The Annals of Probability
1982, Vol. 10, No. 3, 860–862

PROBABILISTIC VERSION OF A CURVATURE FORMULA

BY VLADIMIR DROBOT

University of Santa Clara

Let A be a C^2 curve of length $L(A)$ in some Euclidean space. Let P_n be a sequence of randomly chosen polygons with n vertices which are inscribed in A. It is shown that with probability 1

$$\lim n^2[L(A) - L(P_n)] = \frac{1}{4}\int_A \kappa^2(s)\, ds$$

where κ is the curvature.

Let A be a C^2 curve of finite length $L(A)$ in some Euclidean space E. For each n let P_n be the longest polygon of at most n sides that can be inscribed in A. It was shown in [1] by A. M. Gleason that

$$\lim_{n\to\infty} n^2(L(A) - L(P_n)) = \frac{1}{24}\left(\int_A \kappa^{2/3}\, ds\right)^3 \tag{1}$$

where $\kappa(\cdot)$ is the curvature. The object of this note is to show that a similar result holds when the polygons P_n are chosen at random, the only difference being the value of the limit on the right hand side of (1). Thus, at least as far as the rate of convergence is concerned, approximating by random polygons is as good as approximating by the best fitting polygon. More precisely, we have the following.

THEOREM. *Let $f:[a, b] \to E$ be a C^2 curve parametrized by the arclength so that $\| f'(s) \| = 1$. Let $X_1, X_2, \cdots$ be a sequence of independent random variables, uniformly distributed on $[a, b]$. For each n let*

$$a \equiv X_0(n) \le X_1(n) \le X_2(n) \le \cdots \le X_n(n) \le X_{n+1}(n) \equiv b$$

be the nth *order statistic of $X_1, X_2, \cdots, X_n$; i.e. the values of these variables arranged in a nondecreasing order. Let $V_j = f(X_j(n))$ and let P_n be the polygon with* (*consecutive*) *vertices $V_0, V_1, \cdots, V_{n+1}$. We have, with probability* 1

$$\lim_{n\to\infty} n^2[(b - a) - L(P_n)] = \frac{1}{4}\int_a^b \kappa^2(s)\, ds \tag{2}$$

where $\kappa(s) = \| f''(s) \|$ is the curvature.

The quantity $b - a$ is, of course, the length of the curve. We will present, for convenience, the proof in the case $[a, b] = [0, 1]$, the modifications needed in general case are more or less obvious. We need several lemmas first.

LEMMA 1. *Let $\xi_1, \xi_2, \cdots$ be a sequence of independent, identically distributed random variables with finite third moments ($d = E(|\xi_k|^3) < \infty$). If $\mu = E(\xi_k)$ then*

$$P\left[\left|\frac{1}{n}(\xi_1 + \cdots + \xi_n) - \mu\right| \ge \varepsilon\right] \le Cn^{-2}$$

where C is a constant depending only on ε, μ and d.

Received February 1981; revised May 1981.
AMS 1980 *subject classifications.* Primary 60F15, 53A05.
Key words and phrases. Random polygons, Random partitions of an internal.

PROOF. This is an immediate corollary of a result in [3] (page 54, Theorem 2.6.3).

LEMMA 2. *Let* $X_1, X_2, \cdots$ *be a sequence of independent random variables, uniformly distributed on* $[0, 1]$. *For each* n, *let* $X_1(n) \leq X_2(n) \leq \cdots \leq X_n(n)$ *be the values of* $X_1, X_2, \cdots, X_n$ *arranged in a non-decreasing order* (nth *order statistic*). *Put* $X_0(n) \equiv 0$, $X_{n+1}(n) \equiv 1$ *and define* $U_j(n) = X_{j+1}(n) - X_j(n)$, $j = 0, 1, 2, \cdots, n$. *We have for every* $0 \leq t \leq 1$ *and every* $p > 1$

$$\lim_{n\to\infty}(n+1)^{p-1} \textstyle\sum_{j=0}^{[nt]} (U_j(n))^p = t\Gamma(p+1) \quad \text{almost surely.} \tag{3}$$

