SOME GEOMETRIC APPLICATIONS OF THE BETA DISTRIBUTION

PETER FRANKL\textsuperscript{1} AND HIROSHI MAEHARA\textsuperscript{2}

\textsuperscript{1}CNRS, 15 Quai Anatole France, Paris, France
\textsuperscript{2}College of Education, Ryukyu University, Nishihara, Okinawa 903-01, Japan

(Received March 10, 1989; revised December 25, 1989)

Abstract. Let $\theta$ be the angle between a line and a “random” $k$-space in Euclidean $n$-space $\mathbb{R}^n$. Then the random variable $\cos^2 \theta$ has the beta distribution. This result is applied to show (1) in $\mathbb{R}^n$ there are exponentially many (in $n$) lines going through the origin so that any two of them are “nearly” perpendicular, (2) any $N$-point set of diameter $d$ in $\mathbb{R}^n$ lies between two parallel hyperplanes distance $2d((\log N)/(n - 1))^{1/2}$ apart and (3) an improved version of a lemma of Johnson and Lindenstrauss (1984, Contemp. Math., 26, 189–206). A simple estimate of the area of a spherical cap, and an area-formula for a neighborhood of a great circle on a sphere are also given.

Key words and phrases: Beta distribution, Spherical cap, Johnson-Lindenstrauss Lemma.

1. Introduction

The beta distribution $\text{Beta}(p, q)$ has a continuous probability density inside the interval $(0, 1)$ given by

$$B(p, q)^{-1}x^{p-1}(1 - x)^{q-1},$$

where $p, q > 0$ and

$$B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p + q).$$

If $X$ and $Y$ are independent random variables having the chi-square distributions with degrees of freedom $a$ and $b$, then the random variable $X/(X + Y)$ has the beta distribution $\text{Beta}(a/2, b/2)$ (e.g., Wilks (1962), p. 187). The beta distribution is well known as an appropriate prior distribution for Bayesian inference.

By a random point $v$ in Euclidean $n$-space $\mathbb{R}^n$, we mean a point
\( v = (z_1, \ldots, z_n) \) whose coordinates \( z_1, \ldots, z_n \) are independent normal variables with zero mean and unit variance. In other words, \( v \) is distributed according to the \( n \)-dimensional normal distribution \( N(O, I) \) with mean \( O \) (the origin) and covariance matrix the identity matrix. Let \( v_1, \ldots, v_k \) be \( k \) independent random points in \( R^n \). If \( k < n \) then the vectors \( Ov_i \) (\( i = 1, \ldots, k \)) span a \( k \)-dimensional linear subspace almost surely. We call this subspace a **random \( k \)-space** in \( R^n \).

Let \( L \) be a fixed \( 1 \)-space (line) in \( R^n \), and let \( H \) be a random \( k \)-space in \( R^n \). Let \( \theta \) be the angle between \( L \) and \( H \).

**Theorem 1.1.** The random variables \( \cos^2 \theta \) and \( \sin^2 \theta \) have the beta distributions \( \text{Beta} (k/2, (n-k)/2) \) and \( \text{Beta} ((n-k)/2, k/2) \), respectively.

This theorem seems to have appeared in many places and forms (e.g., Muirhead (1982), p. 39 and Watson (1983)). However, for the convenience of readers, we give here a short proof.

**Proof.** Since the distribution \( N(O, I) \) is "isotropic", we may assume that the \( 1 \)-space \( L \) is a random \( 1 \)-space taken in advance of the random \( k \)-space \( H \). Then we can reverse the order: first take a random \( k \)-space \( H \), and then take a random point \( v \) and determine the random line \( L = Ov \). In this case, we may regard \( H \) as a fixed \( k \)-space, say

\[
H = \{(x_1, \ldots, x_k, 0, \ldots, 0): x_i \in R \}.
\]

If we let \( v = (z_1, \ldots, z_n) \), then since \( z_i \)'s are independent normal variables with zero mean and unit variance, the sums

\[
z_1^2 + \cdots + z_k^2 \quad \text{and} \quad z_{k+1}^2 + \cdots + z_n^2
\]

are independent random variables having the chi-square distributions with degrees of freedom \( k \) and \( n-k \), respectively. Hence

\[
\cos^2 \theta = (z_1^2 + \cdots + z_k^2)/(z_1^2 + \cdots + z_n^2)
\]

has the distribution \( \text{Beta}(k/2, (n-k)/2) \). Similarly, \( \sin^2 \theta \) has the distribution \( \text{Beta}((n-k)/2, k/2) \). □

For a random point \( v \) in \( R^n \), the point \( v/|v| \) is uniformly distributed on the surface of the unit sphere \( S \) centered at the origin in \( R^n \). (Since the probability \( \text{Pr}(v = 0) \) is zero, we may assume \( v \neq 0 \).) Hence we have the following.

