

抽样调查

一、选择题

1. 抽样调查的根本功能是(C)
A. 获取样本资料 B. 计算样本资料
C. 推断总体数量特征 D. 节约费用
2. 概率抽样与非概率抽样的根本区别是(B)
A. 是否能保证总体中每个单位都有完全相同的概率被抽中
B. 是否能保证总体中每个单位都有事先已知或可以计算的 nonzero 概率被抽中
C. 是否能减少调查误差
D. 是否能计算和控制抽样误差
3. 与简单随机抽样进行比较, 样本设计效果系数 $Deff > 1$ 表明(A)
A. 所考虑的抽样设计比简单随机抽样效率低
B. 所考虑的抽样设计比简单随机抽样效率高
C. 所考虑的抽样设计与简单随机抽样效率相同
D. 以上皆对
4. 优良估计量的标准是(B)
A. 无偏性、充分性和一致性 B. 无偏性、一致性和有效性
C. 无误差性、一致性和有效性 D. 无误差性、无偏性和有效性
5. 某乡欲估计今年的小麦总产量进行调查, 已知去年的总产量为12820吨, 全县共123个村, 抽取13个村调查今年的产量, 得到 $\bar{y} = 11863$ 吨, 这些村去年的产量平均为 $\bar{x} = 104.21$ 吨。试采用比率估计方法估计今年该地区小麦总产量(B)
A. 12820.63 B. 14593.96 C. 12817.83 D. 14591.49
6. 抽样标准误差的大小与下列哪个因素无关(C)
A. 样本容量 B. 抽样方式、方法 C. 概率保证程度 D. 估计量
7. 抽样标准误差与抽样极限误差之间的关系是(B)
A. $\Delta = \frac{SE(\hat{\theta})}{\theta}$ B. $\Delta = tSE(\hat{\theta})$ C. $\Delta = \frac{tSE(\hat{\theta})}{\theta}$ D. $\Delta = \frac{SE(\hat{\theta})}{t}$
8. 应用比率估计量能使估计精度有较大改进的前提条件是调查变量与辅助变量之间大致成(A) 关系
A. 正比例 B. 反比例 C. 负相关 D. 以上皆是
9. 能使 $V(\bar{y}_{lr}) = \frac{1-f}{n} (S_Y^2 + \beta^2 S_X^2 - 2\beta S_{YX})$ 达到极小值的 β 值为(B)
A. $\frac{S_Y \cdot S_X}{S_{YX}}$ B. $\frac{S_{YX}}{S_X^2}$ C. $\frac{S_{YX}}{S_Y^2}$ D. $\frac{S_{YX}^2}{S_X}$
10. (B) 是总体里最小的、不可再分的单元。
A. 抽样单元 B. 基本单元 C. 初级单元 D. 次级单元

11. 下面哪种抽样方法是最简单的概率抽样方法 (A)。

A. 简单随机抽样 B. 分层随机抽样 C. 系统抽样 D. 整群抽样

12. 下面关于各种抽样方法的设计效应, 表述错误的是 (B)。

A. 简单随机抽样的 $d_{eff}=1$

B. 分层随机抽样的 $d_{eff}>1$

C. 整群随机抽样的 $d_{eff}>1$

D. 机械随机抽样的 $d_{eff}\approx 1$

13. 假设考虑了有效回答率之外所有其他因素后的初始样本量为400, 而预计有效回答率为80%, 那么样本量则应定为 (B)。
- A. 320 B. 500 C. 400 D. 480
14. 在要求的精度水平下, 不考虑其他因素的影响, 若简单随机抽样所需要的样本量为300, 分层随机抽样的设计效应 $d_{eff}=0.8$, 那么若想达到相同的精度, 分层随机抽样所需要的样本量为 (C)。
- A. 375 B. 540 C. 240 D. 360
15. 分层抽样设计效应满足 (B)。
- A. $d_{eff}=1$ B. $d_{eff}<1$ C. $d_{eff}\approx 1$ D. $d_{eff}>1$
16. 针对总体每一单元都进行信息搜集的调查是 (D)
- A. 抽样调查 B. 典型调查 C. 重点调查 D. 全面调查
17. 调查费是用一个与样本容量有关的函数, 若 C_0 为固定费用, c 为每一个单元的调查费用, 则最简单的线性费用函数为 (D)
- A. $C=C_0+n$ B. $C=C+C_0n$ C. $C=(C_0+C)n$ D. $C=C_0+Cn$
18. 抽样框最直接反映的是 (C)
- A. 目标总体 B. 实际总体 C. 抽样单元 D. 基本单元
19. 在给定费用下使估计量的方差达到最小, 或者对于给定的估计量方差使得总费用达到最小的样本量分配为 (C)
- A. 常数分配 B. 比例分配 C. 最优分配 D. 梯次分配
20. 分层抽样也常被称为 (D)
- A. 整群抽样 B. 系统抽样 C. 组合抽样 D. 类型抽样
21. 整群抽样中群的划分标准为 (A)。
- A. 群的划分尽可能使群间差异小, 群内的差异大
 B. 群的划分尽可能使群间差异大, 群内的差异小
 C. 群的划分尽可能使群间差异大, 群内的差异大
 D. 群的划分尽可能使群间差异小, 群内的差异小
22. 某班级共有六十名学生, 要以直线等距抽样选出15个学生为样本, 调查学生的到课率, 则下列做法正确的是 (D)。
- A. 将60名学生依次编为1~60 B. 计算抽样间距为4
 C. 从1~4随机抽取一个数, 作为抽样的起始单元号, 按每隔4个单元抽取一个, 直至抽出15个样本 D. 以上都正确
23. 初级单元大小不等的多阶段抽样中, 无偏估计量满足自加权的条件是(C)
- A. 第一阶段每个单元被抽中的概率相等
 B. 第二阶段每个单元被抽中的概率相等
 C. 每个基本单元最终被抽中的概率相等
 D. 每个基本单元最终被抽中的概率不等
24. 相对于直线等距抽样, 圆形等距抽样的优点为 (C)
- A. 不用对单元进行编号 B. 随机起点选择范围小
 C. 保证每个单元被抽中的概率严格相等 D. 操作更加简单

