

第 7 章 整體差異試驗(Overall Difference Tests)：
樣品之間的感官差異存在否？

- 1. 前言
- 2. 差異和相似性測試的統一方法 (The Unified Approach to Difference and Similarity Testing)
- 3. 針對差異的三角試驗(Triangle Test for Difference)
- 4. 二-三點試驗(Duo-Trio Test)
- 5. 五取二試驗(Two-out-of-Five Test)
- 6. 同/異試驗(Same/Different Test) 或單純差異試驗(Simple Difference Test)
- 7. A-非 A 試驗 (“A” – Not A” Test)
- 8. 與控制組的差異試驗(Difference-from-Control Test)
- 9. 連續式試驗(Sequential Tests)

- Overall difference tests 整體差異試驗 (第 6 章)：樣品之間感官差異是否存在？(Does a sensory difference exist between samples?)
譬如三角(Triangle)及二-三點(Duo-trio)試驗是設計應用於受試者能否檢測出樣品之間的任何差異。
- Attribute difference tests 屬性差異試驗 (第 7 章)：樣品之間屬性 X 不同？(How does attribute X differ between samples?) 要求受試者只聚焦單項或幾項屬性，譬如根據甜度來排序某些樣品，忽略甜味以外的所有屬性。
包括配對比較試驗(paired comparison tests)、n-替代迫選試驗(n-AFC; Alternative Forced Choice)及各種的多重比較試驗(multiple comparison tests)。

7.1 前言

第 7 及 8 章採用料理書方式介紹各種的差異試驗，並舉例。
基本的理論參照第 5 章 (Measuring Responses)與 13 章 (Basic Statistical Methods)。如何選擇一個特定試驗，其指引請參照每種方法中的“範圍與應用”項，以及第 15 章的摘要式整理，供感官技術上選擇的指南。

差異試驗可以合理地建構成數百種方法，但實用上，本章提出的程序步驟都是常用的。

差異試驗分為兩類，其特點為：

- 所選擇屬性的感知強度，採用第 5 章內的任何方法測定，例如：順位(ranking)、線尺度(line scaling)或大小/量值推估(magnitude estimation, ME)。
- 2- 及 3-AFC 試驗常用於閾值(threshold)的測定，情感型試驗(affective tests) (嗜好/偏好/喜好試驗 preference tests，譬如消費者試驗 consumer tests)也是屬性差異試驗(第 13 章)。

7.2 差異和相似性測試的統一方法

辨識(別)試驗(discrimination tests)用於解決各種實際的目標。

- 某些情況下，研究者興趣於證明兩樣品之間有可感知的差異。
- 其它情況，研究者要確定如果兩樣品十分相似就可互相取代使用。
- 另一情況，有些研究者要證明差異但參與同研究的其他研究者又要證明相似性。

所有這些情況都可透過選擇適當的試驗-靈敏度參數(test-sensitivity parameters) (α 、 β 及 Pd 值)與統一的方法處理。

而 α 、 β 及 Pd 值要如何才是適當的，則取決於該試驗的特定目標。

一致性方法也適用於配對比較試驗(paired comparison tests)，譬如 2-AFC (參照 8.2 項)。

測試相似性(similarity)時，感官分析專家要測定兩樣品是十分地相似而互相替代使用，譬如為降低成本之配方修改及驗證備用的供應商即是常見的兩種範例。

- 針對相似性之試驗設計，專家要測定什麼構成有意義的差異性，選擇一個 Pd 值，然後指定值小的 β -危險率，以確保誤失了差異真的存在只有很小的機率。
- 為使評定者人數落在合理的界限內，允許 α -危險率的值變大。

針對差異之測試，其目的僅要發現兩樣品之間有否存在可感知的差異？

■ 在這隱藏假設(tacit assumption)下，要統計分析的僅有 α -危險率事項(α -risk matters)【 α -危險率：結論出可感知的差異存在，但實際上並不存在之機率】。評估者人數由檢視 α -risk Table 決定，同時並考慮相關的因素，譬如評估者可用度(availability; 即人數)、測試樣品的可用量(即樣品種類)等。

■ 忽略 β -危險率(β -risk)【 β -危險率：結論並無可感知的差異存在，但實際上有差異存在之機率】及品評小組的辨識者比率(the proportion of distinguishers, Pd)，或甚至假定是不重要的。

■ 其結果，在差異測試，研究者為 α -危險率選取小的值，接受任意大值的 β -危險率和 Pd (忽略之)，讓評定者所需的人數落在合理的界限內。

然而有許多情況，重要的是要在誤失差異存在之危險率(β -危險率)和將不存在的差異結論為存在之危險率(α -危險率)之間取得平衡。這樣，專家要為全部三個參數選擇設定值，以達到該試驗要求的靈敏度所需要的評估者人數(參照範例 7.4)。

經驗法則上，統計學顯著的結果(statistically significant result)， **α -危險率**為：

- 10~5% (0.10~0.05)表示中度的証據(moderate evidence)差異是明顯的；
- 5~1% (0.05~0.01)表示強力的証據(strong evidence)差異是明顯的；
- 1~0.1% (0.01~0.001)表示很強的証據(very strong evidence)差異是明顯的；
- < 0.1% (<0.001)表示非常強的証據(extremely strong evidence)差異是明顯的；

對於 **β -危險率**，差異為不明顯的證據強度也採用同上的標準，但**“是不明顯的”** 取代 **“是明顯的”**。

辨識者的最大允許比率(maximum allowable proportion; **Pd**) 分三個範圍：

- **Pd < 25%** 表示小的值(**small** values);
- **25% < Pd < 35%** 表示中等值(**medium-sized** values)
- **Pd > 35%**表示大的值(**large** values)

微軟 Excel 軟體中有建立一個空白表格程式的應用，幫助研究者選擇 α 、 β 及 Pd 值，在所需的試驗靈敏度和可用的資源之間提供最好的妥協(見 14.3.5 項)。這 **Test Sensitivity Analyzer** 允許研究者快速運算由評估者人數 n 、正確回應數 x 和辨識者的最大允許比率 Pd 等不同組合的各種的方案，並且在每種情況下都觀察到對 **α -危險率**和 **β -危險率**的影響。

7.3 三角試驗(Triangle Test)

三角試驗(ASTM 2004; ISO 2004c) 是本書介紹的第一個方法，有些的複雜性，包括許多細節，即(1) 所有感官分析專家都應知道；(2) 也是許多方法共同的，及(3) 故後面介紹的方法中將不再重複列出。統一方法的應用在**範例 7.3** 及 **7.4** 中介紹。

7.3.1 範圍與應用

當試驗目標是要**確定兩產品之間是否存在感官差異**，採用本方法。本法特別適用於**處理效應(treatment effects)** 可能產生產品的變化，但不能只以單一或二項屬性表徵等的情况。

雖然統計學上，較之**配對比較(paired comparison)**和**二-三點試驗**更有效，**三角試驗**使用上有些限制，如產品會有感官疲勞(sensory fatigue)、遲滯(carryover)或順應(adaptation)，以及讓受試者覺得測試三個樣品過於困惑。在某些情況下此方法是有效的：

1. 確定產品差異是否來自組成分、加工、包裝、或貯藏的變化
2. 確定是否整體差異存在，並無特定的屬性被認定受影響
3. 挑選和監控品評員其辨識所給予差異性之能力

7.3.2 試驗原理

1. 呈現給每位受試者 3 個編碼的樣品，
2. 告訴他們其中的 2 個為相同的而另一個是不同的(或單個 odd)，
3. 要求受試者由左至右品嚐 (感覺、檢查)每個樣品，然後選出單個樣品，
4. 計算正確回答的數目，參照表 19.8，解釋顯著性存在有否。

7.3.3 試驗受試者

通常三角試驗採用 20~40 位受試者(subjects)，當差異大而容易挑出時，少至 12 位也可用；另一方面，相似性測試(similarity testing)則需 50 至 100 位受試者。最低限，受試者要熟悉三角試驗(格式、工作、品評步驟)，以及要測試的樣品，特別因為風味記憶(flavor memory)在三角測試中扮有一席之地。

實際品嚐試驗前，建議辦說明會(orientation session)，讓受試者熟悉試驗程序步驟和產品特性。須注意要給予充分的資訊，有啟發性及激勵的，同時不要讓受試者對處理效應和產品認定之特定資訊上產生偏差。

7.3.4 試驗程序

試驗控制(test controls；參見第 3 章的細節)應包括一個分隔的試驗區域，讓每位受試者可以獨立作業而不受干擾。為減少任何的顏色變數，需要照明控制；為所探討的產品類型，在最佳條件下準備和呈現樣品，例如樣品應該是開胃可口且好好擺置。

同時提供樣品，可以的話；若樣品的數量多，留下餘味，或顯出外觀上有些差異，也可以依次先後提供，這不會折損試驗效力。

準備數量相同的六種可能組合(ABB、BAA、AAB、BBA、ABA 及 BAB)，隨機呈現給受試者。

要求受試者從左至右的順序檢查樣品（味道，感覺，氣味等），測試進行中，可以回頭去重新品評每個樣品。

 7.1 所示的評分卷可用於超過一組樣品，但是，只用於感官疲勞很少之時。選出奇數樣品(odd sample)之後，不要另外再詢問有關喜好、接受性、差異程度、或差異的類型等的問題，這是由於受試者對奇數樣品的選擇，會左右他/她對這些問題的回應，要回應此類的問題可透過額外的試驗取得。有關差異的大小或類型(屬性)，參見第 13 章的喜好和接受性試驗，以及第 8 章的差異試驗。

Triangle test		
Name _____	Date _____	
Type of sample _____	_____	
Instructions		
Taste samples from left to right. Two are identical; determine which is the odd sample.		
If no difference apparent, you must guess.		
Sets of three samples	Which is the odd sample?	Comments
_____	_____	_____
_____	_____	_____
_____	_____	_____

圖 7.1 三組的三角試驗用評分表範例

範例 7.1：三角差異試驗：新的麥芽供應來源

試驗啤酒 B 是使用一批新麥芽釀造的，感官分析師想要知道，如果和目前製造的控制組啤酒 A 可被區別，接受誤差為 5%危險率，有 12 位受過訓練的評審員可用；準備 18 杯“A”及 18 杯“B”構成 12 組樣品，隨機分配給受試者，ABB、BAA、AAB、BBA、ABA 與 BAB 組合中的 2 種組合。

有 8 位受試者正確認出單個樣品。從表 19.8，結論：兩種啤酒是不同的，在 5%顯著水準。

7.3.5 結果分析和解釋

計算正確回應(正確認定出奇數樣品)和回應總數。

確定所測試之正確回應的數目，是否等於或大於表 19.8 中所列的數目。

不要計算“無差異(no difference)”回應如同有效回應，告知受試者如果奇數樣品難以偵測時，就用猜的。

表 19.8 – 三角試驗中正確回應的臨界數值(輸入 $X\alpha,n$)

填入數為在所訂 α -水準(欄)及回應者的對應數 n (列)，達顯著時所需的最低正確回應數。拒絕“無差異”的假定，如果正確回應數大於或者等於表列的值。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
3	2	2	3	3	3	—	—	31	12	13	14	15	16	18	20
4	3	3	3	4	4	—	—	32	12	13	14	15	16	18	20
5	3	3	4	4	4	5	—	33	13	13	14	15	17	18	21
6	3	4	4	5	5	6	—	34	13	14	15	16	17	19	21
7	4	4	4	5	5	6	7	35	13	14	15	16	17	19	22
8	4	4	5	5	6	7	8	36	14	14	15	17	18	20	22
9	4	5	5	6	6	7	8	42	16	17	18	19	20	22	25
10	5	5	6	6	7	8	9	48	18	19	20	21	22	25	27
11	5	5	6	7	7	8	10	54	20	21	22	23	25	27	30
12	5	6	6	7	8	9	10	60	22	23	24	26	27	30	33
13	6	6	7	8	8	9	11	66	24	25	26	28	29	32	35
14	6	7	7	8	9	10	11	72	26	27	28	30	32	34	38
15	6	7	8	8	9	10	12	78	28	29	30	32	34	37	40
16	7	7	8	9	9	11	12	84	30	31	33	35	36	39	43
17	7	8	8	9	10	11	13	90	32	33	35	37	38	42	45
18	7	8	9	10	10	12	13	96	34	35	37	39	41	44	48
19	8	8	9	10	11	12	14	102	36	37	39	41	43	46	50
20	8	9	9	10	11	13	14	108	38	40	41	43	45	49	53
21	8	9	10	11	12	13	15	114	40	42	43	45	47	51	55
22	9	9	10	11	12	14	15	120	42	44	45	48	50	53	57
23	9	10	11	12	12	14	16	126	44	46	47	50	52	56	60
24	10	10	11	12	13	15	16	132	46	48	50	52	54	58	62
25	10	11	11	12	13	15	17	138	48	50	52	54	56	60	64
26	10	11	12	13	14	15	17	144	50	52	54	56	58	62	67
27	11	11	12	13	14	16	18	150	52	54	56	58	61	65	69
28	11	12	12	14	15	16	18	156	54	56	58	61	63	67	72
29	11	12	13	14	15	17	19	162	56	58	60	63	65	69	74
30	12	12	13	14	15	17	19	168	58	60	62	65	67	71	76
								174	61	62	64	67	69	74	79
								180	63	64	66	69	71	76	81

Note: For values of n not in the table, compute $z = (k-1(1/3)n)/\sqrt{(2/9)n}$, where k is the number of correct responses. Compare the value of z to the α -critical value of standard normal variable, that is, the values in the last row of Tables 19.3 ($z_\alpha = t_{\alpha,n}$).

範例 7.2：三角差異試驗的詳細範例 — 襯箔與紙包裝糖果棒

問題/情況 — 某家糖果公司包裝部經理想要試驗新的襯箔 (foil-lined) 包裝材質之效果，對照目前採用為糖果棒的紙包裝，初步觀察顯示：紙包裝糖果棒貯藏 3 個月後就開始變硬，襯箔的糖果棒仍維持柔軟。經理覺得如果 3 個月時顯示有顯著差異，他就可以證明為產品移轉包裝。

計畫目標 — 確定如果包裝改變，3 個月貯藏後引起風味及/或質地上的整體差異。

試驗目標 — 測定如果人們透過品嚐，能夠區分兩種 3 個月貯藏的產品。

進行試驗 — 來自表 19.1 的亂數碼，將每群有 54 盤的兩群編碼，從包裝取出樣品，切除丟棄每個糖果棒的頭尾部，再切成一口咬大小，擺放在編碼的小盤上，盤內放紙包裝(P)和襯箔包裝(F)的樣品，並分隔開。為每位的受試者準備一個寫上他/她編號的托盤，擺上三個小盤 P 或者 F，按照圖 7.2 的工作表中所示。記錄受試者投票上的三小盤的編碼(參見圖 7.3)。

試驗設計 — 30 至 36 位受試者的三角差異試驗。在正常的白色光照明下試驗，將外觀上的差異也能列入考慮。安排受試者分為 6 群並確使群內的完全隨機化。差異的顯著性訂為 α 危險率 5%，亦即試驗錯誤結論出有差異存在，僅 5%的機率。

篩選樣品 — 先檢查樣品(包裝之前)，以確保樣品之間並無明顯的感官上的差異。品評 3 個月後的試驗樣品，以確保不會出現總體感官的特徵，使得試驗無效。

表 19.1 1~9 的隨機順序：按三個欄的群組排列

產生一系列三位碼亂數：食指伸入表內的任何位置，譬如眼閉上、食指指向一個位置點，不看數字，決定從停留欄位往上或下移動，記下所需的數目。移除不適當的（超出範圍、重複出現等），如此方式取得數字的序列是為隨機順序。

