

# 农作物产量抽样调查

(电视录象教学片文字稿本)

王广森 晏 森

02

中国统计出版社

电视录象教学片文字稿本  
**农 产 量 抽 样 调 查**  
王广森 晏 森

•  
中国统计出版社出版发行  
西安市第二印刷厂印刷

•  
787×1092毫米 32开 3.625印张 75千字  
1987年3月第1版 1987年3月西安第1次印刷  
印数：1—15,000  
统一书号：4006·142 定价：0.90元



## 前 言

为了满足全国农村抽样调查网点培训基层工作人员的需要，我们根据国家统计局1984年制订的《农村抽样调查网点抽选方案》（试行）摄制了《农作物产量抽样调查》电视录象教学片。内容包括农产量抽样调查的基本理论和抽样方法、田间测产、推算产量等实际操作过程。现将该片的文字稿本整理出来，供教学中参考。由于我们缺乏编写录象教材的经验，加之时间仓促，不妥之处，敬请读者批评指正。

编 者

1987年1月

## 目 录

|                                  |       |
|----------------------------------|-------|
| 第一集 抽样调查的基本理论.....               | (1)   |
| 一、什么是抽样调查.....                   | (1)   |
| 二、什么是随机原则.....                   | (2)   |
| 三、随机抽样的样本平均数分布规律.....            | (12)  |
| 四、概率统计中关于抽样估计的基本公式.....          | (14)  |
| 五、抽样效率.....                      | (20)  |
| 第二集 农作物产量调查方法.....               | (35)  |
| 一、正确选定农作物产量调查点.....              | (35)  |
| 二、用科学的方法搞好农作物的田间测产.....          | (54)  |
| 三、正确运用调查点的资料进行科学推算<br>和综合分析..... | (94)  |
| 附：实测调查结果综合表.....                 | (108) |

## 第一集 抽样调查的基本理论

本片的主要对象为农村抽样调查网点从事农作物产量抽样调查的实际工作人员，为了使这些同志既能正确地贯彻实施国家统计局制订的《农村抽样调查网点抽选方案(试行)》，又能比较全面、深入地理解方案中各项规定的理论依据，本片第一集系统地介绍抽样调查的主要基本概念和基础理论。

### 一、什么是抽样调查

抽样调查是统计调查的一种方法，它的特点是从所要了解的某类现象中抽取一部分现象作为实际调查的对象，通过对这一部分现象调查的结果来对全部对象的某些特征作出判断。为了说明的方便，我们把所要了解的某类现象称之为总体；把抽取出来进行实际调查的这一部分现象称之为样本。通过样本的实际调查结果来对总体的某些特征作出推断是我们组织抽样调查的目的。

统计所调查的某类现象总是由许多单个的对象组成的。例如，乡、镇企业是由许多个企业单位组成的，农村户口是由许多个农户组成的等等，抽样调查就是通过对一部分乡镇企业单位的调查来了解乡镇企业总体的情况，通过对一部分

农户的调查来了解农户总体的情况。换句话说，抽样调查就是对总体中的一部分单位所作的实际调查。我国农村有1.8亿农村住户，8亿农民，要了解农村的情况，就要广泛采用抽样调查的方法。国务院1983年8月批转的国家统计局《关于加强农村统计工作等问题的报告》就提出要求加强农村抽样调查工作，国家统计局于1984年7月制订了《农村抽样调查网点抽选方案》发给各省（市）自治区农村抽样调查队作为试行方案执行，我们制作这部录象片的目的是为了帮助大家了解抽样调查的原理，更好地执行方案中的各项规定，提高抽样调查的质量。

## 二、什么是随机原则

随机原则是指在抽选调查单位时，要排除主观因素的干扰，保证总体中的每个单位都有一定的中选机会。由于排除了主观因素的干扰，才能使抽选出来的样本反映客观实际，避免人为因素造成的偏差；由于每个单位都有一定的中选机会，才能运用数学概率论的方法，计算出产生不同大小抽样误差的可能性。所以说随机原则是有严格的科学性的。《农村抽样调查网点抽选方案》中规定：农村抽样调查网点的抽选，必须坚持随机原则，按照科学的方法抽选调查单位。我们要想了解抽样调查的基本原理，首先要对随机原则有一个正确的理解。

