9.1.1 简单随机抽样

一、学习目标

能结合具体的实际问题情境，说明随机抽样的必要性和重要性，学会用简单随机抽样的方法从总体中抽取样本，能用抽签法和随机数法解决简单的实际问题；

在问题解决与探究活动中，感受样本的随机性和统计结果的不确定性，体会用样本估计总体的统计思想；

通过对实际问题的解决，在获得简单随机抽样的经验的同时感受数学的有用性，体验数学来源于生活、应用于生活.

任务一：认识普查与抽样调查，感受抽查的必要性和重要性，感受数据的收集、整理、分析在社会生活中的作用.

1. 2011 年 10 月 5 日，北京市食品安全办通报有 20 种不合格食品被全市停售，其中有 8 种是蛋糕、牛舌饼、桃酥等老人、小孩们常吃的小糕点；

2. 2010 年 7 月 1 日教育部发布全国高校毕业生就业率 72. 2%；

3. 我国沙漠化土地总面积以及那个超过 1. 74×10^5，并以每年 3. 4×10^3 速度扩张.

4. 2011 年 4 月 28 日中华人民共和国统计局发布，普查以 2010 年 11 月 1 日零时为标准进行，公告显示中国总人口截至当时为 1370536875 人.

问题 1：这些数据是怎样得到的呢？普查与抽样调查的优缺点是什么？

全面调查是对每个调查对象进行调查的方法，其中调查对象的全体叫总体，组成总体的每一个调查对象叫个体；根据一定的目的，从总体中抽取一部分个体进行调查，并以此为依据对总体的情况做出估计的调查方法叫抽样调查，总体中抽取出来的那部分个体叫样本，样本中包含的个体数叫样本量.

相对全面调查而言，抽样调查由于只抽取一部分个体进行调查，因此具有花费少、效率高的特点. 在总体规模比较大的调查中，如果经费、时间上受限，那么抽样调查是比较合适的调查方法. 在有些调查中，抽样调查则具有不可替代的作用. 例如，检测一批灯泡的寿命，或一批种子的发芽率，或一批待售袋装牛奶的细菌数是否超标，这些检测具有毁损性，所以只能用抽样调查.

任务二、体会研究样本的抽取方法的必要性和重要性.

问题 2：某批发市场一个小食品店内的一批袋装牛奶要接受细菌是否超标的检查. 如果你是食品卫生监督员，店主主动把柜台上的牛奶拿出来接受检查. 你同意吗？为什么？

问题 3：你还能举出几个生活中的例子吗？

判断一车或一箱水果的质量，判断一袋米中的含沙量等

问题 4：抽取样本时，应该注意些什么？可靠的统计推断结论需要有代表性好的样本作为基础，样本要能够很好的代表总体

问题 5:假设口袋中有红色和白色共 1000 个小球，除颜色外，小球的大小、质地完全相同，你能通过抽样调查的方法估计袋中红球所占比例吗？

见课本 174 页

任务三、学习简单随机抽样的概念和分类

定义:一般地，设一个总体含有 N（N 为正整数）个个体，从中逐个抽取 $n(1\leqslant n\leqslant N)$ 个个体作为样本.如果抽取是放回的，且每次抽取时总体内的个体被抽到的概率都相等，我们把这样的抽样方法叫做放回简单随机抽样.如果抽取是不放回的，且每次抽取时总体内未进入样本的个体被抽到的概率都相等，我们把这种抽样方法叫做不放回简单随机抽样.放回简单随机抽样和不放回简单随机抽样统称为简单随机抽样.通过简单随机抽样获得的样本称为简单随机抽样的样本.

问题 6：根据你的理解，简单随机抽样有哪些特征？

（1） 总体的个数有限；

（2） 样本是逐个抽取的，每次只能抽取一个个体；

（3） 每个个体被抽到的机会相等，抽样具有公平性.

任务四：学习抽签法和随机数法两种简单随机抽样的方法

问题 7： 元旦联欢活动中，要在 30 人的班级里抽出 5 位幸运者，你有什么办法？

问题 8：一家家具厂要为树人中学高一年级制作课桌椅，他们事先想了解全体高一年级学生的平均身高，以便设定可调节课桌椅的标准高度.已知树人中学高一年级有 712 名学生，如果要通过简单随机抽样的方法调查高一年级学生的平均身高，应该怎么抽取样本?

先给 712 名学生编号，例如按 1～712 选行编号，然后把所有编号写在外观、质地等无差别的小纸片(也可以是卡小球等)上作为号签，并将这些小纸片放在一个不透明的盒里，充分搅拌.最后从盒中不放回地逐个抽取号签，使号签上的编号对应的学生进入样本，直到抽足样本所需要的人数.

