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 #4 – February 21st, 2012

Lecture #4 Outline

- Dirichlet Distribution as a Compound Distribution
- Path Dependence and Spatial Location
  - Binomial Tree (Standard Diffusion Process)
  - Polya Urn Process (Tree Representation)
- A State-Variable Threshold Framework for Understanding the Statistical Implementation of Regime-Switching Models
- Hierarchical Models and the Use of Copulas

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

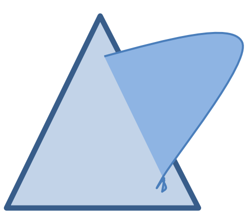
$$p(\theta) \sim \mathrm{D}(\alpha_1, \ldots, \alpha_k) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)} \prod_k p_k^{\alpha_k - 1} \ \ where\ \theta_k > 0\ and\ \sum_k p_k = 1$$

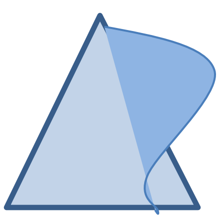
$$\beta(\alpha)^{-1} = \frac{\Gamma(\sum_i \alpha_i)}{\Gamma(\alpha_1) * \ldots * \Gamma(\alpha_n)} = \frac{\Gamma(\alpha_0)}{\prod_i \Gamma(\alpha_i)}\ \ where\ \alpha_0 = \sum_i \alpha_i$$

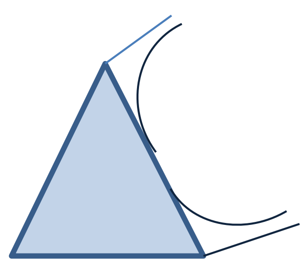
$$\theta = \{\theta_1, \ldots, \theta_n\}, \alpha = \{\alpha_1, \ldots, \alpha_n\}\ and\ \alpha > 0$$

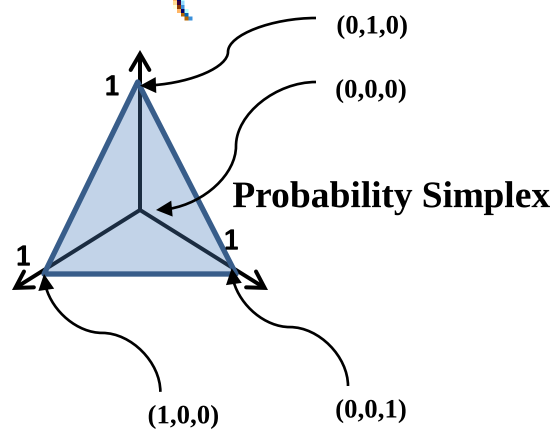
$$p(\theta) = \beta(\alpha)^{-1} \prod_k \theta_i^{\alpha_i - 1}\, \mathrm{I}(\theta \epsilon \mathcal{S})\ where\ \mathcal{S} = \left\{ x \epsilon \mathbb{R}^n : x_i \geq 0, \sum_i^n x_i = 1 \right\}$$

