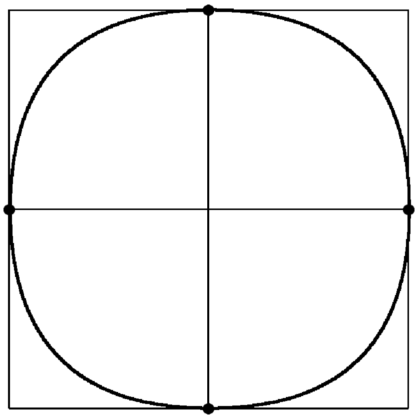
Technische Universität Graz   Institut für Mathematik

A Central Limit Theorem for Convex Chains in the Square

Imre Bárány   Günter Rote   William Steiger   Cun-Hui Zhang

Report 293   May 1998



Technische Universität Graz, Steyrergasse 30, A-8010 Graz, Austria

Version 2

This is a successor of a report with the title *The Limit Shape of Random Convex Polygons* by G. Rote from August 1994.

Title page processed by TeX on May 14, 1998

A Central Limit Theorem for Convex Chains in the Square

Imre Bárány
Mathemetical Institute
Hungarian Academy of Sciences

Günter Rote*
Institut für Mathematik
Technische Universität Graz

William Steiger
Computer Science
Rutgers University

Cun-Hui Zhang
Statistics
Rutgers University

May 1, 1998

Abstract

Points $P_1, \ldots, P_n$ in the unit square define a convex n-chain if they are below $y = x$ and, together with $P_0 = (0,0)$ and $P_{n+1} = (1,1)$, they are in convex position. Under uniform probability, we prove an almost sure limit theorem for these chains that uses only probabilistic arguments, and which strengthens similar limit shape statements established by other authors. An interesting feature is that the limit shape is a direct consequence of the method. The main result is an accompanying central limit theorem for these chains. A weak convergence result implies several other statements concerning the deviations between random convex chains and their limit.

1 Introduction and Summary

Take n points in the unit square in the plane. Write them in order of increasing x-coordinate as $P_1, \ldots, P_n$ and let $P_0 = (0,0)$ and $P_{n+1} = (1,1)$. The points are the vertices of a *convex n-chain* if the vectors $P_{i+1} - P_i$ have increasing slope, $i = 0, \ldots, n$. The chain itself is the set of points on the segments that connect successive vertices. To sample a *random n-chain C*, just take the n points uniformly and independently from the unit square, conditional on the event E that they form a convex chain. This defines the uniform probability on convex n-chains and we refer to it as the *uniform model* for chains. The event E occurs rarely because, as shown in Section 2 by elementary methods,

Theorem 1 *Let $P_1, \ldots, P_n$ be a sample of n points, independently and uniformly distributed in $[0,1]^2$. Then*

$$\text{Prob}(\textit{the sample forms a convex } n\textit{-chain}) = \frac{1}{n!(n+1)!}$$

In deriving this result the vertices on random convex n-chains are revealed as quantiles of n uniform $[0,1]$ random variables.

*Institut für Mathematik, Technische Universität Graz, Steyrergasse 30, A-8010 Graz, Austria; e-mail: `rote@opt.math.tu-graz.ac.at`.

Research of Günter Rote was supported by the Fonds zur Förderung der wissenschaftlichen Forschung, project P8971-PHY.

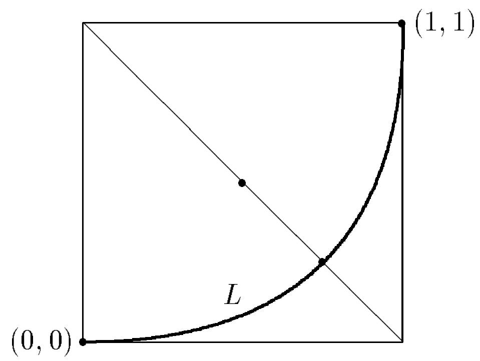


Figure 1: The limit shape L.

Random convex chains have a limit in a rather strong sense. We define the parabolic arc

$$L = \{ (x, y) : \sqrt{y} = 1 - \sqrt{1-x},\ 0 \leq x \leq 1 \}. \tag{1}$$

of points in the square equidistant from $(1/2, 1/2)$ and the line $y = x - 1$, see Figure 1. Denoting the Hausdorff distance by δ, we prove the following statement in Section 3.

Theorem 2 *For each n let C_n be a random convex n-chain. Then*

$$\operatorname{Prob}(\delta(C_n, L) \to 0) = 1.$$

Thus sequences $C_1, C_2, \ldots$ of random, convex chains converge to L with probability 1, an analogue of the strong law of large numbers. The curve L is called the *limit shape*. It is interesting that the proof technique derives L directly.

In Section 4 we prove our main result, which shows that deviations between random chains and the limit shape are asymptotically normally distributed in the following sense. For $t \in [0, 1]$, $x_t = 2t - t^2$ and $y_t = t^2$ are the coordinates of the point on L where the tangent slope is $t/(1-t)$. Then the difference between (x_t, y_t) and the vertex on the random chain where the tangent slope is $t/(1-t)$ converges in distribution to a bivariate normal vector with mean $(0, 0)$.

The technique used to establish these results can be pushed further without much difficulty. In Section 5 we show that random chains converge weakly as stochastic processes, and then use the invariance principle to obtain results for various functionals, e.g., the area between a convex chain and L.

