REU in Random Walk

Xiaoyu Liu

October 14, 2022

Contents

Note 1: Basic definitions and facts

Sun 11 Sep 2022 10:19

1 Random Walk

The probability space of consideration is

$$\begin{aligned}
&\Omega = \{(\omega_1, \omega_2, \dots) : \omega_i \in S\} \\
&\mathcal{F} = \mathcal{S} \times \mathcal{S} \times \dots \\
&P = \mu \times \mu \times \dots \quad \mu \text{ is the distribution of } X_i \\
&X_n(\omega) = \omega_n
\end{aligned} \tag{1}$$

Definition 1 (Recurrence)**.** The number $x \in \mathbf{R}^d$ is said to be a recurrent value for the random walk S_n if for every $\varepsilon > 0, P(\|S_n - x\| < \varepsilon \text{ i.o. }) = 1$. Here $\|x\| = \sup |x_i|$.

A random walk is said to be *recurrent* if the set of recurrent values is nonempty. A random walk is said to be *transient* if it is not recurrent.

ref: Durett, 2019

Theorem 1. The set $\mathcal{V}$ of recurrent values is either $\emptyset$ or a closed subgroup of $\mathbf{R}^d$. In the second case, $\mathcal{V} = \mathcal{U}$, the set of possible values.

Theorem 2 (Hewitt-Savage 0-1 law)**.** Exchangeable events (events unaffected by any finite permutation of $X_1, \ldots, X_n$) are trivial.

If $X_1, X_2, \ldots$ are i.i.d. and $A \in \mathcal{E}$ then $P(A) \in \{0, 1\}$.

Theorem 3. For any random walk, the following are equivalent:
(i) $P(\tau_1 < \infty) = 1$,
(ii) $P(S_m = 0 \text{ i.o. }) = 1$, and
(iii) $\sum_{m=0}^{\infty} P(S_m = 0) = \infty$.

1.0.1 Facts about Recurrence of Random Walk

Definition 2. A *simple symmetric random walk* on $\mathbf{Z}^d$ is a random walk such that

$$P(X_i = e_j) = P(X_i = -e_j) = 1/2d$$

for each of the d unit vectors e_j.

Theorem 4. Simple symmetric random walk is recurrent in $d \leq 2$ and transient in $d \geq 3$.

Remark. Quotation from Durett, 2019: To steal a joke from Kakutani (U.C.L.A. colloquium talk): "A drunk man will eventually find his way home but a drunk bird may get lost forever."

Sufficient conditions for recurrence/transience in 1D, 2D and 3D random walk:

Theorem 5 (Chung-Fuchs theorem)**.** Suppose $d = 1$. If the weak law of large numbers holds in the form $S_n/n \to 0$ in probability, then S_n is recurrent.

Theorem 6. If S_n is a random walk in $\mathbf{R}^2$ and $S_n/n^{1/2} \Rightarrow$ a nondegenerate normal distribution then S_n is recurrent.

Theorem 7. A random walk in $\mathbf{R}^3$ is truly three-dimensional if the distribution of X_1 has $P(X_1 \cdot \theta \neq 0) > 0$ for all $\theta \neq 0$. No truly three-dimensional random walk is recurrent.

1.0.2 Simple Random Walks

Definition 3. A *simple (asymmetric) random walk* on $\mathbb{Z}$ is a random walk such that

$$\begin{aligned} P(X_i = 1) &= p \\ P(X_i = -1) &= 1 - p. \end{aligned}$$

A simple random walk $(W_n)_{n\geq 0}$ can be also view as a Markov chain with transition probabilities

$$p(j,k) = \mathrm{P}\left(W_n = k \mid W_{n-1} = j\right) = \begin{cases} p & k = j+1 \\ 1-p & k = j-1 \\ 0 & \text{otherwise} \end{cases}$$

Definition 4 (Limiting Speed)**.** The limiting speed is defined as

$$\lim_{n\to\infty} \frac{S_n}{n}$$

Theorem 8. The limiting speed of a simple random walk $(W_n)_{n\geq 0}$ with probability p of stepping to the right is $2p-1$.

ref: Cinkoske, Jackson, and Plunkett, 2018

Definition 5 (Limiting Distribution)**.** If S_n converges in distribution to some S, then the distribution function F of S is the *limiting distribution* of the random walk.

