データファイルの読み込み 000000000 

## SAS Enterprise Guide によるデータ解析入門

長島健悟 城西大学 薬学部

#### Kengo NAGASHIMA

Laboratory of Biostatistics, Department of Parmaceutical Technochemistry, Josai University

#### 2011年11月15日

データファイルの読み込み 000000000 

# 本日の内容

- ・ はじめに (SAS / SAS Enterprise Guide について)
- データ入力と読み込み
- 分割表データの解析
- 連続データの解析



はじめに ●00000000000 データファイルの読み込み 000000000 

# SAS とは

SAS (Statistical Analysis System)

- ・統計解析ソフトウェアの最高峰, 127 ヶ国,約 50,000 機関で利用され, 世界でも利用者数が多い統計解析パッケージの一つ
- 教育研究機関, 製薬企業, 医療機関での導入実績が多い
- ・解析結果の信頼性は高いと認知されており、論文執筆の上でも役立つ
- ただし、自由自在に解析するためにはプログラムを覚えなくてはいけない…



データファイルの読み込み

分割表データの解析

連続データの解析

# SAS プログラムの例

189 proc datasets lib = work; delete temp; 190 ods exclude all: 191 proc genmod data = &data.; 192 where TissueArea ^= .: 193 class group; 194 model &y. = group &z. 195 %do j = 1 %to &vnum.; 196 &&vn&i. 197 %end: 198 / link = id dist = n type3; 199 ods output ModelFit = temp; 200 run: 201 data temp; set temp(obs=7 firstobs=7); keep value; 202 ods select all: 204 data &out.; 205 set %if "&i." ^= "1" %then &out.; temp; 206 %end: 207 data &out.; merge &out. allvarcomb; 208 proc datasets lib = work; delete allvarcomb info temp; 209 run: auit: 210 proc rank data = &out. out = &out.; 211 var value: ranks &out.rank: 212 proc sort: by &out.rank: 213 %mend modelselect2: 214 215 %modelselect2(Tb Sp Ob S BS N Ob B Pm ES BS Oc S BS N Oc B Pm, BV TV, bone2, aic1): 216 proc print data = aic1(obs=5); run; 217 %modelselect2(BV\_TV Ob\_S\_BS N\_Ob\_B\_Pm ES\_BS Oc\_S\_BS N\_Oc\_B\_Pm, Tb\_Sp, , bone2, alc2); 218 = proc print data = aic2(obs=5); run;

219 %modelselect2(BV TV Tb Sp N Ob B Pm ES BS Oc S BS N Oc B Pm, Ob S BS, bone2, aic3);

習得には非常に 時間がかかる・・・ はじめに 00●00000000 データファイルの読み込み 0000000000 

## SAS プログラムは大変すぎる...

- 統計解析を行う事が目的なのに、プログラム言語を覚えるために時間 を浪費しすぎるのはもったいない
- 統計解析を行うために学ぶべきことは、データを読むための力を付ける事(グラフ化)と、手法を適切に利用するための知識を付ける事(統計的推定など)と、ソフトウェアパッケージを使った解析方法

データファイルの読み込み 000000000 

#### SAS Enterprise Guide とは

- ・マウス操作だけで解析できる、Windows 用 SAS インターフェース
- プログラムを組む事なしに、SAS による信頼性の高い解析を実施で きる
- ・データを読み込み、手法、変数、オプションをクリックして実行
- ・ 実施した解析内容が記録され,結果のレポートとともに保存できる



はじめに 0000●00000 データファイルの読み込み 000000000 

#### SAS と SAS Enterprise Guide の違い

- SAS: SAS プログラムを自分で記述する
- Enterprise Guide: 操作にしたがって, SAS プログラムを生成してくれる



はじめに 00000●0000 データファイルの読み込み 000000000 

# SAS Enterprise Guideを用いた データ解析手順を理解し実施できること

はし	じめに	
00	0000	•00C

データファイルの読み込み 000000000 

# 起動方法

- ・ SAS Enterprise Guide が適切にインストールされた PC 上で
- スタートメニュー  $\rightarrow$  SAS  $\rightarrow$  Enterprise Guide 4.2 など

前 解析	•
🛅 SAS	🕨 🛅 IntrNet 🔹 🕨
🛅 Software	🕨 🛅 SAS 92 License Renewal & Utilities 🔹 🕨
🛅 アクセサリ	▶ 🥳 SAS 92 (英語 (DBCS))
🛅 スタートアップ	▶ 🥳 SAS 92 (日本語)
🛅 管理ツール	🕨 🔄 SAS GRAPH ODS Graphics Editor 92
🔀 Microsoft Office Excel 2007	ML Studio 32
🌲 Eclipse	👩 Enterprise Guide 42
🦪 JGREP	
😕 Adobe Acrobat 9 Pro	
😕 Acrobat Distiller 9	
🔤 בידע אלעסר 🔤	
🔁 FileZilla	

データファイルの読み込み 000000000 

# 起動ダイアログ

- プロジェクトは、読み込むデータ、解析する方法とオプション、結果を ひとまとめに記録したファイル、行った解析を全て記録しておく事が できる
- ・プロジェクトを開く: 最近使ったプロジェクトが表示される
- 新規作成: 新規にプロジェクト等を作成する
- アシスタンス: チュートリアルを表示する

🐼 SAS Enterprise Guide へようこそ	<
開始するオプションを選択してください:	
プロジェクトを開く	.
🚔 その他のプロジェクト	
新規作成	.
🤯 新規プロジェクト	
💐 新規 SAS プログラム	
122 新しいデータ	
アシスタンス	
3 チュートリアル: SAS Enterprise Guide 入門ガイド	1
□ このウィンドウを次回から表示しない	//

は	じめに
oc	000000

データファイルの読み込み 0000000000 

## 初期画面

- ・ 左上: プロジェクト画面, 複数プロジェクトの切り替えなど
- ・ 左下: リスト, いくつかの操作リストが表示できる, 一番左のタスクリ ストを表示しておくと便利
- ・右:メイン画面,データや解析手法の設定,結果の表示など

SAS Enterprise Guide	
ファイル(E) 編集(E) 表示(V) タスク(E)	70854@ ୬−ルD ヘルプ⊕  ≜・26・54
プロジェクトツリー • ×	702270- •
Seg 70セスフロー	▶ 実行(B) - □ 停止(S)   エウスポード(2) - スケジュール(D) -   <b>ズーム(2) -   () フロジェクトログ(D</b>
J	F
9,791,71 · ×	
カテゴリ別タスク ・	
7-9	
す アイルタと並べ替え	
□□□ クエリビルダ	
135 テーブルの追加	
■ データの並べ替え	
語 出力形式の作成	
2 乾隆	
22 列の分割	
🖳 列の積み上げ	
	1
準備完了	× 未接货

は	じめに	
oc	0000000	ļ

データファイルの読み込み 0000000000 

# 終了方法

- 右上の [×] をクリック
- 未保存のファイル等がある場合警告が表示される
  - ・変更を保存して終了、保存せずに終了、またはキャンセル

TATURE AND AND THE AND	SAS Exterprise Oxide - 7022761aep			
	7741年日 編集(日 表行)(日 5725(8) 70556(日	2-MD ANX8 [월·글·월] 등 전 전 전 전 이 []·[뉴카88278- •		
	70719999- · × 70t270-	•		
	S kg 208220 - ► 309180 -	0 (FESS) 137#-HS - 2691-HD - X-623 - 378919HB933 270/1940 -		
	Concentration			
		一時子一致食むたいとかの常年		
Compare of the second sec		フロジェクト1 の定要を定定すが?		
Image: state in the s	<b></b> -	● 「ハイロフレクトでは、一般アーカビを使用していまった」というになった日本では、一般アール・人体を見たすべ (日本ではなく、一般アーカビネベビした)のペックの情報を引まれ、時期を参加すたには、日本のながたい(ない したしていたい)		
NONITION	× × ×	to the sector to the sector		
Signal 20         Procession         Processi	92.9/2.1 · ×			
(27-26.57      (27-26.57     (27-26.57     (27-26.57     (27-26.57	G 🖬 🛛 🚯	この7003の時間に下き開催に一時で一部を加えて2025日から開催に対応にた。一時で一部1時間時間できます。一時で一次は春日になっな火水がは、した日本1つ10万米~10万米~10万米~10万米~10万米~10万米~10万米~10万米~		
0         0	h+3,64525	21		
Restar         Baseline         Baseline         Baseline           P - State         P - State         Baseline         P - State           P - State         P - State         Baseline         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State           P - State         P - State         P - State         P - State	N 42 E	このプロジョンジャでの一時子一分を取りたと		
0         5-25           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           2         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26           3         5-26	S HONE	DATA1 2012320- Local WORK DATAS		
10         2-5-0           1         7-5-068           2         7-5-068           4         7-5-068           4         7-5-068           4         7-5-068           4         7-5-068           4         7-5-068           4         7-5-068           5         7-5-068           4         7-507           10         7-508           10         7-508           10         7-508           10         132-25           10         14-15	L TORNE			
P - Forders         -	N 7/2			
0	X 7-9502801			
0         7-50-007           0         7-50-007 </td <td>📅 7-0tri-06tt</td> <td></td> <td></td> <td></td>	📅 7-0tri-06tt			
0         -1	9 7-MILE			
G         Main Productions           J RD Production         J RD Production           J RD Production         J ROUTE           J ROUTE         J ROUTE	2 データセットと出た形式であります。			
a         Productorial           b         Productorial	0 9-11-A07-97+18-87+70-			
2 023 (2014)            2 024 (2014)            2 024 (2014)            2 024 (2014)            2 024 (2014)            2 024 (2014)            2 024 (2014)            2 024 (2014)            4 024 (2014)            4 024 (2014)	D BR THERE WAS A DECEMPTION			
2         200-2000/00           BAB         000           BAB         000-000           AVA (100-1000)         100-000           0000-000         100-000           0000	# \$P\$\$ 7r(10)(v#~)			
Display         Display <thdisplay< th=""> <th< td=""><td>Plana 2r(AD()/R=)</td><td></td><td></td><td></td></th<></thdisplay<>	Plana 2r(AD()/R=)			
10         1000000000000000000000000000000000000	2.982			
2. Enclair (10-f-cell) 2. Strain (10-f-cell) 3. Strain (10-f-cell)	1 92.8 (Part of the second sec			
	∑ BINEF® [0/7-Potem] SIGN/9-5 (A15)	115-42	( how	1 tolla
□ 0.0147+ © - 7-50896 ▲ 55 ■5275 - 2	3 SINE (0/9-F0088)	107.76	1.1.4	
iar 2 − Antel Ant iar βα mi − βαθαα	III CARLON-F			
A LOTAL A	6 9m			