In the remainder of this introduction we discuss the context for the above theorems and mention some previous, related results. Most pertain to the *lattice model* of random chains and respond to a question posed by Vershik about 15 years ago: "Is there a limit shape for the set of convex lattice polygons contained in a given convex body $K \subset R^2$?" Let $K = [0, 1]^2$ denote the unit square and $\mathcal{P}_n$, the set of all (upward) convex polygonal paths in K that connect $(0, 0)$ to $(1, 1)$ and whose vertices are in $\frac{1}{n}Z^2$. Bárány [1], Sinai [7], and Vershik [10] each proved theorems giving a positive answer to the question. It is shown for example, that for any $\varepsilon > 0$,

$$\frac{|\{P \in \mathcal{P}_n : \delta(P, L) < \varepsilon\}|}{|\mathcal{P}_n|} \to 1 \tag{2}$$

as $n \to \infty$. In other words a random convex lattice chain is close to the limit shape L with (uniform) probability converging to 1. A main difference between (2) and Theorem 2 is that in the lattice model the number of vertices on a chain $P \in \mathcal{P}_n$ is a random variable. This

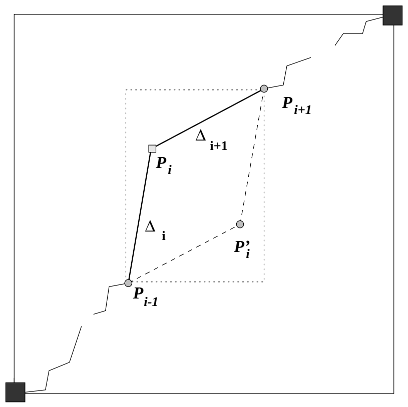


Figure 2: Permuting segments in monotone chains

variable was studied in Sinai's paper [7] where in addition, a central limit theorem for the deviations between P and L was stated. A further development in the lattice model appeared in Bárány [2] where a statement like (2) was shown to hold for *every* compact, convex body $K \subset R^2$ with nonempty interior. In addition he characterized the limit shape as the convex curve with maximal affine perimeter.

Finally, Theorem 2 may be regarded as a strengthening of the following recent result.

Proposition 1 *(Bárány [3]) For every $\varepsilon > 0$,* $\text{Prob}\,(\delta(C_n, L) > \varepsilon) \to 0$*;*

Like (2), this is a weak law of large numbers but here, it pertains to the uniform model for chains.

2 The Uniform Distribution on Chains

From now on $P_i = (x_i, y_i)$, $i = 1, \ldots, n$ will denote a sample of n points taken independently and uniformly from $[0,1]^2$, and numbered so that $x_1 \leq \cdots \leq x_n$. We write $P_0 = (0,0)$ and $P_{n+1} = (1,1)$. The sample space is $S = \{\, z = (x_1, \ldots, x_n; y_1, \ldots, y_n) : x_i, y_i \in [0,1],\ x_i$ increasing$\,\}$; probability is Lebesgue measure, normalized so $\text{Prob}(S) = 1$.

Proof of Theorem 1: By definition, the P_i are vertices on a convex chain C_n only if the slopes of the difference vectors $\Delta_i \equiv P_i - P_{i-1}$ are increasing, $i = 1, \ldots, n+1$. For this it is necessary that the sample define a *monotone chain*; i.e., the y_i are non-decreasing. Otherwise, Δ_1 has positive slope but some Δ_i will have negative slope. The probability that a sample defines a monotone chain is $(n!)^{-1}$ since the subset $M \subset S$ where $y_1 \leq \cdots \leq y_n$ clearly has the same probability as the subset where $y_{\pi_1} \leq \cdots \leq y_{\pi_n}$, for any permutation π.

Now we condition on the event $z \in M$, that the sample defines a monotone chain. We make the following **Claim:** *On the event M, all permutations of the slopes of the segments Δ_i are equally likely.* The proof is based on an idea of Valtr [9] who made a similar statement for increasing paths in a lattice. This fact will complete the proof of the theorem. First note that points in M are in one-to-one correspondence with the set $D = \{\, \underline{\Delta} = (\Delta_1, \ldots, \Delta_{n+1}) : \Delta_i = (u_i, v_i), u_i, v_i \geq 0$, and $\sum_{i=1}^{n+1} \Delta_i = (1,1)\,\}$. For a chain in $z \in M$ with differences $\underline{\Delta}$, if we interchange Δ_i and Δ_{i+1} (see Figure 2), then

1. the vertices $P_0, \ldots, P_{i-1}$ and $P_{i+1}, \ldots, P_{n+1}$ remain fixed.

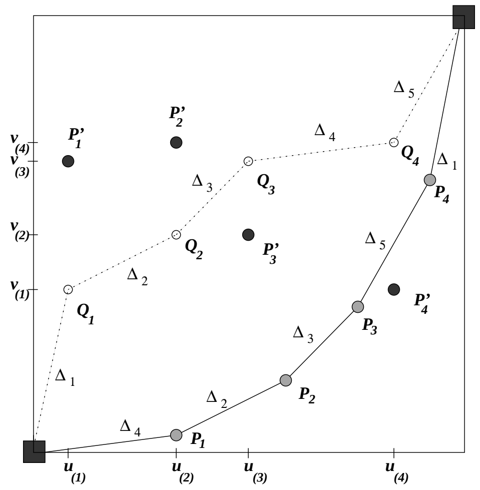


Figure 3: A sample of 4 points, its random monotone chain (dotted), and its convex chain (solid).

2. P_i is reflected at $(P_{i-1} + P_{i+1})/2$ to P'_i.

3. in the new chain, the ranks of the slopes of Δ_i and Δ_{i+1} are interchanged.

Therefore, since P_i is uniform in the rectangle with corners at P_{i-1} and P_{i+1}, given the other points, the chains $z \in M$ whose differences have slopes with ranks given by π have the same probability as the chains whose slopes obey π', a permutation differing from π by a single transposition. Because all permutations may be obtained by a sequence of such transpositions, the claim, and thus the theorem, is proved. ■

Remark 1: By definition, a random convex chain C_n may be generated as a sample of n points in the square, rejecting the sample if $(0,0)$, $(1,1)$ and the points are not in convex position; the P_i in an accepted sample are the internal vertices of the convex chain. Theorem 1 implies that the expected number of samples until one is accepted is $n!(n+1)!$. On the other hand, the proof suggests a more efficient algorithm in which a single random sample is transformed into a convex chain:

1. Generate $P'_i = (u_i, v_i)$, $i = 1, \ldots, n$, a sample of n points uniformly distributed in the square.

2. Writing $u_{(i)}$ for the i^{th} smallest among $u_1, \ldots, u_n$ [it is called the i^{th} *order statistic*] and $v_{(i)}$, the i^{th} smallest among $v_1, \ldots, v_n$ [also an i^{th} order statistic], $Q_i = (u_{(i)}, v_{(i)})$ denotes the n internal vertices on a random, monotone chain in M (dotted line in Figure 3).