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

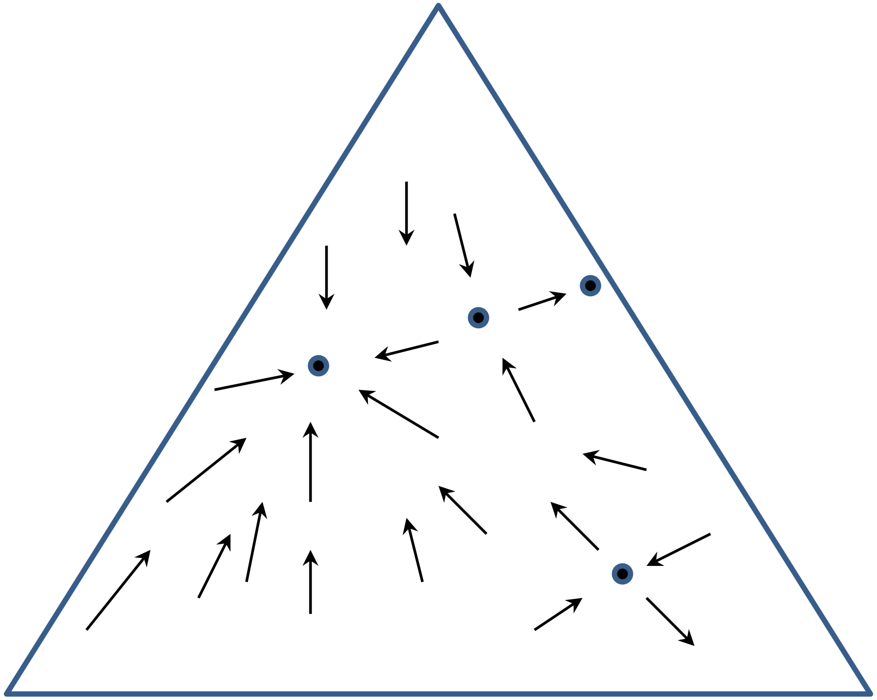






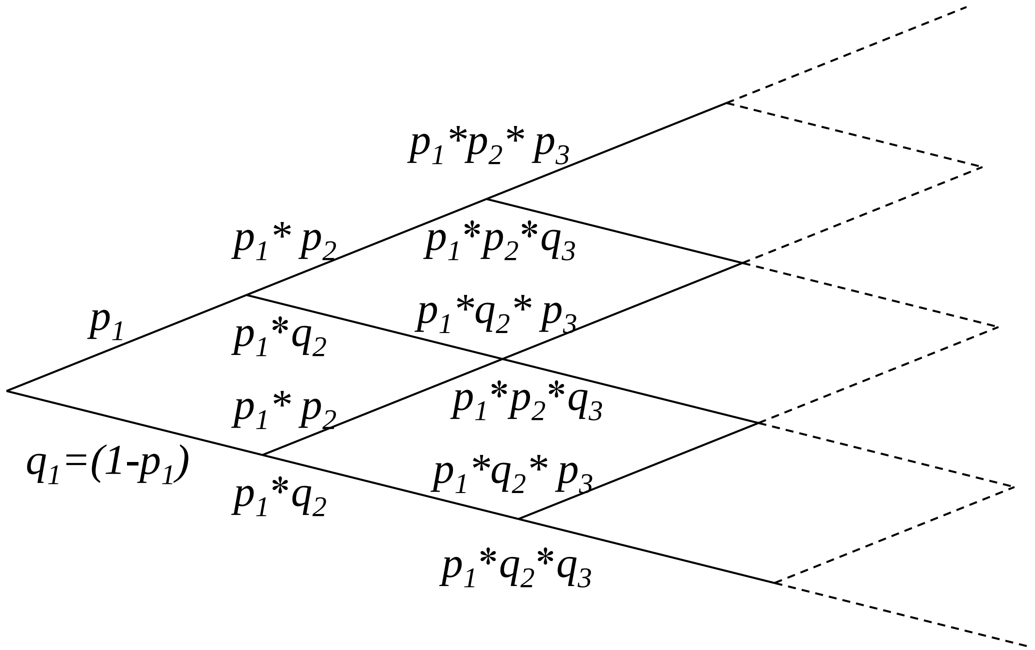


Path Dependence and Spatial Location

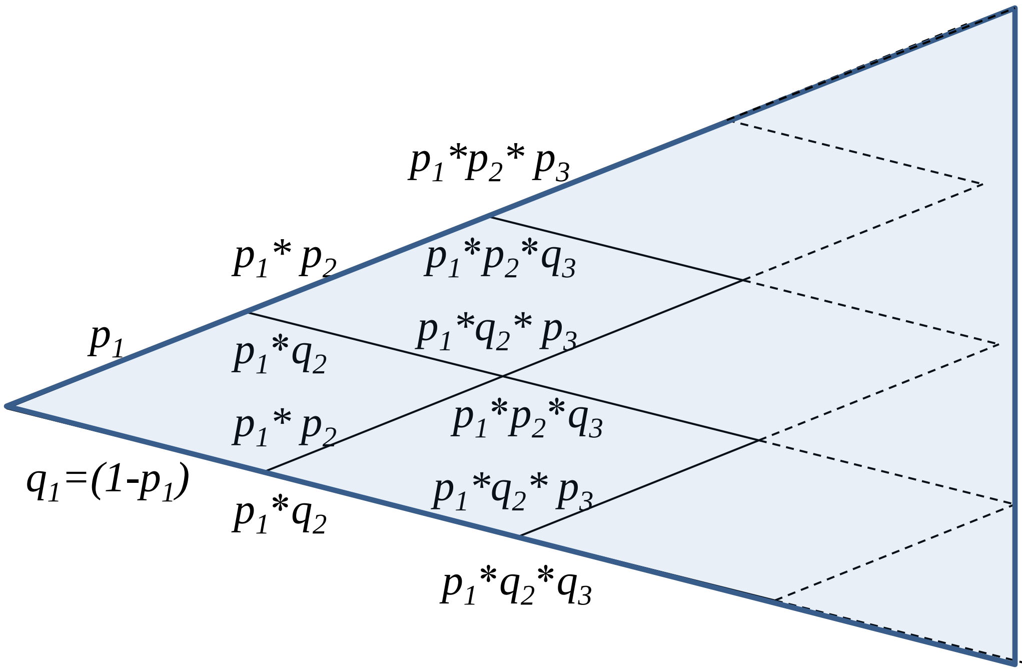


Expected Motions for a Locational Probability Function

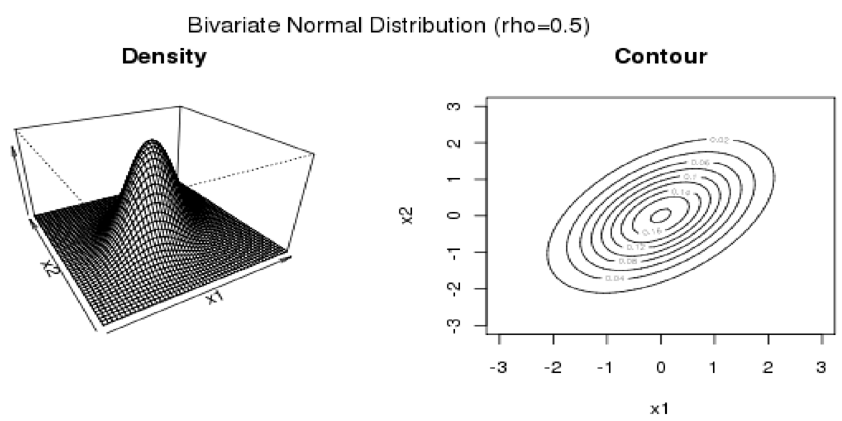
Binomial Tree



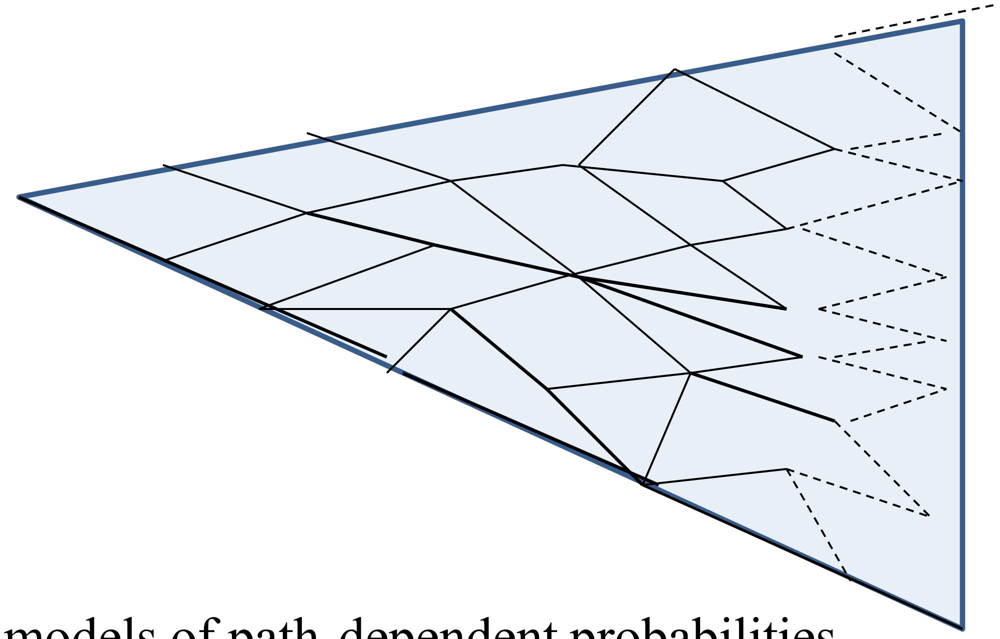
