What I learnt from P.L. Hsu

Kai-Tai Fang

Chinese Academy of Sciences
BNU-HKBU United International College

Probability and Statistics:
an international conference
in honor of Professor P.L. Hsu's 100th birthday
5-7 July, 2010, Beijing

Table of Contents

1. Historical Review

♢ Prof. P.L. Hsu was my supervisor of my graduate thesis during 1962-1963.

♢ Hsu organized a seminar once a week and asked us (nine students) to report the contents of the book "变叙的项的极限分布" section by section.

♢ 在讨论班总结会上，许先生端出了自己的一套思路，若沿用苏联学者的方法，只能作出一维情形的结果，而按许先生的思路推演，可简捷、轻松、自然、流畅地推出一维到多维情形的结果，真是技高一筹！每个学生都受益匪浅。

♢ 我的两位同窗程士宏和马逢时将许先生的思想完整地实现，写成论文发表。随后，许宝騄教授用笔名"班成"又发表另一论文，该论文获得了次序统计量的各种情况下的极限律型，无论是单项的还是多项的，是固定名次的边项还是非固定名次的边项，是正则的还是非正则的中项。

1. Historical Review

♢ 讨论班结束后，许先生分别给每个学生论文问题。他给了我发表在美国杂志《The Annals Of Statistics》上的一篇文章，讲该文的证明有一个漏洞，如果我能发现这个漏洞，并补上有关证明，我的任务就完成了。

♢ 我找到漏洞，应用测度论的典型手法补上证明，并且推广文章的结果至可换随机变量序列，发现了许多意想不到的极限分布，并整理成论文。我的毕业论文被推荐为北大"五四"学术讨论会上宣读，并送北大学报发表。

♢ 1963年春，一场新的政治风暴即将来临，北大处于风口浪尖，"五四科学讨论会"取消了，我的论文虽然已被北大学报录用，但很快北大学报被迫停刊，文革后，我去编辑处询问，编辑处一片混乱，我的原稿也不知下落。我凭记忆，重新发现我丢失的成果，该论文1981年在应用数学学报发表，整整迟了十八年才发表。

1. Historical Review

♢ 1981年夏，我在美国斯坦福大学报告了这论文论文, 这是我第一次在美国作正式的学术的报告。

♢ 文革结束后，我在美国、香港等地，遇到许宝騄当年的学生和朋友，美国斯坦福大学教授钟开莱、安德森（T.W. Anderson) 美国科学院院士、奥肯（I. Olkin）、诺贝尔奖得主杨振宁等。他们告诉我许先生的一些故事。在我的研究中，用到许宝騄的许多思想和结果。在記念许宝騄诞辰一百周年之际，回忆我知道的许宝騄，作为对恩师许先生的怀念。

1. Historical Review

◇ 我在北大求学时，没有机会修"多元统计分析"的课（以后简称多元分析），不知道许先生在这个领域的杰出贡献。文革后期，中国的气象预报，地质勘探，医学研究均迫切需要多元统计中丰富有效的多种方法。在缺少书籍、文献的困难条件下，我和中央气象台、数学地质有关的研究所，中科院计算所及北京医学院共同组织了多元分析讨论班，每周半天，报告文献或新的研究结果，该讨论班持续了好几年。其间我曾邀请张尧庭老师在讨论班上作系列讲演，他的讲演很精彩并提到许先生的一个见解和一个公式。

◇ 许先生的见解：他认为多元的Dirichlet 分布可以认为是许多常见多元分布的中心，它和多元正态分布、多元t分布、多元F分布、多元beta分布、Wishart分布有密切的联系。

1. Historical Review: The Hsu's formula (许氏公式)

Hsu found the following formula in deriving the Wishart distribution:

Hsu's Formula: Let $\mathbf{X}$ be an $n \times p$ matrix and $\mathbf{V}$ be a upper triangular matrix of order p, then

$$\int f(\mathbf{X}^T\mathbf{X})d\mathbf{X} = \frac{2^p \pi^{np/2}}{\Gamma_p(n/2)} \int_{D_p} (\prod_{i=1}^{p} v_{ii}^{n-i} f(\mathbf{V}^T\mathbf{V})d\mathbf{V} \qquad (1)$$

where D_p is the set of upper triangular matrices with positive diagonal elemets and $\Gamma_p(n/2) = \pi^{p(p-1)/2} \prod_{i=1}^{p} \Gamma((n-i+1)/2)$.