3. Compute difference vectors $\Delta_i = Q_i - Q_{i-1}$, $i = 1, \ldots, n+1$, let $\Delta_{(j)}$ be the vector with the j^{th} smallest slope, and compute $P_i = \Delta_{(1)} + \cdots + \Delta_{(i)}$, $i = 1, \ldots, n$. These are the internal vertices — in order of increasing x-coordinate — of a random convex chain C_n (solid line in Figure 3).

3 A Limit Shape Theorem

The proofs of Theorems 2 and 3 are direct, once we have a more convenient representation for the vertices of a convex chain. By definition, the sample space for convex n-chains is $S \subset R^{2n}$ defined by

$$S = \{(x_1, \ldots, x_n; y_1, \ldots, y_n) : x_i, y_i \in [0,1], x_i, y_i \text{ increasing, and } \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \text{ increasing}\};$$

probability is Lebesgue measure, normalized so $\mathrm{Prob}(S) = 1$. Given n, let (Ω, μ) be a probability space on which we define two sequences $X_1, \ldots, X_{n+1}$ and $Y_1, \ldots, Y_{n+1}$ of mutually independent random variables, each exponentially distributed; i.e., $\mu\{w \in \Omega : X_i(w) \leq t\} = F(t) = 1 - e^{-t}$. For each i, write

$$\begin{aligned} R_i &= \frac{Y_i}{X_i} \\ W_i &= X_i + Y_i \\ Z_i &= \frac{R_i}{1 + R_i} = \frac{Y_i}{W_i}. \end{aligned} \tag{3}$$

It is not difficult to verify three known facts concerning exponential variables.

1. Z_i is uniformly distributed in $[0, 1]$.
2. Z_i and W_i are independent.
3. For every $n > 0$, $k \leq n$, and any permutation $(j_1, \ldots, j_{n+1})$ of $1, \ldots, n+1$, the ratio

$$\frac{Y_{j_1} + \cdots + Y_{j_k}}{Y_1 + \cdots + Y_{n+1}} \tag{4}$$

   is distributed like the k^{th} order statistic of n independent, uniform random variables.

Write I_A for the indicator of the event $A \subset \Omega$ and fix $n > 0$. For $t \in (0,1)$ define the functions

$$x_n(t) \equiv \frac{\sum_{i=1}^{n+1} W_i(1 - Z_i) I_{[Z_i \leq t]}}{\sum_{i=1}^{n+1} W_i(1 - Z_i)} = \frac{\sum_{i=1}^{n+1} X_i I_{[Z_i \leq t]}}{\sum_{i=1}^{n+1} X_i} \tag{5}$$

and

$$y_n(t) \equiv \frac{\sum_{i=1}^{n+1} W_i Z_i I_{[Z_i \leq t]}}{\sum_{i=1}^{n+1} W_i Z_i} = \frac{\sum_{i=1}^{n+1} Y_i I_{[Z_i \leq t]}}{\sum_{i=1}^{n+1} Y_i}. \tag{6}$$

These functions describe the vertices of a random convex n-chain.

Lemma 1 *Let $X_1, \ldots, X_{n+1}$ and $Y_1, \ldots, Y_{n+1}$ be i.i.d. exponential random variables on (Ω, μ) with ratios $R_i = Y_i/X_i$ and write $Z_i = R_i/(1 + R_i)$. Let $t_1 < \cdots < t_{n+1}$ be the ordered values of $Z_1, \ldots, Z_{n+1}$. The points*

$$P_i = (x_n(t_i), y_n(t_i)), \ i = 1, \ldots, n$$

are in convex position, and for any measurable $A \subset S$,

$$\mu\{w : (x_n(t_1), \ldots, x_n(t_n); y_n(t_1), \ldots, y_n(t_n))(w) \in A\} = \mathrm{Prob}(A);$$

i.e., the P_i are the n internal vertices of a random, convex n-chain.

Proof: Let $j_1, \ldots, j_{n+1}$ be the permutation that sorts the ratios; i.e., $R_{j_1} < \cdots < R_{j_{n+1}}$. Therefore $t_i = Z_{j_i}$. Observe from (6) that $y_n(t)$ is a step function with a step of size

$$\frac{Y_{j_i}}{Y_1 + \cdots + Y_{n+1}}$$

at t_i and that on $[t_k, t_{k+1})$,

$$y_n(t) = b_k \equiv \frac{Y_{j_1} + \cdots + Y_{j_k}}{Y_1 + \cdots + Y_{n+1}},\ k = 1, \ldots, n.$$

Similarly $x_n(t)$ has a step of size

$$\frac{X_{j_i}}{X_1 + \cdots + X_{n+1}},$$

at t_i and on $[t_k, t_{k+1})$,

$$x_n(t) = a_k \equiv \frac{X_{j_1} + \cdots + X_{j_k}}{X_1 + \cdots + X_{n+1}}.$$

So as t increases from 0 to 1, $(x_n(t), y_n(t))$ "jumps" from $(0,0)$ through the set of points $P_k = (a_k, b_k)$, $k = 1, \ldots, n$, to $(1,1)$. These are the vertices on a random convex chain because from (4),

$$u_{(i)} = \frac{X_1 + \cdots + X_i}{X_1 + \cdots + X_{n+1}} \text{ and } v_{(i)} = \frac{Y_1 + \cdots + Y_i}{Y_1 + \cdots + Y_{n+1}}$$

are both distributed like the i^{th} order statistics from a sample of n independent uniforms. Also the points $Q_i = (u_{(i)}, v_{(i)})$ have differences $\Delta_i = Q_i - Q_{i-1}$ whose slopes are

$$R_i \left[\frac{X_1 + \cdots + X_{n+1}}{Y_1 + \cdots + Y_{n+1}} \right], \tag{7}$$

and they are ordered by the permutation $j_1, \ldots, j_{n+1}$. Therefore using part 2 of Remark 1,

$$P_k = \Delta_{j_1} + \cdots + \Delta_{j_k}$$

is seen to be the k^{th} vertex on a random convex n-chain, and for $t \in [t_k, t_{k+1})$, $P_k = (x_n(t), y_n(t))$. ■

To prove Theorem 2 we need the following statement; here, and throughout, $\|(x,y)\| = \sqrt{x^2 + y^2}$.

