附件2

2021年北京市居民心理健康素养基线

调查抽样方法

北京市居民心理健康素养调查采用多阶段随机抽样方法，在北京市16个区中，每个区抽取5个街道（乡镇），每个街道（乡镇）抽取3个居委会（村），每个居委会（村）抽取50户，每户随机抽取1名18岁以上常住人口作为调查对象。在实际施测时，每个居委会（村）在随机抽取的50人中，完成40份有效调查为止。具体来说，各阶段抽样内容、抽样方法详见下表（表1）。

**表1 各阶段抽样方法**

|  |  |  |
| --- | --- | --- |
| **抽样阶段** | **抽样内容** | **抽样方法** |
| 第一阶段 | 每个区随机抽取5个街道（乡镇） | PPS抽样 |
| 第二阶段 | 每个街道（乡镇）随机抽取3个居委会（村） | PPS抽样 |
| 第三阶段 | 每个居委会（村）随机抽取50个备选家庭户 | 等距抽样 |
| 第四阶段 | 抽取家庭户内1名18岁以上常住人口参与调查 | Kish表抽样 |

**一、抽取街道**

以区为单位，使用PPS抽样方法在每个区随机抽取5个街道（乡镇）。以下介绍什么是PPS抽样，如何进行PPS抽样。

**1. 什么是PPS抽样**

PPS抽样法（Probability Proportionate to Size Sampling）又称按规模大小成比例的概率抽样或按容量比例概率抽样法。它是一种不等概率抽样，每个单元在每次抽选中被抽中的概率与该单元所在的单元规模的大小成比例。PPS抽样的特点是需要使用辅助信息。本调查中主要依据人口规模信息，采用常住人口数。

**2. PPS抽样步骤**

假设某区有10个街道（乡镇），下表列出了每个街道（乡镇）的常住人口数。使用PPS抽样方法抽取5个街道（乡镇），具体的抽样过程为：

第一步，统计所有街道（乡镇）的总人口数，在下表（表2）的例子中，总人口数为142400。

第二步，用总人口数除以抽取的街道（乡镇）数，计算出抽样的间距：142400/5=28480。

第三步，在1到间距28480的范围内随机选取一个整数，例如8800，作为起始的抽样编码。

第四步，在起始编码上累加间距，计算出剩余4个抽样编码。即：8800+28480=37280，37280+28480=65760，65760+28480=94240，94240+28480=122720。这样，一共获得了对应要抽取的5个街道（乡镇）的5个抽样编码，分别是8800，37280，65760，94240，122720。

第五步，在下表（表2）里，找出分别包含这5个抽样编码的街道（乡镇），进入样本。可见，代码为2，5，6，7，9的5个街道（乡镇）被抽中。

**表2街道（乡镇）的PPS抽样示例**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **街道（乡镇）代码** | **常住人口数** | **累计人口数** | **人口数范围** | **抽取的编码** |
| 12345678910 | 11601816083608840123003944012260146801028016920 | 11601932027680365204882088260100520115200125480142400 | 1-11601161-1932019321-2768027681-3652036521-4882048821-8826088261-100520100521-115200115201-125480125481-142400 | 8800372806576094240122720 |

**二、抽取居委会（村）**

以抽中的街道（乡镇）为单位，使用PPS抽样方法在每个街道（乡镇）随机抽取3个居委会（村）。具体的抽样过程为：

第一步，在抽取居委会（村）之前，需将家庭户数在750户以下的居委会（村）与相邻的居委会（村）进行合并，直到所有抽样单位家庭户总数均在750～1500户之间。如果乡镇（街道）所辖居委会（村）人口规模较小，多个居委会（村）合并仍达不到750户的，可将抽样单位户数降低至500户左右。下表（表3）是某街道居委会（村）户数合并过程示例，供参考。

**表3 某街道居委会（村）户数合并示例**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **村（居）名** | **户数** | **合并后总户数** |
| 1 | \*\*店 | 510 | 860 |
| \*\*营 | 350 |
| 2 | \*\*村 | 239 | 773 |
| \*\*村 | 204 |
| \*\*店 | 330 |
| 3 | \*\*村 | 420 | 759 |
| \*\*村 | 200 |
| \*\*村 | 139 |

