Chapter 5
Nonequivalent Groups: Equipercentile Methods

5.1 Frequency Estimation Method

報告人：博三D10711002 蘇恭弘
ghsghs1@gmail.com

Chapter 5
Nonequivalent Groups: Equipercentile Methods

5 Nonequivalent Groups: Equipercentile Methods . . . 143

5.1 Frequency Estimation Method . . . 143

5.1.1 Conditional Distributions . . . 144

5.1.2 Assumptions and Procedures . . . 144

5.1.3 Numerical Example. . . . 147

5.1.4 Estimating the Distributions . . . 150

5.1.5 Special Case: Braun-Holland Linear Method . . . 151

5.1.6 Illustrative Example . . . 152

5.2 Other Methods . . . 158

5.2.1 Modified Frequency Estimation . . . 158

5.2.2 Chained Equipercentile Equating . . . 159

5.2.3 Illustrative Example . . . 164

5.3 Practical Issues. . . . 165

5.4 Exercises . . . 166

References . . . 166

垂直等化

- **Chapter 5**
  Nonequivalent Groups: Equipercentile Methods
  - 5-1 Frequency Estimation Method
    - 5.1.1 Conditional Distributions
    - 5.1.2 Assumptions and Procedures
    - 5.1.3 Numerical Example
    - 5.1.4 Estimating the Distributions
    - 5.1.5 Special Case: Braun-Holland Linear Method
    - 5.1.6 Illustrative Example
  - 5-2 Other Methods
  - 5-3 Practical Issues

Chapter 5
Nonequivalent Groups: Equipercentile Methods
非等同組:相同百分位數等化法
整章內容總覽

為了共同項目非等化群的設計我們已經發展出了相同百分位數等化法，這些方法考慮了總分和分數在共同項目上的分配，而不是僅考慮第四章討論的平均數、標準差和變異數。

如前所述，相同百分位數等化法是從第一章中描述的觀察分數等同性的角度發展起來的觀察分數等化程序。 因此在共同項目非等化群的相等百分位數等化法通常需要有第四章所說的合成群體。

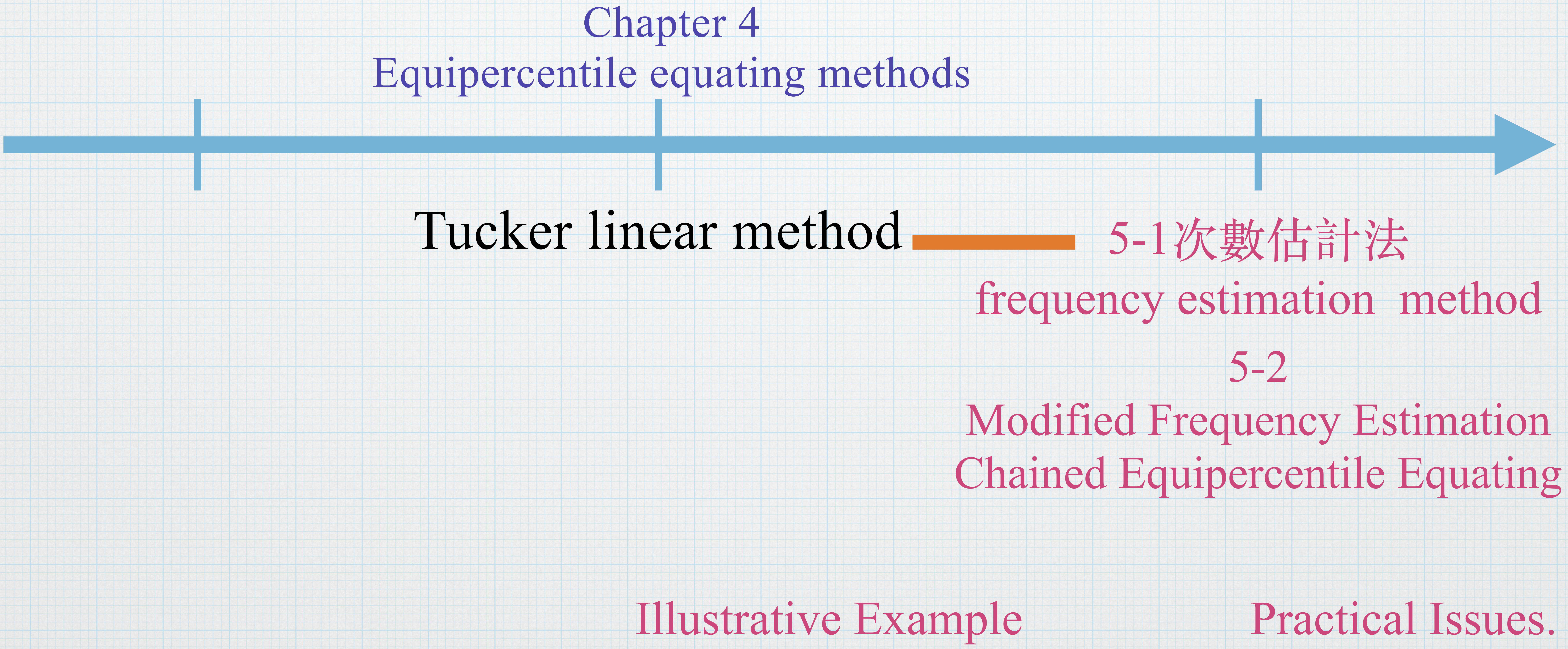
本章先討論第一種方法---次數估計法，它和第四章的Tucker linear method緊密相關，然後我們考慮另外兩種方法。其中一種是次數估計法的修改版。另一種與在Chap4中討論的chained linear method緊密一致。

我們還描述了第三章中描述的平滑的方式，可以在非等化群進行相同百分位數等化法來使用。在本章中使用的方法和第四章使用相同的數據，將結果與第四章的線性結果進行比較。

Chapter 5

Nonequivalent Groups: Equipercentile Methods

整章內容總覽



垂直等化

觀察分數連結法

非線性連結法

5.1 Frequency Estimation Method

The *frequency estimation method (*次數估計法*)*described by Angoff (1971) and Braun and Holland (1982) provides a means for estimating the cumulative distributions(累積分佈) of scores on Form X and Form Y for a synthetic population (合成群體)from data that are collected using the common-item nonequivalent groups design.

5.1 Frequency Estimation Method

5.1.1 Conditional Distribution

Define $f(x, v)$ as the joint distribution of total score and common-item score, so that $f(x, v)$ represents the probability of earning a score of x on Form X and a score of v on the common items. Specifically, $f(x, v)$ is the probability that $X = x$ and $V = v$.

Define $f(x)$ as the marginal distribution of scores on Form X, that is, $f(x)$ represents the probability that $X = x$. $h(v)$ as the marginal distribution of scores on the common items, so that $h(v)$ represents the probability that $V = v$

$f(x \mid v)$ as the conditional distribution of scores on Form X for examinees earning a particular score on the common items. Thus, $f(x \mid v)$ represents the probability that $X = x$ given that $V = v$.

$$f(x|v) = \frac{f(x, v)}{h(v)}. \tag{5.1}$$

From Eq. (5.1), it follows that

$$f(x, v) = f(x|v)h(v). \tag{5.2}$$

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

此法假設，在共同試題連結設計下的兩個群體，在 X 測驗和 Y 測驗得分的條件分配（即當共同試題得分固定時）是相等的；根據此假設可推導出此二群體所構成的合成群體（synthetic population）在 X 測驗和 Y 測驗原始分數的累積分配，然後再將 X 測驗和 Y 測驗分數中擁有相同百分等級者，視為等同。而合成群體在 X 測驗和 Y 測驗之分配，可以下列二公式表示：

$$f_s(x) = w_1 f_1(x) + w_2 f_2(x)$$

$$g_s(y) = w_1 g_1(y) + w_2 g_2(y)$$

(5.3)

其中，**s** 代表合成群體，**1** 和 **2** 分別代表被施測**X**測驗和 **Y** 測驗的群體**1**和群體 **2**；另外，f 和 g 分別代表 **X** 和 **Y** 測驗分數的分配，而 w_1 和 w_2 分別代表群體**1**和 **2** 之加權比重，且 $w_1 + w_2$ **=1**。