**Corollary 1.1.** If \( v = (z_1, \ldots, z_n) \) is a point uniformly distributed on
the surface of the unit sphere $S$ centered at the origin in $R^n$, then
$z_1^2 + \cdots + z_k^2$ is distributed according to Beta $(k/2, (n-k)/2)$.

In this paper we present some geometric applications of Theorem 1.1
and Corollary 1.1.

2. A neighborhood of a great circle on a sphere

Let $S$ be the unit sphere in $R^n$ centered at the origin $O$, and let $G$ be a
great circle on $S$. More precisely, $G$ is the intersection of the sphere $S$ and a
plane (2-space) through the origin $O$. Let $G_\delta$ denote the \( \delta \)-neighborhood of
$G$ in $S$, that is, the set of points on $S$ within angular distance $\delta$ from $G$.
Then the area of $G_\delta$ can be easily calculated from Corollary 1.1.

THEOREM 2.1. \( \text{area} (G_\delta) = (\sin \delta)^{n-2} \text{area} (S). \)

PROOF. Let $H$ be the plane in $R^n$ determined by $G$ and let $\nu$ be a
random point uniformly distributed on $S$. Let $\theta$ be the angle between the
line $O\nu$ and $H$. Then $\sin^2 \theta$ is distributed according to Beta $((n-2)/2, 1)$.
Hence $X := \sin^2 \theta$ has the probability density function

\[
g(x) = B((n-2)/2, 1)^{-1} x^{(n-2)/2 - 1} = [(n-2)/2] x^{(n-2)/2 - 1}.
\]

Therefore,

\[
\Pr (\theta < \delta) = \Pr (X < \sin^2 \delta) = \int_0^{\sin^2 \delta} g(x) dx = (\sin \delta)^{n-2}.
\]

Thus the probability content of $G_\delta$ is $(\sin \delta)^{n-2}$, and hence

\[
\text{area} (G_\delta) = (\sin \delta)^{n-2} \text{area} (S).
\]

3. The angle between two random lines

Let $\theta$ $(0 \leq \theta \leq \pi/2)$ be the angle between a fixed line $L$ passing through
the origin $O$ and a random line $K = O\nu$. Then, by Theorem 1.1, the variable
$X := \sin^2 \theta$ has the probability density function

\[
f(x) = (1/B)x^{(n-1)/2 - 1}(1-x)^{-1/2} \quad (0 < x < 1)
\]

where $B = B((n-1)/2, 1/2)$. Since $\log \Gamma(x) (x > 0)$ is a convex function
(e.g., Artin (1964), p. 17), we have

\[
\log \Gamma((k+1)/2) + \log \Gamma((k-1)/2) > 2 \log \Gamma(k/2)
\]
and hence \( \log \Gamma((k + 1)/2) - \log \Gamma(k/2) > \log \Gamma(k/2) - \log \Gamma((k - 1)/2) \), that is,

\[
\Gamma((k + 1)/2)/\Gamma(k/2) > \Gamma(k/2)/\Gamma((k - 1)/2) .
\]

Therefore, noting that

\[
\{\Gamma((k + 1)/2)/\Gamma(k/2)\}^{\{\Gamma(k/2)/\Gamma((k - 1)/2)\}} = (k - 1)/2
\]

we have

\[
\Gamma(k/2)/\Gamma((k - 1)/2) < \{(k - 1)/2\}^{1/2} < \Gamma((k + 1)/2)/\Gamma(k/2) .
\]

And since \( \Gamma(1/2) = \pi^{1/2} \), we have

\[
(n - 2)/(2\pi)^{1/2} < B((n - 1)/2, 1/2)^{-1} < \{(n - 1)/(2\pi)\}^{1/2} .
\]

