I’m getting too old for heavy duty theorem proving – [digression] what I actually do nowadays is indicated on the next two slides. But I do have some “general abstract” material that I would like some smart young people to think about. This talk is in the “back of an envelope” style – a somewhat fuzzy big picture without defining things carefully. But there are two underlying actual theorems and two specific conjectures.

I occasionally teach a “Probability in the Real World” course, in which I give 20 lectures, ideally “anchored” by some real data.
• Everyday perception of chance
• Ranking and rating
• Risk to individuals: perception and reality
• Luck
• A glimpse at probability research: spatial networks on random points
• Prediction markets, fair games and martingales
• Science fiction meets science
• Coincidences, near misses and one-in-a-million chances.
• Psychology of probability: predictable irrationality
• Mixing: physical randomness, the local uniformity principle and card shuffling
• Game theory
• The Kelly criterion for favorable games: stock market investing for individuals
• Toy models in population genetics: some mathematical aspects of evolution
• Size-biasing, regression effect and dust-to-dust phenomena
• Toy models of human interaction: use and abuse
• Short/Medium term predictions in politics and economics
• Tipping points and phase transitions
• Coding and entropy
This “Real World” project was not intended as research, but the spirit – starting with accessible data of some form that other probabilists don’t think about – does sometimes lead to non-technical/expository papers such as

- The Prediction Tournament Paradox.

and occasionally to more technical work

Today’s topic: Thousands of papers study random processes over networks; typically with a specific model for the process and a specific model for the network. Looking at the theorem-proof math-probability end of this field, and comparing with real examples (e.g. from the Easley - Kleinberg text Networks Crowds and Markets: next page), my reaction is

- the usual “network models of math convenience” are not realistic
- and “network” is better modeled as an edge-weighted graph.

As one small step toward reality . . . . . . what can one say about a specific process over an arbitrary network (envisage as really heterogeneous)? [in the paradigm of size asymptotics: (number of vertices) $n \to \infty$].

Methodology involves compactification and weak concentration, the latter just meaning $X_n/\mathbb{E}[X_n] \to_p 1$. 
Collaboration Graphs
- co-authorships among scientists
- co-appearance in movies by actors
- corporate types on same major board of directors
- Wikipedia editors ever edited same article
- World of Warcraft users ever been allied

Who-talks-to-Whom Graphs
- Microsoft IM graph
- e-mail logs within a company or a university
- phone calls (number-to-number) in given period
- physical proximity (individuals) in given period from cell phone tracking
- buyers and sellers in a market

Information Linkage Graphs
- WWW graph of web pages/links
- linkages among bloggers
- ”friends” on Facebook or MySpace

Technological Networks
- physical Internet (AS graph)
- electricity generating stations in a power grid

Networks in the Natural World
- food webs
- neurons
- biochemical interactions within cells
To me a **network** is a finite \((n \text{ vertices})\) connected edge-weighted undirected graph, vertices \(v, x, y, \ldots\) and edge weights \(w_e = w_{xy}\).

Note two opposite conventions for interpreting weights:
- In TSP-like setting, weight is distance or cost.
- In social networks, weight is strength of relationship (**this talk**).

To me there are three fundamental random processes one can define over a network:
- the (continuous-time) Markov chain.
- **bond percolation**
- first passage percolation.

All of these are (somewhat) amenable to the methodologies in this talk; I will focus on bond percolation because it suggests conjectures for more general epidemic-type models.
Bond percolation.

An edge $e$ of weight $w_e$ becomes open at an $\text{Exponential}(w_e)$ random time.

In this process we can consider

$$C(t) = \text{max size (number of vertices) in a connected component of open edges at time } t$$

and also

$$C^{(k)}(t) = \text{total no. vertices in components of size } \geq k .$$

The former relates to “emergence of the giant component”. Studied extensively on many non-random and specific models of random networks. Can we say anything about $n \to \infty$ asymptotics for (almost) arbitrary networks?

