

統計ソフト R による解析の事例

ロジスティック回帰分析を中心にして

西村 保一郎

大阪医科大学 数学教室

2011 年 11 月 9 日

使用ソフトウェアと参考文献

- ▶ **R ver. 2.13.1, 2011 July**, <http://www.r-project.org/>

使用ソフトウェアと参考文献

- ▶ **R ver. 2.13.1, 2011 July**, <http://www.r-project.org/>
- ▶ **JMP ver. 9.0, 2010, SAS Institute Inc.**

使用ソフトウェアと参考文献

- ▶ **R ver. 2.13.1, 2011 July**, <http://www.r-project.org/>
- ▶ **JMP ver. 9.0, 2010, SAS Institute Inc.**
- ▶ **Kleinbaum D. G., Klein M., Logistic Regression A Self-learning Text 3rd edition, 2010, Springer**
この講演では、本書の「**evans.dat**」, 「**cancer.dat**」を使用する

使用ソフトウェアと参考文献

- ▶ **R ver. 2.13.1, 2011 July**, <http://www.r-project.org/>
- ▶ **JMP ver. 9.0, 2010, SAS Institute Inc.**
- ▶ **Kleinbaum D. G., Klein M., Logistic Regression A Self-learning Text 3rd edition, 2010, Springer**
この講演では、本書の「`evans.dat`」、「`cancer.dat`」を使用する
- ▶ **Chonguvivatwong V., Analysis of epidemiological data using R and epicalc, 2008, Epidemiology Unit Prince of Songkla University THAILAND**
http://cran.r-project.org/doc/contrib/Epicalc_Book.pdf

使用ソフトウェアと参考文献

- ▶ **R ver. 2.13.1, 2011 July**, <http://www.r-project.org/>
- ▶ **JMP ver. 9.0, 2010, SAS Institute Inc.**
- ▶ **Kleinbaum D. G., Klein M., Logistic Regression A Self-learning Text 3rd edition, 2010, Springer**
この講演では、本書の「**evans.dat**」, 「**cancer.dat**」を使用する
- ▶ **Chonguvivatwong V., Analysis of epidemiological data using R and epicalc, 2008, Epidemiology Unit Prince of Songkla University THAILAND**
http://cran.r-project.org/doc/contrib/Epicalc_Book.pdf
- ▶ **Mo Lijia, Examining the reliability of logistic regression estimate software, 116 p., Dissertation, Department of Agricultural Economics, Kansas state university, 2010**
<http://krex.k-state.edu/dspace>
ロジスティック回帰分析について9つのソフトウェア (**Limdep, Matlab, R, Minitab, Eviews, SPSS, SAS, Stata, Shazam**) で色々なテストを行ない、前の6つが信頼できるソフトウェアという結論を得た。意外にも **SAS** は選ばれなかった

Rのインストール

- ▶ 例えば RjpWiki などから, R のインストール用ファイルを download する (R-2.13.1-win.exe の場合,40MB くらい)

Rのインストール

- ▶ 例えば RjpWiki などから, R のインストール用ファイルを download する (R-2.13.1-win.exe の場合,40MB くらい)
- ▶ これを実行するとインストールされる。インストール先を USB メモリにすることも可能。インストーラーの不備を回避するため, インストール時に使用する言語は English にする

Rのインストール

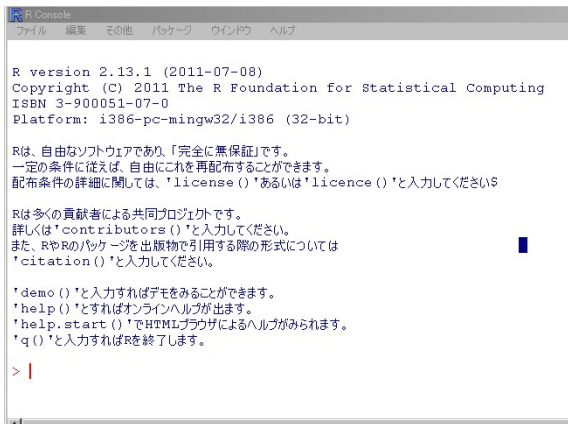
- ▶ 例えば RjpWiki などから, R のインストール用ファイルを download する (R-2.13.1-win.exe の場合,40MB くらい)
- ▶ これを実行するとインストールされる。インストール先を USB メモリにすることも可能。インストーラーの不備を回避するため, インストール時に使用する言語は English にする
- ▶ インストールが終わったら, 「R-2.13.1」「i386」フォルダの Rgui.exe のショートカットを作成し, 便利な場所に移動する

Rのインストール

- ▶ 例えば RjpWiki などから, R のインストール用ファイルを download する (R-2.13.1-win.exe の場合,40MB くらい)
- ▶ これを実行するとインストールされる。インストール先を USB メモリにすることも可能。インストーラーの不備を回避するため, インストール時に使用する言語は English にする
- ▶ インストールが終わったら, 「R-2.13.1」「i386」フォルダの Rgui.exe のショートカットを作成し, 便利な場所に移動する
- ▶ ショートカットの「プロパティ」で, 作業フォルダとして「R-2.13.1」フォルダを指定する

R の起動, コンソールとエディタ

- ▶ Rgui.exe へのショートカットをダブルクリックすると、R が起動してコンソールが開く



```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

R version 2.13.1 (2011-07-08)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)

Rは、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()'あるいは'licence()'と入力してください

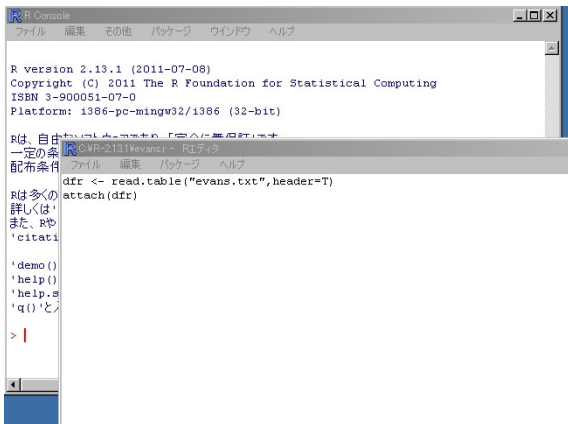
Rは多くの貢献者による共同プロジェクトです。
詳しくは'contributors()'と入力してください。
また、RやRのパッケージを出版物で引用する際の形式については
'citation()'と入力してください。

'demo()'と入力すればデモをみることができます。
'help()'とすればオンラインヘルプが出ます。
'help.start()'でHTMLブラウザによるヘルプがみられます。
'q()'と入力すればRを終了します。

> |
```

R の起動, コンソールとエディタ

- ▶ Rgui.exe へのショートカットをダブルクリックすると、R が起動してコンソールが開く
- ▶ コンソールのファイルメニューから「新しいスクリプト」を選ぶと、エディタが開く



The image shows two overlapping windows from the R software. The background window is the 'R Console', which displays the following text:

```
R version 2.13.1 (2011-07-08)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)
```

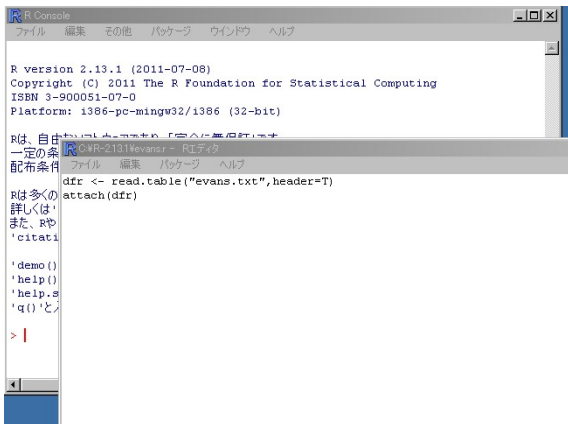
Below this, there is a menu bar with 'ファイル', '編集', 'その他', 'パッケージ', 'ウィンドウ', and 'ヘルプ'. A menu is open, showing options like '新しいスクリプト' (New Script) and '新しいスクリプトを保存' (Save New Script). The foreground window is the R GUI editor, showing a script with the following code:

```
dfr <- read.table("evans.txt", header=T)
attach(dfr)
```

The editor window also has a menu bar with 'ファイル', '編集', 'パッケージ', and 'ヘルプ'. The script is being edited in a text area.

R の起動, コンソールとエディタ

- ▶ Rgui.exe へのショートカットをダブルクリックすると、R が起動してコンソールが開く
- ▶ コンソールのファイルメニューから「新しいスクリプト」を選ぶと、エディタが開く
- ▶ R の 2 つの窓 コンソール、エディタ



```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

R version 2.13.1 (2011-07-08)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)

Rは、自由なソフトウェアであり「完全に無保証」です。
一定の条件下で、Rを複製、再配布、改造、再配布することが
許可されています。Rの複製、再配布、改造、再配布の条件は、
Rのライセンス条件を参照してください。
Rは多くの詳細については、
また、Rや
'citation()' を使って、Rのライセンス条件を参照してください。
'demo()' を使って、Rのデモプログラムを参照してください。
'help()' を使って、Rのヘルプを参照してください。
'help.search()' を使って、Rのヘルプを参照してください。
'q()' を使って、Rを終了してください。

> |
```

```
R Editor
C:\R-2131\evansr - Rエディタ
ファイル 編集 パッケージ ヘルプ

dfr <- read.table("evans.txt", header=T)
attach(dfr)

'citation()' を使って、Rのライセンス条件を参照してください。
'demo()' を使って、Rのデモプログラムを参照してください。
'help()' を使って、Rのヘルプを参照してください。
'help.search()' を使って、Rのヘルプを参照してください。
'q()' を使って、Rを終了してください。

> |
```

R の packages のインストール

R の packages のインストール

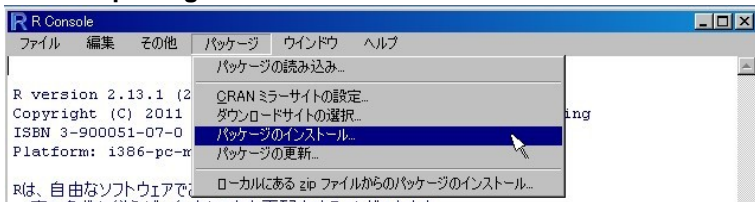
- ▶ R は多数の **packages** の集合体。 **packages** は 3 種類に分類される
 - (1) 初めからインストールされ, R を起動すると利用できる
 - (2) 初めからインストールされ, スクリプトで指定して利用する
 - (3) 後から **download** してインストールし,
スクリプトで指定して利用する

R の packages のインストール

- ▶ R は多数の **packages** の集合体。 **packages** は 3 種類に分類される
 - (1) 初めからインストールされ、R を起動すると利用できる
 - (2) 初めからインストールされ、スクリプトで指定して利用する
 - (3) 後から **download** してインストールし、
スクリプトで指定して利用する
- ▶ この講演では 上の (3) に属する **epicalc**、 **car**、 **ROCR packages** を利用する

packages のインストールは、R の「パッケージ」メニューの「パッケージのインストール」から行う

注. 1 つの **package** をインストールすると、その **package** が依存している **packages** が自動的にインストールされる



ロジスティック回帰 データファイル

目的変数	CHD (冠動脈性心疾患, presence (1) or absence (0))
説明変数	CAT (カテコラミンレベル, high (1) or normal (0))
	AGEGR (年齢, 名義変数, 40-49 (0), 50-59 (1), 60-69 (2), 70- (3))
	CHL (コレステロール, 数値変数)
	SMK (喫煙経験, ever smoked (1) or never smoked (0))
	ECG (心電図異常, presence (1) or absence (0))
	DBP (拡張期血圧, 数値変数)
	SBP (収縮期血圧, 数値変数)
HPT (高血圧症, presence (1) or absence (0))	



	A	B	C	D	E	F	G	H	I	
1	CHD	CAT	AGEGR	CHL	SMK	ECG	DBP	SBP	HPT	
2		0	0	1	270	0	0	80	138	0
3		0	0	0	159	1	0	74	128	0
4		1	1	1	201	1	1	112	164	1
5		0	1	2	179	1	0	100	200	1

冠動脈性心疾患 (coronary heart disease) データ

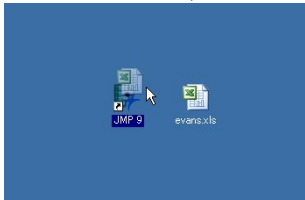
609 名の男性、そのうち CHD presence 71, absence 538

データファイルのソフトへの読み込み

データファイルのソフトへの読み込み

- ▶ JMP の場合

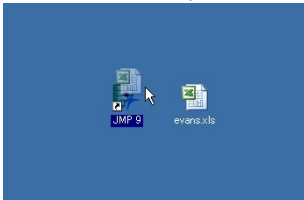
データの Excel のアイコンを JMP のアイコンにドラッグする
JMP が起動して、データを載せたデータテーブルが開く



データファイルのソフトへの読み込み

▶ JMP の場合

データの Excel のアイコンを JMP のアイコンにドラッグする
JMP が起動して、データを載せたデータテーブルが開く



▶ R の場合

Excel でデータをテキスト形式で R フォルダに保存する (evans.txt)
R のエディタの窓につきのスクリプトを入力し実行する

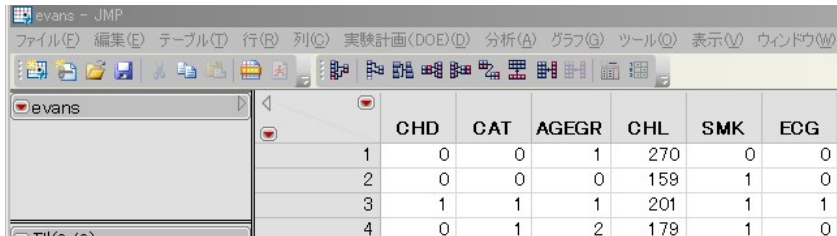
```
dfr <- read.table("evans.txt", header=T)
```

データの行列に、dfr という名前を付けた

読み込んだデータファイルのソフトでの閲覧

読み込んだデータファイルのソフトでの閲覧

- ▶ JMP の場合: Excel ファイルのアイコンを JMP アイコンにドラッグすると JMP が起動してデータテーブルが自動的に開く

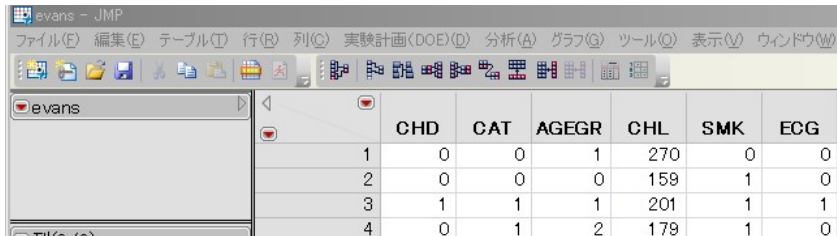


The screenshot shows the JMP software interface. The title bar reads "evans - JMP". The menu bar includes "ファイル(F)", "編集(E)", "テーブル(T)", "行(R)", "列(C)", "実験計画(DOE)(D)", "分析(A)", "グラフ(G)", "ツール(O)", "表示(V)", and "ウィンドウ(W)". The toolbar contains various icons for file operations and data analysis. On the left, a project browser shows a folder named "evans". The main window displays a data table with the following data:

	CHD	CAT	AGEGR	CHL	SMK	ECG
1	0	0	1	270	0	0
2	0	0	0	159	1	0
3	1	1	1	201	1	1
4	0	1	2	179	1	0

読み込んだデータファイルのソフトでの閲覧

- ▶ JMP の場合: Excel ファイルのアイコンを JMP アイコンにドラッグすると JMP が起動してデータテーブルが自動的に開く

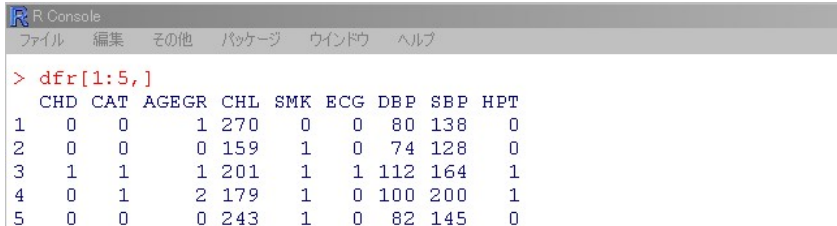


The screenshot shows the JMP software interface. The title bar reads "evans - JMP". The menu bar includes "ファイル(F)", "編集(E)", "テーブル(T)", "行(R)", "列(C)", "実験計画(DOE)(D)", "分析(A)", "グラフ(G)", "ツール(O)", "表示(V)", and "ウインドウ(W)". The toolbar contains various icons for file operations and data analysis. The main window displays a data table with the following data:

	CHD	CAT	AGEGR	CHL	SMK	ECG
1	0	0	1	270	0	0
2	0	0	0	159	1	0
3	1	1	1	201	1	1
4	0	1	2	179	1	0

- ▶ R の場合: R でつぎのスクリプトを入力し実行すると, データの行列の最初の 5 行が表示される

`dfr[1:5,]`



The screenshot shows the R Console interface. The title bar reads "R Console". The menu bar includes "ファイル", "編集", "その他", "パッケージ", "ウインドウ", and "ヘルプ". The console displays the following output for the command `> dfr[1:5,]`:

```
> dfr[1:5,]
  CHD  CAT  AGEGR  CHL  SMK  ECG  DBP  SBP  HPT
1    0    0     1 270    0    0   80 138    0
2    0    0     0 159    1    0   74 128    0
3    1    1     1 201    1    1  112 164    1
4    0    1     2 179    1    0  100 200    1
5    0    0     0 243    1    0   82 145    0
```

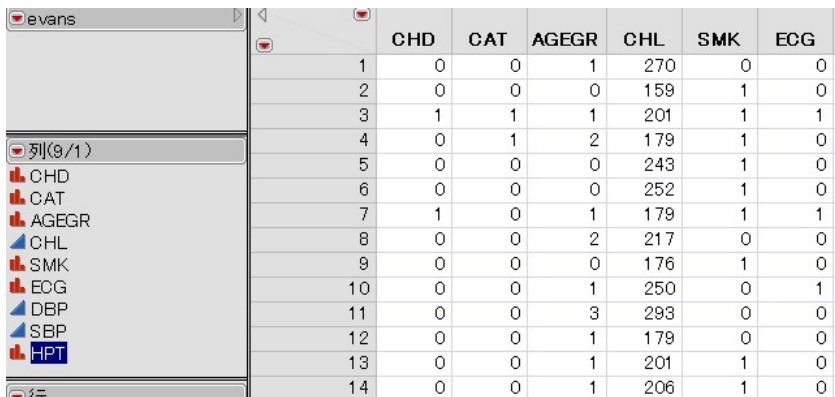
数字で入力されたカテゴリー変数の扱い JMP の場合 1

- ▶ JMP が起動しデータテーブルが開いた

	CHD	CAT	AGEGR	CHL	SMK	ECG
1	0	0	1	270	0	0
2	0	0	0	159	1	0
3	1	1	1	201	1	1
4	0	1	2	179	1	0
5	0	0	0	243	1	0
6	0	0	0	252	1	0
7	1	0	1	179	1	1
8	0	0	2	217	0	0
9	0	0	0	176	1	0
10	0	0	1	250	0	1
11	0	0	3	293	0	0
12	0	0	1	179	0	0
13	0	0	1	201	1	0
14	0	0	1	206	1	0

数字で入力されたカテゴリー変数の扱い JMP の場合 1

- ▶ JMP が起動しデータテーブルが開いた
- ▶ 青い連続変数マークを赤い名義 (カテゴリー) 変数マークへ



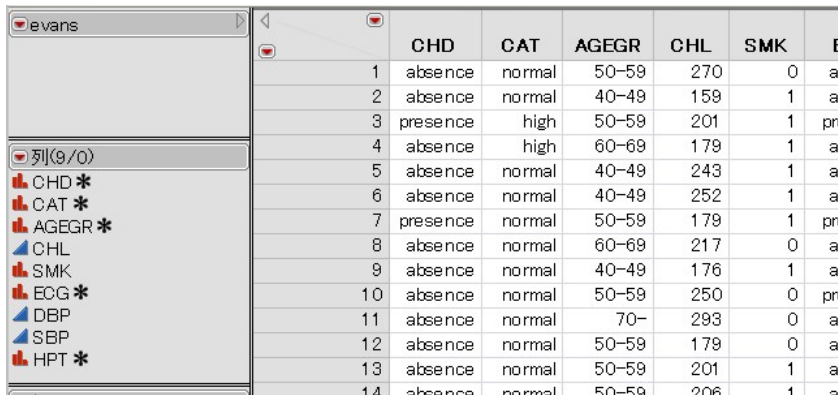
	CHD	CAT	AGEGR	CHL	SMK	ECG
1	0	0	1	270	0	0
2	0	0	0	159	1	0
3	1	1	1	201	1	1
4	0	1	2	179	1	0
5	0	0	0	243	1	0
6	0	0	0	252	1	0
7	1	0	1	179	1	1
8	0	0	2	217	0	0
9	0	0	0	176	1	0
10	0	0	1	250	0	1
11	0	0	3	293	0	0
12	0	0	1	179	0	0
13	0	0	1	201	1	0
14	0	0	1	206	1	0

数字で入力されたカテゴリー変数の扱い JMP の場合 1

- ▶ JMP が起動しデータテーブルが開いた
- ▶ 青い連続変数マークを赤い名義 (カテゴリー) 変数マークへ
- ▶ 数字にラベルを付ける

例. CHD(coronary heart disease) 0 (absence), 1 (presence)

AGEGR(年齢階級) 0 (40-49), 1 (50-59), 2 (60-69), 3 (70-)

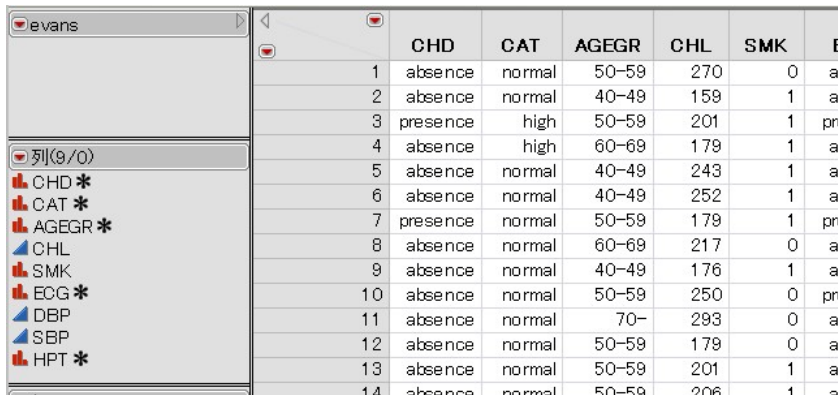


The screenshot shows the JMP interface. On the left, a list of variables is displayed: CHD *, CAT *, AGEGR *, CHL, SMK, ECG *, DBP, SBP, and HPT *. The main window shows a data table with 14 rows and 7 columns. The columns are labeled CHD, CAT, AGEGR, CHL, SMK, and E. The data rows show numerical values for CHD and AGEGR, and categorical values for CAT, CHL, and SMK.

	CHD	CAT	AGEGR	CHL	SMK	E
1	absence	normal	50-59	270	0	a
2	absence	normal	40-49	159	1	a
3	presence	high	50-59	201	1	pn
4	absence	high	60-69	179	1	a
5	absence	normal	40-49	243	1	a
6	absence	normal	40-49	252	1	a
7	presence	normal	50-59	179	1	pn
8	absence	normal	60-69	217	0	a
9	absence	normal	40-49	176	1	a
10	absence	normal	50-59	250	0	pn
11	absence	normal	70-	293	0	a
12	absence	normal	50-59	179	0	a
13	absence	normal	50-59	201	1	a
14	absence	normal	50-59	206	1	a

数字で入力されたカテゴリー変数の扱い JMP の場合 1

- ▶ JMP が起動しデータテーブルが開いた
- ▶ 青い連続変数マークを赤い名義 (カテゴリー) 変数マークへ
- ▶ 数字にラベルを付ける
例. **CHD(coronary heart disease) 0 (absence), 1 (presence)**
AGEGR(年齢階級) 0 (40-49), 1 (50-59), 2 (60-69), 3 (70-)
- ▶ 説明の便宜上、**SMK (喫煙経験の有 (1) 無 (0))** にはラベルを付けない



	CHD	CAT	AGEGR	CHL	SMK	E
1	absence	normal	50-59	270	0	a
2	absence	normal	40-49	159	1	a
3	presence	high	50-59	201	1	pn
4	absence	high	60-69	179	1	a
5	absence	normal	40-49	243	1	a
6	absence	normal	40-49	252	1	a
7	presence	normal	50-59	179	1	pn
8	absence	normal	60-69	217	0	a
9	absence	normal	40-49	176	1	a
10	absence	normal	50-59	250	0	pn
11	absence	normal	70-	293	0	a
12	absence	normal	50-59	179	0	a
13	absence	normal	50-59	201	1	a
14	absence	normal	50-59	206	1	a

数字で入力されたカテゴリー変数の扱い JMP の場合 2

- ▶ 回帰式 (CHD の罹患率の式) を R の出力と揃えるための設定をする

数字で入力されたカテゴリー変数の扱い JMP の場合 2

- ▶ 回帰式 (CHD の罹患率の式) を R の出力と揃えるための設定をする
- ▶ CHD の「リストチェック」を選ぶ

	CHL	SMK	ECG	DBP	SBP	HP
1	absence	ノート	absence	80	138	absence
2	absence	範囲チェック	absence	74	128	absence
3	presence	リストチェック	presence	112	164	presence
4	absence	欠測値のコード	absence	100	200	presence
5	absence	値ラベル	absence	82	145	absence
6	absence	値の順序	absence	88	142	absence
7	presence	値の色	presence	80	128	absence
8	absence	カラーグラデーション	absence	92	135	absence
9	absence	軸	absence	76	114	absence
10	absence	コード変換	presence	114	182	presence
11	absence	配合	absence	100	166	presence
12	absence	データの出現順	absence	90	158	absence
13	absence	仕様限界	absence	86	142	absence
14	absence	管理限界	absence	85	120	absence

数字で入力されたカテゴリー変数の扱い JMP の場合 2

- ▶ 回帰式 (CHD の罹患率の式) を R の出力と揃えるための設定をする
- ▶ CHD の「リストチェック」を選ぶ
- ▶ リストチェックの欄が上から 1、0 の順に変更する

テーブル 'evans1' の 'CHD'

列名

ロック

データタイプ

尺度

表示形式 総桁数

桁区切り(,)を使用

列プロパティ ▾

リストチェック

値ラベル
オプションの項目

リストチェック
列には特定の値だけ入力可能。

HP
abse
abse
prese
prese
abse
abse
abse
abse
abse
prese
prese
abse
abse
abse
prese
prese
abse
abse
prese

数字で入力されたカテゴリー変数の扱い JMP の場合 2

- ▶ 回帰式 (CHD の罹患率の式) を R の出力と揃えるための設定をする
- ▶ CHD の「リストチェック」を選ぶ
- ▶ リストチェックの欄が上から 1、0 の順に変更する
- ▶ パネルの CHD にチェックが付く

	CHD	CAT	AGEGR	CHL	SMK	ECG
1	absence	normal	50-59	270	0	absen
2	absence	normal	40-49	159	1	absen
3	presence	high	50-59	201	1	presen
4	absence	high	60-69	179	1	absen
5	absence	normal	40-49	243	1	absen
6	absence	normal	40-49	252	1	absen
7	presence	normal	50-59	179	1	presen
8	absence	normal	60-69	217	0	absen
9	absence	normal	40-49	176	1	absen
10	absence	normal	50-59	250	0	presen
11	absence	normal	70-	293	0	absen
12	absence	normal	50-59	179	0	absen
13	absence	normal	50-59	201	1	absen
14	absence	normal	50-59	206	1	absen

数字で入力されたカテゴリー変数の扱い JMP の場合 2

- ▶ 回帰式 (CHD の罹患率の式) を R の出力と揃えるための設定をする
- ▶ CHD の「リストチェック」を選ぶ
- ▶ リストチェックの欄が上から 1、0 の順に変更する
- ▶ パネルの CHD にチェックが付く
- ▶ 同様に、CAT、SMK、ECG、HPT は、上から 1、0 の順に変更する
AGEGR は上から 3、2、1、0 の順に変更する

	CHD	CAT	AGEGR	CHL	SMK	ECG	DBP	SBP
1	absence	normal	50-59	270	0	absence	80	
2	absence	normal	40-49	159	1	absence	74	
3	presence	high	50-59	201	1	presence	112	
4	absence	high	60-69	179	1	absence	100	
5	absence	normal	40-49	243	1	absence	82	
6	absence	normal	40-49	252	1	absence	88	
7	presence	normal	50-59	179	1	presence	80	
8	absence	normal	60-69	217	0	absence	92	
9	absence	normal	40-49	176	1	absence	76	
10	absence	normal	50-59	250	0	presence	114	
11	absence	normal	70-	293	0	absence	100	
12	absence	normal	50-59	179	0	absence	90	
13	absence	normal	50-59	201	1	absence	86	
14	absence	normal	50-59	206	1	absence	85	

数字で入力されたカテゴリー変数の扱い

R の場合 1

R でつぎのスクリプトを入力し実行する

数字で入力されたカテゴリー変数の扱い

R の場合 1

R でつぎのスクリプトを入力し実行する

```
dfr <- within(dfr,{  
  CHD<-factor(CHD,labels=c("absence","presence"))  
  CAT<-factor(CAT,labels=c("normal","high"))  
  AGEGR<-factor(AGEGR,labels=c("40-49","50-59","60-69","70-"))  
  SMK<-factor(SMK)  
  ECG<-factor(ECG,labels=c("absence","presence"))  
  HPT<-factor(HPT,labels=c("absence","presence")) })  
attach(dfr)
```

数字で入力されたカテゴリー変数の扱い R の場合 1

R でつぎのスクリプトを入力し実行する

```
dfr <- within(dfr,{  
  CHD<-factor(CHD,labels=c("absence","presence"))  
  CAT<-factor(CAT,labels=c("normal","high"))  
  AGEGR<-factor(AGEGR,labels=c("40-49","50-59","60-69","70-"))  
  SMK<-factor(SMK)  
  ECG<-factor(ECG,labels=c("absence","presence"))  
  HPT<-factor(HPT,labels=c("absence","presence")) })  
attach(dfr)
```