**Theorem 3.1.** For any \( 0 < \alpha < \pi/2 \),

\[
F_n(\alpha)(1 - o(1)) < \Pr(\theta < \alpha) < F_n(\alpha) ,
\]

where \( o(1) \to 0 \) as \( n \to \infty \) and

\[
F_n(\alpha) = \frac{(\sin \alpha)^{n-1}}{(\pi(n - 1)/2)^{1/2} \cos \alpha} .
\]

**Proof.** Letting \( y = \sin^2 \alpha \), we have

\[
\Pr(\theta < \alpha) = \Pr(X < y) = \int_0^y f(x) \, dx 
\]

\[
< (1/B) \int_0^y (1 - y)^{-1/2} x^{(n-1)/2-1} \, dx 
\]

\[
= 2(1/B)(1 - y)^{-1/2}(n - 1)^{-1} y^{(n-1)/2} = \frac{2(\sin \alpha)^{n-1}}{B(n - 1) \cos \alpha} .
\]

This is less than \( F_n(\alpha) \) by (3.1). Letting \( t = 1 - (1/n)^{1/2} \),

\[
\int_0^y f(x) \, dx > \int_{ty}^y f(x) \, dx > (1/B) \int_{ty}^y (1 - ty)^{-1/2} x^{(n-1)/2-1} \, dx 
\]

\[
= \frac{2(1/B)}{(n - 1)(1 - ty)^{1/2}} y^{(n-1)/2} (1 - t^{(n-1)/2})
\]
\[ \frac{2(1/B)}{(n - 1)(1 - y)^{1/2}} \left( 1 - t^{(n-1)/2} \right)^{1/2} \left( 1 - y \right)^{1/2} \left( 1 - t y \right)^{1/2}. \]

Here, \( \left( 1 - y \right)/\left( 1 - t y \right) \left( 1/2 \right) = 1 - o(1) \), and since
\[
 \log t^{(n-1)/2} = \frac{(n - 1)/2}{2} \log (1 - n^{-1/2}) \\
< \frac{(n - 1)/2}{2} \left( - n^{-1/2} \right) < - \frac{(n^{1/2} - 1)}{2} \to - \infty,
\]
we have \( t^{(n-1)/2} = o(1) \). Therefore,

\[
\Pr (\theta < \alpha) > \frac{2(\sin \alpha)^{n-1}}{B(n - 1) \cos \alpha} (1 - o(1)) > \frac{(n - 2)^{1/2}}{(n - 1)^{1/2}} F_n(\alpha)(1 - o(1)) = F_n(\alpha)(1 - o(1)).
\]

\[ \square \]

Note that in the \( n \)-dimensional normal distribution \( N(O, I) \), \( \Pr (\theta < \alpha) \) corresponds to the probability content of the "double" cone with axis \( L \) and angular radius \( \alpha \). Thus we have the following.

**Corollary 3.1.** Let \( C(\alpha) \) \((0 < \alpha < \pi/2)\) be a spherical cap of angular radius \( \alpha \) on the surface \( S \) of a unit sphere in \( R^n \). Then

\[
\text{area } (C(\alpha))/\text{area } (S) < F_n(\alpha)/2 = (\sin \alpha)^{n-1}/\{2\pi(n - 1)^{1/2}\cos \alpha\}
\]

and both sides are asymptotically \((n \to \infty)\) equal.

**Theorem 3.2.** For any \( 0 < \alpha < \pi/2 \), there exist more than

\[
F_n(\alpha)^{-1} = (\pi(n - 1)/2)^{1/2} \cos \alpha(\sin \alpha)^{-(n-1)}
\]

lines in \( R^n \) going through the origin \( O \) such that any two of them determine an angle greater than \( \alpha \).

**Proof.** Consider a collection of lines passing through the origin in \( R^n \) such that

\[(\ast) \quad \text{any two lines in the collection determine an angle } > \alpha.\]

Let \( A = \{L_1, \ldots, L_m\} \) be such a collection of lines which is \textit{maximal} in the sense that no line can be added to \( A \) without violating the condition \((\ast)\). Consider a random line \( K = Ov \), and let \( \theta_i \) be the angle between \( L_i \) and \( K \). Then,
\[ \Pr (\theta_i < \alpha \text{ for some } i) < m \cdot \Pr (\theta_i < \alpha) < m \cdot F_n(\alpha) . \]

Hence, if \( m \leq F_n(\alpha)^{-1} \) then \( \Pr (\theta_i < \alpha \text{ for some } i) < 1 \), that is,

\[ \Pr (\theta_i > \alpha \text{ for all } i) > 0 . \]

Thus, if \( m \leq F_n(\alpha)^{-1} \), then there exists a line \( K \) such that all angles between \( K \) and \( L_i (i = 1, \ldots, m) \) are greater than \( \alpha \), contradicting the maximality of \( \Lambda \). \( \square \)

Thus, in \( \mathbb{R}^n \), there are exponentially many lines going through the origin so that all angles of them are greater than, say \( 89^\circ \). For other related results, see, for example, Erdös and Füredi (1983).