第二步，完成合并后，使用PPS抽样方法，随机抽取3个居委会（村）。PPS抽样步骤参考上述各区抽取街道的方法。

**三、抽取家庭户**

以抽中的居委会（村）为单位，使用等距抽样方法，在每个居委会（村）内随机抽取50户。以下介绍什么是等距抽样，如何进行等距抽样。

**1. 什么是等距抽样**

等距抽样（Systematic Sampling）是指先将总体的全部单元按照一定顺序排列，采用简单随机抽样抽取第一个样本单元（或称为随机起点），再顺序抽取其余样本单元的抽样方法。

**2. 等距抽样步骤**

以抽中的居委会（村）为单位，在每个居委会（村）随机抽取50个家庭户。若抽中的居委会（村）是由多个居委会（村）合并而成，合并后的居委会（村）一共抽取50个家庭户。具体的抽样过程为：

第一步，获取居委会（村）内的家庭户基本信息，在Excel表中制作家庭户基本信息列表。家庭户基本信息可以是门牌号，也可以是户主姓名，能与实际的家庭户对应上即可。

第二步，随机选取一个家庭户作为起点，再按照一定间隔依次抽取其余的家庭户。间隔大小，取决于居委会（村）所辖全部家庭户的数量，间隔=所辖全部家庭户的数量/抽样的家庭户数量。例如，某居委会辖区内共有775户家庭，需要抽取50户备选家庭户参与调查，那么抽样间隔为15（775/50=15.5，取整数）。下表（表4）中，以随机选取1号楼1单元401为起点，间隔15被抽中的家庭户为1号楼2单元502。

**表4 家庭户等距抽样示例**



**四、抽取个体**

获取被抽中50户的家庭成员信息后，通过使用Kish表抽取家庭户内1名18岁以上常住人口作为调查对象。以下介绍什么是Kish表，如何进行Kish表抽样。

**1. 什么是Kish表**

Kish 表是统计学家Leslie Kish针对入户抽样设计的。目前Kish 表已经广泛地应用在世界各国的入户抽样调查中。Kish 表的优点是在理论上坚持随机抽样，而且经过巧妙设计，确保每一位适合的调查候选对象的入选概率。

研究者先将调查表分为（编号为）A、B1、B2、C、D、E1、E2、F八种，同时，印制若干套（一套八种）“选择卡”发给调查员，每人一套。调查员首先对年满18岁的所有家庭成员排序，然后，调查员按照调查表上的编号找出编号相同的那种选择表，根据家庭人口数目从“选择表”中查出该选个体的序号，对应进行访谈。

**2. Kish表抽样步骤**

第一步，根据抽取的家庭户名单，把50个家庭户按顺序填入家庭户代码分配表中（如表6）。每个家庭户对应一个Kish表代码，Kish表的代码分为A、B1、B2、C、D、E1、E2、F八种，50个家庭户每种代码的分配数量如表5所示。

**表5 Kish表八种代码分配数量**

|  |  |  |  |
| --- | --- | --- | --- |
| **Kish表代码** | **数量** | **Kish表代码** | **数量** |
| A | 9 | D | 9 |
| B1 | 4 | E1 | 4 |
| B2 | 4 | E2 | 4 |
| C | 8 | F | 8 |

**表6 家庭户代码分配表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **家庭户编码** | **户主姓名** | **联系电话****（手机号）** | **家庭住址** | **Kish表代码** |
| 12000001 | 赵甲 | \*\*\*\*\*\*\*\* | \*\*\*\*\*\*\*\*\*\*\*\*楼\*\*\*单元\*\*\*室 | F |
| 12000002 | 苏大强 | \*\*\*\*\*\*\*\* | \*\*\*\*\*\*\*\*\*\*\*\*楼\*\*\*单元\*\*\*室 | B1 |
|  |  |  |  | A |
| …… | …… |  | …… | …… |

第二步，填写家庭成员登记表，了解调查户家庭构成，对每户家庭中的所有常住人口进行排序，排序的方法是男性在前，女性在后；在每一个性别内，年龄大的在前，年龄小的在后。例如“苏大强”家庭构成排序（如表7）。

**表7 家庭成员登记表（未编号）**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 姓名 | 性别 | 年龄 | 是否≥18岁 | 家庭成员编号 |
| 苏大强 | 男 | 39 | 是 |  |
| 苏明哲 | 男 | 25 | 是 |  |
| 苏明成 | 男 | 20 | 是 |  |
| 赵美兰 | 女 | 38 | 是 |  |
| 苏明玉 | 女 | 15 | 否 |  |