Binomial Tree



Standard model of diffusion process

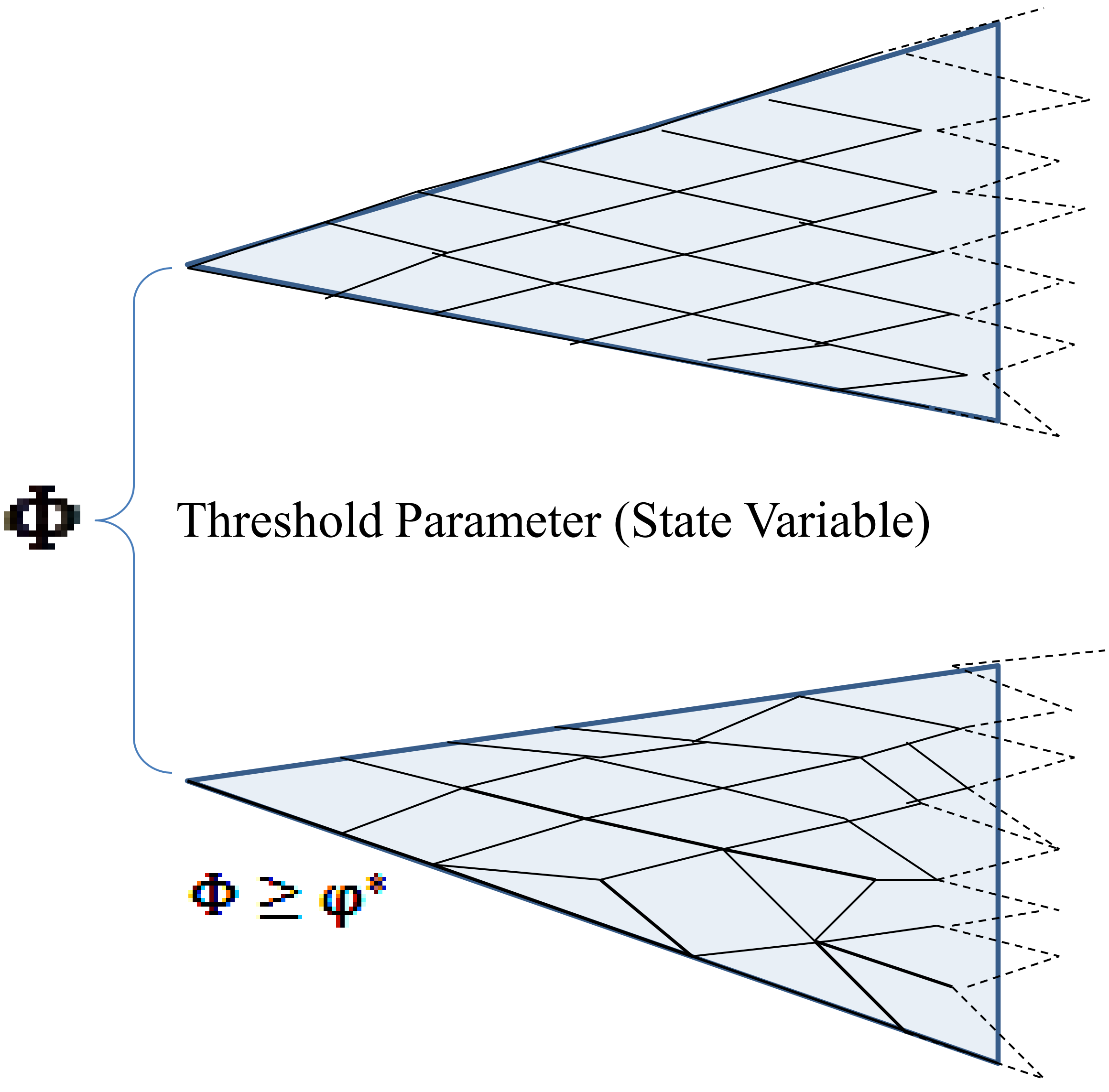


Polya Urn Tree Process



Nonstandard models of path-dependent probabilities

Regime-Switching Model

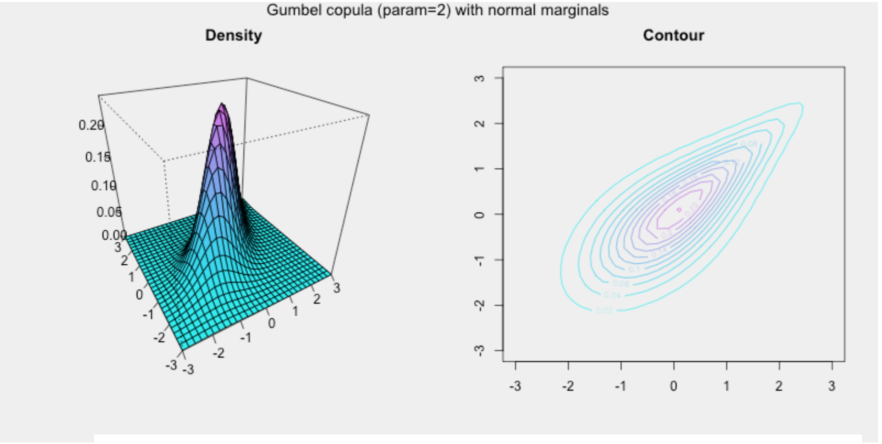


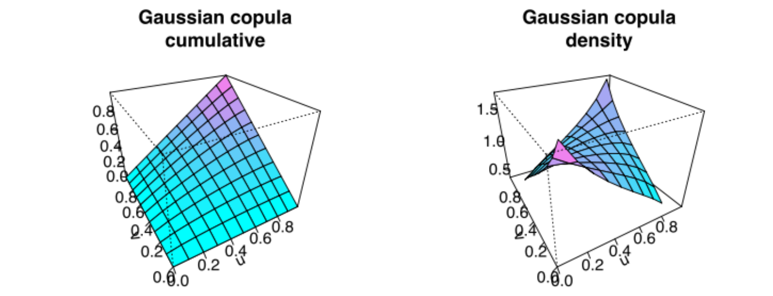
Hierarchical Models and the Use of Copulas

$\theta_i \mid x_i$

