

第4章 統計的学習手法による就業形態の推定と予測¹

第1節 はじめに

この分析では、統計的学習 (statistical learning) の手法をもちいて、わずかなサンプル・サイズの観察データ (training set) から他の観察データ (test set) をどの程度正確に予測できるかという実験をしてみたい。その題材として、就業形態の選択を考えることにする。

より特定の社会的課題に耐える分析結果をえるためには、実際はもっと複雑な状況を考えなければならない。因果性の分析や内生性のバイアスを避けるためには、より精緻なデータが必要となる。ただ、ここでは単年度の横断面データに限った結果に注目していきたい。

いったい何歳までどれくらいの人たちが労働市場に参加して、就業することが可能なのか。そのために必要なのは多様な就業形態なのか、勤続・賃金の関係の修正なのか、健康状態の維持なのか、60歳を超えた年齢で職を転換することなのか、あるいはどれほどの水準の留保賃金率が必要なのか。あるいは、女性の労働力参加に関する問題で、(a)長時間労働が、育児あるいは結婚の妨げになるのではないかという説がある一方で、(b)育児時間が正社員であることを妨げている、結婚をすると正社員を続けることができなくなって労働市場から退出するのではないかと考えられている。同時に、(c)パートタイム労働者はできれば正社員になりたいという希望をもつ比率が高い。それぞれが事の一面をみているためであり、(a)と(b)は因果関係が逆であるとも考えることもできるが、別の共通した言及されていない要因、たとえば夫が家事・育児をしない・できない、結局は正社員としてもオファーされる賃金が低くて働く気がおきないことがあるからだと言説を立てることもできる。

これらの疑問に答えを与えるようなデータや分析をそろえることは難しい。因果の問題やバイアスの問題をクリアしたとしても、たとえば物価の急上昇が起きたり、大きな倒産が増加するなど状況が変化したりすると、平常時に得られた結果も変わる可能性もある。つまり、計測されたときの労働供給の弾力性と、予測するときの労働供給の弾力性が同じ値に保てるのかどうか保証がないからである。過去のデータで分析をしたときの条件付きの推定結果ということに変わりない。

第2節 最近の研究

これまで多くの労働供給を扱った実証的な研究では、パラメーターは固定していた。たとえば、Baker, Gruber, and Milligan (2008) ではカナダの Longitudinal Survey を使って育児ケア政策が悪影響をもたらすことを示しているが、サンプル内ではパラメーターは固定している。Blau and Kahn (2007) ではグループされたデータを使った労働供給の推定を行っ

¹ 本稿を作成するにあたって、独立行政法人 労働政策研究・研修機構 中野 諭氏をはじめとする労働力需給推計研究会マイクロ作業部会のメンバー、および厚生労働省雇用政策課藤井宏一氏から貴重なコメントをいただいた。ここに記して感謝したい。いうまでもなく、本稿に残る誤りは筆者の責任である。

ているが、これもパラメーターは固定である。これに対して、Del Boca, Flinn and Wiswall (2012)、Cuha, Heckman and Schennan (2010) は育児に対する補助が人的資本に与える影響を計測しているが一種のカーネル密度推定であり、Frölich (2006) は局所回帰を使ったパラメーター可変モデルである。Chetty (2012) は Frisch の労働供給パラメーターをノンパラメトリックな方法で推定しているともいえる。

第3節 推定方法

次のような局所尤度を利用するとパラメーターが変化するモデルを推定することになる。第一段階として就業選択を考え、第二段階として労働時間を選択するモデルが考えられる。今回は、労働時間選択のモデルは時間的な余裕がなかったため、行われていない。この逆としては、生活時間の制約から労働時間の選択範囲が狭まり、就業形態が決まるパターンが考えられる。質問票の構造上、逆パターンのモデルを設定することは難しいが、考えられないわけではない。

第一には線形判別関数を利用したもので、属性 x をもつ人がクラス j に分類される確率は、正規分布の密度関数 f_j と事前確率 π_j を使って

$$\Pr(G = j|X = x) = \frac{f_j(x)\pi_j}{\sum_{k=1}^J f_k(x)\pi_k}$$

となる。

サンプル・サイズを N 、クラスの数 J とする。線形判別分析では、結合密度 $P(X, G = j) = f(X; \mu_j, \Sigma)\pi_j$ にもとづく尤度関数を最大化する。つぎの式が線形判別関数である。

$$L_j(x) = x^T \Sigma^{-1} \mu_j - \frac{\mu_j^T \Sigma^{-1} \mu_j}{2} + \log \pi_j$$

クラス分けは、 $G(x) = \operatorname{argmax}_j L_j(x)$ となる。最も高い尤度を実現したクラスに割り振られる。

π_j 、 μ_j と Σ の推定は、標本平均や標本共分散行列である。すなわち、

$$\hat{\pi}_j = N_j / N$$

$$\hat{\mu}_j = \sum_{j \text{ に属する } i} x_i / N_j$$

$$\hat{\Sigma} = \sum_{j=1}^J \sum_{j \text{ に属する } i} (x_i - \hat{\mu}_j)(x_i - \hat{\mu}_j)^T / (N - J)$$

一般には属性 x の次元は高い値が考えられるが、分析を行う上では次元を低くする縮小化の選択を行う。これには縮小回帰の手法が応用可能である。この手法によってサンプル全体でのあてはまりを理論上は限りなく改善することができる。しかも、パラメーターの高次の大きさに応じてペナルティを導入することで、結果がジャンプしないように平滑化することもできる。多項ロジットモデルは線形判別関数と同じ形式になるが、共分散を使って尤度全体を推定することはしない。そのため線形判別関数よりもよく利用されている。

第二にさらに別の手法として、サポート・ベクター・マシンを試みた。この判別方法は、グループを分割する最適な一般的には非線形の境界線を見つけ、それにしたがって分類する方法である。 y_i を*i*番目のデータが y_i に分類されることを示す変数とする。2グループの場合には y_i は-1か1とする。データ行列を X とし p 個の説明変数があり、観察データは $i = 1, \dots, n$ とする。 y を分類する分離超平面は、超平面への p 次元の法線ベクトル W と定数 b を使って、 $W'X + b = 0$ のように決まる。最適な分離超平面は2つのグループの縁を形成する超平面からの距離が最大になるものである。 y_i が-1か1の場合、 $W'X + b = 1$ と $W'X + b = -1$ である。原点から分離超平面までの距離は、 $|b|/||W||$ で定義されるので、縁の超平面の距離は $2/||W||$ となる。これを最大化することは $\frac{||W||}{2} = \sqrt{w_1^2 + w_2^2 + \dots + w_p^2}$ を最小化することになる。完全に分離できる場合には、すべての $i = 1, \dots, n$ について、 $W'X_i + b \geq 1$ か $W'X_i + b \leq -1$ が成立するはずである。これをまとめて記述すると、 $y_i(W'X_i + b) \geq 1$ ($i = 1, \dots, n$)となる。ただし、完全には分離できないのでスラック変数の $\varepsilon_i \geq 0$ とペナルティ $C > 0$ を導入して、 $y_i(W'X_i + b) > 1 - \varepsilon_i$ を制約として、 $\frac{||W||}{2} + C \sum_{i=1}^n \varepsilon_i$ を最小化する。これは X について線形分類をおこなうが、一般にはカーネル関数を使って計算される非線形関数 $\phi(X)$ を利用する。このとき分離平面への法線ベクトル W は、 $\sum_{i=1}^n y_i \alpha_i \phi(X_i)$ と書ける。ここであるグループの集団でもっとも他のグループに近い観測ベクトル X_i をサポート・ベクトルという。このサポート・ベクトルあるいはスラック・ベクトル以外の点では $\alpha_i = 0$ とする。目的関数はカーネル関数 $K(X_i, X_j)$ を使って

$$\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j K(X_i, X_j) + C \sum_{i=1}^n \varepsilon_i$$

と書ける。制約は、 $y_i(W' \phi(X_i) + b) > 1 - \varepsilon_i$ である。カーネル関数の形としては、動径基底関数 $K(X_i, X_j) = \exp(-\gamma ||X_i - X_j||^2)$ などがある。詳細はHastieほか(2009)などを参照。

第4節『社会生活基本調査』で利用する項目

クラスとしては、調査票の「8 くだん仕事をしているか」と就業者については「11 勤めか自営かの別」と「12 勤務形態」の項目があげられる。

属性 x としては、「調査区番号(地域ダミー)」、「1 性(ダミー)」、「2 世帯主との関係(ダミー)」、「3 出生年月」、「4 配偶者の有無(ダミー)」、「5 教育(ダミー)」、さらに世帯主について:「27 住居の種類」、「28 自家用自動車の有無」、「29 世帯の年間収入(税込)」、「30 くだん世帯員以外の人から介護の手助けを受けていますか」、「31 不在者の有無」、がある。以上の変数を利用できる属性として、就業形態の選択モデルを考える。ただし、内生性の問題が避けられないため、単純に説明変数として採用することは難しい。

第5節 変数の妥当性の検討

就業選択の場合、外生的と考えられる世帯属性あるいは労働供給者の属性は非常に限られている。たとえば、地域の区分を選んだとしても、職を得るために引っ越しするということはしばしば行われる。この場合、大都市に移動することが多いとすると、大都市ダミーを使って都会では正規従業員比率が高いことを示したとしても、それは正規従業員の職を得るために大都市に移動しただけの理由である。住居にしても、持ち家に住んでいるから正規従業員なのか、正規従業員だから持ち家になれたのかどちらかわからない。公務員住宅の場合は、世帯のだれかが公務員でなければ住むことができないから、このような情報は就業選択を予測するモデルとしては利用できない。

結局、年齢が一番の独立変数であるということになる。教育年数も、若年層の就業選択の変数としては扱いが難しい。就学中であれば、就業選択の幅が狭まるが、たとえば大学4年生のように就職のために留年するということも考えられる。

同様に、ある意味では有配偶率も内生変数である。正規従業員が続けられるならば結婚はしなくても生活できる。逆の因果として、結婚をしていると無業を選ぶことができる。単身者の有業比率は高く正規従業員の比率も高い。晩婚化が原因なのか、正規従業員比率の上昇が原因で晩婚化したのか、決着をつけることはこのデータでは不可能である。

残念ながら健康状態も重要な説明変数と考えられるが、これは有業者しか調査されていない。そのために就業選択の変数として利用することができなかった。ただし、因果関係については、長時間労働によって健康状態が悪化し、常勤はできなくなるという可能性もないわけではない。就業行動は、生活の一部であるため完全な外生変数を見つけることは難しい。しかも、過去の行動に影響を受けることを考えると、大胆な作業仮説をつけて推計しなければパラメターの解釈をすることも難しい。

第6節 利用したデータについて

この節では『社会生活基本調査』の集計データの特徴を他の統計と比べて、推計結果の解釈の妥当性を示すことが目的である。『社会生活基本調査』は、本来、国民の行動パターンや時間の利用法についての調査である。ただし、サンプル調査としては各年ほぼ20万人前後について調査されており、ここで扱う就業形態、有業・無業に関する分析に耐えうる抽出数であると考えられる（1996年は約26万人、2001年18万6千人、2006・2011年17万8千人）。1996年とその他の年では、『国勢調査』でほぼ50世帯が含まれる調査区を決める方法が異なるが、第1次抽出では6500程度の調査区を抽出。第2次抽出では各調査区から12世帯（1996年は15世帯）を無作為に抽出する層化2段階抽出をおこなっている。変数の誤差率に関しては、ここで利用する就業形態別人口の誤差率は掲載されていない。

基本的な人口構成を調べてみたところ、図4-1の総務省『人口推計』にくらべやや若年層に厚く調査していることがわかる。図4-2では、集計用の推定乗率・比例補正用乗率を

かけずに、調査票そのものの枚数を考えることで、調査した対象の人口ピラミッドを描いている。『社会生活基本調査』は『人口推計』をもとに集計乗率が計算されているため、この加重和をとれば当然その年の『人口推計』と整合的になる。集計乗率をかけずにみることで、調査の特徴がわかり、その結果、若年層がベースとなる人口ピラミッドよりも数が多く調査されているという現象が起きている。統計局より提供されたデータでは1996年は年齢のみ、2001年は年齢と生年月、2006年以降は生年月のみがそれぞれ入力されている。また、10歳以上は各歳で記録があるが、10歳未満については調査が行われていない。

たとえば1996年と2001年では年齢で対象を区切るなので、10歳以上しか回答欄がない。つまり、同じ学年であっても調査月1996年10月以降に生まれていると調査対象とはならない。10歳と回答する場合は、1995年10月から1996年9月に生まれた人が対象となるはずである。

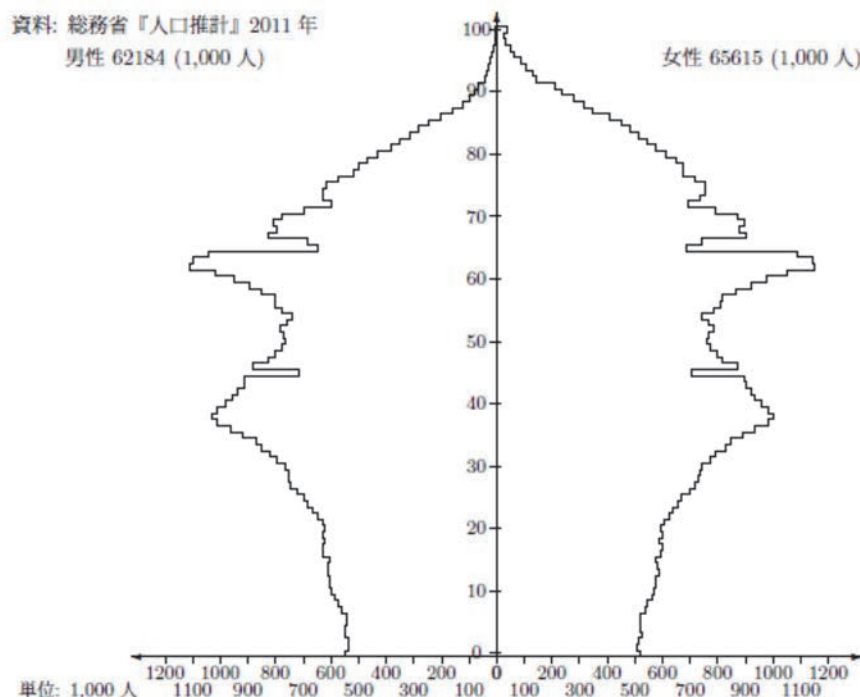
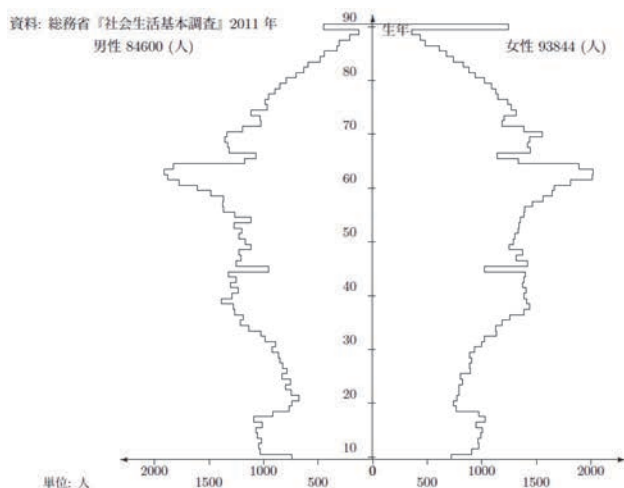


図4-1 総務省『人口推計』2011年の人口ピラミッド

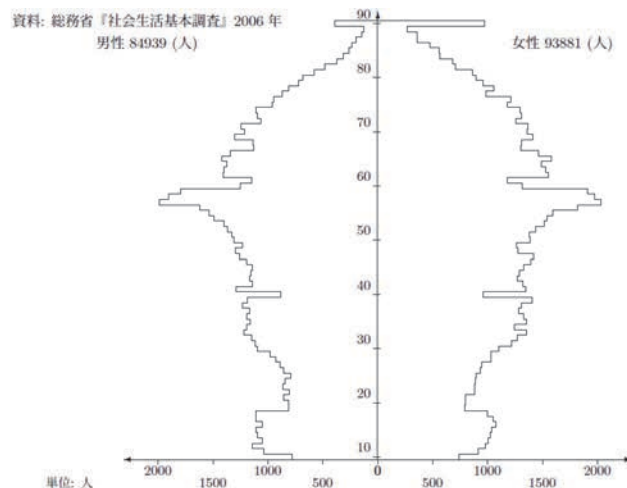
ここでは最後のシミュレーションで実験しているように、1996年のトレーニング・データを使って2011年の就業形態を予測するというをおこなっている。1996年に10歳であった人は15年後の2011年には25歳になっている。5歳刻みで推定をおこなっているため疑似コーホートとしては、1996年に10歳から14歳の方は、2011年には25歳から29歳層に対応することになる。このため主たる分析は15歳以上の就業選択行動であるが、10歳までの人口について傾向を調べている。

2001年では、調査月の10月に9歳であった1991年生まれの人は年齢を回答することが

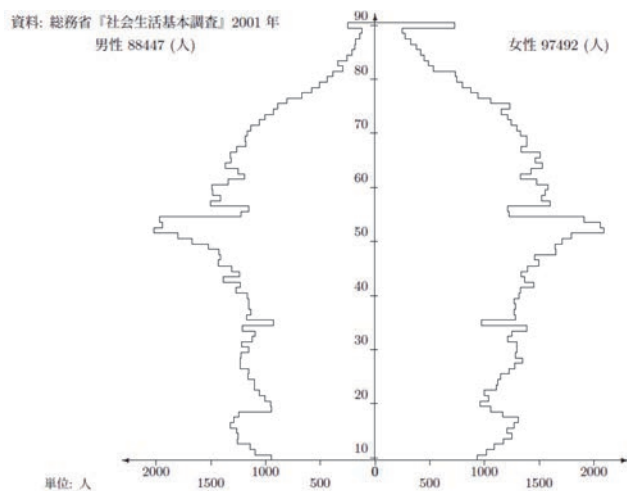
できない。ということで、1991年生まれの人は他の年齢よりも対象者が減るはずである。その他の年の観察は、生年が10年以前にはないため、10歳になるかどうかにかかわらず、2006年調査の場合には、1996年12月生まれから記載することになる。



