proportional to the likelihood times the prior, and this is written as $p(\theta|x) \propto p(x|\theta)p(\theta)$.

The posterior likelihood ratio approach is readily explained in terms of sampling. The LR is considered a function of the parameters under both hypotheses. First, given H_1: $\theta = \theta_1$, the likelihood is a single value $L(\theta_1) = p(x|\theta_1)$. Second, given H_2: $\theta \neq \theta_1$, S parameter values θ^* are sampled from the posterior $p(\theta|x)$ and for each value the likelihood $L(\theta^*)$ is computed. Next, the S ratios $L(\theta_1)/L(\theta^*)$ provide a random sample from the posterior of the LR.

At first sight, the setting in Aitkin [3] is different from the forensic science

setting. For the former, there is a data set and a model, and the hypotheses are about model parameters. For the latter, there is evidence E and background data B, and the hypotheses are about E - not about the model for B.

For the forensic science setting, we can define an LR given an estimate of model parameters for B. This only works if we assume that both the prosecution and the defence accept the same model for B. If the model parameter vector is denoted $\boldsymbol{\theta}$, then we can define a likelihood ratio as the ratio of two probability densities for the evidence. This conditional ratio is given by

$$LR = \frac{p(E|H_p, \boldsymbol{\theta})}{p(E|H_d, \boldsymbol{\theta})}. \tag{5}$$

For the forensic science setting, the BF is defined as

$$BF \quad = \quad \frac{p(E|H_p, B)}{p(E|H_d, B)} = \frac{\int p(E|H_p, \boldsymbol{\theta}_p) p(\boldsymbol{\theta}_p|B) d\boldsymbol{\theta}_p}{\int p(E|H_d, \boldsymbol{\theta}_d) p(\boldsymbol{\theta}_d|B) d\boldsymbol{\theta}_d}, \tag{6}$$

where $p(\boldsymbol{\theta}_p|B)$ and $p(\boldsymbol{\theta}_d|B)$ are posterior densities.

Given these definitions of BF and LR, we can apply the ideas of the posterior likelihood ratio and achieve a middle way between BF and LR such that the uncertainty in the LR is retained by parameter uncertainty derived from the posterior distribution of the model parameter vector for the background data. Thus we see LR as a function of sampled $\boldsymbol{\theta}$, and obtain its posterior by sampling from the posterior $p(\boldsymbol{\theta}|B)$.

The posterior LR distribution is very useful as it can be used to assess the strength of evidence by way of posterior probabilities such as $P(LR > c)$, for any $c > 0$. In this way it is possible to not only have knowledge about the central location of the LR, but also about the precision that is attached. For the end user of the LR (trier of fact) it may be important to know whether for a reported LR of 1000, a 5% lower bound is e.g. 20 or 990.

Care has to be taken not to summarise the posterior distribution of the likelihood ratio by its posterior mean. The posterior mean is not invariant under the

switching of the order of the hypotheses in the sense that

$$\mathbb{E}_{\boldsymbol{\theta}}\left[\frac{p(E|H_p,\boldsymbol{\theta})}{p(E|H_d,\boldsymbol{\theta})}\right] \neq \left(\mathbb{E}_{\boldsymbol{\theta}}\left[\frac{p(E|H_d,\boldsymbol{\theta})}{p(E|H_p,\boldsymbol{\theta})}\right]\right)^{-1}. \tag{7}$$

This is important since the order of the hypotheses should not effect the statistical inference. Instead of assessing the posterior mean, the posterior median and credible intervals can be used for statistical inference.

3   Evaluation of evidence

In this section, the posterior of the likelihood ratio (5) is used for forensic data for height estimation of a perpetrator. A comparison with the Bayes factor (6) is made.

A perpetrator was well visible on a security camera and one image was chosen as the basis of height measuring. Background data B consist of additional measurements of six test persons who were positioned in the same stance as the perpetrator in front of the original camera (Edelman et $al.$ [4]).

We use the following notation. Background data are measurements m_i, for test persons $i = 1, 2, ..., 6$, and known true heights h_i. The model for the height estimation is

$$m_i = \alpha + h_i + \epsilon_i \qquad \text{with} \qquad \epsilon_i \sim N(0, \sigma^2), \tag{8}$$

where α is the systematic measurement error, see Van den Hout and Alberink [5] for an extended model and details of the data. Let $\boldsymbol{\theta} = (\alpha, \log(\sigma))$.

