Skeptical and Optimistic Robust Priors for Clinical Trials

John Cook.*
Jairo Fúquene. †
Luis Pericchi Guerra.‡

September 6, 2010

Abstract

A useful technique from the subjective Bayesian viewpoint, suggested by Spiegelhalter et al. (1994), is to ask the subject matter researchers and other parties involved, such as pharmaceutical companies and regulatory bodies, for reasonable optimistic and pessimistic priors regarding the effectiveness of a new treatment. Up to now, the proposed skeptical and optimistic priors have been limited to conjugate priors, though there is no need for this limitation. The same reasonably adversarial points of view can be taken with robust priors. A recent reference with robust priors usefully applied to clinical trials is in Fuquene, Cook, and Pericchi (2009). Our proposal in this paper is to use Cauchy and intrinsic robust priors for both skeptical and optimistic priors leading to results more closely related with the sampling data when prior and data are in conflict. In other words, the use of robust priors removes the dogmatism implicit in conjugate priors.

Keywords: *Clinical Trials, Skeptical and Optimistic Priors, Robust Priors.*

1 Introduction

Clinical trials are contentious. Pharmaceutical companies are eager to show that their new drug, on which they may have been invested millions of dollars, is a substantial improvement over the current treatment. On the other hand, government regulatory agencies take the opposite view and ask for substantial evidence that the new drug is not less effective than the current standard.

Bayesian statistics provides a useful technique for modeling adversarial positions by using *two* prior distributions on the parameters of interest. One, the optimistic prior, corresponds to positive expectations. The other, the pessimistic prior, corresponds to a more skeptical position. To be more specific, let us consider the following one-sided hypothesis test:

$$H_0 : \theta \leq 0, \ vs \ H_1 : \theta > 0. \tag{1}$$

*Division of Quantitative Sciences, M.D. Anderson Cancer Center, University of Texas, Houston, TX, jdcook@mdanderson.org

†University of Puerto Rico, Rio Piedras. Institute of Statistics, School of Business Administration. jairo.a.fuquene@uprrp.edu, supported by Institute of Statistics, School of Business Administration, UPR-RRP.

‡University of Puerto Rico, Rio Piedras. Department of Mathematics. luarpr@gmail.com, sponsored in part by NIH Grant: P20-RR016470

Part of the clever proposal by Spiegelhalter *et al* (2004) is to assume two priors, π_S and π_O, such that

$$P_S(H_0) = \int_0^{\infty} \pi_S(\theta)d\theta = 0.05, \text{ and, } P_O(H_0) = \int_0^{\infty} \pi_O(\theta)d\theta = 0.95. \quad (2)$$

In words, the skeptical position (say the government regulatory body attitude) is to give only 5% to the improvement (i.e H_1), but the optimistic position gives it as much as 95% probability to the improvement (the investigator position). If under both priors, the posterior probability of H_1 is bigger than 95%, that is if under the skeptical prior $P(H_1|\text{ Data }, \pi_S) > 0.95$, then it is safe to decide in favor of H_1. On the other hand, if both priors conclude that there is no improvement, that is $P(H_0|\text{ Data }, \pi_O) > 0.95$, it is safe to decide in favor of H_O.

This approach is intuitively satisfying. However, this framework may be overly cautious lead to an enormous delay in the decisions. The implementation proposed in Spiegelhalter *et al* (2004) uses conjugate priors, which lead to simple computations. However, we show in this article that we may preserve the useful framework of skeptical and optimistic priors without the relative dogmatism inherent in conjugate priors.

Pericchi an Smith (1992) showed some aspects of the robustness of the Student-t prior for a normal location parameter and provided approximations to the posterior moments in the model Student-t/normal. The Cauchy prior, as a Student-t with one degree of freedom, can be used in this context as well. However, for normal log-odds there is a robust prior proposed by J. O. Berger (Berger 1985) that leads to a closed form posterior and moments, a sort of the "best of both worlds." On the other hand, the "intrinsic prior" was obtained in (Berger and Pericchi 1996) as the implicit prior to which the arithmetic intrinsic Bayes factor converges and is a limiting case of Berger's prior.

