SAS (SAS Enterprise Guide) 使用方法

(2つのグループの1因子(1種類のデータ)の差を検定する)

目次

(1)	各グループ別のヒストグラムおよび、累積度数分布の確率プロットの表示 2
(2)	t 検定 (t-test)を行う5
(3)	データを関数変換して分析を行う12
(4)	Mann-WhitneyのU検定 (Wilcoxon 順位和検定)を行う15
	(補足;「中央値検定(メディアン検定)」)
(5)	記号の順位(順序)尺度データを数字に変換する 21
(6)	χ2検定 (カイ二乗検定)を行う25
(7)	対応のあるデータの差が正規分布しているかどうかを判定する
(8)	対応のある t 検定 (Paired t test)を行う35
(9)	Wilcoxon 符号順位和検定(Wilcoxon signed-ranks test)を行う

- (1) 各グループ別のヒストグラムおよび、累積度数分布の確率プロットの表示
 - 1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。
 - 2) 上部メニュー内の [記述統計]をクリックし、その下に表示されるメニュー内の [分布]
 - をクリックする。 🌀 SAS Enterprise Guide – Sample-Load ファイル(E) 編集(E) 表示(V) コード(C) データ(D) 記述統計(S) グラフ(G) 分析(A) ウィザード(₩) 월•**6]• % | 몸 % 맘 많 X | 9 @ |** Q 📺 IJスト(L)... SASUSERIMPW_0006(データのインポ 🔹 🚬 🕨 🔲 ∑ 要約統計量(S). 8₉₈ 703 プロジェクトエクスプローラ ŧΧ 分布(D)... 💽 Sample-Load 12 E 8 70 270-データの特性分析(日)... 🗄 🎥 Samplexis (Sample\$) 📄 集計表(<u>T</u>)… 2 ⊡ 📲 データのインポート Ⅲ 一元度数表(<u>0</u>)... 3 🏢 最後にサブミットしたコード ログ 岡 сленеер тмрж 0006 分割表分析(A)... 4 Ħ h : LELIFH
 - 3)新しいウィンドウが開くので、ヒストグラムや累積度数分布の確率プロットを表示する 変数(正規分布しているかどうかを確認する変数)を「変数リスト」の下のリストから選んで クリックする。そして、その右隣りにある右向き矢印のボタンをクリックし、新たに表示 されたメニューから「分析変数」をクリックする。

In SAMPLE_0782 の 3 タスクの役割 分布 概要 正規分布 対数正規分布 指数分布 Weibull 分布 パンマ分布 ガンマ分布 核 グラフ 表示 インセット テーブル タイトル	 本 タスクの役割 変数リスト(<u>A</u>) 名前 ② F1 ③ 读字氏名 ④ 性別 ③ 年令階級 ④ 字範(才) ③ 年令階級 ④ 梁運 ⑩ 年数(本) ⑩ 小田田田(mmHg) ● 最低血圧(mmHg) ● 最低血圧(mmHg) ● 最低血圧(mmHg) ● 最低血圧(mmHg) ● 尿 瑪 ● 尿 第 品 ○ 尿 ○ 添加球(万個/mm3) ○ 血色素(g/dl) 	タスクの役書帐E): ③ 分析変数 ③ 分析変数 ◎ 定み (選択の上限:1) ④ 重み (選択の上限:2) ⑦ルーブ分析 ◎ 度数カウント ◎ 重み ⑦数カウント ◎ 重み ⑦数カウント ◎ 重み ⑦対策変数 ⑦	
			4
() 「つードのプレビュー() 「分析変数」役割には、	○)	実行(B) 上書き保存(S) キャンセル ヘル・	<u>, </u>

4) 右の「タスクの役割」の下の「分析変数」の下に、選択した変数名が表示されているこ とを確認する。 5) グループ分けに使用する変数(その変数を使ってグループ分けする変数)を「変数リスト」 の下のリストから選んでクリックする。そして、その右隣りにある右向き矢印のボタンを クリックし、新たに表示されたメニューから「分類変数」をクリックする。

<u>』 </u> _SAMPLE_0782の分	布	×
タスクの役割	<u>タスクの役割</u>	
万布	変数リスト(A) タスクの役割(E):	
正規分布	名前	
対数正規分布	@ F1	
指数分布	▲ 漢字氏名 ⑧ グループ分析	
Weibull 分布	▲ 性別	
ベータ分布	🔞 年齢(才) 🛛 🚽 🚳 重み (選択の上限:1)	
ガンマ分布	□ 毎 毎 日 日 日 日 日 日 日 日 日 日 日 日 日 日 日	
核	▲ 喫煙	
<u> グラフ</u>	(1) 年数(年) (10 - 1) (2) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4	
表示	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
1ノゼット	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
タイトル	1 (1) 体重(Kg)	
21177	⑩ 最高血圧(mmHg) _	
	l 创 最低血圧(mmHg)	
	▲尿	
	(1) 赤血球(万個/mm3)	
	J	7
 」 」 」 ードのプレビュー(9)) 実行(R) 上書き(保存(S) キャンセル ヘルブ	
1		- 111

- 6) 右の「タスクの役割」の下の「分類変数」の下に、選択した変数名が表示されているこ とを確認する。
- 7) 左端のメニューから「分布」の下の「正規分布」をクリックし、新たに表示された右側 の画面で、「正規分布」の前のボックスにチェックを入れる。