Lemma 2 *For each $t \in (0,1)$ and $\varepsilon > 0$,*

$$\text{Prob}(\|(x_n(t), y_n(t)) - (2t - t^2, t^2)\| > \varepsilon) \to 0$$

as $n \to \infty$.

Proof:

Multiply the numerator and denominator of (6) by $1/(n+1)$ and apply the (weak) law of large numbers to each to observe

$$y_n(t) \to \frac{E(WZI_{[Z \le t]})}{1}$$

in probability; here W is the sum of 2 exponentials and Z is uniform on $(0,1)$ and independent of W. Therefore $y_n(t) \to t^2$ in probability. The same steps applied to (5) show that

$$x_n(t) \to [E(WI_{[Z\leq t]}) - E(WZI_{[Z\leq t]})] = 2t - t^2.$$

in probability. ■

Remark 2: For each $t \in [0,1]$ the limits

$$(x_t, y_t) \equiv (2t - t^2, t^2)$$

satisfy $\sqrt{y_t} = 1 - \sqrt{1 - x_t}$, $t \in [0,1]$, because $1 - x_t = (1-t)^2$; therefore $(2t - t^2, t^2)$ is on the limit curve L defined in (2). Since L is the limit of $(x_n(t), y_n(t))$, it has been "discovered" as a consequence of the method of proof. Previous limit shape theorems start with L and show that the difference from a random chain converges to zero.

Remark 3: The tangent to L at $(2t - t^2, t^2)$ has slope $t/(1-t)$. On the other hand for $t \in [t_k, t_{k+1})$, (7) says that $(x_n(t), y_n(t))$ is the vertex on the n-chain supporting the line of slope

$$R_{j_{k+1}} \frac{X_1 + \cdots + X_{n+1}}{Y_1 + \cdots + Y_{n+1}};$$

this quantity $\to t/(1-t)$ in probability as $n \to \infty$ because $t/(1-t) \in [R_{j_k}, R_{j_{k+1}})$ and the ratio of sums converges to 1.

Proof of Theorem 2: For each $n > 0$ let $X_1^{(n)}, \ldots, X_{n+1}^{(n)}$ and $Y_1^{(n)}, \ldots, Y_{n+1}^{(n)}$ be mutually independent exponential variables on Ω and define $x_n(t)$ and $y_n(t)$ as in (5) and (6) except we use $Z_i^{(n)} = Y_i^{(n)}/(X_i^{(n)} + Y_i^{(n)})$ and the formulas

$$x_n(t) = \frac{\sum_{i=1}^{n+1} X_i^{(n)} I_{[Z_i^{(n)} \leq t]}}{\sum_{i=1}^{n+1} X_i^{(n)}}, \tag{8}$$

and

$$y_n(t) = \frac{\sum_{i=1}^{n+1} Y_i^{(n)} I_{[Z_i^{(n)} \leq t]}}{\sum_{i=1}^{n+1} Y_i^{(n)}}. \tag{9}$$

These functions describe a random convex chain C_n. Fix $t \in (0,1)$. Lemma 2 was based on the facts that for large enough n,

$$\mathrm{Prob}(\left|\frac{1}{n+1}\sum_{i=1}^{n+1} X_i^{(n)}(I_{[Z_i^{(n)}\leq t]} - (2t - t^2))\right| > \varepsilon) < \varepsilon$$

and

$$\mathrm{Prob}(\left|\frac{1}{n+1}\sum_{i=1}^{n+1} Y_i^{(n)}(I_{[Z_i^{(n)}\leq t]} - t^2)\right| > \varepsilon) < \varepsilon.$$

In fact a much stronger statement is true because $X_i^{(n)}$ and $Y_i^{(n)}$ have finite variance. According to the complete convergence theorem of Hsu and Robbins [5], (see e.g. [4, p. 375]) the denominators in (8) and (9) satisfy

$$\sum_{n=1}^{\infty} \mathrm{Prob}(\left|\frac{1}{n+1}\sum_{i=1}^{n+1}(X_i^{(n)} - 1)\right| > \varepsilon) < \infty$$

and

$$\sum_{n=1}^{\infty} \mathrm{Prob}(\left|\frac{1}{n+1}\sum_{i=1}^{n+1}(Y_i^{(n)}-1)\right| > \varepsilon) < \infty,$$

and this implies that both $\sum_{i=1}^{n+1} X_i^{(n)}/(n+1)$ and $\sum_{i=1}^{n+1} Y_i^{(n)}/(n+1)$ converge to 1 almost surely. The same result applied to the numerators shows that

$$\sum_{n=1}^{\infty} \mathrm{Prob}(\left|\frac{1}{n+1}\sum_{i=1}^{n+1}[X_i^{(n)}(I_{[Z_i^{(n)}\le t]}-(2t-t^2))]\right| > \varepsilon) < \infty$$

and

$$\sum_{n=1}^{\infty} \mathrm{Prob}(\left|\frac{1}{n+1}\sum_{i=1}^{n+1}[Y_i^{(n)}(I_{[Z_i^{(n)}\le t]}-t^2)]\right| > \varepsilon) < \infty.$$

Together these facts guarantee that for fixed $t \in (0,1)$, $(x_n(t), y_n(t)) \to (2t-t^2, t^2)$ almost surely; i.e., for any $\varepsilon > 0$, and almost all $w \in \Omega$ there is $N(t,\varepsilon,w)$ for which

$$\|(x_n(t), y_n(t)) - (2t-t^2, t^2)\| < \varepsilon,$$

$n > N(t,\varepsilon,w)$.