Theorem 9 (Central Limit Theorem for i.i.d sequences)**.** Let $X_1, X_2, \ldots$ be i.i.d. with $EX_i = \mu, \operatorname{var}(X_i) = \sigma^2 \in (0,\infty)$. If $S_n = X_1 + \cdots + X_n$ then

$$\left(S_n - n\mu\right)/\sigma n^{1/2} \Rightarrow \chi$$

where χ has the standard normal distribution.

Example. (Limiting Distribution for Simple Random Walk)

Since $X_1, \ldots, X_n$ are all i.i.d, we can use central limit theorem for i.i.d sequences. Assume $p \neq \frac{1}{2}$,

$$S_n - n(2p-1)/\sqrt{4p(1-p)n} \Rightarrow \chi.$$

Where $\mathcal{N}(\mu, \sigma^2)$ is normal distribution with mean μ and variance σ^2.

When $p = \frac{1}{2}$, we are in the symmetrical case. The limiting distribution is

$$S_n/\sqrt{n} \Rightarrow \chi.$$

$\diamond$

Note 2: Brownian Motion

Tue 20 Sep 2022 09:49

2 Brownian Motion

Definition 6 (Brownian Motion)**.** A one-dimensional Brownian motion is a real-valued process $B_t, t \geq 0$ that has the following properties:
(a) If $t_0 < t_1 < \ldots < t_n$ then $B(t_0), B(t_1) - B(t_0), \ldots, B(t_n) - B(t_{n-1})$ are independent.
(b) If $s, t \geq 0$ then

$$W_t - W_s \sim \mathcal{N}(0, t-s), (0 \leq s \leq t)$$

(c) With probability one, $t \rightarrow B_t$ is continuous.

Theorem 10 (Levy's Characterization of Brownian Motion)**.** Let $X = (X_1, \ldots, X_n)$ be a continuous stochastic process on a probability space $(\Omega, \Sigma, \mathrm{P})$ taking values in $\mathbf{R}^n$. Then the following are equivalent:

1. X is a Brownian motion with respect to $\mathbf{P}$, i.e., the law of X with respect to $\mathbf{P}$ is the same as the law of an n-dimensional Brownian motion, i.e., the push-forward measure $X_*(\mathbf{P})$ is classical Wiener measure on $C_0([0, +\infty); \mathbf{R}^m)$.
2. both
   (a) X is a martingale with respect to $\mathbf{P}$ (and its own natural filtration); and
   (b) for all $1 \leq i, j \leq n, X_i(t)X_j(t) - \delta_{ij}$ i is a martingale with respect to $\mathbf{P}$ (and its own natural filtration), where δ_{ij} denotes the Kronecker delta.

Translation invariance. $\{B_t - B_0, t \geq 0\}$ is independent of B_0 and has the same distribution as a Brownian motion with $B_0 = 0$.

The **Brownian scaling relation.** If $B_0 = 0$ then for any $t > 0$,

$$\{B_{st}, s \geq 0\} \stackrel{d}{=} \left\{t^{1/2} B_s, s \geq 0\right\}$$

Markov Property. if $s \geq 0$ then $B(t+s) - B(s), t \geq 0$ is a Brownian motion that is independent of what happened before time s.

Reflection Principle. (Example 8.4.1.) Let $a > 0$ and let $T_a = \inf\{t : B_t = a\}$. Then

$$P_0(T_a < t) = 2P_0(B_t \geq a)$$

Intuitive proof. We observe that if B_s hits a at some time $s < t$, then the strong Markov property implies that $B_t - B(T_a)$ is independent of what happened before time T_a. The symmetry of the normal distribution and $P_a(B_u = a) = 0$ for $u > 0$ then imply

$$P_0(T_a < t, B_t > a) = \frac{1}{2} P_0(T_a < t)$$

Rearranging the last equation and using $\{B_t > a\} \subset \{T_a < t\}$ gives

$$P_0\left(T_a < t\right) = 2P_0\left(T_a < t, B_t > a\right) = 2P_0\left(B_t > a\right)$$

□

Theorem 11. (1-d BM is unbounded) (8.2.8) Let B_t be a one-dimensional Brownian motion starting at 0 then with probability 1 ,

$$\limsup_{t\to\infty} B_t/\sqrt{t} = \infty \quad \liminf_{t\to\infty} B_t/\sqrt{t} = -\infty$$

Theorem 12. (1-d BM is recurrent) (8.2.9) Let B_t be a one-dimensional Brownian motion and let $A = \cap_n \{B_t = 0 \text{ for some } t \geq n\}$. Then $P_x(A) = 1$ for all x.