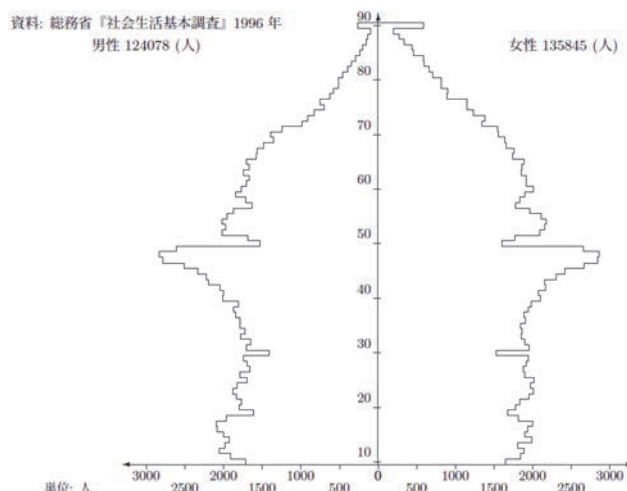
2011年 個票、2001年12月以前の生まれ



2006年 個票、1996年12月以前の生まれ



2001年 個票、1991年12月以前かつ10歳以上



1996年 個票、10歳以上

図 4-2 『社会生活基本調査』による人口ピラミッド(集計乗率は乗じていない。90歳以上は合計)

いずれにしても20代までの調査票の枚数は他の年齢階層に比べて多く調査されている。15歳以上については調査項目が増えるが、その影響で調査対象が減少しているというわけではない。人口ピラミッドの詳細な形状の違いについては、計測結果にどのように反映するかを考えてから再検討することにする。ここでは、20歳以下の標本については、それ以上の場合と別の推計を行うことで対応している。集計用乗率の年齢別平均値を計算してみると、2011年調査に比較して確かに若年の乗率は低くなっている。ここで記されている集計用乗率は、解説にあるように線形推定用乗率(世帯)と比例補正用乗率(地域・性・年齢)の積で、

基準人口である『人口推計』とのギャップが得られると考えられる。

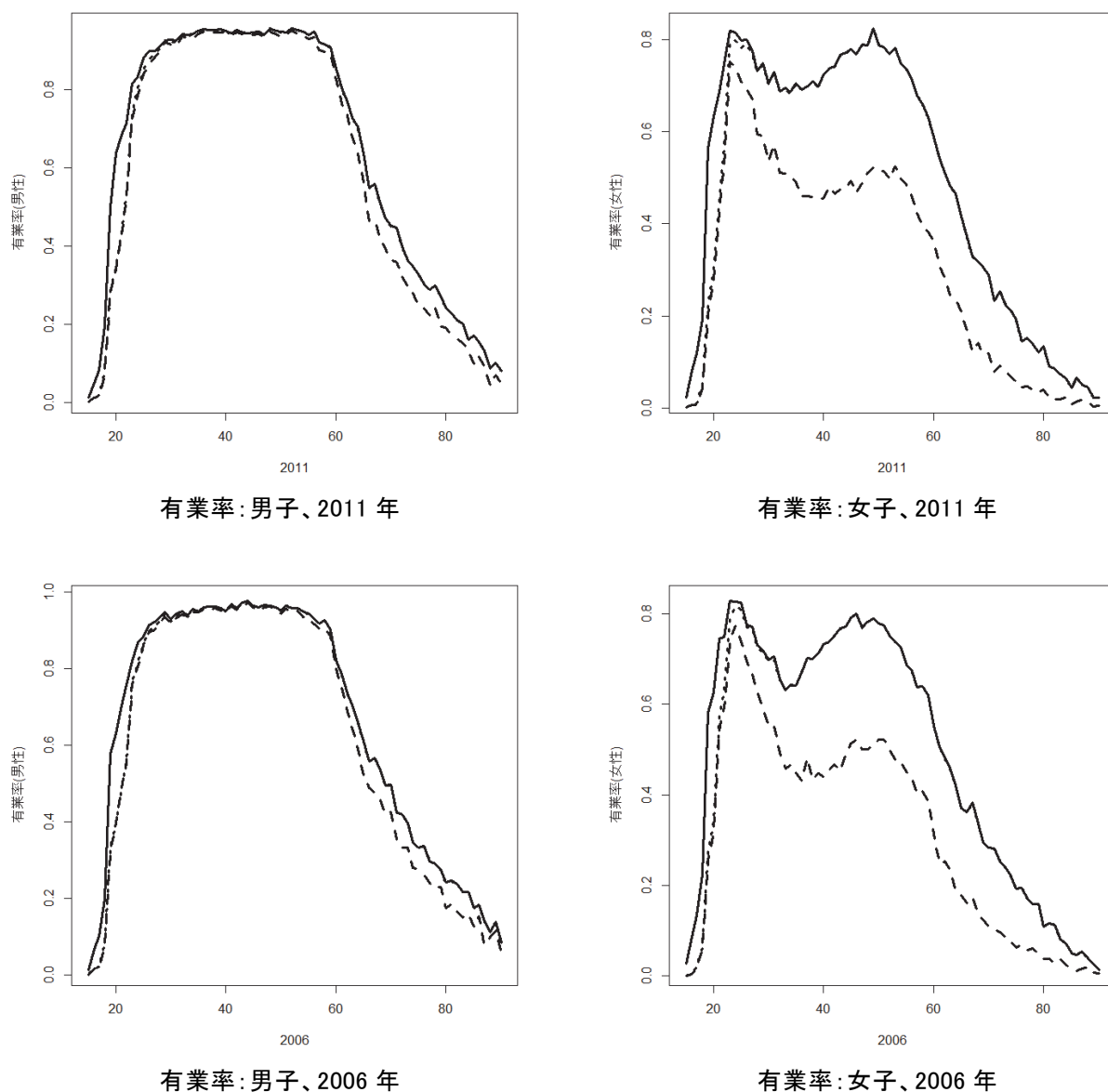


図4-3 『社会生活基本調査』による年齢別有業率(うち点線は正規従業員比率、
実線は家事などのかたわらに仕事の比率を加えたもの。90歳以上は集計)

10歳代の人口分布以外では、大きな食い違いは見られない。そのため、以下の分析で『社会生活基本調査』の集計乗率は使用しなくても一般的に議論することができると判断した。

年齢別有業率の推移を調べたところ、actual状態を調査している総務省『労働力調査』(就業率)に比較して、『社会生活基本調査』では、女性の30代、40代の落ち込みが大きいように観察される。この傾向はusual状態を調査した『就業構造基本調査』でも観察されていた現象である(早見, 1997)。30代の女性の有業率の上昇は2006年以降に顕著にみられるので、21世紀に入って増加してきているものとみられる。分類方法は、『社会生活基本調査』

の項目「ふだん仕事をしていますか」を使って、主に仕事、家事などのかわりに仕事、通学のかたわらに仕事、仕事をしていない・不詳の分類で、仕事をしていない・不詳を無業とした。有業のうち主に仕事の比率を点線で示している。家事などのかわりに仕事は破線で点線の上に加えた値を表示している。

1996年から2011年にかけて、就業形態の上では非正規雇用の増大ということが指摘されているが、図4-3と図4-4を見る限り、正規・非正規雇用ともに女性のM字型の谷の部分は上昇してきている。

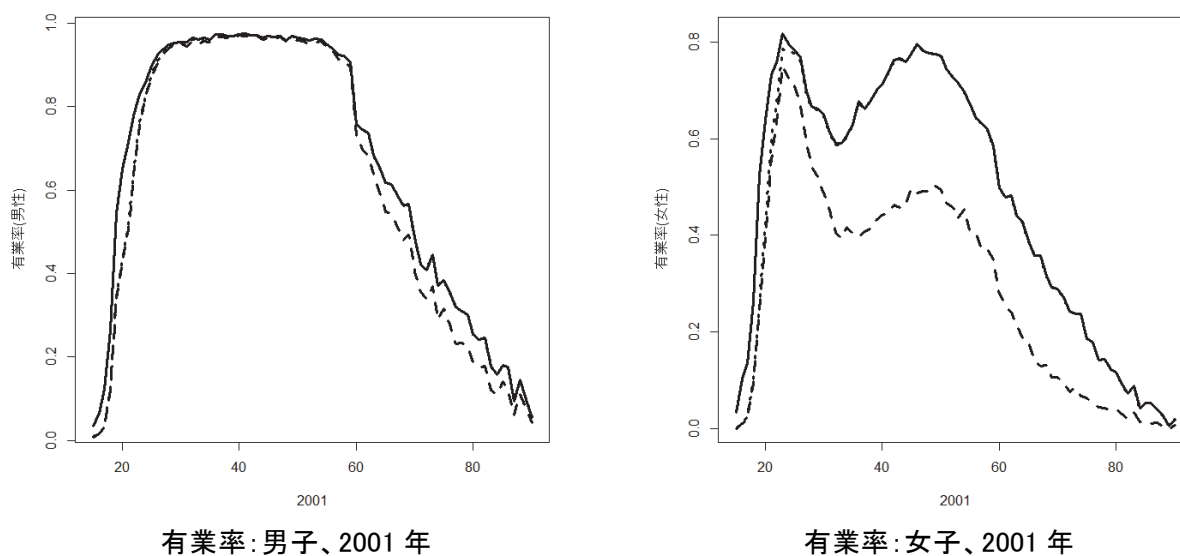


図4-4 『社会生活基本調査』による年齢別有業率(うち点線は正規従業員比率、
実線は家事などのかわりに仕事の比率を加えたもの。90歳以上は集計)

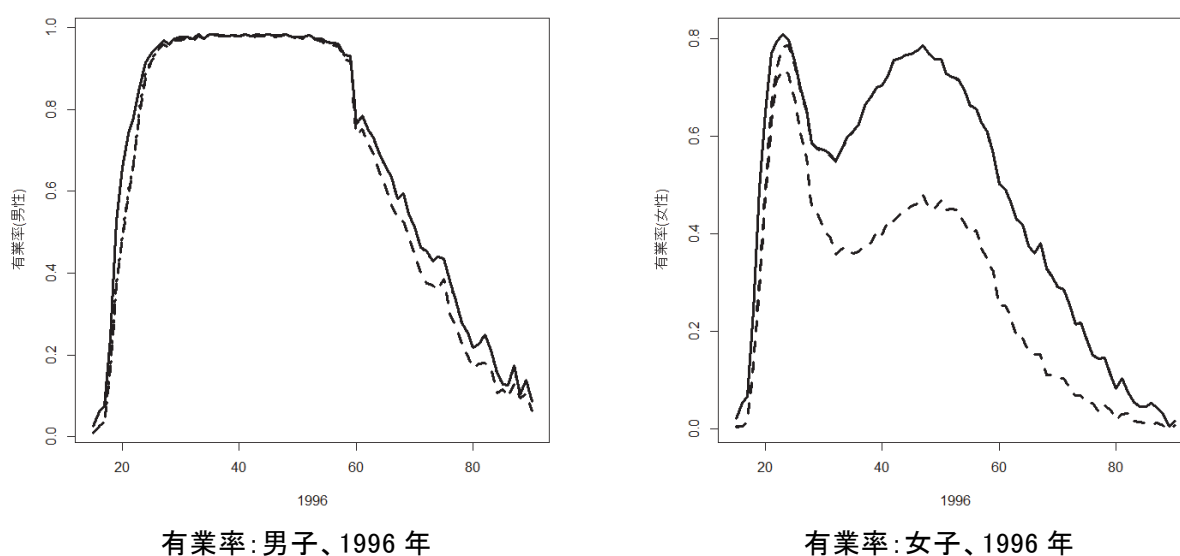


図4-4(続) 『社会生活基本調査』による年齢別有業率(うち点線は正規従業員比率、
実線は家事などのかわりに仕事の比率を加えたもの。90歳以上は集計)

既婚女性、単身女性の有業率の傾向も、2011年の『社会生活基本調査』でも観察されている(図4-5)。これらの傾向は、1977年の総務省『就業構造基本調査』以来の傾向として、過去35年間近くにわたって変化がない(早見 1997)。過去のM字型が解消するプロセスでは、有配偶率のシフト、晩婚化が主因であったと考えられる。M字がはっきりしていたころには、結婚年齢が短期間に集中していた。それが、次第に分散化していった。さらに、最近では以前ほど有配偶と単身の差はなくなってきているといえる。結婚年齢の分散化の影響とともに、高い有業率から低めの有業率へシフトすることは緩慢になりつつあることが示唆される。

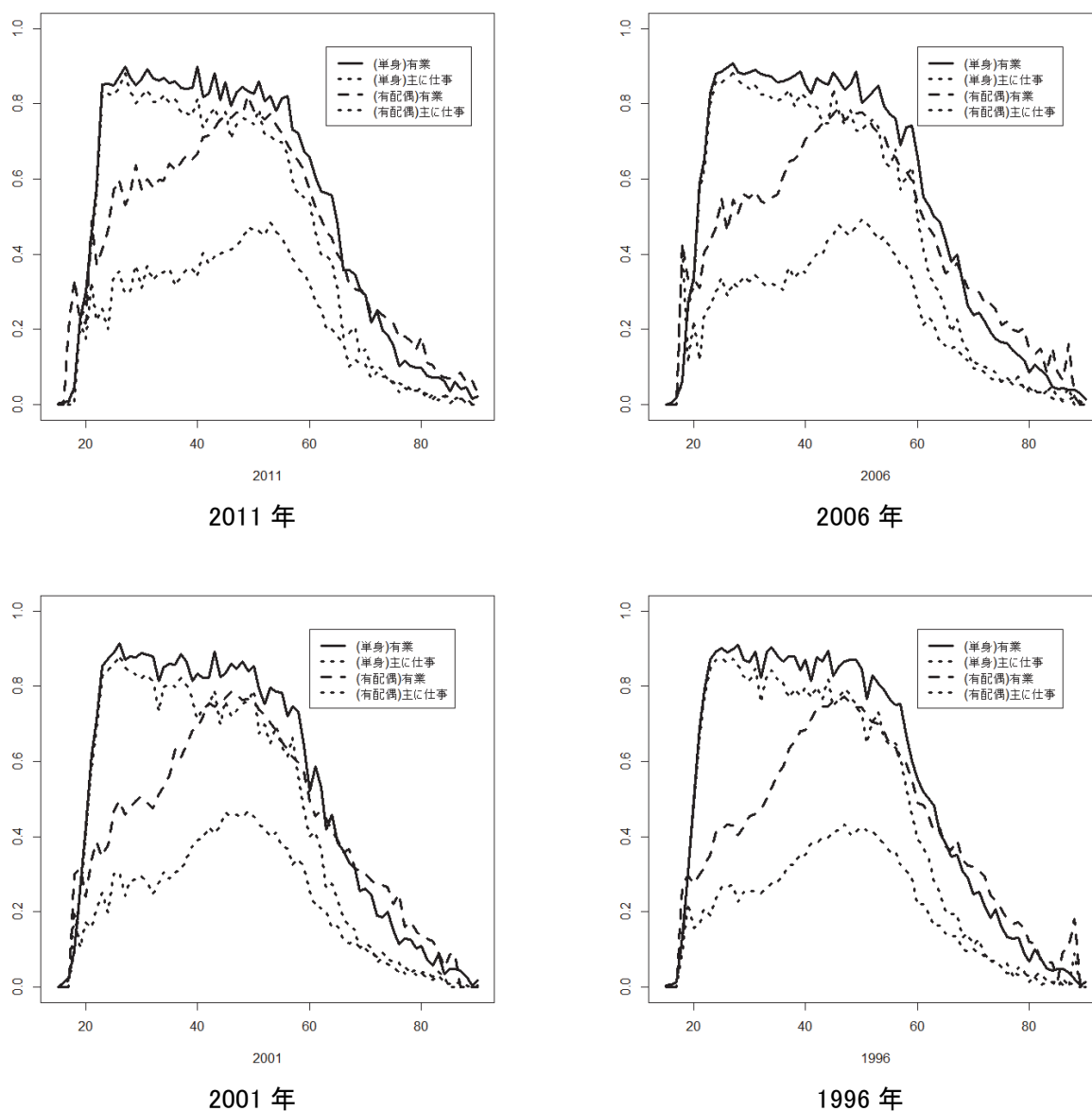


図4-5 女性の有業率と配偶関係(90歳以上は合計した値)

過去の『就業構造基本調査』も『社会生活基本調査』ともに、ほぼ同じ傾向が続いていると考えると、過去のデータによって将来の有業率の傾向を推定することも、ほぼ同じようにできるのではないかと考えたくなる。データの一部を使ってパラメータを推定し、残りのデータの予測ができるかどうかを試すことと、過去のデータを使って推定されたパラメータで後のデータの予測ができるかどうかを試すことができる。

つぎに都道府県別で就業状態の比率に大きな差が表れるかどうかを調べる。かりに、都道府県別でありあまり大きな就業状態の構成比に差がなければ、分析に地域の変数を入れることは不必要である。さらに、都会に住むのは就職を探すためなのか、都会暮らしほど就職を促す傾向にあるかどうかという同時性の問題を考えずにすむ。図 4-6、図 4-7 は都道府県でそれぞれ就業状態の構成比を計算して、47 個の観察値について年齢階層別に box-plot 図を描いたものである。

たとえば、20 歳未満の場合、図 4-6 (男性)、4-7 (女性) でも、box-plot の箱の大きさは小さく、地域別には大きく異なることがわかる。ただし、20 歳未満ではずれ値となっている県がある (図中の小さい○印)。これは神奈川県である。男性 (対象 143 人) も女性 (対象 124 人) も 10 代の有業率が高く、おもに通学による無業比率 (図の x 軸の 5) が低い。その他の年齢階層では、男性 60 歳代の沖縄県 (対象 145 人)、85 歳以上の長野県 (対象 38 人)、女性では 30 歳代の広島県 (対象 117 人)、60 歳代の神奈川県 (264 人) と奈良県 (192 人)、80 歳代の岩手県 (対象 84 人) で他県とはずれた値が観察されている。ただし元になる比率の統計に含まれる誤差は大きいので、都道府県別にみてはずれているかどうかの判断はより詳細な統計が必要である (0.5 で最大になる場合 264 人でも標準偏差が 3%、38 人だと 8%)。

