第17卷第2期　　西北农业大学学报　　Vol.17, No.2
1989年6月　　Acta Univ.Septentrionali Occident.Agric.　　Jun., 1989

χ^2—分布的数学归纳法证明

刘 光 祖

（基础课部）

摘　　要

本文用数学归纳法证明了χ^2分布及其性质、统计量$\frac{(n-1)s^2}{\sigma^2}$的分布，使不具备线性代数和较多微积分知识的读者也易于理解。

关键词：χ^2分布；统计量；数学归纳法

在数理统计中，χ^2分布是几种常用的重要分布之一。χ^2的概率密度在实数范围内一般采用以下两种证法[1-5]：一种是在n维空间上通过求χ^2在n维球壳上的平均概率密度，然后取极限而得到；另一种证法是借助于Γ分布而得到χ^2的概率密度。本文用读者熟悉的数学归纳法及二维随机变量的和分布即可证明。

设随机变量X_1，X_2，$\cdots X_n$，相互独立且均服从$N(0,1)$，则$X_1^2+X_2^2+\cdots+X_n^2$服从自由度为n的χ^2分布，即

$$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$$

其概率密度为

$$f(z)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}z^{\frac{n}{2}-1}e^{-\frac{z}{2}} & z>0\\ 0 & z\leqslant 0\end{cases}$$

证：当$n=1$时，由一维随机变量函数的分布，令$Z=\chi^2=X^2$，则

$$f_1(z)=\begin{cases}\dfrac{1}{2\sqrt{z}}\left[f(\sqrt{z})-f(-\sqrt{z})\right]=\dfrac{1}{\sqrt{2\pi z}}e^{-\frac{z}{2}} & \\ \qquad=\dfrac{1}{2^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)}z^{-\frac{1}{2}}e^{-\frac{z}{2}} & z>0\\ 0 & z\leqslant 0\end{cases}$$

当$n=2$时，$\chi^2=X_1^2+X_2^2$，令$Z=\chi^2$，$X=X_1^2$，$Y=X_2^2$，则$Z=X+Y$的分布为（显然

本文于1988年1月16日收到

z的取值$z>0$）

$$f_2(z)=\int_{-\infty}^{+\infty}f_x(x)f_y(z-x)dx$$

$$=\int_0^z\frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})}x^{-\frac{1}{2}}e^{-\frac{x}{2}}\cdot\frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})}(z-x)^{-\frac{1}{2}}e^{-\frac{z-x}{2}}dx$$

$$=\frac{e^{-\frac{z}{2}}}{2\Gamma^2(\frac{1}{2})}\int_0^z x^{-\frac{1}{2}}(z-x)^{-\frac{1}{2}}dx$$

$$\overset{z-x=zt}{=\!=\!=}\frac{e^{-\frac{z}{2}}}{2\Gamma^2(\frac{1}{2})}\int_0^1(1-t)^{-\frac{1}{2}}t^{-\frac{1}{2}}dt$$

$$=\frac{e^{-\frac{z}{2}}}{2\Gamma^2(\frac{1}{2})}B(\frac{1}{2},\frac{1}{2})^{\bullet}$$

$$=\frac{e^{-\frac{z}{2}}}{2\Gamma^2(\frac{1}{2})}\cdot\frac{\Gamma^2(\frac{1}{2})}{\Gamma(1)}$$

$$=\frac{1}{2}e^{-\frac{z}{2}}$$

即

$$f_2(z)=\begin{cases}\frac{1}{2}e^{-\frac{z}{2}}=\frac{1}{2^{\frac{2}{2}}\Gamma(\frac{2}{2})}z^{\frac{2}{2}-1}e^{-\frac{z}{2}} & z>0\\ 0 & z\leqslant 0\end{cases}$$

由以上证明假设$n=k$结论成立，即

$$f_k(z)=\begin{cases}\frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}x^{\frac{k}{2}-1}e^{-\frac{x}{2}} & x>0\\ 0 & x\leqslant 0\end{cases}$$

•Γ函数与B函数有如下关系：

$$B(\alpha,\beta)\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

证明结论对$n=k+1$也成立。令$Z=\chi^2=X_1^2+X_2^2+\cdots+X_{k+1}^2$，$X=X_1^2+X_2^2+\cdots+X_k^2$，$Y=X_{k+1}^2$，则$Z=X+Y$的分布$(z>0)$

$$f_{k+1}(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx$$

$$=\frac{e^{-\frac{x}{2}}}{2^{\frac{k+1}{2}}\Gamma(\frac{k}{2})\Gamma(\frac{1}{2})}\int_0^z x^{\frac{k}{2}-1}(z-x)^{-\frac{1}{2}}dx$$

$$\overset{z-x=zt}{=\!=\!=\!=}\frac{z^{\frac{k+1}{2}-1}e^{-\frac{z}{2}}}{2^{\frac{k+1}{2}}\Gamma(\frac{k}{2})\Gamma(\frac{1}{2})}\int_0^1(1-t)^{\frac{k}{2}-1}t^{-\frac{1}{2}}dt$$

$$=\frac{z^{\frac{k+1}{2}-1}e^{-\frac{z}{2}}}{2^{\frac{k+1}{2}}\Gamma(\frac{k}{2})\Gamma(\frac{1}{2})}B(\frac{k}{2},\frac{1}{2})$$

$$=\frac{1}{2^{\frac{k+1}{2}}\Gamma(\frac{k+1}{2})}z^{\frac{k+1}{2}-1}e^{-\frac{z}{2}}$$