- ▶ **factor** 関数は、連続変数を名義 (カテゴリー) 変数に変える

数字で入力されたカテゴリー変数の扱い R の場合 1

R でつぎのスクリプトを入力し実行する

```
dfr <- within(dfr,{  
  CHD<-factor(CHD,labels=c("absence","presence"))  
  CAT<-factor(CAT,labels=c("normal","high"))  
  AGEGR<-factor(AGEGR,labels=c("40-49","50-59","60-69","70-"))  
  SMK<-factor(SMK)  
  ECG<-factor(ECG,labels=c("absence","presence"))  
  HPT<-factor(HPT,labels=c("absence","presence")) })  
attach(dfr)
```

- ▶ **factor** 関数は、連続変数を名義 (カテゴリー) 変数に変える
- ▶ 数字にラベルを対応させるには、**factor** 関数の中で、**labels** オプションを用いる。数値の小さい順 (0,1,... の順) に対応するラベルを指定する

数字で入力されたカテゴリー変数の扱い R の場合 1

R でつぎのスクリプトを入力し実行する

```
dfr <- within(dfr,{
  CHD<-factor(CHD,labels=c("absence","presence"))
  CAT<-factor(CAT,labels=c("normal","high"))
  AGEGR<-factor(AGEGR,labels=c("40-49","50-59","60-69","70-"))
  SMK<-factor(SMK)
  ECG<-factor(ECG,labels=c("absence","presence"))
  HPT<-factor(HPT,labels=c("absence","presence")) })
attach(dfr)
```

- ▶ **factor** 関数は、連続変数を名義 (カテゴリー) 変数に変える
- ▶ 数字にラベルを対応させるには、**factor** 関数の中で、**labels** オプションを用いる。数値の小さい順 (0,1,... の順) に対応するラベルを指定する
- ▶ 説明の便宜のため、**SMK**(喫煙経験の有 (1) 無 (0)) には、敢えてラベルを付けない

数字で入力されたカテゴリー変数の扱い R の場合 2

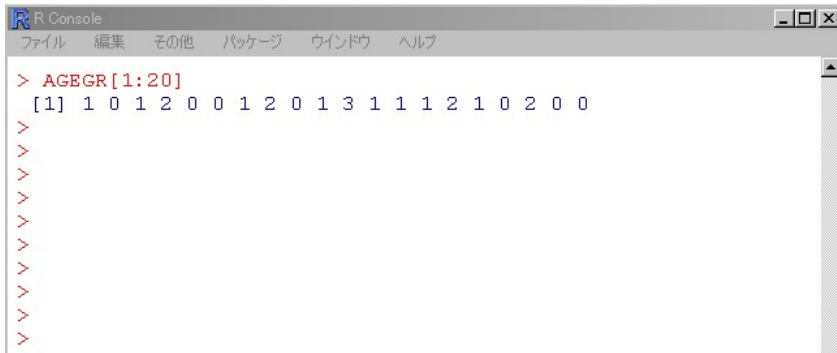
- ▶ 画面にデータ `dfr` の最初の 5 行を表示させる
付けたラベルが反映された表示になる

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> dfr[1:5,]
      CHD      CAT AGEGR CHL  SMK      ECG  DBP  SBP      HPT
1 absence normal 50-59 270    0 absence  80 138 absence
2 absence normal 40-49 159    1 absence  74 128 absence
3 presence  high 50-59 201    1 presence 112 164 presence
4 absence  high 60-69 179    1 absence 100 200 presence
5 absence normal 40-49 243    1 absence  82 145 absence
>
>
>
>
>
>
```

数字で入力されたカテゴリー変数の扱い R の場合 2

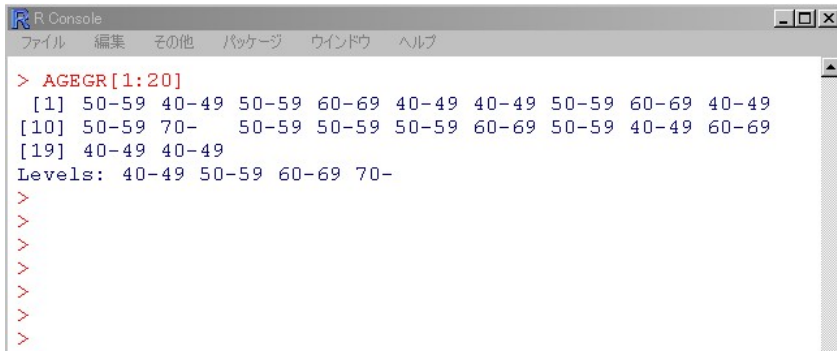
- ▶ 画面にデータ `dfr` の最初の 5 行を表示させる
付けたラベルが反映された表示になる
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行前



```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
> AGEGR[1:20]
[1] 1 0 1 2 0 0 1 2 0 1 3 1 1 1 2 1 0 2 0 0
>
>
>
>
>
>
>
>
>
>
```

数字で入力されたカテゴリー変数の扱い R の場合 2

- ▶ 画面にデータ `dfr` の最初の 5 行を表示させる
付けたラベルが反映された表示になる
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行前
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行後

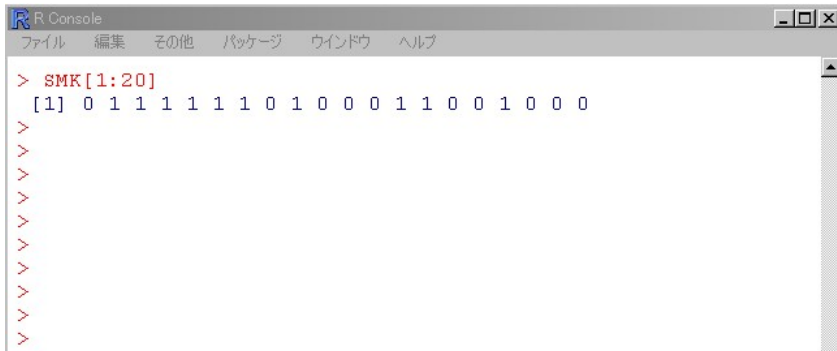


```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> AGEGR[1:20]
 [1] 50-59 40-49 50-59 60-69 40-49 40-49 50-59 60-69 40-49
[10] 50-59 70-   50-59 50-59 50-59 60-69 50-59 40-49 60-69
[19] 40-49 40-49
Levels: 40-49 50-59 60-69 70-
>
>
>
>
>
>
>
```


数字で入力されたカテゴリー変数の扱い R の場合 2

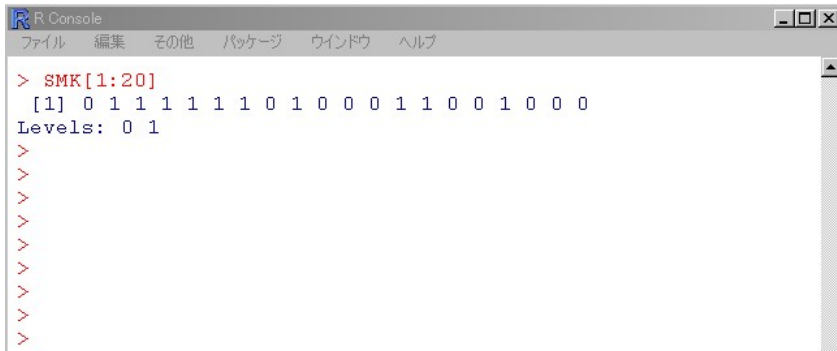
- ▶ 画面にデータ `dfr` の最初の 5 行を表示させる
付けたラベルが反映された表示になる
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行前
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行後
- ▶ **SMK (喫煙経験の有無): factor** 関数のスクリプトの実行前



```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
> SMK[1:20]
[1] 0 1 1 1 1 1 1 1 0 1 0 0 0 1 1 0 0 1 0 0 0
>
>
>
>
>
>
>
>
>
>
```

数字で入力されたカテゴリー変数の扱い R の場合 2

- ▶ 画面にデータ `dfr` の最初の 5 行を表示させる
付けたラベルが反映された表示になる
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行前
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行後
- ▶ **SMK (喫煙経験の有無): factor** 関数のスクリプトの実行前
- ▶ **SMK (喫煙経験の有無): factor** 関数のスクリプトの実行後



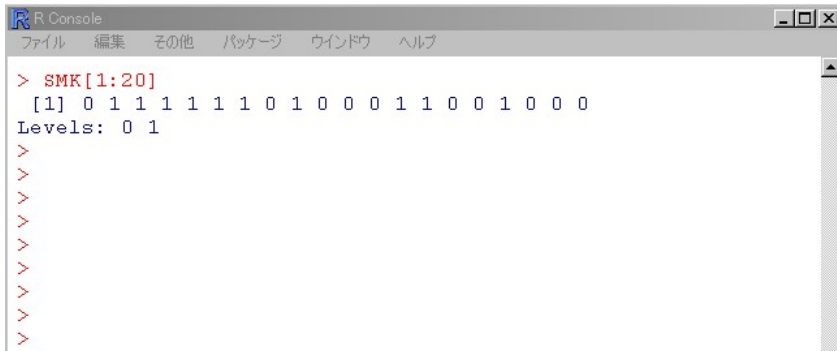
```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> SMK[1:20]
 [1] 0 1 1 1 1 1 1 0 1 0 0 0 1 1 0 0 1 0 0 0
Levels: 0 1

>
>
>
>
>
>
>
>
>
```

数字で入力されたカテゴリー変数の扱い R の場合 2

- ▶ 画面にデータ `dfr` の最初の 5 行を表示させる
付けたラベルが反映された表示になる
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行前
- ▶ **AGEGR (年齢階級): factor** 関数のスクリプトの実行後
- ▶ **SMK (喫煙経験の有無): factor** 関数のスクリプトの実行前
- ▶ **SMK (喫煙経験の有無): factor** 関数のスクリプトの実行後
- ▶ 名義変数では、**Levels** が表示される



```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> SMK[1:20]
 [1] 0 1 1 1 1 1 1 0 1 0 0 0 1 1 0 0 1 0 0 0
Levels: 0 1
>
>
>
>
>
>
>
>
>
```

単変量ロジスティック回帰分析

- ▶ 説明変数

CAT、AGEGR、CHL、SMK、ECG、DBP、SBP、HPT
から目的変数 **CHD** に影響がある変数を選び出す

単変量ロジスティック回帰分析

- ▶ 説明変数
CAT、**AGEGR**、**CHL**、**SMK**、**ECG**、**DBP**、**SBP**、**HPT**
から目的変数 **CHD** に影響がある変数を選び出す
- ▶ 説明変数を 1 つだけ使うロジスティック回帰分析
(=単変量ロジスティック回帰分析)
p 値 < 0.05 ならば、**CHD** に影響があると判断する

単変量ロジスティック回帰分析

- ▶ 説明変数
CAT、AGEGR、CHL、SMK、ECG、DBP、SBP、HPT
から目的変数 CHD に影響がある変数を選び出す
- ▶ 説明変数を 1 つだけ使うロジスティック回帰分析
(=単変量ロジスティック回帰分析)
p 値 < 0.05 ならば、CHD に影響があると判断する
- ▶ CHD に影響がある変数
多変量ロジスティック回帰分析に用いる
CHD に影響がない変数
更なる解析には用いない

単変量ロジスティック回帰分析

- ▶ 説明変数
CAT、AGEGR、CHL、SMK、ECG、DBP、SBP、HPT
から目的変数 **CHD** に影響がある変数を選び出す
- ▶ 説明変数を 1 つだけ使うロジスティック回帰分析
(=単変量ロジスティック回帰分析)
p 値 < 0.05 ならば、**CHD** に影響があると判断する
- ▶ **CHD** に影響がある変数
多変量ロジスティック回帰分析に用いる
CHD に影響がない変数
更なる解析には用いない
- ▶ 連続変数 **CHL、DBP、SBP**
2 値名義変数 **CAT、SMK、ECG、HPT**
4 値名義変数 **AGEGR**
の 3 種類に分けて、単変量ロジスティック分析を説明する

単変量ロジスティック回帰分析 モデル式 1

CHD 有病率 $P(\text{CHD})$ を表す式

注. **CHD = presence** となる確率の意味であるが、 $P(\text{CHD})$ と略記する
指数関数 e^x は、ほとんどいつも $\exp(x)$ と表す

- ▶ 数値説明変数の場合 例. **CHL**
パラメータ 切片 α , 係数 β

$$P(\text{CHD}) = \frac{1}{1+e^{-(\alpha+\beta\text{CHL})}} = \frac{1}{1+\exp(-(\alpha+\beta\text{CHL}))}$$

単変量ロジスティック回帰分析 モデル式 1

CHD 有病率 $P(\text{CHD})$ を表す式

注. **CHD = presence** となる確率の意味であるが、 $P(\text{CHD})$ と略記する
指数関数 e^x は、ほとんどいつも $\exp(x)$ と表す

- ▶ 数値説明変数の場合 例. **CHL**
パラメータ 切片 α , 係数 β

$$P(\text{CHD}) = \frac{1}{1+e^{-(\alpha+\beta\text{CHL})}} = \frac{1}{1+\exp(-(\alpha+\beta\text{CHL}))}$$

- ▶ 2 値名義説明変数の場合 例. **CAT**
パラメータ 切片 α ,
係数 β_{high} , **normal** を **reference state** とする

$$P(\text{CHD}) = \frac{1}{1+\exp(-(\alpha+\beta_{\text{high}}\text{CAT}))}$$

単変量ロジスティック回帰分析 モデル式 1

CHD 有病率 $P(\text{CHD})$ を表す式

注. **CHD = presence** となる確率の意味であるが、 $P(\text{CHD})$ と略記する
指数関数 e^x は、ほとんどいつも $\exp(x)$ と表す

- ▶ 数値説明変数の場合 例. **CHL**
パラメータ 切片 α , 係数 β

$$P(\text{CHD}) = \frac{1}{1+e^{-(\alpha+\beta\text{CHL})}} = \frac{1}{1+\exp(-(\alpha+\beta\text{CHL}))}$$

- ▶ 2 値名義説明変数の場合 例. **CAT**
パラメータ 切片 α ,
係数 β_{high} , **normal** を **reference state** とする

$$P(\text{CHD}) = \frac{1}{1+\exp(-(\alpha+\beta_{\text{high}}\text{CAT}))}$$

- ▶ 4 値名義変数 **AGEGR** の場合
(計算上は **AGEGR**₅₀₋₅₉, **AGEGR**₆₀₋₆₉, **AGEGR**₇₀ の 3 変数)
パラメータ 切片 α ,
係数 β_{50-59} , β_{60-69} , β_{70-} , **40-49** を **reference state** とする

$$P(\text{CHD}) = \frac{1}{1+\exp(-(\alpha+\beta_{50-59}\text{AGEGR}_{50-59}+\beta_{60-69}\text{AGEGR}_{60-69}+\beta_{70-}\text{AGEGR}_{70-}))}$$

単変量ロジスティック回帰分析 モデル式 1

CHD 有病率 $P(\text{CHD})$ を表す式

注. **CHD = presence** となる確率の意味であるが、 $P(\text{CHD})$ と略記する
指数関数 e^x は、ほとんどいつも $\exp(x)$ と表す

- ▶ 数値説明変数の場合 例. **CHL**
パラメータ 切片 α , 係数 β

$$P(\text{CHD}) = \frac{1}{1+e^{-(\alpha+\beta\text{CHL})}} = \frac{1}{1+\exp(-(\alpha+\beta\text{CHL}))}$$

- ▶ 2 値名義説明変数の場合 例. **CAT**
パラメータ 切片 α ,
係数 β_{high} , **normal** を **reference state** とする

$$P(\text{CHD}) = \frac{1}{1+\exp(-(\alpha+\beta_{\text{high}}\text{CAT}))}$$

- ▶ 4 値名義変数 **AGEGR** の場合
(計算上は **AGEGR**₅₀₋₅₉, **AGEGR**₆₀₋₆₉, **AGEGR**₇₀ の 3 変数)
パラメータ 切片 α ,
係数 β_{50-59} , β_{60-69} , β_{70-} , **40-49** を **reference state** とする

$$P(\text{CHD}) = \frac{1}{1+\exp(-(\alpha+\beta_{50-59}\text{AGEGR}_{50-59}+\beta_{60-69}\text{AGEGR}_{60-69}+\beta_{70-}\text{AGEGR}_{70-}))}$$

- ▶ $0 < P(\text{CHD}) < 1$ が成り立つ

単変量ロジスティック回帰分析 JMP 計算

例として CHL を用いるが、他の変数の計算も同様である

- ▶ 分析メニューから「モデルのあてはめ」を選ぶ

The screenshot shows the JMP software interface. The 'Analysis' menu is open, and 'Fit Model' is selected. The main window displays a data table with the following columns: CHD, CAT, AGEGR, CHL, SMK, ECG, and DBP. The data table contains 10 rows of data.

	CHD	CAT	AGEGR	CHL	SMK	ECG	DBP
1	absence	nor					
2	absence	nor					
3	presence	h					
4	absence	high	60-69	179	1	absence	1
5	absence	normal	40-49	243	1	absence	1
6	absence	normal	40-49	252	1	absence	1
7	presence	normal	50-59	179	1	presence	1
8	absence	normal	60-69	217	0	absence	1
9	absence	normal	40-49	176	1	absence	1
10	absence	normal	50-59	250	0	presence	1

単変量ロジスティック回帰分析 JMP 計算

例として CHL を用いるが、他の変数の計算も同様である

- ▶ 分析メニューから「モデルのあてはめ」を選ぶ
- ▶ CHD を Y に、CHL を「モデル効果の構成」に入れて「実行」

The screenshot shows the 'Model Specification' dialog box in JMP. The window title is 'モデルのあてはめ - JMP'. The dialog is titled 'モデルの指定' (Model Specification). It is divided into several sections:

- 列の選択 (Column Selection):** A list of variables on the left: CHD, CAT, AGEGR, CHL, SMK, ECG, DBP, SBP, HPT. CHD is selected as the response variable.
- 役割変数の選択 (Role Variable Selection):** The 'Y' role is assigned to CHD. Other roles like '重み' (Weight), '度数' (Degree), and 'By' are currently empty.
- 手法 (Method):** A dropdown menu is set to '名義ロジスティック' (Nominal Logistic).
- ボタン (Buttons):** 'ヘルプ' (Help), '実行' (Run), '削除' (Delete), and '前回の設定' (Previous Settings) are visible.
- モデル効果の構成 (Model Effects):** A section where 'CHL' is listed as a model effect.

単変量ロジスティック回帰分析 JMP 計算

例として **CHL** を用いるが、他の変数の計算も同様である

- ▶ 分析メニューから「モデルのあてはめ」を選ぶ
- ▶ **CHD** を Y に、**CHL** を「モデル効果の構成」に入れて「実行」
- ▶ 結果出力の窓の赤い三角メニューから、「**Wald 検定**」、「**オッズ比**」、「**信頼区間**」、「**確率の計算式の保存**」を選ぶ

■ evans - 名義ロジスティックのあてはめ - JMP

名義ロジスティックのあてはめ CHD

- ✓ ロジスティックプロット
プロットのオプション ▶
- ✓ 尤度比検定
- ✓ Wald検定
- ✓ 信頼区間
- ✓ オッズ比
逆推定...
- 確率の計算式の保存**
- ROC曲線
- リフトチャート
- 混同行列
- プロフィール
- モデルダイアログ
- スクリプト ▶

自由度	Waldカイ2乗	p値(Prob>ChiSq)
1	5.22544844	0.0223*

自由度	尤度比カイ2乗	p値(Prob>ChiSq)
1	5.13397364	0.0235*

CHD: presence vs absenceのオッズ比に対して

単変量ロジスティック回帰分析 R 計算

例として **CHL** を用いるが、他の変数の計算も同様である

- ▶ **epicalc Package** をロードしておく

```
library(epicalc)
```

単変量ロジスティック回帰分析 R 計算

例として **CHL** を用いるが、他の変数の計算も同様である

- ▶ **epicalc Package** をロードしておく

```
library(epicalc)
```

- ▶ ロジスティック回帰分析は、**glm** 関数に **binomial** オプションを付けて行う
出力を **ul_CHL** と名付ける (**univariate logistic**)

```
ul_CHL<-glm(CHD~CHL,family=binomial)
```


単変量ロジスティック回帰分析 R 計算

例として **CHL** を用いるが、他の変数の計算も同様である

- ▶ **epicalc Package** をロードしておく

```
library(epicalc)
```

- ▶ ロジスティック回帰分析は、**glm** 関数に **binomial** オプションを付けて行う

出力を `ul_CHL` と名付ける (**univariate logistic**)

```
ul_CHL<-glm(CHD~CHL,family=binomial)
```

- ▶ 回帰式パラメータ、**Wald** 検定、尤度比検定の出力

```
summary(ul_CHL)
```

単変量ロジスティック回帰分析 R 計算

例として **CHL** を用いるが、他の変数の計算も同様である

- ▶ **epicalc Package** をロードしておく

```
library(epicalc)
```

- ▶ ロジスティック回帰分析は、**glm** 関数に **binomial** オプションを付けて行う

出力を `ul_CHL` と名付ける (**univariate logistic**)

```
ul_CHL<-glm(CHD~CHL,family=binomial)
```

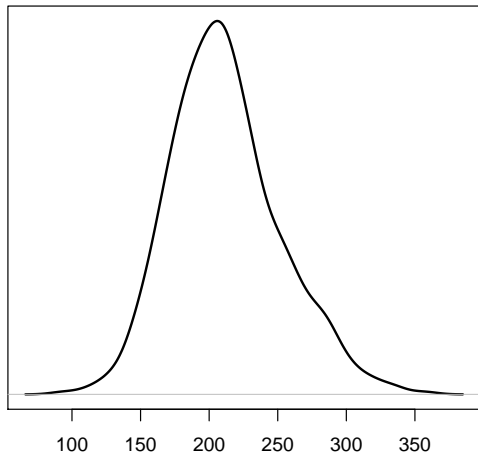
- ▶ 回帰式パラメータ、**Wald** 検定、尤度比検定の出力

```
summary(ul_CHL)
```

- ▶ オッズ比とその信頼区間の出力

```
logistic.display(ul_CHL,decimal=4)
```

609 名の CHL の度数分布



単変量ロジスティック回帰分析 CHL パラメータ

数値説明変数 CHL の場合

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta \text{CHL}))}$$

- ▶ JMP 切片は $\alpha = -3.538$ 、係数は $\beta = 0.007$

名義ロジスティックのあてはめ CHD

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上側95%
切片	-3.5382603	0.6869347	26.53	<.0001*	-4.9031972	-2.204806
CHL	0.00700361	0.0030638	5.23	0.0223*	0.00095472	0.01299527

推定値は次の対数オッズに対するものです:
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.22544844	0.0223*

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.13397364	0.0235*

単変量ロジスティック回帰分析 CHL パラメータ

数値説明変数 CHL の場合

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta \text{CHL}))}$$

- ▶ **JMP** 切片は $\alpha = -3.538$ 、係数は $\beta = 0.007$
- ▶ **R** 切片は $\alpha = -3.538$ 、係数は $\beta = 0.007$

```
> summary(ul_CHL)
```

```
Call:
```

```
glm(formula = CHD ~ CHL, family = binomial)
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-0.7304	-0.5196	-0.4722	-0.4287	2.2803

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-3.538260	0.686879	-5.151	2.59e-07	***
CHL	0.007004	0.003064	2.286	0.0223	*

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta \text{CHL}))} \text{ の計算例}$$

JMP, R とともに, $\alpha = -3.538$, $\beta = 0.007$ と分かった

今のモデルでは $P(\text{CHD})$ は変数 **CHL** の関数
これを明記したいときには、 $P(\text{CHD}|\text{CHL})$ と書く

単変量ロジスティック回帰分析 CHL 確率 1

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta \text{CHL}))} \text{ の計算例}$$

JMP, R ともに, $\alpha = -3.538$, $\beta = 0.007$ と分かった
今のモデルでは $P(\text{CHD})$ は変数 **CHL** の関数
これを明記したいときには、 $P(\text{CHD}|\text{CHL})$ と書く

- ▶ **No.1, No.2** の人はそれぞれ, **CHL = 270, CHL = 159**

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta \text{CHL}))} \text{ の計算例}$$

JMP, R とともに, $\alpha = -3.538$, $\beta = 0.007$ と分かった
今のモデルでは $P(\text{CHD})$ は変数 CHL の関数
これを明記したいときには, $P(\text{CHD}|\text{CHL})$ と書く

- ▶ No.1, No.2 の人はそれぞれ, CHL = 270, CHL = 159
- ▶ CHL = 270 のとき,
$$P(\text{CHD}|\text{CHL} = 270) = \frac{1}{1 + \exp(-(-3.538 + 0.007 \times 270))} = 0.161,$$
- ▶ CHL = 159 のとき,
$$P(\text{CHD}|\text{CHL} = 159) = \frac{1}{1 + \exp(-(-3.538 + 0.007 \times 159))} = 0.0813$$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta \text{CHL}))} \text{ の計算例}$$

JMP, R とともに, $\alpha = -3.538$, $\beta = 0.007$ と分かった
 今のモデルでは $P(\text{CHD})$ は変数 CHL の関数
 これを明記したいときには, $P(\text{CHD}|\text{CHL})$ と書く

- ▶ No.1, No.2 の人はそれぞれ, CHL = 270, CHL = 159
- ▶ CHL = 270 のとき,

$$P(\text{CHD}|\text{CHL} = 270) = \frac{1}{1 + \exp(-(-3.538 + 0.007 \times 270))} = 0.161,$$
- ▶ CHL = 159 のとき,

$$P(\text{CHD}|\text{CHL} = 159) = \frac{1}{1 + \exp(-(-3.538 + 0.007 \times 159))} = 0.0813$$

	CHL	$P(\text{CHD} \text{CHL})$
No. 1	270	0.161
No. 2	159	0.0813

単変量ロジスティック回帰分析 CHL 確率 2

ソフトによる $P(\text{CHD})$ の出力

	CHL	$P(\text{CHD} \text{CHL})$
No. 1	270	0.161
No. 2	159	0.0813

▶ JMP の場合

PT	線形[presence]	確率[presence]	確率[absence]
1 ence	-1.64728582	0.1614761163	0.8385238837
2 ence	-2.424686441	0.0813094989	0.9186905011
3 senc	-2.130534855	0.1061642268	0.8938357732
4 senc	-2.284614257	0.092405243	0.907594757
5 ence	-1.836383268	0.1374795988	0.8625204012
6 ence	-1.773350785	0.1451261208	0.8548738792
7 ence	-2.284614257	0.092405243	0.907594757
8 ence	-2.018477107	0.1172765531	0.8827234469
9 ence	-2.305625085	0.0906581642	0.9093418358
10 senc	-1.787358004	0.1433969455	0.8566030545

単変量ロジスティック回帰分析 CHL 確率 2

ソフトによる $P(\text{CHD})$ の出力

	CHL	$P(\text{CHD} \text{CHL})$
No. 1	270	0.161
No. 2	159	0.0813

- ▶ JMP の場合
- ▶ R の場合

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
>
> CHL[1:2]
[1] 270 159
> fitted(u1_CHL)[1:2]
      1      2
0.1614761 0.0813095
>
>
>
>
>
>
>
```

単変量ロジスティック回帰分析 CHL 確率 3

CHL=270 の集団での CHD の有病率 $P(\text{CHD}|\text{CHL}=270)$

点推定値 JMP,R とも 0.161 と分かった

信頼区間 JMP,R とも出力されない

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算

- ▶ パラメータ (α, β) の分散共分散行列を読む

単変量ロジスティック回帰分析 CHL 確率 3

CHL=270 の集団での CHD の有病率 $P(\text{CHD}|\text{CHL}=270)$

点推定値 JMP,R とも 0.161 と分かった

信頼区間 JMP,R とも出力されない

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算

- ▶ パラメータ (α, β) の分散共分散行列を読む

- ▶ JMP の場合 $\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片	-3.5382603	0.6869347	26.53	<.0001*
CHL	0.00700361	0.0080638	5.23	0.0223*

推定値は次の対数オッズに対するものです:

presence/absence

推定値の共分散

共分散

	切片	CHL
切片	0.47187933	-0.0020684
CHL	-0.0020684	0.00000939

効果の尤度比検定

単変量ロジスティック回帰分析 CHL 確率 3

CHL=270 の集団での CHD の有病率 $P(\text{CHD}|\text{CHL}=270)$

点推定値 JMP,R とも 0.161 と分かった

信頼区間 JMP,R とも出力されない

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算

▶ パラメータ (α, β) の分散共分散行列を読む

▶ JMP の場合
$$\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$$

▶ R の場合 JMP と同じ

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> sum_CHL<-summary(ul_CHL)
> sum_CHL$cov.unscaled
              (Intercept)              CHL
(Intercept)  0.471802131 -0.002068156644
CHL          -0.002068157  0.000009385775
>
>
>
>
>
>
>
```

単変量ロジスティック回帰分析 CHL 確率 4

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算の続き

$$\text{CHL} = 270 \text{ のとき: } h = \alpha + 270\beta, P(\text{CHD}|\text{CHL}=270) = \frac{1}{1+\exp(-h)}$$
$$h \text{ の点推定値は } h = -3.538 + 0.007 \times 270 = -1.648$$

▶ h の分散 (**variance**) の式

$$\text{Var}(h) = \text{Var}(\alpha, \alpha) + 2 \times 270 \text{Covar}(\alpha, \beta) + 270^2 \text{Var}(\beta, \beta)$$

単変量ロジスティック回帰分析 CHL 確率 4

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算の続き

CHL = 270 のとき: $h = \alpha + 270\beta$, $P(\text{CHD}|\text{CHL}=270) = \frac{1}{1+\exp(-h)}$
 h の点推定値は $h = -3.538 + 0.007 \times 270 = -1.648$

- ▶ h の分散 (variance) の式

$$\text{Var}(h) = \text{Var}(\alpha, \alpha) + 2 \times 270 \text{Covar}(\alpha, \beta) + 270^2 \text{Var}(\beta, \beta)$$

