抽樣技術

第一節 抽樣在市場調查中的作用

一、市場普查

（一）市場普查的意義

市場調查所研究對象的全體稱為調查總體，它是市場調查所研究的各個體單位的總和。

市場普查是以市場調查總體中的第一個個體單位作為研究對象的。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

（二）市場普查的局限性

從理論上講，市場普查取得的資料最為全面、最為準確、最有價值。因為，市場普查應該成為市場調查最常用的形式。但實際上只有在總體較小或有某些特殊要求的場合，市場調查才採用普查的形式。這是因為市場普查存在不少局限性。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

（1）市場普查要耗費大量的人力、物力、財力，絕大多數企業和研究機構要經常普遍地採用這種方法是難以承擔的。

（2）市場普查需要花費較長的時間，這和市場調查的時效性往往發生矛盾。

（3）市場普查要受到地理、氣候、政治、經濟、心理等一系列自然的、社會的可變因素的影響。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

（三）市場普查的困難

市場普查還往往在可操作性方面碰到困難。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

二、抽樣調查

（一）抽樣調查的意義

抽樣調查是指從總體中抽取能代表總體的一部分，即樣本，進行調查，然後根據樣本中所包含的資訊對總體的狀況進行估計和推算的一種市場調查方法。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

（二）抽樣調查的優越性

（1）抽樣調查由於只對總體中的部分個體逐一進行調查，所需費用顯然大大減少，為一般企業所能夠承受。

（2）抽樣調查由於樣本容量較小，調查所需的時間也較短，這就能夠在不長的時間內完成資料收集及數據統計工作，保證調查的時效性。

第五章 抽樣技術
第一節 抽樣在市場調查中的作用

（3）抽樣調查由於樣本容量較小，調查工作量較少，在調查實施過程中只需較少的訪問人員。這就便於精心組織，可對訪問員進行更好的訓練，更好地控制調查過程，以確保調查品質。

（4）抽樣調查也易於增加調查深度。在調查時可以對調查對象進行更全面更學入的瞭解和研究，可以獲得更有價值的資訊。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

（三）抽樣調查的基本概念

1、總體和樣本

總體指的是市場調查所研究對象的全體。

樣本由從總體中抽取的部分個體所組成。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

2、總體指標

　總體指標就是調查的目標量，也就是總體的有關參數。這些總體參數在抽樣調查中是可以通過有關樣本指標來估計的。

3、抽樣單元和抽樣框

　抽樣單元是指對總體進行劃分後得到的每一部分。在抽樣時，必須掌握所有抽樣單元的有關資料，如名單、地圖等等，這稱為抽樣框。

第五章　抽樣技術
第一節　抽樣在市場調查中的作用

4、抽樣誤差和非抽樣誤差

抽樣誤差是指用樣本估計總體產生的誤差。抽樣誤差一般用估計量的均方差或方差來表示。

非抽樣誤差是指在抽樣調查中由於人為差錯所造成的誤差。這類誤差應採取一定的管理措施加以控制。

第五章　抽樣技術
第二節　　隨機抽樣技術

一、簡單隨機抽樣

（一）簡單隨機抽樣的方法

1、簡單隨機抽樣的方法

簡單隨機抽樣是指從含有N個抽樣單元的總體中，一次抽取n個單元，使全部可能的種不同的結果，每種被抽到的概率都等於1/的一種抽樣方法。這時所得到的樣本稱為簡單隨機樣本。

第五章　抽樣技術

第二節　　隨機抽樣技術

從總體中逐個有放回地抽取單元，使在每次抽取一個單元之前，總體都恢復到原來的狀態，使每一次抽取與上一次抽取都是相互獨立的，那麼這種抽樣稱為非常簡單隨機抽樣，這是一種有重複的抽樣。這時的樣本稱為非常簡單隨機樣本。