_ 山 IMPW_0006 の分和	म	x
 - タスクの役割	一 分布 〉 正規分布	
: 分布	\frown	
概要	□ 正規分布(N) -	
正規分布	▶ 分布をすべての変数に通	新田する(A)
対数正規分布	I → 75 m ≥ 51 × C053C0×1028	- 狼のオブション
: 指数分布	分析変数(⊻):	平均 (ミュー)(M)
· Weibull 万布		
ガン分布		標準備差(シクマハロ) ▼
核		種類(^):
グラフ		
表示		
インセット		幅(<u>D</u>):
- テーノル		
1 2/11/2		
		推定値の使用(山)
ī.		
-		(指定しない場合、パラメータは推定されます。)
1	1	
1		
-		
		<u> </u>
-		T
1		
□□□ードのプレビュー	(<u>C</u>)	実行(R)上書き保存(S)キャンセル ヘルプ
L		

8) 左端のメニューから「グラフ」の下の「表示」をクリックし、新たに表示された右側の 画面で、「ヒストグラム」と「確率プロット」の前のボックスにチェックを入れ、下にあ る「実行」ボタンをクリックする。

注: インセットは、ヒストグラム、確認 プロットのみ有効です。				
	率ブロット、QQ 軸の色:	背景色:	軸の幅:	
		•	1	•
		•	1	•
)	•	1	Y
● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●		V	1	-
□ テキストベース	オブザベー: のブロット(L) かを、また和 By 変数が	ション数に応じて幹! 箱ひげ図、正規確当 ある場合、プロット(栗図、棒グラフ) ■プロットを作成 ま並列に作成す	のいずれ \$します。 5れます。
連続型の理論分布の確率プロットを タを示す参照線を重ね合わせます。	確作成し、理論分布に対して指	定または推定された	こ位置パラメータ	心尺度パラメー 🧾
, »		書き保存(<u>S</u>)	キャンセル	-
	・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	・ ・ </td <td>・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・</td> <td></td>	・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	

9) 画面に「分布」の分析結果が表示されるが、最下部にヒストグラムおよび、累積度数分 布の確率プロットが、分類変数に設定した変数の値別に表示されている。

10) 最下部に下図のような分類変数に設定した変数の値別の累積度数分布の確率プロットが あるので、分析に使用するグループの確率プロット全てで、青い正方形のプロットが、左 下から右上に伸びている黄色い直線上に乗っているかどうかを確認する。

青いプロットが黄色い直線上にほぼ乗っている場合は、この変数(データ)は正規分布していると判断してよい。もし、青いプロットが黄色い直線から大きくずれている場合は、この変数は正規分布していないと判断できる。



分析: "最低血圧(mmHg)"n

(2) t 検定 (t-test)を行う

t検定は、必ず下記の流れで行うことを、まずしっかりと頭に入れておくこと



1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。

2) t 検定を含めた比較検定を行う場合、その前に必ず、以下の操作で、グループ分けに用い る変数の値またはカテゴリーが、比較検定で用いる2つの値またはカテゴリーであるデー タのみに限定する必要がある。

(a) 上部メニュー内の [データ]をクリックし、その下に表示されるメニュー内の [フィ ルタとクエリ]をクリックする。

🕵 SAS Enterprise Guide - Sample-Load	
ファイル(E) 編集(E) 表示(V) コード(C)	データ(D) 記述統計(S) グラフ(G) 分析(A)
1-6-6 4 4 6 6 X 6	✓ 読み取り専用(E)
SASUSERIMPW_0006(データのインポ 🔹	フィルタとクエリ(E)
プロジェクトエクスプローラ • •	🗮 行(R) 🛛
💽 Sample-Load	III 列(L)
	移動(<u>G</u>) Ctrl+G

(b) 下図のようなダイアログが開くので、左側の変数リストから、グループ分けに用い る変数と、比較する変数を、右側の「データの選択」の下のボックスにドラッグアン ドドロップし、これらの変数名が表示されていることを確認する。

🕷 SASUSERIMPW_0006 のクエリ - クエリビルダ				×
クエリ名(Q): SASUSERIMPW_0006 のク	IU 出力名(<u>0</u>):	SASUSER.QUERY_FOF	.SASUSER_IMPW_0	変更(<u>C</u>)
□ 計算列(M) □ パラメータ(P) ◎ 検証(V)	• □ 〔 ブレビュー(E) □ ゴブ	'9a)(<u>N</u>) ▼		
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	データの選択 フィルタデータ	データの並べ替え		
□ □	<u>列化名</u> ▲ "戰浬" (戰½) ④ '最低血圧(mmHe)' (最_	入力 IMPW_0006段/控 IMPW_0006最低血圧(m_	概要	
	- 要約グルーフ 「 グルーフを自動選択する」 グループが選択されていません 「 重視しない行のみ選択する	(U) 5(S) 実行(B) (保存して閉		20編集(g) メ メ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・