- ▶ (α, β) の分散共分散行列 $\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$ より

単変量ロジスティック回帰分析 CHL 確率 4

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算の続き

CHL = 270 のとき: $h = \alpha + 270\beta$, $P(\text{CHD}|\text{CHL}=270) = \frac{1}{1+\exp(-h)}$
 h の点推定値は $h = -3.538 + 0.007 \times 270 = -1.648$

- ▶ h の分散 (variance) の式

$$\text{Var}(h) = \text{Var}(\alpha, \alpha) + 2 \times 270 \text{Covar}(\alpha, \beta) + 270^2 \text{Var}(\beta, \beta)$$

- ▶ (α, β) の分散共分散行列 $\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$ より

- ▶ $\text{Var}(h) = 0.472 - 2 \times 270 \times 0.00207 + 270^2 \times 0.00000939 = 0.0387$

単変量ロジスティック回帰分析 CHL 確率 4

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算の続き

CHL = 270 のとき: $h = \alpha + 270\beta$, $P(\text{CHD}|\text{CHL}=270) = \frac{1}{1+\exp(-h)}$
 h の点推定値は $h = -3.538 + 0.007 \times 270 = -1.648$

- ▶ h の分散 (variance) の式
 $\text{Var}(h) = \text{Var}(\alpha, \alpha) + 2 \times 270 \text{Covar}(\alpha, \beta) + 270^2 \text{Var}(\beta, \beta)$
- ▶ (α, β) の分散共分散行列 $\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$ より
- ▶ $\text{Var}(h) = 0.472 - 2 \times 270 \times 0.00207 + 270^2 \times 0.00000939 = 0.0387$
- ▶ h の 95%信頼限界 正規分布の両側 95%点 1.96
 $-1.648 \pm 1.96 \sqrt{0.0387} = -1.262, -2.034$
- ▶ $\frac{1}{1+\exp(-(-1.262))} = 0.221$, $\frac{1}{1+\exp(-(-2.034))} = 0.116$

単変量ロジスティック回帰分析 CHL 確率 4

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算の続き

$$\text{CHL} = 270 \text{ のとき: } h = \alpha + 270\beta, P(\text{CHD}|\text{CHL}=270) = \frac{1}{1+\exp(-h)}$$
$$h \text{ の点推定値は } h = -3.538 + 0.007 \times 270 = -1.648$$

- ▶ h の分散 (variance) の式
$$\text{Var}(h) = \text{Var}(\alpha, \alpha) + 2 \times 270 \text{Covar}(\alpha, \beta) + 270^2 \text{Var}(\beta, \beta)$$
- ▶ (α, β) の分散共分散行列 $\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$ より
- ▶ $\text{Var}(h) = 0.472 - 2 \times 270 \times 0.00207 + 270^2 \times 0.00000939 = 0.0387$
- ▶ h の 95%信頼限界 正規分布の両側 95%点 1.96
$$-1.648 \pm 1.96 \sqrt{0.0387} = -1.262, -2.034$$
- ▶ $\frac{1}{1+\exp(-(-1.262))} = 0.221, \frac{1}{1+\exp(-(-2.034))} = 0.116$
- ▶ 求める 95%信頼区間: $0.116 < P(\text{CHD}|\text{CHL}=270) < 0.221$
($P(\text{CHD}|\text{CHL}=270)$ の点推定値は 0.161 であった)

単変量ロジスティック回帰分析 CHL 確率 4

$P(\text{CHD}|\text{CHL}=270)$ の信頼区間の手計算の続き

$$\text{CHL} = 270 \text{ のとき: } h = \alpha + 270\beta, P(\text{CHD}|\text{CHL}=270) = \frac{1}{1+\exp(-h)}$$
$$h \text{ の点推定値は } h = -3.538 + 0.007 \times 270 = -1.648$$

- ▶ h の分散 (variance) の式
$$\text{Var}(h) = \text{Var}(\alpha, \alpha) + 2 \times 270 \text{Covar}(\alpha, \beta) + 270^2 \text{Var}(\beta, \beta)$$
- ▶ (α, β) の分散共分散行列 $\begin{pmatrix} 0.472 & -0.00207 \\ -0.00207 & 0.00000939 \end{pmatrix}$ より
- ▶ $\text{Var}(h) = 0.472 - 2 \times 270 \times 0.00207 + 270^2 \times 0.00000939 = 0.0387$
- ▶ h の 95%信頼限界 正規分布の両側 95%点 1.96
$$-1.648 \pm 1.96 \sqrt{0.0387} = -1.262, -2.034$$
- ▶ $\frac{1}{1+\exp(-(-1.262))} = 0.221, \frac{1}{1+\exp(-(-2.034))} = 0.116$
- ▶ 求める 95%信頼区間: $0.116 < P(\text{CHD}|\text{CHL}=270) < 0.221$
($P(\text{CHD}|\text{CHL}=270)$ の点推定値は 0.161 であった)
- ▶ 以後のモデルで $P(\text{CHD})$ の信頼区間の話題は省く

単変量ロジスティック回帰分析 CHL 検定

$\beta = 0$ の検定

- ▶ JMP p 値は Wald 検定 0.0223、尤度比検定 0.0235

名義ロジスティックのあてはめ CHD

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上側95%
切片	-3.5382603	0.6869347	26.53	<.0001*	-4.9031972	-2.204806
CHL	0.00700361	0.0030638	5.23	0.0223*	0.00095472	0.01299527

推定値は次の対数オッズに対するものです:
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.22544844	0.0223*

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.13397364	0.0235*

単変量ロジスティック回帰分析 CHL 検定

$\beta = 0$ の検定

- ▶ **JMP** p 値は **Wald 検定 0.0223**、**尤度比検定 0.0235**
- ▶ **R** p 値は **Wald 検定 0.0223**、**尤度比検定 0.0235**



```
> logistic.display(ul_CHL, decimal=4)
```

```
Logistic regression predicting CHD : presence vs absence
```

	OR (95%CI)	P(Wald's test)	P(LR-test)
CHL (cont. var.)	1.007 (1.001, 1.0131)	0.02225	0.02346

```
Log-likelihood = -216.712162
```

```
No. of observations = 609
```

```
AIC value = 437.424323
```

```
>  
>
```

単変量ロジスティック回帰分析 CHL 検定

$\beta = 0$ の検定

- ▶ JMP p 値は Wald 検定 0.0223、尤度比検定 0.0235
- ▶ R p 値は Wald 検定 0.0223、尤度比検定 0.0235
- ▶ JMP パラメータ値の出力の表の p 値は Wald 検定の p 値

名義ロジスティックのあてはめ CHD

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上側95%
切片	-3.5382603	0.6869347	26.53	<.0001*	-4.9031972	-2.204806
CHL	0.00700361	0.0030638	5.23	0.0223*	0.00095472	0.01299527

推定値は次の対数オッズに対するものです:
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.22544844	0.0223*

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.13397364	0.0235*

単変量ロジスティック回帰分析 CHL 検定

$\beta = 0$ の検定

- ▶ **JMP** p 値は **Wald 検定 0.0223**、尤度比検定 **0.0235**
- ▶ **R** p 値は **Wald 検定 0.0223**、尤度比検定 **0.0235**
- ▶ **JMP** パラメータ値の出力の表の p 値は **Wald 検定の p 値**
- ▶ **R** パラメータ値の出力の表 p 値は **Wald 検定の p 値**

```
> summary(ul_CHL)
```

```
Call:
```

```
glm(formula = CHD ~ CHL, family = binomial)
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-0.7304	-0.5196	-0.4722	-0.4287	2.2803

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-3.538260	0.686879	-5.151	2.59e-07	***
CHL	0.007004	0.003064	2.286	0.0223	*

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


単変量ロジスティック回帰分析

CHL オッズ比 1

- ▶ 確率 $0 < P < 1$ に対してオッズ $0 < O < \infty$

$$O = \frac{P}{1-P} \iff P = \frac{O}{1+O}$$

単変量ロジスティック回帰分析 CHL オッズ比 1

- ▶ 確率 $0 < P < 1$ に対してオッズ $0 < O < \infty$

$$O = \frac{P}{1-P} \iff P = \frac{O}{1+O}$$

- ▶ 例. $P = 0.25 \iff O = \frac{1}{3}$, $P = 0.5 \iff O = 1$, $P = 0.75 \iff O = 3$

単変量ロジスティック回帰分析 CHL オッズ比 1

- ▶ 確率 $0 < P < 1$ に対してオッズ $0 < O < \infty$

$$O = \frac{P}{1-P} \iff P = \frac{O}{1+O}$$

- ▶ 例. $P = 0.25 \iff O = \frac{1}{3}$, $P = 0.5 \iff O = 1$, $P = 0.75 \iff O = 3$

- ▶ 公式 $P = \frac{1}{1+\exp(-z)} \iff O = \exp(z)$

証明. $O = \frac{\frac{1}{1+\exp(-z)}}{1 - \frac{1}{1+\exp(-z)}} = \frac{1}{1+\exp(-z)-1} = \frac{1}{\exp(-z)} = \exp(z)$

単変量ロジスティック回帰分析 CHL オッズ比 1

- ▶ 確率 $0 < P < 1$ に対してオッズ $0 < O < \infty$

$$O = \frac{P}{1-P} \iff P = \frac{O}{1+O}$$

- ▶ 例. $P = 0.25 \iff O = \frac{1}{3}$, $P = 0.5 \iff O = 1$, $P = 0.75 \iff O = 3$

- ▶ 公式 $P = \frac{1}{1+\exp(-z)} \iff O = \exp(z)$

$$\text{証明. } O = \frac{\frac{1}{1+\exp(-z)}}{1 - \frac{1}{1+\exp(-z)}} = \frac{1}{1+\exp(-z)-1} = \frac{1}{\exp(-z)} = \exp(z)$$

- ▶ CHL = c のとき

CHD の確率 $P(\text{CHD}|\text{CHL} = c)$ とオッズ $O(\text{CHD}|\text{CHL} = c)$

$$P(\text{CHD}|\text{CHL} = c) = \frac{1}{1+\exp(-(\alpha+\beta c))}$$
$$O(\text{CHD}|\text{CHL} = c) = \exp(\alpha + \beta c) = \exp(\alpha) \exp(\beta c)$$

- ▶ 確率 $0 < P < 1$ に対してオッズ $0 < O < \infty$

$$O = \frac{P}{1-P} \iff P = \frac{O}{1+O}$$

- ▶ 例. $P = 0.25 \iff O = \frac{1}{3}$, $P = 0.5 \iff O = 1$, $P = 0.75 \iff O = 3$

- ▶ 公式 $P = \frac{1}{1+\exp(-z)} \iff O = \exp(z)$

$$\text{証明. } O = \frac{\frac{1}{1+\exp(-z)}}{1-\frac{1}{1+\exp(-z)}} = \frac{1}{1+\exp(-z)-1} = \frac{1}{\exp(-z)} = \exp(z)$$

- ▶ CHL = c のとき

CHD の確率 $P(\text{CHD}|\text{CHL} = c)$ とオッズ $O(\text{CHD}|\text{CHL} = c)$

$$P(\text{CHD}|\text{CHL} = c) = \frac{1}{1+\exp(-(\alpha+\beta c))}$$

$$O(\text{CHD}|\text{CHL} = c) = \exp(\alpha + \beta c) = \exp(\alpha) \exp(\beta c)$$

- ▶ CHL が 1 増えたときのオッズ比

$$\frac{O(\text{CHD}|\text{CHL}=c+1)}{O(\text{CHD}|\text{CHL}=c)} = \frac{\exp(\alpha) \exp(\beta c) \exp(\beta)}{\exp(\alpha) \exp(\beta c)} = \exp(\beta)$$

- ▶ **JMP** ではオッズ比の信頼区間は **profile likelihood** 法で計算されている
この方法には本講演では触れない

- ▶ **JMP** ではオッズ比の信頼区間は **profile likelihood** 法で計算されている
この方法には本講演では触れない
- ▶ **R** の **epicalc Package** ではオッズ比の信頼区間は **Wald** 法に基づく
これを説明する

- ▶ **JMP** ではオッズ比の信頼区間は **profile likelihood** 法で計算されている
この方法には本講演では触れない
- ▶ **R** の **epicalc Package** ではオッズ比の信頼区間は **Wald** 法に基づく
これを説明する
- ▶ 「 $\beta = 0$ 」の **Wald** 検定の p 値が $p < 0.05$
⇔ **CHL** のオッズ比の信頼区間が **1** を含まない

- ▶ **JMP** ではオッズ比の信頼区間は **profile likelihood** 法で計算されている
この方法には本講演では触れない
- ▶ **R** の **epicalc Package** ではオッズ比の信頼区間は **Wald** 法に基づく
これを説明する
- ▶ 「 $\beta = 0$ 」の **Wald** 検定の p 値が $p < 0.05$
⇔ **CHL** のオッズ比の信頼区間が **1** を含まない
- ▶ その理由 (ここでは β の点推定値を $\hat{\beta}$ と書く)
「 $\beta = 0$ 」の **Wald** 検定の p 値が $p < 0.05$
⇔ 標準正規分布 z において $P(|z| > |\frac{\hat{\beta}-0}{\beta \text{の標準誤差}}|) < 0.05$
⇔ β の 95%信頼区間: $\hat{\beta} \pm 1.96 \times (\beta \text{の標準誤差})$
が **0** を含まない
⇔ オッズ比の 95%信頼区間: $\exp(\hat{\beta} \pm 1.96 \times (\beta \text{の標準誤差}))$
が **1** を含まない

単変量ロジスティック回帰分析 CHL オッズ比 3

β とその信頼区間 $\xrightarrow{\text{exp}}$ CHL のオッズ比とその信頼区間

R の `epicalc package` の場合

- ▶ β の信頼区間は $0.007 \pm 1.96 \times 0.00306 = (0.00100, 0.0130)$

```
> summary(ul_CHL)

Call:
glm(formula = CHD ~ CHL, family = binomial)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7304 -0.5196 -0.4722 -0.4287  2.2803

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.538260    0.686879  -5.151 2.59e-07 ***
CHL          0.007004    0.003064   2.286  0.0223 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

単変量ロジスティック回帰分析 CHL オッズ比 3

β とその信頼区間 $\xrightarrow{\text{exp}}$ CHL のオッズ比とその信頼区間

R の `epicalc` package の場合

- ▶ β の信頼区間は $0.007 \pm 1.96 \times 0.00306 = (0.00100, 0.0130)$
- ▶ CHL のオッズ比と信頼区間は
 $\exp(0.007) = 1.007, (\exp(0.00100), \exp(0.0130)) = (1.001, 1.013)$

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
```

```
> logistic.display(ul_CHL, decimal=4)
```

```
Logistic regression predicting CHD : presence vs absence
```

	OR (95%CI)	P(Wald's test)	P(LR-test)
CHL (cont. var.)	1.007 (1.001, 1.0131)	0.02225	0.02346

```
Log-likelihood = -216.712162
No. of observations = 609
AIC value = 437.424323
```

```
>
>
```

単変量ロジスティック回帰分析 CHL オッズ比 4

JMP の場合

- ▶ 出力： β の値 0.00700361 と信頼区間 (0.00095472, 0.01299527)
 $(0.00095472 + 0.01299527)/2 = 0.006974995 \neq 0.00700361$
 区間是对称でない： $0.00700361 \pm 1.96 \times 0.0030638$ の式ではない
profile likelihood 法が用いられている

名義ロジスティックのあてはめ CHD

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob)>ChiSq)	下側95%	上側95%
切片	-3.5382603	0.6869347	26.53	<.0001*	-4.9031972	-2.204806
CHL	0.00700361	0.0030638	5.23	0.0223*	0.00095472	0.01299527

推定値は次の対数オッズに対するものです：
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob)>ChiSq)
CHL	1	1	5.22544844	0.0223*

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob)>ChiSq)
CHL	1	1	5.13397364	0.0235*

単変量ロジスティック回帰分析 CHL オッズ比 4

JMP の場合

- ▶ 出力： β の値 0.00700361 と信頼区間 (0.00095472, 0.01299527)
(0.00095472 + 0.01299527)/2 = 0.006974995 \neq 0.00700361
区間是对称でない： $0.00700361 \pm 1.96 \times 0.0030638$ の式ではない
profile likelihood 法が用いられている
- ▶ オッズ比と信頼区間は $\exp(0.00700361) = 1.007028$,
($\exp(0.00095472)$, $\exp(0.01299527)$) = (1.000955, 1.01308)

evans - 名義ロジスティックのあてはめ - JMP

名義ロジスティックのあてはめ CHD

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CHL	1	1	5.13397364	0.0235*

オッズ比

CHD: presence 対 absence のオッズ比に対して

単位オッズ比

連続変数が1単位だけ変化した場合

項	オッズ比	下側95%	上側95%	逆数
CHL	1.007028	1.000955	1.01308	0.9930209

範囲オッズ比

連続変数が範囲全体で変化した場合

項	オッズ比	下側95%	上側95%	逆数
CHL	6.308824	1.285426	30.50092	0.1585082

単変量ロジスティック回帰分析 CAT パラメータ

2 値名義説明変数 CAT の場合
(609 名中 normal 487, high 122)

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))}$$

- ▶ JMP 切片は $\alpha = -1.784$ 、係数は $\beta_{\text{high}} = 0.526$

evans - 名義ロジスティックのあてはめ - JMP

名義ロジスティックのあてはめ CHD

モデル全体の検定

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob)>ChiSq	下側95%	上側95%
切片	-1.7837101	0.1346736	175.42	<.0001*	-2.0560704	-1.5270369
CAT[high]	0.52567006	0.1346736	15.24	<.0001*	0.25742451	0.7872221

推定値は次の対数オッズに対するものです:
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob)>ChiSq
CAT	1	1	15.2356593	<.0001*

効果の尤度比検定

単変量ロジスティック回帰分析 CAT パラメータ

2 値名義説明変数 CAT の場合
(609 名中 normal 487, high 122)

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))}$$

- ▶ JMP 切片は $\alpha = -1.784$ 、係数は $\beta_{\text{high}} = 0.526$
- ▶ R 切片は $\alpha = -2.309$ 、係数は $\beta_{\text{high}} = 1.051$

```
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
> summary(ul_CAT)

Call:
glm(formula = CHD ~ CAT, family = binomial)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7073 -0.4352 -0.4352 -0.4352  2.1928

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.3094     0.1581 -14.610 < 2e-16 ***
CATHigh       1.0513     0.2693   3.903 9.49e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

単変量ロジスティック回帰分析 CAT 確率 1

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))} \text{ の計算例}$$

JMP $\alpha = -1.784, \beta_{\text{high}} = 0.526$; **R** $\alpha = -2.309, \beta_{\text{high}} = 1.051$

単変量ロジスティック回帰分析 CAT 確率 1

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))} \text{ の計算例}$$

JMP $\alpha = -1.784, \beta_{\text{high}} = 0.526$; **R** $\alpha = -2.309, \beta_{\text{high}} = 1.051$

- ▶ No.2, No.3 の人はそれぞれ, CAT=normal, CAT=high

単変量ロジスティック回帰分析 CAT 確率 1

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))} \text{ の計算例}$$

$$\text{JMP } \alpha = -1.784, \beta_{\text{high}} = 0.526; \quad \text{R } \alpha = -2.309, \beta_{\text{high}} = 1.051$$

- ▶ No.2, No.3 の人はそれぞれ, CAT=normal, CAT=high
- ▶ JMP normal \Leftrightarrow CAT = -1, high \Leftrightarrow CAT = 1

$$P(\text{CHD}|\text{CAT} = -1) = \frac{1}{1 + \exp(-(-1.784 - 0.526))} = 0.0903,$$

$$P(\text{CHD}|\text{CAT} = 1) = \frac{1}{1 + \exp(-(-1.784 + 0.526))} = 0.221$$

	CAT		$P(\text{CHD} \text{CAT})$
No. 2	normal	-1	0.0903
No. 3	high	1	0.221

単変量ロジスティック回帰分析 CAT 確率 1

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))} \text{ の計算例}$$

JMP $\alpha = -1.784, \beta_{\text{high}} = 0.526$; **R** $\alpha = -2.309, \beta_{\text{high}} = 1.051$

▶ **No.2, No.3** の人はそれぞれ, **CAT=normal, CAT=high**

▶ **JMP** normal \Leftrightarrow CAT = -1, high \Leftrightarrow CAT = 1

$$P(\text{CHD}|\text{CAT} = -1) = \frac{1}{1 + \exp(-(-1.784 - 0.526))} = 0.0903,$$

$$P(\text{CHD}|\text{CAT} = 1) = \frac{1}{1 + \exp(-(-1.784 + 0.526))} = 0.221$$

	CAT		$P(\text{CHD} \text{CAT})$
No. 2	normal	-1	0.0903
No. 3	high	1	0.221

▶ **R** normal \Leftrightarrow CAT = 0, high \Leftrightarrow CAT = 1

$$P(\text{CHD}|\text{CAT} = 0) = \frac{1}{1 + \exp(-(-2.309))} = 0.0903,$$

$$P(\text{CHD}|\text{CAT} = 1) = \frac{1}{1 + \exp(-(-2.309 + 1.051))} = 0.221$$

	CAT		$P(\text{CHD} \text{CAT})$
No. 2	normal	0	0.0903
No. 3	high	1	0.221

単変量ロジスティック回帰分析

CAT 確率 2

ソフトによる $P(\text{CHD})$ の出力

	CAT	$P(\text{CHD} \text{CAT})$
No. 2	normal	0.0903
No. 3	high	0.221

▶ JMP の場合

evans - JMP

ファイル(F) 編集(E) テーブル(T) 行(R) 列(C) 実験計画(DOE)(D) 分析(A) グラフ(G) ツール(O) 表示(V) ウィンドウ(W)

evans

		線形[presence]	確率[presence]	確率[absence]
1	nce	-2.309380136	0.090349076	0.909650924
2	nce	-2.309380136	0.090349076	0.909650924
3	nce	-1.258040026	0.2213114754	0.7786885246
4	nce	-1.258040026	0.2213114754	0.7786885246
5	nce	-2.309380136	0.090349076	0.909650924
6	nce	-2.309380136	0.090349076	0.909650924
7	nce	-2.309380136	0.090349076	0.909650924
8	nce	-2.309380136	0.090349076	0.909650924
9	nce	-2.309380136	0.090349076	0.909650924
10	nce	-2.309380136	0.090349076	0.909650924

列(13/0)

- CHD ✓ *
- CAT ✓ *
- AGEGR ✓ *
- CHL
- SMK ✓
- ECG ✓ *
- DBP

単変量ロジスティック回帰分析

CAT 確率 2

ソフトによる $P(\text{CHD})$ の出力

	CAT	$P(\text{CHD} \text{CAT})$
No. 2	normal	0.0903
No. 3	high	0.221

- ▶ JMP の場合
- ▶ R の場合

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> CAT[2:3]
[1] normal high
Levels: normal high
> fitted(ul_CAT)[2:3]
      2      3
0.09034908 0.22131148
>
>
>
>
>
>
>
```

$$\beta_{\text{high}} = 0 \text{ の検定}$$

- ▶ JMP Wald 検定 $p < 0.0001$ 、尤度比検定 $p = 0.0002$

ロジスティック回帰分析の結果

モデル全体の検定

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上側95%
切片	-1.7837101	0.1346736	175.42	<.0001*	-2.0560704	-1.5270369
CAT[high]	0.52567006	0.1346736	15.24	<.0001*	0.25742451	0.7872221

推定値は次の対数オッズに対するものです:
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
CAT	1	1	15.2356593	<.0001*

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CAT	1	1	14.1312102	0.0002*

$\beta_{\text{high}} = 0$ の検定

- ▶ **JMP** Wald 検定 $p < 0.0001$ 、尤度比検定 $p = 0.0002$
- ▶ **R** Wald 検定 $p < 0.001$ 、尤度比検定 $p < 0.001$

```

R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> logistic.display(ul_CAT,decimal=4)

Logistic regression predicting CHD : presence vs absence

                OR(95%CI)                P(Wald's test) P(LR-test)
CAT: high vs normal 2.8615 (1.6878,4.8513) < 0.001      < 0.001

Log-likelihood = -212.213543
No. of observations = 609
AIC value = 428.427087

>
>
>
    
```

単変量ロジスティック回帰分析 CAT オッズ比 1

オッズ比とその信頼区間

normal を reference とした high のオッズ比

- ▶ JMP 2.861 (1.673 – 4.828)

名義ロジスティックのあてはめ CHD

パラメータ推定値

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
CAT	1	1	15.2356593	<.0001*

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CAT	1	1	14.1312102	0.0002*

オッズ比

CHD: presence対absenceのオッズ比に対して

CATのオッズ比

水準1 /水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
normal high	0.3494691	0.0002*	0.2071226	0.5975908
high normal	2.8614833	0.0002*	1.6733858	4.8280575

単変量ロジスティック回帰分析 CAT オッズ比 1

オッズ比とその信頼区間

normal を reference とした high のオッズ比

- ▶ **JMP** 2.861 (1.673 – 4.828)
- ▶ **R** 2.862 (1.688 – 4.851)

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> logistic.display(ul_CAT,decimal=4)

Logistic regression predicting CHD : presence vs absence

                OR(95%CI)                P(Wald's test) P(LR-test)
CAT: high vs normal 2.8615 (1.6878,4.8513) < 0.001      < 0.001

Log-likelihood = -212.213543
No. of observations = 609
AIC value = 428.427087

>
>
>
```

単変量ロジスティック回帰分析 CAT オッズ比 1

オッズ比とその信頼区間

normal を reference とした high のオッズ比

- ▶ **JMP** 2.861 (1.673 – 4.828)
- ▶ **R** 2.862 (1.688 – 4.851)
- ▶ 2 値名義変数のオッズ比の信頼区間は、**JMP** と **R** でわずかに違う

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> logistic.display(ul_CAT,decimal=4)

Logistic regression predicting CHD : presence vs absence

                OR(95%CI)                P(Wald's test) P(LR-test)
CAT: high vs normal 2.8615 (1.6878,4.8513) < 0.001      < 0.001