2 Illustration: Skeptical and Optimistic Robust Priors

A useful suggestion under the *subjective Bayesian viewpoint*, taken by Spiegelhalter et al. (1994), is to ask the subject matter researchers, for reasonably optimistic and pessimistic priors (regarding the effectiveness of a new treatment). On the log-odds scale, a skeptical prior on the amount of improvement has mean zero (i.e. no difference between treatments, $\lambda = 0$) and a substantial probability that the new treatment is not better. The prior scale is assessed in reference to an optimistic hypothesis λ_H. Then a small probability ξ, is assessed, for example $\xi = 0.05$, that the effect of the treatment is equal or better than λ_H. For the Cauchy prior, the skeptical parameters are very easy to assess. The location μ is zero and the scale $\beta = \lambda_H/\tan(\pi(\xi - 1/2))$. In Fuquene, Cook and Pericchi (2009) other robust heavy tailed priors are considered, apart from the Cauchy, and several mathematical results are presented. Two such priors are the Intrinsic Prior, discovered by Berger and Pericchi (1996) and the so called Berger's prior, Berger (1985). There is an unexpected and close relationship between the Intrinsic and Berger's prior, and for the most part we present the results for the Intrinsic prior, for which the location is zero and the scale, τ, is found by solving the following equation

$$\int_{-\infty}^{\lambda_H} \frac{1}{2\sqrt{\pi}} \frac{1 - \exp\left[-\frac{(\lambda-\mu)^2}{\tau^2}\right]}{(\lambda-\mu)^2/\tau} d\lambda = 1 - \xi. \tag{3}$$

Example: The following example is adapted from Spiegelhalter, Abrams, and Myles (2004), Wiley, page 69, of normal log odds coupled with a normal sceptical prior (with the current data). Alternatively, we assume heavy-tailed Cauchy and intrinsic priors, with the same location.

Suppose the normal, Cauchy, and intrinsic sceptical priors with mean log(OR)=0 and with a 95% interval running from 50% reduction in odds of death (log(OR) = -0.69) to a 100% increase (log(OR) = 0.69). On a log(OR) scale, this prior has a 95% interval from -0.69 to 0.69 and so has a standard deviation $0.69/1.96 = 0.35$ (for $\xi = 0.025$) and hence the number prior of observations is $n_0 = 4/0.35^2 = 32.3$. For Cauchy and intrinsic priors the scale is 0.05 and 0.06 respectively. We use the normal approximation for binary data for the log-odds with the approximate standard error recommended in Spiegelhalter, Abrams, and Myles (2004) for 2×2 tables, following their suggestion of an standard error of the likelihood normal and normal/normal posterior model equal to $\sigma = 2$.

Suppose that the evidence from study about 30-day mortality was 26/97 on control and 13/193 on new treatment. If the ratio of the odds of death following the new treatment to the odds of death on the conventional: OR< 1 therefore favors the new treatment. We have that the estimated log(OR) is $\bar{X}_n = -1.6$ (OR=0.2 or 80% risk reduction) with estimated standard error 0.36 and $n = 4/0.36^2 = 30.5$ approximately the same weight of the prior. We use the R (R Development Core Team 2009) package named ClinicalRobustPriors, available from the Comprehensive R Archive Network at http://CRAN.R-project.org/package=ClinicalRobustPriors, which can be used to compute probabilities and figures for the prior, likelihood and posterior models. The posterior mean in the normal/normal model is $(n_0\mu + n\bar{X}_n)/(n_0 + n) = -0.77$ with standard deviation $\sigma/\sqrt{n_0 + n} = 0.25$, the estimated odds ratio is $e^{-0.77} = 0.46$ or 54% risk reduction. In the Cauchy/normal and intrinsic/normal models the posterior mean is -1.38 ($e^{-1.48} = 0.22$ or 78% risk reduction) with standard deviation 0.32.