随机原则有两个主要特点：

（1）随机原则的特点之一，排除主观因素的干扰。

抽样调查在抽选调查单位时，应当严格按照事先规定的

程序抽选，不受抽选者个人意图或偏好的影响。

随机原则的上述特点在抽样调查实践中的重要性是容易理解的，因为抽样调查的结果能否反映调查对象的实际情况，关键在于抽选出来的调查单位所构成的样本，对总体是否具有代表性，如果在抽选调查单位时没有一个事先规定的客观程序，而是任凭抽选者个人的意图来决定，就可能出现人为的偏差，保证不了抽选出来的样本的代表性。这里需要强调的是，按照随机原则来抽选样本可以保证样本具有代表性，是指随机样本的全体而言的，并不是说按随机原则抽取的每个样本都有同样的代表性。这一点在我们明瞭了随机原则的第二个特点以后，就可以得到进一步的理解。

（2）随机原则的特点之二，每个调查单位都有一定的中选机会。

这个特点是指在按照事先规定的程序抽选样本时，每个调查单位被选中的机会是一定的，也就是说每个调查单位有一定的被选中的概率，这样才能应用数学方法，对每种样本出现的概率进行计算。随机原则的这个特点是抽样调查能够作出科学推断的依据。

为什么在抽选调查单位时必须坚持随机原则，才能对抽样调查的结果作出科学的推断呢？这个问题最好通过大家一起做个试验来说明。

国家统计局1984年7月发布的《农村抽样调查网点抽选方案（试行）》给出了某省分县的统计资料如下：

表1·1 某省分县资料 (按粮食亩产由低到高顺序排队)

| 县名<br>代号 | 粮食亩产<br>(斤) | 播种面积<br>(万亩) | 播种面积累计<br>(万亩) | 人均收入<br>(元) | 参加分配人口<br>(万人) |
|----------|-------------|--------------|----------------|-------------|----------------|
|          | $X_1$       | $F_1$        | $\Sigma f_i$   | $X_2$       | $F_2$          |
| 1        | 158         | 16.9         | 16.9           | 76          | 4.60           |
| 2        | 207         | 42.9         | 59.8           | 48          | 16.64          |
| 3        | 227         | 136.3        | 196.1          | 50          | 63.56          |
| 4        | 228         | 61.3         | 257.4          | 58          | 82.12          |
| 5        | 228         | 71.2         | 328.6          | 84          | 25.40          |
| 6        | 231         | 46.2         | 374.8          | 66          | 17.31          |
| 7        | 234         | 116.8        | 491.6          | 63          | 41.28          |
| 8        | 239         | 129.5        | 621.1          | 57          | 57.80          |
| 9        | 241         | 40.7         | 661.8          | 75          | 15.74          |
| 10       | 243         | 83.1         | 744.9          | 72          | 31.34          |
| 11       | 247         | 75.7         | 820.6          | 61          | 27.96          |
| 12       | 250         | 100.7        | 921.3          | 64          | 40.40          |
| 13       | 250         | 102.5        | 1023.8         | 68          | 42.32          |
| 14       | 259         | 55.7         | 1079.5         | 55          | 27.18          |
| 15       | 265         | 101.7        | 1181.2         | 71          | 40.95          |
| 16       | 271         | 85.1         | 1266.3         | 62          | 37.27          |
| 17       | 283         | 61.5         | 1327.8         | 60          | 33.46          |
| 18       | 286         | 48.3         | 1376.1         | 62          | 20.85          |
| 19       | 292         | 72.2         | 1448.3         | 70          | 36.44          |
| 20       | 298         | 10.4         | 1552.3         | 71          | 49.95          |
| 21       | 303         | 98.7         | 1651.0         | 64          | 41.93          |
| 22       | 306         | 2.6          | 1653.6         | 174         | 5.32           |
| 23       | 318         | 164.1        | 1817.7         | 60          | 79.46          |
| 24       | 322         | 97.2         | 1914.9         | 64          | 52.34          |



续表

| 县名<br>代号 | 粮食亩产<br>(斤) | 播种面积<br>(万亩) | 播种面积累计<br>(万亩) | 人均收入<br>(元) | 参加分配人口<br>(万人) |
|----------|-------------|--------------|----------------|-------------|----------------|
|          | $X_1$       | $F_1$        | $\Sigma f_1$   | $X_2$       | $F_2$          |
| 25       | 328         | 82.1         | 1997.0         | 83          | 43.61          |
| 26       | 330         | 63.4         | 2060.4         | 72          | 32.11          |
| 27       | 331         | 218.9        | 2279.3         | 103         | 122.08         |
| 28       | 343         | 242.7        | 2522.0         | 88          | 104.81         |
| 29       | 346         | 13           | 2535.0         | 111         | 11.69          |
| 30       | 347         | 1            | 2536.0         | 98          | 3.88           |
| 31       | 350         | 135.6        | 2671.6         | 114         | 39.66          |
| 32       | 352         | 242          | 2913.6         | 86          | 91.96          |
| 33       | 355         | 202.2        | 3115.8         | 89          | 77.89          |
| 34       | 359         | 116.8        | 3232.6         | 103         | 47.00          |
| 35       | 364         | 119.3        | 3351.9         | 72          | 56.01          |
| 36       | 367         | 123.4        | 3475.3         | 72          | 73.11          |