男性の場合、実に 30 代から 60 代まで、都道府県の差はないといえる。したがって、この階層に都道府県の変数を入れることは無駄である。総じて地域差がある比率は女性のおもに仕事している者の比率 (図の x 軸の 1) である。男性のおもに仕事している者の比率が地域で異なるのは、20 代のみである。就職を機に居住地を変える可能性を考えると、説明変数として 20 代の男性に地域を入れることの正当性は失われる。

有業率のグラフから、年齢階層別に推定した場合でも 50 代や 60 代といった年齢によって有業率カーブが変化する年齢階層では、年齢変数と年齢の 2 乗の変数を代入することが考えられる。こうしたことから、統計的学習のモデルでは、年齢変数とその 2 乗を採用することにした。

年齢以外の変数には、学校教育の有無である。さらに介護が必要な家族の有無も式の説明変数として採用することにした。

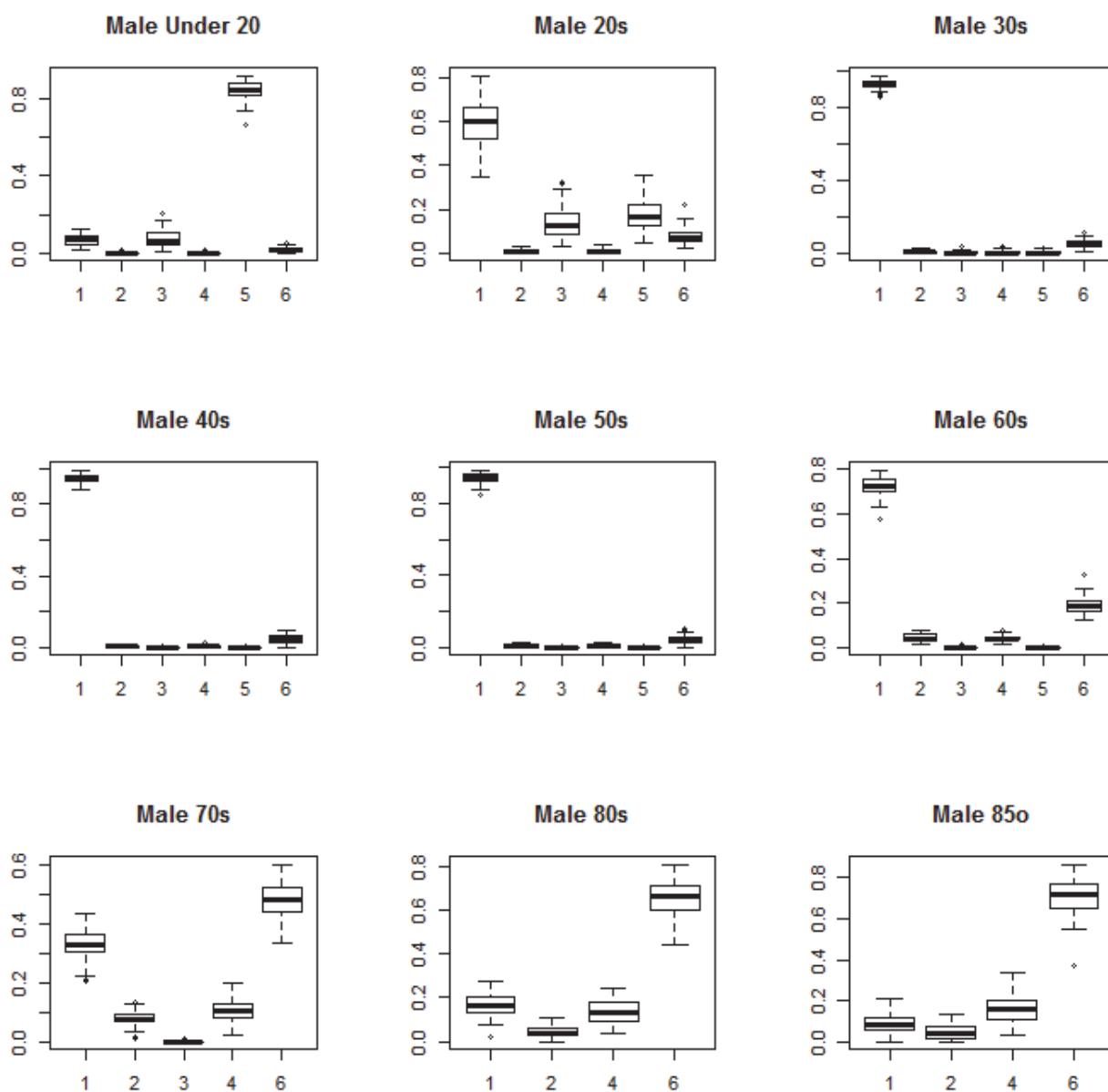


図 4-6 ふだんの就業状態の構成比(男性):年齢階級別、都道府県別集計データ 2011 年

注：1. おもに仕事、2. 家事などのかたわらに仕事、3. 通学のかたわらに仕事、4. 仕事をしていない人のうち主に家事をしている人、5. 仕事をしていない人のうち主に通学している人、6. その他の仕事をしていない人。各都道府県別に 1 から 6 の状態を合計すると 1 になる。図は、都道府県別のばらつきを調べている。上下の線は、第 1・第 4 四分位数の値を示す。箱の中の太線は標本平均、はずれ値は外側の点である。たとえば、30 歳台から 50 歳台のおもに仕事の比率は高く、県別でばらついていないことを示す。

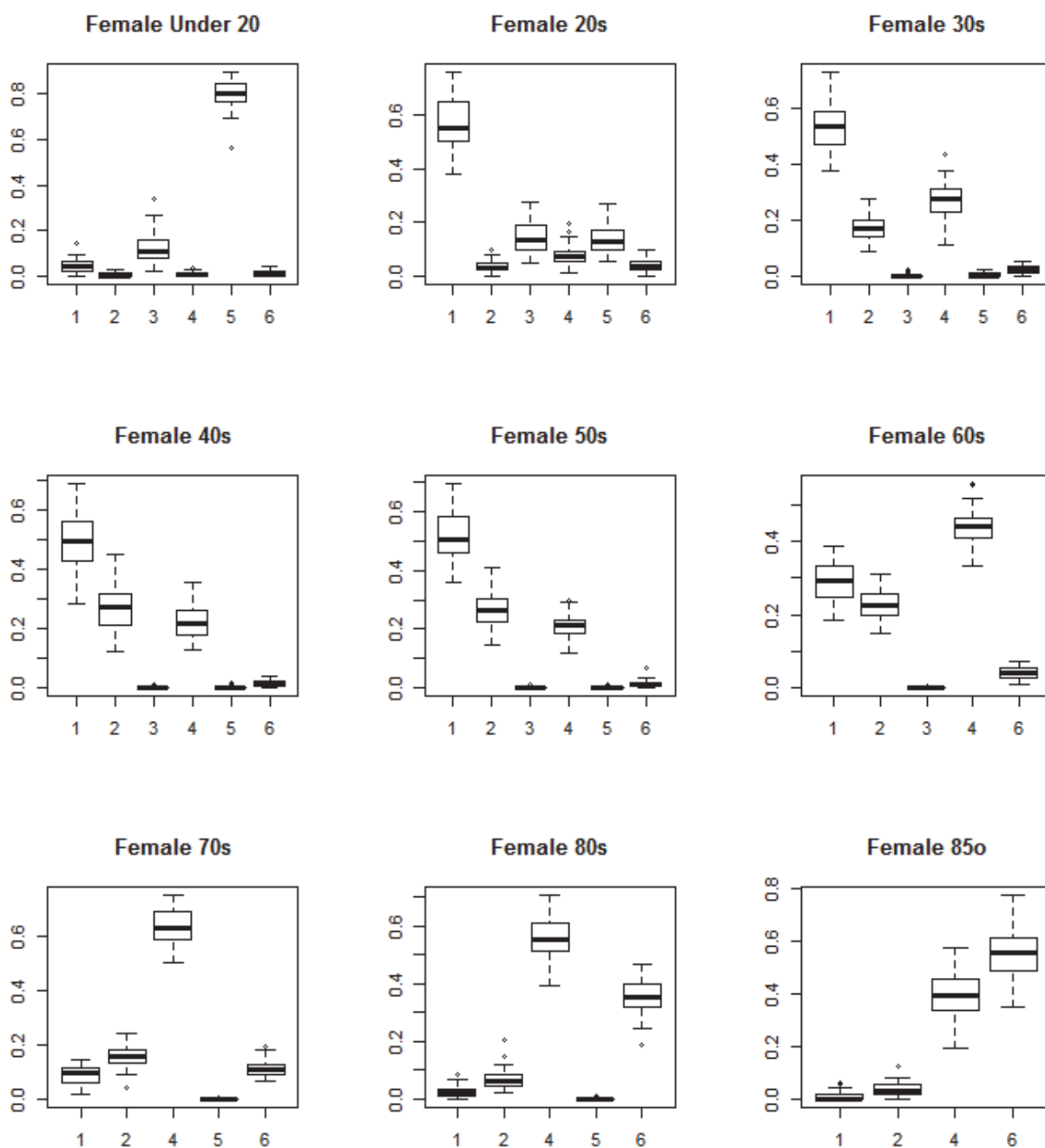


図 4-7 ふだんの就業状態の構成比(女性):年齢階級別、都道府県別集計データ 2011 年

注：1. おもに仕事、2. 家事などのかたわらに仕事、3. 通学のかたわらに仕事、4. 仕事をしていない人のうち主に家事をしている人、5. 仕事をしていない人のうち主に通学している人、6. その他の仕事をしていない人。各都道府県別に 1 から 6 の状態を合計すると 1 になる。図は、都道府県別のばらつきを調べている。上下の線は、第 1・第 4 四分位数の値を示す。箱の中の太線は標本平均。はずれ値は外側の点である。たとえば、男性に比べ 30 歳台から 50 歳台のおもに仕事の比率は低く、県別でばらついている。20 歳未満の 1,3,5 のはずれ値は神奈川県である。他県よりも就業比率が高いことを示している。

第7節 判別分析の結果

計測したパラメータの結果にもとづいて予測を行う。トレーニング・セットとテスト・セットで同じ判別が行われる。ここで利用した従属変数は、6状態ある。この6状態を正しく分類できるかどうかをチェックする。表4-1のように表側に、判別結果で予想された状態をとり、表頭に、観察結果の状態をとる。各観測値について判別結果と観察結果のとの状態の値によってカウントする(人数)。このように行列であらわすと、対角が正しく判別された状態の観察数ということがわかる。

1. 同じ調査年のトレーニングとテスト

トレーニング・データは、サンプルの一部を使って下記の表のパラメータを推定する。テスト・データは、残りのデータの属性のみを使って、下記の表で推定されたパラメータにしたがって予測をする。パラメータの値を得るのに従属変数の結果を使っていない。トレーニング・データとテスト・データについて同じ形式の2つの表が得られる。

表4-1 テスト・データの例:2011年 15歳から19歳(男性)

		観察された分類						計
		仕事あり(有業)			仕事なし(無業)			
		1:おもに仕事	2:家事のかたわら	3:通学のかたわら	4:おもに家事	5:おもに通学	6:その他	
予測した分類	1	202	1	158	4	321	38	724
	2	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0
	5	99	5	217	3	3374	34	3732
	6	0	0	0	0	0	0	0
	計	301	6	375	7	3695	72	4456

注:説明変数に利用したものは、年齢と年齢の2乗。プログラムはRのldaによる。このクロス表は予測値と実際に観察された値とのクロスである。たとえば、「おもに仕事」と推定結果から予測されるが、実際には「仕事ありで家事のかたわら」と回答している人が1名いるということを示す。推定の精度を示す表であり、個人の属性について特定することはできない。同様の表が掲載されているが、同じである。

結果の例を、表4-1に示された15歳から19歳の有業状態の分類:線形判別関数で説明する。説明変数は、年齢と年齢の2乗のみである。トレーニング・セットのサンプル・サイズは220である。つまり、220の観察値で関数のパラメータを推定し、そのパラメータの値を使って、残りの4456の観察値をテストする。つまり、外挿ということである。2011年の15歳から19歳の男性の就業状態について判別分析を行った。

この表4-1から、テスト・データでの的中率は $0.8025=(202+3374)/4456$ である。トレー

ニング・サブサンプルでの的中率は、0.8182 であった。有業か無業かの的中率は、 $0.8465=(361+3411)/4456$ となる。

同様の年齢階層で、2011年の女性については、テスト・データの的中率は0.7927であった。トレーニング・サブサンプルでの的中率は0.8364となっていた。女性の方が当てはまりは低いという傾向が認められた。

男性の場合について、シミュレーションとして5000回異なる220のトレーニング・セットを抽出して、テストを行った。

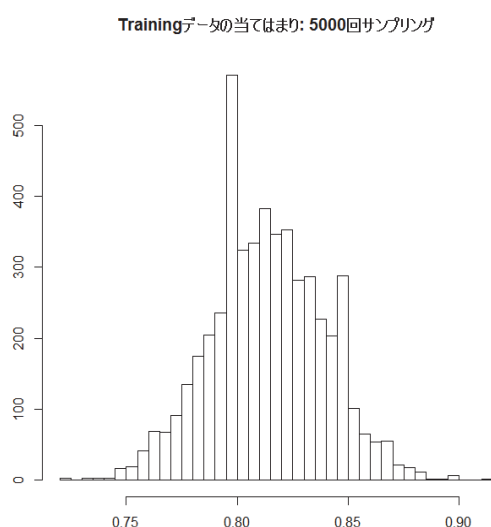


図 4-8 トレーニング・データの当てはまり
:2011年15歳から19歳男性

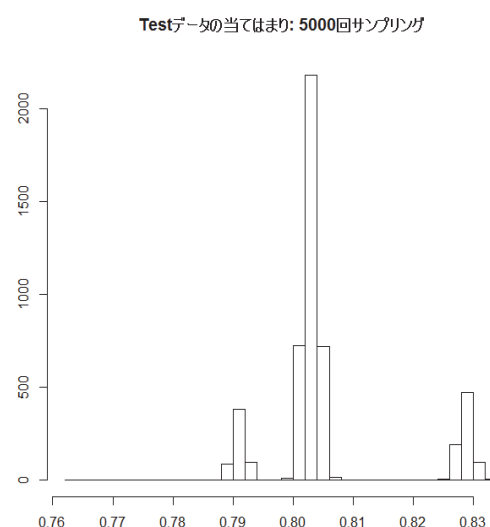


図 4-9 テスト・データの当てはまり
:2011年15歳から19歳男性

注：5000回、各220の観察値をサンプリングして線形判別したときの当てはまりの頻度分布

注：左のサンプリング・データで推計したパラメーターで予測したときの当てはまりの頻度分布

トレーニング・セットを抽出する方法はランダムなので、別のトレーニング・セットでは別の結果が得られる。しかし、その分布はほぼ一定である（図4-8）。このトレーニング・セットによるパラメーターを使って予測したときの当てはまりの分布は、一箇所に固まる傾向が見られる（図4-9）。

このような計算を2011年のデータについて年齢別・性別におこなった。説明変数として、教育ダミーを利用するものと、教育年数を計算してそれを用いるものの2通りを、線形判別分析で比較した。教育ダミーを利用する場合、教育水準に応じて説明変数が1つ増えるので、説明変数の数が増加してしまうという問題がある。しかも、わずかなトレーニング・データで予測すると多重共線性に陥りやすい。教育年数にした場合、各年で同じ限界的効果を与えるという仮定が必要になる。

線形判別関数のほかに代表的な線形ではない境界を選べるサポート・ベクター・マシン

(SVM) による判別を検討してみた。これらの結果をまとめた表が、表 4-2 である。さらに、個別の分類については、付表 4-1、4-2 を参照のこと。

表 4-2 を見て、教育ダミーと教育年数の説明力の違いに注目すると、ダミーではなくてもそれほど差があるようには思えない。トレーニング・データは、全体の半分のデータを利用している。残りの半分のデータをテスト・データとしている。トレーニング・データの当てはまりは、そのデータを使って推定されたパラメーターを利用するので、テスト・データのようないわゆる外挿よりも、よいはずである。しかし、必ずしもそうはなっていない。

男女ともに教育ダミーを使うほうが当てはまりがよいのは、若年層である。それ以降はあまり関係がない。この当てはまりは、1 であれば、すべて観察結果の就業状態（6 種類）と予測結果の就業状態（6 種類）が一致することを示している。

サポート・ベクター・マシン（SVM）との比較を見ると、これもどちらともいえないことがわかる。ここで利用したサポート・ベクター・マシンは、分類モデルで、各グループに所属する確率も計算する。ガウシアン・カーネルを使っているものである。予測には、就業状態に対応して 1 から 6 の値を応答するようにしている。詳細は付表 4-3 を参照。

男女別・年齢階層別の予測結果の当てはまりを見ると、性・年齢別有業率曲線から予想できる結論となっている。すなわち、男性よりも女性のほうが当てはまりが悪く、男性では若年と高齢者、女性ではこれに中年の就業状態を当てはめるのが難しいことがわかる。

サポート・ベクター・マシンによる分類の例を図 4-10 若年女性、図 4-11 高年男性に示している。年齢と教育年数の組み合わせによって就業確率が高い領域（青）と低い領域（赤）が例示されている。

表 4-2 推定法別のあてはまり(2011年)