♢ This formula reduces the integral dimension from np to $p(p+1)/2$.

When $p = 1$ formula (1) reduces

$$\int f(\mathbf{x}^T\mathbf{x})d\mathbf{x} = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2})}\int_0^\infty y^{n/2-1}f(y)dy. \tag{2}$$

♢ It reduces thee integral from n dimension to one dimension. These two formulas have played an important role in development of theory of spherical distributions and generalized multivariate analysis.

[1] Zhang, Y. T. and Fang, K. T. (1982, 1999, 2003), *An Introduction to Multivariate Analysis*, Science Press, Beijing.

2. Hus's Philosophy to My Work

2.1. Stochastic representation method

♢ Prof. Hsu told us that there are three main methods in statistics study:

analytic method including using mathematical analysis, characteristic function, moment generating function, and so on;

algebraic method including using matrix analysis, transformation Jacobian, group theory, and so on; and

probability method. The probability method in general is the best if it can be applied. Here, the probability method is "Stochastic representation method" (SRM). It directly deals with the random variables instead of treating their distribution functions, density functions, or characteristic functions.

♢ I has been benefit from his advice in my research in several areas.

2. Hsu's Philosophy to My Work

2.1. Stochastic representation method

♢ The symbol $X \stackrel{d}{=} Y$ means that two random variables X and Y have the same distribution.

♢ Suppose that X is a random variable of interest and we want to drive its distribution and moments. If we can find $Y_1, \ldots, Y_n$ such that

$$X \stackrel{d}{=} g(Y_1, \ldots, Y_n), \tag{3}$$

that is called as a stochastic representation of X.

♢ For given X it may have many its stochastic representations.

2.1. Stochastic representation method

Example 1: The mean of $X \sim \chi^2_n$, the chi-square distribution with n degrees of freedom, can be directly calculated by

$$E(X) = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty x e^{-x/2} x^{n/2-1} dx = n.$$

♢ In fact, X has a stochastic representation $X \stackrel{d}{=} Z_1^2 + \ldots + Z_n^2$, where $Z_1, \ldots, Z_n$ are i.i.d. follow $N(0,1)$.

♢ It is easy to find $E(X) = E(Z_1^2) + \ldots + E(Z_n^2) = n$ as $E(Z_j^2) = 1$.

2.1. Stochastic representation method

Example 2: Mixed moments of uniform distribution on the unit sphere surface in R^n.

Let $\mathbf{u}^{(n)} = (U_1, \ldots, U_n)^T$ be the random vector uniformly distributed on the unit sphere in R^n. We are requested to calculate the mixed moment $E(\prod_{i=1}^n U_i^{r_i})$. Denote the unit sphere surface in R^n by

$$S_n = \{\mathbf{x} = (x_1, \ldots, x_n)^T : \mathbf{x} \in R^n, \mathbf{x}^T\mathbf{x} = 1\}. \quad (4)$$

The surface area of S_n is $2\pi^{n/2}/\Gamma(n/2)$. Hence the pdf of $\mathbf{u}^{(n)} \sim U(S_n)$ is given by

$$f_{\mathbf{u}^{(n)}}(u_1, \ldots, u_n) = \frac{\Gamma(n/2)}{2\pi^{n/2}} \cdot I_{S_n}(\mathbf{u}^{(n)}),$$

$$E(\prod_{i=1}^n U_i^{r_i}) = \frac{\Gamma(n/2)}{2\pi^{n/2}} \int_{S_n} \prod_{i=1}^n u_i^{r_i} \, dS_n.$$

2.1. Stochastic representation method

Example 2: continuity

This integral on unit sphere surface S_n is not so easy to evaluate.