データファイルの読み込み ●00000000 

## 読み込みの操作

[ファイル] → [開く] → [データ]

	🕑 SA	IS Enter	prise Guide	9						
l	77'	(JV( <u>F</u> )	編集( <u>E</u> )	表示(⊻)	タスク( <u>K</u>	)	プログ	∋ <u>ь(Р</u> )	ツール( <u>T</u> )	ヘルプ(
I	1	新規作	F成( <u>N</u> )			•	170	セスフロー	•	
i	â	開((0	)			۶.	(7	プロジェ	<u> ////////////////////////////////////</u>	rl+O
1		プロジ.	ェクトを閉じる	C			17	データ(	D	
1	<b>C</b>	プロジ	ェクト の保存	( <u>S</u> )	Ctrl+S	τ		プログラ	L(G)	
1		プロジ	ェクト に名前	を付けて保	存( <u>A</u> )			レポート	(D	
1		保存()	Ŵ	Ctrl+	Shift+S		阒	Stored	Process(S	)
1		名前を	付けて保存	( <u>E</u> )				OLAP :	キューブ(山)	
1	8	データ	のインボート(	p			73	Informa	ation Map(]	)
1	₽	プロジ	ェクト の実行	Û				Exchar	i∉e( <u>E</u> )	
1		プロジ	፤ ዕト ወスታን	フ <sub>ユ</sub> ール(Η)				OLEDE	(Q)	
1		エクスァ	#−H( <u>B</u> )			F	-	ODBC(	<u>B</u> )	
1		送信(	D			F		子()(約(	B)	
1	Ш.	パブリ:	シュ(し)				<b>1</b>	COIBS	Ly	
1	Ba	プロセ	スフロー のべ	ージ設定( <u>G</u>	i)					
ĥ		プロセ	λ70 − Ø€Γ	「刷ブレビュー	-(₩)					
	8	プロセ	スフロー のビ	□局(( <u>P</u> )	Ctrl+P					
		プロジ	ェクトプロパテ	-r(J)						
1						_				

データファイルの読み込み ○●○○○○○○○ 

## 読み込み対象ファイルの選択

データを選択 → [開く] (今回は tamo.xls (Excel ファイル) を指定)

🧀 データを開く	×
検索先(0)	🍺 proj 💿 🧔 🗢 🖻 🗮 🗸 🍃 🗐 🔹
	Chamelxis Itemoxifenxis
<u>.</u> 	
SAS フォルダ	
	ファイル名(N): tamoxifenxis 2
	ファイルの種類(I): すべての認識可能なデータファイル (*sas7bdat,*sas7bvew,*sd2,*mdb,*accdb,*dbf; 🔽
	【職K(Q) キャンセル

データファイルの読み込み 00●000000 

## 読み込み対象ファイルの確認

・[次へ] をクリック

🛃 tamoxifen xls 🎊	5のデータのインボート	×
<b>1</b> /4 ₹-	タの指定	<u>s</u> sas
データのインボートタ に変換します。	スクは、非 SAS データを、他のデータ分析タスクとレポーティングタスクで必	要な SAS データファイル
ソースデータファイル 保存場所(L):	□-カルファイルシステム	
ファイルパス(E) データ型(E): SAS 出力 デーカヤッ	OMemositenixte Excel ワークブック	
SAS Server(S) 54/75/(B)	Local SASUSER	参照())
データセット( <u>A</u> ):	tamoxifen2	
	〈 只ぷ(B) <b>* ( 浙へ(版) ) 完了(B) **</b>	ンセル ヘルブ

データファイルの読み込み 000●00000 

## ワークシートのチェックと設定

シート名の選択: [ワークシートを使用する] にチェックし, 解析対象のシート名を選択する

■ temovites via thèm==hmd^at=b.

- ・データの一行目が変数名の場合、チェックする
- ・[次へ] をクリック

<b>P</b>	) <b>(a) = (</b> * - )	■ tamoxifencels [正路モード] -	Microsoft Excel		2 / 4	データソースの選択	C	<u>S</u> sas
	赤ム 挿入	ページ レイアウト 放式 データ 校開	表示 開発 アドイン	Acrobet 🧐 – 🛡	一腑用の漢	IR V	(	
	Al	<ul> <li>・ (*) かいかいがいがいがいがいがいがいがいがいがいがいがいがいがいがいがいがいがいが</li></ul>			@ 17-55	/ートを使用する(W)	17 M	用の共同時についールド名が持入されているのの
- 1	A	D	0	D E	data			
1	リンハ助転移	タモキシフェン使用の有黒	乳かんの再発				E SAS	S市名規則に準拠するように列名を実更する
2	めり	40.9	めツ					
3	あり	あり	あり					
4	あり	あり	あり					
5	あり	89 8	ab 9					
6	あり	あり	あり					
7	あり	あり	あり			- たいートはいの本本でのたい 第二日	は他田はス	
8	あり	あり	あり		⊑ (s)	)	1610033.9.0	
9	あり	あり	あり		883	始-左セル(D)		
10	あり	あり	あり		12	7-521(1):		
11	あり	あり	あり		**C	1.00%		
12	あり	あり	あり			14YH	300世が(8)	
13	あり	あり	あり		<ul> <li>事前定</li> </ul>	職された名前付き範囲を使用	目する(P)	
14	あり	あり	あり		tamo	(data\$A1:C4902)		
15	あり	あり	あり					
16	あり	あり	あり					
17	あり	あり	あり					
18	あり	あり	あり				(3)	
19	あり	あり	あり				<u> </u>	
20	あり	あり	あり		r.	(夏内	(B) 🚽 🖂 (N) > 丨	第7(E) あめおも へます
H 4 777	F Mill data 🖉	7-400	Rh 3 (1911) 130K (-)			1 8484		100

データファイルの読み込み 000000000 

## 変数の確認と型変換

- 読み込まれた変数名の確認と、必要であれば文字型・数値型等を変更できる
- ・[次へ] をクリック

9	ホーム挿入	R-9 14701	敬武 データー	投稿 表示	開発	アドイン	Acrobet		
_	Al	• 0	たりンパ類院	移		_			_
	A		В		C		D	E	_
1	リンパ師伝行	5 タモキシフ	エン使用の有類	ま 男しわ	MOF	19 <u>8</u>			_
2	あり	859		あり		-			_
3	あり	あり		あり					_
4	あり	あり		あり					
5	あり	あり		あり					_
6	あり	あり		あり					_
7	あり	あり		あり					
8	あり	あり		あり					
9	あり	あり		あり					
10	あり	あり		あり					
11	あり	あり		あり					
12	あり	あり		あり					
13	あり	あり		あり					
14	あり	あり		あり					
15	あり	あり		あり					
16	あり	あり		あり					
17	あり	あり		あり					
18	あり	あり		あり					
19	あり	あり		あり					
	赤り	あり		あり					

l	la la	moxifen.xls #	ものデータのイ	ンポート					×
	3	/4 7r	ールド属性の	定義					<u>sas</u>
	nyas	選択と寄性の定	E募(S):						
	1000	ソース名	名前	5KU	種類	ソースの入 力形式	長さ	出力形式	入力形式
l		リンパ節転移	col1	リンパ管轄部	文字列	\$CHAR4.	4	\$CHAR4.	\$CHAR4.
l		ダモキシフェ	col2	タモキシフェン使用	文字列	\$CHAR4.	4	\$CHAR4.	\$CHAR4.
I	V	乳がんの再	col3	乳がんの再発	文字列	\$CHAR4.	4	\$CHAR4.	\$CHAR4.
			LUNCO 1						
	(1)	で選択	勿P©						変更(因)_
				< 戻る(目)  ▼ [])2	∧( <u>N</u> )> ]	売7(E)		×42.00%	へば