	推定方法 教育変数	男性		女性			
		線形判別関数		SVM	線形判別関数		SVM
		ダミー	年数	年数	ダミー	年数	年数
15歳～19歳	Training data	0.8554	0.7917	0.8422	0.8216	0.7878	0.7975
	Test data	0.8614	0.8144	0.8452	0.8189	0.7883	0.8014
20歳～24歳	Training data	0.8892	0.6046	0.6627	0.7085	0.5641	0.6049
	Test data	0.8901	0.6439	0.6563	0.7097	0.5709	0.5883
25歳～29歳	Training data	0.8916	0.8752	0.8733	0.6586	0.6490	0.6641
	Test data	0.8853	0.8892	0.8911	0.6450	0.6478	0.6396
30歳～34歳	Training data	0.9246	0.9274	0.9217	0.5287	0.5295	0.5314
	Test data	0.9262	0.9258	0.9314	0.5387	0.5245	0.5295
35歳～39歳	Training data	0.9424	0.9472	0.9475	0.4798	0.4657	0.4855
	Test data	0.9365	0.9453	0.9450	0.4672	0.4771	0.4636
40歳～44歳	Training data	0.9383	0.9374	0.9423	0.4747	0.4767	0.4738
	Test data	0.9334	0.9348	0.9404	0.4587	0.4671	0.4705
45歳～49歳	Training data	0.9414	0.9416	0.9427	0.4934	0.4978	0.4919
	Test data	0.9422	0.9422	0.9401	0.4984	0.4944	0.5013
50歳～54歳	Training data	0.9393	0.9411	0.9441	0.5133	0.5124	0.5039
	Test data	0.9483	0.9435	0.9405	0.4985	0.5073	0.5127
55歳～59歳	Training data	0.9098	0.9103	0.9181	0.4270	0.4346	0.4396
	Test data	0.9135	0.9117	0.9039	0.4231	0.4101	0.4141
60歳～64歳	Training data	0.7256	0.7314	0.7264	0.4511	0.4436	0.4449
	Test data	0.7237	0.7192	0.7224	0.4327	0.4462	0.4519
65歳～69歳	Training data	0.4746	0.4746	0.4717	0.5809	0.5854	0.5845
	Test data	0.4823	0.4687	0.4799	0.5831	0.5790	0.5784
70歳～74歳	Training data	0.4848	0.4954	0.4979	0.6317	0.6429	0.6432
	Test data	0.4905	0.4841	0.4855	0.6498	0.6392	0.6389
75歳～79歳	Training data	0.5751	0.5794	0.5772	0.6374	0.6467	0.6460
	Test data	0.5722	0.5683	0.5696	0.6306	0.6340	0.6378
80歳～84歳	Training data	0.6438	0.6393	0.6494	0.5708	0.5576	0.5717
	Test data	0.6494	0.6576	0.6488	0.5462	0.5585	0.5539
85歳～	Training data	0.6943	0.6898	0.7238	0.5782	0.5600	0.5882
	Test data	0.7134	0.7100	0.6676	0.5588	0.5843	0.5765

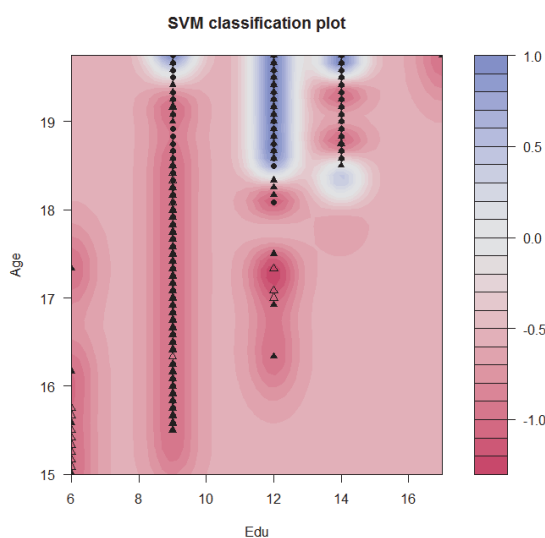


図 4-10 15 歳から 19 歳女性の教育年数・年齢
と就業状態(青は有業)2011 年

注：教育年数が 6 年は、「中学在学中」の者

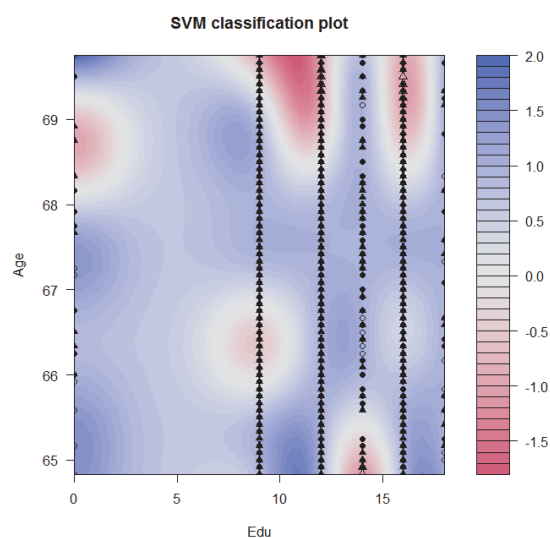


図 4-11 65 歳から 69 歳男性の教育年数・年齢
と就業状態(青は有業)2011 年

注：教育年数が 0 年は、「在学したことがない」者

サポート・ベクター・マシンは、R のパッケージ kernlab にある ksvm である。カーネルにはガウシアン・カーネルの動径基底関数を使っている。たとえば、女性で 40 代の場合、トレーニング・セットのサイズは 3474 で、超平面の σ の値は、0.9287 であり、サポート・ベクターの数は 3352 である。制約に使ったパラメーターは $C=10$ である。説明変数はすべてスケールされている。トレーニング・エラーは 0.5230 と報告されている。分類には有界制約付き (bound-Constraint svm classification) で分類する手法を使っている。

2. 異なる調査年のトレーニングとテスト

つぎの実験として、1996 年のデータをトレーニング・データとして利用して、2011 年の説明変数を代入し、2011 年の就業状態を予測するというテストを行ってみた。その結果は、思ったよりもテストのあてはまりがよいということである。しかし、一方でそれはあたりまえで、あてはまりの悪い年齢階層については、予測すると一層あてはまりは悪くなることもみられた。詳細は付表 4-4 を参照。

表 4-3 で、Training data とあるのは、1996 年のデータで、1996 年の就業状態のあてはまりを計算したものである。Test data とあるのは、1996 年のデータで推計された情報 (サポート・ベクター) を使って 2011 年の年齢と教育年数のデータから就業状態の予測をおこなったものである。時点も異なることにくわえて、調査対象も全く異なるデータに対して、それほど大きな説明力の低下が見られなかったということは、年齢と教育年数でかなりの就業状態の分類ができることを示している。

表 4-3 1996 年のデータでトレーニングして 2011 年を予測する

		男性		女性	
		線形判別	SVM	線形判別	SVM
15 歳～19 歳	Training data	0.7245	0.839	0.7106	0.8142
	Test data	0.6766	0.4767	0.6539	0.6434
20 歳～24 歳	Training data	0.7702	0.7976	0.6993	0.71
	Test data	0.588	0.5898	0.57	0.5402
25 歳～29 歳	Training data	0.9577	0.9578	0.5171	0.5299
	Test data	0.8822	0.8819	0.613	0.5751
30 歳～34 歳	Training data	0.9762	0.9765	0.4079	0.4213
	Test data	0.9266	0.8881	0.2739	0.3174
35 歳～39 歳	Training data	0.9791	0.9793	0.3902	0.3938
	Test data	0.9462	0.9262	0.4102	0.3988
40 歳～44 歳	Training data	0.9792	NA	0.4405	0.4412
	Test data	0.9417	NA	0.4703	0.4637
45 歳～49 歳	Training data	0.9779	0.9779	0.4621	0.4626
	Test data	0.9418	0.9418	0.4958	0.4872
50 歳～54 歳	Training data	0.9681	0.9683	0.4331	0.4371
	Test data	0.9423	0.9274	0.4905	0.4749
55 歳～59 歳	Training data	0.8982	NA	0.4109	0.4128
	Test data	0.911	NA	0.3723	0.368
60 歳～64 歳	Training data	0.6835	0.6837	0.5061	NA
	Test data	0.7253	0.7242	0.445	NA
65 歳～69 歳	Training data	0.5158	0.519	0.554	0.5546
	Test data	0.457	0.4567	0.582	0.5311
70 歳～74 歳	Training data	0.493	NA	0.5556	0.5582
	Test data	0.4882	NA	0.641	0.539
75 歳～79 歳	Training data	0.6259	NA	0.5026	NA
	Test data	0.5737	NA	0.6388	NA
80 歳～84 歳	Training data	0.7236	0.7236	0.5851	NA
	Test data	0.6491	0.6491	0.3846	NA
85 歳～	Training data	0.8229	NA	0.769	0.7747
	Test data	0.695	NA	0.5627	0.5616

注：NA はサポート・ベクター・マシンの 1996 年の推定結果が収束しなかったもの。

その一方で、女性の転換期の有業率、たとえば30歳から34歳層は大きく変化しているが、この階層については、2011年の予測はきわめて低下しているということが出来る。その他の年齢階層たとえば、40歳から44歳層はもともとあてはまりが悪いが、昔のパラメーターを使っても15年後の就業行動は、現在とほぼ同じ程度に予測できることがわかる。

第8節 まとめ

判別分析による結果は、クラスの数が多いにも関わらず、全サンプルの半分程度のサイズのトレーニング・セットで比較的安定的に判別結果が得られていることがわかった。ただし、トレーニング・セットによっては、グループに十分な観察値が入らず、ダミー変数で処理する場合にはしばしば定数項となって推定不可能なケースが生じた。つまり、観察値に十分なばらつきがないということである。その場合にはほとんど年齢で説明してしまうということである。年齢階層としては転換点にあたる若年層と引退する年齢層については、当然であるが他の階層よりもあてはまりが悪くなり、いまよりもさらに詳細な分析が必要である。

人生ではじめて就職する世代である20代のグループは、男女とも都道府県別に異なる構成比も観察されている。20歳未満では、地域差は少ないが神奈川県のようなはずれ値が存在する。他の年齢層でも男性では、70代以上になれば若干あるがそれまで地域差はほとんど観察されていない。男性よりも女性の方が地域差によるばらつきが大きいので、女性の就業機会には地域性と相関する可能性がある。しかし、因果関係の方向性はわからない。そのため地域ダミーを入れて計測しても大きな予測の改善が得られたというわけでもない。

このように詳細なあてはまりについては検討する点が多いが、予想以上にテスト・セットとトレーニング・セットでのあてはまりの差はなく、全体としてのパフォーマンスの良さが確認できた。しかも、今回の場合、より制約の強い線形判別関数を使った分析でも、サポート・ベクター・マシンを使ったフレキシブルな分析でも大きな差は見られなかった。労働供給主体の属性がほぼ正規分布をするということが理由であるかもしれない。

こうした結果は、労働力需給推計の労働供給面での推計作業につぎの3つの点を確認することができた。第一は、労働力率の年齢別推計については、ここでおこなった判別分析は就業形態・無就業形態の合計6形態を分類するものであるが、かなり高いパフォーマンスが得られている。第二は、安定的な年齢階層では、15年前のデータを使って計測されたパラメーターを使ってもほとんど推計結果に大きな違いは現れないことである。1996年のテスト・データでも、2011年の就業形態をかなり高い精度で分類することができる。第三は、あてはまりの悪い年齢階層については、非線形効果を入れても、サポート・ベクター・マシンなどの統計的学習法による分類作業でも、予測を改善することはできなかった。トレーニング・データとテスト・データが同じ年のものを使っても、大きな改善はされていない。就業選択の上で転機にあたる年齢層、女性の30代・40代の就業選択の行動について、より詳細な説明変数の選択が必要になる。ただし、それがどのような変数なのかは、今回の統計では調査項目

の内生性などの問題で十分追求することができなかった。隠れた変数をどう探していくかは今後の課題である。

参考文献

- Baker, M., J. Gruber, and K. Milligan (2008) “Universal Child Care, Maternal Labor Supply, and Family Well-Being.” *Journal of Political Economy*, 116: 709-745.
- Blau, D., and L. M. Kahn (2007) “Changes in the Labor Supply Behavior of Married Women: 1980-2000.” *Journal of Labor Economics* 25: 393-438.
- Chetty, R. (2012) “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply.” *Econometrica* 80: 969-1018.
- Cunha, F. and J. J. Heckman, and S. Schennan. (2010). “Estimating the Technology of Cognitive and Non-Cognitive Skill Formation.” *Econometrica* 78: 883-931.
- Del Boca, D., C. Flinn, and M. Wishwall (2012). “Transfers to Households with Children and Child Development.”
- Del Boca, D. and C. Flinn (2009). “Engogeneous Household Interaction.” IZA DP No. 4377.
- Frölich, M. (2006) “Non-parametric Regression for Binary Dependent Variables.” *Econometrics Journal*, 9: 511-540.
- Hastie, T., R. Tibshirani, and J. Friedman (2009) *The Elements of Statistical Learning*, second edition, Springer-Verlag.
- 早見 均 (1997) 「女性就業行動の変化: 1977-82-87-92 年の比較分析」労働省婦人局委託調査財団法人労働問題リサーチセンター『女性労働者の雇用と賃金に関する調査研究』1997年3月, 123-152.

付表 4-1 線形判別関数による分類:説明変数は、年齢、年齢の2乗、性別ダミー(2011年)

男性	観察値						女性	観察値																			
	1	2	3	4	5	6		計	1	2	3	4	5	6	計												
15歳~19歳	15歳~19歳												15歳~19歳														
トレーニング	トレーニング												トレーニング														
おもに仕事	おもに仕事												おもに仕事														
家事などのかたわらに仕事	家事などのかたわらに仕事												家事などのかたわらに仕事														
通学のかたわらに仕事	通学のかたわらに仕事												通学のかたわらに仕事														
仕事をしていない人のうち主に家事をしている人	仕事をしていない人のうち主に家事をしている人												仕事をしていない人のうち主に家事をしている人														
仕事をしていない人のうち主に通学している人	仕事をしていない人のうち主に通学している人												仕事をしていない人のうち主に通学している人														
その他の仕事をしていない人	その他の仕事をしていない人												その他の仕事をしていない人														
計	147						7	187						3	1949						45						
Training dataの あてはまり	0.798												Training dataの あてはまり						0.804								
テスト	テスト												テスト						テスト								
おもに仕事	59												0	48						1	99						8
家事などのかたわらに仕事	0												0	3						0	19						0
通学のかたわらに仕事	23												0	25						1	44						5
仕事をしていない人のうち主に家事をしている人	6												0	6						0	54						1
仕事をしていない人のうち主に通学している人	71												0	116						3	1700						18
その他の仕事をしていない人	13												0	3						0	12						0
計	172						0	201						5	1928						32						
Test dataの あてはまり	0.763												Test dataの あてはまり						0.775								
20歳~24歳	20歳~24歳												20歳~24歳						20歳~24歳								
トレーニング	トレーニング												トレーニング						トレーニング								
おもに仕事	1												2	3						4	5						6
家事などのかたわらに仕事	954												19	167						15	209						114
通学のかたわらに仕事	39												3	84						1	65						12
仕事をしていない人のうち主に家事をしている人	5												0	0						2	1						0
仕事をしていない人のうち主に通学している人	54												2	36						3	69						5
その他の仕事をしていない人	0												0	0						0	0						0
計	1052						24	287						21	344						131						
Training dataの あてはまり	0.597												Training dataの あてはまり						0.588								
テスト	テスト												テスト						テスト								
おもに仕事	953												14	193						11	224						118
家事などのかたわらに仕事	0												0	0						0	0						0
通学のかたわらに仕事	63												1	70						3	47						10
仕事をしていない人のうち主に家事をしている人	9												0	0						0	0						1
仕事をしていない人のうち主に通学している人	30												0	50						0	53						9
その他の仕事をしていない人	0												0	0						0	0						0
計	1055						15	313						14	324						138						
Training dataの あてはまり	0.579												Training dataの あてはまり						0.588								
テスト	テスト												テスト						テスト								
おもに仕事	934												55	160						127	130						54
家事などのかたわらに仕事	0												0	0						0	0						0
通学のかたわらに仕事	78												4	101						11	82						10
仕事をしていない人のうち主に家事をしている人	24												0	0						5	1						2
仕事をしていない人のうち主に通学している人	56												3	55						5	55						8
その他の仕事をしていない人	0												0	0						0	0						0
計	1092						62	316						148	268						74						
Test dataの あてはまり	0.559												Test dataの あてはまり						0.559								

付表 4-1 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、県別ダミー(2011年)

性別	年齢	観察値						観察値	計						
		1	2	3	4	5	6								
男性	25歳~29歳	1	2	3	4	5	6	1	2	3	4	5	6	計	
	トレーニング	1763	11	29	12	28	160	2003	1406	250	15	395	21	84	2171
	予測1 おもに仕事	27	2	1	0	0	2	32	0	0	0	0	0	0	0
	予測2 家事などのかたわらに仕事	0	0	0	0	0	0	0	0	1	0	2	0	0	3
	予測3 通学のかたわらに仕事	27	0	0	2	1	1	31	7	3	0	4	0	0	14
	予測4 仕事をしない人のうち主に家事をしている人	9	0	0	0	0	0	9	0	0	0	0	0	0	0
	予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	予測6 その他の仕事をしない人	1826	13	30	14	29	163	0	1413	254	15	401	21	84	0
	計	0.852							0.644						
	Training dataの あてはまり														
	テスト														
予測1 おもに仕事	1769	18	23	19	48	128	2005	1416	231	19	383	35	77	2161	
予測2 家事などのかたわらに仕事	32	0	0	1	0	0	33	0	0	0	0	0	0	0	
予測3 通学のかたわらに仕事	0	0	0	0	0	0	0	8	1	0	0	1	0	10	
予測4 仕事をしない人のうち主に家事をしている人	31	1	1	0	0	1	34	10	3	0	5	0	0	18	
予測5 仕事をしない人のうち主に通学している人	3	0	0	0	0	0	3	0	0	0	0	0	0	0	
予測6 その他の仕事をしない人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
計	1835	19	24	20	48	129	0	1434	235	19	388	36	77	0	
	0.853							0.649							
	Test dataの あてはまり														
男性	30歳~34歳	1	2	3	4	5	6	1	2	3	4	5	6	計	
	トレーニング	2250	12	7	13	16	137	2435	1249	366	4	604	11	48	2282
	予測1 おもに仕事	0	0	0	0	0	0	0	12	17	0	13	0	3	45
	予測2 家事などのかたわらに仕事	0	0	0	0	0	0	0	31	4	1	12	0	2	50
	予測3 通学のかたわらに仕事	39	1	0	3	0	1	44	66	46	1	80	0	2	195
	予測4 仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	18	7	0	12	1	2	40
	予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	予測6 その他の仕事をしない人	2289	13	7	16	16	138	0	1376	440	6	721	12	57	0
	計	0.909						0.516							
	Training dataの あてはまり														
	テスト														
予測1 おもに仕事	2249	17	10	15	7	122	2420	1239	393	7	588	9	58	2294	
予測2 家事などのかたわらに仕事	0	0	0	0	0	0	0	22	8	0	15	0	1	46	
予測3 通学のかたわらに仕事	0	0	0	0	0	0	0	23	12	1	17	0	0	53	
予測4 仕事をしない人のうち主に家事をしている人	56	0	0	0	1	2	59	71	36	0	68	3	4	182	
予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	14	8	0	13	0	2	37	
予測6 その他の仕事をしない人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
計	2305	17	10	15	8	124	0	1369	457	8	701	12	65	0	
	0.907							0.504							
	Test dataの あてはまり														