5.1 Frequency Estimation Method

如何找出$f_2(x)$與$g_1(y)$?

$$f_s(x) = w_1 f_1(x) + w_2 f_2(x)$$

$$g_s(y) = w_1 g_1(y) + w_2 g_2(y) \qquad (5.3)$$

$f_1(x)$

$f_2(x)$

X測驗

群體1

w_1

共同試題V

$h_1(v)$、$h_2(v)$

$w_1 + w_2 = 1$

$g_1(y)$

$g_2(y)$

Y測驗

群體2

w_2

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

延續第4章的概念，我們只能獲得群體1施測X測驗的$f_1(x)$ 與群體1施測Y測驗的$g_2(y)$ ，因為群體2未施測X測驗，群體2施測X測驗，故$g_2(x)$與$f_1(y)$為未知。

因此有關有以下的 X 和 Y 測驗得分的條件分配假設，以下述的公式表示：

$$f_1(x|v) = f_2(x|v)$$
$$g_1(y|v) = g_2(y|v) \quad \textbf{(5.4)}$$

其中，$f(x|v)$ 和 $g(y|v)$ 分別代表 **X** 和 **Y** 測驗分數在共同試題V的條件分配。

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

由條件機率導出的式子(5.2)

$$f(x, v) = f(x|v)h(v). \tag{5.2}$$

$$f_2(x, v) = f_2(x|v)h_2(v) \quad \text{and} \quad g_1(y, v) = g_1(y|v)h_1(v). \tag{5.5}$$

組合式子(5.4)與(5.5)，得到式子(5.6)如下

$$\begin{aligned} f_2(x,v) &= f_2(x|v)h_2(v) \\ &= f_1(x|v)h_2(v) \\ g_1(y,v) &= g_1(y|v)h_1(v) \\ &= g_2(y|v)h_1(v) \end{aligned} \tag{5.6}$$

$$\begin{aligned} f_1(x|v) &= f_2(x|v) \\ g_1(y|v) &= g_2(y|v) \end{aligned} \tag{5.4}$$

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

本來因為群體2未施測X測驗，群體1施測Y測驗，故$f_2(x)$與$g_1(y)$為未知。
接下來說明如何求出 $f_2(x)$ 與 $g_1(y)$。
藉由式子(5.6)，我們可利用已知的$f_1(x|v)$、$g_2(y|v)$、$h_1(v)$、$h_2(v)$
求得$f_2(x,v)$、$g_1(y,v)$

$$
\begin{aligned}
f_2(x,v) &= f_2(x|v)h_2(v) \\
&= f_1(x|v)h_2(v)
\end{aligned}
\tag{5.6}
$$

$$
\begin{aligned}
g_1(y,v) &= g_1(y|v)h_1(v) \\
&= g_2(y|v)h_1(v)
\end{aligned}
$$

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

由定義，Define f (x) as the marginal distribution of scores on Form X

因此可得式子(5.7)

$$f_2(x) = \sum_v f_2(x, v) = \sum_v f_1(x|v)h_2(v) \quad \text{and}$$
$$g_1(y) = \sum_v g_1(y, v) = \sum_v g_2(y|v)h_1(v). \tag{5.7}$$

$$f_2(x, v) = f_1(x|v)h_2(v)$$
$$g_1(y, v) = g_2(y|v)h_1(v) \tag{5.6}$$

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

$$f_2(x) = \sum_v f_2(x,v) = \sum_v f_1(x|v)h_2(v) \quad \text{and}$$

$$g_1(y) = \sum_v g_1(y,v) = \sum_v g_2(y|v)h_1(v). \tag{5.7}$$

將式子(5.7) 代入式子(5.3)中，可得

$$f_s(x) = w_1 f_1(x) + w_2 f_2(x) \tag{5.3}$$

$$g_s(y) = w_1 g_1(y) + w_2 g_2(y),$$

$$f_s(x) = w_1 f_1(x) + w_2 \sum_v f_1(x|v)h_2(v) \quad \text{and}$$

$$g_s(y) = w_1 \sum_v g_2(y|v)h_1(v) + w_2 g_2(y). \tag{5.8}$$

合成群體在 X 測驗和 Y 測驗之分配均可求出

5.1 Frequency Estimation Method

5.1.2 Assumptions and Procedures

對於合成群體，$f_s(x)$ 可以累加到它的值來產生累積分佈 $F_s(x)$。累積分佈 $G_s(y)$ 也可由相同方法得出。

定義 P_s 為X測驗的百分等級函數，使用在第二章中談及的百分等級的定義，定義為Y測驗的百分等級函數 Q_s。類似地，$P^{-1}{}_s$和 $Q^{-1}{}_s$為百分位函數。

合成群體的等分函數為

$$e_{Ys}(x) = Q_s^{-1}[P_s(x)], \tag{5.9}$$

它類似於等式2.17中的隨機群的等化關係

$$e_Y(x) = y = Q^{-1}[P(x)], \qquad -.5 \le x \le K_X + .5. \tag{2.17}$$

$$\begin{aligned} e_Y(x) &= Q^{-1}[P(x)] \\ &= \frac{P(x)/100 - G(y_U^* - 1)}{G(y_U^*) - G(y_U^* - 1)} + (y_U^* - .5), \qquad 0 \le P(x) < 100, \\ &= K_Y + .5, \qquad\qquad P(x) = 100. \end{aligned} \tag{2.18}$$

5.1.2 Assumptions and Procedures

公式（5.4）的次數估計假設不能使用使用共同項非等效分組設計收集的數據進行檢驗。為了驗證這一假設，一組來自群體1的代表性考生需要採用Y試題，來自群體2的代表性考生群體需要採用X試題。

不幸的是，這些數據在實際中無法獲得。如果群體1和群體2相同，則滿足等式（5.4）中的假設。從邏輯上講，群體1和群體2越相似，這個假設就越有可能成立。因此，只有當兩個總體合理相似時，才應進行次數估計等化。“合理相似”的相似程度取決於等化的背景和所需相似程度的經驗證據。當總體差異較大時，應考慮基於真實分數模型的方法，如本章後面描述的修正次數估計方法(Modified Frequency Estimation)或第6章中描述的IRT方法，儘管當總體差異很大時，不可能進行充分的等化。第8章將進一步討論這個問題。

5.1.3 Numerical Example

接下來用一個基於合成群體的數值例子來幫助理解這種方法。

在這個例子中，X測驗有5個項目，Y測驗也有5個項目，並且有3個共同試題。

假設共同試題是外部的。

Table 5.1 Form X and common-item distributions for population 1 in a hypothetical example

	v					
x	0	1	2	3	$f_1(x)$	$F_1(x)$
0	.04	.04	.02	.00	.10	.10
1	.04	.08	.02	.01	.15	.25
2	.06	.12	.05	.02	.25	.50
3	.03	.12	.05	.05	.25	.75
4	.02	.03	.04	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_1(v)$	.20	.40	.20	.20		

Note Values shown in the body of table are for $f_1(x, v)$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

表5.1給出了假設示例中群體1的數據。表格中的數字代表聯合機率分配$f_1(x,v)$。例如，左上角的值是.04。該值表示第1組的考生在X測驗中獲得0分，在共同試題中獲得0分的機率。表5.1中所有值之和為1。表格底部的數值是共同試題$h_1(v)$的邊際分配。

例如，在群體1中，獲得共同試題分數為0的機率為0.20。$f_1(x)$列中的數值代表X測驗中總分的邊際分配。表格中每行數值之和等於$f_1(x)$所示的邊際值，$f_1(x)$的邊際分配之值總和等於1。最右邊的一列是X測驗分數的累積分佈$F_1(x)$。該列中的值是通過對$f_1(x)$列的機率進行累積而得到的。

