# Mathematical Statistics

Velen Kong

2019 年 8 月 7 日

**摘要**

出于下学期的课程以及数学建模的需要，只好暑假自学数理统计，自学首选课本为

韦来生.著. 数理统计(第二版) [M] 北京:科学出版社 2015

由于本人水平限制，大概率不会涉及以下参考资料

陈希孺.著. 高等数理统计学 [M] 合肥:中国科学技术大学出版社 2009

(美) Hogg R.V. & Craig A.T.著. Introduction to Mathematical Statistics:Fifth Edition 数理统计学导论：第5版(影印版) [M] 北京:高等教育出版社 2004

(美) Casella G. & Berger R.L.著. Statistical Inference 统计推断(原书第二版) [M] 张忠占 & 傅莺莺.译. 北京:机械工业出版社 2009

## 1 绪论及预备知识

几个基本概念，总体(population)，个体(individual)，样本(sample)，抽样(sampling)，随机变量(random variable)，观察值(observation)。总体又分为有限总体(finite population)和无限总体(infinite population)。

然后又有样本空间(sample space)，简单随机样本，联合分布函数和联合密度函数

$$F(x_1,\ldots,x_n)=\prod_{i=1}^{n}F(x_i)$$

$$f(x_1,\ldots,x_n)=\prod_{i=1}^{n}f(x_i)$$

对于多维样本同样有

$$F(X,Y)=\prod_{i=1}^{n}F(x_i,y_i)$$

$$f(X,Y)=\prod_{i=1}^{n}f(x_i,y_i)$$

确定统计模型(statistical model)后，从总体中抽取一定的样本来推断总体模型被称为统计推断(statistical inference)。需要估计的是参数向量(parameter vector)，参数的取值范围就是参数空间(parameter space)。除此以外还有未知样本分布的非参数统计推断。

而由于参数的不确定性，统计模型是一个样本分布族(distribution family of the sample)。

统计量(statistic)则是样本算出的值，例如样本均值(mean)和方差(variance)。下面介绍均值和方差的推广——样本矩(sample moments)。

设$X_1,\dots,X_n$是总体中抽取的样本，则称

$$a_{n,k}=\frac{1}{n}\sum_{i=1}^{n}X_i^k,\qquad k=1,2,\dots$$

为样本$k$阶原点矩。

$$m_{n,k}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k$$

为样本$k$阶中心矩。

对于二维总体而言，称

$$S_{XY}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{X})$$

为$X$和$Y$的样本协方差(sample covariance)。

然后我们介绍次序统计量(order statistics)。将$X_1,\dots,X_n$按递增次序排列为$X_{(1)},\dots,X_{(n)}$就是样本的次序统计量。

利用次序统计量可以定义样本中位数(sample median)，样本极值(extremum of sample)，样本$p$分位数(sample p-fractile)，样本极差(sample range)。其中$p$分位数定义为$X_{(m)},m=[(n+1)p]$。

样本变异系数(sample coefficient of variation)则是对于总体变异系数(population coefficient of variation)的估计。总体变异系数衡量总体分布的散布程度，定义为

$$\nu = \frac{SD(X)}{E(X)}$$

而样本变异系数则为

$$\nu = \frac{S_n}{\overline{X}}$$

样本偏度(sample skewness)反映了总体偏度的信息，定义是

$$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

其中正态分布的偏度为0。

而样本偏度则定义为

$$\hat{\beta}_1 = \frac{m_{n,3}}{m_{n,2}^{3/2}}$$

样本峰度(sample kurtosis)则是用来反映总体峰度的信息，总体峰度表示密度函数在最大值附近的集中程度，正态分布的峰度为0。总体峰度定义为

$$\beta_2 = \frac{\mu_4}{\mu_2^2} - 3$$

而样本峰度则定义为

$$\hat{\beta}_2 = \frac{\mu_4}{\mu_2^2} - 3$$

接下来介绍经验分布函数，对于一组次序统计量，我们定义

$$F_n(x) = \begin{cases} 0, & x \le X_{(1)}, \\ \frac{k}{x}, & X_{(k)} < x \le X_{(k+1)} \\ 1, & X_{(n)} < x \end{cases}$$

为经验分布函数(empirical distibution function)。易见经验分布函数是单调不减左连续的，它可以看作总体分布函数的一个估计量。

若记示性函数为

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & others \end{cases}$$

那么$F_n(x)$可以表示为

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,x)}(X_i)$$

所以有

$$nF_n(x) = \sum_{i=1}^{n} Y_i\, b(n, F(x))$$

由中心极限定理当$n \to \infty$时有

$$\frac{\sqrt{n}(F_n(x) - F(x))}{\sqrt{F_n(x) - F(x)}} \xrightarrow{\mathcal{L}} N(0,1)$$

其中$\xrightarrow{\mathcal{L}}$表示依分布收敛。

又由Bernoulli大数定律得

$$F_n(x) \xrightarrow{P} F(x) \qquad n \to \infty$$

由Borel大数定律得

$$P(\lim_{n\to\infty} F_n(x) = F(x)) = 1$$

更进一步，有下列格里汶科定理(Glivenko-Cantelli Theorem)。

定理设$F(x)$为$r.v.\ X$的分布函数，$\{X_i\}$为简单随机样本，$F_n(x)$为其经验分布函数，记$D_n = \sup\limits_{-\infty<x<+\infty} |F_n(x) - F(x)|$，则有

$$P(\lim_{n\to\infty} D_n = 0) = 1$$

其表明在$n$足够大时，对于所有的$x$，经验分布函数的误差都很小。

# 2 抽样分布

## 2.1 正态分布

## 2.2 次序统计量

## 2.3 $\chi^2$分布

## 2.4 $t$分布

## 2.5 $F$分布

## 2.6 重要推论

## 2.7 极限分布

## 2.8 指数族

## 2.9 充分统计量

## 2.10 完全统计量

# 3 点估计