当$z\leqslant 0$时$f_{k+1}(z)=0$。至此结论全部获证。

对于χ^2分布的性质以及某些统计量的分布，也可用数学归纳法证明。

1．若$\chi_i^2\sim\chi_i^2(n_i)$，$i=1,2,\cdots,k$，且相互独立，则

$$\sum_{i=1}^{k}\chi_i^2\sim\chi^2(\sum_{i=1}^{k}n_i)$$

证：当$k=2$时，设$X^2\sim\chi^2(n_1)$，$\chi_2^2\sim\chi^2(n_2)$，与χ^2分布的证明类似，令$Z=\chi_1^2+\chi_2^2=X+Y$，即得

$$f_1(Z)=\begin{cases}\dfrac{1}{2^{\frac{n_1+n_2}{2}}\Gamma(\frac{n_1+n_2}{2})}Z^{\frac{n_1+n_2}{2}}e^{-\frac{z}{2}} & z>0\\ 0 & z\leqslant 0\end{cases}$$

假设对于$k-1$性质成立，即

$$f_{k-1}(x)=\begin{cases}\dfrac{1}{2^{\frac{1}{2}\sum_{i=1}^{k-1}n_i}\Gamma(\frac{1}{2}\sum_{i=1}^{k-1}n_i)}X^{\frac{1}{2}\sum_{i=1}^{k-1}n_i-1}e^{-\frac{X}{2}} & X>0\\ 0 & X\leqslant 0\end{cases}$$

证明对于k性质也成立。$Z=(X_1^2+X_2^2+\cdots+X_{k-1}^2)+X_k^2=X+Y$，当$z>0$时

$$f_k(z)=\int_0^z \frac{1}{2^{\frac{1}{2}\sum_{i=1}^{k-1}n_i}\Gamma(\frac{1}{2}\sum_{i=1}^{k-1}n_i)}X^{\frac{1}{2}\sum_{i=1}^{k-1}n_i}e^{-\frac{x}{2}}\cdot\frac{1}{\frac{n_k}{2}\Gamma(\frac{n_k}{2})}e^{-\frac{z-x}{2}}dx$$

$$\xlongequal{z-x=zt}\frac{1}{2^{\frac{1}{2}\sum_{i=1}^{k}n_i}\Gamma(\frac{1}{2}\sum_{i=1}^{k-1}n_i)\Gamma(\frac{n_k}{2})}\int_0^1(1-t)^{\frac{1}{2}\sum_{i=1}^{k-1}n_i-1}t^{\frac{n_k}{2}-1}dt$$

$$=\frac{1}{2^{\frac{1}{2}\sum_{i=1}^{k}n_i}\Gamma(\frac{1}{2}\sum_{i=1}^{k}n_i)}z^{\frac{1}{2}\sum_{i=1}^{k}n_i-1}e^{-\frac{z}{2}}$$

当$z\leqslant 0$时，$f_k(z)=0$。χ^2分布的这一性质得到了证明。

2．统计量$\frac{(n-1)S^2}{\sigma^2}$分布的证明

设$X_1,X_2,\cdots,X_n$为正态总体$N(\mu,\sigma^2)$的一个样本，

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2,\text{ 则 }\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

证：当$n=2$时，$\overline{X}=\frac{1}{2}(X_1+X_2)$

$$\frac{S^2}{\sigma^2}=\frac{1}{\sigma^2}(\sum_{i=1}^{2}X_i^2-2\overline{X}^2)$$

$$=\frac{2}{\sigma^2}(\frac{X_1-X_2}{2})^2$$

$$=(\frac{X_1-X_2}{\sqrt{2}\sigma})^2$$

显然$\frac{X_1-X_2}{\sqrt{2}\sigma}\sim N(0,1)$，故$\frac{S^2}{\sigma^2}\sim\chi^2(1)$。

假设$n=k$结论成立，即

$$\frac{(k-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}(\sum_{i=1}^{k}X_i^2-k\overline{X}'^2)\sim\chi^2(k-1)$$

其中$\overline{X}'=\frac{1}{k}\sum_{i=1}^{k}X_i$，证明$n=k+1$结论成立。

$$\frac{kS^2}{\sigma^2}=\frac{1}{\sigma^2}\left[\sum_{i=1}^{k+1}X_i^2-(k+1)\overline{X}^2\right]$$

$$=\frac{1}{\sigma^2}\left[\sum_{i=1}^{k}X_i^2-k\overline{X'}^2+k\overline{X'}^2+X_{k+1}-(k+1)\overline{X}^2\right]$$

$$=\frac{1}{\sigma^2}\left(\sum_{i=1}^{k}X_i^2-k\overline{X'}^2\right)+\frac{1}{\sigma^2}\left[k\overline{X'}^2+X^2_{k+1}-\frac{1}{k-1}(k\overline{X'}^2+X_{k+1})^2\right]$$

$$=\frac{1}{\sigma^2}\left(\sum_{i=1}^{k}X_i^2-k\overline{X'}^2\right)+\frac{1}{\sigma^2}\cdot\frac{k}{k+1}(X_{k+1}-\overline{X'})^2$$