Yes: there are two parallel results.
Suppose (after time-scaling) there exist constants $t_* > 0$, $t^* < \infty$ such that

$$\lim_{n} \mathbb{E}C_n(t_*)/n = 0; \quad \lim_{n} \mathbb{E}C_n(t^*)/n > 0. \quad (1)$$

In the language of random graphs, this condition says a giant component emerges (with non-vanishing probability) at some random time of order 1.

**Proposition (1)**

*Given a sequence of networks satisfying (1), there exist constants $\tau_n$ such that, for every sequence $\varepsilon_n \downarrow 0$ sufficiently slowly, the random times

$$T_n := \inf \{ t : C_n(t) \geq \varepsilon_n n \}$$

satisfy

$$T_n - \tau_n \rightarrow_p 0.$$*

The Proposition asserts, informally, that the “incipient” time at which a giant component starts to emerge is deterministic to first order.
Suppose (after time-scaling) there exist constants $t^* > 0$, $t^* < \infty$ such that

$$\lim_{k \to \infty} \limsup_n \mathbb{E}[n^{-1} C_n^{(k)}(t^*)] = 0 \quad \lim_{k \to \infty} \liminf_n \mathbb{E}[n^{-1} C_n^{(k)}(t^*)] > 0. \quad (2)$$

**Proposition (2)**

Given a sequence of networks satisfying (2), there exist constants $\tau_n$ such that, for every sequence $k_n \uparrow \infty$ sufficiently slowly, and every $\varepsilon > 0$,

$$n^{-1} C_n^{(k_n)}(\tau_n - \varepsilon) \to 0 \quad \text{in probability}$$

$$n^{-1} C_n^{(k_n)}(\tau_n + \varepsilon) \quad \text{bounded away from 0 in probability}$$

The Proposition asserts, informally, that the time at which a non-vanishing proportion of vertices are not in bounded size components is deterministic to first order.
An instructive example

In nice networks the two \((\tau_n)\) will be the same, but we want to handle really heterogeneous networks.
Curiously, the proofs are very different. Each involves a different **magic trick** which doesn’t seem to generalize.

- Proposition 1: proof is hard but not sophisticated. See *The Incipient Giant Component in Bond Percolation* … (2016).
- Proposition 2: will outline next a proof – soft but sophisticated. Details left to some smart young person.

**Local weak convergence.**

LWC usually interpreted as Benjamini-Schramm convergence for sparse unweighted graphs, but (Aldous-Steele 2003) extends to weighted graphs, under the opposite convention (weight $w_e$ re-interpreted as length $1/w_e$, and distance is minimum route-length).
The space $\mathbf{N}$ of rooted locally finite networks has a natural topology.

Given a (random or deterministic) network, pick uniform random vertex to be root. For a sequence of networks we may have convergence in distribution (of the randomly-rooted networks) in space $\mathbf{N}$ to a limit random network – infinite but locally finite.

This theory only directly useful for sparse graphs – compactness corresponds (unweighted case) roughly to bounded average degree. If not sparse then we don’t have a limit in this space.
But . . . . . . we are considering random models over a network and these often (as in bond percolation) involve attaching Exponential($w_e$) or a Poisson (rate $w_e$) process of “event times” to edges $e$. Consider the behavior of such a process up to a fixed time $t$. Only edges which have had “events” are relevant. In this context the relevant condition on the underlying network is much weaker. All we need is that vertex-weights $w_v = \sum_y w_{vy}$ have the same order of magnitude (then scaled to order 1). Then there will only be order 1 relevant events at a typical vertex. So in a sequence of networks such that

vertex-weights $w_v$ are order 1

the subnetworks of edges on which an event has occurred by time $t$ form a compact sequence, so we can assume they have a local weak limit.
Back to bond percolation – what does this tell us? Suppose we have LWC of the process of open edges.

