A Note on Moon's Problem -- Crossings in Random Graphs

A Note on Moon's Problem -- Crossings in Random Graphs

Hiroshi MAEHARA*

1. Introduction

Let G be an (abstract) simple graph. Place the vertices of G randomly on the surface of a unit sphere S so that all vertices of G are distributed independently and uniformly on S. Connect two vertices a, b by the shortest arc on S whenever $\{a, b\}$ is an edge of G. The resulting configuration is called a *random drawing of G on S*. A random drawing of G on a hemisphere H of S is defined similarly. The *crossing number* of a random drawing of G is the number of pairs of arcs that interscet each other in a point interior to both. (All 'singular' cases of special position may be ignored as they occur with probability zero.).

Moon studied the crossing number $c(K_n : S)$ in a random drawing of the complete n-graph K_n on S. In [2] he stated that the distribution of $c(K_n : S)$ is asymptotically normal as n tends to infinity. However the argument to show the asymptotic normality of $c(K_n : S)$ was incorrect [3] .

We show here that the "skewness" of the distribution of $c(K_n : S)$ tends to a positive constant as n tends to infinity. Hence the distribution of $c(K_n : S)$ is never asymptotically normal. On the other hand, it is proved that the distribution of the crossing number $c(K_n : H)$ in a random drawing of K_n on a hemisphere H is asymptotically normal. It is also shown that among all graphs G with n vertices and m edges, the expected value of the crossing number in a random drawing of G on S (or H) takes the largest value when the degrees of the vertices of G are as equal as possible.

2. Geometric probability on the sphere

We recall here some results on geometric probability on a unit sphere S for later use (see [4]). For non-antipodal points a, b of S, ab denotes the shortest arc (and its length) joining them. A subset K of S is *convex* if K is hemispherical and $ab \subset K$ for every non-antipodal a, b of K.

(2.1) The probability density function of the length $s = ab$ for two random points a, b on the unit sphere S is $(1/2)\ \sin s$.

(2.2) The probability that a "random great circle" intersect a convex set K of perimeter L is $L / (2\pi)$.

(2.3) The mean distance between two points on the unit hemisphere H is $4/\pi$.

(2.4) The probability that four random points on the unit hemisphere H form a convex spherical quadrilateral is $3-24/\pi^2$.

---

*Dept. of Math., Coll. of Educ., Univ. of the Ryukyus

3. A complete graph on a unit sphere

Consider a random drawing of K_n on S and let V be the vertex set of the drawing. Let $x(abcd)$ be the number of crossings in the six arcs ab, ac, ad, bc, bd, cd. Then $x(abcd)$ is a random (0, 1) -variable, and the crossing number $c(K_n : S)$ is written as

$$c(K_n : S) = \Sigma\, x(abcd)\ ,$$

where the summation is taken over all 4-subsets $\{a, b, c, d\}$ of V . The conditional probability that cd crosses ab given $ab = s$ follows easily from (2.2) :

$$\mathrm{Prob}[cd \text{ crosses } ab \mid ab = s] = s/(4\pi).$$

Then by (2.1) we have the expected values of $x(abcd)$ and $c(K_n : S)$:

$$\mathrm{E}[x(abcd)] = 3/8,\ \ \mathrm{E}[c(K_n : S)] = \binom{n}{4}(3/8).$$

Three points a, b, c, determine three great circles of the sphere S, and they divide the surface into eight spherical triangles almost surely : the triangle T_{abc} enclosed by ab, bc, ca; the triangles T_{ab}, T_{bc}, T_{ca}, each having one side in common with T_{abc}; the triangles T_a, T_b, T_c, each having one point in common with T_{abc}; and the triangle T having no point in common with T_{abc}. It is easily seen that the arcs ab and cd intersect each other if and only if the point d is in the triangle T_{ab}. Since the probability that cd crosses ab under the condition $ab = s$ is $s/(4\pi)$, we have $\mathrm{E}[\mathrm{area}(T_{ab}) \mid ab = s]/(4\pi) = s/(4\pi)$, where $\mathrm{E}[\ \mid **]$ denotes the conditional expectation under the condition $**$. Since $x(abcd)$ takes the value 1 if and only if d falls in one of T_{ab}, T_{bc}, T_{ca}, and since $\mathrm{area}(T_{ab}) = \mathrm{area}(T_c), \ldots, \mathrm{area}(T_{abc}) = \mathrm{area}(T)$, we have.

