7章

（森 棟） 単回帰　$y_i = \beta X_i + u_i$　においてβの最小２乗推定量は

$$\hat{\beta} = \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \sum_{i=1}^{n} X_i y_i$$

となるが、この推定量の最良線形不偏性を証明してみよう．

最良線形不偏性

証明１：　推定量の展開は $\hat{\beta} = \beta + \frac{1}{\Sigma X_i^2} \Sigma X_i u_i$ となる．w_i の定義を用いれば $\hat{\beta} = \sum_{i=1}^{n} w_i y_i$ となる．Johnston 流の証明では $\hat{\beta}$ と異なる推定量は $\tilde{\beta} = \sum_{i=1}^{n} (w_i + d_i) y_i$ とする．ここで $\tilde{\beta}$ の不偏性の条件より

$$(1) \qquad \sum_{i=1}^{n} d_i X_i = 0$$

となる．さて $\tilde{\beta}$ の分散を求めると（１）より

$$E\left\{ \sum_{i=1}^{n} (w_i + d_i) u_i \right\}^2 = \sigma^2 \Sigma (w_i^2 + d_i^2)$$

となる．したがってすべての d が 0 の時に分散が最小になる．（一意性）

証明２：ラグランジュ未定定数法による証明を考えると、$\tilde{\beta} = \sum_{i=1}^{n} c_i y_i$ が線型推定量で、不偏性の条件より $\sum_{i=1}^{n} c_i X_i = 1$ が制約となる．分散は $\sigma^2 \sum_{i=1}^{n} c_i^2$ だから

$$J = \sum_{i=1}^{n} c_i^2 + \lambda \left(\sum_{i=1}^{n} c_i X_i - 1 \right)$$

を最小化するためのすべての c について一次の条件を導く．

$$(2) \qquad \frac{\partial J}{\partial c_i} = 2c_i + \lambda X_i = 0, \qquad i = 1, \cdots, n$$

$$(3) \qquad \frac{\partial J}{\partial \lambda} = \Sigma c_i X_i - 1 = 0$$

が一次の条件だが、(2) より $c_i = -(\lambda/2) X_i$ だから、これを (3) に代入して λ を求めると、

(4)　　$\lambda = -2\left(\sum_{i=1}^{n} X_i^2\right)^{-1}$

となる.だから $c_i = \left(\sum_{i=1}^{n} X_i^2\right)^{-1} X_i$ と求まり、w_i と等しいことがわかる.

証明3：第３の方法は Cauchy-Schwarz の不等式による.一般の線型推定量の係数を $c_i(i=1,\cdots,n)$ とするが、Cauchy- Schwarz の不等式により

$(\sum_{i=1}^{n} c_i^2)(\sum_{i=1}^{n} X_i^2) \geq (\sum_{i=1}^{n} c_i X_i)^2$ となる.（この証明は$\Sigma(t\,c_i + X_i)^2 \geq 0$であることから、t についての２次式の判別式より導かれる.）ところが、不偏性により右辺は１であるから $\Sigma c_i^2 \geq \dfrac{1}{\Sigma X_i^2}$ となる.左辺からは一般の線型推定量の分散が導かれ、右辺からは最小２乗推定量の分散が導かれる.

練習問題(7章)

1.　任意の座標値を(2,3)とする.　(7.6)式に代入すると,y－3＝$\dfrac{7-3}{5-2}(x-2)$より

a＝3－$\left(\dfrac{4}{3}\right)\times 2 = \dfrac{1}{3}$、b＝$\dfrac{4}{3}$. 他方、 (2,3)と(5,7)を正規方程式に代入すると、

2α＋7β－10＝0、29β＋7α－41＝0,となり、α＝$\dfrac{1}{3}$、β＝$\dfrac{4}{3}$.　(7.6)式の a と b の係数値に一致する.

2.　大気中の二酸化炭素濃度を従属変数、化石燃料の消費量を独立変数として線形回帰式を推定すると、

二酸化炭素濃度＝292.42(257.37)＋0.844(21.88)化石燃料の消費量

と計算される. R^2＝0.976 だから、二酸化炭素濃度の変化の 97.6%が化石燃料消費量の変化によって説明できるという結果を得る.

