行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　　　　※
※　　　　李克尺度資料平均數區間檢定之初探　　　　※
※　　　　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：☑個別型計畫　　☐整合型計畫
計畫編號：NSC 90-2416-H-009-008
執行期間：90 年 8 月 1 日至 91 年 7 月 31 日

計畫主持人：丁　承

本成果報告包括以下應繳交之附件：
☐赴國外出差或研習心得報告一份
☐赴大陸地區出差或研習心得報告一份
☐出席國際學術會議心得報告及發表之論文各一份
☐國際合作研究計畫國外研究報告書一份

執行單位：國立交通大學經營管理研究所

中　華　民　國　91　年　10　月　11　日

行政院國家科學委員會專題研究計畫成果報告

李克尺度資料平均數區間檢定之初探

Interval Testing About Means for the Likert-scale Data

計畫編號：NSC 90-2416-H-009-008

執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：丁 承　國立交通大學經營管理研究所

一、 中文摘要

在行為科學研究中常使用李克尺度 (Likert scale) 來量測觀點、認知、滿意度等，若欲檢視某變量平均值是否為「無意見」(中立)，區間假設應是較合理的表達方式，本研究以單一母體單變量平均數之檢定為對象，提出二項分配供作母體分配，並發展大樣本下基於常態理論與非中心性卡方分配之近似檢定法，我們以蒙地卡羅模擬評估該檢定法的有效性，結果顯示，該法之檢定力與二項母體分配下之模擬檢定力相當接近，且小樣本時亦然！再者，該法之使用十分簡易。針對李克尺度資料關於平均數為「無意見」的區間檢定，我們強烈推薦所提之檢定方法。

關鍵詞：李克尺度、區間檢定、二項分配、非中心性卡方分配、檢定力、蒙地卡羅模擬

Abstract

In behavior research, the Likert scale is often used to measure perception, opinion, satisfaction, etc. To examine if the average satisfaction for a variate is neutral, the interval hypothesis should be more reasonable. In this study, the binomial distribution is proposed to serve as a population distribution. A test for the interval hypothesis under the binomial population has been developed. The test is based on large-sample normal approximation, and requires the evaluation of the noncentral chi-square distribution. Monte Carlo simulation was used to evaluate the effectiveness of the test. The results indicated that the power of the test and the simulated power under the binomial population are very close, even for small samples. In addition, the test, illustrated with examples, is easy to perform. The test is strongly recommended for interval testing about means regarding neutrality for the Likert-scale data.

Keywords: Likert Scale, Interval Testing, Binomial Distribution, Noncentral Chi-square Distribution, Power, Monte Carlo Simulation

二、緣由與目的

在行為科學研究中常使用李克尺度 (Likert scale) 來量測觀點、認知、滿意度等，並以統計方法進行後續推論[3]。舉例而言，針對某一特定服務項目，若以七分尺度量測其滿意度，吾人通常以 1、2、3、4、5、6、7 分別表示「非常不滿意」、「不滿意」、「稍不滿意」、「無意見」、「稍滿意」、「滿意」及「非常滿意」。就某一特定族群，平均而言，對該項目是否滿意常是吾人關心的第一點，若以μ表滿意度之母體平均值，則待檢定之假設在傳統上的表達方式為 $H_0: \mu = 4$ vs. $H_1: \mu \neq 4$。然而，由於李克尺度係順序尺度，屬間斷性質，故以「$H_0: \mu = 4$」表達「無意見」並無法反映不確定性，吾人可將所示之單點假設修訂為一區間型態，即 $H_0: 4-\delta \le \mu \le 4+\delta$ $(\delta > 0)$ (例如：$3.8 \le \mu \le 4.2$) [4]以反映「無意見」之無差異區域 (indifference zone)，該表達

方式似較具說服力。另一方面，由於李克尺度下之母體分配非常態，僅能於樣本數大時藉助中央極限定理採常態檢定，然相關之檢定問題在文獻上似未見具體的著墨，故擬於本研究中提出初步的探討，並藉供應用參考。