Instructions		
(1)	To generate a sequence of three-digit random numbers, enter the table at any location, e.g., closing the eyes and pointing. Without inspecting the numbers, decide whether to move up or down the column entered. Record as many numbers as needed. Discard any numbers that are unsuitable (out of range, came up before, etc.). The sequence of numbers obtained in this manner is in random order.	
(2)	To generate a sequence of two-digit random numbers, proceed as in (1), but first decide, e.g., by coin toss, whether to use the first two or last two digits of each number taken from the table. Treat each three-digit number in the same manner, i.e., discard the same digit from each. If two-digit number comes up more than once, retain only the first.	
(3)	Random number tables are impractical for problems such as: "place the numbers from 15 to 50 in random order." Instead, write each number on a card and draw the cards blindly from a bag or use a computerized random number generator such as PROC PLAN from SAS. [®]	
862	245 458 396 522 498 298 665 635 665 113 917 365 332 896 314 688 468 663 712 585 351 847	
223	398 183 765 138 369 163 743 593 252 581 355 542 691 537 222 746 636 478 368 949 797 295	
756	954 266 174 496 133 759 488 854 187 228 824 881 549 759 169 122 919 946 293 874 289 452	
544	537 522 459 984 585 946 127 711 549 445 793 734 855 121 885 595 152 237 574 611 145 784	
681	829 614 547 869 742 822 554 448 813 976 688 959 714 912 646 873 397 159 155 136 463 363	
199	113 941 933 375 651 414 891 129 938 862 572 698 128 363 478 214 841 314 437 792 874 926	
918	481 797 621 743 827 377 916 966 429 657 246 423 277 685 533 937 223 582 946 323 626 519	
335	662 875 282 617 274 635 379 287 791 334 139 117 963 448 957 451 585 821 829 267 512 638	
477	776 339 818 251 916 581 232 372 374 799 461 276 486 274 791 369 774 795 681 458 938 171	
653	489 538 216 446 849 914 337 993 459 325 614 771 244 429 874 557 119 122 417 882 714 769	
749	824 721 967 287 556 628 843 725 731 553 253 183 653 988 431 788 426 875 838 457 927 475	
522	967 259 532 618 624 396 562 134 563 932 441 834 787 231 958 232 537 439 956 531 345 352	
475	172 986 859 925 932 282 924 842 642 797 565 399 896 596 282 441 784 258 684 625 662 291	
894	333 612 728 869 487 741 259 476 127 286 736 257 168 847 316 969 692 786 549 949 559 526	
116	218 464 191 132 218 573 786 258 296 471 372 618 935 353 747 123 863 644 161 793 196 847	
381	641 393 375 354 193 165 615 587 384 119 187 965 572 112 695 615 941 361 375 376 871 633	
968	755 847 643 773 765 439 478 611 978 868 898 546 319 775 169 896 275 513 222 114 233 184	
742	421 226 286 522 618 471 218 397 745 461 477 478 535 957 674 132 228 442 225 444 171 151	
859	878 392 311 659 772 935 447 834 117 658 161 754 654 176 883 855 195 637 751 586 948 513	
964	593 137 574 288 994 582 961 746 336 983 782 611 988 833 265 969 584 564 683 197 214 326	
177	636 674 897 167 157 856 524 662 598 145 926 362 777 415 931 313 317 195 137 959 536 985	
228	755 915 955 946 233 647 653 425 674 719 543 549 826 669 429 576 773 756 392 632 725 879	
591	214 851 669 384 349 299 192 179 264 332 294 896 299 782 397 791 659 921 569 811 683 762	
636	167 789 438 413 565 118 889 253 452 577 859 125 141 241 746 444 841 313 446 225 362 248	
415	982 543 743 835 826 364 776 988 923 224 615 283 462 328 512 228 466 278 874 373 499 437	
383	349 468 122 771 481 723 335 511 889 896 338 937 313 594 158 687 932 889 918 768 857 694	

Source: From W.G. Cochran and G.M. Cox, 1957. *Experimental Design*, John Wiley & Sons, New York.

Date	6-2-99	Worksheet	Test code	587 FF03
Post this sheet in the area where trays are prepared. Code scoresheets ahead of time. Label serving containers ahead of time.				
Type of samples:	Candy bars			
Type of test:	Triangle test			
Sample identification	Code			
Pkg 4736 (paper)	P			
Pkg 3987 (foil)	F			
Code serving containers as follows:				
Panelist #	Order of presentation			
1,7,13,19,25,31	P - F - F			
2,8,14,20,26,32	F - P - F			
3,9,15,21,27,33	F - F - P			
4,10,16,22,28,34	P - F - P			
5,11,17,23,29,35	P - P - F			
6,12,18,24,30,36	P - P - F			
1. Place stickers with panelist's number on tray.				
2. Select plates "P" of "F" from those previously coded and place on tray from left to right.				
3. Write codes selected on panelist's scoresheet.				
4. Place tray in serving container.				
5. Receive filled-in scoresheet and note on it the order of presentation used, and whether reply was correct (c) or incorrect (i).				

圖 7.2－三角試驗用工作表。範例 7.2：襯箔與紙包裝糖果棒。

分析結果－30 位受試者出席此項試驗，有 17 位正確認出奇數樣品(odd sample)。

受試者數 30
正確者數 17

查看檢定表 Table 19.8，此差異為顯著的，在 α -危險率 1% (機率 $p \leq 0.01$)。

試驗報告－完整的報告應包括計畫目標、試驗目標和試驗設計，如前述。工作表和評分表範例可以包含在內，給予受試者之任何訊息或建議(例如，關於樣品的來源)都要報告，附上表格化結果(30 之中 17 正確)和 α -危險率(符合 5%的目標)。在結論，結果要綁住計畫目標：發現紙和襯箔包裝糖果之間有顯著差異，襯箔組產生可感知的影響，10 條評論表達襯箔包裝樣品的質地較軟。

Triangle test			Test code:
Taster no.	Name:	Date:	
Type of sample:			
Instruction Taste the samples on the tray from left to right. Two samples are identical; one is different. Select the odd/different sample and indicate by placing an X next to the code of the odd sample.			
Samples on tray	Indicate odd sample	Remarks	
	☐		
	☐		
If you wish to comment on the reasons for your choice or if you wish to comment on the product characteristics, you may do so under Remarks.			

圖 7.3－三角試驗用評分表。範例 7.2：襯箔與紙包裝糖果棒，受試者選三空格的其中之一畫上 X 號，寫備註欄的意見可以超過一列。

表 19.8－三角試驗中正確回應的臨界值(輸入 $X\alpha,n$)

填入數為在所列 α -水準(欄)及回應者的對應數 n (列)，達顯著所需的正確回應的最低數。如果正確回應數大於或者等於表格的值，拒絕“無差異”假定。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
3	2	2	3	3	3	—	—	31	12	13	14	15	16	18	20
4	3	3	3	4	4	—	—	32	12	13	14	15	16	18	20
5	3	3	4	4	4	5	—	33	13	13	14	15	17	18	21
6	3	4	4	4	5	5	6	34	13	14	15	16	17	19	21
7	4	4	4	5	5	6	7	35	13	14	15	16	17	19	22
8	4	4	5	5	6	7	8	36	14	14	15	17	18	20	22
9	4	5	5	6	6	7	8	42	16	17	18	19	20	22	25
10	5	5	6	6	7	8	9	48	18	19	20	21	22	25	27
11	5	5	6	7	7	8	10	54	20	21	22	23	25	27	30
12	5	6	6	7	8	9	10	60	22	23	24	26	27	30	33
13	6	6	7	8	8	9	11	66	24	25	26	28	29	32	35
14	6	7	7	8	9	10	11	72	26	27	28	30	32	34	38
15	6	7	8	8	9	10	12	78	28	29	30	32	34	37	40
16	7	7	8	9	9	11	12	84	30	31	33	35	36	39	43
17	7	8	8	9	10	11	13	90	32	33	35	37	38	42	45
18	7	8	9	10	10	12	13	96	34	35	37	39	41	44	48
19	8	8	9	10	11	12	14	102	36	37	39	41	43	46	50
20	8	9	9	10	11	13	14	108	38	40	41	43	45	49	53
21	8	9	10	11	12	13	15	114	40	42	43	45	47	51	55
22	9	9	10	11	12	14	15	120	42	44	45	48	50	53	57
23	9	10	11	12	12	14	16	126	44	46	47	50	52	56	60
24	10	10	11	12	13	15	16	132	46	48	50	52	54	58	62
25	10	11	11	12	13	15	17	138	48	50	52	54	56	60	64
26	10	11	12	13	14	15	17	144	50	52	54	56	58	62	67
27	11	11	12	13	14	16	18	150	52	54	56	58	61	65	69
28	11	12	12	14	15	16	18	156	54	56	58	61	63	67	72
29	11	12	13	14	15	17	19	162	56	58	60	63	65	69	74
30	12	12	13	14	15	17	19	168	58	60	62	65	67	71	76
								174	61	62	64	67	69	74	79
								180	63	64	66	69	71	76	81

Note: For values of n not in the table, compute $z = (k-1(1/3)n)/\sqrt{(2/9)n}$, where k is the number of correct responses. Compare the value of z to the α -critical value of standard normal variable, that is, the values in the last row of Tables 19.3 ($z_{\alpha} = t_{\alpha,n}$).

範例 7.3：相似性三角試驗— α 、 β 及 Pd 決定品評小組人數：混合糖漿

問題/情況 — 某糖漿工廠得悉他的玉米糖漿供應商正打算提高價格，研究小組已確定另一家高品質玉米糖漿的供應商，價格更可以接受。因此，要求感官分析專家測試兩種樣品的相等性，其一為現行供應商產品的配方，另者來自另一供應商的較便宜玉米糖漿的配方。

計畫目標 — 確定如果公司的混和糖漿可改以另一供應商較便宜玉米糖漿為配方，但風味沒有可感知的變化。

試驗目標 — 測試來自目前與另一供應商的玉米糖漿所製成混和糖漿的相似性。

表 19.7 – 三角試驗中最低的評定數(輸入 $X_{\alpha,\beta,Pd}$)

填入數為三角試驗中，要提供由 α 、 β 及 Pd 的選定值所定義的靈敏度，所需的樣品大小 n 。輸入選定 Pd 值對應的項目，從所選 β 值所對應的欄，讀取所需的樣品大小 n 。

α	β									
	0.50	0.40	0.30	0.20	0.10	0.05	0.01	0.001		
$P_d = 50\%$										
0.40	3	3	3	6	8	9	15	26		
0.30	3	3	3	7	8	11	19	30		
0.20	4	6	6	7	10	12	20	33		
0.10	7	9	8	12	15	23	35	48		
0.05	13	15	11	16	20	23	35	48		
0.01	22	26	20	30	43	48	62	81		
$P_d = 40\%$										
0.40	3	3	3	6	9	15	26	41		
0.30	3	3	3	7	8	11	19	30		
0.20	6	10	15	12	17	25	36	55		
0.10	11	15	16	23	30	40	57	79		
0.05	21	26	30	40	56	66	97	136		
0.01	36	48	53	76	102	138	233	396		
$P_d = 30\%$										
0.40	3	6	6	9	15	26	44	73		
0.30	3	8	8	16	22	30	53	84		
0.20	7	12	17	20	28	39	64	97		
0.10	15	25	30	40	54	66	98	119		
0.05	16	23	30	40	53	66	98	136		
0.01	33	40	52	62	82	97	131	181		
$P_d = 20\%$										
0.40	6	9	12	18	35	50	94	153		
0.30	8	11	13	20	37	52	97	167		
0.20	12	20	28	39	64	86	140	212		
0.10	25	33	46	62	89	119	178	260		
0.05	40	48	66	87	117	147	213	305		
0.01	72	92	110	136	176	211	292	397		
0.001	130	148	176	207	257	302	396	513		

評定者人數和 α 、 β 及 Pd 值的選擇 — 感官分析專家和計畫經理查看表 19.7，注意到：避免錯誤結論出相似性而可做最大的保護，譬如設定 β 為 0.1% ($\beta = 0.001$) 相對於另一假設(alternative hypothesis)，即能夠測出樣品之間的差異的人群其真正比率至少 20% ($P_d = 0.20$)，然後保留適度的 α -危險率 0.10 時，他們需要至少有 260 位評定者。

最後，他們妥協改設定為 $\alpha = 0.20$ 、 $\beta = 0.01$ 、 $P_d = 30\%$ ，這需要 64 位評定者。

試驗設計 — 感官分析專家根據為混合糖漿所建立的試驗方案進行 66-回應的三角試驗。感官品評小間(booth)備有紅色濾波器，以遮蔽顏色上的差異。12 位品評員安排進行五個連續品評會的每一次，以及 6 位安排做第六次最後的品評會。圖 7.4 顯示專家為典型的一次會期所做的工作表。

Date	11-5-98	Worksheet	No.	95-0032-31
Post this sheet in the area where trays are prepared. Code sheets ahead of time. Label serving containers ahead of time.				
Type of samples :	Blended table syrups			
Type of test :	Triangle similarity test			
Sample identification:		Codes used for :		
A:	Lab code 47-3651	Sets with 2 A's	Sets with 2 B's	
B:	Lab code 026 (Control)	587 894	246 413	
Codes serving containers as follows:		Underlying pattern*		
Subject #	Codes in order			
1	587	246	894	AAA
2	3	413	365	BBB
3	751	413	365	BAB
4	246	587	894	AAA
5	587	894	246	ABA
6	587	894	246	ABB
7	413	751	365	AAA
8	584	587	246	BAA
9	365	751	413	BBA
10	584	587	246	BAB
11	365	413	751	BBB
12	365	413	751	BAB
*Each pattern is repeated twice to allow for each code in each position.				
Underlying pattern*				
AAA				
BBB				
BAA				
ABA				
BAB				

圖 7.4 – 相似性三角試驗。範例 7.3：混合餐桌用糖漿。

分析結果 – 66 位回應者之中，**21 位**正確找出奇數樣品，參照表 19.8，在 $n = 66$ 所對應列和 $\alpha = 0.20$ 對應欄的交叉點，找到：顯著性所需的最低正確回應是 26，因此，正確回應僅 21 即結論：兩種糖漿之間的任何感官差異非常小而可忽略之，也就是兩種樣品有充分地相似而可互相取代使用。

解釋結果 – 專家告知計畫經理在試驗結果為 66 回應中，22 位正確回應，99%信賴度顯示能夠感知差異之人群的比率低於 30%，很可能更低，可接受另一個供應商的產品。

pd 的信賴區間 – 如需要，分析專家計算可區分樣品之人群比率的信賴區間(confidence limits)，計算如下：

$$\begin{aligned} p_c \text{ (proportion correct)} &= c/n \\ p_d \text{ (proportion distinguishers)} &= 1.5p_c - 0.5 \\ s_d \text{ (standard deviation of } p_d) &= 1.5 \sqrt{p_c(1 - p_c)} / n \\ \text{one-sided upper confidence limit} &= p_d + z_{\beta} s_d \\ \text{one-sided lower confidence limit} &= p_d - z_{\alpha} s_d \end{aligned}$$

其中， c 為正確回應數； n 為評定者總數； **z_{θ}** 和 **s_d** 為標準常態分佈的臨界值，單尾(one-sided)信賴區間的常用 **z** 值：

信賴區間	z
75%	0.674
80%	0.842
85%	1.036
90%	1.282
95%	1.645
99%	2.236

表 19.8 – 在三角試驗中正確回應的臨界數(輸入為 $X_{\alpha,n}$)。

填入數為在所定 α -水準(欄)下針對回應者的對應數 n (列)，連顯著所需的最低正確回應數。如果正確回應數大於或者等於表列的值，拒絕“無差異”假定。

n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	π	0.40	0.30	0.20	0.10	0.05	0.01	0.001
3								31	12	13	14	15	16	18	20
4	2							32	12	13	14	15	16	18	20
5	3	3						33	13	13	14	15	17	19	21
6	3	4	4					34	13	14	15	16	17	19	21
7	4	4	5	5				35	14	14	15	16	17	20	22
8	4	4	5	6	6	7	8	36	14	15	16	17	19	20	22
9	4	5	6	6	7	8	9	37	15	16	17	18	20	22	25
10	5	5	6	7	8	9	10	38	16	17	18	19	20	22	27
11	5	6	7	7	8	10	11	39	16	18	19	20	22	25	30
12	5	6	7	8	9	10	11	40	17	18	19	20	22	27	33
13	6	6	7	8	9	10	11	41	17	19	20	21	23	27	35
14	6	7	7	8	9	11	12	42	18	19	20	21	23	30	38
15	6	7	8	9	10	11	12	43	18	20	21	22	25	32	40
16	7	7	8	9	11	12	13	44	19	20	21	22	25	34	43
17	7	8	8	9	10	11	12	45	19	21	22	23	27	36	45
18	7	8	9	10	10	12	13	46	20	21	22	23	27	37	48
19	8	8	9	10	11	12	13	47	20	22	23	24	27	39	50
20	8	9	9	10	11	12	14	48	21	22	23	24	28	40	53
21	8	9	10	11	12	13	14	49	21	23	24	25	28	41	55
22	9	9	10	11	12	13	15	50	22	23	24	25	29	42	57
23	9	10	11	12	12	14	16	51	22	24	25	26	30	43	60
24	10	10	11	12	13	15	16	52	23	24	25	26	31	44	62
25	10	11	11	12	13	15	17	53	23	25	26	27	32	45	64
26	10	11	12	13	14	16	17	54	24	25	26	27	33	46	67
27	11	11	12	13	14	16	18	55	24	26	27	28	34	47	69
28	11	12	12	13	14	16	18	56	25	26	27	28	35	48	72
29	11	12	13	14	15	17	19	57	25	27	28	29	36	49	74
30	12	12	13	14	15	17	19	58	26	27	28	30	37	50	76

Note: For values of n not in the table, compute $z = (k - 1)(1/3)n^{1/3} / \sqrt{(2/9)n}$, where k is the number of correct responses and n is the number of subjects. The value of z to the α -critical value of standard normal variable, that is, the values in the last row of Tables 19.3 ($\alpha = \alpha_{nom}$).