$$\begin{aligned} &\mathrm{E}[x(abcd) \mid ab = s) = \mathrm{Prob}[x(abcd) = 1 \mid ab = s] \\ (3.1) \qquad &= \mathrm{E}[\mathrm{area}(T_{ab}) + \mathrm{area}(T_{bc}) + \mathrm{area}(T_{ca}) \mid ab = s]/(4\pi) \\ &= 1/2 - \mathrm{E}[\mathrm{area}(T_{abc}) \mid ab = s]/(4\pi) = 1/2 - s/(4\pi). \end{aligned}$$

Hence, for different a, b, c, d, e, f, we have

$$\begin{aligned} \mathrm{E}[x(abcd)\ x(abef)] &= \mathrm{E}[\ (1/2 - s/(4\pi))^2] \\ &= (5\pi^2 - 4)/(32\pi^2). \end{aligned}$$

Let $y(abcd) = x(abcd) - 3/8$. Then

$$\mathrm{E}[y(abcd)\ y(cdef)] = (\pi^2 - 8)/(64\pi^2).$$

(Note that $y(abcd)$ and $y(defg)$ are mutually independent as well as $y(abcd)$ and $y(efgh)$ are.) Hence the variance of $c(K_n)$ is

$$\sigma(n)^2 = \mathrm{E}[(\Sigma y(abcd))^2] = \binom{n}{4}\binom{n-4}{2}\binom{4}{2}(\pi^2-8)/(64\pi^2)+O(n^5)$$

$$= [(\pi^2-8)/(2^9\pi^2)]\, n^6 + O(n^5).$$

4. The skewness

We want to estimate the third central moment $\mu_3(n)$ of $c(K_n : S)$ when n is large. First we consider the expected value of the product $z = x(abcd)\ x(defg)\ x(ghia)$. From (3.1) it follows that

$$\mathrm{E}[z \mid ad=s,\ dg=t,\ ga=u] = [1/(4\pi)]^3\ (2\pi-s)(2\pi-t)(2\pi-u)$$

and hence

$$\mathrm{E}[z] = 7/2^7 - [1/(4\pi)]^3\ \mathrm{E}[(ad)(dg)(ga)] .$$

Let $f(s, t, u)$ be the joint probability density function of $s=ad$, $t=dg$, $u=ga$, and let $f_{\mathrm{H}}(s, t, u)$ be the joint probability density function of s, t, u when the random three points a, d, g are chosen independently and uniformly on a fixed hemisphere H of S. Then

$$f_{\mathrm{H}}(s, t, u)\ \mathrm{Prob}\,(a, d, g \in H) = f(s, t, u)\ \mathrm{Prob}\,(\Delta adg \cap G = \phi)/2,$$

where G is the great circle bounding H, and Prob ($\Delta adg \cap G = \phi$) is the probability that G does not cut the triangle Δadg provided that the perimeter of Δadg is $s+t+u$. Then from (2.2)

$$\mathrm{Prob}\,(\Delta adg \cap G = \phi) = 1-(s+t+u)/(2\pi).$$

Hence we have

$$f_{\mathrm{H}}(s, t, u) = 4f(s, t, u) - [2(s+t+u)/\pi]\ f(s, t, u).$$

Multiplying both sides by $(s)(t)=(ad)(dg)$ and integrate (in full range of s, t, u such that s, t, u form a spheirical triangle), we get

$$\mathrm{E}[(ad)(dg) \mid a, d, g \in H] = 4\mathrm{E}[(ad)(dg)] - (4/\pi)\ \mathrm{E}[(ad)^2(dg)]$$

$$-(2/\pi)\ \mathrm{E}[(ad)(dg)(ga)] .$$

Since $\mathrm{E}[(ad)(dg)] = \mathrm{E}[ad]^2 = (\pi/2)^2$ and $\mathrm{E}[(ad)^2 dg] = \mathrm{E}[(ad)^2]\,\mathrm{E}[dg] = (\pi^3 - 4\pi)/4$, we have

$$\mathrm{E}[(ad)(dg)(ga)] = 2\pi - (\pi/2)\,\mathrm{E}[(ad)(dg) \mid a, d, \mathrm{d} \in H] .$$