表5.2給出了群體2中Y測驗和共同試題分數的聯合分配和邊際分配。

5.1 Frequency Estimation Method

P. 147

5.1.3 Numerical Example

Table 5.1 Form X and common-item distributions for population 1 in a hypothetical example

	v					
x	0	1	2	3	$f_1(x)$	$F_1(x)$
0	.04	.04	.02	.00	.10	.10
1	.04	.08	.02	.01	.15	.25
2	.06	.12	.05	.02	.25	.50
3	.03	.12	.05	.05	.25	.75
4	.02	.03	.04	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_1(v)$	.20	.40	.20	.20		

Note Values shown in the body of table are for $f_1(x, v)$

Table 5.2 Form Y and common-item distributions for population 2 in a hypothetical example

	v					
y	0	1	2	3	$g_2(y)$	$G_2(y)$
0	.04	.03	.01	.00	.08	.08
1	.07	.05	.07	.01	.20	.28
2	.03	.05	.12	.02	.22	.50
3	.03	.04	.13	.05	.25	.75
4	.02	.02	.05	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_2(v)$	.20	.20	.40	.20		

Note Values shown in the body of table are for $g_2(y, v)$

5.1.3 Numerical Example

表5.1和表5. 2中所列分配的估計值可從共同試題非等效組設計中獲得。由於群體2中沒有施測X測驗，因此無法估計X測驗在群體2中的分布情況。同樣地，在群體1中Y測驗的分配估計也不可用。然而，式子（5.4）中的次數估計假設仍然可以進行等化。

$$f_1(x|v)=f_2(x|v)$$
$$g_1(y|v)=g_2(y|v) \quad (5.4)$$

$$f_s(x) = w_1 f_1(x) + w_2 \sum_v f_1(x|v)h_2(v) \quad \text{and}$$
$$g_s(y) = w_1 \sum_v g_2(y|v)h_1(v) + w_2 g_2(y). \quad (5.8)$$

為了簡化示例，

假設w_1=1，式子（5.8）有了以下的簡化：

w_2=0

$$f_s(x) = f_1(x) \quad \text{and} \quad g_s(y) = \sum_v g_2(y|v)h_1(v). \quad (5.10)$$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

（5.10）的第一個式子 $f_s(x) = f_1(x)$說明，合成群體的X測驗的分數分配與群體1中的分配相同。因此，表5.1中 $f_1(x)$的最右邊一列也給出了w_1=1 的$F_s(x)$。

因為在示例中w1=1，合成群體是群體1，因此，式子（5.10）中的第二個式子提供了群體1中受試者Y形分數累積分佈的運算式。由於群體1中沒有施測Y測驗，因此有必要使用群體2中給定共同試題分數的Y測驗的條件分配，並假設該條件分配也適用於群體1中的所有共同試題得分[見式子（5.4）]。

表5.3給出了群體2的Y測驗的條件分配。為了計算表中的值，取表5.2中的聯合機率，並除以其在共同試題上的相關邊際機率。具體來說，

$$f_1(x|v) = f_2(x|v)$$
$$g_1(y|v) = g_2(y|v) \qquad (5.4)$$

$$g_2(y|v) = \frac{g_2(y, v)}{h_2(v)}, \qquad (5.11)$$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

Table 5.2 Form Y and common-item distributions for population 2 in a hypothetical example

y	v = 0	1	2	3	$g_2(y)$	$G_2(y)$
0	.04	.03	.01	.00	.08	.08
1	.07	.05	.07	.01	.20	.28
2	.03	.05	.12	.02	.22	.50
3	.03	.04	.13	.05	.25	.75
4	.02	.02	.05	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_2(v)$	.20	.20	.40	.20		

Note Values shown in the body of table are for $g_2(y, v)$

$$\frac{.04}{.2} = .2$$

Table 5.3 Conditional distributions of Form Y given common-item scores for population 2 in a hypothetical example

y	v = 0	1	2	3
0	.20	.15	.025	.00
1	.35	.25	.175	.05
2	.15	.25	.30	.10
3	.15	.20	.325	.25
4	.10	.10	.125	.30
5	.05	.05	.05	.30
$h_2(v)$	.20	.20	.40	.20

Note Values in the body of the table are for $g_2(y|v) = \frac{g_2(y,v)}{h_2(v)}$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

相乘後再加總

Table 5.1 Form X and common-item distributions for population 1 in a hypothetical example

	v					
x	0	1	2	3	$f_1(x)$	$F_1(x)$
0	.04	.04	.02	.00	.10	.10
1	.04	.08	.02	.01	.15	.25
2	.06	.12	.05	.02	.25	.50
3	.03	.12	.05	.05	.25	.75
4	.02	.03	.04	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_1(v)$	.20	.40	.20	.20		

Note Values shown in the body of table are for $f_1(x, v)$

Table 5.3 Conditional distributions of Form Y given common-item scores for population 2 in a hypothetical example

	v			
y	0	1	2	3
0	.20	.15	.025	.00
1	.35	.25	.175	.05
2	.15	.25	.30	.10
3	.15	.20	.325	.25
4	.10	.10	.125	.30
5	.05	.05	.05	.30
$h_2(v)$	.20	.20	.40	.20

Note Values in the body of the table are for $g_2(y|v) = \frac{g_2(y,v)}{h_2(v)}$

$$f_s(x) = f_1(x) \quad \text{and} \quad \boxed{g_s(y) = \sum_v g_2(y|v)h_1(v).} \qquad (5.10)$$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

Table 5.4 Calculation of distribution of Form Y and common-item scores for population 1 using frequency estimation assumptions in a hypothetical example

	v					
y	0	1	2	3	$g_1(y)$	$G_1(y)$
0	.20(.20) = .04	.15(.40) = .06	.025(.20) = .005	.00(.20) = .00	.105	.105
1	.35(.20) = .07	.25(.40) = .10	.175(.20) = .035	.05(.20) = .01	.215	.320
2	.15(.20) = .03	.25(.40) = .10	.30(.20) = .06	.10(.20) = .02	.210	.530
3	.15(.20) = .03	.20(.40) = .08	.325(.20) = .065	.25(.20) = .05	.225	.755
4	.10(.20) = .02	.10(.40) = .04	.125(.20) = .025	.30(.20) = .06	.145	.900
5	.05(.20) = .01	.05(.40) = .02	.05(.20) = .01	.30(.20) = .06	.100	1.000
$h_1(v)$	.20	.40	.20	.20		

Note Values in the body of the table are for $g_1(y, v) = g_2(y|v)h_1(v)$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

Table 5.5 Cumulative distributions and finding equipercentile equivalents for $w_1 = 1$

x	$F_1(x)$	$P_1(x)$	y	$G_1(y)$	$Q_1(y)$	x	$e_{Y_s}(x)$
0	.100	5.0	0	.105	5.25	0	−.02
1	.250	17.5	1	.320	21.25	1	.83
2	.500	37.5	2	.530	42.50	2	1.76
3	.750	62.5	3	.755	64.25	3	2.92
4	.900	82.5	4	.900	82.75	4	3.98
5	1.000	95.0	5	1.000	95.00	5	5.00

5.1 Frequency Estimation Method

5.1.3 Numerical Example

表5.5列出了累積分配、百分等級和等值的百分位數。這些值可以使用第2章中描述的計算程式進行驗證。

$$e_Y(x) = G^{-1}[F(x)], \tag{2.11}$$

表5.4，計算了群體1的Y測驗的總分和共同試題分數的聯合分配。如前所述，群體1沒有進行Y測驗。計算此表中數值的方法是通過與次數估計相關的統計假設來計算的。為了估計這一聯合分配，假設在群體2中觀察到的條件分配在所有共同試題得分下對群體1適用。郡體2的條件分配乘以群體1共同試題的邊際分配，形成表5.4所示的聯合機率。但是在共同試題上的分配可以被視為提供權重，在共同試題的每個分數上乘以群體2的條件分配。

5.1 Frequency Estimation Method

5.1.3 Numerical Example

如何找出$f_2(x)$與$g_1(y)$?

$$f_s(x) = w_1 f_1(x) + w_2 f_2(x)$$

$$g_s(y) = w_1 g_1(y) + w_2 g_2(y) \qquad (5.3)$$

$f_1(x)$

$f_2(x)$

X測驗

群體1

w_1

共同試題V

$h_1(v)$、$h_2(v)$

$w_1 + w_2 = 1$

$g_1(y)$

$g_2(y)$

Y測驗

群體2

w_2

5.1 Frequency Estimation Method

5.1.3 Numerical Example

如何找出$f_2(x)$與$g_1(y)$?