25. 某学院共有500名学生，依次编号为1~500,要从中抽取50名学生调查学生的到课率，首先从1~10号中随机抽取一个数，作为抽样的起始单元号，然后每隔10个单元抽取一个，直到抽足50个单元。这种抽样方法是 (C)

A. 简单随机抽样 B. 分层抽样 C. 系统抽样 D. 整群抽样

26. 非概率抽样与概率抽样的主要区别为 (D)

√ 18分层后总体各层的方差是不同的，为了提高估计的精度，通常的做法是在方差较大的层多抽一些样本。

√ 19在不同的层中每个单位的抽样费用可能是不等的。

×20 在分层抽样的条件下，样本容量的确定与简单随机抽样的共同点都是取决于总体的方差。

√21 多主题抽样中，不同的主题对样本量大小的要求不同。在费用允许的情况下，应尽可能地选择较大的样本量。

√22 有时在抽样时无法确定抽样单位分别属于哪一层，只有在抽取样本之后才能区分。

×23 比例分配指的是按各层的单元数占样本单元数的比例进行分配。

×24 等容量分配时各层的样本单元数与各层的层权是相同的。

√25 所谓最优分配是指给定估计量方差的条件下，使总费用最小。

√26 在奈曼分配时，如果某一层单元数较多，内部差异较大，费用比较省，则对这一层的样本量要多分配一些。

√27 在实际工作中如果第k层出现 k_n 超过 kN ，最优分配是对这个层进行100%的抽样。

√28 在实际工作中，如果要给出估计量方差的无偏估计，则每层至少2个样本单元，层数不能超过 $n/2$ 。

×29 无论层的划分与样本量的分配是否合理，分层抽样总是比简单随机抽样的精度要高。

×30 即使层权与实际情况相近，利用事后分层技术也难以达到提高估计精度的目的。

×31. 在任何条件下，估计量的方差都与估计量的均方差相等，因此一般所讲的估计误差也就是指估计量的方差。

×32. 在多阶段抽样中，各阶段只能采用同一种抽样方法。

×33. 总样本量在各层间按内曼分配的结果可以形成自加权的估计量。

×34. 估计抽样误差时，在各种抽样技术条件下都可以用样本方差代替总体方差。

×35. 比估计就是比例估计。

×36. 随机原则就是要使得总体中的每一个抽样单元都有相等的可能性被抽中。

√37. 整群抽样可以被理解为是第二阶段抽样比为100%时的一种特殊的两阶段抽样。

×38. 分层抽样可被理解为是第二阶段抽样比为100%时的一种特殊的两阶段抽样。

×39. 比估计与回归估计都充分利用了有关辅助变量，因此一般情况下都较简单估计的精度要高。

√40. 当第一重样本量等于总体容量时，二重分层抽样与一般分层抽样具有相同的估计精度。

三、名词解释

1. 滚雪球抽样

答：滚雪球抽样是指利用样本点(构成样本的单元)寻找样本点，即由目前的受访者去寻找新的具有新的具有某一特征的受访者。

2. 分别比估计

答：分别比估计是指利用将比率估计的思想和技术用于分层随机样本时，两种可行的办法之一：对每层样本分别考虑比估计量，然后对各层的比估计量进行加权平均，此时所得的估计量称为分别比估计。

3. PPS 抽样

答:

4. 配额抽样

5. 概率抽样

6. 不等概率抽样

7. π PS 抽样的 Brewer 方法

8. 最优分配

9. 比率估计

四、计算题

1、(简单随机抽样的均值、比例估计和样本量的确定)某住宅区调查居民的用水情况, 该区共有 $N=1000$ 户, 调查了 $n=100$ 户, 得 $T=12.5$ 吨, $s^2=1252$, 有40户用水超过了规定的标准。

要求计算:

- ①该住宅区总的用水量及95%的置信区间;
- ②若要求估计的相对误差不超过10%, 应抽多少户作为样本?
- ③以95%的可靠性估计超过用水标准的户数;

2、(内曼分配和按比例分配的均值和比例估计)有下列数据

层	W_h	y_h	S_h	P_h
1	0.35	3.1	2	0.54
2	0.55	3.9	3.3	0.39
3	0.1	7.8	11.3	0.24

设 $n=1000$

①采用按比例分层抽样的方法估计 Y 和 P 并计算其标准误;

②采用奈曼分配的方法估计 Y 和 P 并计算标准误;

3、(两阶段抽样)某市为了了解职工收入情况, 从该市的630个企业中随机抽取了5个企业, 在中选的企业中对职工再进行随机抽样, 有关数据如下:

企业号	M_i	m_i	y_i (元)	S_2
1	520	10	328	400.056
2	108	10	400	301.134
3	1400	20	310	1303.158
4	1200	20	370	1205.786
5	9000	100	420	4200.000

其中, M_i 为企业职工数, m_i 为样本量; y_i 为样本均值, S_2 为样本方差。试估计该市职工平均收入及标准差。

4、（比率估计）某养兔场共有100只兔子，上月末称重一次对每只兔的重量作了纪录，并计算平均重量为3.1磅，一个月后随机抽取10只兔子标重如下：

序 号	1	2	3	4	5	6	7	8	9	10
--------	---	---	---	---	---	---	---	---	---	----

上次	3.2	3	2.9	2.8	2.8	3.1	3	3.2	2.9	2.8
本次	4.1	4	4.1	3.9	3.7	4.1	4.2	4.1	3.9	3.9

- ①估计这批兔子较上月末增重的比率及其标准误差;
- ②估计现有兔子的平均重量及其标准误差;
- ③将比估计方法与均值估计法进行比较,哪一种方法效率高?分析其原因。

5、为调查某5443户城镇居民服装消费情况,采用简单随机不重复抽样调查了36户进行调查,得到平均消费支出 $y=64972$ 元, $s^2=304802$ 试根据此估计:

- (1)该地区居民服装消费支出总额,并给出置信水平为95%的置信区间。
- (2)如果希望服装平均消费支出的相对误差限不超过5%,则样本量至少应为多少?

6、某地区10000名群众,现欲估计在拥有本科学历及以上的群众所占的比例,随机不重复抽取了300名群众进行调查,得到 $p=0.25$, 试估计该地区群众拥有本科以上学历的比例,并以正态分布近似给出其95%的置信区间。

7、对某地区171980户居民家庭收入进行调查,以居民户为抽样单位,根据城镇和乡村将居民划为2层,每层按简单随机抽样抽取300户,经整理得如下数据:

层	N_1	y_n	W_4	Sh
城镇	23560	15180	0.137	2972
乡村	148420	9856	0.863	2546

试根据此估计:

- (1)居民平均收入及其95%的置信区间。
- (2)若是按比例分配和奈曼分配时,各层样本量分别应为多少?

8、某居民小区共有600个单元,每个单元均居住15户,现以单元为群进行整群抽样,随机抽取8个单元,调查每户每周的食品支出费用,调查结果经整理,各单元样本均值和标准差如下表所示:

y_i	205	219	202	218	212	217	208	220
S_i	30.61	32.14	29.62	28.36	25.84	33.59	34.20	26.84

试求:

- (1)该居民小区平均每户每周食品支出费用,并给出其置信水平为95%的置信区间。

(2) 计算以单元为群的群内相关系数与设计效应

9、某县有300个村，小麦播种面积为23434亩。全部村子按地势分为平原和山区两种类型，各按10%的抽样比抽样，调查亩产量，经整理得到下表结果，以大写字母表示总体数据，小写字母表示样本数据， Y 代表调查变量，为今年的总产量， X 代表辅助变量，为去年的总产量，相应的均值为平均亩产量。

类型	N_1	W_4	Y_1	xn	X
平原	102	0.34	583	561	568
山区	198	0.66	290	274	271

试分别对全县今年的平均亩产量构建分别比率估计量和联合比率估计量。

10、一个由 $N=1000$ 个人构成的总体被划分为两层：第一层由 $N_1=400$ 名男性组成，第二层由 $N_2=600$ 名女性组成。从中抽取一个样本量为 $n=250$ 的样本，将样本等比例地分配给各层，使得两层的抽样比都等于 $n/N=1/4$ 。求各层的样本量分别是多少？