资料来源：《农村抽样调查网点抽选方案》（试行），第13、14页，国家统计局1984年7月

这个表的第一栏，是各县的代号，从1到36，一共是36个县。

第二栏，是各县的平均粮食亩产水平，以斤为单位，这一栏数字用 $X_1$ 表示。

第三栏，是各县的粮食播种面积（万亩），用 $F_1$ 来表示。

第四栏，是各县的粮食播种面积累计数，用 $\Sigma f_1$ 来表示。

第五栏，是各县农户的人均年纯收入（元），用 $X_2$ 来表示。

第六栏，是各县参加分配的人口（万人），用 $F_2$ 来表示。

为了使我们的试验尽量简化，我们只利用表上第一、二两栏的资料，即各县的代号和各县的粮食平均亩产（斤），不考虑各县粮食播种面积的差别。这样一来，我们的调查对象就是由代号1到36的36个县的粮食平均亩产量所构成的总体，每个县的粮食平均亩产就是一个总体单元。表上的36个总体单元是按粮食亩产水平由低到高的顺序排列的。这个有序的总体单元序列就是我们抽取样本的根据，在统计学上把它叫做抽样框。

抽样调查的工作程序就是①从抽样框中抽取样本，②对抽取的样本进行实际调查，③根据样本调查的结果对总体作出推断。

上表就是我们进行抽样试验的抽样框。

现在让我们根据这个抽样框进行一次不重置的随机抽样的试验。

要从这个抽样框中按随机原则抽取样本，首先要规定样本的容量和实现随机原则的具体抽样方法。

这里我们规定样本容量为11，即抽取11个县的粮食亩产作为一个样本。

实现随机原则的抽样方法，过去一般采用“抽签法”或“随机数查表法”，现在在电子计算机上利用统计软件包可以根据使用者的要求产生随机数组，这种方法最为简便，我们这次抽样试验就是用在计算机上产生随机数组的方法。

假定在电子计算机上完成产生随机数组的程序操作后，屏幕上显示所产生的第一个随机数组为：

12, 11, 17, 34, 24, 2, 20, 16, 29, 7, 3（由于是不重置的随机抽样，在一个随机数组中不含相同的数）。

屏幕上显示的这第一个随机数组就是我们按随机原则选出的第一个容量为11的随机样本，这11个数分别代表11个选中的县，它们的粮食平均亩产就被用来代表全省的粮食亩产水平。

现在让我们看看这个随机样本的粮食平均亩产同全省的粮食平均亩产相差有多大。为了简便起见，在我们的抽样试验中，计算平均亩产都采用简单算术平均数的方法，不用粮食播种面积加权，因为我们的目的是通过试验来说明抽样调查的基本原理，不论计算平均数用的是简单算术平均法还是加权算术平均法，在说明抽样调查的基本原理方面的作用是一样的。

第一个随机样本

| 代号 | 粮食平均亩产（斤） |
|----|-----------|
| 12 | 250       |
| 11 | 247       |
| 17 | 283       |
| 34 | 359       |
| 24 | 322       |
| 2  | 207       |
| 20 | 298       |
| 16 | 271       |
| 29 | 346       |
| 7  | 234       |
| 3  | 227       |

$$\text{样本平均数} = \frac{250 + 247 + 283 + \cdots + 227}{11} = 276.7 \text{ 斤/亩}$$

根据全省36个县的资料用简单算术平均法计算得到的总体平均数 =  $\frac{158 + 207 + 227 + \cdots + 367}{36} = 287.7 \text{ 斤/亩}$

样本平均数和总体平均数相差11斤/亩，相当于总体平均数的3.82%。

我们按照随机原则抽取了11个县，用它们来代表全省36个县的粮食平均亩产，从计算结果来看，这11个县的粮食平均亩产的简单算术平均数是276.7斤，而全省36个县的粮食平均亩产（简单算术平均）是287.7斤，相差11斤或相差3.82%，应该说，用这个随机样本来推断全省的粮食亩产，误差是不大的。

