審査学位論文

企業調査における経営特性の測定と仮説検定の方法

森　治憲

1997年7月17日

学習院大学

審査学位論文

企業調査における経営特性の測定と仮説検定の方法

森　治憲

1997年7月17日

学習院大学

はじめに

企業の経営活動を特徴付ける概念の多くは定量的に測定することが難しい。例えば、製造企業の持つ競争力に関する次の仮設を考えてみよう。

「競争力のある企業は環境変化に対する適応力が高い。その適応力は製品開発から生産ラインに乗るまでの期間が短いほど高くなる。そして、その期間を短縮させるには、製品設計において開発部門と生産部門における職能間の協力が必要となる。」

この仮説の検証には、競争力の高さで分類した二つの企業間で職能間協力に差があるか否かを検定するという方法と、それとは逆に、職能間協力の程度に関して分類した企業間で競争力の差を検定するという二つの方法が考えられる。

いずれの場合も、客観的な指標のない「製品設計での職能間協力の程度」を数量的に評価しなければならない。そのため、一般的には「職能間協力の程度」を従業員に数量化させるという方法が用いられる。しかし、統計学的な観点から見ると、個人の主観に依存した測定値による仮説検定には考慮すべき問題が多く、通常の経済データのように扱うことはできない。

測定の信頼性に関係して生じる問題を考慮した仮説検定のフレームワークを示すことが本論文の目的である。冒頭の例に即して言うと、具体的には、測定値を「職能間協力の程度」の尺度として捉える測定モデルと、このときの仮説検定の方法が提案される。仮説の立て方に関しては、競争力で分類した企業間で「職能間協力の程度」を検定するという一番目の方法を前提とする。企業の分類が確定的となるから、その統計的取り扱いが容易になるためである。

仮説検定で個人の主観が問題になるかどうかは、統計的推測の対象となる概念の捉え方次第である。従業員の感じた「職能間協力の程度」を考察の対象とするのなら、従業員の主観を反映した測定値を用いて検定しても問題はない。しかし、この検定の目的は「職能間協力の程度」という現実の世界における状態を検証することであり、従業員の心理的状態を検証することではない。従って、ある従業員が「我々は実に良く協力している」と感じて測定（回答）しても、それが常識的に考えて普通の状態のことであれば、この過大評価を修正した後で仮説検定を行わなければならない。測定値から個人差（質問の差や所属する企業の差もあるだろう）を補正するという考え方は、一般化可能性理論における測定値の表現に外ならない。そして、この補正を考慮した検定方法が所謂分散分析である。

測定値を一般化可能性理論の枠組みの中で捉え、分散分析を適用すれば、すべての問題が解決されるというわけではない。測定値の比較に意味を持たせるには、同じ測定値に対応した状態はすべての従業員で共通していなければならないが、「職能間協力の程度」の状態と実数との対応関係は従業員に依存しているからである。従って、測定値を比較するのであれば、特定の対応関係による尺度に変換してから行わなければならない。過大評価や過小評価などの個人差は状態の特定化に際しての傾向だから、その補正は共通の対応関係で表した尺度へ変換した後で行う必要もある。変換した値に一般化可能性理論と分散分析を適用すべきということである。

しかし、従業員が「職能間協力の程度」を数値化するのに用いた、本人ですら分からないであろう対応関係を分析者が知ることなどできるはずがない。測定値を特定の対応関係による尺度へ変換することができないのであれば、個人差の補正に加え、対応関係の違いも考慮した測定モデルは、測定値と共通化した尺度の関係を示す「関数」ではなく、共通化した尺度の候補値の集合が対応する所謂「対応」として考える必要がある。一つの測定値に複数の候補値が対応するため、各候補値に対して、それが真の尺度である確率が仮定される。真の標本が存在する事前分布を除いて、その値が未知という形式の検定問題は汎検定問題と呼ばれる。従って、競争力で分類した企業間で職能間協力に差があるか否かという仮説検定は汎検定問題として扱わなければならないことが分かる。

この論文は三つの部分から構成されている。その中心となるのは、汎検定問題を解説した第3章と、経営特性の測定モデルを導出する第4章である。そして、この仮説検定のフレームワークは第5章で応用される。ここでは企業調査で必要になることが多い、質問項目の選択と業種間比較を具体例として取り上げた。最後は、これらの議論の理論的基礎をまとめた部分である。第1章では分散分析、第2章では一般化可能性理論を説明する。

平成9年6月21日
筆者

目次

1 分散分析の理論

第４章で経営特性の測定値を表現する数理（測定）モデルを導出する。この測定モデルを前提とした業種間比較などの比較研究では、分散分析に標本の持つ信頼性を考慮した検定方法の適用が提案される。その検定方法については第３章で解説するので、ここでは分散分析の理論を詳しく説明していく。

まず、関連する予備知識を１-１節でまとめた後で、分散分析の理論を１-２節[1]で、検定方法としての性質を１-３節[2]で説明する。

1－1 関連する基本的な結果

任意の行列 $\mathbf{X}$ の列ベクトルが張る部分空間を $\mathrm{V}(\mathbf{X})$ と表せば；

定理 1-1.1

$$\mathrm{V}(\mathbf{X}) = \mathrm{V}(\mathbf{XX}')$$

が成立する。証明は Rao(1973,p27)を参照。

定理 1-1.2

$\mathbf{X}$ と $\mathbf{H}$ を列数の等しい行列とする。$L = \left\{\mathbf{X}\boldsymbol{\gamma} \middle| \mathbf{H}\boldsymbol{\gamma} = \mathbf{0}\right\}$ とすれば、L は $L \subseteq V(\mathbf{X})$ を満たす部分空間であり、その次元は $\dim L = \mathrm{rank}(\mathbf{H}'\ \mathbf{X}') - \mathrm{rank}\,\mathbf{H}$ である。

（証明）L が $L \subseteq V(\mathbf{X})$ を満たす部分空間であることは容易に示される。部分空間 L は行列 $\mathbf{H}$ の核の行列 $\mathbf{X}$（線形写像 X）による像空間のことだから、$\mathrm{X}\colon\left\{\boldsymbol{\gamma} \middle| \mathbf{H}\boldsymbol{\gamma} = \mathbf{0}\right\} \to L$ と表すことができる。一般に、線形写像 $\mathrm{F}\colon V \to \mathrm{F}(V)$ の表現行列を $\mathbf{F}$ で表すと；

$$\begin{aligned}\dim V &= \dim\left\{\mathbf{y} \middle| \mathbf{y} = \mathbf{Fx}, \mathbf{x} \in V\right\} + \dim\left\{\mathbf{x} \middle| \mathbf{Fx} = \mathbf{0}, \mathbf{x} \in V\right\} \\ &= \dim\mathrm{F}(V) + \dim\left(V \cap \left\{\mathbf{x} \middle| \mathbf{Fx} = \mathbf{0}\right\}\right)\end{aligned}$$

[1] Rao(1973)の記述に従っている。

[2] Lehmann(1986)の記述に従っている。

が成立する[3]。この結果から部分空間 L の次元を求めることができる。

$$\dim\{\gamma|\mathbf{H}\gamma=\mathbf{0}\}=m-\operatorname{rank}\mathbf{H}$$

$$\dim\left(\{\gamma|\mathbf{H}\gamma=\mathbf{0}\}\cap\{\gamma|\mathbf{X}\gamma=\mathbf{0}\}\right)=\dim\left\{\gamma\left|\begin{pmatrix}\mathbf{H}\\ \mathbf{X}\end{pmatrix}\gamma=\mathbf{0}\right.\right\}=m-\operatorname{rank}(\mathbf{H}'\ \mathbf{X}')$$

$$\therefore\ \dim L=\dim\{\gamma|\mathbf{H}\gamma=\mathbf{0}\}-\dim\left(\{\gamma|\mathbf{H}\gamma=\mathbf{0}\}\cap\{\gamma|\mathbf{X}\gamma=\mathbf{0}\}\right)=\operatorname{rank}(\mathbf{H}'\ \mathbf{X}')-\operatorname{rank}\mathbf{H}$$

ただし、m は $\mathbf{X}$ と $\mathbf{H}$ の列数である。（終）

行列 $\mathbf{X}$ の任意の一般逆行列を $\mathbf{X}^-$ と表そう。

定理 1-1.3

(i) $(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^-\mathbf{X}'=\mathbf{X}'$

(ii) $\mathbf{X}(\mathbf{X}'\mathbf{X})^-(\mathbf{X}'\mathbf{X})=\mathbf{X}$

(iii) $\mathbf{G}$ を $\mathbf{X}'\mathbf{X}$ の一般逆行列の一つとする。このとき、$\mathbf{XGX}'$ は一般逆行列の選び方に依存しない。

（証明）定理 1-1.1 から $\mathrm{V}(\mathbf{X}')=\mathrm{V}(\mathbf{X}'\mathbf{X})$ であり、$(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^-$ は $\mathbf{R}^m$ から $\mathrm{V}(\mathbf{X}'\mathbf{X})$ への射影行列[4]だから、(i)式は明らか。

$\mathbf{XX}^-\mathbf{X}=\mathbf{X}\Leftrightarrow\mathbf{X}'(\mathbf{X}^-)'\mathbf{X}'=\mathbf{X}'$ だから、$(\mathbf{X}^-)'$ は $\mathbf{X}'$ の一般逆行列の一つである。ここから、$((\mathbf{X}'\mathbf{X})^-)'$ が $\mathbf{X}'\mathbf{X}$ の一般逆行列であることが分かる。この $((\mathbf{X}'\mathbf{X})^-)'$ を(i)式に代入して両辺を転置すれば、(ii)式を得ることができる。

$\mathrm{V}(\mathbf{X}')=\mathrm{V}(\mathbf{X}'\mathbf{X})$ より $\mathbf{X}'=\mathbf{X}'\mathbf{XC}$ なる行列 $\mathbf{C}$ を取れば、$\mathbf{XGX}'=\mathbf{C}'\mathbf{X}'\mathbf{XC}=\mathbf{C}'\mathbf{X}'$ と表すことができる。しかし、行列 $\mathbf{C}$ は一意に定められないから、$\mathbf{XGX}'$ が一意であることを示すには、$\mathbf{C}_1\neq\mathbf{C}_2$ に対して $\mathbf{XC}_1=\mathbf{XC}_2$ を示さなければならない。行列 $\mathbf{C}$ の定義から $\mathbf{X}'\mathbf{X}(\mathbf{C}_1-\mathbf{C}_2)=\mathbf{0}$ に注意すると；

[3] 小山(1994,p*425*)

[4] 小山(1994,p*559*)

$$(\mathbf{C}_1-\mathbf{C}_2)'\mathbf{X}'\mathbf{X}(\mathbf{C}_1-\mathbf{C}_2)=\|\mathbf{X}(\mathbf{C}_1-\mathbf{C}_2)\|^2=0 \Leftrightarrow \mathbf{XC}_1=\mathbf{XC}_2$$

であることが分かる。（終）

この結果は以下の議論で頻繁に使用される一般逆行列の性質である。次に、二次形式の分布に関する定理をまとめておく。ただし、非心χ^2分布に従う二次形式$\mathbf{y}'\mathbf{A}\mathbf{y}$の非心度は$\mathrm{E}(\mathbf{y}')\mathbf{A}\,\mathrm{E}(\mathbf{y})$と定義される。以下の定理の証明はRao(1973,3b.4)を参照。

定理 1-1.4

$\mathbf{y}\sim \mathrm{N}(\boldsymbol{\mu},\mathbf{I})$に関する二次形式$\mathbf{y}'\mathbf{A}\mathbf{y}$が非心$\chi^2$分布に従う必要十分条件は、行列$\mathbf{A}$がベキ等行列となることである。また、自由度は$\mathrm{rank}\,\mathbf{A}=\mathrm{tr}\,\mathbf{A}$である。

定理 1-1.5

$Q=Q_1+Q_2, Q\sim\chi^2(f,\lambda), Q_1\sim\chi^2(f_1,\lambda_1), Q_2\geq 0$ならば、$Q_2\sim\chi^2(f-f_1,\lambda-\lambda_1)$であり、$Q_2$と$Q_1$は独立に分布する。

定理 1-1.6

$\mathbf{y}\sim \mathrm{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$とする。このとき、二次形式$(\mathbf{y}-\boldsymbol{\mu})'\mathbf{A}(\mathbf{y}-\boldsymbol{\mu})$が非心$\chi^2$分布に従う必要十分条件は$\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}=\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}$が成立することである。このときの自由度は$\mathrm{tr}\,\mathbf{A}\boldsymbol{\Sigma}$である。

定理 1-1.7

$\mathbf{y}\sim \mathrm{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$とする。二つの二次形式$(\mathbf{y}-\boldsymbol{\mu})'\mathbf{A}(\mathbf{y}-\boldsymbol{\mu})$と$(\mathbf{y}-\boldsymbol{\mu})'\mathbf{B}(\mathbf{y}-\boldsymbol{\mu})$が独立に分布する必要十分条件は$\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\Sigma}=\mathbf{0}$が成立することである。

最後に二次形式の期待値を与えておく。これは佐和(1979,p42)による。

定理 1-1.8

$\mathrm{E}(\mathbf{y})=\boldsymbol{\mu}, \mathrm{V}(\mathbf{y})=\boldsymbol{\Sigma}$とする。このとき、$\mathrm{E}(\mathbf{y}'\mathbf{A}\mathbf{y})=\mathrm{tr}\,\mathbf{A}\boldsymbol{\Sigma}+\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$である。

1－2 分散分析の理論

1－2－1 分散分析の基本定理

本節では回帰モデル $\mathbf{y} \sim N(\mathbf{X\beta}, \sigma^2 \mathbf{I}), \mathbf{y} \in \mathbf{R}^n, \mathbf{X}: n \times m, \operatorname{rank} \mathbf{X} = r \le m$ について考察していく。ここで、行列 $\mathbf{X}$ に最大階数 m を仮定しないのは、分散分析を視野に入れた統一的な議論を可能にするためである。実際、分散分析における加法モデルでは $\operatorname{rank}\mathbf{X} = m$ が成立しない[5]。しかし、この仮定が満たされないと行列 $\mathbf{X'X}$ が特異となるため、最小二乗推定量 $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-}\mathbf{X'Y}$ を一意に定めることができない。そのため、一般化した理論では母数 $\boldsymbol{\beta}$ ではなく線形関数 $\mathbf{H\beta}$ が考察の対象となり、その最小二乗推定量 $\mathbf{H}\hat{\boldsymbol{\beta}}$ が一意に定まる必要十分条件；

$$V(\mathbf{H'}) \subseteq V(\mathbf{X'})$$

が与えられる[6]。この条件を満たす線形関数 $\mathbf{H\beta}$ は推定可能(Estimable)であるという。

定理 1-2.1

$$R_0^2 = \min_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta}) \sim \sigma^2 \chi^2 (n - r)$$

（証明）一般性を失うことなく $\sigma^2 = 1$ を仮定する。ノルム $(\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta}) = \|\mathbf{y} - \mathbf{X\beta}\|^2$ は $\mathbf{X\beta}$ が $\mathbf{y}$ の $V(\mathbf{X})$への直交射影となるとき最小値を取る。そこで、$\mathbf{R}^n$ から $V(\mathbf{X})$への直交射影行列を $\mathbf{P}$ とすれば；

$$R_0^2 = \min_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta}) = (\mathbf{y} - \mathbf{Py})'(\mathbf{y} - \mathbf{Py}) = \mathbf{y}'(\mathbf{I} - \mathbf{P})\mathbf{y}$$

と表すことができる。直交射影行列はベキ等行列だから、この二次形式が非心 χ^2 分布に従うことは定理 1-1.4 から分かる。まず、非心度は $E(\mathbf{y}) = \mathbf{X\beta} \Rightarrow (\mathbf{X\beta})'(\mathbf{I} - \mathbf{P})(\mathbf{X\beta}) = 0$ よりゼロである。その自由度は $\operatorname{rank} \mathbf{P} = \dim V(\mathbf{X})$ だから $n - r$ となる。（終）

---

[5] 5-5節で具体例が示される。

[6] Rao(1973,p*223*)

次に、行列$\mathbf{H}: s \times m, \mathrm{rank}\,\mathbf{H} = k, \dim\left(\mathrm{V}(\mathbf{X}') \cap \mathrm{V}(\mathbf{H}')\right) = t$を取り、線形制約$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi}$の下での残差平方和；

$$R_1^2 = \min_{\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

について考えていく。このとき、前定理に対応した次の定理が成立する。

定理 1-2.2

$R_1^2 - R_0^2$はR_0^2と独立に分布し、自由度tの非心χ^2分布に従う。仮説$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi}$が真であれば、$R_1^2 - R_0^2 \sim \sigma^2 \chi^2(t)$であり、$\dfrac{R_1^2 - R_0^2}{t} \div \dfrac{R_0^2}{n-r} \sim \mathrm{F}(t, n-r)$となる。

（証明）ここでも$\sigma^2 = 1$とする。連立方程式$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi}$の特殊解を$\boldsymbol{\beta}_0$、同時方程式$\mathbf{H}\boldsymbol{\beta} = \mathbf{0}$の任意の解を$\boldsymbol{\gamma}$とすれば、その一般解は$\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{\gamma}$と表すことができる。前定理の証明と同様に考えれば；

$$R_1^2 = \min_{\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \min_{\mathbf{H}\boldsymbol{\gamma}=\mathbf{0}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0 - \mathbf{X}\boldsymbol{\gamma})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0 - \mathbf{X}\boldsymbol{\gamma})$$

が達成されるのは、$\mathbf{X}\boldsymbol{\gamma}$が$\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0$の$L = \left\{\mathbf{X}\boldsymbol{\gamma} \middle| \mathbf{H}\boldsymbol{\gamma} = \mathbf{0}\right\} \subseteq \mathrm{V}(\mathbf{X})$への直交射影になるときである。そこで、$\mathbf{R}^n$から$L$への直交射影行列を$\mathbf{U}$とすれば、定理1-1.4から；

$$\begin{aligned} R_1^2 &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0 - \mathbf{U}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0))'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0 - \mathbf{U}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0)) \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{I} - \mathbf{U})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0) \end{aligned}$$

が非心度$\lambda = (\mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{I} - \mathbf{U})(\mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta}_0)$の非心$\chi^2$分布に従うことが分かる。その自由度は定理1-1.2より；

$$\mathrm{rank}\mathbf{U} = \dim L = \mathrm{rank}(\mathbf{H}'\ \mathbf{X}') - \mathrm{rank}\,\mathbf{H} = (r + k - t) - k = r - t$$

だから $n-r+t$ となる。$R_1^2 \geq R_0^2$ だから、定理1-1.5より、$R_1^2 - R_0^2 \sim \chi^2(t,\lambda)$ が R_0^2 と独立に分布することも分かる。仮説 $\mathbf{H\beta}=\xi$ が真であれば、$\mathbf{X\beta}-\mathbf{X\beta}_0=\mathbf{X\gamma}, \mathbf{H\gamma}=\mathbf{0}$ となるから、$\lambda=(\mathbf{X\gamma})'(\mathbf{I}-\mathbf{U})(\mathbf{X\gamma})=0 \Rightarrow R_1^2-R_0^2 \sim \chi^2(t)$ となる。（終）

これは次の二点において一般化した定理である。まず、行列 $\mathbf{H}$ に最大階数を仮定していないことである。確かに、行列 $\mathbf{H}$ に含まれる線形従属な行に関する検定は理論的に無意味であるが、理解し易いように行列表記した仮説 $\mathbf{H\beta}=\xi$ の係数行列から敢えて線形従属な行を取り除き変形する必要はないと思われる。この操作上の利点が、行列 $\mathbf{H}$ に最大階数を仮定しない理由である。

二番目は仮説 $\mathbf{H\beta}=\xi$ に推定可能であるための必要十分条件 $\mathrm{V}(\mathbf{H}') \subseteq \mathrm{V}(\mathbf{X}')$ を仮定していないことである。説明変数行列 $\mathbf{X}$ の行とは線形独立な行列 $\mathbf{H}$ の行が意味する仮説が検定の対象とはならないことは、$R_1^2-R_0^2$ の自由度が $\dim\left(\mathrm{V}(\mathbf{X}') \cap \mathrm{V}(\mathbf{H}')\right)=t$ という定理の主張から明らかであろう。*1-2-4*で具体的に説明するが、モデルに仮説ではない線形制約を含む場合が想定されるため、この二番目の一般化は非常に重要である。このために定理が一般化されてるといっても過言ではない。

1-2-2 分散分析

モデル $\mathbf{y} \sim \mathrm{N}(\mathbf{X\beta}, \sigma^2\mathbf{I})$ において、帰無仮説 $\mathbf{H\beta}=\xi$ の下では；

$$F=\frac{R_1^2-R_0^2}{t} \div \frac{R_0^2}{n-r} \sim \mathrm{F}(t, n-r)$$

が成立する。このとき、定理1-1.8から；

$$\begin{aligned}
\mathrm{E}(R_0^2) &= \mathrm{E}(\mathbf{y}'(\mathbf{I}-\mathbf{P})\mathbf{y}) = \sigma^2 \operatorname{tr}(\mathbf{I}-\mathbf{P}) + (\mathbf{X\beta})'(\mathbf{I}-\mathbf{P})(\mathbf{X\beta}) = (n-r)\sigma^2 \\
\mathrm{E}(R_1^2) &= \mathrm{E}\left((\mathbf{y}-\mathbf{X\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{y}-\mathbf{X\beta}_0)\right) \\
&= \sigma^2 \operatorname{tr}(\mathbf{I}-\mathbf{U}) + (\mathbf{X\beta}-\mathbf{X\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{X\beta}-\mathbf{X\beta}_0) \\
&= (n-r+t)\sigma^2 + (\mathbf{X\beta}-\mathbf{X\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{X\beta}-\mathbf{X\beta}_0) \\
\therefore\ \mathrm{E}(R_1^2-R_0^2) &= t\sigma^2 + (\mathbf{X\beta}-\mathbf{X\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{X\beta}-\mathbf{X\beta}_0)
\end{aligned}$$

となるが、$\mathbf{I}-\mathbf{U}$ は非負値定符号行列だから、$(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)\geq 0$ であることが分かる。この結果を F 統計量に代入すると；

$$\frac{\mathrm{E}(R_1^2-R_0^2)}{t}\div\frac{\mathrm{E}(R_0^2)}{n-r}=1+\frac{(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)}{t\sigma^2}\geq 1$$

となる。行列 $\mathbf{U}$ は $L=\left\{\mathbf{X}\boldsymbol{\gamma}\middle|\mathbf{H}\boldsymbol{\gamma}=\mathbf{0}\right\}$ への直交射影行列だから、二次形式；

$$\begin{aligned}(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{I}-\mathbf{U})(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)&=(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{I}-\mathbf{U})^2(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)\\&=\left\|(\mathbf{I}-\mathbf{U})(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)\right\|^2\end{aligned}$$

がゼロとなるのは、$(\mathbf{I}-\mathbf{U})(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}_0)=\mathbf{0}$ 、すなわち、$\mathbf{X}\boldsymbol{\beta}=\mathbf{X}\boldsymbol{\beta}_0$ か $\mathbf{X}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\in L$ のいずれかの場合のみである。これは仮説が真であることと同値である。

このように、平均的に考えると F 統計量は常に 1 以上だから、その大きな値は仮説からのズレを反映していることが分かる。この性質に基づいて F 統計量を検定統計量とした仮説 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$ の検定を分散分析(Analysis of Variance) という。また、検定方法としての性質については次節で説明される。

1-2-3 残差平方和 R_0^2 と R_1^2 の具体的な表現

最初に、R_0^2 を具体的に求めよう。よく知られているように、$\mathbf{R}^n$ から $\mathrm{V}(\mathbf{X})$ への直交射影行列は $\mathbf{P}=\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ と表すことができる。また、この行列が一般逆行列 $(\mathbf{X}'\mathbf{X})^-$ の選び方に依存しないことは、定理 1 - 1 . 3 (iii)により保証されている。よって；

$$R_0^2=\mathbf{y}'(\mathbf{I}-\mathbf{P})\mathbf{y}=\mathbf{y}'(\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')\mathbf{y}$$

という表現が得られる。

これに対し、R_1^2 を具体的に表現することは、係数行列 $\mathbf{H}$ と説明変数行列 $\mathbf{X}$ の関係に依存するため非常に面倒である。そこで、仮説 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$ が検定可能（$\mathbf{H}\boldsymbol{\beta}$ が推定可能）となる条件 $\mathrm{V}(\mathbf{H}')\subseteq\mathrm{V}(\mathbf{X}')$ を満たす場合から説明していこう。まず、$\mathrm{V}(\mathbf{X}')=\mathrm{V}(\mathbf{X}'\mathbf{X})$ に注意して、$\mathbf{H}'=\mathbf{X}'\mathbf{X}\mathbf{C}$ となる $m\times s$ 型行列 $\mathbf{C}$ を取ると、行列 $\mathbf{XC}$ の階数は；

$$\mathrm{rank}\,\mathbf{H} = \mathrm{rank}\,\mathbf{X}'\mathbf{X}\mathbf{C} \le \mathrm{rank}\,\mathbf{X}\mathbf{C} = \mathrm{rank}\,\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C} \le \mathrm{rank}\,\mathbf{X}'\mathbf{X}\mathbf{C} = \mathrm{rank}\,\mathbf{H}$$
$$\therefore \quad \mathrm{rank}\,\mathbf{X}\mathbf{C} = \mathrm{rank}\,\mathbf{H} = k$$

となる。次に、k 次元部分空間 $\mathrm{V}(\mathbf{XC}) \subseteq \mathrm{V}(\mathbf{X})$ には含まれない、$\mathrm{V}(\mathbf{X})$の直交基底を列とする $n \times (r-k)$型行列 $\mathbf{D}$ を取る。このとき、$\mathrm{V}(\mathbf{I}-\mathbf{P}) = \mathrm{V}(\mathbf{X}')^{\perp}$ という直交射影行列 $\mathbf{P}$ に関する性質から、三つの行列 $\mathbf{D},\mathbf{I}-\mathbf{P},\mathbf{XC}$ の列ベクトルは、それが異なる行列の列であれば互い直交することが分かる。

そこで、$(r-k+n+s) \times n$ 型行列 $\mathbf{T}$；

$$\mathbf{T} = \begin{pmatrix} \mathbf{D}' \\ \mathbf{I}-\mathbf{P} \\ \mathbf{C}'\mathbf{X}' \end{pmatrix} \begin{matrix} r-k \\ n \\ s \end{matrix} \quad ,\mathrm{rank}\,\mathbf{T} = n$$

による線形写像 $\mathbf{z} = (\mathbf{z}_1' \; \mathbf{z}_2' \; \mathbf{z}_3') = \mathbf{Ty}$ を考える。ただし、ベクトル $\mathbf{z}$ の分割は行列 $\mathbf{T}$ の分割に対応している。

R_1^2 の表現を求める前に、行列 $\mathbf{T}$ に関する二つの性質を挙げておく。一つ目は、最大階数 n を持つ行列 $\mathbf{T}$ による線形写像の核が$\{\mathbf{0}\}$であることから、任意の n 次元ベクトル $\mathbf{x}$ について $\mathbf{T}^-\mathbf{T}\mathbf{x} = \mathbf{x}$ が成立することである。二つ目は、$(\mathbf{T}')^-\mathbf{T}^-$ が $\mathbf{TT}'$ の一般逆行列になっていることである。最初の性質を使うと；

$$(\mathbf{TT}')((\mathbf{T}')^-\mathbf{T}^-)(\mathbf{TT}') = (\mathbf{TT}')(\mathbf{T}')^-\mathbf{T}' = \mathbf{TT}'$$

が成立する。これは$(\mathbf{T}')^-\mathbf{T}^-$ が $\mathbf{TT}'$ の一般逆行列であることに外ならない。

元の議論に戻る。一つ目の性質から；

$$\mathbf{y} - \mathbf{X}\boldsymbol{\beta} = \mathbf{T}^-\mathbf{T}\mathbf{z} - \mathbf{T}^-\mathbf{T}\mathbf{X}\boldsymbol{\beta}$$

が一般逆行列 $\mathbf{T}^-$ の選び方とは無関係に成立する。従って、残差平方和は；

$$
\begin{aligned}
&(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) \\
&=(\mathbf{T}^{-}\mathbf{z}-\mathbf{T}^{-}\mathbf{T}\mathbf{X}\boldsymbol{\beta})'(\mathbf{T}^{-}\mathbf{z}-\mathbf{T}^{-}\mathbf{T}\mathbf{X}\boldsymbol{\beta}) \\
&=(\mathbf{z}-\mathbf{T}\mathbf{X}\boldsymbol{\beta})'(\mathbf{T}^{-})'\mathbf{T}^{-}(\mathbf{z}-\mathbf{T}\mathbf{X}\boldsymbol{\beta}) \\
&=(\mathbf{z}-\mathbf{T}\mathbf{X}\boldsymbol{\beta})'(\mathbf{T}')^{-}\mathbf{T}^{-}(\mathbf{z}-\mathbf{T}\mathbf{X}\boldsymbol{\beta}) \\
&=(\mathbf{z}-\mathbf{T}\mathbf{X}\boldsymbol{\beta})'(\mathbf{T}\mathbf{T}')^{-}(\mathbf{z}-\mathbf{T}\mathbf{X}\boldsymbol{\beta}) \\
&=\left(\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\\ \mathbf{z}_3\end{pmatrix}-\begin{pmatrix}\mathbf{D}'\mathbf{X}\boldsymbol{\beta}\\ \mathbf{0}\\ \mathbf{C}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}\end{pmatrix}\right)'\begin{pmatrix}\mathbf{I}_{r-k} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & (\mathbf{I}-\mathbf{P})^{-} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & (\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C})^{-}\end{pmatrix}\left(\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\\ \mathbf{z}_3\end{pmatrix}-\begin{pmatrix}\mathbf{D}'\mathbf{X}\boldsymbol{\beta}\\ \mathbf{0}\\ \mathbf{C}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}\end{pmatrix}\right) \\
&=(\mathbf{z}_1-\mathbf{D}'\mathbf{X}\boldsymbol{\beta})'(\mathbf{z}_1-\mathbf{D}'\mathbf{X}\boldsymbol{\beta})+\mathbf{z}_2'(\mathbf{I}-\mathbf{P})^{-}\mathbf{z}_2+(\mathbf{z}_3-\mathbf{H}\boldsymbol{\beta})'(\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C})^{-}(\mathbf{z}_3-\mathbf{H}\boldsymbol{\beta})
\end{aligned}
$$

と書き換えることができる。ただし、三番目の等号は$(\mathbf{T}^{-})'$が$\mathbf{T}'$の一般逆行列の一つであるという性質から[7]、四番目の等号は前述した二つ目の性質から導かれる。

ここで、二番目の二次形式について以下の関係は明らか。

$$
\mathbf{z}_2'(\mathbf{I}-\mathbf{P})^{-}\mathbf{z}_2=\mathbf{y}'(\mathbf{I}-\mathbf{P})(\mathbf{I}-\mathbf{P})^{-}(\mathbf{I}-\mathbf{P})\mathbf{y}=\mathbf{y}'(\mathbf{I}-\mathbf{P})\mathbf{y}=R_0^2
$$

一つ目の二次形式の、線形制約$\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$の下での最小値は次のようになる。まず、各行列の定義から連立方程式；

$$
\begin{pmatrix}\mathbf{D}'\mathbf{X}\\ \mathbf{C}'\mathbf{X}'\mathbf{X}\end{pmatrix}\boldsymbol{\beta}=\begin{pmatrix}\mathbf{D}'\mathbf{X}\\ \mathbf{H}\end{pmatrix}\boldsymbol{\beta}=\begin{pmatrix}\mathbf{z}_1\\ \boldsymbol{\xi}\end{pmatrix}
$$

の中で線形独立な方程式はr個[8]であり；

$$
r=\operatorname{rank}\mathbf{X}=\operatorname{rank}\mathbf{T}\mathbf{X}=\operatorname{rank}\begin{pmatrix}\mathbf{D}'\\ \mathbf{G}\\ \mathbf{C}'\mathbf{X}'\end{pmatrix}\mathbf{X}=\operatorname{rank}\begin{pmatrix}\mathbf{D}'\mathbf{X}\\ \mathbf{0}\\ \mathbf{H}'\end{pmatrix}=\operatorname{rank}\begin{pmatrix}\mathbf{D}'\mathbf{X}\\ \mathbf{H}'\end{pmatrix}
$$

だから、この連立方程式は解を持つ。ここから、$\min_{\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}}(\mathbf{z}_1-\mathbf{D}'\mathbf{X}\boldsymbol{\beta})'(\mathbf{z}_1-\mathbf{D}'\mathbf{X}\boldsymbol{\beta})=0$である

[7] 定理1-1-3

[8] 方程式数は$(r-k+s)$個であるが、線形仮説$\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$で線形独立な式は$k$個である。

ことが分かる。また、$V(\mathbf{H}')\subseteq V(\mathbf{X}')$ より線形関数 $\mathbf{H\beta}$ は推定可能だから、その最小二乗推定量は；

$$\mathbf{H}\hat{\boldsymbol{\beta}}=\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}=\mathbf{C}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}=\mathbf{C}'\mathbf{X}'\mathbf{y}=\mathbf{z}_3$$

となる。この結果を三番目の二次形式に代入することで、$V(\mathbf{H}')\subseteq V(\mathbf{X}')$ が成立するときの残差平方和 R_1^2 は；

$$\begin{aligned}R_1^2&=\min_{\mathbf{H\beta}=\boldsymbol{\xi}}(\mathbf{y}-\mathbf{X\beta})'(\mathbf{y}-\mathbf{X\beta})\\&=R_0^2+(\mathbf{H}\hat{\boldsymbol{\beta}}-\boldsymbol{\xi})'(\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C})^{-}(\mathbf{H}\hat{\boldsymbol{\beta}}-\boldsymbol{\xi})\\&=R_0^2+(\mathbf{H}\hat{\boldsymbol{\beta}}-\boldsymbol{\xi})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H})^{-}(\mathbf{H}\hat{\boldsymbol{\beta}}-\boldsymbol{\xi})\end{aligned}$$

と表現できることが分かる。また、$(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H})^{-}=(\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C})^{-}$ を一意に定めることはできないが、二次形式 $(\mathbf{H}\hat{\boldsymbol{\beta}}-\boldsymbol{\xi})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H})^{-}(\mathbf{H}\hat{\boldsymbol{\beta}}-\boldsymbol{\xi})$ の一意性は、$\mathbf{H}'=\mathbf{X}'\mathbf{X}\mathbf{C}$ という関係と定理 1 - 1 - 3 (iii)により保証される。

次に、$V(\mathbf{H}')\cap V(\mathbf{X}')=\{\mathbf{0}\}$ の場合を考える。この場合の線形制約 $\mathbf{H\beta}=\boldsymbol{\xi}$ が検定すべき仮説ではなく母数 $\boldsymbol{\beta}$ に課せられた制約であることは、$R_1^2-R_0^2$ の自由度がゼロになることから明らかであろう（実際、$\mathbf{H\beta}$ は推定可能ではない）。この条件の下では行列 $\mathbf{T}$ を構成できないので、ここではモデルから線形制約 $\mathbf{H\beta}=\boldsymbol{\xi}$ を消去する方法を採用する。

そこで、同時方程式 $\mathbf{H\beta}=\mathbf{0}$ の任意の解 $(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})\boldsymbol{\theta}$ を用いて、連立方程式 $\mathbf{H\beta}=\boldsymbol{\xi}$ の一般解を $\boldsymbol{\beta}=\boldsymbol{\beta}_0+(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})\boldsymbol{\theta}$ と表し、$\mathbf{y}^*=\mathbf{y}-\mathbf{X}\boldsymbol{\beta}_0,\mathbf{X}^*=\mathbf{X}(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})$ とおけば、このモデルは $\mathbf{y}^*\sim N(\mathbf{X}^*\boldsymbol{\theta},\sigma^2\mathbf{I})$ と制約なしのモデルに変換することができる。ここから；

$$\begin{aligned}R_1^2&=(\mathbf{y}^*)'(\mathbf{I}-\mathbf{X}^*((\mathbf{X}^*)'\mathbf{X}^*)^{-}(\mathbf{X}^*)')\mathbf{y}^*\\&=(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}_0)'\times\\&\quad\left(\mathbf{I}-\{\mathbf{X}(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})\}\left(\{\mathbf{X}(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})\}'\{\mathbf{X}(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})\}\right)^{-}\{\mathbf{X}(\mathbf{I}-\mathbf{H}^{-}\mathbf{H})\}'\right)(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}_0)\end{aligned}$$

と表現できることが分かる。また、これが係数行列 $\mathbf{H}$ の行と説明変数行列 $\mathbf{X}$ の行との関係に依存しない残差平方和の統一的な表現である。

1-2-4 線形制約 $\mathbf{H}\beta=\xi$ の下での線形仮説 $\mathbf{G}\beta=\zeta$ の検定

ここで扱う回帰モデルは；

$$
\begin{aligned}
&\mathbf{y} \sim \mathrm{N}(\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I}), s.t.\, \mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}\\
&\operatorname{rank}\mathbf{X}=r, \operatorname{rank}\mathbf{H}=k, \dim\left(\mathrm{V}(\mathbf{X}')\cap\mathrm{V}(\mathbf{H}')\right)=t
\end{aligned}
$$

と線形制約が含まれたモデルであり、線形仮説は；

$$
\mathbf{G}\boldsymbol{\beta}=\boldsymbol{\zeta}, \dim\left(\left(\mathrm{V}(\mathbf{X}')-\mathrm{V}(\mathbf{H}')\right)\cap\mathrm{V}(\mathbf{G}')\right)=u
$$

である。ただし、$\left(\mathrm{V}(\mathbf{X}')-\mathrm{V}(\mathbf{H}')\right)$ は $\mathrm{V}(\mathbf{H}')$ に含まれない $\mathrm{V}(\mathbf{X}')$ の基底により張られる部分空間を意味している。この場合はモデルから線形制約 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$ を消去して、検定問題を制約なしの分散分析に帰着させて考えればよい。

ところで、係数行列 $\mathbf{H}$ の行の一部が $\mathrm{V}(\mathbf{X}')$ に含まれる場合は、行列 $\mathbf{X}$ の行と線形独立な部分を再び線形制約 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$ と表し、残りを線形仮説 $\mathbf{G}\boldsymbol{\beta}=\boldsymbol{\zeta}$ と表すことで、この問題が上記の検定問題で；

$$
\mathrm{V}(\mathbf{X}')\cap\mathrm{V}(\mathbf{H}')=\{\mathbf{0}\}, \mathrm{V}(\mathbf{G}')\subseteq\mathrm{V}(\mathbf{X}')
$$