A line going through the center of a sphere in \( \mathbb{R}^n \) meets the surface of the sphere at two points. Hence, letting \( \alpha = 2\delta \) in the above theorem, we have the following.

**COROLLARY 3.2.** On the surface of a sphere in \( \mathbb{R}^n \), more than

\[ 2F_n(2\delta)^{-1} = (2\pi(n-1))^{1/2} \cos 2\delta (\sin 2\delta)^{(n-1)} \]

caps of angular radius \( \delta \) \((0 < \delta < \pi/4)\) can be packed.

4. A random hyperplane

Let \( L \) be a fixed 1-space, and \( K \) be a random 1-space, \( H \) a random \((n-1)\)-space in \( \mathbb{R}^n \). Let \( \theta_K \) be the angle between the two lines \( L \) and \( K \), and \( \theta_H \) be the angle between \( L \) and the hyperplane \( H \). Then by Theorem 1.1, \( \sin^2 \theta_K = 1 - \cos^2 \theta_K \) has the same distribution as \( \cos^2 \theta_H \). Therefore,

\[ (4.1) \quad \Pr (\theta_H > \delta) = \Pr (\cos^2 \theta_H < \cos^2 \delta) \\
= \Pr (\sin^2 \theta_K < \sin^2 (\pi/2 - \delta)) \\
< F_n(\pi/2 - \delta) = \frac{(\cos \delta)^{n-1}}{(\pi(n-1)/2)^{1/2} \sin \delta} . \]

**THEOREM 4.1.** Any \( N \)-point set \( V \) in a ball of radius \( r \) in \( \mathbb{R}^n \) lies between some two parallel hyperplanes at distance

\[ 2r(2 \cdot \log N)/(n - 1))^{1/2} \]

apart from each other.
For example, if $N = O(n)$ and $n$ tends to infinity, then

$$(2 \cdot \log N)/(n - 1))^{1/2} \to 0.$$  

Thus, for any $N = O(n)$ points in a ball of radius $r$ in $R^n$, there exists a hyperplane which is close to all these $N$ points, provided that $n$ is sufficiently large.

**Proof.** Let $V = \{v_1, \ldots, v_N\}$. By translating $V$, if necessary, we may suppose the ball of radius $r$ containing $V$ is centered at the origin. If $2 \log N \geq n - 1$, then the theorem is trivial. So we consider the case $2 \log N < n - 1$. Let $L_i$ be the line $Ov_i$, $i = 1, \ldots, N$. Let $H$ be a random $(n - 1)$-space and let $\theta_i$ be the angle between $L_i$ and $H$. Suppose, for a moment, $N \cdot F_n(\pi/2 - \delta) < 1$, where $\delta$ is defined by

$$\sin \delta = (2 \cdot \log N)/(n - 1))^{1/2}, \quad 0 < \delta < \pi/2.$$  

Then applying (4.1),

$$\Pr(\theta_i > \delta \text{ for some } i) < N \cdot F_n(\pi/2 - \delta) < 1,$$  

and hence $\Pr(\theta_i < \delta \text{ for all } i = 1, \ldots, N) > 0$. Therefore, there exists a hyperplane $H$ such that

$$\theta_i < \delta \quad \text{for all } i = 1, \ldots, N.$$  

In this case the distance between $v_i$ and $H$ is less than

$$r \sin \delta < r(2 \cdot \log N)/(n - 1))^{1/2}.$$  

Hence $V$ is sandwiched between the two hyperplanes, each parallel to $H$, at distance $r(2 \cdot \log N)/(n - 1))^{1/2}$ apart from $H$.

