A Phase Transition In Random Coin Tossing

David Levin, Robin Pemantle, Yuval Peres

June 20, 1998

http://www.stat.berkeley.edu/∼peres

Kakutani's Theorem (1948).

μ_0 = law of i.i.d. fair coin tosses $\{X_n\}_{n=0}^{\infty}$:

$$X_n \sim \begin{cases} +1 & \text{with prob. } \frac{1}{2} \\ -1 & \text{with prob. } \frac{1}{2} \end{cases}.$$

ν_θ = law of independent coin tosses $\{X_n\}_{n=0}^{\infty}$ with biases $\{\theta_n\}_{n=0}^{\infty}$:

$$X_n \sim \begin{cases} +1 & \text{with prob. } \frac{1+\theta_n}{2} \\ -1 & \text{with prob. } \frac{1-\theta_n}{2} \end{cases}.$$

(i) $\sum_{n=1}^{\infty} \theta_n^2 = \infty \implies \nu_\theta \perp \mu_0$.

(ii) $\sum_{n=1}^{\infty} \theta_n^2 < \infty \implies \nu_\theta \ll \mu_0$ and $\mu_0 \ll \nu_\theta$.

The Model.
First analyzed by Harris and Keane (1997).

- Coins are $\{-1, 1\}$-valued r.v.s.

$$\textbf{fair coin}: \quad Prob(+1) = \frac{1}{2}$$

$$\textbf{coin with bias } \theta: \quad Prob(+1) = \frac{1+\theta}{2}$$

- Suppose $\{\Gamma_n\}$ is a recurrent, countable state Markov chain started at $\Gamma_0 = o$.

- When $\Gamma_n = o$, a θ-biased coin is tossed.

- When $\Gamma_n \neq o$, a fair coin is tossed.

- The observer only sees the results of the tosses $\{X_n\}_{n=0}^{\infty}$. The law of $\{X_n\}$ is μ_θ.

$\Delta_n = \mathbf{1}_{\{\Gamma_n = o\}}$ is a renewal sequence.
Any renewal sequence can be represented this way.

Question:

When can a sample from μ_θ be distinguished from i.i.d. fair coins $\sim \mu_0$?

Standard 0-1 Laws imply

$$\mu_\theta \perp \mu_0$$

or

$$\mu_\theta \ll \mu_0 \text{ and } \mu_0 \ll \mu_\theta\,.$$

So we can rewrite our question:

When is $\mu_\theta \perp \mu_0$?

Related models analyzed by Benjamini and Kesten(1996) and Howard(1996).

Theorem (Harris and Keane (1997))

(i) $\sum_{n=1}^{\infty} u_n^2 = \infty \implies \mu_\theta \perp \mu_0$.

(ii) $\sum_{n=1}^{\infty} u_n^2 = ||u||^2 < \infty$
$\implies \mu_\theta \ll \mu_0$ for $\theta < \frac{1}{||u||}$.

Harris and Keane conjectured that

$$\sum_{n=0}^{\infty} u_n^2 < \infty \implies \mu_\theta \ll \mu_0 \, ,$$

and in particular that for $\theta \neq 0$, the value of θ plays no role in deciding whether $\mu_\theta \perp \mu_0$.

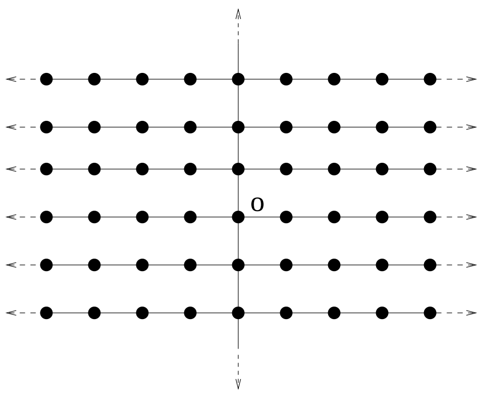
Theorem 1 *If* $u_n \asymp n^{-\gamma}$ *for* $1/2 < \gamma < 1$ *and* u_1 *is large, then*

(i) $\theta > \frac{2^\gamma}{u_1} - 1 \implies \mu_\theta \perp \mu_0$.