付表 4-1 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、県別ダミー(2011年)

性別	年齢	観察値						観察値	計				
		1	2	3	4	5	6						
男性	35歳~39歳	2948	17	0	14	7	126	3112	0	738	5	6	計
	トレーニング												
	おもに仕事	0	0	0	0	0	0	0	0	0	17	0	81
	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	34	0	4	133
	通学のかたわらに仕事	46	0	0	0	0	0	0	0	119	1	9	304
	仕事をしない人のうち主に家事をしている人	0	0	0	1	2	1	50	0	0	0	0	0
	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
	その他の仕事をしない人	2994	17	0	0	0	0	0	0	0	0	0	0
	計	0.933					9	127		2	908	9	69
	Training dataの あてはまり												
	テスト												
予測1	おもに仕事	2954	22	2	18	4	123	3123	0	795	2	57	2844
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	11	0	0	58
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	29	0	1	124
予測4	仕事をしない人のうち主に家事をしている人	37	1	0	0	0	2	40	0	113	2	7	315
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしない人	2991	23	2	18	4	125		0	0	0	0	0
	計	0.934							5	948	4	65	
	Training dataの あてはまり												
	テスト												
予測1	おもに仕事	1324	661	5	795	2	57	2844	1324	661	2	57	2844
予測2	家事などのかたわらに仕事	31	16	0	11	0	0	58	31	16	0	0	58
予測3	通学のかたわらに仕事	67	27	0	29	0	1	124	67	27	0	1	124
予測4	仕事をしない人のうち主に家事をしている人	119	74	0	113	2	7	315	119	74	2	7	315
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしない人	1541	778	5	948	4	65		1541	778	5	65	
	計	0.435							5	948	4	65	
	Training dataの あてはまり												
	テスト												
予測1	おもに仕事	1313	721	1	575	1	30	2641	1313	721	1	30	2641
予測2	家事などのかたわらに仕事	76	83	0	67	1	5	232	76	83	0	5	232
予測3	通学のかたわらに仕事	93	35	3	37	1	2	171	93	35	3	2	171
予測4	仕事をしない人のうち主に家事をしている人	110	119	0	131	0	6	366	110	119	0	6	366
予測5	仕事をしない人のうち主に通学している人	35	12	0	13	2	2	64	35	12	0	2	64
予測6	その他の仕事をしない人	1627	970	4	823	5	45		1627	970	4	45	
	計	0.441							4	823	5	45	
	Training dataの あてはまり												
	テスト												
予測1	おもに仕事	1350	714	5	604	1	41	2715	1350	714	5	41	2715
予測2	家事などのかたわらに仕事	65	75	0	69	1	0	210	65	75	0	0	210
予測3	通学のかたわらに仕事	91	35	0	34	1	3	164	91	35	0	3	164
予測4	仕事をしない人のうち主に家事をしている人	105	111	0	101	1	4	322	105	111	0	4	322
予測5	仕事をしない人のうち主に通学している人	30	15	0	16	0	3	64	30	15	0	3	64
予測6	その他の仕事をしない人	1641	950	5	824	4	51		1641	950	5	4	51
	計	0.439							5	824	4	51	
	Training dataの あてはまり												
	テスト												

付表 4-1 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、県別ダミー(2011年)

性別	年齢	観察値						観察値	計							
		1	2	3	4	5	6									
男性	45歳~49歳	1	2	3	4	5	6	1	2	3	4	5	6	計		
	トレーニング	2562	16	0	23	0	120	2721	0	643	0	462	0	42	2438	
	予測1 おもに仕事	0	0	0	0	0	0	0	0	182	260	0	111	1	12	566
	予測2 家事などのかたわらに仕事	97	1	1	0	0	4	103	33	40	33	1	30	0	2	106
	予測3 通学のかたわらに仕事	0	0	0	0	0	0	0	0	12	10	0	9	0	1	32
	予測4 仕事をしない人のうち主に家事をしている人	50	0	0	0	1	3	54	10	24	10	0	9	1	6	50
	予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	予測6 その他の仕事をしない人	2709	17	1	23	1	127	3081	956	1549	621	1	621	2	63	
	計	0.891							0.489							
	Training dataの あてはまり															
	テスト															
男性	50歳~54歳	1	2	3	4	5	6	1	2	3	4	5	6	計		
	トレーニング	2556	20	0	17	1	121	2715	652	2	425	8	30	2417		
	予測1 おもに仕事	0	0	0	0	0	0	0	226	203	0	127	2	9	567	
	予測2 家事などのかたわらに仕事	105	0	0	3	0	3	111	35	35	0	27	0	0	97	
	予測3 通学のかたわらに仕事	0	0	0	0	0	0	0	21	14	0	12	0	0	47	
	予測4 仕事をしない人のうち主に家事をしている人	51	0	0	0	0	1	52	34	13	0	13	1	3	64	
	予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	予測6 その他の仕事をしない人	2712	20	0	20	1	125	3088	917	1616	604	11	42			
	計	0.888							0.475							
	Training dataの あてはまり															
	テスト															
男性	50歳~54歳	1	2	3	4	5	6	1	2	3	4	5	6	計		
	トレーニング	2698	18	0	28	0	119	2863	722	0	605	0	29	2846		
	予測1 おもに仕事	0	0	0	0	0	0	0	53	61	0	43	0	1	158	
	予測2 家事などのかたわらに仕事	0	0	0	0	0	0	0	30	18	1	13	0	0	62	
	予測3 通学のかたわらに仕事	0	0	0	0	0	0	0	42	42	0	41	0	0	125	
	予測4 仕事をしない人のうち主に家事をしている人	118	2	0	2	2	3	127	40	20	0	11	1	2	74	
	予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	25	7	0	8	0	5	45	
	予測6 その他の仕事をしない人	2816	20	0	30	2	122	3088	870	1680	721	1	37			
	計	0.903							0.483							
	Training dataの あてはまり															
	テスト															
男性	50歳~54歳	1	2	3	4	5	6	1	2	3	4	5	6	計		
	トレーニング	2705	27	1	21	1	116	2871	736	0	601	2	32	2845		
	予測1 おもに仕事	0	0	0	0	0	0	0	68	49	0	26	0	3	146	
	予測2 家事などのかたわらに仕事	0	0	0	0	0	0	0	36	9	0	6	1	2	54	
	予測3 通学のかたわらに仕事	0	0	0	0	0	0	0	64	53	0	36	0	1	154	
	予測4 仕事をしない人のうち主に家事をしている人	115	2	0	1	0	2	120	36	21	0	24	0	0	81	
	予測5 仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	17	8	0	5	0	0	30	
	予測6 その他の仕事をしない人	2820	29	1	22	1	118	3083	876	1695	698	3	38			
	計	0.904							0.471							
	Training dataの あてはまり															
	テスト															

付表 4-1 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、県別ダミー(2011年)

性別	55歳～59歳 トレーニング	観察値						観察値	計					
		1	2	3	4	5	6							
男性	3117	57	0	54	0	205	3433	1281	653	0	729	0	45	2708
予測1	0	0	0	0	0	0	0	76	83	0	62	0	4	225
予測2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	0	0	0	0	0	0	0	244	209	0	323	0	17	793
予測4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	3117	57	0	54	0	205	3433	1601	945	0	1114	0	66	66
Training dataの あてはまり														
テスト														
予測1	3138	44	2	47	1	201	3433	1267	658	0	762	1	56	2744
予測2	0	0	0	0	0	0	0	58	64	0	78	0	4	204
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	255	229	1	282	0	12	779
予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	3138	44	2	47	1	201	3433	1580	951	1	1122	1	72	72
Test dataの あてはまり														
60歳～64歳														
男性	3112	177	0	181	1	768	4239	269	118	0	199	0	21	607
予測1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測2	100	5	2	10	0	44	161	45	36	1	57	0	4	143
予測3	0	0	0	0	0	0	0	1016	948	0	1825	0	156	3945
予測4	70	2	0	1	1	16	90	0	0	0	0	0	0	0
予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	3282	184	2	192	2	828	4239	1330	1102	1	2081	0	181	181
Training dataの あてはまり														
テスト														
予測1	3112	177	0	181	1	768	4239	216	150	0	243	0	17	626
予測2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	100	5	2	10	0	44	161	31	22	0	59	0	3	115
予測4	70	2	0	1	1	16	90	1094	887	2	1793	0	179	3955
予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	3282	184	2	192	2	828	4239	1341	1059	2	2095	0	199	199
Test dataの あてはまり														

付表 4-1 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、県別ダミー(2011年)

性別	65歳～69歳 トレーニング	観察値						観察値	1	2	3	4	5	6	計
		1	2	3	4	5	6								
男性	65歳～69歳 トレーニング	951	177	0	143	0	625	1896	0	0	0	0	0	0	0
予測1	おもに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	40	11	2	5	0	35	93	0	0	0	0	0	0	48
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	58	11	0	7	1	48	125	0	0	0	0	0	0	241
予測5	仕事をしない人のうち主に通学している人	346	71	1	82	0	496	996	0	0	0	0	0	0	3338
予測6	その他の仕事をしない人	1395	270	3	237	1	1204		0	0	0	0	0	0	0
計		0.466							518	654	0	1969	0	245	
	Training dataの あてはまり								0.581						
	テスト														
予測1	おもに仕事	938	155	3	145	0	671	1912	0	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	12	14	0	30	0	6	62
予測3	通学のかたわらに仕事	36	10	0	6	0	34	86	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	489	659	1	1943	0	233	3325
予測5	仕事をしない人のうち主に通学している人	57	8	0	7	0	41	113	0	0	0	0	0	0	0
予測6	その他の仕事をしない人	419	79	0	91	1	410	1000	0	0	0	0	0	0	0
計		1450	252	3	249	1	1156		501	673	1	1973	0	239	
	Training dataの あてはまり								0.578						
	テスト														
予測1	おもに仕事	196	42	0	43	0	157	438	0	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	18	3	1	4	0	27	53	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	286	512	0	2053	0	380	3231
予測5	仕事をしない人のうち主に通学している人	718	186	2	245	0	1186	2337	5	14	0	60	1	7	87
予測6	その他の仕事をしない人	932	231	3	292	0	1370		291	526	0	2113	1	387	
計		0.49							0.619						
	Training dataの あてはまり														
	テスト														
予測1	おもに仕事	159	27	0	47	0	207	440	0	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	20	5	0	9	0	25	59	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	292	488	0	2062	0	365	3207
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	10	13	0	79	0	9	111
予測6	その他の仕事をしない人	741	173	2	247	0	1166	2329	0	0	0	0	0	0	0
計		920	205	2	303	0	1398		302	501	0	2141	0	374	
	Training dataの あてはまり								0.621						
	テスト														

付表 4-1 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、性別ダミー(2011年)

男性	観察値						観察値	1	2	3	4	5	6	計
	75歳~79歳 トレーニング	75歳~79歳 トレーニング	75歳~79歳 トレーニング	75歳~79歳 トレーニング	75歳~79歳 トレーニング	75歳~79歳 トレーニング								
予測1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	23	13	2	14	0	36	88	予測3	2	0	5	0	2	12
予測4	0	0	0	0	0	0	0	予測4	0	0	0	0	0	0
予測5	0	0	0	0	0	0	0	予測5	111	297	0	1864	0	606
予測6	487	137	0	278	0	1321	2223	予測6	1	3	0	23	1	14
計	510	150	2	292	0	1357		計	115	302	0	1892	1	622
Training dataの あてはまり														
テスト														
予測1	0	0	0	0	0	0	0	予測1	0	0	0	0	0	0
予測2	0	0	0	0	0	0	0	予測2	2	3	0	11	0	3
予測3	24	5	0	16	0	37	82	予測3	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	予測4	149	304	0	1834	2	579
予測5	0	0	0	0	0	0	0	予測5	3	6	0	26	0	10
予測6	535	155	0	282	0	1258	2230	予測6	0	0	0	0	0	0
計	559	160	0	298	0	1295		計	154	313	0	1871	2	592
Training dataの あてはまり														
テスト														
予測1	0	0	0	0	0	0	0	予測1	0	0	0	0	0	0
予測2	0	0	0	0	0	0	0	予測2	0	0	0	0	0	0
予測3	24	5	0	16	0	37	82	予測3	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	予測4	149	304	0	1834	2	579
予測5	0	0	0	0	0	0	0	予測5	3	6	0	26	0	10
予測6	535	155	0	282	0	1258	2230	予測6	0	0	0	0	0	0
計	559	160	0	298	0	1295		計	154	313	0	1871	2	592
Test dataの あてはまり														
テスト														
予測1	1	2	3	4	5	6	計	予測1	1	2	3	4	5	6
予測2	14	1	0	5	0	11	31	予測2	7	3	0	19	0	19
予測3	10	11	0	6	0	28	55	予測3	1	10	0	17	0	17
予測4	0	0	0	0	0	0	0	予測4	0	0	0	0	0	0
予測5	0	0	0	0	0	0	0	予測5	51	109	0	1050	0	608
予測6	233	63	0	192	0	1012	1500	予測6	0	0	0	0	0	0
計	257	75	0	203	0	1051		計	68	139	0	1206	0	784
Training dataの あてはまり														
テスト														
予測1	5	1	0	4	0	21	31	予測1	1	3	0	25	0	20
予測2	10	0	0	4	0	31	45	予測2	1	7	0	16	0	15
予測3	0	0	0	0	0	0	0	予測3	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	予測4	36	135	0	1069	2	607
予測5	0	0	0	0	0	0	0	予測5	0	0	0	0	0	0
予測6	253	78	0	223	0	956	1510	予測6	5	11	0	128	0	116
計	268	79	0	231	0	1008		計	43	156	0	1238	2	758
Test dataの あてはまり														
テスト														
予測1	0.65							予測1	0.637					
予測2	0.544							予測2	0.626					
予測3	0.544							予測3	0.549					
予測4	0.65							予測4	0.543					
予測5	0.65							予測5	0.543					
予測6	0.65							予測6	0.543					
計	0.65							計	0.543					

付表 4-1 (続) 線形判別関数による分類: 説明変数は、年齢、年齢の2乗、性別ダミー(2011年)

男性	85歳～ トレーニング						女性						85歳～ トレーニング						
	1	2	3	4	5	6 計	1	2	3	4	5	6 計	1	2	3	4	5	6 計	
予測1	11	2	0	3	0	16	32	予測1	0	0	0	15	4	4	0	15	0	24	47
予測2	1	3	0	4	0	5	13	予測2	0	0	0	0	0	0	0	0	0	0	0
予測3	0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0	0	0	0	0
予測4	0	0	0	11	0	12	23	予測4	4	19	0	184	4	19	0	184	0	146	353
予測5	57	31	0	94	0	456	638	予測5	0	0	0	0	0	0	0	0	0	0	0
予測6	0	0	0	0	0	0	0	予測6	3	33	0	312	3	33	0	312	0	527	875
計	69	36	0	112	0	489	0.68	計	11	56	0	511	11	56	0	511	0	697	
Training dataの あてはまり																			
テスト																			
予測1	2	1	0	4	0	18	25	予測1	2	3	0	18	2	3	0	18	0	34	57
予測2	1	0	0	2	0	12	15	予測2	0	0	0	0	0	0	0	0	0	0	0
予測3	0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0	0	0	0	0
予測4	2	4	0	1	0	23	30	予測4	5	9	0	147	5	9	0	147	0	158	319
予測5	54	29	0	107	0	447	637	予測5	0	0	0	0	0	0	0	0	0	0	0
予測6	0	0	0	0	0	0	0	予測6	12	23	0	318	12	23	0	318	0	546	899
計	59	34	0	114	0	500	0.636	計	19	35	0	483	19	35	0	483	0	738	
Test dataの あてはまり																			

付表 4-2 線形判別関数による分類: 説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

男性	15歳～19歳 トレーニング						女性						15歳～19歳 トレーニング					
	1	2	3	4	5	6 計	1	2	3	4	5	6 計	1	2	3	4	5	6 計
予測1	123	0	0	3	34	22	186	予測1	84	11	7	8	18	12	140	0	0	0
予測2	0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0	0	0	0
予測3	0	0	88	0	120	0	208	予測3	1	0	138	0	166	0	305	0	166	0
予測4	0	0	0	0	0	0	0	予測4	20	5	1	8	2	10	46	2	10	46
予測5	2	1	108	2	1767	7	1887	予測5	1	0	141	1	1644	3	1790	1	1644	3
予測6	31	1	0	2	1	22	57	予測6	0	0	0	0	0	0	0	0	0	0
計	156	3	199	7	1922	51	2338	計	106	16	287	17	1830	25	2281	16	287	17
Training dataの あてはまり																		
テスト																		
予測1	118	3	2	4	5	6	9	予測1	85	7	5	11	13	13	134	1	2	3
予測2	0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0	0	0	0
予測3	0	0	73	0	115	0	188	予測3	0	0	158	0	160	0	318	0	160	0
予測4	0	0	0	0	0	0	0	予測4	22	5	1	9	7	13	57	5	1	9
予測5	2	1	112	0	1810	4	1929	予測5	0	0	147	2	1616	7	1772	0	147	2
予測6	43	0	2	0	4	13	62	予測6	0	0	0	0	0	0	0	0	0	0
計	163	4	189	1	1955	26	2338	計	107	12	311	22	1796	33	2281	12	311	22
Test dataの あてはまり																		