Theorem 13. (Relating zero and infinity) If B_t is a Brownian motion starting at 0, then so is the process defined by $X_0 = 0$ and $X_t = tB(1/t)$ for $t > 0$.

Corollary 1. B_t reaches 0 infinitely many times by time $\varepsilon > 0$.

Properties of zero set. The zero set $\mathcal{Z}(\omega) \equiv \{t : B_t(\omega) = 0\}$ has no isolated points, is uncountable, is of measure zero, and has Hausdorff dimension $\frac{1}{2}$. (Compare to Cantor's Set with Hausdorff dimension $\log 2/\log 3$.

2.0.1 Brownian Motion and Random Walk

We can construct brownian motion by taking scaling limit of random walk:

In probability theory, Donsker's theorem (also known as Donsker's invariance principle, or the functional central limit theorem), named after Monroe D. Donsker, is a functional extension of the central limit theorem.

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance 1. Let $S_n := \sum_{i=1}^n X_i$. The stochastic process

$S := (S_n)_{n\in\mathbb{N}}$ is known as a random walk. Define the diffusively rescaled random walk (partialsum process) by

$$W^{(n)}(t) := \frac{S_{\lfloor nt\rfloor}}{\sqrt{n}}, \quad t \in [0,1]$$

The central limit theorem asserts that $W^{(n)}(1)$ converges in distribution to a standard Gaussian random variable $W(1)$ as $n \to \infty$. Donsker's invariance principle extends this convergence to the whole function $W^{(n)} := \left(W^{(n)}(t)\right)_{t\in[0,1]}$. More precisely, in its modern form, Donsker's invariance principle states that:

Theorem 14 (Donsker's Invariance Principle)**.** As random variables taking values in the Skorokhod space $\mathcal{D}[0,1]$ (right continuous and have left limit everywhere), the random function $W^{(n)}$ converges in distribution to a standard Brownian motion $W := (W(t))_{t\in[0,1]}$ as $n \to \infty$.

source: *Donsker's Theorem* 2021

We can also embed a random walk inside a brownian motion:

Skorohod Embedding. Suppose we are given a standard Brownian motion (B_t), and a stopping time T. Then, so Iong as T satisfies one of the regularity conditions under which the Optional Stopping Theorem applies, we know that $\mathbb{E}\left[B_T\right] = 0$. Furthermore, since $B_t^2 - t$ is a martingale, $\mathbb{E}\left[B_T^2\right] = \mathbb{E}[T]$, so if the latter is finite, so is the former.

Now, using the strong Markov property of Brownian motion, we can come up with a sequence of stopping times $0 = T_0, T_1, T_2, \ldots$ such that the increments $T_k - T_{k-1}$ are IID with the same distribution as T. Then $0, B_{T_1}, B_{T_2}, \ . \ .$ is a centered random walk. By taking T to be the hitting time of $\{-1, +1\}$, it is easy to see that we can embed simple random walk in a Brownian motion using this approach.

copied from: dominicyeo, 2016. For more information see: Obloj, 2004

Note 3: Autocorrelation Function

Wed 12 Oct 2022 18:36

3 Autocorrelation Function

Definitions and examples are mostly copied from Brockwell and Davis, 2016. Essential definitions are given below.

Definition 7. (Mean and Covariance) Let $\{X_t\}$ be a time series with $E\left(X_t^2\right) < \infty$. The mean function of $\{X_t\}$ is

$$\mu_X(t) = E\left(X_t\right).$$

The covariance function of $\{X_t\}$ is

$$\gamma_X(r,s) = \operatorname{Cov}\left(X_r, X_s\right) = E\left[\left(X_r - \mu_X(r)\right)\left(X_s - \mu_X(s)\right)\right]$$

for all integers r and s.