但是，按照随机原则从一定的总体中能够抽取的样本是很多的，在我们的例子中，总体是36个县，从中抽取11个县构成一个样本，总共可以抽取多少个不同的样本呢？用代数中从n个数中抽r个数的组合公式 $C_r^n$ 可以很容易地计算出来：

$${}_{36}C_{11} = \frac{36!}{11!(36-11)!} = 600,805,296 \text{ 个组合，即可以}$$

抽得600,805,296个不同的样本，我们所抽得的随机样本只不过是6亿多个中的一个而已。现在让我们再抽一个样本看看。

第二个随机样本：

随机数组： 9    20    4    10    19    34    12    8    6

X<sub>1</sub>        241 298 228 243 292 359 250 239 231

随机数组：17 12

$X_1$  283 250

$$\text{样本平均数} = \frac{241 + 298 + \cdots + 250}{11} = 264.9 \text{ 斤/亩}$$

总体平均数为287.7斤/亩

样本平均数与总体平均数相差22.8斤/亩，相当于总体平均数的7.92%。

这一次我们仍然用电子计算机产生随机数组的方法，抽取11个县的粮食亩产构成一个随机样本。根据这个随机样本计算的样本平均数为264.9斤/亩，用它来作为总体的代表值，和实际的总体平均数（287.7斤/亩）相差22.8斤/亩，相当于总体平均数的7.92%。同前一次抽取的随机样本比较起来，这一次抽样的误差大了一倍。可见随机抽样所得的样本，一次跟一次不同，有的误差小些，有的误差大些。了解到这一点是很重要的，因为我们在实际工作中总是只能抽取一个样本做调查，然后就根据这一个样本来对总体作出推论，而这一个调查样本不过是许许多多可能样本中的一个而已。

为了对随机抽样的这个特点有一个更深入、更全面的认识，我们按照前面抽取随机样本的程序，继续进行随机抽样的试验，连同前面已经抽取的两个样本，一共抽取50个样本，各个样本的代号、计算出来的样本平均粮食亩产、样本平均亩产与总体平均亩产相差的绝对数如下表（见表1·2）

表1·2

| 样本号数 | 样本平均粮食亩产<br>(斤) | 样本平均粮食亩产与总体平均<br>粮食亩产的差数 (斤) |
|------|-----------------|------------------------------|
| 1    | 276.7           | 11                           |
| 2    | 264.9           | 22.8                         |
| 3    | 269.4           | 18.3                         |
| 4    | 257.9           | 29.8                         |
| 5    | 318.4           | —30.7                        |
| 6    | 275.3           | 12.4                         |
| 7    | 292.5           | — 4.8                        |
| 8    | 298.9           | —11.2                        |
| 9    | 272.6           | 15.1                         |
| 10   | 298.7           | —11.0                        |
| 11   | 276.4           | 11.3                         |
| 12   | 275.1           | 12.6                         |
| 13   | 268.0           | 19.7                         |
| 14   | 277.1           | 10.6                         |
| 15   | 280.6           | 7.1                          |
| 16   | 289.9           | — 2.2                        |
| 17   | 276.6           | 11.1                         |
| 18   | 265.3           | 22.4                         |
| 19   | 269.6           | 18.1                         |
| 20   | 302.4           | —14.7                        |
| 21   | 288.6           | — 0.9                        |
| 22   | 294.7           | — 7.0                        |
| 23   | 277.7           | 10                           |
| 24   | 296.4           | — 8.7                        |
| 25   | 271.6           | 16.1                         |

续表

| 样本号数 | 样本平均粮食亩产<br>(斤) | 样本平均粮食亩产与总体平均<br>粮食亩产的差数 (斤) |
|------|-----------------|------------------------------|
| 26   | 302.3           | —14.6                        |
| 27   | 310.8           | —23.1                        |
| 28   | 278.6           | 9.1                          |
| 29   | 310.5           | —22.8                        |
| 30   | 292.7           | 5.0                          |
| 31   | 300.8           | —13.1                        |
| 32   | 315.3           | —27.6                        |
| 33   | 300.6           | —12.9                        |
| 34   | 288.9           | — 1.2                        |
| 35   | 325.7           | —38.0                        |
| 36   | 302.4           | —14.7                        |
| 37   | 290.8           | — 3.1                        |
| 38   | 277.9           | 9.8                          |
| 39   | 304.1           | —16.4                        |
| 40   | 293.1           | — 5.4                        |
| 41   | 294.4           | — 6.7                        |
| 42   | 272.0           | 15.7                         |
| 43   | 290.1           | — 2.4                        |
| 44   | 280.5           | 7.2                          |
| 45   | 284.2           | 3.5                          |
| 46   | 279.7           | 8.0                          |
| 47   | 252.0           | 35.7                         |
| 48   | 289.9           | — 2.2                        |
| 49   | 279.7           | — 2.0                        |
| 50   | 262.5           | 25.2                         |