三、結果與討論

由於李克尺度最常採用七分尺度和五分尺度[e.g., 1]，亦有採用九分尺度者，故所涉及之檢定問題應一併考量。設隨機變數 X 代表某項目之滿意度，並以李克尺度進行量測。在從事統計檢定前，母體分配應先行設定，雖然二項分配通常係扮演抽樣分配的角色，但在此卻能適當地擔任 X 之母體分配，設 k 為李克尺度之尺度分級數，則 X 之母體分配可設為 $b(k-1, p) + 1$。若 $k = 7$ (七分尺度)，則母體分配可設為 $b(6, p) + 1$；若 $k = 5$ (五分尺度)，則設為 $b(4, p) + 1$；若 $k = 9$ (九分尺度)，則為 $b(8, p) + 1$，分配型態隨 p 值不同而變化。

我們將分別就七分尺度、五分尺度及九分尺度進行檢定之探討，並歸納檢定通則。

(一) 七分尺度

在七分尺度下，由於 $E(X) = \mu = 6p+1$，$Var(X) = \sigma^2 = 6p(1-p)$，故單點假設 $H_{01}: \mu = 4$ 與區間假設 $H_{02}: 4-\delta \le \mu \le 4+\delta$ 可藉 p 分別表達如下：$H_{01}: p = 0.5$ 及 $H_{02}: 0.5-\delta/6 \le p \le 0.5+\delta/6$。以下將分別探討檢定方式及其檢定力。鑒於區間假設之檢定方式係植基於單點假設檢定之特性，故先探討單點檢定，再論區間檢定。

1. 單點檢定

欲檢定 $H_{01}: \mu = 4$ vs. $H_{11}: \mu = \mu_1 \neq 4$，當樣本數大時，可藉由中央極限定理，採檢定統計量 $Z = (\bar{X}-4)/\sqrt{6\times0.5\times0.5/n} = (\bar{X}-4)/\sqrt{6\times0.25/n} = (\bar{X}-4)/\sqrt{1.5/n}$ 或 $Z^2 = n(\bar{X}-4)^2/1.5$。當 H_{01} 為真，$Z \sim N(0,1)$，$Z^2 \sim \chi^2_1$，故在顯著水準α下，檢定之臨界值可採用 $\chi^2_{1,\alpha}$ (表自由度為 1 之卡方分配的第 100(1−α)百分位數)，若 $Z^2 \ge \chi^2_{1,\alpha}$，則拒絕 H_{01}；另一方面，當 H_{11} 為真，$p = p_1 = (\mu_1-1)/6$，

$$\bar{X} \sim N(\mu_1, 6p_1(1-p_1)/n).$$

$$=> (\bar{X}-4)/\sqrt{6\times0.25/n} \sim N((\mu_1-4)/\sqrt{6\times0.25/n},\ p_1(1-p_1)/0.25).$$

$$=> (\bar{X}-4)/\sqrt{(1.5/n)(4p_1(1-p_1))} \sim N((\mu_1-4)/\sqrt{6p_1(1-p_1)/n},\ 1).$$

$$=> n(\bar{X}-4)^2/1.5 \sim 4p_1(1-p_1)\chi^2_{1,\lambda},\quad \lambda = n(\mu_1-4)^2/6p_1(1-p_1) \quad \text{[e.g., 4].}$$

因此，當 H_{01} 為真，χ^2 檢定之檢定力 (power) = $\Pr(Z^2 \ge \chi^2_{1,\alpha} \mid H_{01}) = \Pr(n(\bar{X}-4)^2/1.5 \ge \chi^2_{1,\alpha} \mid H_{01}) = \alpha$；而當 H_{11} 為真，

