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 #1 – January 24th, 2012

Focus of the Course

- To explore the (natural) social computing paradigm within the context of naturally social systems
  - "social computing": adaptive, distributed learning, based upon local interaction between those agents comprising any collective system

- Exploration via mapping landscapes (i.e. surfaces): aggregation and examination of the statistical properties of those local interactions in space and time, employing fundamental computational operations:
  - Counting and measuring (including sorting and mixing; sampling/resampling)
  - Addition (sums), subtraction (differences), multiplication (products), division (ratios)
  - Geometry of length, area, volume, and curvature (i.e. lines, triangles, squares, and circles)

Relevant domains at population and cohort level

- Examples:
  - Socioeconomic
  - Geopolitical
  - Legal
  - Linguistic
  - Epidemiological
  - Demographic
- Relevant Phenomena:
  - Information dissemination/aggregation
  - technological diffusion/adoption
  - cultural transmission/propagation (herds, fads, fashions);
  - learning/adaptation
  - play (imitation, bargaining, coordination, competition, searching-matching, sorting-mixing)
  - socioeconomic and geopolitical phenomena (predation, conflict, integration/segregation)

Large Population Games

- A General Framework for Simulation and Inference in N-player population games: what can be elements and choices?
  - States of nature *X* (Observed or Hidden)
  - Agents *N*
  - Types *K* (expectations or beliefs; utilities or preferences: wealth, technology, constraints, tastes, attitudes, habits, propensities, predilections, etc)
  - Neighborhoods and/or Cohorts *L* (geographies; groupings)
  - Information or innovations *F* (regarding states or types)
  - Actions *A* (choices, strategies)
  - Payoffs observable *Y* (returns, outcomes)

Fundamental notion of counting processes

- Distributions with useful statistical properties for spatio-temporal (state-space) modeling
  - Relevance as a means of characterizing both path-dependence (i.e. memory) and local interactions (i.e. network effects)
- Bernoulli distribution and urn sampling (either with and without replacement)
  - A sequence or set of independent (with replacement) Bernoulli observations:
  - Basis for the Polya family (of binomial, geometric, Pascal, and negative binomial distributions), in which the sequence or set is truncated after either n trials of x positive (negative) outcomes (or Poisson in which the sequence or set is arbitrarily large, i.e. approaches infinity), e.g. in a case where 'true'=1, 'false'=0:

Some Useful Distributions

- Bernoulli probability describes the likelihood of a 'true' outcome in a single observation;
  - Binomial probability is the likelihood (a ratio) of i>1 true observations relative to a sequence or set of n observations
  - Geometric probability describes the ratio of 'false' observations relative to the a sequence of n total observations before the first 'true' observation (i.e., the number of observations relative to a single true observation)
  - Negative Binomial probability describes the ratio of 'false' observations occurring before a specified i true observations relative to the sequence of n total observations
  - Poisson probability describes the ratio for an arbitrarily large sequence of n total observations, such that the product of the n total observations times the ratio of the probability of untrue observations (prior to the first true observation) relative to sequence of n total observations converges towards a constant.
    - Pascal distribution describes an integer-valued stopping-time parameter r
    - Polya distribution for the real-valued case more accurately describes occurrences of “contagious” discrete events than the Poisson distribution

Binomial Distribution (and Some Useful Variants)

