Describing phenology with beta, beta-binomial, and beta-hypergeometric distributions

Geoffrey Legault

University of British Columbia

Characterizing phenology

CONCEPTS & SYNTHESIS

EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

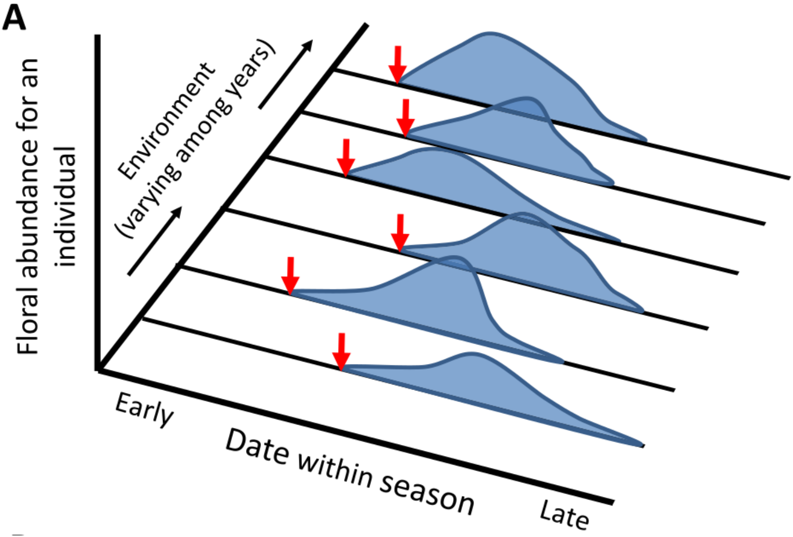
Ecological Monographs, 89(2), 2019, e01352

Phenology as a process rather than an event: from individual reaction norms to community metrics

Brian D. Inouye,[1,2,3,5] Johan Ehrlén,[2,4] and Nora Underwood[1,2,3]

- Phenology is usually characterized with discrete measures such as *first flowering*
- This approach can obscure many important features
- More useful to describing phenology as a distribution

Phenology as a distribution



Source: *Inouye et al. 2019*

Advantages of phenology as a distribution

- More comprehensive (accurate?) description
- Use more of the available data/information
- Fit the distribution to data
- Parameters of the distribution will (hopefully) be interpretable
- Given a distribution, we can estimate measures like mean, variance, correlation
- May allow comparisons over time/environments

Objective

Find an appropriate, general, and above all interpretable distribution for describing phenology