Let $\mathbf{z} = (Z_1, \dots, Z_n)^T \sim N_n(\mathbf{0}, \mathbf{I}_n)$. It can be verified from Fang, Kotz and Ng (1990) that

$$\mathbf{z} = \|\mathbf{z}\|_2 \cdot \frac{\mathbf{z}}{\|\mathbf{z}\|_2} \stackrel{d}{=} \|\mathbf{z}\|_2 \cdot \mathbf{u}^{(n)}, \tag{5}$$

where the L_2-norm is used and $\|\mathbf{z}\|_2^2 = \mathbf{z}^T\mathbf{z} \sim \chi^2_{(n)}$, and $\|\mathbf{z}\|_2$ is independent of $\mathbf{u}^{(n)}$. Hence

$$E\left(\prod_{i=1}^n Z_i^{r_i}\right) = E(\|\mathbf{z}\|_2^r) \cdot E\left(\prod_{i=1}^n U_i^{r_i}\right), \quad r = \sum_{i=1}^n r_i. \tag{6}$$

2.1. Stochastic representation method

Example 2: continuity The moments of the standard univariate normal distribution and chi-square distribution are well-known and from (5)(6) we have

$$E\left(\prod_{i=1}^{n} U_i^{r_i}\right) = \begin{cases} \dfrac{1}{(n/2)^{[\ell]}} \displaystyle\prod_{i=1}^{n} \dfrac{(2\ell_i)!}{4^{\ell_i}(\ell_i)!}, & \text{if } r_i = 2\ell_i \text{ are even, } i = 1,\ldots,n,\ r = 2\ell; \\ 0, & \text{if at least one of the } r_i \text{ is odd,} \end{cases} \tag{7}$$

where $x^{[\ell]} = x(x+1)\cdots(x+\ell-1)$, $\ell = \sum_{i=1}^{n} \ell_i$.

2. Hsu's Philosophy to My Work

2.1. Stochastic representation method

Example 3: Spherical distributions

♢ The spherical distribution can be defined by many ways. One way is to use its characteristic function. A $n \times 1$ random vector $\mathbf{x}$ is said to have an spherical distribution if its characteristic function is of the form $\phi(\mathbf{t}^T \mathbf{t})$.

♢ Another definition is by the stochastic representation method:

Let $\mathbf{u}^{(n)}$ be uniformly distributed on the unit sphere in R^n and R be a non-negative random variable and is independent of $\mathbf{u}^{(n)}$. The distribution of

$$\mathbf{x} \overset{d}{=} R\mathbf{u}^{(n)} \tag{8}$$

is called a spherical distribution. If $R \sim \chi_n$, the corresponding $\mathbf{x} \sim N_n(\mathbf{0}, \mathbf{\Sigma})$.

2. Hsu's Philosophy to My Work

Example 3: Spherical distributions, continuity

♢ It can be shown that these two definition are equivalent. But the latter can easily obtain many properties of the spherical distribution.

♢ Let **P** be an orthogonal matrix. The invariance of the distribution of **x** under any orthogonal transformation can be seen that

$$\mathbf{Px} \stackrel{d}{=} \mathbf{P}R\mathbf{u}^{(n)} = R\mathbf{Pu}^{(n)} \stackrel{d}{=} R\mathbf{u}^{(n)} \stackrel{d}{=} \mathbf{x}.$$

♢ By theory of the maximal invariant under a group, it is true that the c.f. of a spherical distribution should has of the form $\phi(\mathbf{t}^T\mathbf{t})$ and the distribution has of the form $F(\mathbf{x}^t\mathbf{x})$.

♢ From (7) and (8) We can easily to find the mixed moment

$$E(\prod_{i=1}^{n} X_i^{r_i}) = E(R^r)E(\prod_{i=1}^{n} U_i^{r_i}),$$

Example 3: Spherical distributions, continuity

♢ If we regard components $X_1, \ldots, X_n$ to be a sample, they have the identical distribution, but they are not in general independent.

♢ Let $\tau(\mathbf{x}) = \tau(X_1, \ldots, X_n)$ be a statistic, It is difficult to find the distribution of $\tau(\mathbf{x})$ directly. If we can change R in (8), a special $R^* \sim \chi_n$ can instead of R, we can find the distribution of spherical under the normality.