Binomial

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

Geometric

$$(1-p)^k p$$ for $$k = 0,1,2,3, \ldots$$

Negative Binomial

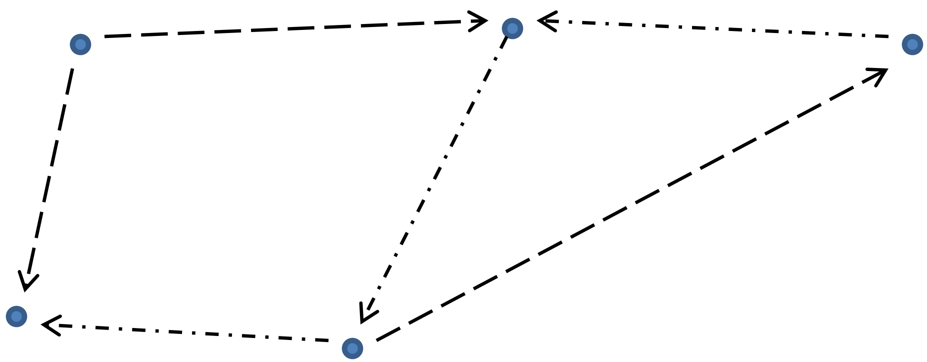
$$\binom{k+r-1}{k} (1-p)^r p^k$$ for $$k = 0,1,2,3, \ldots$$

Poisson

$$\frac{\lambda^k e^{-\lambda}}{k!}$$ for $$k = 0,1,2,3, \ldots$$

Pascal

$$\frac{\Gamma(k+r)}{k!\,\Gamma(r)}$$ If r is a real valued integer



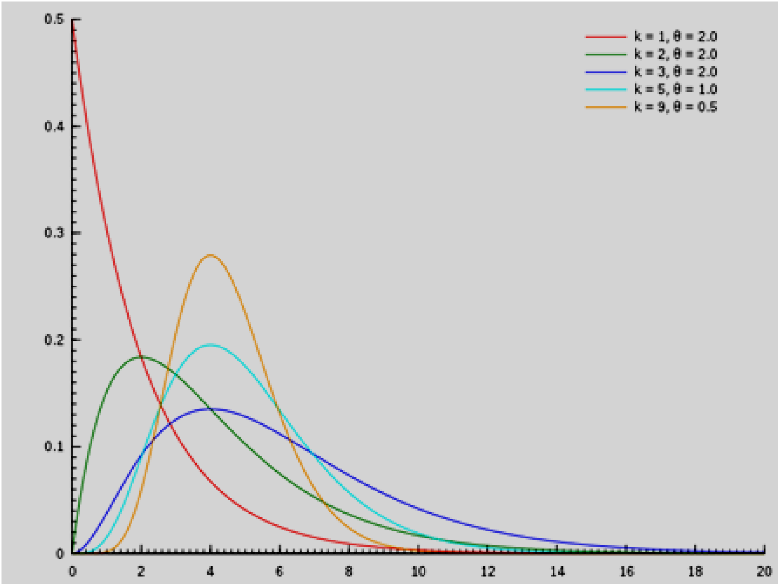
N agents
Each endowed with some technology
(heterogeneous type ➔ their costs)
Each choose to "write" some quantity of forwards q_i^f