$$\text{power} = \Pr(Z^2 \ge \chi^2_{1,\alpha} \mid H_{11})$$

$$= \Pr(4p_1(1-p_1)\chi^2_{1,\lambda} \ge \chi^2_{1,\alpha})$$

$$= \Pr(\chi^2_{1,\lambda} \ge \chi^2_{1,\alpha}/4p_1(1-p_1)),$$

$$\lambda = n(\mu_1-4)^2/6p_1(1-p_1). \quad (1)$$

檢定力可採式(1)求算，與常態母體下逕以非中心性卡方分配求取有些許差異，而係非中心性卡方分配之調整式！

以上論點係當樣本數大時之結果，但樣本數究應多大方可適用則可以蒙地卡羅模擬 (Monte Carlo simulation) 進行檢視。在母體分配為 $b(6,p) + 1$ 的前提假設下，我們就不同樣本數 (檢視 n=10, 20, 30, 50, 100)，針對不同母體平均數 (取至所對應之檢定力值接近 1.0)，以二項分配隨機數值產生器抽取樣本，重複執行 5000 次，可得樣本平均數 $\bar{X}$ 之估計抽樣分配，接著取在 $\mu = 4$ 下 $\bar{X}$ 模擬抽樣分配之第 2.5 及第 97.5 百分位數 (顯著水準 α 設為 0.05) 作為檢定臨界值，分別以 crit1 及 crit2 表之，檢定力可以下式近似之：

$\widehat{\text{power}} = \#\{\bar{x}_i \le \text{crit1}$ 或 $\bar{x}_i \ge \text{crit2}\}/5000.$ (2)

由於在常態母體下之小樣本檢定係採用 $t = \sqrt{n}(\bar{X}-4)/S$，$S^2 = \sum_{i=1}^{n}(X_i-\bar{X})^2/n-1$ (服從自由度 $n-1$ 之 t 分配)，故在模擬時亦同時求取 t 檢定之檢定力近似值 (以 $\widehat{\text{power}}_t$ 表之)以資比較。

模擬結果顯示，以基於常態理論之(1)式所得之檢定力與二項母體下之模擬檢定力相當接近，縱使當樣本數小至 $n = 10$ 亦然！但 $\widehat{\text{power}}_t$ 在 $n \le 30$ 時呈明顯低估之態勢 (樣本數 ≥50 時方近似於 $\widehat{\text{power}}$)，檢定力函數係以 $\mu = 4$ 為中心呈現對稱型態，當 $\mu = 4$ 時，模擬檢定力十分靠近 $\alpha = 0.05$。因此，在二項母體分配下， H_{01}：$\mu = 4$ vs. H_{11}：$\mu \ne 4$ 之檢定可採上述基於常態理論之 χ^2 檢定獲充分支持，臨界值係基於中心性 χ^2_1，檢定力則利用式(1)所示調整方式求取。由於方法簡易，故在相關的推論上(如樣本數之決定)甚具應用價值，該項優點可延伸至區間檢定。至於 t 檢定，因樣本數小 (≤ 30) 時檢定力被低估，故不建議使用。

2. 區間檢定

欲檢定 H_{02}: $4-\delta \le \mu \le 4+\delta$ (即 $0.5-\delta/6 \le p \le 0.5+\delta/6$) ($0 < \delta$) vs. H_{12}: $\mu < 4-\delta$ or $\mu > 4+\delta$，仍採 $Z^2 = n(\bar{X}-4)^2/1.5$ 作為檢定統計量，如前所述，當樣本數大時，$Z^2 \sim 4p(1-p)\chi^2_{1,\lambda}$, $p = (\mu-1)/6$, $\lambda = n(\mu-4)^2/6p(1-p)$。由於檢定之臨界值是在 $\mu = 4 \pm \delta$ 下(即 H_{02} 的兩個端點)求取 ([4], [5])，而當 $\mu = 4 \pm \delta$ (即 $p = 0.5 \pm \delta/6$)時，$Z^2 \sim 4p'(1-p')\chi^2_{1,\lambda'}$，$\lambda' = n\delta^2/[6p'(1-p')]$，$p' = 0.5 + \delta/6$，故在顯著水準 α 下，檢定之臨界值 $c = 4p'(1-p')\chi^2_{1,\lambda',\alpha}$ ($\chi^2_{1,\lambda',\alpha}$ 表自由度為 1 且非中心性參數為 λ' 之非中心性卡方分配的第 $100(1-\alpha)$ 百分位數)，若 $z^2 \ge c$，則拒絕 H_{02}。當 H_{02} 為真，該檢定之檢定力 $\le \alpha$；當 H_{12} 為真 (此時 $\mu = \mu_1$, $p = p_1 = (\mu_1-1)/6$, μ_1, $p_1 \in H_{12}$)，檢定力