(c)「データの選択」の右隣の「フィルタデータ」のタグをクリックし、その下に表示 されたボックスに、グループ分けに用いる変数をドラッグアンドドロップする。

置 SASUSERIMPW_0006 のクエリ - クエリビルダ		×
クエリ名(Q): SASUSERIMPW_0006のク	EU 出力名(Q): SASUSER(QUERY_FOR_SASUSER_IMPW_0 変更(_)
🎰 計算列(M) 🛛 🥫 パラメータ(P) 🦻 検証(V)	• 🗟 ブレビュー(E) 🔮 オブション(N) •	
 □ 追加(① ※ 削除(②) □ 結合(④) □ 副 IMPW_0006 □ ③ F1(F1) □ ④ F1(F1) □ ④ 注影(才)(年齢(才)) □ ④ 年齢(才)(年齢(才)) □ ④ 年齢(才)(年齢(才)) □ ④ 年齢(次)(年齢(次)) □ ④ 年齢(年)(年齢(年)) □ ● 年数(年)(年数(年)) □ ● 身長(cm)(身長(cm)) □ ● (年(Kg)((年重(Kg))) 	データの選択 フィルタデータ ータの並べ替え 未加エデータ(こフィルタを適用 (オプション)データこフィルタを適用す	ここに列をドロップしてください。

(d)「フィルタの編集」というダイアログが開くので、「値(V)」の行の右端の黒い逆三

角形ボタンをクリックする。

列:	IMPW_0006 喫煙
演算子(0):	次の値に等しい
值色在	
12(1)	
- フィルタ定義	
- フィルタ定義	Ē = "
- フィルタ定義 IMPW_0006 喫烟	Ē = "
- フィルタ定義	Ē = " "囲む(E)
 フィルタ定義 IMPW_0006 喫閉 IMPW_0006 厚閉 IMPW_0006 厚別 IMPW_0006 FN IMPW_000	E = " *囲む(E) 適用された日 付を使用する(U)
7-7ルタ定義 「MPW_0006 時期 「 値を引用符で 「 出力形式が近	*囲む(E) 適用された日付を使用する(①)

(e)新しいダイアログが更に表示されるので、その下部にある「値の取得」ボタンをク リックする。



(f) グループ分けに用いる変数の値やカテゴリーのリストが表示されるので、まず、2 つのグループの一方の値またはカテゴリーをクリックする。

		×
データ値列	パラメータ	
値	出力形式が適用された	:値 [
<u>A</u>	A	
Ċ	Č	
<u> </u>		
値の追加取得	(v)	
		11

(g)「フィルタの編集」のダイアログに戻るので、「値」の右に、選択した値またはカ テゴリーが表示されていることを確認し、下の「OK」ボタンをクリックする。

鬱 フィルタの編集	×
歹!):	IMPW_0006 喫煙
演算子(0):	次の値に等しい
値()):	A
- フィルタ定義 IMPW_0006 喫煙 =	'A'
▼ 値を引用符で囲る	b(E)
☑ 出力形式が適用	された日付を使用する(旦)

(h) 最初のダイアログに戻り、選択した値またはカテゴリーが表示されていることを確 認する

🚡 SASUSERIMPW_0006 のクエリ - クエリビルダ 🛛 🔀							
クエリ名(Q):	SASUSER.	IMPW_0006 のク:	ΕIJ	出力名(_):	SASUSER.QUERY_FOR_SASUSER_IMPW_0	変更(<u>C</u>)	
⊞計算列(M)	🕞 パラメータ(<u>P</u>	り 🔊 検証(У)	- B 7	⁷ レビュー(E) 📓 オプシ	ν∃)(<u>N</u>) ▼		
田 追加(I) □ 副 IMPW(□ □ □ 目 □ □ □ 第 □ □ □ 10 10 10 10 10 10 10 10 10 10 10 10 10 1	★ 削除(D) 0006 (F1) 字氏名(漢字氏: 別(性別) (性別)	[■] 結合(」) ▲ 名)	データの選 「未加工デ 「IMPW」	択 フィルタデータ テ ータにフィルタを適用 - 0006 喫煙 = " A"	*-タの並べ替え	V 2	

(i) (c)~(h)の操作を繰り返し、2つのグループのもう一方の値またはカテゴリーを設 定する。 (j) 2つのグループの値またはカテゴリーが選択されていることを確認した後、その下の「フィルタの組み合せ」ボタンをクリックする。

	A CONTRACTOR OF	1.6.4		40.10	
🐻 SASUSER IM	PW_0006 のクエリー・	ウエリビルダ			×
クエリ名(Q):	SASUSERIM	PW_0006 のクエリ	出力名(0):	SASUSER.QUERY_FOR_SASUSER_IMPW_0	〕 変更(<u>C</u>)
⊞ 計算列(M)	🎲 パラメータ(P)	② 検証(⊻) ・!	🕄 ブレビュー(E) 🔛 オブ	ý∋∑(<u>N</u>) +	
	★ 町(R金/ D)		の避壊「フィルタデーター」	デー (2) (1) (1) (1) (1) (1) (1) (1) (1) (1) (1	
	A HUMADY		エデータにフィルタを適用・		1
□ 🖃 🚟 IMPW_0)006 (F1)	- I I	IPW_0006.喫煙 = " A"		
漢 漢 漢	, 字氏名(漢字氏名))	IPW_0006.喫煙 = "C"		
(住)	问(性別) 渝(才)(年齡(才))				<u></u>
6 年4	令階級(年令階級)			×
- (1) 年3	罜(喫/罜) 汝(年)(年数(年))				
一適林	波(本/日)(本数(本	(日))			
	を(cm)(身長(cm), 重(Kg)(体重(Kg))	, w	HERE IMPW_0006.喫煙	= ″A″ AND IMPW_0006.喫煙 = ″C″	<u>^</u>
- 6 最	ー 高血圧(mmHg)(最 ダカロ(リン/目	高血圧()			
- (2) 取	5.血圧(mmHg)(策 		フィルタの組み合せ(F)		
- ▲ 尿	蛋白(尿蛋白)				
脉?	1911日(水清川)	「要約	テータにフィルタを適用―		

