Computational Statistical Modeling of Dynamic Socioeconomic, Geopolitical and Financial Systems

David K. A. Mordecai

NYU Courant Institute of Mathematical Sciences

Applied Mathematics Advanced Topics Course

Lecture #6 – March 6th, 2012

Lecture #6 Outline

- Relationship between the *Beta* and *Gamma* functions
  - *Beta* is a ratio of the product of two factorials divided by a factorial
  - *Beta* can also be expressed as the ratio of the product of two gammas (each with a different parameter x, y) divided by the gamma of the sum $x+y$
- The *Beta* function has many other useful forms
- How is the *Beta* function Analogous to the *Gamma* function for integers?
- *Branching Processes* and the Role of the *Beta* function in the *Yule-Simon process* (*linear preferential attachment)* of *cumulative advantage*
  - *Preferential attachment* exhibits which type of stochastic urn process?
  - Role of preferential attachment in scale-free networks (under suitable circumstances)
    - Species Propagation; Urban population densities; Income/wealth distributions
    - Links on the world wide web; cultural/linguistic propagation; technology adoption
    - Lotka’s law of scientific productivity; Bradford’s law of citations

Beta Function as a Ratio of *Gamma* Functions

When x and y are positive integers, it follows trivially (from the definition of the gamma function) that Beta is a ratio of factorials:

$$\mathrm{B}(x,y) = \frac{(x-1)!\,(y-1)!}{(x+y-1)!}$$

Beta has many other forms, including a ratio of *Gamma* functions (product over sum):

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}$$

The *Beta* Function Analogous to *Gamma* Function for Factorials

Just as the gamma function for integers describes factorials, the beta function can define a binomial coefficient as follows:

$$\binom{n}{k} = \frac{1}{(n+1)\mathrm{B}(n-k+1, k+1)}.$$

Note: For a given integer *n*, *Beta* can be integrated to give a closed form to interpolate continuous values of k:

$$\binom{n}{k} = (-1)^n n! \frac{\sin(\pi k)}{\pi \prod_{i=0}^{n}(k-i)}.$$

Estimating A Hierarchical Model From Data

Beta-Binomial PMF

$$\binom{n}{k}\frac{B(k+\alpha, n-k+\beta)}{B(\alpha,\beta)}$$

As 2-Stage Hierarchical Model

$$k_i \sim \mathrm{Bin}(n_i, \theta_i)$$

$$\theta_i \sim \mathrm{Beta}(\mu, M), \text{ i.i.d.}$$

$$p(\theta|x) \propto p(x|\theta)p(\theta)$$

$$p(\theta, \varphi|x) \propto p(x|\theta)p(\theta|\varphi)p(\varphi).$$

$$\varphi|\psi \text{ where } \psi \sim \ldots$$

Conditional Probability Expressed In Terms of Properties of Beta Distribution

$$f(k|\alpha,\beta) = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)}\frac{\Gamma(\alpha+k)\Gamma(n+\beta-k)}{\Gamma(\alpha+\beta+n)}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}.$$

The *Beta* Function Has Many Useful Forms

$$\mathrm{B}(x,y) = \sum_{n=0}^{\infty} \frac{\binom{n-y}{n}}{x+n},$$

$$\mathrm{B}(x,y) = \frac{x+y}{xy} \prod_{n=1}^{\infty} \left(1 + \frac{xy}{n(x+y+n)}\right)^{-1},$$

$$\mathrm{B}(x,y) = \int_0^{\infty} \frac{t^{x-1}}{(1+t)^{x+y}}\, dt, \qquad \mathrm{Re}(x) > 0,\ \mathrm{Re}(y) > 0$$

$$\mathrm{B}(x,y) \cdot (t \mapsto t_+^{x+y-1}) = (t \to t_+^{x-1}) * (t \to t_+^{y-1}) \qquad x \geq 1, y \geq 1,$$

$$\mathrm{B}(x,y) \cdot \mathrm{B}(x+y, 1-y) = \frac{\pi}{x \sin(\pi y)},$$

$$\mathrm{B}(x,y) = 2 \int_0^{\pi/2} (\sin\theta)^{2x-1} (\cos\theta)^{2y-1}\, d\theta, \qquad \mathrm{Re}(x) > 0,\ \mathrm{Re}(y) > 0$$

The *Yule-Simon Process and the Beta* Function

- in the theory of the *preferential attachment process*, the *Yule-Simon process* (linear preferential attachment) of *cumulative advantage*
  - a type of *stochastic urn process*
  - under suitable circumstances, can generate power law distributions from *branching processes*
- Commonly observed within species distributions, income and wealth distributions, various *scale-free* networks
  - Speciation; Urban population densities; Income/Wealth Distributions
  - Cultural and Language Propagation
  - Technology Adoption/Innovation; Financial Contagion (?)
  - Lotka's law of scientific productivity, Bradford's law of citations
  - World Wide Web