3.　オゾン濃度＝373.81(9.36)－1.31(-2.06)フロンガス, R^2＝0.516 と計算される.

4.　$\sum(y_i - a)^2$ の最小化より、$3a = 3\bar{y}$.　a＝$\bar{y}$と求まる.

5.　回帰式の形より、RSS=TSS になる.$R^2=0$.　問４の結果より，$a=\bar{y}$.　$\bar{y}=\alpha+\bar{U}$ より，$E(\bar{U})=\sum_{i=1}^{n}\frac{1}{n}E(U_i)=0$ に注意すると、$E(a)=E(\alpha+\bar{U})=\alpha$.分散は $\sigma_a^2=E(\alpha-a)^2=E(\bar{U})^2$. $V(\bar{U})=\sum_{i=1}^{n}\frac{1}{n^2}V(U_i)=\sum_{i=1}^{n}\frac{\sigma^2}{n^2}=\frac{\sigma^2}{n}$ より $\sigma_a^2=\frac{\sigma^2}{n}$ と求まる.

6.　標本相関係数に $\frac{S_y}{S_x}$ を掛けると、

$$r\times\frac{S_y}{S_x}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}\frac{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=b$$

7.　(7.37)　$\sum_{i=1}^{n}y_i-an-b\sum_{i=1}^{n}x_i-c\sum_{i=1}^{n}z_i=0$,

(7.38)　$\sum_{i=1}^{n}y_ix_i-a\sum_{i=1}^{n}x_i-b\sum_{i=1}^{n}x_i^2-c\sum_{i=1}^{n}z_ix_i=0$

(7.39)　$\sum_{i=1}^{n}y_iz_i-a\sum_{i=1}^{n}z_i-b\sum_{i=1}^{n}x_iz_i-c\sum_{i=1}^{n}z_i^2=0$

と書きなおせる.行列式でまとめると、

$$\begin{pmatrix}\sum_{i=1}^{n}y_i\\ \sum_{i=1}^{n}x_iy_i\\ \sum_{i=1}^{n}y_iz_i\end{pmatrix}-\begin{pmatrix}n & \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}z_i\\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_iz_i,\\ \sum_{i=1}^{n}z_i & \sum_{i=1}^{n}x_iz_i & \sum_{i=1}^{n}z_i^2\end{pmatrix}\begin{pmatrix}a\\ b\\ c\end{pmatrix}=0.$$

8.　$\Psi=\sum(x_i-a-dy_i)^2$ を、a と d について１階微分すると、正規方程式は $\frac{\partial}{\partial a}\Psi=an-d\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}x_i=0$, $\frac{\partial}{\partial d}\Psi=d\sum_{i=1}^{n}y_i^2-\sum_{i=1}^{n}x_iy_i-a\sum_{i=1}^{n}y_i=0$．２式より、a と d を求める． $d=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{y}\bar{x}}{\left(\sum_{i=1}^{n}y_i^2-n\bar{y}^2\right)}$. したがって、$\hat{d}\times\hat{b}=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{y}\bar{x}}{\left(\sum_{i=1}^{n}y_i^2-n\bar{y}^2\right)}\times\frac{\sum_{i=1}^{n}x_iy_i-n\bar{y}\bar{x}}{\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)}$

$=r^2$.

9. (i) $X^TX=\begin{pmatrix}1,\cdots,1\\ x_1,\cdots,x_n\end{pmatrix}\begin{pmatrix}1,x_1\\ \vdots,\vdots\\ 1,x_n\end{pmatrix}=\begin{pmatrix}1+1+\cdots 1, x_1+x_2+\cdots x_n\\ x_1+x_2+\cdots x_n, x_1^2+x_2^2+\cdots x_n^2\end{pmatrix}=\begin{pmatrix}n,\sum_{i=1}^n x_i\\ \sum_{i=1}^n x_i,\sum_{i=1}^n x_i^2\end{pmatrix}$

$$X^T y=\begin{pmatrix}1,\cdots,1\\ x_1,\cdots,x_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\begin{pmatrix}y_1+y_2+\cdots y_n\\ x_1y_1+x_2y_2+\cdots x_ny_n\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^n y_i\\ \sum_{i=1}^n x_iy_i\end{pmatrix}$$