(ii) The bias θ *can be recovered from data if it is large enough:* $\exists\, g$ *so that*

$$\theta > \frac{2^\gamma}{u_1} - 1 \implies \theta = g(X)\, \mu_\theta - a.s.$$

Remark. There exist such $\{u_n\}$. For instance, if Γ_n is the "'"Fishbone" depicted below, then $u_n = \mathbf{P}[\Gamma_n = o] \asymp n^{-3/4}$.



Definition: There is a **phase transition** in θ when

$\exists\, \theta_c \in (0,1)$ so that

$$\begin{aligned} \theta < \theta_c &\implies \mu_\theta \ll \mu_0 \\ \theta > \theta_c &\implies \mu_\theta \perp \mu_0 \,. \end{aligned}$$

$\theta_s = (\sum_{n=1}^{\infty} u_n^2)^{-1/2}$ is the critical parameter for μ_θ to have a square-integrable density with respect to μ_0.

asymptotics of u_n	**critical parameters**
$u_n \asymp n^{-1/2}$	$0 = \theta_s = \theta_c$
$u_n \asymp n^{-\gamma},\ \frac{1}{2} < \gamma < 1$	$0 < \theta_s \le \theta_c \le u_1^{-1} 2^\gamma - 1$
$u_n \asymp n^{-1}$	$0 < \theta_s \le \theta_c \stackrel{?}{=} 1$
$u_n = o\left(\frac{1}{n \log\log n}\right)$	$0 < \theta_s \le \theta_c = 1$

Sometimes $\theta_s < \theta_c$.

Harris and Keane asked if it is possible to reconstruct θ when $\sum u_n^2 = \infty$.

We prove

Theorem 2 *If* $\sum_n u_n^2 = \infty$, *then*

$$\exists h \;\; \textit{with} \;\; \theta = h(X) \;\; \mu_\theta - a.s. \;\; \forall\, \theta\,.$$

Under a further hypothesis, this was obtained independently by Klaasen.

Proof of Harris and Keane (ii):

$$\sum_{n=0}^{\infty} u_n^2 < \infty \quad \text{and} \quad \theta \text{ small } \Rightarrow \mu_\theta \ll \mu_0.$$

Let $\{\Delta'_n\}$ be an independent copy of $\{\Delta_n\}$, and

$$J = \sum_{n=0}^{\infty} \Delta_n \Delta'_n \, .$$

J, the number of joint renewals, is geometric with mean $\sum u_n^2$.

If $\rho_n = \frac{d\mu_\theta}{d\mu_0}|\mathcal{F}_n$,

$$\rho_n(x) = \int_{\{0,1\}^\infty} \prod_{k=0}^{n} (1 + \theta x_k \Delta_k) dP(\Delta_k) \,.$$

By Fubini, $\int_{\{-1,1\}^\infty} \rho_n^2(x) d\mu_0(x)$

$$\begin{aligned}
&= \int\int \prod_{k=0}^{n} (1 + \theta^2 \Delta_k \Delta_k') d\mathbf{P}(\Delta) d\mathbf{P}(\Delta') \\
&\leq \sum_{j=1}^{\infty} (1+\theta^2)^j \mathbf{P}\left[J = j\right] \\
&\leq \sum_{j=1}^{\infty} (1+\theta^2)^j \mathbf{P}\left[J > 1\right]^{j-1} \\
&< \infty \,,
\end{aligned}$$

provided that

$$1 + \theta^2 < \frac{1}{\mathbf{P}[J > 1]} = \sum_{n=0}^{\infty} u_n^2 \,.$$

$\square$

Proof Ideas for Theorem 1 (i):

Assumption: $u_n \sim n^{-\gamma}$ for some $1/2 < \gamma < 1$

$R_n =$ length of run of '$+1$'s starting at n

Weak form of Erdős and Rényi: For i.i.d. unbiased sequences,

$$\limsup_n \frac{R_n}{\log_2 n} = 1 \qquad \mu_0 - a.s.$$

We show that for sequences under μ_θ:

$$\limsup_n \frac{R_n}{\log_2 n} = \hat{R}(\theta) > 1 \qquad \mu_\theta - a.s.$$

provided $\theta > \frac{2^\gamma}{u_1} - 1$.