11、一公司希望估计某一个月内由于事故引起的工时损失。因工人、技术人员及行政管理人员的事故率不同，因而采用分层抽样。已知下列资料

工人	技术人员	行政管理人员
$N_1=132$	$N_2=92$	$N_3=27$
$s^1=36$	$s_2=25$	$s_3=9$

若样本量 $n=30$ ，试用奈曼分配确定各层的样本量。

12、某工厂生产的新产品供应国内市场的300家用户，试销售满一年后，现欲请用户对该厂的新产品进行评价。现把这些用户分成本地区、本省外地区、外省三层。现有资料如下：

本地区	本省外地区	外省
$N_1=154$	$N_2=93$	$N_3=53$
$S_1^2 = 2.25$	$S_2^2 = 3.24$	$S^3 = 3.24$
$C_1=9$	$C_2=25$	$C_3=36$

若要求估计评价成绩均值的方差 $V(yx)=0.1$ ，并且费用最省(假定费用为线性形式)，求样本量 n 在各层的分配。

13、某林业局欲估计植树面积，该局共辖240个林场，按面积大小分为四层，用等比例抽取40个林场，取得下列资料(单位：公顷)

第一层	第二层	第三层	第四层
$N_1=86$	$N_2=72$	$N_3=52$	$N_4=30$
$n_1=14$	$n_2=12$	$n_3=9$	$n_4=5$

Y _{1i}					Y _{2i}				Y _{3i}		Y _{4i}
97	67	42	125	25	125	155		256			
92	86	27	43	45	47	310		220	142		167
5953	52				352		236		440	256	310
					142	190			495		22054
									396	510	320
									196		780
											0

试估计该林业局总的植树面积及95%的置信区间。

14. 一个县内所有农场按规模大小分层，各层内平均每个年农场谷物(玉米)的英亩数列在下表中。

农场规模(英亩)	农场数 N	平均每一农场的玉米面积 Y	标准差 S,
0-40	394	5.4	8.3
41-80	461	16.3	13.3
81-120	391	24.3	15.1
121-160	334	34.5	19.8
161-200	169	42.1	24.5
201-240	113	50.1	26.0
>241	148	63.8	35.2
总和或均值	2010	26.3	—

现要抽出一个包含100个农场的样本，目的是估计该县平均每个农场的玉米面积，请问：

(1) 按比例分配时，各层的样本量为多少？

(2) 按最优分配时，各层的样本量为多少？（假定各层的单位调查费用相等）

15. 某县欲调查某种农作物的产量，由于平原、丘陵和山区的产量有差别，故拟划分为平原、丘陵和山区三层采用分层抽样。平原区共有150个村庄，丘陵区共有100个村庄，山区共有250个村庄。按照各种地形等比例各抽取5%样本，进行实割实测产量，结果计算如下表。

(1) 在95.45%的概率保证程度下，试估计该县农作物平均每村产量的区间范围。

(2) 若村庄的农作物产量低于150吨，县政府并将其归为低产量村，从而对其加强农业补贴政策。因此，试图在95.45%的概率保证程度下估计该县低产量村比例的区间范围。

地形	村庄总个数	样本村个数	样本平均产量(吨)	样本产量标准差	样本低产量村个数
平原	140	7	202.5	82.2C	2

丘陵	100	5	147	36.84	1
山区	240	12	121.11	47.35	7

16. 某城市居民小区的食品消费量调查，以每个楼层为群进行整群抽样，每个楼层都有8个住户。用简单随机抽样在 $N=510$ 个楼层中抽取 $n=12$ 个楼层，得

到96个样本户人均食品消费额，及按楼层的平均数和标准差，如表。请估计该小区人均食品消费额的户平均值，并给出95%置信区间。

17. 邮局欲估计每个家庭的平均订报份数，该辖区共有4000户，划分为400个群，每群10户，现随机抽取4个群，取得资料如下表所示：

群	各户订报数 y	y_i
1	1, 2, 1, 3, 3, 2, 1, 4, 1, 1	19
2	1, 3, 2, 2, 3, 1, 4, 1, 1, 2	20
3	2, 1, 1, 1, 1, 3, 2, 1, 3, 1	16
4	1, 1, 3, 2, 1, 5, 1, 2, 3, 1	20

试估计平均每户家庭订报份数及总的订报份数，以及估计量的方差。

18. 某工业系统准备实行一项改革措施。该系统共有87个单位，现采用整群抽样，用简单随机抽样抽取15个单位做样本，征求入选单位中每个工人对政策改革措施的意见，结果如下：

单位	总人数	赞成人数
1	51	42
2	62	53
3	49	40
4	73	45
5	101	63
6	48	31
7	65	38
8	49	30
9	73	54
10	61	45

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：
<https://d.book118.com/516135102034010113>