X^TX に右側から d を掛けると、(7.10)式になる.(ii)は

$$X^TX=\begin{pmatrix}1&\cdots&1\\ x_1&\cdots&x_n\\ z_1&\cdots&z_n\end{pmatrix}\begin{pmatrix}1&x_1&z_1\\ \vdots&\vdots&\vdots\\ 1&x_n&z_n\end{pmatrix}=\begin{pmatrix}1+\cdots+1 & x_1+\cdots+x_n & z_1+\cdots+z_n\\ x_1+\cdots+x_n & x_1^2+\cdots x_n^2 & x_1z_1+\cdots+x_nz_n\\ z_1+\cdots+z_n & z_1x_1+\cdots+z_nx_n & z_1^2+\cdots+z_n^2\end{pmatrix}$$

$$=\begin{pmatrix}n & \sum_{i=1}^n x_i & \sum_{i=1}^n z_i\\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_iz_i\\ \sum_{i=1}^n z_i & \sum_{i=1}^n x_iz_i & \sum_{i=1}^n z_i^2\end{pmatrix}.$$

$$X^Ty=\begin{pmatrix}1,\cdots,1\\ x_1,\cdots,x_n\\ z_1,\cdots,z_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\begin{pmatrix}y_1+\cdots+y_n\\ x_1y_1+\cdots+x_ny_n\\ z_1y_1+\cdots+z_ny_n\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^n y_i\\ \sum_{i=1}^n x_iy_i\\ \sum_{i=1}^n z_iy_i\end{pmatrix}.$$

X^TX に右側から d を掛けると、(7.41)式になる.

10. アーティフィシァルに次の 2 乗和を考える。

$$\sum_{i=1}^n (x_i-ty_i)^2=t^2\sum_{i=1}^n (y_i^2)-2t\sum_{i=1}^n (x_iy_i)+\sum_{i=1}^n (x_i^2).$$

この **t に関する２次式**は常に非負であるから、判別式は正か負でないといけない. だから

$$D=(\sum_{i=1}^n x_iy_i)^2-\sum_{i=1}^n (x_i^2)\sum_{i=1}^n (y_i^2)\le 0$$

が成立する.これがコーシー・シュワルツの不等式である。

他方,判別式が０になるならば,$\sum_{i=1}^{n}(x_i - ty_i)^2 = 0$となる重根 t が存在する.（２次式が横軸に接する.）したがって,この重根 t^*について、$\sum_{i=1}^{n}(x_i - t^* y_i)^2 = 0$である.

したがって,すべての i について　$x_i = ty_i$　が成立する.

8 章

全体として統計学辞典（東洋経済）の 124 ページあたりが役に立つ

順位和検定

（Mann-Witney の検定）（統計数値表　解説編の 143P を参照.）

符号検定

成功確率が$\frac{1}{2}$の二項確率関数からの自明（統計数値表 140P を参照.）

符号つき順位和検定

（Wilcox の検定）　（統計数値表 141P を参照.）

クラスカル･ワリス検定

第 2 項の 3 はクラスの数と無関係.

スピアマン順位相関係数

(8.22)式で数値計算するか、(8.23)式で数値計算するかは好みによるが、Excel を使う場合は(8.22)式を用いればよい。標本相関係数のルーチンが使えるからである。

Excel が使えるなら、順位間の相関係数を求めるだけであるから、(8.23)式は不要になる。(8.22)式は、

$$\sum_{i=1}^{n}(R_i - \frac{n+1}{2})^2 = \sum_{i=1}^{n} R_i^2 - \frac{1}{4}n(n+1)^2 = \frac{1}{6}n(n+1)(2n+1) - \frac{1}{4}n(n+1)^2 = \frac{1}{12}n(n+1)(n-1)$$