Log-likelihood = -212.213543
No. of observations = 609
AIC value = 428.427087

>
>
>
```

単変量ロジスティック回帰分析 CAT オッズ比 2

- ▶ CAT に対する CHD の確率 $P(\text{CHD}|\text{CAT})$ とオッズ $O(\text{CHD}|\text{CAT})$

$$P(\text{CHD}|\text{CAT}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))}$$

公式 $P = \frac{1}{1 + \exp(-z)}$ $\Leftrightarrow O = \exp(z)$ より

$$O(\text{CHD}|\text{CAT}) = \exp(\alpha + \beta_{\text{high}} \text{CAT}) = \exp(\alpha) \exp(\beta_{\text{high}} \text{CAT})$$

- ▶ CAT に対する CHD の確率 $P(\text{CHD}|\text{CAT})$ とオッズ $O(\text{CHD}|\text{CAT})$

$$P(\text{CHD}|\text{CAT}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))}$$

公式 $P = \frac{1}{1 + \exp(-z)}$ $\Leftrightarrow O = \exp(z)$ より

$$O(\text{CHD}|\text{CAT}) = \exp(\alpha + \beta_{\text{high}} \text{CAT}) = \exp(\alpha) \exp(\beta_{\text{high}} \text{CAT})$$

- ▶ normal を reference とした high のオッズ比 $\frac{O(\text{CHD}|\text{CAT}=\text{high})}{O(\text{CHD}|\text{CAT}=\text{normal})}$
JMP normal のとき $\text{CAT} = -1$, high のとき $\text{CAT} = 1$

$$\frac{O(\text{CHD}|\text{CAT}=1)}{O(\text{CHD}|\text{CAT}=-1)} = \frac{\exp(\alpha) \exp(\beta_{\text{high}})}{\exp(\alpha) \exp(-\beta_{\text{high}})} = \exp(2 \times \beta_{\text{high}})$$

- R** normal のとき $\text{CAT} = 0$, high のとき $\text{CAT} = 1$

$$\frac{O(\text{CHD}|\text{CAT}=1)}{O(\text{CHD}|\text{CAT}=0)} = \frac{\exp(\alpha) \exp(\beta_{\text{high}})}{\exp(\alpha)} = \exp(\beta_{\text{high}})$$

- ▶ CAT に対する CHD の確率 $P(\text{CHD}|\text{CAT})$ とオッズ $O(\text{CHD}|\text{CAT})$

$$P(\text{CHD}|\text{CAT}) = \frac{1}{1 + \exp(-(\alpha + \beta_{\text{high}} \text{CAT}))}$$

公式 $P = \frac{1}{1 + \exp(-z)}$ $\Leftrightarrow O = \exp(z)$ より

$$O(\text{CHD}|\text{CAT}) = \exp(\alpha + \beta_{\text{high}} \text{CAT}) = \exp(\alpha) \exp(\beta_{\text{high}} \text{CAT})$$

- ▶ normal を reference とした high のオッズ比 $\frac{O(\text{CHD}|\text{CAT}=\text{high})}{O(\text{CHD}|\text{CAT}=\text{normal})}$
JMP normal のとき $\text{CAT} = -1$, high のとき $\text{CAT} = 1$

$$\frac{O(\text{CHD}|\text{CAT}=1)}{O(\text{CHD}|\text{CAT}=-1)} = \frac{\exp(\alpha) \exp(\beta_{\text{high}})}{\exp(\alpha) \exp(-\beta_{\text{high}})} = \exp(2 \times \beta_{\text{high}})$$

- R** normal のとき $\text{CAT} = 0$, high のとき $\text{CAT} = 1$

$$\frac{O(\text{CHD}|\text{CAT}=1)}{O(\text{CHD}|\text{CAT}=0)} = \frac{\exp(\alpha) \exp(\beta_{\text{high}})}{\exp(\alpha)} = \exp(\beta_{\text{high}})$$

- ▶ 例.

JMP $\beta_{\text{high}} = 0.526$ $\exp(2 \times 0.526) = 2.863$

R $\beta_{\text{high}} = 1.051$ $\exp(1.051) = 2.861$

丸め誤差を除いて両者は一致

単変量ロジスティック回帰分析 AGEGR パラメータ

4 値名義説明変数 AGEGR の場合

(40 - 49 247 名, 50 - 59 203 名, 60 - 69 115 名, 70- 44 名)

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}))}$$

計算上は 3 個の変数 AGEGR_{50-59} , AGEGR_{60-69} , AGEGR_{70-}

- ▶ **JMP** $\alpha = -1.910$, $\beta_{50-59} = -0.304$, $\beta_{60-69} = 0.577$, $\beta_{70-} = 0.245$

名義ロジスティックのあてはめ CHD

モデル全体の検定

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上側95%
切片	-1.9101963	0.1441374	175.63	<.0001*	-2.2071195	-1.6
AGEGR[70-]	0.2451885	0.3251426	0.57	0.4508	-0.4520055	0.84
AGEGR[60-69]	0.57739059	0.2170333	7.08	0.0078*	0.14586624	1.00
AGEGR[50-59]	-0.3035576	0.220245	1.90	0.1681	-0.7467791	0.12

推定値は次の対数オッズに対するものです:
presence/absence

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
AGEGR	3	3	10.163822	0.001201

単変量ロジスティック回帰分析 AGEGR パラメータ

4 値名義説明変数 AGEGR の場合

(40 - 49 247 名, 50 - 59 203 名, 60 - 69 115 名, 70- 44 名)

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}))}$$

計算上は 3 個の変数 AGEGR₅₀₋₅₉, AGEGR₆₀₋₆₉, AGEGR₇₀₋

- ▶ **JMP** $\alpha = -1.910, \beta_{50-59} = -0.304, \beta_{60-69} = 0.577, \beta_{70-} = 0.245$
- ▶ **R** $\alpha = -2.429, \beta_{50-59} = 0.216, \beta_{60-69} = 1.096, \beta_{70-} = 0.764$

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.6842  -0.4555  -0.4555  -0.4109   2.2422

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.4292     0.2332  -10.415 < 2e-16 ***
AGEGR50-59    0.2155     0.3315   0.650 0.515671
AGEGR60-69    1.0964     0.3272   3.351 0.000806 ***
AGEGR70-     0.7642     0.4736   1.614 0.106604
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}))}$$

40-49, 50-59, 60-69, 70-のそれぞれに対する $P(\text{CHD})$ の手計算

JMP の場合

$$\alpha = -1.910, \beta_{50-59} = -0.304, \beta_{60-69} = 0.577, \beta_{70-} = 0.245$$

- ▶ 40-49, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (-1, -1, -1)$
 $P(\text{CHD} | \text{AGEGR} = 40 - 49) = \frac{1}{1 + \exp(-(-1.91 + 0.304 - 0.577 - 0.245))} = 0.081$
- 50-59, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$
 $P(\text{CHD} | \text{AGEGR} = 50 - 59) = \frac{1}{1 + \exp(-(-1.91 - 0.304))} = 0.0985$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}))}$$

40-49, 50-59, 60-69, 70-のそれぞれに対する $P(\text{CHD})$ の手計算

JMP の場合

$$\alpha = -1.910, \beta_{50-59} = -0.304, \beta_{60-69} = 0.577, \beta_{70-} = 0.245$$

- ▶ 40-49, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (-1, -1, -1)$
 $P(\text{CHD} | \text{AGEGR} = 40 - 49) = \frac{1}{1 + \exp(-(-1.91 + 0.304 - 0.577 - 0.245))} = 0.081$
- 50-59, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$
 $P(\text{CHD} | \text{AGEGR} = 50 - 59) = \frac{1}{1 + \exp(-(-1.91 - 0.304))} = 0.0985$
- ▶ JMP

	$(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-})$	$P(\text{CHD})$
40 - 49	$(-1, -1, -1)$	0.0810
50 - 59	$(1, 0, 0)$	0.0985
60 - 69	$(0, 1, 0)$	0.209
70-	$(0, 0, 1)$	0.159

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}))}$$

40-49, 50-59, 60-69, 70-のそれぞれに対する $P(\text{CHD})$ の手計算

R の場合

$$\alpha = -2.429, \beta_{50-59} = 0.216, \beta_{60-69} = 1.096, \beta_{70-} = 0.764$$

- ▶ 40-49, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (0, 0, 0)$

$$P(\text{CHD} | \text{AGEGR} = 40 - 49) = \frac{1}{1 + \exp(-(-2.429))} = 0.081$$

- 50-59, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$

$$P(\text{CHD} | \text{AGEGR} = 50 - 59) = \frac{1}{1 + \exp(-(-2.429 + 0.216))} = 0.0985$$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(\alpha + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}))}$$

40-49, 50-59, 60-69, 70-のそれぞれに対する $P(\text{CHD})$ の手計算

R の場合

$$\alpha = -2.429, \beta_{50-59} = 0.216, \beta_{60-69} = 1.096, \beta_{70-} = 0.764$$

- ▶ 40-49, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (0, 0, 0)$

$$P(\text{CHD} | \text{AGEGR} = 40 - 49) = \frac{1}{1 + \exp(-(-2.429))} = 0.081$$

- ▶ 50-59, $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$

$$P(\text{CHD} | \text{AGEGR} = 50 - 59) = \frac{1}{1 + \exp(-(-2.429 + 0.216))} = 0.0985$$

- ▶ R

	$(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-})$	$P(\text{CHD})$
40 - 49	(0, 0, 0)	0.0810
50 - 59	(1, 0, 0)	0.0985
60 - 69	(0, 1, 0)	0.209
70-	(0, 0, 1)	0.159

単変量ロジスティック回帰分析 AGEGR 確率 3

手計算の結果は

40 - 49	0.0810
50 - 59	0.0985
60 - 69	0.209
70-	0.159

ソフトによる $P(\text{CHD})$ の出力

▶ JMP の場合

The screenshot shows the JMP software interface with a table of results. The table has 7 columns: P, HPT, 線形[presence], AGEGR, 確率[presence], and 確率[absence]. The rows represent different AGEGR categories and HPT status.

P	HPT	線形[presence]	AGEGR	確率[presence]	確率[absence]
1 38	absence	-2.213753879	50-59	0.0985221675	0.9014778325
2 28	absence	-2.429217742	40-49	0.08097166	0.91902834
3 34	presence	-2.213753879	50-59	0.0985221675	0.9014778325
4 00	presence	-1.332805676	60-69	0.2086956522	0.7913043478
5 15	absence	-2.429217742	40-49	0.08097166	0.91902834
6 12	absence	-2.429217742	40-49	0.08097166	0.91902834
7 28	absence	-2.213753879	50-59	0.0985221675	0.9014778325
8 35	absence	-1.332805676	60-69	0.2086956522	0.7913043478
9 4	absence	-2.429217742	40-49	0.08097166	0.91902834
10 32	presence	-2.213753879	50-59	0.0985221675	0.9014778325
11 36	presence	-1.665007764	70-	0.1590909091	0.8409090909

単変量ロジスティック回帰分析 AGEGR 確率 3

手計算の結果は

40 - 49	0.0810
50 - 59	0.0985
60 - 69	0.209
70 -	0.159

ソフトによる $P(\text{CHD})$ の出力

- ▶ JMP の場合
- ▶ R の場合

```
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
>
> AGEGR[1:11]
[1] 50-59 40-49 50-59 60-69 40-49 40-49 50-59 60-69 40-49 50-59 70-
Levels: 40-49 50-59 60-69 70-
> fitted(ul_AGEGR)[1:11]
      1      2      3      4      5      6
0.09852217 0.08097166 0.09852217 0.20869565 0.08097166 0.08097166
      7      8      9     10     11
0.09852217 0.20869565 0.08097166 0.09852217 0.15909091
>
>
>
>
>
>
>
```

単変量ロジスティック回帰分析 AGEGR 確率 3

手計算の結果は

40 - 49	0.0810
50 - 59	0.0985
60 - 69	0.209
70 -	0.159

ソフトによる $P(\text{CHD})$ の出力

- ▶ JMP の場合
- ▶ R の場合
- ▶ 手計算とソフトの出力は一致した

```
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
>
> AGEGR[1:11]
[1] 50-59 40-49 50-59 60-69 40-49 40-49 50-59 60-69 40-49 50-59 70-
Levels: 40-49 50-59 60-69 70-
> fitted(ul_AGEGR)[1:11]
      1          2          3          4          5          6
0.09852217 0.08097166 0.09852217 0.20869565 0.08097166 0.08097166
      7          8          9         10         11
0.09852217 0.20869565 0.08097166 0.09852217 0.15909091
>
>
>
>
>
>
>
```

単変量ロジスティック回帰分析 AGEGR 検定 1

$$\beta_{50-59} = \beta_{60-69} = \beta_{70-} = 0 \text{ の検定}$$

- ▶ **JMP** Wald 検定 $p = 0.0042$ 尤度比検定 $p = 0.0055$

パラメータ推定値						SBP
項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)		
切片	-1.9101963	0.1441374	175.63	<.0001*	-2.2	138 at
AGEGR[70-]	0.2451885	0.3251426	0.57	0.4508	-0.4	128 at
AGEGR[60-69]	0.57739059	0.2170333	7.08	0.0078*	0.14	164 pr
AGEGR[50-59]	-0.3035576	0.220245	1.90	0.1681	-0.7	200 pr
推定値は次の対数オッズに対するものです: presence/absence						145 at
▷ 推定値の共分散						142 at
効果に対するWald検定						128 at
要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)		
AGEGR	3	3	13.187323	0.0042*		135 at
効果の尤度比検定						114 at
要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)		
AGEGR	3	3	12.6472571	0.0055*		182 pr
						166 pr
						158 at
						142 at
						120 at

単変量ロジスティック回帰分析 AGEGR 検定 1

$\beta_{50-59} = \beta_{60-69} = \beta_{70-} = 0$ の検定

- ▶ **JMP** Wald 検定 $p = 0.0042$ 尤度比検定 $p = 0.0055$
- ▶ **R** Wald 検定 * * * * * 尤度比検定 $p = 0.0055$

```
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ
>
> logistic.display(ul_AGEGR,decimal=4)

Logistic regression predicting CHD : presence vs absence

                OR(95%CI)                P(Wald's test) P(LR-test)
AGEGR: ref.=40-49                0.00546
  50-59                1.2404 (0.6478,2.3753)  0.51567
  60-69                2.9934 (1.5763,5.6843)  < 0.001
  70-                  2.1473 (0.8487,5.4327)  0.1066

Log-likelihood = -212.95552
No. of observations = 609
AIC value = 433.91104
>
```


単変量ロジスティック回帰分析

AGEGR 検定 2

- ▶ 4 値名義説明 **AGEGR** に、3 つの 2 値名義変数 β_{50-59} , β_{60-69} , β_{70-}

単変量ロジスティック回帰分析 AGEGR 検定 2

- ▶ 4 値名義説明 **AGEGR** に、3 つの 2 値名義変数 β_{50-59} , β_{60-69} , β_{70-79}
- ▶ 1 つの変数 **AGEGR** には、
 - 計 3 個の **Wald** 検定の p 値
 - 1 個の尤度比検定の p 値
 - 1 個の **generalized Wald** 検定の p 値

単変量ロジスティック回帰分析 AGEGR 検定 2

- ▶ 4 値名義説明 **AGEGR** に、3 つの 2 値名義変数 β_{50-59} , β_{60-69} , β_{70-}
- ▶ 1 つの変数 **AGEGR** には、計 3 個の **Wald** 検定の p 値
1 個の尤度比検定の p 値
1 個の **generalized Wald** 検定の p 値
- ▶ 3 個の **Wald** 検定の p 値はそれぞれ
「 $\beta_{50-59} = 0$ 」、 「 $\beta_{60-69} = 0$ 」、 「 β_{70-} 」 の検定
尤度比検定の p 値、 **generalized Wald** 検定の p 値は
「 $\beta_{50-59} = \beta_{60-69} = \beta_{70-} = 0$ 」 の検定

単変量ロジスティック回帰分析 AGEGR 検定 2

- ▶ 4 値名義説明 **AGEGR** に、3 つの 2 値名義変数 β_{50-59} , β_{60-69} , β_{70-}
 - ▶ 1 つの変数 **AGEGR** には、計 3 個の **Wald 検定** の p 値
1 個の **尤度比検定** の p 値
1 個の **generalized Wald 検定** の p 値
- ▶ 3 個の **Wald 検定** の p 値はそれぞれ
「 $\beta_{50-59} = 0$ 」、 $\beta_{60-69} = 0$ 、 β_{70-} の検定
尤度比検定の p 値、**generalized Wald 検定** の p 値は
「 $\beta_{50-59} = \beta_{60-69} = \beta_{70-} = 0$ 」の検定

- ▶

	Wald 検定	generalized Wald 検定	尤度比検定
JMP			
R epicalc		×	

単変量ロジスティック回帰分析 AGEGR オッズ比 1

AGEGR のオッズ比と信頼区間の出力
40-49 を reference とした 50-59, 60-69, 70- のオッズ比

▶ JMP の場合

50-59 / 40-49	1.240	0.645-2.386
60-69 / 40-49	2.993	1.579-5.734
70- / 40-49	2.147	0.797-5.233

右表ロジスティックのめではめCHD

▲ オッズ比
CHD: presence対absenceのオッズ比に対して

▲ AGEGRのオッズ比

水準1 /水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
60-69 70-	1.3940345	0.4729	0.5762047	3.7524568
50-59 70-	0.5776737	0.2639	0.2364946	1.5602405
50-59 60-69	0.4143898	0.0073*	0.2156778	0.7880965
40-49 70-	0.4657017	0.1246	0.1910967	1.2554013
40-49 60-69	0.3340675	0.0008*	0.1743962	0.6332902
40-49 50-59	0.8061674	0.5160	0.419154	1.5503367
70- 60-69	0.7173423	0.4729	0.2664921	1.7354942
70- 50-59	1.7310811	0.2639	0.6409268	4.228426
60-69 50-59	2.4131868	0.0073*	1.2688802	4.6365463
70- 40-49	2.1472973	0.1246	0.796558	5.232952
60-69 40-49	2.9934066	0.0008*	1.5790549	5.7340706
50-59 40-49	1.2404372	0.5160	0.6450212	2.3857583

単変量ロジスティック回帰分析 AGEGR オッズ比 2

AGEGR のオッズ比と信頼区間の出力
40-49 を reference とした 50-59, 60-69, 70- のオッズ比

▶ R の場合

50-59 / 40-49	1.240	0.648-2.375
60-69 / 40-49	2.993	1.576-5.684
70- / 40-49	2.147	0.849-5.433

```
>  
> logistic.display(ul_AGEGR, decimal=4)  
  
Logistic regression predicting CHD : presence vs absence  
  
AGEGR: ref.=40-49  
50-59      OR(95%CI)          P(Wald's test) P(LR-test)  
60-69      1.2404 (0.6478,2.3753) 0.51567  
70-        2.9934 (1.5763,5.6843) < 0.001  
           2.1473 (0.8487,5.4327) 0.1066
```

```
Log-likelihood = -212.95552  
No. of observations = 609  
AIC value = 433.91104
```

単変量ロジスティック回帰分析 AGEGR オッズ比 2

AGEGR のオッズ比と信頼区間の出力
40-49 を reference とした 50-59, 60-69, 70- のオッズ比

▶ R の場合

50-59 / 40-49	1.240	0.648-2.375
60-69 / 40-49	2.993	1.576-5.684
70- / 40-49	2.147	0.849-5.433

▶ JMP と R でオッズ比は一致、信頼区間はわずかに違う

```
>
> logistic.display(ul_AGEGR, decimal=4)

Logistic regression predicting CHD : presence vs absence


```

	OR(95%CI)	P(Wald's test)	P(LR-test)
AGEGR: ref.=40-49			0.00546
50-59	1.2404 (0.6478, 2.3753)	0.51567	
60-69	2.9934 (1.5763, 5.6843)	< 0.001	
70-	2.1473 (0.8487, 5.4327)	0.1066	

```
Log-likelihood = -212.95552
No. of observations = 609
AIC value = 433.91104
```

40-49 を reference とした 50-59 のオッズ比の手計算

オッズ $O(\text{CHD}|\text{AGEGR})$

$$= \exp(\alpha + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-})$$

40-49 を reference とした 50-59 のオッズ比の手計算

オッズ $O(\text{CHD}|\text{AGEGR})$

$$= \exp(\alpha + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-})$$

▶ JMP

40-49 のとき

$$(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (-1, -1, -1)$$

50-59 のとき $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \frac{\exp(\alpha) \exp(\beta_{50-59})}{\exp(\alpha) \exp(-\beta_{50-59} - \beta_{60-69} - \beta_{70-})}$$
$$= \exp(2 \times \beta_{50-59} + \beta_{60-69} + \beta_{70-})$$

▶ JMP

$$\beta_{50-59} = -0.304, \beta_{60-69} = 0.577, \beta_{70-} = 0.245$$

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \exp(-2 \times 0.304 + 0.577 + 0.245) = 1.239$$

単変量ロジスティック回帰分析 AGEGR オッズ比 4

40-49 を reference とした 50-59 のオッズ比の手計算

$O(\text{CHD}|\text{AGEGR})$

$$= \exp(\alpha + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-})$$

JMP $\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = 1.239$

40-49 を reference とした 50-59 のオッズ比の手計算

$O(\text{CHD}|\text{AGEGR})$

$$= \exp(\alpha + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-})$$

JMP $\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = 1.239$

▶ R

40-49 のとき $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (0, 0, 0)$

50-59 のとき $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \frac{\exp(\alpha) \exp(\beta_{50-59})}{\exp(\alpha)} = \exp(\beta_{50-59})$$

▶ R

$$\beta_{50-59} = 0.216$$

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \exp(0.216) = 1.241$$

40-49 を reference とした 50-59 のオッズ比の手計算

$O(\text{CHD}|\text{AGEGR})$

$$= \exp(\alpha + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-})$$

$$\text{JMP} \quad \frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = 1.239$$

▶ R

40-49 のとき $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (0, 0, 0)$

50-59 のとき $(\text{AGEGR}_{50-59}, \text{AGEGR}_{60-69}, \text{AGEGR}_{70-}) = (1, 0, 0)$

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \frac{\exp(\alpha) \exp(\beta_{50-59})}{\exp(\alpha)} = \exp(\beta_{50-59})$$

▶ R

$$\beta_{50-59} = 0.216$$

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \exp(0.216) = 1.241$$

▶ オッズ比の手計算とソフトの出力は、丸め誤差を除いて一致

単変量ロジスティック回帰分析 AGEGR オッズ比 5

JMP では 50-59 を reference とした 60-69 のオッズ比なども出力される
60-69 / 50-59 2.413 (1.269 – 4.636)

- ▶ R で、50-59 を reference とするオッズ比の計算

単変量ロジスティック回帰分析 **AGEGR** オッズ比 5

JMP では **50-59** を **reference** とした **60-69** のオッズ比なども出力される
60-69 / 50-59 2.413 (1.269 – 4.636)

- ▶ **R** で、**50-59** を **reference** とするオッズ比の計算
- ▶ **50-59** を **reference** とする変数 **AGE_50GR** の作成

50-59 が最初にくる **levels** の順番を指定

```
AGE_50GR<-factor(AGEGR,  
                 levels=c("50-59", "40-49", "60-69", "70-"))
```

単変量ロジスティック回帰分析 **AGEGR** オッズ比 5

JMP では **50-59** を **reference** とした **60-69** のオッズ比なども出力される
60-69 / 50-59 2.413 (1.269 – 4.636)

- ▶ **R** で、**50-59** を **reference** とするオッズ比の計算
- ▶ **50-59** を **reference** とする変数 **AGE_50GR** の作成
50-59 が最初にくる **levels** の順番を指定
`AGE_50GR<-factor(AGEGR,
 levels=c("50-59", "40-49", "60-69", "70-"))`
- ▶ **AGE_50GR** を説明変数とした分析とオッズ比の出力
`ul_AGE_50GR<-glm(CHD~AGE_50GR, family=binomial)
logistic.display(ul_AGE_50GR, decimal=4)`

単変量ロジスティック回帰分析 AGEGR オッズ比 5

JMP では 50-59 を reference とした 60-69 のオッズ比なども出力される
60-69 / 50-59 2.413 (1.269 – 4.636)

- ▶ R で、50-59 を reference とするオッズ比の計算
- ▶ 50-59 を reference とする変数 AGE_50GR の作成

50-59 が最初にくる levels の順番を指定

```
AGE_50GR<-factor(AGEGR,  
                 levels=c("50-59","40-49","60-69","70-"))
```

- ▶ AGE_50GR を説明変数とした分析とオッズ比の出力
ul_AGE_50GR<-glm(CHD~AGE_50GR,family=binomial)
logistic.display(ul_AGE_50GR,decimal=4)
- ▶ 60-69 / 50-59 2.413 (1.267 – 4.597)

```
> logistic.display(ul_AGE_50GR,decimal=4)
```

```
Logistic regression predicting CHD : presence vs absence
```

	OR(95%CI)	P(Wald's test)	P(LR-test)
AGE_50GR: ref.=50-59			0.00546
40-49	0.8062 (0.421, 1.5437)	0.51567	
60-69	2.4132 (1.2668, 4.5971)	0.00738	
70-	1.7311 (0.6827, 4.3893)	0.2477	

多変量ロジスティック回帰分析

説明変数 **CAT**、**AGEGR**、**CHL**、**SMK**、**ECG**、**DBP**、**SBP**、**HPT** はすべて $P(\text{CHD})$ に影響があることが分かった

- ▶ つぎに複数の説明変数による $P(\text{CHD})$ の回帰式を作る。
少ない変数による効率の良い式を求める

多変量ロジスティック回帰分析

説明変数 **CAT**、**AGEGR**、**CHL**、**SMK**、**ECG**、**DBP**、**SBP**、**HPT** はすべて $P(\text{CHD})$ に影響があることが分かった

- ▶ つぎに複数の説明変数による $P(\text{CHD})$ の回帰式を作る。
少ない変数による効率の良い式を求める
- ▶ 説明変数の選択には種々の方法がある。
ここではステップワイズ法の変数増加法で、
最終的に残す説明変数の選択をする

多変量ロジスティック回帰分析

説明変数 **CAT**、**AGEGR**、**CHL**、**SMK**、**ECG**、**DBP**、**SBP**、**HPT** はすべて $P(\text{CHD})$ に影響があることが分かった

- ▶ つぎに複数の説明変数による $P(\text{CHD})$ の回帰式を作る。
少ない変数による効率の良い式を求める
- ▶ 説明変数の選択には種々の方法がある。
ここではステップワイズ法の変数増加法で、
最終的に残す説明変数の選択をする
- ▶ 選ばれた説明変数を用いた回帰式を、最終的なモデル式とする

多変量ロジスティック回帰分析

説明変数 CAT、AGEGR、CHL、SMK、ECG、DBP、SBP、HPT はすべて $P(\text{CHD})$ に影響があることが分かった

- ▶ つぎに複数の説明変数による $P(\text{CHD})$ の回帰式を作る。
少ない変数による効率の良い式を求める
- ▶ 説明変数の選択には種々の方法がある。
ここではステップワイズ法の変数増加法で、
最終的に残す説明変数の選択をする
- ▶ 選ばれた説明変数を用いた回帰式を、最終的なモデル式とする
- ▶ 多変量ロジスティック回帰分析の目標 ([Kleinbaum, Klein, p.169])
 - 目標 1. 1組の変数による $P(\text{CHD})$ の説明
**to obtain a valid estimate of
an exposure-disease relationship**
 - 目標 2. $P(\text{CHD})$ のモデル式による CHD の鑑別
to obtain a good predictive model

ステップワイズ・ロジスティック回帰分析 JMP の場合

- ▶ JMP の分析メニューから「モデルのあてはめ」を選ぶ

The screenshot shows the JMP software interface. The 'Analyze' menu is open, and 'Fit Model' is selected. The main data table is visible in the background.

		CAT	AGEG				DBP	SBP
1	e	normal	50-59			absence	80	138
2	e	normal	40-49			absence	74	128
3	e	high	50-59			absence	112	164
4	e	high	60-69	179	1	absence	100	200
5	e	normal	40-49	243	1	absence	82	145
6	e	normal	40-49	252	1	absence	88	142
7	e	normal	50-59	179	1	presence	80	128
8	e	normal	60-69	217	0	absence	92	135
9	e	normal	40-49	176	1	absence	76	114
10	e	normal	50-59	250	0	presence	114	182

ステップワイズ・ロジスティック回帰分析 JMP の場合

- ▶ JMP の分析メニューから「モデルのあてはめ」を選ぶ
- ▶ CHD を Y に、すべての説明変数を「モデル効果の構成」に入れて「手法」を「ステップワイズ法」にして「実行」を押す

The screenshot shows the 'Stepwise Logistic Regression' dialog box in JMP. The 'Method' is set to 'Stepwise Method'. The response variable 'Y' is 'CHD'. The explanatory variables are 'CAT', 'AGEGR', 'CHL', 'SMK', 'ECG', 'DBP', 'SBP', and 'HPT'. The 'Model Effect Structure' is set to 'Add', and the 'Order' is 2.

列の選択

- CHD
- CAT
- AGEGR
- CHL
- SMK
- ECG
- DBP
- SBP
- HPT

(説明変数の)選択

Y: CHD
オプション

重み: オプション(数値)

度数: オプション(数値)

By: オプション

手法: ステップワイズ法

ヘルプ 実行

前回の設定 ダイアログを開いたままにする

削除

モデル効果の構成

追加 CAT

交差 AGEGR

枝分かれ CHL

マクロ SMK

ECG

DBP

次数 2 SBP

HPT

ステップワイズ・ロジスティック回帰分析 JMP の場合

- ▶ JMP の分析メニューから「モデルのあてはめ」を選ぶ
- ▶ CHD を Y に、すべての説明変数を「モデル効果の構成」に入れて「手法」を「ステップワイズ法」にして「実行」を押す
- ▶ 開いた窓で「停止ルール」を「最小 AICc」に、
(修正赤池情報量基準 **Corrected Akaike Information Criterion**)
「方向」を「変数増加」にして、「実行」を押す

The screenshot shows the 'CHDのステップワイズ' dialog box in JMP. The 'ステップワイズ回帰の設定' section is active, showing '停止ルール' set to '最小AICc', '方向' set to '変数増加', and 'ルール' set to '組み合わせ'. Below the dialog, a table shows the results of the stepwise regression process.

(-1)*対数尤度	p	R2乗	AICc	BIC
219.27915	11	0.0000	440.565	444.97

Below the table, the '現在の推定値' section is visible, showing a table of parameter estimates:

ロック追加	パラメータ	推定値	自由度	Wald/スコア	カイ2乗	“有意確率”
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> 切片[absence]	-2.0251787	1	0		1
<input type="checkbox"/>	<input type="checkbox"/> CAT{high-normal}	0	1	16.24646		5.56e-5
<input type="checkbox"/>	<input type="checkbox"/> AGEGR{60-69&70--50-59&40-49}	0	1	12.83662		0.00034

ステップワイズ・ロジスティック回帰分析 JMP の場合

- ▶ JMP の分析メニューから「モデルのあてはめ」を選ぶ
- ▶ CHD を Y に、すべての説明変数を「モデル効果の構成」に入れて「手法」を「ステップワイズ法」にして「実行」を押す
- ▶ 開いた窓で「停止ルール」を「最小 AICc」に、
(修正赤池情報量基準 **Corrected Akaike Information Criterion**)
「方向」を「変数増加」にして、「実行」を押す
- ▶ CAT、AGEGR、CHL、SMK、HPT が選択された

現在の推定値

ロック追加パラメータ		推定値	自由度	Wald/ スコアカイ2乗	“有意確率”
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> 切片[absence]	-3.7631598	1	0	1
<input type="checkbox"/>	<input checked="" type="checkbox"/> CAT{high-normal}	0.37158089	1	3.631895	0.05668
<input type="checkbox"/>	<input checked="" type="checkbox"/> AGEGR{60-69&70--50-59&40-49}	0.37554059	3	8.423385	0.03803
<input type="checkbox"/>	<input type="checkbox"/> AGEGR{60-69-70-}	0	1	0.711776	0.39885
<input type="checkbox"/>	<input type="checkbox"/> AGEGR{50-59-40-49}	0	1	0.224261	0.63581
<input type="checkbox"/>	<input checked="" type="checkbox"/> CHL	0.00866602	1	6.854316	0.00884
<input type="checkbox"/>	<input checked="" type="checkbox"/> SMK{1-0}	0.41864792	1	8.76476	0.00307
<input type="checkbox"/>	<input type="checkbox"/> ECG{presence-absence}	0	1	1.845516	0.17431
<input type="checkbox"/>	<input type="checkbox"/> DBP	0	1	0.502159	0.47855
<input type="checkbox"/>	<input type="checkbox"/> SBP	0	1	1.111783	0.29169
<input type="checkbox"/>	<input checked="" type="checkbox"/> HPT{presence-absence}	0.24524283	1	2.126048	0.14481

ステップ履歴

ステップ	パラメータ	アクション	尤度比カイ2乗	“有意確率”	R2乗
1	CAT{high-normal}	追加	14.13121	0.0002	0.0322
2	CHL	追加	8.919006	0.0028	0.0526

ステップワイズ・ロジスティック回帰分析 R の場合

- ▶ **epicalc Package** のロード、データの読み込み、名義変数の指示

```
R C:\R-2.13.1\evans.r - R1データ
ファイル 編集 パッケージ ヘルプ
library(epicalc)

dfr <- read.table("evans.txt",header=T)

dfr <- within(dfr,{
  CHD<-factor(CHD,labels=c("absence","presence"))
  CAT<-factor(CAT,labels=c("normal","high"))
  AGEGR<-factor(AGEGR,labels=c("40-49","50-59","60-69","70-"))
  SMK<-factor(SMK)
  ECG<-factor(ECG,labels=c("absence","presence"))
  HPT<-factor(HPT,labels=c("absence","presence")) })
attach(dfr)
```

ステップワイズ・ロジスティック回帰分析 R の場合

- ▶ **epicalc Package** のロード、データの読み込み、名義変数の指示
- ▶ つぎのスクリプトを実行