當總體的單元數N非常大，所抽取的樣本單元數n相對又比較小，有放回和無放回的柚樣幾乎是一樣的，否則兩者還是存在差異。

第五章　抽樣技術
第二節　隨機抽樣技術

2、簡單隨機抽樣的抽籤法

抽籤法可以如下操作：先把總體中的每個單元編上不同的號碼，寫在簽上。將簽充分攪拌均勻，從中任意抽取一個號碼，將對應的單元選入樣本，直到抽足預先規定的樣本數目n為止。實際操作時也可以同時抽取幾個簽。

如果抽取一個簽後，記下號碼後又放回總體，再隨機抽取下一個，這樣每次抽取都是相互獨立的，這時得到的將是非常簡單隨機樣本，其中可能出現同樣的樣本單元。

第五章　抽樣技術
第二節　　隨機抽樣技術

3、簡單隨機抽樣的亂數表法

亂數表又稱亂數表，是指含有整齊排列的亂數字的表格。亂數表法可以如下操作：先將總體中的每個單元編上不同的號碼，根據最大編號的位數，確定使用若十位數字，然後查亂數表，先在亂數表中任意選定一個數字作為開始數字，隨後向任何一個方向連續摘取數字，得到一系列規定位數的數，凡編號範圍內的數對應的單元即被入樣。如果不是重複抽樣，碰上重複的數應舍去，直到抽足預定的樣本數目n為止。

第五章　抽樣技術
第二節　　隨機抽樣技術

例5-1

要從80戶居民中抽取10戶，先將80戶居民由01到80（或從00至79）編號，然後把亂數表的任意一行一列的一個數字為起始數，假定本例以第3行第5列的數字4為起始數，再按任意方向起讀，如本例由左向右起讀。即得:47、33、84、51、67、47、97、19、98、40、07、17、66、23，其中，47、33、51、67、19、40、07、17、66、23即為抽中的單元編號。

第五章　抽樣技術
第二節　　隨機抽樣技術

4、簡單隨機抽樣的電腦隨機程式法

簡單隨機抽樣還可以方便地應用電腦亂數程式。起動這一程式，在電腦螢幕上將連續出現亂數。由此可以確定入樣單元的編號。

第五章　抽樣技術
第二節　　隨機抽樣技術

（二）實例----地區居民資料庫的建立

這裏以上海地區為例，建立區、街道、居委會的三級地區居民資料庫。

上海市市區有：（1）黃浦　（2）盧灣　（3）靜安　（4）楊浦　（5）虹口　（6）閘北　（9）普陀　（8）長寧　（9）徐匯　（10）寶山　（11）浦東新區　（12）金山　（13）青浦　（14）南匯　（15）嘉定　（16）閔行　（17）奉賢　（18）松江等18個區。

以普陀區為例有：（1）長壽　（2）曹楊　（3）長風　（4）長征鎮　（5）甘泉　（6）石泉　（7）真如鎮　（8）宜川　（9）桃浦鎮等9個街道。

第五章 抽樣技術
第二節 隨機抽樣技術

以長風街道為例有：（1）長風一村 （2）長風二村一委 （3）長風二村二委 （4）長鳳三村 （5）長風四村一委 （6）長風四村二委 （7）師大一村 （8）師大二村 （9）師大三村 （10）曹家巷 （11）中山橋 （12）新渡口 （13）海鑫 （14）大渡河路95弄 （15）白玉新村一委 （16）白玉新村二委 （17）金沙新村 （18）光復西路1091弄 （19）曹家村 （20）普陀二村 （21）普陀四村一委 （22）普陀四村二委 （23）錦綠新城 （24）中江等24個居委會。