$$f_s(x) = w_1 f_1(x) + w_2 f_2(x)$$

$$g_s(y) = w_1 g_1(y) + w_2 g_2(y)$$

(5.3)

Assumptions

$$f_1(x|v) = f_2(x|v)$$

$$g_1(y|v) = g_2(y|v)$$

(5.4)

$$f_2(x,v) = f_1(x|v)h_2(v)$$
$$g_1(y,v) = g_2(y|v)h_1(v) \qquad (5.6)$$

透過條件機率

$$f_2(x) = \sum_v f_2(x,v) = \sum_v f_1(x|v)h_2(v) \quad \text{and}$$
$$g_1(y) = \sum_v g_1(y,v) = \sum_v g_2(y|v)h_1(v). \qquad (5.7)$$

$$f_s(x) = w_1 f_1(x) + w_2 \sum_v f_1(x|v)h_2(v) \quad \text{and}$$
$$g_s(y) = w_1 \sum_v g_2(y|v)h_1(v) + w_2 g_2(y). \qquad (5.8)$$

5.1 Frequency Estimation Method

5.1.3 Numerical Example

Table 5.1 Form X and common-item distributions for population 1 in a hypothetical example

x	v = 0	1	2	3	$f_1(x)$	$F_1(x)$
0	.04	.04	.02	.00	.10	.10
1	.04	.08	.02	.01	.15	.25
2	.06	.12	.05	.02	.25	.50
3	.03	.12	.05	.05	.25	.75
4	.02	.03	.04	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_1(v)$	.20	.40	.20	.20		

Note Values shown in the body of table are for $f_1(x, v)$

Table 5.2 Form Y and common-item distributions for population 2 in a hypothetical example

y	v = 0	1	2	3	$g_2(y)$	$G_2(y)$
0	.04	.03	.01	.00	.08	.08
1	.07	.05	.07	.01	.20	.28
2	.03	.05	.12	.02	.22	.50
3	.03	.04	.13	.05	.25	.75
4	.02	.02	.05	.06	.15	.90
5	.01	.01	.02	.06	.10	1.00
$h_2(v)$	.20	.20	.40	.20		

Note Values shown in the body of table are for $g_2(y, v)$

$$f_s(x) = w_1 f_1(x) + w_2 \sum_v f_1(x|v)h_2(v) \quad \text{and}$$

$$g_s(y) = w_1 \sum_v g_2(y|v)h_1(v) + w_2 g_2(y). \tag{5.8}$$

5.1 Frequency Estimation Method

$$f_1(x|v) = f_2(x|v) \qquad (5.4)$$

$$g_1(y|v) = g_2(y|v)$$

P. 150

5.1.4 Estimating the Distributions

在實際中使用次數估計時，可用分配估計來代替參數。然而，當一個組中沒有考生獲得某個特定的共同試題分數，而另一組中有一些考生獲得了該分數，則出現了一個問題。

當估計群體1中的Y測驗的分配時，式子（5.4）中假設所有v，$g_1(y|v) = g_2(y|v)$。如果沒有群體2的受試者在樣本中獲得關於v的特定分數，則該v不存在$g_1(y|v)$的估計值。

但是，如果群體1中的一些考生獲得了v. Jarjoura和Kolen（1985）建議使用接近v的條件分配（例如，v+1）作為條件分配在v處的估計值，則需要這樣的估計來進行等化，他們認為，這種替代會在實踐中造成不明顯的偏差，在這種情況下，一個群體中很少有考生的分數在另一個群體中的次數為0。

一個實際的解決方案是使用非零次數的v的條件分配，當我們向v分佈的中位數移動時，這個分配最接近所討論的v。

5.1 Frequency Estimation Method

5.1.4 Estimating the Distributions

平滑方法也可以與次數估計方法一起使用。Holland和Thayer（1987,1989, 2000)，von Davier等人（2004a），Rosenbaum和Thayer（1987）在次數估計的背景下描述了對數線性前平滑方法的延伸。在這延伸中，非共同試題的分數和共同試題的分數的聯合分配使用對數線性模型進行擬合。然後使用本章所述的次數估計方法等化平滑聯合分配。

使用這種方法進行模型擬合需要擬合一個聯合分配，這使得該方法的動差比隨機分組設計更複雜。為了計算聯合分配，需要指定與觀測分配相同的各擬合邊際分配的動差。此外，還需要規定與觀測分配相同的擬合聯合分佈的乘積。例如，可以指定一個模型，使每個邊際分配的前四個動差以及擬合和觀測分配的變異數相等。該模型的適配可以與其他更複雜和更不複雜的模型進行比較。

Moses和Holland（2010a，b）研究了使用對數線性前平滑聯合分佈的不同模型選擇方法。

5.1 Frequency Estimation Method

5.1.4 Estimating the Distributions

Lord's（1965）beta4方法也可用於計算總分和共同試題分數的聯合分配。在這個應用中，假設共同試題的真實分數和總測試的真實分數是函數相關的。也就是說，總測驗和共同試題量測的是完全相同的結構。Hanson(1991)、Livingston和Feryok（1987）、Liou和Cheng(1995)指出，二元平滑技術可以提高共同試題非等價群設計的等化精度。Kolen和Jarjoura(1987)描述的三次樣條函數後平滑方法是第3章所述隨機群方法的直接延伸。

在這種方法中，使用本章所述的次數估計來估計非平滑的相等百分位數法。然後實現第三章中描述的三次樣條函數法。方法上的唯一區別是，Jarjoura和Kolen（1985）提出的次數估計等化的標準誤差來代替隨機群的標準誤差。Kolen和Jarjoura（1987）報告了次數估計中使用的三次樣條函數法提高了等化精度。

5.1 Frequency Estimation Method

5.1.5 Special Case: Braun-Holland Linear Method

Braun和Holland(1982) 提出了一種線性方法，它利用次數估計假設產生的平均值和標準差進行線性等化。這方法與第四章提出的Tucker linear method密切相關。在式子（5.4）中的次數估計假設下，合成群體的X測驗分數的平均值和標準差可以表示為

$$\mu_s(X) = \sum_x x f_s(x), \tag{5.12}$$

$$\sigma_s^2(X) = \sum_x [x - \mu_s(X)]^2 f_s(x), \tag{5.13}$$

5.1 Frequency Estimation Method

5.1.5 Special Case: Braun-Holland Linear Method

$$\mu_s(X) = \sum_x x f_s(x), \tag{5.12}$$

$$\sigma_s^2(X) = \sum_x [x - \mu_s(X)]^2 f_s(x), \tag{5.13}$$

$$f_s(x) = w_1 f_1(x) + w_2 \sum_v f_1(x|v) h_2(v) \quad \text{and}$$

$$g_s(y) = w_1 \sum_v g_2(y|v) h_1(v) + w_2 g_2(y). \tag{5.8}$$

$f_s(x)$ 取自式子(5.8)，Y測驗的合成群體平均值和標準差表示類似。

用第4章中描述的共同項非等效群設計，得到的平均值和標準差可以用一般形式的線性等值關係代替，在這裡被稱為Braun and Holland linear method。

Braun和Holland（1982）指出，當下列條件成立時，使用Braun-Holland線性方法得出的值與第4章中描述的Tucker線性方法相同：

（1）*X*對*V*和*Y*對*V*的迴歸是線性的。

（2）X對V和Y對V的迴歸均為齊次回歸。這個性質意味著*X*給定*V*的變異數，對於所有*v*是相同的，*Y*給定*V*的變異數對於所有v都是相同的。

5.1 Frequency Estimation Method

5.1.5 Special Case: Braun-Holland Linear Method

當共同試題總檢驗回歸為非線性時，*Braun-Holland*方法可以看作是*Tucker*方法的推廣。

*Braun*和*Holland*（*1982*）建議群體1的*X*對*V*和群體2的*Y*對*V*的迴歸進行非線性檢驗。*Braun-Holland*方法在計算上比*Tucker*方法複雜，而且在實際應用中應用較少。