- (k)「フィルタの組み合せ」というダイアログが新たに開くので、2つの漏斗の絵の間
 にある「AND ▼」をクリックし、その下に表示されたメニューから「OR に変更」をク
 - リックする。



- (1)「AND ▼」が「OR ▼」に変ったことを確認した後、最下部の「OK」ボタンをクリッ クする。
- (m) 最初のダイアログに戻るので「実行」ボタンをクリックする。
- (n) 画面上に、グループ分けに使用する変数と、比較する変数のみの表が、新たに表示 されたことを確認する。
- 3)上部メニュー内の[分析]をクリックし、その下に表示されるメニュー内の[分散分析] をクリックし、更にその右に表示されたメニュー内の「t 検定」をクリックする。

My on o Enterprise dande Gampie Edda			
̄ ファイル(E) 編集(E) 表示(V) コード(C)	データ(<u>D</u>) 記述統計(<u>S</u>) グラフ(<u>G</u>) [1	分析(<u>A)</u> アドイン(<u>1</u>) OLAP(<u>O</u>)	_ ツール(I) ウィンドウ(W) ヘルプ(H)
11-6-5 B+ BC X 0) (ล 🔍 📜 🎭 วิธวีรรีหารีรี	分散分析(<u>A</u>)	赵 t 検定(丁)
		回帰分析(<u>R</u>)	→ 元配置分散分析(0)
		多変量解析(<u>M</u>)	△ ノンパラメトリックな一元配置分散分析(N)
	🏭 🎭 プロジェクトデザイナ 📗 SASU	生存時間分析(S)	
-LOAD	🔊 喫煙 😥 最低血		送 混合モデル(<u>M</u>)
SamplexIs (Sample\$)	1 0	空理図(0)	
■ データのインボート	2 0		
■	3 C	™ パレート図(P)	
	4 C	時系列分析(工)	-
SASUSERIMPW_0006	5 C		-
SASUSERIMPW_0006	6 A	💱 モデルのスコアリング(<u>L</u>)	

4) t 検定のダイアログが開くので、「t 検定の種類」で「2 標本に対する検定」の前にマー クが入っていることを確認する。もし、マークが付いていない場合は、その前の丸をクリッ クしてマークを付ける。



5) 左端のリストから「タスクの役割」をクリックし、その右に新たに表示された「変数リ スト」からグループ分けに用いる変数をクリックした後、その右の右向き矢印のボタンを クリックし、新たに表示されたメニューから「グループ化する基準」をクリックし、「タ スクの役割」の下の「グループ化する基準」の下に、その変数名が表示されたことを確認 する。



6)「変数リスト」から比較する変数をクリックした後、その右の右向き矢印のボタンをクリックし、新たに表示されたメニューから「分析変数」をクリックし、「タスクの役割」の下の「分析変数」の下に、その変数名が表示されたことを確認する。

Ŀ	_QUERY_FOR_S	ASUSER_IMPW_0006_0003のt検定	
•	t 検定の種類	タスクの役割	
1	タスクの役割 分析	変数リスト(<u>A</u>)	タスクの役割(E):
•	グラフ	名前	🞯 グループ化する基準 (選択の上限
•	タイトル	▲ 喫煙	喫煙
•		😡 最低血圧(mmHg)	◎ 分析変数
1			
1			(2) グループ分析 (2) ウループ分析
1			◎ 度钡カリント (選択の上限:1)
1			◎ 里の (選択の上版:1)
1			
1			- <u>御</u> ジルーフ1L9の卒年
1			分析変数
1			_ 一〇 グループ分析
			◎ 度数カウント
•			 動 重み
•			
•			
•		1	

7) 左端のリストから「グラフ」をクリックし、その右に新たに表示された「種類」で、「平均」の前のボックスにチェックを入れ、「平均のプロットオプション」では「標準偏差」にマークを付け、「描画するばらつきの大きさ」では「2標準偏差」の前にマークを付けた後、最下部の「実行」ボタンをクリックする。

タスクの役割 分析 グラフ タイトル 種類 平均のプロットオブション 平均変数の種類 ● ● 「 箱ひげ図(B) ● ● 「 箱ひげ図(B) ● ● 「 和切(標準認差(E) ● ● ● ● ● □ 「 フールされた分散を使用する(U) ● ● ● ● ● □ ● ● ● □ ● ● ● □ ● ● ● □ ● ● ● □ ● ● ● □ ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●	QUERY_FOR_S t検定の種類	ASUSER_IMPW_0006_0003 の t 検定 グラフ	x
	9人200代書) 分析 グラフ タイトル	種類 自 日 下 和 正 和 田 (B) (B) (D)	平均のブロットオブション 平均の種類 ・ 平均の標準誤差(E) ・ 標準偏差(D) 「 フールされた分散を使用する(U) 描画するばらっきの大きさ ・ 1 標準偏差(標準誤差)(1) ・ 2 確準偏差(標準誤差)(2) ・ 3 標準偏差(標準誤差)(3) 「 垂直軸を 0 から開始する(D)
平均ブロットを生成します。平均ブロットには、独立変数の平均と標準誤差が示されます。グラフ上の点は、ばらつきを示す ▲ バー内に存在します。	二二一下のプレビ	平均プロットを生成します。平均プロットには、独立変表 バー内に存在します。 ュー(<u>C)</u>	如の平均と標準誤差が示されます。 グラフ上の点は、 ばらつきを示す 「(B) 」 上書き保存(S) キャンセル ヘルプ

