Vote Bounded above Quantitative Model

Lu Shengge
Statistics Department, Zhongnan University of Finance and Law,
Wuhan, Hubei, 430064, China
E-mail: jtx@znufe.edu.cn

In judgment work, suppose we have m judges and n objects. And the objects can only belong to category A or $\overline{A}$. An object will belong to category A only when there are more than αm-1 (α is a definite figure, 0<α≤1) judges decide the object is of category A. We define the objects of category A is less than h+1, and vote bounded above x of each judge is to be determined. Assume that n objects are selected, that is, each object might be divided into category A. Further more, mutual influence doesn't exist among the judges. Because the judges hold different understandings and opinions to the judgment standard, we can take it at random that which object will be divided into category A or $\overline{A}$. Hence, we obtain the random variables:

$$Y_{ij}=\begin{cases}1\\0\end{cases}(j=1,2,\cdots,m;\ i=1,2,\cdots,n)$$

we can believe random variables $Y_{i1}, Y_{i2},\cdots Y_{im}(i=1,2,\cdots,n)$ are mutually independent, and they all submit to (0-1) distribution. So their probability distributions are

$$P\{Y_{ij}=1\}=x/n\,,\ P\{Y_{ij}=0\}=1-x/n\,.$$

Sign the random variable as

$$Y_i=Y_{i1}+Y_{i2}+\cdots+Y_{im}(i=1,2,\cdots,n)$$

Obviously, Y_i is votes of the ith objects belonging to category A. Y_i submits to the binominal distribution b (m, x/n) with the parameters m and x/n , mathematical expectation $E(Y_i)= mx/n$, Variance $Var(Y_i)=mx/n\,(1-x/n)\approx 0.25m$, $\sqrt{Var(Y_i)}\approx 0.5\sqrt{m}$ $(i=1,2,\cdots,n)$. Define

$$r=\begin{cases}\alpha m\\ [\alpha m]+1\end{cases}$$

where [αm] is the largest integer of αm. The probability

P{the ith object is divided into category A}

$$=P\{Y_i\ge r\}=P\{(Y_i-mx/n)/0.5\sqrt{m}\ge(r-mx/n)/0.5\sqrt{m}\}\triangleq p(i=1,2,\cdots,n)(1)$$

According to the central limit theorem of De Moivre-Laplace in probability theory, when m is comparatively large, $(Y_i-mx/n)/0.5\sqrt{m}$ approximately obeys standard normal distribution, the Formula (1) can be denoted by

$$1-\Phi((nr-mx)/0.5n\sqrt{m})\approx p\ (i=1,2,\cdots,n)(2)$$

where $\Phi(z)$ is the distribution function of standard normal distribution. Sign the random variables as

$$Y_i'=\begin{cases}1\\0\end{cases}(i=1,2,\cdots,n)$$

Hence, $Y_1', Y_2', \cdots Y_n'$ are mutually independent and all of them submit to (0-1) distribcction. Their probability distributions are

$P\{Y_i'=1\}=P\{Y_i \geq r\}=p, P\{Y_i'=0\}=P\{Y_i<r\}=1-p (i=1,2,\cdots n)$

Finally, Y out of n objects have been divided into category A, then the random variable $Y=Y_1'+Y_2'+\cdots+Y_n'$ submits to binomial distribution b(n, p) with the parameters n and p, mathematical expectation E(Y)=np, variance Var(Y)=np(1-p)≈0.25n, $\sqrt{\mathrm{Var}(Y)} \approx 0.5\sqrt{n}$. According to the central limit theorem of De Moivre-Laplace, when n is comparatively large, (Y-np)/0.5 $\sqrt{n}$ will approximatively submit to standard normal distribution, Let E(Y)=np=h-k(generally, 0≤k<h), then p=(h-k)/n. From Formula (2), when p=(h-k)/n≤0.5, We have

$$\Phi^{-1}(1-(h-k)/n) \approx (nr-mx)/0.5n\sqrt{m} \quad (3)$$

when $p=(h-k)/n>0.5$, we obain

$$\Phi^{-1}((h-k)/n) \approx -(nr-mx)/0.5n\sqrt{m} \quad (4)$$

Synshesizing Formula (3) and (4), we arrive at the vote bounded above quantitative model:

$$x=\begin{cases}[x']+1 & \text{if } (x')^2>[x']([x']+1) \\ [x'] & \text{if not}\end{cases} \quad (5)$$

where

$$x'=\begin{cases}[r-0.5\sqrt{m}\Phi^{-1}(1-(h-k)/n)]n/m & \text{if (h-k)/n}\leq 0.5 \\ [r+0.5\sqrt{m}\Phi^{-1}((h-k)/n)]n/m & \text{if (h-k)/n}>0.5\end{cases}$$

According to Formula (5) and (2) we obtain p and np=h-k. Then we can show

$P\{(h-k)-0.5\sqrt{n} \leq Y \leq (h-k)+0.5\sqrt{n}\} \approx 68.27\%$

In practice, if only the total votes approximate to mx and each object is having been competitive, then we will almost get

$P\{Y \leq (h-k)-0.5\sqrt{n}\} \leq 5\%$ and $P\{Y>(h-k)+0.5\sqrt{n}\} \leq 5\%$ (6)

then the estimation interval of Y

$((h-k)-0.5\sqrt{n}, (h-k)+0.5\sqrt{n})$ (7)

Its guarantee probability isn't less than 90%.

In practice, at first we can properly select several values of k. From model (5) and (7) we can respectively get several corresponding values of x and estimation interval of Y. Then according to the concrete conditions in judgment work, we can determine the votes bounded above x and the corresponding estimation interval of Y.