brr: Basic usage.

Stéphane Laurent

2015-07-10

The **brr** package performs Bayesian inference on the rate ratio $\phi = \frac{\lambda}{\mu}$ in the *two Poisson samples model* given by two independent observations

$$\begin{cases} x \sim \mathcal{P}(\lambda S) \\ y \sim \mathcal{P}(\mu T) \end{cases}$$

where λ and μ are the unknown incidence rates and S and T are the known observation-opportunity sizes, or, for short, the sample sizes. S and T are also called the times at risk when they represent some durations.

The **brr** package implements the *semi-conjugate family of prior distributions*. Precisely, for positive numbers a, b, c and d, to be set by the user, the following independent prior distributions are assigned on μ and $\phi = \lambda/\mu$:

$$\mu \sim \mathcal{G}(a, b) \quad \text{and} \quad \phi \sim \frac{T+b}{S} \times \mathcal{B}'(c, d),$$

Then the joint posterior on (μ, ϕ) is given by

$$(\mu \mid \phi, x, y) \sim \mathcal{G}(a + x + y, b + \phi S + T) \quad \text{and} \quad (\phi \mid x, y) \sim \frac{T+b}{S} \times \mathcal{B}'(c + x, a + d + y).$$

In particular :

- when $a = c = 0.5$ and $b = d = 0$, the prior is the reference prior, also called the *non-informative prior*;
- when $a, b > 0$, $c = 0.5$ and $d = 0$, the prior is the semi-reference prior, also called the *semi-informative prior*, that is to say the reference prior after the arbitrary Gamma prior distribution $\mathcal{G}(a, b)$ is assigned on μ.

Setting parameters with `brr`

Use the **Brr** function to set the prior parameters, the sample sizes, and the observed counts. One can proceed step by step, for example below we start by supplying the parameters a and b of the prior Gamma distribution on μ:

```
library(brr)
model <- Brr(a=2,b=3)
summary(model)
```

```
## Type of prior distribution: semi-informative prior
##
## *Prior distribution on μ*:  Gamma(a=2,b=3)
##
## +--------+--------+--------+--------+--------+--------+
## |  mode  |  mean  |   sd   |   Q1   |   Q2   |   Q3   |
## +========+========+========+========+========+========+
## | 0.3333 | 0.6667 | 0.4714 | 0.3204 | 0.5594 | 0.8975 |
## +--------+--------+--------+--------+--------+--------+
##
## *Prior distribution on ϕ*:  Non-informative prior
##
## *Sample sizes*
##   S (treated group): not supplied yet
##   T (control group): not supplied yet
##
## *Observed counts*
##   x (treated group): not supplied yet
##   y (control group): not supplied yet
##
## *Posterior distribution on ϕ*:
##   a, b, c, d, S, T, x and y must be supplied
```

Since c and d were not supplied, **brr** automatically considers the non-informative prior on ϕ. Equivalently the same **brr** object can be defined by **Brr(a=2, b=3, c=NULL, d=NULL)** or **Brr(a=2, b=3, c=0.5, d=0)**.

The **brr** object is a function which can be used to update itself with new parameters, for example:

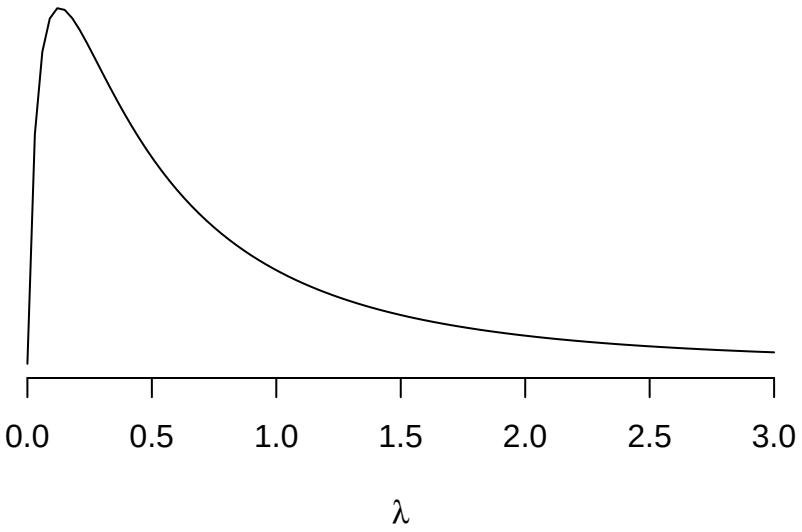
```
model <- model(c=3, d=3, S=100, T=100)
summary(model)
```

```
## Type of prior distribution: informative prior
##
## *Prior distribution on μ*:  Gamma(a=2,b=3)
##
## +--------+--------+--------+--------+--------+--------+
## |  mode  |  mean  |   sd   |   Q1   |   Q2   |   Q3   |
## +========+========+========+========+========+========+
## | 0.3333 | 0.6667 | 0.4714 | 0.3204 | 0.5594 | 0.8975 |
## +--------+--------+--------+--------+--------+--------+
##
## *Prior distribution on ϕ*:  Beta2(c=3,d=3,scale=1.03)
##
## +-------+-------+------+------+-----+------+
```

```
## |  mode   |  mean   |  sd    |  Q1    |  Q2   |  Q3    |
## +========+========+=======+=======+======+=======+
## | 0.515  | 1.545  | 1.995 | 0.578 | 1.03 | 1.836 |
## +--------+--------+-------+-------+------+-------+
##
## *Sample sizes*
##   S (treated group): 100
##   T (control group): 100
##
## *Observed counts*
##   x (treated group): not supplied yet
##   y (control group): not supplied yet
##
## *Posterior distribution on ϕ*:
##   a, b, c, d, S, T, x and y must be supplied
```