の成立する特殊な場合に過ぎないことが分かる。

まず、*1-2-3* の最後の議論から；

$$
\begin{aligned}
R_0^2 &= \min_{\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\\
&= \mathbf{y}^{*\prime}\left(\mathbf{I}-\mathbf{X}^*((\mathbf{X}^*)'\mathbf{X}^*)^-(\mathbf{X}^*)'\right)\mathbf{y}^* \sim \sigma^2\chi^2(n-r+t)\\
\mathbf{y}^* &= \mathbf{y}-\mathbf{X}\boldsymbol{\beta}_0, \mathbf{X}^*=\mathbf{X}(\mathbf{I}-\mathbf{H}^-\mathbf{H})
\end{aligned}
$$

が成立する。しかし、この場合も $\mathrm{V}(\mathbf{G}')\subseteq\mathrm{V}(\mathbf{X}')\cup\mathrm{V}(\mathbf{H}')$ が成立しなければ、$R_1^2-R_0^2$ を

陽表的に表現することはできない。そこで、$\hat{\boldsymbol{\beta}}^* = \mathbf{G}^*((\mathbf{X}^*)'\mathbf{X}^*)^-(\mathbf{X}^*)'\mathbf{y}^*$ とおけば；

$$
\begin{aligned}
R_1^2 &= \min_{\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi},\mathbf{G}\boldsymbol{\beta}=\boldsymbol{\zeta}} (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) \\
&= \min_{\mathbf{G}^*\boldsymbol{\theta}=\boldsymbol{\zeta}^*} (\mathbf{y}^*-\mathbf{X}^*\boldsymbol{\beta})'(\mathbf{y}^*-\mathbf{X}^*\boldsymbol{\beta}), \mathbf{G}^* = \mathbf{G}(\mathbf{I}-\mathbf{H}^-\mathbf{H}), \boldsymbol{\zeta}^* = \boldsymbol{\zeta}-\mathbf{G}\boldsymbol{\beta}_0 \\
&= R_0^2 + (\mathbf{G}^*\hat{\boldsymbol{\beta}}^* - \boldsymbol{\zeta}^*)\left(\mathbf{G}^*((\mathbf{X}^*)'\mathbf{X}^*)^-(\mathbf{G}^*)'\right)^-(\mathbf{G}^*\hat{\boldsymbol{\beta}}^* - \boldsymbol{\zeta}^*)
\end{aligned}
$$

と表すことができる。条件が成立しない場合は、R_0^2 を求めたように二つの制約をモデルから消去した統一的な表現を使えばよい。R_1^2 の自由度は、この条件とは無関係に；

$$
\begin{aligned}
\mathrm{rank}\left((\mathbf{G}^*)'\ (\mathbf{H}^*)'\right) &= \mathrm{rank}\left(\begin{pmatrix}\mathbf{G}\\ \mathbf{X}\end{pmatrix}(\mathbf{I}-\mathbf{H}^-\mathbf{H})\right) \\
&= \dim\left(\begin{pmatrix}\mathbf{G}\\ \mathbf{X}\end{pmatrix}\boldsymbol{\gamma}\,\middle|\,\mathbf{H}\boldsymbol{\gamma}=\mathbf{0}\right) \\
&= \mathrm{rank}(\mathbf{G}'\ \mathbf{X}'\ \mathbf{H}') - \mathrm{rank}\,\mathbf{H} \\
&= r+k-t+\dim\left(\mathrm{V}(\mathbf{G}') - \left(\mathrm{V}(\mathbf{X}')\cup\mathrm{V}(\mathbf{H}')\right)\right) - k \\
\mathrm{rank}\,\mathbf{G}^* &= \mathrm{rank}\left(\mathbf{G}(\mathbf{I}-\mathbf{H}^-\mathbf{H})\right) \\
&= \mathrm{rank}(\mathbf{G}'\ \mathbf{H}') - \mathrm{rank}\,\mathbf{H} \\
&= \mathrm{rank}\,\mathbf{G} + k - \dim\left(\mathrm{V}(\mathbf{G}')\cap\mathrm{V}(\mathbf{H}')\right) - k \\
R_1^2\text{の自由度} &= n - \left(\mathrm{rank}\left((\mathbf{G}^*)'\ (\mathbf{H}^*)'\right) - \mathrm{rank}\,\mathbf{G}^*\right) \\
&= n-r+t+\mathrm{rank}\,\mathbf{G} - \dim\left(\mathrm{V}(\mathbf{G}') - \left(\mathrm{V}(\mathbf{X}')\cup\mathrm{V}(\mathbf{H}')\right)\right) \\
&\quad - \dim\left(\mathrm{V}(\mathbf{G}')\cap\mathrm{V}(\mathbf{H}')\right) \\
&= n-r+t+\dim\left(\left(\mathrm{V}(\mathbf{X}') - \mathrm{V}(\mathbf{H}')\right)\cap\mathrm{V}(\mathbf{G}')\right) \\
&= n-r+t+u
\end{aligned}
$$

であることが分かる。ここから、$R_1^2 - R_0^2 \sim \sigma^2\chi^2(u)$ が導かれる。

前述した、$\mathrm{V}(\mathbf{X}')\cap\mathrm{V}(\mathbf{H}') = \{\mathbf{0}\}, \mathrm{V}(\mathbf{G}') \subseteq \mathrm{V}(\mathbf{X}')$ である場合は；

$$
R_0^2 \sim \sigma^2\chi^2(n-r+k), R_1^2 - R_0^2 \sim \sigma^2\chi^2(\mathrm{rank}(\mathbf{G}))
$$

となる。もし、$V(\mathbf{X}')\cap V(\mathbf{H}')=\{\mathbf{0}\}, V(\mathbf{G}')\subseteq V(\mathbf{H}')$ であれば、$u=0$ となるから線形仮説 $\mathbf{G}\boldsymbol{\beta}=\boldsymbol{\zeta}$ の検定はできない。これは当然であろう。

1-3 検定論から見た分散分析の理論

1-2節では分散分析の理論を最小二乗法の視点から論じてきた。この議論は非常に明快であるが、線形仮説 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$ の検定に F 検定を用いる根拠や、仮説検定としての分散分析の良さについては明らかにされなかった。そこで、本節では分散分析の良さを検定論の立場から理論的に保証しておく。

本節で扱う検定問題も回帰モデル $\mathbf{y}\sim N(\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I}), \mathbf{y}\in\mathbf{R}^n, \mathrm{rank}\,\mathbf{X}=r$ における帰無仮説 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}, \mathrm{rank}\,\mathbf{H}=k$ と対立仮説 $\mathbf{H}\boldsymbol{\beta}\neq\boldsymbol{\xi}$ の検定である。ただし、ここでは条件；

$$V(\mathbf{H}')\subseteq V(\mathbf{X}')$$

の成立を仮定する。本節の議論の目的を考えれば、検定可能な仮説に対象を制限するのは当然であろう。さらに、仮説は $\boldsymbol{\xi}=\mathbf{0}$ の場合だけ考えれば十分である。なぜなら、連立方程式 $\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}$ の特殊解 $\boldsymbol{\beta}_0$ を用いて、$\boldsymbol{\beta}^*=\boldsymbol{\beta}-\boldsymbol{\beta}_0, \mathbf{Y}^*=\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}_0$ とすれば、モデルと帰無仮説は $\mathbf{y}^*\sim N(\mathbf{X}\boldsymbol{\beta}^*,\sigma^2\mathbf{I}), \mathbf{H}\boldsymbol{\beta}^*=\mathbf{0}$ に変換されるからである。

さて、所与の r 次元部分空間 Π_Ω の元である母数が $(r-k)$ 次元部分空間 $\Pi_\omega\subseteq\Pi_\Omega$ に含まれるか否かを検定するとき、この仮説を線形仮説(Linear Hypothesis)という。分散分析では母数 $\mathbf{X}\boldsymbol{\beta}$ に対して $\Pi_\Omega=V(\mathbf{X}), \dim\Pi_\Omega=r$ である。一方、$\mathbf{H}'=\mathbf{X}'\mathbf{X}\mathbf{C}$ を満たす行列 $\mathbf{C}$ を取れば、仮説 $\mathbf{H}\boldsymbol{\beta}=\mathbf{0}$ は $\mathbf{H}\boldsymbol{\beta}=\mathbf{C}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}=\mathbf{0}$ と表すことができる。従って；

$$\Pi_\omega=\left\{\mathbf{X}\boldsymbol{\beta}\,\middle|\,\mathbf{C}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}=\mathbf{0}\right\}$$

となる。ここで、$\mathrm{rank}\,\mathbf{H}=\mathrm{rank}\,\mathbf{X}\mathbf{C}=k$ が成立するから[9]、定理1-1.2より；

$$\dim\Pi_\omega=\dim\Pi_\Omega-\dim V(\mathbf{C}'\mathbf{X}')=r-k$$

[9] *1-2-3*参照。

であることが分かる。

次に、最初の r 行が $\Pi_\Omega = \mathrm{V}(\mathbf{X})$ の直交基底で、残りの $n-r$ 行が $(\Pi_\Omega)^\perp$ の直交基底となる n 次直交行列 $\mathbf{Q}$ を取る。ただし、最初の r 行のうち、第$(k+1)$行目から第 r 行目までは $\Pi_\omega \subseteq \Pi_\Omega$ の直交基底であるものとする[10]。この直交行列の定義から；

$$\mathbf{QX\beta} = (\eta_1 \cdots \eta_r\ 0 \cdots\ 0)'$$

となる。そして、帰無仮説 $\mathbf{H\beta = 0} \Leftrightarrow \mathbf{X\beta} \in \Pi_\omega$ が真のときに限り、$\boldsymbol{\eta}_k = (\eta_1 \cdots \eta_k)' = \mathbf{0}$ となることも分かる。従って、$\mathbf{Qy}$ を新しく $\mathbf{y}$ と書き直せば、モデル $\mathbf{y} \sim \mathrm{N}(\mathbf{X\beta}, \sigma^2\mathbf{I})$ における帰無仮説 $\mathbf{H\beta = 0}$ の両側検定は；

$$\mathbf{y} \sim \mathrm{N}(\boldsymbol{\eta}, \sigma^2\mathbf{I}), \mathbf{y} \in \mathbf{R}^n, \text{帰無仮説：}\ \boldsymbol{\eta}_k = \mathbf{0}, \text{対立仮説：}\ \boldsymbol{\eta}_k \neq \mathbf{0}$$

という両側検定に変換することができる。

しかし、複数母数の両側検定には一様最強力不偏検定(Uniformly Most Powerful Unbiased Test)が存在しないため[11]、こうした複数母数の同時検定は一次元に縮約した母数の関数についての検定問題として扱われる。

定理 1-3.1

上記の線形仮説の検定には以下の変換群；

$$\begin{aligned}
&G_1 : y_i' = y_i + c_i; i = k+1, \ldots, r, y_i' = y_i; otherwise \\
&G_2 : y_1, \ldots, y_k \text{の直交変換} \\
&G_3 : y_i' = c y_i; i = 1, \ldots, n
\end{aligned}$$

に不変な統計量 $W = \sum_{i=1}^{k} y_i^2 \Big/ \sum_{i=r+1}^{n} y_i^2$ による一様最強力不変検定(Uniformly Most Powerful Invariant Test)が存在する。そして、この検定は自由度 $k,n-r$ の F 検定となる。

---

[10] *1-2-3*で定義した行列 $\mathbf{T}$ と同じ構成となっている。

[11] $2 \le k$ ならば一様最強力不偏検定が存在する。

(証明) まず、上記の変換群全体に関する最大不変量(Maximal Invariant)が統計量 W であることは、各変換群に関する最大不変量を順に求めることで容易に示される。これらの変換群から母数空間に誘導される変換群は；

$$\overline{G}_1 : \eta_i' = \eta_i + c_i; i = k+1,\ldots,r, \eta_i' = \eta_i; i = 1,\ldots,k, \sigma' = \sigma$$
$$\overline{G}_2 : \eta_1,\ldots,\eta_k \text{の直交変換}, \sigma' = \sigma$$
$$\overline{G}_3 : \eta_i' = c\eta_i; i = 1,\ldots,n, \sigma' = |c|\sigma$$

であり、その最大不変量は $\delta^2 = \sum_{i=1}^{k} \eta_i^2 \Big/ \sigma^2$ である。最大不変量の一対一変換もまた最大不変量だから、ここでは統計量 $W^* = \dfrac{n-r}{k} W$ を基に議論を進めていく。統計量 W^* が非心度 δ^2、自由度 $k,n-r$ の非心 F 分布に従うことから分かるように、その分布は未知母数である非心度 δ^2 のみに依存する。従って、仮説 $\boldsymbol{\eta}_k = \mathbf{0}$ の両側検定は；

$$W^* \sim \mathrm{F}(k, n-r, \delta^2), \text{帰無仮説}: \delta = 0, \text{対立仮説}: \delta > 0$$

という母数 δ に関する検定問題に帰着する。任意の $\delta = \delta_1 > 0$ 対する尤度比(Likelihood Ratio)が W^* の単調増加関数であることは、F 分布と非心 F 分布の密度関数から容易に示すことができる。これは、統計量 W^* による母数 δ の検定が一様最強力検定であることに外ならない。以上から、モデル $\mathbf{y} \sim \mathrm{N}(\boldsymbol{\eta}, \sigma^2\mathbf{I})$ における線形仮説 $\boldsymbol{\eta}_k = \mathbf{0}$ の両側検定が、自由度 $k,n-r$ の F 検定と定義される一様最強力不変検定を持つことが示された。(終)

分散分析は次の定理によって検定論の立場から保証される。

定理 1-3.2

帰無仮説 $\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi}$ が真ならば、$\dfrac{R_1^2 - R_0^2}{k} \div \dfrac{R_0^2}{n-r} = W^* \sim \mathrm{F}(k, n-r)$ である。

(証明) 一般性を失うことなく $\boldsymbol{\xi} = \mathbf{0}$ と仮定して、検定問題を $\mathbf{y} \sim \mathrm{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}), \mathbf{H}\boldsymbol{\beta} = \mathbf{0}$ から $\mathbf{Q}\mathbf{y} \sim \mathrm{N}(\boldsymbol{\eta}, \sigma^2\mathbf{I}), \boldsymbol{\eta}_k = \mathbf{0}$ に変換したときの直交行列 $\mathbf{Q}$ を取る。ここで、この直交行列は

最初の r 行が $\Pi_\Omega = V(\mathbf{X})$ の、残りの $n-r$ 行が $V(\mathbf{X})^\perp$ の直交基底であることを注意しておく。まず、$R_0^2 = \mathbf{y}'(\mathbf{I}-\mathbf{P})\mathbf{y}$ の行列 $\mathbf{I}-\mathbf{P}$ は $V(\mathbf{X})^\perp$ への直交射影行列だから、この二次形式は直交変換 $\mathbf{y} = \mathbf{Q}'\tilde{\mathbf{y}}$ により；

$$R_0^2 = \tilde{\mathbf{y}}'\mathbf{Q}(\mathbf{I}-\mathbf{P})\mathbf{Q}'\tilde{\mathbf{y}} = \tilde{y}_{r+1}^2 + \cdots + \tilde{y}_n^2$$

と表すことができる。これに対し、$R_1^2 = \mathbf{y}'(\mathbf{I}-\mathbf{U})\mathbf{y}$ の行列 $\mathbf{U}$ は $L = \{\mathbf{X}\boldsymbol{\beta} | \mathbf{H}\boldsymbol{\beta} = \mathbf{0}\} = \Pi_\omega$ への直交射影行列だから、仮説 $\mathbf{H}\boldsymbol{\beta} = \mathbf{0} \Leftrightarrow \mathbf{X}\boldsymbol{\beta} \in \Pi_\omega$ が真ならば；

$$\begin{aligned} &\tilde{\mathbf{y}}'\mathbf{QUQ}'\tilde{\mathbf{y}} = \tilde{y}_{k+1}^2 + \cdots + \tilde{y}_r^2 \\ &\therefore\ R_1^2 = \tilde{\mathbf{y}}'\mathbf{Q}(\mathbf{I}-\mathbf{U})\mathbf{Q}'\tilde{\mathbf{y}} = \tilde{y}_1^2 + \cdots + \tilde{y}_k^2 + \tilde{y}_{r+1}^2 + \cdots + \tilde{y}_n^2 \end{aligned}$$

となることが分かる。以上から、$\dfrac{R_1^2 - R_0^2}{k} \div \dfrac{R_0^2}{n-r} = W^* \sim \mathrm{F}(k, n-r)$ は明らか。（終）

ところで、複数母数に関する検定を前述した変換群に対して不変な検定に制限することは自然な制約と考えられる。しかし、不変性とは、この検定問題を検定の良さが明確に示される一次元母数の検定問題に帰着させるために導入した人為的な制約である。検出力が有意水準以上となる検定に対象を制限する不偏性の条件と比べ、不変性とは根拠の弱い基準と言わざるをえない。実際、一様最強力不偏検定は許容的(Admissible)[12]であるが、一様最強力不変検定が許容的であるとは限らない。しかし、一様最強力不変検定である分散分析の許容性は幸運なことに示されている[13]。

[12] 一様最強力不偏検定 ϕ_0 より許容的な検定、すなわち、すべての対立仮説で検出力が大きく、すべての帰無仮説でサイズが小さい検定があったとする。この検定は不偏性の条件を満たす。これは ϕ_0 の定義と矛盾である。

[13] Lehmann(1986,p*370*)

2 一般化可能性理論の概要

従業員の判断に基づく企業の経営活動（経営特性）に関する測定は、人間の主観を通した測定という意味で、テスト理論が扱う対象となる。しかし、経営特性の測定が持つ性質を考慮すると、測定値の分析は一般化可能性理論（テスト理論の一つ）を拡張した体系の中で行わなければならない。その分析枠組みについては第４章で解説するので、ここでは前提となる一般化可能性理論を中心にテスト理論を説明していく。

次節でテスト理論を簡単に紹介した後で、この理論を２-２節で説明する。２-３節ではテストの信頼性、２-４節ではテストの妥当性を取り上げる。２-５節では項目応答理論という別のテスト理論を、２-６節ではα係数と呼ばれる信頼性の指標を補足しておく。

2-1 テスト理論の概要

最初に、いくつかの用語を定義しておく。まず、テストを構成する課題を項目(Item)と呼ぶことにする。これは、テストという言葉を所謂学力テストだけではなく、直接観測できない特性を測定する道具として一般的な意味で用いるからである。この測定値を本章では得点[1]という言葉で表すことにする。得点に影響を与える受験者や項目などの要因はテスト計画の相(Facet)と呼ばれる。

さて、テスト理論には一般化可能性理論(Generalizability Theory)と項目応答理論(Item Response Theory)という二つの考え方がある。一般化可能性理論とは、得点を真の値と誤差の和として捉える古典的テスト理論(Classical Test Theory)を相に関する分散成分モデル(Variance Components Model)[2]として体系化した理論である。従って、この理論では受験者の相だけでなく、項目や採点者などのテスト計画の相が同時に考慮される。また、こうした相の効果は最適テスト計画立案における指標となる。

しかし、テストによる測定を基本測定(Fundamental Measurement)として扱う古典的テスト理論を前提とした一般化可能性理論では、テストに回答するという行為自体が考察されることはない。さらに、基本測定の結果として得点の連続性が仮定されるため、多段階評定値への適用には論理的整合性の面で問題がある。

一方、項目応答理論とはコンピューターの発達に伴い 1970 年代以降急速に発展してきた現在における中心的なテスト理論である。得点を直接尺度と見なす一般化可能性理論とは異なり、項目応答理論ではテストによる回答を、テスト固有の母数と回答者が持つ心理学的潜在特性により定義された確率現象として扱う。従って、その測定は確率分布を媒介とした潜在特性の誘導測定(Derived Measurement)ということになる。当然、ここでは得点の離散性は問題とならない。そして、多段階評定値に対応した段階応答モデル(Graded

---

[1] 経営特性の測定を考える第４章では、連続的な値には尺度、離散的な場合は得点という用語が用いられる。これは測定という行為の意味を考慮したためである。

[2] Rao(1973,4f.2)参照。

Response Model)から連続的な尺度を表す連続応答モデル(Continuous Response Model)まで、項目の性質に応じて様々なモデルが提案されている。

2-2 一般化可能性理論

本節と次節では一般化可能性理論を要約する。この理論は、予備テストからテスト計画に含まれる相の効果を抽出するG-研究(Generalizability Study)と、そこで得られた情報を基に最適なテスト計画を立案するD-研究(Decision Study)から構成される。しかし、質問項目の選択を扱う5-2節で説明するように、経営特性の分析でD-研究が適用できる余地は少ない。そこで、本節ではG-研究の考え方について説明していく。

2-2-1 真の得点と誤差

一般化可能性理論の解説に先立ち、その理論の前提となる古典的テスト理論以来の得点の捉え方を説明していく。古典的テスト理論では、受験者 a の項目 g による得点 y_{ga} を仮想的な試行における実現値と考えて確率変数 Y_{ga} として扱う。そこで、相を表す添え字をベクトル $\mathbf{F}$、その得点を確率変数 $Y_{\mathbf{F}}$ で表し、この考え方を一般的に説明しよう。一般化可能性理論では、テスト（項目）の試行に関する得点 $Y_{\mathbf{F}}$ の期待値 $T_{\mathbf{F}} = \mathrm{E}(Y_{\mathbf{F}})$ を相ベクトル $\mathbf{F}$ の下での特殊真値(Specific True-Score)と呼び、得点と特殊真値との差 $E_{\mathbf{F}} = Y_{\mathbf{F}} - T_{\mathbf{F}}$ を特殊誤差(Specific Error)と呼ぶ。

固定された相ベクトル $\mathbf{F}$ の下で確率変数である特殊誤差 $E_{\mathbf{F}}$ の持つ性質は、その相ベクトルに依存する。そこで；

仮定 2-2.1

同じ相で構成される任意の相ベクトル $\mathbf{F}_1, \mathbf{F}_2$ について、 $E_{\mathbf{F}_1}$ と $E_{\mathbf{F}_2}$ は独立である。

と仮定する。ただし、必ずしも通常の意味での独立性が必要となるわけはでなく、古典的テスト理論では項目の相に関して線形実験的独立(Linear Experimental Independent) が保障されていれば十分である[3]。

[3] Lord and Novick(1968,p*45*)

2-2-2 相Fに関する説明

特殊真値$T_{\mathbf{F}} = \mathrm{E}(Y_{\mathbf{F}})$は固定された相ベクトル**F**の下では定数であるが、異なる相ベクトル（構成する相は等しい）に関しては異なる値となる。例えば、企業調査で異なる企業の特殊真値が異なる値となるのは当然であろう。一般化可能性理論の基本的な考え方は、特殊真値を相ベクトルを構成する各相の効果に分解し、関心の対象である相の純粋な効果を抽出するというものである。この考え方を具体的に説明する前に、相ベクトルに関する議論をまとめておく。

最初に、二つの相 A と B に関するネスト(Nest)とクロス(Cross)という分散分析の用語について説明しておこう。まず、相Bの各要素に対し相Aの複数かつ異なる要素が対応するとき、相Aは相Bの中にネストされているという。これを$A:B$と表す。相Aの共通する要素が対応する場合はクロスしていると呼び、$A \times B$ と表す。従業員による回答しか得られないという企業調査の制約は、回答者が相は企業の相の中にネストされていることを意味している。回答者と質問項目の相はクロスするから、一般的な企業調査とは質問項目×（回答者：企業）計画と言うことができる。

一般化可能性理論では原則として相の具体的な要素（企業の相ならば具体的な企業）を仮想母集団からの無作為標本と考える。企業の例では、各企業は企業の母集団からの無作為標本として扱われる。このように、特殊真値は相に関する確率変数となるが、無作為性を仮定するのは特殊真値間の独立性を保証するためである。業種や国籍のように母集団からの無作為標本とは想定し難い相は固定された相(Fixed Facet)と呼ばれ、特殊真値をこの相に関する確率変数として扱う必要はなくなる。実際、統計量を推定量ではなく母数として解釈する以外に、固定された相を通常の相と区別する必要なほとんどない。

また、予備テストで得られた情報を用いてテスト計画を最適化するときに制御可能となる相を一般化の母域(Universe of Generalization)と呼ぶ。

2-2-3 弱真値モデル

固定された相やネストされた相を含むテスト計画に関して特殊真値の分解を一般的に記述することは非常に面倒である。そこで、特定の項目を固定して受験者の相のみを対象とする古典的テストモデル(Classical Test Theory Model)と、複数項目のテスト計画を扱う弱真値モデル(Weak True-Score Model)を用いて、この分解を説明していく。

まず、弱真値モデルから説明していこう。弱真値モデルが対象とするテスト計画は（受験者×項目）計画であり、受験者は受験者の母集団から、実施される項目は同じ測定目的を持つ項目の母集団からの無作為標本であることが前提である。そして、項目と受験者を添え字g,aで表すことにする[4]。

得点に含まれる項目と受験者に基づく効果を；

---

[4] この場合は$\mathbf{F} = (g,a)$である。

$$\mathrm{E}_g\mathrm{E}_a(T_{ga}) = \mu$$
$$\mathrm{E}_a(T_{ga}) = \mu_g \;\Rightarrow\; \nu_g = \mu_g - \mu\text{; 項目効果}$$
$$\mathrm{E}_g(T_{ga}) = \mu_a \;\Rightarrow\; \nu_a = \mu_a - \mu\text{; 受験者効果}$$

と定義する。これらの効果で説明できない特殊真値 T_{ga} の変動を、項目×受験者の交互作用効果 ν_{ga} と表せば、受験者 a が項目 g を受けたときの得点 Y_{ga} は；

$$Y_{ga} = T_{ga} + E_{ga} = \mu + \nu_g + \nu_a + \nu_{ga} + E_{ga}$$

と恒等的に分解することができる。ここで、$\sigma_e^2 = \mathrm{E}_g\mathrm{E}_a\mathrm{E}(E_{ga}^2)$ とおけば、項目と受験者間での変動を含む得点 Y_{ga} の全分散は；

$$\sigma_{ga}^2(Y_{ga}) = \mathrm{E}_g\mathrm{E}_a\mathrm{E}(Y_{ga} - \mu)^2 = \sigma_g^2 + \sigma_a^2 + \sigma_{ga}^2 + \sigma_e^2$$

と各相に対応した分散成分の和として表現することができる。

（受験者×項目）計画における一般化の母域は項目の相である。このとき、受験者 a の母得点(Universe Score) τ_a は特殊真値から項目による影響を除いた総平均と受験者効果の和 $\tau_a = \mu + \nu_a = \mu_a$ と定義される。そして、得点と母得点との差 $\Delta_{ga} = Y_{ga} - \tau_a$ は絶対誤差(Absolute Error)[5]と呼ばれる。この母得点を用いた得点の表現；

$$Y_{ga} = \tau_a + \Delta_{ga} = \mu_a + \Delta_{ga}$$

が弱真値モデルである。一般化可能性理論において母得点と絶対誤差の定義がテスト計画に依存することは、これらの定義から明らかであろう。また、この母得点を弱真値理論では受験者 a の一般真値(Generic True-Score)、絶対誤差を受験者 a の項目 g に対する一般誤差(Generic Error)と呼ぶ。

[5] ⇔相対誤差(Relative Error)、Shavelson and Webb(1991,p*84*)

2-2-4 古典的テストモデル

古典的テストモデルは特定の項目 g を固定した（受験者）計画を対象とする。項目が固定されるので、受験者 a の得点は；

$$Y_{ga} = T_{ga} + E_{ga} = \mu + \nu_a + E_{ga}$$

と分解することができる。$\sigma_e^2 = \mathrm{E}_a\mathrm{E}(E_{ga}^2)$ とすれば、得点 Y_{ga} の全分散は；

$$\sigma_a^2(Y_{ga}) = \mathrm{E}_a\mathrm{E}(Y_{ga} - \mu)^2 = \sigma_a^2 + \sigma_e^2$$

となる。一般化の母域はないため、その母得点は特殊真値 $\tau_a = \mu + \nu_a = T_{ga}$ となり、絶対誤差は特殊誤差になる。この分解による得点の表現；

$$Y_{ga} = \tau_a + \Delta_{ga} = T_{ga} + E_{ga}$$

を古典的テストモデルと呼ぶ。ただし、この説明は一般化可能性理論の視点から記述したものであり、通常は、特殊真値と特殊誤差による得点の分解として定義される。

2-2-5 補足：古典的テスト理論と合成テスト、平行性

測定値の信頼性を高めるには複数回の測定が必要という主張は、多くの場合、統計理論により保証される。しかし、記憶や学習効果から、実際に同一項目を複数回実施して意味のある測定値を得ることは難しい。そのため、「同じ」項目と見なされる複数項目からなる合成テスト(Composite Test)が代替案として実施される。一般化可能性理論が合成テストを前提としていることは弱真値モデルの考え方から明らかであろう。

ところで、古典的テスト理論が体系化された時点では、項目を相として扱う考え方は採られていなかったため、G 個の項目による合成テストは合計得点 $Z_a = \sum_{g=1}^{G} Y_{ga}$ を得点とする一つの項目として扱われた。そして、この考え方の下で平行性(Parallelism)と α 係数という古典的テスト理論独自の概念が導入された。まず、G 個の項目が厳密に同一項目の繰り返しと考えられるとき、これらの項目は平行(Parallel)であるという。次節で説明するように、信頼性の指標である特殊信頼性係数(Specific Reliability Coefficient)を推定するには繰り返しか項目間の平行性が必要となる。合成テストの特殊信頼性係数を推定する場合

も合成テスト自体の繰り返しか項目間の平行性が要請される。
しかし、平行性は繰り返しとほとんど同じ概念だから、実際に平行テストを用意すること自体が難しい。この問題を解決するために開発されたモデルが繰り返しと平行性を要求しない弱真値モデルである。一方、古典的テスト理論の枠内では推定に前述した条件を必要としないα係数（特殊信頼性係数の下限）が考案された。そこで、2-6節でα係数と特殊信頼性係数との関係を記述するのに必要な概念を最後に説明しておこう。
独立かつ同一の分布に従う確率変数が得られたとき、統計学では当該確率モデルからの繰り返し(Replications)が得られたという。しかし、古典的テスト理論は特定の分布を仮定しないため、任意の受験者 a と R 個の項目について；

$$T_{1a} = T_{2a} = \cdots = T_{Ra}$$
$$\sigma^2(Y_{1a}) = \sigma^2(Y_{2a}) = \cdots = \sigma^2(Y_{Ra})$$

が成立し、R 個の項目が独立（線形実験的独立）であれば十分である。そこで、この条件を満たす項目を同等(Homogeneous)であるとし、同一の項目 g から R 個の繰り返しが得られたと考える。この繰り返しを添え字 r により Y_{gar}；$r = 1,\ldots,R$ と表し、これらの項目を平行テストと呼ぶ。ここから、古典的テスト理論において繰り返しと平行テストが同じ概念であることが分かる。
ところで、上の特殊真値に関する条件；

$$T_{1a} = T_{2a} = \cdots = T_{Ra}$$

と線形実験的独立性が成立するとき、R 個の項目はτ等価(τ-equivalent)であるという。この条件を緩めて；

$$T_{ga} = c_{gh} + T_{ha}$$

と平行移動を認めた場合は本質的にτ等価(Essentially τ-equivalent)であるという。弱真値モデル $T_{ga} = \mu + \nu_g + \nu_a + \nu_{ga}$ において、項目と受験者の母集団で $\nu_{ga} = 0$（交互作用効果が存在しない）が成立すれば、各項目は本質的にτ 等価である。さらに、受験者の母集団で $\nu_g = 0$（項目効果がない）であればτ等価となる。

2-3 テストの信頼性

2-3-1 信頼度指数

テストの信頼性(Reliability)とは測定値の精度や正確さを反映させる概念であり、その指標はモデルに依存する。一般化可能性理論では、誤差分散との関連で定義された信頼度指数(Index of Dependability)が信頼性の指標として用いられる。

信頼度指数Φは母得点τの分散$\sigma^2(\tau)$と絶対誤差Δの分散$\sigma^2(\Delta)$による比；

$$\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau) + \sigma^2(\Delta)}$$

と定義される。もちろん、各分散成分は対象とする相ベクトル$\mathbf{F}$について適当に定義されているものとする。

ところで、信頼度指数を一つの項目が持つ信頼性の指標として定義する場合は、この式の分母について$\sigma^2(\tau) + \sigma^2(\Delta) = \sigma^2_{\mathbf{F}}(Y_{\mathbf{F}})$が成立する。しかし、合成テストの場合は母得点と絶対誤差の分散の和が合計得点の全分散となり、この関係は成立しない。これらの違いについては*2-3-2*で説明される。

2-3-2 一般信頼性係数

弱真値モデルにおける受験者aの母得点は$\tau_a = \mu_a$だから、その信頼度指数は；

$$\Phi = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{ga}^2(\Delta_{ga})} = \frac{\sigma_a^2}{\sigma_a^2 + (\sigma_g^2 + \sigma_{ga}^2 + \sigma_e^2)}$$

と定義される。この指標は弱真値理論の用語で一般信頼性係数(Generic Reliability Coefficient)と呼ばれる。一般信頼性係数が、得点Y_{ga}と母得点$\tau_a = \mu_a$の間の、テストの試行と相（項目と受験者）に関する相関係数の二乗、すなわち；

定理 2-3.1

$$\Phi = \rho_{ga}^2(Y_{ga}, \mu_a)$$

であることは容易に示される。

ところで、G 個の項目による合成テストの母得点は $\tau_a = G\mu_a$ と定義される。合計得点の全分散は各相の分散成分により；

$$\sigma_{ga}^2\left(\sum_{g=1}^{G} Y_{ga}\right) = G\sigma_g^2 + G^2\sigma_a^2 + G\sigma_{ga}^2 + G\sigma_e^2$$

と分解されるから、合成テストの信頼度指数は次のようになる。

$$\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\Delta)} = \frac{\sigma_a^2}{\sigma_a^2 + \left(\dfrac{\sigma_g^2+\sigma_{ga}^2+\sigma_e^2}{G}\right)}$$

当然、この値は一般信頼性係数より大きい。この関係式は目標とする信頼度指数の達成に必要な項目数 G を求めるのに使うことができる。これは D-研究の一例である。

2-3-3 特殊信頼性係数

古典的テストモデルにおける受験者 a の母得点は特殊真値 $\tau_a = T_{ga}$ だから、固定した項目 g に関する信頼度指数は；

$$\Phi = \frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2} = \frac{\sigma_a^2}{\sigma_a^2(Y_{ga})}$$

と定義される。これを古典的テスト理論ではテスト g の特殊信頼性係数と呼ぶ。

二つの平行テスト g,h が与えられたとき、次の定理が成立する。

定理 2-3.2

$$\Phi = \rho_a(Y_{ga}, Y_{ha}) = \rho_a^2(Y_{ga}, T_{ga})$$

この証明も容易である。これは、特殊信頼性係数が二つの平行テストによる得点間の相関係数であり、得点と特殊真値との相関係数の二乗に等しいことを意味している。また、相関係数はテストの試行と受験者に関して取られている。

2-3-4 合成テストの特殊信頼性係数

この特殊信頼性係数の定義は合成テストにも適用できる。しかし、合成テストの信頼度指数については既に *2-3-2* で定義したように、ここで定義される特殊信頼性係数はもはや合成テストの信頼度指数ではない。合成テストの特殊信頼性係数とは、2-6 節で説明する α 係数に繋がる古典的テスト理論独自の概念である。

合成テスト Z の特殊信頼性係数を ρ_Z^2 と表せば、この特殊信頼性係数は；

$$\rho_Z^2 = \frac{\sigma_a^2(T_{Za})}{\sigma_a^2(Z_a)} = \frac{\sigma_a^2\left(\sum_{g=1}^{G} T_{ga}\right)}{\sigma_a^2\left(\sum_{g=1}^{G} Y_{ga}\right)} = 1 - \frac{\sum_{g=1}^{G}\sigma_a^2(E_{ga})}{\sigma_a^2\left(\sum_{g=1}^{G} Y_{ga}\right)}$$

と定義することができる。

G 個の項目が平行であれば、得点と特殊真値の分散は項目に依存しないから、各項目の特殊信頼性係数も項目に依存しないことが分かる。これらの値を順に；

$$\begin{aligned}
\sigma_Y^2 &= \sigma_a^2(Y_{ga}); g = 1,\ldots,G \\
\sigma_{T_Y}^2 &= \sigma_a^2(T_{ga}) = \sigma_a(Y_{ga}, Y_{ha}); g,h = 1,\ldots,G \\
\rho_Y^2 &= \sigma_a^2(T_{ga})\big/\sigma_a^2(Y_{ga}) = \sigma_{T_Y}^2\big/\sigma_Y^2 ; g = 1,\ldots,G
\end{aligned}$$

として、定理 2-3.2 の第一式を適用すると、平行テストによる合成テストの特殊信頼性係数が；

定理 2-3.3

$$\rho_Z^2 = \frac{G^2\sigma_{T_Y}^2}{G\sigma_Y^2\left(1+(G-1)\rho_Y^2\right)} = \frac{G\rho_Y^2}{1+(G-1)\rho_Y^2}$$

と表せることが分かる。これを Spearman-Brown の公式という。この定理は平行テストによる特殊信頼性係数 ρ_Z^2 が項目数の増加関数という性質だけではなく、その推定で項目間の平行性が重要な意味を持つことを示している。この話題については *2-3-5* と 2-6 節で再び取り上げる。

2-3-5 信頼度指数の推定

本章では信頼度指数の推定に、その定義式に各分散成分の不偏推定量を代入するという方法を採用する。分散成分の不偏推定量は、分散分析における平均平方和のテストの試行と相に関する期待値が分散成分の線形関数になるという性質から、各平均平方和を連立方程式として解くことで求められる。この方法を用いるのは、経営特性の分析では分散分析が用いられるので、その適用が容易なためである。また、信頼度指数が必要となる質問項目の選択では、分散成分に関する議論は各相の効果に関する議論に転化されるため[6]、分散成分の推定値自体にそれほどの重要性はない。従って、敢えて面倒な最尤法や MINQUE (Minimum Norm Quadratic Unbiased Estimator)法のような分散成分の推定法を用いる必要はないと思われる。ただし、平均平方和から逆算する推定方法では、分散成分の推定値に非負性を保証できないことを注意しておく[7]。