显然 $y=\frac{1}{\sigma^2}\left(\sum_{i=1}^{k}X_i^2-k\overline{X'}^2\right)\sim\chi^2_{(k-1)}$。因 $\sqrt{\frac{1}{\sigma^2}\cdot\frac{k}{k+1}}(X_{k+1}-\overline{X'})\sim N(0,1)$，$X_{k+1}$是与$\overline{x}'$无关的独立变量，所以

$$Z=\frac{1}{\sigma^2}\cdot\frac{k}{k+1}(X_{k+1}-X'^2)\sim\chi^2_{(1)}$$

且Y与Z相互独立，由1的证明知$Y+Z\sim\chi^2_{(k)}$，亦即

$$\frac{kS^2}{\sigma^2}\sim\chi^2_{(k)}$$ 证毕。

由以上对χ^2分布及其性质、统计量$(n-1)s^2/\sigma^2$分布的数学归纳法证明可以看出，二维随机变量，尤其是二维随机变量和分布的重要。仿此方法并借助于二维随机变量的商分布，可容易证明t分布，F分布。因此在概率论中只要重点讲清楚二维随机变量的和、商分布，则数理统计中的几个重要分布就能容易得到证明。另外还可以看出，对于某些随机变量的分布，特别是服从同一分布的n个随机变量的和分布，可以考虑用数学归纳法试求其分布。如

设$X_1,X_2,\cdots,X_n$相互独立且服从参数分别为$\lambda,\lambda_2,\cdots,\lambda_n$的泊松分布，则$Z=X_1+X_2+\cdots+X_n$服从参数为$\lambda_1+\lambda_2+\cdots+\lambda_n$的泊松分布。

证：当$n=2$时，$z=X_1+X_2$，

$$p_{(X_1=k)}=\frac{\lambda_1^k e^{-\lambda 1}}{k!}\qquad k=0,1,2,\cdots$$

$$p_{(X_2=i)}=\frac{\lambda_2^i e^{-\lambda 2}}{i!}\qquad i=0,1,2,\cdots$$

则

$$P_{(Z=X_1+Y_2=i)}=\sum_{k=0}^{i}\frac{\lambda_1^k e^{-\lambda 1}}{k!}\cdot\frac{\lambda_2^{i-k}e^{-\lambda 2}}{(i-k)!}$$

$$=e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{i}\frac{\lambda_1^k}{k!}\cdot\frac{\lambda_2^{i-k}}{(i-k)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{i!} \sum_{i=0}^{k} C_i^{\,k} \lambda_1^{\,k} \lambda_2^{\,i-k}$$

$$= \frac{(\lambda_1+\lambda_2)^i}{i!} e^{-(\lambda_1+\lambda_2)} \qquad i=0,\ 1,\ 2,\ \cdots$$

假设$X=X_1+X_2+\cdots+X_{n-1}$服从参数为$\lambda_1+\lambda_2+\cdots+\lambda_{n-1}$的松泊分布，$Z=(X_1+X_2+\cdots+X_{n-1})+X_n=X+X_n$，令$\lambda=\lambda_1+\lambda_2+\cdots+\lambda_{n-1}$，则

$$P_{(Z=i)} = \sum_{k=0}^{i} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\lambda_n^{\,i-k}}{(i-k)!} e^{-\lambda n}$$

$$= \frac{(\lambda+\lambda n)^i}{i!} e^{-(\lambda+\lambda n)} \qquad i=0,\ 1,\ 2,\ \cdots$$

即$Z=X_1+X_2+\cdots+X_n$服从参数为$\lambda=\lambda_1+{}_2+\cdots+\lambda_n$的松泊分布。

对于$X_i\sim N(\mu_i,\sigma_i^2)$，$i=1,2,\cdots,n$及$X_i\sim B(n,\ p)$，$i=1,2,\cdots,n$等，仿上述证明；其和分布是不难求得的。

参 考 文 献

1 中山大学编.概率论与数理统计（上册）.高等教育出版社,1980.
2 陈希孺编.数理统计引论.科学出版社，1981
3 西北农业大学编.概率基础与数理统计.农业出版社，1988
4 复旦大学编.概率论（第二分册）.人民教育出版社，1979
5 王福保等编.概率论及数理统计.同济大学出版社，1984

A PROOF OF X^2 DISTRIBUTION BY MATHEMATICAL INDUCTION

Liu Guangzu

(*Department of Basic and Common Courses*)

Asbtract

This paper proves χ^2 distribution and its nature, and the statistical quantity of the distribution of $\frac{(n-1)s^2}{\sigma^2}$ by mathematical induction, which is easily comprehensible to teaders knowing little about infinitesimal calculus or ignorant of linear algebra.

Key Words: χ^2 distribution; statistical quantity; mathematical induction