The evidence is the measured height m_p of the perpetrator. The height of the suspect is h_s. The prosecution's hypothesis is H_p: perpetrator is suspect ($h_p = h_s$). The defence's hypothesis is H_d: perpetrator is not suspect ($h_p \neq h_s$). Assume that both the prosecution and the defence accept model (8). The BF is

given by

$$BF \;\; = \;\; \frac{p(m_p|H_p, B)}{p(m_p|H_d, B)} = \frac{p(m_p|h_p = h_s, B)}{\int p(m_p|h_p = h, B)p(h)dh} \tag{9}$$

$$= \;\; \frac{\int p(m_p|\boldsymbol{\theta}, h_p = h_s)p(\boldsymbol{\theta}|B)d\boldsymbol{\theta}}{\int \left[\int p(m_p|\boldsymbol{\theta}, h_p = h)p(h)dh\right] p(\boldsymbol{\theta}|B)d\boldsymbol{\theta}}. \tag{10}$$

Let us assume that the height distribution of the population is given by $p(h|\mu_h, \sigma_h)$, a normal distribution with known mean μ_h and known standard deviation σ_h. The conditional LR is given by

$$LR \;\; = \;\; \frac{p(m_p|h_p = h_s, \boldsymbol{\theta})}{\int p(m_p|h_p = h, \boldsymbol{\theta})p(h|\mu_h, \sigma_h)dh}. \tag{11}$$

The numerator of (11) is a normal density and is given by

$$p(m_p|h_p = h_s, \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\frac{(m_p - \alpha - h_s)^2}{\sigma^2}\right] \tag{12}$$

Since $p(h|\mu_h, \sigma_h)$ is a normal distribution, there is a closed-form solution for the integral in the denominator of (11). The integrand is a convolution of two normal distributions and the denominator is given by

$$\int p(m_p|h_p = h, \boldsymbol{\theta})p(h|\mu_h, \sigma_h)dh = \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma_h^2)}} \exp\left(-\frac{1}{2}\frac{(m_p - \alpha - \mu_h)^2}{\sigma^2 + \sigma_h^2}\right), \tag{13}$$

see, e.g., Gelman *et al.* [6] (Section 2.6) for a similar computation. If $\boldsymbol{\theta}$ is treated as a fixed value, then there is no uncertainty associated with LR.

For the posterior of LR, firstly, we sample $\boldsymbol{\theta}^*$ from the posterior $p(\boldsymbol{\theta}|B)$. Secondly we compute LR for each sampled $\boldsymbol{\theta}^*$.

To obtain the posterior $p(\boldsymbol{\theta}|B)$, we have to specify the prior of the model parameter vector $\boldsymbol{\theta}$. Gelman *et al.* [6] discuss the definition of the prior density in the context of the normal distribution, and also the sampling from the resulting posterior. Various levels of informativeness and conjugacy are presented by Gelman *et al.*

Table 1: Background data on measured heights and true heights of test persons, and measured height of perpetrator.

	Test persons						Perpetrator
Measured height	1.964	1.832	1.900	1.780	1.937	1.865	1.885
True height	1.950	1.795	1.865	1.755	1.910	1.825	-

For the evaluation of evidence in the present setting, we specify an informative proper prior $p(\boldsymbol{\theta})$ without worrying about conjugacy as we will rely on the automatic MCMC procedures in WinBUGS to do the sampling.

To compare the posterior likelihood ratio approach with the Bayes Factor (10), we approximate the integrals in the latter by using the trapezoidal rule (with 500 nodes). This computation includes the estimation of the marginal density $p(B)$ since the posterior for $\boldsymbol{\theta}$ is given by $p(\boldsymbol{\theta}|B) = p(B|\boldsymbol{\theta})p(\boldsymbol{\theta})/p(B)$. In general, the estimation of marginal density can be complex (Carlin and Louis [7]). Since $\boldsymbol{\theta}$ consists of only two parameters, numerical approximation of the integrals works fine. Sampling from the posterior of LR is undertaken in WinBUGS (Lunn et al. [2]). WinBUGS is freely available software for the Bayesian analysis of statistical models using Markov chain Monte Carlo (MCMC) methods, see also www.mrc-bsu.cam.ac.uk/bugs. Code is provided in the Appendix. For the inference in this application, the MCMC consisted of two chains, each with a burn-in of 10000, and a further 10000 updates for inference. Convergence of the MCMC was checked by using the diagnostic tools provided within WinBUGS.