データファイルの読み込み 0000000000 

# 保存オプションと読み込み操作の完了

- 「生成した SAS コードにデータを埋め込む」にチェック
   ・埋め込んでおくと再解析しやすい
- 完了をクリック



は	じめに	
00	000000	boc

データファイルの読み込み 00000000000 

#### 読み込み結果

- 正常に読み込みが完了すると、以下のような画面が表示されます
- これで、Excel データの読み込み作業が終了しました

🐼 SAS Ente	rprise Guid	e							
ファイル(E)	編集( <u>E</u> )	表示(⊻)	タスク( <u>K</u> )	プログラ	ム(P) ツール(I)	ヘルプ(円)	🗎 • 🗳 • 🐔	📇 94 📭	$\mathbb{B} \times$
プロジェクトツリ	-		- ×	データ	(tamoxifen.xls[da	ita]) のインポー	<b>۲.</b>		
プロセスフロー					コード) 📋 Dグ	📰 出力データ	1		
SAS Enterpris     ファイル(E) 編     プロジェクトツリー     プロジェクトツー     プロシスロー     ゴーゴントンフロー     ゴーゴントンフロー     ゴーゴントンフロー     ゴーゴントンフロー     ゴーゴントンフロー     マーン (tai     スクリスト     マーコー ショ	n xls 7. (tamoviter	n vie[data])	のインポート	\$5	、 、 タスクの変更( <u>y</u> )	7771192:	並べ替え( <u>L</u> ) 🛄 ケ	エリビルダ(Q)	データ( <u>D</u> )
	(Cambxile)	12(10)(00(0))	012/6 1		💧 リンパ節転	移 🔌 矢モキ	シフェン使用の有無	🔌 乳がん	の再発
				1	あり	あり		あり	
				2	あり	あり		あり	
				3	あり	あり		あり	
				4	あり	あり		あり	
				5	あり	あり		あり	
				6	あり	あり		あり	
				7	あり	あり		あり	
				8	あり	あり		あり	
				9	あり	あり		あり	
				10	あり	あり		あり	
				11	あり	あり		あり	
1				12	あり	あり		あり	
				13	あり	あり		あり	
				14	あり	あり		あり	
タスクリスト			- X	15	あり	あり		あり	
	0		_	16	あり	あり		あり	
🕰 Li 💷	<b>(1)</b>			17	あり	あり		あり	

データファイルの読み込み 00000000000 

## メイン画面に戻る

データ (ファイル名.xls) のインポートと表示されている部分をクリック → [プロセスフロー]

🐼 SAS Enterprise G	uide					
ファイル(E) 編集(	E) 表示( <u>V</u> )	タスク( <u>K</u> )	プログラ	ム(P) ツール(T)		📬 🚳   🔒 🛩 🗈 🖄 🗙
プロジェクトツリー		- x	データ	(tamoxifenxls[da	ta]) のインポート 🔹	
プロセスフロー			géő	プロセスフロー		
🛐 tamoxifen.xls	ifen vis[data])	መረጉሥታ	\$5		売フィルタと並べ替	え(L) = クエリビルダ(Q)   データ(D)
- y comos	incirio (data)	012/01		💧 リンパ節転	移 🔌 タモキシフェン	使用の有無 📐 乳がんの再発
			1	あり	あり	あり
			2	あり	あり	あり
			3	あり	あり	あり
			4	あり	あり	あり
			5	あり	あり	あり
			6	あり	あり	あり
			7	あり	あり	あり
			8	あり	あり	あり
			9	あり	あり	あり
			10	あり	あり	あり
			11	あり	あり	あり
			12	あり	あり	あり
			13	あり	あり	あり
			14	あり	あり	あり
タスクリスト		- x	15	あり	あり	あり
			16	あり	あり	あり
🖳 La 📃 🕙			17	あり	あり	あり

データファイルの読み込み 00000000● 

## メイン画面で結果を確認

- Excel ファイルのデータを読み込んで変換し、
- SAS データセット (SAS で扱うことのできるデータ形式) に変換されているか確認できます
- 下図一番右のアイコンが表示されていれば OK

🐼 SAS Enterprise Guide	
ファイル(E) 編集(E) 表示(V) タスク(K)	プロጛ҄ラム(Ⴒ) ツール(D) ヘルフ(฿)   🗎 • 🝰 • 💁   📇 🌳 🗅 🛝 🗙
プロジェクトツリー • ×	プロセスフロー -
ダロセス20 著 Tamoxiferxls 一週 データ (tamoxiferxls[data]) のインボート	$ \begin{array}{c} \flat \hspace{0.5mm} \begin{array}{c} \flat \hspace{0.5mm} \begin{array}{c} \blacksquare \hspace{0.5mm} \left\langle \blacksquare \hspace{0.5mm} \left( \blacksquare $
۲ (۲۰۰۰) ۶۷۵/۱۵۶ - ۲ ۲۰۰۰ - ۲	

## 先ほど読み込んだ分割表データ<u>について</u>

 佐藤俊哉. 交絡事実と反事実の比較. 岩波科学 2008 年 4 月号. より 引用

1982年から1990年にかけて日本の9施設で手術を受けた乳がん患者 4901名について,手術後にタモキシフェンという女性ホルモンによく似 た乳がんの治療薬を使用したかどうかと,その後の乳がんの再発との関係 を1996年まで追跡して調べた結果の一部である

 ・ 変数: タモキシフェン使用の有無,乳がん再発の有無,リンパ節転移の 有無

データファイルの読み込み 000000000  

# 単変量解析

- 目的 タモキシフェン使用の有無によって、乳がんの再発割合に違いがある かどうかを検討する事
- 目的に対応する解析内容
   二元分割表の集計と、再発割合に違いがあるかどうかの仮説検定、再発割合の差の信頼区間によって評価を行う

データファイルの読み込み 000000000 分割表データの解析 ○○●○○○○○○○○○○○○○○○○○○ 

# 二元分割表の作成(1)

- [タスク] → [記述統計] → [分割表分析]
- ・ 変数 [タモキシフェン使用の有無] [乳がんの再発] を表変数に設定 (ドラッグ& ドロップ) → 表の設定

				Local SASUSER, TAM	0XFEN2 (0) 分割表分析1	×
<u>ታ</u> ዞ፣	:¢P			<b>F</b> - <b>3</b>	- データ	
女: 	(水気) プログラム(P) ツ データ(D) 記述統計(P) グラス(Q) 分散分析(A) 回帰分析(P) 冬麦量解析(M) 生存時間分析(S) 工程能力分析(f) 管理[D(C) 以, 」(F(R))	-JI(I		でいみは1年度 実統は1号 のまま しまま が良まプジョン 料量が付けてれた差 が自まプジョン 科集 それが日金の結果 それが日金の結果 がは1000年 クイトル プロパディ	データンー2: LocalSASUSERTAMONDEN2 23/27x1/2 だし SEED:01/0910 SEED:01/0910 SEED:02/04 (SEEN0.LR: SEED:02/04 (SEEN0.LR:	·編集(£)
5.	時系列分析(1) モデルのスコアリング(1)…	٠	Ⅲ 一元度致表(②) Ⅲ 分割表分析(A)		a	
	OLAP(Q)	•				
-	タスクテンプレート( <u>K</u> )	•			x	
					コンテキストヘルプを表示する役割を選択します。	×
				CO-Kのプレビュー(C)	天行(1) - 保存(1) 年代也は	~117
				[テーブル] ページで少なくと		.:

データファイルの読み込み 000000000  

# 二元分割表の作成(2)

- ・ 変数 [タモキシフェン使用の有無] を表の行に, [乳がんの再発] を表の列に設定します
- 右図の様になっていれば OK です



データファイルの読み込み

分割表データの解析 ○○○○●○○○○○○○○○○○○○○○○ 

# 二元分割表

- ・[行のパーセント], [セルの度数] にチェック → [実行]
- 二元分割表の結果が表示されます

🖬 Loo	calSASUSER TAMOX	FEN2 の 分割表分析1		×						
デ表	-9	セル統計量			SAS レポ	(ト 🕑 結果 - RTF				
起表		使用できる統計係			₹ポート(⊠	) • 送信( <u>N</u> ) • 作成( <u>C</u> ) • パブ	ゆシュ(D)   🖸 プロ/	রিন⊅		
	一致 順序付けされた養 純句検定 計算オプション	F 1507-世界版 F 1507-世界版 F 2507-世界版				5	制表分析			
864	* セル統計量の結果 表統計量の結果	F セルのパーセンド(E) F 欠損価の度数(y)					綶			
力	11/77 11/77	<ul> <li>「 セルビとの x2 乗続計量(L)</li> <li>「 セル成計の期待億からの爆差(L)</li> <li>「 期待たい成時(E)</li> </ul>				FRI	E <b>Q プロシジ</b> ャ			
Ŀ		■ 会社の社内パーセンド日 ■ 会社の社内パーセンドで含める(	9			+			<b>- - - -</b>	_
						表:タセキシノエン	12日の有黒 * 3	しかんの	円光	
								乳かんの	り円発	
								あり	なし	合計
						タモキシフェン使用の有無				
						<b>5</b> 0	度数	464	2085	2549
		出力に含めるセル統計量を選択します。		-		.00	行のパーセント	18.20	81.80	
		セルごとに行のパーセントを表示します。2		_		71.	度数	424	1928	2352
1.000	- 100-0112 (0)		anarcon I I marcon I	التعديد المعدية ا		40	行のパーセント	18.03	81.97	
	-1007/C140		MITES - 1994(2)	49200 000		合計	度数	888	4013	4901

再発割合

タモキシフェン使用 464/2549=18.2%, 非使用 424/2352=18.0%

データファイルの読み込み

分割表データの解析 ○○○○○●○○○○○○○○○○○○○○○ 

## 仮説検定

- プロセスフローに戻る → [分割表分析 1] を右クリック → [分割表分析 1 の変更]
- 表統計量] → [関連] → [関連性の検定] → [χ<sup>2</sup> 検定] → [実行] → 結果 の置き換え [はい]



データファイルの読み込み 000000000  

## 仮説検定の結果

 再発確率が異なるかどうかを、ピアソンのカイニ乗検定を用いて検定 する (α = 0.05 とする)

 $\begin{cases} H_0: タモキシフェン使用の有無で再発確率は変わらない \\ H_1: タモキシフェン使用の有無で再発確率が異なる \end{cases}$ 

統計量	自由度	値	p 値
力イ2乗値	1	0.0255	0.8730
尤度比力f2 乗値	1	0.0255	0.8730
連続性補正力イ2乗値	1	0.0151	0.9023
Mantel-Haenszel の力イ2 乗値	1	0.0255	0.8730
JPイ係数		0.0023	
一致係数		0.0023	
Cramer の V 統計量		0.0023	

Fisherの正確	剣定
セル (1,1) 度数 (F)	464
左側 Pr <= F	0.5779
右側 Pr >= F	0.4513
表の確率 (P)	0.0292
両側 Pr <= P	0.8820

*P*-value=0.873 であり, *a* = 0.05 のもとで帰無仮説は棄却されない

データファイルの読み込み 000000000  

# リスク差の推定

- プロセスフローに戻る → [分割表分析 1] を右クリック → [分割表分析 1 の変更]
- [表統計量] → [関連] → [関連性の指標] → [2×2表に対するリスクの 差] → [実行] → 結果の置き換え [はい]



データファイルの読み込み 0000000000  

# リスク差の推定結果

	列1リスクの推定値											
	リスク	漸近標準誤差	( <b>漸</b> 近) 信頼[	95% 区間	(直接確 信葉	率) 95% 限界						
行1	0.1820	0.0076	0.1671	0.1970	0.1672	0.1976						
行2	0.1803	0.0079	0.1647	0.1958	0.1649	0.1964						
合計	0.1812	0.0055	0.1704	0.1920	0.1705	0.1923						
差	0.0018	0.0110	-0.0198	0.0233								
		îт1-	行 2 のえ	Ê								

リスク差=0.0018 であり、95%信頼区間も0をまたぐため、再発確率の差はそれほど大きくない

データファイルの読み込み 000000000 分割表データの解析 ○○○○○○○○●○○○○○○○○○○○○ 

# 多変量解析

- 目的
  - リンパ節転移の有無を考慮した上で、タモキシフェン使用の有無に よって、乳がんの再発確率に違いがあるかどうかを検討する事
- 目的に対応する解析内容
   多元分割表の集計と、ロジスティック回帰モデルを用いた調整オッズ
   比の信頼区間を求めて評価を行う

データファイルの読み込み 000000000 分割表データの解析 ○○○○○○○○○●○○○○○○○○○○ 

# 多元分割表の作成(1)

- [タスク] → [記述統計] → [分割表分析]
- ・ 変数 [タモキシフェン使用の有無] [乳がんの再発] を表変数に, [リン パ節転移] をグループ変数に設定 (ドラッグ& ドロップ) → [表] の 設定





データファイルの読み込み 000000000 分割表データの解析 ○○○○○○○○○○○●○○○○○○○○ 

# 多元分割表の作成(2)

- ・ 変数 [タモキシフェン使用の有無] を表の行に, [乳がんの再発] を表の列に設定します
- [行のパーセント], [セルの度数] にチェック → [実行]

■LocalSASUSERTAMD/0FEN2 の分割表分析1	m Local SASUSER TAMOXIFEN2 の 分割表分析1	×
Provide and the second	CAREFA	
□□-F07122-Q 2010 Y (870) 49228 へルフ	□□□=F071/ピュー() 第(7(8) - (将符(5) 4+>とせん _	A\$7
ローフルドページでジルなども「シルケーフルを定義していたという		

データファイルの読み込み

 

## 多元分割表

'リン/	(節転移'n=あり			リンパ節転移	多'n=なし					
表:タモキシフェン	使用の有無 * 캵	しがんの	再発	表 : タモキシフェン 使用の有黒 * 乳がんの 再発						
	乳がんの	の再発			1	乳がんの	)再発			
	あり	なし	合計			あり	なし	合計		
タモキシフェン使用の有無					タモキシフェン使用の有無					
<b>*</b> n	度数	368	847	1215	tan 應	該数	96	1238	1334	
100	行のパーセント	30.29	69.71		<sup>のワ</sup> 行のパ	パーセント	7.20	92.80		
	度数	253	507	760	たし 歴	該数	171	1421	1592	
40	行のパーセント	33.29	66.71		ゆし 行のパ	パーセント	10.74	89.26		
合計	度数	621	1354	1975	合計 度	き数	267	2659	2926	

- リンパ節転移ありの再発割合 タモキシフェン使用 368/1215=30.3%, 非使用 253/760=33.3%
- ・リンパ節転移なしの再発割合 タモキシフェン使用 96/1334=7.2%, 非使用 171/1592=10.7%

データファイルの読み込み

 

## ロジスティック回帰モデルによる解析(1)

- [タスク] → [回帰分析] → [ロジスティック回帰分析]
- ・ 変数 [乳がんの再発] を従属変数に, [タモキシフェン使用の有無] [リンパ節転移] を分類変数に設定 (ドラッグ& ドロップ) → [モデル] → [効果] の設定



はじめに	
0000000000	

データファイルの読み込み 000000000  

## ロジスティック回帰モデルによる解析 (2)

- ・ 変数 [タモキシフェン使用の有無] [リンパ節転移] を選択し, [主効果]
   をクリックして効果を設定
- [モデル] → [オプション] の設定 → [条件付きオッズ比] → [Wald 検 定に基づく] にチェック → [実行]

	III LocalSASUSER.TAMO/UFEN2 の ロジスティック回帰分析	×
	データ モデル モデル マンチ	
		1
1000000000000000000000000000000000000	シイトム フレパラ・         モデルの含ては中の大学をみ作用 下 後回した。         このまたのでは、中の大学・シート についた。         このまたのでは、中の大学・シート についた。           ド None - Landon の (地名美国学校) ド オピアン 入おお (New York) の 22 条約日素(Q) ド ー 特社 FC 条(Q)         シントラーム ・シート ・シート ・シート         シントラーム ・シート ・シート ・シート	ي ت
	(1488) (155-5) (155-5) 「フランドル大型に基づ(E) 「 Wea時間に基づ(B) 「 Wea時間に基づ(B)	
「戸 切片を含める印     「戸 切片を含める印     」     、      、      、     、     、      、     、      、     、     、     、      、     、     、     、     、     、     、     、      、      、	(896/v94(E) 202 -	
E	モデルの当てはか手法 当てはか単注(40) (自動(1919年1.)	-
<u>→</u> →→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→	2	×
25		7
データファイルの読み込み 0000000000  

### ロジスティック回帰モデルの解析結果(1)

• ロジスティックモデル

logit(p乳がんの再発) = $\beta$ 切片 + $\beta$ リンパ節転移の有無 $x_1$ +  $\beta$ タモキシフェン使用の有無 $x_2$ 

