

Bayesian regions of evidence

(for normal distributions)

OISP meeting, May 12th, 2021

Michael Höfler
Robert Miller

Open Science issue of priors in Bayesian analysis:

- Imagine you submit paper with a Bayesian analysis and a **specific (pre-chosen!) prior**
- (because of the benefits of Bayes: probability that hypothesis is true, more stable estimates ...)
- **But a reviewer wants you to use another prior!**
- But this can not be a real prior!

- **Dilemma:** stick with your prior vs. satisfy the reviewer

We suggest a solution that builds on “reverse Bayesian analysis”

Bayesian analysis usually **looks forward**:

Choose **prior**

Collect data

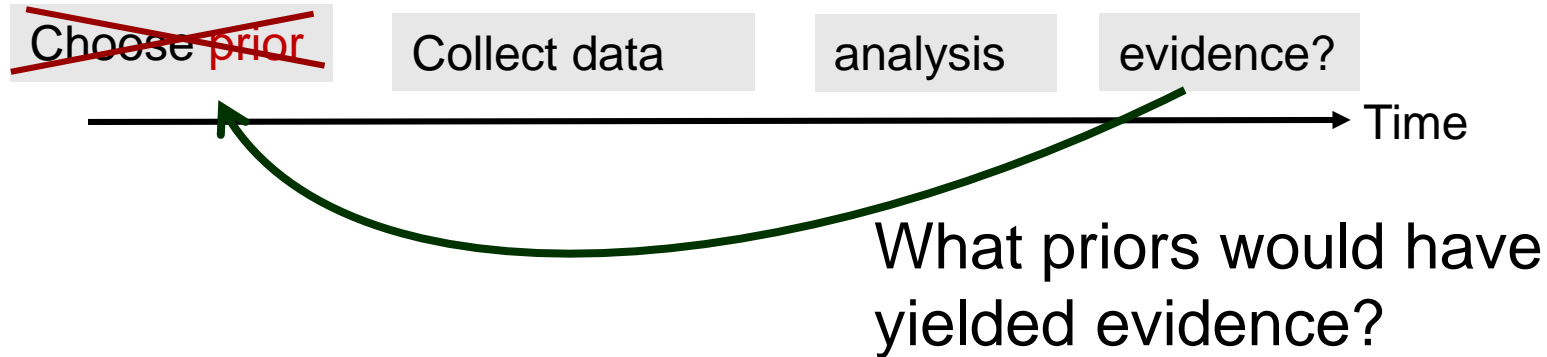
analysis

evidence?

Time

- Criticized for being **subjective** through choosing the **prior**
- (But a specific prior may be more defensible than the frequentist “I know nothing before seeing the data“)

Reverse Bayes looks backward



- Avoids being subjective
- By e.g. asking: what is the **most pessimistic prior** that allows concluding $effect > 0$ (Δ)?
- This is easy if both the prior and the data contribution are **normally distributed** (see appendix)

Origin and use of reverse Bayes

Good (1950):

PROBABILITY AND THE WEIGHING OF EVIDENCE

By

I. J. GOOD, M.A., PH.D.
FORMER LECTURER IN MATHEMATICS
AT THE UNIVERSITY OF MANCHESTER

Instead of regarding this case as being included under heading (iii) it may be more convenient to make direct judgments about the initial distribution of the chance. For example, if this distribution is uniform the sample frequency, m/n , is the “most probable value” † of the chance. (Whatever the initial distribution the sample frequency is the maximum likelihood value of the chance.)

If n were large, adherents of the frequency approach would *judge* that the chance x was approximately m/n . (They would not usually judge that the proportional accuracy was good if m was small.) If they would define the degree

† See the index.

80

STATISTICS AND PROBABILITY

7.3

of the approximation then Bayes' theorem (in **reverse**) could be used for obtaining information about the initial probability distribution of the chance.

(v) Finally, suppose that N is unknown. As before you can use judgments about the initial distribution of the chance. (Or you could work with the distribution of N and the distribution of M for each N .)