- 8) 分析結果が表示されるので、まず、F検定の結果を確認し、2つのグループ間で、比較す る変数の分散(ばらつき)が異なっているかどうかを見る。
 - (注;現在行っている分析は、2つのグループ間である変数の平均値に差があるかどうかを調べるものであるが、具体的な差の検定方法は、比較する変数の「ばらつき」が異なる場合と、異ならない場合で、別々の検定方法を使う必要がある。このため、まず最初に、分散(ばらつき)に差があるかを調べる「F検定」を行っている。しかし、本来の目的は、2つのグループ間である変数の平均値に差があるかどうかを調べることなので注意すること)

表示されている表の中から、下図のような「等分散性」と書かれた表を見る。この表は、 2つのグループ間で比較する変数の分散(ばらつき)が異なっているかどうかを調べるF検 定の結果を表示したものであるが、その中から、「Pr>F」と書かれたところの下の数値 を確認する。この値が、「F検定の危険率」(2つのグループの分散に差が無い」ことが起 こりうる確率の値)なので、この値が5%(0.05)未満の場合は、「2つのグループの分散 には差がある」とみなすことができる。この値が5%(0.05)以上の場合は、2つのグルー プの分散に差があるかどうかの判定は「保留」されるが、この場合は<u>便宜上「2つのグルー プの分散には差が無い」とみなす</u>。

等分散性						
変数	手法	分子の自由度	分母の自由度	F/I	Pr > F	N
最低血圧(mmHg)	Folded F	697	280	1.96	0.5634	V

9) 8) で、もし、(便宜上)「2つのグループの分散(ばらつき)には差が無い」とみなすこと ができた場合、10)の処理を行う。

もし、「2つのグループの分散(ばらつき)には差がある」とみなすことができた場合は、 11)の処理を行う。 10) 8)の「F 検定」の結果が、(便宜上)「2つのグループの分散(ばらつき)には差が無い」 とみなすことができた場合、表示されている表の中から、下図のような「t 検定」と書かれ た表を見る。その表中から、「分散」の下に「Equal」と書かれた行で、「Pr>|t|」の下 の数値を確認する。この値が、「t 検定の危険率」(2つのグループの間で比較した変数の 平均値に差が無い」ことが起こりうる確率の値)なので、この値が<u>5%(0.05)未満</u>の場合は、 帰無仮説を棄却できる(否定できる)ため、5%の有意水準で「<u>比較した変数の平均値は2</u> つのグループの間で有意な差がある」とみなすことができる。

これに対し、この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できない)ため、2つのグループの間に差があるかどうかの判定は「<u>保留</u>」となり、「<u>2つのグ</u> ループの間で比較した変数の平均値に有意な差があるとはいえない」あるいは「<u>2つのグ</u> ループの間で比較した変数の平均値に有意な差は見られなかった」という結論になる。決 して「2つのグループ間に<u>差が無い」という結論は出せない</u>。これは、危険率が「帰無仮 説が正しいのに、間違って棄却してしまう確率」であるためである。また現実的にも、サ ンプル数を増加させた場合や他の検定方法を使用した場合、有意差が検出される可能性が あるためである。

t 検定					
変数	手法	分散	自由度	t値	Pr > t
最低血圧(mmHg)	Pooled	Equal	977	-2.08	0.0376
最低血圧(mmHg)	Satterthwaite	Unequal	531	-2.11	0.0354

11) 8)の「F検定」の結果が、「2つのグループの分散(ばらつき)には差がある」とみなすことができた場合、表示されている表の中から、下図のような「t検定」と書かれた表を見る。その表中から、「分散」の下に「Unequal」と書かれた行で、「Pr>|t|」の下の数値を確認する。この値が、「t検定の危険率」(2つのグループの間で比較した変数の平均値に差が無い」ことが起こりうる確率の値)なので、この値が<u>5%(0.05)未満</u>の場合は、帰無仮説を棄却できる(否定できる)ため、5%の有意水準で「比較した変数の平均値は2つの グループの間で有意な差がある」とみなすことができる。

これに対し、この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できない)ため、2つのグループの間に差があるかどうかの判定は「<u>保留</u>」となり、「<u>2つのグ</u> ループの間で比較した変数の平均値に有意な差があるとはいえない」あるいは「<u>2つのグ</u> ループの間で比較した変数の平均値に有意な差は見られなかった」という結論になる。決 して「2つのグループ間に<u>差が無い」という結論は出せない</u>。これは、危険率が「帰無仮 説が正しいのに、間違って棄却してしまう確率」であるためである。また現実的にも、サ ンプル数を増加させた場合や他の検定方法を使用した場合、有意差が検出される可能性が あるためである。

t検定					
変数	手法	分散	自由度	t値	Pr > t
最低血圧(mmHg)	Pooled	Equal	977	-2.08	0.0376
最低血圧(mmHg)	Satterthwaite	Unequal	531	-2.11	0.0354