由範例的數據，辨識者比率其上限 99%單尾的信賴區間，計算如下：

$$\begin{aligned} p_{\max} = p_d + z_{\beta} s_d &= [1.5(21/66) - 0.5] + (2.326)(1.5) \sqrt{(21/66)(1 - (21/66)) / 66} \\ &= [-0.023] + 2.326(1.5)(0.05733) \\ &= 0.177 \text{ or } 18\% \end{aligned}$$

while the lower 80% one-sided confidence limit falls at

$$p_{\min} = p_d - z_{\alpha} s_d = [-0.023] - 0.842(1.5)(0.05733) = -0.095 \text{ (i.e., 0.0, it cannot be negative)}$$

或換言之，感官分析專家是 99%確信能夠辨識樣品之人群的真正比率為不超過 18%，可能低至 0%*。

*統一方法相對於相似性檢定表：注意第五版採用的統一方法並不含相似性檢定表，如第二版中，如本範例所示，表 19.8 僅顯示相似性的證明存在，要瞭解其證據多強，亦即“Pd 不超過 18%，並可能低至 0%”，專家需計算信賴區間，參照 14.2.3 項的信賴區間的導出。

範例 7.4：平衡 α 、 β 及 P_d 的設定。設定軟性飲料組合的食用期限

問題/情況 — 某家軟性飲料組合的製造商想選出建議的食用期限列印在瓶裝軟性飲料上，已知冷藏(2°C)樣品能貯存超過一年而無任何風味的改變，而在高溫下，風味貯藏壽命就變短。進行試驗：樣品於高室溫(30°C)下貯存 6、8 及 12 個月，然後做差異測試。

計畫目標 — 選出組合物製成的瓶裝產品的有效期限。

試驗目標 — 確定冷藏產品與較高溫度下貯存的三個產品之間，感官差異是否明顯。

評定者人數和 α 、 β 及 P_d 值的選擇 — 製造人員希望看到最慢的可能到期日，並決定僅願意冒 5% 機率當並不存在而結論出有差異 (即 $\alpha = 0.05$)，而另一方面，品質保證經理希望合理地確定消費者在到期日前不會測出陳味(“aged flavor”)，因而，他同意接受：90% 確定性 (即 $\beta = 0.01$) 不超過 30% 的受試群體 (即 $P_d = 30\%$) 能偵測出差異。

製造人員：要看到最遲的可能食用期限，因而決定當無差異卻結論出有差異，僅為 5% 機率。品質保證經理：要合理地確定消費者在食用期限之內都不會感覺到陳味。

參照表 19.7 中 $\beta = 0.10$ 欄和 $P_d = 30\%$ 區，感官專家發現此試驗需要 53 人的品評小組，但試驗期間只有 30 位品評員可用。因此，他們三人妥協試驗靈敏度的參數，在可用的評定者人數下，提供可能的風險保護的最大值；再參照表 19.7，他們決定的妥協為 $P_d = 30\%$ 、 $\beta = 0.20$ 和 $\alpha = 0.10$ ，可用評定人數下可接受的靈敏度。

表 19.7 – 三角試驗中最低的評定數(輸入 X_{α,β,P_d})

填入數為三角試驗中，要提供由 α 、 β 及 P_d 的選定值所定義的靈敏度，所需的樣品大小(n)。輸入選定 P_d 值對應的項目，從所選 β 值所對應的欄，讀取所需的樣品大小 n 。

α	β							
	0.50	0.40	0.30	0.20	0.10	0.05	0.01	0.001
$P_d = 50\%$								
0.40	3	3	3	6	8		15	26
0.30	3	3	3	7	8	11	19	30
0.20	4	6	7	7	12	16	25	36
0.10	7	8	8	12	15	20	30	43
0.05	7	9	11	16	20	23	35	48
0.01	13	15	19	25	30	35	47	62
0.001	22	26	30	36	43	48	62	81
$P_d = 40\%$								
0.40	3	3	6	6	9	15	26	41
0.30	3	3	7	8	11	19	30	47
0.20	6	7	7	12	17	25	36	55
0.10	8	10	15	17	25	30	46	67
0.05	11	15	16	23	30	40	57	79
0.01	21	26	30	35	47	56	76	102
0.001	36	39	48	55	68	76	102	130
$P_d = 30\%$								
0.40	3	6	6	9	15	26	44	73
0.30	3	8	8	16	22	30	53	84
0.20	7	12	17	20	28	39	64	97
0.10	15	15	20	30	43	54	81	119
0.05	16	23	30	40	53	66	98	136
0.01	33	40	52	62	82	97	131	181
0.001	61	69	81	93	120	138	181	233
$P_d = 20\%$								
0.40	6	9	12	18	35	50	94	153
0.30	8	11	19	30	47	67	116	183
0.20	12	20	28	39	64	86	140	212
0.10	25	33	46	62	89	119	178	260
0.05	40	48	66	87	117	147	213	305
0.01	72	92	110	136	176	211	292	397
0.001	130	148	176	207	257	302	396	513

表 19.8 – 在三角試驗中正確回應的臨界數(輸入為 $X_{\alpha,n}$)。

填入數為在所定 α -水準(欄)下針對回應者的對應數 n (列)，達顯著所需的最低正確回應數。如果正確回應數大於或者等於表列的值，拒絕“無差異”假定。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
3	2	2	3	3	3			31	12	13	14	15	16	18	20
4	3	3	3	4	4			32	12	13	14	15	16	18	20
5	3	3	4	4	4	5		33	13	13	14	15	17	18	21
6	3	4	4	5	5	6		34	13	14	15	16	17	19	21
7	4	4	4	5	5	6	7	35	13	14	15	16	17	19	22
8	4	4	5	5	6	7	8	36	14	14	15	17	18	20	22
9	4	5	5	6	6	7	8	42	16	17	18	19	20	22	25
10	5	5	6	6	7	8	9	48	18	19	20	21	22	25	27
11	5	5	6	7	7	8	10	54	20	21	22	23	25	27	30
12	5	6	6	7	8	9	10	60	22	23	24	26	27	30	33
13	6	6	7	8	8	9	11	66	24	25	26	28	29	32	35
14	6	7	7	8	9	10	11	72	26	27	28	30	32	34	38
15	6	7	8	8	9	10	12	78	28	29	30	32	34	37	40
16	7	7	8	9	9	11	12	84	30	31	33	35	36	39	43
17	7	8	8	9	10	11	13	90	32	33	35	37	38	42	45
18	7	8	9	10	10	12	13	96	34	35	37	39	41	44	48
19	8	8	9	10	11	12	14	102	36	37	39	41	43	46	50
20	8	9	9	10	11	13	14	108	38	40	41	43	45	49	53
21	8	9	10	11	12	13	15	114	40	42	43	45	47	51	55
22	9	9	10	11	12	14	15	120	42	44	45	48	50	53	57
23	9	10	11	12	12	14	16	126	44	46	47	50	52	56	60
24	10	10	11	12	13	15	16	132	46	48	50	52	54	58	62
25	10	11	11	12	13	15	17	138	48	50	52	54	56	60	64
26	10	11	12	13	14	15	17	144	50	52	54	56	58	62	67
27	11	11	12	13	14	16	18	150	52	54	56	58	61	65	69
28	11	12	12	13	14	16	18	156	54	56	58	61	63	67	72
29	11	12	13	14	15	17	19	162	56	58	60	63	65	69	74
30	12	12	13	14	15	17	19	168	58	60	62	65	67	71	76
								174	61	62	64	67	69	74	79
								180	63	64	66	69	71	76	81

Note: For values of n not in the table, compute $z = (k-1(1/3)n)/\sqrt{(2/9)n}$, where k is the number of correct responses. Compare the value of z to the α -critical value of standard normal variable, that is, the values in the last row of Tables 19.3 ($z_{\alpha} = t_{\alpha,n}$).

7.4 二-三點試驗 Duo-Trio Test

7.4.1 範圍與應用

和三角試驗比較，二-三點試驗在統計上是效率很低的，因取得正確結果之機會為二分之一。另一方面，此試驗簡單、易懂。和配對比較試驗(paired Comparison test)比較，其優點為呈現一個參考樣品(reference sample)，這避免是甚麼構成差異這方面的困擾，但有一個缺點即必須嚐試三個樣品，而非二個樣品。

當試驗目標是要確定兩樣品間是否感官差異存在，使用這方法。此法特別適用於下列情況：

- 1. 確定產品差異是否來自組成分、加工、包裝、或貯藏的改變
- 2. 確定是否有整體差異存在，並無特定的屬性被認定受影響

試驗設計 — 專家準備和執行品評員 30 位的三角試驗。
分析結果 — 正確選擇人數的結果：貯存 6 個月 11 位；8 個月 13 位；12 個月 15 位。查表 19.8，感官分析專家結論：貯存 6 個月，並無差異存在之證實，在 8 個月，差異變更大，從表 19.8 顯示，採 $\alpha = 0.20$ 之較高值時，就證實有差異存在，最後，在 12 個月，檢定表顯示在 $\alpha = 0.05$ ，證實有差異存在。

解釋 — 這群主管決定：食用期限為 8 個月，對於未達期限的產品就發生有陳味，可提供充分的保證。額外檢定其結論，當計算每個試驗的 80%單尾的信賴界限(confidence limit)，可發現 80%確定不超過 16%的消費者能偵測出貯存 6 個月樣品的差異，對 8 個月樣品為 26%，對 12 個月樣品可能高達 37%，在 $P_d = 30\%$ 的限制下，8 個月的產品是安全的。

*6 個月樣品為例，其信賴界限的計算：

$$P_d = (1.5(11/30) - 0.5) + 0.84(1.5)\sqrt{(11/30)(1 - (11/30))} / 30 = 0.16$$

二-三點試驗的一般應用，超過 15 位，最好能超過 30 位的試驗受試者可用。此試驗有兩種形式：

- **恆定參考品模式(constant reference mode)**：通常取自正常生產之相同的樣品，一直作為參考品。採用經訓練的受試者，每當他們熟知的產品可以作為參考品。
- **平衡參考品模式(balanced reference mode)**：要比較的樣品兩者都隨機當作參考品。要測試樣品兩者都未知或者採用未訓練的受試者。

如果有明顯的餘味(aftertastes)，比起**配對比較試驗(Paired Comparison test)**；參見 7.2 項)，更不適合使用**二-三點試驗**。

Duo-trio test		Test No.
Taster no.	Name:	Date:
Type of sample		
Instructions: Taste samples from left to right. The left hand sample is a reference. Determine which of the two samples matches the reference and indicate by placing an X.		
If no difference is apparent between the two unknown samples, you must guess.		
Reference	Code	Code
☒	☐	☐
Comments:		

圖 7.5 二-三點試驗的評分卷

7.4.2 試驗原理

給每位受試者一個已知的參考樣品，然後兩個編碼樣品，其中一個**匹配參考品**，詢問受試者指出哪個編碼的樣品和參考品匹配。計算正確回答數，並參考表 19.10 作解釋。

7.4.3 試驗受試者

挑選、訓練及指導受試者如 7.3.3 項所述，一般而言，最少 16 位受試者，但若低於 28 位， β -誤差(β -error)高，如果 32、40 位或能採用更多的人數，可大為提升辨識(discrimination)。

7.4.4 試驗程序

試驗管制及產品管制參見第 3 章。可能的話，同時提供出樣品，否則連續地。可能的組合（參見範例）都準備相同數量，並隨機分配這些組別給受試者。評分卷（在恆定和平衡參考品模式都相同）範例列於圖 7.5。

評分卷提供數次二-三點試驗夠用的空間，但不要詢問另追加的問題(例如，差異的類型或幅度、或者受試者的喜好/偏好)，因受試者的選擇匹配樣品會造成對這些額外問題的回應造成的偏見。

計算正確回應數和總回應數，參照表 19.10 檢定。

不要計算“無差異”之回應，受試者如有疑惑還是必須猜。以下的範例都採用統一的方法。

範例 7.5：平衡型參考品－面紙盒用香水

問題/情況 — 某產品開發香水化學家想知道，如果盒裝面紙的兩種香水傳送的方法，香水直接傳送至面紙或香水傳送至盒內部，在感知的香水品質或濃度上，會產生產品的差異？

計畫目標 — 確定如香水傳送的兩種方法，在儲存一段時間和正常的產品產期比較使用時，會產生這兩種面紙所感知香氣的任何差異。

試驗目標 — 確定兩種面紙樣品貯存三個月後，能否感知香氣的差異。

Duo-trio test		Test no.
		230S
Panelist no.	21	Name:
Type of Sample:	Facial tissue in a glass jar	
Date:		
Instructions		
1. Please sniff each sample, starting at the left. Remove the cap only briefly and take short, shallow sniffs.		
2. The left hand sample is a reference. Determine which of the two coded samples matches the fragrance of the reference.		
3. Indicate the matching sample by placing an X in the corresponding box.		
If no difference is apparent between the two unknown samples, you must guess.		
Reference	Code	Code
	☐	☐
Comments:		

圖 7.6 二-三點試驗評分卷；範例 6.5：平衡型參考品模式

試驗設計 — 當對象是複雜的刺激物，二-三點試驗比起三角試驗或屬性差異測試(attribute difference testing)，樣品的嗅聞較少重複，這減少氣味順應及/或整理三個樣品的相互比較時所引起可能的困惑。

試驗由具備氣味品評經驗的 40 位受試者來執行。

由香水化學家製備樣品，採用相同的香水和同一天生產的面紙。

然後，於同樣的條件下面紙盒貯存 3 個月，從盒的中間 50%處取出試驗用面紙；品評前 1 小時每組面紙放入密封玻璃罐內，使香氣能擴散至上部間隙，採用密封罐可減少測試小間內的香氣濃度。

兩組樣品的每種都用作一半品評(20)的參考品，圖 7.6 為使用的分卷。

分析結果 — 40 位受試者中僅 21 位選對設定參考品的正確匹配，依照表 19.10，5%危險率所需的正確回應為 26，再者，審視每個樣品都作為參考品其位置上的可能效應，結果顯示正確回應的分佈均等(10 和 11)，這表示兩種香氣的質及/或量，如有的話，額外對結果的偏愛影響很少。

解釋結果 — 感官分析專家告知香水化學家在研究所採用的貯存時間，兩種香水包裝系統的面紙彼此間，氣味二-三點試驗無法測出任何顯著的氣味差異。

試驗的靈敏度 — 為考量此類型在以後的研究，注意二-三點試驗選 40 位受試者，產生下列的試驗靈敏度參數的值：

Proportion of Distinguishers (p_d) %	Probability of Detecting	
	$(1-\beta) @ \alpha=0.05$	$(1-\beta) @ \alpha=0.10$
10	0.13	0.21
15	0.21	0.32
20	0.32	0.44
25	0.44	0.57
30	0.57	0.69
35	0.70	0.80
40	0.81	0.88
45	0.89	0.94
50	0.95	0.97

若採用 40 人及設定 $\alpha = 0.05$ ，此測試當群體的 25%能測出差異 (Pd=25%)，測得此狀況的機率為 44% ($1 - \beta = 0.44$)。增加測試者人數，可提高任何 Pd 值下測出差異的可能性，而提高 α 值，也提高任何 Pd 值下測出差異的機率。

