

姓名

日期

时期

家庭辅助学习资料

统计推断

在本单元中，学生将运用小型数据样本来估计有关一个较大群组（称为**群体**）的信息，并利用模拟来确定估计值的范围。群体是对某个问题而言，相关对象的整个集合，样本是该群体中的较小群组。



例如，我们可能想计算出美国家庭每月在食品上花费的**平均**金额。群体包括美国的所有家庭，但从每个人那里收集信息会非常困难，而且会花费很多钱，因此，我们可能会开始从50个家庭的样本中收集数据。

开始从样本中收集信息时，要考虑的一个重要问题是如何选择样本。询问在当地杂货店购物的家庭与询问在高档餐厅外的人们，收集到的数据可能会非常不同。同样，旧金山的食品支出可能与爱荷华州农村地区的食品支出有很大不同。甚至在我们还没有想到的地方，可能还有一些隐藏的消费习惯。那么，如何确保你的样本能够代表美国家庭，而不会包含来自非典型支出群体的过多家庭呢？

姓名

日期

时期

解决方案是利用*随机性*。我们可以运用随机过程选择 50 个家庭，例如，使用计算机从数据库中随机选择家庭，而不考虑其他因素。这样应该会减少人们在试图获取有关人的信息时可能引入的*偏见*，并且纳入的美国不同类型家庭的比例可能更准确。虽然随机性可能无法完全消除样本选择中的偏差，但与没有随机性的选择相比，它可以显著减少存在的偏差。

研究人员进行了此类研究，发现了每个月在食品上的平均支出。一份报告称，每月的平均食品支出为 600 美元，*误差幅度 (MoE)* 为 150 美元。误差幅度用于表示我们并不预期样本中的每个家庭都花费 600 美元。

留意统计结果中的误差幅度很重要。在不提供误差幅度（描述该值预计会发生多大偏差）的情况下，讨论统计数据是不负责任的。新闻报道中用到的许多图表都会用小字在图表上说明误差幅度。在下次选举期间，留意官员支持率或民意调查图表中的 $\pm 3\%$ 等。这意味着图表中显示的百分比实际上可能比所示数字低或高 3%。

你可以和学生一起尝试这个任务：

一个城镇即将进行投票，决定是否将企业所得税提高 2% 以增加对公立学校的资助。当地新闻展示了一张图片，显示 52% 的投票人口支持增税，角落里显示了“误差幅度 $\pm 3.5\%$ ”。记者听起来对增加企业所得税充满信心，因为只要赞成票超过 50%，法律就会予以通过。

1. 记者驱车前往城镇周围 20 个不同社区中的 4 个，询问居民的意见，得出了 52% 的数字。这样做有什么问题吗？你能想出更好的方法来收集数据吗？
2. 此图片中的误差幅度是什么意思？
3. 你应该对增税充满信心吗？解释你的推理。

解：

1. 只走访镇里的 4 个街区，可能会漏掉记者没有前往的其他街区中许多选民的意见。随机选择镇上的几户家庭，调查他们的意见可能是收集民意的更好方法。随机选择更有可能避免记者对走访哪些社区存在的任何偏见。
2. 误差幅度意味着支持增税的实际百分比可能比样本报告的 52% 高 3.5% 或低 3.5%。这意味着实际百分比将落在 48.5% 至 55.5% 之间。
3. 示例回答：
 - 我认为增税的可能性还是很大的。尽管根据误差幅度，实际百分比可能低至 48.5%，但可能高达 55.5%。大多数可能的百分比都在 50% 以上，所以我认为增税将会发生。

姓名

日期

时期

- 我认为目前还不清楚是否会增税。根据误差幅度，实际百分比可能低至48.5%，这将导致增税不会发生。我也不太清楚这份报告中记者的样本采集方法，所以报告可能不太准确。



© CC BY 2019 by Illustrative Mathematics®