第三步，对已经排序的家庭成员登记表上年满18岁的家庭成员从1开始编号（如表8）。

**表8 家庭成员登记表（已编号）**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 姓名 | 性别 | 年龄 | 是否≥18岁 | 家庭成员编号 |
| 苏大强 | 男 | 39 | 是 | 1 |
| 苏明哲 | 男 | 25 | 是 | 2 |
| 苏明成 | 男 | 20 | 是 | 3 |
| 赵美兰 | 女 | 38 | 是 | 4 |
| 苏明玉 | 女 | 15 | 否 | 无 |

第四步，根据第一步中调查户的Kish表代码，找出相应的Kish表。例如，苏大强家预先分配的Kish表代码为B1，找出对应的表B1如下：

|  |
| --- |
| **表B1** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6或者以上 | 2 |

第五步，根据年满18岁的家庭成员人数，在Kish表上找出相应的成员编号。例如，苏大强家年满18岁的有4个人，在B1表上找出的家庭成员编号是1，也就是说，应选择家庭成员编号为1的苏大强作为调查对象。

|  |
| --- |
| **表B1** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| **4** | **1** |
| 5 | 2 |
| 6或者以上 | 2 |

附录1 PPS抽样表

附录2等距抽样表

附录3 家庭户代码分配表

附录4 家庭成员登记表

附录5 Kish表

**附录1**

**PPS抽样表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **街道（乡镇）** | **常住人口数** | **累计人口数** | **人口数范围** | **抽取的编码** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**附录2**

**等距抽样表（以居委会为例）**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **居委会名称** | **小区名称** | **楼号** | **单元** | **户** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**附录3**

**家庭户代码分配表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **家庭户编码** | **户主姓名** | **联系电话****（手机号）** | **家庭住址** | **Kish表代码** |
|  |  |  |  | F |
|  |  |  |  | B1 |
|  |  |  |  | A |
|  |  |  |  | D |
|  |  |  |  | E1 |
|  |  |  |  | B2 |
|  |  |  |  | C |
|  |  |  |  | E2 |
|  |  |  |  | A |
|  |  |  |  | D |
|  |  |  |  | C |
|  |  |  |  | A |
|  |  |  |  | F |
|  |  |  |  | D |
|  |  |  |  | F |
|  |  |  |  | B1 |
|  |  |  |  | A |
|  |  |  |  | D |
|  |  |  |  | E1 |
|  |  |  |  | B2 |
|  |  |  |  | C |
|  |  |  |  | E2 |
|  |  |  |  | C |
|  |  |  |  | A |
|  |  |  |  | F |
|  |  |  |  | D |
|  |  |  |  | F |
|  |  |  |  | B1 |
|  |  |  |  | A |
|  |  |  |  | D |
|  |  |  |  | E1 |
|  |  |  |  | B2 |
|  |  |  |  | C |
|  |  |  |  | E2 |
|  |  |  |  | C |
|  |  |  |  | A |
|  |  |  |  | F |
|  |  |  |  | D |
|  |  |  |  | F |
|  |  |  |  | B1 |
|  |  |  |  | A |
|  |  |  |  | D |
|  |  |  |  | E1 |
|  |  |  |  | B2 |
|  |  |  |  | C |
|  |  |  |  | E2 |
|  |  |  |  | C |
|  |  |  |  | A |
|  |  |  |  | F |
|  |  |  |  | D |

**附录4**

**家庭成员登记表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **姓名** | **性别** | **年龄** | **是否≥18岁** | **家庭成员编号** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**附录5**

**Kish表**

|  |
| --- |
| **表A** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6或者以上 | 1 |

|  |
| --- |
| **表B1** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6或者以上 | 2 |

|  |
| --- |
| **表B2** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 2 |
| 5 | 2 |
| 6或者以上 | 2 |

|  |
| --- |
| **表C** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 1 |
| 3 | 2 |
| 4 | 2 |
| 5 | 3 |
| 6或者以上 | 3 |

|  |
| --- |
| **表D** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 2 |
| 3 | 2 |
| 4 | 3 |
| 5 | 4 |
| 6或者以上 | 4 |

|  |
| --- |
| **表E1** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 3 |
| 5 | 3 |
| 6或者以上 | 5 |

|  |
| --- |
| **表E2** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6或者以上 | 5 |

|  |
| --- |
| **表F** |
| ≥18岁常住居民人数 | 家庭成员编号 |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6或者以上 | 6 |