$$\text{power} = \Pr(\chi^2_{1,\lambda} \ge c/4p_1(1-p_1)),$$

$$\lambda = n(\mu_1-4)^2/6p_1(1-p_1). \quad (3)$$

接著仍就不同樣本數，以蒙地卡羅法模擬母體分配為 $b(6,p)+1$ 假設下上述區間檢定之檢定力。檢定係以 $\bar{X}$ 為基礎，在顯著水準 $\alpha = 0.05$ 及重複次數 5000 之下，檢定下臨界值(crit1)與上臨界值 (crit2) 之決定係滿足以下條件：

$$\#\{\bar{x}_i \le \text{crit1} \text{ 或 } \bar{x}_i \ge \text{crit2} \mid \mu = 4 \pm \delta\}/5000 \approx 0.05.$$

而檢定力仍以式(2)近似之。我們針對 $\delta = 0.1, 0.2, 0.3, 0.4$, 及 0.5 進行模擬分析，發現區間檢定之模擬檢定力與利用式(3)所求得之檢定力皆十分接近，縱使當樣本數小至 $n = 10$ 亦然！因此，在母體分配為二項分配下，H_{02}: $4-\delta \le \mu \le 4+\delta$ 之檢定可採基於常態理論之 χ^2 檢定亦獲強力支持，與單點檢定之差異僅在於臨界值 c 係求自於非中心性 $\chi^2_{1,\lambda'}$ 之調整式。

蒙地卡羅模擬係利用統計套裝軟體 SAS [6]執行，二項分配隨機數值產生器利用軟體中之 RANBIN 函數，非中心性卡方分配之計算則藉助 PROBCHI 及 CINV 二函數。SAS 程式及所得結果限於篇幅不做列式，但歡迎來函索取。

(二) 五分尺度

在五分尺度下，X 之母體分配設為 $b(4,p)+1$，故 $E(X) = \mu = 4p+1$，$\text{Var}(X) = 4p(1-p)$，當樣本數大時，區間假設 H_{02}: $3-\delta \le \mu \le 3+\delta$ ($0.5-\delta/4 \le p \le 0.5+\delta/4$) vs. H_{12}: $\mu < 3-\delta$ or $\mu > 3+\delta$ 之檢定可採統計量 $Z = (\bar{X}-3)/\sqrt{4\times 0.5\times 0.5/n} = \sqrt{n}(\bar{X}-3)$ 或 $Z^2 = n(\bar{X}-3)^2$，此時，$p = (\mu-1)/4$，

$$\bar{X} \sim N(\mu, 4p(1-p)/n).$$

$$\Rightarrow \sqrt{n}(\bar{X}-3) \sim N(\sqrt{n}(\mu-3), 4p(1-p)).$$

$\Rightarrow \sqrt{n}\,(\bar{X}-3)/\sqrt{4p(1-p)}$
$\quad \sim N(\sqrt{n}\,(\mu-3)/\sqrt{4p(1-p)}\,,\,1).$

$\Rightarrow Z^2 = n(\bar{X}-3)^2 \sim 4p\,(1-p)\,\chi^2_{1,\lambda}$,
$\quad \lambda = n(\mu-3)^2/4p(1-p).$

當 $\mu = 3 \pm \delta$ (即 $p = 0.5 \pm \delta/4$) 時，$Z^2 \sim 4p'(1-p')\,\chi^2_{1,\lambda'}$，$\lambda' = n\delta^2/[4p'(1-p')]$，$p' = 0.5 + \delta/4$，故在顯著水準$\alpha$下，檢定之臨界值 $c = 4p'(1-p')\ \chi^2_{1,\alpha,\lambda'}$，若 $z^2 \geq c$，則拒絕 H_{02}。當 H_{02} 為真，檢定力≤ α；當 H_{12} 為真 (此時$\mu = \mu_1$, $p = p_1 = (\mu_1 - 1)/4$, $\mu_1, p_1 \in H_{12}$)，檢定力

power = $\Pr(\chi^2_{1,\lambda} \geq c/4p_1(1-p_1))$,
$\quad \lambda = n(\mu_1-3)^2/4p_1(1-p_1).$