```
ml_all<-glm(CHD~CAT+AGEGR+SMK+ECG+HPT+CHL+DBP+SBP,  
            family=binomial,data=dfr)  
ml_null<-glm(CHD~1,family=binomial,data=dfr)  
ml<-step(ml_null,scope=formula(ml_all),direction="forward")  
summary(ml)
```

ステップワイズ・ロジスティック回帰分析 R の場合

- ▶ **epicalc Package** のロード、データの読み込み、名義変数の指示
- ▶ つぎのスクリプトを実行

```
ml_all<-glm(CHD~CAT+AGEGR+SMK+ECG+HPT+CHL+DBP+SBP,  
            family=binomial,data=dfr)  
ml_null<-glm(CHD~1,family=binomial,data=dfr)  
ml<-step(ml_null,scope=formula(ml_all),direction="forward")  
summary(ml)
```

- ▶ **CAT、AGEGR、CHL、SMK、HPT** が選択された

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.0635	-0.5255	-0.4034	-0.3055	2.5577

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-5.214315	0.803389	-6.490	8.56e-11	***
CAT	0.761161	0.336069	2.265	0.02352	*
CHL	0.008669	0.003255	2.663	0.00774	**
SMK	0.804696	0.303270	2.653	0.00797	**
AGEGR	0.291635	0.145592	2.003	0.04517	*
HPT	0.488420	0.289337	1.688	0.09140	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

ステップワイズ・ロジスティック回帰分析 JMP と R

- ▶ **JMP** 「最小 **AICc**」以外に「最小 **BIC**」、「閾値 **p** 値」
R の **step** 関数 「最小 **AIC**」以外に「最小 **BIC**」

ステップワイズ・ロジスティック回帰分析 JMP と R

- ▶ **JMP** 「最小 **AICc**」以外に「最小 **BIC**」、「閾値 **p** 値」
R の **step** 関数 「最小 **AIC**」以外に「最小 **BIC**」
- ▶ **JMP** の「最小 **AICc**」と、R の **step** 関数の「最小 **AIC**」

変数増加法

step	R		JMP	
	enter	AIC	enter	AICc
1	CAT	428.4	CAT	428.4
2	CHL	421.5	CHL	421.5
3	SMK	417.4	SMK	417.4
4	AGEGR	415.5	AGEGR	413.0
5	HPT	414.6	HPT	412.2
6	これ以降、AIC、AICc は増加する			

ステップワイズ・ロジスティック回帰分析 JMP と R

- ▶ **JMP** 「最小 **AICc**」以外に「最小 **BIC**」、「閾値 **p** 値」
R の **step** 関数 「最小 **AIC**」以外に「最小 **BIC**」
- ▶ **JMP** の「最小 **AICc**」と、R の **step** 関数の「最小 **AIC**」

変数増加法

step	R		JMP	
	enter	AIC	enter	AICc
1	CAT	428.4	CAT	428.4
2	CHL	421.5	CHL	421.5
3	SMK	417.4	SMK	417.4
4	AGEGR	415.5	AGEGR	413.0
5	HPT	414.6	HPT	412.2
6	これ以降、AIC、AICc は増加する			

- ▶ **AIC**, **AICc** は値がわずかに異なるが
選択される変数には違いがなかった

ステップワイズ法の変数増加法で選ばれた
説明変数 **CAT**、**AGEGR**、**CHL**、**SMK**、**HPT** の回帰式を扱う

▶ 変数

CAT, **CHL**, **SMK**, **HPT**, **AGEGR**₅₀₋₅₉, **AGEGR**₆₀₋₆₉, **AGEGR**₇₀₋

ステップワイズ法の変数増加法で選ばれた
説明変数 **CAT**、**AGEGR**、**CHL**、**SMK**、**HPT** の回帰式を扱う

- ▶ 変数

CAT, **CHL**, **SMK**, **HPT**, **AGEGR**₅₀₋₅₉, **AGEGR**₆₀₋₆₉, **AGEGR**₇₀₋

- ▶ パラメータ

切片 α

係数 β_{high} , β_{CHL} , β_1 , β_{presence} , β_{50-59} , β_{60-69} , β_{70-}

ステップワイズ法の変数増加法で選ばれた
説明変数 **CAT**、**AGEGR**、**CHL**、**SMK**、**HPT** の回帰式を扱う

▶ 変数

CAT, **CHL**, **SMK**, **HPT**, **AGEGR**₅₀₋₅₉, **AGEGR**₆₀₋₆₉, **AGEGR**₇₀₋

▶ パラメータ

切片 α

係数 β_{high} , β_{CHL} , β_1 , β_{presence} , β_{50-59} , β_{60-69} , β_{70-}

▶ 回帰式 (**CHD** 有病率 $P(\text{CHD})$)

$h = \alpha + \beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL} + \beta_1\text{SMK} + \beta_{\text{presence}}\text{HPT} + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-}$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}$$

多変量ロジスティック回帰分析

JMP 計算

- ▶ JMP の分析メニューから「モデルのあてはめ」を選ぶ

The screenshot shows the JMP software interface. The 'Analysis' menu is open, and 'Fit Model' is selected. The data table is visible, showing columns for CAT, AGEGR, and DBP/SBP. The 'Fit Model' dialog box is also visible, showing the model structure.

	CAT	AGEGR		DBP	SBP
1	normal	50-59	presence	80	138
2	normal	40-49	presence	74	128
3	high	50-59	presence	112	164
4	high	60-69	absence	100	200
5	normal	40-49	absence	82	145
6	normal	40-49	absence	88	142
7	normal	50-59	presence	80	128
8	normal	60-69	absence	92	135
9	normal	40-49	absence	76	114
10	normal	50-59	presence	114	182

多変量ロジスティック回帰分析

JMP 計算

- ▶ JMP の分析メニューから「モデルのあてはめ」を選ぶ
- ▶ CHD を Y に、すべての説明変数を「モデル効果の構成」に入れて「手法」を「名義ロジスティック」にして「実行」を押す

▼ モデルの指定

列の選択

- CHD
- CAT
- AGEGR
- CHL
- SMK
- ECG
- DBP
- SBP
- HPT

役割変数の選択

Y: CHD
オプション

重み: オプション(数値)

度数: オプション(数値)

By: オプション

手法: 名義ロジスティック

ヘルプ 実行

前回の設定 ダイアログを開いたままにする

削除

モデル効果の構成

追加: CAT, AGEGR, CHL, SMK, HPT

交差

枝分かれ

- ▶ 単変量と同様に、多変量ロジスティック回帰分析も `glm` 関数に **binomial** オプションを付けて行う
出力を `m1_5_vars` と名付ける

```
m1_5_vars<-glm(CHD~CAT+AGEGR+SMK+HPT+CHL, family=binomial)
```

- ▶ 単変量と同様に、多変量ロジスティック回帰分析も `glm` 関数に **binomial** オプションを付けて行う
出力を `ml_5_vars` と名付ける

```
ml_5_vars<-glm(CHD~CAT+AGEGR+SMK+HPT+CHL,family=binomial)
```

- ▶ 回帰式パラメータ、**Wald** 検定、尤度比検定の出力
`summary(ml_5_vars)`

- ▶ 単変量と同様に、多変量ロジスティック回帰分析も **glm** 関数に **binomial** オプションを付けて行う
出力を `ml_5_vars` と名付ける

```
ml_5_vars<-glm(CHD~CAT+AGEGR+SMK+HPT+CHL, family=binomial)
```

- ▶ 回帰式パラメータ、**Wald** 検定、尤度比検定の出力

```
summary(ml_5_vars)
```

- ▶ オッズ比とその信頼区間の出力

```
logistic.display(ml_5_vars, decimal=4)
```

多変量ロジスティック回帰分析 パラメータ 1

▶ JMP の回帰式の切片と係数

名義ロジスティックのあてはめ CHD							
あてはまりの悪さ(LOF)							
パラメータ推定値							
項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上側	
切片	-3.8316527	0.7287841	27.64	<.0001*	-5.2859024	-2.4215	
CAT[high]	0.35567098	0.1705126	4.35	0.0370*	0.0204609	0.690	
AGEGR[70-]	0.06984564	0.3445566	0.04	0.8394	-0.6614159	0.7058	
AGEGR[60-69]	0.56995236	0.2328616	5.99	0.0144*	0.10799289	1.0249	
AGEGR[50-59]	-0.2602029	0.2276036	1.31	0.2529	-0.7171865	0.180	
CHL	0.00865341	0.0032728	6.99	0.0082*	0.00222038	0.0150	
SMK[1]	0.4316216	0.1539592	7.86	0.0051*	0.14015025	0.7467	
HPT[presence]	0.26137501	0.1452361	3.24	0.0719	-0.0233104	0.5480	

推定値は次の対数オッズに対するものです：
presence/absence

▶ 推定値の共分散

多変量ロジスティック回帰分析

パラメータ 1

- ▶ JMP の回帰式の切片と係数
- ▶ R の回帰式の切片と係数

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.1968	-0.5256	-0.3936	-0.2908	2.6398

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-5.259916	0.821852	-6.400	1.55e-10	***
CAThigh	0.711342	0.341023	2.086	0.03699	*
AGEGR50-59	0.119392	0.352330	0.339	0.73471	
AGEGR60-69	0.949548	0.373202	2.544	0.01095	*
AGEGR70-	0.449441	0.518080	0.868	0.38566	
SMK1	0.863243	0.307914	2.804	0.00505	**
HPTpresence	0.522750	0.290470	1.800	0.07191	.
CHL	0.008653	0.003273	2.644	0.00819	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- ▶ 回帰式 (CHD 有病率 $P(\text{CHD})$)

$$h = \alpha + \beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL} + \beta_1\text{SMK} + \beta_{\text{presence}}\text{HPT} + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-}$$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}$$

- ▶ 回帰式 (CHD 有病率 $P(\text{CHD})$)

$$h = \alpha + \beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL} + \beta_1\text{SMK} + \beta_{\text{presence}}\text{HPT} + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-}$$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}$$

変数	JMP 係数	R 係数	JMP p 値	R p 値
α	-3.831	-5.260	< 0.0001	< 0.0001
β_{high}	0.356	0.711	0.0370	0.0370
β_{50-59}	-0.260	0.119	0.253	0.735
▶ β_{60-69}	0.570	0.950	0.0144	0.011
β_{70-}	0.070	0.449	0.839	0.386
β_{CHL}	0.00865	0.00865	0.0082	0.0082
β_1	0.432	0.863	0.0051	0.0051
β_{presence}	0.261	0.523	0.0719	0.0719

p 値は Wald 検定の p 値

多変量ロジスティック回帰分析 確率 1

CHD の有病率 $P(\text{CHD})$ の手計算

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}, \text{ここに } h \text{ は}$$

$$h = \alpha + \beta_{\text{high}} \text{CAT} + \beta_{\text{CHL}} \text{CHL} + \beta_1 \text{SMK} + \beta_{\text{presence}} \text{HPT} + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}$$

多変量ロジスティック回帰分析 確率 1

CHD の有病率 $P(\text{CHD})$ の手計算

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}, \text{ここに } h \text{ は}$$

$$h = \alpha + \beta_{\text{high}} \text{CAT} + \beta_{\text{CHL}} \text{CHL} + \beta_1 \text{SMK} + \beta_{\text{presence}} \text{HPT} + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}$$

▶ No. 2 のデータ

	CAT	AGEGR	CHL	SMK	HPT
	normal	40 - 49	159	1	absence
JMP	-1	(-1, -1, -1)	159	1	-1
R	0	(0, 0, 0)	159	1	0

多変量ロジスティック回帰分析 確率 1

CHD の有病率 $P(\text{CHD})$ の手計算

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}, \text{ここに } h \text{ は}$$

$$h = \alpha + \beta_{\text{high}} \text{CAT} + \beta_{\text{CHL}} \text{CHL} + \beta_1 \text{SMK} + \beta_{\text{presence}} \text{HPT} + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}$$

▶ No. 2 のデータ

	CAT	AGEGR	CHL	SMK	HPT
	normal	40 - 49	159	1	absence
JMP	-1	(-1, -1, -1)	159	1	-1
R	0	(0, 0, 0)	159	1	0

- ▶ JMP $\alpha = -3.831, \beta_{\text{high}} = 0.356, \beta_{50-59} = -0.260, \beta_{60-69} = 0.570, \beta_{70-} = 0.070, \beta_{\text{CHL}} = 0.00865, \beta_1 = 0.432, \beta_{\text{presence}} = 0.261$
 $h = -3.831 - 0.356 - (-0.260 + 0.570 + 0.070) + 159 * 0.00865 + 0.432 - 0.261 = -3.021$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(-3.021))} = 0.0465$$

多変量ロジスティック回帰分析 確率 1

CHD の有病率 $P(\text{CHD})$ の手計算

$$P(\text{CHD}) = \frac{1}{1 + \exp(-h)}, \text{ここに } h \text{ は}$$

$$h = \alpha + \beta_{\text{high}} \text{CAT} + \beta_{\text{CHL}} \text{CHL} + \beta_1 \text{SMK} + \beta_{\text{presence}} \text{HPT} + \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}$$

▶ No. 2 のデータ

	CAT	AGEGR	CHL	SMK	HPT
	normal	40 - 49	159	1	absence
JMP	-1	(-1, -1, -1)	159	1	-1
R	0	(0, 0, 0)	159	1	0

- ▶ JMP $\alpha = -3.831, \beta_{\text{high}} = 0.356, \beta_{50-59} = -0.260, \beta_{60-69} = 0.570, \beta_{70-} = 0.070, \beta_{\text{CHL}} = 0.00865, \beta_1 = 0.432, \beta_{\text{presence}} = 0.261$
 $h = -3.831 - 0.356 - (-0.260 + 0.570 + 0.070) + 159 * 0.00865 + 0.432 - 0.261 = -3.021$

$$P(\text{CHD}) = \frac{1}{1 + \exp(-(-3.021))} = 0.0465$$

- ▶ R $\alpha = -5.260, \beta_{\text{CHL}} = 0.00865, \beta_1 = 0.863,$
 $h = -5.260 + 159 * 0.00865 + 0.863 = -3.022$

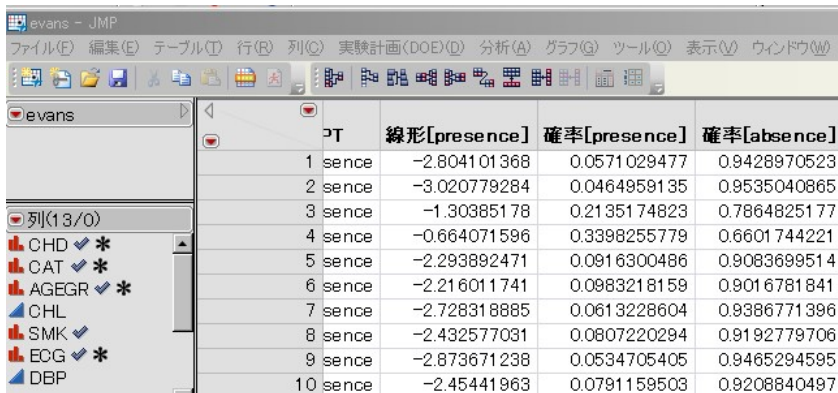
$$P(\text{CHD}) = \frac{1}{1 + \exp(-(3.022))} = 0.0464$$

多変量ロジスティック回帰分析 確率 2

ソフトによる $P(\text{CHD})$ の出力

No.2 の人は、**JMP**、**R** とも手計算では $P(\text{CHD}) = 0.0465$

▶ JMP の場合



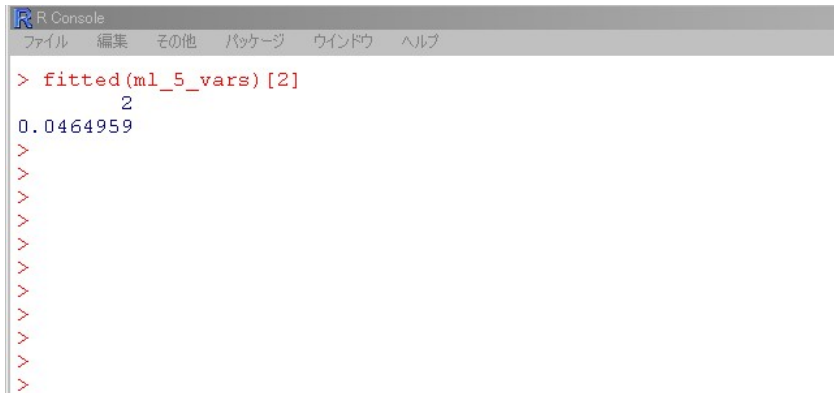
PT	線形[presence]	確率[presence]	確率[absence]
1 sence	-2.804101368	0.0571029477	0.9428970523
2 sence	-3.020779284	0.0464959135	0.9535040865
3 sence	-1.30385178	0.2135174823	0.7864825177
4 sence	-0.664071596	0.3398255779	0.6601744221
5 sence	-2.293892471	0.0916300486	0.9083699514
6 sence	-2.216011741	0.0983218159	0.9016781841
7 sence	-2.728318885	0.0613228604	0.9386771396
8 sence	-2.432577031	0.0807220294	0.9192779706
9 sence	-2.873671238	0.0534705405	0.9465294595
10 sence	-2.45441963	0.0791159503	0.9208840497

多変量ロジスティック回帰分析 確率 2

ソフトによる $P(\text{CHD})$ の出力

No.2 の人は、**JMP**、**R** とも手計算では $P(\text{CHD}) = 0.0465$

- ▶ **JMP** の場合
- ▶ **R** の場合



```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> fitted(ml_5_vars)[2]
      2
0.0464959
>
>
>
>
>
>
>
>
>
>
>
```


- ▶ JMP の Wald 検定の p 値、尤度比検定の p 値

△ 効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
CAT	1	1	4.35094789	0.0370*
AGEGR	3	3	8.03244383	0.0453*
CHL	1	1	6.99104116	0.0082*
SMK	1	1	7.85950202	0.0051*
HPT	1	1	3.23876309	0.0719

△ 効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
CAT	1	1	4.32149817	0.0376*
AGEGR	3	3	7.6659267	0.0534
CHL	1	1	6.91909103	0.0085*
SMK	1	1	8.66900577	0.0032*
HPT	1	1	3.23989481	0.0719

△ オッズ比

CHD: presence対absenceのオッズ比に対して

多変量ロジスティック回帰分析 検定 1

- ▶ JMP の Wald 検定の p 値、尤度比検定の p 値
- ▶ R の Wald 検定の p 値、尤度比検定の p 値

```

                                P(Wald's test) P(LR-test)
CAT: high vs normal            0.03699      0.03763
AGEGR: ref.=40-49
  50-59                        0.73471
  60-69                        0.01095
  70-                          0.38566
SMK: 1 vs 0                    0.00505      0.00324
HPT: presence vs absence      0.07191      0.07187
CHL (cont. var.)              0.00819      0.00853

Log-likelihood = -199.424602
No. of observations = 609
AIC = 414.849202
```

	Wald 検定の p 値		尤度比検定の p 値	
	JMP	R	JMP	R
CAT	0.0370	0.0370	0.0376	0.0376
AGEGR	0.0453	***	0.0534	0.0534
SMK	0.0051	0.0051	0.0032	0.0032
HPT	0.0719	0.0719	0.0719	0.0719
CHL	0.0082	0.0082	0.0085	0.0085

- ▶ AGEGR の Wald 検定は generalized Wald 検定
R の `epicalc` はこれを出力しない

	Wald 検定の p 値		尤度比検定の p 値	
	JMP	R	JMP	R
CAT	0.0370	0.0370	0.0376	0.0376
AGEGR	0.0453	***	0.0534	0.0534
SMK	0.0051	0.0051	0.0032	0.0032
HPT	0.0719	0.0719	0.0719	0.0719
CHL	0.0082	0.0082	0.0085	0.0085

- ▶ AGEGR の Wald 検定は generalized Wald 検定
R の `epicalc` はこれを出力しない
- ▶ AIC, AICc で変数選択すると必ずしも p 値 < 0.05 に揃わない
 p 値を基準にした変数選択をすると p 値 < 0.05 に揃えられる

	Wald 検定の p 値		尤度比検定の p 値	
	JMP	R	JMP	R
CAT	0.0370	0.0370	0.0376	0.0376
AGEGR	0.0453	***	0.0534	0.0534
SMK	0.0051	0.0051	0.0032	0.0032
HPT	0.0719	0.0719	0.0719	0.0719
CHL	0.0082	0.0082	0.0085	0.0085

- ▶ AGEGR の Wald 検定は generalized Wald 検定
R の `epicalc` はこれを出力しない
- ▶ AIC, AICc で変数選択すると必ずしも p 値 < 0.05 に揃わない
 p 値を基準にした変数選択をすると p 値 < 0.05 に揃えられる
- ▶ JMP では p 値を基準にした変数選択をすることができる
R には p 値を基準にした変数選択をする Package が無い

▶ JMP CHL, CAT

オッズ比

CHD: presence 対 absence のオッズ比に対して

△ 単位オッズ比

連続変数が1単位だけ変化した場合

項	オッズ比	下側95%	上側95%	逆数
CHL	1.008691	1.002223	1.015205	0.9913839

△ 範囲オッズ比

連続変数が範囲全体で変化した場合

項	オッズ比	下側95%	上側95%	逆数
CHL	9.736172	1.793127	52.91983	0.1027098

△ CATのオッズ比

水準1 / 水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
normal high	0.4909849	0.0376*	0.2511941	0.9599042
high normal	2.0367226	0.0376*	1.0417706	3.9809847

多変量ロジスティック回帰分析 オッズ比 1

オッズ比の出力

- ▶ JMP CHL, CAT
- ▶ JMP AGEGR

△ CATのオッズ比

△ AGEGRのオッズ比

水準1	/水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
60-69	70-	1.6488972	0.2989	0.6524357	4.6178579
50-59	70-	0.7188888	0.5155	0.2802503	2.0287157
50-59	60-69	0.4359816	0.0165*	0.2192456	0.8589937
40-49	70-	0.6379849	0.3958	0.2396159	1.8655758
40-49	60-69	0.3869161	0.0113*	0.1852372	0.8051051
40-49	50-59	0.8874598	0.7349	0.4439233	1.7809829
70-	60-69	0.6064659	0.2989	0.2165506	1.532718
70-	50-59	1.3910357	0.5155	0.4929227	3.5682393
60-69	50-59	2.2936749	0.0165*	1.1641529	4.5610952
70-	40-49	1.5674353	0.3958	0.5360275	4.173346
60-69	40-49	2.5845397	0.0113*	1.2420738	5.398484
50-59	40-49	1.1268117	0.7349	0.5614877	2.2526416

△ SMKのオッズ比

多変量ロジスティック回帰分析

オッズ比 1

オッズ比の出力

- ▶ JMP CHL, CAT
- ▶ JMP AGEGR
- ▶ JMP SMK, HPT

△ AGEGRのオッズ比

水準1 / 水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
70- 60-69	0.6064659	0.2989	0.2165506	1.532718
70- 50-59	1.3910357	0.5155	0.4929227	3.5682393
60-69 50-59	2.2936749	0.0165*	1.1641529	4.5610952
70- 40-49	1.5674353	0.3958	0.5360275	4.173346
60-69 40-49	2.5845397	0.0113*	1.2420738	5.398484
50-59 40-49	1.1268117	0.7349	0.5614877	2.2526416

△ SMKのオッズ比

水準1 / 水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
0 1	0.4217919	0.0032*	0.2246006	0.7555567
1 0	2.3708373	0.0032*	1.3235275	4.4523476

△ HPTのオッズ比

水準1 / 水準2	オッズ比	p値(Prob>Chisq)	下側95%	上側95%
absence presence	0.5928878	0.0719	0.3341484	1.0477247
presence absence	1.6866596	0.0719	0.9544492	2.9926827

多変量ロジスティック回帰分析 オッズ比 1

オッズ比の出力

- ▶ JMP CHL, CAT
- ▶ JMP AGEGR
- ▶ JMP SMK, HPT
- ▶ R crude OR は単変量のオッズ比。adjusted OR の方を見る

Logistic regression predicting CHD : presence vs absence

	crude OR(95%CI)	adj. OR(95%CI)
CAT: high vs normal	2.8615 (1.6878, 4.8513)	2.0367 (1.0439, 3.9738)
AGEGR: ref.=40-49		
50-59	1.2404 (0.6478, 2.3753)	1.1268 (0.5649, 2.2478)
60-69	2.9934 (1.5763, 5.6843)	2.5845 (1.2437, 5.371)
70-	2.1473 (0.8487, 5.4327)	1.5674 (0.5678, 4.3269)
SMK: 1 vs 0	1.9555 (1.1035, 3.4653)	2.3708 (1.2966, 4.3351)
HPT: presence vs absence	2.3615 (1.4232, 3.9185)	1.6867 (0.9545, 2.9804)
CHL (cont. var.)	1.007 (1.001, 1.0131)	1.0087 (1.0022, 1.0152)

オッズ比の出力



変数	JMP		R	
	オッズ比	信頼区間	オッズ比	信頼区間
CAT	2.037	1.042 – 3.981	2.037	1.044 – 3.974
AGEGR₅₀₋₅₉	1.127	0.561 – 2.253	1.127	0.565 – 2.248
AGEGR₆₀₋₆₉	2.585	1.242 – 5.398	2.585	1.244 – 5.371
AGEGR₇₀₋	1.567	0.536 – 4.173	1.567	0.568 – 4.327
CHL	1.0087	1.002 – 0.0152	1.0087	1.002 – 1.0152
SMK	2.371	1.324 – 4.452	2.371	1.297 – 4.335
HPT	1.687	0.954 – 2.993	1.687	0.954 – 2.980

オッズ比の出力



変数	JMP		R	
	オッズ比	信頼区間	オッズ比	信頼区間
CAT	2.037	1.042 – 3.981	2.037	1.044 – 3.974
AGEGR₅₀₋₅₉	1.127	0.561 – 2.253	1.127	0.565 – 2.248
AGEGR₆₀₋₆₉	2.585	1.242 – 5.398	2.585	1.244 – 5.371
AGEGR₇₀₋	1.567	0.536 – 4.173	1.567	0.568 – 4.327
CHL	1.0087	1.002 – 0.0152	1.0087	1.002 – 1.0152
SMK	2.371	1.324 – 4.452	2.371	1.297 – 4.335
HPT	1.687	0.954 – 2.993	1.687	0.954 – 2.980

- ▶ オッズ比は **JMP** と **R** で、**AGEGR** も込めてすべての変数で一致

オッズ比の出力



変数	JMP		R	
	オッズ比	信頼区間	オッズ比	信頼区間
CAT	2.037	1.042 – 3.981	2.037	1.044 – 3.974
AGEGR ₅₀₋₅₉	1.127	0.561 – 2.253	1.127	0.565 – 2.248
AGEGR ₆₀₋₆₉	2.585	1.242 – 5.398	2.585	1.244 – 5.371
AGEGR ₇₀₋	1.567	0.536 – 4.173	1.567	0.568 – 4.327
CHL	1.0087	1.002 – 0.0152	1.0087	1.002 – 1.0152
SMK	2.371	1.324 – 4.452	2.371	1.297 – 4.335
HPT	1.687	0.954 – 2.993	1.687	0.954 – 2.980

- ▶ オッズ比は **JMP** と **R** で、**AGEGR** も込めてすべての変数で一致
- ▶ 名義変数のオッズ比の信頼区間は、**JMP** と **R** でわずかに違う

多変量ロジスティック回帰分析 オッズ比 3

- ▶ $h = \alpha + \beta_{\text{high}} \text{CAT} + \beta_{\text{CHL}} \text{CHL} + \beta_1 \text{SMK} + \beta_{\text{presence}} \text{HPT}$
 $+ \beta_{50-59} \text{AGEGR}_{50-59} + \beta_{60-69} \text{AGEGR}_{60-69} + \beta_{70-} \text{AGEGR}_{70-}$
確率 $P(\text{CHD}) = \frac{1}{1 + \exp(-h)}$, オッズ $O(\text{CHD}) = \frac{P(\text{CHD})}{1 - P(\text{CHD})} = \exp(h)$

多変量ロジスティック回帰分析 オッズ比 3

- ▶ $h = \alpha + \beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL} + \beta_1\text{SMK} + \beta_{\text{presence}}\text{HPT}$
 $+ \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-}$
確率 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$, オッズ $O(\text{CHD}) = \frac{P(\text{CHD})}{1-P(\text{CHD})} = \exp(h)$
- ▶ オッズ比は各変数について、**reference** 以外の値ごとに定められる
AGEGR(ref. は 40-49) では各 **50-59**、**60-69**、**70-**にオッズ比がある

多変量ロジスティック回帰分析 オッズ比 3

- ▶ $h = \alpha + \beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL} + \beta_1\text{SMK} + \beta_{\text{presence}}\text{HPT} + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-}$
確率 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$, オッズ $O(\text{CHD}) = \frac{P(\text{CHD})}{1-P(\text{CHD})} = \exp(h)$
- ▶ オッズ比は各変数について、reference 以外の値ごとに定められる
AGEGR(ref. は 40-49) では各 50-59、60-69、70-にオッズ比がある
- ▶ 40-49 を reference とした 50-59 のオッズ比 $\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)}$

(CAT, CHL, SMK, HPT の値は分子、分母で共通)

分子、分母の $\beta_{\text{high}}\text{CAT}$ などの項は約分され消える

残るのは AGEGR だけの単変量解析のオッズ比と同じ式

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \begin{cases} \exp(2 \times \beta_{50-59} + \beta_{60-69} + \beta_{70-}) & \text{JMP} \\ \exp(\beta_{50-59}) & \text{R} \end{cases}$$

多変量ロジスティック回帰分析 オッズ比 3

- ▶ $h = \alpha + \beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL} + \beta_1\text{SMK} + \beta_{\text{presence}}\text{HPT} + \beta_{50-59}\text{AGEGR}_{50-59} + \beta_{60-69}\text{AGEGR}_{60-69} + \beta_{70-}\text{AGEGR}_{70-}$
確率 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$, オッズ $O(\text{CHD}) = \frac{P(\text{CHD})}{1-P(\text{CHD})} = \exp(h)$
- ▶ オッズ比は各変数について、reference 以外の値ごとに定められる
AGEGR(ref. は 40-49) では各 **50-59**、**60-69**、**70-**にオッズ比がある
- ▶ **40-49** を reference とした **50-59** のオッズ比 $\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)}$

(CAT, CHL, SMK, HPT の値は分子、分母で共通)

分子、分母の $\beta_{\text{high}}\text{CAT}$ などの項は約分され消える

残るのは **AGEGR** だけの単変量解析のオッズ比と同じ式

$$\frac{O(\text{CHD}|\text{AGEGR}=50-59)}{O(\text{CHD}|\text{AGEGR}=40-49)} = \begin{cases} \exp(2 \times \beta_{50-59} + \beta_{60-69} + \beta_{70-}) & \text{JMP} \\ \exp(\beta_{50-59}) & \text{R} \end{cases}$$

- ▶ 単変量と多変量では β_{50-59} , β_{60-69} , β_{70-} の値が異なる
オッズ比の値も異なる

単変量 R 50-59/40-49 1.240 (0.648 – 2.375)

多変量 R 50-59/40-49 1.127 (0.565 – 2.248)

鑑別 (Discrimination)

	CHD +	CHD -	total
logistic +	a	b	$a + b$
logistic -	c	d	$c + d$
total	71	538	609

- ▶ cutoff 値 c_0 により $P(\text{CHD}) \geq c_0$ のとき, logistic + とする

鑑別 (Discrimination)

	CHD +	CHD -	total
logistic +	a	b	$a + b$
logistic -	c	d	$c + d$
total	71	538	609

- ▶ cutoff 値 c_0 により $P(\text{CHD}) \geq c_0$ のとき, logistic + とする
- ▶ cutoff 値 c_0 は感度 (sensitivity) $\frac{a}{a+c}$ 、特異度 (specificity) $\frac{d}{b+d}$ がともに大きくなるように採る

鑑別 (Discrimination)

	CHD +	CHD -	total
logistic +	a	b	$a + b$
logistic -	c	d	$c + d$
total	71	538	609

- ▶ cutoff 値 c_0 により $P(\text{CHD}) \geq c_0$ のとき, **logistic +** とする
- ▶ cutoff 値 c_0 は感度 (sensitivity) $\frac{a}{a+c}$ 、特異度 (specificity) $\frac{d}{b+d}$ がともに大きくなるように採る
- ▶ **ROC 曲線 (Receiver Operating Characteristics curve)** 解析
 cutoff 値 c_0 を少しずつ変化させ、感度、特異度の変化をみる
 横軸に 1- 特異度、縦軸に感度
 - 目的 1 最も良い cutoff 値を探す
 - 目的 2 **AUC(area under the curve)** を計算する
 (AUC が 1 に近いほど鑑別能が高い)

JMP の場合

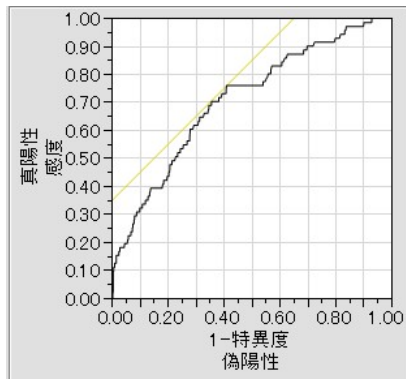
- ▶ 出力のメニューから「ROC 曲線」を選ぶ

The screenshot shows the JMP interface with the 'Nominal Logistic Regression' menu open. The 'ROC Curve' option is highlighted. A tooltip explains that the ROC curve plots sensitivity and 1-specificity. Below the menu, a table displays model fit statistics for various predictors.

Model	Deviance	ChiSq	Prob	ChiSq	Lower 95%
CHL	0.00865341	0.0032728	6.99	0.0082*	0.0022203
SMK[1]	0.4316216	0.1539592	7.86	0.0051*	0.1401502
HPT[presence]	0.26137501	0.1452361	3.24	0.0719	-0.023310

JMP の場合

- ▶ 出力のメニューから「ROC 曲線」を選ぶ
- ▶ 「ROC 曲線」が作られる。AUC=0.710



CHD='presence'を陽性としています。

曲線の下面積= 0.70988

4	presence
0	presence
5	absence
2	absence
8	absence
5	absence
4	absence
2	presence
6	presence
8	absence
2	absence
0	absence
8	presence
5	presence

JMP の場合

- ▶ 出力のメニューから「ROC 曲線」を選ぶ
- ▶ 「ROC 曲線」が作られる。AUC=0.710
- ▶ 「ROC テーブル」に JMP が * の印を表示する
cutoff = 0.0962, a = 54, d = 318 など



確率	1-特異度	感度	感度-(1-特異度)	真陽性	真陰性	偽陽性
0.1014	0.3903	0.7183	0.3280	51	328	21
0.1010	0.3922	0.7324	0.3402	52	327	21
0.1006	0.3941	0.7324	0.3383	52	326	21
0.0997	0.3959	0.7324	0.3365	52	325	21
0.0989	0.3978	0.7324	0.3346	52	324	21
0.0983	0.4033	0.7324	0.3290	52	321	21
0.0981	0.4052	0.7324	0.3272	52	320	21
0.0968	0.4071	0.7324	0.3253	52	319	21
0.0964	0.4071	0.7465	0.3394	53	319	21
0.0962	0.4089	0.7606	0.3516	54	318	21
0.0960	0.4108	0.7606	0.3498	54	317	21
0.0959	0.4126	0.7606	0.3479	54	316	21
0.0958	0.4182	0.7606	0.3423	54	313	21

R の場合 **ROCR Package** を用いる

▶ **ROC** 解析の実行