- A LWC limit infinite random network automatically has a property ("unimodular") which is an analog of (spatial) stationarity.
- Our vague notion of finite networks being "really heterogeneous" corresponds to the limit process being not ergodic, instead a mixture of ergodic slices.
- In the ergodic case, bond percolation on limit network has (excluding trivial cases) some constant critical time $t_{\text{crit}}$ for emergence of infinite components.
- In the non-ergodic case we can define a constant $\tau$ as the ess. $\inf.$ of the $t_{\text{crit}}$ of the mixed-over ergodic networks.
Given this background set-up (sophisticated in detail) Proposition 2 falls out.

- If we have LWC of the open subnetworks, then we can use the limit $\tau$ above, which is finite by hypothesis on $t^*$.
- The hypothesis on $t_*$ implies compactness at time $t_*$. If compact at all $t$ then take subsequence and use $\tau$ above.
- If not then use the time at which compactness fails. At that time a non-vanishing proportion of vertices have degree $\Omega(1)$, implying the property we seek. This is a magic trick which fails in the “giant component” setting.
Reformulation as epidemics (well known but subtle).

An \textit{SI} model refers to a model in which individuals are either \textit{infected} or \textit{susceptible}. In the network context, individuals are represented as vertices of an edge-weighted graph, and the model is

\textit{for each edge (vy), if at some time one individual (v or y) becomes infected while the other is susceptible, then the other will later become infected with some transmission probability $p_{vy}$.}

These transmission events are independent over edges. Regardless of details of the time for such transmissions to occur, this \textbf{SI model} is related to the \textbf{random graph model} defined by

\textit{edges $e = (vy)$ are present independently with probabilities $p_e = p_{vy}$.}

The relation is:

\textit{(*) The set of ultimately infected individuals in the SI model is, in the random graph model, the union of the connected components which contain initially infected individuals.}
In modeling an SI epidemic within a population with a given graph structure, we regard edge-weights \( w_e = w_{vy} \) as indicating relative frequency of contact. Introduce a virulence parameter \( \theta \), and define transmission probabilities

\[
p_e = 1 - \exp(-w_e \theta).
\]  

Note this allows completely arbitrary values of \( (p_e) \), by appropriate choice of \( (w_e) \). Now the point of the parametrization (3) is that the set of potential transmission edges is exactly the same as the time-\( \theta \) configuration in the bond percolation model.

Even though this is mathematically trivial, it is **conceptually subtle**. A real-world flu epidemic proceeds in real-world time; instead we just consider the set of ultimately infected people and actual transmission edges; this structure, as a process parametrized by \( \theta \), is a nice stochastic process (bond percolation).
So we can translate our two Propositions into statements about whether the SI epidemic model is pandemic (has $\Theta(n)$ vertices ultimately infected) in terms of the number $\kappa_n$ of initially infected vertices.

- If $\kappa_n \uparrow \infty$ slowly then for a pandemic we need the largest component size to be almost $\Theta(n)$
- If $\kappa_n/n \downarrow 0$ slowly then for a pandemic we need $\Theta(n)$ vertices in components of size of order $n/\kappa_n$

Say a sequence of non-negative random variables $(Y_n)$ is \textit{bounded away from 0 in probability} if

$$\lim_{\delta \downarrow 0} \limsup_{n} \mathbb{P}(Y_n \leq \delta) = 0$$

and write this as $Y_n \gg_p 0$. 

David Aldous

Limits for processes over general networks
Proposition (3)

Take edge-weighted graphs with $n \to \infty$, consider the SI epidemics with transmission probabilities of form (3), and write $C'_{n,\kappa}(\theta)$ for the number of ultimately infected individuals in the epidemic started with $\kappa$ uniformly random infected individuals. Suppose there exist some $0 < \theta_1 < \theta_2 < \infty$ such that, for all $\kappa_n \to \infty$ sufficiently slowly,

$$\lim_n n^{-1} \mathbb{E} C'_{n,\kappa_n}(\theta_1) = 0; \quad \liminf_n n^{-1} \mathbb{E} C'_{n,\kappa_n}(\theta_2) > 0.$$  \hspace{1cm} (4)

Then there exist deterministic $\tau_n \in [\theta_1, \theta_2]$ such that, for all $\kappa_n \to \infty$ sufficiently slowly,

$$n^{-1} C'_{n,\kappa_n}(\tau_n - \delta) \to_p 0, \quad n^{-1} C'_{n,\kappa_n}(\tau_n + \delta) \not\to_p 0$$

for all fixed $\delta > 0$. 