付表 4-2 (続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

性別	年齢	観察値						観察値						計			
		1	2	3	4	5	6	計	1	2	3	4	5		6	計	
男性	20歳~24歳	女性												0.889			
	トレーニング	トレーニング												0.709			
予測1	おもに仕事	1822	17	5	15	13	149	2021	20	1	1022	46	10	128	4	69	1279
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	24	10	0	15	0	6	55
予測3	通学のかたわらに仕事	5	0	12	0	11	0	28	28	0	4	0	256	0	190	3	453
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	0	8	5	0	14	0	6	33
予測5	仕事をしない人のうち主に通学している人	7	0	7	0	11	1	26	26	0	3	0	50	0	86	0	139
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計		1834	17	24	15	35	150	2075	計		1061	61	316	157	280	84	1959
	Training dataのあてはまり	Training dataのあてはまり												0.709			
	テスト	テスト												0.710			
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計		
予測2	家事などのかたわらに仕事	1820	14	11	19	10	140	2014	20	1062	67	3	102	13	61	1308	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	14	3	0	21	0	6	44		
予測4	仕事をしない人のうち主に家事をしている人	4	0	13	0	18	2	37	4	0	235	1	188	1	429		
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	19	7	0	16	0	5	47	
予測6	その他の仕事をしていない人	3	1	6	0	14	0	24	4	1	52	0	75	0	132		
計		1827	15	30	19	42	142	2075	計		1103	78	290	140	276	73	1960
	Test dataのあてはまり	Test dataのあてはまり												0.710			
男性	25歳~29歳	女性												0.659			
	トレーニング	トレーニング												0.710			
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計		
予測2	家事などのかたわらに仕事	1823	18	9	19	12	140	2021	20	1386	210	7	347	14	84	2048	
予測3	通学のかたわらに仕事	3	0	12	0	17	0	32	0	6	0	11	0	12	0	29	
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	37	22	0	41	0	3	103	
予測5	仕事をしない人のうち主に通学している人	3	1	2	0	15	0	21	0	0	0	2	0	3	0	5	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	3	0	0	0	0	0	3	
計		1830	19	23	19	44	140	2075	計		1432	232	20	388	29	87	2188
	Training dataのあてはまり	Training dataのあてはまり												0.659			
	テスト	テスト												0.645			
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計		
予測2	家事などのかたわらに仕事	1815	13	8	15	11	151	2013	20	1367	220	6	368	12	72	2045	
予測3	通学のかたわらに仕事	4	0	12	0	12	0	28	0	8	1	6	0	9	0	24	
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	37	36	0	32	0	2	107	
予測5	仕事をしない人のうち主に通学している人	7	0	11	0	10	1	29	0	1	0	2	0	7	0	10	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	2	0	0	1	0	0	3	
計		1831	13	31	15	33	152	2075	計		1415	257	14	401	28	74	2189
	Test dataのあてはまり	Test dataのあてはまり												0.645			

付表 4-2(続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

性別	年齢	年齢の2乗	教育ダミー	育児補助ダミー	観察値	1	2	3	4	5	6	計	観察値	1	2	3	4	5	6	計
男性	30歳~34歳					1	2	3	4	5	6	計		1	2	3	4	5	6	計
	トレーニング					2286	19	2	18	4	129	2458	予測1	1291	380	3	667	7	56	2404
	おもに仕事					0	0	0	0	0	0	0	予測2	20	25	0	27	0	2	74
	家事などのかたわらに仕事					5	0	3	0	4	0	12	予測3	1	0	3	0	1	0	5
	通学のかたわらに仕事					0	0	0	0	0	0	0	予測4	30	30	0	54	0	115	
	仕事をしない人のうち主に家事をしている人					4	0	2	0	3	0	9	予測5	1	2	2	0	6	0	11
	仕事をしない人のうち主に通学している人					0	0	0	0	0	0	0	予測6	1	0	0	0	0	2	3
	その他の仕事をしていない人					2295	19	7	18	11	129	2479	計	1344	437	8	748	14	61	2612
	Training dataのあてはまり					0.925							テスト	0.529						
	おもに仕事					1	2	3	4	5	6	計		1	2	3	4	5	6	計
	家事などのかたわらに仕事					2291	11	5	13	6	132	2458	予測1	1346	418	4	615	5	56	2444
	通学のかたわらに仕事					0	0	0	0	0	0	0	予測2	25	20	0	21	0	4	70
	仕事をしない人のうち主に家事をしている人					8	0	2	0	4	0	14	予測3	3	0	2	0	3	0	8
	仕事をしない人のうち主に通学している人					0	0	0	0	0	0	0	予測4	27	22	0	37	0	86	
	その他の仕事をしていない人					0	0	3	0	3	1	7	予測5	0	0	0	1	2	1	4
	Training dataのあてはまり					2299	11	10	13	13	133	2479	計	1401	460	0	674	10	61	2612
	テスト					0.926							Training dataのあてはまり	0.539						
男性	35歳~39歳					1	2	3	4	5	6	計		1	2	3	4	5	6	計
	トレーニング					3969	27	0	16	4	168	4184	予測1	1531	741	2	869	4	53	3200
	おもに仕事					0	1	0	0	0	0	0	予測2	0	0	0	0	0	0	0
	家事などのかたわらに仕事					3	0	1	1	2	0	7	予測3	0	0	1	0	0	1	0
	通学のかたわらに仕事					17	0	0	1	0	1	19	予測4	26	40	0	68	0	134	
	仕事をしない人のうち主に家事をしている人					3	0	0	0	1	1	5	予測5	3	0	0	0	2	0	5
	仕事をしない人のうち主に通学している人					0	0	0	0	0	0	0	予測6	0	0	0	0	0	1	1
	その他の仕事をしていない人					3992	28	1	18	7	170	4216	計	1560	781	3	937	6	54	3341
	Training dataのあてはまり					0.942							Training dataのあてはまり	0.480						
	おもに仕事					1	2	3	4	5	6	計		1	2	3	4	5	6	計
	家事などのかたわらに仕事					1975	12	1	15	6	82	2091	予測1	1506	763	2	863	4	79	3217
	通学のかたわらに仕事					0	0	0	0	0	0	0	予測2	0	0	0	1	0	0	1
	仕事をしない人のうち主に家事をしている人					3	0	0	0	0	0	3	予測3	0	0	0	0	0	0	0
	仕事をしない人のうち主に通学している人					13	0	0	0	0	0	13	予測4	32	28	0	54	2	117	
	その他の仕事をしていない人					2	0	0	0	0	0	2	予測5	1	1	2	1	1	0	6
	Training dataのあてはまり					1993	12	1	15	6	82	2109	計	1539	792	4	919	7	80	3341
	テスト					0.937							Training dataのあてはまり	0.467						

付表 4-2(続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

性別	年齢	観察値						観察値	観察値								
		1	2	3	4	5	6		計	1	2	3	4	5	6	計	
男性	40歳~44歳	女性						女性									
	トレーニング	トレーニング						トレーニング									
予測1	おもに仕事	3996	17	3	27	2	191	4236	0	0	0	0	0	0	0	0	24
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	2	予測2	家事などのかたわらに仕事	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	2	0	0	0	0	0	予測3	通学のかたわらに仕事	1	1	1	0	0	4
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	2	5	0	2	0	9
予測5	仕事をしない人のうち主に通学している人	22	0	0	0	1	1	24	予測5	仕事をしない人のうち主に通学している人	18	5	0	7	2	0	32
予測6	その他の仕事をしない人	0	0	0	0	0	0	0	0	予測6	その他の仕事をしない人	3	0	0	0	0	4
計		4018	17	5	27	3	192	4262	計	1661	936	4	818	6	49	3474	
	Training dataのあてはまり	0.938							Training dataのあてはまり	0.475							
	テスト	1	2	3	4	5	6	計	テスト	1	2	3	4	5	6	計	
予測1	おもに仕事	1989	15	0	11	0	102	2117	予測1	おもに仕事	1580	969	1	802	2	46	3400
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	4	8	1	11	0	0	24
予測3	通学のかたわらに仕事	1	0	0	0	0	0	0	予測3	通学のかたわらに仕事	4	1	1	2	0	0	8
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	3	1	0	4	0	0	8
予測5	仕事をしない人のうち主に通学している人	12	0	0	0	0	1	13	予測5	仕事をしない人のうち主に通学している人	15	5	2	10	1	1	34
予測6	その他の仕事をしない人	0	0	0	0	0	0	0	予測6	その他の仕事をしない人	1	0	0	0	0	0	1
計		2002	15	0	11	0	103	2131	計	1607	984	5	829	3	47	3475	
	Test dataのあてはまり	0.933							Test dataのあてはまり	0.459							
男性	45歳~49歳	女性						女性									
	トレーニング	トレーニング						トレーニング									
予測1	おもに仕事	3610	25	0	26	2	169	3832	予測1	おもに仕事	1568	940	0	620	5	48	3181
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	2	0	1	0	0	0	3	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	2	0	0	2	0	0	4
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	1	0	0	1	4	0	6
予測6	その他の仕事をしない人	1	0	0	0	0	1	2	予測6	その他の仕事をしない人	0	0	0	0	0	1	1
計		3613	25	1	26	2	170	3837	計	1571	940	0	623	9	49	3192	
	Training dataのあてはまり	0.941							Training dataのあてはまり	0.493							
	テスト	1	2	3	4	5	6	計	テスト	1	2	3	4	5	6	計	
予測1	おもに仕事	1808	11	0	17	0	82	1918	予測1	おもに仕事	1591	932	2	602	3	56	3186
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	2	0	1	0	1	0	4
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	1	1	0	0	0	0	2
予測6	その他の仕事をしない人	0	1	0	0	0	0	1	予測6	その他の仕事をしない人	0	0	0	0	0	0	0
計		1808	12	0	17	0	82	1919	計	1594	933	3	602	4	56	3192	
	Test dataのあてはまり	0.942							Test dataのあてはまり	0.488							

付表 4-2(続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

性別	50歳～54歳 トレーニング	50歳～54歳 トレーニング	観察値						観察値						6計							
			1	2	3	4	5	6計	1	2	3	4	5	6計								
男性	3745	30	0	35	2	175	3987	予測1	おもに仕事	0	0	0	0	0	0	1694	853	0	695	2	36	3280
予測2	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	1	0	0	1	0	0	2
予測5	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	4	0	0	0	0	1	0	5
予測6	0	0	0	0	0	0	0	予測6	その他の仕事をしない人	10	5	0	0	0	10	5	0	5	0	3	23	
計	3745	30	0	35	2	175	3987	計		1709	858	0	701	3	39	3310						
	0.939							Training dataのあてはまり		0.513												
男性	1891	19	1	17	1	65	1994	予測1	おもに仕事	1647	876	0	714	0	35	3272						
予測2	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0						
予測3	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0						
予測4	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	2	0	0	1	0	0	3						
予測5	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	1	1	1	1	0	4						
予測6	0	0	0	0	0	0	0	予測6	その他の仕事をしない人	17	11	0	2	0	1	31						
計	1891	19	1	17	1	65	1994	計		1666	888	1	718	1	36	3310						
	0.948							Test dataのあてはまり		0.499												
男性	4164	68	1	73	1	270	4577	予測1	おもに仕事	1575	917	0	1125	1	65	3683						
予測2	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	8	14	0	9	0	1	32						
予測3	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0						
予測4	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	1	0	0	2	0	0	3						
予測5	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0						
予測6	0	0	0	0	0	0	0	予測6	その他の仕事をしない人	5	1	0	2	0	0	8						
計	4164	68	1	73	1	270	4577	計		1589	932	0	1138	1	66	3726						
	0.910							Training dataのあてはまり		0.427												
男性	2091	33	1	28	0	136	2289	予測1	おもに仕事	1568	954	1	1085	0	70	3678						
予測2	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	13	9	0	10	0	0	32						
予測3	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0						
予測4	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	4	0	0	0	0	2	6						
予測5	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0						
予測6	0	0	0	0	0	0	0	予測6	その他の仕事をしない人	7	1	0	3	0	0	11						
計	2091	33	1	28	0	136	2289	計		1592	964	1	1098	0	72	3727						
	0.914							Test dataのあてはまり		0.423												

付表 4-2(続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

性別	60歳～64歳 トレーニング						65歳～69歳 トレーニング						観察値					
	1	2	3	4	5	6 計	1	2	3	4	5	6 計	1	2	3	4	5	6 計
男性	60歳～64歳 トレーニング						65歳～69歳 トレーニング						観察値					
予測1	3253	173	0	0	0	0	3259	173	0	0	0	0	0	0	0	0	0	0
予測2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	4	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
予測4	1	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	1	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0
計	3259	173	2	196	2	858	4490	計	1308	1075	3	2128	0	181	4695	0.451		
	Training dataのあてはまり						Training dataのあてはまり						観察値					
予測1	1	2	3	4	5	6 計	1	2	3	4	5	6 計	1	2	3	4	5	6 計
予測2	3248	219	0	0	0	0	3248	219	0	0	0	0	0	0	0	0	0	0
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	2	0	0	0	0	1	2	0	0	0	0	0	17	11	0	15	0	4
予測5	0	0	0	0	0	0	0	0	0	0	0	0	1341	1068	0	2027	0	193
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	3255	219	2	184	3	828	4491	計	1363	1086	0	2048	0	199	4696	0.433		
	Test dataのあてはまり						Test dataのあてはまり						観察値					
男性	65歳～69歳 トレーニング						65歳～69歳 トレーニング						観察値					
予測1	1979	356	1	0	0	0	1979	356	1	0	0	0	1	2	0	0	0	0
予測2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
予測3	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	0	0	0	0	0	514	676	0	1967	0	229
予測5	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
予測6	605	114	1	129	0	702	1551	計	514	676	0	1967	0	229	3386	0	0	0
計	2585	471	4	439	2	2154	5655	計	514	676	0	1967	0	229	3386	0.581		
	Training dataのあてはまり						Training dataのあてはまり						観察値					
予測1	1	2	3	4	5	6 計	1	2	3	4	5	6 計	1	2	3	4	5	6 計
予測2	204	36	0	0	0	0	204	36	0	0	0	0	0	0	0	0	0	0
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測5	0	0	0	0	0	0	0	0	0	0	0	0	505	651	1	1975	0	255
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	56	15	0	14	0	69	154	計	505	651	1	1975	0	255	3387	0.482		
	Test dataのあてはまり						Test dataのあてはまり						観察値					

付表 4-2(続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

男性	70歳~74歳						75歳~79歳						70歳~74歳						75歳~79歳					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
男性	Training dataのあてはまり												Test dataのあてはまり											
予測1	おもに仕事												おもに仕事											
予測2	家事などのかたわらに仕事												家事などのかたわらに仕事											
予測3	通学のかたわらに仕事												通学のかたわらに仕事											
予測4	仕事をしない人のうち主に家事をしている人												仕事をしない人のうち主に家事をしている人											
予測5	仕事をしない人のうち主に通学している人												仕事をしない人のうち主に通学している人											
予測6	その他の仕事をしない人												その他の仕事をしない人											
計	106	18	0	29	0	107	260	260	0	107	260	260	1	2	3	4	5	6	304	537	0	2096	0	381
	0.485												0.632						0.632					
予測1	おもに仕事												おもに仕事											
予測2	家事などのかたわらに仕事												家事などのかたわらに仕事											
予測3	通学のかたわらに仕事												通学のかたわらに仕事											
予測4	仕事をしない人のうち主に家事をしている人												仕事をしない人のうち主に家事をしている人											
予測5	仕事をしない人のうち主に通学している人												仕事をしない人のうち主に通学している人											
予測6	その他の仕事をしない人												その他の仕事をしない人											
計	100	23	0	32	0	107	262	262	0	107	262	262	1	2	3	4	5	6	289	488	0	2154	1	380
	0.491												0.650						0.650					
男性	Training dataのあてはまり												Test dataのあてはまり											
予測1	おもに仕事												おもに仕事											
予測2	家事などのかたわらに仕事												家事などのかたわらに仕事											
予測3	通学のかたわらに仕事												通学のかたわらに仕事											
予測4	仕事をしない人のうち主に家事をしている人												仕事をしない人のうち主に家事をしている人											
予測5	仕事をしない人のうち主に通学している人												仕事をしない人のうち主に通学している人											
予測6	その他の仕事をしない人												その他の仕事をしない人											
計	524	162	1	288	0	1322	2297	2311	0	1322	2297	2311	1	2	3	4	5	6	136	304	0	1890	2	600
	0.575												0.637						0.637					
予測1	おもに仕事												おもに仕事											
予測2	家事などのかたわらに仕事												家事などのかたわらに仕事											
予測3	通学のかたわらに仕事												通学のかたわらに仕事											
予測4	仕事をしない人のうち主に家事をしている人												仕事をしない人のうち主に家事をしている人											
予測5	仕事をしない人のうち主に通学している人												仕事をしない人のうち主に通学している人											
予測6	その他の仕事をしない人												その他の仕事をしない人											
計	535	145	1	299	0	1318	2298	2312	0	1318	2298	2312	1	2	3	4	5	6	133	311	0	1873	1	614
	0.572												0.631						0.631					

付表 4-2(続) 線形判別関数による分類・説明変数は、年齢、年齢の2乗、教育ダミー、育児補助ダミー(2011年)