上記のモデルを用いれば、リンパ節転移の有無で調整した、タモキシフェン使用の有無のオッズ比を求める事ができる



データファイルの読み込み 0000000000  

### ロジスティック回帰モデルの解析結果(2)

- タモキシフェン使用の効果についての推定値は -0.25,
   *P*-value=0.0017 であり, α = 0.05 のもとで帰無仮説は棄却された
- タモキシフェン使用の効果についての調整オッズ比は 0.781, 95%信頼区間は [0.669, 0.912] であった
- したがって、リンパ節転移を考慮するとタモキシフェン使用の有無で 再発オッズが異なる

データファイルの読み込み

 

### 単変量・多変量解析の結果のまとめ

タモキシフェン	再発あり	再発なし	対象者数		
使用	464 (18.2%)	2085	2549		
非使用	424 (18.0%)	1928	2352		
合計	888	4013	4901		
	再発割合の差= 0.2% [-0.02, 0.02]				

表 1. タモキシフェン使用と乳がんの再発

表 2. リンパ節転移で層別したタモキシフェン使用と乳がんの再発

タモキシ	リン	パ節転移あり	J	リンパ節転移なし			
フェン	再発あり	再発なし	対象者数	再発あり	再発なし	対象者数	
使用	368 (30.3%)	847	1215	96 (7.2%)	1238	1334	
非使用	253 (33.3%)	507	760	171 (10.7%)	1421	1592	
合計	621	1354	1975	267	2659	2926	
	再発害	合の差= - :	3%	再発割	<b>合の差=</b> -3	.5%	

調整オッズ比=0.781 [0.669, 0.912]

データファイルの読み込み 000000000 分割表データの解析 ○○○○○○○○○○○○○○○○○○○○○○ 

### 交絡因子

- 表2を横方向にみると…
  - タモキシフェン使用群, 1215:1334 でリンパ節転移あり (高リス ク集団) が多い
  - タモキシフェン非使用群, 760:1592 でリンパ節転移あり (高リス ク集団) が少ない
  - タモキシフェンに効果があるにも関わらず、高リスクの人が多いので、見かけ上効果が薄まって見えていた
- リンパ節転移の様な因子は交絡因子であり、交絡因子を調整しない推定結果は正しくない
- 多変量解析を用いて、交絡因子の状態を仮想的に同じとした場合の、 タモキシフェン使用の再発予防効果をみるべき

#### タモキシフェンの乳がん再発予防効果はある

データファイルの読み込み

分割表データの解析 ○○○○○○○○○○○○○○○○○○○○○ 

### プロジェクトの保存 / プロジェクトを開く

- [ファイル] → [プロジェクトに名前を付けて保存] で今まで行った全ての解析をプロジェクトに保存します
- [ファイル] → [開く] → [プロジェクト] で保存しておいた解析を開く ことができます

S one enciprise data
ファイル(E) 編集(E) 表示(V) タスグ(K) プログラム(E) ツール(T) ヘルブ(J
新規作成(N)
局(① ・ 「 て プロジェクト(P) OtrHO
ブロジェクトを開じる(公)         デーシ(D).           グロジェクトの採用(S)         CPH-S           ブロジェクトの採用(S)         レポート(D).           (第四)         CPH-S           (日本)
プロジェクトのあがりュール(出)     「ロシェクトのステジュール(出)     「ロシェクトのステジュール(出)     「ロシェクトのステジュール(出)     「ロショントのステント(出)     「ロショントのステント)     「ロショントのステント(出)     「ロショントのステント(出)     「ロショントのステント(出)     「ロショントのステント)     「ロショントのステントのステントのステントのステントのステントのステントのステントのステ
<ul> <li>プロセスフロ・のページ総定(3)</li> <li>プロセスフロ・の年間(7)レビューへ処</li> <li>プロセスフロ・の年間(1)レビューへ処</li> <li>プロセスフロ・の年間(1)レニー CH+P</li> </ul>
2

データファイルの読み込み

# 新規プロジェクトの追加

• [ファイル] → [新規作成] → [プロジェクト]

😿 SA	S Enterprise Guide		×
79-	イル(E) 編集(E) 表示(y) タスク(b)	プログラム(円) ツール(円) ヘルプ(円)	📔 • 🚰 • 🗳 🚨 🖗 🖻 🖄 🗙 🕬 🖉 • 🖆 •
1	新規作成(№)	🌝 プロジェクト(P)	
	開い(の) プロジェクトを開じる(C)	<ul> <li>デーダD)</li> <li>ズログラム(G)</li> </ul>	
<b>%</b>	プロジェクト の保存(S) Ctrl+S プロジェクト に名前を付けて保存(A)		
	保存(型) Ctrl+Shift+S  名前を付けて(保存(E)	<ol> <li>メモ(N)</li> <li>プロセスフロー(E)</li> </ol>	
Þ	テータロインボートロ プロジェクト の実行(U) プロジェクト のスケジュール(H)	」順序付けされたリスト(②)	
550	エクスポート(B) ・ 送信(D) ・ パブリッシュ(L)		
82a 144	プロセスフロー のページ設定(G) プロセスフロー の印刷ブレビュー(W) プロセスフロー の印刷(P) Cer+P		
	プロジェクトプロパティ(」)	-	
	1 projegp		
	終了(2)	•	×
準備研	ŧī		× 未接続

データファイルの読み込み 000000000 連続データの解析 ○●○○○○○○○○○○○○○○○

## 連続データの概要

- 2007年の都道府県別の人口 10万対病院数・歯科診療所数・一般診 療所数<sup>[4]</sup>と人口推計<sup>[5]</sup>
- 高齢者の割合が多い県と少ない県で分けた [高齢者] という変数が含 まれる
- 目的

高齢者の割合によって、人口 10 万対病院数・歯科診療所数・一般診 療所数は異なるかどうかを検討する

- 目的に対応する解析内容 要約統計量、ヒストグラムや箱ひげ図の作成と、2群の母平均に対す る仮説検定と信頼区間の計算
- ・新しく作成したプロジェクトで、データ [byouin.xls] を読み込みます
- スライド 13のデータ読み込みの手順を参照

データファイルの読み込み 000000000 連続データの解析 ○○●○○○○○○○○○○○○○○○

### 要約統計量の計算とヒストグラム・箱ひげ図の出力(1)

- [タスク] → [記述統計] → [要約統計量ウィザード]
- ・[次へ]

ファイル(E) 編集(E) 表示(V)	タスク	(E) プログラム(E) "	- IKI)	∿11700   🎦 • 🚔 • 强   📇	3
プロジェクトツリー		データ(D)	•		
		記述統計(图)	• 🗉	92HD-	Į.
🖻 🚮 byouinxis		グラフ( <u>G</u> )	1 2	要約統計量ウィザード(層)-	iĽ.
<ul> <li>一番 byouinxls からインボートされ</li> </ul>		分散分析(点)	+ Σ	要約統計量(S)	14
🍈 🏡 栗約統計量		回帰分析(B)	• 🗖	事計表ウィザード(B)	ı,
		冬麦量解析(图)	· ) 🚆	泰計表(T).	r
		生存時間分析(S)	· • 🚆	1011 L 8 10 46 100	÷
		工程能力分析(y)	•	テームの時後公共(1)	E
		管理图(C)	+ 100 16	分布(D)	E
	lí	パレート国(型)_		→元度敬表(0)_	E
		時系列分析(T)	• •	分割表分析(A)_	E
	η.	モデルのスコアリング(L)…			4
		OLAP(Q)	•		
•		タスクテンプレード(K)	•		

LocalSASUSER.B	NOUIN の 要約統計量	×
<b>1</b> /4 <i>7</i> −9	ētējī.at.	<u>s</u> sas
要約統計量ウィザード	では、記述統計量の表、ヒストグラム、箱ひげ図を作成できます。また、グループ実験の使用	も可能です。
データ		(
CAC Count	Least	編集(1)-
54J5U	SASUSER	
データセット	BAOTIN	
タスクフィルダ なし		
	(REVE) (XA(M)) (X7(E)	- キャンセル ヘルブ

データファイルの読み込み 000000000 連続データの解析 ○○○●○○○○○○○○○○○○○○

### 要約統計量の計算とヒストグラム・箱ひげ図の出力(2)

- ・ 変数 [人口 10 万対病院・歯科・一般診療所数] を [分析変数] に, 変数 [高齢者] を [分類変数] にドラッグ&ドロップ → [次へ]
- [統計量を表示する], [ヒストグラム], [箱ひげ図] にチェック → [次へ]

※ Local SASUSER BYOUDN の 要約統計量 🛛 🗶	≥ LocalSASUSERBYOUDI の 裏約統計量 🛛 🗶
2 / 4 皮肤在技事化制的当てます。	3/4 #目#248##61#21.その他の要約統計量の路定
(전) 2013 FRARE (1) 2013 FRARE	REC.         REC.           P Mid Star 7420         P Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420           P Provide Mid Star 7420         P Provide Mid Star 7420