然而，當懷疑存在非線性迴歸時，應考慮使用*Braun-Holland*方法。

5.1 Frequency Estimation Method

5.1.5 Special Case: Braun-Holland Linear Method

Table 5.6 Computation of equating relationship for Braun-Holland method in a hypothetical example

From Table 5.1		From Table 5.4	
x	$f_1(x)$	y	$g_1(y)$
0	.100	0	.105
1	.150	1	.215
2	.250	2	.210
3	.250	3	.225
4	.150	4	.145
5	.100	5	.100
$\mu_1(X)$	2.5000	$\mu_1(Y)$	2.3900
$\sigma_1(X)$	1.4318	$\sigma_1(Y)$	1.4792

$$slope = \frac{1.4792}{1.4318} = 1.0331$$

$$intercept = 2.3900 - 1.0331(2.5000) = -.1927$$

$$l_{Y_s}(x = 0) = -.1927, l_{Y_s}(x = 1) = .8404, l_{Y_s}(x = 2) = 1.8735,$$
$$l_{Y_s}(x = 3) = 2.9066, l_{Y_s}(x = 4) = 3.9397, l_{Y_s}(x = 5) = 4.9728$$

5.1 Frequency Estimation Method

5.1.5 Special Case: Braun-Holland Linear Method

表*5.6* 給出了在w_1=1的次數估計範例中，使用*Braun-Holland*方法和假設數據的結果。在本表中，*X* 測驗的分配取自表 *5.1*。表*5.4*中採用了次數估計假設計算出的*Y*測驗分佈。平均值和標準差用等式（*5.12*）和（*5.13*）計算。根據平均值和標準差計算斜率和截距。利用該斜率和截距計算線性等化。注意，線性等化與表*5.5*中所示的相同百分位數有些不同，這表明當使用次數估計假設時，等化關係不是線性的。

5.1 Frequency Estimation Method

5.1.5 Special Case: Braun-Holland Linear Method

Table 5.6 Computation of equating relationship for Braun-Holland method in a hypothetical example

From Table 5.1			From Table 5.4
x	$f_1(x)$	y	$g_1(y)$
0	.100	0	.105
1	.150	1	.215
2	.250	2	.210
3	.250	3	.225
4	.150	4	.145
5	.100	5	.100
$\mu_1(X)$	2.5000	$\mu_1(Y)$	2.3900
$\sigma_1(X)$	1.4318	$\sigma_1(Y)$	1.4792

$$slope = \frac{1.4792}{1.4318} = 1.0331$$

$intercept = 2.3900 - 1.0331(2.5000) = -.1927$

$l_{Y_s}(x=0) = -.1927$, $l_{Y_s}(x=1) = .8404$, $l_{Y_s}(x=2) = 1.8735$,

$l_{Y_s}(x=3) = 2.9066$, $l_{Y_s}(x=4) = 3.9397$, $l_{Y_s}(x=5) = 4.9728$

Table 5.5 Cumulative distributions and finding equipercentile equivalents for $w_1 = 1$

x	$F_1(x)$	$P_1(x)$	y	$G_1(y)$	$Q_1(y)$	x	$e_{Y_s}(x)$
0	.100	5.0	0	.105	5.25	0	−.02
1	.250	17.5	1	.320	21.25	1	.83
2	.500	37.5	2	.530	42.50	2	1.76
3	.750	62.5	3	.755	64.25	3	2.92
4	.900	82.5	4	.900	82.75	4	3.98
5	1.000	95.0	5	1.000	95.00	5	5.00

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

第4章的實際數據例子說明了次數估計等值的一些方面。如該章所述，本例中使用的測試是36個項目的多項選擇測試。採用了兩種測驗—X測驗和Y測驗。測驗上的每三個項目都是一個共同試題，共同試題在每個表格上的位置相同。因此，第3、6、9項。每張表格上36個代表12個常見項目。1655名受試者使用X測驗，1638名受試者使用Y測驗。

Table 5.7 Moments for equating Form X and Form Y in the common-item nonequivalent groups design

Group	Score	$\hat{\mu}$	$\hat{\sigma}$	$\widehat{sk}$	$\widehat{ku}$	Correlation
1	X	15.8205	6.5278	.5799	2.7217	$\hat{\rho}_1(X, V) =$
1	V	5.1063	2.3760	.4117	2.7683	.8645
2	Y	18.6728	6.8784	.2051	2.3028	$\hat{\rho}_2(Y, V) =$
2	V	5.8626	2.4515	.1072	2.5104	.8753

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

Table 5.7 Moments for equating Form X and Form Y in the common-item nonequivalent groups design

Group	Score	$\hat{\mu}$	$\hat{\sigma}$	$\widehat{sk}$	$\widehat{ku}$	Correlation
1	X	15.8205	6.5278	.5799	2.7217	$\hat{\rho}_1(X,V) =$
1	V	5.1063	2.3760	.4117	2.7683	.8645
2	Y	18.6728	6.8784	.2051	2.3028	$\hat{\rho}_2(Y,V) =$
2	V	5.8626	2.4515	.1072	2.5104	.8753

本例統計如表*5.7*所示（*sk*表示估計的偏度，*ku*表示估計的峰度）。
接受*X*測驗的受試者在共同試題上的正確分數平均值為*5.1063*，標準差為*2.3760*。
採用*Y*測驗的受試者在共同試題上的正確分數平均值為*5.8626*，標準差為*2.4515*。

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

Table 4.3 Directly observable statistics for an illustrative example of equating forms X and Y using the common-item nonequivalent groups design

Group	Score	$\hat{\mu}$	$\hat{\sigma}$	Covariance	Correlation
1	X	15.8205	6.5278		
1	V	5.1063	2.3760	13.4088	.8645
2	Y	18.6728	6.8784		
2	V	5.8626	2.4515	14.7603	.8753

Note $N_1 = 1,655$ and $N_2 = 1,638$

Table 5.7 Moments for equating Form X and Form Y in the common-item nonequivalent groups design

Group	Score	$\hat{\mu}$	$\hat{\sigma}$	$\widehat{sk}$	$\widehat{ku}$	Correlation
1	X	15.8205	6.5278	.5799	2.7217	$\hat{\rho}_1(X, V) =$
1	V	5.1063	2.3760	.4117	2.7683	.8645
2	Y	18.6728	6.8784	.2051	2.3028	$\hat{\rho}_2(Y, V) =$
2	V	5.8626	2.4515	.1072	2.5104	.8753

根據共同試題的統計，施測*Y*試題的小組似乎比採用*X*試題的小組成績更高。本表所示的統計資料也用於計算第*4*章所述的塔克和萊文等值函數。表*5.7*所示的一些統計資料也列於表*4.3*。

分析使用附錄*B*中描述的*CIPE*電腦程式進行。

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

利用次數估計進行等化時，還需要考慮總分和共同試題得分的聯合分配。如本章前面所述，次數估計相等的假設，要求給定共同試題得分的總得分在兩個群體中的分配相同。然而，從收集到的數據來看，沒有數據可以直接解決這個假設。然而，一般項目總檢驗迴歸的線性關係可以解決。

如果迴歸是非線性的，那麼*Tucker*方法的使用可能是有問題的，而*Braun-Holland*方法可能是首選。

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

第一行列出了共同試題的可能得分。
第二行列出群體1中在共同試題中獲得每一分的考生人數。
第三行列出了在共同試題得分下*X*測驗的平均總分。例如，14名共同試題得分為零的考生在*X*測驗上的平均總分為*6.2143*。
第四行是標準差
第五行是基於使用標準線性迴歸估計*X*測驗給定共同試題*v*的平均值。