第五章　抽樣技術
第二節　　隨機抽樣技術

（三）總體均值和比例的估計及抽樣誤差的估算

1、總體均值和比例的估計

採用簡單隨機抽樣。可以用樣本的特徵值為估計總體的目標值，並且可以對抽樣誤差進行估算。

（1）總體均值的估計。設總體X含 有N個單元，總體均值為。在簡單隨機抽樣中，樣本含量為n，則可用樣本平均值

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad （5\text{-}1）$$

作為總體均值的估計量。

第五章　抽樣技術
第二節　　隨機抽樣技術

（2）總體比例的估計。設總體X含有N個單元，其中有個單元具有某種特性，那麼總體中含有該種特徵的單元的比例

$$\pi = \frac{\phi}{N} \qquad (5\text{-}2)$$

在簡單隨機抽樣中，樣本含量為n，其中具有該種特性的單元數為f，則可以用

$$P = \frac{f}{n} \qquad (5\text{-}3)$$

作為總體比例的估計量。

第五章　抽樣技術
第二節　　隨機抽樣技術

例5-2

從某區50家超市中隨機抽取10家超市，他們的日銷售量分別為51、67、45、37、81、49、72、65、43、75萬元。

（1）試求該區超市平均日銷售量的估計值；

（2）試求該區銷售量低於60萬元的超市比例的估計值。

第五章　抽樣技術
第二節　　隨機抽樣技術

解：（1）樣本的平均日銷售量X的平均值

$$\overline{X}=\frac{1}{10}\sum_{i=1}^{10}X_i=\frac{1}{10}(51+67+\cdots+75)=58.5$$

因此，可用X=58.5（萬元）作為該區超市平均日銷售量的估計值。

（2）樣本中日銷售量低於60萬元的超市有5家，因此，樣本比例

$$P=\frac{f}{n}=\frac{5}{10}=50\%$$

因此，可用P=50%作為日銷售量低於50萬元的超市比例的估計值。

第五章　抽樣技術
第二節　　隨機抽樣技術

2、抽樣誤差的估算

（1）均值的抽樣誤差。在簡單抽樣情況下，均值的抽樣誤差有如下的計算公式：

$$\mu_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \bullet \frac{S}{\sqrt{n}} \qquad (5\text{-}4)$$

其中：$\mu_{\bar{x}}$表示均值的抽樣誤差，S表示樣本標準差，N表示總體單元數，n表示樣本單元數。

第五章 抽樣技術
第二節 隨機抽樣技術

在非常簡單隨機抽樣情況下，均值的抽樣誤差有如下的計算公式：

$$\mu_{\bar{x}} = \frac{S}{\sqrt{n}} \qquad （5\text{-}5）$$

第五章　抽樣技術
第二節　　隨機抽樣技術

例5-3

　　某飲料公司進行一次居民戶平均飲料消費量的抽樣調查。在總體10萬戶居民家庭中，抽選樣本2000戶。已知，樣本標準差為2.5升，試求在簡單隨機抽樣和非常簡單隨機抽樣條件下的抽樣誤差。

解：N=100000戶） n = 2000（戶），S = 2.5升

設：$\mu \bar{x}$ 為抽樣誤差

第五章　抽樣技術
第二節　　隨機抽樣技術

①在簡單隨機抽樣條件下

$$\mu_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \bullet \frac{S}{\sqrt{n}} = \sqrt{\frac{100000-2000}{100000-1}} \times \frac{2.5}{\sqrt{2000}} = 0.99 \times \frac{2.5}{44.72} = 0.0553$$

②在非常簡單隨機抽樣條件下

$$\mu_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{2.5}{\sqrt{2000}} = 0.0559$$

第五章　抽樣技術
第二節　　隨機抽樣技術

（2）比例的抽樣誤差。比例的抽樣誤差和均值的抽樣誤差有類似的公式，不同之處是用 $\sqrt{P(1-P)}$ 代替S，其中P是樣本比例。

在簡單隨機抽樣情況下，比例的抽樣誤差有如下的計算公式：

$$\mu p=\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{P(1-P)}{n}} \qquad （5\text{-}6）$$