Here $[\cdot]$ *is the greatest integer function.*

PROOF. Let $Y_0, Y_1, \cdots$ be a sequence of independent random variables, all exponentially distributed ($P(Y_k \leq x) = 1 - e^{-x}$ if $x \geq 0$). Let $S_n = Y_0 + Y_1 + \cdots + Y_n$. It is well known that for each n the vectors $(U_0(n), U_1(n), \cdots, U_n(n))$ and $(Y_0/S_n, Y_1/S_n, \cdots, Y_n/S_n)$ are identically distributed (see [2], page 242). To show (3) it is enough to establish that for each $\varepsilon > 0$

$$P[(n+1)^{p-1} \textstyle\sum_{j=0}^{[nt]} (U_j(n))^p > t(p+1) + \varepsilon] = O\left(\frac{1}{n^2}\right) \tag{4}$$

$$P[(n+1)^{p-1} \textstyle\sum_{j=0}^{[nt]} (U_j(n))^p < t(p+1) - \varepsilon] = O\left(\frac{1}{n^2}\right). \tag{5}$$

The result will then follow by a standard application of Borel-Cantelli theorem. We will show (4) only, the case of (5) is completely analogous. By the preceding remarks we must show that

$$P[(n+1)^{p-1} \textstyle\sum_{j=0}^{[nt]} Y_j^p S_n^{-p} > t\Gamma(p+1) + \varepsilon] = O(n^{-2}).$$

Now

$$(n+1)^{p-1} \textstyle\sum_{j=0}^{[nt]} Y_j^p S_n^{-p} = \left(\frac{[nt]+1}{n+1}\right)\left(\frac{1}{[nt]+1} \sum_{j=0}^{[nt]} Y_j^p\right)\left(\frac{1}{n+1} S_n\right)^{-p}.$$

Thus we must show that

$$P[A_n] = P\left[\left(\frac{1}{[nt]+1} \textstyle\sum_{j=0}^{[nt]} Y_j^p\right) \Big/ \left(\frac{1}{n+1} S_n\right)^p > \left(\frac{n+1}{[nt]+1}\right) (t\Gamma(p+1) + \varepsilon)\right] = O(n^{-2}).$$

For n sufficiently large we have

$$\left(\frac{n+1}{[nt]+1}\right) (t\Gamma(p+1) + \varepsilon) > \Gamma(p+1) + \varepsilon/2.$$

Let $\zeta > 0$ be such that $(\Gamma(p+1) + \varepsilon/2)(1 - \zeta) > \Gamma(p+1) + \varepsilon/4$. Let the sets B_n, C_n be defined by the conditions

$$\begin{aligned} B_n &= \left[\frac{1}{[nt]+1} \textstyle\sum_{j=0}^{[nt]} Y_j^p > \Gamma(p+1) + \varepsilon/4\right] \\ C_n &= \left[\frac{1}{n+1} S_n > (1-\zeta)^{1/p}\right]. \end{aligned} \tag{6}$$

Since $E(Y_j^p) = \Gamma(p+1)$, the direct application of Lemma 1 gives $P[B_n] = O([nt]^{-2}) = O(n^{-2})$ and $P(\bar{C}_n) = O(n^{-2})$ ($\bar{C}_n$ denotes the complement of C_n). Since it is clear that $P[A_n] \leq P[B_n] + P[\bar{C}_n]$ the result follows.

LEMMA 3. *Let* $g(x)$ *be a continuous, real valued function on* $[0, 1]$, *let* $X_1, X_2, \cdots$ *be a sequence of independent random variables, uniformly distributed on* $[0, 1]$. *Let* $X_j(n)$ *and* $U_j(n)$ *be defined as in the statement of Lemma* 2. *Then with probability* 1

(7) $\lim_{n\to\infty}(n+1)^{p-1}\sum_{j=0}^{n} g(X_j(n))(U_j(n))^p = \lim_{n\to\infty} A(n,p,g) = \Gamma(p+1)\int_0^1 g(x)\,dx.$