となる. $\sum_{i=1}^{n}(R_i - \frac{n+1}{2})^2 = \sum_{i=1}^{n}(Q_i - \frac{n+1}{2})^2$　も明らかであるから、この逆数が統計

量に入ってくる．次に分子は、

$$\sum_{i=1}^{n}(R_i - \frac{n+1}{2})(Q_i - \frac{n+1}{2}) = \sum_{i=1}^{n}(R_i - Q_i)(Q_i - \frac{n+1}{2}) + \sum_{i=1}^{n}(Q_i - \frac{n+1}{2})^2$$

であり、第 2 項は分母に等しい.第 1 項は$\sum_{i=1}^{n}(R_i - Q_i)Q_i$、である.構成からして第 1 項は、$\frac{1}{2}\{\Sigma(R_i - Q_i)Q_i + \Sigma(Q_i - R_i)R_i\} = -\frac{1}{2}\sum_{i=1}^{n}(R_i - Q_i)^2$となる．以上を合成して、(8.23)式が導かれる。

例 8.4 については、E 表および$(E-O)^2/E$ 表は上の様になる。E 表は、「行和」×

E			$(E-O)^2/E$		
6.382579	40.79675	53.82067	3.340433	1.649475	3.054022
9.668659	61.80102	81.53032	1.940343	0.010382	0.152865
14.4082	92.09564	121.4962	2.029998	0.281942	0.9081
16.93595	108.2528	142.8113	2.080517	0.62916	1.409693
14.53459	92.9035	122.5619	0.859556	0.164012	0.451406
8.215201	52.51067	69.27412	0.943995	2.095306	2.543553
3.854825	24.63962	32.50555	2.566167	0.016604	0.19313
74	473	624	13.76101	4.846882	8.712768

「列和」÷「総和」による。最終行は列和で、統計量の総和は 27.3 である。19

E

71.9	323.1	237.4	58.8	33.4	724.5
64.4	289.7	212.9	52.7	29.9	649.7
64.7	291.0	213.9	53.0	30.1	652.7
64.4	289.3	212.6	52.7	29.9	648.7
59.3	266.6	195.9	48.5	27.6	597.8

$(E-O)^2/E$

3.5	0.0	10.0	25.2	19.3	58.0
0.2	0.0	6.0	4.6	8.2	19.0
0.2	0.8	4.1	2.9	0.7	8.7
0.1	0.8	2.7	0.3	2.0	5.9
0.1	0.0	2.0	5.8	2.9	10.8
					102.0

歳以下を除くと、E 表及び$(E-O)^2/E$ 表は次の様になり、検定統計量は 19.2 になる。

E			$(E-O)^2/E$		
9.008411	60.62804	83.36355	2.765855	0.002282	0.345087
13.4243	90.34766	124.228	1.458134	0.124041	0.48623
15.77944	106.1981	146.0224	1.447646	0.361747	0.825264
13.54206	91.14019	125.3178	0.477184	0.050257	0.174942
7.654206	51.51402	70.83178	1.462508	2.561007	3.10569
3.591589	24.17196	33.23645	3.234576	0.001223	0.315154
63	424	583	10.8459	3.100557	5.252368

例 8.5

　この例では数値の微妙な不整合があり、計算の仕方によって答えが変わる。上の表では「行和」×「列和」によってE表を作成した。E表から分かるように、行和を再計算しても列和は元のデータをもたらさない。そのまま計算すると、上の表の結果をもたらす。

沖縄	熊本	新潟	香川	神奈川	大分	宮城	山形	宮崎	群馬	千葉	東京	北海道	山口
81.2	80.9	80.3	80.3	80.3	80.2	80.2	80.1	80.1	80.1	80.0	80.0	80.0	80.0
S	S	N	S	N	S	N	N	S	N	N	N	N	S
1	2	3	4	5	6	7	8	9	10	11	12	13	14
埼玉	岩手	高知	愛媛	佐賀	福岡	鹿児島	福島	徳島	長崎	茨城	秋田	栃木	青森
79.9	79.9	79.9	79.9	79.9	79.8	79.8	79.7	79.7	79.7	79.6	79.5	79.4	78.6
N	N	S	S	S	S	S	N	S	S	N	N	N	N
15	16.5	16.5	18	19	20	21	22	23.5	23.5	25	26	27	28