計畫目標 — 確定如果室溫 8 週架售貯存後，這兩種啤酒能被感知任何的感官差異。

評定者數目 — 釀酒商從過去的經驗得知，如果不過品評小組的 Pd = 30% 能測出差異，他可假定在市場上就無有意義的風險。比起錯過空罐 B 所提供稍微延長的架售期，他略為更在意導入不想要的差異。因此，他決定設定 β -危險率 0.05 和 α -危險率 0.10，參考表 19.9 中的 Pd = 30%項， $\beta = 0.05$ 欄和 $\alpha = 0.10$ 列，發現此試驗需要 96 個回應。

試驗設計 — 採用恆定參考品模式的二-三點試驗是適合的，因公司的啤酒以空罐 A 裝填是試吃者所熟悉。以分隔試驗在釀酒商的三個試驗點分開進行，每一試驗配備 32 位受試者，以“A”為參考品；準備 64 杯啤酒“A”和 32 杯啤酒“B”，提供 16 個組合 AAB 和 16 個組合 ABA，左手邊樣品為參考品。

表 19.10 – 在二-三點和單尾定向差異試驗的正確回應臨界數 (輸入為 $X_{\alpha,n}$)

填入數為所訂 α -水準(欄)下回應者的對應數 n (列)達顯著所需的最低正確回應數。如果正確回應數大於或者等於表列的值，拒絕“無差異”假定。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
2	2	2	—	—	—	—	—	31	17	18	19	20	21	23	25
3	3	3	3	—	—	—	—	32	18	18	19	21	22	24	26
4	3	4	4	4	—	—	—	33	18	19	20	21	22	24	26
5	4	4	4	5	5	—	—	34	19	20	20	22	23	25	27
6	4	5	5	6	6	—	—	35	19	20	21	22	23	25	27
7	5	5	6	6	7	7	—	36	20	21	22	23	24	26	28
8	5	6	6	7	7	8	—	40	22	23	24	25	26	28	31
9	6	6	7	7	8	9	—	44	24	25	26	27	28	31	33
10	6	7	7	8	9	10	10	48	26	27	28	29	31	33	36
11	7	7	8	9	9	10	11	52	28	29	30	32	33	35	38
12	7	8	8	9	10	11	12	56	30	31	32	34	35	38	40
13	8	8	9	10	10	12	13	60	32	33	34	36	37	40	43
14	8	9	10	10	11	12	13	64	34	35	36	38	40	42	45
15	9	10	10	11	12	13	14	68	36	37	38	40	42	45	48
16	10	10	11	12	12	14	15	72	38	39	41	42	44	47	50
17	10	11	11	12	13	14	16	76	40	41	43	45	46	49	52
18	11	11	12	13	13	15	16	80	42	43	45	47	48	51	55
19	11	12	12	13	14	15	17	84	44	45	47	49	51	54	57
20	12	12	13	14	15	16	18	88	46	47	49	51	53	56	59
21	12	13	13	14	15	17	18	92	48	50	51	53	55	58	62
22	13	13	14	15	16	17	19	96	50	52	53	55	57	60	64
23	13	14	15	16	16	18	20	100	52	54	55	57	59	63	66
24	14	14	15	16	17	19	20	104	54	56	57	60	61	65	69
25	14	15	16	17	18	19	21	108	56	58	59	62	64	67	71
26	15	15	16	17	18	20	22	112	58	60	61	64	66	69	73
27	15	16	17	18	19	20	22	116	60	62	64	66	68	71	76
28	16	16	17	18	19	21	23	122	63	65	67	69	71	75	79
29	16	17	18	19	20	22	24	128	66	68	70	72	74	78	82
30	17	18	19	20	22	24	24	134	69	71	73	75	78	81	86
31	17	18	20	20	22	24	24	140	72	74	76	79	81	85	89

Note: For values of n not in the table, compute $z = (k - 0.5n) / \sqrt{0.25n}$ where k is the number of correct responses. Compare the value of z to the α -critical value of a standard normal variable, that is, the values in the last row of Table 19.3 ($z_\alpha = t_{\alpha,n}$).

範例 7.6: 恆定參考品 — 新罐頭塗料

問題/情況 — 一位釀酒商正面對兩種空罐來源，編號 A 是已使用多年的固定供應來源，編號 B 是新推薦的，宣稱在架售期上有些好處，他想要知道是否這兩種罐之間可測出任何的差異。釀酒商覺得導入啤酒有不要的改變相對於錯過空罐 B 提供的展延架售期的風險，之間取得平衡是重要的。

計畫目標 — 確定如果包裝改變而引起架售期貯存後啤酒任何的可感知的差異，如同貿易上所經歷的哪樣。

分析結果 — 18、20 及 19 位受試者正確認出匹配參考品之樣品。依照表 19.10，在 10%水準的顯著性需要 21 個正確數。

表 19.10 – 在二-三點和單尾定向差異試驗的正確回應臨界數
(輸入為 $\chi_{\alpha,n}$)

填入數為所訂 α -水準(欄)下回應者的對應數 n (列)達顯著所需的最低正確回應數。如果正確回應數大於或者等於表列的值，拒絕“無差異”假定。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
2	2	2	—	—	—	—	—	31	17	18	19	20	21	23	25
3	3	3	3	—	—	—	—	32	18	18	19	21	22	24	26
4	3	4	4	4	—	—	—	33	18	19	20	21	22	24	26
5	4	4	4	5	5	—	—	34	19	20	20	22	23	25	27
6	4	5	5	6	—	—	—	35	19	20	21	22	23	25	27
7	5	5	6	6	7	7	—	36	20	21	22	23	24	26	28
8	5	6	6	7	7	8	—	40	22	23	24	25	26	28	31
9	6	6	7	7	8	9	—	44	24	25	26	27	28	31	33
10	6	7	7	8	9	10	10	48	26	27	28	29	31	33	36
11	7	7	8	9	9	10	11	52	28	29	30	32	33	35	38
12	7	8	8	9	10	11	12	56	30	31	32	34	35	38	40
13	8	8	9	10	10	12	13	60	32	33	34	36	37	40	43
14	8	9	10	10	11	12	13	64	34	35	36	38	40	42	45
15	9	10	10	11	12	13	14	68	36	37	38	40	42	45	48
16	10	10	11	12	12	14	15	72	38	39	41	42	44	47	50
17	10	11	11	12	13	14	16	76	40	41	43	45	46	49	52
18	11	11	12	13	13	15	16	80	42	43	45	47	48	51	55
19	11	12	12	13	14	15	17	84	44	45	47	49	51	54	57
20	12	12	13	14	15	16	18	88	46	47	49	51	53	56	59
21	12	13	13	14	15	17	18	92	48	50	51	53	55	58	62
22	13	13	14	15	16	17	19	96	50	52	53	55	57	60	64
23	13	14	15	16	17	19	20	100	52	54	55	57	59	63	66
24	14	14	15	16	17	19	20	104	54	56	57	60	61	65	69
25	14	15	16	17	18	19	21	108	56	58	59	62	64	67	71
26	15	15	16	17	18	20	22	112	58	60	61	64	66	69	73
27	15	16	17	18	19	20	22	116	60	62	64	66	68	71	76
28	16	16	17	18	19	20	22	122	63	65	67	69	71	75	79
29	16	17	18	19	21	23	24	128	66	68	70	72	74	78	82
30	17	18	19	20	22	24	24	134	69	71	73	75	78	81	86
31	17	18	20	20	22	24	24	140	72	74	76	79	81	85	89

Note: For values of n not in the table, compute $z = (k - 0.5n)/\sqrt{0.25n}$ where k is the number of correct responses. Compare the value of z to the α -critical value of a standard normal variable, that is, the values in the last row of Table 19.3 ($z_{\alpha} = t_{\alpha,n}$).

範例 7.7: 二-三點相似性試驗(Duo-Trio Similarity Test) — 取代調豆咖啡

問題/情況 — 某家咖啡製造商得知長期以來用來調豆的一種主要咖啡豆即將在未來二年後短缺，研究小組已完成三種新的調豆配方，他們覺得在風味上和現行的調豆相同。研究小組要求感官分析專家測試這些新調豆配方和現有的產品的等同性(equivalency)。

- 計畫目標 — 確定這三種調豆的何者是最佳用來取代現有調豆。
- 試驗目標 — 測試現行調豆和計畫調豆的每一種之間的相似性。

Table 19.9 Minimum Number of Assessments in a Duo-Trio or One-Sided Directional Difference Test (Entries are n_{α,β,p_d})
Entries are the sample sizes (n) required in duo-trio or one-sided directional difference test to deliver the sensitivity defined by the values chosen for α , β , and p_d . Enter the table in the section corresponding to the chosen value of p_d for duo-trio test or $p_{\max}$ for a directional difference test and the row corresponding to the chosen value of α . Read the required sample size, n , from the column corresponding to the chosen value of β .

α	β							
	0.50	0.40	0.30	0.20	0.10	0.05	0.01	0.001
$p_d = 50\%$								
$p_{\max} = 75\%$								
0.40	2	4	4	6	10	14	27	41
0.30	2	5	7	9	13	20	30	47
0.20	5	5	10	12	19	26	39	58
0.10	9	9	14	19	26	33	48	70
0.05	13	16	18	23	33	42	58	82
0.01	22	27	33	40	50	59	80	107
0.001	38	43	51	61	71	83	107	140
$p_d = 40\%$								
$p_{\max} = 70\%$								
0.40	4	4	6	8	14	25	41	70
0.30	5	7	9	13	22	28	49	78
0.20	5	10	12	19	30	39	60	94
0.10	14	19	21	28	39	53	79	113
0.05	18	23	30	37	53	67	93	132
0.01	35	42	52	64	80	96	130	174
0.001	61	71	81	95	117	135	176	228
$p_d = 30\%$								
$p_{\max} = 65\%$								
0.40	4	6	8	14	29	41	76	120
0.30	7	9	13	24	39	53	88	144
0.20	10	17	21	32	49	68	110	166
0.10	21	28	37	53	72	96	145	208
0.05	30	42	53	69	93	119	173	243
0.01	64	78	89	112	143	174	235	319
0.001	107	126	144	172	210	246	318	412

(Continued)

註：許多案例為了提高辨識的結果，允許合併二或三個試驗。在目前的案例，罐頭是同一批的樣品，受試者來自相同的品評小組，因此合併是允許的。

3 x 32 = 96 試驗中，得到正確數 18 + 20 + 19 = 57，參照表 19.10，正確回答的臨界數：在 10%顯著水準為 55，在 5%顯著水準為 57。

解釋結果 — 結合三次試驗的數據，在 5%顯著水準下，結論出：有差異存在。其次，檢討品評員做出的注釋，來描述這差異。如果沒有，提送樣品給描述性品評小組，最後，如果這差異既非不受喜歡的也非受喜歡的，可能就需要消費者試驗來確定是否有偏好於某一罐頭或另一罐頭。

試驗設計 — 預備試驗的結果顯示：差異都小，也沒有特別關連於某一特定的屬性。因此，採用相似性的二-三點試驗是適合的，而為降低誤失可感知的差異之風險，感官分析專家建議每次試驗採用 60 位品評員(比常態用於測試差異性之 36 位，再增加)。

利用軟體 Excel 中的 spreadsheet test-sensitivity analyzer (第 13 章 13.3.5)，決定 60 位回應者的二-三點試驗：在品評員的 $Pd = 25\%$ 能測出差異， α -危險率 0.25 之情況，有 90% (即 $\beta = 0.10$) 機率測出。

感官分析專家接受 α -危險率的值大，乃更在意於錯誤地贊成新調豆配方異於控制組，且僅有 60 位品評員可用於試驗。

Date3-4-99

Cell no. 3

WORKSHEET

No. 2803-30

Post this sheet in the area where trays are prepared. Code score sheets ahead of time. Label serving containers ahead.

Type of samples: cups of coffee

Type of test: Duo-trio similarity test (balanced reference)

Samples served

A = Control

B = Blend 62-A

C = Blend 223B

D = Blend 211

Codes Used:

For B versus A

Sets w/ 2 A's

Sets w/ 2 B's

For C versus A

Sets w/ 2 A's

Sets w/ 2 C's

For D versus A

Sets w/ 2 A's

Sets w/ 2 D's

Sample A

317 543

986

Sample B

314

393 737

Sample C

Sample D

866 581

674

373 158

121 225

221

499 134

Code serving containers as follows:

Subject

Pat-tern

Codes in order

Pat-tern

Codes in order

37

ABA

R

314-543

AAC

R

581-674

DAD

R

965-134

38

BBA

R

737-986

CCA

R

674-866

AAD

R

225-221

39

BAB

R

986-393

CAA

R

158-541

ADA

R

221-121

40

AAB

R

317-314

CAC

R

541-373

DDA

R

499-965

41

ABA

R

314-543

AAC

R

581-674

DAD

R

965-134

42

BBA

R

737-986

CCA

R

674-866

AAD

R

225-221

43

BAB

R

986-393

CAA

R

158-541

ADA

R

221-121

44

AAB

R

317-314

CAC

R

541-373

DDA

R

499-965

45

ABA

R

314-543

AAC

R

581-674

DAD

R

965-134

46

BBA

R

737-986

CCA

R

674-866

AAD

R

225-221

47

BAB

R

986-393

CAA

R

158-541

ADA

R

221-121

48

AAB

R

317-314

CAC

R

541-373

DDA

R

499-965

圖 7.7 - 二-三點相似性試驗的工作表，範例 7.7：取代咖啡調豆

對每一調豆配方，感官分析專家計畫在一周內進行 60 個回應的咖啡試驗，製備和放置時間會影響風味而是為關鍵因素，受試者必須小心排定時程，在產品製備好的 10 分鐘內。使用感官實驗室的 12 間品評小間，燈光採用褐色濾光器，每次試驗安排 12 位不同受試者，每會期 12 位品評員可使每樣品的呈現平衡，如同參考樣品，同樣在每品評小間兩試驗樣品呈現的順序也平衡。圖 7.7 為專家的工作表。

樣品呈現不加奶精和糖，咖啡壺保溫在 175°F，倒入預熱(130°F)的陶器杯，如工作表所述編碼，並按順序排列。評分卷(圖 7.8) 事先準備以解省時間，當受試者已作在品評小間內，再倒入樣品。

Taster no. _____		Name: _____		Date: _____					
Type of sample: _____		cups of coffee, freshly brewed							
Duo-trio test									
Test no. 28 03-03									
Instructions: Taste sample from left to right The left hand sample is a reference. Determine which of the two axis samples matches the reference and indicate by placing an X									
If no difference is apparent between the two unknown samples, you must guess									
Reference		Code		Code					
☒		☐		☐					
Comments: _____									

圖 7.8 - 二-三點相似性試驗的評分卷，範例 7.7：取代咖啡調豆

分析結果 — 三個試驗調豆的正確回應數目分別為

Cell No. (of 12 Subjects)	Blend B	Blend C	Blend D
1	3	6	8
2	4	5	8
3	5	7	5
4	7	7	7
5	5	5	7
Total	24	30	35

Output from Test-Sensitivity Analyzer

Inputs			Output		
Number of Respondents (n)	Probability of Correct Guess (p ₀)	Probability of a Correct Response @ p _d (p _{max})	Type I Error (α-risk)	Type II Error (β-risk)	Power (1 - β)
60	0.50	0.625	0.25	0.2595	0.9077

$$\begin{aligned} P_{\max(90\%)} &= [2(x/n) - 1] + z_{\beta} \sqrt{[4(x/n)(1 - (x/n))]/n} \\ &= [2(35/60) - 1] + 1.282 \sqrt{[4(35/60)(1 - (35/60))]/60} \\ &= [0.1667] + 1.282(0.1273) \\ &= 0.33, \text{ or } 33\% \end{aligned}$$

感官分析專家結論：有 90%信賴區間，能辨識調豆 D 和控制組之群體的真正比率可能大至 33%，超出原設定的臨界界限(Pd)的 25%，多出 8%。

感官分析專家可能有另外的顧慮，60 位回應者中僅 24 位正確答對調豆 B，在 60 位回應者的二-三點試驗法，當全部都用猜(Pd = 0)，正確選擇的期望數是 n/2 = 30，低於正確回應的期望數可能指出試驗調豆 B 期間，有些額外的因素左右回應者偏離做正確的選擇，例如，編碼錯誤或提供品評之前樣品的製備或處置不好。

輸入 spreadsheet test-sensitivity analyzer 的結果，專家知道：在本試驗所選定 α-危險率(大体為 0.25)下，要結論有顯著差異存在，需要 33 個正確回應，因此從 60 回應者試驗，32 或更少的正確回應是為足夠的相似性的證明。

解釋：

- 33 或更多的正確回應，是顯著水準 α = 0.26 下，有差異之證實。
- 32 或更少的正確回應，指你能 91%確定不超過 25%的品評員能測出差異，也就是，相對於在 β = 0.09 顯著水準下的 P_d = 25%，相似性之證實。

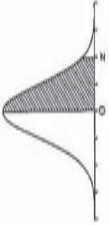
因此，結論為：試驗調豆 B 和 C 充分類似於控制組，擔保作進一步的考慮，但試驗調豆 D 不是，因有 35 個正確回答。對調豆 D (根據二-三點試驗法) 的辨識者之真正比率其 90%上界單尾信賴區間為

感官分析專家測試假設，正確回應的真正機率是至少 50% (虛無假設 null hypothesis；H₀: p ≥ 0.5) 相對於另一假設為低於 50% (alternative hypothesis: H_a: p < 0.5)，採用單尾信賴區間設定為 95% (即 α = 0.05，lower tail) 的二項式的常態近似值，試驗的統計式為

$$\begin{aligned} z &= [(x/n) - p_o] / \sqrt{p_o(1 - p_o)} / n \\ &= [(24/60 - 0.50)] / \sqrt{0.50(1 - 0.50)} / 60 \\ &= [-0.10] / (0.06455) \\ &= -1.55 \end{aligned}$$

查表 19.2 (注意 Pr [z < -1.55] = Pr (z > 1.55))，感官分析專家發現，試驗統計不大於 -1.55 所觀察的值其機率為 (0.5 - 0.4394) = 0.0606，此機率大於 α = 0.05 的值；故專家結論：並無足夠的證據拒 95%水準的絕虛無假設，24 位正確回應並非充分地偏離 30 位期望值，而讓我們下結論有額外的因素存在。

Table 19.2 Standard Normal Distribution



Instruction: See the Examples in Chapter 14.