Where to start

- With the available data, of course

  e.g., percent budburst over time

Where to start

- With the available data, of course

    e.g., percent budburst over time

- Probability theory

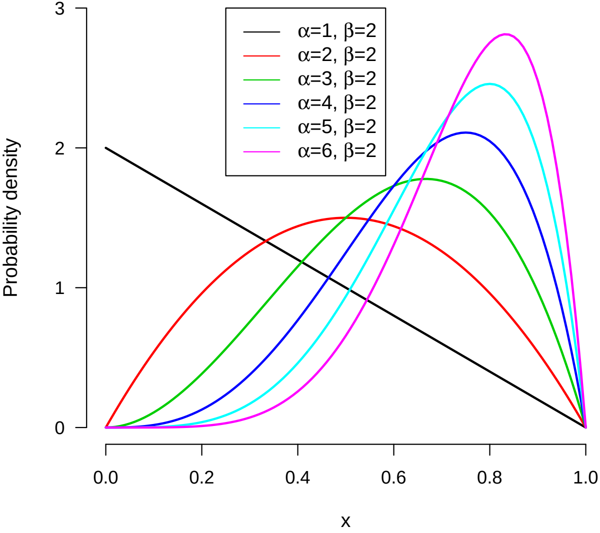
Where to start

- With the available data, of course

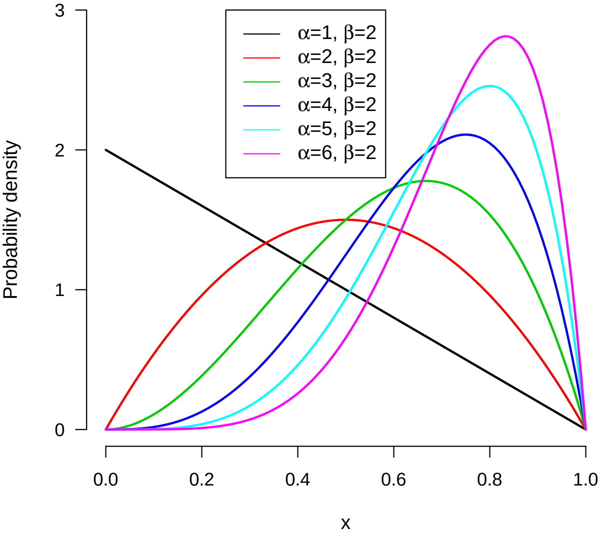
    e.g., percent budburst over time

- Probability theory
- *Beta distribution*

Beta distribution



Beta distribution



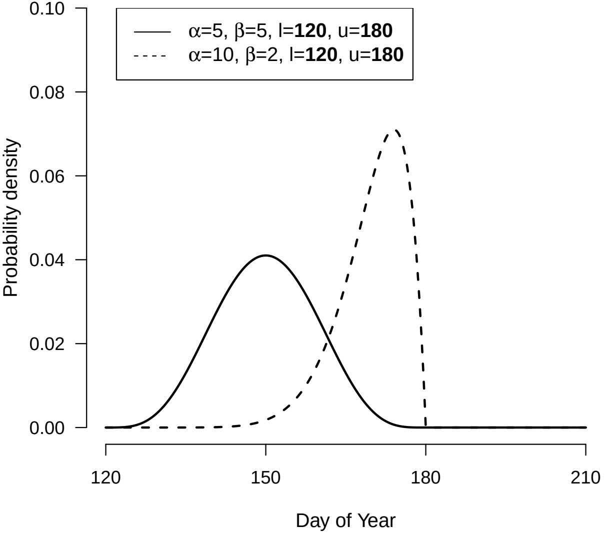
Beta distribution

- Only two parameters (that's nice)
- Effects of α and β are fairly interpretable

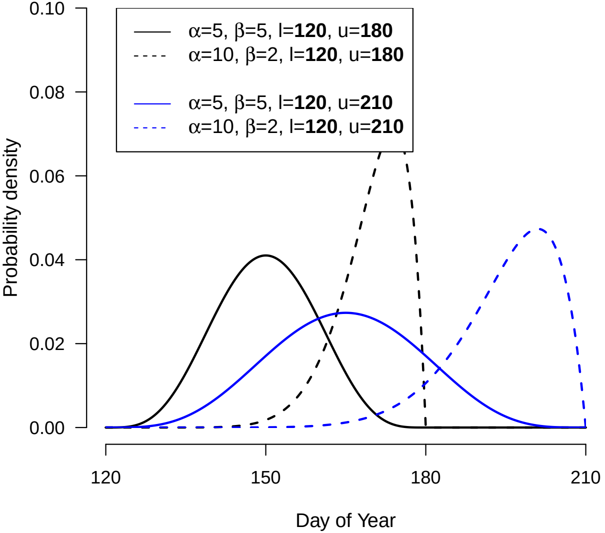
Beta distribution

- Only two parameters (that's nice)
- Effects of α and β are fairly interpretable
- **Problem 1: x ranges between 0 and 1**
- **Problem 2: Area under the curve is 1**

Beta distribution (non-standardized)



Beta distribution (non-standardized)



Beta distribution (non-standardized)

- 2 shape parameters, 2 "interval" parameters
- Parameters are still fairly interpretable

Beta distribution (non-standardized)

- 2 shape parameters, 2 "interval" parameters
- Parameters are still fairly interpretable
- **Problem 1:** x **ranges between 0 and 1**

    Solved!