 ・青線部分に示された統計量以外を出力したい場合,右上から設定で
 きる

データファイルの読み込み 000000000 連続データの解析 ○○○○●○○○○○○○○○○○○

### 要約統計量の計算とヒストグラム・箱ひげ図の出力(3)



$\Sigma$ LocalSASUSER.BYO	DUIN の 要約統計量	×
<b>4</b> /4 9-11-118	とフットノートを指定します。	<u>s</u> .sas.
分析( <u>A</u> ):	[更约統計量 結果	04261-( <u>B</u> )
ヒストグラム(出):	要約統計量 ヒストグラム	Ut20H(E)
箱O(F図( <u>W</u> ):	栗約統計量 箱0份图	Utzyh(S)
フットノート( <u>Q</u> ):	(SAS (& SASSERVERNAME, &SYSSOPU) (_E-T, #TRIM(MOGYSFUNC(DATE(), NLDATE20)) # TRIM(#SYSFUNC(TIME(), NLTIMAP20)) (ZE#\$24(#U,C,	Utzah(D
	<戻る(日) ※へ(日) ↓ キャンセル	

データファイルの読み込み 000000000 

高齢者	オブザベーション教 変数	平均	標準偏差	最小値	最大値	Ν
	人口10万対病院数	6.3130435	1.7535171	3.9000000	10.8000000	23
	人口10万対歯科診療所数	50.3956522	8.7888089	39.000000	82.7000000	23
少ない	23 人口10万対一般診療所数	74.8826087	12.9442740	55.4000000	99.1000000	23
	人口10万対病院数	10.0958333	3.2468686	5.7000000	17.6000000	24
	人口10万対歯科診療所数	45.6250000	4.8172922	33.5000000	55.1000000	24
多い	24 人口10万対一般診療所数	81.6666667	11.2479435	68.3000000	106.4000000	24

- 人口 10 万対病院数
   高齢者が少ない:平均 6.3 施設,標準偏差 1.75 施設
   高齢者が多い:平均 10.1 施設,標準偏差 3.24 施設
- 人口 10 万対歯科診療所数
   高齢者が少ない:平均 50.4 施設,標準偏差 8.79 施設
   高齢者が多い:平均 45.6 施設,標準偏差 4.81 施設
- 人口 10 万対一般診療所数
   高齢者が少ない:平均 74.9 施設,標準偏差 12.9 施設
   高齢者が多い:平均 81.7 施設,標準偏差 11.2 施設

データファイルの読み込み 000000000 連続データの解析 ○○○○○○●○○○○○○○○○○

## ヒストグラム



- ・人口 10 万対病院数
   ピーク位置とばらつきが異なりそう → 対数変換して処理
- 人口 10 万対歯科診療所数 似た形状,高齢者が少ない群に外れた値がある
- 人口 10 万対一般診療所数 似た形状,ピーク位置が異なりそう

データファイルの読み込み 000000000 



- ・人口 10 万対病院数
   ピーク位置とばらつきが異なりそう → 対数変換して処理
- 人口 10 万対歯科診療所数 似た形状,高齢者が少ない群に外れた値がある
- 人口 10 万対一般診療所数 似た形状,ピーク位置が異なりそう

データファイルの読み込み

連続データの解析 ○○○○○○○○●○○○○○○○

# 二標本 t 検定の実行 (1)

- [タスク] → [分散分析] → [t 検定]
- [2標本に対する t 検定] → [データ] 設定画面へ

ファイル(E) 編集(E) 表示(Y)	52	5(K) 70354(P) 9	-MD	~	UJJ(B)  }	ù• 🧀 🚳 I 📇 🕫 🗅 :
10ジュクトンリー プロセスフロー 動 byouinxts 」 愛 データ (byouinxts/byouin)		データ( <u>0</u> ) 記述統計( <u>8</u> ) グラフ( <u>0</u> )		5)	エクスポード	20 • スケジュール(D) •   ズーム(Z) •
1975 Shore	6	3版546(B) 日滞分析(B) 多支量解析(M) 生存時間分析(S) 工程能力分析(Y) 管理団(S) パレート国(P)-		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	・ 税金(1). ・ 元配置分 ノンパラメトリ・ 線形モデル(1) 混合モデル(1)	教分析(1) 約5年一元起量分数分析(1) ) り
	ų,	時糸列分析(① モデルのスコアリング(1)…	,			
C		0LAP(0) タスクテンプレード(K)	•			

topal/SASUSER BY	UIN2 の + 検定 🛛 🛛 🔀
第二章 デース 第二章 プラ プラ プラ プラ プラ プラ プラ プラ プラ プラ	
()コードのプレビュー	) 東洋(8) × (水井(6) キャンセル ヘルブ
対象変数"役割には3	放を割り当てる必要があります。

データファイルの読み込み 000000000 連続データの解析 ○○○○○○○○○○○○○○○○○○○

## 二標本 t 検定の実行 (2)

- ・ 変数 [人口 10 万対歯科療所数・一般診療所数,病院数 (Log)] を [分析 変数] へ,変数 [高齢者] を [分類変数] ヘドラッグ& ドロップ → [グラ フ] の設定画面へ
- [要約プロット] および [正規 Q–Q プロット] にチェック → [実行]

2757 5719	
#389.2.14(2) #2209/#891D F 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
Control and a control a	
	×
	~87

データファイルの読み込み 000000000 

# ニ標本 t 検定の結果 (人口 10 万対病院数 (Log))

- (H<sub>0</sub>: 高齢者の割合によって,人口 10万対病院数 (Log) に違いはない
- $H_1$ : 高齢者の割合によって、人口 10 万対病院数 (Log) は異なる



 P-value<0.0001であり、帰無仮説は棄却される したがって、高齢者の割合によって、人口10万対病院数(Log)は異な ると考えられる

データファイルの読み込み

# 二標本 / 検定の結果 (人口 10 万対歯科診療所数)



 Q-Q プロットから、はずれ値の影響で正規性の仮定が満たされない 可能性が示唆される よってノンパラメトリックな方法を用いて解析しなおすことにする

データファイルの読み込み

## ニ標本 t 検定の結果 (人口 10 万対一般診療所数)

- $ig( _{H_0}:$  高齢者の割合によって, 人口 10 万対一般診療所数に違いはない
- H1: 高齢者の割合によって,人口10万対一般診療所数は異なる



 P-value=0.06 であり、帰無仮説は棄却されない したがって、高齢者の割合によって、人口 10 万対一般診療所数が異な るかどうかは分からない

54/83

データファイルの読み込み

## ノンパラメトリック検定の実行(1)

- [タスク] → [分散分析] → [ノンパラメトリックな一元配置分散分析]
- ・ 変数 [人口 10 万対歯科療所数] を [分析変数] へ, 変数 [高齢者] を [分 類変数] ヘドラッグ& ドロップ → [分析] の設定画面へ

ファイル(E) 編集(E) 表示(y)	タスカ(K) プログラム(P) ツ	-x0 vx20 (治 🧀 🖓 🖓 🖓 🖓 🔿
カイロの 単本位 それの オリビスコー 第 DownAll Art (14) 美国などの水ートではたす 美国などの水ートではたす した19831 人口のグライリックセーアを変更	デーダ①     総注核計(8)     ガラブ(9)     分類が休(8)     の)ゆう竹(8)     ちラ短編新作(9)     な理解新作(9)     生存時間分析(5)     工程能力分析(5)     工程能力分析(5)     工程能力分析(7)     管理(9)     パレード致史)	・         1         122パード(ジ - 2752 440) -   スーム(ジ -
	時代列的新D 時代列の第2000年(1)- 0LAP(2)	•
		* 25/1014/H10072
	タスクテンプレード(K)	-元配蛋分数分析



データファイルの読み込み

連続データの解析 ○○○○○○○○○○○○○○○○○

ノンパラメトリック検定の実行(2)

• [検定に用いるスコア] → [Wilcoxon] のみにチェックが入るように変 更 → [実行]

LocalSASUSER.BYOUIN2のノンパラメトリックな一元配置分離	教分析	×
データ         第二           第二         第二           148         ア           タイトル         ア           フリバディ         第二           第二         第二           11         Strate(3)           マーキ共同のあまり(4)         「キーキ共同のあまり(4)           マーキ共同のあまり(4)         「ホーホートをおける(4)           マートホーホートをおける(4)         「ホーホートをおける(4)           「たちじい)         「たちじい)           「たちじい)         「まテーメ(2)           第二         デーメ(2)           第二         デーメ(2)	■ 経験分布理解紙計量(EDD 転計算する(D) ■ 大相撲を分類や本としてきめる(D) ■ 連続性の神運をあ示しない(U) 注(正確な p(0) ペーンで、正確な p(0) を使用可能(2)、(D20) ペーンで、な好量です。 を使用可能(2) たい(2) (2) (2) (2) (2) (2) (2) (2) (2) (2)	
F07V21-(©)	実行(B) ▼ 保存(S) キャンセル ヘルプ	<u> </u>