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3415	0.3438	0.3461	0.3483	0.3505	0.3526	0.3547	0.3568	0.3588	0.3607
1.1	0.3625	0.3643	0.3661	0.3679	0.3695	0.3712	0.3729	0.3745	0.3761	0.3777
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4098	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4464	0.4474	0.4484	0.4494	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4685	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4908	0.4910	0.4912	0.4914
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4980	0.4981	0.4982
2.9	0.4982	0.4983	0.4984	0.4985	0.4986	0.4987	0.4988	0.4989	0.4990	0.4991
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4988	0.4989	0.4989	0.4990	0.4990

7.5 五取二試驗 (Two-out-of-Five Test)

7.5.1 範圍與應用

本方法在統計上的效率高，正確猜對 5 樣品中的 2 個，機會是 10 對 1，對照三角試驗的 3 對 1。但此法也受到感官疲勞(sensory fatigue)和記憶效應(memory effects)的強烈影響，因而其主要的用途在於視覺的、聽覺的與觸覺的應用，不用於風味測試。

當試驗目標是要確定是否兩樣品間存在感官差異，特別當只有少數的受試者可用時(例如，10 人)，採用本方法。

如同三角試驗，五取二試驗在某些情況是有效的：

- 1. 確定產品差異是否來自組成分、加工、包裝、或貯藏的變化
- 2. 確定是否整體差異存在，並無特定的屬性被認定受影響
- 3. 挑選和監控品評員辨識在感官疲勞效應小的試驗情況下所給予差異性之能力

7.5.4 試驗程序

有關試驗控制和產品控制，參照如 3.2 及 3.3 項所示。可以的話，同時提供樣品；但是，樣品如量多、或外觀上略有不同，可以先後依序提供，這樣不會影響有效性。如受試者超過 20 位，從下列組合中隨機挑選，且讓有 3 A's 與 3 D's 的組合數相等：

AAAB B ABABA BBBAA BABAB
AABAB BAABA BBABA ABBAB
ABAA B ABBAA BABBA BAABB
BAAAB BABAA ABBBA ABABB
AABBA BBAAB BBAAB AABBB

評分卷範例如圖 7.9。計算總回應數和正確回應數，參照表 19.14。不要計算“無差異”之回應；要求受試者如有疑惑時，就用猜的。

7.5.2 試驗原理

呈現給每位受試者 5 個編碼樣品，告知有兩個同屬同一種而其餘三個種是另外一種，要求他們品嚐(感覺、視、檢查) 從左至右的每樣品，並挑選出和其餘三個有所不同的兩個樣品。計算正確回答的數目，參照表 19.14 做解釋。

7.5.3 試驗受試者

如 3.3 項所示，挑選、訓練和指導受試者。通常，採用 10 至 20 位受試者，少至 5 至 6 位也可採用，是當差異大而容易找出，僅採用訓練的受試者。

Two-out-of-five test				
Name: _____		Date: _____		
Type of sample: _____				
Instructions				
1. Examine the samples from left to right. Two are of one type, and the other three of another.				
2. Identify the group of two samples by placing an X in the corresponding boxes.				
	Test 1	Test 2	Test 3	
Left	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
Right	_____	_____	_____	
Comments				
Left	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
Right	_____	_____	_____	

圖 7.9 – 三次五取二試驗的評分卷

範例 7.8: 比較布料的粗糙度

問題/情況 — 某紡織品製造商希望以聚乙烯/尼龍混織布取代現有的聚乙烯布料，但他收到一則抱怨，提及聚乙烯/尼龍混織布的表面較粗糙且有些扎刺感。

計畫目標 — 確定是否聚乙烯/尼龍混織布由於過於粗糙而需做修改。

試驗目標 — 取得兩種布料之間表面感覺的相對差異之測定。

試驗設計 — 由於感官疲勞並非考慮的因素，就評定差異，五取二試驗是最有效的，12 人品評小組也能夠測出相當小的差異。從前列的 20 個組合中，隨機選出這兩種布料之 12 個組合，詢問受試者：哪些樣品感覺相同但不同於其餘的三種？

使用的評分卷，如圖 7.11。

Two-out-of-five test		Test code
Name:		
Type of sample:	Date:	
Type of difference:		
Instructions		
1. Examine the samples in the order listed below. Two are of one type and the other three of another. Feel the surface gently with fingers or palm of hand.		
2. Identify the two samples that feel the same by placing an X in the corresponding boxes.		
Codes	x	Comments
	☐	
	☐	
	☐	
	☐	

圖 7.11 – 五取二試驗用評分卷。範例 7.8：比較織物的粗糙度

進行試驗 — 將布料夾入遮蔽紙板架內，直線面向每位品評員（見圖 7.10），必須讓品評員能夠感觸布料但看不到布料，從隨機三位數字碼（見表 19.1）的列表，分派給樣品編碼。

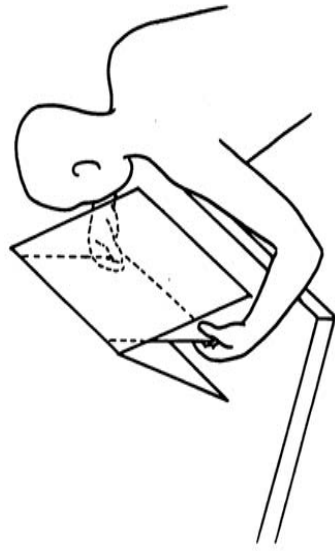


圖 7.10 – 五取二試驗。範例 7.8：品評員面前布料樣品的陳列

分析結果 — 12 位受試者之中，有 9 位能夠正確找出兩群布料樣品，參照表 19.14，顯示表面感覺上的差異是可測出的，在 $\alpha = 0.001$ 顯著水準。

解釋結果 — 告知織品製造商這兩種布料的表面感覺的差異很容易被察覺出來。

範例 7.9：面霜中的潤膚劑

- 問題/情況 — 由於明顯節省製造成本，有需要更換面霜配方中的潤膚劑，但這樣的取代似也減低產品的表面光澤。
- 計畫目標 — 市場部門希望在到消費者確定任何對接受性的影響之前，確定是否兩種配方間存在可看出的差異。
- 試驗目標 — 確定兩種面霜配方之間，外觀上是否存在統計學上顯著的差異。

Date	3-05-99	WORKSHEET	Test code	TO-AF88
Post this sheet in the area where trays are prepared. Code sheets ahead of time. Label serving containers ahead of time.				
Type of samples :	Face cream for Juehwa			
Type of test :	Two-out-of-five test			
Sample Identification	Code			
Px-2316 (control)	A			
Px-2602 (new emollient)	B			
Arrange samples as follows in the front of each subject:				
Judge No.	Order of Samples			
1	A	A	B	B
2	A	B	A	B
3	B	A	B	B
4	B	A	B	A
5	B	B	A	A
6	B	A	A	A
7	A	A	A	B
8	A	B	A	A
9	A	A	B	B
10	A	A	B	B

圖 7.12 – 五取二試驗的工作表。範例 7.9：面霜中的潤膚劑，察看樣品排列

分析結果 — 5 位受試者正確地分群樣品，從表 19.14，相當於 1% 顯著水準。

表 19.14 – 在五取二試驗的正確回應臨界數(輸入為 $\chi_{\alpha,n}$)

填入數為所述 α -水準(欄)下回應者的對應數 n (列)達顯著所需的最低正確回應數。如果正確回應數大於或者等於表列的值，拒絕“無差異”假定。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
3	1	1	2	2	2	3	3	31	4	5	5	6	7	8	10
4	1	2	2	2	3	3	4	32	4	5	6	6	7	9	10
5	2	2	2	2	3	3	4	33	5	5	6	7	7	9	11
6	2	2	2	3	3	4	5	34	5	5	6	7	7	9	11
7	2	2	2	3	3	4	5	35	5	5	6	7	8	9	11
8	2	2	2	3	3	4	5	36	5	5	6	7	8	9	11
9	2	2	3	3	4	4	5	37	5	6	6	7	8	9	11
10	2	2	3	3	4	5	6	38	5	6	6	7	8	10	11
11	2	3	3	3	4	5	6	39	5	6	6	7	8	10	12
12	2	3	3	4	4	5	6	40	5	6	7	7	8	10	12
13	2	3	3	4	4	5	6	41	5	6	7	8	8	10	12
14	3	3	3	4	4	5	7	42	6	6	7	8	9	10	12
15	3	3	3	4	5	6	7	43	6	6	7	8	9	10	12
16	3	3	4	4	5	6	7	44	6	6	7	8	9	11	12
17	3	3	4	4	5	6	7	45	6	6	7	8	9	11	13
18	3	3	4	4	5	6	8	46	6	7	7	8	9	11	13
19	3	3	4	5	5	6	8	47	6	7	7	8	9	11	13
20	3	4	4	5	5	7	8	48	6	7	8	9	9	11	13
21	3	4	4	5	6	7	8	49	6	7	8	9	10	11	13
22	3	4	4	5	6	7	8	50	6	7	8	9	10	11	14
23	4	4	4	5	6	7	9	51	7	7	8	9	10	12	14
24	4	4	5	5	6	7	9	52	7	7	8	9	10	12	14
25	4	4	5	5	6	7	9	53	7	7	8	9	10	12	14
26	4	4	5	6	6	8	9	54	7	7	8	9	10	12	14
27	4	4	5	6	6	8	9	55	7	8	8	9	10	12	14
28	4	5	5	6	7	8	10	56	7	8	8	9	10	12	14
29	4	5	5	6	7	8	10	57	7	8	8	9	10	11	15
30	4	5	5	6	7	8	10	58	7	8	9	10	11	13	15
								59	7	8	9	10	11	13	15
								60	7	8	9	10	11	13	15

Note: For values of n not in the table compute $z = (k - 0.1n) / \sqrt{0.09n}$, where k is the number of correct responses. Compare the value of z to the α -critical value of standard normal variable, that is, the values in the last row of Tables 19.3 ($z_{\alpha} = t_{\alpha,n}$).

試驗設計/篩選樣品 — 採用經過色盲和視力缺損篩選的 10 位

受試者，白熾燈下，測試表玻璃上 2 毫升樣品，並以白色為背景。樣品先做預備試驗，確定其表面在暴露 30 分鐘內(每次試驗的最長時間)不會變化，例如乾涸、流動、變色等。

進行試驗 — 按照工作表(見圖 7.12)所示的計畫，樣品從左至右排列成一直線；使用的評分券類似於圖 7.11 所示。詢問受試者“辨認哪兩樣品是有同樣的外觀，並和其餘三個樣品不同”。

7.6 同/異試驗(Same/Different Test)或單純差異試驗(Simple Difference Test)

7.6.1 範圍與應用

當試驗目標是要確定兩產品間是否存在感官差異，特別當不適合於三重或多重呈現時，例如當三角和二-三點試驗不能使用。這樣情況的例子譬如帶有強烈或持久風味的樣品之間的比較，半臉測試需要敷在皮膚上的樣品，以及具有很複雜的刺激並造成品評員心理上的迷惑。

如同其它整體差異試驗，這同/異試驗在下列情況是有效用的：

1. 確定是否產品差異是來自組成分、製程、包裝、或貯藏的改變。
2. 確定是否有整體差異存在，即並無特定的屬性確認被受影響。

7.6.2 試驗原理

給每位受試者 2 個樣品(一配對 a pair)，詢問這些樣品是相同或不同。一半的配對組是呈現 2 個不同的樣品，另一半配對組是 2 個匹配的樣品(同一種樣品，重複)。比較匹配對組的不同回應數目與相異對組的不同回應數目，採用卡方試驗(χ^2 -test)檢定。

7.6.3 試驗受試者

為了確定差異，通常需要 4 種樣品組合(A/A、B/B、A/B、B/A)，每種組合各呈現 20 至 50 次，採用多至 200 位受試者，或 100 位受試者各給 2 個對組。如果同/異試驗是因為刺激物的複雜性而被選定，在同一時間內不要多於 1 個對組呈現給任何一位受試者，受試者可以是以訓練或未受訓練者，但品評員不可混合組成。

表 19.14 – 在五取二試驗的正確回應臨界數(輸入為 $X_{\alpha,n}$)
填入數為所述 α -水準(欄)下回應者的對應數 n (列)達顯著所需的最低正確回應數。如果正確回應數大於或者等於表列的值，拒絕“無差異”假定。

α								α							
n	0.40	0.30	0.20	0.10	0.05	0.01	0.001	n	0.40	0.30	0.20	0.10	0.05	0.01	0.001
3	1	1	2	2	2	3	3	31	4	5	5	6	7	8	10
4	1	2	2	2	3	3	4	32	4	5	6	6	7	9	10
5	2	2	2	2	3	3	4	33	5	5	6	7	7	9	11
6	2	2	2	3	3	4	5	34	5	5	6	7	7	9	11
7	2	2	2	3	3	4	5	35	5	5	6	7	8	9	11
8	2	2	2	3	3	4	5	36	5	5	6	7	8	9	11
9	2	2	3	3	4	4	5	37	5	6	6	7	8	9	11
10	2	2	3	3	4	5	6	38	5	6	6	7	8	10	11
11	2	3	3	3	4	5	6	39	5	6	6	7	8	10	12
12	2	3	3	4	4	5	6	40	5	6	7	7	8	10	12
13	2	3	3	4	4	5	6	41	5	6	7	8	8	10	12
14	3	3	3	4	4	5	7	42	6	6	7	8	9	10	12
15	3	3	3	4	4	5	7	43	6	6	7	8	9	10	12
16	3	3	4	4	5	6	7	44	6	6	7	8	9	11	12
17	3	3	4	4	5	6	7	45	6	6	7	8	9	11	13
18	3	3	4	4	5	6	8	46	6	7	7	8	9	11	13
19	3	3	4	5	5	6	8	47	6	7	7	8	9	11	13
20	3	4	4	5	5	7	8	48	6	7	8	9	9	11	13
21	3	4	4	5	6	7	8	49	6	7	8	9	10	11	13
22	3	4	4	5	6	7	8	50	6	7	8	9	10	11	14
23	4	4	4	5	6	7	9	51	7	7	8	9	10	12	14
24	4	4	5	5	6	7	9	52	7	7	8	9	10	12	14
25	4	4	5	5	6	7	9	53	7	7	8	9	10	12	14
26	4	4	5	6	6	8	9	54	7	7	8	9	10	12	14
27	4	4	5	6	6	8	9	55	7	8	8	9	10	12	14
28	4	5	5	6	7	8	10	56	7	8	8	10	10	12	14
29	4	5	5	6	7	8	10	57	7	8	9	10	11	12	15
30	4	5	5	6	7	8	10	58	7	8	9	10	11	13	15
30	4	5	5	6	7	8	10	59	7	8	9	10	11	13	15
30	4	5	5	6	7	8	10	60	7	8	9	10	11	13	15

Note: For values of n not in the table compute $z = (k - 0.1n)/\sqrt{0.09n}$, where k is the number of correct responses. Compare the value of z to the α -critical value of standard normal variable, that is, the values in the last row of Tables 19.3 ($z_{\alpha} = t_{\alpha,n}$).