例 8.6

この表では、平均余命順に２８県を並べてある。（データは小数一桁まで。）あとは指示に従って統計量を計算する。北は228.5、南は177.5となる。

例 8.9

大岡川の順位が26.5になっているが、これは27.5の間違い。従って T_1=342、T_2=247.5、T_3=190.5となる。K=6.93、P=0.031。

例 8.10

　各国の出生率と幼児死亡率の順位を次の表で与える。順位間の相関係数を計算すれば、r=0.86、t=6.5となる。

国番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
B出生率	10	15	21	22	33	27	24	13	17	17	22	28	36	29	21	26	15
順位	1	3.5	7.5	9.5	16	13	11	2	5.5	5.5	9.5	14	17	15	7.5	12	3.5
D幼児死亡率	4	7	7	32	112	71	47	4	33	32	40	75	95	35	21	11	9
順位	1.5	3.5	3.5	8.5	17	14	13	1.5	10	8.5	12	15	16	11	7	6	5

練習問題 (8 章)

1. (1,1)セルであれば、理論値は$\frac{101\times18}{162}$=11.22.　だから$(O-E)^2/E$ ＝0.0044.

以下全てのセルについて同じ計算をすると、総和は 1.36.　自由度 2 のχ^2分布の上側 5%点は 5.99 だから 10 代の年齢層の行動が 70 代のと同じであるという帰無仮説は棄却できない.

O				E			$(O-E)^2/E$		
11	49	41	101	11.22	45.51	44.27	0.00	0.27	0.24
7	24	30	61	6.78	27.49	26.73	0.01	0.44	0.40
18	73	71	162	18	73	71	0.01	0.71	0.64

2. (1,1)セルであれば､理論値は$\frac{654\times131.07}{1302.70}$=65.7.　だから$(O-E)^2/E$ ＝0.131.
以下全てのセルについて同じ計算をすると､総和は 13.7.　自由度 4 のχ^2分布の上側 5%点は 9.49 だから､第 3 階級と第 4 階級の貯蓄行動が同じであるという帰無仮説は棄却される.

3. 東北地方だけの順位を書き出す. 13、6、3、12、4、10 これらの順位数の和は$T_1=48,\ n_1=6,\ n_2=7$, だから(8.16), (8.17)から$E(T_1)=42,\ \sigma^2=49$.従って

$$z_1=\frac{T_1-E(T_1)}{\sigma}=0.86$$

有意水準が 5%の両側検定では、 Z_1の値が絶対値で 1.96 よりも小だから、帰無仮説は棄却できない.二グループ間の平均寿命に差異はないと判断される.

青森	岩手	宮城	秋田	山形	福島	
78.61	79.88	80.16	79.52	80.11	79.7	
13	6	3	12	4	10	
福岡	佐賀	長崎	熊本	大分	宮崎	鹿児島
79.78	79.85	79.69	80.85	80.22	80.1	79.75
8	7	11	1	2	5	9

4. $T_1=79,\ T_2=65,\ T_3=87$, だから

$$K = \frac{12}{n(n+1)}(\frac{1}{n_1}T_1^2 + \frac{1}{n_2}T_2^2 + \frac{1}{n_3}T_3^2) - 3(n+1) = \frac{12(79^2 + 65^2 + 87^2)}{21 \times 22 \times 7} - 3(22) = 0.92$$

自由度が 2 の χ^2 分布の 5%点は 5.99 だから帰無仮説は棄却でかない. 異なる 3 種の効果は異なるとはいえない.

	順位							順位和
肥料1	2	20	7	1	16	21	12	79
肥料2	4.5	10	14	4.5	6	8	18	65
肥料3	13	19	17	11	3	15	9	87

5.　出生率と所得の順位は以下のようになる。順位間の相関係数を計算すればよい。$r_S = 0.68$、t 値は 3.5.　自由度 14 の t 分布より、相関がないという帰無仮説は棄却される.