データファイルの読<u>み込み</u>

分割表データの解析

連続データの解析 000000000000000000

## ノンパラメトリック検定の結果

# $iggle H_0:$ 高齢者の割合によって,人口 10 万対歯科診療所数に違いはない $iggree H_1:$ 高齢者の割合によって,人口 10 万対歯科診療所数は異なる

Wilcoxon の順位和検定( 2 標本)				
統計量	654.0000			
止現近似				
Z	2.1604			
片側 Pr > Z	0.0154			
<b>酉側</b> Pr >  Z	0.0307			
t 分布で近似				
片側 Pr > Z	0.0180			
両側 Pr >  Z	0.0360			
Z には 0.5 の連続性の補証	Eが含まれて <mark>います。</mark>			

Kruskal-Wallis 検定		
カイ 2 乗	4.7133	
自由度	1	
Pr > Chi-Square	0.0299	

 P-value=0.04 であり、帰無仮説は棄却される したがって、高齢者の割合によって、人口10万対歯科診療所数は異な ると考えられる

### 参考文献

- [1] Der G, Everitt BS. Basic statistics using SAS "Enterprise Guide": a primer. SAS Publishing, 2007.
- [2] 高柳良太 (著), SAS Institute Japan (監修). SAS による統計分析–SAS Enterprise Guide ユーザーズ ガイド. オーム社, 2008.
- [3] 佐藤俊哉. 交絡事実と反事実の比較. 岩波科学 2008 年 4 月号.
- [4] **厚生労働省.** 平成 19 年医療施設 (動態) 調査. 2007. http://www.mhlw.go.jp/toukei/list/79-1.html
- [5] 総務省統計局.人口推計平成 19 年 10 月 1 日現在人口. 2007. http://www.stat.go.jp/data/jinsui/2007np/index.htm

追加解析 00000

## グラフの作成と読み方

- 離散変数(性別,重症度など)
  - 一変数: 棒グラフ, 層別: 積み上げ棒グラフ, 三次元棒グラフ
- 連続変数 (年齢, 臨床検査値など)
  - 一変数/層別: ヒストグラム, 箱ひげ図 (スライド 44), 二変数: 散 布図
- 生存時間変数
  - 一変数/層別: Kaplan-Meier プロット

追加情報 ○●○○○○○○○○○○○○○○○○○○○ 追加解析 00000

# 棒グラフ/積み上げ棒グラフ (1)

- [タスク] → [グラフ] → [棒グラフウィザード]
- データを選択し、[次へ]



A 1.1		<u> </u>	
1/4 7-5	を相任します。	<u>S</u> :	d
レクラフウィザードでは	、縦梯グラフや横梯グラフを 20	)や SD で作成でをます。また、バーのグループ化や分割や行えます。	
1-9			(85)
040.0		. <u></u>	<u></u>
	SASIRER		
データだかと	TAMOOFEN		
2 Amonth fol			
0.074092.40			

追加解析 00000

### 棒グラフ/積み上げ棒グラフ (2)

- [バー]: 横軸の変数
- [分割の基準]: 積み上げ棒グラフ
- [Z軸]: 層別因子が二つ以上の場合, 三次元棒グラフを作成する
- ・ [グループ化する基準]: 棒グラフを並べる
- 変数を選択し, [完了]

In Local SASUSER TAMOXIFEN の種グラフ2	×	X	
2 / 4 实现在说明口都的当てます。	<u>s</u> sas	IS 3000-	
			3.の再発 ■ あり ■ ねし
<22(0) 32(0)	<u>実T(E)</u> ▼ <u>キャンセル</u> へルフ	タモキシフェン使用の有無	

## 散布図(1)

- ・ [タスク] → [グラフ] → [散布図]
- [2D 散布図] をダブルクリック





### 散布図 (2)

- *x* 軸, *y* 軸の変数をドラッグアンドドロップして設定する
- [補完] をクリックし, 補完法とオプションを設定し, [実行]



## 散布図 (3)

- ・ 散布図からは、二変数の関連を読み取る事ができる
- 二変数間に直線的な関連,二次関数的な関連(減ってから増えたり), 周期的な関連,または何も関連が無いのかを視覚的に評価する



### Kaplan-Meier プロット (1)

- [タスク] → [生存時間分析] → [ノンパラメトリック法による生命表 分析]
- 生存時間変数,打ち切り変数と打ち切りを表わす水準,必要に応じて 層別変数をドラッグアンドドロップして設定する





### Kaplan-Meier プロット (2)

- [グラフ] をクリックし、[生存関数プロットを表示する] にチェックを 入れる、信頼区間を表示する場合は [ポイントワイズ信頼限界を取り 込む] にチェックを入れる
- Log-rank 検定などの結果と Kaplan-Meier プロットが出力される



追加情報 ○○○○○○○○●**○○**○○○○○○○○○ 追加解析 00000

## チュートリアルの表示

・ 起動時に表示されるようこそ画面, または [ヘルプ] → [チュートリアル]



ページが表示されました 😜 インターネット

• 操作方法を忘れた時に読むと便利です

- 100% ·

### タスクリスト

・ 左下: サーバリストからタスクリストに変更しておくと, 解析手法一覧が表示される



## タスクステータスの表示

実行中・実行予定の処理内容を表示できる





## 出力形式の追加

- [ツール] → [オプション]
- [結果一般] → [結果ファイルの形式] から [RTF] を探してチェック



• RTF 形式で出力すると, Word 等にコピー& ペーストして編集しや すい

追加解析 00000

## 出力デザインの変更

[ツール] → [オプション]

[RTF] → [スタイル]





追加解析 00000

# 出力デザインサンプル (一部)

### デザインリスト:http://www.josai.ac.jp/~nagasima/contents/ sas/odsstyle/ods\_style.html

ods rtf style = BarrettsBlue The FREQ Procedure

Table of x by y			
x(薬剤の投 与)	y(治癒)		
Frequency Percent Row Pct Col Pct	あり	なし	Total
あり	63 21.00 42.00 60.00	87 29.00 58.00 44.62	150 50.00
なし	42 14.00 28.00 40.00	108 36.00 72.00 55.38	150 50.00
Total	1 05 35.00	195 65.00	300 1 00.00

ods rtf style = Listing The FREQ Procedure

Table of x by y			
x(薬剤の投与)	y(治癒)		
Frequency Percent Row Pct Col Pct	あり	なし	Total
あり	63	87	150
	21.00	29.00	50.00
	42.00	58.00	
	60.00	44.62	
なし	42	108	150
	14.00	36.00	50.00
	28.00	72.00	
	40.00	55.38	
Total	105	195	300
	35.00	65.00	100.00

ods rtf style = Journal The FREQ Procedure

Table of x by y					
x(薬剤の投与)		y(治癒)			
Frequency					
Percent					
Row Pct					
Col Pct	あり	なし	Total		
あり	63	87	150		
	21.00	29.00	50.00		
	42.00	58.00			
	60.00	44.62			
なし	42	108	150		
	14.00	36.00	50.00		
	28.00	72.00			
	40.00	55.38			
Tota/	105	195	300		
	35.00	65.00	100.00		

BarrettsBlue

Listing

Journal
追加解析 00000

# 出力デザインのカスタマイズ (1)

[ツール] → [スタイルマネージャ]
 カスタマイズしたいスタイルを選択 → [コピーの作成]

ツー	ル(D) ヘルプ(出)   🎦 - 🚰 - 强   📇	0							
	アドイン( <u>D</u> )								
Ð	HTML ドキュメントの作成(©)	ļ							
	スタイルマネージャ( <u>M</u> )								
<u>R</u>	SAS Enterprise Guide Explorer(G)	1							
<b>1</b>	プロジェクトライブラリの割り当て( <u>A</u> )								
Ð	ライブラリメタデータの更新(山)	l							
Ē¥-	JMP Stored Process Packager(J)								
	プロジェクトのメンテナンス( <u>P</u> )								
	開かれているデータセットの表示(⊻)								
	オプション( <u>0</u> )	1							

· 177 7711 (2)-			Come of Fick .	
マタイル	位置	UR 🔺		- 4
g Festival	ENHIDAGIN	C#	SAS System Title	
FestivalPrinter	ENHIDIA	0.9	OAO Drassdura Tida	
g Gears	ENH SAMIN	0.4	SAS Procedure The	
Grayscalerrinter	E/FM 2X91/F	0.8		
# Harvest	ENH DARTIN	0.4	Column 1 Column 2 Column 3	
HighContrast	E// M 2//9/1//	0.1		
Journal	ビルトインスタイル	~~ I	Row 1 Data (Num) Data (Char)	
gournai2	ビルドインスタイル	0.#	Row 2 Data (Num) Data (Char)	
g courness	C/VP12X91/V	0.1		
LISTING	ビルドインスタイル	0.4		
g magniny Maadam	ビルドリンスタリル	0.#	GRAPH	
Meauow	C/0P1 2X91/0	0.4	RESULTS	
meadowrrinter	ビルドリンスタイル	0.4		
g minimai Maaaa	ビルトインスタイル	0.0	SAS System Footnote	
g Moneckrowe Printer	ビルトロンスタイル	~~	SAS System Foundle	
Monochromerrinter	ビルトリンスタリル	0.#		
Moroniueraun	ビルトインスタイル	0.4		
Mormal Printer	ビルトロンスタイル	0.4		
One	ビルトロンスクロル	0.4		
Plateau	ビルトインフタイル	0*		
# printer	ビルトインスタイル	04		
P Dana	Gal A. T. S. A.			
		•	(2)	
四字(-45字(C)	(F#/r)	1 3	the (A) B(Da(n) C) B(Da(n) C)	
MARICANAELO/	THENCLE			