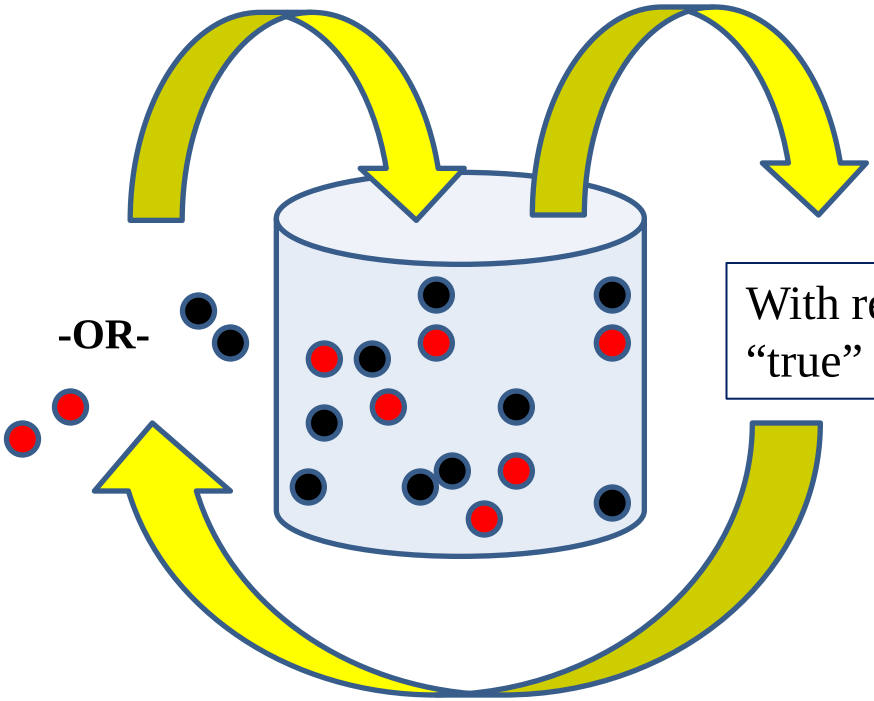
Beta-Binomial

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

Where: *p ~ Beta(α,β)*
and

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

For positive integer values *α , β:*



With replacement x 2 (if either "true" or "not true"): *Beta-Binomial*

Linear Preferential Attachment Processes

- Linear preferential attachment processes in which the number of urns increases are known to produce a distribution of balls over the urns following the so-called
- Yule-Simon distribution (also a continuous mixture of geometric distributions, e.g. exponential-geometric mixture).
- In the most general form of the process, balls are added to the system at an overall rate of m new balls for each new urn.
- Each newly created urn starts out with k_0 balls and further balls are added to urns at a rate proportional to the number k that they already have plus a constant $a > -k_0$.

The fraction $P(k)$ of urns having k balls in the limit of long time is given by:

$$P(k) = \frac{\mathrm{B}(k+a, \gamma)}{\mathrm{B}(k_0+a, \gamma-1)},$$

with $\Gamma(x)$ being the standard gamma function, and

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

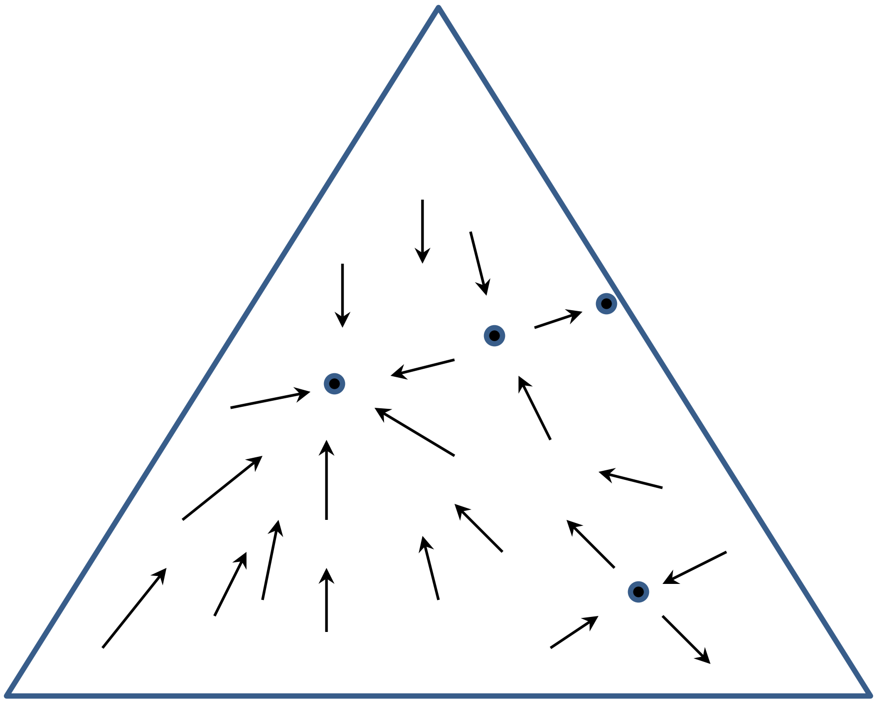
for $k \geq k_0$ (and zero otherwise), where $\mathrm{B}(x, y)$ is the Euler beta function:

$$\gamma = 2 + \frac{k_0 + a}{m}.$$

The beta function behaves asymptotically as $\mathrm{B}(x, y) \sim x^{-y}$ for large x and fixed y, which implies that for large values of k, the preferential attachment process generates a "long-tailed" distribution (a realization of Zipf's law) following a Pareto distribution (i.e. power law in the tail of the distribution):

$$P(k) \propto k^{-\gamma}.$$

Preferential Attachment, Path Dependence, and Spatial Location



Expected Motions for a Locational Probability Function

For example, in the case of species propagation, new "urns" are added to a class in a taxonomy whenever a newly appearing species is considered sufficiently different from its predecessors that it does not belong in any of the current classes. New species "balls" are added as prior generations branch out (i.e. speciate). Assuming that new species belong to the same genus as their parent (except for those that start new genera), the probability that a species is added to a new genus will be proportional to the number of species the genus already has. This process of speciation (*assortative mixing* process) first studied by Yule, is a linear preferential attachment process, since the rate at which classes accrue new species is linear in their prior number.

Hence:

$$\Gamma(x)\,\Gamma(y) = \Gamma(x+y)\mathrm{B}(x,y).$$

$$\Gamma(x)\Gamma(y) = \left(\int_{\mathbb{R}} f(u)\,du\right)\left(\int_{\mathbb{R}} g(u)\,du\right) = \int_{\mathbb{R}}(f*g)(u)\,du = \mathrm{B}(x,y)\,\Gamma(x+y)$$

$$\frac{\partial}{\partial x}\mathrm{B}(x,y) = \mathrm{B}(x,y)\left(\frac{\Gamma'(x)}{\Gamma(x)} - \frac{\Gamma'(x+y)}{\Gamma(x+y)}\right) = \mathrm{B}(x,y)(\psi(x) - \psi(x+y)),$$

where $\psi(x)$ is the digamma function defined as the logarithmic derivative of the gamma function

$$\psi(x) = \frac{d}{dx}\ln\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

But, remember from this note: Some Other Useful Distributions Related to the Multinomial:

- When k = 2, the multinomial distribution is the Binomial distribution.
- The continuous analogue is Multivariate Normal distribution.
- Categorical distribution (for k = 2 is Bernoulli)
- **The Dirichlet distribution (the Bayesian *conjugate prior* of the multinomial)**
- **Multivariate Pólya distribution.**
- **Beta-binomial model.**

So, Is Dirichlet A Multivariate Generalization of Beta?

$$\theta \sim \mathcal{D}(\alpha)$$

$$p(\theta) \sim \mathrm{D}(\alpha_1, \ldots, \alpha_k) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)} \prod_k p_k^{\alpha_k - 1} \; where \; \theta_k > 0 \; and \; \sum_k p_k = 1$$

$$\beta(\alpha)^{-1} = \frac{\Gamma(\sum_i \alpha_i)}{\Gamma(\alpha_1) * \ldots * \Gamma(\alpha_n)} = \frac{\Gamma(\alpha_0)}{\prod_i \Gamma(\alpha_i)} \; where \; \alpha_0 = \sum_i \alpha_i$$

$$\theta = \{\theta_1, \ldots, \theta_n\}, \alpha = \{\alpha_1, \ldots, \alpha_n\} \; and \; \alpha > 0$$

$$p(\theta) = \beta(\alpha)^{-1} \prod_k \theta_i^{\alpha_i - 1} \mathrm{I}(\theta \epsilon \mathcal{S}) \; where \; \mathcal{S} = \left\{ x \epsilon \mathbb{R}^n : x_i \geq 0, \sum_i^n x_i = 1 \right\}$$

$$E[\theta_i] = \frac{\alpha_i}{\alpha_0}$$