In the normal/normal an 95% credible interval on the log(OR) scale is between -1.27 and -0.28 that corresponds to odds ratios from 0.28 to 0.75, or a 95% probability that the true risk reduction lies between 25% and 72%. For the Cauchy/normal and intrinsic/normal posterior models the 95% credible interval show that the true risk reduction lies between 57% and 88%. On the other hand, the likelihood shows a risk reduction between 60% and 90%. In Figure 1 we can see that dispite the large discrepancy between the prior and data, with the Cauchy and intrinsic priors the posterior distributions are very similar than the current data. In other words, the normal sceptical prior is more dogmatic than the Cauchy and intrinsic skeptical priors.

As a counterbalance to the skeptical priors, Spiegelhalter et al. (1994) suggest an "enthusiastic" (or "optimistic" as we call it here) prior centered on the alternative hypothesis and with low chance that the true treatment benefit negative. In this example the alternative hypothesis is $\lambda_H = -0.69$ (50% risk reduction) and the low chance is $\xi = 0.025$.

The scale and prior sample size are the same as with the skeptical priors. Figure 2 display the results. The posterior mean in the normal/normal conjugate model is -1.13 ($e^{-1.13} = 0.32$ or 68% risk reduction), much closer than the -1.39 ($e^{-1.39} = 0.24$ or 76% risk reduction) of the Cauchy/normal or intrinsic/normal models. The scale for the posterior models are 0.25 and 0.34 respectively with normal and Cauchy (i.e intrinsic) priors. However, with the Cauchy and intrinsic optimistic robust priors the results are more closer to the likelihood.