1.抽签法

第一步：将总体的所有 N 个个体从 0 到 (N-1)编号；

第二步：准备 N 个号签分别标上这些编号，将号签放在容器中搅拌均匀后，每次抽取一个号签；

第三步：将取出的 n 个号签上的号码所对应的 n 个个体作为样本.

优点：*抽签法能够保证每个个体入选的机会都相等(得到的样本是简单随机样本)*

缺点：①当总体中的个数较多时，制作号签的成本会增加，使得抽签法成本（费时、费力）②号签很多时，“搅拌均匀”就非常困难，结果很难保证每个个体入选的机会都相等，从而产生坏样本的可能性增加.

2.随机数法（阅读课本 175-176 页）

先给 712 名学生编号，例如按 1～712 选行编号，y 随机数工具产生 1～712 范围内的正整数，把产生的数作为抽中的编号，使与编号对应的学生进入样本，重复上述过程，直到抽足样本所需要的人数.

（1）用随机试验生成随机数；

（2）用信息技术生成随机数：用计算器生成随机数，用电子表格软件生成随机数，用 R 统计软件生成随机数课本.

262	598	101	127	322	424	305	668	402	415
7	137	512	69	316	546	457	354	75	162
362	126	286	287	119	449	516	325	644	344
356	71	360	122	702	362	651	532	681	271
40	349	338	557	475	702	39	121	72	179

随着信息技术的发展，人们越来越多的利用计算器，数学软件、统计软件等工具来产生随机数，用信息技术工具产生随机数的优点就是方便、快捷.

任务四：学习总体均值和样本均值两个概念，体会二者的联系与区别.

下面是用随机数法从树人中学高一年级学生中抽取的一个容量为 50 的简单随机样本他们的身高变量值(单位:cm)如

156.0 166.0 157.0 155.0 162.0 168.0 173.0 155.0 157.0 160.0

175.0 177.0 158.0 155.0 161.0 158.0 161.5 166.0 174.0 170.0

162.0 155.0 156.0 158.0 183.0 164.0 173.0 155.5 176.0 171.0

164.5 160.0 149.0 172.0 165.0 176.0 176.0 168.5 171.0 169.0

156.0 171.0 151.0 158.0 156.0 165.0 158.0 175.0 165.0 171.0

样本的平均数为 164.3cm，估计树人中学高一年级学生的平均身高为 164.3cm 左右

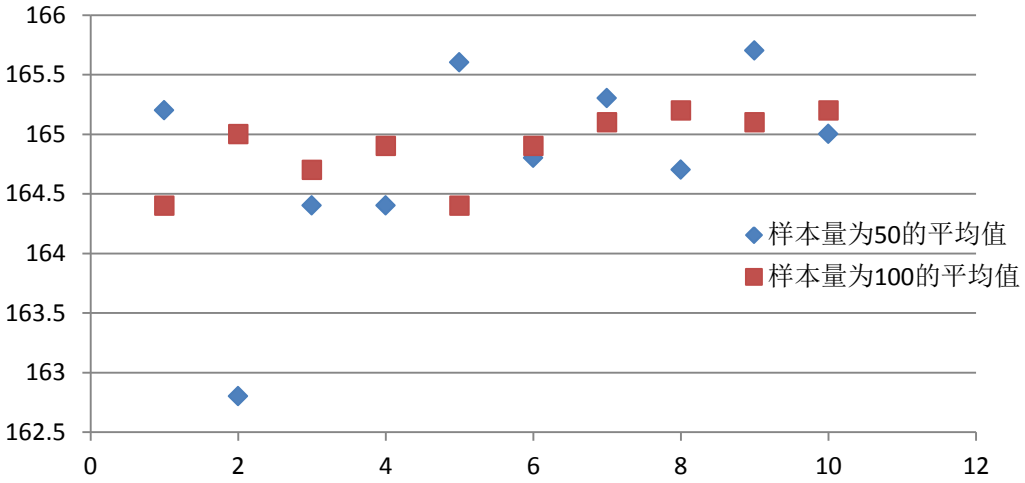
总体均值:$\overline{Y}=\frac{Y_1+Y_2+\cdots+Y_N}{N}=\frac{1}{N}\sum_{i=1}^{N}Y_i$