Now take $t_i = i/(m+1)$, $i = 1,\ldots,m$ and apply the previous fact to each t_i. For almost all $w \in \Omega$ there is $N(\varepsilon,w)$ such that

$$\|(x_n(t_i), y_n(t_i)) - (x_{t_i}, y_{t_i})\| < \varepsilon, \quad \text{for all } i = 1,\ldots,m,$$

when $n > N(\varepsilon,w)$. If m is sufficiently large, this condition for the m points of the chain is sufficient to ensure that $\delta(C_n, L) < 2\varepsilon$ when $n > N(\varepsilon,w)$, since the curves are convex. ■

Remark 4: Lemma 2 says that a random, convex n-chain C_n is likely to be close to L when n is large. Theorem 2 says that in a sequence $C_1, C_2, \ldots$ of chains, C_j having j vertices, the chains are sure to be close to L *and they remain close.* Note also that the chains need not be independent.

4 A Central Limit Theorem

For each $n > 0$, we have mutually independent exponential variables $X_1,\ldots,X_{n+1}$ and $Y_1,\ldots,Y_{n+1}$, and use the definitions in (3), (5), and (6). We will show that for any $t \in [0,1]$, the deviations

$$D_n(t) \equiv \left(x_n(t) - (2t-t^2), y_n(t) - t^2\right)$$

have a limiting normal distribution, a fact responsible for the heading of this section. Consider first

$$\sqrt{n+1}\left[y_n(t) - t^2\right] = \frac{\frac{1}{\sqrt{n+1}}\sum_{i=1}^{n+1}\left[Y_i(I_{[Z_i\le t]} - t^2)\right]}{\frac{1}{n+1}\sum_{i=1}^{n+1} Y_i}.$$

The random variables in the numerator sum are independent with mean zero and variance $\sigma_y^2(t) = 2t^3(1-t)(1+2t)$ so by the central limit theorem

$$\mathrm{Prob}\left\{y_n(t) - t^2 \le \frac{v}{\sqrt{n+1}}\right\} \to \frac{1}{\sqrt{2\pi\sigma_y^2(t)}}\int_{-\infty}^{v} e^{-\frac{w^2}{2\sigma_y^2(t)}}\,dw.$$

Similarly the quantity

$$\sqrt{n+1}\left[x_n(t)-(2t-t^2)\right] = \frac{\frac{1}{\sqrt{n+1}}\sum_{i=1}^{n+1}\left[X_i\left(I_{[Z_i\le t]}-(2t-t^2)\right)\right]}{\frac{1}{n+1}\sum_{i=1}^{n+1}X_i}$$

has a limiting normal distribution because the numerator is the sum of random variables with mean zero and variance $\sigma_x^2(t)=2(1-t)^3t(3-2t)$. Thus

$$\text{Prob}\left\{x_n(t)-(2t-t^2)\le\frac{v}{\sqrt{n+1}}\right\}\to\frac{1}{\sqrt{2\pi\sigma_x^2(t)}}\int_{-\infty}^{v}e^{-\frac{w^2}{2\sigma_x^2(t)}}dw.$$

Clearly $\sigma_x^2(t)=\sigma_y^2(1-t)$. Not only does the central limit theorem hold independently for each coordinate of the point $(x_n(t),y_n(t))$ representing vertices of a random convex chain, but also

Theorem 3 *For each* $t\in[0,1]$

$$P(u,v)=\text{Prob}\{x_n(t)\le(2t-t^2)+\frac{u}{\sqrt{n+1}}\ \textit{and}\ y_n(t)\le t^2+\frac{v}{\sqrt{n+1}}\}$$

converges to the bivariate normal distribution with mean $(0,0)$ *and covariance matrix*

$$K_t=\begin{pmatrix}\sigma_x^2(t) & \sigma_{x,y}(t)\\ \sigma_{x,y}(t) & \sigma_y^2(t)\end{pmatrix};$$

$\sigma_x^2(t)=2(1-t)^3t(3-2t)$, $\sigma_y^2(t)=2t^3(1-t)(1+2t)$, *and* $\sigma_{x,y}(t)=3t^2(1-t)^2$. *Thus*

$$P(u,v)\to\int_{-\infty}^{u}\int_{-\infty}^{v}\phi(r,s)drds,$$

where $\phi(r,s)=\frac{1}{2\pi\sqrt{\det(K_t)}}\exp(-\frac{1}{2}(r,s)K_t^{-1}(r,s)^T)$.

Proof: Using (6) and (7) we write $\sqrt{n+1}D_n(t)$ as

$$\frac{\frac{1}{\sqrt{n+1}}\sum_{i=1}^{n+1}\left(X_i(I_{[Z_i\le t]}-(2t-t^2)),A_n[Y_i(I_{[Z_i\le t]}-t^2)]\right)}{\frac{1}{n+1}\sum_{i=1}^{n+1}X_i}, \tag{10}$$

where $A_n=(\sum_{i=1}^{n+1}X_i)/(\sum_{i=1}^{n+1}Y_i)$. Concentrating on the numerator, we estimate the probability that it is componentwise less than (u,v), an arbitrary pair of reals. This is

$$\text{Prob}\left[\frac{1}{\sqrt{n+1}}\sum_{i=1}^{n+1}\left(X_i(I_{[Z_i\le t]}-(2t-t^2)),Y_i(I_{[Z_i\le t]}-t^2)\right)\le(u,\frac{v}{A_n})\right].$$

The sum adds independent random vectors, each of mean $(0,0)$; the expectation of the product of the components of these vectors is easily verified to be $\sigma_{x,y}(t)=3t^2(1-t)^2$. Therefore the sum has a limiting normal distribution with mean $(0,0)$ and covariance matrix K_t. The asserted limit statement holds because both A_n and the denominator of (10) converge to 1. ∎

In fact all finite dimensional distributions along a random chain are asymptotically normal. Suppose we are given $s_1<\cdots<s_k$ in $[0,1]$. An argument similar to the previous one leads to the conclusion that $(x_n(s_1),y_n(s_1),\ldots,x_n(s_k),y_n(s_k))$ converges to a certain $2k$-dimensional normal random variable.