- (3) データを関数変換して分析を行う
 - 1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。
 - 2) 上部メニュー内の [データ]をクリックし、その下に表示されるメニュー内の [フィルタ とクエリ]をクリックする。

🕵 SAS Enterprise Guide - Sample-Load				
ファイル(E) 編集(E) 表示(V) コード(C)	デー	-久(<u>D</u>) 記述統計	H(<u>S</u>) グラフ(<u>G)</u> 分析(<u>A</u>)
1 • G • ¶ 4 % • 6 C × 4		読み取り専用(<u>E</u>))	
SASUSERIMPW_0006(プロセスフロー) 🔹		フィルタとクエリ(<u>E</u>))	
プロジェクトエクスプローラ 🛛 🖡 🗙		行(<u>R</u>)		PN
Sample-Load		列(<u>L</u>)		10
		移動(<u>G</u>)		Ctrl+G
E Bamplexis (Sample\$)	FTE	テニゴル の注意われ()	0)	

3)新しいウィンドウが開くので、左側の変数のリストからグループ分けに用いる変数を選 び、ダブルクリックし、右側の枠内にその変数が表示されていることを確認する。

📓 SASUSERIMPW_0006 のク	エリー クエリビルダ				x
カエリ名(Q): SASUS	ERIMPW_0006 のクエリ	出力名(0):	SASUSER.QUERY_FOR	SASUSER_IMPW_0	変更(<u>C</u>)
🏼 計算列(M) 🛛 💀 パラメー	- 匆史 🏂 検証(V) 🔹 🗟	プレビュー(E) 🔮 オプシ	V∃ン(<u>N</u>) ▼		
 □ 記 追加(① ※ 削除(① ※ 削除(①) □ 記 IMPW_00066 □ □ F1(F1) □ □ F1(F1)<td>② <u> </u></td><td>繋択 │ フィルタデータ う 煙' (吵煙)</td><td>"一タの並べ替え】 入力 IMPW_0006 喫煙</td><td>_ 概要</td><td></td>	② <u> </u>	繋択 │ フィルタデータ う 煙' (吵煙)	"一タの並べ替え】 入力 IMPW_0006 喫煙	_ 概要	
1 1 1 1 1 1 1 1 1 1 1 1 1 1	(NG) / g)(最高血圧() g)(最低血圧()				

4) 左上にある「計算列」をクリックする。

SASUSERIMPW_	0006 のクエリ -	クエリビルダ	
クエリ名(<u>Q</u>):	SASUSERIM	IPW_0006 のク:	τIJ
Ⅲ計算列(M)	♪ パラメータ(<u>P</u>)	☑ 検証(⊻)	• 🗟 7
🔠 追加(I) 💙	K 削除(<u>D</u>)	■結合(J)	データの選び
	2		和女

5)「計算列」というダイアログが開くので、右上の「新規作成」ボタンをクリックし、その 下に新たに表示されたメニュー内の「式の作成」をクリックする。

121 計算列	×
	新規作成(N)▼
	列の再コード化(<u>R</u>)
:	式(01年成(图)
	肖·耶余(<u>D</u>)
	名前の変更(B)
	閉じる
▼ クエリに新しい計算列を追加する(A)	
	11

6)「拡張式エディタ」というダイアログが開くので、左中央の「関数」タグをクリックし、 右下の「関数」の下のリストから、変換に用いる関数を選び、ダブルクリックする。

(例;対数変換を行う場合は、	「LOG」の関数を選択する)
----------------	----------------

S号 拡張式工ディタ 式テキスト(丁): + - × / × AND OR NOT =	$\langle \rangle \langle \langle = \rangle \rangle = =^{x}$	× × × ×
	式のクリア(<u>C</u>)	式へ追加(<u>D</u>)
カテゴリ: ま計 算術 ビットごとの論理演算 文字 文字列の合致 条件 Data Quality (SAS/DQUALITY が必要です) 日付と時間 DBOS 記述統計量 外部ファイル	関数: ABS AIRY ARCOS ARSIN ATAN ATTRC ATTRC ATTRN AVG {All} AVG {Distinct} BAND BETAINV BLSHIFT	
ABS(引数) ABS 関数は、引数と同じ大きさの正の数を返します。		
L	ОК	*+>>t=/

7)「式テキスト」の下のボックスに選択した関数が表示されるので、その中の「<numValue>」 をマウスで選択した後、「BackSpace」キーか「Delete」キーで削除する。

Σ厚拡張式エディタ
式テキスト(<u>T</u>):
LOG(KnumValue>)
+ - * / ** AND (

8) 左中央の「データ」タグをクリックし、「変数値」の下のリストから変換する変数を選んでダブルクリックし、「式テキスト」の下の関数の中にその変数が表示されたことを確認した後、最下部の「0K」ボタンをクリックする。

■■数字の目的である。	X
式テキスト(工):	
LOG(IMPW_0006',GPT(IU/I)'n)	X
+ - * / ** AND OR NOT =	
\frown	式のりリア(©) 式へ追加(D)
データレ 勝知	
使用できる変数(<u>A</u>):	変数値(<u>R</u>):
 □ 読 IMPW 0006 □ 録 FI □ 換字氏名 □ ④ 律部(才) □ ⊕ 年齢能級 □ ● 年令階級 □ ● ■ 	
(逾 年数(年) (逾 本数(本/日) (动 身長(cm)	● Griter(GPT) ● 「APT(GPT) 「 種類の値を取得(G)
l	OK キャンセル ヘルプ