表*5.8* 顯示群體1施測*X*試題對*V*迴歸的相關統計資料

Table 5.8 Analysis of residuals from the linear regression of total score on common-item score for group 1

v	Number of examinees	Mean X given v	Standard Deviation X given v	Mean X given v, Linear regression	Residual mean
0	14	6.2143	2.2097	3.6923	2.5220
1	54	7.5741	2.2657	6.0674	1.5067
2	142	9.1901	2.6429	8.4425	.7476
3	249	10.8032	2.9243	10.8177	−.0145
4	274	12.7628	3.1701	13.1928	−.4300
5	247	15.1377	3.3302	15.5680	−.4303
6	232	16.9957	3.6982	17.9431	−.9474
7	173	20.5260	3.5654	20.3182	.2078
8	118	23.1610	3.5150	22.6934	.4676
9	75	25.6533	2.8542	25.0685	.5848
10	42	28.5000	3.4658	27.4436	1.0564
11	27	31.1852	2.1780	29.8188	1.3664
12	8	33.2500	1.6394	32.1939	1.0561

平均值隨著 *v* 共同試題分數的增加而增加。

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

Table 4.3 Directly observable statistics for an illustrative example of equating forms X and Y using the common-item nonequivalent groups design

Group	Score	$\hat{\mu}$	$\hat{\sigma}$	Covariance	Correlation
1	X	15.8205	6.5278		
1	V	5.1063	2.3760	13.4088	.8645
2	Y	18.6728	6.8784		
2	V	5.8626	2.4515	14.7603	.8753

Note $N_1 = 1,655$ and $N_2 = 1,638$

$$\textit{regression slope} = \hat{\rho}_1(X,V)\frac{\hat{\sigma}_1(X)}{\hat{\sigma}_1(V)} = .8645\frac{6.5278}{2.3760} = 2.3751.$$

$$\textit{regression intercept} = \hat{\mu}_1(X) - (\textit{regression slope})\,\hat{\mu}_1(V)$$
$$= 15.8205 - (2.3751)\,5.1063 = 3.6923,$$

平均殘差等於第三行減去第五行。

ex: 6.2143-3.6923=2.5220

斜率和截距可用於生成第五行中的值

平均殘差的數值代表使用線性迴歸預測的平均與觀察到的平均不同的程度。

Table 5.8 Analysis of residuals from the linear regression of total score on common-item score for group 1

v	Number of examinees	Mean X given v	Standard Deviation X given v	Mean X given v, Linear regression	Residual mean
0	14	6.2143	2.2097	3.6923	2.5220
1	54	7.5741	2.2657	6.0674	1.5067
2	142	9.1901	2.6429	8.4425	.7476
3	249	10.8032	2.9243	10.8177	−.0145
4	274	12.7628	3.1701	13.1928	−.4300
5	247	15.1377	3.3302	15.5680	−.4303
6	232	16.9957	3.6982	17.9431	−.9474
7	173	20.5260	3.5654	20.3182	.2078
8	118	23.1610	3.5150	22.6934	.4676
9	75	25.6533	2.8542	25.0685	.5848
10	42	28.5000	3.4658	27.4436	1.0564
11	27	31.1852	2.1780	29.8188	1.3664
12	8	33.2500	1.6394	32.1939	1.0561

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

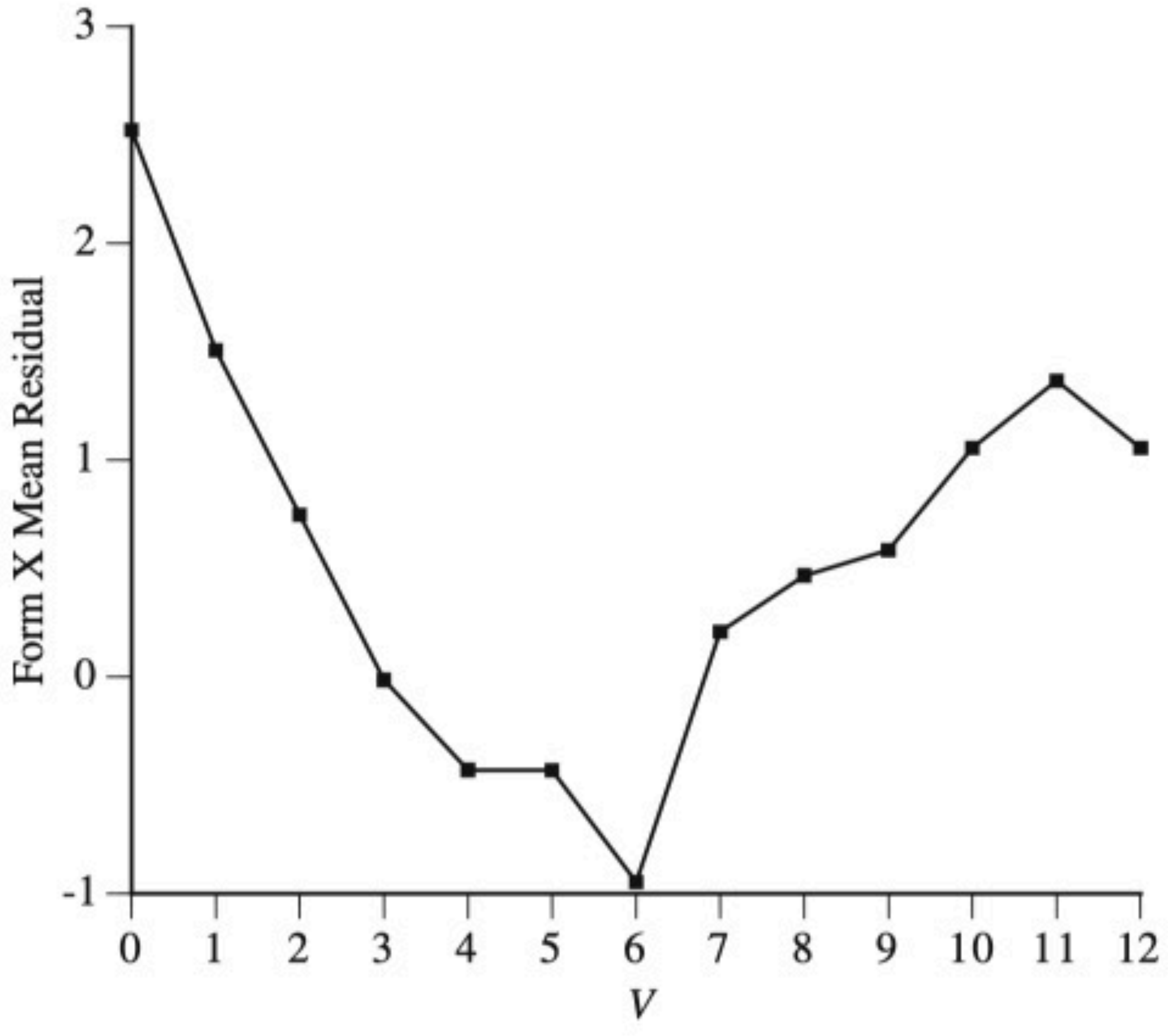


Fig. 5.1 Form X mean residual plot

*X*測驗中平均殘差的圖示如*5.1*

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

如果迴歸真的是線性的，那麼平均殘差會在*0*左右隨機變化。然而，殘差平均對於*v*的低分和高分是正的，對於*3*到*6*分的分數是負的。這種模式表明迴歸不是線性的。

也可以使用更複雜的方法來測試關於迴歸線性的假設（例如，見*Draper*和*Smith 1998*）。群體*2*的*Y*對*V*的迴歸如表*5.9*所示，平均殘差如圖*5.2*所示。這種迴歸似乎也有點非線性。這些非線性迴歸說明 *Braun-Holland Linear Method*可能比*Tucker method* 更適合。

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

Table 5.9 Analysis of residuals from the linear regression of total score on common-item score for group 2

v	Number of examinees	Mean Y given v	Standard Deviation Y given v	Mean Y given v, Linear regression	Residual mean
0	11	6.2727	2.1780	4.2740	1.9988
1	36	8.0000	2.2361	6.7300	1.2700
2	88	9.6023	3.0359	9.1860	.4162
3	159	12.1195	3.2435	11.6421	.4774
4	213	13.9202	3.3929	14.0991	−.1779
5	240	16.0750	3.4234	16.5541	−.4791
6	232	18.3147	1.5623	19.0101	−.6955
7	246	21.2073	3.4854	21.4662	−.2588
8	161	24.1801	3.3731	23.9222	.2579
9	120	27.3333	2.9533	26.3782	.9551
10	85	29.1294	2.8811	28.8343	.2952
11	34	31.8235	1.8396	31.2903	.5332
12	13	33.6154	1.7338	33.7463	−.1309