5 Weak Convergence

Again, for each $n > 0$ we have mutually independent exponential variables $X_1, \ldots, X_{n+1}$ and $Y_1, \ldots, Y_{n+1}$, and use (5) and (6) to define $x_n(t)$ and $y_n(t)$, $t \in [0,1]$. In Section 2 we showed that $(x_n(t), y_n(t))$ describes the *vertices* of a random convex chain C_n. Here we study the chain itself and show that it converges as a stochastic process. This allows us to invoke the invariance principle to study various functionals of the chain, for example $A(C_n, L)$ and $\delta(C_n, L)$, respectively the area and Hausdorff distance between the chain and the limit shape.

Under the notation of Lemma 1, $W_i = X_i + Y_i$, $Z_i = Y_i/W_i$ and $t_1 < \cdots < t_{n+1}$ denotes the ordered values of $Z_1, \ldots, Z_{n+1}$. For each $t \in [t_k, t_{k+1}]$ define

$$C_n(t) = \frac{t_{k+1} - t}{t_{k+1} - t_k}(x_n(t_k), y_n(t_k)) + \frac{t - t_k}{t_{k+1} - t_k}(x_n(t_{k+1}), y_n(t_{k+1})),$$

$k = 0, \ldots, n$, where $t_0 = 0$. By Lemma 1, this function interpolates linearly between vertices on a random n-chain C, so $C_n(t)$, $0 \leq t \leq 1$ provides a parametrization of the chain. Write

$$L(t) = (x(t), y(t)) = (2t - t^2, t^2)$$

and define

$$\xi_n(t) = \sqrt{n+1}\,(C_n(t) - L(t))\,. \tag{11}$$

Since $|C_n(t) - (x_n(t), y_n(t))| < c \log n / n$ almost surely, we can write

$$\xi_n(t) = \sqrt{n+1}((x_n(t), y_n(t)) - L(t)) + o(1),$$

a fact we will use repeatedly.

Let $C_0^2[0,1]$ be the Banach space of all continuous functions $g(t)$ from $[0,1]$ to R^2 under the sup-norm $\|g\|_\infty = \sup_{0 \leq t \leq 1} \|g(t)\|$. Define

$$f(z;t) = (f_1(z;t), f_2(z;t))\,,$$

where

$$f_1(z;t) = \sqrt{6}(1-z)(I_{[z \leq t]} - (2t - t^2))\,; \quad f_2(z;t) = \sqrt{6}z(I_{[z \leq t]} - t^2),$$

and, letting $B(z)$ be a standard Wiener process, define

$$\xi(t) = \int_0^1 f(z;t)\,dB(z). \tag{12}$$

From (10)

$$\xi_n(t) = \frac{\frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} \frac{W_i}{\sqrt{6}} \left(f_1(Z_i;t), A_n f_2(Z_i;t)\right)}{\frac{1}{n} \sum_{i=1}^{n+1} X_i} + o(1),$$

and this has a normal limit. Calculation of the covariance operator of ξ_n shows that

$$\lim_{n \to \infty} \mathrm{Cov}(\xi_n(t), \xi_n(s)) = K(t,s) = \int f(z;t)^T f(z;s)\,dz,$$

which is identical to that of $\xi(\cdot)$. (In fact $K_t = K(t,t)$ is explicitly given in Theorem 3.) This is the intuition behind the following statement which gives much more information about the convergence.

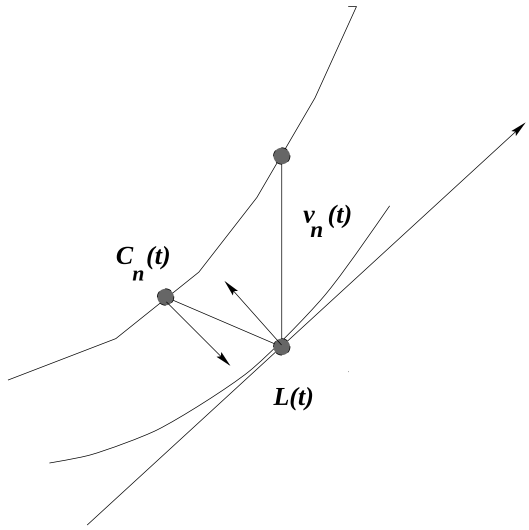


Figure 4: Area and Hausdorff distance.

Theorem 4 *The stochastic process $\xi_n(\cdot), n \geq 1$, converges weakly (in distribution) to the Gaussian process $\xi(\cdot)$ in (12) as random elements in $C_0^2[0,1]$; i.e.,*

$$\lim_{n\to\infty} Eh(\xi_n) = Eh(\xi) \tag{13}$$

for all bounded continuous mappings h from $C_0^2[0,1]$ to the reals. The covariance operator of $\xi(\cdot)$ is $K(t,s) = \int f(z;t)^T f(z;s)\,dz$. In addition,

$$\sup_n E\exp[\lambda\|\xi_n\|_\infty] < \infty$$

for all $\lambda < \infty$.

Proof: (The argument is straightforward, but somewhat technical, so it appears in the appendix.)

Theorem 4 can be used to investigate the convergence of many functionals of the random chain C_n. Perhaps the easiest example is the coordinate functional $h_t(\xi) = \xi(t)$ which gives Theorem 3.

The boundedness of the moment generating function implies that (13) also holds for all continuous mappings h from $C_0^2[0,1]$ to the reals which, for some λ satisfy $|h(g)| \leq \exp[\lambda\|g\|_\infty]$, $g \in C_0^2[0,1]$, even for unbounded ones. This property is needed in some of the following applications.