(三) 九分尺度

在九分尺度下，$E(X) = \mu = 8p+1$，$Var(X) = 8p(1-p)$，區間假設為 H_{02}：$5-\delta \leq \mu \leq 5+\delta$ $(0.5-\delta/8 \leq p \leq 0.5+\delta/8)$ vs. H_{12}：$\mu < 5-\delta$ or $\mu > 5+\delta$。檢定統計量採 $Z = (\bar{X}-5)/\sqrt{8\times 0.25/n} = (\bar{X}-5)/\sqrt{2/n}$ 或 $Z^2 = n(\bar{X}-5)^2/2$。此時，$p = (\mu-1)/8$，

$\bar{X} \sim N(\mu, 8p(1-p)/n).$

$\Rightarrow (\bar{X}-5)/\sqrt{8\times 0.25/n}$
$\quad \sim N((\mu-5)/\sqrt{8\times 0.25/n},\, p(1-p)/0.25).$

$\Rightarrow (\bar{X}-5)/\sqrt{(2/n)(4p(1-p))}$
$\quad \sim N((\mu-5)/\sqrt{8p(1-p)/n},\, 1).$

$\Rightarrow n(\bar{X}-5)^2/2 \sim 4p(1-p)\,\chi^2_{1,\lambda}$,
$\quad \lambda = n(\mu-5)^2/8p(1-p).$

當 $\mu = 5 \pm \delta$ (即 $p = 0.5 \pm \delta/8$) 時，$Z^2 \sim 4p'(1-p')\,\chi^2_{1,\lambda'}$，$\lambda' = n\delta^2/[8p'(1-p')]$，$p' = 0.5 + \delta/8$，故在顯著水準$\alpha$下，檢定之臨界值 $c = 4p'(1-p')\ \chi^2_{1,\alpha,\lambda'}$，若 $z^2 \geq c$，則拒絕 H_{02}。當 H_{02} 為真，檢定力≤ α；當 H_{12} 為真(此時$\mu = \mu_1$, $p = p_1 = (\mu_1 - 1)/8$, $\mu_1, p_1 \in H_{12}$)，檢定力

power = $\Pr(\chi^2_{1,\lambda} \geq c/4p_1(1-p_1))$,
$\quad \lambda = n(\mu_1-5)^2/8p_1(1-p_1).$

我們仿照七分尺度所進行之蒙地卡羅模擬方式，對五分尺度以及九分尺度從事分析，在母體具二項分配的前提假設下，無論樣本數大小，模擬檢定力與基於常態理論所得者十分接近，故結論與七分尺度所獲者相同。

(四) 區間檢定之通則

綜合以上討論，針對李克尺度資料，單一母體單變量平均數為「無意見」的區間檢定通則可歸納如下：

設 k 為李克尺度之尺度分級數，母體分配設為二項分配 $b(k-1, p) + 1$，故平均數 $\mu = (k-1)\,p + 1$，而 $p = (\mu-1)/(k-1)$，區間假設型態為 H_0：$(k+1)/2 - \delta \leq \mu \leq (k+1)/2 + \delta$ $(\delta > 0)$，亦可藉 p 表達成 H_0：$0.5 - \delta/(k-1) \leq p \leq 0.5 + \delta/(k-1)$。欲檢定 H_0，可採檢定統計量

$Z^2 = n[\bar{X} - (k+1)/2]^2/[0.25(k-1)]$，

其抽樣分配為 $4p(1-p)\,\chi^2_{1,\lambda}$,
$\quad \lambda = n[\mu - (k+1)/2]^2/[(k-1)p(1-p)].$