从上面所抽取的50个随机样本来看，每个样本的具体组成都不一样，因而它们的平均粮食亩产也各不一样。我们把这50个样本的平均粮食亩产制成一个频数分布表，就可以清楚地看出，绝大部分样本平均数在250斤到300斤之间，包括我们最初抽取的那两个样本（见表1·3）。换句话说，尽管随机抽样所抽取的样本一次和一次不同，但是样本平均数的分布却是比较集中的，这就是我们前面所说，从总体上看，随机抽样是有代表性的。随机抽样的理论就是要说明产生这种现象的原因和条件，并对随机样本的分布规律作出定量的描述，这样我们就可以运用随机抽样的理论知识来改进抽样调查工作并对抽样调查的结果作出有科学根据的推断。

表1·3      五十个随机样本平均数的频数分布

| 组        别<br>(斤/亩) | 样本平均数的分布 |           |
|---------------------|----------|-----------|
|                     | 次    数   | 频   率 (%) |
| 150—199.99          | 0        | 0         |
| 200—249.99          | 0        | 0         |
| 250—299.99          | 39       | 78        |
| 300—349.99          | 11       | 22        |
| 350—399.99          | 0        | 0         |
| 合        计          | 50       | 100       |

三、随机抽样的样本平均数分布规律

粗略地说：按照随机原则从一个总体中抽取的许多样本，若总体是正态的，或样本足够大，其平均数集中在总体平



均数的近旁两侧呈对称分布。

随机抽样样本平均数的这种分布规律可以从我们所做的抽样试验结果中加以观察和验证。

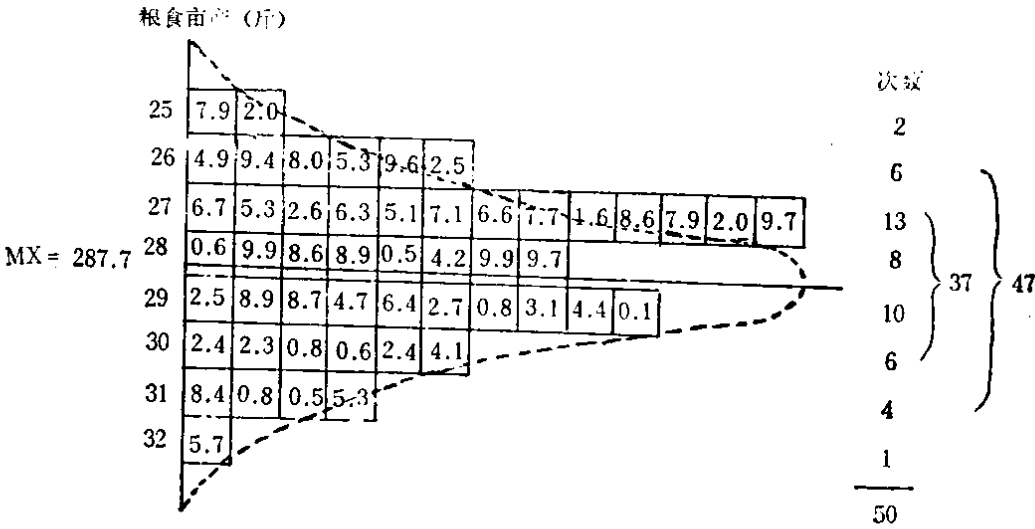


图 1—1 50个随机样本平均数的次数分布

注：座标线左侧为粮食亩产的百位数和十位数，右侧为各样本平均粮食亩产的个位数及其小数，如257.9斤、252.0斤……等等。

图 1—1 形象地表现出随机样本平均数的分布规律：样本平均数落在270—309.9斤之间的有37个，落在260—319.9斤之间的有47个，只有三个样本的平均数值小于260斤或大于319.9斤，可见样本平均值的分布是比较集中的。如果以总体平均数287.7斤为标准，那么，样本平均数小于287.7斤的有24个，大于287.7斤的有26个，基本上是对称的。

上面这张图给了我们一个关于样本平均数分布规律的直观概念，但是要对抽样推断作出具体的估计，还必需运用概率统计的知识。这里我们只举出概率统计中关于抽样估计所用到的几个基本公式，不对它们作进一步的解释和证明。