それでは、これまで紹介してきた一般信頼性係数と特殊信頼性係数、合成テストの特殊信頼性係数の推定方法を順に説明していく。まず、一般信頼性係数は項目×受験者の二元配置分散分析表（表2-3.1）を用いて推定することができる。この表については5-5節で説明されるので、ここでは次の二点だけ補足しておく。一点目は、期待値 E(平方和÷自由度)がテストの試行だけではなく、項目と受験者に関して取られていることである。二点目は、添え字“・”の付いた変数が対応する添え字に関する標本平均を表すという表記上の約束である。

さて、一般信頼性係数は；

$$\Phi = \sigma_a^2 \Big/ \left(\sigma_a^2 + \sigma_{ga}^2(\Delta_{ga})\right), \sigma_{ga}^2(\Delta_{ga}) = \sigma_g^2 + \sigma_{ga}^2 + \sigma_e^2$$

と定義されるから、表2-3.1より次のように推定すればよい。

$$\Phi = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_{ga}^2(\Delta_{ga})}$$

$$\hat{\sigma}_a^2 = \frac{\text{Between Persons}}{G(A-1)R} - \frac{\text{Interaction}}{G(G-1)(A-1)R}$$

$$\hat{\sigma}_{ga}^2(\Delta_{ga}) = \frac{\text{Between Test} \times \text{Person Cells}}{GA-1} - \hat{\sigma}_a^2; R > 1$$

$$\hat{\sigma}_{ga}^2(\Delta_{ga}) = \frac{1}{A(G-1)}(\text{ Between Persons} - \text{ Interaction}); R = 1$$

ここから、一般信頼性係数の推定には繰り返しを必要としないことも分かる。

---

[6] 5-2節参照。

[7] 非負条件を付けない限り、通常の方法では非負性は満たされない。

特殊信頼性係数は固定した項目 g の下での受験者に関する一元配置分散分析により推定することができる。これは表2-3.2にまとめてある。ただし、R 回の繰り返しは R 個の平行な項目と考えてもよい。特殊信頼性係数の定義式は $\Phi=\sigma_a^2/(\sigma_a^2+\sigma_e^2)$ だから；

$$\Phi=\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2}$$

$$\hat{\sigma}_a^2=\frac{\text{Between Persons}}{(A-1)R}-\frac{\text{Within Cells}}{AR(R-1)}$$

$$\hat{\sigma}_e^2=\frac{\text{Within Cells}}{A(R-1)}$$

と推定される。この式から明らかなように、特殊信頼性係数の推定には繰り返しか、項目間の平行性が必要となる。

最後に、合成テスト Z の特殊信頼性係数 ρ_Z^2 の推定方法を説明する。もし、合成テスト自体の繰り返しが得られている場合は、項目 g ではなく合成テスト Z に関して上の方法を適用すればよい。繰り返しがなくても項目間の平行性が満たされていれば項目の特殊信頼性係数 ρ_Y^2 は推定できるから、定理2-3.3を用いて特殊信頼性係数 ρ_Z^2 を推定することができる。しかし、平行性も成立しなければ推定は不可能だから、この場合は特殊信頼性係数の下限として定義される α 係数[8]で代用せざるをえない。

2-4 テストの妥当性

テストの妥当性(Validity)とは信頼性とならぶ重要な概念である。これは、当該テストが何かある対象を測定しようとするものならば、本当にその「何か」を測定しているか否かを示す概念である。この説明からも明らかなように、妥当性とは間接測定の妥当性を意味する概念であり、間接測定が中心となる心理学特有の考え方である。そのため、直接測定が可能な場合に問題とされることはない。

妥当性の概念は様々な方法で表現されるが、次の二つの概念により補完的に捉えれば十分であろう。一つ目は、内容的妥当性(Content Validity)と呼ばれる概念である。これは分析者がテストの目的を鑑み系統的にテストの取捨選択を行った結果として得られる妥当性である。予め内容的妥当性を保証しておくことは最低限の要件であり、系統的な手続きを取らず思い付きで編集されたテストとは区別すべきである[9]。当然、内容的妥当性は明確に数量化できる概念ではないが、二つ目の構成的妥当性(Construct Validity)は指標として表

---

[8] 2-6節参照。

[9] 池田(1980,p*56*)

すことができる。

構成的妥当性は理論的に妥当性を持つと考えられる基準変数の構成概念（潜在特性）との関連として定義される。直接の関連で定義される妥当性が経験的妥当性(Empirical Validity)だから、構成的妥当性とは経験的妥当性を一般化した概念と言うことができる。経験的妥当性の指標には後述する妥当性係数などがある。

しかし、テストの実施前に基準変数を見つけることは難しく、別のテストの得点が基準変数として用いられることが多い。こうした場合に構成的妥当性による妥当性の検証は有効な方法であるが、その結果を盲目的に受け入れることには注意が必要である。内容的妥当性を持つテストの構成的妥当性の高さがテストの妥当性を保証する有力な根拠であることは事実でも、テストの得点と想定したモデルの適合度が意味することは、妥当性を仮定したテストの信頼性に過ぎないからである。そこで、妥当性と信頼性の関係を示す次の定理を最後に紹介しておく。この定理は質問項目の選択を議論する5-2節で必要となる。

項目 g の特殊信頼性係数を Φ とし、任意の基準変数を Z_a と表そう。このとき、以下の不等式が成立する。

定理 2-4.1

$$\rho_a(Y_{ga}, Z_a) \le \sqrt{\Phi}$$

（証明） *2-2-1* の議論から $\sigma_a(Y_{ga}, T_{ga}) = \sigma_a^2(T_{ga}), \sigma_a(Y_{ga}, T_{Za}) = \sigma_a(T_{ga}, T_{Za})$ は容易に示される。このとき；

$$\begin{aligned}\rho_a(Y_{ga}, T_{Za}) &= \frac{\sigma_a(Y_{ga}, T_{Za})}{\sigma_a(Y_{ga})\sigma_a(T_{Za})} = \frac{\sigma_a^2(T_{ga})}{\sigma_a(Y_{ga})\sigma_a(T_{ga})} \frac{\sigma_a(Y_{ga}, T_{Za})}{\sigma_a(T_{ga})\sigma_a(T_{Za})} \\ &= \rho_a(Y_{ga}, T_{ga})\rho_a(T_{ga}, T_{Za}) \\ &\therefore\ \rho_a(Y_{ga}, T_{ga}) \ge \rho_a(Y_{ga}, T_{Za})\end{aligned}$$

が成立する。さらに、$\sigma_a^2(Y_{ga}, Z_a) = \sigma_a^2(Y_{ga}, T_{Za})$ に注意すれば；

$$\begin{aligned}\rho_a^2(Y_{ga}, Z_a) &= \frac{\sigma_a^2(Y_{ga}, T_{Za})}{\sigma_a^2(Y_{ga})\sigma_a^2(Z_a)} = \frac{\sigma_a^2(Y_{ga}, T_{Za})}{\sigma_a^2(Y_{ga})} \frac{\sigma_a^2(Z_a, T_{Za})}{\sigma_a^2(T_{Za})} \\ &= \rho_a^2(Y_{ga}, T_{Za})\sigma_a^2(Y_{ga}, T_{Za}) \\ &\therefore\ \rho_a^2(Y_{ga}, T_{Za}) \ge \rho_a^2(Y_{ga}, Z_a) \\ &\therefore\ -\rho_a(Y_{ga}, T_{Za}) \le \rho_a(Y_{ga}, Z_a) \le \rho_a(Y_{ga}, T_{Za})\end{aligned}$$

となるから、以上をまとめることで；

$$\sqrt{\Phi} = \rho_a(Y_{ga}, T_{ga}) \geq \rho_a(Y_{ga}, Z_a)$$

が得られる。（終）

相関係数$\rho_a(Y_{ga}, Z_a)$を妥当性係数という。この値が経験的妥当性の指標であることは既に述べた通りである。この定理で基準変数Z_aは任意だから、いかなる基準変数の妥当性係数も特殊信頼性係数の平方根を超すことはできない。これは信頼性の高さが妥当性を保証する必要条件であることを意味している。

2-5 補足1：項目応答理論

現代のテスト理論の主流は項目応答理論である。得られた得点を直接扱う一般化可能性理論とは異なり、項目応答理論は得点が生起する確率を基に理論が展開される。この確率は回答者が対象を測定するという行為をモデル化することで導かれる。範囲や離散性など制約の多い得点を単純に分解していく一般化可能性理論に比べて、この項目応答理論は優れた考え方と言うことができる。しかし、本論文で扱う経営特性の測定に項目応答理論を適用することは難しい。4-5節で説明するように、このモデルの枠内では、通常の企業調査で得られる測定値に基づいた仮説検定ができないからである。そこで、この議論の準備として、本節では段階応答モデルを用いて項目応答理論の簡単な説明を与えておく。

2-5-1 項目応答モデルの考え方

n個の項目が与えられたとき、少なくとも一つの項目の応答に影響を与える心理学上の構成概念を潜在特性(Latent Trait)θと表すことにする。潜在特性とは、その構成概念を説明するために導入した人為的な変数であり、その直接観測できない尺度である。n個の項目に関する潜在特性の全体は完備潜在空間[10](Complete Latent Space)を構成するが、ここでは簡単化のため完備潜在空間を一次元と仮定する。もちろん、この仮定により以下の議論が一般性を失うことはない。また、項目応答理論では受験者に関する情報はすべて潜在特性θに含まれるため、受験者を表す添え字aは原則として省略される。

潜在特性θを与えたとき、n個の項目の得点に関する条件付き同時分布Fが各得点の条件付き周辺分布F_gの積になっているとき、すなわち；

[10] 厳密な定義はLord and Novick(1968,p*359*)を参照。

$$F(y_1,\ldots,y_G|\theta)=\prod_{g=1}^{G}F_g(y_g|\theta)$$

であるとき、これらの項目は局所独立(Local Independence)であるという。項目応答理論では仮定２-２.１に対応して以下の仮定をおく。

仮定 2-5.1

対象とする n 個の項目は局所独立である。

次に、潜在特性 θ の受験者が項目 g で得点 s を取る確率を；

$$P_g(s|\theta)=P(Y_g=s|\theta);s=0,\ldots,m$$

と定義する。この確率モデルで潜在特性以外の母数は項目の特性を規定する母数としてテスト母数と呼ばれる。また、この受験者の得点が s 点を超える確率；

$$P_g^*(s|\theta)=\sum_{j=s}^{m}P_g(j|\theta)$$

を得点 s に対応する項目応答曲線(Item Characteristic Curve)と呼ぶ。項目応答曲線は潜在特性の単調増加関数と仮定されるが、これはテストの性質を考えれば当然であろう。

確率 $P_g(s|\theta)$ は連続応答モデルへの拡張と操作性を考慮して密度関数による表記で定義されることが多い。この密度関数を ψ とすれば；

$$P_g(s|\theta)=\int_{a_g(\theta-b_s)}^{a_g(\theta-b_{s-1})}\psi_g(y)dy=a_g\int_{b_{s-1}}^{b_s}\psi_g(a_g(\theta-y))dy$$

と定義される。ここで、母数 b_s は得点 s に対応した困難度水準(Difficulty Level)の指標であり、$m+1$ 個の得点に対応したカテゴリーの境界値を表している。母数 a_g は項目 g の持つ識別力(Discrimination Power)の指標である。

2-5-2 テストの信頼性

項目応答理論では、項目 g の信頼性を潜在特性 θ の受験者が項目 g に回答したときの信

頼性として考え、潜在特性θに依存した項目情報関数(Item Information Function)；

$$\mathrm{I}(\theta, Y_g) = \mathrm{E}_{Y_g}\left(\frac{d}{d\theta}\log \mathrm{P}_g\left(Y|\theta\right)\right)^2$$

が指標として用いられる。これは統計理論におけるフィッシャー情報量(Fisher Information)に外ならず、この値を信頼性の指標として用いるのは、項目情報関数（フィッシャー情報量）が以下に述べる性質を持つためである。

潜在特性θの任意の不偏推定量$\hat{\theta}$について、Cramér-Rao の不等式；

$$\sigma^2(\hat{\theta}) \geq \frac{1}{\mathrm{I}(\theta, Y)}$$

が成立する。厳密な言い方ではないが、この不等式は潜在特性θの不偏推定量の分散を項目情報関数の逆数まで小さくできることを意味している。従って、項目情報関数が大きいほど分散の小さな不偏推定量を得ることができる。信頼性の高い項目とは潜在特性が正確に反映される項目のことだから、その信頼性は推定した潜在特性の分散として示すことができる。これが、信頼性の指標として項目情報関数を用いる理由である。

2-5-3 潜在特性とテスト母数の推定

潜在特性とテスト母数の推定には通常の最尤法ではなく後述する周辺最尤法(Marginal Maximum Likelihood Estimation)の適用が一般的である[11]。しかし、本論文が対象とする企業調査への適用には問題が多い。4-5節で示すように、これが企業調査に項目応答理論を適用できない最大の理由となっている。

さて、確率モデルの母数は観測値に依存する付随母数(Incidental Parameter)と、すべての観測値に共通な構造母数(Structural Parameter)の二つに分類される。項目応答モデルでは付随母数が潜在特性に該当し、構造母数が各項目のテスト母数（ベクトル）に対応することは明らかであろう。この分類を与えたNeyman and Scott(1948)は、構造母数の最尤推定量が一致推定量になるとは限らないことと、たとえ一致性が成立しても漸近有効性(Asymptotic Efficiency)が保証されないことを示した。このように、受験者数を増やしてもテスト母数の最尤推定量に一致性を保証することができない。テスト理論の目的はテスト作成にあるから、これは最尤法の持つ大きな問題と言うことができる。

これに対し、周辺最尤法では比較的重要性の低い潜在特性θを受験者の母集団で定義された確率変数として扱う。そこで、観測された得点の持つ確率の潜在特性に関する期待値を取れば、潜在特性に依存しないテスト母数の周辺尤度関数；

---

[11] 池田(1994,p *78*)

$$\mathrm{E}_{\theta}\left(\prod_{a}\prod_{g}\mathrm{P}_{g}\left(s_{a}\middle|\theta_{a}\right)\right)$$

を求めることができる。この周辺尤度関数に関する最尤法が周辺最尤法である。潜在特性の分布を適当に仮定することで、周辺最尤法が最尤法の望ましい性質を持つことも知られている[12]。テスト母数の周辺最尤推定量が一致推定量であれば、この推定値を基にした最尤法により潜在特性を推定することも可能となる。

2-6 補足2：α係数

*2-3-5*で述べたように、合成テストZの特殊信頼性係数ρ_Z^2の推定には、項目間の平行性か合成テスト自体の繰り返しが必要となる。項目を相の一つとして捉えることがない古典的テスト理論で、この問題を解決するために導入した信頼性の指標が、合成テストの特殊信頼性係数の下限として定義されるα係数である。そして、経営学の分野で最も広く用いられている信頼性の指標がα係数である。平行性や繰り返しが得られないという操作上の理由のほか、一般化可能性理論自体が普及していないこともあろう。そこで、本章の最後にα係数を簡単に説明しておく。

定理 2-6.1

$$\rho_Z^2 \geq \alpha = \frac{G}{G-1}\left(1-\frac{\sum_{g=1}^{G}\sigma_a^2(Y_{ga})}{\sigma_a^2(Z_a)}\right)$$

（証明）Cauchy-Schwarz の不等式から次式が成立する。

$$\left(\sigma_a(T_{ia})-\sigma_a(T_{ja})\right)^2 \geq 0 \;\Rightarrow\; \sigma_a^2(T_{ia})+\sigma_a^2(T_{ja}) \geq 2\sigma_a(T_{ia})\sigma_a(T_{ja}) \geq 2\sigma_a(T_{ia},T_{ja})$$

この式を$i \neq j$について加えることで；

[12] Lehmann(1983,p*450*)

$$\sum_{i=1}^{G}\sum_{j=1}^{G}\left(\sigma_a^2(T_{ia})+\sigma_a^2(T_{ja})\right)=2G\sum_{i=1}^{G}\sigma_a^2(T_{ia})$$

$$=2\sum_{i=1}^{G}\sigma_a^2(T_{ia})+\sum\sum_{i\neq j}\sigma_a(T_{ia})\sigma_a(T_{ja})$$

$$\therefore\ \sum_{i=1}^{G}\sigma_a^2(T_{ia})=\frac{\sum\sum_{i\neq j}\sigma_a(T_{ia})\sigma_a(T_{ja})}{G-1}\geq\frac{\sum\sum_{i\neq j}\sigma_a(T_{ia},T_{ja})}{G-1}$$

となる。一方、$\sigma_a^2(T_{Za})=\sigma_a^2\left(\sum T_{ia}\right)=\sum\sigma_a^2(T_{ia})+\sum\sum_{i\neq j}\sigma_a(T_{ia},T_{ja})$ だから、上の式をまとめることで；

$$\sigma_a^2(T_{Za})\geq\frac{G}{G-1}\sum\sum_{i\neq j}\sigma_a(T_{ia},T_{ja})$$

であることが分かる。最後に、この式と；

$$\sigma_a^2(Z_a)-\sigma_a^2\left(\sum Y_{ia}\right)=\sum\sum_{i\neq j}\sigma_a(Y_{ia},Y_{ja})=\sum\sum_{i\neq j}\sigma_a(T_{ia},T_{ja})$$

を特殊信頼性係数の定義式 $\rho_Z^2=\sigma_a^2(T_{Za})/\sigma_a^2(Z_a)$ に代入すれば証明は完了する。（終）

α 係数は通常の分散の不偏推定量を用いることで推定できる。この推定量を定義式に代入することで；

$$\hat{\alpha}=\frac{G}{G-1}\left(1-\frac{\sum_{g=1}^{G}\hat{\sigma}_a^2(Y_{ga})}{\hat{\sigma}_a^2(Z_a)}\right)$$

$$\hat{\sigma}_a^2(Y_{ga})=\frac{1}{G-1}\sum_{a=1}^{A}(y_{ga}-y_{g\bullet})^2$$

$$\hat{\sigma}_a^2(Z_a)=\frac{1}{A-1}\sum_{a=1}^{A}(z_a-z_{\bullet})^2$$

と推定できることが分かる。

ところで、$R=1$ とした表２-３.１を用いることで、この推定式は次のように書き換える

ことができる[13]。ただし、$\hat{\sigma}_a^2(Z_a)$ は上で定義した通りである。

$$\hat{\alpha} = 1 - \frac{G(\text{Interaction})}{(G-1)(A-1)\hat{\sigma}_a^2(Z_a)}$$

この α 係数の推定式において；

$$\frac{G}{(G-1)(A-1)}\mathrm{E}(\text{Interaction}) = G\sigma_e^2 + G\sigma_{ga}^2$$

であることは表2-3.1に示されている。従って、各項目が本質的に τ 等価であれば、この期待値は $G\sigma_e^2$ となる[14]。ここで、$\sigma_e^2 = \mathrm{E}_g\mathrm{E}_a\mathrm{E}(E_{ga}^2) = \mathrm{E}_g\,\sigma_a^2(E_{ga})$ に注意すると；

$$\sigma_a^2(E_{Za}) = \sum_{g=1}^{G}\sigma_a^2(E_{ga}) \Rightarrow \mathrm{E}_g\,\sigma_a^2(E_{Za}) = G\sigma_e^2$$

だから、$\dfrac{G(\text{Interaction})}{(G-1)(A-1)}$ が $\sigma_a^2(E_{Za})$ の推定値となることが分かる。これは、各項目が本質的に τ 等価であれば、合成テストの特殊信頼性係数 $\rho_Z^2 = 1 - \sigma_a^2(E_{Za})/\sigma_a^2(Z_a)$ を α 係数の推定値で代替できることを意味している。

この結果は次の定理によって理論的に保証される。

定理 2-6.2

$\rho_Z^2 = \alpha$ となる必要十分条件は各項目が本質的に τ 等価なことである。

（証明）これは、α 係数の証明で不等式を導く関係式；

---

[13] Lord and Novick(1968,p*204*)

[14] 本質的に τ 等価ならば $\nu_{ga} = 0$ である。*2-2-5*参照。

$$\sigma_a^2(T_{ia})+\sigma_a^2(T_{ja})\geq 2\sigma_a(T_{ia})\sigma_a(T_{ja})\geq 2\sigma_a(T_{ia},T_{ja})$$

の等号が成立する必要十分条件が、$T_{ia}=c_{ij}+T_{ja}$、すなわち、各項目が本質的にτ等価であることから明らかである。（終）

各項目が平行テストかτ等価であれば本質的にτ等価だから、こうした場合も合成テストの特殊信頼性係数はα係数と一致する。これは繰り返しが得られないときにα係数を用いる理論的根拠となっている。しかし、平行テストの得点から本質的にτ等価であるための条件$\nu_{ga}=0$が実際に推定されるとは限らない。この条件が成立しなければ、α係数の推定値と特殊信頼性係数の推定値が一致することはない。従って、項目間の平行性を事前に仮定できる場合、特殊信頼性係数の推定には平行性の性質から特殊信頼性係数を直接導いた定理2-3.3の関係式を使うべきであろう。

最後に、α係数の値は共分散$\sigma_a(T_{ia},T_{ja})$に依存するため、項目間の平行性が満たされない一般の場合には、その非負性を保証できないことを注意しておく。

表 2 - 3 . 1

		平方和	自由度	E(平方和÷自由度)
$R^2_{T_{ga}=\nu_a} - R^2_{T_{ga}=\nu_g+\nu_a}$	Between Tests	$AR\sum_g y^2_{g\bullet\bullet} - GARy^2_{\bullet\bullet\bullet}$	$G-1$	$AR\sigma^2_g + R\sigma^2_{ga} + \sigma^2_e$
$R^2_{T_{ga}=\nu_g} - R^2_{T_{ga}=\nu_g+\nu_a}$	Between Persons	$GR\sum_a y^2_{\bullet a\bullet} - GARy^2_{\bullet\bullet\bullet}$	$A-1$	$GR\sigma^2_a + R\sigma^2_{ga} + \sigma^2_e$
$R^2_{T_{ga}=\nu_g+\nu_a} - R^2_{T_{ga}}$	Interaction	Between Test×Person Cells－Between Tests－Between Persons	$(G-1)(A-1)$	$R\sigma^2_{ga} + \sigma^2_e$
$R^2_{T_{ga}=\mu} - R^2_{T_{ga}}$	Between Test× Person Cells	$R\sum_g\sum_a y^2_{ga\bullet} - GARy^2_{\bullet\bullet\bullet}$	$GA-1$	$\sigma^2_g + \sigma^2_a + \sigma^2_{ga} + \sigma^2_e$
$R^2_{T_{ga}}$	Residual	Total－Between Test×Person Cells	$GA(R-1)$	σ^2_e
$R^2_{T_{ga}=\mu}$	Total	$\sum_g\sum_a\sum_r y^2_{gar} - GARy^2_{\bullet\bullet\bullet}$	$GAR-1$	

表2-3.2

		平方和	自由度	E(平方和÷自由度)
$R^2_{T_{ga}=\mu} - R^2_{T_{ga}}$	Between Persons	$R\sum_a y^2_{ga\bullet} - ARy^2_{g\bullet\bullet}$	$A-1$	$R\sigma^2_a + \sigma^2_e$
$R^2_{T_{ga}}$	Within Cells	Total－Between Persons	$A(R-1)$	σ^2_e
$R^2_{T_{ga}=\mu}$	Total	$\sum_a \sum_r y^2_{gar} - ARy^2_{g\bullet\bullet}$	$AR-1$	

3 標本の持つ信頼性を考慮した仮説検定

経営特性の測定では、経営特性の状態と実数との対応関係が回答者や質問項目に依存するため、同じ経営特性の尺度が同じ数値となる保証が得られない。尺度の比較を可能にするためには、観測値を共通の対応関係による尺度に変換しなければならないが、その変換が不可能であることは言うまでもない。この場合、経営特性の分析（仮説検定）をどのように考えたらよいだろうか。

こうした状況における一般的な検定方法を3-1節で解説した後で、分散分析への適用例が3-2節で説明される。

3-1 汎検定問題の定義

3-1-1 問題の設定と分類

通常の検定問題では観測された標本は既知として扱われる。それでは、その標本を完全に知ることができず；

- 真の標本が存在する定義域
- 定義域内の元が真の標本である確率

という情報のみが与えられた場合、その仮説検定をどのように行えばよいだろうか。この検定問題を定式化することから議論を始めよう。

前提となる検定問題を確率モデル$\mathbf{y} \sim \mathrm{F}(\mathbf{y}|\boldsymbol{\theta}), \mathbf{y} \in \mathbf{R}^n$の母数ベクトル$\boldsymbol{\theta}$に関する検定問題$\mathrm{H}_0$ *vs.* H_1とし、この検定問題を原問題と呼ぶことにする。そして、真の標本が存在する定義域の元$\mathbf{w}$を標本$\mathbf{y}$の候補値と呼ぶことにする。次に、任意の候補値$\mathbf{w}$を固定した候補値$\boldsymbol{\omega}$の平行移動；

$$\mathbf{w} = \boldsymbol{\omega} + \boldsymbol{\pi}(\mathbf{w})$$

により表し、この固定した候補値$\boldsymbol{\omega}$を定義域の代表値と呼ぶことにしよう。そこで、候補値$\mathbf{w}$の定義域を$\mathrm{D}(\mathbf{w}|\boldsymbol{\omega})$とし、その候補値が真の標本$\mathbf{y}$である確率の分布関数を$\mathrm{G}(\mathbf{w}|\boldsymbol{\omega})$と代表値$\boldsymbol{\omega}$を明示して表すことにする。このとき；

$$\mathbf{y} \sim \mathrm{F}(\mathbf{y}|\boldsymbol{\theta}), \mathbf{y} \in \mathrm{D}(\mathbf{w}|\boldsymbol{\omega}), \mathbf{w} \sim \mathrm{G}(\mathbf{w}|\boldsymbol{\omega}), \mathrm{H}_0 \ \textit{vs.}\ \mathrm{H}_1$$

と定義される検定問題の形式を形式Ⅰと呼ぶ。ただし、形式Ⅰは次の条件を満たしている

ことが前提である。原問題の標本 $\mathbf{y}$ の定義域[1]を $\mathbf{S}_{\mathrm{y}} \subseteq \mathbf{R}^n$ とすれば；

$$\mathrm{D}(\mathbf{w}|\omega) \subseteq \mathbf{S}_{\mathrm{y}}$$

が成立していることである。未知の標本 $\mathbf{y}$ の候補値が標本 $\mathbf{y}$ の定義域に含まれないことは候補値の意味に矛盾するためである。後で必要となるため、ベクトル $\boldsymbol{\pi}(\mathbf{w})$ の定義域を；

$$\mathbf{S}_{\pi|\omega} = \left\{\boldsymbol{\pi}(\mathbf{w}) \middle| \boldsymbol{\pi}(\mathbf{w}) = \mathbf{w} - \boldsymbol{\omega}, \mathbf{w} \in \mathrm{D}(\mathbf{w}|\omega)\right\}$$

と定義しておく。

次に、形式Ⅰとは表裏一体の関係にある形式Ⅱを定義しよう。形式Ⅱとは、原問題の確率変数 $\mathbf{y}$ と、この確率変数とは独立に分布する確率変数 $\boldsymbol{\pi}^{(-)} \sim \mathrm{H}(\boldsymbol{\pi}^{(-)})$ との和 $\boldsymbol{\omega}$ が観測値となる場合の検定問題である。すなわち；

$$\boldsymbol{\omega} = \mathbf{y} + \boldsymbol{\pi}^{(-)}, \mathbf{y} \sim \mathrm{F}(\mathbf{y}|\theta), \boldsymbol{\pi}^{(-)} \sim \mathrm{H}(\boldsymbol{\pi}^{(-)}), \mathrm{H}_0 \ \textit{vs.}\ \mathrm{H}_1$$

と定義される検定問題が形式Ⅱである。ただし、形式Ⅱを適用する場合、この観測値を未知の標本 $\mathbf{y}$ と誤差項 $\boldsymbol{\pi}^{(-)}$ の和として考えるのではなく、理論的に想定した標本 $\mathbf{y}$ を表す便宜上の値と解釈していることが前提である。この解釈の違いは形式Ⅱを形式Ⅰへ変換することで明確なものとなる。

そこで、確率変数 $\boldsymbol{\pi}^{(-)}$ の定義域を $\mathbf{S}_{\pi^{(-)}} \subseteq \mathbf{R}^n$ とし、確率変数 $\mathbf{y} \in \mathbf{S}_{\mathrm{y}}$ との和が標本 $\boldsymbol{\omega}$ となる $\boldsymbol{\pi}^{(-)} \in \mathbf{S}_{\pi^{(-)}}$ の集合を；

$$\mathbf{S}_{\pi^{(-)}|\omega} = \left\{\boldsymbol{\pi}^{(-)} \middle| \mathbf{y} + \boldsymbol{\pi}^{(-)} = \boldsymbol{\omega}, \mathbf{y} \in \mathbf{S}_{\mathrm{y}}, \boldsymbol{\pi}^{(-)} \in \mathbf{S}_{\pi^{(-)}}\right\}$$

と定義する。このとき、標本 $\boldsymbol{\omega}$ と確率変数 $\boldsymbol{\pi}^{(-)} \in \mathbf{S}_{\pi^{(-)}|\omega}$ との差 $\boldsymbol{\omega} - \boldsymbol{\pi}^{(-)}$ は未知の標本 $\mathbf{y}$ の候補値 $\mathbf{w}$ と考えることができる。従って、標本 $\boldsymbol{\omega}$ を代表値としたときの、候補値 $\mathbf{w}$ が従う分布関数 G は；

---

[1] 標本空間(Sample Space)、または、分布関数 $\mathrm{F}(\mathbf{y}|\theta)$ の台(Support)のことである。

$$G(\mathbf{w}|\omega)=\frac{P_H\left(\mathbf{W}=\omega-\Pi^{(-)}<\mathbf{w},\Pi^{(-)}\in\mathbf{S}_{\pi^{(-)}|\omega}\right)}{P_H\left(\mathbf{S}_{\pi^{(-)}|\omega}\right)}=\frac{P_H\left(\Pi^{(-)}>\omega-\mathbf{w},\Pi^{(-)}\in\mathbf{S}_{\pi^{(-)}|\omega}\right)}{P_H\left(\mathbf{S}_{\pi^{(-)}|\omega}\right)}$$

と定義することができる。分布関数 G の台である候補値 $\mathbf{w}$ の定義域は；

$$D(\mathbf{w}|\omega)=\left\{\mathbf{w}\middle|\mathbf{w}=\omega-\pi^{(-)},\pi^{(-)}\in\mathbf{S}_{\pi^{(-)}|\omega}\right\}$$

と表すこともできる。

この結果は、標本ω を未知の標本 $\mathbf{y}$ を表す便宜上の値と考える形式Ⅱの解釈が、この値を代表値とした形式Ⅰの解釈と同じというだけでなく、標本ω を与えたときの形式Ⅱが対応する形式Ⅰと数学的に同じ構造であることを示している。実際、確率変数 $\pi^{(-)}$ と形式Ⅰで導入した確率変数ベクトル$\pi(\mathbf{w})$との間には；

$$\begin{aligned}\mathbf{S}_{\pi^{(-)}|\omega}&=\left\{\pi^{(-)}\middle|\mathbf{y}+\pi^{(-)}=\omega,\mathbf{y}\in\mathbf{S}_{\mathbf{y}},\pi^{(-)}\in\mathbf{S}_{\pi^{(-)}}\right\}\\&=\left\{-\pi(\mathbf{w})\middle|\pi(\mathbf{w})=\mathbf{w}-\omega,\mathbf{w}\in D(\mathbf{w}|\omega)\right\}=\left\{-\pi(\mathbf{w})\middle|\pi(\mathbf{w})\in\mathbf{S}_{\pi|\omega}\right\}\end{aligned}$$

という関係が成立している。しかし、二つの標本 $\omega_1\neq\omega_2$ に対応した確率変数 $\pi^{(-)}$ の定義域 $\mathbf{S}_{\pi^{(-)}|\omega_1},\mathbf{S}_{\pi^{(-)}|\omega_2}$ が等しくなる保証はないため、これらの標本から同じ構造[2]の形式Ⅰが導かれるとは限らない。ただし、原問題の定義域 $\mathbf{S}_{\mathbf{y}}$ が $\mathbf{R}^n$ の場合は、$\mathbf{S}_{\pi^{(-)}|\omega_1}=\mathbf{S}_{\pi^{(-)}|\omega_2}$ が成立するため、任意の標本ω に対応する形式Ⅰはすべて同じ構造となる。このように、形式Ⅱは形式Ⅰを一般化した表現と言うことができる。

最後に、形式Ⅱの標本 $\mathbf{y}+\pi^{(-)}$ に対して構成した形式Ⅰ；

$$\mathbf{y}\sim F(\mathbf{y}|\theta),\pi^{(-)}\sim H(\pi^{(-)}),\mathbf{y}+\pi^{(-)}\in D^*(\omega^*|\omega),\omega^*\sim G^*(\omega^*|\omega),H_0\ \textit{vs.}\ H_1$$

を形式Ⅲとして紹介しておく。ただし、この場合も $D^*(\omega^*|\omega)\subseteq\mathbf{S}_{\mathbf{y}}$ が前提である。

---

[2] 二つの定義域 $D(\mathbf{w}|\omega_1),D(\mathbf{w}|\omega_2)$ が平行移動の関係にあるという意味で用いている。

形式IIの標本$\mathbf{y}+\boldsymbol{\pi}^{(-)}$の候補値として適当な値$\boldsymbol{\omega}^* \in D^*\left(\boldsymbol{\omega}^*\middle|\boldsymbol{\omega}\right)$を取り、この値に前述した方法を適用すれば、原問題の標本$\mathbf{y}$に対する候補値$\mathbf{w}$の事前分布$H\left(\mathbf{w}\middle|\boldsymbol{\omega}^*\right)$を求めることができる。この便宜上の代表値$\boldsymbol{\omega}^*$には事前分布$G^*$が与えられているから、形式IIIの代表値$\boldsymbol{\omega}$を代表値として持つ形式Iの事前分布Gを；

$$G\left(\mathbf{w}\middle|\boldsymbol{\omega}\right)=\int_{D^*\left(\boldsymbol{\omega}\middle|\boldsymbol{\omega}^*\right)} H\left(\mathbf{w}\middle|\boldsymbol{\omega}^*\right) dG^*\left(\boldsymbol{\omega}\middle|\boldsymbol{\omega}^*\right)$$

と定義することができる。このように、形式IIIも形式Iに変換できることが分かる。

3-1-2 形式IIと通常の検定方法

*3-1-1*では、形式Iの派生形として形式IIと形式IIIを紹介した。この章の議論は原則として標本$\mathbf{y}$が未知という形式Iを前提とするが、ここでは例外的に標本が二つの確率変数の和として定義される形式II；

$$\boldsymbol{\omega}=\mathbf{y}+\boldsymbol{\pi}^{(-)}, \mathbf{y}\sim F\left(\mathbf{y}\middle|\theta\right), \boldsymbol{\pi}^{(-)}\sim H\left(\boldsymbol{\pi}^{(-)}\right), H_0 \ \ vs.\ \ H_1$$

を想定して議論を進めていく。さて、確率変数$\boldsymbol{\omega}$の分布関数を求めることで、形式II（形式Iと形式IIIも）は通常の検定問題として扱うことができる。この一見当然とも思える検定方法を採用しないのは、確率変数$\boldsymbol{\pi}^{(-)}$の持つ意味を考えた場合、その検定方法が必ずしも合理的とは言い切れないからである。所与の標本$\boldsymbol{\omega}$に対して確率変数$\boldsymbol{\pi}^{(-)} \in \mathbf{S}_{\boldsymbol{\pi}^{(-)}|\boldsymbol{\omega}}$が持つ確率とは、候補値$\mathbf{w}=\boldsymbol{\omega}-\boldsymbol{\pi}^{(-)} \in \mathbf{S}_{\mathbf{y}}$が真の標本$\mathbf{y}$である可能性を意味している。この確率の意味に注意して、次の数値例を考えてみよう。

母数θと誤差項εとの和$y=\theta+\varepsilon$で定義された確率モデルにおける検定問題；

$$H_0:\theta=0 \ \ vs.\ \ H_1:\theta=1$$

を原問題とする。誤差項εの分布は；

ε	-3	-2.5	-2	-1.5	-1	-0.5
確率	0.01	0.03	0.04	0.07	0.10	0.15
0	0.5	1	1.5	2	2.5	3
0.20	0.15	0.10	0.07	0.04	0.03	0.01

とする。確率変数$\pi^{(-)}$には；

$$\Pi^{(-)} = (-0.5 \quad 0 \quad 0.5)$$

という三通りの値を想定する。そして、$\omega = \theta + \varepsilon + \pi^{(-)} = 1.5$を観測値としよう。

確率変数$\pi^{(-)}$の持つ確率はまだ定義していないが、真の標本yは候補値；

$$w = 1, w = 1.5, w = 2$$

のいずれかである。まず、これらの候補値を真の標本と考えて、原問題での検定結果を調べてみる。この結果は各仮説の下で標本yが生じる確率と尤度比$P(y|\theta=1)/P(y|\theta=0)$を載せた次の表；

y	$P(y\|\theta=0)$	$P(y\|\theta=1)$	尤度比
−3.0	0.01	0.00	0.00000
−2.5	0.03	0.00	0.00000
−2.0	0.04	0.01	0.25000
−1.5	0.07	0.03	0.42857
−1.0	0.10	0.04	0.40000
−0.5	0.15	0.07	0.46667
0.0	0.20	0.10	0.50000
0.5	0.15	0.15	1.00000
1.0	0.10	0.20	2.00000
1.5	0.07	0.15	2.14286
2.0	0.04	0.10	2.50000
2.5	0.03	0.07	2.33333
3.0	0.01	0.04	4.00000
3.5	0.00	0.03	∞
4.0	0.00	0.01	∞

から求めることができる。Neyman-Pearsonの補題から、各候補値が真の標本であるときに帰無仮説を棄却できる尤度比の最小値と、そのときの有意水準は次のようになる。