```
library(ROCR)
```

```
pred <- prediction(fitted(ml_5_vars), CHD)
```

```
perf <- performance(pred, "tpr", "fpr")
```

```
perf1 <- performance(pred, "auc")
```

R の場合 ROCR Package を用いる

- ▶ ROC 解析の実行

```
library(ROCR)
```

```
pred <- prediction(fitted(ml_5_vars), CHD)
```

```
perf <- performance(pred, "tpr", "fpr")
```

```
perf1 <- performance(pred, "auc")
```

- ▶ 最良の **cutoff** 値の番号を求める (246 が出力される)

```
M <- max(perf@y.values[[1]]-perf@x.values[[1]])
```

```
which(perf@y.values[[1]]-perf@x.values[[1]]==M)
```


R の場合 ROCR Package を用いる

- ▶ ROC 解析の実行

```
library(ROCR)
pred <- prediction(fitted(ml_5_vars), CHD)
perf <- performance(pred, "tpr", "fpr")
perf1 <- performance(pred, "auc")
```

- ▶ 最良の **cutoff** 値の番号を求める (**246** が出力される)

```
M <- max(perf@y.values[[1]]-perf@x.values[[1]])
which(perf@y.values[[1]]-perf@x.values[[1]]==M)
```

- ▶ **cutoff** 値、感度、特異度、**AUC** の出力

```
perf@alpha.values[[1]][246]
perf@y.values[[1]][246]
1-perf@x.values[[1]][246]
perf1@y.values
```

0.0962, 0.761, 0.591, 0.710 が出力される

R の場合 **ROCR Package** を用いる

- ▶ **ROC** 解析の実行

```
library(ROCR)
```

```
pred <- prediction(fitted(ml_5_vars), CHD)
```

```
perf <- performance(pred, "tpr", "fpr")
```

```
perf1 <- performance(pred, "auc")
```

- ▶ 最良の **cutoff** 値の番号を求める (**246** が出力される)

```
M <- max(perf@y.values[[1]]-perf@x.values[[1]])
```

```
which(perf@y.values[[1]]-perf@x.values[[1]]==M)
```

- ▶ **cutoff** 値、感度、特異度、**AUC** の出力

```
perf@alpha.values[[1]][246]
```

```
perf@y.values[[1]][246]
```

```
1-perf@x.values[[1]][246]
```

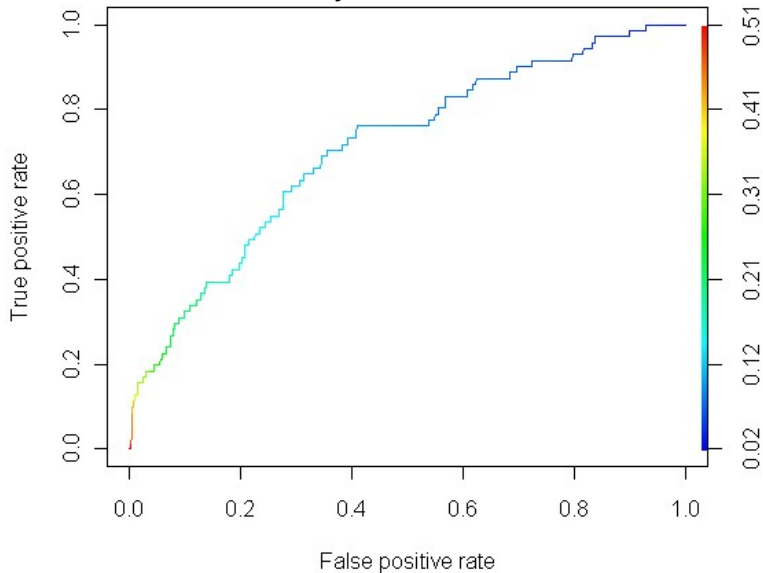
```
perf1@y.values
```

0.0962, 0.761, 0.591, 0.710 が出力される

- ▶ **ROC** 曲線の描画

```
plot(perf, colorize=TRUE)
```

ROCR library の描いた ROC 曲線



JMP と R の ROC 解析の結果は一致

▶ **AUC= 0.710**

0.90-1.0	excellent discrimination
0.80-0.90	good
0.70-0.80	fair
0.60-0.70	poor
0.50-0.60	failed

[Kleinbaum, Klein, p.357]

JMP と R の ROC 解析の結果は一致

- ▶ **AUC= 0.710**

0.90-1.0	excellent discrimination
0.80-0.90	good
0.70-0.80	fair
0.60-0.70	poor
0.50-0.60	failed

[Kleinbaum, Klein, p.357]

- ▶ **cutoff 値 = 0.0962**

JMP と R の ROC 解析の結果は一致

- ▶ **AUC= 0.710**

0.90-1.0	excellent discrimination
0.80-0.90	good
0.70-0.80	fair
0.60-0.70	poor
0.50-0.60	failed

[Kleinbaum, Klein, p.357]

- ▶ **cutoff 値 = 0.0962**
- ▶ **感度 = 0.761, 特異度 = 0.591**

	CHD +	CHD -	total
logistic +	54	220	274
logistic -	17	318	335
total	71	538	609

- ▶ モデル式 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$ の h に
例えば 定数 \times CAT \times CHL のような項を含めるとき
(2 次の) 交互作用項 (interaction term) という
1 次の項 $\beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL}$ では表せない効果があるとき
 h にこのような項を含めたモデルが必要になる

- ▶ モデル式 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$ の h に
例えば 定数 \times CAT \times CHL のような項を含めるとき
(2 次の) 交互作用項 (interaction term) という
1 次の項 $\beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL}$ では表せない効果があるとき
 h にこのような項を含めたモデルが必要になる
- ▶ モデルが 2 次の変数 $A \times B$ を含めば
1 次の変数 A と B も含むという条件をみたすとき
hierarchically well-formulated model, HWF モデルという

- ▶ モデル式 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$ の h に
例えば 定数 \times CAT \times CHL のような項を含めるとき
(2 次の) 交互作用項 (interaction term) という
1 次の項 $\beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL}$ では表せない効果があるとき
 h にこのような項を含めたモデルが必要になる
- ▶ モデルが 2 次の変数 $A \times B$ を含めば
1 次の変数 A と B も含むという条件をみたすとき
hierarchically well-formulated model, HWF モデルという
- ▶ JMP と R で名義変数の扱いが違ってても
(例えば 2 値変数は JMP では $-1, 1$ を R では $0, 1$ を当てはめる)
HFW モデルでは両者の解析結果は一致する
[Kleinbaum, Klein, p.183]

- ▶ モデル式 $P(\text{CHD}) = \frac{1}{1+\exp(-h)}$ の h に
例えば 定数 \times CAT \times CHL のような項を含めるとき
(2 次) 交互作用項 (interaction term) という
1 次の項 $\beta_{\text{high}}\text{CAT} + \beta_{\text{CHL}}\text{CHL}$ では表せない効果があるとき
 h にこのような項を含めたモデルが必要になる
- ▶ モデルが 2 次の変数 $A \times B$ を含めば
1 次の変数 A と B も含むという条件をみたすとき
hierarchically well-formulated model, HWF モデルという
- ▶ JMP と R で名義変数の扱いが違ってても
(例えば 2 値変数は JMP では $-1, 1$ を R では $0, 1$ を当てはめる)
HFW モデルでは両者の解析結果は一致する
[Kleinbaum, Klein, p.183]
- ▶ 実際、4 値変数 AGEGR、2 値変数 HPT と AGEGR \times HPT の
3 つの説明変数をもつモデルなどで
JMP と R の計算結果が一致することを確認した

- ▶ 1 次の項と 2 次の交互作用項を含む h によりデータにより適合したモデルが作れるか検討する

- ▶ 1 次の項と 2 次の交互作用項を含む h によりデータにより適合したモデルが作れるか検討する
- ▶ 1 次の項 : **CAT, AGEGR, CHL, SMK, ECG, DBP, SBP, HPT** の 8 個
2 次の交互作用項 : ${}_8C_2 = 28$ 個
合計 36 個の項から
JMP では「最小 **AICc**」を、**R** では「最小 **AIC**」を基準にしたステップワイズ法の変数増加法で項の選択をする

- ▶ 1 次の項と 2 次の交互作用項を含む h によりデータにより適合したモデルが作れるか検討する
- ▶ 1 次の項 : **CAT, AGEGR, CHL, SMK, ECG, DBP, SBP, HPT** の 8 個
2 次の交互作用項 : ${}_8C_2 = 28$ 個
合計 36 個の項から
JMP では「最小 **AICc**」を、**R** では「最小 **AIC**」を基準にしたステップワイズ法の変数増加法で項の選択をする
- ▶ ただし、**HWF** モデルに限定した変数選択を行なう

多変量ロジスティック回帰分析 交互作用 3

- ▶ JMP での計算 「モデルの指定」で、手法に「ステップワイズ」
Y に CHD を次数に 2 を入れ、説明変数 8 個を選択反転表示して

列の選択

- CHD
- CAT
- AGEGR
- CHL
- SMK
- ECG
- DBP
- SBP
- HPT

役割変数の選択

Y: CHD
オプション

重み: オプション(数値)

度数: オプション(数値)

By: オプション

手法: ステップワイズ法

ヘルプ 実行

前回の設定 ダイアログを開いたままにする

削除

モデル効果の構成

追加

交差

枝分かれ

マクロ

次数: 2

属性

多変量ロジスティック回帰分析 交互作用 3

- ▶ JMP での計算 「モデルの指定」で、手法に「ステップワイズ」
Y に CHD を次数に 2 を入れ、説明変数 8 個を選択反転表示して
- ▶ マクロから「設定された次数まで」を選ぶ

列の選択

- CHD
- CAT
- AGEGR
- CHL
- SMK
- ECG
- DBP
- SBP
- HPT

役割変数の選択

Y: CHD
オプション

重み: オプション(数値)

度数: オプション(数値)

By: オプション

手法: ステップワイズ法

ヘルプ 実行

前回の設定 ダイアログを開いたままにする

削除

モデル効果の構成

追加

交差

枝分かれ

マクロ ▾

完全実施要因
設定された次数まで
すべての組み合わせ

多変量ロジスティック回帰分析 交互作用 3

- ▶ JMP での計算 「モデルの指定」で、手法に「ステップワイズ」
Y に CHD を次数に 2 を入れ、説明変数 8 個を選択反転表示して
- ▶ マクロから「設定された次数まで」を選ぶ
- ▶ これで 36 個の項が「モデル効果の構成」に入る。「実行」を押す

列の選択

- CHD
- CAT
- AGEGR
- CHL
- SMK
- ECG
- DBP
- SBP
- HPT

役割変数の選択

Y: CHD
オプション

重み: オプション(数値)

度数: オプション(数値)

By: オプション

手法: ステップワイズ法

ヘルプ

実行

前回の設定 ダイアログを開いたままにする

削除

モデル効果の構成

追加: CAT

交差: AGEGR

枝分かれ: CHL

マクロ: SMK

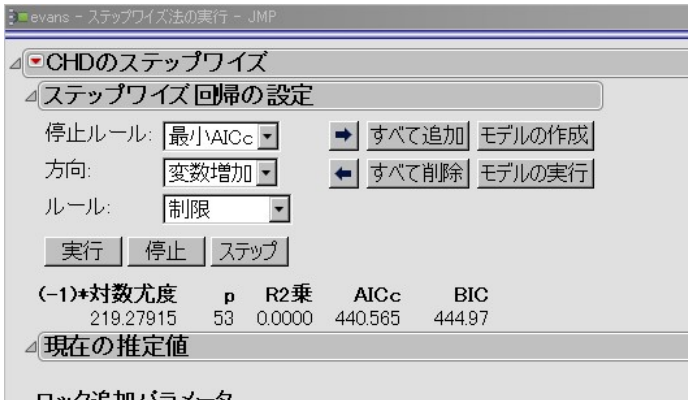
次数: 2

属性: ECG

CAT*AGEGR

多変量ロジスティック回帰分析 交互作用 4

- ▶ 「最小 AICc」、「変数増加」、「制限」を入れ実行を押す
(「制限」は HWF モデルに限定するため)



evans - ステップワイズ法の実行 - JMP

CHDのステップワイズ

ステップワイズ回帰の設定

停止ルール: 最小AICc → すべて追加 モデルの作成

方向: 変数増加 ← すべて削除 モデルの実行

ルール: 制限

実行 停止 ステップ

(-1)*対数尤度	p	R2乗	AICc	BIC
219.27915	53	0.0000	440.565	444.97

現在の推定値

ロック追加パラメータ

多変量ロジスティック回帰分析 交互作用 4

- ▶ 「最小 AICc」、「変数増加」、「制限」を入れ実行を押す
(「制限」は HWF モデルに限定するため)
- ▶ 1step ごとに、それまでの step で入った項に
残りの項のうち AICc が最小となる項が 1 つ追加される

SMK{1-0}	8	362.888
AGEGR{60-69&70--50-59&40-49}*SMK{1-0}	9	360.773
(CHL-211.739)*HPT{presence-absence}	10	358.8
AGEGR{60-69&70--50-59&40-49}*HPT{presence-absence}	11	357.167
AGEGR{60-69&70--50-59&40-49}*(CHL-211.739)	12	358.103
(CHL-211.739)*SMK{1-0}	13	358.499
CAT{high-normal}*AGEGR{60-69&70--50-59&40-49}	14	359.95
DBP	15	361.632
(CHL-211.739)*(DBP-91.1806)	16	361.569
SBP	17	362.101
(CHL-211.739)*(SBP-145.476)	18	352.993
AGEGR{60-69&70--50-59&40-49}*(SBP-145.476)	19	353.131
CAT{high-normal}*(DBP-91.1806)	20	353.148
AGEGR{60-69&70--50-59&40-49}*(DBP-91.1806)	21	353.393
(DBP-91.1806)*HPT{presence-absence}	22	354.093
(SBP-145.476)*HPT{presence-absence}	23	353.877

出力の上部下部を省いた抜粋

多変量ロジスティック回帰分析 交互作用 4

- ▶ 「最小 AICc」、「変数増加」、「制限」を入れ実行を押す
(「制限」は HWF モデルに限定するため)
- ▶ 1step ごとに、それまでの step で入った項に
残りの項のうち AICc が最小となる項が 1 つ追加される
- ▶ 方法 1. AICc が減少から増加に転じたら停止 10 項が入る
方法 2. 36 項すべてを追加し最小 AICc を探す 17 項が入る

SMK{1-0}	8	362.888
AGEGR{60-69&70--50-59&40-49}*SMK{1-0}	9	360.773
(CHL-211.739)*HPT{presence-absence}	10	358.8
AGEGR{60-69&70--50-59&40-49}*HPT{presence-absence}	11	357.167 ●
AGEGR{60-69&70--50-59&40-49}*(CHL-211.739)	12	358.103
(CHL-211.739)*SMK{1-0}	13	358.499
CAT{high-normal}*AGEGR{60-69&70--50-59&40-49}	14	359.95
DBP	15	361.632
(CHL-211.739)*(DBP-91.1806)	16	361.569
SBP	17	362.101
(CHL-211.739)*(SBP-145.476)	18	352.993 ●
AGEGR{60-69&70--50-59&40-49}*(SBP-145.476)	19	353.131
CAT{high-normal}*(DBP-91.1806)	20	353.148
AGEGR{60-69&70--50-59&40-49}*(DBP-91.1806)	21	353.393
(DBP-91.1806)*HPT{presence-absence}	22	354.093
(SBP-145.476)*HPT{presence-absence}	23	353.877

出力の上部下部を省いた抜粋

▶ Rでのステップワイズ法のスクリプト

```
degree2<-glm(CHD ~ CAT*AGEGR + CAT*SMK + CAT*ECG
+ CAT*HPT + CAT*CHL + CAT*DBP + CAT*SBP + AGEGR*SMK
+ AGEGR*ECG + AGEGR*HPT + AGEGR*CHL + AGEGR*DBP
+ AGEGR*SBP + SMK*ECG + SMK*HPT + SMK*CHL + SMK*DBP
+ SMK*SBP + ECG*HPT + ECG*CHL + ECG*DBP + ECG*SBP
+ HPT*CHL + HPT*DBP + HPT*SBP + CHL*DBP + CHL*SBP
+ DBP*SBP,
              family=binomial,data=dfr)

null<-glm(CHD~1,family=binomial,data=dfr)

ml_degree2<-step(null,scope=formula(degree2),
                  direction="forward")
```

▶ Rでのステップワイズ法のスクリプト

```
degree2<-glm(CHD ~ CAT*AGEGR + CAT*SMK + CAT*ECG  
+ CAT*HPT + CAT*CHL + CAT*DBP + CAT*SBP + AGEGR*SMK  
+ AGEGR*ECG + AGEGR*HPT + AGEGR*CHL + AGEGR*DBP  
+ AGEGR*SBP + SMK*ECG + SMK*HPT + SMK*CHL + SMK*DBP  
+ SMK*SBP + ECG*HPT + ECG*CHL + ECG*DBP + ECG*SBP  
+ HPT*CHL + HPT*DBP + HPT*SBP + CHL*DBP + CHL*SBP  
+ DBP*SBP, family=binomial,data=dfr)
```

```
null<-glm(CHD~1,family=binomial,data=dfr)
```

```
ml_degree2<-step(null,scope=formula(degree2),  
direction="forward")
```

- ▶ HWF モデルに限定して選択が行われる
AIC は 11 個の項が入るまで減少し 12 項目に増加に転じる
11 個の項が入って停止する

**JMP, R に共通に入るのはつぎの 9 項
 CAT, CHL, AGEGR, SMK, HPT,
 CAT×HPT, CAT×CHL, CHL×HPT, AGEGR×SMK**

▶ **JMP の選択** **R の選択**

項	JMP AICc	R AIC
9 項 + AGEGR×HPT の 10 項	357.2	362.7
9 項 + AGEGR×CHL + SMK×CHL の 11 項	370.2	360.9

**JMP, R に共通に入るのはつぎの 9 項
 CAT, CHL, AGEGR, SMK, HPT,
 CAT×HPT, CAT×CHL, CHL×HPT, AGEGR×SMK**

▶ **JMP の選択 R の選択**

項	JMP AICc	R AIC
9 項 + AGEGR×HPT の 10 項	357.2	362.7
9 項 + AGEGR×CHL + SMK×CHL の 11 項	370.2	360.9

▶ **AUC, 感度, 特異度**

10 項モデル			
	AUC	感度	特異度
JMP	0.813	0.690	0.805
R	0.807	0.718	0.762

11 項モデル			
	AUC	感度	特異度
JMP	0.811	0.690	0.829
R	0.808	0.746	0.770

**JMP, R に共通に入るのはつぎの 9 項
 CAT, CHL, AGEGR, SMK, HPT,
 CAT×HPT, CAT×CHL, CHL×HPT, AGEGR×SMK**

- ▶ **JMP の選択** **R の選択**

項	JMP AICc	R AIC
9 項 + AGEGR×HPT の 10 項	357.2	362.7
9 項 + AGEGR×CHL + SMK×CHL の 11 項	370.2	360.9

- ▶ **AUC, 感度, 特異度**

10 項モデル			
	AUC	感度	特異度
JMP	0.813	0.690	0.805
R	0.807	0.718	0.762

11 項モデル			
	AUC	感度	特異度
JMP	0.811	0.690	0.829
R	0.808	0.746	0.770

- ▶ 交互作用をもつモデルでは、**JMP** と **R** の結果に違いがある

多変量ロジスティック回帰分析 交互作用 7

10 項モデルの JMP, R での計算を比較する (パラメータ総数:17)

変数	パラメータの数	計
切片, CAT, CHL, SMK, HPT	1	5
CAT×HPT, CAT×CHL, CHL×HPT	1	3
AGEGR, AGEGR×SMK, AGEGR×HPT	3	9

- ▶ JMP による No.1 ~ No.5 の $P(\text{CHD})$ の計算値

	線形[presence]	確率[presence]	確率[absence]
1	-3.266317378	0.0367449516	0.9632550484
2	-2.773017602	0.0587997897	0.9412002103
3	-2.815462058	0.056494332	0.943505668
4	-2.955403583	0.0494817436	0.9505182564
5	-2.446218802	0.0797155009	0.9202844991
6	-2.411204645	0.082322267	0.917677733

多変量ロジスティック回帰分析 交互作用 7

10 項モデルの JMP, R での計算を比較する (パラメータ総数:17)

変数	パラメータの数	計
切片, CAT, CHL, SMK, HPT	1	5
CAT×HPT, CAT×CHL, CHL×HPT	1	3
AGEGR, AGEGR×SMK, AGEGR×HPT	3	9

- ▶ JMP による No.1 ~ No.5 の $P(\text{CHD})$ の計算値
- ▶ R による No.1 ~ No.5 の $P(\text{CHD})$ の計算値

```

R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> ml_10<-glm(CHD ~ CAT + CHL + SMK + HPT + AGEGR + CAT:CHL +
+           CAT:HPT + CHL:HPT + SMK:AGEGR + HPT:AGEGR,
+           family=binomial,data=df)
> fitted(ml_10) [1:5]
      1          2          3          4          5
0.03466004 0.06763080 0.09334879 0.03166595 0.08645357
>
>
>

```

多変量ロジスティック回帰分析 交互作用 7

10 項モデルの JMP, R での計算を比較する (パラメータ総数:17)

変数	パラメータの数	計
切片, CAT, CHL, SMK, HPT	1	5
CAT×HPT, CAT×CHL, CHL×HPT	1	3
AGEGR, AGEGR×SMK, AGEGR×HPT	3	9

- ▶ JMP による No.1 ~ No.5 の $P(\text{CHD})$ の計算値
- ▶ R による No.1 ~ No.5 の $P(\text{CHD})$ の計算値
- ▶ JMP と R の結果は両者は非常に異なる

JMP	0.0367	0.0588	0.0565	0.0495	0.0797
R	0.0347	0.0676	0.0933	0.0317	0.0865