Definition 8 (Stationary Time Series)**.** $\{X_t\}$ is (weakly) stationary if (i) $\mu_X(t)$ is independent of t, and (ii) $\gamma_X(t+h,t)$ is independent of t for each h.

Definition 9 (Autocorrelation Function)**.** Let $\{X_t\}$ be a stationary time series. The autocovariance function (ACVF) of $\{X_t\}$ at lag h is

$$\gamma_X(h) = \operatorname{Cov}\left(X_{t+h}, X_t\right).$$

The autocorrelation function (ACF) of $\{X_t\}$ at lag h is

$$\rho_X(h) \equiv \frac{\gamma_X(h)}{\gamma_X(0)} = \operatorname{Cor}\left(X_{t+h}, X_t\right)$$

Example. 1. (i.i.d. noise) If $\{X_t\}$ is iid noise and $E\left(X_t^2\right) = \sigma^2 < \infty$,we have

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2, & \text{if } h = 0 \\ 0, & \text{if } h \neq 0 \end{cases}.$$

Thus

$$\rho_X(h) = \begin{cases} 1 & h = 0 \\ 0 & h \neq 0 \end{cases}.$$

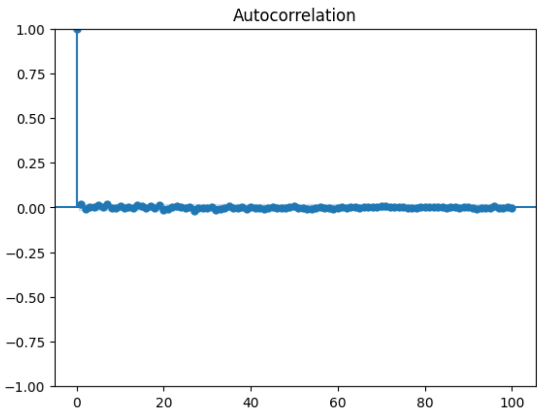

Figure 1: Example of a calculated ACF for i.i.d process

2. (First-Order Moving Average or MA(1) Process) Consider the series defined by the equation

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \ldots,$$

where $\{Z_t\} \sim \mathrm{WN}\left(0, \sigma^2\right)$ and θ is a real-valued constant. From above we see that $EX_t = 0, EX_t^2 = \sigma^2\left(1+\theta^2\right) < \infty$, and

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2\left(1+\theta^2\right), & \text{if } h = 0 \\ \sigma^2\theta, & \text{if } h = \pm 1 \\ 0, & \text{if } |h| > 1 \end{cases}$$

Thus $\{X_t\}$ is stationary. The autocorrelation function of $\{X_t\}$ is

$$\rho_X(h) = \begin{cases} 1, & \text{if } h = 0 \\ \theta/\left(1+\theta^2\right), & \text{if } h = \pm 1 \\ 0, & \text{if } |h| > 1 \end{cases}$$

◇

Definition 10 (Strictly Stationary)**.** (2.1.2) $\{X_t\}$ is a strictly stationary time series if

$$(X_1, \ldots, X_n)' \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})'$$

for all integers h and $n \geq 1$. (Here $\stackrel{d}{=}$ is used to indicate that the two random vectors have the same joint distribution function.)

Proposition 1. Properties of a Strictly Stationary Time Series $\{X_t\}$:
a. The random variables X_t are identically distributed.
b. $(X_t, X_{t+h})' \stackrel{d}{=} (X_1, X_{1+h})'$ for all integers t and h.
c. $\{X_t\}$ is weakly stationary if $E\left(X_t^2\right) < \infty$ for all t.
d. Weak stationarity does not imply strict stationarity.
e. An iid sequence is strictly stationary.

Note 4: Concepts of Continuity

Wed 12 Oct 2022 18:59

4 Continuity and Hölder continuity test

4.1 Definitions and Examples

The main reference in this note is Durett, 2019.

Definition 11 (Lipschitz continuous)**.** f is said to be *Lipschitz continuous* if there is a constant C so that $|f(x) - f(y)| \leq C\rho(x, y)$.

Lipschitz continuity is a stronger notion of continuity than classical continuity. Lipschitz continuity implies continuity.

Geometrically, Lipchitz condition puts a finite bound on the slope of any secant line one can get from the graph of the function.