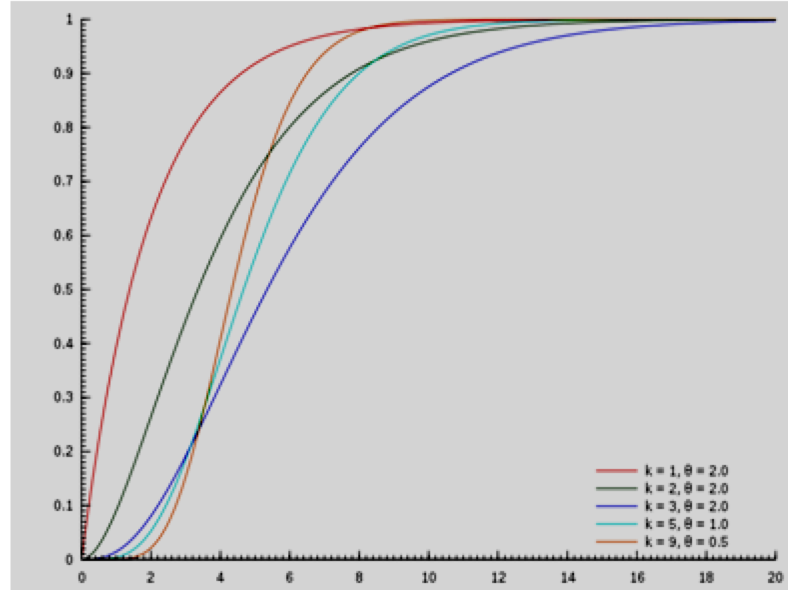
$$Q_i^f = \Sigma q_i^f$$

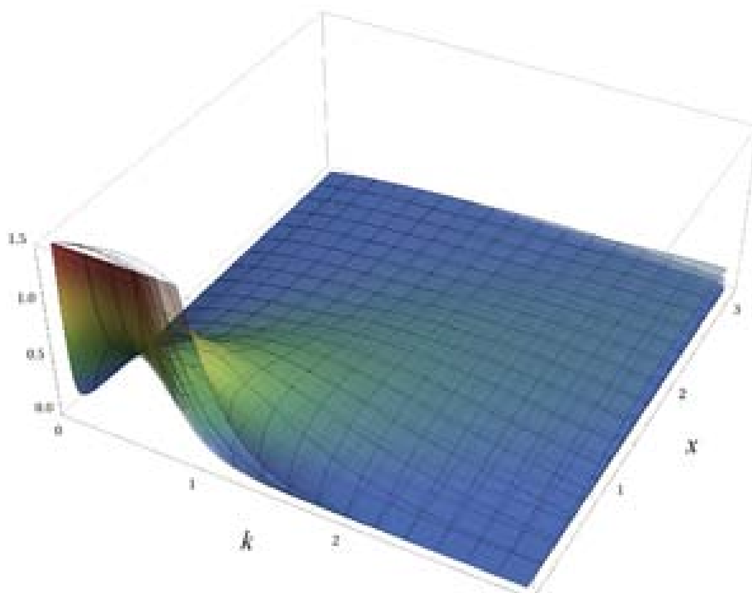
Spot Price Process $\boldsymbol{S}$ ~ *Jump Diffusion*

Jump Hazard: Negative Binomial Distributed
Overdisperse Poisson (λ)
Where λ ~ Gamma

t_0 t_1 t_2 t_3 … t_n T







3D Representation of a Gamma Probability Distribution. With superimposed surfaces (each for a different value of θ)

If X_i has a $\Gamma(k_i, \theta)$ distribution for $i = 1, 2, ..., N$ (i.e., all distributions have the same scale parameter θ), then provided all $X_{i'}$ *are i.i.d.*

$$\sum_{i=1}^{N} X_i \sim \mathrm{Gamma}\left(\sum_{i=1}^{N} k_i, \theta\right)$$

The gamma distribution exhibits infinite divisibility.

Note: the computational intensity of computing the median depends upon the α parameter and does not have an easy closed-form solution (in contrast to the median and mode)

$$\frac{1}{\Gamma(a)\theta^a}\int_0^{x_0} x^{a-1}e^{-x/\theta}dx = \frac{1}{2}$$

… and hence is best simulated

Negative Binomial is equivalent to a Poisson(λ) distribution, where λ is itself a random variable, distributed according to *Gamma*(r, $p/(1-p)$)

$$f(k) = \int_0^\infty f_{\mathrm{Poisson}(\lambda)}(k) \cdot f_{\mathrm{Gamma}\left(r,\, \frac{p}{1-p}\right)}(\lambda)\, \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} \cdot \lambda^{r-1} \frac{e^{-\lambda(1-p)/p}}{\left(\frac{p}{1-p}\right)^r \Gamma(r)}\, \mathrm{d}\lambda$$

$$= \frac{(1-p)^r p^{-r}}{k!\,\Gamma(r)} \int_0^\infty \lambda^{r+k-1} e^{-\lambda/p}\, \mathrm{d}\lambda$$

$$= \frac{(1-p)^r p^{-r}}{k!\,\Gamma(r)}\, p^{r+k}\, \Gamma(r+k)$$

$$= \frac{\Gamma(r+k)}{k!\,\Gamma(r)}\, (1-p)^r p^k.$$

Also known as the gamma–Poisson (mixture) distribution, a continuous mixture of Poisson distributions where the mixing distribution of the Poisson rate is a gamma distribution