David Aldous
Limits for processes over general networks
Proposition (4)

Take edge-weighted graphs with \( n \to \infty \), consider the SI epidemics with transmission probabilities of form (3), and write \( C'_{n,\kappa}(\theta) \) for the number of ultimately infected individuals in the epidemic started with \( \kappa \) uniformly random infected individuals. Suppose there exist some \( 0 < \theta_1 < \theta_2 < \infty \) such that, for all \( \kappa_n \) with \( \kappa_n/n \to 0 \) sufficiently slowly,

\[
\lim_{n} n^{-1} \mathbb{E} C'_{n,\kappa_n}(\theta_1) = 0; \quad \liminf_{n} n^{-1} \mathbb{E} C'_{n,\kappa_n}(\theta_2) > 0.
\]

Then there exist deterministic \( \tau_n \in [\theta_1, \theta_2] \) such that, for all \( \kappa_n \to \infty \) sufficiently slowly,

\[
n^{-1} C'_{n,\kappa_n}(\tau_n - \delta) \to_p 0, \quad n^{-1} C'_{n,\kappa_n}(\tau_n + \delta) \gg_p 0
\]
for all fixed \( \delta > 0 \).
Propositions 3 and 4 provide a subcritical/supercritical dichotomy for the SI epidemics under consideration. The conceptual point is that, for virulence parameter $\theta$ not close to the critical value $\tau_n$, either almost all or almost none of the realizations of the epidemic affect a non-negligible proportion of the population. It really is a phase transition.

Note: for epidemics the “few initial infectives” seems more realistic; this is not the “giant component” version familiar in the math theory.

The central point of this talk is that the format of Propositions 3 and 4 suggest conjectures for analogous “general network” results in other sub/supercritical settings, such as SIS epidemics (next slides). But different proofs are apparently required: the magic tricks don’t extend.
An SIS model (contact process): Given a network (finite connected edge-weighted graph) and a rate function $\mu_v$ on vertices $v$. Introduce a parameter $0 < \theta < \infty$ and a (small) parameter $\varepsilon > 0$.

- Each $v$ is in state $S$ (susceptible) or $I$ (infected); transition rates at $v$ as follows.
  - $I \rightarrow S$ at rate $\mu_v$.
  - $S \rightarrow I$ at rate $\varepsilon + \theta \sum \{ w_{vy} : y \text{ infected} \}$.

Conceptually, you get infected by your contacts with “virulence” parameter $\theta$, or from “outside” with low probability.

Mathematically this is a finite state Markov chain and so has a stationary distribution; we study $X_{\theta, \varepsilon} = \text{number of infected vertices, at stationarity}$.
Now consider a sequence of such networks/rate functions, indexed by $n = \text{number of vertices}$. The basic assumption we will make is: there exist $0 < \theta_* < \theta^* < \infty$ such that, for every sequence $\epsilon_n \downarrow 0$ sufficiently slowly,

$$n^{-1} X_{\theta^*, \epsilon_n}^{(n)} \to 0 \text{ in probability}; \quad n^{-1} X_{\theta^*, \epsilon_n}^{(n)} \gg p 0. \quad (6)$$

**Conjecture**

*Under assumption (6) (and perhaps further but weak assumptions), there exist $\theta_n \in [\theta_*, \theta^*]$ such that, for all $\epsilon_n \downarrow 0$ sufficiently slowly,*

$$n^{-1} X_{\theta_n - \delta, \epsilon_n}^{(n)} \to 0 \text{ in probability}; \quad n^{-1} X_{\theta_n + \delta, \epsilon_n}^{(n)} \gg p 0 \ \forall \delta > 0.$$ 

It is not clear whether this can be proved via the LWC technique.
To me a **network** is a finite \((n\) vertices) connected edge-weighted undirected graph, vertices \(v, x, y, \ldots\) and edge weights \(w_e = w_{xy}\).

