

# Note on Randomized Response Technique, RRT

by 邢超

xingnix@live.com

对敏感性问题的调查方案,为了能保护个人秘密,使被调查者愿意做出真实回答,产生了相应的统计分析方法——随机化回答技术,包括:

- Warner model
- Simmons model

## 1 Warner model

由美国统计学家Warner提出,针对要调查的敏感性问题的,列出正反两个问题。如:调查考试作弊问题时,出示两个问题:

A:你考试作弊?

B:你考试没有作弊?

然后由被调查者按照给定的概率随机选择一个问题,与自己情况一致则回答"是",否则回答"否"。调查者不知道被调查者在回答哪个问题,保证了被调查者个人秘密不被泄露。需要注意的是:选择问题A的概率是确定的。

理论分析如下:

设:

$P(A)$ :选择问题A的概率,

$P(B)$ :选择问题B的概率,

则有:

$$P(A) + P(B) = 1$$

设:

$P(T)$ :回答"是"的概率,

$P(T|A)$ :选择问题A并回答"是"的概率,

$P(T|B)$ :选择问题B并回答"是"的概率,

$$P(T|A) + P(T|B) = 1$$

如果对n个人进行调查,调查结果中有m个人回答"是",有n-m个人回答"否",可得:

$$\begin{aligned} P(T|A)P(A) + P(T|B)P(B) &= P(T) \\ P(T|A)P(A) + (1 - P(T|A))(1 - P(A)) &= P(T) \\ P(T|A) &= \frac{P(T) + P(A) - 1}{2P(A) - 1} \\ \hat{P}(T|A) &= \frac{\hat{P}(T) + P(A) - 1}{2P(A) - 1} \\ &= \frac{\frac{m}{n} + P(A) - 1}{2P(A) - 1} \end{aligned}$$

根据二项分布质可得:  
期望

$$\begin{aligned} E[\hat{P}(T)] &= P(T) \\ E[\hat{P}(T|A)] &= P(T|A) \end{aligned}$$

方差

$$\begin{aligned}\text{Var}[\hat{P}(T)] &= \frac{P(T)(1-P(T))}{n} \\ \text{Var}[\hat{P}(T|A)] &= \frac{P(T)(1-P(T))}{n(2P(A)-1)^2}\end{aligned}$$

## 2 Simmons model

Warner的方法中回答A的人数比例不能为 $\frac{1}{2}$ 。1967年Simmons对Warner模型进行了改进。调查人员提出两个不相关的问题,其中一个为敏感性问题A,另一个为非敏感性问题B。如:调查考试作弊问题时,出示两个问题:

A:你考试作弊?

B:在心里想一个数字,它是否是奇数?

然后由被调查者按照给定的概率随机选择一个问题。

理论分析如下:

设:

$P(A)$ :选择问题A的概率,

$P(B)$ :选择问题B的概率,

得:

$$P(A) + P(B) = 1$$

设:

$P(T)$ :回答"是"的概率,

$P(T|A)$ :选择问题A并回答"是"的概率,

$P(T|B)$ :选择问题B并回答"是"的概率,

得:

$$\begin{aligned}P(T|A)P(A) + P(T|B)P(B) &= P(T) \\ P(T|A)P(A) + P(T|B)(1-P(A)) &= P(T) \\ P(T|A) &= \frac{P(T) - P(T|B)(1-P(A))}{P(A)} \\ \hat{P}(T|A) &= \frac{\hat{P}(T) - P(T|B)(1-P(A))}{P(A)}\end{aligned}$$

二项分布质可得:

期望

$$E[\hat{P}(T|A)] = P(T|A)$$

方差

$$\begin{aligned}\text{Var}[\hat{P}(T)] &= \frac{P(T)(1-P(T))}{n} \\ \text{Var}[\hat{P}(T|A)] &= \frac{P(T)(1-P(T))}{n(P(A))^2}\end{aligned}$$