這方法有些費時，因關於可能的產品差異之訊息，是來自比較從相異對組(different pairs; A/B 與 B/A)和從匹配對組(matched pairs; A/A 與 B/B)的回應，匹配對組的呈現能使感官分析專家就單純詢問差異之問題時，評定安慰劑效應(placebo effect)的幅度。

7.6.4 試驗程序

有關試驗控制和產品控制，參照 3.2 和 3.3 項。如果可以的話，同時提供樣品，或者依次前後提供。4 個配對組準備相同的數量，如果每人只品評 1 個對組，隨機分配給受試者；如果試驗設計為每人品評多於 1 個對組(一個相異對組和一個匹配對組，或者全部 4 個組合)。然後，保留每位受試者的試驗得分紀錄。常用的工作表和評分卷列於範例 7.10。

7.6.5 分析與解釋結果
參見範例 7.10。

試驗設計 – 烤肉醬有辣味，品嚐時會引起滯延效應 (carryover effects)，因此，也用平淡無味的載體譬如白土司，進行同/異試驗是一項合適的試驗。收集自 60 位受試者之 60 回應總數，30 個是匹配對組(matched pairs)和 30 個是非匹配對組(unmatched pairs)。在一次品評會期，每位受試者品評 1 個匹配對組(A/A 或 B/B)或者 1 個相異對組(A/B 或 B/A)，此試驗的工作表和評分卷示於圖 7.13 及圖 7.14。試驗在品評小間(booth)區域進行，並配紅色光線以遮蔽任何顏色的差異。

範例 7.10：更換烤肉醬製造的加熱釜

問題/情況 – 在試圖將調味醬工廠現代化，廠家必須更換生產烤肉醬的老舊加熱釜，工廠經理想知道如果新加熱釜所做的產品嚐起來是否如同老機器做的。

計畫目標 – 確定新加熱釜能取代舊機器而加入工廠的生產行列。

試驗目標 – 確定不同加熱釜生產的兩種烤肉醬產品，在滋味上可辨識出來。

Date	2-26-99	Worksheet	Test code	84-46F09
Post this sheet in the area where trays are prepared. Code scoresheets ahead of time. Label serving containers ahead of time.				
Type of samples :	<u>Barbecue sauce on white bread pieces</u>			
Type of test :	<u>Same/Different test</u>			
Sample identification				
5-117-36 (old cooker)			Code	
5-117-39 (new cooker)			36	
			39	
Code serving containers with 3-digit random numbers and divide into two lots, one lot to receive sample 36, the other sample 39.				
When preparing panelists' trays, place samples from left to right in the following order :				
Panelist code				Sample order
1 - 15				36 - 36
16 - 30				36 - 39
31 - 45				39 - 36
46 - 60				39 - 39

圖 7.13 – 同/異試驗的工作表。範例 7.10：更換烤肉醬製造的加熱釜

Same/different test		Test no. 84-4639
Taster no. _____	Name: _____	Date: _____
Type of sample: <u>Barbecue sauce on white bread pieces</u>		
Instructions		
1. Taste the two samples from left to right.		
2. Determine if samples are the same/identical or different.		
3. Mark your response below.		
Note that some of the sets consist of two identical samples.		
_____ Products are the same		
_____ Products are different		
Comments: _____		

圖 7.14 – 同/異試驗的評分卷。範例 7.10：更換烤肉醬製造的加熱釜

	Subjects Received			欄： 試驗 樣品 數
	Matched Pair AA or BB	Unmatched Pair AB or BA	Total	
Subjects said	17	9	26	
Same	13	21	34	
Different	30	30	60	
Total				

採用卡方(chi-analysis)分析(參見 14.3 項，範例 14.6)，比較安慰劑效應(17/13)與處理效應(9/21)。計算如下：

$$\chi^2 = \sum \frac{(O - E)^2}{E},$$

O 觀察數；E 預期數(same/matched、same/unmatched、different/matched、different/unmatched 等框架)的每框架。舉例，框架 same/matched 的計算：

$$E = (26 \times 30)/60 = 13, \text{ and,}$$
$$\chi^2 = \frac{(17 - 13)^2}{13} + \frac{(9 - 13)^2}{13} + \frac{(13 - 17)^2}{17} + \frac{(21 - 17)^2}{17} = 4.34,$$

篩選樣品 — 先用 5 位經驗型品嚐員做預備實驗，確定樣品不加料或塗在載體如白吐司上何者較容易品嚐，載體使得比較更容易、也不會導入額外的感官因素。預試驗也有助於確定試驗中相對於土司(大小計)，適合的樣品量(重量或體積計)。

進行試驗 — 受試者品嚐之前，將預秤重的醬料塗在切好的吐司片(先貯存於密封罐冷藏)上，依照工作表中為每位品評員所列的順序，將樣品擺在編碼的盤上。

分析結果 — 下表中，欄部分指試驗的樣品數目；列部分指有少數目是被受試者確認：

	Subjects Received		
	Matched Pair AA or BB	Unmatched Pair AB or BA	Total
Subjects said	17	9	26
Same	13	21	34
Different	30	30	60
Total			

same/matched : $E = (26 \times 30)/60 = 13$

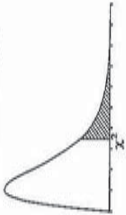
same/unmatched: $E = (26 \times 30)/60 = 13$

different/matched: $E = (34 \times 30)/60 = 17$

different/unmatched: $E = (34 \times 30)/60 = 17$

卡方值大於表 19.5 中(df = 1, probability = 0.05, $\chi^2 = 3.84$)所示的值，即顯著差異存在。

Table 19.5 Upper-α Probability Points of χ^2 -distribution (Entries are $\chi^2_{\alpha, v}$)



Instructions:

(1) Enter the row of the table corresponding to the number of degrees of freedom (v) for χ^2 .

(2) Pick the value of χ^2 in that row, from the column that corresponds to the predetermined α -level.

	0.995	0.990	0.975	0.950	0.900	0.800	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
v	0.0000993	0.000157	0.000262	0.000393	0.000582	0.000857	0.00118	0.00171	0.00253	0.00377	0.00541	0.00756	0.0106	0.0143
1	0.0000993	0.000157	0.000262	0.000393	0.000582	0.000857	0.00118	0.00171	0.00253	0.00377	0.00541	0.00756	0.0106	0.0143
2	0.0100	0.0201	0.0384	0.0541	0.0717	0.0900	0.106	0.136	0.175	0.216	0.260	0.302	0.354	0.405
3	0.0717	0.115	0.216	0.352	0.584	0.853	1.10	1.36	1.63	1.92	2.20	2.48	2.77	3.08
4	0.207	0.297	0.484	0.711	1.06	1.52	1.92	2.36	2.78	3.18	3.58	3.94	4.28	4.61
5	0.412	0.554	0.831	1.15	1.61	2.20	2.67	3.18	3.66	4.10	4.55	4.96	5.40	5.82
6	0.676	0.872	1.24	1.69	2.28	2.83	3.35	3.85	4.31	4.75	5.19	5.61	6.05	6.46
7	0.989	1.24	1.69	2.17	2.73	3.25	3.74	4.21	4.64	5.06	5.48	5.89	6.32	6.73
8	1.34	1.65	2.18	2.70	3.23	3.74	4.17	4.59	4.99	5.39	5.79	6.18	6.60	7.00
9	1.73	2.09	2.70	3.23	3.74	4.17	4.59	4.99	5.39	5.79	6.18	6.57	6.97	7.36
10	2.16	2.56	3.25	3.94	4.67	5.41	6.17	6.94	7.71	8.49	9.24	9.99	10.7	11.4
11	2.60	3.05	3.82	4.57	5.38	6.21	7.06	7.91	8.68	9.46	10.2	10.9	11.6	12.3
12	3.07	3.57	4.40	5.23	6.13	7.04	7.98	8.84	9.61	10.3	11.0	11.7	12.4	13.1
13	3.57	4.11	5.01	5.89	6.86	7.84	8.84	9.71	10.4	11.1	11.8	12.5	13.2	13.9
14	4.07	4.66	5.63	6.57	7.59	8.64	9.71	10.6	11.3	12.0	12.7	13.4	14.1	14.8
15	4.60	5.23	6.26	7.26	8.34	9.46	10.6	11.6	12.3	13.0	13.7	14.4	15.1	15.8
16	5.14	5.81	6.91	7.96	9.11	10.3	11.5	12.5	13.2	13.9	14.6	15.3	16.0	16.7
17	5.70	6.41	7.56	8.67	9.86	11.1	12.3	13.3	14.0	14.7	15.4	16.1	16.8	17.5
18	6.26	7.01	8.23	9.39	10.6	11.9	13.1	14.1	14.8	15.5	16.2	16.9	17.6	18.3
19	6.84	7.63	8.91	10.1	11.4	12.7	13.9	14.9	15.6	16.3	17.0	17.7	18.4	19.1

(Continued)

受試者給 A/B 或 B/A 和回應為

受試者給 A/A 或 B/B	Same	Same	Same	Different
和回應為	a = 2	b = 15	c = 7	d = 6

試驗統計量為 $McNemar's T = (b - c)^2 / (b + c)$

對於 $(b + c) \geq 20$ ，無差異的假設被拒絕，如果 T 值大於表 19.5 中的 χ^2 (自由度 1)，對於 $(b + c) < 20$ ，應用二項式步驟(參見 Conover, 1980)。

目前的範例：

$$McNemar's T = (15 - 7)^2 / (15 + 7) = 2.91$$

其值小於 $\chi^2_{1,0.05} = 3.84$ ，故我們不能結論這些樣品是不同的。

如果處理來自 30 位品評員的配對數據如同來自 60 位品評員個別的觀察，我們會得到如 分析結果 項範例 7.6 之數據，這標準的卡方分析導致樣品間存在有統計上顯著差異之錯誤結論。

解釋結果 — 結果顯示兩種不同加熱釜所製的烤肉醬之間有顯著的差異，感官分析專家告知工廠經理新加熱釜的供應商的宣稱不是真實的，兩產品之間測出差異；專家提議如果新加熱釜的更換在工廠有其重要的成本/效率的考量，兩種烤肉醬應該以使用者測試嗜好(喜好/偏好)，若消費者試驗結果為兩種醬料同位或偏好新加熱釜的醬料，就允許工廠實施更換的過程。

註： 假設範例 7.10 是用 30 位而非 60 位受試者，每人在分開的會期都品評一個匹配對組和一個相異對組，結果可能如上面一樣，但卡方分析不合適而應採用 McNemar test (Conover, 1980)。為執行 McNemar 檢定程序，專家須追蹤從每位品評員的兩個回應，依下列公式計算：

7.7 A 或非 A 試驗("A" – "NOT A" Test)

7.7.1 範圍與應用

當試驗目標要確定兩樣品間是否有感官的差異存在，尤其那些不適合以 2 或 3 個呈現，例如當不能使用二-三點試驗和三角試驗。

其情況的範例為帶有強烈及/或滯留的風味之產品的比較，在半臉試驗(half-face tests)須用在皮膚上之樣品，外觀上些許不同的樣品，以及很複雜的刺激且使品評員心理上易生混淆之樣品。

採用 A 或非 A 試驗而優先於同/異試驗，當兩產品其中之一有其作為標準品或參考產品之重要性，受試者所熟識，或要測定目前樣品對所有其他樣品而為計畫所必要的。

如同其它整體差異試驗，在下列情況下 **A 或非 A 試驗**是有效用的：

1. 確定是否產品差異來自於組成分、製程、包裝或貯藏中的改變。
2. 確定是否整體差異存在，並無特定的屬性被認定已受影響。

本試驗也適用於**品評員篩選**，例如，確定是否有試驗受試者(或一群受試者)從相對於其它甜味劑而識別出某特定甜味劑，以及能用以訊號偵測法(Signal Detection method)確定感官閾值。

7.7.2 試驗原理

讓品評員熟悉樣品“A”和“非 A”。呈現樣品給品評員，其中有些是產品“A”，而其餘為產品“非 A”；就每一樣品，受試者判定是“A”，還是“非 A”。通過**卡方試驗**比較正確的識別和不正確的識別，確定受試者的辨識能力。

註：在該程序的標準版本中，可看到下列的協定：

1. 產品“A”和“非 A”在試驗開始之前才呈現給受試者。
2. 每次的試驗只有一個“非 A”樣品。
3. 每次的試驗 “A”和“非 A”樣品數量相同。

這些協定可能在任何給定的試驗中會被改變，但試驗開始之前必須告知受試者。在第 2 條，如果“非 A”樣品超過 1 個，在試驗前每一個都要現給受試者知道。

7.7.5 結果分析與解釋

4 種不同的樣品組合相對於回應之數據分析有些繁瑣，可透過**範例 7.11** 獲得最好的瞭解。

7.7.3 試驗受試者

訓練 **10** 至 **50** 位受試者認定 “A” 和 “非 A” 樣品。研究中每個樣品採用 **20** 至 **50** 次呈現，每位受試者可給予只一個樣品(“A” 或“非 A”)、兩個樣品(一個“A”和一個“非 A”)、或者每位受試者在一系列中測試至 10 個樣品。允許的樣品數取決於樣品所產生對受試者的身體及/或精神上疲勞的程度。

註： 本方法的一個變種，受試者未熟識“非 A”樣品，但並不推薦這樣的做法，此乃欠缺參考品的框架，受試者可能到處用猜的，而產生偏見的結果。

7.7.4 試驗程序

有關試驗控制和產品控制，參照 **3.2** 和 **3.3** 項。同時呈現樣品和評分卷，樣品都以亂數編碼，隨機提出，不能讓受試者察覺出在任何系列中 “A”樣品 對照 “非 A”樣品的樣式。

在完成試驗系列之前，不要透露樣本的身份。

範例 7.11：新鮮甜味劑和蔗糖比較

問題/情況 — 一位產品開發化學家正在為目前配方使用 5%蔗糖的飲料尋找代用的甜味料，預備的滋味試驗已建立 0.1%新甜味料的甜度等同於 5%蔗糖，但也得知若同一時間呈現一個以上的樣品，辨識上會遇到甜味及其它滋味的遲滯(carryover)以及口感因素的影響，化學家希望知道是否這兩種飲料因滋味而可辨別的。

- 計畫目標** — 確定代用甜味料 0.1%能否用來取代 5%蔗糖。
- 試驗目標** — 減低遲滯和疲勞效應，直接比較兩種甜味料。

試驗設計 — **A-非 A 試驗**允許樣品以非直接方式比較，並允許受試者清楚認識新甜味劑的風味。新甜味劑 0.1%溶液當作“A”，重複給受試者嚐，5%蔗糖溶液則為“非 A”；20 位受試者在 20 分鐘的試驗期間每位給 10 個樣品，一次品嚐 1 個樣品，記錄回應(“**A**”或“**非 A**”)，用白開水漱口，等待 1 分鐘後，再嚐下一個樣品。**圖 7.15** 為試驗工作表，**圖 7.16** 為評分卷。

"A"..."Not A" Test						Test code
Taster no: _____		Name: _____		Date: _____		
Type of sample: _____		Sweetened beverage				
Instructions						
1. Before taking this test, familiarize yourself with the flavor of the samples "A" and "Not A" which are available from the attendant.						
2. Taste the test samples from left to right. After each sample, record your response below, rinse your palate with water, and wait one full minute between samples.						
Note: You have received approximately equal numbers of "A" and "Not A" samples.						
Sample No.	Code	The sample is: "A" "Not A"	Sample No.	Code	The sample is: "A" "Not A"	
1	_____	☐ ☐	6	_____	☐ ☐	
2	_____	☐ ☐	7	_____	☐ ☐	
3	_____	☐ ☐	8	_____	☐ ☐	
4	_____	☐ ☐	9	_____	☐ ☐	
5	_____	☐ ☐	10	_____	☐ ☐	
Comments: _____						