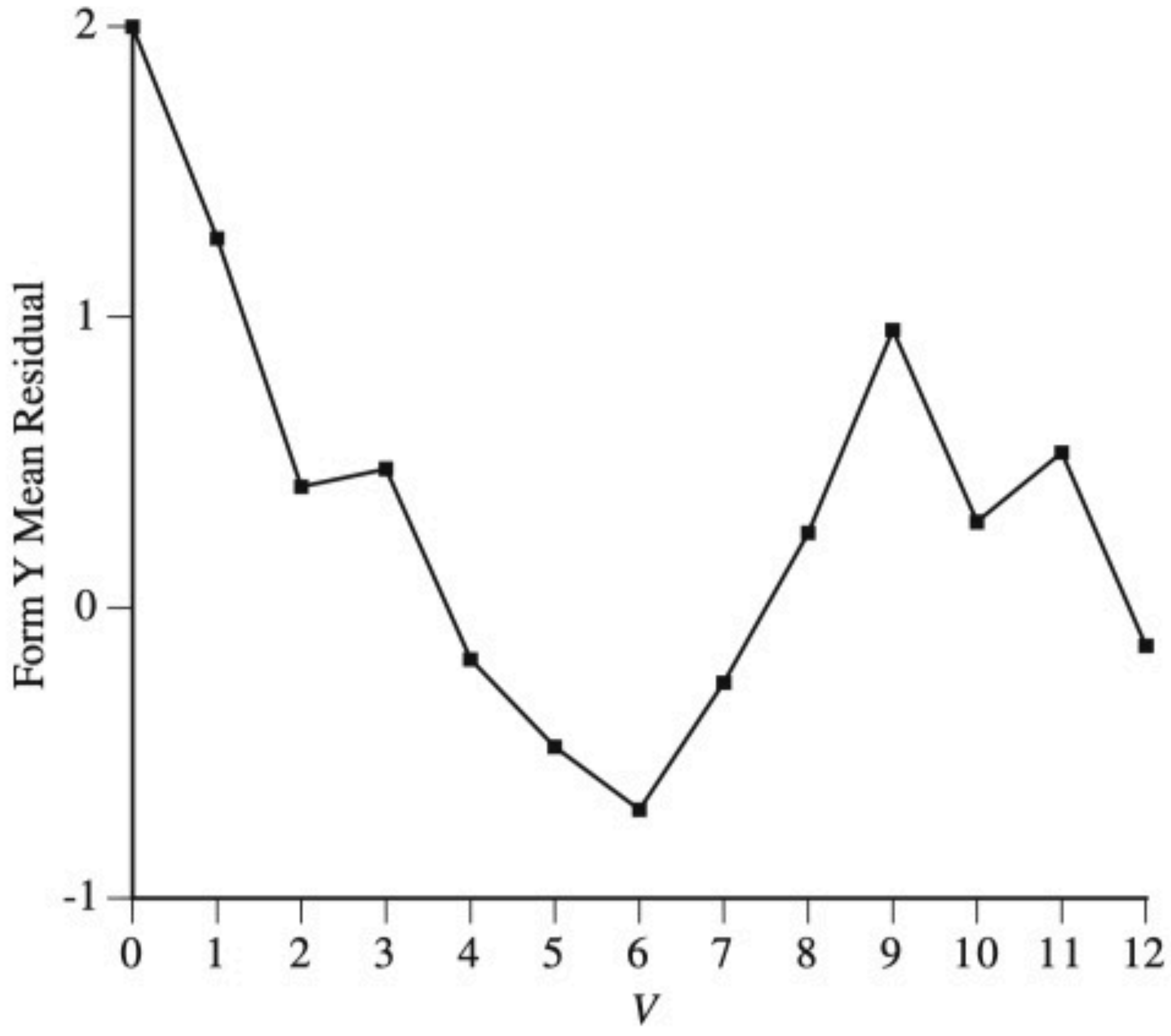


Fig. 5.2 Form Y mean residual plot

5.1 Frequency Estimation Method

5.1.6 Illustrative Example Comparison Among Methods

這些數據均採用*Tucker*和*Braun-Holland*線性化方法和三次樣條平滑相同百分位數次數估計方法。在同類模型下採用*Levine*觀察評分法。
所得動差如表*5.10*所示，等化關係如圖*5.3*所示。

Table 5.10 Moments of Form X scores converted to Form Y scores using various methods for examinees from population 1

Method	$\hat{\mu}$	$\hat{\sigma}$	$\widehat{sk}$	$\widehat{ku}$
Tucker linear	16.8153	6.7168	.5799	2.7217
Levine linear	16.2485	6.6007	.5799	2.7217
Braun-Holland linear	16.8329	6.6017	.5799	2.7217
Equipercentile				
Unsmoothed	16.8329	6.6017	.4622	2.6229
$S = .10$	16.8334	6.5983	.4617	2.6234
$S = .25$	16.8333	6.5947	.4674	2.6249
$S = .50$	16.8192	6.5904	.4983	2.6255
$S = .75$	16.8033	6.5858	.5286	2.6503
$S = 1.00$	16.7928	6.5821	.5501	2.6745

