Describing phenology as a distribution

Geoffrey Legault

University of British Columbia

Phenology as a multinomial with beta-binomial **p**

- Bursting is a sample from a multinomial
- Probability of bursting is a vector $\mathbf{p} = \{p_1, p_2, ..., p_N\}$
- $p_k = (k = d - l; n = u - l + 1, \alpha, \beta) = \binom{n}{k}\frac{B(k+\alpha, n-k+\beta)}{B(\alpha,\beta)}$
    - d is day of year
    - l is first possible day of phenology
    - u is last possible day of phenology
    - α and β are shape parameters

Discrete parameters

- Both l and u are discrete parameters
- Stan cannot directly estimate discrete parameters
- However, this can be done indirectly by considering *marginal* and *conditional* distributions

Definitions

- Conditional distribution: Probability contingent upon specific values of other variables
- Marginal distribution: Probability after considering all possible values of other variables.
  - Other variables have been "marginalized out"

Leaf eating example

Suppose a caterpillar can encounter two types of leaves ("fresh" and "dried out"). How much they eat depends on the type of leaf.

For a fresh leaf, the amount eaten is:
$N \sim Normal(\mu_F = 25, \sigma_F^2 = 5)$

For a dried leaf, the amount eaten is:
$N \sim Normal(\mu_D = 10, \sigma_D^2 = 5)$

Leaf eating example

The probability that a caterpillar eats *N* amount of leaf materials is therefore a *conditional* probability.

$$f(N = n | leaf = \text{"}fresh\text{"}) = f_{Normal}(\mu_F, \sigma_F^2)$$

$$f(N = n | leaf = \text{"}dried\text{"}) = f_{Normal}(\mu_D, \sigma_D^2)$$

Leaf eating example

$$f(N = n | leaf = \text{"}fresh\text{"}) = f_{Normal}(\mu_F, \sigma_F^2)$$
$$f(N = n | leaf = \text{"}dried\text{"}) = f_{Normal}(\mu_D, \sigma_D^2)$$

Given a probability, *p*, of encountering a fresh leaf:

Joint probability of eating *N* amount of a fresh leaf is:

$$f(N = n, leaf = \text{"}fresh\text{"}) = f_{Normal}(\mu_F, \sigma_F^2) * p$$

Joint probability of eating *N* amount of a dried leaf is:

$$f(N = n, leaf = \text{"}dried\text{"}) = f_{Normal}(\mu_D, \sigma_D^2) * (1 - p)$$

Leaf eating example

$$f(N = n, leaf = "fresh") = f_{Normal}(\mu_F, \sigma_F^2) * p$$
$$f(N = n, leaf = "dried") = f_{Normal}(\mu_D, \sigma_D^2) * (1 - p)$$

Suppose we wish to know only probability of eating *N* leaf

- What is probability of eating *N regardless* of leaf type?
- We need to "marginalize out" the leaf probability

Leaf eating example

Probability of eating N leaf is:

$$f(N=n) = f(N=n|\text{''}fresh\text{''})p + f(N=n|\text{''}dried\text{''})(1-p)$$

Leaf eating example

Let's simulate

```
## Generate leaf type
leaf <- rbinom(n = 1000, size = 1, prob = .5)
## Conditional sampling
data <- dim(length(leaf))
for(i in 1:length(leaf)){
    if(leaf[i] == 0){
        data[i] <- rnorm(n = 1, mean = 10, sd =
            sqrt(5))
    } else{
        data[i] <- rnorm(n = 1, mean = 25, sd =
            sqrt(5))
      }
}
```