$x_i \sim N(\theta_i, \sigma^2)$ $\theta_i \sim N(\varphi, \tau^2)$

What if: $\varphi_i \sim N(0,1)$, *where* $\tau^2 \sim U(0,1)$?





Sklar's theorem: $H(x_1, \ldots, x_d) = \mathbb{P}[X_1 \leq x_1, \ldots, X_d \leq x_d]$ Multivariate CDF

random vector $(X_1, X_2, \ldots, X_d)$ margins $F_i(x) = \mathbb{P}[X_i \leq x]$

copula ***C*** $H(x_1, \ldots, x_d) = C\left(F_1(x_1), \ldots, F_d(x_d)\right),$

Hierarchical Models and the Use of Copulas (cont'd)

If copula and margins have been estimated, the expectation of an arbitrary response function $g : \mathbb{R}^d \to \mathbb{R}$ can be approximated through the application of a Monte Carlo algorithm:

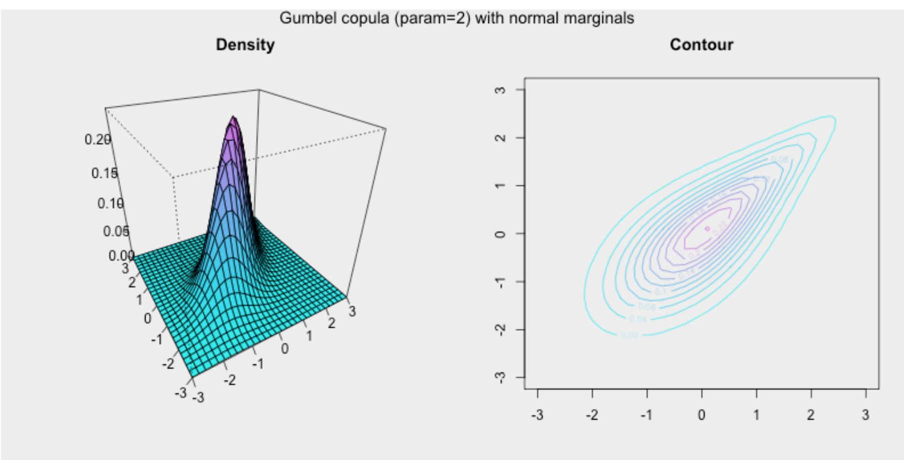
1. Draw a sample of size **n** from the copula **C**