5.1 Frequency Estimation Method

5.1.6 Illustrative Example Comparison Among Methods

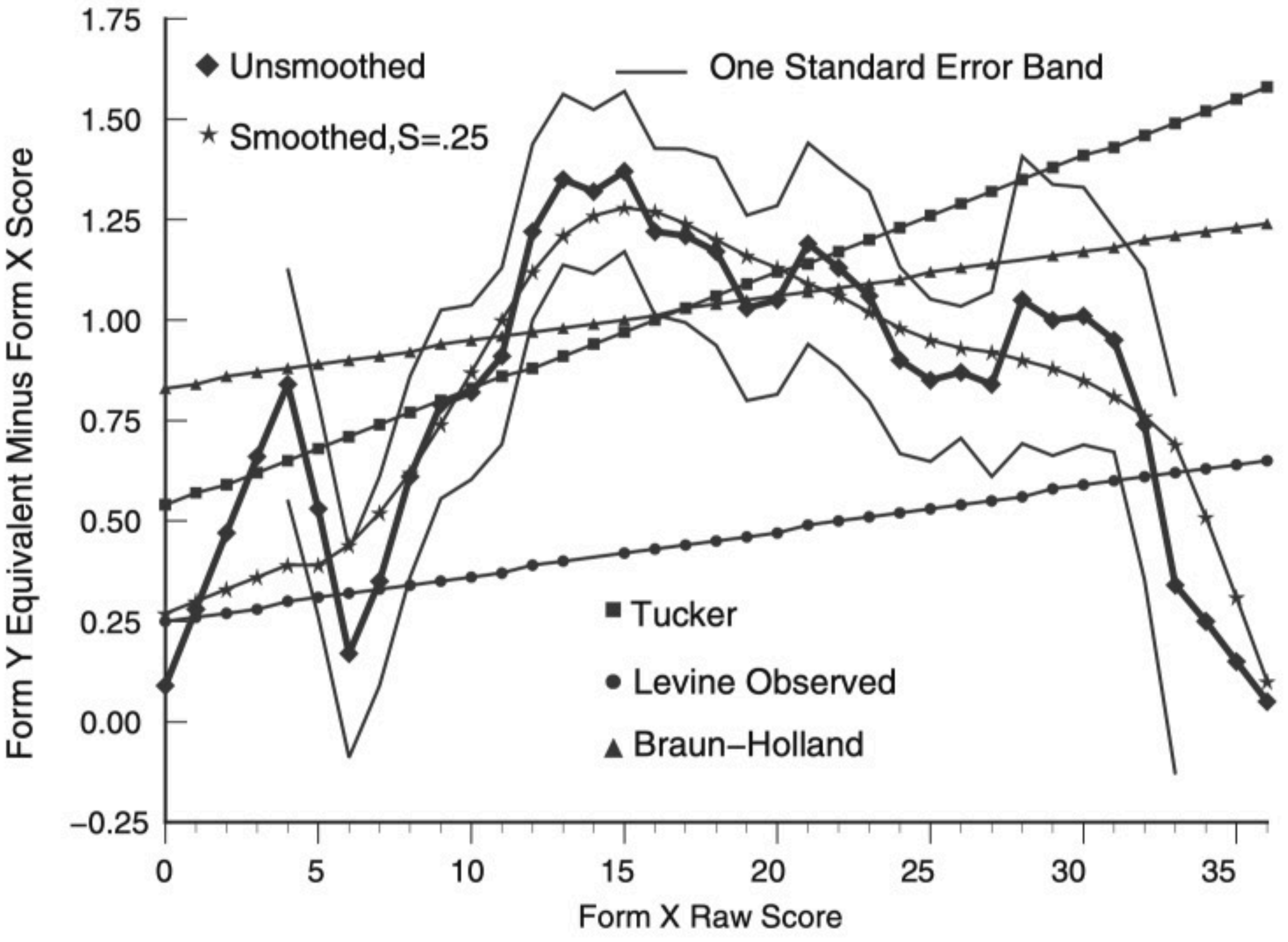


Fig. 5.3 Equating relationships for frequency estimation equipercentile equating and linear methods

5.1 Frequency Estimation Method

5.1.6 Illustrative Example　Comparison Among Methods

首先，圖*5.3*，*The Levine relationship*似乎與其他人不同。如第*4*章所述，*Levine*方法是基於對真實分數的假設，而其他方法是對觀察分數進行假設。假設的差異很可能是造成這種差異的原因。不幸的是，無法得到數據，無法判斷*Levine*方法假設（可能的線性迴歸除外）是否比本例中其他方法的假設更適合。

Tucker、*Braun-Holland* 和 *frequency estimation methods* 都要求假設對總得分與共同試題得分之間的關係，在兩個群體中是相同的。首先考慮*Tucker*和*Braun-Holland*方法。這些方法之間的主要區別在於迴歸線性的假設。因此，範例中兩種方法之間的相對較小差異是由於假設差異所致。最好使用*Braun-Holland*方法，因為該迴歸被認為是非線性的。

5.1 Frequency Estimation Method

5.1.6 Illustrative Example　Comparison Among Methods

接下來，在表*5.10*和圖*5.3*中比較*Braun-Holland*和*frequency estimation methods*（稱為不平滑方法）。該關係似乎是非線性的。在分數範圍的部分上，*Braun-Holland*關係超出了*frequency estimation methods* 的標準誤差範圍。因此，與本例中的*Braun-Holland*方法相比，*frequency estimation methods* （不平滑法）似乎可以更準確地反映不同測驗間相同百分位數的關係。

表*5.10* 給出了不同程度的三次樣條平滑的結果。*S*值大於*.25*的動差似乎與不平滑的等式所期望的動差相差更大。

Table 5.10 Moments of Form X scores converted to Form Y scores using various methods for examinees from population 1

Method	$\hat{\mu}$	$\hat{\sigma}$	$\widehat{sk}$	$\widehat{ku}$
Tucker linear	16.8153	6.7168	.5799	2.7217
Levine linear	16.2485	6.6007	.5799	2.7217
Braun-Holland linear	16.8329	6.6017	.5799	2.7217
Equipercentile				
Unsmoothed	16.8329	6.6017	.4622	2.6229
$S = .10$	16.8334	6.5983	.4617	2.6234
$S = .25$	16.8333	6.5947	.4674	2.6249
$S = .50$	16.8192	6.5904	.4983	2.6255
$S = .75$	16.8033	6.5858	.5286	2.6503
$S = 1.00$	16.7928	6.5821	.5501	2.6745

5.1 Frequency Estimation Method

5.1.6 Illustrative Example

Comparison Among Methods

圖5.3中繪製了$S = .25$的關係。這種關係保持在標準誤差範圍內，並且看起來很平滑，並且與未平滑的值相差不大。

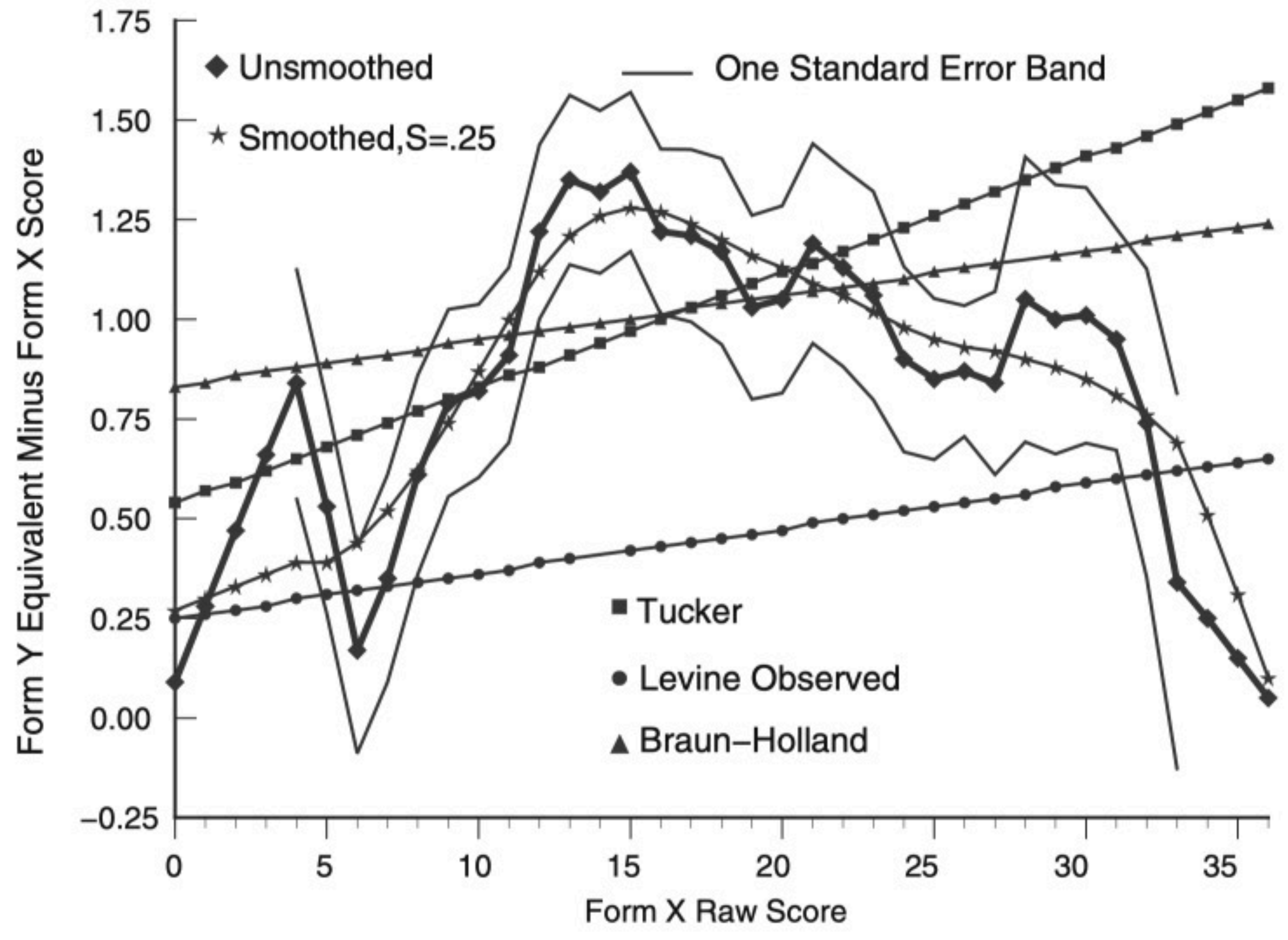


Fig. 5.3 Equating relationships for frequency estimation equipercentile equating and linear methods