Definition 12. A real or complex valued function f on d-dimensional Euclidean space satisfies a Hölder condition with exponent α, or is α-Hölder continuous, when there are nonnegative real constants $C, a > 0$, such that

$$|f(x) - f(y)| \leq C\|x - y\|^{\alpha}$$

for all x and y in the domain of f.

- When $\alpha > 1$, an α-Hölder continuous function is *constant*.

- When $\alpha = 1$, an α-Hölder continuous function is *Lipchitz continuous.*
- When $\alpha > 0$, an α-Hölder continuous function is *uniformly continuous.*
- Whenver $0 < \alpha \leq \alpha'$, α'-Hölder continuity implies α-Hölder continuity.

Remark. (from wikipedia) We have the following chain of strict inclusions for functions over a closed and bounded non-trivial interval of the real line:

Continuously differentiable $\subset$ Lipschitz continuous $\subset$ α-Hölder continuous $\subset$ uniformly continuous $\subset$ continuous

where $\alpha \in (0, 1]$.

Theorem 15. 8.1.5. Brownian paths are Hölder continuous for any exponent $\gamma < 1/2$.

Theorem 16. 8.1.6. With probability one, Brownian paths are not Lipschitz continuous (and hence not differentiable) at any point.

4.2 Test for hölder continuity exponent

From the definition of hölder continuity:

$$|f(x) - f(y)| \leq C\|x - y\|^{\alpha}$$

Our goal is to find smallest possible α such that the equation above holds true empirically.

We take logarithm on both sides (using the fact that log is monotone increasing):

$$\log|f(x) - f(y)| \leq \log C + \alpha \log\|x - y\|$$

rearrange and assuming $\log\|x - y\| < 0$ for small $\|x - y\|$:

$$\alpha \leq \frac{\log|f(x) - f(y)| - \log C}{\log\|x - y\|} \sim \frac{\log|f(x) - f(y)|}{\log\|x - y\|}. \tag{2}$$

The smallest such α is then

$$\alpha = \inf_{x,y} \frac{\log|f(x) - f(y)| - \log C}{\log\|x - y\|}. \tag{3}$$

The term involving $\log C$ should fade away when $\|x - y\|$ is sufficiently small.

4.2.1 Compute α by finding a limit

Let $x_1, \ldots, x_i, \ldots, x_n$ be random points in the domain of f. Let $\varepsilon_1, \ldots, \varepsilon_j, \ldots, \varepsilon_m$ be small changes.

Denote

$$\hat{a}_i^{(j)} := \frac{\log|f(x_i) - f(x_i + \varepsilon_j)| - \log C}{\log \varepsilon_j}.$$

We know $\hat{a}_i^{(j)}$ is always an **upper estimate** of α. Thus a reasonable guess of α would be:

$$\hat{\alpha} = \inf_{i,j} \hat{a}_i^{(j)}.$$

Since $\hat{a}_i^{(j)} \sim a_i^{(j)}$ as $j \to \infty$, we can replace our estimate of α by

$$\hat{\alpha} = \lim_{j\to\infty} \inf_i \hat{a}_i^{(j)}.$$

The limit, however, converges very slowly. As ε_j decreases exponentially, $|\log \varepsilon_j|$ only increases linearly, so the error terms decreases very slowly.

4.2.2 Find α and C by Curve Fitting

Rewrite equation (3) as

$$\alpha = \inf_{x,\varepsilon} \frac{\log|f(x) - f(x+\varepsilon)| - \log C}{\log \varepsilon} \quad \text{where } \varepsilon > 0.$$

Hence

$$\alpha = \inf_{\varepsilon} \left(\inf_x \frac{\log|f(x) - f(x+\varepsilon)|}{\log \varepsilon} - \frac{\log C}{\log \varepsilon} \right).$$

We claim:

Proposition 2.

$$\alpha = \lim_{\delta\to 0} \inf_{\varepsilon<\delta} \left(\inf_x \frac{\log|f(x) - f(x+\varepsilon)|}{\log \varepsilon} - \frac{\log C}{\log \varepsilon} \right).$$

This is basically saying that all roughness must remain in arbitrarily small detail.