```

R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> ml_10<-glm(CHD ~ CAT + CHL + SMK + HPT + AGEGR + CAT:CHL +
+           CAT:HPT + CHL:HPT + SMK:AGEGR + HPT:AGEGR,
+           family=binomial,data=df)
> fitted(ml_10) [1:5]
      1          2          3          4          5
0.03466004 0.06763080 0.09334879 0.03166595 0.08645357
>
>
>

```

- ▶ ロジスティック回帰分析では、最大尤度法によりパラメータを未知数とする方程式 (最尤方程式) を作る

- ▶ ロジスティック回帰分析では、最大尤度法によりパラメータを未知数とする方程式 (最尤方程式) を作る
- ▶ 最尤方程式は代数的計算で解くことが出来ない
解の初期候補値から始めてニュートン法による反復計算により第 2 値、第 3 値、... を求めこれが収束したら極限值を解とする

- ▶ ロジスティック回帰分析では、最大尤度法によりパラメータを未知数とする方程式 (最尤方程式) を作る
- ▶ 最尤方程式は代数的計算で解くことが出来ない
解の初期候補値から始めてニュートン法による反復計算により第2値、第3値、... を求めこれが収束したら極限值を解とする
- ▶ ニュートン法では異なる初期候補値から始めると異なる極限值が得られることもある

- ▶ ロジスティック回帰分析では、最大尤度法によりパラメータを未知数とする方程式 (最尤方程式) を作る
- ▶ 最尤方程式は代数的計算で解くことが出来ない
解の初期候補値から始めてニュートン法による反復計算により第2値、第3値、... を求めこれが収束したら極限值を解とする
- ▶ ニュートン法では異なる初期候補値から始めると異なる極限值が得られることもある
- ▶ ロジスティック回帰を扱う統計ソフトの良否は極限值として最大尤度解が得られるかどうかにかかっている

- ▶ ロジスティック回帰分析では、最大尤度法によりパラメータを未知数とする方程式 (最尤方程式) を作る
- ▶ 最尤方程式は代数的計算で解くことが出来ない
解の初期候補値から始めてニュートン法による反復計算により第2値、第3値、... を求めこれが収束したら極限值を解とする
- ▶ ニュートン法では異なる初期候補値から始めると異なる極限值が得られることもある
- ▶ ロジスティック回帰を扱う統計ソフトの良否は極限值として最大尤度解が得られるかどうかにかかっている
- ▶ 10 項モデル (17 パラメータ) では **JMP** の計算では最大尤度解でない値に収束していると思われる

ordinal または polytomous

ロジスティック回帰 データファイル 1

目的変数は 3 値以上、
値に順序があるとき ordinal、順序がないとき polytomous

目的変数	ordinal	GRADE (tumor grade of endometrial cancer, well (0), moderate (1) or poor (2))
目的変数	polytomous	SUBTYPE (adenocarcinoma (0), adenosquamous (1) or other (2))
説明変数	共通	ESTROGEN (never (0) or ever used (1)) AGE (50-64 (0) or 65-79 (1))



cancer.xls [互換モード] - Microsoft Excel

	A	B	C	D	E	F	G
1	GRADE	SUBTYPE	ESTROGEN	AGE			
2	1	1	0	0			
3	0	2	1	0			
4	1	1	0	1			
5	0	0	0	1			
6	0	0	1	0			

子宮内膜癌 (endometrial cancer) データ, 標本数は 286

ordinal または polytomous ロジスティック回帰 データファイル 2

▶ **GRADE** の内訳

well	moderate	poor
128	105	53

ordinal または polytomous ロジスティック回帰 データファイル 2

▶ **GRADE** の内訳

well	moderate	poor
128	105	53

▶ **SUBTYPE** の内訳

adenocarcinoma	adenosquamous	other
184	45	57

ordinal または polytomous ロジスティック回帰 データファイル 2

▶ **GRADE** の内訳

well	moderate	poor
128	105	53

▶ **SUBTYPE** の内訳

adenocarcinoma	adenosquamous	other
184	45	57

▶ **AGE** の内訳

50-64	65-79
106	180

ordinal または polytomous ロジスティック回帰 データファイル 2

▶ **GRADE** の内訳

well	moderate	poor
128	105	53

▶ **SUBTYPE** の内訳

adenocarcinoma	adenosquamous	other
184	45	57

▶ **AGE** の内訳

50-64	65-79
106	180

▶ **ESTROGEN** の内訳

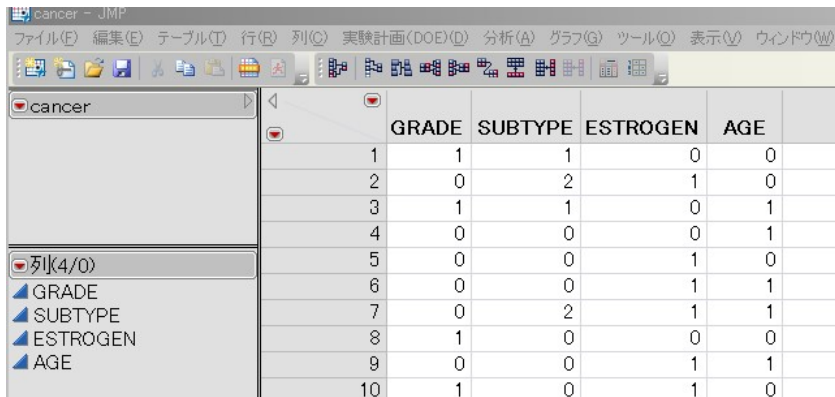
never	ever
135	151

OrdinalまたはPolytomous

ロジスティック回帰分析

JMP 計算の準備

- ▶ Excel データの読み込み



The screenshot shows the JMP interface with a data table. The table has 10 rows and 4 columns: GRADE, SUBTYPE, ESTROGEN, and AGE. The left sidebar shows the column list with GRADE, SUBTYPE, ESTROGEN, and AGE selected.

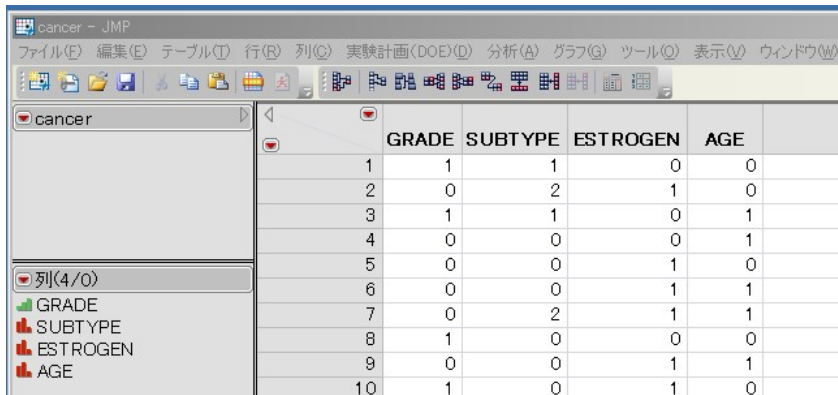
	GRADE	SUBTYPE	ESTROGEN	AGE
1	1	1	0	0
2	0	2	1	0
3	1	1	0	1
4	0	0	0	1
5	0	0	1	0
6	0	0	1	1
7	0	2	1	1
8	1	0	0	0
9	0	0	1	1
10	1	0	1	0

Ordinal または Polytomous

ロジスティック回帰分析

JMP 計算の準備

- ▶ Excel データの読み込み
- ▶ **GRADE** は順序変数、他は名義変数



The screenshot shows the JMP interface with a data table named 'cancer'. The table has 10 rows and 4 columns: GRADE, SUBTYPE, ESTROGEN, and AGE. The GRADE column contains values 1, 2, 1, 0, 0, 0, 0, 1, 0, 1. The SUBTYPE column contains values 1, 2, 1, 0, 0, 0, 2, 0, 0, 0. The ESTROGEN column contains values 0, 1, 0, 0, 1, 1, 1, 0, 1, 1. The AGE column contains values 0, 0, 1, 1, 0, 1, 1, 0, 1, 0.

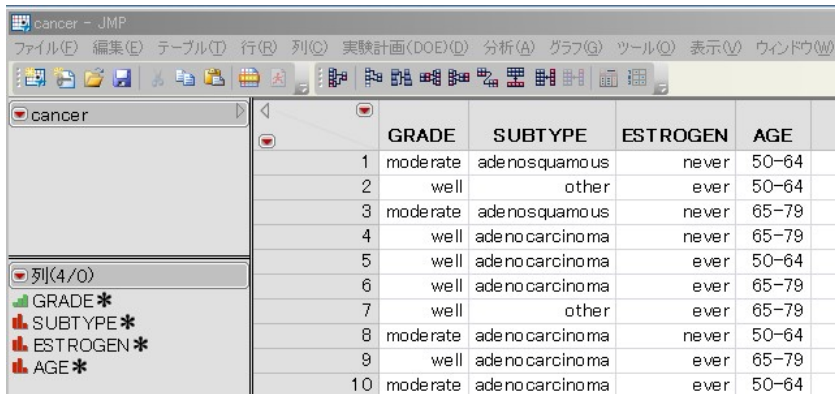
	GRADE	SUBTYPE	ESTROGEN	AGE
1	1	1	0	0
2	0	2	1	0
3	1	1	0	1
4	0	0	0	1
5	0	0	1	0
6	0	0	1	1
7	0	2	1	1
8	1	0	0	0
9	0	0	1	1
10	1	0	1	0

Ordinal または Polytomous

ロジスティック回帰分析

JMP 計算の準備

- ▶ Excel データの読み込み
- ▶ **GRADE** は順序変数、他は名義変数
- ▶ 4 つの変数のそれぞれに、値にラベルを付ける



The screenshot shows the JMP software interface with a data table named 'cancer'. The table has 10 rows and 4 columns: GRADE, SUBTYPE, ESTROGEN, and AGE. The GRADE column contains values 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. The SUBTYPE column contains values 'moderate', 'well', 'adenosquamous', 'adenocarcinoma', 'adenocarcinoma', 'adenocarcinoma', 'well', 'adenocarcinoma', 'adenocarcinoma', and 'adenocarcinoma'. The ESTROGEN column contains values 'never', 'ever', 'never', 'never', 'ever', 'ever', 'ever', 'never', 'ever', and 'ever'. The AGE column contains values '50-64', '50-64', '65-79', '65-79', '50-64', '65-79', '65-79', '50-64', '65-79', and '50-64'. The left sidebar shows the variable list with GRADE, SUBTYPE, ESTROGEN, and AGE marked with an asterisk.

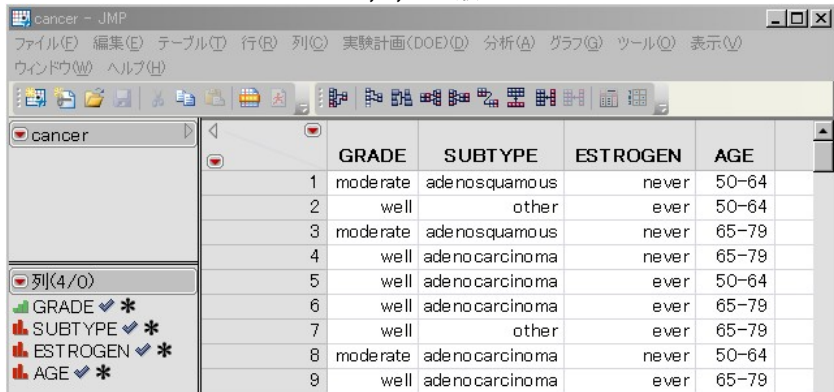
	GRADE	SUBTYPE	ESTROGEN	AGE
1	moderate	adenosquamous	never	50-64
2	well	other	ever	50-64
3	moderate	adenosquamous	never	65-79
4	well	adenocarcinoma	never	65-79
5	well	adenocarcinoma	ever	50-64
6	well	adenocarcinoma	ever	65-79
7	well	other	ever	65-79
8	moderate	adenocarcinoma	never	50-64
9	well	adenocarcinoma	ever	65-79
10	moderate	adenocarcinoma	ever	50-64

OrdinalまたはPolytomous

ロジスティック回帰分析

JMP 計算の準備

- ▶ Excel データの読み込み
- ▶ **GRADE** は順序変数、他は名義変数
- ▶ 4 つの変数のそれぞれに、値にラベルを付ける
- ▶ R の出力に合わせるために、**AGE** と **ESTROGEN** は 1, 0 の順
GRADE と **SUBTYPE** は 2, 1, 0 の順にリストチェック



	GRADE	SUBTYPE	ESTROGEN	AGE
1	moderate	adenosquamous	never	50-64
2	well	other	ever	50-64
3	moderate	adenosquamous	never	65-79
4	well	adenocarcinoma	never	65-79
5	well	adenocarcinoma	ever	50-64
6	well	adenocarcinoma	ever	65-79
7	well	other	ever	65-79
8	moderate	adenocarcinoma	never	50-64
9	well	adenocarcinoma	ever	65-79

Variable Properties (Left Sidebar):

- GRADE: Order (1, 2, 3), List (2, 1, 0), Scale *
- SUBTYPE: List (2, 1, 0), Scale *
- ESTROGEN: List (1, 0), Scale *
- AGE: List (1, 0), Scale *

Ordinal または Polytomous

ロジスティック回帰分析 R 計算の準備

新たに R を起動するところから始めるものとする

- ▶ 必要な **library** をロードする

```
library(epicalc)
```

```
library(car)
```

Ordinal または Polytomous

ロジスティック回帰分析 R 計算の準備

新たに R を起動するところから始めるものとする

- ▶ 必要な **library** をロードする

```
library(epicalc)
```

```
library(car)
```

- ▶ テキスト形式データを読み込む

```
dfr<-read.table("cancer.txt",header=T)
```

Ordinal または Polytomous

ロジスティック回帰分析

R 計算の準備

新たに R を起動するところから始めるものとする

- ▶ 必要な **library** をロードする

```
library(epicalc)
library(car)
```

- ▶ テキスト形式データを読み込む

```
dfr<-read.table("cancer.txt",header=T)
```

- ▶ **GGADE** は **ordered** 関数で順序変数に、他は **factor** で名義変数にそれぞれにラベルも付ける

```
dfr <- within(dfr, {
  GRADE<-ordered(GRADE,labels=c("well","moderate","poor"))
  ESTROGEN<-factor(ESTROGEN,labels=c("never","ever"))
  SUBTYPE<-factor(SUBTYPE,labels=c("adenocarcinoma",
                                   "adenosquamous","other"))
  AGE<-factor(AGE,labels=c("50-64","65-79"))
}); attach(dfr)
```

Ordinal ロジスティック回帰分析 モデル式

- ▶ 目的変数 **GRADE (well (0), moderate (1) or poor (2))**

Ordinal ロジスティック回帰分析 モデル式

- ▶ 目的変数 **GRADE (well (0), moderate (1) or poor (2))**
- ▶ パラメータ 切片 $\alpha_{\text{moderate}}, \alpha_{\text{poor}}$, 係数 $\beta_{65-79}, \beta_{\text{ever}}$

Ordinal ロジスティック回帰分析 モデル式

▶ 目的変数 **GRADE (well (0), moderate (1) or poor (2))**

▶ パラメータ 切片 α_{moderate} , α_{poor} , 係数 β_{65-79} , β_{ever}

$$\begin{aligned} P(\text{GRADE} \geq \text{moderate}) &= \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))} \\ P(\text{GRADE} \geq \text{poor}) &= \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))} \end{aligned}$$

- ▶ 目的変数 **GRADE (well (0), moderate (1) or poor (2))**

- ▶ パラメータ 切片 $\alpha_{\text{moderate}}, \alpha_{\text{poor}}$, 係数 $\beta_{65-79}, \beta_{\text{ever}}$

- ▶
$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))}$$
$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))}$$

- ▶ この2つの確率が分かれば、以下の3つが分かる
 $P(\text{GRADE} = \text{well}), P(\text{GRADE} = \text{moderate}),$
 $P(\text{GRADE} = \text{poor})$

Ordinal ロジスティック回帰分析 JMP 計算

- ▶ 「分析」の「モデルのあてはめ」

The screenshot shows the JMP software interface with the 'Analyze' menu open. The 'Fit Model' option is highlighted. The data table below shows the following columns and rows:

	GRADE	SU		AGE
1	moderate	aden		
2	well			
3	moderate	adenosquamous	never	65-7
4	well	adenocarcinoma	never	65-7
5	well	adenocarcinoma	ever	50-6
6	well	adenocarcinoma	ever	65-7
7	well	other	ever	65-7
8	moderate	adenocarcinoma	never	50-6
9	well	adenocarcinoma	ever	65-7

Ordinal ロジスティック回帰分析 JMP 計算

- ▶ 「分析」の「モデルのあてはめ」
- ▶ Y に **GRADE**、「モデル効果の構成」に **ESTROGEN**、**AGE**
手法は「順序ロジスティック」

▼ モデルの指定

列の選択

- GRADE
- SUBTYPE
- ESTROGEN
- AGE

役割変数の選択

Y: GRADE オプション

重み: オプション(数値)

度数: オプション(数値)

By: オプション

手法: 順序ロジスティック

ヘルプ 実行

前回の設定 ダイアログを開いたままにする

削除

モデル効果の構成

追加: ESTROGEN
交差: AGE

枝分かれ

マクロ

Ordinal ロジスティック回帰分析 JMP 計算

- ▶ 「分析」の「モデルのあてはめ」
- ▶ Y に **GRADE**、「モデル効果の構成」に **ESTROGEN**、**AGE** 手法は「順序ロジスティック」
- ▶ 赤い三角のメニューから、「Wald 検定」、「信頼区間」、「確率の計算式の保存」を選ぶ
注. メニューに「オッズ比」が無い

cancer - 順序ロジスティックのあてはめ - JMP

順序ロジスティックのあてはめ GRADE

検定

尤度比検定
Wald検定
信頼区間
ROC曲線
リフトチャート
混同行列
プロフィール

保存 ▶

モデルダイアログ
スクリプト ▶

599.169
(または重みの合計) 286

学習 定義

0.0309 $1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
0.0711 $(1 - (L(0) / L(\text{model}))^{2/n}) / (1 - L(0)^{2/n})$

確率の計算式の保存 $[i] / n$
分位点の保存 $\phi [i]^2 / n$
期待値の保存 $[i] / n$

0.4965 $\sum (\rho [i] \neq \rho \text{Max}) / n$

N 286 n

あてはまりの悪さ(LOF)

要因	自由度	(-1)*対数尤度	カイ2乗
あてはまりの悪さ(LOF)	4	2.08778	4.17556

- ▶ **polr** 関数で計算
出力を例えば **mo (multivariate ordinal)** と名付ける

```
mo<-polr(GRADE~AGE+ESTROGEN)
```

Ordinal ロジスティック回帰分析 R 計算

- ▶ `polr` 関数で計算
出力を例えば `mo` (**m**ultivariate **o**rdinal) と名付ける

```
mo<-polr(GRADE~AGE+ESTROGEN)
```

- ▶ 切片や係数の出力

```
summary(mo)
```

- ▶ **polr** 関数で計算
出力を例えば **mo (multivariate ordinal)** と名付ける

```
mo<-polr(GRADE~AGE+ESTROGEN)
```

- ▶ 切片や係数の出力

```
summary(mo)
```

- ▶ 尤度比検定の出力

```
Anova(mo) #Anova は car library の関数
```

- ▶ **polr** 関数で計算
出力を例えば **mo (multivariate ordinal)** と名付ける

```
mo<-polr(GRADE~AGE+ESTROGEN)
```

- ▶ 切片や係数の出力

```
summary(mo)
```

- ▶ 尤度比検定の出力

```
Anova(mo) #Anova は car library の関数
```

- ▶ オッズ比の出力

```
ordinal.or.display(mo,decimal=4)
```

Ordinal ロジスティック回帰分析 パラメータ

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

	α_{moderate}	α_{poor}	β_{65-79}	β_{ever}
JMP	0.213	-1.565	0.123	-0.465
R	0.555	-1.224	0.246	-0.929

R の切片は正負を変えて読む

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%
切片[poor]	-1.5654374	0.1600141	95.71	<.0001*	.
切片[moderate]	0.21295885	0.1259267	2.86	0.0908	.
ESTROGEN[ever]	-0.4645796	0.1142337	16.54	<.0001*	-0.6903971
AGE[65-79]	0.12277735	0.1170866	1.10	0.2944	-0.1046846

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
ESTROGEN	1	1	16.5398541	<.0001*
AGE	1	1	1.09956845	0.2944

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
ESTROGEN	1	1	16.9690855	<.0001*
AGE	1	1	1.11660886	0.2906

評価が完了しました。



Ordinal ロジスティック回帰分析 パラメータ

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

	α_{moderate}	α_{poor}	β_{65-79}	β_{ever}
JMP	0.213	-1.565	0.123	-0.465
R	0.555	-1.224	0.246	-0.929

R の切片は正負を変えて読む

```
> summary(mo)
```

```
Re-fitting to get Hessian
```

```
Call:
```

```
polr(formula = GRADE ~ AGE + ESTROGEN)
```

```
Coefficients:
```

```
                Value Std. Error t value
AGE65-79         0.2456    0.2328   1.055
ESTROGENever -0.9292    0.2284  -4.069
```

```
Intercepts:
```

```
                Value Std. Error t value
well|moderate -0.5548    0.2257  -2.4578
moderate|poor  1.2236    0.2371   5.1612
```

Ordinal ロジスティック回帰分析 確率 1

出力された切片、係数から **GRADE** の確率を手計算してみる

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

- ▶ **No.2** のデータを使う

	AGE	ESTROGEN
No. 2	50-64	poor
JMP	-1	1
R	0	1

Ordinal ロジスティック回帰分析 確率 1

出力された切片、係数から **GRADE** の確率を手計算してみる

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

- ▶ No.2 のデータを使う

	AGE	ESTROGEN
No. 2	50-64	poor
JMP	-1	1
R	0	1

- ▶ JMP

$$\alpha_{\text{moderate}} = 0.213, \alpha_{\text{poor}} = -1.565, \beta_{65-79} = 0.123, \beta_{\text{ever}} = -0.465$$

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(- (0.213 - 0.123 - 0.465))} = 0.407$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(- (-1.565 - 0.123 - 0.465))} = 0.104$$

Ordinal ロジスティック回帰分析 確率 1

出力された切片、係数から **GRADE** の確率を手計算してみる

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

- ▶ No.2 のデータを使う

	AGE	ESTROGEN
No. 2	50-64	poor
JMP	-1	1
R	0	1

- ▶ JMP

$$\alpha_{\text{moderate}} = 0.213, \alpha_{\text{poor}} = -1.565, \beta_{65-79} = 0.123, \beta_{\text{ever}} = -0.465$$

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(- (0.213 - 0.123 - 0.465))} = 0.407$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(- (-1.565 - 0.123 - 0.465))} = 0.104$$

- ▶ R

$$\alpha_{\text{moderate}} = 0.555, \alpha_{\text{poor}} = -1.224, \beta_{65-79} = 0.246, \beta_{\text{ever}} = -0.929$$

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(- (0.555 - 0.929))} = 0.408,$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(- (-1.224 - 0.929))} = 0.104$$

Ordinal ロジスティック回帰分析 確率 1

出力された切片、係数から **GRADE** の確率を手計算してみる

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79} \text{AGE} + \beta_{\text{ever}} \text{ESTROGEN}))}$$

- ▶ No.2 のデータを使う

	AGE	ESTROGEN
No. 2	50-64	poor
JMP	-1	1
R	0	1

- ▶ JMP

$$\alpha_{\text{moderate}} = 0.213, \alpha_{\text{poor}} = -1.565, \beta_{65-79} = 0.123, \beta_{\text{ever}} = -0.465$$

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(- (0.213 - 0.123 - 0.465))} = 0.407$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(- (-1.565 - 0.123 - 0.465))} = 0.104$$

- ▶ R

$$\alpha_{\text{moderate}} = 0.555, \alpha_{\text{poor}} = -1.224, \beta_{65-79} = 0.246, \beta_{\text{ever}} = -0.929$$

$$P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(- (0.555 - 0.929))} = 0.408,$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(- (-1.224 - 0.929))} = 0.104$$

- ▶ 故に JMP, R 両方とも $P(\text{GRADE} = \text{well}) = 0.592,$

$$P(\text{GRADE} = \text{moderate}) = 0.303, P(\text{GRADE} = \text{poor}) = 0.104$$

Ordinal ロジスティック回帰分析 確率 2

No.2 の人の確率のソフトでの出力

手計算では JMP でも R でも $P(\text{GRADE} = \text{well}) = 0.592$,
 $P(\text{GRADE} = \text{moderate}) = 0.303$, $P(\text{GRADE} = \text{poor}) = 0.104$,

▶ JMP の場合

	確率 derate]	確率[poor]	確率 [moderate]	確率[well]	最 GRA
1	0.2272973755	0.4079421233	0.3647605011		
2	0.1040703936	0.303408318	0.5925212884		
3	0.273272841	0.4167691898	0.3099579692		
4	0.273272841	0.4167691898	0.3099579692		
5	0.1040703936	0.303408318	0.5925212884		
6	0.1292912839	0.3385423469	0.5321663692		
7	0.1292912839	0.3385423469	0.5321663692		
8	0.2272973755	0.4079421233	0.3647605011		
9	0.1292912839	0.3385423469	0.5321663692		
10	0.1040703936	0.303408318	0.5925212884		
11	0.1292912839	0.3385423469	0.5321663692		
12	0.1040703936	0.303408318	0.5925212884		
13	0.1292912839	0.3385423469	0.5321663692		

Ordinal ロジスティック回帰分析 確率 2

No.2 の人の確率のソフトでの出力

手計算では JMP でも R でも $P(\text{GRADE} = \text{well}) = 0.592$,
 $P(\text{GRADE} = \text{moderate}) = 0.303$, $P(\text{GRADE} = \text{poor}) = 0.104$,

- ▶ JMP の場合
- ▶ R の場合

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> AGE[2]
[1] 50-64
Levels: 50-64 65-79
> ESTROGEN[2]
[1] ever
Levels: never ever
> fitted(mo)[2,1]
[1] 0.5925216
> fitted(mo)[2,2]
[1] 0.3034082
> fitted(mo)[2,3]
[1] 0.1040702
>
>
```

Ordinal ロジスティック回帰分析 検定

$\beta_{65-79} = 0$ の検定、 $\beta_{\text{ever}} = 0$ の検定

▶ JMP の場合

Wald 検定	AGE $p = 0.2944$	ESTROGEN $p < 0.0001$
尤度比検定	AGE $p = 0.2906$	ESTROGEN $p < 0.0001$

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%
切片[poor]	-1.5654374	0.1600141	95.71	<.0001*	.
切片[moderate]	0.21295885	0.1259267	2.86	0.0908	.
ESTROGEN[ever]	-0.4645796	0.1142337	16.54	<.0001*	-0.6903971
AGE[65-79]	0.12277735	0.1170866	1.10	0.2944	-0.1046846

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
ESTROGEN	1	1	16.5398541	<.0001*
AGE	1	1	1.09956845	0.2944

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
ESTROGEN	1	1	16.9690855	<.0001*
AGE	1	1	1.11660886	0.2906

評価が完了しました。



Ordinal ロジスティック回帰分析 検定

$\beta_{65-79} = 0$ の検定、 $\beta_{\text{ever}} = 0$ の検定

- ▶ JMP の場合
 - Wald 検定 AGE $p = 0.2944$ ESTROGEN $p < 0.0001$
 - 尤度比検定 AGE $p = 0.2906$ ESTROGEN $p < 0.0001$
- ▶ R の場合
 - Wald 検定 計算されない
 - 尤度比検定 AGE $p = 0.2906$ ESTROGEN $p < 0.0001$

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> Anova(mo)
Analysis of Deviance Table (Type II tests)

Response: GRADE
          LR Chisq Df Pr(>Chisq)
AGE          1.1166  1   0.2906
ESTROGEN    16.9691  1 3.799e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

>
>
>
>
>
```

Ordinal ロジスティック回帰分析 オッズ比 1

- ▶ JMP オッズ比と信頼区間は出力されない

Wald 法の手計算はできる

$$\text{AGE} \quad \exp(0.123 \times 2) = 1.279$$

$$\exp((0.123 \pm 0.117 \times 1.96) \times 2) = 2.023, 0.808$$

$$\text{ESTROGEN} \quad \exp(-0.465 \times 2) = 0.395$$

$$\exp((-0.465 \pm 0.114 \times 1.96) \times 2) = 0.617, 0.252$$

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%
切片[poor]	-1.5654374	0.1600141	95.71	<.0001*	.
切片[moderate]	0.21295885	0.1259267	2.86	0.0908	.
ESTROGEN[ever]	-0.4645796	0.1142337	16.54	<.0001*	-0.6903971
AGE[65-79]	0.12277735	0.1170866	1.10	0.2944	-0.1046846

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
ESTROGEN	1	1	16.5398541	<.0001*
AGE	1	1	1.09956845	0.2944

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
ESTROGEN	1	1	16.9690855	<.0001*
AGE	1	1	1.11660886	0.2906

評価が完了しました。



Ordinal ロジスティック回帰分析 オッズ比 1

- ▶ JMP オッズ比と信頼区間は出力されない

Wald 法の手計算はできる

$$\text{AGE} \quad \exp(0.123 \times 2) = 1.279$$
$$\exp((0.123 \pm 0.117 \times 1.96) \times 2) = 2.023, 0.808$$

$$\text{ESTROGEN} \quad \exp(-0.465 \times 2) = 0.395$$
$$\exp((-0.465 \pm 0.114 \times 1.96) \times 2) = 0.617, 0.252$$

- ▶ R
- | | |
|----------|-----------------------|
| AGE | 1.278 (0.811 - 2.023) |
| ESTROGEN | 0.395 (0.251 - 0.616) |

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> ordinal.or.display(mo, decimal=4)

Re-fitting to get Hessian

Waiting for profiling to be done...

Re-fitting to get Hessian

      Ordinal OR lower95ci upper95ci P value
AGE65-79  1.2783    0.8111    2.0227  1.463e-01
ESTROGENever 0.3949    0.2514    0.6159  3.073e-05
>
>
>
```

Ordinal ロジスティック回帰分析 オッズ比 2

- ▶ $P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))}$
 $P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))}$
両式で β_{65-79} , β_{ever} は共通

Ordinal ロジスティック回帰分析 オッズ比 2

$$\blacktriangleright P(\text{GRADE} \geq \text{moderate}) = \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))}$$

$$P(\text{GRADE} \geq \text{poor}) = \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))}$$

両式で β_{65-79} , β_{ever} は共通

$$\blacktriangleright \text{例えば } P(\text{GRADE} \geq \text{moderate})$$

$$\begin{aligned} \text{オッズは } O(\text{GRADE} \geq \text{moderate}) &= \frac{P(\text{GRADE} \geq \text{moderate})}{1 - P(\text{GRADE} \geq \text{moderate})} \\ &= \exp(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}) \end{aligned}$$

Ordinal ロジスティック回帰分析 オッズ比 2

$$\begin{aligned} \blacktriangleright P(\text{GRADE} \geq \text{moderate}) &= \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))} \\ P(\text{GRADE} \geq \text{poor}) &= \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))} \end{aligned}$$

両式で β_{65-79} , β_{ever} は共通

$$\begin{aligned} \blacktriangleright \text{例えば } P(\text{GRADE} \geq \text{moderate}) \\ \text{オッズは } O(\text{GRADE} \geq \text{moderate}) &= \frac{P(\text{GRADE} \geq \text{moderate})}{1 - P(\text{GRADE} \geq \text{moderate})} \\ &= \exp(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}) \end{aligned}$$

▶ 例えば AGE 50-64 を reference とした 65-79 のオッズ比

$$\frac{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=50-64}}$$

(ESTROGEN の値は分子、分母で共通)

分子、分母の $\beta_{\text{ever}}\text{ESTROGEN}$ などの項は約分され消える

$$\frac{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=50-64}} = \begin{cases} \exp(2 \times \beta_{65-79}) & \text{JMP} \\ \exp(\beta_{65-79}) & \text{R} \end{cases}$$

Ordinal ロジスティック回帰分析 オッズ比 2

$$\begin{aligned} \text{▶ } P(\text{GRADE} \geq \text{moderate}) &= \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))} \\ P(\text{GRADE} \geq \text{poor}) &= \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))} \end{aligned}$$

両式で β_{65-79} , β_{ever} は共通

$$\begin{aligned} \text{▶ 例えば } P(\text{GRADE} \geq \text{moderate}) \\ \text{オッズは } O(\text{GRADE} \geq \text{moderate}) &= \frac{P(\text{GRADE} \geq \text{moderate})}{1 - P(\text{GRADE} \geq \text{moderate})} \\ &= \exp(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}) \end{aligned}$$

▶ 例えば AGE 50-64 を reference とした 65-79 のオッズ比

$$\frac{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=50-64}}$$

(ESTROGEN の値は分子、分母で共通)

分子、分母の $\beta_{\text{ever}}\text{ESTROGEN}$ などの項は約分され消える

$$\frac{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=50-64}} = \begin{cases} \exp(2 \times \beta_{65-79}) & \text{JMP} \\ \exp(\beta_{65-79}) & \text{R} \end{cases}$$

$$\text{▶ 同様に } \frac{O(\text{GRADE} \geq \text{poor})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{poor})|_{\text{AGE}=50-64}} = \begin{cases} \exp(2 \times \beta_{65-79}) & \text{JMP} \\ \exp(\beta_{65-79}) & \text{R} \end{cases}$$

