关于平均数和数学期望的应用

□ 苏均和

在统计学中常用平均数来表述数据(或数列)集中趋势的测度。平均数有:算术平均数、调和平均数、几何平均数、中位数及众数。各种平均数有各自的定义和它们应用的范围。

算术平均数是最常用的一种。它有简单算术平均数与加权算术平均数之分。简单算术平均数是将总体的各个单位标志值简单相加,然后除以单位个数,求出平均标志值;加权算术平均数是用各组标志值乘以相应的总体单位数来计算的,由此,它的大小不仅取决于总体各单位的变量值,而且受单位变量重复出现的次数影响,各组出现的次数在这里面起了权衡轻重的作用。

调和平均数是各个变量值倒数的算术平均数,又称倒数平均数。在统计工作中应用机会很少,往往是作为算术平均数的变形来使用的。

几何平均数是计算平均比率和平均速度时适用的一种方法。几何平均数也有简单几何平均数和加权几何平均数之分。只有当变量值总量表现为各变量值的连乘积时,才适用几何平均数方法来计算平均值。一般说来,计算社会经济现象在各个时期的平均发展速度时,则采用几何平均数。

中位数是一种按其在数据中的特殊位置而决定的平均数。它将总体中某数量标志的各个数值按大小顺序排列,形成一个数列,处在数列中点位次的数值即为中位数。中位数的最大特点是:它是数据序列中间一项或两项的平均数,不受极端值影响,所以当一个数据序列中含有特大值或特小值的情况下,采用中位数较为适宜。

众数也是一种位置平均数,是指在总体中出现次数最多的那个变量值。众数的计算只适用于单位数较多,且有明显的集中趋势。否则计算众数是没有意义的。

为了使得平均值更精确、更科学化,还可以用另一种语言来进行描述,即数学期望。

在自然界和社会生活中,一般有两种现象:一种是确定性的,如太阳从东方升起等等;另一种是不确定现象,如抛掷一枚硬币,出现正面还是反面,在抛掷以前是无法知道的。这类现象称为随机现象。在客观世界中随机现象是普遍存在的。

随机现象和随机试验是密不可分的,所谓随机试验就是一个试验满足下述条件:

(1)试验可以在相同的情形下重复进行;

(2)试验的所有可能结果是明确可知的,并且不止一个;

(3)每次试验恰好出现这些可能结果的一个,但试验前不能肯定会出现那一个结果。

随机试验每一个可能结果,称为基本事件。由多个基本事件所组成的事件称为随机事件。比如一个盒子中有十个完全相同的球,分别标以号码1,2,…,10,令i={取得球的标号为i},则从盒子中任取一球标号为i为一基本事件,而从盒中任取一球标号为偶数为一随机事件。

度量一个随机事件出现可能性大小的量常用概率来刻画,而一个随机事件总是和它的概率紧密相联的。

随机事件虽然与实数之间没有直接自然关系,但我们常可以人为地给它们建立起一个对应关系,比如抛掷一枚硬币,约定 $\xi=\begin{cases}1 & \text{出现正面}\\ 0 & \text{出现反面}\end{cases}$,$\xi$是一个变量,它取什么值在每次试验以前是不能确定的,也就是说,它的取值是随机的,我们一般把这种变量称为随机变量。随机变量就它们的取值可分为离散型和连续型。我们把下表称为是离散型随机变量ξ的分布列。

表4

ξ	$a_1, a_2, a_3, \cdots a_n \cdots$
P_i	$P_1, P_2, P_3, \cdots P_n \cdots$

其中 $P_i = P(\xi = a_i)$

而把随机变量ξ的数学期望定义为:

$E\xi = \sum\limits_{i=1} a_i P_i$,从它的定义我们可以看出,它是以概率为权数的平均数,因此它较前面的平均数而言就更为精确而科学。设ξ和η为两个随机变量,则有$(\xi、\eta)$的联合分布列及条件期望等等一些问题。

现在我们已经知道数学期望的含义，但对于它的作用，尤其是关于它在各方面的应用，迄今未能引起人们重视。当前，在社会主义市场经济条件下，如何利用数学期望来解决一些经济问题是很有必要的。由于篇幅关系下面仅举一例，说明数学期望的应用。

例，(抽样信息的价值)某公司拟扩展其生产线，有三项备选方案。方案Ⅰ，建造大型厂房；方案Ⅱ，建造中型厂房；方案Ⅲ，租赁合适的设备。产品的需求情况分为高、低两档，以 $S=1$ 表示需求量高，$S=2$ 表示需求量低。以 R_1、R_2、R_3 分别表示方案Ⅰ、Ⅱ、Ⅲ的收益，它依赖于需求情况 S，其在各情况下的收益值如表所示：