As a sum of geometric random variables, if $Y_r \sim NegBin(r, p)$, with support $\{0, 1, 2, ...\}$, then Y_r is a sum of r independent variables following the geometric distribution (on $\{ 0, 1, 2, 3, ... \}$) with parameter $1 - p$ ➔ CLT: Y_r (properly scaled and shifted) is therefore approximately normal for sufficiently large r.

$$\begin{aligned}\Pr(Y_r \le s) &= 1 - I_p(s+1, r) \\ &= 1 - I_p((s+r)-(r-1), (r-1)+1) \\ &= 1 - \Pr(B_{s+r} \le r-1) \\ &= \Pr(B_{s+r} \ge r) \\ &= \Pr(\text{after } s+r \text{ trials, there are at least } r \text{ successes}).\end{aligned}$$

Note: Another useful property is that *NegBin* is infinitely divisible such that for any positive integer n, there exist independent identically distributed random variables $Y_1, ..., Y_n$ whose sum has the same distribution as does Y.

Also, if B_{s+r} is a random variable following the binomial distribution with parameters $s + r$ and $1 - p$, then the *NegBin* distribution is the "inverse" of the binomial distribution, and the sum of independent negative-binomially distributed random variables r_1 and r_2 with the same value for parameter p is negative-binomially distributed with the same p but with "r-value" $r_1 + r_2$.

$$f(k) = \frac{-p^k}{k \ln(1-p)}, \qquad k \in \mathbb{N}.$$

$$X = \sum_{n=1}^{N} Y_n$$

$NegBin(r, p)$ can be also represented as a *compound (overdisperse) poisson*: Let $\{Y_n, n \in \mathbb{N}_0\}$ denote a sequence of *i.i.d*, each one having the logarithmic distribution $log(p)$, with probability mass function and let N be a random variable (independent of the sequence). If N has a Poisson distribution with parameter $\lambda = -r \ln(1 - p)$. Then the random sum is NB(r,p)-distributed.

The binomial theorem implies that $Y \sim \mathrm{Bin}(n,p)$ and $p + q = 1$ with $p, q \geq 0$.

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

Which by Newton's binomial theorem can be expressed as the following (for which the upper bound of summation is infinite)

$$(p+q)^n = \sum_{k=0}^{\infty} \binom{n}{k} p^k q^{n-k},$$

In certain cases, the binomial coefficient can be defined when n is a real number, (vs. a positive integer), or the binomial distribution can be zero when k > n.

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

If one supposes r > 0 and uses a negative exponent, Then all of the terms are positive, and the probability that the number of failures before the rth success is equal to k, provided r is an integer.

$$1 = p^r \cdot p^{-r} = p^r (1-q)^{-r} = p^r \sum_{k=0}^{\infty} \binom{-r}{k} (-q)^k. \quad \Rightarrow \quad p^r \binom{-r}{k} (-q)^k$$

The Newton-Raphson method in one variable: Given a function $f(x)$ and its derivative $f'(x)$, we begin with a first guess x_0 for a root of the function and iterate (provided the function is reasonably well-behaved) to find approximation $x_1 \ldots x_n$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Geometrically, x_1 is the intersection with the x-axis of a line tangent to f at $f(x_0)$. Repeat until a sufficiently accurate value is reached:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

To simulate a multinomial:

(1) reorder the parameters such that they are sorted in descending order (this is only to speed up computation and not strictly necessary).

(2) for each trial, draw X from a uniform (0, 1). The resulting outcome is the component for the multinomial distribution with n = 1.

$$j = \arg\min_{j'=1}^{k} \left(\sum_{i=1}^{j'} p_i \geq X \right).$$

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 conjugate prior of the multinomial in Bayesian statistics)
- Multivariate Pólya distribution.
- Beta-binomial model.