To me there are three fundamental random processes one can define over a network:

- the (continuous-time) Markov chain.
- bond percolation
- first passage percolation.

All of these are (somewhat) amenable to the methodologies in this talk; I have focused on bond percolation because it suggests conjectures for more general epidemic-type models.

In fact there is an analogous “concentration” result for first-passage percolation times on general networks – see *Weak Concentration for First Passage Percolation Times on Graphs ... (2016)*. Instead I’ll describe a loosely-similar problem involving continuous-time Markov chains.
The Markov chain on a network simply uses edge-weights as transition rates:

\[ x \rightarrow y \text{ rate } w_{xy} \]

and has uniform stationary distribution; indeed this is the general form of a reversible chain with uniform stationary distribution.

A “compactification” result conjectured by me and proved in a weak form by Henry Towsner (Limits of sequences of Markov chains, Electron. J. Probab. 2015).

**Theorem**

An arbitrary sequence of networks with \( n \rightarrow \infty \) has a subsequence in which (after time-scaling) the Markov chain either

- has the \( L^2 \) cutoff property
- or converges (in a certain subtle sense) to a limit Markov process of the form described below.
The form of the limit process?

Important note: this here is purely Measure Theoretic – no topology. So we can take state space as \([0, 1], \mathcal{B}, \text{Leb}\). Consider measurable functions \(p^\infty(x, y, t)\) for \(x, y \in [0, 1]\) and \(t > 0\) such that

- \(p^\infty(x, y, t) \equiv p^\infty(y, x, t)\).
- \(y \to p^\infty(x, y, t)\) is a probability density function.
- \(p^\infty(x, z, t + s) = \int p^\infty(x, y, t)p^\infty(y, z, s)dy\). (Chapman-Kolmogorov)
- some \(t \downarrow 0\) pinning.

This specifies the finite-dimensional distributions of a symmetric Markov process on \([0, 1]\) started at \(x\).
The proof uses the 1980s Hoover-Aldous-Kallenberg work on structure of random arrays with an exchangeability property

\[(Z_{ij}) = d (Z_{\pi(i)\pi(j)})\] for all permutations \(\pi\).

For the chain on a finite network define

\[p^n(x, y, t) = n P(X_t = y | X_0 = x).\]

Take \(V_i\) IID uniform on the \(n\) states and define

\[Z_{ij}^n\] is the random function \(t \rightarrow p^n(V_i, V_j, t)\) \hspace{1cm} (7)

**Towsner's theorem**: for arbitrary sequence of networks there is a subsequence with either the \(L^2\) cut-off property or with \((Z_{ij}^n) \rightarrow (Z_{ij}^\infty)\), the limit defined as at (7) with IID uniform\([0,1]\) \((V_i)\).
Not easy to see what this means . . . . .

but the point is that the sampled transition densities \((Z_{ij}^n)\) in the finite case identify the Markov chain “up to relabeling of states”, that is up to a bijection \(\phi : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}\).

So the limit \((Z_{ij}^\infty)\) identifies a process on states \([0, 1]\) “up to measure-preserving transformation of \([0, 1]\)”.

As mentioned before, this is all measure-theoretic, and (in the spirit of the famous quote *History doesn’t repeat itself but it often rhymes*) Towsner repeats Hoover in giving a “logic” proof.

It would be nice to have a “standard” proof.
Conjecture  There is some natural way to define a topology, for instance
\[
d(x_1, x_2) := \sqrt{\int \int (p^\infty(x_1, y, t) - p^\infty(x_2, y, t))^2 e^{-t} dy dt}
\]
which makes \([0, 1]\) into a complete separable metric space and makes the
Markov process have the Feller property.

This in turn could used to define a distance function of the vertices of
approximating networks.