单位：万元

方案 \ 收益 \ 需求	Ⅰ	Ⅱ	Ⅲ
	R_1	R_2	R_3
S=1	18	12	10
S=2	-15	-5	-1

今考虑聘请市场顾问对市场需求情况进行预测。在此之前我们已知需求随机变量 S 的概率分布为

$P(S=1)=0.6, P(S=2)=0.4$

这一概率分布常称为事前分布(或先验分布)。以 $C=1$ 和 $C=2$

$$P(S=2 \mid C=1)=\frac{8}{29}$$

$$P(S=1 \mid C=2)=\frac{P(S=1, C=2)}{P(C=2)}=\frac{0.18}{0.18+0.14}=\frac{3}{7}$$

$$P(S=2 \mid C=2)=\frac{4}{7}$$

这两个条件分布是在已知预测信息后得到的称为事后分布(或后验分布)，关于该后验分布求条件期望称为后验期望。分别求出三个方案在 $C=1$ 和 $C=2$ 下的后验期望为

$$\begin{pmatrix} E(R_1 \mid C=1) \\ E(R_2 \mid C=1) \\ E(R_3 \mid C=1) \end{pmatrix}=\begin{pmatrix} 18 & -15 \\ 12 & -5 \\ 10 & -1 \end{pmatrix}\begin{pmatrix} \frac{21}{29} \\ \frac{8}{29} \end{pmatrix}=\begin{pmatrix} 8.9 \\ 7.3 \\ 7.0 \end{pmatrix}$$

$$\begin{pmatrix} E(R_1 \mid C=2) \\ E(R_2 \mid C=2) \\ E(R_3 \mid C=2) \end{pmatrix}=\begin{pmatrix} 18 & -15 \\ 12 & -5 \\ 10 & -1 \end{pmatrix}\begin{pmatrix} \frac{3}{7} \\ \frac{4}{7} \end{pmatrix}=\begin{pmatrix} -0.81 \\ 2.29 \\ 3.71 \end{pmatrix}$$

$$\max\{E(R_1 \mid C=1), E(R_2 \mid C=1), E(R_2 \mid C=1)\}=E(R_1 \mid C=1)=8.9$$

$$\max\{E(R_1 \mid C=2), E(R_2 \mid C=2), E(R_2 \mid C=2)\}=E(R_3 \mid C=2)=3.71$$

于是按后验期望收益最大原则的决策规则为：当预测需求量 $C=1$ 时，执行方案Ⅰ，当预测需求量 $C=2$ 时，执行方案Ⅲ，不论 C 值如何均不执行方案Ⅱ。此时决策的收益是 S 和 C 的函数，记为 $R^*(S,C)$，即

$$R^*(S,C)=\begin{cases} 18, & 当\ C=1, S=1 \\ -15, & 当\ C=1, S=2 \\ 10, & 当\ C=2, S=1 \\ -1, & 当\ C=2, S=2 \end{cases}$$

由(S,C)的联合分布，可得

$$E[R^*(S,C)]=18\times0.42+(-15)\times0.16+10\times0.18+(-1)\times0.24=6.72$$

聘请顾问得到的预测信息是不完全信息或称抽样信息，它的价值为

$$E[R^*(S,C)]-E(R_3)=6.72-5.6=1.12$$

即聘请顾问比不聘请顾问平均说来可净增收益 1.12 万元，而聘请费用仅需 5000 元，因而聘请该顾问是值得的。

在信息市场中信息是有售价的，仅当该抽样信息的价值大于其售价时，该信息才值得购买。一般说来计算抽样信息的价值，其步骤如下：

(1)利用市场需求随机变量 S 的先验分布求出各个方案的期望收益，期望收益最大的方案即为在没有咨询信息情况下的决策方案，该方案的期望收益记为 E(R)。

(2)由咨询信息 C 计算 S 的后验分布，它依赖于 C。

(3)由后验分布计算各方案的后验期望收益，它是 C 的函数。

(4)对于 C 的每一个值，选出后验期望收益最大的方案，这就是利用咨询信息 C 的情况下的决策规则，此时收益是 S 和 C 的函数，记为 $R^*(S,C)$。

(5)计算信息的价值：$E[R^*(S,C)]-E(R)$。

(作者单位：上海财经大学统计系)