$$=\frac{1}{N}\sum_{i=i}^{k}f_iY_i(k\le N,f_1+f_2+\cdots f_k=N)$$

样本均值:$\overline{Y}=\frac{Y_1+Y_2+\cdots+Y_n}{n}=\frac{1}{n}\sum_{i=1}^{n}Y_i$

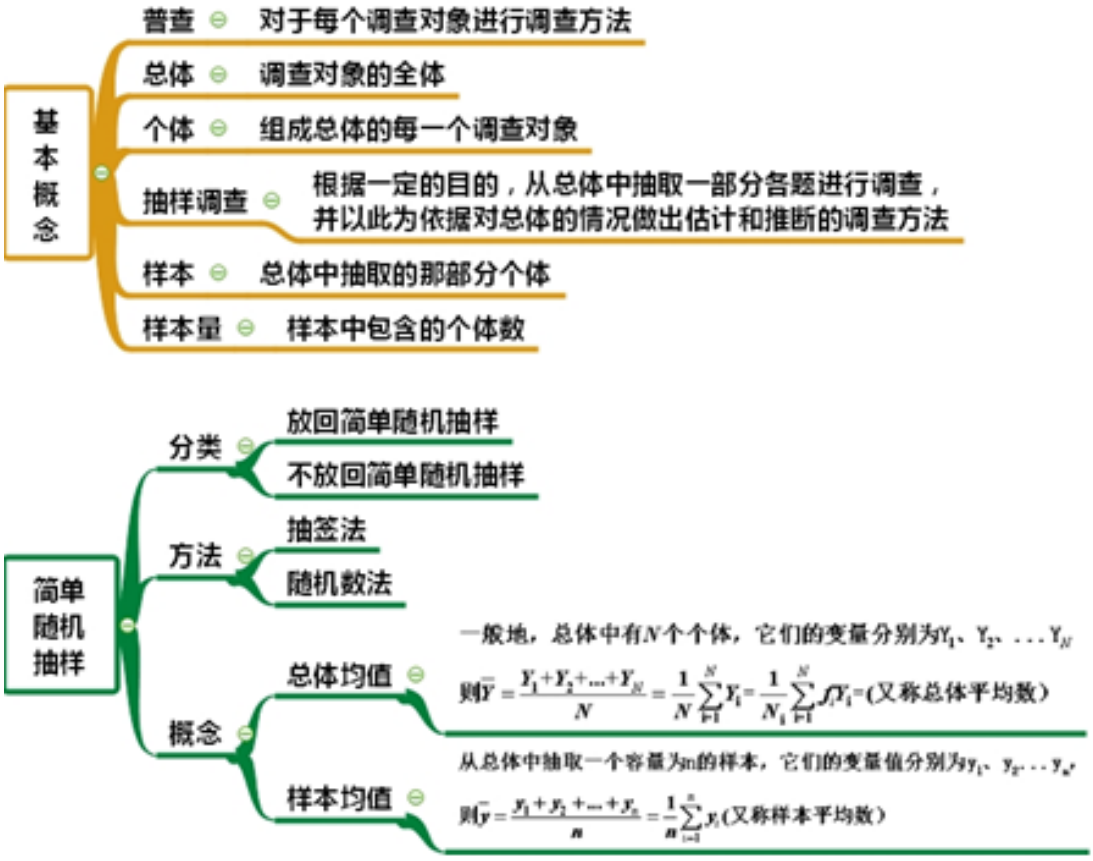
探究：小明想考察一下简单随机抽样的估计效果，他从树人中学医务室得到了高一年级学生身高的所有数据，计算出整个年级学生的平均身高为 165.0cm，然后，小明用简单随机抽样的方法，从这些数据中抽取了样本量为 50 和 100 的样本各 10 个，分别计算出样本平均数，如表 9，1-1 所示. 从小明多次抽样所得的结果中，你有什么发现?

	抽样序号									
	1	2	3	4	5	6	7	8	9	10
样本量为 50 的平均值	165.2	162.8	164.4	164.4	165.6	164.8	165.3	164.7	165.7	165.0
样本量为 100 的平均值	164.4	165.0	164.7	164.9	164.4	164.9	165.1	165.2	165.1	165.2



从试验的结果来看，不管样本量为 50 还是 100，不同样本的平均数往往是不同的. 由于样本的选取是随机的，因此样本平均数也是随机的，这与总体平均数是一个确定的数是不同的. 另一方面，在所有的样本平均数中，除样本为 50 的第 2 个外，样本平均数偏离总体平均数都不超过 cm，即大部分样本平均数都在总体平均数附近波动，这些体现了统计思想的随机性和规律性. 然后，比较样本量为 50 和样本量为 100 的样本平均数，还可以发现样本量为 100 的波动幅度明显小于样本量为 50 的样本平均数，这与我们对增加样本量可以提高估计效果的认识是一致的.

小结：



1．简单随机抽样包括抽签法和随机数法，它们都是等概率抽样，从而保证了抽样的公平性；

2．简单随机抽样有操作简便易行的优点，在总体个数较小的情况下是行之有效的抽样方法；

3．简单随机抽样的局限性：当总体差异明显时不适合用简单随机抽样.