We first study the limiting Hausdorff distance. From (11) we write

$$\xi_n(t) = \sqrt{n+1}((x_n(t), y_n(t)) - L(t)) + o(1) = (\xi_{1,n}(t), \xi_{2,n}(t))$$

and note that $(t, t-1)/\sqrt{t^2+(t-1)^2}$ is the unit normal to the tangent line at $L(t)$. Then (see Fig. 4) $d(C_n(t), L)$ and $d(C_n, L(t))$ [distance from $C_n(t)$ to the limit shape and distance from $L(t)$ to the random chain, respectively] are both

$$(1+o(1))\left|\frac{t\xi_{1,n}(t) + (t-1)\xi_{2,n}(t)}{\sqrt{t^2+(t-1)^2}}\right|.$$

Therefore $\sqrt{n}\delta(C_n, L) = h(\xi_n) + o(1)$ where

$$h(\xi_n) = \sup_t (|(t\xi_{1,n}(t) + (t-1)\xi_{2,n}(t))|/\sqrt{t^2 + (t-1)^2}).$$

Clearly h is continuous and $h(g) \leq 2\|g\|_\infty \leq \exp[\lambda\|g\|_\infty]$, $\lambda = 2$. Now,

$$E(h(\xi)) = E\sup_t \left| \int_0^1 \frac{(tf_1(z;t) + (t-1)f_2(z;t))}{\sqrt{t^2 + (t-1)^2}} dB(z) \right|,$$

and Theorem 4 implies

Corollary 1 *The Hausdorff distance between the random convex chain C_n and its limit L satisfies*

$$\lim_{n\to\infty} \sqrt{n}E\delta(C_n, L) = E \sup_{0\leq t\leq 1} \left| \int_0^1 f^*(z;t)dB(z) \right| < \infty,$$

where

$$f^*(z;t) = \frac{\sqrt{6}\{(t-z)I_{[z\leq t]} - t^2(2 - 3z - t + 2tz)\}}{\sqrt{t^2 + (t-1)^2}}.$$

Next let $A(C_n, L)$ denote the area between the random convex chain C_n and its limit L and $v_n(t)$ the vertical distance from $L(t)$ to C_n (Fig. 4). Then

$$\sqrt{n}A(C_n, L) = \sqrt{n}\int_0^1 v_n(t)dx(t) = 2\sqrt{n}\int_0^1 (1-t)v_n(t)dt.$$

Since (see Figure 4) $\sqrt{n}(1-t)v_n(t) = |t\xi_{1,n}(t) + (t-1)\xi_{2,n}(t)| + o(1)$,

$$\sqrt{n}A(C_n, L) = \int_0^1 |t\xi_{1,n}(t) + (t-1)\xi_{2,n}(t)|dx(t) + o(1).$$

The integrand converges uniformly on $0 \leq t \leq 1 - \varepsilon$ so we write

$$\sqrt{n}A(C_n, L) = h_\varepsilon(\xi_n) + 2\sqrt{n}\int_{1-\varepsilon}^1 (1-t)v_n(t)dt + o(1), \tag{14}$$

where

$$h_\varepsilon(\xi_n) = \int_0^{1-\varepsilon} |t\xi_{1,n}(t) + (t-1)\xi_{2,n}(t)|dx(t).$$

$\sqrt{n}v_n(t)(1-t)$ converges in distribution to $|\int_0^1[tf_1(z;t) + (t-1)f_2(z;t)]dB(z)|$, which is the absolute value of a normal random variable with mean zero and variance $V_t = \int_0^1[tf_1(z;t) + (t-1)f_2(z;t)]^2dz$, by Theorem 3. A simple calculation shows V_t to be $2t^3(1-t)^3$. In addition, $|h_\varepsilon(\xi_n)| \leq \|\xi_n\|_\infty$, so that by Theorem 4,

$$Eh_\varepsilon(\xi_n) \to Eh_\varepsilon(\xi) = \int_0^{1-\varepsilon} |\mathrm{Normal}(0, 2t^3(1-t)^3)|\mathrm{dt}.$$

It is easy to show that $\lim_{\varepsilon\to 0}\lim_{n\to\infty}$ of the last two terms in (14) is zero. Therefore, since the L_1-norm of $\mathrm{Normal}(0, 2t^3(1-t)^3)$ is $4\sqrt{(t^3(1-t)^3/\pi)}$,

Corollary 2 *The area $A(C_n, L)$ between C_n and L satisfies*

$$\lim_{n\to\infty} \sqrt{n}E[A(C_n, L)] = 2\sqrt{\frac{2}{\pi}}\int_0^1 \sqrt{2t^3(1-t)^3}dt = \frac{4}{\sqrt{\pi}}\frac{\Gamma^2(5/2)}{\Gamma(5)} = \frac{3\sqrt{\pi}}{32}.$$

Appendix — Proof of Theorem 4

Weak convergence is easy to prove. From (5), (6) and the definition of $(f_1(z;t), f_2(z;t))$ it follows that

$$\frac{(x_n(t), y_n(t)) - L(t)}{(n+1)^{-1/2}} = \Big(\frac{\sum_{i=1}^{n+1}(W_i/\sqrt{6})f_1(Z_i;t)}{(n+1)^{-1/2}\sum_{i=1}^{n+1}X_i}, \frac{\sum_{i=1}^{n+1}(W_i/\sqrt{6})f_2(Z_i;t)}{(n+1)^{-1/2}\sum_{i=1}^{n+1}Y_i}\Big). \tag{15}$$

By the strong law of large numbers $\sum_1^n X_i/n \to 1$, $\sum_1^n Y_i/n \to 1$ and $\sum_1^n W_i/n \to 2$. By the Borel-Cantelli lemma, $\limsup_n \max_{1\le i\le n} W_i/\log n = 1$. Thus, by (11) and (15), the weak convergence of ξ_n follows from that of $\xi_n'(t) = (n+1)^{-1/2}\sum_{i=1}^n (W_i/\sqrt{6})f(Z_i;t)$ under the $\|\cdot\|_\infty$ norm, and the two should share the same limiting distribution if the weak convergence holds. Since $\mathcal{F} = \{(w/\sqrt{6})f(z;t) : 0 \le t \le 1\}$ is a Vapnik-Červonenkis class of functions of (w,z), the weak convergence of ξ_n' follows from standard results in the empirical process theory, e.g. van der Vaart and Wellner [8, Theorems 2.6.7 and 2.5.2]. The limiting covariance operator $E(W_1/\sqrt{6})^2 f^T(Z_1;t)f(Z_1;s)$ of ξ_n' is clearly identical to $K(t,s)$ as $E(W_1/\sqrt{6})^2 = 1$ and Z_1 is independent of W_1 and uniformly distributed on $[0,1]$, so (13) holds.