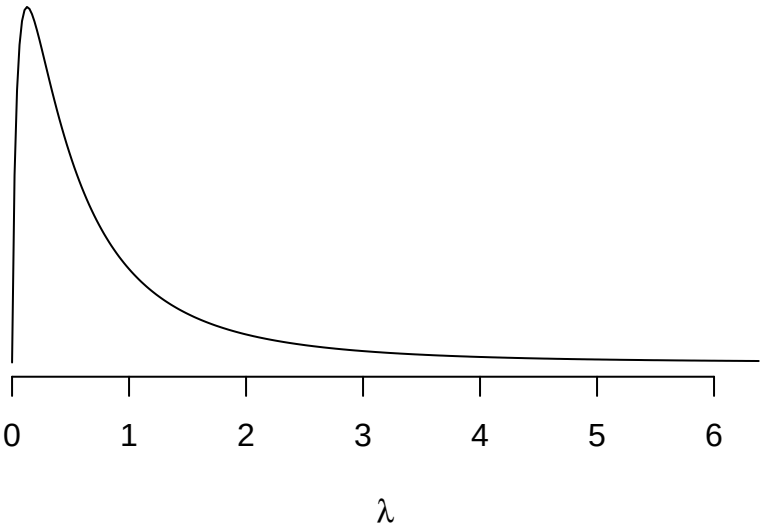
Now that a, b, c, d, S and T have been supplied, the user can play with all functions related to the prior distributions.
For example, **dprior(model, "lambda", .)** evaluates the density of the prior distribution of λ. This is convenient to generate a graphic with the **curve** function:

```
par(mar=c(4, 3, 1, 1))
curve(dprior(model, "lambda", x), from=0, to=3, axes=FALSE,
      xlab=expression(lambda), ylab=NA)
axis(1)
```



The **brr** package also provides a way to generate a plot with automatic aesthetics:

```
par(mar=c(4, 3, 1, 1))
plot(model, dprior(lambda))
```



If you are not pleased with the automatic bounds of the interval over which the curve is plotted, set your bounds in the **bounds** argument of the **plot** function. You can also set usual graphical parameters such as **lwd**, **lty**, …

Posterior inference with `brr`

Posterior inference is available after, obviously, setting the observed counts x and y:

```
model <- model(x=14, y=20)
summary(model)
```

```
## Type of prior distribution: informative prior
##
## *Prior distribution on μ*:  Gamma(a=2,b=3)
##
## +--------+--------+--------+--------+--------+--------+
## |  mode  |  mean  |   sd   |   Q1   |   Q2   |   Q3   |
## +========+========+========+========+========+========+
## | 0.3333 | 0.6667 | 0.4714 | 0.3204 | 0.5594 | 0.8975 |
## +--------+--------+--------+--------+--------+--------+
##
## *Prior distribution on ϕ*:  Beta2(c=3,d=3,scale=1.03)
##
## +--------+--------+-------+-------+------+-------+
## |  mode  |  mean  |  sd   |  Q1   |  Q2  |  Q3   |
## +========+========+=======+=======+======+=======+
## | 0.515  | 1.545  | 1.995 | 0.578 | 1.03 | 1.836 |
## +--------+--------+-------+-------+------+-------+
##
```

```
## *Sample sizes*
##   S (treated group): 100
##   T (control group): 100
##
## *Observed counts*
##   x (treated group): 14
##   y (control group): 20
##
## *Posterior distribution on ϕ*:  Beta2(17,25,scale=1.03)
##
## +--------+--------+--------+--------+-------+--------+
## |  mode  |  mean  |   sd   |   Q1   |  Q2   |   Q3   |
## +========+========+========+========+=======+========+
## | 0.6338 | 0.7296 | 0.2363 | 0.5613 | 0.696 | 0.8605 |
## +--------+--------+--------+--------+-------+--------+
##
##  Pr('relative risk is greater than 1') = 0.876132340775555
```

Estimates are provided by the **coef** function:

```
coef(model)
```

```
## Estimates of ϕ
##
## mode : 0.6338462
## mean : 0.7295833
## median : 0.6959761
## intrinsic : 0.6966669
## intrinsic2 : 0.7033364
```

Posterior credibility intervals are provided by the **confint** function:

```
confint(model)
```

```
## 95%-credibility intervals about ϕ
##
## +--------------+--------+-------+
## |   interval   |  lwr   |  upr  |
## +==============+========+=======+
## | equi-tailed  | 0.3679 | 1.284 |
## +--------------+--------+-------+
## | equi-tailed* | 0.3679 | 1.284 |
## +--------------+--------+-------+
```

```
## |     hpd      | 0.3232 | 1.202 |
## +--------------+--------+-------+
## |  intrinsic   | 0.3624 | 1.268 |
## +--------------+--------+-------+
## |  intrinsic2  | 0.3522 | 1.244 |
## +--------------+--------+-------+
```

Predictions are provided by the **predict** function after adding the sample sizes of the future experiment:

```
model <- model(Snew=500, Tnew=500)
predict(model)
```

```
## Predictions and 95%-credibility prediction intervals
##
## +-------+--------+----------+-------+-------+
## |  obs  |  size  |  median  |  lwr  |  upr  |
## +=======+========+==========+=======+=======+
## | xnew  |  500   |    71    |  38   |  117  |
## +-------+--------+----------+-------+-------+
## | ynew  |  500   |   102    |  62   |  156  |
## +-------+--------+----------+-------+-------+
```