Preventive Medicine

Volume 53, Issues 4-5, October–November 2011, Pages 225-228



Null misinterpretation in statistical testing and its
impact on health risk assessment

Sander Greenland



Preventive Medicine

Volume 53, Issues 4-5, October–November 2011, Pages 225-228



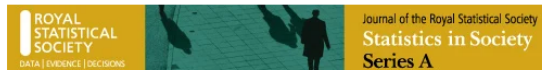
Null misinterpretation in statistical testing and its
impact on health risk assessment

Sander Greenland

Special case: “sceptical prior”, prior expectation $e_o = 0$: “no effect”

With normal prior only question left: what prior standard deviation sd_o ?
How strong pre-belief in values around 0?

„Sufficiently sceptical prior“ :
How sceptical may it be? = **lowest SD** that allows for inferring
effect > 0 (Δ)



Original Article | [Open Access](#) |

A new standard for the analysis and design of replication studies

Leonhard Held

First published: 26 December 2019 | <https://doi.org/10.1111/rssa.12493> | Citations: 6

We first need to compute the variance of the sufficiently sceptical prior. Matthews (2001a) has shown that the equitailed credible interval of the sufficiently sceptical prior at level $1-\alpha$ has limits $\pm S$ where

$$S = \frac{(U - L)^2}{4\sqrt{(UL)}}$$

is the *scepticism limit* (Matthews, 2018). Note that equation 1 holds for any value of α , not just for the traditional 5% level. The sufficiently sceptical prior variance τ^2 can be derived from equation 1 and expressed as a function of the variance σ_o^2 (the squared standard error, which is assumed to be known) of the estimate $\hat{\theta}_o$, the corresponding test statistic $t_o = \hat{\theta}_o/\sigma_o$ and $z_{\alpha/2}$, the $(1-\alpha/2)$ -quantile of the standard normal distribution (Held (2019), appendix):

$$\tau^2 = \frac{\sigma_o^2}{t_o^2/z_{\alpha/2}^2 - 1},$$

Own approach: “Regions on evidence“

- We want to extend that by presenting **all** such priors
- Presenting all possibilities also avoids being subjective!
- Covers all possible priors like our own and the one of the reviewer
- → Who would conclude an effect and who would not?

Hypothetical example: **effect** of intelligence training on measured IQ

- Two randomly assigned groups
- training versus no training, n = 30 each
- Choose $\Delta = 3.75$ (= SD/4 on the IQ scale)
- In a sample those with training score 6.0 IQ points higher

```
. tabstat IQ , stat(n mean sd) by(x)
```

```
Summary for variables: IQ  
by categories of: x
```

| x | N | mean | sd |
|-------------|----|----------|----------|
| no training | 30 | 101.0445 | 16.33471 |
| training | 30 | 107.0169 | 14.17013 |
| Total | 60 | 104.0307 | 15.45682 |

- In the example the estimate of effect \approx normally distributed with expectation 6.0 and variance $3.9 \cdot 3.9 = 15.6$, $\sim N(6.0, 15.6)$

. regress y x

| Source | SS | df | MS | Number of obs | = | 60 |
|----------|------------|----|------------|---------------|---|--------|
| Model | 535.046963 | 1 | 535.046963 | F(1, 58) | = | 2.29 |
| Residual | 13560.8451 | 58 | 233.807674 | Prob > F | = | 0.1358 |
| Total | 14095.8921 | 59 | 238.913425 | R-squared | = | 0.0380 |
| | | | | Adj R-squared | = | 0.0214 |
| | | | | Root MSE | = | 15.291 |

| y | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|----------|-----------|-------|-------|----------------------|
| x | 5.97242 | 3.94806 | 1.51 | 0.136 | -1.930481 13.87532 |
| _cons | 101.0445 | 2.7017 | 36.19 | 0.000 | 95.45627 106.6327 |

- Einseitiger p-Wert = .068, also nach üblichem frequentistischen Standard **keine Evidenz für effect > 0**
- (A bit confusing: standard error = standard deviation of the estimate's distribution)

With normal prior and a normally distributed estimate, we easily get the posterior distribution:

The posterior then is also a normal distribution:

$$(*) \text{ effect posterior} \sim N(e_p, sd_p)$$

$$\text{With } e_p = \frac{e_0/sd_0^2 + e_{\text{obs}}/sd_{\text{obs}}^2}{\frac{1}{sd_0^2} + \frac{1}{sd_{\text{obs}}^2}}$$

$$\text{And } sd_p^2 = \frac{1}{\frac{1}{sd_0^2} + \frac{1}{sd_{\text{obs}}^2}}$$

inverse-variance weighted
average of prior and observed effect
the higher the precision (inverse variance, $\frac{1}{sd_p^2}$)
in the prior = **the stronger the prior belief**, the more is e_{obs}
shrunk toward e_0

Bayesian perspectives for epidemiological research:
I. Foundations and basic methods 

Sander Greenland

International Journal of Epidemiology, Volume 35, Issue 3, June 2006, Pages 765–775,

<https://doi.org/10.1093/ije/dyi312>



Regions of evidence

Now use the formula * $\text{effect posterior} \sim N(e_p, sd_p)$
to determine relevant **regions of evidence** and
display with which prior **one may conclude that:**

$$\text{effect} > 0$$

$$\text{effect} > \Delta$$

$$|\text{effect}| < \Delta$$

E.g. show that $effect > 0$ (Δ)

Calculate z-statistics $\mathbf{z_p = e_p / sd_p}$

usual idea: divide an estimate by it's standard error. Technically, the standard error of an estimate is the standard deviation of the estimate's distribution.

Probability that $effect > 0 = \Phi(\mathbf{z_p})$
 $> \Delta = \Phi(\mathbf{e_p - \Delta / sd_p})$
... must be $> 1 - \alpha$ (equation *)

Φ = cumulative of the standard normal distribution.

For the common $\alpha = .05$, $\Phi^{-1}(1 - .05) = 1.64$

Unfortunately, the formula * is complicated to solve, but it is computational easy

For any e_{obs} and sd_{obs} , simulate 1m data points,
determine for each point whether it is inside or outside

```
// simulate
clear
set obs 500000
gen x=uniform()*15-5
gen y=uniform()*10

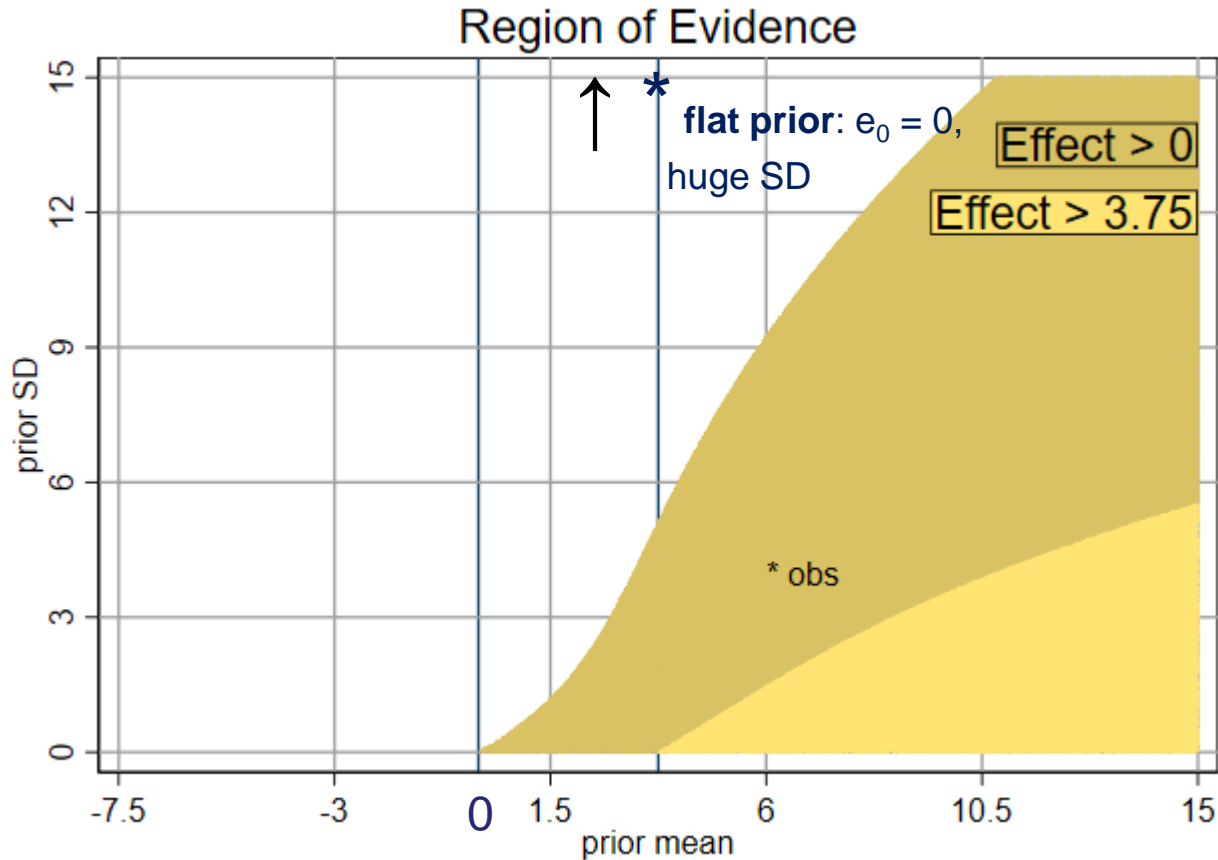
// Posterior mean
gen ep= ( x/y^2 + `e1'/v1) / (1/y^2 + 1/v1)

// Posterior SD
gen sdp = sqrt(1/ (1/y^2 + 1/v1))

// posterior z for delta=0
gen z0=ep/sdp
// posterior z for delta=3.75
gen z375=(ep-3.75)/sdp

// evidence, binary
gen ev0 = (z0 > `phi')
gen ev375= (z375 > `phi')
```