国	出生率	順位	国	所得	順位
1	10	1	1	38160	1
2	15	3.5	2	29080	3
3	21	7.5	3	16180	4
4	22	9.5	4	1780	9
5	27	13	5	370	14
6	24	11	6	1110	11
7	13	2	7	32810	2
8	17	5.5	8	2740	8
9	17	5.5	9	860	12
10	22	9.5	10	3130	7
11	28	14	11	360	15
12	36	16	12	500	13
13	29	15	13	1200	10
14	21	7.5	14	310	16
15	26	12	15	4530	6
16	15	3.5	16	10550	5

6. 順位相関係数などを計算する.

7. 観測値から言って、アルカリ性の結果が全くないのだから、符号検定の結果を待たずとも分布の非対称性は明らかであろう.　符号統計量(8.19)では、$Z = \frac{1}{\sqrt{n}/2}(X - \frac{n}{2}) = \frac{1}{\sqrt{92/4}}(91 - \frac{92}{2}) = 9.38$.　分布が対称であるという帰無仮

説が棄却される．しかし、これでは分析にはならない． 実際酸性雨とは ph が 5.6 以下を言う。境界を 5.5 とすると分析はどうなるか。厳しい基準としては pH4.5 がとられることもある。この場合、帰無仮説は pH4.5 以下の確率は 0、となろう。相対頻度は 0.315 だから、やはり帰無仮説は棄却される。（符号検定ではプラス符号の数＝対立仮説のもとでの数）

解説　　(247 頁、F=t^2 について)

(8.5)式は、

$$\text{級間変動} = \frac{1}{n} n_1 n_2 (\bar{x}_1 - \bar{x}_2)^2 \qquad (8.5)'$$

と表してもよい。一般の場合では、グループ数を m として

$$\text{級内変動} = \sum_{i=1}^{m} \sum_{j} (x_{ij} - \bar{x}_i)^2 \qquad (8.4)'$$

$$\text{級間変動} = \sum_{i=1}^{m} n_i (\bar{x}_i - \bar{x})^2 = \sum_{j=i+1}^{m} \sum_{i=1}^{m-1} \frac{n_i n_j}{n} (\bar{x}_i - \bar{x}_j)^2 \qquad (8.5)''$$

と書くことができる。この表現は「級間」という意味がよく出ていて理解しやすい。

247 頁の下から 5 行目以下については、F=t^2 になることを証明しよう。F 統計量は

$$F = \frac{\dfrac{n_1 n_2}{n} (\bar{x}_1 - \bar{x}_2)^2}{\dfrac{1}{n-2} \sum_{i=1}^{2} \sum_{j} (x_{1j} - \bar{x}_1)^2} = \frac{(\bar{x}_1 - \bar{x}_2)^2}{\dfrac{n_1 + n_2}{n_1 n_2} S^2} = \frac{(\bar{x}_1 - \bar{x}_2)^2}{\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right) S^2}$$

となり、189 頁から 190 頁にかけて説明された Z_0 の 2 乗になっている。これが統計量間の関係である。

(8.5)”式の展開は、2 グループについては

$$\bar{x} = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2$$

であるので

$$n_1 (\bar{x}_1 - \bar{x})^2 = n_1 \left(\frac{n_2}{n} \bar{x}_1 - \frac{n_2}{n} \bar{x}_2 \right)^2 = \frac{n_1 n_2^2}{n^2} (\bar{x}_1 - \bar{x}_2)^2 .$$

同様に

$$n_2(\bar{x}_2 - \bar{x})^2 = \frac{n_2 n_1^2}{n^2}(\bar{x}_1 - \bar{x}_2)^2$$

となり、

$$級間変動 = \frac{n_1 n_2}{n}\left(\frac{n_1}{n} + \frac{n_2}{n}\right)(\bar{x}_1 - \bar{x}_2)^2 = \frac{n_1 n_2}{n}(\bar{x}_1 - \bar{x}_2)^2$$

となる。一般型も同様に展開できる。