追加解析 00000

## 出力デザインのカスタマイズ (2)

- 作成したコピーに適切な名前を付けて [保存] → 作成したコピーを選 択 → [編集]
- スタイルエディタで確認しながら編集 → [OK]

			10 スタイルエディタ	×
<b>マ</b> スタイルマネージャ			要素を選択するには、クリックします	Body の開始:
291A 9210		Copy of Listing (071/21~	CAS Sustam Title	テキスト 野線 イメージ カスタム
29/16 29 lock Print block Foreen	位置 UR * Eルーインスタイル C4 Eルーインスタイル C4 Eルーインスタイル C4 Eルーインスタイル C4	SAS System Title SAS Procedure Title	SAS System The SAS Procedure Title	満定式ありまといい  wish#hetics.serare.exif テキストのサイズ(S) テキストのスタイル(2)  マinall ♥ (酒) ♥
Core of Using Core Care Core Core Core Core Core Core Core Co		416542        ● See        ● See        ■ See	Column Column Column 1 2 33 Row 1 2410 Data Row 2 Data Data (Nam)(Char) Row 2 Data Data (Nam)(Char) SAS System Footnote	Particle      Particle        Partit      Particle
MELINE(S)		isto(4)	istorWIRKE	्र   194 (7). 🗹
		0K 492/2% ^%		OK #+>tz#

## データのエクスポート

- SAS 形式のデータを選択 → [エクスポート] → [ファイル名] のエクス ポート
- ・保存先とファイル名および形式を指定→[保存]



- ・ 左の画像の下部のアイコンは SAS 形式のデータ
- SAS 形式以外も可

#### SAS / SAS Enterprise Guide インストールの注意点 I

SAS のシステム必要要件 (Windows):

Windows XP Professional (Service Pack 2 を適用), Windows Vista Enterprise Edition / Business Edition / Ultimate Edition, Windows XP Professional x64 Edition, Windows Vista Enterprise x64 Edition / Business x64 Edition / Ultimate x64 Edition, Windows 7 Enterprise / Ultimate / Professional / Home Premium Windows XP Home ではインストールも不可能

- SAS と SAS Enterprise Guide を別々にインストールすると SAS Enterprise Guide はうまく動作しませんので、同時にインストールし て下さい。
- ③ ユーザアカウント名、PC名、Workgroup名が日本語の場合インストールできない可能性があるので、管理者権限を持ったユーザアカウント名、PC名、Workgroup名をアルファベットで新規作成してください。

### SAS / SAS Enterprise Guide インストールの注意点 II

- ④ 以下のエラーメッセージが表示される場合 (Windows 7):
  - SAS Enterprise Guide にエラーが発生しました。作業中のファイルを 保存して SAS Enterprise Guide を再起動してください。 エラー詳細:
    - System.IO.FileNotFoundException ファイルまたはアセンブリ 'Microsoft.Web.Services3, Version=3.0.0.0, Culture=neutral, PublicKeyToken=31bf3856ad364e35'、またはその依存関係の1つが 読み込めませんでした。指定されたファイルが見つかりません。 → Web Services Enhancements (WSE) 3.0 for Microsoft .NET をイン ストールしてください。

http://www.microsoft.com/download/en/details.aspx?id=14089

#### SAS / SAS Enterprise Guide インストールの注意点 III

 6 その他にもエラーが表示される場合、該当バージョンに対応する アップデートファイルをダウンロードしてインストールすることで 問題を解決できる場合があります。
 http://fm ass.com/tachaup/download/batfur/HE2/A20.html

http://ftp.sas.com/techsup/download/hotfix/HF2/A20.html http://ftp.sas.com/techsup/download/hotfix/HF2/A32.html http://ftp.sas.com/techsup/download/hotfix/HF2/A53.html

追加解析 ●0000

#### Mantel-Haenszel 検定

- 目的
  - リンパ節転移の有無を考慮した上で,タモキシフェン使用の有無に よって,乳がんの再発オッズに違いがあるかどうかを検討する事
- プロセスフローに戻る → [分割表分析 2] を右クリック → [分割表分析 2 の変更]
- [表統計量] → [関連] → [CMH 統計量] → [実行] → 結果の置き換え [はい]





### 仮説検定の結果

- ・リンパ節転移を考慮した上で再発オッズが異なるかどうかを、 Mantel-Haenszel を用いて検定する ( $\alpha = 0.05$ とする)
- *H*<sub>0</sub>: リンパ節転移のありの群でも,なしの群でもタモキシフェン使用の有無で再発オッズは変わらない
- *H*<sub>1</sub>: リンパ節転移のありの群でも,なしの群でもタモキシフェン使用 の有無で再発オッズは異なる

Cochran	-Mantel-Haenszel	<del>結】重(</del> テ・	ーブルスコア	に基づく)
統計量	対立仮説	自由度	値	p値
1	相関統計量	1	10.9946	0.0009
2	ANOVA 統計量	1	10.9946	0.0009
3	一般関連統計量	1	10.9946	0.0009

相対リスクの推定値 (行 1/行 2)									
研究の種類	調整方法	値	95%信	頼限界					
ケースコントロール研究	Mantel-Haenszel	0.6444	0.4962	0.8368					
(オッズ比)	ロジット	0.6444	0.4962	0.8368					
コーホート研究	Mantel-Haenszel	0.6700	0.5275	0.8510					
(列1のリスク)	ロジット	0.6700	0.5275	0.8510					
コーホート 研究	Mantel-Haenszel	1.0397	1.0164	1.0635					
(列2のリスク)	ロジット	1.0397	1.0164	1.0635					

- *P*-value=0.0009であり, α = 0.05のもとで帰無仮説は棄却される
- リンパ節転移を考慮するとタモキシフェン使用の有無で再発オッズ が異なる



追加解析 00●00

## 分割表データの手入力(1)

- [ファイル] → [新規作成] → [データ]
- ファイル名を指定 → 保存先として [WORK] をクリック → [次へ]





追加解析 0000●0

## 分割表データの手入力(2)

- 列を三つにし, 名前をそれぞれ [x] [y] [n] に設定 → [完了]
- 表を右図の様に入力し、左上のプロセスフローをダブルクリックして 戻る

「新しいデー	<i>φ</i>			×												
	The factor and the state			6000	6	😿 SAS Ente	rprise Guide	e - วัดวีร!	ኃኑ1.eep							
2/2	9921年から、フロバティを指定しまう			<u>y</u> .3a3.		(The second	編集( <u>E</u> )	表示(⊻)	タスク( <u>K</u> )	プログ	ラム( <u>P</u> ) ツール(	ローハルプ	(H) 📔	· 👝 · 🤇	5 I 📇 I	
別(():		3	りのプロパティ(①):		1 1	プロジェクトツ!	-		+ ×	DAT	<b>A</b> 1 -					
名前	長さ()(イト)		- 名前	n	l î	E 808 70.	2220-				イルタと並べ替え	յ երջու	ビルダ(Q)	データ(D	) - 記述	
A× Av	12	÷	ラベル Sett	<b>*</b> 定		8	amoxiten xl 夏 データ (tu	ls amovifen vli	k[data]) () (		o) x	1	у	1	n	
۸	12	<u>.</u>	グループ	文字		ė-81	amoxifenxl	ls からインボ	ートされたデー	1		1	1			
	- 2	× .	長さ	12			🔠 分割表分	分析1		2		1	2			
			表示形式	\$12.		ė 🟭	DATA1			3		2	1			
		1	読み込み形式	\$12.		i	🔜 分割表分	分析		4		2	2			
										5						
										6						
										-						
										0						
										10						
\$6:07	et (n)   mP_(n)   Represent	1			I I					11						
4/1/261 1		~								12						
	< 戻る(B) (かへ)	00.2	完了(F)	**\/h/ \/J/	L 1											1



追加解析 0000●

## 分割表データの手入力(3)

- [タスク] → [記述統計] → [分割表分析]
- ・ 変数 [x] [y] を表変数に設定 → 変数 [n] を度数カウントに設定 → [表]
  をクリックしてスライド 3 と同様に設定する