9)「拡張式エディタ」が閉じて、「計算式」ダイアログに戻るので、左のボックス内に表示 されている「Calculation1」をクリックした後、「名前の変更」ボタンをクリックする。

鬱 計算列	×
Calculation1	新規作成(N)▼
	編集(<u>E</u>)
	削除(<u>D</u>)
	名前の変更(<u>R</u>)
I	閉じる
✓ クエリに新しい計算列を追加する(A)	

10)変換後の変数の名前を適当に入力した後、「閉じる」ボタンをクリックする。



- 11)新しい変数名が、グループ分けに用いる変数の下に表示されていることを確認した後、 「実行」ボタンをクリックする。
- 12) グループ分けに用いる変数と関数変換後の比較する変数の表が表示されるので、これら を用いて、この資料の「(1) 各グループ別のヒストグラムおよび、累積度数分布の確率プ ロットの表示」の操作を行い、関数変換後の比較する変数が、2つのグループのどちらで も正規分布しているかどうかを判定する。
 - (注;なお、対応のある2つのデータの差を関数変換した場合は、配布資料「SAS (SAS Enterprise Guide)使用方法(1因子(1種類のデータ)の概要を見る)」の「(1)ヒストグラムおよび、累積度数分布の確率プロットの表示」の操作を行い、関数変換後の対応のある2つのデータの差の値が正規分布しているかどうかを判定する。

もし、正規分布していた場合は、13)の操作に進むが、もし、関数変換後も正規分布していない場合は、配布資料「統計学基礎(3) 2つのグループの1因子(1種類のデータ)の差を検定する」の(12)を参照する。

- 13) もし、関数変換後の比較する変数が、2つのグループのどちらでも正規分布している場合は、この資料の「(2) t 検定(t-test)を行う」の操作を行い、2つのグループ間で比較 する変数の関数変換後の平均値に差があるかどうかを判定する。
- (4) Mann-Whitney のU検定 (Wilcoxon 順位和検定)を行う
 - 1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。
 - 2) Mann-Whitney の U 検定 (Wilcoxon 順位和検定)を含めた比較検定を行う場合、その前に 必ず、以下の操作で、グループ分けに用いる変数の値またはカテゴリーが、比較検定で用 いる2つの値またはカテゴリーであるデータのみに限定する必要がある。
 - (a) 上部メニュー内の [データ]をクリックし、その下に表示されるメニュー内の [フィ ルタとクエリ]をクリックする。

🌠 SAS Enterprise Guide - Sample-Load	
ファイル(E) 編集(E) 表示(V) コード(C)	「データ(<u>D</u>) 記述統計(<u>S</u>) グラフ(<u>G</u>) 分析(<u>A</u>)
1 1 • 6 • 5 4 + 6 6 X •	□ 読み取り専用(E)
SASUSERIMPW_0006(データのインポ 🔹	フィルタとクエリ(E)
プロジェクトエクスプローラ + *) 🚍 行(<u>R</u>)
Sample-Load	■ 歹(L)
	移動(<u>G</u>) Ctrl+G

(b) 下図のようなダイアログが開くので、左側の変数リストから、グループ分けに用い る変数と、比較する変数を、右側の「データの選択」の下のボックスにドラッグアン ドドロップし、これらの変数名が表示されていることを確認する。

SASUSER IMP	W_0006 のクエリ - クエリビルタ	ž				
クエリ名(Q):	SASUSERIMPW_0006	のクエリ	出力名(0):	SASUSER.QUERY_FC	R_SASUSER_IMPW_0	変更(<u>C</u>)
⊞ 計算列(M)	🐻 パラメータ(P) 🔊 検証	w •!	🗟 プレビュー(E) 🙆 オブ	9∋>(<u>N</u>) ▼		
		_				
🖽 追加(T)	×前除(D) 助結合(J) データ	の選択 フィルタデータ :	データの並べ替え		
🖃 I IMPW 00	006	▲ 列:	名	入力	概要	
🗍 🗍 🔞 F1 (F1)	- 🔺	'喫煙' (喫煙)	IMPW_0006.喫煙		
🍈 漢字	(漢字氏名)	1	'GOT(IU/I)' (GOT(IU/	IMPW_0006.GOT(IU/I)		
——————————————————————————————————————](性別)					
一個年齡	(才)(年齡(才))					<u>-</u> (†
(1) 牛労	階級(牛令階級) i(ntalme)					
() 年期	(年)(年期(年))					
	(本/日)(本教(本/日))					
- 🔞 身長	(cm)(身長(cm))					
	(Kg)(体 <u>重</u> (Kg))					
- 🔞 最高	i血圧(mmHg)(最高血圧()					
	:血圧(mmHg)(最低血圧()					
- 原本						
小田	(日(水重日) 9面(昆湖面)					
·····································	134(万個/mm3)(赤血球()					
	(g/dl)(血色素(g/dl))		6 L M			
-0.421	5Jット(%)(ヘマトクリット(%))	要	約クループ			
-@ GO	t(IU/I) (GOT(IU/I))		グループを自動選択する(U)	グルー	·プの編集(<u>G</u>)…