Evidence m_p and background data for the height estimation are presented in Table 1. The population distribution of Dutch Caucasian men is assumed to be normal with mean $\mu_h = 1.806$ and standard deviation $\sigma_h = 0.1$ (Statistics Netherlands, www.cbs.nl, 2006). This specifies $p(h|\mu_h, \sigma_h)$. For the prior of $\boldsymbol{\theta}$ we
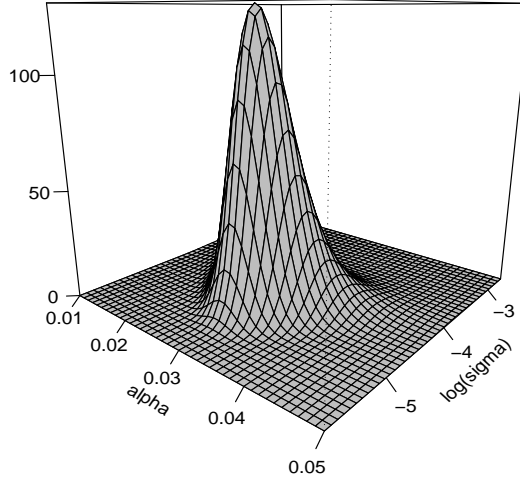
Figure 1: Posterior density $p(\boldsymbol{\theta}|B) = p(\alpha, \log(\sigma)|B)$.

assume $p(\boldsymbol{\theta}) = p(\alpha, \log(\sigma)) = p(\alpha)p(\log(\sigma))$, and furthermore $\alpha \sim N(0, 0.1)$ and $\log(\sigma) \sim U(-10, 0)$. These priors are informative and take into account that the measurements are in meters.

Bayesian inference using WinBUGS yields a posterior mean 0.029 for α with 95% credible interval (CI) $(0.017, 0.042)$. So there is a systematic overestimation of the height of about 3cm. For σ the figures are 0.012 $(0.006, 0.024)$. The posterior density $p(\boldsymbol{\theta}|B)$ has a regular shape and is depicted in Figure 1.

We will illustrate the evaluation of the evidence $m_p = 1.885$ for various values of the height of the suspect h_s. Say that the suspect has the same height as the perpetrator. The height of the perpetrator is of course unknown, but this is the situation which would pertain if the suspect were the perpetrator. In that case $m_p - \alpha \approx 1.885 - 0.029 = 1.856 = h_s$. The posterior distribution of the likelihood ratio for this situation is depicted in Figure 2. If the suspect has the same height
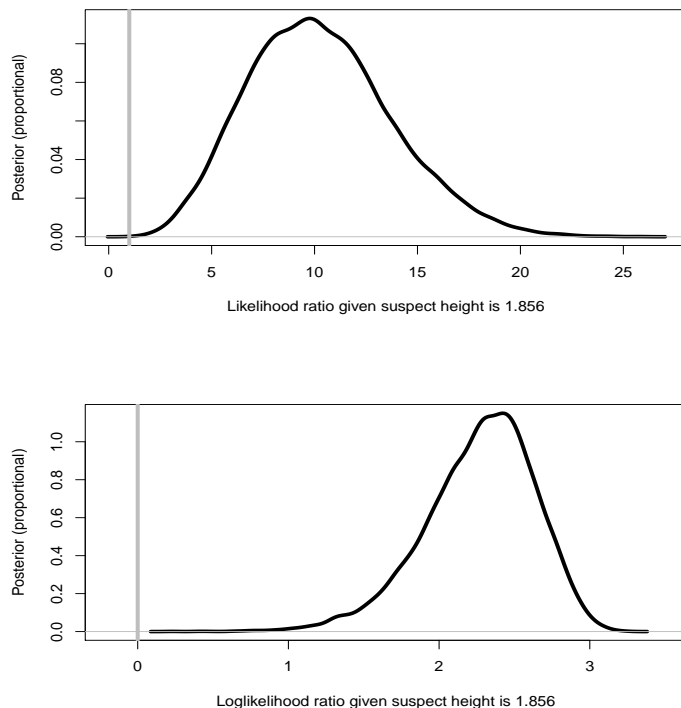
Figure 2: For the situation where the suspect has the same height as the perpetrator ($h_s = 1.856$), the posterior distribution of the likelihood ratio (top panel) and the posterior distribution of the logarithm of the likelihood ratio (bottom panel). Reference values 1 and 0 indicated by grey lines.