其中：　μp 表示比例抽樣抽差，P表示樣本比例，N表示總體單元數，n表示樣本單元數。

第五章 抽樣技術
第二節　隨機抽樣技術

在非常簡單隨機抽樣情況下，比例的抽樣誤差有如下的計算公式：

$$\mu p = \sqrt{\frac{P(1-P)}{n}} \qquad (5\text{-}7)$$

例5-4

從總體10萬戶居民中抽取2000戶調查，發現飲用果汁飲料的戶數為450戶，求其抽樣誤差。

解：N＝100 000（戶）　　n＝2 000（戶）

$$P = \frac{450}{2000} = 0.225$$

第五章　抽樣技術
第二節　　隨機抽樣技術

設：μp 為抽樣誤差，則：

在簡單隨機抽樣情況下：

$$\mu p=\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{P(1-P)}{n}}=\sqrt{\frac{100000-2000}{100000-1}}\sqrt{\frac{0.225\times 0.775}{2000}}=0.99\times 0.0093=0.0092$$

在非常簡單隨機抽樣情況下：

$$\mu p=\sqrt{\frac{P(1-P)}{n}}=\sqrt{\frac{0.225\times 0.775}{2000}}=0.0093$$

第五章　抽樣技術
第二節　隨機抽樣技術

（3）抽樣調查的容許誤差。

在數理統計中已經證明，在簡單隨機抽樣情況下，總體均值 μ 落在下列區間中的可能性為 $1-\alpha$。

$$\mu = \overline{X} \pm t_{a/2}(n-1)\sqrt{\frac{N-n}{N-1}} \bullet \frac{S}{\sqrt{n}} \qquad \text{（5-8）}$$

其中：μ 為總體均值，$\overline{X}$ 為樣本均值，$t_{a/2}(n-1)$ 為 t 統計量值，α 為置信度，N為總體單元數，n為樣本單元數。

第五章 抽樣技術
第二節 隨機抽樣技術

這一區間也稱為總體均值的置信度為 $1-\alpha$的置信區間。從另一角度理解式（5-8），可以認為對於一次隨機抽樣所得的樣本均值$\overline{X}$，總體均值 μ與 $\overline{X}$ 的實際差異的絕對值有的$1-\alpha$ 可能性小於

$$t_{a/2}(n-1)\sqrt{\frac{N-n}{N-1}}\bullet\frac{S}{\sqrt{n}} \qquad （5\text{-}9）$$

為樣本平均值$\overline{X}$在置信度$1-\alpha$下的最大容許絕對誤差，常常簡稱為最大容許誤差。因此，

$r\overline{x}=\Delta\overline{x}/\overline{X}$ 為最大容許相對誤差。（5-10）

第五章　抽樣技術
第二節　　隨機抽樣技術

對於總體比例的估計，在簡單隨機抽樣情況下也有類似的結果。

總體比例π的置信度為 $1-\alpha$的置信區間為：

$$\pi = P \pm t_{a/2}(n-1)\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{P(1-P)}{n}} \qquad （5\text{-}11）$$

其中：π 為總體比例，P為樣本比例，$t_{a/2}(n-1)$ 為t統計量值，a為置信水準，n為總體單元數，N為樣本單元數。

因此，$\Delta p = t_{a-2}(n-1)\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{P(1-P)}{n}} \qquad (5\text{-}12)$

第五章　抽樣技術
第二節　　隨機抽樣技術

為樣本比例P在置信度 $1-\alpha$下的最大容許絕對誤差，簡稱為最大容許誤差，因此，

$$rp = \Delta p / P \qquad （5\text{-}13）$$

為最大容許相對誤差。

對於非常簡單隨機抽樣，有完全類似的公式，

只是 $\sqrt{\dfrac{N-n}{n-1}}$ 取為1。

在一般市場調查中，n均較大這時 $t_{a/2}(n-1)$可用 $Z_{a/2}$代替。