Ordinal ロジスティック回帰分析 オッズ比 2

$$\begin{aligned} P(\text{GRADE} \geq \text{moderate}) &= \frac{1}{1 + \exp(-(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))} \\ P(\text{GRADE} \geq \text{poor}) &= \frac{1}{1 + \exp(-(\alpha_{\text{poor}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN}))} \end{aligned}$$

両式で β_{65-79} , β_{ever} は共通

▶ 例えば $P(\text{GRADE} \geq \text{moderate})$
オッズは $O(\text{GRADE} \geq \text{moderate}) = \frac{P(\text{GRADE} \geq \text{moderate})}{1 - P(\text{GRADE} \geq \text{moderate})}$
 $= \exp(\alpha_{\text{moderate}} + \beta_{65-79}\text{AGE} + \beta_{\text{ever}}\text{ESTROGEN})$

▶ 例えば AGE 50-64 を reference とした 65-79 のオッズ比

$$\frac{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=50-64}}$$

(ESTROGEN の値は分子、分母で共通)

分子、分母の $\beta_{\text{ever}}\text{ESTROGEN}$ などの項は約分され消える

$$\frac{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{moderate})|_{\text{AGE}=50-64}} = \begin{cases} \exp(2 \times \beta_{65-79}) & \text{JMP} \\ \exp(\beta_{65-79}) & \text{R} \end{cases}$$

▶ 同様に $\frac{O(\text{GRADE} \geq \text{poor})|_{\text{AGE}=65-79}}{O(\text{GRADE} \geq \text{poor})|_{\text{AGE}=50-64}} = \begin{cases} \exp(2 \times \beta_{65-79}) & \text{JMP} \\ \exp(\beta_{65-79}) & \text{R} \end{cases}$

▶ モデル全体で、オッズ比は AGE で 1 個、ESTROGEN で 1 個

Polytomous ロジスティック回帰分析 モデル式

- ▶ 目的変数 **SUBTYPE**
(adenocarcinoma (0), adenosquamous (1), other (2))

Polytomous ロジスティック回帰分析 モデル式

- ▶ 目的変数 **SUBTYPE**
(adenocarcinoma (0), adenosquamous (1), other (2))

- ▶ パラメータ

切片 α_1 α_2

係数 $\beta_{1, 65-79}$ $\beta_{2, 65-79}$ $\beta_{1, \text{ever}}$ $\beta_{2, \text{ever}}$



$$h_1 = \alpha_1 + \beta_{1, 65-79}\text{AGE} + \beta_{1, \text{ever}}\text{ESTROGEN}$$

$$h_2 = \alpha_2 + \beta_{2, 65-79}\text{AGE} + \beta_{2, \text{ever}}\text{ESTROGEN}$$

$$P(\text{SUBTYPE} = 0) = \frac{1}{1 + \exp(h_1) + \exp(h_2)}$$

$$P(\text{SUBTYPE} = 1) = \frac{\exp(h_1)}{1 + \exp(h_1) + \exp(h_2)}$$

$$P(\text{SUBTYPE} = 2) = \frac{\exp(h_2)}{1 + \exp(h_1) + \exp(h_2)}$$

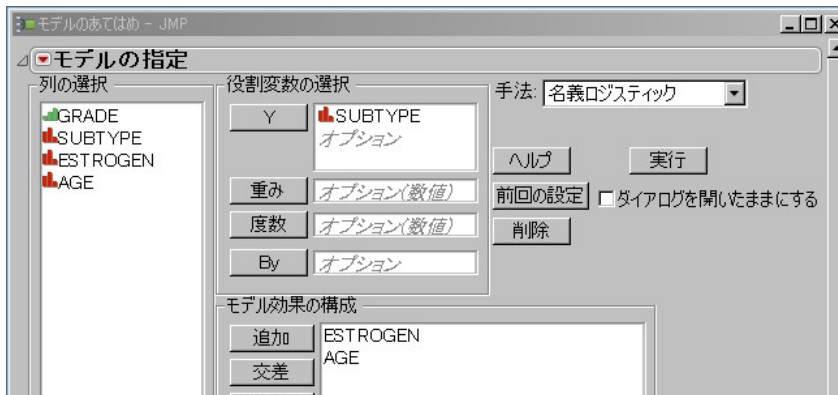
- ▶ 「分析」の「モデルのあてはめ」

The screenshot shows the JMP software interface with the 'Analyze' menu open. The 'Fit Model' option is highlighted. The data table below shows the following columns and rows:

	GRADE	SUBTYPE	ESTROGEN	AGE
1	moderate	adenocarcinoma	never	65-7
2	well	adenocarcinoma	never	65-7
3	moderate	adenosquamous	never	65-7
4	well	adenocarcinoma	never	65-7
5	well	adenocarcinoma	ever	50-6
6	well	adenocarcinoma	ever	65-7
7	well	other	ever	65-7
8	moderate	adenocarcinoma	never	50-6
9	well	adenocarcinoma	ever	65-7

Polytomous ロジスティック回帰分析 JMP 計算

- ▶ 「分析」の「モデルのあてはめ」
- ▶ Y に **SUBTYPE**、「モデル効果の構成」に **ESTROGEN, AGE**
手法は「名義ロジスティック」



Polytomous ロジスティック回帰分析 JMP 計算

- ▶ 「分析」の「モデルのあてはめ」
- ▶ Y に **SUBTYPE**、「モデル効果の構成」に **ESTROGEN, AGE**
手法は「名義ロジスティック」
- ▶ 赤い三角メニューから、
「**Wald 検定**」、「**信頼区間**」、「**確率の計算式の保存**」を選ぶ
注. メニューに「オッズ比」が無い

名義ロジスティックのあてはめ SUBTYPE

学習 定義

0180 $1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$

0380 $(1 - (L(0) / L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$

8801 $\sum -\text{Log}(\rho [j]) / n$

5614 $\sqrt{\sum (y[j] - \rho [j])^2 / n}$

5122 $\sum |y[j] - \rho [j]| / n$

3566 $\sum (\rho [j] \neq \rho \text{ Max}) / n$

286 n

LOF

要因	自由度	(-1)*対数尤度	カイ2乗
あてはまりの悪さ(LOF)	2	0.34482	0.689635
飽和モデル	6	251.36210	p値(Prob)ChiSq
あてはめたモデル	4	251.70691	0.7083

- ▶ **multinom** 関数で計算
出力を例えば **mp (multivariate polytomous)** と名付ける
`mp<-multinom(SUBTYPE~AGE+ESTROGEN)`

Polytomous ロジスティック回帰分析 R 計算

- ▶ **multinom** 関数で計算
出力を例えば **mp (multivariate polytomous)** と名付ける

```
mp<-multinom(SUBTYPE~AGE+ESTROGEN)
```

- ▶ 切片や係数の出力

```
summary(mp)
```

- ▶ **multinom** 関数で計算
出力を例えば **mp (multivariate polytomous)** と名付ける
`mp<-multinom(SUBTYPE~AGE+ESTROGEN)`
- ▶ 切片や係数の出力
`summary(mp)`
- ▶ 尤度比検定の出力
`Anova(mp)` #Anova は `car library` の関数

- ▶ **multinom** 関数で計算
出力を例えば **mp (multivariate polytomous)** と名付ける

```
mp<-multinom(SUBTYPE~AGE+ESTROGEN)
```

- ▶ 切片や係数の出力

```
summary(mp)
```

- ▶ 尤度比検定の出力

```
Anova(mp) #Anova は car library の関数
```

- ▶ オッズ比の出力

```
mlogit.display(mp,decimal=4)
```

Polytomous ロジスティック回帰分析 パラメータ

$$h_1 = \alpha_1 + \beta_{1, 65-79} \text{AGE} + \beta_{1, \text{ever}} \text{ESTROGEN}$$

$$h_2 = \alpha_2 + \beta_{2, 65-79} \text{AGE} + \beta_{2, \text{ever}} \text{ESTROGEN}$$

	α_1	α_2	$\beta_{1, 65-79}$	$\beta_{2, 65-79}$	$\beta_{1, \text{ever}}$	$\beta_{2, \text{ever}}$
JMP	-1.556	-1.226	0.390	0.220	-0.312	-0.061
R	-1.633	-1.385	0.780	0.440	-0.624	-0.122

名義ロジスティックのあてはめ SUBTYPE

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上
切片	-1.2255827	0.1613709	57.68	<.0001*	-1.5537959	-0.91
ESTROGEN[ever]	-0.0610699	0.1525905	0.16	0.6890	-0.3603726	0.239
AGE[65-79]	0.22018174	0.1609853	1.87	0.1714	-0.0894984	0.54
切片	-1.5556874	0.1912813	66.15	<.0001*	-1.953664	-1.19
ESTROGEN[ever]	-0.3121876	0.1706212	3.35	0.0673	-0.653147	0.019
AGE[65-79]	0.38999775	0.1898368	4.22	0.0399*	0.0312857	0.78

推定値は次の対数オッズに対するものです:

other/adenocarcinoma,

adenosquamous/adenocarcinoma

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
ESTROGEN	2	2	0.04707000	0.1075

Polytomous ロジスティック回帰分析 パラメータ

$$h_1 = \alpha_1 + \beta_{1, 65-79} \text{AGE} + \beta_{1, \text{ever}} \text{ESTROGEN}$$

$$h_2 = \alpha_2 + \beta_{2, 65-79} \text{AGE} + \beta_{2, \text{ever}} \text{ESTROGEN}$$

	α_1	α_2	$\beta_{1, 65-79}$	$\beta_{2, 65-79}$	$\beta_{1, \text{ever}}$	$\beta_{2, \text{ever}}$
JMP	-1.556	-1.226	0.390	0.220	-0.312	-0.061
R	-1.633	-1.385	0.780	0.440	-0.624	-0.122

```
> summary(mp)
```

```
Call:
```

```
multinom(formula = SUBTYPE ~ AGE + ESTROGEN)
```

```
Coefficients:
```

```
(Intercept) AGE65-79 ESTROGENever  
adenosquamous -1.633484 0.7799977 -0.6244115  
other -1.384702 0.4403656 -0.1221339
```

```
Std. Errors:
```

```
(Intercept) AGE65-79 ESTROGENever  
adenosquamous 0.3575824 0.3796739 0.3412432  
other 0.3116559 0.3219708 0.3051813
```

```
Residual Deviance: 503.4138
```

```
AIC: 515.4138
```

```
> |
```

Polytomous ロジスティック回帰分析 確率 1

出力された切片、係数から **SUBTYPE** の確率を手計算
まず h_1, h_2 を求める

No.2 のデータを使う $AGE[2] = 0, ESTROGEN[2] = 1$

$$h_1 = \alpha_1 + \beta_{1, 65-79}AGE + \beta_{1, ever}ESTROGEN$$

$$h_2 = \alpha_2 + \beta_{2, 65-79}AGE + \beta_{2, ever}ESTROGEN$$

	α_1	α_2	$\beta_{1, 65-79}$	$\beta_{2, 65-79}$	$\beta_{1, ever}$	$\beta_{2, ever}$
JMP	-1.556	-1.226	0.390	0.220	-0.312	-0.061
R	-1.633	-1.385	0.780	0.440	-0.624	-0.122

▶ **JMP** $AGE = -1, ESTROGEN = 1$

$$h_1 = -1.556 - 0.390 - 0.312 = -2.258,$$

$$h_2 = -1.226 - 0.220 - 0.061 = -1.507$$

出力された切片、係数から **SUBTYPE** の確率を手計算
まず h_1, h_2 を求める

No.2 のデータを使う $AGE[2] = 0, ESTROGEN[2] = 1$

$$h_1 = \alpha_1 + \beta_{1, 65-79}AGE + \beta_{1, ever}ESTROGEN$$

$$h_2 = \alpha_2 + \beta_{2, 65-79}AGE + \beta_{2, ever}ESTROGEN$$

	α_1	α_2	$\beta_{1, 65-79}$	$\beta_{2, 65-79}$	$\beta_{1, ever}$	$\beta_{2, ever}$
JMP	-1.556	-1.226	0.390	0.220	-0.312	-0.061
R	-1.633	-1.385	0.780	0.440	-0.624	-0.122

- ▶ **JMP** $AGE = -1, ESTROGEN = 1$

$$h_1 = -1.556 - 0.390 - 0.312 = -2.258,$$

$$h_2 = -1.226 - 0.220 - 0.061 = -1.507$$

- ▶ **R** $AGE = 0, ESTROGEN = 1$

$$h_1 = -1.633 - 0.624 = -2.257,$$

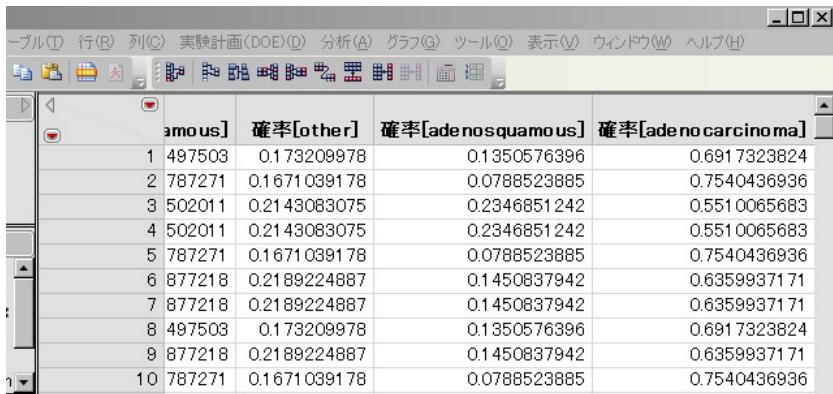
$$h_2 = -1.385 - 0.122 = -1.507$$

Polytomous ロジスティック回帰分析 確率 2

No.2 の SUBTYPE の確率計算の続き $h_1 = -2.257, h_2 = -1.507$

$$P(\text{SUBTYPE} = \text{adenocarcinoma}) = \frac{1}{1 + \exp(-2.257) + \exp(-1.507)} = 0.754$$
$$P(\text{SUBTYPE} = \text{adenosquamous}) = \frac{\exp(-2.257)}{1 + \exp(-2.257) + \exp(-1.507)} = 0.079$$
$$P(\text{SUBTYPE} = \text{other}) = \frac{\exp(-1.507)}{1 + \exp(-2.257) + \exp(-1.507)} = 0.167$$

▶ JMP の場合



The screenshot shows the JMP software interface with a data table. The table has four columns: 'amous', '確率[other]', '確率[adenosquamous]', and '確率[adeno carcinoma]'. The rows contain numerical data for 10 different cases.

	amous]	確率[other]	確率[adenosquamous]	確率[adeno carcinoma]
1	497503	0.173209978	0.1350576396	0.6917323824
2	787271	0.1671039178	0.0788523885	0.7540436936
3	502011	0.2143083075	0.2346851242	0.5510065683
4	502011	0.2143083075	0.2346851242	0.5510065683
5	787271	0.1671039178	0.0788523885	0.7540436936
6	877218	0.2189224887	0.1450837942	0.6359937171
7	877218	0.2189224887	0.1450837942	0.6359937171
8	497503	0.173209978	0.1350576396	0.6917323824
9	877218	0.2189224887	0.1450837942	0.6359937171
10	787271	0.1671039178	0.0788523885	0.7540436936

Polytomous ロジスティック回帰分析 確率 2

No.2 の SUBTYPE の確率計算の続き $h_1 = -2.257, h_2 = -1.507$

$$P(\text{SUBTYPE} = \text{adenocarcinoma}) = \frac{1}{1 + \exp(-2.257) + \exp(-1.507)} = 0.754$$
$$P(\text{SUBTYPE} = \text{adenosquamous}) = \frac{\exp(-2.257)}{1 + \exp(-2.257) + \exp(-1.507)} = 0.079$$
$$P(\text{SUBTYPE} = \text{other}) = \frac{\exp(-1.507)}{1 + \exp(-2.257) + \exp(-1.507)} = 0.167$$

▶ JMP の場合

▶ R の場合

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> AGE[2]
[1] 50-64
Levels: 50-64 65-79
> ESTROGEN[2]
[1] ever
Levels: never ever
> fitted(mp)[2,1]
[1] 0.7540452
> fitted(mp)[2,2]
[1] 0.07885078
> fitted(mp)[2,3]
[1] 0.1671041
>
>
```

Polytomous ロジスティック回帰分析 検定

$\beta_{1, 65-79} = \beta_{2, 65-79} = 0$ の検定、 $\beta_{1, ever} = \beta_{2, ever} = 0$ の検定

▶ JMP の場合

Wald 検定 AGE $p = 0.0731$ ESTROGEN $p = 0.188$

尤度比検定 AGE $p = 0.0640$ ESTROGEN $p = 0.182$

cancer - 名義ロジスティックのあてはめ - JMP

名義ロジスティックのあてはめ SUBTYPE

パラメータ推定値
推定値は次の対数オッズに対するものです:
other/adenocarcinoma,
adenosquamous/adenocarcinoma

推定値の共分散

効果に対するWald検定

要因	パラメータ数	自由度	Waldカイ2乗	p値(Prob>ChiSq)
ESTROGEN	2	2	3.34787989	0.1875
AGE	2	2	5.23081563	0.0731

効果の尤度比検定

要因	パラメータ数	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
ESTROGEN	2	2	3.41091646	0.1817
AGE	2	2	5.49832964	0.0640

Polytomous ロジスティック回帰分析 検定

$\beta_{1, 65-79} = \beta_{2, 65-79} = 0$ の検定、 $\beta_{1, ever} = \beta_{2, ever} = 0$ の検定

▶ JMP の場合

Wald 検定 AGE $p = 0.0731$ ESTROGEN $p = 0.188$

尤度比検定 AGE $p = 0.0640$ ESTROGEN $p = 0.182$

▶ R の場合

Wald 検定 計算されない

尤度比検定 AGE $p = 0.0640$ ESTROGEN $p = 0.182$

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> Anova(mp)
Analysis of Deviance Table (Type II tests)

Response: SUBTYPE
          LR Chisq Df Pr(>Chisq)
AGE          5.4983  2  0.06398 .
ESTROGEN     3.4109  2  0.18169
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

>
>
>
>
>
```

Polytomous ロジスティック回帰分析 オッズ比 1

JMP オッズ比と信頼区間は出力されないが Wald 法の手計算はできる

- AGE**
- $\exp(0.390 \times 2) = 2.181$
 $\exp((0.390 \pm 0.190 \times 1.96) \times 2) = 4.594, 1.036$
 - $\exp(0.220 \times 2) = 1.553$
 $\exp((0.220 \pm 0.161 \times 1.96) \times 2) = 2.919, 0.826$
- ESTROGEN**
- $\exp(-0.312 \times 2) = 0.536$
 $\exp((-0.312 \pm 0.171) \times 2) = 1.047, 0.274$
 - $\exp(-0.061 \times 2) = 0.885$
 $\exp((-0.061 \pm 0.153 \times 1.96) \times 2) = 1.612, 0.486$

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)	下側95%	上
切片	-1.2255827	0.1613709	57.68	<.0001*	-1.5537959	-0.91
ESTROGEN[ever]	-0.0610699	0.1525905	0.16	0.6890	-0.3603726	0.23
AGE[65-79]	0.22018174	0.1609853	1.87	0.1714	-0.0894984	0.5
切片	-1.5556874	0.1912813	66.15	<.0001*	-1.953664	-1.1
ESTROGEN[ever]	-0.3121876	0.1706212	3.35	0.0673	-0.653147	0.01
AGE[65-79]	0.38999775	0.1898368	4.22	0.0399*	0.0312857	0.7

推定値は次の対数オッズに対するものです:

other/adenocarcinoma,
adenosquamous/adenocarcinoma

Polytomous ロジスティック回帰分析 オッズ比 2

	AGE	adenosquamous	2.182 (1.037 – 4.591)
		other	1.553 (0.826 – 2.920)
R	ESTROGEN	adenosquamous	0.536 (0.274 – 1.045)
		other	0.885 (0.487 – 1.610)

```
R Console
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

> mlogit.display(mp, decimal=4)

Outcome =SUBTYPE; Referent group = adenocarcinoma
      adenosquamous                other
      Coeff./SE                Coeff./SE
(Intercept) -1.6335/0.35758*** - -1.3847/0.31166***
AGE65-79     0.78/0.37967*      2.1815 (1.0365, 4.5912) 0.4404/0.32197
ESTROGENever -0.6244/0.34124      0.5356 (0.2744, 1.0454) -0.1221/0.30518

      RRR (95%CI)
(Intercept) -
AGE65-79     1.5533 (0.8264, 2.9195)
ESTROGENever 0.885 (0.4866, 1.6096)

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

AGE の 50-64 を reference とする 65-79 のオッズ比を考える
SUBTYPE (adenocarcinoma (0), adenosquamous (1), other (2))

▶ $P(\text{SUBTYPE} = 1) = \frac{\exp(h_1)}{1 + \exp(h_1) + \exp(h_2)}$ が $\frac{1}{1 + \exp(-z)}$ の形でないので

$P = \frac{1}{1 + \exp(-z)} \Leftrightarrow O = \exp(z)$ が適用できず

オッズ $O(\text{SUBTYPE} = 1)$ や

オッズ比 $\frac{O(\text{SUBTYPE}=1|\text{AGE}=65-79)}{O(\text{SUBTYPE}=1|\text{AGE}=50-64)}$ は複雑な式になる

AGE の 50-64 を reference とする 65-79 のオッズ比を考える
 SUBTYPE (adenocarcinoma (0), adenosquamous (1), other (2))

- ▶ $P(\text{SUBTYPE} = 1) = \frac{\exp(h_1)}{1 + \exp(h_1) + \exp(h_2)}$ が $\frac{1}{1 + \exp(-z)}$ の形でないので

$$P = \frac{1}{1 + \exp(-z)} \Leftrightarrow O = \exp(z) \text{ が適用できず}$$

オッズ $O(\text{SUBTYPE} = 1)$ や

オッズ比 $\frac{O(\text{SUBTYPE}=1|\text{AGE}=65-79)}{O(\text{SUBTYPE}=1|\text{AGE}=50-64)}$ は複雑な式になる

- ▶ オッズ $O(\text{SUBTYPE} = 1) = \frac{P(\text{SUBTYPE}=1)}{1 - P(\text{SUBTYPE}=1)}$
 $= \frac{P(\text{SUBTYPE}=1)}{P(\text{SUBTYPE}=0) + P(\text{SUBTYPE}=2)}$
 $= \frac{\exp(h_1)}{1 + \exp(h_2)}$
 $= \frac{\exp(\alpha_1 + \beta_1, 65-79 \text{ AGE} + \beta_1, \text{ ever ESTROGEN})}{1 + \exp(\alpha_2 + \beta_2, 65-79 \text{ AGE} + \beta_2, \text{ ever ESTROGEN})}$

複雑な式である

$$P(\text{SUBTYPE} = 0) = \frac{1}{1 + \exp(h_1) + \exp(h_2)}$$

$$P(\text{SUBTYPE} = 1) = \frac{\exp(h_1)}{1 + \exp(h_1) + \exp(h_2)}$$

$$P(\text{SUBTYPE} = 2) = \frac{\exp(h_2)}{1 + \exp(h_1) + \exp(h_2)}$$

オッズ $O(\text{SUBTYPE} = 1) = \frac{P(\text{SUBTYPE}=1)}{P(\text{SUBTYPE}=0) + P(\text{SUBTYPE}=2)}$

$$O(\text{SUBTYPE} = 2) = \frac{P(\text{SUBTYPE}=2)}{P(\text{SUBTYPE}=0) + P(\text{SUBTYPE}=1)}$$

- ▶ SUBTYPE で値 0 (adenocarcinoma) を reference として odds-like expression \tilde{O} を定義する

$$\tilde{O}(\text{SUBTYPE} = 1) = \frac{P(\text{SUBTYPE}=1)}{P(\text{SUBTYPE}=0)} = \exp(h_1)$$

$$\tilde{O}(\text{SUBTYPE} = 2) = \frac{P(\text{SUBTYPE}=2)}{P(\text{SUBTYPE}=0)} = \exp(h_2)$$

$$P(\text{SUBTYPE} = 0) = \frac{1}{1 + \exp(h_1) + \exp(h_2)}$$

$$P(\text{SUBTYPE} = 1) = \frac{\exp(h_1)}{1 + \exp(h_1) + \exp(h_2)}$$

$$P(\text{SUBTYPE} = 2) = \frac{\exp(h_2)}{1 + \exp(h_1) + \exp(h_2)}$$

オッズ $O(\text{SUBTYPE} = 1) = \frac{P(\text{SUBTYPE}=1)}{P(\text{SUBTYPE}=0)+P(\text{SUBTYPE}=2)}$
 $O(\text{SUBTYPE} = 2) = \frac{P(\text{SUBTYPE}=2)}{P(\text{SUBTYPE}=0)+P(\text{SUBTYPE}=1)}$

- ▶ SUBTYPE で値 0 (adenocarcinoma) を reference として odds-like expression \tilde{O} を定義する

$$\tilde{O}(\text{SUBTYPE} = 1) = \frac{P(\text{SUBTYPE}=1)}{P(\text{SUBTYPE}=0)} = \exp(h_1)$$

$$\tilde{O}(\text{SUBTYPE} = 2) = \frac{P(\text{SUBTYPE}=2)}{P(\text{SUBTYPE}=0)} = \exp(h_2)$$

- ▶ [Kleinbaum, Klein], p.436,
Polytomous ロジスティック回帰分析では
odds-like expression をオッズと呼んでしまう

- ▶ odds-like expression $\tilde{O}(\text{SUBTYPE} = 1) = \exp(h_1)$
ここに、 $h_1 = \alpha_1 + \beta_{1, 65-79}\text{AGE} + \beta_{1, \text{ever}}\text{ESTROGEN}$

- ▶ odds-like expression $\tilde{O}(\text{SUBTYPE} = 1) = \exp(h_1)$
ここに、 $h_1 = \alpha_1 + \beta_{1, 65-79}\text{AGE} + \beta_{1, \text{ever}}\text{ESTROGEN}$
- ▶ AGE の 50-64 を reference とする 65-79 のオッズ比

ratio of odds-like expression

$$\frac{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=65-79)}{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=50-64)}$$

(分子、分母の **ESTROGEN** の値は共通とする)

- ▶ odds-like expression $\tilde{O}(\text{SUBTYPE} = 1) = \exp(h_1)$
ここに、 $h_1 = \alpha_1 + \beta_{1, 65-79}\text{AGE} + \beta_{1, \text{ever}}\text{ESTROGEN}$
- ▶ AGE の 50-64 を reference とする 65-79 のオッズ比

ratio of odds-like expression

$$\frac{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=65-79)}{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=50-64)}$$

(分子、分母の ESTROGEN の値は共通とする)

- ▶ JMP 50-64 のとき AGE = -1, 65-79 のとき AGE = 1
$$\begin{aligned}\frac{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=65-79)}{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=50-64)} &= \frac{\exp(\alpha_1 + \beta_{1, 65-79} + \beta_{1, \text{ever}}\text{ESTROGEN})}{\exp(\alpha_1 - \beta_{1, 65-79} + \beta_{1, \text{ever}}\text{ESTROGEN})} \\ &= \exp(2 \times \beta_{1, 65-79})\end{aligned}$$

- ▶ odds-like expression $\tilde{O}(\text{SUBTYPE} = 1) = \exp(h_1)$
ここに、 $h_1 = \alpha_1 + \beta_{1, 65-79}\text{AGE} + \beta_{1, \text{everESTROGEN}}$
- ▶ AGE の 50-64 を reference とする 65-79 のオッズ比

ratio of odds-like expression

$$\frac{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=65-79)}{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=50-64)}$$

(分子、分母の ESTROGEN の値は共通とする)

- ▶ JMP 50-64 のとき AGE = -1, 65-79 のとき AGE = 1

$$\begin{aligned} \frac{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=65-79)}{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=50-64)} &= \frac{\exp(\alpha_1 + \beta_{1, 65-79} + \beta_{1, \text{everESTROGEN}})}{\exp(\alpha_1 - \beta_{1, 65-79} + \beta_{1, \text{everESTROGEN}})} \\ &= \exp(2 \times \beta_{1, 65-79}) \end{aligned}$$
- ▶ R 50-64 のとき AGE = 0, 65-79 のとき AGE = 1

$$\begin{aligned} \frac{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=65-79)}{\tilde{O}(\text{SUBTYPE}=1|\text{AGE}=50-64)} &= \frac{\exp(\alpha_1 + \beta_{1, 65-79} + \beta_{1, \text{everESTROGEN}})}{\exp(\alpha_1 + \beta_{1, \text{everESTROGEN}})} \\ &= \exp(\beta_{1, 65-79}) \end{aligned}$$

Ordinal または Polytomous ロジスティック回帰分析

結果

- ▶ **JMP, R** の解析は一致した。

Ordinal または Polytomous ロジスティック回帰分析

結果

- ▶ JMP, R の解析は一致した。
- ▶ Ordinal ロジスティック回帰分析では、
ESTROGEN 使用経験有りでは **GRADE** 良好という傾向が見られ
(オッズ比 **0.395 (0.251-0.616)**)
AGE は **GRADE** に影響しなかった
(オッズ比 **1.278 (0.811-2.023)**)

Ordinal または Polytomous ロジスティック回帰分析

結果

- ▶ **JMP, R** の解析は一致した。
- ▶ **Ordinal** ロジスティック回帰分析では、**ESTROGEN** 使用経験有りでは **GRADE** 良好という傾向が見られ
(オッズ比 **0.395 (0.251-0.616)**)
AGE は **GRADE** に影響しなかった
(オッズ比 **1.278 (0.811-2.023)**)
- ▶ **Polytomous** ロジスティック回帰分析では、**AGE** が高いと **adenocarcinoma** に比べて **adenosquamous** が多かった
(オッズ比 **2.182 (1.037-4.591)**)
ESTROGEN は **SUBTYPE** に影響しなかった
(オッズ比 対 **adenosquamous 0.536 (0.274-1.045)**,
対 **other 0.885 (0.487-1.610)**)