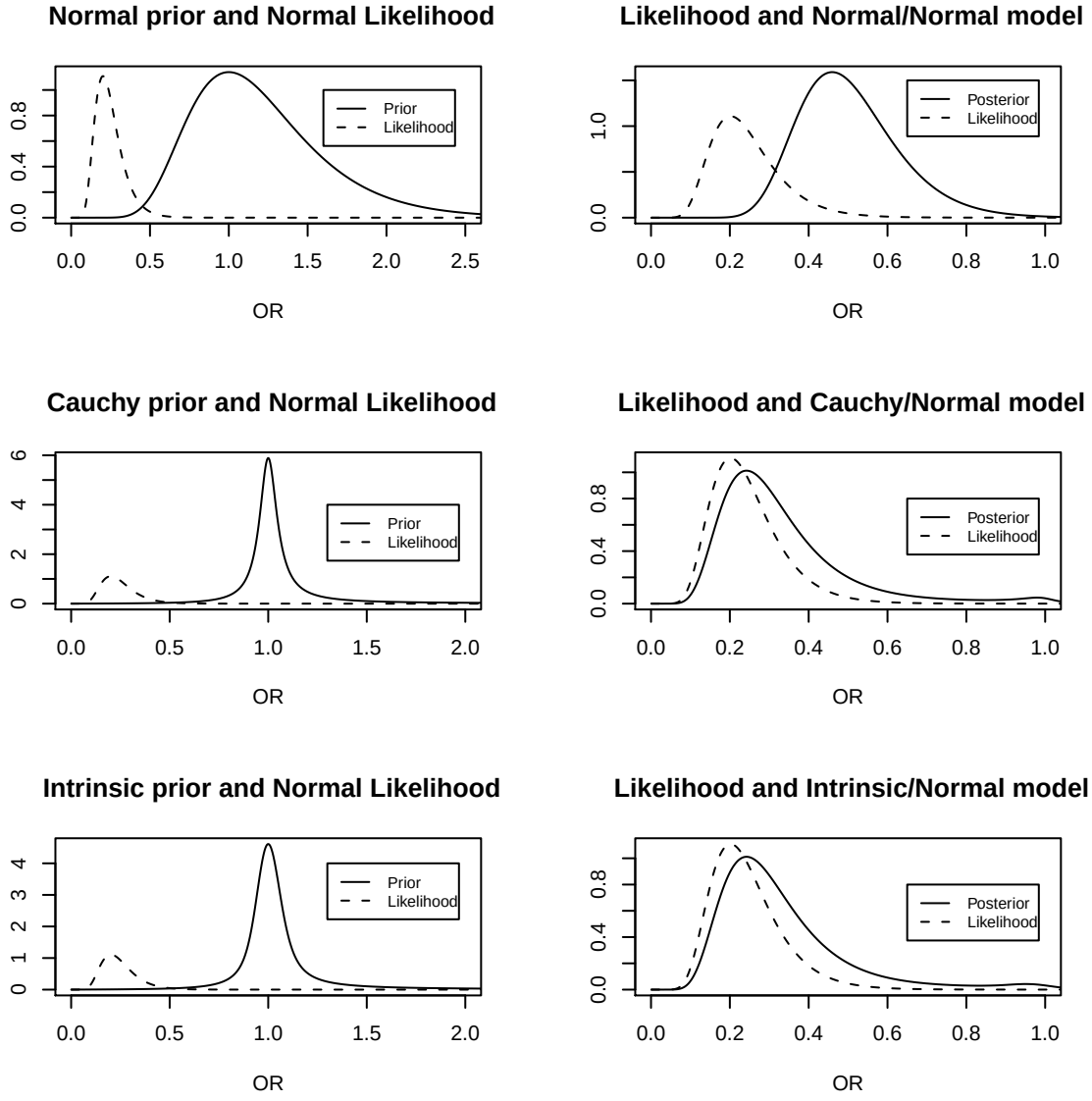

Figure 1: Skeptical priors, likelihoods and posterior models: normal/normal, cauchy/normal and intrinsic/normal.