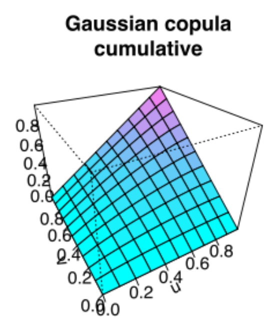
$$(U_1^k, \ldots, U_d^k) \sim C \quad (k = 1, \ldots, n)$$

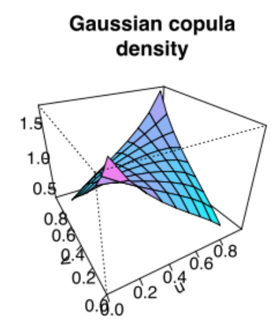
2. Set $(X_1^k, \ldots, X_d^k) = (F_1^{-1}(U_1^k), \ldots, F_d^{-1}(U_d^k)) \sim H \quad (k = 1, \ldots, n)$ and by applying the inverse marginal CDF's, produce a sample of $(X_1, \ldots, X_d)$
3. Use

$$\mathbb{E}\left[g(X_1, \ldots, X_d)\right] \approx \frac{1}{n} \sum_{k=1}^{n} g(X_1^k, \ldots, X_d^k)$$

to approximate: $\mathbb{E}\left[g(X_1, \ldots, X_d)\right]$







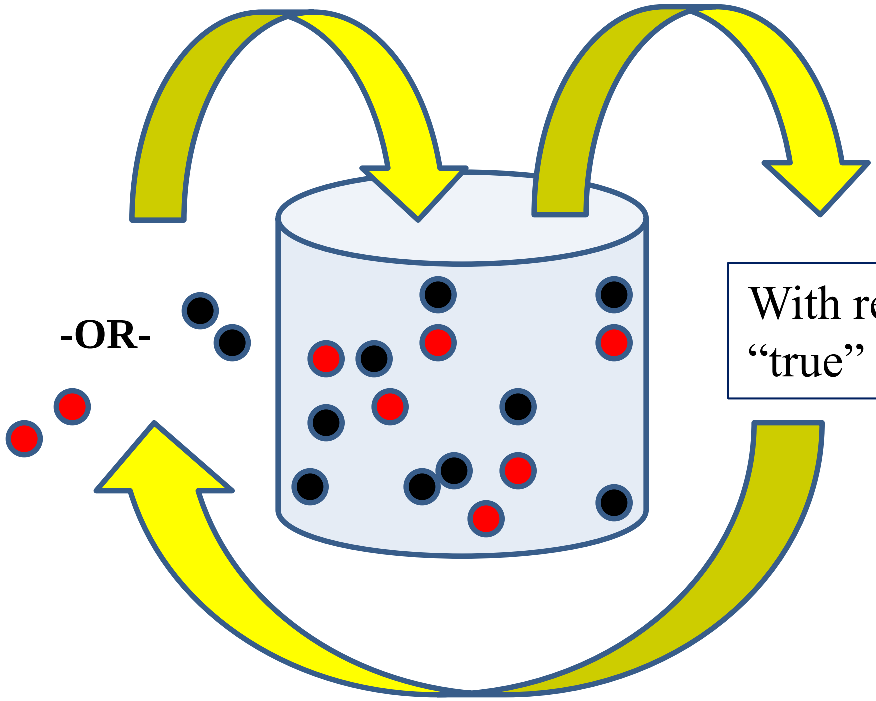
Beta-Binomial

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

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*

Estimating A Hierarchical Model From Data

Beta-Binomial PMF

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

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

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

$\varphi|\psi$ *where* $\psi \sim$ ...

As 2-Stage Hierarchical Model

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

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

Conditional Probability Expressed In Terms of Properties of Beta Distribution

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