圖 7.15 – “A”–“非 A”試驗用評分卷。範例 7.11：新甜味料和蔗糖比較

Date	1-15-99	Worksheet	Test code	612A83					
Post this sheet in the area where trays are prepared. Code scoresheets ahead of time. Label serving containers ahead of time.									
Type of samples: _____	Sweetened beverage								
Type of test: _____	"A" – "Not A" test								
Sample Identification									
Beverage with 0.1% sweetener ("A")		Code A							
Beverage with 5% sucrose ("Not A")		B							
Code 200 6-oz cups with random 3-digit numbers and divide into two lots of 100 each. Use sample "A" for the first 100 cups and sample "Not A" for the second 100 cups. When preparing panelists' trays, place samples from left to right in the following order:									
Panelist									
Sample order									
1 - 5	A	A	B	A	B	A	B	B	A
6 - 10	B	A	B	A	A	B	A	B	B
11 - 15	A	B	A	B	B	A	B	A	A
16 - 20	B	B	A	A	B	A	B	A	B

圖 7.15 – “A”–“非 A”試驗。範例 7.11：新甜味料和蔗糖比較

分析結果 — 下表中，欄：所呈現的樣品，列：受試者認出：

Subject Received			Total		
		A	Not A		
Subject said	A	60	35	95	
	Not A	40	65	105	
	Total	100	100	200	

The χ^2 -statistic is calculated as in Section 6.6:

$$\chi^2 = \frac{(60 - 47.5)^2}{47.5} + \frac{(35 - 47.5)^2}{47.5} + \frac{(40 - 52.5)^2}{52.5} + \frac{(65 - 52.5)^2}{52.5} = 12.53,$$

框架 same/matched : $E = (95 \times 100) / 200 = 47.5$
 same/unmatched : $E = (95 \times 100) / 200 = 47.5$
 different/matched : $E = (105 \times 100) / 200 = 52.5$
 different/unmatched : $E = (105 \times 100) / 200 = 52.5$

所得的 $\chi^2 = 12.53$ ，大於表 19.5 (df = 1, α -risk = 0.05, $\chi^2 = 3.84$)，即顯著差異存在。

註：剛列出的卡方分析並非完全適合的，因為每位回應者都做**多重**的評定，但是，目前沒有計算上便利的代替方法可用，從本試驗得到的顯著水準應考慮其適當的值。

解釋結果 — 所得結果指出，0.1%甜味溶液顯著地不同於 5% 蔗糖溶液，感官分析專家告知開發化學家，這特別的甜味料很可能引起飲料風味上可測出的改變，接下來可能是做描述性分析 (descriptive analysis)，為了表徵差異為何。

有人可能會問：怎麼樣使差異變成不顯著的？如果結果是如此，情況就是這樣：

60	50
40	50

其 χ^2 等於 2.02，小於 3.84。參見 ISO (1987) 中更多類似的範例。

由於疲勞或遲滯效應，使得多重樣品試驗不合適的情況下，與控制組差異試驗可用在二樣品試驗。與控制組差異試驗本質上是一個單純差異試驗，多加了差異大小的評估。

7.8.2 試驗原理

呈現給每位受試者 1 個控制組樣品加上 1 個或更多的樣品，要求受試者 **評級 (rate)** 每個樣品和控制組之間的差異的大小，為此並提供一個尺度。透過比較它們和以**盲控制組 (blind controls)** 取得的與控制組差異，評估所產生的平均與控制組差異估量*。

* 由盲控制組所取得的估量，之所以使用是要取得**安慰劑效應 (placebo effect)** 的數量。這估量代表單純詢問多少的差異問題時的數字效應 (numerical effect)，當事實上並無差異存在。

7.8 與控制組差異試驗 (Difference-from-control Test)

7.8.1 範圍與應用

當計畫或試驗目標是雙重的，(1) 確定一個或更多樣品和控制組之間是否有差異存在、(2) 推估這樣差異的大小幅度。通常，一個樣品指定為控制組 “control”、參考品 “reference” 或標準品 “standard”，品評其餘所有樣品和控制組之間**如何的不同 (how different)**。

在差異可能是可測出的、但差異大小會影響試驗目標的決定之情況下，**與控制組差異試驗**是有用的。品質保證/品質控制和貯藏研究就是案例，和控制組的差異的相對大小對於決策是重要的，在二-三點和三角試驗不能夠使用時，**與控制組差異試驗**是合適的，因為產品的正常非均一性 (normal heterogeneity)，譬如肉類、沙拉類及烘焙食品。

7.8.3 試驗受試者

通常，每樣品呈現 20 至 50 次，標記為控制組之盲控制組 (blind control) 是要確定差異的程度 (degree of difference)，如果是因為複雜的比較或疲勞因素，而選擇了**與控制組差異試驗**，在同一時間應不超過 1 個配對樣品提供給任何一位受試者。受試者可以是訓練型或未訓練者，但品評小組不要由兩種人組成，所有的受試者要熟悉試驗格式、尺度的涵義、和部分的試驗樣品成為盲控制組。

7.8.4 試驗程序

關於試驗管制及產品管制，參見 3.2 及 3.3 項。可能的話，同時提供樣品，先品評標記為控制組。為每位受試者準備 1 個標記為控制組樣品 (control sample)，加上額外的控制組但標記為試驗樣品 (test sample)。如果試驗是設計要所有的受試者最終都測試所有的樣品，但無法在一個會期完成，必須保留按樣品記載的受試者記錄，確保剩餘的樣品在隨後的會期呈現出來。

使用的尺度可以是 5.6 項中討論的任何一種，例如，

言詞類別尺度	數字類型尺度
No difference	0 = No difference
Very slight difference	1
Slight/moderate difference	2
Moderate difference	3
Moderate/large difference	4
Large difference	5
Very large difference	6
	7
	8
	9 = Very large difference

(當計算言詞類型尺度的結果，轉換為對面的數字，即 large difference = 5)

7.8.5 結果分析與解釋

計算每樣品和盲控制組的與控制組差異平均(mean difference-from-control)，透過變異數分析(或 paired t-test，如果只有一個樣品和控制組比較) 評定結果，如範例中所示。

試驗設計 一預秤重的每樣品放在編號的表玻璃片上，每個樣品都秤同樣的重量(產品重量是正常塗在 10 cm² 面積)，受試者的手背上畫出 10 cm² 面積的線跡。試驗採用 42 為受試者，每會期需連續的 3 天，在每一天，受試者觀看一配對組，可能是

- 控制組 vs. 產品 F
- 控制組 vs. 產品 N
- 控制組 vs. 盲控制組(blind control)

(參見工作表，圖 7.17)，所有受試者都先給標示控制組之樣品，以及第二個為試驗樣品。受試者坐品評小間內，室內通風良好，以減少氣味積累，並光線充足，讓視覺上提示能助於評定。

範例 7.12：止痛膏 — 黏度提高

問題 / 情況 — 一家製藥公司的家庭健康照護部門計畫要提高止痛膏基底的黏度，兩種建議原型其質地在儀器上比控制組更為濃稠，樣品 F 需要較大力道使其流動/移動，而樣品 N 最初較易流動但其整體的黏度較高。產品研究人員想知道這兩樣品和控制組是如何不同，這類試驗最好是在手背上面做，並限定兩樣品同時間品評。

計畫目標 — 確定樣品 F 或者樣品 N 其整體上是最接近於目前產品。

試驗目標 — 測定兩原型和常規止痛膏之間的可感知整體感官差異。

Date	10-2-98	WORKSHEET	No.	13-625
Post this sheet in the area where trays are prepared. Code scoresheets ahead of time. Label serving containers ahead of time.				
Type of samples : Analgesic cream				
Type of test : Difference from Control test				
Sample	Description	Sample Code		
C	Control	C		
F	Experimental 10A3 (thixotropic)	Random #s under 500		
N	Experimental 2-6X (high viscosity)	Random #s over 500		
Serve in the following order:				
Subject #	Day 1	Day 2	Day 3	
1 - 7	C - F	C - N	C - C	
8 - 14	C - N	C - F	C - C	
15 - 21	C - F	C - C	C - N	
22 - 28	C - N	C - C	C - F	
29 - 35	C - C	C - N	C - F	
36 - 42	C - C	C - F	C - N	
Hour	Subject #			
9:00	1,8,15,22,29,36			
9:45	2,9,16,23,30,37			
10:30	3,10,17,24,31,38			
11:15	4,11,18,25,32,39			
1:00	5,12,19,26,33,40			
1:45	6,13,20,27,34,41			
2:30	7,14,21,28,35,42			

圖 7.17 — 與控制組差異試驗的工作表。範例 7.12：止痛膏

進行試驗 — 每次試驗在 15 分鐘內秤完樣品，將要呈現的兩樣品貼上三位數編碼，採用易於擦拭的標記，在每位受試者的手背上畫上 10-cm² 面積的線跡，告訴受試者小心依循評分卷上的指示(見圖 7.18)。

圖 7.18 – 與控制組差異試驗的評分表。範例 7.12：止痛膏

Difference-from-control test

Name: _____ Date: _____

Type of sample: _____ Code of test sample _____

Instructions

1. You have received two samples, a control sample labeled C and a test sample labeled with a three-digit number.

2. Remove all of the control sample from the watch glass using your right Index and middle fingers.

3. Using the index and middle fingers, spread the control product around the area traced on the back of your left hand.

4. Wipe finger tips with cloth on tray.

5. Pick up all of the test sample from the labeled watch glass using your left index and middle fingers.

6. Using the index and middle fingers, spread the product across the area traced on your right hand.

7. Indicate the size of the difference in skinfeel of the sample, relative to the control, on the scale below.

0 = no difference

1 =

2 =

3 =

4 =

5 =

6 =

7 =

8 =

9 =

10 = extreme difference

Remember that a duplicate control is the sample some of the time.

Comments : _____

國立台灣海洋大學食品科學系：食品品評學上課資料 食品感官品評 第 7 章 整體差異試驗：邱思魁整理 (47/74)

分析結果 — 所得結果列表 7.1，採用適合於隨機完全區集設計 (randomized (complete) block design) 的變異數分析(ANOVA 或 AOV)步驟來分析數據。在設計中，42 位判定為區集 “blocks”，三個樣品為處理 “treatments” (或更適當地，是處理的三階 the three levels of the treatment)。(參見 14.5 項，ANOVA 及區集設計的一般討論)

表 7.1 – 範例 7.12：與控制組差異試驗 — 止痛膏的結果

Judge	Blind Control	Product F	Product N	Judge	Blind Control	Product F	Product N
1	1	4	5	22	3	6	7
2	4	6	6	23	4	5	6
3	1	4	7	24	4	6	6
4	4	8	7	25	0	3	3
5	2	4	3	26	2	5	1
6	1	4	5	27	2	5	5
7	3	3	6	28	2	6	4
8	0	2	4	29	3	5	6
9	6	2	9	30	1	4	7
10	7	7	2	31	4	4	7
11	0	1	9	32	1	4	5
12	1	3	6	33	3	5	4
13	4	5	7	34	1	4	4
14	1	6	5	35	4	6	6
15	1	7	5	36	2	3	3
16	2	6	6	37	4	4	6
17	2	6	7	38	0	4	4
18	4	4	7	39	4	8	7
19	0	5	4	40	0	5	6
20	5	3	5	41	1	5	5
21	2	3	3	42	3	4	4

國立台灣海洋大學食品科學系：食品品評學上課資料 食品感官品評 第 7 章 整體差異試驗：邱思魁整理 (48/74)

表 7.2：本試驗的統計量結果。總變異 total variability 劃分為三個獨立的變異來源，即由於品評員之間的變異(區集效應 block effect)，由於樣品之間的變異(所興趣的處理效應 treatment effect)，及已交代的兩變異來源之後，仍存在不明的變異(實驗誤差 experimental error)。

表 7.2 – 範例 7.12：與控制組差異試驗 — 止痛膏的變異數分析

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	p
Total	125	545.78			
Judges	41	247.11	6.03	6.8	0.0001
Samples	2	225.78	112.89	127	0.0001
Error	82	72.89	0.89		
Sample Means with Dunnett's Multiple Comparisons					
Sample Mean response					
Blind control					
2.4a					
Product F					
4.8b					
Sample Mean response					
Blind control					
2.4a					
Product N					
5.5b					

Note: Within a row, means not followed by the same letter are significantly different at the 95% confidence level. Dunnett's $d_{0.05}=0.46$. Product N is significantly more different from the control than product F ($LSD_{0.05}=0.4$).

對樣品的 F -統計量(F -statistic)是高顯著的(表 19.6); $F_{2,82} = 127.0, p < 0.0001$ 。 F -統計量是一項比率：樣品的均方 (mean square)除以誤差項的均方，適當的自由度為那些與 F -統計量的分子與分母的均方關聯(分別是 2 及 82)。樣品均值(sample means)採用含控制組的多重比較用的 Dunnett 試驗做檢定，顯示兩個試驗樣品都顯著不同於盲控制組，也結論根據 LSD 多重比較，比起產品 F，產品 N 是顯著地($p < 0.05$)更加不同於控制組(LSD 值= 0.4; least significant difference)。

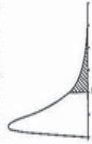
解釋 — 兩種樣品都測出有顯著的差異，**結論**：兩種配方充分地不同於控制組，有必要進行屬性差異試驗(attribute difference tests)(參見第 6 章表 6.3)或描述性分析(descriptive tests)(參見第 12 章附錄 12.2E，202-205 頁)，如黏度/濃稠性、皮膚熱感、皮膚涼感及延後感等。

Table 19.6 (Continued) Upper- α Probability Points of F -distribution (Entries are F_{α, v_1, v_2})

↑

v_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.94	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.61	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.55	1.51	1.47
29	2.88	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.50	1.46	1.42	1.38
60	2.79	2.39	2.20	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.61	1.54	1.48	1.44	1.40	1.35	1.29	1.25
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

Instructions: (1) Enter the section of the table corresponding to the predetermined α -level.
(2) Enter the row that corresponds to the denominator degrees of freedom (v_2).
(3) Pick the value of F in that row from the column that corresponds to the numerator degrees of freedom (v_1).