as the perpetrator, we would expect the likelihood ratio value 1 to be located in the left tail of the density of LR because it is likely that the suspect is the perpetrator and hence the mean of LR should be larger than 1. In other words, $P(LR < 1)$ should be small. For the same reason, we would expect BF to be larger than 1. This is indeed the case, the posterior median of LR is 10.03 with 95% CI $(4.25, 17.69)$, and the BF is estimated at 10.28 . The big advantage of the posterior of LR is clear: we have estimated its distribution. The 95% CI for instance shows immediately that $LR = 1$ is not very likely. This is information

that the computation of BF does not provide. The evidence is clearly in favour of the prosecution's hypothesis.

Next we consider two values of h_s that are clearly in favour of the defence's hypothesis. Both values $h_s = 1.7$ and $h_s = 2.0$ yield a posterior median of LR smaller than 0.01 and $P(LR < 1) > 0.99$. The corresponding BF are both smaller than 0.001. For both theses values of the suspect's height, the evidence is clearly in favour of the defence's hypothesis.

The value $h_s = 1.825$ illustrates a situation where the extra information of the posterior of the LR is of particular use. The BF is estimated at 0.53. This is dissimilar to the posterior median 0.15 of the LR, whereas the posterior mean 0.521 of the LR is close to the BF. Where the BF gives no uncertainty information, the sampled values of the LR allow many possible quantities to be estimated to assess whether the evidence is in favour of the defence's hypothesis. The latter is not the case. The 95% CI for the LR is $(< 0.01, 2.86)$ which includes the value 1. Probability $P(LR < 1)$ is estimated at 0.82.

For the value $h_s = 1.825$, we investigate the sensitivity of the results with regard to the specification of the prior $p(\boldsymbol{\theta}) = p(\alpha, \log(\sigma))$. First, we use priors which are less informative. We specify $\alpha \sim N(0, 1)$ and $\log(\sigma) \sim U(-10, 5)$. Given that measurements are in meters, these priors do not contain much information. For the LR, we obtain median 0.150 and 95% CI $(< 0.01, 3.03)$, the BF is estimated at 0.54. Next we specify $\alpha \sim N(0, 0.05)$ and $\log(\sigma) \sim U(-10, -3)$. The prior for α implies that about 95% of the systematic error falls with the interval (-10cm, 10cm), the prior for σ implies that σ is less than 10cm. These priors are informative, but are still reasonable for this case. For the LR, we obtain median 0.145 and CI $(< 0.01, 2.80)$, the BF is estimated at 0.51. Given these alternative specifications of the priors, results are very similar to the previous results.

4 Conclusion

A fully Bayesian evaluation of evidence requires the computation of a Bayes factor. For complex models, this factor may be hard to compute. Using the ideas in Dempster [8] and Aitkin [3], the posterior distribution of the likelihood ratio is used in a forensic science setting as an alternative to the Bayes factor. Using the posterior likelihood ratio is not frequentist as sampling from a posterior is required, but it is also not fully Bayesian since it does not use the Bayes factor for hypothesis testing.

The application discussed forensic data where heights were estimated on the bases of images from a security camera. The posterior mean of the likelihood ratio was similar to the Bayes factor. With samples available from the posterior of the likelihood ratio, an all-round inference was possible by investigating posterior percentiles and credible intervals.

Gelman *et al.* [6] criticise the posterior likelihood ratio approach by arguing that it is incompatible with a Bayesian perspective, and that it does not seem to be useful for common applications in statistics. We hope to have shown in this paper that forensic science is an area where the approach seems useful. The points raised by Gelman *et al.* with respect to using vague priors, comparing discrete hypotheses, and the problem with product of posteriors, are not applicable in our setting: In forensic science, it make sense to use vague prior densities for the parameters in the model for the background data, researchers are interested in comparing discrete hypotheses, and – at least in the current application – there is no assessment of a product of posteriors.