候補値w	確率変数$\pi^{(-)}$	尤度比の最小値	有意水準
1	0.5	2.00000	0.25
1.5	0	2.14286	0.15
2	−0.5	2.50000	0.08

，

次に、確率変数$\pi^{(-)}$に定義された確率が；

$$P\left(\Pi^{(-)}\right) = P(-0.5 \quad 0 \quad 0.5) = \begin{cases} \mathbf{P}_1 = (1/3 \quad 1/3 \quad 1/3) \\ \mathbf{P}_2 = (0.495 \quad 0.01 \quad 0.495) \\ \mathbf{P}_3 = (0.01 \quad 0.98 \quad 0.01) \end{cases}$$

という三通りの場合で、形式IIの最強力検定を求めてみる。各確率において標本ωが生じる確率と尤度比は表3-1.1の通りである。この表から、標本$\omega = 1.5$で帰無仮説を棄却することができる有意水準が、上記の確率の順で；

$$\mathbf{P}_1 \to 0.16, \mathbf{P}_2 \to 0.16485, \mathbf{P}_3 \to 0.1503$$

であることが分かる。有意水準はいずれの場合も約16%だから、これは通常の検定方法が確率$\mathbf{P}_1 \sim \mathbf{P}_3$の違いを区別していないことを意味している。確率変数$\pi^{(-)}$を誤差項として考えれば、この結果は標本ωが生じる確率（そのように計算されたのだから）に基づいた検定結果として何の問題もない。

しかし、確率$\mathbf{P}_1 \sim \mathbf{P}_3$を候補値$\omega - \pi^{(-)}$の持つ信頼性と捉えた場合、この検定結果は果たして妥当なものだろうか。そこで、これらの確率が表す状況を確認しておこう。確率$\mathbf{P}_1$は三つの候補値$w = 1, w = 1.5, w = 2$のどれが真の標本であるのか全く分からないという状況を意味している。確率$\mathbf{P}_2$では等確率で$w = 1$か$w = 2$というところまで候補値を絞ることができる。確率$\mathbf{P}_3$の場合は実質的に$w = 1.5$を真の標本と考えることができる。この場合は形式IIを原問題と区別する必要はほとんどないから、有意水準16%で帰無仮説を棄却できるという決定は妥当なものであろう。しかし、伝統的な検定論が帰無仮説を誤って棄却してしまう第一種の過誤に重点を置いていることを考えると、真の標本が未知という状況では、「棄却」という決定には一層の慎重さが必要ではないだろうか。この慎重さは真の標本と紛れたデータ（候補値）との関係が希薄であるほど要請されるはずである。この考え方を尊重するのであれば、確率$\mathbf{P}_1, \mathbf{P}_2$の場合に帰無仮説の「棄却」を決定すべきか否かが微妙な問題であることが分かる。

このように、候補値を標本としたときに帰無仮説を棄却できる有意水準の、その候補値が真の標本である信頼性に関する平均を検定の有意水準（この数値例では確率$\mathbf{P}_1 \sim \mathbf{P}_3$のいずれの場合も約16%である）とする考え方では、事前分布Hの違いを反映させることができない。従って、形式IIを未知の標本$\mathbf{y}$に想定される候補値$\mathbf{w}$を一つの式で表現した確率モデルと考えるのならば、標本$\boldsymbol{\omega}$の確率を基に統計的決定を行うのではなく、各候補値の検定結果と信頼性に基づいて統計的決定を行うべきではないだろうか。

3-1-3 汎検定問題と受容リスク

ここからの議論は再び形式Ｉ；

$$\mathbf{y} \sim \mathrm{F}(\mathbf{y}|\theta), \mathbf{y} \in \mathrm{D}(\mathbf{w}|\omega), \mathbf{w} \sim \mathrm{G}(\mathbf{w}|\omega), \mathrm{H}_0 \ \ vs. \ \ \mathrm{H}_1$$

を前提とする。ただし、表記上の理由から確率化検定の議論を避けるため、ここでは原問題の分布関数 F を絶対連続型と仮定する。形式Ｉ（形式Ⅱと形式Ⅲも）では各候補値の検定結果と信頼性に基づいて仮説検定が行われるので、原問題の検定方法を事前に決めておかなければならない。そこで、原問題における有意水準αの検定関数(Test Function)を一次元の統計量 $T(\mathbf{y})$を基に；

$$\delta_\alpha(\mathbf{y}) = \begin{cases} 1; T(\mathbf{y}) > c_\alpha \\ 0; T(\mathbf{y}) < c_\alpha \end{cases}$$

と表すことにする。この検定に不偏性や不変性が課せられていても構わないが、その制限された検定方法の中では一様最強力であることが前提である。

各候補値の検定結果を基に帰無仮説の棄却と受容を決定する場合、候補値に付与された確率と検定結果はすべて既知だから、その決定が候補値 $\mathbf{w}$ に対してではなく、候補値の定義域$\mathrm{D}(\mathbf{w}|\omega)$に対して定められることが分かる。この決定を定義域の代表値ω を明示した決定関数(Decision Function)；

$$d(\omega) = \begin{cases} 1; \mathrm{H}_0 \text{ を棄却} \\ 0; \mathrm{H}_0 \text{ を受容} \end{cases}$$

で表現する。この決定問題の母数は真の標本 $\mathbf{y}$ における検定結果だから、ここでは損失関数(Loss Function)として；

$$\mathrm{L}(\delta_\alpha(\mathbf{y}), d(\omega)) = \begin{cases} \mathrm{L}(1,1) = 0 \\ \mathrm{L}(1,0) = 1 \\ \mathrm{L}(0,0) = 0 \\ \mathrm{L}(0,1) = L_{01} \end{cases}$$

を採用しよう。この損失関数は検定論における 0-1 損失の流用である。従って、標本 $\mathbf{y}$ を棄却できる有意水準に基づいた損失関数を考えても構わない。また、この決定問題では決

定関数 $d(\omega)$が確率変数として扱われないので[3]、損失関数とリスク関数(Risk Function)が一致することを注意しておく。

さて、母数に当たる検定結果に対して（事前）確率；

$$\mathrm{P}\left(\delta_\alpha(\mathbf{y})=1\middle|\omega\right)=\int_{D(\mathbf{w}|\omega)}\delta_\alpha(\mathbf{w})d\,\mathrm{G}\left(\mathbf{w}\middle|\omega\right)=\mathrm{P}_{\mathrm{G}(\mathbf{w}|\omega)}\left(T(\mathbf{w})>c_\alpha\right)$$
$$\mathrm{P}\left(\delta_\alpha(\mathbf{y})=0\middle|\omega\right)=1-\mathrm{P}\left(\delta_\alpha(\mathbf{y})=1\middle|\omega\right)$$

が定義できるので、ここでは決定関数の選択にベイズ基準を採用しよう。このときのベイズリスク(Bayes Risk)は；

$$\mathrm{BR}\left(d(\omega)=1\right)=L_{01}\left(1-\mathrm{P}\left(\delta_\alpha(\mathbf{y})=1\middle|\omega\right)\right)$$
$$\mathrm{BR}\left(d(\omega)=0\right)=\mathrm{P}\left(\delta_\alpha(\mathbf{y})=1\middle|\omega\right)$$

となる。従って、帰無仮説を棄却できるのは；

$$\mathrm{BR}\left(d(\omega)=1\right)<\mathrm{BR}\left(d(\omega)=0\right)\Leftrightarrow\mathrm{P}\left(\delta_\alpha(\mathbf{y})=1\middle|\omega\right)>\frac{L_{01}}{L_{01}+1}$$

が成立する場合であり、不等号が逆の場合は受容ということになる。ところで、検定統計量としての事前確率；

$$\mathrm{R}_\alpha(\omega)=\mathrm{P}\left(\delta_\alpha(\mathbf{y})=1\middle|\omega\right)=\int_{D(\mathbf{w}|\omega)}\delta_\alpha(\mathbf{w})d\,\mathrm{G}\left(\mathbf{w}\middle|\omega\right)$$

は受容という決定のベイズリスクである。そこで、この事前確率 $\mathrm{R}_\alpha(\omega)$ を受容リスクと呼ぶことにする。

以上の議論をまとめよう。形式 I に対応した検定関数は受容リスクの関数として；

$$\bar{\delta}(\omega)=\begin{cases}1;\mathrm{R}_\alpha(\omega)>r\\0;\mathrm{R}_\alpha(\omega)<r\end{cases}$$

---

[3] 所与の定義域に与えられる決定 $d(\omega)$が確定的という意味である。この検定の有意水準を求める場合は、定義域自体に確率が定義されるため、決定 $d(\omega)$は確率変数となる。この話題は *3-1-4* で議論される。

と定義することができる。ただし、定数 r は誤って棄却を決定したときの損失 L_{01} の相対的な大きさ $L_{01}/(L_{01}+1)$ のことである。この検定関数を原問題の検定関数 $\delta_\alpha(\mathbf{y})$ と区別するため、原問題の検定結果に関する検定という意味で汎検定関数と呼ぶことにする。この検定問題（形式Ⅰ～形式Ⅲ）も同様に汎検定問題と呼ぶことにしよう。

ところで、受容リスクとは真の標本 $\mathbf{y}$ における検定結果が帰無仮説の棄却である事前確率のことであるが、これを「事前」確率というのはベイズ基準を採用したためで、実際には各候補値について検定した後に求めた「事後」確率を表している。一方、汎検定関数の定義から、受容を誤って棄却と決定したときの損失 L_{01} が相対的に大きいほど、棄却という決定には大きな受容リスクが必要であることが分かる。これは、損失 L_{01} が大きいほど大きな「事後」確率が要求されるということだから、汎検定関数とは非常に合理的な決定と言うことができる。

この方法を *3-1-2* の数値例に適用してみよう。原問題の有意水準を 15%とすると、三通りの確率について受容リスクは；

$$\mathbf{P}_1 \to \mathrm{R}_{0.15}(\omega) = 2/3, \mathbf{P}_2 \to \mathrm{R}_{0.15}(\omega) = 0.505, \mathbf{P}_3 \to \mathrm{R}_{0.15}(\omega) = 0.99$$

という値となる。確率 $\mathbf{P}_3$ の場合に棄却という決定は妥当だが、確率 $\mathbf{P}_1,\mathbf{P}_2$ の場合は損失の値 L_{01} 如何ということになるだろう。

3-1-4 汎検定関数の有意水準

汎検定関数の定義から明らかなように、この検定方法では予め設定した受容リスクが有意水準より優先される。これは汎検定関数の持つ実用上の利点と言うことができる。原問題の確率モデル F と候補値の事前分布 G の組み合わせに対し、受容リスクを求めることは比較的容易だが、対応する有意水準の算出は非常に面倒だからである。しかし、形式Ⅱの標本 ω に適用した検定関数と汎検定関数との比較など、有意水準を求めることは理論的に興味のあるところである。汎検定関数の有意水準とは受容リスクに定義された確率のことだから、ここでは受容リスクの持つ確率について説明していこう。

その定義から分かるように、受容リスク；

$$\mathrm{R}_\alpha(\omega) = \mathrm{P}(\delta_\alpha(\mathbf{y}) = 1|\omega) = \int_{\mathrm{D}(\mathbf{w}|\omega)} \delta_\alpha(\mathbf{w}) d\,\mathrm{G}(\mathbf{w}|\omega)$$

とは候補値 $\mathbf{w}$ の定義域 $\mathrm{D}(\mathbf{w}|\omega)$ の集合関数となっている。従って、受容リスクに確率を定義するには、候補値の定義域に確率を定めなければならないことが分かる。

まず、形式Ⅰにおいて代表値 ω_0 を固定しよう。これは定義域 $\mathrm{D}(\mathbf{w}|\omega_0)$ が与えられたことを意味している。この定義域の比較対象となるのは、ベクトル $\boldsymbol{\pi}(\mathbf{w})$ の定義域；

$$\mathbf{S}_{\pi|\omega_0} = \left\{ \pi(\mathbf{w}) \middle| \pi(\mathbf{w}) = \mathbf{w} - \omega_0, \mathbf{w} \in \mathrm{D}(\mathbf{w}|\omega_0) \right\}$$

が共通となる定義域 $\mathrm{D}(\mathbf{w}|\omega) = \left\{ \mathbf{w} \middle| \mathbf{w} = \omega + \pi, \pi \in \mathbf{S}_{\pi|\omega_0} \right\} \subseteq \mathbf{S}_{\mathbf{y}}$ に限られる。そこで、この条件を満たす代表値 ω の集合を；

$$\mathbf{S}_{\omega|\omega_0} = \left\{ \omega \middle| \left\{ \mathbf{w} \middle| \mathbf{w} = \omega + \pi, \pi \in \mathbf{S}_{\pi|\omega_0} \right\} \subseteq \mathbf{S}_{\mathbf{y}} \right\}$$

とおくことで、対象とする定義域の全体を集合族 $\mathbf{S}(\mathrm{D}(\mathbf{w}|\omega_0)) = \left\{ \mathrm{D}(\mathbf{w}|\omega) \middle| \omega \in \mathbf{S}_{\omega|\omega_0} \right\}$ で表すことにしよう。この集合族が汎検定問題における標本空間となる。

やや面倒な議論が続いたが、これは標本 $\mathbf{y}$ の定義域 $\mathbf{S}_{\mathbf{y}}$ が有界集合の場合を考慮して正確に記述したためである。この議論は *3-1-2* の数値例を考えると分かり易い。このときの標本 $\omega = 1.5$ を代表値とした定義域は $\mathrm{D}(w|1.5) = \{1,1.5,2\}$ だから、この定義域が起こる確率を考える場合は左から順に確率 0.495,0.01,0.495（この説明では確率 $\mathbf{P}_2$ を用いている）が付与される定義域を対象とすべきである。例えば、代表値が $\omega = 3$ の場合、標本 y の定義域に含まれない値 $w = 3 - (-0.5) = 3.5$ が真の標本である確率は 0.495 という命題は偽だから、この値を候補値と考えることはできない。従って、この場合は $w = 3.5$ を外した集合 $\{2.5,3\}$ が定義域 $\mathrm{D}(w|3)$ となる。三つの候補値を持つ定義域が得られたときに、候補値が二つの定義域を場合分けの対象とする必要のないことは明らかである。

さて、この数値例で帰無仮説 $\mathrm{H}_0 : \theta = 0$ が真のときに、定義域 $\mathrm{D}(w|1.5) = \{1,1.5,2\}$ が生じる確率は；

$$\begin{aligned}
\mathrm{P}(\mathrm{D}(w|1.5)|\theta = 0) &\propto \mathrm{P}(w = 1 \ or \ w = 1.5 \ or \ w = 2|\theta = 0) \\
&= \mathrm{P}(\pi^{(-)} = 0.5)\mathrm{P}(w = 1|\theta = 0) + \cdots + \mathrm{P}(\pi^{(-)} = -0.5)\mathrm{P}(w = 2|\theta = 0) \\
&= 0.495 \times 0.1 + 0.01 \times 0.07 + 0.495 \times 0.04 \\
&= 0.07
\end{aligned}$$

に比例した値と考えればよい。第一式で等号が成立しないのは、汎検定問題の標本空間が制限（この例では $\{2.5,3,3.5\}$ は含まれない）されているからである。比例定数は；

$$\mathrm{P}\left(w=-3\ or\ w=-2.5\ or\ w=-2\middle|\theta=0\right)+\cdots+\mathrm{P}\left(w=2\ or\ w=2.5\ or\ w=3\middle|\theta=0\right)=0.9602$$

の逆数だから、問題の確率は $\mathrm{P}\left(\mathrm{D}\left(w|1.5\right)\middle|\theta=0\right)=0.07\div0.9602\fallingdotseq0.073$ となる。

それでは、この確率を一般的に定義する。*3-1-3*で原問題の分布関数 F を絶対連続型と仮定しているので、その密度関数を $\mathrm{f}(\mathbf{w}|\theta)$ と表そう。このとき、定義域 $\mathrm{D}(\mathbf{w}|\omega)$ に付与される確率 $\overline{\mathrm{P}}$ は；

$$\overline{\mathrm{P}}\left(\mathrm{D}\left(\mathbf{w}|\omega\right)\right)=C\int_{\mathrm{D}(\mathbf{w}|\omega)}\mathrm{f}\left(\mathbf{w}|\theta\right)d\,\mathrm{G}\left(\mathbf{w}|\omega\right)$$

と定義することができる。ここで、$\mathrm{D}(\mathbf{w}|\omega)\in\mathbf{S}\left(\mathrm{D}\left(\mathbf{w}|\omega_0\right)\right)\Leftrightarrow\omega\in\mathbf{S}_{\omega|\omega_0}$ だから；

$$\frac{1}{C}=\int_{\mathbf{S}_{\omega|\omega_0}}\int_{\mathrm{D}(\mathbf{w}|\omega)}\mathrm{f}\left(\mathbf{w}|\theta\right)d\,\mathrm{G}\left(\mathbf{w}|\omega\right)d\mu(\omega)$$

である。従って、受容リスク r が起こる確率は確率 $\overline{\mathrm{P}}$ を用いることで；

$$\overline{\mathrm{P}}\left(\mathrm{R}_\alpha(\omega)=r\right)=\overline{\mathrm{P}}\left(\mathrm{D}\left(\mathbf{w}|\omega\right)\middle|\mathrm{R}_\alpha(\omega)=r,\omega\in\mathbf{S}_{\omega|\omega_0}\right)$$

と表すことができる。この確率により汎検定関数における有意水準が定義される。

これが候補値の定義域 $\mathrm{D}(\mathbf{w}|\omega)$ に与えられる確率 $\overline{\mathrm{P}}$ の定義であるが、標本空間が集合族である場合の議論は数学的にやや面倒である。そこで、定義域 $\mathrm{D}(\mathbf{w}|\omega)$ と代表値 ω の対応関係は固定されているから[4]、確率 $\overline{\mathrm{P}}$ を代表値 ω に定義された確率として捉え直そう。この対応は一対一であるから、任意の代表値 $\omega\in\mathbf{S}_{\omega|\omega_0}$ に対して $\overline{\mathrm{P}}(\omega)=\overline{\mathrm{P}}\left(\mathrm{D}\left(\mathbf{w}|\omega\right)\right)$ と確率 $\overline{\mathrm{P}}$ を定義すればよい。この代表値は実数ベクトルだから、分布関数；

---

[4] 観測値としての定義域における代表値 ω_0 の選び方は任意であるが、この定義域を基に標本空間 $\mathbf{S}\left(\mathrm{D}\left(\mathbf{w}|\omega_0\right)\right)$ を定義するときは、その対応関係は固定されている。

$$\overline{F_G}(\omega) = \overline{P}(\Omega|\Omega < \omega) = \overline{P}\left(D(\mathbf{w}|\Omega)\middle|\Omega < \omega\right)$$

を定義することが可能となる。そこで、次節以降は代表値$\boldsymbol{\omega}$ に定義された確率$\overline{P}(\omega)$ と分布関数$\overline{F_G}(\omega)$を基に議論していくことを約束する。

ところで、形式Ⅰでは代表値$\boldsymbol{\omega} \in \mathbf{S}_{\omega|\omega_0}$ に対して確率変数ベクトル$\boldsymbol{\pi}(\mathbf{w}) \in \mathbf{S}_{\pi|\omega_0}$ が矛盾なく定義されているため、これを形式Ⅱに変形することは容易である。いま、代表値ω に対して確率変数$\boldsymbol{\pi}^{(-)}$を；

$$\pi^{(-)} = -\pi(\mathbf{w}), \omega = \mathbf{w} + \pi^{(-)}, \mathbf{w} \in D(\mathbf{w}|\omega), \pi^{(-)} \in \mathbf{S}_{\pi^{(-)}|\omega_0} = \left\{-\pi\middle|\pi \in \mathbf{S}_{\pi|\omega_0}\right\}$$

と定義すると、確率変数$\boldsymbol{\pi}^{(-)}$の分布関数 H と事前分布 G の間には；

$$\begin{aligned} G(\mathbf{w}|\omega) &= P_G(\mathbf{W} = \omega + \Pi < \mathbf{w}) \\ &= P_G(-\Pi > \omega - \mathbf{w}) \\ &= 1 - P_G\left(\Pi^{(-)} < \omega - \mathbf{w}\right) \\ &= 1 - H(\omega - \mathbf{w}) \end{aligned}$$

という関係が成立する。この分布関数$H\left(\pi^{(-)}\right) = H(\omega - \mathbf{w})$が変数 $\mathbf{w}$ の減少関数であることに注意すると、$\overline{P}(D(\mathbf{w}|\omega))$の積分部分は確率変数$\boldsymbol{\pi}^{(-)}$についての積分；

$$\begin{aligned} \int_{D(\mathbf{w}|\omega)} f(\mathbf{w}|\theta)\, d\, G(\mathbf{w}|\omega) &= \int_{D(\mathbf{w}|\omega)} f(\mathbf{w}|\theta)\, d(1 - H(\omega - \mathbf{w})) \\ &= \int_{D(\mathbf{w}|\omega)} f(\mathbf{w}|\theta)\, d(-H(\omega - \mathbf{w})) \\ &= \int_{S_{\pi^{(-)}|\omega_0}} f\left(\omega - \pi^{(-)}\middle|\theta\right) d\, H\left(\pi^{(-)}\right) \end{aligned}$$

として書き換えられることが分かる。この式で、最後の積分は確率変数 $\mathbf{w}$ と$\boldsymbol{\pi}^{(-)}$の畳み込みだから、この値は形式Ⅱの標本$\boldsymbol{\omega} = \mathbf{w} + \boldsymbol{\pi}^{(-)}$が持つ確率を意味している。そして、これが形式Ⅰで定義した確率$\overline{P}(\omega) = \overline{P}(D(\mathbf{w}|\omega))$の別解釈である。

3-2 具体例：原問題が分散分析の場合

3-2-1 分散分析の要約

この具体例では原問題として第１章で詳しく説明した分散分析を取り上げる。まず、その内容を要約しておこう。検定問題は回帰モデル；

$$\mathbf{y} \sim \mathrm{N}(\mathbf{X\beta}, \sigma^2 \mathbf{I}), \mathbf{y} \in \mathbf{R}^n, \mathbf{X}: n \times m, \mathrm{rank}\, \mathbf{X} = r$$

における線形仮説 $\mathrm{H}_0: \mathbf{H\beta} = \boldsymbol{\xi}$ *vs.* $\mathrm{H}_1: \mathbf{H\beta} \neq \boldsymbol{\xi}$ の検定である。ただし、ここでは未知母数としての分散 σ^2 を期待値が $\overline{\sigma}^2$ の確率変数として扱う。

さて、$\mathbf{R}^n$ から $\mathrm{V}(\mathbf{X})$ への直交射影行列を $\mathbf{P}$、$\mathbf{R}^n$ から $\{\mathbf{X\gamma} | \mathbf{H\gamma} = \mathbf{0}\} \subseteq \mathrm{V}(\mathbf{X})$ への直交射影行列を $\mathbf{U}$ としよう。連立方程式 $\mathbf{H\beta} = \boldsymbol{\xi}$ の特殊解を $\boldsymbol{\beta}_0$ とすれば；

$$\begin{aligned} R_0^2 &= \mathbf{y}'(\mathbf{I} - \mathbf{U})\mathbf{y} \sim \sigma^2 \chi^2(n-r) \\ R_1^2 - R_0^2 &= (\mathbf{y} - \mathbf{X\beta}_0)'(\mathbf{P} - \mathbf{U})(\mathbf{y} - \mathbf{X\beta}_0) \sim \sigma^2 \chi^2(k, \lambda) \\ \sigma^2 \lambda &= (\mathbf{X\beta} - \mathbf{X\beta}_0)'(\mathbf{P} - \mathbf{U})(\mathbf{X\beta} - \mathbf{X\beta}_0) \end{aligned}$$

が成立し、二つの統計量は独立に分布する。従って；

$$F(\mathbf{y}) = \frac{R_1^2 - R_0^2}{k} \div \frac{R_0^2}{n-r} \sim \mathrm{F}(k, n-r, \lambda)$$

であることが分かる。これは 1 - 2 節で述べた通りである。

回帰モデル $\mathbf{y} \sim \mathrm{N}(\mathbf{X\beta}, \sigma^2 \mathbf{I})$ を $F(\mathbf{y}) \sim \mathrm{F}(k, n-r, \lambda)$ に変換することで、検定問題を F 分布の非心度 λ に関する検定問題 $\mathrm{H}_0: \lambda = 0$ *vs.* $\mathrm{H}_1: \lambda > 0$ に帰着させることができる。この変換後の検定問題における一様最強力検定が分散分析である。そこで、有意水準 α の分散分析を検定関数；

$$\delta_\alpha(\mathbf{y}) = \begin{cases} 1; F(\mathbf{y}) > c_\alpha \\ 0; F(\mathbf{y}) < c_\alpha \end{cases}$$

で表すことにする。

また、F 統計量は1-3節で述べた変換群 $G_1 \sim G_3$ に関する最大不変量だから、これらの変換群に不変な一様最強力不変検定が分散分析ということになる。特に、確率変数 $\mathbf{y}$ の定数倍という変換群 G_3 は未知母数である分散 σ^2 に依存しない確率モデル（F 分布は正規分布の分散に依存しない）を導くために要請される。この値が既知の場合は変換群 G_3 に関する不変性は不要となるから、検定問題は確率モデル $R_1^2 - R_0^2 \sim \sigma^2 \chi^2(k, \lambda)$ の非心度 λ に関する仮説 $\mathrm{H}_0 : \lambda = 0$ *vs.* $\mathrm{H}_1 : \lambda > 0$ の検定へと変換される。

3-2-2 分散分析を原問題とした汎検定問題

この原問題に対し、事前分布 $\mathrm{G}(\mathbf{w}|\boldsymbol{\omega})$ が分散既知の正規分布 $\mathrm{N}(\boldsymbol{\omega}, \tau^2 \mathbf{I})$ である場合を汎検定問題として考えていこう。この場合は；

$$\mathbf{S}_{\mathbf{y}} = \mathrm{D}(\mathbf{w}|\boldsymbol{\omega}) = \mathbf{S}_{\boldsymbol{\omega}|\boldsymbol{\omega}_0} = \mathbf{R}^n$$

となるから、汎検定問題の標本空間を制限する際の面倒な議論は不要となる。また、候補値 $\mathbf{w} \sim \mathrm{N}(\boldsymbol{\omega}, \tau^2 \mathbf{I})$ に関する二つの統計量；

$$\begin{aligned} R_0^2(\mathbf{w}) &= \mathbf{w}'(\mathbf{I} - \mathbf{P})\mathbf{w} \sim \tau^2 \chi^2(k_1, \lambda_1) \\ k_1 &= n - r, \tau^2 \lambda_1 = \boldsymbol{\omega}'(\mathbf{I} - \mathbf{P})\boldsymbol{\omega} \\ R_1^2(\mathbf{w}) - R_0^2(\mathbf{w}) &= (\mathbf{w} - \mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{P} - \mathbf{U})(\mathbf{w} - \mathbf{X}\boldsymbol{\beta}_0) \sim \tau^2 \chi^2(k_2, \lambda_2) \\ k_2 &= k, \tau^2 \lambda_2 = (\boldsymbol{\omega} - \mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{P} - \mathbf{U})(\boldsymbol{\omega} - \mathbf{X}\boldsymbol{\beta}_0) \end{aligned}$$

が独立に分布することを注意しておく。

さて、$c'_\alpha = kc_\alpha / (n - r)$ とおけば、このときの受容リスクは；

$$\mathrm{R}_\alpha(\boldsymbol{\omega}) = \mathrm{P}_{\mathrm{G}(\mathbf{w}|\boldsymbol{\omega})}\left(F(\mathbf{w}) > c_\alpha\right) = \mathrm{P}_{\mathrm{G}(\mathbf{w}|\boldsymbol{\omega})}\left(R_1^2(\mathbf{w}) - R_0^2(\mathbf{w}) > c'_\alpha R_0^2(\mathbf{w})\right)$$

と定義することができる[5]。ここで、Γ 分布の密度関数と分布関数をそれぞれ；

[5] この F 統計量は二重非心 F 分布に従うが、この分布を直接用いる方法は賢明でない。

$$\gamma(x|\alpha,p)=\frac{\alpha^{p}}{\Gamma(p)}e^{-\alpha x}x^{p-1},\Gamma(x|\alpha,p)=\int_0^x\gamma(t|\alpha,p)dt$$

とすれば、括弧内右辺の統計量$c'_\alpha R_0^2$の密度関数ψは；

$$\begin{aligned}\psi_{c'_\alpha R_0^2}(x)&=e^{-\frac{\lambda_1}{2}}\sum_{s=0}^{\infty}\frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s\frac{1}{2^{s+\frac{k_1}{2}}\Gamma\left(s+\frac{k_1}{2}\right)}e^{-\frac{x}{2c'_\alpha}}\left(\frac{x}{c'_\alpha}\right)^{s+\frac{k_1}{2}-1}\frac{1}{c'_\alpha}\\&=e^{-\frac{\lambda_1}{2}}\sum_{s=0}^{\infty}\frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s\frac{1}{(2c'_\alpha)^{s+\frac{k_1}{2}}\Gamma\left(s+\frac{k_1}{2}\right)}e^{-\frac{x}{2c'_\alpha}}x^{s+\frac{k_1}{2}-1}\\&=e^{-\frac{\lambda_1}{2}}\sum_{s=0}^{\infty}\frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s\gamma\left(x\left|\frac{1}{2c'_\alpha},s+\frac{k_1}{2}\right.\right)\end{aligned}$$

となる。以上の結果から、受容リスクは次のように表すことができる。

$$\begin{aligned}\mathrm{R}_\alpha(\omega)&=\mathrm{P}_{\mathrm{G}(\mathbf{w}|\omega)}\left(R_1^2(\mathbf{w})-R_0^2(\mathbf{w})>c'_\alpha R_0^2(\mathbf{w})\right)\\&=\int_0^\infty \mathrm{f}_{c'_\alpha R_0^2}(x)\mathrm{P}_{R_1^2-R_0^2}\left(R_1^2(\mathbf{w})-R_0^2(\mathbf{w})>x\right)dx\\&=\int_0^\infty e^{-\frac{\lambda_1}{2}}\sum_{s=0}^{\infty}\frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s\gamma\left(x\left|\frac{1}{2c'_\alpha},s+\frac{k_1}{2}\right.\right)\left\{1-\int_0^x e^{-\frac{\lambda_2}{2}}\sum_{t=0}^{\infty}\frac{1}{t!}\left(\frac{\lambda_2}{2}\right)^t\gamma\left(z\left|\frac{1}{2},t+\frac{k_2}{2}\right.\right)dz\right\}dx\\&=\int_0^\infty e^{-\frac{\lambda_1}{2}}\sum_{s=0}^{\infty}\frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s\gamma\left(x\left|\frac{1}{2c'_\alpha},s+\frac{k_1}{2}\right.\right)\left\{1-e^{-\frac{\lambda_2}{2}}\sum_{t=0}^{\infty}\frac{1}{t!}\left(\frac{\lambda_2}{2}\right)^t\Gamma\left(x\left|\frac{1}{2},t+\frac{k_2}{2}\right.\right)\right\}dx\\&=1-e^{-\frac{\lambda_1+\lambda_2}{2}}\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s\left(\frac{\lambda_2}{2}\right)^t\int_0^\infty\gamma\left(x\left|\frac{1}{2c'_\alpha},s+\frac{k_1}{2}\right.\right)\Gamma\left(x\left|\frac{1}{2},t+\frac{k_2}{2}\right.\right)dx\end{aligned}$$

この展開における積分と総和の順序交換は被積分関数が$[0,\infty)$で可測な正値関数であるため保証されている。

このように、受容リスクは二つの非心度λ_1,λ_2の関数となっている。そして、これらの非心度は事前分布Gの平均ω、すなわち、代表値ωの関数だから、以下の議論では、この関係を明示して$\lambda_1(\omega),\lambda_2(\omega)$と表すことにする。

さて、この汎検定問題は；

$$\boldsymbol{\omega} = \mathbf{y} + \boldsymbol{\pi}^{(-)}, \mathbf{y} \sim \mathrm{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}), \boldsymbol{\pi}^{(-)} \sim \mathrm{N}(\mathbf{0}, \tau^2\mathbf{I}), \mathrm{H}_0 : \mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi} \ vs.\ \mathrm{H}_1 : \mathbf{H}\boldsymbol{\beta} \neq \boldsymbol{\xi}$$

と形式IIに変換することができる。標本$\boldsymbol{\omega}$の分布関数$\overline{\mathrm{F_G}}(\boldsymbol{\omega})$は確率変数$\mathbf{y}$と$\boldsymbol{\pi}^{(-)}$の畳み込みとして定義されるから、この場合は正規分布$\mathrm{N}\left(\mathbf{X}\boldsymbol{\beta}, (\sigma^2 + \tau^2)\mathbf{I}\right)$となる。この正規分布に対して*3-2-1*の変換を適用すると；

$$\begin{aligned} R_0^2(\boldsymbol{\omega}) &= \boldsymbol{\omega}'(\mathbf{I} - \mathbf{P})\boldsymbol{\omega} \sim (\sigma^2 + \tau^2)\chi^2(k_1) \\ R_1^2(\boldsymbol{\omega}) - R_0^2(\boldsymbol{\omega}) &= (\boldsymbol{\omega} - \mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{P} - \mathbf{U})(\boldsymbol{\omega} - \mathbf{X}\boldsymbol{\beta}_0) \sim (\sigma^2 + \tau^2)\chi^2(k_2, \bar{\lambda}) \\ (\sigma^2 + \tau^2)\bar{\lambda} &= (\mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta}_0)'(\mathbf{P} - \mathbf{U})(\mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta}_0) \end{aligned}$$

が成立する。この関係を用いることで、代表値$\boldsymbol{\omega}$に関して確率変数である受容リスクの母数$\lambda_1(\boldsymbol{\omega}), \lambda_2(\boldsymbol{\omega})$が；

$$\begin{aligned} \tau^2\lambda_1(\boldsymbol{\omega}) &= (\sigma^2 + \tau^2)R_0^2(\boldsymbol{\omega}) \sim (\sigma^2 + \tau^2)\chi^2(k_1) \\ &\therefore\ \lambda_1(\boldsymbol{\omega}) \sim \left(1 + \sigma^2/\tau^2\right)\chi^2(k_1) \\ \tau^2\lambda_2(\boldsymbol{\omega}) &= (\sigma^2 + \tau^2)\left(R_1^2(\boldsymbol{\omega}) - R_0^2(\boldsymbol{\omega})\right) \sim (\sigma^2 + \tau^2)\chi^2(k_2, \bar{\lambda}) \\ &\therefore\ \lambda_2(\boldsymbol{\omega}) \sim \left(1 + \sigma^2/\tau^2\right)\chi^2(k_2, \bar{\lambda}) \end{aligned}$$

に従うことが分かる。当然、二つの統計量の比$\lambda_1(\boldsymbol{\omega})/\lambda_2(\boldsymbol{\omega}) = F^*(\boldsymbol{\omega})$は未知母数$\sigma^2$に依存しない。この統計量を受容リスクに代入すると；

$$\mathrm{R}_\alpha(\boldsymbol{\omega}) = 1 - e^{-\frac{(1+F^*)}{2}\lambda_2} \sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \frac{1}{s!t!}\left(F^*\right)^s \left(\frac{\lambda_2}{2}\right)^{s+t} \int_0^\infty \gamma\left(x \middle| \frac{1}{2c'_\alpha}, s + \frac{k_1}{2}\right)\Gamma\left(x \middle| \frac{1}{2}, t + \frac{k_2}{2}\right)dx$$

と変形することができる。このように、受容リスクは統計量$F^*(\boldsymbol{\omega})$だけの関数として表すことができない。統計量$\lambda_2(\boldsymbol{\omega})$は未知母数$\sigma^2$に依存しているから、これは受容リスクの確率が実際には求められないことを意味している。すなわち、汎検定関数の有意水準を求めることはできないということである。

汎検定関数の棄却限界$r = L_{01}/\left(L_{01} + 1\right)$は損失関数を基に定義するため、有意水準が求

められなくても特に問題はない。しかし、この値が分析における重要な指標であることに疑いはない。そこで、次に汎検定関数の有意水準を詳しく調べていくことにする。

3-2-3 有意水準に関する考察

本節の初めに述べたように、ここでは回帰モデルの分散 σ^2 を期待値 $\overline{\sigma}^2$ の確率変数として扱っている。この分散が未知母数であることに変わりはないから、分散 σ^2 に依存しない分散分析が検定方法として採用されている。しかし、分散が未知の場合は汎検定関数の有意水準を求めることができないので、ここでは分散を期待値 $\overline{\sigma}^2$ と見なしたときの有意水準について考えていく。

まず、原問題の分散を期待値 $\overline{\sigma}^2$ したときの代表値 $\boldsymbol{\omega}$ の分布 $\mathrm{N}\left(\mathbf{X\beta},(\overline{\sigma}^2+\tau^2)\mathbf{I}\right)$ と、検定問題 $\mathrm{H}_0{:}\mathbf{H\beta}=\xi$ *vs.* $\mathrm{H}_1{:}\mathbf{H\beta}\neq\xi$ との関係をまとめておく。前述したように、二つの統計量 $R_0^2(\boldsymbol{\omega}), R_1^2(\boldsymbol{\omega})-R_0^2(\boldsymbol{\omega})$ について；

$$
\begin{aligned}
R_0^2(\boldsymbol{\omega}) &= \boldsymbol{\omega}'(\mathbf{I}-\mathbf{P})\boldsymbol{\omega} \sim (\overline{\sigma}^2+\tau^2)\chi^2(k_1) \\
R_1^2(\boldsymbol{\omega})-R_0^2(\boldsymbol{\omega}) &= (\boldsymbol{\omega}-\mathbf{X\beta}_0)'(\mathbf{P}-\mathbf{U})(\boldsymbol{\omega}-\mathbf{X\beta}_0) \sim (\overline{\sigma}^2+\tau^2)\chi^2(k_2,\overline{\lambda}) \\
(\overline{\sigma}^2+\tau^2)\overline{\lambda} &= (\mathbf{X\beta}-\mathbf{X\beta}_0)'(\mathbf{P}-\mathbf{U})(\mathbf{X\beta}-\mathbf{X\beta}_0)
\end{aligned}
$$

