



一、二元資料（二項分配）樣本數

二元分配資料的母群很大時，分配的性質很接近標準化的常態分配，因此可以使用數理方式，推求出各種抽樣出入、抽樣把握、及母群變異數下所需要的樣本數。

二元分配資料通常都是用來推求某種現象的百分比，這2類資料的百分比分別為 p 和 q ，而 $p+q=100\%$ ，同時，其變異數與 $p*q$ 成正比。所以，當 p 和 q 的樣本統計量愈趨近50%時， $p*q$ 會愈大，根據抽樣結論下判斷所需要的樣本數愈大，反之愈趨近0或100%時，樣本數可以較少。

譬如張三、李四兩人競選臺北縣長，如果雙方掌握的選民均近50%，即雙方愈「相峙不下」、選民意見愈紛歧、母群變異數大，那麼預測誰會當選的抽樣調查，所需要的樣本就要大；如果選情呈「一面倒」，彼此「相去懸殊」，即選民的意見很一致，那麼不需要很大的樣本，也可以預測勝負了。

二項分配的樣本數決策公式如：

公式5.1 二項分配的樣本數

$$n = pq \left(\frac{Z}{d} \right)^2$$

d : 抽樣出入

Z : 常態分配的標準差。由於常用的抽樣把握(β)為.95, .99，它們分別對應的 Z 值則為：1.96和2.575。

這個公式其實非常好計算，但還是有許多學者將使用二元資料時，如何決定樣本數的數據列表，以供快速查詢使用(Backstrom [88], Monroe [248], Parten [267])。



表格5-1 二元分配的樣本數

把握	0.95						0.99					
比例	10	20	30	40	45	50	10	20	30	40	45	50
出入	90	80	70	60	55	50	90	80	70	60	55	50
±0.5%	13,829	24,585	32,268	36,878	38,016	38,415	23,886	42,463	55,733	63,695	65,340	66,349
1	3,457	6,146	8,067	9,220	9,504	9,604	5,971	10,616	13,933	15,924	16,335	16,587
2	864	1,537	2,017	2,305	2,376	2,401	1,493	2,654	3,483	3,981	4,084	4,147
3	384	683	896	1,024	1,056	1,067	633	1,180	1,548	1,769	1,185	1,843
4	216	384	504	576	594	600	373	664	871	995	1,021	1,037
5	138	246	323	369	380	384	239	425	557	637	653	663
6	96	171	224	256	264	267	166	295	387	442	454	461
7	71	125	165	188	194	196	122	217	284	325	333	339
8	54	96	126	144	148	150	93	166	218	249	255	259
9	48	76	100	114	117	119	74	131	172	197	202	205
10	35	61	81	92	95	96	60	106	139	159	163	166

來源：Parten (1950)

這種表有兩種使用的方法：¹

a. 先行使用法

即在研究經費、人力許可下，研究者先行預定研究需要幾成把握？可以容忍多大的出入？然後決定樣本數。譬如某電子商務公司計畫向既有客戶推出一項新產品，決策者認為需要九成五的把握，允許3%的出入，而根據過去的經驗，客戶對新產品的接受度在三成以上才有生產價值。因此，查表的方法是：

(a) 先找到「把握=.95」的表；

(b) 找到「出入=±3.0」的那一列；

(c) 在「比例」一欄中，找到「30-70」那一行（「30-70」表示調查問題中二元分配的可能比例，在此即為會接受新產品的客戶，至少達到30%的比例）。

(d) 「出入」列和「比例」行的交叉點，就是應有的樣本數，查出來是八百九十六個。

同理，研究者如果希望把握提高到九成九，那麼就換一個表查，決定樣本是一千五百四十八個。

b. 回溯使用法

在和上例相同的調查中，如果研究者的人力、物力，頂多只能抽查五百個樣本，那麼可以就這個樣本數先調查，回頭再來檢查這個抽樣有幾成把握，可能有多大出入。假如樣本反應出來，有32%的客戶會接受新產品，那麼解釋的方法是：

(注¹) 國內也有別的书本曾經輾轉引用Parten的表，或許因為忽略失察，無論解釋如何使用這個表，以及表中數據，均有嚴重錯誤，需要特別注意。



(a) 在「把握 = .95」的表中，找到比例為「30-70」那一行，查出接近於樣本數「500」的數據「504」。

(b) 這個數據所對應的「出入」是「 $\pm 4\%$ 」亦即在九成五的把握下，會接受新產品的客戶的約在28%至36%之間。

同理，也可以推求九成五把握的情形。

值得特別注意的是，一般文獻中所提供的「抽樣數據」均為二元分配的資料而設，並不適用於多元資料，與後續將談到的「連續性資料」。