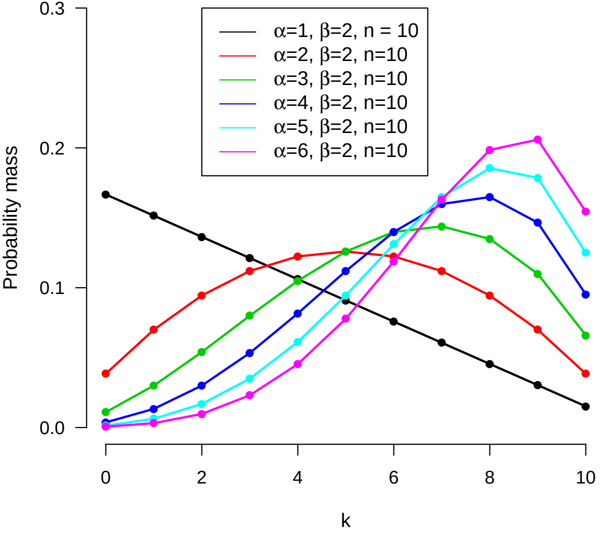
Beta distribution (non-standardized)

- 2 shape parameters, 2 "interval" parameters
- Parameters are still fairly interpretable
- **Problem 1: *x* ranges between 0 and 1**

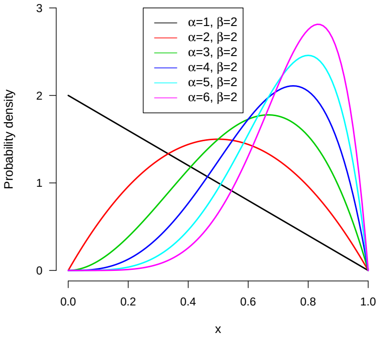
  Solved!
- **Problem 2: Area under the curve is 1**
- **Problem 3: Assumes data are continuous**

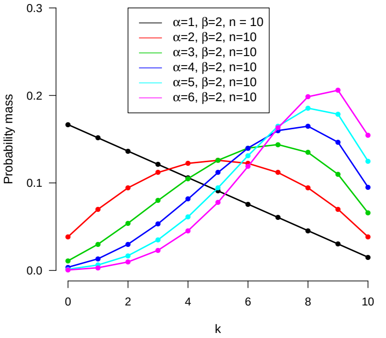
  Hmmm...

Beta-binomial distribution

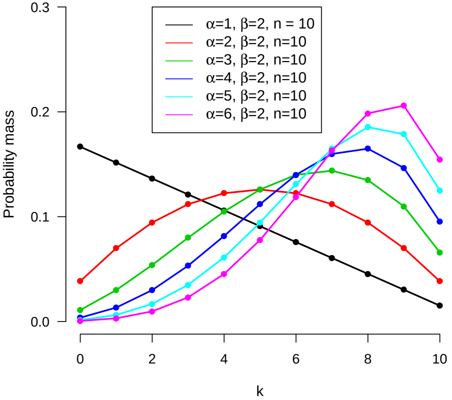


Beta-binomial distribution





Beta-binomial distribution



- Compound distribution
- Binomial random variable, $B(n, p)$, p is beta-distributed

Beta-binomial distribution

- PMF of binomial random variable:

$$P(X = k; p, n) = \binom{n}{k} p^k (1-p)^{n-k}$$

Beta-binomial distribution

- PMF of binomial random variable:

$$P(X = k; p, n) = \binom{n}{k} p^k (1 - p)^{n-k}$$

- PDF of a beta random variable:

$$f_P(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$$

Beta-binomial distribution

- PMF of binomial random variable:

$$P(X = k; p, n) = \binom{n}{k} p^k (1 - p)^{n-k}$$

- PDF of a beta random variable:

$$f_P(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$$

- PMF of compound distribution:

$$Pr(k; n, \alpha, \beta) = \binom{n}{k} \frac{1}{B(\alpha,\beta)} \int_0^1 p^{k+\alpha-1}(1 - p)^{n-k+\beta-1} dp$$

Beta-binomial distribution

- PMF of binomial random variable:

$$P(X = k; p, n) = \binom{n}{k} p^k (1 - p)^{n-k}$$

- PDF of a beta random variable:

$$f_P(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$$

- PMF of compound distribution:

$$Pr(k; n, \alpha, \beta) = \binom{n}{k} \frac{1}{B(\alpha,\beta)} \int_0^1 p^{k+\alpha-1}(1 - p)^{n-k+\beta-1} dp$$

- *Problem: This uses the standardized beta distribution (range 0 to 1)*

- **We probably want the beta-hypergeometric (which doesn't exist yet, but soon will!)**