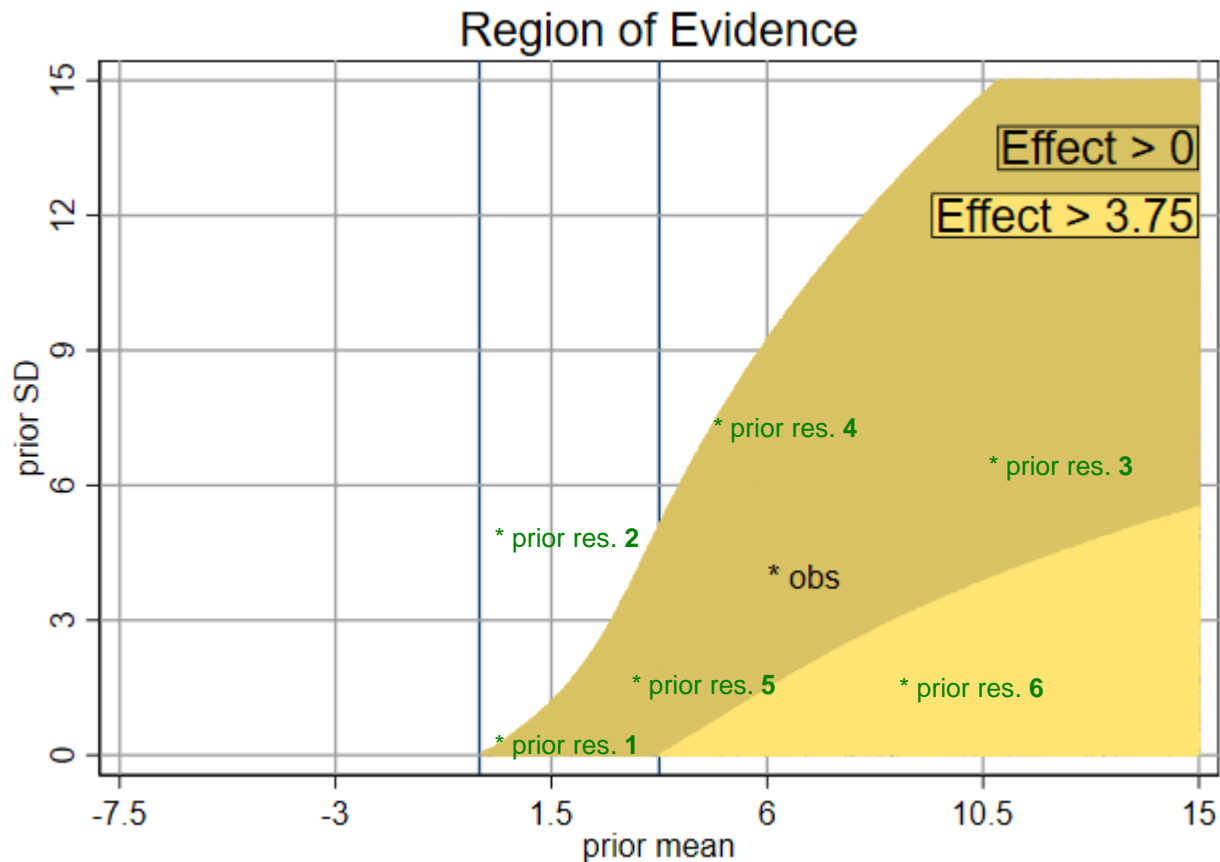
In the example $* = e_{\text{obs}} = 6.0$, $\text{sd}_{\text{obs}} = 3.9$, $\Delta = 3.75$