性別	年齢	観察値						観察値									
		1	2	3	4	5	6	計	1	2	3	4	5	6	計		
男性	80歳~84歳	Training dataのあてはまり															
	トレーニング	テスト															
	予測1	1	2	3	4	5	6	計	0.644	1	2	3	4	5	6	計	
	予測2	4	0	0	2	0	0	2	0.644	4	2	0	0	0	3	9	計
	予測3	0	0	0	0	0	0	0	0.644	0	0	0	0	0	0	0	計
	予測4	1	0	0	4	0	0	6	0.644	1	0	0	4	0	0	11	計
	予測5	0	0	0	0	0	0	0	0.644	0	0	0	0	0	0	0	計
	予測6	261	72	0	221	0	1013	1567	0.644	261	72	0	221	0	90	111	計
	計	266	72	0	227	0	1021	1586	0.644	266	72	0	227	0	90	111	計
		Training dataのあてはまり															
	テスト	テスト															
	予測1	1	2	3	4	5	6	計	0.571	1	2	3	4	5	6	計	
	予測2	4	2	0	0	0	0	3	0.571	4	2	0	0	0	1	11	計
	予測3	0	0	0	0	0	0	0	0.571	0	0	0	0	0	0	0	計
	予測4	0	0	0	0	0	0	0	0.571	0	0	0	0	0	0	0	計
	予測5	4	0	0	2	0	11	17	0.571	4	0	0	2	0	0	17	計
	予測6	0	0	0	0	0	0	0	0.571	0	0	0	0	0	0	0	計
	計	251	80	0	205	0	1024	1560	0.571	251	80	0	205	0	117	118	計
		259	82	0	207	0	1038	1586	0.571	259	82	0	207	0	1204	1204	計
		Test dataのあてはまり															
男性	85歳~	Training dataのあてはまり															
	トレーニング	テスト															
	予測1	1	2	3	4	5	6	計	0.546	1	2	3	4	5	6	計	
	予測2	1	0	0	0	0	0	0	0.546	1	0	0	0	0	0	0	計
	予測3	0	0	0	0	0	0	0	0.546	0	0	0	0	0	0	0	計
	予測4	0	0	0	0	0	0	0	0.546	0	0	0	0	0	0	0	計
	予測5	0	0	0	0	0	0	0	0.546	0	0	0	0	0	0	0	計
	予測6	77	45	0	166	0	653	941	0.546	77	45	0	166	0	542	861	計
	計	78	45	0	166	0	653	942	0.546	78	45	0	166	0	664	955	計
		Training dataのあてはまり															
	テスト	テスト															
	予測1	1	2	3	4	5	6	計	0.578	1	2	3	4	5	6	計	
	予測2	0	0	0	0	0	0	0	0.578	0	0	0	0	0	0	0	計
	予測3	0	0	0	0	0	0	0	0.578	0	0	0	0	0	0	0	計
	予測4	0	0	0	0	0	0	0	0.578	0	0	0	0	0	0	0	計
	予測5	0	0	0	0	0	0	0	0.578	0	0	0	0	0	0	0	計
	予測6	50	25	0	60	0	336	471	0.578	50	25	0	60	0	274	419	計
	計	50	25	0	60	0	336	471	0.578	50	25	0	60	0	330	480	計
		Test dataのあてはまり															

付表 4-3 サポート・ベクトル・マシンによる分類:説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

性別	年齢	観察値						観測値											
		1	2	3	4	5	6	計	1	2	3	4	5	6	計				
男性	15歳~19歳													観測値					
	トレーニング													1 2 3 4 5 6 計					
	おもに仕事	115	3	30	0	86	18	252	0	0	0	0	0	0	0	0	0	0	
	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	通学のかたわらに仕事	0	0	22	0	20	0	42	0	0	0	74	0	68	0	142	0	0	
	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	仕事をしていない人のうち主に通学している人	21	1	171	2	1832	17	2044	0	0	0	238	12	1745	30	2139	0	0	
	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
計		136	4	223	2	1988	35	2338	0.842	103	11	312	12	1813	30	2281	0.798		
	Training dataのあてはまり													Training dataのあてはまり					
	テスト													テスト					
	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計	0	0	0	
	家事などのかたわらに仕事	147	1	26	4	109	27	314	0	0	0	0	0	0	0	0	0	0	
	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	仕事をしていない人のうち主に家事をしている人	0	0	22	0	23	0	45	0	0	0	70	0	55	0	126	1	0	
	仕事をしていない人のうち主に通学している人	36	2	117	2	1807	15	1979	0.845	109	17	216	27	1758	28	2155	0.801		
	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
計		183	3	165	6	1939	42	2338	0.845	110	17	286	27	1813	28	2281	0.801		
	Test dataのあてはまり													Test dataのあてはまり					
	トレーニング													トレーニング					
男性	20歳~24歳													観測値					
	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計	0	0	0	
	家事などのかたわらに仕事	981	17	57	14	97	135	1301	0	0	0	0	0	0	0	0	0	0	
	通学のかたわらに仕事	75	2	240	1	219	5	542	79	2	180	4	120	6	391	0	1	0	
	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	8	9	0	17	1	7	42	0	0		
	仕事をしていない人のうち主に通学している人	0	0	5	0	10	0	15	0	0	0	0	0	1	1	0	0		
	その他の仕事をしていない人	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0		
計		1056	19	302	15	326	141	1859	0.663	1072	73	306	146	280	82	1959	0.605		
	Training dataのあてはまり													Training dataのあてはまり					
	テスト													テスト					
	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計	0	0	0	
	家事などのかたわらに仕事	979	19	57	20	102	121	1298	972	61	135	129	153	66	1516	0	0	0	
	通学のかたわらに仕事	72	1	233	0	232	7	545	108	2	161	7	118	5	401	0	0	0	
	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	12	3	1	15	0	4	35	0	0		
	仕事をしていない人のうち主に通学している人	0	0	8	0	8	0	16	0	0	3	0	5	0	8	0	0		
	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
計		1051	20	298	20	342	128	1859	0.656	1092	66	300	151	276	75	1960	0.588		
	Test dataのあてはまり													Test dataのあてはまり					

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

性別	年齢	観察値						観察値	観察値										
		1	2	3	4	5	6		計	1	2	3	4	5	6	計			
男性	25歳~29歳	女性												25歳~29歳					
	トレーニング	トレーニング												トレーニング					
	予測1	1812	15	34	11	43	160	2075	1416	220	20	344	26	79	2105				
	予測2	0	0	0	0	0	0	0	1	7	0	2	0	0	10				
	予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測4	0	0	0	0	0	0	0	21	15	1	30	1	5	73				
	予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	計	1812	15	34	11	43	160	2075	1438	242	21	376	27	84	2188				
		0.873	Training dataのあてはまり												0.664				
	テスト	テスト												テスト					
	予測1	1849	17	20	23	34	132	2075	1372	230	13	384	30	73	2102				
	予測2	0	0	0	0	0	0	0	6	4	0	5	0	0	15				
	予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測4	0	0	0	0	0	0	0	31	13	0	24	0	4	72				
	予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	計	1849	17	20	23	34	132	2075	1409	247	13	413	30	77	2189				
		0.891	Test dataのあてはまり												0.640				
男性	30歳~34歳	女性												30歳~34歳					
	トレーニング	トレーニング												トレーニング					
	予測1	2285	19	10	17	15	133	2479	1323	408	6	662	16	61	2476				
	予測2	0	0	0	0	0	0	0	16	28	0	15	0	4	63				
	予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測4	0	0	0	0	0	0	0	19	15	0	37	1	1	73				
	予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	計	2285	19	10	17	15	133	2479	1358	451	6	714	17	66	2612				
		0.922	Training dataのあてはまり												0.531				
	テスト	テスト												テスト					
	予測1	2309	11	7	14	9	129	2479	1331	406	7	653	6	46	2449				
	予測2	0	0	0	0	0	0	0	24	18	1	21	1	5	70				
	予測3	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測4	0	0	0	0	0	0	0	32	22	0	34	0	5	93				
	予測5	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	予測6	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	計	2309	11	7	14	9	129	2479	1387	446	8	708	7	56	2612				
		0.931	Test dataのあてはまり												0.530				

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

性別	年齢	観察値						観察値	観察値										
		1	2	3	4	5	6		計	1	2	3	4	5	6	計			
男性	35歳~39歳	女性												35歳~39歳					
	トレーニング	トレーニング												トレーニング					
	予測1	予測1												予測1					
	予測2	予測2												予測2					
	予測3	予測3												予測3					
	予測4	予測4												予測4					
	予測5	予測5												予測5					
	予測6	予測6												予測6					
	計	計												計					
	Training dataのあてはまり	Training dataのあてはまり												Training dataのあてはまり					
	テスト	テスト												テスト					
	予測1	予測1												予測1					
	予測2	予測2												予測2					
	予測3	予測3												予測3					
	予測4	予測4												予測4					
	予測5	予測5												予測5					
	予測6	予測6												予測6					
	計	計												計					
	Test dataのあてはまり	Test dataのあてはまり												Test dataのあてはまり					
	40歳~44歳	女性												40歳~44歳					
	トレーニング	トレーニング												トレーニング					
	予測1	予測1												予測1					
	予測2	予測2												予測2					
	予測3	予測3												予測3					
	予測4	予測4												予測4					
	予測5	予測5												予測5					
	予測6	予測6												予測6					
	計	計												計					
	Training dataのあてはまり	Training dataのあてはまり												Training dataのあてはまり					
	テスト	テスト												テスト					
	予測1	予測1												予測1					
	予測2	予測2												予測2					
	予測3	予測3												予測3					
	予測4	予測4												予測4					
	予測5	予測5												予測5					
	予測6	予測6												予測6					
	計	計												計					
	Test dataのあてはまり	Test dataのあてはまり												Test dataのあてはまり					

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

性別	年齢	観察値						観察値	1	2	3	4	5	6	計
		1	2	3	4	5	6								
男性	45歳~49歳							1566	951	3	621	5	46	3192	
	トレーニング							0.943							
	テスト							1804	920	0	603	8	59	3188	
予測1	おもに仕事	0	0	0	0	0	0	1	2	3	4	5	6	計	
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	1598	920	0	603	8	59	3188	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	1	2	0	1	0	0	4	
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	0	0	0	0	0	
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0	0	
計		3617	24	0	29	0	167	3837	計						
Training dataのあてはまり															
	テスト							0.940							
予測1	おもに仕事	1	2	3	4	5	6	計							
予測2	家事などのかたわらに仕事	1804	13	1	14	2	85	1919	予測1	0	603	8	59	3188	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測2	1598	920	0	603	8	
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測3	1	2	0	1	0	
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	予測4	0	0	0	0	0	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	予測5	0	0	0	0	0	
計		1804	13	1	14	2	85	1919	予測6	0	0	0	0	0	
Test dataのあてはまり															
	トレーニング							1648	870	0	720	2	27	3267	
予測1	おもに仕事	1	2	3	4	5	6	計							
予測2	家事などのかたわらに仕事	2823	21	0	28	0	118	2990	予測1	1648	870	0	720	2	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測2	10	19	0	9	0	
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測3	0	0	0	0	0	
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	予測4	0	0	0	0	0	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	予測5	0	0	0	0	0	
計		2823	21	0	28	0	118	2990	予測6	0	0	0	0	0	
Training dataのあてはまり															
	テスト							1658	889	0	730	2	31	3310	
予測1	おもに仕事	1	2	3	4	5	6	計							
予測2	家事などのかたわらに仕事	2813	28	1	24	3	122	2991	予測1	1658	889	0	730	2	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測2	0.504					
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測3	1684	843	1	674	2	
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	予測4	32	13	0	15	0	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	予測5	0	0	0	0	0	
計		2813	28	1	24	3	122	2991	予測6	0	0	0	0	0	
Test dataのあてはまり															
	トレーニング							1717	857	1	689	2	44	3310	
予測1	おもに仕事	1	2	3	4	5	6	計							
予測2	家事などのかたわらに仕事	1717	857	1	689	2	44	3310	予測1	1	2	3	4	5	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測2	1564	948	3	617	5	
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測3	2	3	0	1	0	
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	予測4	0	0	0	0	0	
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	予測5	0	0	0	0	0	
計		1717	857	1	689	2	44	3310	予測6	0	0	0	0	0	
Test dataのあてはまり															
	テスト							0.513							

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

性別	年齢	観察値						観察値	1	2	3	4	5	6	計				
		1	2	3	4	5	6												
男性	55歳~59歳	女性						55歳~59歳						観察値					
	トレーニング	トレーニング						トレーニング						観察値					
	おもに仕事	おもに仕事						おもに仕事						観察値					
	予測1	予測1						予測1						観察値					
	予測2	予測2						予測2						観察値					
	予測3	予測3						予測3						観察値					
	予測4	予測4						予測4						観察値					
	予測5	予測5						予測5						観察値					
	予測6	予測6						予測6						観察値					
	計	計						計						計					
	Training dataのあてはまり	Training dataのあてはまり						Training dataのあてはまり						Training dataのあてはまり					
	テスト	テスト						テスト						テスト					
	おもに仕事	おもに仕事						おもに仕事						おもに仕事					
	予測1	予測1						予測1						予測1					
	予測2	予測2						予測2						予測2					
	予測3	予測3						予測3						予測3					
	予測4	予測4						予測4						予測4					
	予測5	予測5						予測5						予測5					
	予測6	予測6						予測6						予測6					
	計	計						計						計					
	Test dataのあてはまり	Test dataのあてはまり						Test dataのあてはまり						Test dataのあてはまり					
	60歳~64歳	女性						60歳~64歳						観察値					
	トレーニング	トレーニング						トレーニング						観察値					
	おもに仕事	おもに仕事						おもに仕事						観察値					
	予測1	予測1						予測1						予測1					
	予測2	予測2						予測2						予測2					
	予測3	予測3						予測3						予測3					
	予測4	予測4						予測4						予測4					
	予測5	予測5						予測5						予測5					
	予測6	予測6						予測6						予測6					
	計	計						計						計					
	Training dataのあてはまり	Training dataのあてはまり						Training dataのあてはまり						Training dataのあてはまり					
	テスト	テスト						テスト						テスト					
	おもに仕事	おもに仕事						おもに仕事						おもに仕事					
	予測1	予測1						予測1						予測1					
	予測2	予測2						予測2						予測2					
	予測3	予測3						予測3						予測3					
	予測4	予測4						予測4						予測4					
	予測5	予測5						予測5						予測5					
	予測6	予測6						予測6						予測6					
	計	計						計						計					
	Test dataのあてはまり	Test dataのあてはまり						Test dataのあてはまり						Test dataのあてはまり					

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

性別	年齢	観察値						観察値	観察値												
		1	2	3	4	5	6		計	1	2	3	4	5	6	計					
男性	65歳~69歳	Training dataのあてはまり												0.472							
	トレーニング	テスト												0.585							
予測1	おもに仕事	1111	198	3	177	2	830	2321	0	0	0	0	0	0	5	0	0	0	5		
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	4	10	0	6	0	1	21
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	1	0	1	0	0	2	2	2	2	2	2	2	482	664	0	1964	0	250	3360
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしない人	302	58	1	71	0	355	787	0	0	0	0	0	0	0	0	0	0	0	0	0
計		1413	257	4	249	2	1185	3110	計	491	674	0	1970	0	251	3386					
	65歳~69歳	Training dataのあてはまり												0.472							
	テスト	テスト												0.585							
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計						
予測2	家事などのかたわらに仕事	1122	201	2	170	0	804	2299	1	2	0	6	0	0	9						
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	3	3	0	11	0	2	19						
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
予測6	その他の仕事をしない人	310	64	0	67	0	371	812	524	648	1	1955	0	231	3359						
計		1432	265	2	237	0	1175	3111	計	528	653	1	1972	0	233	3387					
	70歳~74歳	Training dataのあてはまり												0.480							
	トレーニング	テスト												0.578							
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計						
予測2	家事などのかたわらに仕事	31	2	0	2	0	16	51	0	0	0	0	0	0	0						
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	5	0	1	0	1	7						
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	291	502	0	2129	0	389	3311						
予測6	その他の仕事をしない人	914	212	3	271	0	1377	2777	0	0	0	0	0	0	0						
計		945	214	3	273	0	1393	2828	計	291	507	0	2130	0	390	3318					
	70歳~74歳	Training dataのあてはまり												0.498							
	テスト	テスト												0.643							
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計						
予測2	家事などのかたわらに仕事	14	1	0	5	0	16	36	0	0	0	0	0	0	0						
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	2	2	0	6	0	3	13						
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	300	518	0	2118	1	368	3305						
予測6	その他の仕事をしない人	893	221	2	317	0	1359	2792	0	0	0	0	0	0	0						
計		907	222	2	322	0	1375	2828	計	302	520	0	2124	1	371	3318					
	70歳~74歳	Training dataのあてはまり												0.486							
	テスト	テスト												0.639							