が成立している。ここで、二つの統計量は独立に分布し、統計量 $R_0^2(\boldsymbol{\omega})$ に関する第一式は仮説とは無関係に成立する。これに対し、非心 χ^2 分布に従う統計量 $R_1^2(\boldsymbol{\omega})-R_0^2(\boldsymbol{\omega})$ の非心度 $\overline{\lambda}$ は帰無仮説が真のときに限りゼロとなる。ここから、代表値 $\boldsymbol{\omega}$ を標本とした線形仮説の検定は、統計量 $R_1^2(\boldsymbol{\omega})-R_0^2(\boldsymbol{\omega})$ の非心度 $\overline{\lambda}$ に関する仮説 $\mathrm{H}_0{:}\overline{\lambda}=0$ *vs.* $\mathrm{H}_1{:}\overline{\lambda}>0$ の検定として考えればよいことが分かる。実際、この検定方法は1-3節で定義した二つの変換群 G_1, G_2 に関して不変な一様最強力不変検定となっている。

それでは、汎検定関数の有意水準を求めてみよう。分散比を $V=\overline{\sigma}^2/\tau^2$ とおけば、受容リスクの母数 $\lambda_1(\boldsymbol{\omega}),\lambda_2(\boldsymbol{\omega})$ が、代表値 $\boldsymbol{\omega}\sim\mathrm{N}\left(\mathbf{X\beta},(\overline{\sigma}^2+\tau^2)\mathbf{I}\right)$ に関して；

$$
\lambda_1(\boldsymbol{\omega}) \sim (1+V)\chi^2(k_1),\ \lambda_2(\boldsymbol{\omega}) \sim (1+V)\chi^2(k_2,\overline{\lambda})
$$

と分布することは*3-2-2*で説明している。ここで、受容リスクは固定した母数$\lambda_1(\omega)$の下で母数$\lambda_2(\omega)$の増加関数だから[6]、所与の受容リスク$\mathrm{R}_\alpha(\omega)=r_0$に対し；

$$\lambda_{10}\in\left\{\lambda_1(\omega)\middle|\mathrm{R}_\alpha(\omega)=r_0\right\}$$

を条件として与えれば、統計量$\lambda_2(\omega)=\lambda_{20}$を一意に定めることができる。すなわち；

$$\left\{\omega\middle|\mathrm{R}_\alpha(\omega)=r_0\middle|\lambda_1(\omega)=\lambda_{10}\right\}=\left\{\omega\middle|\lambda_2(\omega)=\lambda_{20}\middle|\lambda_1(\omega)=\lambda_{10}\right\}$$

が成立する。統計量$\lambda_1(\omega)=\omega'(\mathbf{I}-\mathbf{P})\omega/\tau^2$は常に既知だから、この受容リスク$r_0$に対して統計量$\lambda_2(\omega)=\lambda_{20}$が一意に定められることが分かる。

ここで、任意の対立仮説$\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}^*\neq\boldsymbol{\xi}$を取り、この連立方程式の特殊解を$\boldsymbol{\beta}^*$とし、対応する非心度を$(\overline{\sigma}^2+\tau^2)\lambda^*=(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}^*)'(\mathbf{P}-\mathbf{U})(\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}^*)$とする。このとき、既知である統計量$\lambda_1(\omega)=\lambda_{10}$を与えたとき、受容リスク$r_0$の持つ確率は；

$$\begin{aligned}\mathrm{P}_{\overline{\mathrm{F}_\mathrm{G}}}\left(\mathrm{R}_\alpha(\omega)=r_0\middle|\lambda_1(\omega)=\lambda_{10},\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}^*\right)&=\mathrm{P}_{\overline{\mathrm{F}_\mathrm{G}}}\left(\lambda_2(\omega)=\lambda_{20}\middle|\lambda_1(\omega)=\lambda_{10},\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}^*\right)\\&=\mathrm{P}_{\overline{\mathrm{F}_\mathrm{G}}}\left(\lambda_2(\omega)=\lambda_{20}\middle|\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}^*\right)\\&=\mathrm{P}_{(1+V)\chi^2(k_2,\lambda)}\left(\lambda_{20}\middle|\lambda=\lambda^*\right)\end{aligned}$$

と表すことができる。ここから、受容リスクを用いた検定方法が非心χ^2統計量；

$$\lambda_2(\omega)=(1+V)\left(R_1^2(\omega)-R_0^2(\omega)\right)\sim(1+V)\chi^2(k_2,\overline{\lambda})$$

による一様最強力不変検定と同値であることが分かる。

以上の議論から、汎検定関数；

[6] *3-2-5*参照。

$$\bar{\delta}(\omega)=\begin{cases}1;\mathrm{R}_{\alpha}(\omega)>r_0\\0;\mathrm{R}_{\alpha}(\omega)<r_0\end{cases}$$

の有意水準α_0は、所与の分散比Vと統計量$\lambda_1(\omega)=\lambda_{10}$に対して；

$$\begin{aligned}\mathrm{P}_{\overline{\mathrm{F}_\mathrm{G}}}\left(\mathrm{R}_{\alpha}(\omega)>r_0\middle|\lambda_1(\omega)=\lambda_{10},\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}\right)&=\mathrm{P}_{(1+V)\chi^2(k_2,\lambda)}\left(\chi^2>\lambda_{20}\middle|\lambda=0\right)\\&=\mathrm{P}_{(1+V)\chi^2(k_2)}\left(\chi^2>\lambda_{20}\right)\\&=\alpha_0\end{aligned}$$

と求めることができる。

ところで、この汎検定関数は未知母数σ^2を期待値$\bar{\sigma}^2$として構成したものだから、F統計量による一様最強力不変検定（分散分析）とは一致しない。実際に分散が期待値$\bar{\sigma}^2$に近い値であれば、分散分析（変換群$G_1\sim G_3$に不変）より不変性の条件が緩い受容リスクによる汎検定関数（変換群G_1,G_2に不変）の方が、任意の対立仮説$\mathbf{H}\boldsymbol{\beta}=\boldsymbol{\xi}^*\neq\boldsymbol{\xi}$に対して高い検出力を持つはずである。しかし、期待値$\bar{\sigma}^2$からの偏差が大きい場合の優劣は必ずしも明らかではない。

3-2-4 数値例

この数値例の目的は、汎検定関数の有意水準に対応する受容リスクの大きさを実際に求めることである。回帰モデルと線形仮説の自由度は$k_1=n-r=47, k_2=k=3$、分散分析の有意水準αは0.1、事前分布Gの分散τ^2は25と固定した。これに対し、汎検定関数の有意水準α_0と回帰モデルの分散の期待値$\bar{\sigma}^2$、統計量$\lambda_1(\omega)$の値は複数の組み合わせについて計算した。

まず、分散の期待値$\bar{\sigma}^2$を16,25,36,49とする。これらの値に応じた分散比$V=\bar{\sigma}^2/25$と三通りの確率$p=0.3,0.5,0.7$を基に、統計量$\lambda_1(\omega)$の値$\lambda_{1,p}$を；

$$\mathrm{P}_{(1+V)\chi^2(3)}(\lambda\geq\lambda_{1,p})=p$$

を満たすように求めた。そして、0.75から0.99まで0.01刻みで調べた汎検定関数の有意水準α_0と分散比Vの各組み合わせに対し、統計量$\lambda_2(\omega)$の値λ_{2,α_0}を；

$$\mathrm{P}_{(1+V)\chi^2(47)}(\lambda \geq \lambda_{2,\alpha_0}) = \alpha_0$$

となるように求めた。

最後に、この汎検定問題における受容リスクは次のように計算した。F(3,47)分布の上側 10%点 $c_{0.1} = 2.204182$ に対して $c'_{0.1} = 3c_{0.1}/47$ とおけば、受容リスクは互いに独立な確率変数 $R_0^2, R_1^2 - R_0^2$ により；

$$\begin{aligned} &\mathrm{R}_{\alpha}(\omega) = \mathrm{P}\left(R_1^2 - R_0^2 > c'_{0.1} R_0^2\right) \\ &R_0^2 \sim \tau^2 \chi^2(47, \lambda_{1,p}), R_1^2 - R_0^2 \sim \tau^2 \chi^2(3, \lambda_{2,\alpha_0}) \end{aligned}$$

と定義することができる。そこで、上の分布に従う二つの乱数 $R_0^2, R_1^2 - R_0^2$ を 10000 回発生させたときに、括弧内の不等式が成立した回数の比率を受容リスクの近似値とした。

これらの設定の下で計算した結果が表 3-2.1 ～ 4 である[7]。この表から；

- 二つの有意水準 α と α_0 を固定しても、受容リスクは分散比 V と統計量 $\lambda_1(\omega)$ に依存して様々な値を取りうる
- 代表値による検定で、非常に小さい有意水準で帰無仮説が棄却できるとき、有意水準10%の分散分析に対応した受容リスクも大きい値となる

という二つの含意を導くことができる。さらに、有意水準 5%の汎検定関数で帰無仮説を棄却できても、原問題で棄却されていた可能性は約 50%以下（受容リスクが 0.5 以下）に過ぎないことも分かる。一般には真の分散比 V は未知という点を考慮すると、この結果は代表値に直接適用した検定結果を盲目的に受け入れることの危険性を指摘していると言えよう。

3-2-5 補足：受容リスクが非心度 λ_2 の増加関数であることの証明

3-2-3 で説明した汎検定関数の有意水準を導く過程では、受容リスク；

---

[7] 固定した統計量 $\lambda_{1,p}$ の下で受容リスクは有意水準 α_0 の増加関数であるが、5 個所で逆転が起きていたので、これらについては再計算した。

$$\mathrm{R}_\alpha(\omega) = 1 - e^{-\frac{\lambda_1+\lambda_2}{2}} \sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t \Gamma^*(s,t)$$

$$\Gamma^*(s,t) = \int_0^\infty \gamma\left(x \left| \frac{1}{2c'_\alpha}, s + \frac{k_1}{2}\right.\right) \Gamma\left(x \left| \frac{1}{2}, t + \frac{k_2}{2}\right.\right) dx$$

の持つ次の性質が用いられた。すなわち；

定理 3-2.1

受容リスク $\mathrm{R}_\alpha(\omega)$ は固定した母数 $\lambda_1(\omega)$ の下で母数 $\lambda_2(\omega)$ の単調増加関数である。

である。そこで、この定理の証明をもって本節の終わりとしよう。まず、その証明で必要となる二つの定理を補題として載せておく。

定理 3-2.2

区間 $[0,\infty)\times[0,\infty)$ で定義された関数；

$$\mathrm{f}_{s,t}(\lambda_1,\lambda_2) = \frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t e^{-\frac{\lambda_1+\lambda_2}{2}} \Gamma^*(s,t)$$

による無限級数 $\sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \mathrm{f}_{s,t}(\lambda_1,\lambda_2)$ は当該区間で一様収束する。

（証明）任意の s,t に対して $\Gamma^*(s,t) < 1$ だから、区間 $[0,\infty)\times[0,\infty)$ で；

$$\mathrm{f}_{s,t}(\lambda_1,\lambda_2) < \frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t e^{-\frac{\lambda_1+\lambda_2}{2}}$$

が成立する。この関数列による無限級数は当該区間で；

$$\sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t e^{-\frac{\lambda_1+\lambda_2}{2}} = \left(e^{-\frac{\lambda_1}{2}} \sum_{s=0}^{\infty} \frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s\right) \times \left(e^{-\frac{\lambda_2}{2}} \sum_{t=0}^{\infty} \frac{1}{t!}\left(\frac{\lambda_2}{2}\right)^t\right) = 1$$

と収束するから、Weierstrass の優級数定理より、無限級数$\sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \mathrm{f}_{s,t}(\lambda_1, \lambda_2)$が当該区間で一様収束することが分かる。（終）

定理 3-2.3

任意の s,t に対し、$\Gamma^*(s,t) - \Gamma^*(s,t+1) > 0$ である。

（証明）二つの独立な確率変数 $x_1 \sim \Gamma(1/2, t + k_2/2), x_2 \sim \Gamma(1/2, 1)$ を考える。ここで、Γ 分布の持つ再生性と、その非負性に注意すると；

$$\Gamma\left(x\left|\frac{1}{2}, t+\frac{k_2}{2}+1\right.\right) = \mathrm{P}_{X_1,X_2}(X_1 + X_2 < x)$$
$$= \mathrm{P}_{X_1,X_2}(X_1 < x - X_2) < \mathrm{P}_{X_1}(X_1 < x) = \Gamma\left(x\left|\frac{1}{2}, t+\frac{k_2}{2}\right.\right)$$

であることが分かる。この関係から；

$$\Gamma^*(s,t) - \Gamma^*(s,t+1) = \int_0^{\infty} \gamma\left(x\left|\frac{1}{2c'_\alpha}, s+\frac{k_1}{2}\right.\right)\left\{\Gamma\left(x\left|\frac{1}{2}, t+\frac{k_2}{2}\right.\right) - \Gamma\left(x\left|\frac{1}{2}, t+\frac{k_2}{2}+1\right.\right)\right\}dx > 0$$

と命題を証明することができる。（終）

それでは、定理 3-2.1 を証明しよう。まず、受容リスク；

$$\mathrm{R}_\alpha(\omega) = 1 - e^{-\frac{\lambda_1+\lambda_2}{2}} \sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t \Gamma^*(s,t)$$
$$= 1 - e^{-\frac{\lambda_1+\lambda_2}{2}} \sum_{s=0}^{\infty}\left\{\sum_{t=1}^{\infty} \frac{1}{s!t!}\left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t \Gamma^*(s,t) + \frac{1}{s!}\left(\frac{\lambda_1}{2}\right)^s \Gamma^*(s,0)\right\}$$

を母数 λ_2 で偏微分すると、定理 3-2.2 より項別微分が可能だから、その偏導関数は次のようになる。

$$
\begin{aligned}
\frac{\partial \mathrm{R}_{\alpha}}{\partial \lambda_2} &= \frac{1}{2} e^{-\frac{\lambda_1}{2}} \sum_{s=0}^{\infty} \left[\sum_{t=1}^{\infty} \left\{ \frac{1}{s!t!} \left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t e^{-\frac{\lambda_2}{2}} \Gamma^*(s,t) - \frac{1}{s!(t-1)!} \left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^{t-1} e^{-\frac{\lambda_2}{2}} \Gamma^*(s,t) \right\} \right. \\
&\quad \left. + \frac{1}{s!} \left(\frac{\lambda_1}{2}\right)^s e^{-\frac{\lambda_2}{2}} \Gamma^*(s,0) \right] \\
&= \frac{1}{2} e^{-\frac{\lambda_1+\lambda_2}{2}} \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} \frac{1}{s!t!} \left(\frac{\lambda_1}{2}\right)^s \left(\frac{\lambda_2}{2}\right)^t \left(\Gamma^*(s,t) - \Gamma^*(s,t+1)\right) \\
&> 0
\end{aligned}
$$

ここで、最後の不等号は定理3-2.3の結果である。この不等式は、受容リスクR_{α}が固定した母数λ_1の下で母数λ_2に関する単調増加関数であることを意味している。

表３-１.１

	$\mathbf{P}_1=(1/3\quad 1/3\quad 1/3)$			$\mathbf{P}_2=(0.495\quad 0.01\quad 0.495)$			$\mathbf{P}_3=(0.01\quad 0.98\quad 0.01)$		
ω	$P(\omega\|\theta=0)$	$P(\omega\|\theta=1)$	尤度比	$P(\omega\|\theta=0)$	$P(\omega\|\theta=1)$	尤度比	$P(\omega\|\theta=0)$	$P(\omega\|\theta=1)$	尤度比
−3.5	0.00333	0.00000	1.00000	0.00495	0.00000	0.00000	0.00010	0.00000	0.00000
−3.0	0.01333	0.00000	0.00000	0.01495	0.00000	0.00000	0.01010	0.00000	0.00000
−2.5	0.02667	0.00333	0.12500	0.02505	0.00495	0.19760	0.02990	0.00010	0.00334
−2.0	0.04667	0.01333	0.28571	0.04990	0.01495	0.29960	0.04020	0.01010	0.25124
−1.5	0.07000	0.02667	0.38095	0.07000	0.02505	0.35786	0.07000	0.02990	0.42714
−1.0	0.10667	0.04667	0.43750	0.10990	0.04990	0.45405	0.10020	0.04020	0.40120
−0.5	0.15000	0.07000	0.46667	0.15000	0.07000	0.46667	0.15000	0.07000	0.46667
0.0	0.16667	0.10667	0.64000	0.15050	0.10990	0.73023	0.19900	0.10020	0.50352
0.5	0.15000	0.15000	1.00000	0.15000	0.15000	1.00000	0.15000	0.15000	1.00000
1.0	0.10667	0.16667	1.56250	0.10990	0.15050	1.36943	0.10020	0.19900	1.98603
1.5	0.07000	0.15000	2.14286	0.07000	0.15000	2.14286	0.07000	0.15000	2.14286
2.0	0.04667	0.10667	2.28571	0.04990	0.10990	2.20240	0.04020	0.10020	2.49254
2.5	0.02667	0.07000	2.62500	0.02505	0.07000	2.79441	0.02990	0.07000	2.34114
3.0	0.01333	0.04667	3.50000	0.01495	0.04990	3.33779	0:01010	0.04020	3.98020
3.5	0.00333	0.02667	8.00000	0.00495	0.02505	5.06061	0.00010	0.02990	299.000
4.0	0.00000	0.01333	∞	0.00000	0.01495	∞	0.00000	0.01010	∞
4.5	0.00000	0.00333	∞	0.00000	0.00495	∞	0.00000	0.00010	∞

表3-2.1

$\bar{\sigma}^2 = 16, \tau^2 = 25$				
		受容リスク		
有意水準 α_0	λ_{2,α_0}	$\lambda_{1,0.3}$	$\lambda_{1,0.5}$	$\lambda_{1,0.7}$
0.75	6.74	0.14	0.12	0.10
0.76	6.90	0.15	0.13	0.10
0.77	7.07	0.15	0.13	0.10
0.78	7.24	0.17	0.14	0.11
0.79	7.42	0.17	0.14	0.12
0.80	7.61	0.18	0.15	0.12
0.81	7.81	0.19	0.16	0.13
0.82	8.02	0.21	0.17	0.14
0.83	8.24	0.22	0.18	0.15
0.84	8.47	0.23	0.19	0.15
0.85	8.72	0.24	0.20	0.17
0.86	8.98	0.25	0.21	0.18
0.87	9.26	0.27	0.22	0.18
0.88	9.57	0.28	0.24	0.20
0.89	9.89	0.29	0.25	0.20
0.90	10.25	0.31	0.26	0.22
0.91	10.65	0.32	0.28	0.24
0.92	11.08	0.35	0.31	0.25
0.93	11.58	0.37	0.32	0.28
0.94	12.15	0.40	0.36	0.31
0.95	12.82	0.44	0.39	0.34
0.96	13.63	0.47	0.42	0.36
0.97	14.67	0.53	0.47	0.42
0.98	16.13	0.60	0.55	0.49
0.99	18.61	0.70	0.66	0.60

表3-2.2

$\overline{\sigma}^2 = 25, \tau^2 = 25$				
		受容リスク		
有意水準 α_0	λ_{2,α_0}	$\lambda_{1,0.3}$	$\lambda_{1,0.5}$	$\lambda_{1,0.7}$
0.75	8.22	0.15	0.11	0.09
0.76	8.41	0.15	0.12	0.09
0.77	8.62	0.16	0.13	0.10
0.78	8.83	0.16	0.13	0.10
0.79	9.05	0.18	0.14	0.11
0.80	9.28	0.19	0.15	0.12
0.81	9.53	0.20	0.16	0.13
0.82	9.78	0.21	0.17	0.13
0.83	10.05	0.22	0.18	0.14
0.84	10.33	0.24	0.19	0.15
0.85	10.63	0.25	0.21	0.16
0.86	10.95	0.27	0.22	0.17
0.87	11.30	0.27	0.22	0.18
0.88	11.67	0.30	0.25	0.20
0.89	12.07	0.31	0.26	0.21
0.90	12.50	0.33	0.28	0.22
0.91	12.98	0.35	0.30	0.24
0.92	13.52	0.38	0.32	0.27
0.93	14.12	0.40	0.34	0.28
0.94	14.81	0.45	0.39	0.33
0.95	15.63	0.47	0.41	0.35
0.96	16.62	0.53	0.46	0.40
0.97	17.89	0.58	0.52	0.45
0.98	19.67	0.66	0.59	0.53
0.99	22.69	0.76	0.71	0.65

表3-2.3

$\overline{\sigma}^2 = 36, \tau^2 = 25$				
		受容リスク		
有意水準 α_0	λ_{2,α_0}	$\lambda_{1,0.3}$	$\lambda_{1,0.5}$	$\lambda_{1,0.7}$
0.75	10.02	0.14	0.11	0.08
0.76	10.26	0.15	0.12	0.09
0.77	10.51	0.15	0.12	0.09
0.78	10.77	0.17	0.13	0.09
0.79	11.04	0.18	0.14	0.10
0.80	11.33	0.19	0.14	0.11
0.81	11.62	0.20	0.15	0.11
0.82	11.93	0.22	0.17	0.12
0.83	12.26	0.23	0.18	0.13
0.84	12.61	0.24	0.19	0.14
0.85	12.97	0.25	0.20	0.16
0.86	13.36	0.27	0.21	0.17
0.87	13.78	0.28	0.23	0.18
0.88	14.23	0.30	0.25	0.20
0.89	14.72	0.33	0.27	0.22
0.90	15.25	0.34	0.29	0.23
0.91	15.84	0.37	0.31	0.24
0.92	16.49	0.40	0.34	0.28
0.93	17.23	0.44	0.36	0.30
0.94	18.07	0.46	0.40	0.33
0.95	19.07	0.51	0.44	0.37
0.96	20.28	0.57	0.49	0.43
0.97	21.83	0.63	0.56	0.49
0.98	24.00	0.69	0.63	0.56
0.99	27.68	0.81	0.76	0.70

表3-2.4

$\overline{\sigma}^2 = 49, \tau^2 = 25$				
		受容リスク		
有意水準α_0	λ_{2,α_0}	$\lambda_{1,0.3}$	$\lambda_{1,0.5}$	$\lambda_{1,0.7}$
0.75	12.16	0.14	0.11	0.07
0.76	12.45	0.15	0.11	0.07
0.77	12.75	0.16	0.11	0.08
0.78	13.07	0.17	0.12	0.09
0.79	13.40	0.18	0.13	0.09
0.80	13.74	0.19	0.14	0.10
0.81	14.10	0.20	0.15	0.11
0.82	14.48	0.21	0.16	0.11
0.83	14.87	0.23	0.17	0.13
0.84	15.29	0.25	0.19	0.14
0.85	15.74	0.25	0.21	0.15
0.86	16.21	0.29	0.22	0.17
0.87	16.72	0.30	0.24	0.17
0.88	17.27	0.32	0.26	0.19
0.89	17.86	0.35	0.27	0.21
0.90	18.50	0.37	0.30	0.24
0.91	19.21	0.40	0.32	0.26
0.92	20.01	0.42	0.35	0.28
0.93	20.90	0.46	0.39	0.32
0.94	21.92	0.51	0.43	0.35
0.95	23.13	0.56	0.48	0.40
0.96	24.60	0.61	0.53	0.45
0.97	26.48	0.67	0.60	0.51
0.98	29.12	0.75	0.68	0.60
0.99	33.58	0.85	0.81	0.74

4 経営特性の測定モデル

本論文の目的は、企業調査で得られた経営特性の尺度を業種間や国際間で比較研究していく統一的な分析枠組みを提示することである。この分析枠組みの前提となるのが測定値を表現する測定モデルと第3章で解説した検定方法である。そこで、本章では経営特性の測定モデルについて説明していく。

まず、経営特性の測定という「テスト」を4-1節で特徴付けしておく。4-2節で連続的な尺度を表す測定モデルを提案し、このときの分析枠組みを4-3節で説明する。離散的な得点（Likert 尺度）を表す測定モデルは4-4節で提案する。最後に、既存のテスト理論では十分な対応ができない理由を4-5節で説明する。

4-1 経営特性の特徴と測定について

4-1-1 本論文が扱う経営特性

本論文では、企業の経営活動を特徴付ける経営学上の構成概念(Constructs)を経営特性と便宜的に定義しておく。その概念に応じて様々な経営特性が考えられるが、ここでは次の三つの条件を満たす経営特性に対象を限定する

(1) 理論的には間隔尺度(Interval Scale)を定義することが可能
(2) 現時点では、客観的な指標は存在しない
(3) 企業の客観的な分類を可能にする状態は存在しない

まず、間隔尺度が定義できる経営特性に限定するのは、尺度間の距離に意味のある間隔尺度には通常の統計手法が適用できるという分析上の都合のためである。しかし、状態の違いを「程度」で表現した尺度は間隔尺度と考えられるため、条件を満たす経営特性は案外多いのではないだろうか。もし、この経営特性に客観的な指標が存在すれば、既存の統計手法や分析枠組みが適用できるため、本論文で議論する必要はない。これが条件(2)を課す理由である。最後に、条件(3)は条件(2)の補足である。客観的な分類が可能な経営特性は離散的な尺度を構成することが可能なためである。

これらの条件を簡単に言い直すと、「連続的な状態を持つが基準となる状態はなく、理論的な測定方法も確立していない経営特性」と言うことができる。例えば、「監督者と従業員が現場で接触している程度」や「段取り時間短縮を努力している程度」といった経営特性が該当する。「接触」や「努力」の程度は連続的でも、その状態を分類するための基準は存在しないため、これらの程度を客観的に数量化することは難しいからである。

これに対し、製造現場で特定の工作機械を「使用してるか否か」が意味する経営特性には二種類の状態しか存在しない。この場合、ノンパラメトリック法を直接適用して比較研究を行うか、ロジット・モデルのような離散選択モデル(Discrete Choice Model)を介して分析を進めればよい。また、不良品率や営業利益のように定義が明確で連続的な値を持つ

場合は、適当なモデルを介して分散分析を適用すれば十分である。前述したように、こうした経営特性を本論文で取り上げる必要はない。

4-1-2 経営特性の測定の特徴

さて、これらの特徴を持つ経営特性を測定するには、操作的定義を示すことができないという条件(2)から、その経営特性の状態を指定された範囲の中に主観的判断で位置付けるという方法を取らざるをえない。そこで、本論文では；

- 回答者に経営特性を説明する文言
- その尺度の定義域を規定する範囲

の組を質問項目と呼ぶことにする。

この尺度の定義域としては実数の区間や数値のない線分（グラフ評定法）など連続的な尺度を構成するものと、評定尺度値（離散的な得点）を構成する離散的な数字の集合の二種類を考えることができる。しかし、評定尺度値について考察する4-4節を除いて、ここでは連続的な尺度を前提に議論を進めていく。評定尺度値を表現する測定モデルは連続的な尺度を扱う測定モデルから直接導かれ、いずれの場合も実質的に同じ分析方法が提案されるからである。

このように、回答者の判断に依存した経営特性の測定は第2章で要約したテスト理論の扱う対象と考えることができる。しかし、経営特性の測定モデルをテスト理論における測定モデル（テストモデル）と同一視することは必ずしも妥当ではない。これらの測定では測定対象と測定装置の捉え方が異なるからである。

この違いは小論文を用いたテストを例に考えると分かり易い。言うまでもなく、小論文の水準に反映された受験者の能力がテスト理論の測定対象がであり、この能力は仮説的な心理学的連続体の上で定義される。そして、ここでの測定装置は小論文という形式のテストだから、この測定を特徴付ける性質はテストの持つ性質に集約される。一方、経営特性の測定に該当する部分は採点者による小論文の評価である。測定対象となる小論文の水準は観測可能な状態であり、この場合の測定装置は採点者の思考となる。このように、小論文の評価を対象とする場合は、心理学的連続体とテストではなく、小論文の水準と採点者の思考について考察していく必要がある。

経営特性の測定に関する議論に戻り、以上の結果を表にまとめておこう。

	テストモデル	経営特性の測定モデル
測定対象	心理学的連続体の上で定義された潜在特性	観測可能な状態
測定装置	テスト（質問項目）	回答者の思考

ここから、経営特性の測定をモデル化するには経営特性の状態と回答者が持つ「心の物差し」を考察しなければならないことが分かる。また、個体差の測定では回答者の思考を測

定装置と考えることができる。しかし、この場合も測定対象は心理学的連続体の上に定義されるから、ここが経営特性の測定との違いとなる。

4-1-3 経営特性の状態

本論文で扱う経営特性は間隔尺度で表現できることが前提である。ここで、間隔尺度とは傾きが正の任意の線形変換を許容する尺度（変換された尺度も間隔尺度となる）と定義される。そして、測定される状態から実数への写像を考えたとき、その写像による対象とする状態の像が尺度である。この写像は状態間の関係系(Relational System)を保存する準同型写像(Homomorphism)でなければならない。

さて、状態間の順序関係と点列の収束、距離が生成する関係系から **R** における同型の関係系の中への準同型写像は間隔尺度を定義することができる[1]。そこで、本論文では経営特性の状態の集合を状態集合 $\overline{\boldsymbol{\Omega}}$ と表し、そこには距離が定義されていることを前提に議論を進めていく。もちろん、この距離は客観的な距離ではない。客観的な距離が定義できれば客観的な指標を定義することが可能となり、ここで扱う経営特性の条件と矛盾するからである。この理由から、本論文では人間の主観的判断に基づいた主観的距離を導入する。この主観的距離とは、状態集合 $\overline{\boldsymbol{\Omega}}$ の元である二つの状態を同程度とみなす主観的な近さを反映した距離のことである。二つの状態を同程度とみなすことができれば、主観的に状態間で差はないから、これらの主観的距離は小さくなる。反対に、同程度と考えることができなければ主観的距離は大きくなる。この議論を仮定としてまとめておこう。

仮定 4-1.1

経営特性の間隔尺度は $\overline{\boldsymbol{\Omega}}$ 上の順序関係、点列の収束、主観的距離が生成する関係系を保存する準同型写像により定義される。

間隔尺度が主観的距離に基づいているという仮定は重要である。尺度の差で表される経営特性の違いは回答者が感じる主観的な違いであり、理論的に導かれる客観的な違いを意味するものではないからである。

*4-1-2*で説明した経営特性の測定では、質問項目で指定される区間が準同型写像の値域であり、尺度の定義域となる。この区間を **I** とし、区間 **I** に対応した状態集合の部分集合を $\boldsymbol{\Omega} \subseteq \overline{\boldsymbol{\Omega}}$ とすれば、仮定 4-1.1 の準同型写像 h は；

$$h(\mathbf{I}) = \boldsymbol{\Omega}$$

を満たす準同型写像として一意に定めることができる。この集合 $\boldsymbol{\Omega}$ が実質的な準同型写像の定義域である。状態集合 $\overline{\boldsymbol{\Omega}}$ を用いないのは、写像の定義域が規定されていると、経営特

[1] Pfanzagl(1968, p *147*)

性に関する測定モデルを考えていくときに具合が悪いためである。状態集合を測定の目的からは対象とならないような広い範囲を含む状態の集合と定義すれば、このように考えることに問題はない。

4-1-4 二種類の誤差

ここで扱う経営特性の測定とは、回答者が指定された経営特性の状態を明確な基準のない中で特定化し、その特定化した状態に自分自身の思考を通して数値を与えるというものである。比喩的に言うと、目印のない対象に指定された位置を特定化し、そこに物差しを当てて目盛りを読み取るという測定である。この場合は、指定された位置を特定化するときと、物差しを当てるとき（その点以外に目印はない）に誤差が発生する。従って、経営特性の測定をモデル化するには二種類の誤差；

誤差Ⅰ：状態を特定化するときの誤差
誤差Ⅱ：特定化した状態を数値化するときの誤差

を考慮する必要がある。

これらの誤差については測定モデルを構築する次節で厳密に定義されるので、ここでは解説的な説明に止めておく。いま、ある回答者は指定された経営特性の状態$\overline{\omega} \in \overline{\Omega}$を異なる状態$\omega \in \overline{\Omega}$として認識したとしよう。このとき、特定の準同型写像hによる各状態の尺度を$\mathrm{h}(\overline{\omega}) = T, \mathrm{h}(\omega) = Z$とすれば；

$$Z = T + E$$

と表現することができる。このズレを表すEが誤差Ⅰであり、通常考えられる誤差に外ならない。しかし、この回答者が測定時に準同型写像hを用いる保証はない。測定時に準同型写像$\mathrm{h}_t \neq \mathrm{h}$を想定したとき、この写像による尺度を$\mathrm{h}_t(\omega) = Z_t$とすれば；

$$Z_t = Z + \pi_t^{(-)}$$

という平行移動の関係が成立する。この尺度の差$\pi_t^{(-)}$が誤差Ⅱである。しかし、誤差Ⅱを通常の誤差項と考えることはできない。誤差Ⅱは尺度Zに加えられた誤差項ではなく、観測された尺度Z_tを共通の尺度Zに変換するために分析者が導入した補正項を意味しているからである。もちろん、実際には誤差Ⅱの値は未知だから、尺度間の比較が可能となる尺度Zを観測値Z_tから求めることはできない。分析者に与えられる情報は、誤差Ⅱの持つ確率として定義される、観測値と誤差Ⅱとの差$Z_t - \pi_t^{(-)}$が補正された尺度Zとなる確率だけである。この解釈は3-1節で解説した形式Ⅱの考え方だから、観測された尺度Z_tに関する仮説検定は汎検定問題として扱わなければならないことが分かる。このように、測定

モデルに誤差IIを含めることは非常に重要である。ここで提案する測定モデルと分析枠組みを一般化可能性理論と区別する相違点は誤差IIの有無にあるからである。

4-2 経営特性の測定値を表現するモデル

4-2-1 測定に関する仮定

ここでは以下の議論で必要となる仮定をまとめておく。ただし、回答者を添え字 a、質問項目を添え字 g、企業を添え字 c で表すことにする。

まず、議論の前提として質問項目の妥当性；

仮定 4-2.1

状態集合 $\overline{\mathbf{\Omega}}$ は質問項目と回答者に依存しない。

を仮定する。前述した比喩を用いると、この妥当性の仮定は物差しを当てる対象が質問項目と回答者によらず共通ということを述べている。この測定対象に関する仮定の次に、物差しに関する仮定が要請されるのは自然なことであろう。

最初に要請されることは、ある質問項目と回答者の組が二つの物を測定して同じ長さである場合は、別の組が測定しても同じ長さになることである。そこで；

仮定 4-2.2

回答者間で $\overline{\mathbf{\Omega}}$ 上の主観的距離は定数倍を除いて等しい。

と仮定する。主観的距離の定数倍を許容するのは間隔尺度を前提にした結果である。なぜなら、間隔尺度は状態集合上の主観的距離が定義する関係系（二点間の距離に関する大小関係）を保存するが、その関係系は距離の定数倍に不変なためである。

もちろん、この仮定が厳密な意味で成立するはずはない。まず、その定義から主観的距離は個人に依存しているからである。さらに、経営特性を説明する質問項目の文言による影響も存在するだろう。従って、現実の企業調査では、回答者と質問項目による影響が小さくなるような調査計画に従うことが重要となる。具体的には、同質な回答者を選び、経営特性の説明が理解し易い質問項目を用意することである。

物差しに関する要請はこれだけではない。測定値を比較して意味があるのは、物差しの一目盛りの長さが共通の場合に限られるためである。ただし、目盛りの個数（区間 $\mathbf{I}$）は指定されているから、これは物差しの長さが共通という条件と同値である。この条件を本節の議論に即して述べると、測定時に想定する状態集合 $\mathbf{\Omega}$ は質問項目と回答者によらず同じ大きさを持つということである。この条件を正確に記述しよう。

まず、測定の度に異なる状態集合を明示するため、測定の試行を添え字 t により表すことを約束する。そして、回答者 a が質問項目 g に従い企業 c の経営特性を測定するときに

想定する状態集合の全体を集合族；

$$\mathbf{B}_{gca} = \left\{ \Omega_{gcat} \middle| \Omega_{gcat} \subset \overline{\Omega}, t = 1,2,\ldots \right\}$$

と表すことにしよう。そこで、仮定４-２.２から主観的距離は質問項目と回答者に依存しないので、適当な主観的距離 d を取り；

仮定 ４-２.３

回答者と質問項目、企業によらず任意の$\Omega \in \mathbf{B}_{gca}$の主観的距離 $d(\Omega)$は等しい。

を仮定する。これが問題の条件である。ただし、主観的距離 $d(\Omega)$とは；

$$d(\Omega) = d\left(\inf\left\{\omega \middle| \omega \in \Omega\right\}, \sup\left\{\omega \middle| \omega \in \Omega\right\}\right)$$

の略記である[2]。

仮定４-２.２と同じように、これも強い仮定と言わざるをえない。従って、この場合も回答者と質問項目への依存度を低める調査計画が要請される。同質な回答者を選び、文言の理解し易い質問項目を用いるというだけでなく、事前に調査目的や想定する企業を明確にしておく必要もあろう。

4-2-2 測定モデルの定式化

まず、経営特性の状態が特定化されたときに、この誤差 I を含む状態に尺度を与えるプロセスから考えていく。このプロセスで、測定の度に状態集合が異なることは既に述べた通りである。尺度の定義域は区間 **I** と固定されているため、各状態集合に対応した準同型写像は一意に定められるが、その尺度は状態集合に依存して異なる値となる。しかし、仮定４-２.３の下では、任意の尺度は平行移動により特定の状態集合による尺度で表現することができる。このとき、全く同質な二つの組では同じ状態は同じ値となるように、この変換はすべての組で共通する基準の下で行わなければならない。

そこで、状態集合$\Omega_{gca} \in \mathbf{B}_{gca}$を適当に固定して、この状態集合を定義域とする準同型写像をh_{gca}としよう。このとき、任意の$\Omega_{gcat} \in \mathbf{B}_{gca}$に対応した$h_{gcat}$に対し；

[2] 仮定４-１.１から、状態空間には順序関係と点列の収束が定義されている。

$$\mathrm{h}_{gcat}(\omega) = \mathrm{h}_{gca}(\omega) + \pi_{gcat}^{(-)}, \omega \in \overline{\Omega}$$

という$\pi_{gcat}^{(-)}$だけ平行移動した関係が成立する。この$\pi_{gcat}^{(-)}$が測定tにおいて生じた誤差IIを表している。ここで、次の仮定をおく。