在顯著水準α下，檢定之臨界值 $c = 4p'(1-p')\,\chi^2_{1,\alpha,\lambda'}$，$\lambda' = n\delta^2/[(k-1)\,p'\,(1-p')]$，$p' = 0.5 + \delta/(k-1)$，若 $z^2 \geq c$，則拒絕 H_0。無論樣本大小，該檢定之有效性已經蒙地卡羅模擬檢驗通過，且該法之使用十分簡易(見下節之操作範例)，針對上述型態之區間假設，我們強烈推薦該檢定法。

值得注意的是，該檢定法亦適用於五分、七分和九分以外之其他李克尺度資料。

(五) 範例

本小節舉例說明上述區間假設之檢定操作方式。在李克七分尺度下，已知顯著

水準α= 0.05，樣本數為 30，現擬檢定 H_0：$3.8 \le \mu \le 4.2$ ($\delta = 0.2$)。此例 $k = 7$, 檢定統計量 $Z^2 = 30(\bar{X} - 4)^2 / 1.5$，$p' = 0.5 + 0.2/6 = 0.5333$，非中心性參數 ¯ ’= $30*(0.2)^2$ / (6*0.5333*0.4667) = 0.8036，$\chi^2_{1,0.05,0.8036} =$ 6.4725，故檢定之臨界值 c = 4*0.5333* 0.4667* 6.4725 = 6.44。因此，所採用之檢定為：若 $z^2 = 30\,(\bar{x} - 4)^2/1.5 \ge 6.44$，則拒絕 H_0。針對不同 δ 值與樣本數，其對應之檢定臨界值彙總於表 1，該表顯示，在樣本數固定下，臨界值隨 δ 之增加而增大；在 δ(>0)固定下，臨界值亦隨樣本數之增加而增大。至於檢定之樣本數如何決定可參考 Ding [2] 所提示的方法，利用 iteration 的方式求取。

表 1 李克七分尺度下不同 δ 值與樣本數 *n* 之區間檢定臨界值彙總表

$n \backslash \delta$	0	0.1	0.2	0.3	0.4	0.5
10	3.84	4.09	4.80	5.85	7.10	8.49
20	3.84	4.34	5.67	7.47	9.55	11.89
30	3.84	4.58	6.44	8.87	11.69	14.88
50	3.84	5.04	7.82	11.35	15.52	20.33
100	3.84	6.08	10.72	16.70	23.97	32.54

註：1. 虛無假設為 H_0：$3.8 \le \mu \le 4.2$。
2. 顯著水準採 0.05。
3. $\delta = 0$ 時相當於單點檢定。

(六) 研究限制與後續研究建議

母體分配設定為二項分配係本研究之研究限制，應用時須先檢測該前題假設是否成立。另反映不確定性之 δ 值須事先給定，如何客觀決定 δ 值是有趣的後續研究課題。後續研究亦可延伸至不同母體間比較之區間檢定，而針對不同母體分配應如何進行檢定及針對任意平均數之區間檢定法亦有待進一步思考。

四、計畫成果自評

研究成果在應用上具參考價值，擬投稿學術期刊發表。

五、參考文獻

[1] Collins, W. C., Raju, N. S., and Edwads, J. E. (2000), “Assessing Differential Functioning in a Satisfaction Scale,” *Journal of Applied Psychology*, 85, 451-461.

[2] Ding, C. G. (1999), “An Efficient Algorithm for Computing Quantiles of the Noncentral Chi-Squared Distri- bution,” *Computational Statistics & Data Analysis*, 29, 253-259.

[3] Judd, C. M., Smith, E. R., and Kidder, L. H. (1991), *Research Methods in Social Relations* (6th ed.), Orlando, FL: Holt Rinehart and Winston.

[4] Lehmann, E. L. (1986), *Testing Statistical Hypotheses* (2nd ed.), New York: Wiley.

[5] Roussas, G. G. (1973), *A First Course in Mathematical Statistics*, Reading, MA: Addison-Wesley.

[6] SAS Institute Inc. (1990), *SAS Language: Reference* (Version 6, First edition), Cary, NC: Author.