PROOF. First we show that (7) holds for the function $1_{[0,t]}(x)$, the characteristic function of the interval $[0, t]$. It follows from Lemma 2 that if $0 \le t < s \le 1$ then with probability 1

(8) $$\lim_{n\to\infty}(n+1)\sum_{j=[nt]}^{[ns]}(U_j(n))^p = (s-t)\Gamma(p+1).$$

Let $0 \le t \le 1$ be fixed and let $\varepsilon > 0$ be given. Let $j_t(n)$ be the largest subscript j for which $X_j(n) \le t$. The SLLN implies that $(1/n)j_t(n) \to t$ a.e. Thus for n large enough we have $[n(t-\varepsilon)] \le j_t(n) \le [n(t+\varepsilon)]$. Hence

(9) $$|A(n,p,1_{[0,t]}) - (n+1)^{p-1}\sum_{j=0}^{[nt]}(U_j(n))^p| \le (n+1)^{p-1}\sum_{j=[n(t-\varepsilon)]}^{[n(t+\varepsilon)]}(U_j(n))^p.$$

It follows from (8) that the right hand side of (9) converges a.e. to $2\,\varepsilon\,\Gamma(p+1)$ and Lemma 2 implies that the sum inside the absolute values converges a.e. to $t\Gamma(p+1)$. Since ε was arbitrary, we see that (7) holds for $g = 1_{[0,t]}$. It is clear that if (7) holds for functions $g_1(x)$ and $g_2(x)$ (not necessarily continuous) then it also holds for their linear combination $\alpha_1 g_1(x) + \alpha_2 g_2(x)$. It is also clear that if $g_1(x) \le g_2(x)$ then $A(n,p,g_1) \le A(n,p,g_2)$. We have just shown that (7) is true for the function $g(x) = 1_{[0,t]}(x)$. Thus (7) is true for a characteristic function of any interval $[a, b]$ since $1_{[a,b]} = 1_{[0,b]} - 1_{[0,a]}$ and thus for any simple function $s(x) = \sum_1^m \alpha_j 1_{I_j}(x)$ where I's are disjoint intervals. If $g(x)$ is a continuous function, choose $s_1(x)$, $s_2(x)$ to be two such functions so that $s_1(x) \le g(x) \le s_2(x)$. We have then

$$\Gamma(p+1)\int_0^1 s_1(x)\,dx \le \liminf_n A(n,p,g) \le \limsup_n A(n,p,g) \le \Gamma(p+1)\int_0^1 s_2(x)\,dx.$$

The result now follows, since $s_1(x)$ and $s_2(x)$ can be chosen arbitrarily close to $g(x)$.

We are now ready to prove the main theorem. We recall a result from [1]. If $f:[0,1] \to E$ is a curve parametrised by the arclength then

(10) $$h - \|f(t+h) - f(t)\| = \frac{1}{24}\kappa^2(t)h^3(1+o(1)) \quad \text{as} \quad h \to 0$$

where $o(1)$ is uniform in t. Hence

$$\begin{aligned} n^2[(b-a) - L(P_n)] &= n^2\sum_{j=0}^{n}\{(x_j(n+1) - x_j(n)) - \|V_{j+1} - V_j\|\} \\ &= n^2\sum_{j=0}^{n}\frac{1}{24}\kappa^2(X_j(n))[U_{j+1}(n) - U_j(n)]^3(1+o(1)). \end{aligned}$$

The last equality is justified by (10) if we take $h = X_{j+1}(n) - X_j(n)$. By Lemma 3 this last expression approaches $(1/24)\Gamma(3+1)\int_0^1 \kappa^2(s)\,ds = \frac{1}{4}\int_0^1 \kappa^2(s)\,ds$ with probability 1.

REFERENCES

[1] GLEASON, A. M. (1979). A curvature formula. *Amer. J. Math.* **101** pt. 1 86–93.
[2] KARLIN, S. (1968). *A First Course in Stochastic Processes.* Academic, New York.
[3] RÉVÉSZ, P. (1968). *The Laws of Large Numbers.* Academic, New York.

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF SANTA CLARA
SANTA CLARA, CALIFORNIA 95053