On the other hand

$$\mathrm{E}[(ad)(dg) \mid a, d, g \in H] = \mathrm{E}[w(d)^2 \mid d \in H] ,$$

where $w(d) = \mathrm{E}[(ad) \mid a \in H$ with d fixed$]$. Since $w(d)$ is continuous in $d \in H$ and not constant (because : by (2.3), $\mathrm{E}[\,w(d) \mid d \in H\,] = 4/\pi$, however, if d is the "center" of H then $w(d) = 1$ by (2.1)), we must have

$$\mathrm{E}[\,w(d)^2 \mid d \in H] > \mathrm{E}[w(d) \mid d \in H]^2 = (4/\pi)^2.$$

Thus we have

$$\mathrm{E}\,[(ad)(dg)(ga)] < 2\pi - (\pi/2)(4/\pi)^2 = 2\pi - 8/\pi = 3.7367\ldots$$

$$< (\pi/2)^3 = 3.8757.\ldots$$

and $\mathrm{E}[z] - (3/8)^3 > 7/2^7 - [1/(4\pi)]^3 (\pi/2)^3 - (3/8)^3 = 0$. Hence

$$p := \mathrm{E}[y(abcd)\,y(defg)\,y(ghia)] = \mathrm{E}[z] - (3/8)^3 > 0.$$

Now it is not difficult to see that the third central moment $\mu_3(n)$ is

$$\mu_3(n) = \mathrm{E}[\{\,\Sigma\, y\,(abcd)\,\}^3] = \binom{n}{9}\binom{9}{4}\binom{5}{3}(36)\,p + O(n^8)$$

$$= (p/8)\, n^9 + O(n^8).$$

Thus the skewness of $c(K_n : S)$ is

$$\mu_3(n) \,/\, \sigma(n)^3 = (p/8)\,[2^9\,\pi^2 / (\pi^2 - 8)]^{3/2} + o(1),$$

which tends to a positive constant as n tends to infinity.

5. A complete graph on a hemisphere

Here we prove the asymptotic normality of the crossing number $c\,(K_n : H)$ on a hemisphere H. This is a simple application of a limit lemma proved in [1]. First we state the lemma.

Let N be the set of natural numbers and r a positive integer. Suppose that for every r-

element subset A of N, there corresponds a random variable $x(A)$ defined on a common probability space and having the same mean θ. We impose the following three conditions.

(5.1) For any finite number of r-subsets $A, B, \ldots, D \subset N$, the expected value $E[x(A) \ldots x(D)]$ exists, and for any bijection $\tau : N \to N$, $E[x(\tau A) \ldots x(\tau D)] = E[x(A) \ldots x(D)]$.

(5.2) If $(A \cup \ldots \cup B) \cap (C \cup \ldots \cup D) = \phi$, then

$$E[x(A) \ldots x(D)] = E[x(A) \ldots x(B)]\ E[x(C) \ldots x(D)].$$

Under the condition (5.1), the covariance $cov\ [x(A), x(B)]$ of $x(A)$ and $x(B)$ depends only on $| A \cap B |$, the number of elements in $A \cap B$. Let $c(m) = cov\ [x(A), x(B)]$ if $| A \cap B | = m$. Let t be the minimum value of m such that $c(m) \neq 0$.

(5.3) If $| A \cap (B \cup \ldots \cup D) | \leqq t$ and $|A \cap B| < t, \ldots, | A \cap D | < t$, then $E[x(A) \ldots x(D)] = E[x(A)]\ E[x(B) \ldots x(D)]$.

Note that if $t = 1$ then (5.3) automatically follows from (5.2).