Note that $\hat{R}(\theta)$ is constant μ_θ-a.s. for each θ.

$\exists$ around $\sum_{k=1}^{n} k^{-\gamma} \asymp n^{1-\gamma}$ renewals before n.

Each starts run of renewals of length k with probability u_1^k.

$\Rightarrow$ about

$$n^{1-\gamma}\left[u_1 \frac{1+\theta}{2}\right]^k \tag{1}$$

runs of "+1"s of length k before time n.

The $\log_2$ of (1) is positive when

$$k < \frac{1-\gamma}{-\log_2(u_1 \frac{1+\theta}{2})} \log_2 n = c(\theta) \log_2 n \,.$$

When the hypothesis $\theta > \frac{2^\gamma}{u_1} - 1$ is satisfied, $c(\theta) > 1$. $\square$

This is not how longest runs arise. They come from long intervals with high density of renewals.

Reconstruction of θ when $u_n \asymp n^{-\gamma}$
(Recall $\frac{1}{2} < \gamma < 1$):

We show that

$$\widehat{R}(\theta) := \limsup_{n\to\infty} \frac{R_n}{\log_2 n}$$

is a strictly monotone function of θ for large θ.

Reconstruction when $\sum_{n=1}^{\infty} u_n^2 = \infty$:
Natural linear estimators:

$$T_n := \frac{\sum_{i=1}^{n} u_i X_i}{\sum_{i=1}^{n} u_i^2} .$$

Note: T_n is Minimum Variance Unbiased Estimator in i.i.d. case.

- $\mathbf{E}_0 T_n = 0$ and $\mathtt{Var}_0(T_n) \to 0$
- $\mathbf{E}_\theta T_n = \theta$ and $\mathtt{Var}_\theta(T_n) < B$ for B not depending on n.
- T_n is bounded sequence in $L^2(\mu_\theta) \Rightarrow T_n$ has a L^2-weakly convergent subsequence.
- L^2 limit T is a tail function $\Rightarrow T = \theta$ a.s.
- $\exists$ convex combination of T_n that tend in L^2 and a.s. to θ.

Comments:

- Argument only works for fixed θ. (Subsequence may depend on θ.)

- We find an explicit combination of linear estimators on disjoint segments that works simultaneously for all θ.

Connection to Percolation.

- Assume transitive Markov chain Γ.

- Form bonds between $m, \ell \in \mathbf{Z}^+$ iff

$$\Gamma_m = \Gamma_\ell, \text{ but } \Gamma_j \neq \Gamma_m \text{ for } m < j < \ell.$$

- Assign a coin with bias θ to cluster of 0, fair coin to all other clusters.

- Observer does not see bonds, must determine if $\theta \neq 0$.

- Example of a 1-dimensional, long-range, dependent percolation model with phase transition (c.f. M. Aizenman, J.T. Chayes, L. Chayes, and C.M. Newman (1988)).

Almost transient case.

We prove that

$$u_n = o\left(\frac{1}{n \log \log n}\right) \Rightarrow \mu_\theta \ll \mu_0.$$

Note: Delays will not make $\mu_\theta \perp \mu_0$.

Recall $J = \sum_{n=0}^{\infty} \Delta_n \Delta'_n$, where Δ' is an independent copy of Δ.

Proof uses

$$\mathbf{E}[r^J \mid \Delta] < \infty \Rightarrow \mu_\theta \ll \mu_0$$

for $r = (1 + \theta^2)$.

Open Problems

1. When $u_n \asymp \frac{1}{n}$ for $n \geq 1$, for which r is

$$\mathbf{E}[r^J \mid \Delta] < \infty\ ? \tag{2}$$

   (2) would imply $\mu_\theta \ll \mu_0$ for $r = 1 + \theta^2$. We conjecture that (2) holds for all r.

2. Does $\mathbf{E}[(1+\theta^2)^J \mid \Delta] = \infty$ imply that $\mu_\theta \perp \mu_0$?

3. Does $\mu_{\theta_1} \perp \mu_0$ imply that $\mu_{\theta_1} \perp \mu_{\theta_2}$ for all $\theta_2 \neq \theta_1$?

4. When $\theta_c \in (0,1)$, is $\mu_{\theta_c} \perp \mu_0$?