Proof. Fix some ε' such that for some $\delta > 0$ and all $\varepsilon < \varepsilon'$,

$$\inf_x \frac{\log|f(x) - f(x+\varepsilon)|}{\log \varepsilon} - \frac{\log C}{\log \varepsilon} > \alpha + \delta.$$

We claim that f is actually $(\alpha+\delta)$-Hölder continuous, which results in a contradiction. It suffices to show that there exists some constant C such that for all ε where $\varepsilon \geq \varepsilon'$,

$$|f(x) - f(x+\varepsilon)| \leq C\varepsilon^{\alpha+\delta}.$$

But since f is α-Hölder continuous, there exists some $C' > 0$ such that

$$|f(x) - f(x+\varepsilon)| \leq C'\varepsilon^{\alpha}.$$

Take $C = \frac{C'}{(\varepsilon')^{-\delta}}$, and we have

$$|f(x) - f(x+\varepsilon)| \leq C(\varepsilon')^{\delta}\varepsilon^{\alpha} \leq C\varepsilon^{\delta}\varepsilon^{\alpha} = C'\varepsilon^{\alpha+\delta}.$$

□

This claim allows us to take smaller and smaller ε, and be convinced that the limit infimum will be our desired α.

Proposition 3. We have the bound

$$\inf_x \frac{\log|f(x)-f(x+\varepsilon)|}{\log\varepsilon} - \frac{\log C}{\log\varepsilon} \leq \inf_\varepsilon \left(\inf_x \frac{\log|f(x)-f(x+\varepsilon)|}{\log\varepsilon} - \frac{\log C}{\log\varepsilon} \right) + O(\varepsilon)$$

under the following assumptions:

1. f has bounded total variation.

Note: I don't know how to prove this theorem. I discussed this with Zijie, and have some sketchy idea on how to prove it for brownian motions.

Using the approximation above, for some specific ε, we can approximate α by

$$\alpha \approx \inf_x \frac{\log|f(x)-f(x+\varepsilon)|}{\log\varepsilon} - \frac{\log C}{\log\varepsilon}.$$

We then use $\inf_i a_i^{(j)}$ as an upper estimation:

$$\inf_i a_i^{(j)} \approx \inf_x \frac{\log|f(x)-f(x+\varepsilon)|}{\log\varepsilon}.$$

We finally have a curve to fit:

$$\alpha + \frac{\log C}{\log\varepsilon_j} = \inf_i a_i^{(j)} + O(\varepsilon). \tag{4}$$

Note: We still need to bound the errors of the approximations above.

Thus our testing strategy is:

1. Pick some small ε_1 and randomly pick $x_1^{(1)}, \ldots, x_n^{(1)}$ in the domain of interested function.
2. Calculate
$$a_i^{(j)} := \frac{\log\left|f(x_i^{(j)}) - f(x_i^{(j)}+\varepsilon_j)\right|}{\log\varepsilon_j}.$$
3. Calculate $\inf_i a_i^{(j)}$
4. Do the same for successively smaller $\varepsilon_2, \ldots, \varepsilon_m, \ldots$ that approach 0.
5. Fit the curve (4). Then read the α and C.

References

Brockwell, Peter J. and Richard A. Davis (2016). *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Cham: Springer International Publishing. ISBN: 978-3-319-29852-8 978-3-319-29854-2. DOI: 10.1007/978-3-319-29854-2.

Cinkoske, Mike, Joe Jackson, and Claire Plunkett (2018). "On the Speed of an Excited Asymmetric Random Walk". In: *Rose-Hulman Undergraduate Mathematics Journal Rose-Hulman Undergraduate Mathematics Journal* 19, p. 25.

dominicyeo (Dec. 12, 2016). *Skorohod Embedding*. Eventually Almost Everywhere. URL: https://eventuallyalmosteverywhere.wordpress.com/2016/12/12/skorohod-embedding/ (visited on 09/20/2022).

Donsker's Theorem (Sept. 27, 2021). In: *Wikipedia*.

Durett, Rick (2019). *Probability : Theory and Examples*. 5th ed. 432 pp.

Obloj, Jan (Jan. 1, 2004). "The Skorokhod Embedding Problem and Its Offspring". In: *Probability Surveys* 1 (none). ISSN: 1549-5787. DOI: 10.1214/154957804100000060.