LEMMA. Suppose $x(A)$ (A runs over all r-subsets of N) satisfy (5.1), (5.2), (5.3), and let $s(n)$ be the sum of $x(A)$ for all r-subsets A of $\{1, 2, \ldots, n\}$. Then $[s(n) - \mu]/\sigma$ tends to the normal distribution with zero mean and unit variance as n tends to infinity, where

$$\mu = \binom{n}{r}\theta,\ \sigma^2 = [c(t)\, n^{2r-t}] / \{t! [(r-t)!]^2\}.$$

Now we proceed to the proof of asymptotic normality of the distribution of $c(K_n : H)$. Consider a countably infinite number of random points on the unit hemisphere H, distributed independently and uniformly on H. Label these points by natural numbers. For any 4-subset $A = \{a, b, c, d\}$ of the natural numbers, let $x(A) = x(abcd)$, the number of crossings in six arcs ab, ac, ad, bc, bd, cd. Then $x(abcd) = 1$ if four points a, b, c, d, form a convex spherical quadrilateral, and $= 0$ otherwise. Thus $\theta := E[x(A)] = 3 - 24/\pi^2$ by (2.4), and $x(A)$'s clearly satisfy the conditions (5.1), (5.2). Furthermore, $c(K_n : H) = s(n)$, the sum of $x(A)$ for all 4-subsets of $\{1, 2, \ldots, n\}$.

Let $v(a) = E[x(abcd) \mid a : \text{fixed}]$. Then as a function of a, $v(a)$ is not constant. This is seen as follows. Suppose a is fixed on the boundary of the hemisphere H, and b, c be random points on H. Then H is divided by the three great circles determined by a, b, c, into six triangles (almost surely): the triangle T_{abc} enclosed by ab, bc, ca; the triangles T_{ab}, T_{bc}, T_{ca}, each having one side in common with T_{abc}; and triangles T_b, T_c, each having one point in common with T_{abc}. Further, these six triangles have the same expected area, as easily seen. Since $x(abcd) = 1$ if and only if d falls in T_{ab} or T_{bc} or T_{ca}, we have $E[x(abcd)] = v(a) = 1/2 \neq \theta$. Hence, when a varies in H, $v(a)$ also varies, and hence $E[x(abcd)\, x(aefg)] = E[v(a)^2] > \theta^2$. Thus $c(1) = cov[x(abdd), x(aefg)] > 0$, and hence we can apply the lemma. Therefore, the distribution of $c(K_n : H)$ is asymptotically normal as $n \to \infty$.

6. Crossings in a general graph

Let G be a simple graph with n vertices and m edges. Let V be the vertex set of G. Denote by $c(G)$ the number of crossings in a random drawing of G on a unit sphers S or on a hemisphere H. We show here that the expected value of $c(G)$ is

$$(6.1) \qquad \mathrm{E}[\ c(G)\ ] = (\theta/2)[m^2+m-\sum_{a\in V} deg\ (a)^2\ ],$$

where θ is the probability that two non-adjacent arcs ab and cd cross each other, and $deg\ (a)$ is the degree of the vertex a of G.

Let $\{a, b\}$ be any edge of G. Then there are

$$m-(deg(a)+deg(b))+1$$

edges not incident to a or b. Hence

$$\mathrm{E}[\ c(G)\ ] = \Sigma\ [\ m-(deg(a)+deg(b))+1\ ]\ \theta/2$$

(where the summation is over all edges $\{a, b\}$ of G)

$$= (\theta/2)\ [m^2+m-\ \Sigma\ (deg\ (a)+deg(b))] .$$

In the summation $\Sigma\ (deg(a)+deg(b))$, each $deg(a)$ appears exactly $deg(a)$ times. Hence

$$\Sigma\ (deg(a)+deg(b)) = \sum_{a\in V} deg(a)^2.$$

This proves (6.1).

Let $\bar{d}$ be the average degree of G. Then the "variance" of $deg\ (a)$ $(a \in V)$ is $(\Sigma\ deg(a)^2)/n-(\bar{d})^2$. Therefore, from (6.1) it follows that among all graphs with n vertices and m edges, $\mathrm{E}[\ c(G)\ ]$ takes the largest value when G has the minimum variance of $deg\ (a)$.

References

〔1〕 H. Maehara, On random simplices in product distributions, J. Appl. Prob. 17, 553-558 (1980).

〔2〕 J. W. Moon, On the distribution of crossings in random complete graphs, SIAM J., 13, 506-510 (1965).

〔3〕 J. W. Moon, Erratum : On the distribution of crossings in random complete graphs, SIAM J. Appl. Math. 32 (3), 706 (1977).

〔4〕 L. A. Santalo, *Integral Geometry and Geometric Probability*, Addison-Wesley, Reading Massachusetts (1976).