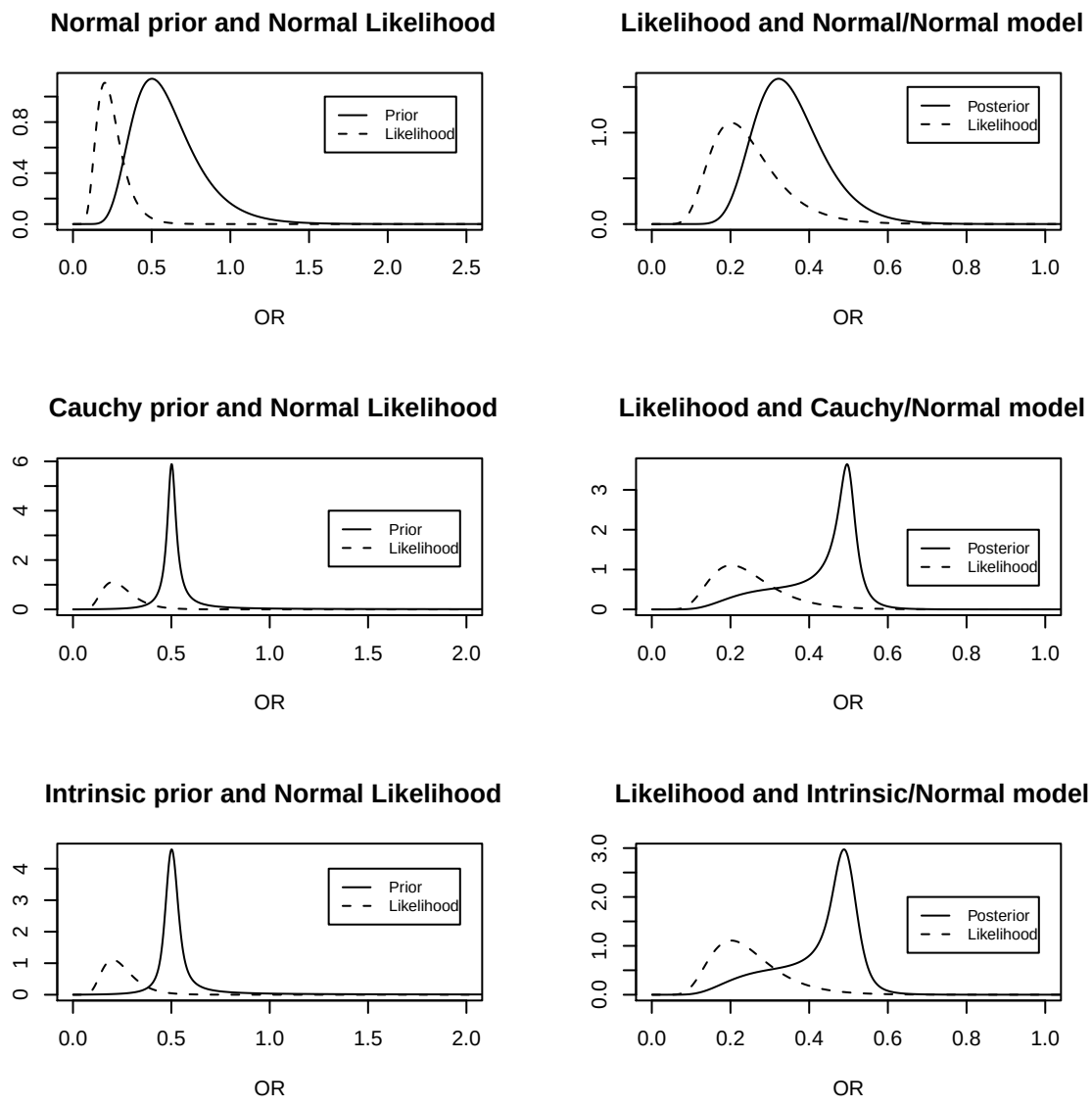

Figure 2: Optimistic priors, likelihoods and posterior models: normal/normal, cauchy/normal and intrinsic/normal.

3 Asymptotic Results

In this section we show that robust priors can "change their mind" more readily than conjugate priors by looking at asymptotic properties of posterior means under each type of prior. When both the skeptical and optimistic prior are robust, they may reach agreement more quickly than if the two priors were not robust.

Consider a single sample y from a normal(θ, σ^2) distribution where θ has a conjugate normal(0, τ^2) prior. It is well-known that the posterior distribution on θ has mean

$$\frac{\tau^2}{\tau^2 + \sigma^2} y.$$

If we use a Cauchy(0, 1) prior rather than a normal(0, τ^2) prior on θ above, the posterior mean of θ is

$$y - \mathcal{O}\left(\frac{1}{y}\right)$$

as $y \to \infty$. See Cook (2010) for the derivation of this asymptotic result.

Under the normal prior, the posterior mean of θ under-estimates the data y by a constant ratio which depends on the relative scales of the sampling distribution and the prior distribution. Under the Cauchy prior, however, the posterior mean of θ asymptotically approaches the value of y, independent of the scales of the sampling and prior distributions. This is illustrated by Figure 3.

Next we consider the case of n samples y_i with mean $\overline{y}$ and consider the behavior of the posterior mean as $n \to \infty$. Under a normal(0, τ^2) prior, the posterior mean is given by

$$\frac{1}{1 + \frac{\sigma^2}{n\tau^2}} \overline{y} = \left(1 - \frac{\sigma^2}{n\tau^2} + \mathcal{O}\left(\frac{1}{n^2}\right)\right) \overline{y}.$$

Under a Cauchy prior, however, the posterior mean is

$$\overline{y} + \frac{(\overline{y}^2 - 3)\overline{y}}{(1 + \overline{y}^2)^2} \frac{\sigma^2}{n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

The rate at which the posterior mean converges to $\overline{y}$ depends on τ in the case of the normal prior and on $\overline{y}$ in the case of the Cauchy prior. For any value of τ, the convergence is faster under the Cauchy prior for sufficiently large values of $\overline{y}$.