Nevertheless, we acknowledge that there are still important issues in the posterior likelihood ratio approach that need further attention. Using the posterior distribution of LR for evidence evaluation can be seen as a hybrid of Bayesian and frequentist methods. It is not fully Bayesian, but it is also not a frequentist

analysis. This ambiguity causes interpretation problems. For example, in a fully Bayesian framework, a 95% credible interval of a parameter means that the posterior probability that the parameter lies in that interval is 0.95. A frequentist 95% confidence interval means that given a large number of repeated samples, 95% of the estimated confidence intervals includes the true value of the parameter. What are the properties of the credible intervals for LR that we computed in the current application?

In this paper the situation is considered in which there is only one piece of evidence. If there is more than one piece of evidence, a posterior distribution may be determined of the LR of the combination of the evidence. This topic may be explored elsewhere.

Using the posterior likelihood ratio has a wide range of possible applications in forensic practice. Computationally it is a feasible method to evaluate evidence. It takes into account the uncertainty with regard to inference from background and at the same time allows to model prior knowledge.

Appendix

WinBUGS code used in the evaluation of evidence. For more information on the software and MCMC sampling see `www.mrc-bsu.cam.ac.uk/bugs`.

```
## Data:
list(h = c(1.950, 1.795, 1.865, 1.755 ,1.910, 1.825) ,
     m = c(1.964, 1.832, 1.900, 1.780, 1.937, 1.865))

## Inits:
list(alpha=0,    logsigma= -4)
list(alpha=0.02, logsigma= -5)

## Model:
model{
 # Model for measurement:
 for(i in 1:6){
   mu[i] <- h[i]+alpha
   m[i] ~ dnorm(mu[i], tau)
```

```
}
# Evaluation of evidence:
h_s <- 1.825
m.p <- 1.885
# Under H_p:
pi   <- 3.141593
p_Hp <- 1/(sqrt(2*pi)*sigma)*exp(-1/2*tau*pow(m.p-(h_s+alpha),2))
# Under H_d:
mu_h.pop  <- 1.806
var_h.pop <- 0.01
tau_h.pop <- 1/var_h.pop
p_Hd <- 1/sqrt(2*pi*(var+var_h.pop) )*exp(-1/(2*(var+var_h.pop))
       *pow(m.p-alpha-mu_h.pop,2))
# LR:
LR <- p_Hp/p_Hd
# Strength of evidence:
c   <- 1
pprob <- step(c-LR)
# Converting precision to sd and var:
tau <- pow(sigma,-2)
var <- pow(tau,-1)
# Priors:
alpha ~ dnorm(0,0.1)
logsigma ~ dunif(-10,0)
sigma <- exp(logsigma)
}
```

References

[1] D. Lucy (2005). *Introduction to Statistics for Forensic Scientists*, Wiley, Chichester, UK.

[2] D. Lunn, C. Jackson, N. Best, A. Thomas, D. Spiegelhalter (2013) *The BUGS Book*, Chapman & Hall/CRC, Boca Raton, FL.

[3] M. Aitkin (2010). *Statistical Inference. An Integrated Bayesian/Likelihood Approach*, Chapman & Hall/CRC, Boca Raton, FL.

[4] G. Edelman, I. Alberink, B. Hoogeboom (2010). Comparison of the performance of two methods for height estimation, *Journal of Forensic Science*, *55*, 358–365.

[5] A. Van den Hout, A., I. Alberink (2010). A hierarchical model for body height estimation in images, *Forensic Science International, 197*, 48-53.

[6] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin (2004), *Bayesian data analysis*, Chapman & Hall, London.

[7] B.P. Carlin, T.A. Louis (2009). *Bayesian Methods and Data Analysis. Third Edition*, Chapman & Hall, London.

[8] A.P. Dempster (1974). *The direct use of likelihood for significance testing*, in O. Barndorff-Nielsen, P. Bleasild and G. Sihon (eds.), Proc. Conf. Foundational Questions in Statistical Inference, University of Aarhus, 335–352.