Now we show that $N \cdot F_n(\pi/2 - \delta) < 1$.

$$F_n(\pi/2 - \delta)^{-1} = (\frac{\pi (n - 1)}{2})^{1/2} \sin \delta \ (1 - \sin^2 \delta)^{-(n-1)/2}$$

$$= (\pi \log N)^{1/2} (1 - \sin^2 \delta)^{-(n-1)/2}$$

$$= (\pi \log N)^{1/2} \exp \{- ((n - 1)/2) \log (1 - \sin^2 \delta)\}.$$

Since $\log (1 - t) < -t$ for $0 < t < 1$,

$$F_n(\pi/2 - \delta)^{-1} > (\pi \log N)^{1/2} \exp \{(n - 1)(\sin \delta)^2/2\}$$

$$= (\pi \log N)^{1/2} \exp (\log N).$$
\[ = N(\pi \log N)^{1/2} > N. \]

Hence \( N \cdot F_n(\pi/2 - \delta) < 1. \square \)

For a finite point set in Euclidean space, the distance between the farthest pair in the set is called the diameter of the set. Jung’s theorem (e.g., Danzer et al. (1963)) asserts that if a point set \( V \) in \( \mathbb{R}^n \) has diameter \( d \), then \( V \) is contained in a ball of radius \( d/[n/(2n + 2)]^{1/2} \). Hence we have the following.

**Corollary 4.1.** Let \( V \) be an \( N \)-point set of diameter \( d \) in \( \mathbb{R}^n \). Then \( V \) lies between some two parallel hyperplanes at distance \( 2d/[\log N]/(n - 1)]^{1/2} \) apart from each other.

5. Tails of Beta \((p, q)\)

Let \( X \) be distributed according to the beta distribution Beta \((p, q)\). The mean and variance of \( X \) are

\[
\mu = p/(p + q), \quad \sigma^2 = pq/[(p + q)(p + q + 1)],
\]

respectively. Then \((X - \mu)/\sigma\) is asymptotically normally distributed with zero mean and unit variance, when \( p \) and \( q \) both tend to infinity (e.g., Moran (1968), p. 329). Apart from the asymptotic normality, it seems to be useful to evaluate the probabilities

\[
\Pr(X/\mu < 1 - \varepsilon) \quad \text{and} \quad \Pr(X/\mu > 1 + \varepsilon)
\]

for a given constant \( \varepsilon > 0 \). In this section, we consider this problem for large \( p, q \). The result (Theorem 5.1) will be used in the next section.

We start with evaluating the probability density function \( f(x) \) of the beta distribution Beta \((p, q)\):

\[
f(x) = B(p, q)^{-1}x^{p-1}(1-x)^{q-1} \quad (0 < x < 1).
\]

By Stirling’s formula,

\[
\Gamma(s) = (2\pi/s)^{1/2}(s/e)^e \varepsilon(s) \quad \text{for} \quad s > 0,
\]

where

\[
\varepsilon(s) = \sum_{n=0}^{\infty} \{(s + n + 1/2) \log (1 + (s + n)^{-1}) - 1\} < 1/(12s),
\]
see, for example, Artin ((1964), p. 24). Since the function \( \xi(s) \) is monotone decreasing,

\[
B(p, q)^{-1} = \frac{\Gamma(p + q)}{\{\Gamma(p)\Gamma(q)\}} \\
= \{pq/(2\pi(p + q))^1/2\}((p + q)/p)^{\mu/2}((p + q)/q)^q \exp((p + q)^{1/2}((p + q)/p)^{(p - q)/2}) \\
< \{pq/(2\pi(p + q))^1/2\}((p + q)/p)^{\mu/2}((p + q)/q)^q \\
= (p + q)^3/2(2\pi pq)^{-1/2}((p + q)/p)^{(p - 1)/2}((p + q)/q)^{(q - 1)/2} \\
= (p + q)^3/2(2\pi pq)^{-1/2}(1/\mu)^{(p - 1)/2}((p/q)/\mu)^{(q - 1)/2},
\]

where \( \mu = p/(p + q) \). Hence

\[
f(x) < C\{x/\mu\}^{p-1}((p/q)(1-x)/\mu)^{q-1}, \quad C := (p + q)^3/2(2\pi pq)^{-1/2}.
\]

Letting \( x = (1 + t)\mu \) \( (-1 < t < q/p) \), we have

\[
f((1 + t)\mu) < C(1 + t)^{p-1}(1 - pt/q)^{q-1}.
\]

In the following we suppose \( p, q \geq 1 \).