2.2. Some properties on $\stackrel{d}{=}$ operator

♢ Let $\mathbf{x} \stackrel{d}{=} \mathbf{y}$ and $h_j(\cdot)$ for $j = 1, \ldots, m$ be measurable functions. Then

$$(h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})) \stackrel{d}{=} (h_1(\mathbf{y}), \ldots, h_m(\mathbf{y})).$$

♢ Let random variable Z be independent of X and Y, respectively. Then (i) $X \stackrel{d}{=} Y$ implies $ZX \stackrel{d}{=} ZY$;

(ii) If $\Pr(X > 0) = \Pr(Y > 0) = \Pr(Z > 0) = 1$ and the cf of $\log Z$ satisfies $\varphi_{\log Z}(t) \neq 0$ for almost all t, then $ZX \stackrel{d}{=} ZY$ implies $X \stackrel{d}{=} Y$;

(iii) If $\Pr(Z > 0) = 1$ and the cf of $\log Z$ satisfies $\varphi_{\log Z}(t) \neq 0$ for almost all t, then $ZX \stackrel{d}{=} ZY$ implies $X \stackrel{d}{=} Y$.

2.2. Some properties on $\stackrel{d}{=}$ operator

For proving the above properties we need Hsu's paper.

Definition: We say that a c.f. belong to the class $(\bar{U})$ if it can equal another c.f. in the neighborhood of zero without equalling identically. A c.f. is said to belong to (U) if it does not belong to $(\bar{U})$.

♢ Essen (1945) proved that if the distribution has finite moments and is uniquely determined by these moments, it belong to (U);

♢ Zygmund (1951) pointed out some sufficient condition for belonging (U);

♢ Hsu (1954) reviewed the known examples of c.f.'s in $(\bar{U})$ and gave new examples and theorems.

2.3. Hsu's Contributions Appeared in My work

[2] Fang, K. T. and Zhang, Y. T. (1990), *Generalized Multivariate Analysis* Science Press and Springer-Verlag, Beijing and Berlin. Cited by 308 - Related articles Feb. 2010

[3] Fang, K. T., Kotz, S., and Ng, K. W. (1990), *Symmetric Multivariate and Related Distributions*, Chapman and Hall Ltd., London and New York. Cited by 830 - Related articles Feb. 2010

[4] Fang, K. T. and Anderson, T. W. (eds) (1990), *Statistical Inference in Elliptically Contoured and Related Distributions*, Allenton Press Inc., New York. Cited by 118 - Related articles Feb. 2010

These three books published 20 years now.

Prof. T.W. Anderson's Tribute

Dear Kai-Tai:

I want to congratulate you on your upcoming 65th Birthday. There has been a big poster in Sequoia Hall with the announcement of the celebration in Hong Kong next month. Of course, I was aware earlier of the occasion.

You and I have enjoyed a long period of cooperation in research, writing, and editing. I remember you coming here almost a quarter of a century ago to renew your academic career; I was glad to help you in that respect. Our collaboration over the ensueing years has been very productive and satisfying for me.

Prof. T.W. Anderson's Tribute

You have had a very impressive and useful career. Your energy and initiative and accomplishments have been outstanding. It has been a pleasure to collaborate with you.

Incidentally, we have more in common - the birthday of June 5.

With warm regards,

Ted

Hsu's Contributions Appeared in My work

美国IMS和ASA Fellow，美国乔治亚技术大学董永良教授，赞美我们这一段合作， 在他的回忆录(《回首学算路- 一个旅美学算者的故事》，2007，台湾商务印书馆)中按照杨振宁教授詠陈诗的格式送给我一首诗

文名驰中外，学誉在香江。
耕耘深亦广，桃李满门墙。
寻觅代表点，处处费思量。
统计加中药，深入资料矿。
千古寸心事，多维许安方。

其中最后一句中，"多维"指多元分析，"许"指许宝騄、"安"指安德森，"方"指方开泰。用来表示三人薪火相传之意。我和安德森教授带领年轻同事和研究生进行探索，在八年的研究工作中，发表了论文六十多篇，我们选择了四十篇，编辑成论文集，由位于纽约Allerton Press 出版社出版.

Conclusion

THANK YOU!

Kai-Tai Fang: Email: ktfang@uic.edu.hk

Beijing Normal University and Hong Kong Baptist University
United International College (UIC), Zhuhai, China
and
Chinese Academy of Sciences, Beijing