仮定 4-2.4

誤差II $\pi_{gcat}^{(-)}$は分布関数$\mathrm{H}_{gca}\left(\pi^{(-)}\right)$に従う確率変数である。

しかし、固定した状態集合$\boldsymbol{\Omega}_{gca}$は任意に選ばれているため、この誤差IIの期待値がゼロになるとは限らない。そこで、回答者と企業、質問項目の各組に対し；

$$\mathrm{E}\left(\pi_{gcat}^{(-)}\right) = 0$$

を満たす状態集合を基準として用いることにする。

このように、回答者aが質問項目gにより測定した企業cの経営特性の尺度を特定の状態集合$\boldsymbol{\Omega}_{gca}$に基づく尺度で表現できることが示された。しかし、回答者a'と企業c'、質問項目g'の組の状態集合が$\boldsymbol{\Omega}_{g'c'a'} \neq \boldsymbol{\Omega}_{gca}$であれば、同じ状態を特定化しても異なる値の尺度が定義されてしまう。この場合も再び共通の尺度による表現が要請される。

そこで、適当な状態集合；

$$\boldsymbol{\Omega}_{\bullet\bullet\bullet} \subset \overline{\boldsymbol{\Omega}}, \mathrm{d}(\boldsymbol{\Omega}_{\bullet\bullet\bullet}) = \mathrm{d}(\boldsymbol{\Omega}), \forall \boldsymbol{\Omega} \in \mathbf{B}_{gca}$$

を取り、対応する準同型写像を$\mathrm{h}_{\bullet\bullet\bullet}$と表そう。この状態集合の選び方から、任意の準同型写像h_{gca}に対して平行移動の関係；

$$\mathrm{h}_{gca}(\omega) = \mathrm{h}_{\bullet\bullet\bullet}(\omega) + V_{gca}, \omega \in \overline{\Omega}$$

が成立する。そこで、最初の特定化と同じように、回答者と企業、質問項目の母集団における期待値（平均値）について；

$$\mathrm{E}_g \mathrm{E}_c \mathrm{E}_a (V_{gca}) = 0$$

となる状態集合$\mathbf{\Omega}_{\bullet\bullet\bullet}$を用いることにする。ところで、定数$V_{gca}$が誤差IIと異なる点は、この値が回答者と企業、質問項目の特性を表していることである。実際、これは第５章で説明する比較研究で重要な母数となる。

尺度$\mathrm{h}_{gca}(\omega)$に状態を特定化するときの誤差Ｉが含まれたままであることは、議論の始めに述べた通りである。そこで、誤差Ｉを組み込むことを定式化の最後に考えていく。

いま、企業cの経営特性の状態$\omega_c \in \overline{\Omega}$を回答者$a$が質問項目$g$に従い$\omega_{gcat} \in \overline{\Omega}$と特定化したとする。これを前述した準同型写像$\mathrm{h}_{\bullet\bullet\bullet}$により；

$$\begin{aligned} Z_{\bullet\bullet\bullet t} &= \mu_c + E_{gcat} \\ \mu_c &= \mathrm{h}_{\bullet\bullet\bullet}(\omega_c), Z_{\bullet\bullet\bullet t} = \mathrm{h}_{\bullet\bullet\bullet}(\omega_{gcat}) \end{aligned}$$

と表現したときの、尺度の差E_{gcat}が誤差Ｉとなる。そこで、誤差Ｉについて；

仮定 ４-２.５

$$E_{gcat} \sim \mathrm{N}\left(0, \sigma^2_{gca}\right)$$

を仮定する。

このときの尺度をZ_{gcat}（準同型写像h_{gcat}による尺度）とすれば、これまで導入してきた準同型写像を順に適用することで；

$$\begin{aligned} Z_{gcat} &= Z_{gca} + \pi^{(-)}_{gcat} \\ &= Z_{\bullet\bullet\bullet t} + V_{gca} + \pi^{(-)}_{gcat} \\ &= \mu_c + V_{gca} + \pi^{(-)}_{gcat} + E_{gcat} \end{aligned}$$

と表現することができる。これが本論文で提案する測定モデルである。

最後に、仮定４-２.５について補足しておく。測定誤差を意味する誤差Ｉに正規分布を仮定すること自体に無理はないが、尺度の定義域は区間$\mathbf{I}$と固定されているため；

$$\mathrm{P}(Z_{gcat} \in \mathbf{I}) = \frac{1}{\sqrt{2\pi\sigma^2_{gca}}} \int_{\mathbf{I}} \exp\left(-\frac{(z-\mu_c - V_{gca} - \pi^{(-)}_{gcat})^2}{2\sigma^2_{gca}} \right) dz < 1$$

が成立してしまう。しかし、この区間 **I** は質問項目で指定されているのだから、平均的には $\mathrm{P}(Z_{gcat} \in \mathbf{I}) \fallingdotseq 1$ と考えて問題ないだろう。

4‐3　企業調査への適用

4‐3‐1　企業調査の形式

ここでは、前節で提案した経営特性の測定モデルを現実の企業調査と比較研究に適用することを考えていく。まず、具体的な分析枠組みを提示する前に、本論文で対象とする企業調査について説明しておこう。

本論文では、企業調査を各企業の持つ経営特性を業種間や国際間で比較研究するために行われる調査活動と考えることにする。そして、ここで扱う経営特性は；

(1) 理論的には間隔尺度を定義することが可能
(2) 現時点では、客観的な指標は存在しない
(3) 企業の客観的な分類を可能にする状態は存在しない

を満たすことが前提である。この経営特性の測定には、その状態を指定された区間に含まれる実数で表すという方法が用いられる。この区間と当該経営特性を説明する文言の組を質問項目と呼ぶ。

また、通常の企業調査には；

(1) 回答者が回答する企業は一社だけである
(2) 回答者と質問項目の各組み合わせに測定の繰り返しは得られない

というデータ収集上の制約がある。この制約の下で、最尤法を前提とした検定方法を適用することが不適切であるのは明らかであろう。この問題は4‐5節で取り上げる。

それでは、本論文が想定する企業調査を簡単に紹介しておく。この調査は、前述した条件を満たす経営特性に関する業種間比較を目的（国際間や格付けの違う企業間での比較を目的に加えても構わない）としている。この目的のため、対象とする複数の業種から無作為に複数の企業が選ばれ、各企業からは複数の従業員が無作為に選ばれる。そして、用意された内容的妥当性を持つ複数の質問項目に従い、従業員は所属する企業の経営特性について回答する。ただし、この企業調査では、分析の前提となる測定モデルの妥当性を確保するため、仮定4‐2.1～3の説明で指摘した要請；

(1) 回答者は当該経営特性の知識を持っている
(2) 経営特性の説明が理解し易く、内容的妥当性を持つ質問項目を用意する
(3) 同質な回答者を選ぶ
(4) 調査目的や想定する企業を明確にしておく

が尊重されているものとする。

4-3-2 測定モデルを適用した分析方法

前節で提案した測定モデルによれば、回答者 a が質問項目 g に従い測定した企業 c の経営特性の尺度 Z_{gcat} は；

$$Z_{gcat} = T_{gca} + \pi^{(-)}_{gcat} + E_{gcat}$$

と表すことができた。ただし、$T_{gca} = \mu_c + V_{gca}$ である。この尺度を回答者と企業、質問項目の各組み合わせについて分析していくには、以下の仮定が要請される。

仮定 4-3.1

(i) 任意の回答者と企業、質問項目について誤差Ⅰ E_{gcat} は独立である。

(ii) 任意の回答者と企業、質問項目について誤差Ⅱ $\pi^{(-)}_{gcat}$ は独立である。

(iii) 誤差Ⅰと誤差Ⅱは独立である。

この仮定は一般化可能性理論における仮定 2-2.1 や項目応答理論の仮定 2-5.1 に対応している。さらに、現実の分析における要請として次の条件も仮定しよう。

仮定 4-3.2

(i) 誤差Ⅰが従う正規分布の分散 σ^2_{gca} は回答者と企業、質問項目に依存しない。

(ii) 誤差Ⅱが従う分布関数 $\mathrm{H}_{gca}\left(\pi^{(-)}\right)$ は回答者と企業、質問項目に依存しない。

企業調査には回答者と質問項目（と企業）の組に繰り返しが得られないというデータ収集

上の制約があるため、これは実際の分析において不可欠な条件となる。分散分析を実際に適用するには各組における等分散性が必要となるが、この等分散性を検定するには各組に二個以上の観測値が必要なためである。ただし、この仮定も仮定４-２.２～３と同様にやや無理があるため、理解し易い質問項目を用意して同質な回答者を選ぶという企業調査の原則を尊重することが重要となる。

それでは、測定モデルを前提とした分析方法を具体的に説明していく。まず、回答者と企業、質問項目の組に対して定められる準同型写像$\mathbf{h}_{gca}$による経営特性の尺度T_{gca}を適当な母数ベクトル$\boldsymbol{\beta}$の線形結合；

$$T_{gca} = \mathbf{x}_{gca}\boldsymbol{\beta}$$

として表そう。仮定４-３.２から二つの誤差に回答者と企業、質問項目の添え字が不要となるため、このときの尺度は$Z_{gcat} = T_{gca} + \pi_t^{(-)} + E_t$と表すことができる。そこで、この尺度を適当な順で並べたベクトルにより；

$$\mathbf{Z} = \mathbf{T} + \Pi^{(-)} + \mathbf{E} = \mathbf{X}\boldsymbol{\beta} + \Pi^{(-)} + \mathbf{E}$$

と表すことにする。ただし、回答者と質問項目、企業と質問項目の組み合わせはすべて存在するが、回答者と企業の組み合わせは不完全であることに注意する。このとき、二つの誤差ベクトルが、仮定４-３.１から；

$$\mathbf{E} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I}), \Pi^{(-)} \sim \prod_{g,c,a} \mathrm{H}\left(\pi^{(-)}\right)$$

であることが分かる。

企業調査における相の効果はすべて母数$\boldsymbol{\beta}$に集約されるから、経営特性に関する仮説の多くは線形仮説$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi}$として表すことができる。誤差Ⅱ（尺度$Z$）の分布に正規分布を仮定すれば、これは回帰モデルにおける線形仮説の検定問題に帰着する。しかし、その検定には分散分析ではなく、汎検定問題の形式Ⅱとして受容リスクによる検定方法を適用すべきということは既に述べた通りである。そこで、この検定問題を汎検定問題として扱う理由について確認しておこう。

経営特性の状態ω_{gcat}に尺度Z_{gcat}が与えられる過程を考えれば、誤差Ⅱを用いた尺度の分解$Z_{gcat} = Z_{gca} + \pi_{gcat}^{(-)}$が極めて人為的な表現であることが分かる。この表現は、回答者

が測定した尺度Z_{gca}に誤差項$\pi_{gcat}^{(-)}$が加えられた値Z_{gcat}が観測されたことを意味するものではない。状態ω_{gcat}の尺度は測定時に想定した状態集合$\mathbf{\Omega}_{gcat} \in \mathbf{B}_{gca}$に依存して一意に定められないため、尺度を揃えるために分析者が導入した尺度がZ_{gca}であり、そのための補正項が誤差Ⅱ$\pi_{gcat}^{(-)}$である。誤差Ⅱ$\pi_{gcat}^{(-)}$に定義された確率とは、$Z_{gcat} - \pi_{gcat}^{(-)}$が基準とする尺度$Z_{gca}$である確率を意味している。この解釈は3-1節で定義した形式Ⅱの解釈に外ならない。そして、真の標本（真の尺度Z_{gca}）が未知という汎検定問題に、通常の検定方法を適用したときの問題も3-1節で説明した通りである。

このように、経営特性に関する線形仮説$\mathbf{H\beta} = \xi$の検定は、原問題を分散分析とした汎検定問題の形式Ⅱとして行うことが望ましい。この具体的な方法については第3章で詳しく説明した通りである。また、4-4節で議論するように、受容リスクによる検定方法は尺度の離散性とは無関係に適用できるため、Likert尺度を同じ分析枠組みの中で議論することが可能となる。このときの検定問題は汎検定問題の形式Ⅲとなる。

4-3-3 誤差Ⅱの分布形に関する一つの提案

これまでの議論は誤差Ⅰの正規性を仮定していた。これは、原問題として分散分析を想定しているからである。これに対し、受容リスクに関する議論は特定の分布形を仮定する必要がないため、誤差Ⅱが従う確率分布については言及してこなかった。しかし、実際に受容リスクを求めるには分布関数を特定化する必要があるので、議論の最後に誤差Ⅱが従う確率分布について一つの提案をしておこう。

この測定モデルは；

$$Z_{gcat} \left| \pi_t^{(-)} \sim \mathrm{N}\left(T_{gca} - \pi_t^{(-)}, \sigma^2\right), \pi_t^{(-)} \sim \mathrm{H}\left(\pi^{(-)}\right)\right.$$

と表現することができる。これは形式的に統計的決定問題のベイズ基準による定式化と一致する。このモデルの中で誤差Ⅱ$\pi^{(-)}$は未知母数と見なすことができる。そこで、この母数空間を区間$\mathbf{I}$の長さを7で割った値を基に；

$$(-u, u), u = |\mathbf{I}|/14$$

と定義する。これは、Likert 尺度による測定は 7 段階程度が限界という Symonds(1924)[3] の指摘を考慮した結果である。そこで、誤差Ⅱの分布を事前分布と考えて；

$$\pi_t^{(-)} \sim \mathrm{U}(-u,u)$$

と一様分布を仮定しよう。

上記の定式化において一様分布による事前分布は最も不利な分布(Least Favorable Distribution)となっている[4]。この分布によるベイズ決定関数はミニマックス決定関数と一致する。ミニマックス基準とは最悪の場合を想定して、その中で最も望ましい統計的決定を行うという非常に保守的な基準である。

誤差Ⅱの従う確率分布としてベイズ基準における最も不利な分布に該当する一様分布を提案するのは、積み重ねた仮定が現実を反映しない結論を導くこと（最悪の場合）を押し止める必要があるのではないかと筆者には思えるからである。もちろん、これは一つの提案に過ぎないから、分析者は適当な判断に基づき分布関数を仮定すればよい。

4-4 Likert 尺度が用いられた場合

前節で示した分析枠組みは尺度の連続性を前提としていたが、その内容は離散的な尺度の場合でもほとんど成立する。それは、離散的な尺度を表すモデルが 4-2 節で提案した測定モデルの派生型となっているためである。

具体的な議論の前に、いくつかの点を補足しておく。まず、本節では m 段階評定法を考察していくから、その尺度の定義域は区間 **I** ではなく；

- 自然数の集合 $\mathbf{S} = \{1,2,\ldots,m\}$

となる。この離散的な Likert 尺度 $s \in \mathbf{S}$ に対し、本来の意味で関心の対象は連続的な経営特性の状態である。そのため、分析では Likert 尺度から特定化された状態を表す間隔尺度を推測するという要請が生じてしまう。その完全な復元は不可能だから、ここでは対応する間隔尺度が存在する範囲を与えるという方法が提案される。このときの分析枠組みは汎検定問題の形式Ⅲとなる。連続的な尺度の場合は形式Ⅱであったから、これが尺度の違いによる分析上の相違点と言うことができる。

このように、本節では分析者が導入する間隔尺度とデータとしての Likert 尺度を平行して議論していかなければならない。これらを共に尺度と呼ぶのは紛らわしいので；

---

[3] Ramsay(1973)からの引用。Symonds,P.M.(1924),On the loss of reliability in ratings due to coarseness of the scale.*Journal of Experimental Psychology*,7,456-461.

[4] 竹村(1991,p*333*)

- 間隔尺度を「尺度」、Likert 尺度を「得点」

と呼ぶことを約束する。

4-4-1 Likert 尺度を表現する測定モデル

目印のない所に物差しを当てて目盛りを読むという比喩を用いると、これまでは連続的な目盛りのある物差しを前提としてきたが、多段階評定法とは目盛りが段階数しかない物差しによる測定と言うことができる。この測定では指定した点に最も近い目盛りを得点として読み取るはずである。これが多段階評定法で回答者に要求される測定行為のすべてである。以下、この考え方に沿って測定モデルを定式化していく。

経営特性の状態集合 $\mathbf{\Omega} \subset \overline{\mathbf{\Omega}}$ は無限集合だから、状態集合に得点の集合 $\mathbf{S}$ を対応させるということは、その部分集合に得点を対応させることに外ならない。そこで、回答者 a が企業 c の経営特性を質問項目 g に従い測定したときの状態集合 $\mathbf{\Omega}_{gcat} \in \mathbf{B}_{gca}$ を；

$$\mathbf{\Omega}_{gcat} = \mathbf{\Omega}_{gcat}^{1} \cup \mathbf{\Omega}_{gcat}^{2} \cup \cdots \cup \mathbf{\Omega}_{gcat}^{m}$$

と直和分解する。この各部分状態集合と特定化した経営特性の状態 $\omega_{gcat} \in \mathbf{\Omega}_{gcat}$ 対し；

$$\omega_{gcat} \in \mathbf{\Omega}_{gcat}^{j} \Leftrightarrow s_{gcat} = j; j = 1, \ldots, m$$

と得点が定義される。また、各部分状態集合は連結集合(Connected Set)であり、この分解は状態間の順序関係により矛盾なく定義されているものとする。

次に、経営特性の多段階評定法が意味を持つための条件を考えていく。多段階評定法は連続的な測定を前提に議論されるため、そのときの仮定；

(1) 状態集合 $\overline{\mathbf{\Omega}}$ は質問項目と回答者に依存しない
(2) 回答者間で $\overline{\mathbf{\Omega}}$ 上の主観的距離は定数倍を除いて等しい
(3) 回答者と質問項目、企業によらず任意の $\mathbf{\Omega} \in \mathbf{B}_{gca}$ の主観的距離 $\mathrm{d}(\mathbf{\Omega})$ は等しい

がそのまま要請される。この三つの仮定に加え、次の仮定が必要となる。

仮定 4-4.1

得点 $s \in \mathbf{S}$ について、回答者と質問項目、企業によらず $\mathrm{d}(\boldsymbol{\Omega}^{s}_{gcat})$ は等しい。

この仮定が成立しない限り得点に基づく意味のある比較研究は不可能であるが、前述した三つの仮定と同様に、その厳密な成立は疑わしい。そのため、仮定からの逸脱を小さくするような調査計画が必要となろう。それには調査に先立ち、各回答者に；

- 得点と経営特性の対応関係を明らかにしておく

ことが重要となる。例えば、7 段階評定で上位 5%を得点 7 にするという対応関係や、均等に得点を割り振るといったことである。この要請を言い換えると、各得点に対応した部分状態集合の主観的距離の比；

$$\mathrm{d}(\boldsymbol{\Omega}^{1}_{gcat}):\mathrm{d}(\boldsymbol{\Omega}^{2}_{gcat}):\cdots:\mathrm{d}(\boldsymbol{\Omega}^{m}_{gcat})$$

を事前に与えておくということになる。

さて、Likert 尺度による分析では回答者が特定化した経営特性の状態を間隔尺度で表現しなければならない。その尺度を観測された得点から復元することは不可能だが、間隔尺度の定義域を分析者が決めることで、その状態の間隔尺度が存在する範囲を特定化することは可能である。そこで、この定義域を；

$$\mathbf{I} = [b_0, b_m)$$

としよう。そうすれば、回答者は測定時に状態集合 $\boldsymbol{\Omega}_{gcat}$ を想定するから、間隔尺度を定義する準同型写像 h_{gcat} を一意に定めることが可能となる。そして、得点 s に対応した部分状態集合 $\boldsymbol{\Omega}^{s}_{gcat}$ の準同型写像 h_{gcat} による像を；

$$\mathrm{h}_{gcat}(\boldsymbol{\Omega}^{s}_{gcat}) = [b_{s-1}, b_s); s = 1, \ldots, m$$

と表すことにする。従って、特定化された状態 $\omega_{gcat} \in \boldsymbol{\Omega}_{gcat}$ の尺度を $Z_{gcat} = \mathrm{h}_{gcat}(\omega_{gcat})$ とすれば、前述した状態と得点との対応関係は；

$$\omega_{gcat} \in \Omega^j_{gcat} \Leftrightarrow Z_{gcat} \in [b_{j-1}, b_j) \Leftrightarrow s_{gcat} = j; j = 1, \ldots, m$$

と間隔尺度を用いて表現することが可能となる。もちろん、回答者が特定化したのは経営特性の状態 ω_{gcat} だけである。従って、このように定義された間隔尺度 Z_{gcat} は測定により得られた値ではないことを注意しておく。

ここまで来れば連続的な尺度を扱った4-2節の議論と同じである。*4-2-2*で説明した方法に従い、回答者と企業、質問項目に依存した尺度 Z_{gcat} を各組みで共通する尺度に変換していけばよい。この結果をまとめると次のようになる。

$$\begin{aligned} Z_{gcat} &= T_{gca} + \pi^{(-)}_t + E_t \in [b_{j-1}, b_j) \Leftrightarrow s_{gcat} = j; j = 1, \cdots, m \\ &E_t \sim \mathrm{N}(0, \sigma^2), \pi^{(-)}_t \sim \mathrm{H}\left(\pi^{(-)}\right) \end{aligned}$$

これが Likert 尺度を表現する測定モデルである。

また、正規性の仮定から；

$$\mathrm{P}(Z_{gcat} \in \mathbf{I}) = \sum_{j=1}^{m} \mathrm{P}(s_{gcat} = j) < 1$$

が必然的に成立する。しかし、*4-2-2*の最後で述べた理由から、この場合も現実の企業調査では $\mathrm{P}(Z_{gcat} \in \mathbf{I}) \fallingdotseq 1$ が成立すると考えられる。

4-4-2 Likert 尺度を基にした分析

ここでは企業調査で得られた Likert 尺度を本節で示した測定モデルを基に分析する一般的な方法を提案する。もちろん、*4-3-1*で設定した企業調査が前提である。

この Likert 尺度を表現する測定モデル；

$$\begin{aligned} Z_{gcat} &= T_{gca} + \pi^{(-)}_t + E_t \in [b_{j-1}, b_j) \Leftrightarrow s_{gcat} = j; j = 1, \cdots, m \\ &E_t \sim \mathrm{N}(0, \sigma^2), \pi^{(-)}_t \sim \mathrm{H}\left(\pi^{(-)}\right) \end{aligned}$$

は3-1節で示した形式Ⅲに外ならない。この形式に沿って分析を進めるには、尺度 Z の事前分布 G^* と各得点に対応した半開区間を決めなければならない。まず、得点 s を条件とした尺度 Z の事前分布に；

$$Z|s \sim \mathrm{U}[b_{s-1}, b_s)$$

と一様分布を提案しよう。この提案は*4-3-3*で誤差Ⅱの分布に一様分布を提案したのと同じ理由による。半開区間$[b_{j-1}, b_j)$は部分状態集合に関する次の性質；

$$\mathrm{d}(\boldsymbol{\Omega}^1_{gcat}):\mathrm{d}(\boldsymbol{\Omega}^2_{gcat}):\cdots:\mathrm{d}(\boldsymbol{\Omega}^m_{gcat}) = b_1 - b_0 : b_2 - b_1 : \cdots : b_m - b_{m-1}$$

を用いて求めればよい。この関係は距離が生成する関係系を保存するという準同型写像の持つ性質から明らかであろう。当然、仮定4-4.1に関して説明したように、左辺の比を事前に明らかにしておくことが必要となる。

また、半開区間$[b_{j-1}, b_j)$の特定化には離散的な得点を間隔尺度に変換する尺度構成法を用いることができる。この方法を適用しない理由については、経営特性の測定と分析に既存のテスト理論や尺度構成法を適用したときの問題を取り上げた次節で説明する。

4-5 補足：既存のテスト理論を適用したときの問題

これまで説明してきた経営特性の測定モデルは、テスト理論が想定する測定とは異なる測定を表現したものであった。比較研究の分析枠組みを考えたとき、その違いは検定問題を汎検定問題として扱うということに帰着した。この検定方法の違いを除けば、本章で提案した分析枠組みは連続的な尺度も離散的な得点の場合も一般化可能性理論の考え方に基づいたものである。ただし、離散的な得点の分析に連続的な尺度を対象とした一般化可能性理論を適用できるのは、比較研究が間隔尺度T_{gca}に関する仮説$\mathbf{H\beta} = \boldsymbol{\xi}$の検定を原問題とした汎検定問題として扱われるためである。

さて、Likert尺度を表す測定モデルでは、誤差Ⅰ Eと誤差Ⅱ $\pi^{(-)}$の確率分布を基に各得点に対応した確率を；

$$\mathrm{P}(s_{gcat} = j) = \mathrm{P}\left(Z_{gcat} = T_{gca} + \pi^{(-)}_t + E_t \in [b_{j-1}, b_j)\right)$$

と定義することができる。ここから、Likert尺度を表す測定モデルは2-5節で紹介した段階応答モデルとして扱えることが分かる。しかし、前節で提案した分析枠組みは、項目応答理論ではなく、一般化可能性理論に基づいたものである。そこで、通常の企業調査には項目応答理論を適用できない理由から説明していこう。

まず、経営特性の測定モデルを段階応答モデルと見なした場合の母数の捉え方について

確認しておく。*4-1-2*で述べたように、経営特性の測定では観測可能な経営特性の状態が測定対象となり、回答者の思考が測定装置となる。従って、項目応答理論における潜在特性には各企業の経営特性の尺度；

$$\mu_c = T_{gca} - V_{gca}$$

が該当し、それ以外の母数は回答者に関する母数を含めテスト母数となる。ただし、テスト母数として扱われる半開区間$[b_{j-1}, b_j)$の端点（カテゴリー境界値）は回答者と質問項目に依存しないことを仮定しておく[5]。

このように、段階応答モデルの下で、経営特性の比較研究は潜在特性（経営特性μ_c）に関する仮説検定となる。項目応答理論における実用的な検定方法は大標本を前提とした尤度比検定(Likelihood Ratio Test)とスコア検定(Scoring Test)、Wald の検定(Wald Test)のみである。これらの検定に用いられる統計量は共通の漸近分布を持ち、その成立は最尤推定量の持つ一致性に依存する。しかし、テスト母数と潜在特性の最尤推定量が一致性を持つ保証のないことは2-5節で説明した通りである。そして、この問題を解決するために導入された推定方法が周辺最尤法であった。経営特性を確率変数と考えて周辺最尤法を適用すればテスト母数の一致推定量が得られるため、この推定値から求めた経営特性により仮説検定を行えばよい。

しかし、業種間比較など経営特性に関する仮説検定を想定した場合、この論法の正当性は疑わしいと言わざるをえない。検定すべき経営特性の分布を検定前に仮定することはできないからである。実際、周辺最尤法を適用する背後には確率変数として扱われる母数が局外母数(Nuisance Parameter)という暗黙の前提がある。また、最尤法に基づく検定方法は汎検定関数の考え方を反映していないので、形式Ⅲを通常の検定問題として扱った場合の問題が解決できないことも注意しておく。

*4-3-1*で述べたように、企業調査には回答者と質問項目の組に繰り返しは得られないというデータ収集上の制約がある。この制約の下では、カテゴリー境界値が回答者と質問項目に依存しないという前記の仮定を外すことはできない。経営特性だけでなく境界値自体が付随母数[6]となり、これらを推定することが不可能になるからである。そこで、この仮定を認めた上で、半開区間$[b_{j-1}, b_j)$の特定化について補足しておく。

前節の最後に提案した方法は、部分状態集合間の主観的距離の比を基に各半開区間を求めるという非常に単純なものであった。半開区間が回答者と質問項目に依存しないという仮定をおいたとしても、この半開区間を実際に推定することができれば、それが最善の方法であろう。しかし、残念ながら、この方法が使えないことは既に述べた通りである。

これ以外の方法としては Likert 法や数量化Ⅲ類などが考えられる。まず、Likert 法と

---

[5] 仮定4-4.1

[6] 2-5参照。

は尺度Z_{gcat}が企業の母集団の中で正規分布に従うと仮定して、観測された得点の分布から半開区間を逆算するという方法である。この方法を適用しない理由は周辺最尤法を適用しない理由と全く同じであり、分析前に経営特性の分布を仮定できないからである。

さて、こうした考え方はデータに加える操作は最小限に止めたいという本論文の立場を反映したものである。これは、本論文が得られた得点（尺度）を基にした仮説の検証を想定して議論を進めているからである。この点を考慮すると、数量化III類のようなデータ解析手法の適用には問題があるのではないだろうか。こうした手法は内的整合性を高める最適（この場合は企業間分散を大きくし企業内分散を小さくする）な尺度構成法となっているからである。確かに、次章で提案する質問項目の選択方法も内的整合性を基準にしたものである。しかし、これは質問項目の取捨選択による非常に消極的な方法であり、積極的なデータの変換は一切行われない。少なくとも筆者には、仮説検定を目的とした分析で企業間の差を強調するような尺度構成が適切な方法とは思えない。

もちろん、こうした手法を否定しているのではない。分析の目的が経営特性に関する仮説を見出すことにあるのなら、データ解析的な尺度構成は極めて有効な方法である。この場合は、半開区間$[b_{j-1},b_j)$ではなく間隔尺度Z_{gcat}が特定化されるので、第３章で示した形式IIに基づいて分析を進めればよい。

5 質問項目の選択と業種間比較

第４章では経営特性に関する分析枠組みを提示した。前提となる測定モデルは一般化可能性理論を拡張したものであり、仮説検定は一般化可能性理論に対応する分散分析を原問題とした汎検定問題になることが示された。そこで、本章では具体例として質問項目の選択と業種間比較を取り上げる。

まず、5-1節で企業調査を一般化可能性理論の視点から記述した後で、質問項目の選択と業種間比較の方法を5-2節と5-3節で解説し、実証例を5-4節で紹介する。5-5節では分析に用いた分散分析表を説明し、最後の5-6節では因子分析を用いた項目選択の持つ問題点を指摘する。

5-1 企業調査と一般化可能性理論

5-1-1 議論の前提

ここでの議論も第４章で考察した経営特性の測定や企業調査を前提とするので、そこで説明した条件や性質を確認しておこう。まず、ここでも4-1節で述べた条件；

(1) 理論的には間隔尺度を定義することが可能
(2) 現時点では、客観的な指標は存在しない
(3) 企業の客観的な分類を可能にする状態は存在しない

を満たす経営特性を対象とする。この測定が行われる企業調査には4-3節で指摘した；

(1) 回答者が回答する企業は一社だけである
(2) 回答者と質問項目の各組み合わせに測定の繰り返しは得られない

という分析上の制約がある。さらに、測定モデルと分析枠組みが前提とする仮定の妥当性を確保するため；

(1) 回答者は当該経営特性の知識を持っている
(2) 経営特性の説明が理解し易く、内容的妥当性を持つ質問項目を用意する
(3) 同質な回答者を選ぶ
(4) 調査目的や想定する企業を明確にしておく

という要件が尊重されているものとする。また、2-3節で紹介した信頼度指数がテスト計画と目的に応じて定義されることから明らかなように、信頼度指数に基づく項目選択の方法は分析目的に依存する。そこで、ここで扱う企業調査の目的として；

- 企業毎に得られる尺度に異なる企業属性間で差があるか否かを検証する

を前提とする。本章では企業属性として業種の相を想定するが、このことにより5-3節で説明する業種間比較の方法が一般性を失うことはない。この方法を国籍や企業の格付けの相について適用しても全く問題はない。

5-1-2 一般化可能性理論による記述

4-2節で示した連続的な尺度を表す測定モデルは；

$$Z_{gcat} = T_{gca} + \pi_t^{(-)} + E_t, E_t \sim \mathrm{N}(0,\sigma^2), \pi_t^{(-)} \sim \mathrm{H}\left(\pi^{(-)}\right)$$

であり、m 段階 Likert 尺度の場合は4-4節で；

$$\begin{aligned} Z_{gcat} &= T_{gca} + \pi_t^{(-)} + E_t \in [b_{j-1}, b_j) \Leftrightarrow s_{gcat} = j; j = 1,\cdots,m \\ & E_t \sim \mathrm{N}(0,\sigma^2), \pi_t^{(-)} \sim \mathrm{H}\left(\pi^{(-)}\right) \end{aligned}$$

と定義された。そして、二つの測定モデルは共通の回帰モデル；

$$Y_{gcat} = T_{gca} + E_t, E_t \sim \mathrm{N}(0,\sigma^2)$$

を原問題とした汎検定問題である。連続的な尺度の場合は形式II、Likert 尺度の場合は形式IIIであることも前章で述べた通りである。汎検定問題の検定方法については第3章で解説してあるので、ここでは原問題に関して議論すれば十分であろう。従って、本章では尺度の形式を区別せず、未知の尺度 Y を用いて記述していく。

さて、この回帰モデルは特殊誤差（誤差I）に正規性が仮定された古典的テストモデルと考えることができる。実際、質問項目の選択と業種間比較の方法は第2章で説明した一般化可能性理論の考え方に基づいて導かれる。そこで、その方法を説明する前に、ここでは企業調査を一般化可能性理論におけるテスト計画として捉え直しておこう。

繰り返しはなく回答者は特定の企業についてのみ回答するという *5-1-1* で指摘した制約の下では、ここで対象とする企業調査は一般化可能性理論の用語で繰り返しのない質問項目×（回答者：企業）計画[1]と言うことができる。しかし、一企業当たりの回答者数が少ない場合、このテスト計画における分散成分の推定では、回答者効果を含む企業効果分散から回答者効果分散を分離しなければならないため、企業効果分散の不偏推定量が過小推定になりやすいという問題がある。そこで、ここでは回答者の相を無視し、企業調査を同

---

[1] *2-2-2*参照。

答者数の繰り返しがある（企業×質問項目）計画と扱うことにする。ただし、回答者の相を無視するには条件；

$$\mathrm{E}(Y_{gcat}) = T_{gc}$$

の成立が必要である。この条件は任意の尺度 Y が企業と質問項目の違いで説明できることを意味している。言い換えると、回答者に起因する変動は企業と質問項目による変動と比べて相対的に小さいということである。また、企業調査を（企業×質問項目）計画と捉えた場合、業種や国籍など企業属性の相が考慮されないが、これらの相を含めるべきでないことは *5-2-3* で説明される。

それでは、（企業×質問項目）計画における尺度を各相の効果に分解しよう。まず、このときの殊真値 $\mathrm{E}(Y_{gcat}) = T_{gc}$ を；

$$\begin{aligned}
&\mathrm{E}_g \mathrm{E}_c (T_{gc}) = \mu \\
&\quad \mathrm{E}_c (T_{gc}) = \mu_g \;\Rightarrow\; \nu_g = \mu_g - \mu \text{；質問項目効果} \\
&\quad \mathrm{E}_g (T_{gc}) = \mu_c \;\Rightarrow\; \nu_c = \mu_c - \mu \text{；企業効果}
\end{aligned}$$

と企業と質問項目の効果に分解する。これらの効果で説明できない部分を質問項目×企業の交互作用効果 ν_{gc} と表せば、未知の尺度 Y は；

$$Y_{gcat} = T_{gc} + E_t = \mu + \nu_g + \nu_c + \nu_{gc} + E_t$$

と表現することができる。そして、尺度 Y_{gcat} の全分散は；

$$\sigma^2_{gcat}(Y_{gcat}) = \sigma^2_g + \sigma^2_c + \sigma^2_{gc} + \sigma^2$$

と各効果に対応した分散成分の和に分解される。

また、企業調査を表す（企業×質問項目）計画が、2-2節で紹介した弱真値モデルが対象とする（受験者×項目）計画と同じ構造であることを指摘しておく。

5-2 質問項目の選択

5-2-1 要約

質問項目の選択という行為が、測定対象についての比較研究の形式に依存することは言うまでもない。こうした調査目的に応じて信頼度指数が定義され、具体的な項目選択の手続きが導かれるからである。そして、本論文では企業の経営特性を測定対象とし、その尺度による企業属性に関する比較研究を想定している。ここでは本節で提案する項目選択の考え方を要約しておく。

まず、質問項目の選択を行うときの一般的な原則は；

(1) 十分な質問項目の個数を確保する
(2) 信頼度指数が高くなるようにする

である。しかし、この二つの原則は互いに矛盾した関係となっている。なぜなら、信頼度指数を最大にする質問項目の組み合わせを選ぶと、少数個の質問項目しか残らない可能性が大きいからである。そこで、本論文では、この矛盾を調整するため；

- 選択された質問項目による仮説検定の精度を高める

という基準を導入した。この基準を考慮することで、3段階の項目選択の方法；

段階1：すべての質問項目を採用する
段階2：企業との交互作用効果の小さい質問項目を選ぶ
段階3：質問項目効果が等しい質問項目を選ぶ

が導かれる。各段階で要請される条件は次の段階へ進むための必要条件となっている。この意味では系統的な選択方法と言えるだろう。しかし、前述した原則間の矛盾が完全に解決されたわけではない。段階を進むにつれ項目数は確実に減少するが、それに伴う信頼度指数の向上は保証できないからである。このため、各段階では二つの原則を確認せざるをえない。こうした内容は*5-2-4*で詳しく説明される。

以上の議論から分かるように、二つの原則間の矛盾が項目選択を考える上での最大の問題となっている。これは、質問項目の選択という操作を；

- 質問項目の取捨選択により異質な項目を排除する操作

と捉えた結果である。これに対し、一般化可能性理論ではD-研究の一つとして；

- 指定した信頼度指数を達成するのに必要な項目数を確保する操作

と考えるため、二つの原則間に何の矛盾も生じない。比較研究を一般化可能性理論の考え方に沿って進めるにも関わらず、本論文が後者の考え方を取らないのは；

(1) 実施された企業調査に対して質問項目の選択が要求される
(2) 代替的な質問項目の作成が難しい

という企業調査の持つ性質のためである。これが質問項目の選択に項目数の最適化を適用できない理由である。これらの結果は*5-2-2*で議論される。

この項目選択の基準は企業の相を基に定義した信頼度指数となる。これは、経営特性の比較研究という本論文が想定した企業調査の目的から導かれる必然的な結果である。この点については*5-2-3*で議論する。

5-2-2 質問項目選択の原則

項目選択の原則が；

(1) 十分な質問項目の個数を確保する
(2) 信頼度指数が高くなるようにする

であることは言うまでもないだろう。まず、項目数の確保という原則(1)は、測定の信頼性は項目数が多いほど高くなるという主張に外ならない。これは第２章で繰り返し説明した通りである。そして、ここでの議論（尺度 Y に関する原問題）は一般化可能性理論を前提にしているので、2-3節で定義した信頼度指数が信頼性の指標となる。このとき、信頼度指数を高めるという原則(2)は、経験的妥当性の見地から理論的に保証される。