To prove the boundedness $E\exp[\lambda\|\xi_n\|_\infty]$ we shall compare $\xi_n(t)$ and (15) with

$$\xi_n''(t) = \sqrt{n+1}\frac{\sum_{i=1}^n W_i f(Z_i;t)}{\sum_{i=1}^n W_i}. \tag{16}$$

By the large deviation results for gamma-distributions,

$$\frac{1}{n}\log P\Big\{\sum_{i=1}^n \frac{X_i}{n} \le c\Big\} \to I(c)\ \forall 0 < c < 1,$$

$$\frac{1}{2n}\log P\Big\{\sum_{i=1}^n \frac{W_i}{2n} > c\Big\} \to I(c)\ \forall c > 1,$$

where $I(c) = 1 - c + \log c$. Since $I(c) \to -\infty$ as $c \to 0$ or $c \to \infty$ and $\|\xi_n\|_\infty \le \sqrt{n+1}$, the boundedness of $E\exp[\lambda\|\xi_n\|_\infty]$ for all λ follows from that of $E\exp[\lambda\|\xi_n''\|_\infty]$ for all λ. Note here that by (11), the maximum of each component of $\xi_n(t)$ in absolute value over $0 < t < 1$ is identical to those of (15). Since $\{Z_i\}$ are independent of $\{W_i\}$, (16) and the standard symmetrization methods imply

$$E\exp[\lambda\|\xi_n''\|_\infty] \le E\exp\left[2\lambda\Big\|\sum_{i=1}^{n+1} a_i\epsilon_i\Big\|_\infty\right], \tag{17}$$

where $a_i = a_i(t) = \sqrt{n+1}W_i f(Z_i;t)/\sum_{i=1}^n W_i$ and $\{\epsilon_i\}$ are Rademacher variables (i.e., $\varepsilon_i = \pm 1$, each with probability 1/2), independent of $\{(W_i, Z_i)\}$. Let $\tilde{E}$ be the expectation with respect to $\{\epsilon_i\}$ given $\{(W_i, Z_i)\}$. Since $\mathcal{F} = \{(w/\sqrt{6})f(z;t) : 0 \le t \le 1\}$ is a Vapnik-Červonenkis class of functions, by the Dudley-Pisier and Hoeffding inequalities, (cf. van der Vaart and Wellner [8, Corollary 2.2.8]).

$$\tilde{\mu}_n \stackrel{\text{def}}{=} \tilde{E}\Big\|\sum_{i=1}^{n+1} a_i\epsilon_i\Big\|_\infty \le KJ(\tilde{\tau}_n) \tag{18}$$

for some finite constant K, where $J(c) < \infty$ is the entropy integral of $\mathcal{F}$ and $\tilde{\tau}_n^2 = \sum_{i=1}^{n+1}\|a_i\|_\infty^2$. We apply Talagrand's deviation inequalities for product measures, (cf. e.g. Ledoux [6, top of

p. 70]) to $f = (\sum_{i=1}^{n+1} a_i \epsilon_i)/\tilde{\tau}_n$ (but using $\lambda\tilde{\tau}_n$ instead of λ) to see

$$\tilde{E} \exp\left[\lambda \left\| \sum_{i=1}^{n+1} a_i \epsilon_i \right\|_\infty \right] \leq \exp\left[\lambda \tilde{\mu}_n + \lambda^2 \tilde{\tau}_n^2/2\right] \tag{19}$$

for all $\lambda > 0$. It follows from inequalities (17)–(19) that, for any M,

$$Ee^{\lambda\|\xi_n''\|_\infty} \leq e^{2K\lambda J(M)+2\lambda^2 M^2} + e^{2\lambda\sqrt{12(n+1)}} \operatorname{Prob}\left\{ \frac{12(n+1)\sum_{i=1}^{n+1} W_i^2}{\left(\sum_{i=1}^{n+1} W_i\right)^2} > M^2 \right\},$$

since $\|\xi_n''\|_\infty \leq \sqrt{12(n+1)}$ and $\tilde{\tau}_n^2 \leq 12(n+1)\sum_{i=1}^{n+1} W_i^2/(\sum_{i=1}^{n+1} W_i)^2$. For each λ and as $n \to \infty$, the probability in the above expression is of an order smaller than $\exp\{-2\lambda\sqrt{12(n+1)}\}$ for large M. ■

References

[1] Imre Bárány. The limit shape of convex lattice polygons. *Discrete Comp. Geom.* **13** (1995), 270–295.

[2] Imre Bárány. Affine perimeter and limit shape. *J. reine angew. Mathematik* **484** (1997), 71–84.

[3] Imre Bárány. Sylvester's question: the probability that n points are in convex position. Preprint, 1997.

[4] Y. S. Chow and H. Teicher. *Probability Theory.* Springer-Verlag, New York, 1978.

[5] P. L. Hsu and H. Robbins. Complete Convergence and the Law of Large Numbers. *Proc. Nat. Acad. Sci.* **33** (1947), 229–240.

[6] M. Ledoux. On Talagrand's deviation inequalities for product measures. *ESAIM: Probab. Statist.* **1** (1995/97), Art. 4, 63–87 (electronic).
`http://www.emath.fr/Maths/Ps/Articles/psEng-Vol1.4.html`

[7] Ya. G. Sinai. Probabilistic approach to the analysis of statistics for convex polygonal curves. *Funkts. Anal. Prilozh.* **28**, No. 2 (1994), 41–48 (in Russian). English translation in *Funct. Anal. Appl.* **28** (1994), 108–113.

[8] A. van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes.* Springer, New York 1996.

[9] Pavel Valtr. Probability that n random points are in convex position. *Discrete Comput. Geom.* **13**, (1995), 637–643.

[10] A. M. Vershik. The limit shape for convex lattice polygons and related topics. *Funkts. Anal. Prilozh.* **28**, No. 1 (1994), 16–25 (in Russian). English translation in *Funct. Anal. Appl.* **28** (1994), 13–20.