* **obs** determines the entire graph!

Every „sceptical prior“ ($e_0=0$) is outside

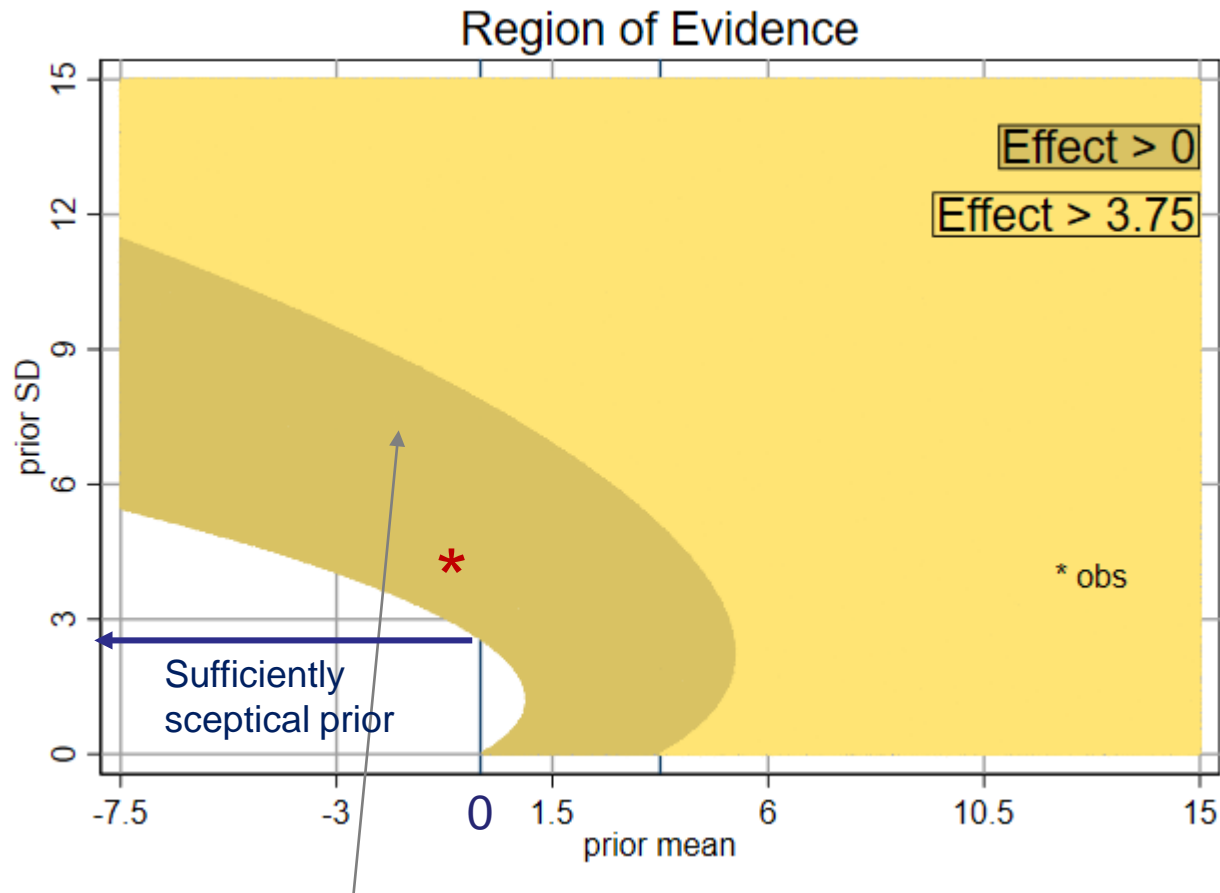
Integration of several researchers **priors**:
Who would conclude an effect after
observing the same?
Dilemma solved