そこで、項目 g を固定した古典的テストモデルを考える。定理2-4.1によれば、経験的妥当性の指標である妥当性係数と特殊信頼性係数 $\boldsymbol{\Phi}$ の間には；

$$\rho_a(Y_{ga}, Z_a) \leq \sqrt{\boldsymbol{\Phi}}$$

が成立している。この不等式は任意の基準変数による妥当性係数が特殊信頼性係数の平方根を超せないことを示している。すなわち、高い特殊信頼性係数が妥当性を保証するための必要条件となっている。この定理は十分条件ではないが、内容的妥当性が満たされた質問項目の持つ特殊信頼性係数の高さが、その質問項目の妥当性を保証する重要な根拠であることに疑いはない。

この結果は一般のテスト計画に拡張することができる。母得点が受験者の相を基に定義されている場合を考えると、特殊信頼性係数とは受験者以外の相に起因する変動がゼロのときに達成される信頼度指数の上限である。ここから、信頼度指数を高くする質問項目を選ぶことで、妥当性を確保するための必要条件が間接的に満たされることが分かる。以上が信頼度指数を項目選択の基準として用いる理論的根拠である。

しかし、この信頼度指数を用いた項目選択については；

(1) 質問項目の取捨選択により異質な項目を排除する操作
(2) 指定した信頼度指数を達成するのに必要な項目数を確保する操作

という正反対の考え方がある。既に述べたように、本論文で提案する方法は最初の考え方に基づいている。一般化可能性理論が提唱する二番目の方法を採用しないのは、その考え方が企業調査にそぐわないためである。それには二つの理由がある。

第一は、質問項目の選択が適用される場面の違いである。一般化可能性理論で項目選択とはD-研究の一つであり、次回のテスト作成のために行われる。一方、企業調査では調査結果を分析するための要請だから、再調査を前提に項目選択を考えることはできない。二番目の理由は、質問項目とテストの作成し易さの違いである。文言の違いだけで同等な質問項目を多数用意することは難しいため、指定した項目数を確保するという方法は現実的でないからである。

しかし、異質な項目の排除という考え方を前提とすると、二つの項目選択の原則間に矛盾が生じてしまう。質問項目効果の標本分散（信頼度指数の分母の一部）が項目数の増加関数となっているため、質問項目の取捨選択により信頼度指数を最大化すると、少数個の質問項目しか残らない可能性が大きいからである。この矛盾を調整し、段階的な項目選択が可能となるように導入された基準が；

- 選択された質問項目による仮説検定の精度を高める

である。企業属性間の比較研究という前提を考えれば、この基準を導入する妥当性については問題ないであろう。

ところで、取捨選択による異質な項目の排除と一般化可能性理論が仮定する標本抽出の無作為性は矛盾している。このため、選択された質問項目は適当な条件を満たす部分母集団からの無作為標本として再定義しなければならない。このように考えると、項目選択の指標として実際に求める信頼度指数とは、その部分母集団における信頼度指数の推定値ということになる。

5-2-3 経営特性の測定と信頼度指数

信頼度指数が項目選択の指標となることは既に説明した。この信頼度指数が企業の相を基に定義されることも述べた通りである。そこで、2-3節の議論に従い定義式と推定方法を示した後で、この信頼度指数が採用される理由を説明していく。

まず、企業の相が比較研究の単位となるから、その母得点は企業効果と総平均の和；

$$\tau_c = \mu + \nu_c (= \mu_c)$$

と定義される。このときの絶対誤差は$\Delta_{gcat} = Y_{gcat} - \tau_c$だから、信頼度指数$\Phi$は；

$$\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\Delta)} = \frac{\sigma_c^2}{\sigma_{gcat}^2(Y_{gcat})} = \frac{\sigma_c^2}{\sigma_c^2+\left(\sigma_g^2+\sigma_{gc}^2+\sigma^2\right)}$$

となる。この信頼度指数は、回答者数の繰り返しがある質問項目×企業の二元配置分散分析を示した表5-2.1を用いて推定することができる。また、この表については5-5節で詳細に解説される。この表から；

$$\hat{\Phi} = \frac{\hat{\sigma}_c^2}{\hat{\sigma}_c^2 + \hat{\sigma}_{gc}^2(\Delta_{gc})}$$

$$\hat{\sigma}_c^2 = \frac{\text{Between Companies}}{G(C-1)A} - \frac{\text{Interaction}}{G(G-1)(C-1)A}$$

$$\hat{\sigma}_{gc}^2(\Delta_{gc}) = \frac{\text{Between Item} \times \text{Company Cells}}{GC-1} - \hat{\sigma}_c^2$$

と推定できることが分かる。

さて、信頼度指数を企業の相を基に定義するのは、企業毎に測定される経営特性の比較研究を想定して議論しているからである。この場合は、尺度の企業間分散が企業内分散より大きくなる質問項目ほど良い質問項目と言うことができる。ここで、企業間分散とは母得点分散$\sigma^2(\tau)$のことであり、企業内分散は同一企業内での回答者による変動として絶対誤差分散$\sigma^2(\Delta)$の大きさに反映する。これは、項目選択には信頼度指数Φを用いるべきという主張に外ならない。

以上の議論から、一般信頼性係数やα係数の適用が誤りであることも分かる。これらの指標は企業の相を無視して推定されるため、企業間変動は小さいが回答者間変動は大きい質問項目と、企業間変動は大きいが回答者間変動は小さい質問項目を区別できないからである。この場合は最初の項目を削除すべきであることは明らかであろう。

最後に、企業効果の分散を純粋な企業効果ではなく不特定の企業属性の効果を含む分散として定義する理由について説明しておく。これには次の三つの理由がある。まず、企業属性の相は企業の相をネストするため、純粋な企業効果分散を推定できないという操作上の理由である。第二の理由は母得点の特定化の困難さである。例えば、企業の相をネストする業種と国籍の相が互いにクロスしているように、通常の企業調査では複数の企業属性の相が錯綜している。さらに、これらの相に関する様々な仮説検定を目的としている。このような場合に、統一的な母得点の特定化は事実上無理であろう。三番目は尺度構成にデータ解析手法を適用しない4-5節で述べた理由と同じで、比較研究の公正さを保証するた

めである。そこで、業種間比較のため母得点を企業効果ではなく業種効果を基に定義した場合を例に考えてみよう。信頼度指数の分子は業種効果の分散となるから、項目選択では業種間で差が出るような質問項目が選ばれることになる。しかし、比較研究の目的が業種間における差の検証にあるのなら、その差の有無は検証すべき仮説のはずである。このように業種間で「差がある」と結論できるような質問項目を選ぶことは、分析として公正さに欠くと言わざるをえない。

5-2-4 質問項目の選択方法

本節で提案する項目選択の方法は異質な項目を排除する3段階のプロセス；

段階1：すべての質問項目を採用する
段階2：企業との交互作用効果の小さい質問項目を選ぶ
段階3：質問項目効果が等しい質問項目を選ぶ

からなる。そして、この方法は項目選択の原則；

(1) 十分な質問項目の個数を確保する
(2) 信頼度指数が高くなるようにする

に新しく導入した基準；

- 選択された質問項目による仮説検定の精度を高める

を照らして具体化した手続きである。そこで、この導出過程を説明していこう。
まず、質問項目の選択で用いられる信頼度指数の定義式；

$$\Phi = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_g^2 + \sigma_{gc}^2 + \sigma^2}$$

から、この値（推定値）を大きくする二つの必要条件；

(1) 質問項目×企業の交互作用効果の分散 σ_{gc}^2 の推定値を小さくする質問項目を選ぶ

(2) 質問項目効果の分散 σ_g^2 の推定値を小さくする質問項目を選ぶ

を導くことができる。ここで、5-1節で述べた尺度 Y_{gcat} の分解を用いると；

(1) $\sigma_{gc}^2 = 0 \Leftrightarrow \nu_{gc} = 0$

(2) $\sigma_g^2 = 0 \Leftrightarrow \nu_g = 0$

が成立する。ここから、分散に関する二つの条件が質問項目×企業の交互作用効果 ν_{gc} と質問項目効果 ν_g に関する条件と同値であることが分かる。

次に、二つの条件の優先順位について考えていく。選択後に行われる仮説検定の精度を事前に高めておきたい場合、その尺度では企業による効果を質問項目の影響から完全に分離できることが最初に期待される。これは二つの効果の交互作用効果がないことと同値である。さらに、質問項目による差がなければ、これらの項目は平行テスト[2]となり、一層望ましい状況となる。この直観的な意味での優先順位は統計学的に保証される。質問項目効果がゼロという二番目の仮説の検定は、交互作用効果がゼロという最初の条件の成立を前提とするからである。これが冒頭で示した段階 2 と段階 3 の条件である。

この二つの項目選択の段階は、信頼度指数を高めるという原則(2)に異質な項目の排除という考え方を適用して導いたものだから、この項目選択では少数個の質問項目しか採用さない可能性が大きい。従って、項目数の確保という原則(1)を考えた場合、対象とする質問項目が既に十分高い信頼度指数を持つのであれば、必ずしも項目選択を段階 2 以降へ進める必要のないことが分かる。これが項目選択における段階 1 である。

項目選択を異質な項目を排除する操作として捉えたとき、二つの原則が互いに矛盾した関係にあることは既に述べた通りである。当然、ここで提案した 3 段階の選択手続きは不完全なものである。正確に述べると、この方法は信頼度指数を高める必要条件を基に構成したため、信頼度指数の向上を保証できないということである。しかし、各段階の導出過程から明らかなように、ここで提案した方法により比較研究の目的や用意された質問項目の個数などを考慮した段階的な項目選択が可能となる。また、比較研究と項目選択の段階との関係については業種間比較を取り上げた次節で詳しく説明される。

5-2-5 段階 1：すべての質問項目を採用する

項目選択の最初の段階が、すべての質問項目を採用するという段階 1 である。これは項目数の確保という項目選択の原則(1)を反映させたものである。しかし、言葉通りにすべての質問項目を対象とするわけではない。実際には、質問項目が次の二つの条件を満たしていることが必要となる。第一は、既に十分高い信頼度指数を持つことである。この条件については説明するまでもないだろう。

二番目の条件は、経営特性は企業と質問項目の違いで説明できるという 5-1 節で述べた

---

[2] 2-2 節参照。交互作用効果がない場合は、これらの項目は本質的に τ 等価となる。

条件 $\mathrm{E}(Y_{gcat})=T_{gc}$ の成立である。この条件の成立しない尺度は経営特性の尺度として無意味だから、これは信頼度指数に関する条件に比べて重要な条件と言うことができる。この重要性に関する統計学的な側面は具体的な検定方法を示した後で説明する。

さて、この条件の成立は；

$$\text{モデル：}\mathrm{E}(Y_{gcat})=T_{gc},\ \text{帰無仮説：}T_{gc}=\mu$$

という検定問題として検証することができる。そして、この帰無仮説を棄却できることが段階1で要請される条件である。表5-2.1から、この検定問題は F 統計量；

$$\frac{R^2_{T_{gc}=\mu}-R^2_{T_{gc}}}{GC-1}\div\frac{R^2_{T_{gc}}}{GC(A-1)}=\frac{\dfrac{\text{Between Item}\times\text{Company Cells}}{GC-1}}{\dfrac{\text{Residual}}{GC(A-1)}}\sim\mathrm{F}(GC-1,GC(A-1))$$

により検定することができる。

さて、帰無仮説 $T_{gc}=\mu$ が棄却できなければ、企業と質問項目に関する分散分析を構成することができない。この場合は、業種間比較や国際間比較など、企業の相に関係する一切の仮説検定が不可能となる。ここから、条件 $\mathrm{E}(Y_{gcat})=T_{gc}$ の成立が極めて重要であることが分かる。

5-2-6 段階2：企業との交互作用効果の小さい質問項目を選ぶ

質問項目の選択として段階2以降まで進むのは次の二つの場合である。用意された質問項目による信頼度指数が小さい場合と、選択した質問項目による分析の精度を事前に高めておきたい場合である。この後者の目的を考えたときの弱い基準が、企業との交互作用効果の小さい質問項目を選ぶという段階2の条件である。

段階2の条件を統計的に検証するには、段階1の条件 $\mathrm{E}(Y_{gcat})=T_{gc}$ が成立していなければならない。このモデルの下で、質問項目×企業の交互作用効果 ν_{gc} がゼロという条件は質問項目効果と企業効果の加法モデル；

$$\mathrm{E}(Y_{gcat})=T_{gc}=\nu_g+\nu_c$$

を帰無仮説として検定することができる。この検定問題をまとめると；

$$\text{モデル：}\ \mathrm{E}(Y_{gcai}) = T_{gc},\ \text{帰無仮説：}\ T_{gc} = \nu_g + \nu_c$$

となる。この場合は、帰無仮説を受容することで、与えられた質問項目に対して段階2の条件を受け入れることができる。また、この加法モデルと一般化可能性理論における本来のモデル $T_{gc} = \mu + \nu_g + \nu_c$ が同値であることは5-5節で示される。

表5-2.1から、この仮説は F 統計量；

$$\frac{R^2_{T_{gc}=\nu_g+\nu_c} - R^2_{T_{gc}}}{(G-1)(C-1)} \div \frac{R^2_{T_{gc}}}{GC(A-1)} = \frac{\dfrac{\text{Interaction}}{(G-1)(C-1)}}{\dfrac{\text{Residual}}{GC(A-1)}} \sim \mathrm{F}((G-1)(C-1), GC(A-1))$$

により検定することができる。

これは G 個の質問項目すべてについての加法性の検定である。従って、各質問項目に対して項目選択を個別に行う場合は、一つずつ質問項目を外して満足のいく結果が得られるまで検定を続ければよい。また、段階3は段階2の条件 $T_{gc} = \nu_g + \nu_c$ の成立を前提としているため、満足のいく結果が得られない場合は段階1で項目選択は打ち切りとなる。

5-2-7 段階3：質問項目効果が等しい質問項目を選ぶ

分析者にとって、新しくデータが追加されたとき、追加後の信頼度指数の低下や採用する質問項目の変更は避けたいところである。このような状況は、日本企業のデータにアメリカ企業のデータが追加される場合など容易に想定することができる。こうした段階的なデータ収集における分析の安定性を確保するための条件が、質問項目効果の均質性という段階3の条件である。この条件を満たすことで分析の安定性が確保される理由は最後に説明する。また、項目選択を段階3まで進めた場合の仮説検定の精度が段階2で終了させた場合より高まるのは明らかであろう。

段階3の条件を検定するには段階2の条件 $T_{gc} = \nu_g + \nu_c$ の成立を前提とする。質問項目効果 ν_g が等しいという帰無仮説は、この加法モデルの下で；

$$T_{gc} = \nu_g + \nu_c, \nu_g = \nu \Leftrightarrow T_{gc} = \nu + \nu_c \Leftrightarrow T_{gc} = \nu_c$$

と表現できる。従って、段階３の検定問題は；

$$\text{モデル：} T_{gc} = \nu_g + \nu_c, \quad \text{帰無仮説：} T_{gc} = \nu_c$$

と表される。段階３まで項目選択を進めるには、この帰無仮説を受容する質問項目を選べばよい。この検定問題も表５-２.１を用いて；

$$\frac{R^2_{T_{gc}=\nu_c} - R^2_{T_{gc}=\nu_g+\nu_c}}{G-1} \div \frac{R^2_{T_{gc}=\nu_g+\nu_c}}{\text{自由度}} = \frac{R^2_{T_{gc}=\nu_c} - R^2_{T_{gc}=\nu_g+\nu_c}}{G-1} \div \frac{\left(R^2_{T_{gc}=\nu_g+\nu_c} - R^2_{T_{gc}}\right) + R^2_{T_{gc}}}{(G-1)(C-1)+GC(A-1)}$$

$$= \frac{\dfrac{\text{Between Items}}{G-1}}{\dfrac{\text{Interaction + Residual}}{(G-1)(C-1)+GC(A-1)}} \sim \mathrm{F}(G-1,(G-1)(C-1)+GC(A-1))$$

と検定することができる。

ところで、段階２で要請される加法性の検定では各質問項目の影響を表す指標を導くことができなかった。しかし、段階３の条件を検証する場合は、こうした指標を質問項目効果を基に定義することが可能である。そこで、質問項目効果から全質問項目効果の平均値を引いて相対化した値；

$$\nu_{g_j} - \frac{1}{G}\sum_{k=1}^{G}\nu_{g_k}\,; j = 1,\ldots,G$$

を各質問項目の影響を表す指標とする。質問項目効果を直接用いないのは、加法モデルにおいて個々の母数は推定可能ではないからである。

この指標の推定方法を示すため、５-５-２で定義する行列を用いる。ここでは加法モデル$T_{gc} = \nu_g + \nu_c$の成立が前提だから、説明変数行列は$\mathbf{X}_2$、母数ベクトルは$\boldsymbol{\beta}_2$となる。表記の繁雑さ避けるため、これらを単に$\mathbf{X},\boldsymbol{\beta}$と表そう。そこで、各質問項目に対応する係数ベクトルを；

$$\mathbf{h}_j = -\frac{1}{G}\left(\overbrace{1 \cdots (1-G) \cdots 1}^{G\text{個}}\Big|\mathbf{0}'_C\right); j = 1,\ldots,G$$

$$\uparrow j\text{番目}$$

と定義すれば、問題の指標は；

$$\mathbf{h}_j\boldsymbol{\beta}; j=1,\ldots,G \Leftrightarrow \mathbf{H}\boldsymbol{\beta}$$

と母数の線形関数として表すことができる。ただし、行列 $\mathbf{H}$ は行ベクトル $\mathbf{h}$ を縦に重ねた G 行の行列である。条件 $V(\mathbf{H}') \subseteq V(\mathbf{X}')$ は明らかに成立するから、1-2 節より、この線形関数 $\mathbf{H}\boldsymbol{\beta}$ は確かに推定可能である。従って、その値は；

$$\mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}$$

と推定すればよい。また、個別（第 j 質問項目）の質問項目効果を推定するには、行ベクトル $\mathbf{h}$ を$(G + C)$次元の第 j 単位行ベクトルとしなければならない。しかし、この場合は推定可能であるための条件 $V(\mathbf{h}') \subseteq V(\mathbf{X}')$ は成立しない。

最後に、質問項目効果の均質性が分析の安定性を保証する理由について説明する。段階的にデータが追加される場合、分析は手元のデータから逐次に進められる。データの追加による分析結果の変動は避けられないが、こうした修正は最小限にしたいというのが分析者の希望であろう。

そこで、日米での比較を例に考える。両国とも項目選択は段階 2 で終了し、5 つの共通する質問項目が採用されたとしよう。前述した質問項目効果の指標は；

	項目 1	項目 2	項目 3	項目 4	項目 5
日本	2	−0.5	−0.5	−0.5	−0.5
アメリカ	0.5	0.5	0.5	0.5	-2

であり、このときの信頼度指数も同程度の大きさであったとする。

これらの尺度による日米比較が意味を持つための条件は、二カ国のデータを合わせたときの信頼性が低下しないことである。この例では、最初と最後の質問項目で企業との交互作用効果が発生するため、その信頼度指数の推定値は確実に低下する。従って、項目選択を段階 2 で打ち切る場合でも項目 1 と 5 の削除は避けられない。

これに対し、両国とも項目選択を段階 3 まで進めていれば、日本では項目 2～5 が、アメリカでは項目 1～4 が採用されるはずである。この場合は両国で共通する項目 2～4 が日米比較の対象となる。既に質問項目効果の均質性が満たされているから、データを合わせて項目選択を進めても項目 2～4 の残る可能性は大きい。

以上の説明から分かるように、予め項目選択を段階 3 まで進めておくことで、採用する質問項目の変更を押さえることが可能となる。この例では、5 個から 3 個への項目数の減少を 4 個から 3 個への減少に押さえることができた。採用する質問項目の違いは比較研究の結果に大きく影響する。データの追加により分析を逐次に進めていくのであれば、項目選択を段階 3 まで進めておく方が無難であろう。

5-2-8 問題点

これまで説明してきた3段階の項目選択の手続きが持つ操作上の問題は*5-2-4*で述べた通りである。しかし、より重要なことは、この方法を導く前提；

- 比較研究で統計的に望ましい性質を持つ質問項目を選ぶ
- 用意された質問項目は妥当性を持つ

が現実の企業調査に合致するとは言い切れないことである。そこで、ここでは統計的利便性を優先した項目選択の持つ問題について考えていく。

さて、この方法が統計学的に最善であることは言うまでもない。しかし、項目選択の結果として残された同質かつ少数個の質問項目が、企業調査の目的と照らして望ましい尺度になっている保証はない。なぜなら、後述する妥当性の問題とも関係しているが、外された質問項目（異質な項目）が当該経営特性を測定していた可能性を否定できないからである。さらに、ある経営活動の意図を従業員が理解している程度を分析したい場合は、様々な回答のパターンにこそ特徴が現れているかもしれない。当然、同質な質問項目を選択しても適切な尺度を構成することはできない。

こうした問題の多くは妥当性の問題に帰着される。*5-2-2*で述べた異質な項目の排除という考え方の背後には、与えられた質問項目は妥当性を持つという暗黙の仮定があるからである。しかし、妥当性を厳密に検証することは難しく、信頼度指数を大きくする質問項目が異なる経営特性を測定している可能性を否定することはできない。

このように、集めた質問項目の妥当性に疑問が残る場合や、項目選択後に想定している分析の目的を特定化していない場合は、ここで提案した方法が必ずしも望ましい質問項目を選択する保証はない。従って、3段階の選択手続きを実際に適用する場合は、得られた結果をデータに照らして十分に検討するという慎重さが必要であろう。

5-3 業種間での比較

5-3-1 項目選択の手続きと仮説検定との関連

前節で提案した項目選択の手続きは信頼度指数の向上と項目数の確保という相反する条件を調整することで具体化された。そして、この調整は選択された項目を用いた仮説検定の精度を事前に高めておくという目的の下で行われた。そこで、具体的な業種間比較の方法を解説する前に、項目選択の段階を進めることで仮説検定の有効性が高められる統計学上の理由を説明しておく。

まず、採用する質問項目が段階1に基づいているのであれば尺度Yは；

$$\mathrm{E}(Y'_{gcat}) = T_{gc}$$

と質問項目と企業の交互作用効果を含むモデルとして扱わなければならない。段階 2 で打ち切られている場合は；

$$\mathrm{E}(Y_{gcat}) = T_{gc} = \nu_g + \nu_c$$

と加法モデルとなる。さらに、段階 3 の条件も満たせば、尺度 Y は；

$$\mathrm{E}(Y_{gcat}) = T_{gc} = \nu_c$$

と企業効果のみで表現することが可能となり、企業に関する一元配置分散分析を用いた分析が可能となる。

この項目選択の方法は異質な項目の排除という考え方に基づいているため、段階を進むにつれ、採用される項目数が減少していく可能性は高くなる。しかし、尺度のモデル化に必要な母数は確実に減少していくため；

- 統計的処理の単純化
- 仮説の解釈の容易さ
- 検出力の向上

を期待することができる。このうち、最初の二つは明らかだから、ここでは母数の減少に伴う検出力の向上について説明していく。

そこで、モデル $\mathbf{y} \sim \mathrm{N}(\mathbf{X\beta}, \sigma^2\mathbf{I}), \mathbf{y} \in \mathbf{R}^n, \mathbf{X}{:}\, n \times m, \mathrm{rank}\,\mathbf{X} = r$ と $\mathrm{V}(\mathbf{H}') \subseteq \mathrm{V}(\mathbf{X}')$ を満たす階数 k の $s \times m$ 型行列 $\mathbf{H}$ による線形仮説 $\mathbf{H\beta} = \xi$ の検定を考えよう。ここで：

$$\begin{aligned}
R_0^2 &= \min_{\beta}(\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta}) = \mathbf{y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}')\mathbf{y} \\
R_1^2 &= \min_{\mathbf{H\beta}=\xi}(\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta}) \\
&= R_0^2 + (\mathbf{H}\hat{\beta} - \xi)'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}(\mathbf{H}\hat{\beta} - \xi) \\
\mathbf{H}\hat{\beta} &= \mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}
\end{aligned}$$

とすれば、帰無仮説 $\mathbf{H\beta} = \xi$ が真であれば；

$$\frac{R_1^2 - R_0^2}{k} \div \frac{R_0^2}{n - r} \sim \mathrm{F}(k, n - r)$$

が成立する。この F 統計量による線形仮説の検定が第 1 章で説明した分散分析である。

この分散分析において、不要な母数をモデルから削除しても残差平方和 R_0^2 がほとんど変化しないことは明らかである。線形仮説の表現方法はモデルに依存するが、行列 $\mathbf{H}$ の階数 k も変化せず[3]、$R_1^2 - R_0^2$ の大きさも変わらないであろう。そして、母数の個数と説明変数行列 $\mathbf{X}$ の階数 r の増減は一致しているから、分母の自由度 $n-r$ も大きくなる。以上のことから、不要な母数の削除により F 統計量が大きくなることが分かる。F 統計量は帰無仮説からのズレを反映した値だから、これは分散分析の検出力が向上することを意味している。また、分母の自由度 $n-r$ が大きくなるほど F 分布の密度関数は原点方向に集中した形状となるため、棄却限界も小さい値となる。これは比較的小さな F 統計量の値でも帰無仮説を棄却できるということだから、この性質も不要な母数の削除による検出力向上を意味している。もちろん、項目数の減少は標本数の減少となるため、このときの検出力は低下する。しかし、不要な母数の削除による検出力の向上が F 統計量の値と F 分布の性質の双方から導かれるという事実は考慮すべきであろう。

ところで、質問項目と回答者に関する平均値を企業の経営特性を表す尺度とし、この尺度を業種毎に平均の異なる正規母集団からの標本と考えれば、一元配置分散分析による業種間比較が可能となる。簡便さだけでなく、離散的な得点を集計により連続変数で近似できるという特徴から実際に適用されることが多いが、この方法には；

(1) 標本数の減少による検出力の低下
(2) 相の違いを無視することによる検定の不正確さ

という二つの問題がある。この場合は企業数×質問項目数×回答者数だけ存在した標本が集計により企業数にまで減少してしまう。標本数の減少は検出力を低下させるから、これは非常に大きな情報の損失と言わざるをえない。

相に関する情報の無視という第二の問題については、C.R.Rao がその著書で実例として挙げた分散分析（3 人の観測者により異なる 3 つの地層から発掘された頭蓋骨の鼻の高さに関する分析）でのコメントを紹介しておく。

"Suppose we ignored the observers and examined stratum differences. The analysis would then be....., leading to the conclusion that stratum differences exist. A closer analysis reveals that observer differences are important, so some caution is necessary in combining the results of the three different investigators. The stratum differences are not very prominent when the differences due to observers are corrected for."[4]

この説明から分かるように、企業と質問項目の相が無視される簡便法では、検出力というより仮説検定の正確さ自体が低下してしまう。これに対し、本節で提案する業種間比較の方法では、尺度 Y を表す各モデルから明らかなように、統計的に必要な相の影響は常に母数として考慮される。

---

[3] 3 段階の項目選択に対応したモデルでは実際に変わらない。

[4] Rao,(1973,p*257*)

5-3-2 表記上の約束

*5-3-1*で要約したように；

- 得点を表す被説明変数ベクトル $\mathbf{y}$
- モデルを規定する説明変数行列 $\mathbf{X}$ と階数 r
- 母数ベクトル $\boldsymbol{\beta}$
- $\mathrm{V}(\mathbf{H}') \subseteq \mathrm{V}(\mathbf{X}')$ を満たす線形制約 $\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\xi}$ と係数行列 $\mathbf{H}$ の階数 k

が定義されれば、分散分析を構成することができる。従って、項目選択の各段階に応じた業種間比較の方法を説明するには、これらの行列と階数、条件 $\mathrm{V}(\mathbf{H}') \subseteq \mathrm{V}(\mathbf{X}')$ を示せば十分であることが分かる。

また、以下の議論では回答者数を A 人、質問項目数を G 個、企業数を C 社と表すことを約束する。そして、企業は M 業種から選ばれ、企業を表す添え字 c について；

$$c = \underbrace{1,\ldots,C_1}_{\text{業種1}},\underbrace{C_1+1,\ldots,C_2}_{\text{業種2}},\ldots\ldots,\underbrace{C_{M-1}+1,\ldots,C}_{\text{業種M}}$$

$$\text{業種1}: N_1 = C_1 \text{個}$$

$$\text{業種2}: N_2 = C_2 - C_1 \text{個}$$

$$\ldots\ldots$$

$$\text{業種M}: N_M = C - C_{M-1} \text{個}$$

という対応関係が成立しているものとする。

5-3-3 質問項目の選択が段階1で終了している場合

ここでは質問項目の選択が段階1で終了している場合を考える。このときの尺度は質問項目×企業の交互作用効果を含むモデル；

$$\mathrm{E}(Y_{gcat}) = T_{gc}$$

と定義される。*5-5-1*から、被説明変数ベクトルには $\mathbf{y}$ を、説明変数行列と母数ベクトルには $\mathbf{X} = \mathbf{X}_1, \boldsymbol{\beta} = \boldsymbol{\beta}_1$ を用いればよい。この場合は $\mathrm{V}(\mathbf{X}') = \mathbf{R}^{GC}$ だから、任意の線形仮説が検定可能となる。

このモデルの下で、業種間に差はないという帰無仮説は；

$$\frac{1}{N_1}\sum_{c=1}^{C_1}\sum_{g=1}^{G}T_{gc} = \frac{1}{N_2}\sum_{c=C_1+1}^{C_2}\sum_{g=1}^{G}T_{gc} = \cdots = \frac{1}{N_M}\sum_{c=C_{M-1}+1}^{C}\sum_{g=1}^{G}T_{gc}$$

と表される。これは企業毎の真の尺度T_{gc}の項目に関する合計（平均でもよい）を業種毎に平均した値が全業種で等しいという仮説を意味している。この仮説で質問項目の影響を除けないのは質問項目×企業の交互作用効果が存在しているためである。また、業種毎に平均を取るのは各業種で企業数が異なるためであり、業種効果だけで表せないのは業種の相が企業の相をネストしているためである。このように、質問項目に課せられた条件が弱い場合は非常に曖昧な意味の「業種間に差はない」という仮説となる。

さて、この仮説で質問項目 g に関する母数$T_{g1}\cdots T_{gC}$の係数行列を $\mathbf{Q}$ とすれば、この行列は次のように表すことができる。

$$\mathbf{Q} = \begin{array}{c} \begin{array}{cccc} \text{業種1} & \text{業種2} & \text{業種3} & \text{業種M} \end{array} \\ \left(\begin{array}{c|c|c|c|c} \frac{1}{N_1}\cdots\frac{1}{N_1} & -\frac{1}{N_2}\ \cdots\ -\frac{1}{N_2} & \mathbf{0}'_{N_3} & \cdots & \mathbf{0}'_{N_M} \\ \frac{1}{N_1}\cdots\frac{1}{N_1} & \mathbf{0}'_{N_2} & -\frac{1}{N_3}\ \cdots\ -\frac{1}{N_3} & \cdots & \mathbf{0}'_{N_M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{N_1}\cdots\frac{1}{N_1} & \mathbf{0}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & -\frac{1}{N_M}\ \cdots\ -\frac{1}{N_M} \end{array}\right) \end{array}$$

$$= \left(\begin{array}{c|c|c|c|c} \frac{1}{N_1}\mathbf{E}'_{N_1} & -\frac{1}{N_2}\mathbf{E}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & \mathbf{0}'_{N_M} \\ \frac{1}{N_1}\mathbf{E}'_{N_1} & \mathbf{0}'_{N_2} & -\frac{1}{N_3}\mathbf{E}'_{N_3} & \cdots & \mathbf{0}'_{N_M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{N_1}\mathbf{E}'_{N_1} & \mathbf{0}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & -\frac{1}{N_M}\mathbf{E}'_{N_M} \end{array}\right) : (M-1)\times C$$

ここで、第 1 行は$\frac{1}{N_1}\sum_{c=1}^{C_1}T_{gc} = \frac{1}{N_2}\sum_{c=C_1+1}^{C_2}T_{gc}$を、第 2 行は業種 1 と業種 3 について、最後の第$(M-1)$行は業種 1 と業種 M に関する等号を表している。従って、行列 $\mathbf{Q}$ は質問項目に依存していないから、係数行列 $\mathbf{H}$ は行列 $\mathbf{Q}$ を質問項目数だけ横に並べた行列；

$$\mathbf{H} = (\overbrace{\mathbf{Q}\cdots\mathbf{Q}}^{G}) : (M-1)\times GC$$

として表すことができる。ここから、定数ベクトルが$\boldsymbol{\xi} = \mathbf{0}_{M-1}$であることも分かる。係数行列$\mathbf{H}$の階数は；

$$\operatorname{rank}\mathbf{H} = \operatorname{rank}\mathbf{Q} = \operatorname{rank}\left(\overbrace{\left.\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ & & \ddots & & \\ 1 & 0 & 0 & \cdots & -1 \end{pmatrix}\right\} M-1}^{M}\right) = M-1$$

である。これらの設定の下で、F統計量は$\mathrm{F}(M-1, GCA-GC)$に従う。

5-3-4 質問項目の選択が段階2で終了している場合

質問項目が段階2の条件を満たす場合、尺度は質問項目×企業の交互作用効果が存在しない加法モデル；

$$\mathrm{E}(Y_{gcat}) = T_{gc} = \nu_g + \nu_c$$

で表現される。まず、被説明変数ベクトル$\mathbf{y}$は当然共通である。そして、このときの説明変数行列と母数ベクトルには、*5-5-3*で定義する$\mathbf{X} = \mathbf{X}_2$と$\boldsymbol{\beta} = \boldsymbol{\beta}_2$を用いればよい。行列$\mathbf{X}$の階数が$G + C - 1$であることも*5-5-3*で示される。

加法モデルの下で業種間で差がないという帰無仮説は；

$$\frac{1}{N_1}\sum_{k=1}^{C_1}\nu_{c_k} = \frac{1}{N_2}\sum_{k=C_1+1}^{C_2}\nu_{c_k} = \cdots = \frac{1}{N_M}\sum_{k=C_{M-1}+1}^{C}\nu_{c_k}$$

と企業効果を業種毎に平均した値が等しいという仮説となる。純粋な業種効果による仮説とならないのは、業種の相が企業の相をネストしているためである。それでも、この式は企業効果だけで表されているから、段階1で終了した場合に比べて解釈し易い仮説と言うことができる。

*5-3-3*の場合と同様に、係数行列$\mathbf{H}$は業種1と業種2についての等号を第1行に、業種1と業種3についての等号を第2行という具合に；

$$\mathbf{H}=\left(\mathbf{0}_{(M-1)\times G}\;\middle|\;\begin{matrix}\frac{1}{N_1}\mathbf{E}'_{N_1} & -\frac{1}{N_2}\mathbf{E}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & \mathbf{0}'_{N_M}\\ \frac{1}{N_1}\mathbf{E}'_{N_1} & \mathbf{0}'_{N_2} & -\frac{1}{N_3}\mathbf{E}'_{N_3} & \cdots & \mathbf{0}'_{N_M}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \frac{1}{N_1}\mathbf{E}'_{N_1} & \mathbf{0}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & -\frac{1}{N_M}\mathbf{E}'_{N_M}\end{matrix}\right);\ (M-1)\times(G+C)$$

と定義すればよい。この場合も、$\mathrm{rank}\,\mathbf{H}=M-1$ であり、$\xi=\mathbf{0}_{M-1}$ である。

最後に、線形仮説 $\mathbf{H}\beta=\xi$ が検定可能であることを示そう。まず、行列 $\mathbf{X}$ は *5-5-2* で定義する $GC\times(G+C)$ 型行列 $\mathbf{Z}=\left(\mathbf{I}_G\otimes\mathbf{E}_C\ \ \mathbf{E}_G\otimes\mathbf{I}_C\right)$ の各行を A 回ずつ繰り返した行列であるから、$\mathrm{V}(\mathbf{X}')=\mathrm{V}(\mathbf{Z}')$ は明らか。一方、行列 $\mathbf{H}$ の第 1 行は行列 $\mathbf{Z}$ の第 1 行～第 C_1 行の $1/N_1$ 倍を加えたベクトルから、第 C_1+1 行～第 C_2 行の $1/N_2$ 倍を加えたベクトルを引いて作ることができる。この操作を繰り返すことで、$\mathrm{V}(\mathbf{H}')\subseteq\mathrm{V}'(\mathbf{Z}')=\mathrm{V}(\mathbf{X}')$ であることが分かる。これは線形仮説 $\mathbf{H}\beta=\xi$ が検定可能である必要十分条件に外ならない。

以上から、分散比を表す F 統計量が項目選択を段階 1 で打ち切った場合と比べて分母の自由度が大きい $\mathrm{F}(M-1, GCA-G-C+1)$ に従うことが示された。

5-3-5 質問項目の選択が段階 3 で終了している場合

質問項目が段階 3 の条件で選択されると尺度は；

$$\mathrm{E}(Y_{gcai})=T_{gc}=\nu_c$$

と企業効果だけで表現され、仮説の表現や統計的処理はさらに簡潔となる。このモデルの説明変数行列と母数ベクトルは；

$$\mathbf{X}=\left(\left.\begin{matrix}\mathbf{E}_{GA}\cdots\ \mathbf{0}_{GA}\\ \ddots\\ \mathbf{0}_{GA}\cdots\ \mathbf{E}_{GA}\end{matrix}\right\}C\right)=\mathbf{I}_C\otimes\mathbf{E}_{GA}\ ;\ GCA\times C, \mathrm{rank}\,\mathbf{X}=C$$

$$\beta'=(\nu_{c_1}\cdots\nu_{c_C})$$

と定義される。$\mathrm{V}(\mathbf{X}')=\mathbf{R}^C$ だから、任意の線形仮説に対し条件 $\mathrm{V}(\mathbf{H}')\subseteq\mathrm{V}(\mathbf{X}')$ が保証される。被説明変数ベクトル $\mathbf{y}$ はこれまでと共通である。

業種間に差がないという帰無仮説は *5-3-4* の場合と共通で；

$$\frac{1}{N_1}\sum_{k=1}^{C_1} v_{c_k} = \frac{1}{N_2}\sum_{k=C_1+1}^{C_2} v_{c_k} = \cdots = \frac{1}{N_M}\sum_{k=C_{M-1}+1}^{C} v_{c_k}$$