**Lemma 5.1.** Suppose \( 0 < \varepsilon < 1, p\varepsilon^2 > 6 \). Then

1. \( f((1 - t)\mu) < C\exp(-pt^2/2) < C\exp(-pet/2) \) for \( \varepsilon < t < 1 \), and
2. \( f((1 + t)\mu) < C\exp(-pa_\varepsilon t) \) for \( \varepsilon < t < q/p \),

where \( \mu = p/(p + q) \) and \( a_\varepsilon = \varepsilon/2 - \varepsilon^3/3 \).

**Proof.** (1) Using the inequality \( \log(1 - t) \leq -t - t^2/2 - t^3/3 \) \( (t < 1) \),

\[
\log f((1 - t)\mu) < \log C + (p - 1)\log(1 - t) + (q - 1)\log(1 + pt/q) \\
< \log C + (p - 1)(-t - t^2/2 - t^3/3) + (q - 1)pt/q \\
< \log C + p(-t - t^2/2 - t^3/3) + 2t + pt \\
< \log C - pt^2/2 - pt^3/3 + 2t.
\]

Since \( pt^2 > 6 \) for \( t > \varepsilon \), this is less than

\[
\log C - pt^2/2 = \log \{C\exp(-pt^2/2)\}.
\]

(2) First, let \( G(t) = (1 + t)^{p-1}(1 - pt/q)^{q-1} \exp\{p(t^2/2 - t^3/3)\} \). Then \( \log G(0) = 0 \), and

\[
\{\log G(t)\}' = (p - 1)/(1 + t) - p(q - 1)/(q - pt) + pt - pt^3.
\]
\[ <p/(1+t) - p + pt - pt^2 = pt^4/(1+t) - pt^2 < 0. \]

Hence \( G(t) < 1 \). Next, let

\[ H(t) = (1+t)^{p-1}(1-qt/q)^n \exp(pat). \]

Then \( H(\alpha) = G(\alpha) < 1 \). And

\[
\begin{align*}
\{\log H(t)\}' &= (p-1)/(1+t) - p(q-1)/(q-pt) + p(\alpha/2 - \alpha^2/3) \\
&< p/(1+t) - p + p(\alpha/2 - \alpha^2/3) \\
&= p\{3\alpha(1+t) - 6t - 2\alpha^2(1+t)/6(1+t)\}. 
\end{align*}
\]

Since \( 3\alpha(1+t) - 6t < 0 \) for \( t > \alpha \), we have \( \log H(t) < 0 \). Hence (2) follows.

\[ \square \]

**THEOREM 5.1.** Let \( X \) be a random variable having the beta distribution \( \text{Beta}(p, q) \). For an \( \alpha, 0 < \alpha < 1 \), suppose \( p\alpha^2 > 6 \). Then

\[
\begin{align*}
\Pr\{X < (1-\alpha)\mu\} &< (2/\alpha)[(p+q)/(2\pi pq)]^{1/2} \exp(-p\alpha^2/2), \\
\Pr\{X > (1+\alpha)\mu\} &< (1/\alpha)[(p+q)/(2\pi pq)]^{1/2} \exp\{-p(\alpha^2/2 - \alpha^3/3)\},
\end{align*}
\]

and hence

\[
\Pr\{|X - \mu| > \alpha\mu\} < (2/\alpha)[(p+q)/(2\pi pq)]^{1/2} \exp\{-p(\alpha^2/2 - \alpha^3/3)\},
\]

where \( \mu = p/(p+q) \) and \( a_\alpha = \alpha/2 - \alpha^2/3 \).

**PROOF.** By Lemma 5.1(1), \( \Pr\{X < (1-\alpha)\mu\} \) is

\[
\int_0^{(1-\alpha)\mu} f(x)dx < \mu \int_\alpha^1 f((1-t)\mu)dt < C\mu \int_\alpha^\infty \exp(-p\alpha t/2)dt \\
< (2/\alpha)[(p+q)/(2\pi pq)]^{1/2} \exp(-p\alpha^2/2).
\]

Similarly, using Lemma 5.1(2), \( \Pr\{X > (1+\alpha)\mu\} \) is less than

\[
(1/\alpha)[(p+q)/(2\pi pq)]^{1/2} \exp\{-p(\alpha^2/2 - \alpha^3/3)\}.
\]

Thus we have the theorem. \( \square \)

6. Johnson-Lindenstrauss lemma

Johnson and Lindenstrauss (1984) proved an interesting lemma (see
Theorem 6.1) concerning a nearly isometric embedding of \( n \) point set in \( \mathbb{R}^n \) into surprisingly lower dimensions. A slightly improved version of this lemma was given in Frankl and Maehara (1988). Here we present a short proof of this lemma in more improved form.