We now obtain results for Berger's prior that are quite similar to those above for the Cauchy prior. If we observe n samples from a normal(θ, σ^2) distribution and θ has Berger's prior with location 0, the posterior mean of θ is

$$\frac{2\sigma^2\overline{y}}{(\sigma^2 + n\beta^2)\left(\exp\left(\frac{n\overline{y}}{\sigma^2 + n\beta^2}\right) - 1\right)} - \frac{2\sigma^2}{n\overline{y}}$$

where $\overline{y}$ is the sample mean of the observations as shown in Fúquene, Cook, and Pericchi (2009). Therefore as a single sample $y \to \infty$, the posterior mean of θ is

$$y - \frac{2\sigma^2}{y} + \mathcal{O}(\exp(-y^2)).$$

Also, as $n \to \infty$, the posterior mean of θ is

$$\overline{y} - \frac{\sigma^2}{n} \frac{2}{\overline{y}} + \mathcal{O}(\exp(-n)).$$

These results show that for Cauchy and Berger priors, the influence of the prior diminishes more quickly when the data are far from the location parameter of the prior. When skeptical and optimistic priors using either of these robust distributions, the priors will reach consensus more quickly than corresponding conjugate priors when the data are in conflict with one or other of the priors.

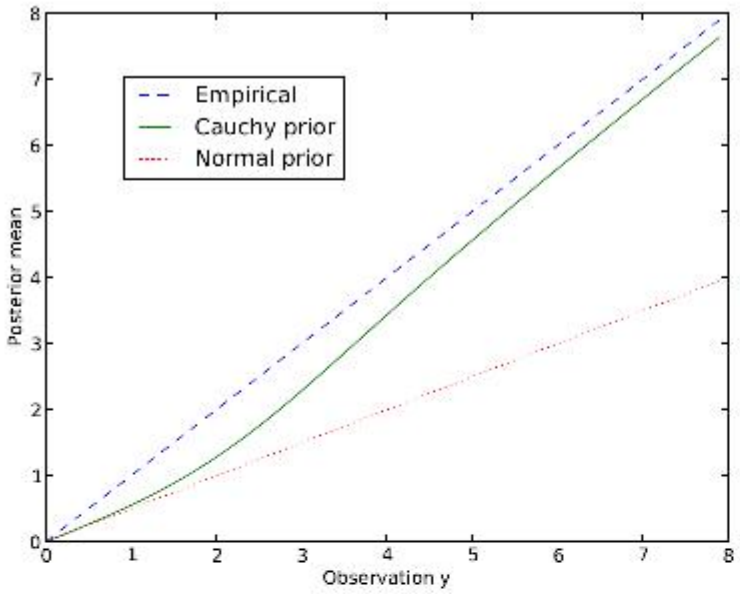

Figure 3: Posterior mean of θ under robust and conjugate priors compared to empirical value

References

[1] Berger, J. O. (1985), Statistical Decision Theory and Bayesian Analysis, second edition, Springer-Verlag.

[2] Berger, J. O. and Pericchi, L. R. (1996). "The Intrinsic Bayes Factor for Model Selection and Prediction." JASA, 91: 112-115.

[3] Cook, J. D. (2010) "Asymptotic results for Normal-Cauchy model." UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series. Working Paper 61. http://www.bepress.com/mdandersonbiostat/paper61

[4] Fúquene, J., Cook J. and Pericchi L.R. (2009) "A Case for Robust Bayesian Priors with Applications to Clinical Trials", *Bayesian Analysis*, 4, Number 4, pp. 817-846.

[5] Pericchi, L. R. and Smith, A. F. M. (1992). "Exact and Approximate Posterior Moments for a Normal Location Parameter." JRSSB, 54: 793-804.

[6] Spiegelhalter, D. J., Abrams, K. R., and Myles, J. P. (2004). Bayesian Approaches to Clinical Trials and Health-Care Evaluation. London: Wiley.

[7] Spiegelhalter, D. J., Freedman, L. S. and Parmar, M. K. B. (1994), "Bayesian approaches to randomized trials (with discussion)", Journal of the Royal Statistical Society, 157: 357-416.