と定式化される。その係数行列は質問項目効果に対応した左側の$(M-1)\times G$型ゼロ行列を取り除いた$(M-1)\times C$型行列となる。

$$\mathbf{H}=\begin{pmatrix} \frac{1}{N_1}\mathbf{E}'_{N_1} & -\frac{1}{N_2}\mathbf{E}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & \mathbf{0}'_{N_M} \\ \frac{1}{N_1}\mathbf{E}'_{N_1} & \mathbf{0}'_{N_2} & -\frac{1}{N_3}\mathbf{E}'_{N_3} & \cdots & \mathbf{0}'_{N_M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{N_1}\mathbf{E}'_{N_1} & \mathbf{0}'_{N_2} & \mathbf{0}'_{N_3} & \cdots & -\frac{1}{N_M}\mathbf{E}'_{N_M} \end{pmatrix}$$

$$\mathrm{rank}\,\mathbf{H} = M-1$$

と表される。当然、定数ベクトルは$\xi=\mathbf{0}_{M-1}$である。

質問項目の選択が段階3まで進むと、このときのF分布は段階2で打ち切った場合と比較して分母の自由度がさらに大きな$\mathrm{F}(M-1, GCA-C)$となる。

5-3-6 補足：部分仮説の検定

これまでの議論は全業種を同時に比較する仮説検定を対象としてきたが、実際の比較研究では一部の業種に関する部分比較の問題が当然予想される。そこで、この問題を本節の最後に補足しておく。また、一般性を失うことなく、対象とする部分比較をM業種のうち最初のK業種間での比較と仮定しておく。

さて、質問項目が選ばれた段階によらず、どの帰無仮説も各業種の効果を反映した母数の線形関数が全業種で等しいという形式で表された。そして、その係数行列$\mathbf{H}$の階数はすべて$M-1$であり、各行は業種1との等号関係を示すものであった。従って、選択された段階に応じて、係数表列の最初の$K-1$行を部分仮説の係数行列$\mathbf{H}$とすればよい。その階数はすべて$K-1$であり、定数ベクトルは$\xi=\mathbf{0}_{K-1}$となる。各場合におけるF統計量の自由度は次のようになる。

段階1 $\Rightarrow$ $\mathrm{F}(K-1, GCA-GC)$
段階2 $\Rightarrow$ $\mathrm{F}(K-1, GCA-G-C+1)$
段階3 $\Rightarrow$ $\mathrm{F}(K-1, GCA-C)$

5-4 数値例

5-4-1 データの説明

ここでは、5-2節と前節で示した質問項目の選択と業種間比較の方法を現実の企業調査に適用した結果を紹介しておく。分析で用いたデータは 1991 年に始められた企業調査からの抜粋である。この企業調査は製造業を対象としたもので、各企業は機械と電機、自動車の三業種から優良性と国籍を考慮して選ばれている。このうち、具体例として業種間比較を取り上げるため、本節では日本の優良企業に関するデータのみを用いた。

分析された経営特性とデータの構成は以下の通りである。

経営特性 1：製造戦略と事業戦略の適合性
　項目 1：我々は製造戦略を持ち、これを積極的に追求している。
　項目 2：製造投資案は事業戦略との整合性によって選別される。
　項目 3：我々の工場では、製造は事業戦略と歩調を合わせている。
　項目 4：製造部門側は事業戦略に無頓着である。
　項目 5：我々の事業戦略は製造上の意味合いがわかるようになっている。
機械業界：8 社，電機業界：11 社，自動車業界：8 社，計 27 社
各企業に回答者数は 3 人

経営特性 2：アクティビティ会計の採用度
　項目 1：間接費用を直接費用化してきた。
　項目 2：費用をその決定要因に帰属させる能力を持っている。
　項目 3：原価会計システムによって付加価値を生まない活動を排除している。
　項目 4：アクティビティ基準の原価計算を行っている。
機械業界：11 社，電機業界：8 社，自動車業界：6 社，計 25 社
各企業に回答者数は 3 人

経営特性 3：品質改善における供給者との協力度
　項目 1：工場の成功にとって設備供給業者と共に働くことは肝要なことである。
　項目 2：新工程技術の開発では我々は供給業者と密接に働く。
　項目 3：我々は新工程設備を既製品として買う。
　項目 4：我々は自前の工程設備を作らないが、その設計には強く影響する。
機械業界：9 社，電機業界：11 社，自動車業界：8 社，計 28 社
各企業に回答者数は 3 人

経営特性 4：外部取引者についての品質情報の把握度
　項目 1：供給業者の品質適格認定のためのシステムを持っている。
　項目 2：供給業者は原材料に関して指定された試験や検査の結果を証明する記録を我々に送らなければならない。

項目 3：購入を検討している部品に関する品質のデータを容易に入手できる。
項目 4：供給業者や独立の研究所が行った品質検査データを容易に利用できる。
項目 5：重要な部品については統計的工程管理のデータを供給業者に要求する。

機械業界：12 社，電機業界：10 社，自動車業界：7 社，計 29 社
各企業に回答者数は 3 人

これらの質問項目に対して 5 段階評定法が用いられた。ただし、否定的に質問されている経営特性 1 の項目 4 と経営特性 3 の項目 3 は、6 点から回答された得点を引いた値を当該質問項目の得点とした。また、経営特性により企業数が異なるのは、欠損値を含む企業と回答者数の異なる企業を分析から外したためである。

5-4-2 汎検定問題の定式化

ここで扱う尺度は 5 段階 Likert 尺度だから、測定モデルは；

$$
\begin{gathered}
Z_{gcat} = T_{gca} + \pi_t^{(-)} + E_t \in [b_{j-1}, b_j) \Leftrightarrow s_{gcat} = j; j = 1, \cdots, 5 \\
E_t \sim \mathrm{N}(0, \sigma^2), \pi_t^{(-)} \sim \mathrm{H}\left(\pi^{(-)}\right)
\end{gathered}
$$

と表される。この測定モデルを前提とした分析は、汎検定問題の形式Ⅲとして扱わなければならないから；

- 未知の尺度 Z の定義域 $\mathbf{I}$
- 得点 s に対応した尺度 Z の存在する半開区間 $[b_{s-1}, b_s)$
- Z が真の $T+E+\pi^{(-)}$ である事前分布 G^*
- Z から誤差 II $\pi^{(-)}$ を除いた真の尺度 $Y = T + E$ の事前分布 G

を特定化しなければならない。そこで、得られた得点に加える操作を最小限にするという考え方を尊重して、ここでは尺度 Z の定義域と得点 s に対応した半開区間を；

$$
\begin{gathered}
\mathbf{I} = [0.5, 5.5) \\
[b_{s-1}, b_s) = [s - 0.5, s + 0.5); s = 1, \cdots, 5
\end{gathered}
$$

とした。さらに、二種類の事前分布 G^* と H には一様分布を仮定した。

$$
\begin{gathered}
Z|s \sim \mathrm{U}[s - 0.5, s + 0.5) \\
\pi^{(-)} \sim \mathrm{U}(-5/14, 5/14)
\end{gathered}
$$

この仮定がミニマックス的な考え方を反映させた結果であり、最も弱い仮定であることは既に述べた通りである。この二つの分布関数から、候補値 w が真の尺度 $Y=T+E$ である事前分布 G を求めると次のようになる。この場合は密度関数を持つから、得点 s を代表値としたときの密度関数を $\mathrm{g}(w|s)$ とすれば；

$$\mathrm{g}(w|s)=\begin{cases}\frac{7}{5}(w-s)+\frac{6}{5}\ ;\ -\frac{6}{7}+s<w\leq-\frac{1}{7}+s\\ 1\ ;\ -\frac{1}{7}+s<w\leq\frac{1}{7}+s\\ -\frac{7}{5}(w-s)+\frac{6}{5}\ ;\ \frac{1}{7}+s<w\leq\frac{6}{7}+s\end{cases}$$

となる。これは、半開区間 $[s-0.5,s+0.5)$ の中央付近が一様で、両端から隣接する区間の一部にかけて確率が小さくなる台形型の密度関数である。ここから、すべての得点 s_{gcat} に関する候補値ベクトル $\mathbf{w}$ の定義域が；

$$\mathrm{D}(\mathbf{w}|\mathbf{s})=\prod_{g,c,a}\left[s_{gcat}-\frac{6}{7},s_{gcat}+\frac{6}{7}\right)$$

であることが分かる。

このように、汎検定問題を形式Ⅲから定義域 $\mathrm{D}(\mathbf{w}|\mathbf{s})$ と事前分布 $\mathrm{g}(\mathbf{w}|\mathbf{s})=\prod_{g,c,a}\mathrm{g}(w|s_{gcat})$ を基に定義される形式Ｉに変換することができた。従って、原問題は分散分析だから、有意水準 α の分散分析に対応した受容リスク；

$$\mathrm{R}_{\alpha}(\mathbf{s})=\int_{\mathrm{D}(\mathbf{w}|\mathbf{s})}\delta_{\alpha}(\mathbf{w})d\,\mathrm{G}(\mathbf{w}|\mathbf{s})$$

により汎検定関数を構成すればよい。ただし、この受容リスクは陽表的に求めることができないので、分布関数 G に従う乱数ベクトル $\mathbf{w}$ を 300 回発生させたモンテカルロ法による数値積分で算出した。

5-4-3 質問項目の選択

まず、5-2 節で提案した 3 段階の項目選択の手続きを確認しておこう。各段階で要請さ

段階１：「モデル：$\mathrm{E}(Y_{gcat}) = T_{gc}$，帰無仮説：$T_{gc} = \mu$」で帰無仮説を棄却

段階２：「モデル：$\mathrm{E}(Y_{gcat}) = T_{gc}$，帰無仮説：$T_{gc} = \nu_g + \nu_c$」で帰無仮説を受容

段階３：「モデル：$T_{gc} = \nu_g + \nu_c$，帰無仮説：$T_{gc} = \nu_c$」で帰無仮説を受容

である。そして、企業の相を基に定義された信頼度指数；

$$\Phi = \frac{\sigma_c^2}{\sigma_c^2 + \left(\sigma_g^2 + \sigma_{gc}^2 + \sigma^2\right)}$$

が質問項目の持つ信頼性の指標となる。

ところで、この項目選択の方法は信頼度指数を基に構成したものだから、その適用時における信頼度指数とは単なる信頼性の指標に過ぎず、この値が統計的決定の指標として用いられることはない。従って、信頼度指数の推定には通常の点推定で十分であり、汎検定問題の考え方を踏襲する必要はないだろう。$\mathrm{E}(w|s) = s$ であり、*4-2-3*で説明した推定方法は分布関数に依存しないから、この方法を得点 s に直接適用することで信頼度指数を推定すればよい。

この項目選択を *5-4-1*のデータに適用した結果を説明していこう。最初は全質問項目を対象とした段階１と段階２に関する検定結果である。原問題である分散分析の有意水準は比較的緩い 10%としている。

		経営特性 1	経営特性 2	経営特性 3	経営特性 4
段階 1	信頼度指数	0.2133	0.1449	0.0048	0.0766
	受容リスク	1	0.83	1	0.77
段階 2	受容リスク	0	0	0.89	0

この結果を基に次の決定をした。まず、段階２の条件を満たし信頼度指数が比較的高い経営特性１と２は段階３まで項目選択を進める。信頼度指数が小さく段階２の条件を満たさない経営特性３は各質問項目を順に外して段階２をもう一度確認する。段階２の条件は満たしているが信頼度指数の小さい経営特性４も同様にする。

まず、経営特性３と４に関する段階２の結果は以下の通りである[5]。

[5] 負の信頼度指数は企業効果分散の推定値が負になった結果である。この推定方法で推定値の非負性を保証できないことは２-３節で説明した。しかし、母企業効果分散が小さいことは間違いなく、この結果が、これ以後の分析結果に影響を与えることはない。

	外した項目	信頼度指数	受容リスク
経営特性 3	項目 1	−0.0081	0.98
	項目 2	−0.0160	0.94
	項目 3	0.0213	0.30
	項目 4	0.0482	0.08
	項目 3 と 4	0.1634	0
経営特性 4	項目 1	0.1008	0
	項目 2	0.0814	0
	項目 3	0.0741	0
	項目 4	0.0686	0
	項目 5	0.0635	0

ここから、経営特性 3 では項目 3 と 4 を、経営特性 4 は項目 1 を削除すれば信頼度指数が高くなることが分かる。経営特性 1 と 2 に加え、これらの経営特性にも段階 3 の条件を検定した結果が次の表である。また、既に求めた信頼度指数と合わせて、2-3 節で紹介した合成テストの信頼度指数 Φ_Z ；

$$\Phi_Z = \frac{\sigma_c^2}{\sigma_c^2 + \left(\dfrac{\sigma_g^2 + \sigma_{gc}^2 + \sigma^2}{G}\right)}$$

も載せておこう。こうした企業調査では、2-6 節で補足した α 係数を信頼性の指標として用いることが多いためである。考え方自体が異なるため比較することはできないが、共に合成テストの信頼性という意味では参考になるだろう。

	経営特性 1	経営特性 2	経営特性 3	経営特性 4
項目数	5	4	2	4
受容リスク	1	1	0	1
信頼度指数	0.2123	0.1449	0.1634	0.1008
Φ_Z	0.5741	0.4039	0.2809	0.3096

このように、項目選択を段階 3 まで進めることができるのは経営特性 3 だけである。残りの経営特性については段階 2 で打ち切りである。信頼性の基準としては慣例的に 0.7 以上の α 係数が必要とされているが、こうした指摘を受けるまでもなく、二つの信頼度指数は決して高いとは言えない。この程度の信頼性の下では、業種間比較のような複数企業間での比較が限界であり、企業を個別に比較することは避けた方が無難であろう。

5-4-4 業種間比較

それでは、各経営特性について業種間で差があるか否かを検定してみよう。ただし、この業種とは機械と電機、自動車の三業種である。まず、経営特性 1 と 2、4 は段階 2 で項目選択が終了しているので、尺度 $Y = T + E$ は加法モデル；

$$\mathrm{E}(Y_{gca}) = T_{gc} = \nu_g + \nu_c$$

と表すことができる。経営特性 3 は段階 3 の条件も満たしているので；

$$\mathrm{E}(Y_{gca}) = T_{gc} = \nu_c$$

と企業効果 ν_c だけのモデルとなる。従って、すべての経営特性について、業種間に差がないという帰無仮説は；

$$\frac{1}{N_1}\sum_{k=1}^{C_1}\nu_{c_k} = \frac{1}{N_2}\sum_{k=C_1+1}^{C_2}\nu_{c_k} = \frac{1}{N_3}\sum_{k=C_2+1}^{C}\nu_{c_k}$$

と企業効果だけで表現することが可能となる。

分散分析の有意水準は項目選択のときより厳しい 5%とした。また、汎検定問題による仮説検定と比較するため、同じ形式のモデルと仮説に対し、得点のみに基づく通常の分散分析も行われた。この結果は以下の通りである。

	経営特性 1	経営特性 2	経営特性 3	経営特性 4
受容リスク	1	0.81	0.01	0.18
F 値	18.3946	4.5439	0.9588	2.5181
P(*F*>*F* 値)	0	0.0115	0.3859	0.0819

3-1 節で定義した汎検定関数の棄却限界とは、誤って棄却を決定したときの損失 L_{01} の相対的な大きさ $L_{01}/(L_{01}+1)$ である。しかし、これまでの検定結果についても当てはまることだが、この表に関する限りは棄却限界を明確にしておく必要はないだろう。汎検定問題として業種間に差がないという帰無仮説を棄却できるのは、明らかに受容リスクが大きい経営特性 1 と 2 である。これらは通常の分散分析でも帰無仮説を有意水準 5%で棄却することができる。帰無仮説を棄却できないのは経営特性 3 と 4 であり、経営特性 3 は分散分析でも棄却することができない。

この表で注目すべき点は経営特性 4 についての結果である。分析者が観測された尺度を

比較可能にするため導入した尺度 Y で帰無仮説が棄却されていた確率が 0.18 に過ぎないという状況では、確かに帰無仮説を棄却することはできない。ところが、分散分析では棄却を決定してもおかしくない結果が得られている。離散的な得点の分散分析自体に無理があるため厳密な比較はできないが、仮説検定で第一種の過誤に重点を置く検定論の考え方を尊重するのであれば、この結果は得点に基づいた一回限りの検定に結論を委ねてしまう危険性を指摘していると言うことができる。

最後に、*5-3-1* で紹介した一元配置分散分析による簡便法を適用した結果も参考までに載せておこう。この方法では、企業毎に取った回答者と質問項目に関する平均値が経営特性を表す尺度となる。業種間で差はないという帰無仮説に対して行った一元配置分散分析の結果は以下の通りである。

		経営特性 1	経営特性 2	経営特性 3	経営特性 4
機械	平均	3.9417	3.2424	3.7778	3.7847
	標準偏差	0.3991	0.4908	0.4082	0.3932
電機	平均	4.3030	3.5729	3.8485	3.7083
	標準偏差	0.3740	0.3346	0.3761	0.3099
自動車	平均	4.2417	3.4861	3.8750	4.0476
	標準偏差	0.6089	0.4928	0.6487	0.1854
F 値		1.53	1.39	0.10	2.26
P(F>F 値)		0.2370	0.2708	0.9068	0.1148

業種間で差はないという帰無仮説を棄却できる経営特性は一つもない。しかし、各業種の平均値と標準偏差の大きさ、集計したことで標本が企業数にまで減少したという状況を考えれば、この結果は驚くに当たらないだろう。

5-5　補足１：分散分析表、表５-２-１の説明

これまでの議論は回答者数を繰り返しと見なした質問項目×企業の二元配置分散分析に基づいて進められた。この分散分析の結果が表５-２.１[6]である。そこで、本節では、この表の導出と構造について説明していく。最初に、この表の見方について次の二点を注意しておく。第一は、期待値 E(平方和÷自由度)が測定の試行 t（回答者 a と同値）に加え質問項目 g と企業 c についても取られていることである。この値を基に信頼度指数が推定されることは既に述べた通りである。第二は、添え字に“・”の付いた変数が対応する添え字についての標本平均を表していることである。

[6] この表は弱真値モデルに対応した表２-３.１と同じ形式である。

5-5-1 4～6行目：Between Item×Company Cells ～ Total

表5-2.1の4～6行目は、項目選択の段階1で要請される仮説検定；

$$\text{モデル：} \mathrm{E}(Y_{gcat}) = T_{gc},\ \text{帰無仮説：} T_{gc} = \mu$$

に対応している。まず、被説明変数ベクトル$\mathbf{y}$は次のように定義される。

$$\mathbf{y}'_{gc} = (y_{gc1} \cdots y_{gcA})$$
$$\mathbf{y}' = \left(\mathbf{y}'_{11} \cdots \mathbf{y}'_{1C} \mid \mathbf{y}'_{21} \cdots \mathbf{y}'_{2C} \mid \cdots \mid \mathbf{y}'_{G1} \cdots \mathbf{y}'_{GC}\right)\ ;\ GCA \times 1$$

モデル$\mathrm{E}(Y_{gcat}) = T_{gc}$に対応した説明変数行列$\mathbf{X}_1$と母数ベクトル$\boldsymbol{\beta}_1$は；

$$\mathbf{X}_1 = \left(\left. \begin{matrix} \mathbf{E}_A \cdots \mathbf{0}_A \\ \ddots \\ \mathbf{0}_A \cdots \mathbf{E}_A \end{matrix} \right\} GC \right) = \mathbf{I}_{GC} \otimes \mathbf{E}_A\ ;\ GCA \times GC, \mathrm{rank}\,\mathbf{X}_1 = GC$$
$$\boldsymbol{\beta}'_1 = \left(T_{11} \cdots T_{1C} \mid \cdots \mid T_{G1} \cdots T_{GC}\right)\ ;\ GC \times 1$$

と定義される。ただし、$\mathbf{0}_n$はn次元ゼロ列ベクトル、$\mathbf{E}_n$はすべての要素が1のn次元列ベクトル、演算子⊗はクロネッカー積[7]である。

この仮説$T_{gc} = \mu$を線形仮説$\mathbf{H}_1\boldsymbol{\beta}_1 = \boldsymbol{\xi}_1$と行列表現したとき、その係数行列$\mathbf{H}_1$と定数ベクトル$\boldsymbol{\xi}_1$は次のようになる。

$$\mathbf{H}_1 = \left(\left. \begin{matrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ & & \ddots & & \\ 1 & 0 & 0 & \cdots & -1 \end{matrix} \right\} GC-1 \right)\ ;\ (GC-1) \times GC, \mathrm{rank}\,\mathbf{H}_1 = GC-1$$
$$\boldsymbol{\xi}_1 = \mathbf{0}_{GC-1}$$

---

[7] 行列$\mathbf{A}{:}n \times m$と任意の行列$\mathbf{B}$に対し、$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1m}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & \cdots & a_{nm}\mathbf{B} \end{pmatrix}$と定義される。

5-5-2 3行目：Interaction

表の3行目は項目選択の段階2に対応した仮説検定；

$$\text{モデル：}\ \mathrm{E}(Y_{gcai}) = T_{gc}\text{，帰無仮説：}\ T_{gc} = \nu_g + \nu_c$$

における残差平方和を示している。ところで、5-1節で述べたように、本来の帰無仮説は $T_{gc} = \mu + \nu_g + \nu_c$ である。検定で総平均 μ を省略できることは、具体的な行列を定義した後で説明する。

ここでのモデルは *5-5-1* で説明したモデルと同じだから、そこで定義した被説明変数ベクトルと説明変数行列を用いればよい。帰無仮説 $T_{gc} = \nu_g + \nu_c$ は行列 $\mathbf{Z},\boldsymbol{\beta}_2$；

$$\boldsymbol{\beta}_2' = \left(\nu_{g_1}\cdots\nu_{g_G} \mid \nu_{c_1}\cdots\nu_{c_C}\right);\ (G+C)\times 1$$

$$\mathbf{Z} = \left(\begin{array}{c|c} \overbrace{1\,0\cdots 0}^{G} & \left.\begin{matrix}1\cdots 0\\ \ddots\\ 0\cdots 1\end{matrix}\right\}C \\ \cdots\cdots & \\ 1\,0\cdots 0 & \\ \hline \cdots & \cdots \\ \hline 0\,0\cdots 1 & 1\cdots 0 \\ \cdots\cdots & \ddots \\ 0\,0\cdots 1 & 0\cdots 1 \end{array}\right) = \left(\mathbf{I}_G \otimes \mathbf{E}_C\ \ \mathbf{E}_G \otimes \mathbf{I}_C\right);\ GC\times(G+C)$$

$$\operatorname{rank}\mathbf{Z} = G+C-1$$

により、$\boldsymbol{\beta}_1 = \mathbf{Z}\boldsymbol{\beta}_2$ と表すことができる。しかし、ベクトル $\boldsymbol{\beta}_2$ は未知だから；

$$\boldsymbol{\beta}_1 = \mathbf{Z}\boldsymbol{\beta}_2 \ \Leftrightarrow\ \boldsymbol{\beta}_2 = \mathbf{Z}^{-}\boldsymbol{\beta}_1 \ \Leftrightarrow\ (\mathbf{I}_{GC\times GC} - \mathbf{Z}\mathbf{Z}^{-})\boldsymbol{\beta}_1 = \mathbf{0}$$

において、$\mathbf{H}_2 = \mathbf{I}_{GC\times GC} - \mathbf{Z}\mathbf{Z}^{-}, \boldsymbol{\xi}_2 = \mathbf{0}_{GC}$ とおくことで、この仮説を $\mathbf{H}_2\boldsymbol{\beta}_1 = \boldsymbol{\xi}_2$ と、係数行列と定数ベクトルが既知の通常の線形仮説として表せばよい。この係数行列 $\mathbf{H}_2$ の階数は一般逆行列の性質から；

$$\operatorname{rank}\mathbf{H}_2 = \operatorname{rank}(\mathbf{I}_{GC\times GC} - \mathbf{Z}\mathbf{Z}^{-}) = GC - \operatorname{rank}\mathbf{Z} = (G-1)(C-1)$$

となる。また、$\mathrm{V}(\mathbf{X}') = \mathbf{R}^{GC}$ だから、仮説 $T_{gc} = \nu_g + \nu_c$ は検定可能である。

ここで扱った仮説 $T_{gc} = \nu_g + \nu_c$ と本来の仮説 $T_{gc} = \mu + \nu_g + \nu_c$ が同値であることを最後に説明しておく。まず、仮説 $T_{gc} = \nu_g + \nu_c$ は $\boldsymbol{\beta}_1 = \mathbf{Z}\boldsymbol{\beta}_2$ と表される。これは母数ベクトル $\boldsymbol{\beta}_1$ に関する条件 $\boldsymbol{\beta}_1 \in \mathrm{V}(\mathbf{Z})$ を意味している。次に、$\boldsymbol{\beta}^* = \begin{pmatrix} \mu \\ \boldsymbol{\beta}_2 \end{pmatrix}, \mathbf{Z}^* = \begin{pmatrix} \mathbf{E}_{GC} & \mathbf{Z} \end{pmatrix}$ と定義すれば、総平均 μ を加えた仮説 $T_{gc} = \mu + \nu_g + \nu_c$ は $\boldsymbol{\beta}_1 = \mathbf{Z}^*\boldsymbol{\beta}^* \Leftrightarrow \boldsymbol{\beta}_1 \in \mathrm{V}(\mathbf{Z}^*)$ と表すことができる。このとき、$\mathrm{V}(\mathbf{Z}^*) = \mathrm{V}(\mathbf{E}_{GC}\ \mathbf{Z}) = \mathrm{V}(\mathbf{Z})$ が成立する。これは二つの仮説が同値であることに外ならない。

5-5-3 1～2 行目 : Between Tests ～ Between Companies

表の 1 行目は項目選択の段階 3 に関する検定問題 ;

$$\text{モデル : } T_{gc} = \nu_g + \nu_c,\ \text{帰無仮説 : } T_{gc} = \nu_c$$

の、2 行目は帰無仮説 $T_{gc} = \nu_g$ の検定に対応した残差平方和を示している。二つの仮説は同じ構造なので、ここでは最初の仮説について説明する。

まず、加法モデル $T_{gc} = \nu_g + \nu_c$ の下での説明変数行列 $\mathbf{X}_2$ は次のようになる。

$$\mathbf{X}_2 = \left(\begin{array}{c|c} \overbrace{\mathbf{E}_A\ \mathbf{0}_A \cdots \mathbf{0}_A}^{G} & \left.\begin{matrix} \mathbf{E}_A \cdots \mathbf{0}_A \\ \ddots \\ \mathbf{0}_A \cdots \mathbf{E}_A \end{matrix}\right\} C \\ \cdots\cdots & \\ \mathbf{E}_A\ \mathbf{0}_A \cdots \mathbf{0}_A & \\ \hline \cdots & \cdots \\ \hline \mathbf{0}_A\ \mathbf{0}_A \cdots \mathbf{E}_A & \mathbf{E}_A \cdots \mathbf{0}_A \\ \cdots\cdots & \ddots \\ \mathbf{0}_A\ \mathbf{0}_A \cdots \mathbf{E}_A & \mathbf{0}_A \cdots \mathbf{E}_A \end{array}\right) = \left(\left(\begin{matrix} \overbrace{\mathbf{E}_A \cdots \mathbf{0}_A}^{G} \\ \ddots \\ \mathbf{0}_A \cdots \mathbf{E}_A \end{matrix} \right) \otimes \mathbf{E}_C \quad \mathbf{E}_G \otimes \left(\begin{matrix} \overbrace{\mathbf{E}_A \cdots \mathbf{0}_A}^{C} \\ \ddots \\ \mathbf{0}_A \cdots \mathbf{E}_A \end{matrix} \right) \right)$$

$$= \big((\mathbf{I}_G \otimes \mathbf{E}_A) \otimes \mathbf{E}_C \quad \mathbf{E}_G \otimes (\mathbf{I}_C \otimes \mathbf{E}_A) \big) ;\ GCA \times (G + C)$$

これは*5-5-2*で定義した行列$\mathbf{Z}$の各行をA回繰り返して得られる行列だから、その階数は$G+C-1$である。母数ベクトルは*5-5-2*で定義した$\boldsymbol{\beta}_2$である。また、被説明変数ベクトル$\mathbf{y}$はこれまでと共通である。このとき、仮説$T_{gc}=\nu_c$は；

$$\mathbf{H}_3=\left(G-1\left\{\begin{bmatrix}1 & -1 & 0 & \cdots & 0\\ 1 & 0 & -1 & \cdots & 0\\ & & \ddots & & \\ 1 & 0 & 0 & \cdots & -1\end{bmatrix}\right. \mathbf{0}_{(G-1)\times C}\right);\ (G-1)\times(G+C), \operatorname{rank}\mathbf{H}_3=G-1$$

$$\boldsymbol{\xi}_3=\mathbf{0}_{G-1}$$

により$\mathbf{H}_3\boldsymbol{\beta}_2=\boldsymbol{\xi}_3$と表すことができる。また、条件$\mathrm{V}(\mathbf{H}_3')\subseteq\mathrm{V}(\mathbf{X}_3')$が成立することは容易に示される。

5-6 補足2：質問項目の選択と因子分析

おそらく、企業調査における質問項目の選択で現在最も普及している項目選択の方法は因子分析によるものである。この方法は合成尺度の特殊信頼性係数を高めるように構成した項目選択の方法と理論的に同値であり、5-2節で提案したのと同じ3段階の選択手続きを導くことができる。しかし、実際に因子分析を適用することは薦められない。以下、因子分析による項目選択の方法と問題点について説明していく。

因子分析モデルとは2-4節で説明した質問項目の構成的妥当性の考え方に基づいた尺度の表現である。因子分析では、構成的妥当性を持つ項目の尺度は共通する構成概念θにより説明されると考え、この構成概念を共通因子(Common Factor)と呼ぶ。そこで、共通因子の項目gに対する効果を因子負荷量(Factor Loading)λ_gとし、回答者aの項目gによる尺度Y_{ga}を；

$$Y_{ga}=\mu_g+\lambda_g\theta_a+E_{ga}$$

と表現する。これを因子分析モデルという。ただし、$\mathrm{E}_a\mathrm{E}(Y_{ga})=\mu_g$である。$E_{ga}$は特殊誤差のことであるが、因子分析では固有因子(Unique Factor)と呼ばれる。

さて、構成的妥当性を持たない異質な質問項目は相対的に小さな因子負荷量として検出できるはずである。ここから、因子負荷量の大きい項目を共通因子により説明される分散

を勘案して選択するという方法が導かれる。具体的な選択手続きを述べる前に、この方法と信頼性との関係から説明していく。

合成テストの特殊信頼性係数は；

$$\rho_Z^2 = \frac{\sigma_a^2(T_{Za})}{\sigma_a^2(Z_a)} = \frac{\sigma_a^2\left(\sum_{g=1}^{G} T_{ga}\right)}{\sigma_a^2\left(\sum_{g=1}^{G} Y_{ga}\right)} = 1 - \frac{\sum_{g=1}^{G}\sigma_a^2(E_{ga})}{\sum_{g=1}^{G}\sigma_a^2(Y_{ga}) + \sum\sum_{g \neq h}\sigma_a(Y_{ga}, Y_{ha})}$$

だから、互いに相関が高く特殊誤差分散の小さい質問項目を選ぶことで、この値を大きくすることができる。まず、固有因子と特殊誤差は同じ変数だから、特殊誤差分散の小さい質問項目を選ぶことは共通因子が説明する分散を大きくする質問項目を選ぶことと同値である。さらに、項目間の相関係数について関係式；

$$\rho_a(Y_{ga}, Y_{ha}) \propto \sigma_a(Y_{ga}, Y_{ha}) = \mathrm{E}(\lambda_g\theta_a + E_{ga})(\lambda_h\theta_a + E_{ha}) = \lambda_g\lambda_h\sigma_a^2(\theta_a)$$

が成立するから、互いに相関係数の大きな質問項目を選ぶことが因子負荷量の大きい質問項目を選ぶことと同値であることも分かる。

このように、因子分析に基づく項目選択と質問項目の信頼性に基づいた項目選択が理論的に等しいことが示された。従って、因子分析による方法も回答者毎ではなく、企業毎に適用される限りは正しいと言うことができる。以上の議論から；

段階 1：共通因子により説明される分散が大きければ質問項目の取捨選択は行わない
段階 2：因子負荷量が大きく均等な質問項目を選ぶ
段階 3：$\mathrm{E}_a\mathrm{E}(Y_{ga}) = \mu_g$ が等しい質問項目を選ぶ

が具体的な項目選択の方法となる。この方法は本論文で提案した 3 段階の選択手続きと同じものである。しかし、以下で述べる二つの理由から、質問項目の選択で因子分析を実際に適用することは薦められない。

第一は因子分析による項目選択が必要以上に厳しくなることである。前章で提案した測定モデルは一般化可能性理論と同様に尺度 Y を；

$$\mathrm{E}(Y_{gcat}) = \mu + \nu_g + \nu_c + \nu_{gc}$$

と表現することができた。これはモデルというより恒等的な分解を表している。この表現と因子分析のモデルを比較することで；

$$\nu_c + \nu_{gc} = \lambda_g \theta_c$$

の成立が必要となることが分かる。しかし、この関係が成立するのは、質問項目が項目選択の段階 2 の条件を完全に満たしている場合に限られる。この場合は質問項目×企業の交互作用効果が存在せず、因子負荷量も質問項目に依存しないため、$\nu_c = \lambda\theta_c$ が成立するからである。一方、条件が完全に満たされない場合、その不整合は因子分析の推定において固有因子にすべて吸収されてしまう。固有因子と特殊誤差は同じものだが、その分散が本論文で提案した測定モデルの誤差分散より大きく推定されることは明らかである。これは因子分析に基づく方法が分析者にとって不利であることを意味している。

第二は因子分析を用いる必要性がないためである。選択された質問項目を用いての分散分析は尺度を表現するモデルに依存するが、それは 5-2 節で提案した項目選択の結果として特定化されるからである。さらに、段階 3 を検証するには *5-2-6* で示した分散分析に頼らざるをえない。以上から、因子分析を敢えて項目選択の手法とする必要のないことが分かる。

表 5 - 2 - 1

		平方和	自由度	E(平方和÷自由度)
$R^2_{T_{gc}=\nu_c} - R^2_{T_{gc}=\nu_g+\nu_c}$	Between Items	$CA\sum_g y^2_{g\bullet\bullet} - GCAy^2_{\bullet\bullet\bullet}$	$G-1$	$CA\sigma^2_g + A\sigma^2_{gc} + \sigma^2$
$R^2_{T_{gc}=\nu_g} - R^2_{T_{gc}=\nu_g+\nu_c}$	Between Companies	$GA\sum_c y^2_{\bullet c\bullet} - GCAy^2_{\bullet\bullet\bullet}$	$C-1$	$GA\sigma^2_c + A\sigma^2_{gc} + \sigma^2$
$R^2_{T_{gc}=\nu_g+\nu_c} - R^2_{T_{gc}}$	Interaction	Between Item×Company Cells－Between Items－Between Companies	$(G-1)(C-1)$	$A\sigma^2_{gc} + \sigma^2$
$R^2_{T_{gc}=\mu} - R^2_{T_{gc}}$	Between Item×Company Cells	$A\sum_g\sum_c y^2_{gc\bullet} - GCAy^2_{\bullet\bullet\bullet}$	$GC-1$	$\sigma^2_g + \sigma^2_c + \sigma^2_{gc} + \sigma^2$
$R^2_{T_{gc}}$	Residual	Total－Between Item×Company Cells	$GC(A-1)$	σ^2
$R^2_{T_{gc}=\mu}$	Total	$\sum_g\sum_c\sum_a y^2_{gca} - GCAy^2_{\bullet\bullet\bullet}$	$GCA-1$	

参考文献

池田 央.(1980).調査と測定.新曜社.
池田 央.(1994).現代テスト理論.朝倉書店.
伊藤清三.(1963).ルベーグ積分入門.裳華房.
小山昭雄.(1994).経済数学教室 2.岩波書店.
小山昭雄.(1994).経済数学教室 4.岩波書店.
小山昭雄.(1995).経済数学教室 6.岩波書店.
佐和隆光.(1979).回帰分析.朝倉書店.
竹村彰通.(1991).現代数理統計学.創文社.
田中 豊,脇本和昌.(1983).多変量統計解析法.現代数学社.
田中良久.(1973).尺度構成. 心理学研究法 16.東京大学出版会.
西里静彦.(1975).応用心理尺度構成法.誠信書房.

Cliff,N.(1989).Ordinal consistency and ordinal true scores. *Psychometrika*,54,75-91.
Koch,W.R.(1983).Likert scaling using the graded response latent trait model. *Applied Psychological Measurement*,7,15-32.
Lehmann,E.L.(1983).*Theory of point estimation*. John Wiley.
Lehmann,E.L.(1986).*Testing Statistical Hypotheses*(2nd edition). John Wiley.
Lehmann,D.R. and Hulbert,J.(1972).Are three-point scales always good enough ?. *Journal of Marketing Research*,9,444-446.
Lord,F.M. and Novick,M.R.(1968).*Statistical theories of mental test scores*, Addison-Wesley.
Masters,G.N.(1985).A comparison of latent trait and latent class analyses of Likert-type data. *Psychometrika*,50,69-82.
Neyman,J. and Scott,E.L.(1948).Consistent estimates based on partially consistent observations. *Econometrica*,16,1-32.
Pfanzagl,J.(1971).*Theory of measurement*. John Wiley.
Ramsay,J.O.(1973).The effect of number of categories in rating scales on precision of estimation of scale values. *Psychometrika*,38,513-532
Rao,C.R.(1973).*Liner statistical inference and its applications*(2nd edition). John Wiley.
Samejima,F.(1969).Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*,No.17.
Samejima,F.(1973).Homogeneous case of the continuous response model. *Psychometrika*,38,203-219.
Sanders,P.F.,Theunissen,T.J.J.M. and Baas,S.M.(1991).Maximizing the coefficient of generalizability under the constraint of limited resources. *Psychometrika*,56,87-96.
Scheffé,H.(1959).*The analysis of variance*. John Wiley.
Schulman,R.S. and Haden,R.L.(1975).A test theory model for ordinal measurements. *Psychometrika* ,40,455-472.

Shavelson,R.J. and Webb,N.M.(1991).*Generalizability theory : A primer.* Sage Publications.

Woodward,J.A. and Joe,G.W.(1973).Maximizing the coefficient of generalizability in multi-facet decision studies. *Psychometrika*,**38**,173-181.