**Lemma 6.1.** For an \( \varepsilon, 0 < \varepsilon < 1 \), let \( p, q \) be positive numbers such that

\[
q \geq p \geq 2(\varepsilon^2/2 - \varepsilon^3/3)^{-1} \log (2q).
\]

If \( X \) is a random variable having the distribution Beta \((p,q)\) then

\[
\Pr \{|X - \mu| > \varepsilon\mu\} < 2(\log 2q)^{-1/2}(2q)^{-2},
\]

where \( \mu = p/(p + q) \), the mean of Beta \((p,q)\).

**Proof.** From the condition imposed on \( p, q \), it follows easily that \( p\varepsilon^2 > 6 \). Hence applying Theorem 5.1,

\[
\Pr \{|X - \mu| > \varepsilon\mu\} < (2/a_\varepsilon)(p + q)/(2\pi pq)\exp\left\{-p(\varepsilon^2/2 - \varepsilon^3/3)\right\}
\]

\[
< (2/a_\varepsilon)\exp\left\{\log (2q)^{-2}\right\}
\]

\[
< (2/a_\varepsilon)\exp\left\{(\varepsilon^2/2 - \varepsilon^3/3)^{-1/2}(2q)^{-2}\right\}
\]

\[
= 2\exp\left\{(\varepsilon^2/2 - \varepsilon^3/3)^{-1/2}(2q)^{-2}\right\}
\]

\[
< 2(\log 2q)^{-1/2}(2q)^{-2}.
\]

**Theorem 6.1.** (Johnson-Lindenstrauss lemma) For an \( \varepsilon (0 < \varepsilon < 1) \) and an integer \( n \), let \( k \) be positive integers such that

\[
k \geq 4(\varepsilon^2/2 - \varepsilon^3/3)^{-1} \log n.
\]

Then for any \( n \)-point set \( V \) in Euclidean space, there exists a map \( f \) from \( V \) to a \( k \)-space such that

\[
1 - \varepsilon < \frac{|f(u) - f(v)|^2}{|u - v|^2} < 1 + \varepsilon \quad \text{for all } u, v \text{ of } V, u \neq v.
\]

**Remark.** In Johnson and Lindenstrauss (1984), the function \( 4(\varepsilon^2/2 - \varepsilon^3/3) \) of \( \varepsilon \) is not specified, and in Frankl and Maehara (1988), an extra condition \( (n > k^2) \) is imposed.

**Proof.** If \( n \leq k \), then the theorem is trivial. Hence we may suppose that \( n > k \). Let \( V = \{v_1, \ldots, v_n\} \) be any \( n \) point set in \( \mathbb{R}^{n+k} \). Let \( H \) be a
random $k$-space in $\mathbb{R}^{n+k}$, and $w_i$ be the projection of $v_i$ on $H$. Then for any $i, j$ ($i \neq j$),

$$X_{ij} := \frac{|w_i - w_j|^2}{|v_i - v_j|^2}$$

is distributed according to Beta $(k/2, n/2)$. Hence by Lemma 6.1,

$$\Pr \{ |X_{ij} - k/(n + k)| > \varepsilon k/(n + k) \} < 2(\log n)^{-1/2} n^{-2}.$$ 

Therefore, the probability that

$$|X_{ij} - k/(n + k)| > \varepsilon k/(n + k) \quad \text{for some} \quad i, j \quad (i \neq j)$$

is less than $\binom{n}{2} 2(\log n)^{-1/2} n^{-2} < 1$. Thus there exists a $k$-space $H$ for which

$$(1 - \varepsilon)k/(n + k) < |w_i - w_j|^2/|v_i - v_j|^2 < (1 + \varepsilon)k/(n + k) \quad (i \neq j).$$

Then letting $f(v_i) = \{(n + k)/k\}^{1/2}w_i$, we have a desired map $f$. \qed

**REFERENCES**