To show **sceptical prior**:

What if we had observed a twice as large effect?

$$e_{\text{obs}} = 12.0, \text{sd}_{\text{obs}} = 3.9$$



In these areas we would assume a **negative effect**, $e_0 < 0$, but SD_0 is large, so the weight of this prior is small

Implementation

Self-written command `arev` in Stata: Robert Miller has written a **R** function.

```
. program define arevi
  1.
  . #delimiter;
delimiter now ;
  . syntax [anything] [using/] ,
  >
  > [IMmediate(numlist min=2 max=2)]
  > [par(namelist min=1 max=1)]
  > [ PRIor(numlist min=2 max=2)]
  > [ DELta(numlist min=1 max=1)]
  > [rope(numlist min=1 max=1)]
  > [ ALpha(numlist min=1 max=1)]
  > [ n(numlist min=1 max=1)]
  > [erange(numlist min=2 max=2)]
  > [sdmax(numlist min=1 max=1)]
  > [epoint(numlist min=1 max=1)]
  > [sdpoint(numlist min=1 max=1)]
  > [xline(numlist min=1 max=3)]
  > [conflict]
  > [prpoint]
  > [postpoint]
```

Can be used as **immediate command**; i.e., you don't need data, just enter any values for **e_{obs}** and **sd_{obs}**, maybe from a paper.

```
. arevi , immediate(5.97 3.95) prior(0 3.75) rope(3.75)
```

```
Results of AREVI, area of evidence for Reverse Bayes analysis
```

```
alpha = .05
```

```
Prior normal distribution, ~ N(expectation,SD) , with 0 and 3.75
```

```
Model-based, observed, ~ N(expectation,SD) , with 5.97 and = 3.95
```

```
Posterior distribution, ~ N(expectation,SD) , with 2.8300396 and 2.7196072
```

```
z-stat for effect > 0 = 1.040606 , one-tailed Bayesian 95% credibility interval = (1.185186, ∞)
```

```
Probability that effect > 0 = .85097078
```

```
(1,000,000 missing values generated)
```

Appendix

The easy and relevant case of

normal distributions

- **Estimates** from data have distributions (although `effect` itself is just an **unknown constant** in frequ. stats)
- „Maximum likelihood“ Estimates are often **normally distributed**
- E.g. estimates of regression coefficients in many kinds of models (any “maximum likelihood estimate“, e.g. $\ln(\beta)$ in logistic regression; at least approximately with increasing sample size)

How to choose a normal prior? “Percentile method”

- With **expectation e_0** : bet as much on **effect** $< e_0$ as you bet on **effect** $> e_0$
- **Variance**: uncertainty on **effect**, how large is the variance around the expectation e_0 ? Huge variance \rightarrow **flat prior**
- $SD = \sqrt{\text{Variance}}$ translates into intervals of belief:

- Choose the variance e.g. via the value sd_0 for which you would **bet 2:1** that the true value is within $e_0 - sd_0$ and $e_0 + sd_0$
- Or bet **95:5** that true value is within $e_0 - 2 * sd_0$ and $e_0 + 2 * sd_0$ (similar to 95% confidence interval)
- In normal distributions this interval covers $\approx 2/3$ probability mass
- Thus you get the prior $N(e_0, sd_0^2)$
- Here, we might choose $e_0 = 0$ (see below)
- Assume that huge **effect** in both directions is unlikely:
- E.g. $2 * sd_0 = 3.75 \rightarrow$ **95% interval $\approx -7.5, 7.5$** (**effect** is likely to be $\leq 1/2$ standard deviation on IQ scale)
- \rightarrow **prior = $N(0, 14.1)$**