↑

v_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.94	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.61	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.55	1.51	1.47
29	2.88	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.50	1.46	1.42	1.38
60	2.79	2.39	2.20	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.61	1.54	1.48	1.44	1.40	1.35	1.29	1.25
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

$\alpha = 0.05$

1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.47	19.48	19.49	19.50	19.50

Table 19.6 Upper- α Probability Points of F -distribution (Entries are F_{α, v_1, v_2})



- ↑
- (1) 進入表內對應於所預定 α -水準的項。

(2) 進入列中對應於分母自由度 (v_2)。

(3) 從欄中對應於分子自由度 (v_1)之處，取在列之中的 F 值。

↑

v_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.11	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66

(Continued)

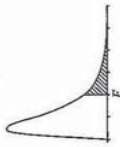
↑

v_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.05	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.00	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.73	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.16	2.13
15	4.54	3.68	3.28	3.05	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.15	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.48	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.26	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
30	4.16	3.31	2.91	2.68	2.52	2.41	2.32	2.25	2.19	2.14	2.07	2.00	1.92	1.88	1.83	1.78	1.73	1.67	1.61
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.36	1.25
∞	3.94	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

(Continued)

Table 19.6 (Continued) Upper- α Probability Points of F-distribution (Entries are $F_{\alpha;v_1,v_2}$)

輸入為 $F_{\alpha;v_1,v_2}$



(1) 進入表內對應於所預定 α -水準的項。(2) 進入列中對應於分母自由度 (v_2)。(3) 從欄中對應於分子自由度(v_1)之處，取在列之中的 F 值。

$v_2 \backslash v_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
$\alpha = 0.01$																			
1	4082	4999.50	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60

範例 7.13: 調味花生

問題/情況 — 有間大的糕餅製造工廠的品質保證經理需要監控調味花生在生產線中感官上的變動，各種添加風味(蜂蜜、香料、調味醬等)在**批次之間(among batches)**會有自然的變動，因而排除採用三角試驗、二三點試驗、或相同/不同試驗。大多數的整體差異試驗，譬如像這些如果受試者能測出**批次內(within a batch)**的變動，這就嚴重地減弱試驗會測出**批次對批次(batch-to-batch)**的差異的機會。因此，這需要的是**一個試驗允許從批次之間的變化中分離批次內的變化**。

12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.76	2.67	2.58
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

計畫目標 — 建立適合在調味花生製造中，**監控批次對批次變動**之試驗方法，最後制定品質保證/品質管制(QA/QC)的感官規格。

試驗目標 — 已知來源的調味花生在**批次內和批次間**，測定可感知的差異。

試驗設計 — 從倉庫取出一批**新近的控制組批次(control batch)** (正常生產)的樣品，來自**兩生產線**的每條的調味花生罐都採樣，標示為**Control-A** 與 **Control-B**，這些樣品代表**批次內的變動**；也從大量生產中抽取樣品，以不同批次的花生作為原料，這樣品標示為**“test”**。建立要測試 3 個配對之**與控制組差異試驗設計**：

- **Control A vs. Control A (blind control 盲控制組)**
- **Control A vs. Control B (within batch measure 批次內測量)**
- **Control A vs. Test (between batch measure 批次間測量)**

安排 50 位受試者在 3 天內參加 3 次獨立的測試(C_A vs. C_A ; C_A vs. C_B ; C_A vs. C_B)。配對樣品隨機分給受試者，所有的配對都是第一個給 C_A ，作為控制組，受試者在 0 至 10 的尺度上，評級 配對的成員之間的差異。結果的分析透過 Aust et al. (1985)方法，在方法中，為了確定差異的統計顯著性，**批次測量**之間的差異要 減去 **盲控制組和批次內測量分數**之間的差異 (*the difference between the score for the blind control and that for the within batch measure is subtracted from the between batch measure in order to determine statistical significance for a difference*)。

篩選樣品 — 來自生產、QA、市場及 R&D 等熟悉產品的人員預篩選樣品的風味、質地及外觀，為了確定每個樣品都是產品在批次內和批此間變動之代表，夥同感官分析專家一起決定，此次的試驗只有整顆花生才採樣和測試。

DIFFERENCE-FROM-CONTROL TEST

Name: _____ Date: 8-7-98 Test # 1103-68
Type of Sample: Flavored peanut snacks

Instructions

1. Taste the sample marked "Control" first.

2. Taste the sample marked with the three digit code.

3. Assess the overall sensory difference between the two samples using the scale below.

4. Mark the scale to indicate the size of the overall difference.

No difference

1
2
3
4
5
6
7
8
Extremely different 10

SCALE MARK TO INDICATE DIFFERENCE

REMEMBER THAT A DUPLICATE CONTROL IS THE SAMPLE SOME OF THE TIME.

COMMENTS : _____

圖 17.9 與控制組差異試驗之評分卷。範例 7.13：調味花生

進行試驗 — 每個樣品數 15 顆完整的花生，放入編碼的杯中。放在第一位置的 **Control-A** 標示為 **Control**；其餘的樣品都有三位數編碼：

Pair 1: Control A vs. Control A

標示：**“Control”** vs. [編碼]

Pair 2: Control A vs. Control B

標示：**“Control”** vs. [編碼]

Pair 3: Control A vs. Test Sample

標示：**“Control”** vs. [編碼]

評分卷列圖 7.19。

表 7.3 範例 7.13 的結果：與控制組差異試驗 -- 調味花生

Judge	Control A	Control B	Test
1	2	1	5
2	0	2	7
3	1	2	5
4	1	3	7
5	0	3	6
6	2	2	6
7	2	1	6
8	2	3	6
9	2	2	6
10	3	4	5
11	1	2	7
12	0	1	4
13	0	2	8
14	0	1	8
15	0	1	8
16	1	4	7
17	1	1	6
18	1	1	6
19	1	3	6
20	0	1	6
21	0	1	4
22	1	1	4
23	1	1	4
24	1	1	6

分析結果 — 品評結果(見表 7.3)的分析根據 Aust et al. (1985) 所述的程序。這程序測試是否試驗樣品的評分數顯著不同於兩控制組樣品的平均數。虛無和另一假設為

$$H_0: \mu_T = (\mu_{c_A} + \mu_{c_B})/2 \text{ vs. } H_1: \mu_T > (\mu_{c_A} + \mu_{c_B})/2$$

用來測試這假設的誤差(error)一詞，稱為純誤差均方 (“**pure error mean square**”) (見表 7.4)，是將所有品評員對兩個控制組樣品之間的差異的平方加總之，再除以品評員數的 2 倍。ANOVA 顯示試驗樣品和控制組樣品之間的差異是高顯著的。

$$F\text{-test } (F_{1,24} = MS_T \text{ vs. } R/MS_{\text{pure error}} = 326.54)$$

表 7.4 範例 7.13：調味花生的數據依據 Aust et al. (1985)的與控制組差異試驗所得的變異數分析表

Source	Degrees-of-Freedom	Sum of Squares	Mean Squares	F	p
Total	71	456.61			
Test vs. references	1	367.36	367.36	326.54	<0.0001
Pure error	24	27	1.13		
Residual	46	62.25			

7.9.1 範圍與應用

連續式試驗是一個可以節省取得結論所需品評數量的方法，例如在品評小組是要接受或拒絕訓練者，或大量生產商品是要裝運或銷毀。

不同於本章前述的試驗法，是為了**固定的 α** 與判定數量(n)，而將第二型誤差 Type II error (β) 最小化。在連續式試驗， α 及 β 值事先就決定， n 則透過評估每次感官品評的結果而決定。同時，由於 α 及 β 值事先決定，連續式試驗提供一個直接的方法，同時測試兩個樣本之間的差異或相似性。

7.9 連續式試驗(Sequential Tests)

連續式試驗非常實用及有效率，因為考慮到最初幾個評定的證據可能就足以(**固定的 α 及 β 值**)得出結論之可能性，任何再進一步的測試可能浪費時間與金錢。事實上，連續式試驗能夠減少所需品評數量多至 50%。

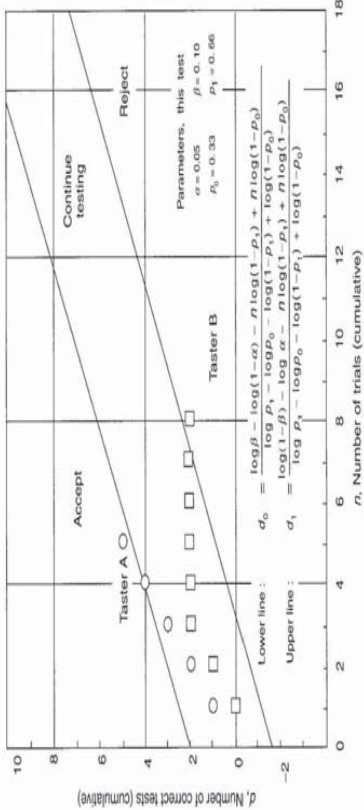
連續式方法可用於那些有正確及不正確答案的**差異存在著試驗 (existence-of-difference tests)**，例如三角、五取二及二-三點等試驗。

7.9.2 試驗原理

依照適合於所選擇方法之程序步驟，進行一系列品評，每一完成試驗的結果就輸入如圖 7.20 中的圖，其中可識別出三個區域：**接受區** (acceptance region)、**拒絕區** (rejection region) 和 **繼續測試區** (continue-testing region)。

解釋 — 專家結論出：即使控制組樣品中存在著變化性，試驗樣品是顯著不同於控制組的平均，他建議在下個步驟，用消費者確定是否試驗批次不同於喜好(preference)或接受性(acceptance)。這樣的決定能使公司確定品評小組所感知的差異對消費者是有意義的。透過消費者試驗搭配與控制組差異試驗之進一步的研究，就可以確定 QA 的實際規格。

圖 7.20：將試驗次數繪於水平(x)軸，正確回應總數繪於縱(y)軸。輸入第一次試驗的結果，如正確，為(x,y) = (1,1)；若不正確，為(x,y) = (1,0)。每一後續的試驗，正確回答者將 x 增加 1、y 也增加 1，若不正確回答，改 y 增加 0。持續測試直到出現一個點觸及或跨越與猶豫區接壤的線，這表示可取得結論(即接受或拒絕)。



α 聲明差異存在但實際上不存在之機率； β 聲明差異不存在但實際上存在之機率； p_0 當樣品是相同的，正確判斷的預期比率； p_1 當奇數樣品被檢測(猜的除外)到總場合數的一半，正確判斷的預期比率

圖 7.20 以三角試驗作為品評小組訓練員挑選之連續式方法範例

範例 7.14：在品評小組兩位訓練者的接受或拒絕

計畫目標 — 基於他們對一系列試驗樣品的靈敏度而選擇或拒絕訓練員。

試驗目標 — 確定每位訓練員是否其正確回答的長期比率 (p) 是足夠大而允入品評小組。

試驗設計 — 樣品配對以三角試驗的形式一次提出一個，試驗之間間隔要夠長而避免疲勞，每次三角試驗完了，結果輸入 **圖 7.20**，試驗系列持續直到訓練員是被接受或被拒絕。

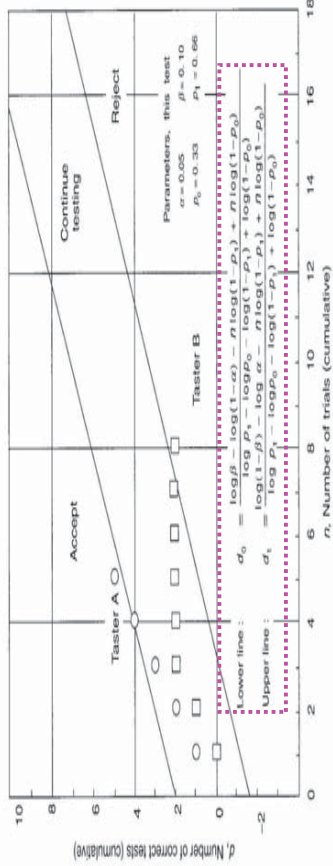
分析結果、試驗參數 — 四項參數值是由品評小組領導指派：

- α 是選出了一位不能接受的訓練者之機率
- β 是拒絕了一位可接受的訓練者之機率
- p_0 是最大不能接受的能力；以正確答對的比率測量
- p_1 是最小可接受的能力；以正確答對的比率測量

7.9.3 結果的分析與解釋：試驗的參數

本處採用的連續式試驗版本是 ISO (1983)，試驗本身是根據 Wald (1947)，Rao (1950)提出一個替代試驗，Bradley (1953)詳細說明這兩種試驗，並給予去計算達到結論所需的預期品評數量之方法，以及選擇與該方法有關的參數之規則，如範例 7.14 和 7.15。

圖 7.20：將圖分割為接受區等的線，其方程式取決於 α 、 β 、 p_0 、及 p_1 。目前的範例，訓練者 A 在所有的試驗都是正確的，五次三角試驗後就可被接受。訓練者 B 在第一次三角試驗失敗，第二及三次三角試驗成功，但之後的每次三角試驗都失敗，故第 8 次時被拒絕。



α 聲明差異存在但實際上不存在之機率； β 聲明差異不存在但實際上存在之機率； p_0 當樣品是相同的，正確判斷的預期比率； p_1 當奇數樣品被檢測(猜的除外)到總場合數的一半，正確判斷的預期比率

圖 7.20 以三角試驗作為品評小組訓練員挑選之連續式方法範例

四項參數可使用各種的值，當 p_0 靠近 p_1 ，所需測試的數量增加。幾種方法可減少所需的測試平均數量，第一，使用本章的三角試驗範例，將測出差異的最小可接受的機率設定為較高值，例如從現行範例的 50%提高至 67%，如此， $p_1 = 0.78$ [來自 $p_1 = 0.67 + (1 - 0.67)(1/3)$]，第二，如果可取得的訓練員夠多， α 及 β 要指派給較大值(例如 $\alpha > 0.05$ 及/或 $\beta > 0.10$)

三個樣品配對(控制組 vs. 1-day 組、控制組 vs. 3-day 組、控制組 vs. 5-day 組) 提供在分開的 **二-三點試驗**，控制組與樣品每隔受試者呈現作為參考品。每位受試者完成一次試驗，結果加入前面的回應，累計的結果繪圖，試驗序列繼續一直到貯藏樣品被宣告相似於或不同於控制組。

分析結果 — 所得結果如表 7.5 所示。在此 α 是宣告樣品不同於控制組之機率； β 是當實際上不同，但宣告樣品相似於控制組之機率。

範例 7.15：連續式二-三點試驗—牛肉餅的油騷味(warmed-over flavor)

計畫目標 — 軍食品工程站例行品管品評小組偵測出冷藏 5 天再復熱的牛肉餅有油騷味(WOF)，計畫負責人了解”軍隊靠肚子行軍”，想要為牛肉餅能夠冷藏的天數設定可信的最大值。

試驗目標 — 確定貯藏 1 天、3 天及 5 天的樣品，對照剛煎好的控制組，是否可測出差異。

試驗設計 — 預備試驗顯示在**二-三點試驗**，5-天組牛肉餅有強油騷味，而 1-天組則無，因此連續式試驗設計是合適的；這兩個樣本的決定可能會出現很少的回應。

(表 7.5) 來自範例 7.15 的結果：連續式二-三點試驗—牛肉餅的油騷味

Subject No.	Test A Control vs. 1 Day	Test B Control vs. 3 Day	Test C Control vs. 5 Day
1	I	0	0
2	I	0	I
3	I	0	I
4	C	I	C
5	I	1	I
6	C	I	C
7	I	2	I
8	C	3	C
9	I	3	C
10	I	4	C
11			I
12			C
13			C
14			C
15			C
16			C
17			C
18			C
19			C
20			C
21			C
22			C
23			C
24			C
25			C
26			C
27			C
28			C
29			C
30			C

Note: Column 1: I, incorrect; C, correct; Column 2: cumulative correct.

感官分析專家與計畫負責人決定設定 $\alpha = 0.10$ 及 $\beta = 0.10$ ，設定 $p_0 = 0.50$ ，二-三點試驗的虛無假設 p -值。又他們決定群體能辨識新鮮和貯藏樣品的最大比率不超過 40%，因此， p_1 值為

$$p_1 = (0.40) (1.0) + (0.60) (0.50) = 0.70$$

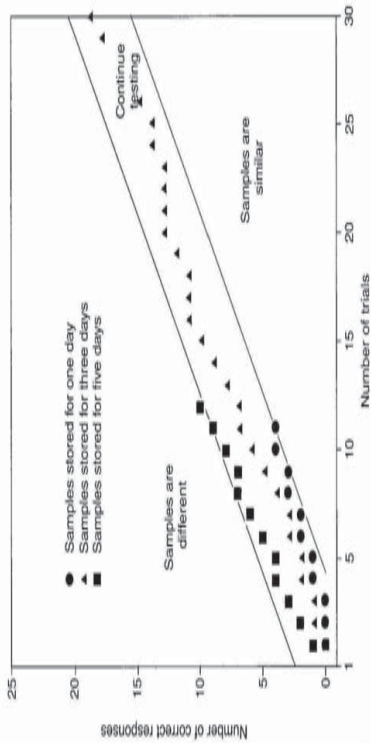
(從： p_1 = 辨識者機率 x 辨識者給的正確回應機率 + 非辨識者機率 x 非辨識者給的正確回應機率)

形成接受、拒絕與繼續-測試區域的界線之兩條線的方程式為

$$d_0 = -2.59 + 0.60n$$

$$d_1 = 2.59 + 0.60n$$

將兩條線、三個貯藏樣品(表 7.5)的每個在二-三點試驗證確回應數累計繪於圖 7.21，貯藏 1 天樣品宣告相似於控制組，貯藏 5 天樣品宣告顯著不同於控制組，而 30 次試驗後，貯藏 3 天樣品未被宣告顯著相似或不同於控制組。



(圖 7.21) 範例 7.15 結果的試驗圖：連續式二-三點試驗—牛肉餅的油騷味

解釋結果 — 計畫負責人接受 1-天及 5-天樣品的決定性結果，並被告知 3-天樣品其結果經 30 次測試後仍無法定案，他要接受 3 天為規格或者選擇繼續測試直到有了確定結論的結果。