The Beta-Binomial Distribution

Kevin R. Coombes

August 4, 2017

Contents

1 Introduction

This vignette documents the beta-binomial distribution, which is included in the *TailRank* package

```
> library(TailRank)
```

Mathematically, the beta-binomial distribution has parameters N, u, and v that determine the density function

$$\binom{N}{x} Beta(x+u, N-x+v)/Beta(u,v).$$

Statistically, one can think of this distribution as a hierarchical model, starting with a binomial distribution $Binom(x, N, \theta)$ whose success parameter θ comes from a beta distribution, $\theta \sim Beta(x, u, v)$. This distribution has a larger variance than the binomial distribution with a fixed (known) parameter θ.

We provide the usual set of functions to implement a distribution:

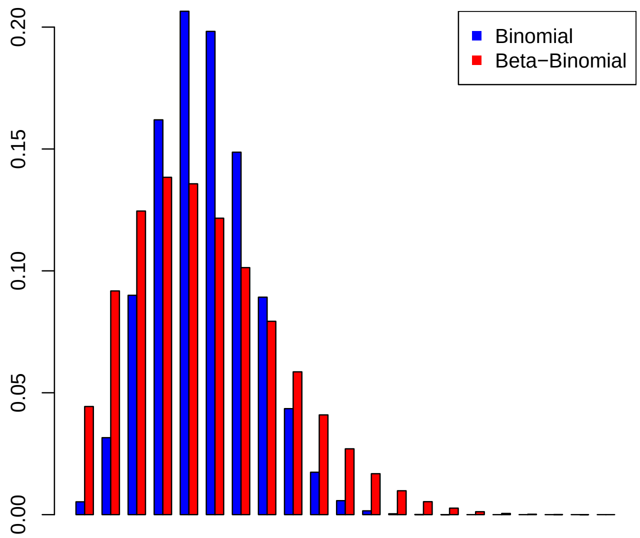
- `dbb` is the distribution function.
- `pbb` is the cumulative distribution function.
- `qbb` is the quantile function.
- `rbb` is the random-sample function.

We start by comparing the distributions of a binomial distribution and a beta-binomial distribution.

```
> N <- 20
> u <- 3
> v <- 10
```

```
> p <- u/(u+v)
> x <- 0:N
> y <- dbinom(x, N, p)
> yy <- dbb(x, N, u, v)
>

> barplot(t(matrix(c(y, yy), ncol=2)), beside=TRUE, col=c("blue", "red"))
> legend("topright", c("Binomial", "Beta-Binomial"), col=c("blue", "red"), pch=15)
```



Now we sample data from each of these distributions.

```
> set.seed(561662)
> r <- rbinom(1000, N, p)
> rr <- rbb(1000, N, u, v)
> mean(r)

[1] 4.599

> mean(rr)

[1] 4.482

> var(r)

[1] 3.345545
```

```
> var(rr)
[1] 7.741417
> sd(r)
[1] 1.829083
> sd(rr)
[1] 2.78234
```