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

男性		75歳～79歳						75歳～79歳						観察値							
予測1		トレーニング						トレーニング						観察値							
予測2		おもに仕事						おもに仕事						観察値							
予測3		家事などのかたわらに仕事						家事などのかたわらに仕事						観察値							
予測4		通学のかたわらに仕事						通学のかたわらに仕事						観察値							
予測5		仕事をしない人のうち主に家事をしている人						仕事をしない人のうち主に家事をしている人						観察値							
予測6		仕事をしない人のうち主に通学している人						仕事をしない人のうち主に通学している人						観察値							
計		0.575						0.575						0.646							
		Training dataのあてはまり						Training dataのあてはまり						観察値							
予測1		テスト						テスト						観察値							
予測2		おもに仕事						おもに仕事						観察値							
予測3		家事などのかたわらに仕事						家事などのかたわらに仕事						観察値							
予測4		通学のかたわらに仕事						通学のかたわらに仕事						観察値							
予測5		仕事をしない人のうち主に家事をしている人						仕事をしない人のうち主に家事をしている人						観察値							
予測6		仕事をしない人のうち主に通学している人						仕事をしない人のうち主に通学している人						観察値							
計		0.575						0.575						0.638							
		Test dataのあてはまり						Test dataのあてはまり						観察値							
予測1		トレーニング						トレーニング						観察値							
予測2		おもに仕事						おもに仕事						観察値							
予測3		家事などのかたわらに仕事						家事などのかたわらに仕事						観察値							
予測4		通学のかたわらに仕事						通学のかたわらに仕事						観察値							
予測5		仕事をしない人のうち主に家事をしている人						仕事をしない人のうち主に家事をしている人						観察値							
予測6		仕事をしない人のうち主に通学している人						仕事をしない人のうち主に通学している人						観察値							
計		0.649						0.649						0.572							
		Training dataのあてはまり						Training dataのあてはまり						観察値							
予測1		テスト						テスト						観察値							
予測2		おもに仕事						おもに仕事						観察値							
予測3		家事などのかたわらに仕事						家事などのかたわらに仕事						観察値							
予測4		通学のかたわらに仕事						通学のかたわらに仕事						観察値							
予測5		仕事をしない人のうち主に家事をしている人						仕事をしない人のうち主に家事をしている人						観察値							
予測6		仕事をしない人のうち主に通学している人						仕事をしない人のうち主に通学している人						観察値							
計		0.649						0.649						0.554							
		Test dataのあてはまり						Test dataのあてはまり						観察値							
1	2	3	4	5	6	計	1	2	3	4	5	6	計	1	2	3	4	5	6	計	
7	2	0	0	0	2	11	予測1	おもに仕事	0	0	0	0	0	0	1	0	0	0	0	1	
0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	128	311	0	1893	2	597	2931	0	0	0	0	0	
0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	
541	152	0	286	0	1321	2300	予測6	その他の仕事をしていない人	129	311	0	1893	2	597	2932	0	0	0	0	0	
548	154	0	286	0	1323	2311	計	0.575	0.575	0.646											
1	2	3	4	5	6	計	予測1	おもに仕事	1	2	3	4	5	6	計	予測2	家事などのかたわらに仕事	0	0	0	1
4	1	0	0	0	3	8	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	140	304	0	1870	1	616	2931	0	0	0	0	0	
0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	
517	155	2	304	0	1326	2304	予測6	その他の仕事をしていない人	140	304	0	1870	1	616	2932	0	0	0	0	0	
521	156	2	304	0	1329	2312	計	0.575	0.575	0.638											
1	2	3	4	5	6	計	予測1	おもに仕事	1	2	3	4	5	6	計	予測2	家事などのかたわらに仕事	0	0	0	1
2	1	0	0	0	1	4	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	47	136	0	1126	0	656	1965	0	0	0	0	0	
0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	
260	72	0	222	0	1028	1582	予測6	その他の仕事をしていない人	4	8	0	90	0	129	231	0	0	0	0	0	
262	73	0	222	0	1029	1586	計	0.649	0.649	0.572											
1	2	3	4	5	6	計	予測1	おもに仕事	1	2	3	4	5	6	計	予測2	家事などのかたわらに仕事	0	0	0	0
1	0	0	0	0	2	3	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	55	136	0	1121	2	661	1975	0	0	0	0	0	
0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	
262	81	0	212	0	1028	1583	予測6	その他の仕事をしていない人	5	14	0	107	0	96	222	0	0	0	0	0	
263	81	0	212	0	1030	1586	計	0.649	0.649	0.554											

付表 4-3(続) サポート・ベクトル・マシンによる分類・説明変数は、年齢、年齢の2乗、教育年数、育児補助ダミー(2011年)

男性	85歳～ トレーニング	観察値						女性	85歳～ トレーニング	観察値												
		1	2	3	4	5	6			計	1	2	3	4	5	6	計					
予測1	おもに仕事	1	0	0	0	0	0	1	予測1	おもに仕事	1	0	0	0	0	0	1	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	1	0	0	3	0	2	6	予測4	仕事をしない人のうち主に家事をしている人	4	10	0	176	0	159	349					
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	
予測6	その他の仕事をしない人	56	32	0	104	0	507	699	予測6	その他の仕事をしない人	9	36	0	307	0	574	926					
計		58	32	0	107	0	509	706	計		13	46	0	483	0	733	1275					
Training dataの あてはまり		0.588																				
テスト		0.577																				
予測1	おもに仕事	1	2	3	4	5	6	計	1	2	3	4	5	6	計							
予測2	家事などのかたわらに仕事	0	1	0	0	0	1	2	予測1	おもに仕事	0	0	0	0	0	0						
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0						
予測4	仕事をしない人のうち主に家事をしている人	1	0	0	0	0	7	8	予測3	通学のかたわらに仕事	0	0	0	0	0	0						
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	7	16	0	172	0	139	334					
予測6	その他の仕事をしない人	69	37	0	119	0	472	697	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0						
計		70	38	0	119	0	480	707	計		10	29	0	339	0	563	941					
Test dataの あてはまり		0.668																				

付表 4-4 線形判別関数による将来予測・説明変数は、年齢、年齢の2乗、教育ダミー

男性	15歳～19歳 トレーニング	観察値						女性	15歳～19歳 トレーニング	観察値							
		1	2	3	4	5	6			計	1	2	3	4	5	6	計
予測1	おもに仕事	1034	11	312	8	556	140	2061	予測1	おもに仕事	804	20	286	58	521	43	1732
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	895	10	527	3	5888	170	7493	予測4	仕事をしない人のうち主に家事をしている人	127	25	2	70	9	43	276
予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	予測5	仕事をしない人のうち主に通学している人	667	23	702	67	5666	71	7196
予測6	その他の仕事をしない人	1929	21	839	11	6444	310	9554	予測6	その他の仕事をしない人	1598	68	990	195	6196	157	9204
計		0.725	0.711														
Training dataの あてはまり		0.711															
2011年 テスト		0.654															
予測1	おもに仕事	115	1	47	2	197	28	390	予測1	おもに仕事	84	9	69	8	96	17	283
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	9	0	3	6	636	3	657
予測5	仕事をしない人のうち主に通学している人	204	6	338	5	3047	47	3647	予測5	仕事をしない人のうち主に通学している人	120	19	526	25	2893	38	3621
予測6	その他の仕事をしない人	0	0	3	1	633	2	639	予測6	その他の仕事をしない人	0	0	0	0	0	0	0
計		319	7	388	8	3877	77	4676	計		213	28	598	39	3626	58	4562
Test dataの あてはまり		0.677															

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性 20歳～24歳													
1996年 トレーニング													
観察値													
1	2	3	4	5	6	計	1	2	3	4	5	6	計
6936	44	627	14	1030	354	9005	予測1	0	0	0	0	0	6653
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	193	73	3	193	1	18
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0
6936	44	627	14	1030	354	9005	計	6846	456	391	1307	546	244
0.770	Training dataの あてはまり												
2011年 テスト													
1	2	3	4	5	6	計	1	2	3	4	5	6	計
2033	36	429	35	515	252	3300	予測1	2089	131	446	281	421	140
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	6	6	0	10	0	4
74	3	171	0	153	17	418	予測5	69	2	160	6	135	13
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0
2107	39	600	35	668	269	3718	計	2164	139	606	297	556	157
0.588	Test dataの あてはまり												
男性 25歳～29歳													
1996年 トレーニング													
観察値													
1	2	3	4	5	6	計	1	2	3	4	5	6	計
7941	26	42	10	54	219	8292	予測1	3721	731	32	2253	34	122
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	835	362	0	1023	2	60
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0
7941	26	42	10	54	219	8292	計	4556	1093	32	3276	36	182
0.958	Training dataの あてはまり												
2011年 テスト													
1	2	3	4	5	6	計	1	2	3	4	5	6	計
3661	32	54	34	77	292	4150	予測1	2504	371	30	610	51	126
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	343	118	4	179	6	35
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0
3661	32	54	34	77	292	4150	計	2847	489	34	789	57	161
0.882	Test dataの あてはまり												

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性 30歳~34歳		観察値						観察値					
1996年 トレーニング		1	2	3	4	5	6	1	2	3	4	5	6
予測1	おもに仕事	8444	12	7	11	12	164	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	33	18	0	31	1	4
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	3452	1919	12	3816	17	97
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0
計		8444	12	7	11	12	164	3485	1937	12	3847	18	101
Training dataの あてはまり		0.976											
2011年 テスト		1	2	3	4	5	6	1	2	3	4	5	6
予測1	おもに仕事	4594	30	17	31	24	262	21	2	1	12	0	1
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	13	7	0	7	0	1
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	2711	888	13	1403	24	120
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	0
計		4594	30	17	31	24	262	2745	897	14	1422	24	122
Test dataの あてはまり		0.927											
男性 35歳~39歳		1	2	3	4	5	6	1	2	3	4	5	6
予測1	おもに仕事	9123	13	9	12	4	157	3379	2550	3	2667	6	55
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	121	120	0	118	0	8
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	282	174	2	334	0	4
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしていない人	0	0	0	0	0	0	0	0	0	0	0	1
計		9123	13	9	12	4	157	3782	2844	5	3120	6	67
Training dataの あてはまり		0.979											
2011年 テスト		1	2	3	4	5	6	1	2	3	4	5	6
予測1	おもに仕事	5985	40	2	33	13	252	2271	1215	6	1379	6	110
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	15	8	0	13	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしていない人のうち主に家事をしている人	0	0	0	0	0	0	802	349	1	462	7	24
予測5	仕事をしていない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしていない人	0	0	0	0	0	0	11	1	0	2	0	14
計		5985	40	2	33	13	252	3099	1573	7	1856	13	134
Test dataの あてはまり		0.946											

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性 40歳~44歳														
1996年 トレーニング														
観察値														
1	2	3	4	5	6	計	1	2	3	4	5	6	計	
10575	29	2	13	3	178	10800	予測1	4896	3496	4	2620	3	95	11114
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
10575	29	2	13	3	178	10800	計	4896	3496	4	2620	3	95	11114
0.979	Training dataの あてはまり													
2011年 テスト														
6020	32	5	38	3	295	6393	予測1	3268	1920	9	1647	9	96	6949
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
6020	32	5	38	3	295	6393	計	3268	1920	9	1647	9	96	6949
0.942	Test dataの あてはまり													
男性 45歳~49歳														
1996年 トレーニング														
観察値														
1	2	3	4	5	6	計	1	2	3	4	5	6	計	
11996	30	2	22	4	213	12267	予測1	5838	3893	3	2796	0	104	12634
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
11996	30	2	22	4	213	12267	計	5838	3893	3	2796	0	104	12634
0.978	Training dataの あてはまり													
2011年 テスト														
5421	37	1	43	2	252	5756	予測1	3165	1873	3	1225	13	105	6384
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
5421	37	1	43	2	252	5756	計	3165	1873	3	1225	13	105	6384
0.942	Test dataの あてはまり													
女性 40歳~44歳														
1996年 トレーニング														
観察値														
1	2	3	4	5	6	計	1	2	3	4	5	6	計	
10575	29	2	13	3	178	10800	予測1	4896	3496	4	2620	3	95	11114
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
10575	29	2	13	3	178	10800	計	4896	3496	4	2620	3	95	11114
0.979	Training dataの あてはまり													
2011年 テスト														
6020	32	5	38	3	295	6393	予測1	3268	1920	9	1647	9	96	6949
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
6020	32	5	38	3	295	6393	計	3268	1920	9	1647	9	96	6949
0.942	Test dataの あてはまり													
女性 45歳~49歳														
1996年 トレーニング														
観察値														
1	2	3	4	5	6	計	1	2	3	4	5	6	計	
11996	30	2	22	4	213	12267	予測1	5838	3893	3	2796	0	104	12634
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
11996	30	2	22	4	213	12267	計	5838	3893	3	2796	0	104	12634
0.978	Training dataの あてはまり													
2011年 テスト														
5421	37	1	43	2	252	5756	予測1	3165	1873	3	1225	13	105	6384
0	0	0	0	0	0	0	予測2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	予測6	0	0	0	0	0	0	0
5421	37	1	43	2	252	5756	計	3165	1873	3	1225	13	105	6384
0.942	Test dataの あてはまり													

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性		50歳～54歳						観察値					
1996年		1996年						1996年					
トレーニング		トレーニング						トレーニング					
予測		予測						予測					
1		1						1					
2		2						2					
3		3						3					
4		4						4					
5		5						5					
6		6						6					
計		計						計					
男性		女性						女性					
50歳～54歳		50歳～54歳						50歳～54歳					
トレーニング		トレーニング						トレーニング					
予測1		予測1						予測1					
予測2		予測2						予測2					
予測3		予測3						予測3					
予測4		予測4						予測4					
予測5		予測5						予測5					
予測6		予測6						予測6					
計		計						計					
Training dataの あてはまり													
2011年													
テスト													
予測1													
予測2													
予測3													
予測4													
予測5													
予測6													
計													
Test dataの あてはまり													
55歳～59歳													
トレーニング													
予測1													
予測2													
予測3													
予測4													
予測5													
予測6													
計													
Training dataの あてはまり													
2011年													
テスト													
予測1													
予測2													
予測3													
予測4													
予測5													
予測6													
計													
Test dataの あてはまり													

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性		60歳～64歳						女性		60歳～64歳					
1996年		トレーニング						1996年		トレーニング					
1	2	3	4	5	6	計	1	2	3	4	5	6	計		
予測1	5798	341	0	262	2	2080	8483	0	0	0	0	0	0		
予測2	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測4	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測5	0	0	0	0	0	0	0	1921	2192	3	4775	2	541		
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0		
計	5798	341	0	262	2	2080	8483	1921	2192	3	4775	2	541		
Training dataの あてはまり															
2011年 テスト															
予測1	6514	392	4	380	5	1686	8981	1	2	3	4	5	6		
予測2	0	0	0	0	0	0	0	8	1	1	5	0	15		
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測4	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測5	0	0	0	0	0	0	0	2663	2160	2	4171	0	380		
予測6	0	0	0	0	0	0	0	0	0	0	0	0	0		
計	6514	392	4	380	5	1686	8981	2671	2161	3	4176	0	380		
Test dataの あてはまり															
65歳～69歳															
1996年 トレーニング															
予測1	3701	437	3	348	0	2662	7151	1	2	3	4	5	6		
予測2	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測4	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測5	0	0	0	0	0	0	0	1067	1736	7	4625	2	911		
予測6	85	11	0	16	0	94	206	0	0	0	0	0	0		
計	3786	448	3	364	0	2756	7357	1067	1736	7	4625	2	911		
Training dataの あてはまり															
2011年 テスト															
予測1	0.516	1	2	3	4	5	6	計	1	2	3	4	5		
予測2	2819	522	6	483	2	2336	6168	0.554	1	2	3	4	5		
予測3	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測4	0	0	0	0	0	0	0	0	0	0	0	0	0		
予測5	0	0	0	0	0	0	0	1019	1327	1	3942	0	484		
予測6	26	0	0	3	0	24	53	0	0	0	0	0	0		
計	2845	522	6	486	2	2360	6221	1019	1327	1	3942	0	484		
Test dataの あてはまり															
0.457															

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性 70歳~74歳		観察値						女性 70歳~74歳						観察値									
1996年 トレーニング		1	2	3	4	5	6	1996年 トレーニング						1	2	3	4	5	6				
予測1	おもに仕事	0	0	0	0	0	0	0	予測1	おもに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	516	1038	0	3692	0	1399	6645	0	0	0	0	0	0
予測5	仕事をしない人のうち主に通学している人	1772	308	4	281	1	2301	4667	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしない人	1772	308	4	281	1	2301	4667	予測6	その他の仕事をしない人	516	1038	0	3692	0	1399	6645	0	0	0	0	0	0
計		0.493							計	0.556													
Training dataの あてはまり		2011年 テスト						Training dataの あてはまり						2011年 テスト									
予測1	おもに仕事	3	1	0	1	0	8	13	予測1	おもに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	0
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	593	1027	0	4254	1	761	6636	0	0	0	0	0	0
予測5	仕事をしない人のうち主に通学している人	5	0	0	0	0	2	7	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	0
予測6	その他の仕事をしない人	1844	435	5	594	0	2758	5636	予測6	その他の仕事をしない人	593	1027	0	4254	1	761	6636	0	0	0	0	0	0
計		1852	436	5	595	0	2768	5656	計	0.641													
Test dataの あてはまり		75歳~79歳						Test dataの あてはまり						75歳~79歳									
予測1	おもに仕事	1	2	3	4	5	6	計	予測1	おもに仕事	1	2	3	4	5	6	計	0	0	0	0	0	0
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	143	323	1	1762	0	1182	3411	0	0	0	0	0	0
予測5	仕事をしない人のうち主に通学している人	719	183	1	217	0	1874	2994	予測5	仕事をしない人のうち主に通学している人	36	85	0	504	0	536	1161	0	0	0	0	0	0
予測6	その他の仕事をしない人	719	183	1	217	0	1874	2994	予測6	その他の仕事をしない人	179	408	1	2266	0	1718	4572	0	0	0	0	0	0
計		0.626							計	0.503													
Training dataの あてはまり		2011年 テスト						Training dataの あてはまり						2011年 テスト									
予測1	おもに仕事	0	0	0	0	0	0	0	予測1	おもに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	予測2	家事などのかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
予測3	通学のかたわらに仕事	1	0	0	0	0	0	1	予測3	通学のかたわらに仕事	0	0	0	0	0	0	0	0	0	0	0	0	
予測4	仕事をしない人のうち主に家事をしている人	0	0	0	0	0	0	0	予測4	仕事をしない人のうち主に家事をしている人	265	609	0	3731	3	1199	5807	0	0	0	0	0	0
予測5	仕事をしない人のうち主に通学している人	1068	310	2	590	0	2652	4623	予測5	仕事をしない人のうち主に通学している人	0	0	0	0	0	0	0	0	0	0	0	0	
予測6	その他の仕事をしない人	1069	310	2	590	0	2652	4623	予測6	その他の仕事をしない人	4	6	0	32	0	15	57	0	0	0	0	0	0
計		0.574							計	0.639													

付表 4-4 (続) 線形判別関数による将来予測:説明変数は、年齢、年齢の2乗、教育ダミー

男性 80歳～84歳													
1996年 トレーニング													
観察値													
1	2	3	4	5	6	計	1	2	3	4	5	6	計
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
270	82	0	124	0	1246	1722	0	136	0	1033	1	1745	2980
0.724	82	0	124	0	1246	1722	計	137	0	1044	1	1753	3001
Training dataの あてはまり													
2011年 テスト													
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
525	154	0	434	0	2059	3172	計	256	0	2219	2	1465	4037
0.649	154	0	434	0	2059	3172	計	295	0	2444	2	1542	4394
Test dataの あてはまり													
男性 85歳～													
1996年 トレーニング													
観察値													
1	2	3	4	5	6	計	1	2	3	4	5	6	計
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
68	18	0	29	1	539	655	計	39	0	272	0	1072	1394
0.823	18	0	29	1	539	655	計	39	0	272	0	1072	1394
Training dataの あてはまり													
2011年 テスト													
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	4	0	7	16	計	91	0	994	0	1435	2550
123	70	0	222	0	982	1397	計	91	0	994	0	1435	2550
0.695	70	0	222	0	982	1397	計	91	0	994	0	1435	2550
Test dataの あてはまり													