(c)「データの選択」の右隣の「フィルタデータ」のタグをクリックし、その下に表示 されたボックスに、グループ分けに用いる変数をドラッグアンドドロップする。

🐮 SASUSERIMPW_0006 のクエリ - クエリビルダ		×
クエリ名(Q): SASUSERIMPW_0006 のク	UU 出力名(O): SASUSERQUERY_FOR_SASUSER_IMPW_0	変更(<u>C</u>)
🏢 計算列(M) 🛛 👼 パラメータ(P) 💋 検証(V)		
団 追加(刀 ※ 削除(①) 町 結合(J) □ 副 IMPW_0006 ▲ □ ③ F1 (F1) ▲ □ ④ 年齢(才) (年齢(才)) ● □ ④ 年齢(オ) (年齢(オ)) ● □ ○ 年数(年) (年数(年)) ● □ ● 年数(年) (年数(年)) ● □ ● 長(cm) (身長(cm)) ● □ ● 長(cm) (月長(cm)) ●	データの選択 フィルタデータ ータの並べ替え 未加工データ(こフィルタを適用するには、ここに列をドロップしてください。	

(d)「フィルタの編集」というダイアログが開くので、「値(V)」の行の右端の黒い逆三

角形ボタンをクリックする。

列:	IMPW_0006 喫煙
演算子(0):	次の値に等しい
值色在	
12(1)	
- フィルタ定義	
- フィルタ定義	Ē = "
- フィルタ定義 IMPW_0006 喫烟	Ē = "
- フィルタ定義	Ē = " "囲む(E)
 フィルタ定義 IMPW_0006 喫閉 IMPW_0006 厚閉 IMPW_0006 厚別 IMPW_0006 FN IMPW_000	E = " *囲む(E) 適用された日 付を使用する(U)
 フィルタ定義 「MPW_0006 時間 「MPW_0006 # [MPW_0006 # [MP	『囲む(E) 適用された日付を使用する(①)

(e)新しいダイアログが更に表示されるので、その下部にある「値の取得」ボタンをク リックする。



(f) グループ分けに用いる変数の値やカテゴリーのリストが表示されるので、まず、2 つのグループの一方の値またはカテゴリーをクリックする。

		×
データ値列	パラメータ	
値	出力形式が適用された	:値 [
<u>A</u>	A	
Ċ	Č	
<u> </u>		
値の追加取得	(v)	
		11

(g)「フィルタの編集」のダイアログに戻るので、「値」の右に、選択した値またはカ テゴリーが表示されていることを確認し、下の「OK」ボタンをクリックする。

📓 フィルタの編集	×
列:	IMPW_0006.1契/里
演算子(<u>O</u>):	次の値に等しい
値(⊻):	A
フィルタ定義 IMPW_0006 喫煙 = 'A'	
☑ 値を引用符で囲む(Ð
▶ 出力形式が適用され	1た日付を使用する(旦)

(h) 最初のダイアログに戻り、選択した値またはカテゴリーが表示されていることを確 認する

SASUSERIM	PW_0006 のクエリ	- クエリビルダ					×
クエリ名(Q):	SASUSER.	IMPW_0006 のク:	ΕIJ	出力名(_):	SASUSER.QUERY_FOR_SASUSER_IMPW_0	変更(<u>C</u>)	
⊞計算列(M)	🕞 パラメータ(<u>P</u>	り 🔊 検証(У)	- B 7	⁷ レビュー(E) 📓 オプシ	ν∃)(<u>N</u>) ▼		
田 追加(I) □ 副 IMPW(□ □ □ 目 □ □ □ 第 □ □ □ 10 10 10 10 10 10 10 10 10 10 10 10 10 1	★ 削除(D) 0006 (F1) 字氏名(漢字氏: 別(性別) (性別)	[■] 結合(」) ▲ 名)	データの選 「未加工デ 「IMPW」	択 フィルタデータ テ ータにフィルタを適用 - 0006 喫煙 = " A"	*-タの並べ替え	V 2	

(i) (c)~(h)の操作を繰り返し、2つのグループのもう一方の値またはカテゴリーを設 定する。 (j) 2つのグループの値またはカテゴリーが選択されていることを確認した後、その下の「フィルタの組み合せ」ボタンをクリックする。

	A CONTRACTOR OF	1.6.4		4010	
🐻 SASUSER IM	PW_0006 のクエリー {	ウエリビルダ			×
クエリ名(<u>Q</u>):	SASUSERIM	PW_0006 のクエリ	出力名(0):	SASUSERQUERY_FOR_SASUSER_IMPW_0	変更(<u>C</u>)
⊞ 計算列(M)	🔜 パラメータ(P) ,	≫)検証(⊻) ・	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	ション(<u>N</u>) ・	
🔡 追加(T)	★削除(<u>D</u>) ■	■ 結合(」) デー	2の選択 フィルタデータ	データの並べ替え	
_ □ IMPW_0	0006		ロ上ナータにフィルタを適用		
60 F1	(F1) 新成点()港南原点:		MPW_0006鸣卿生 = "A" MPW_0006鸣槊性 = "C"		<u> </u>
▲ 世知 一 ▲ 世知	子氏名(漢子氏名) 別(性別)	,	-		2
- @ 年齢	論(才)(年齡(才))				
- (1) 年4	令階級(年令階級) 栗(喫煙))			
@ 年	敖(年) (年数(年))				
	波(本/日)(本数(本 ミ(cm)(身長(cm))	(日))			
	重(Kg)(体重(Kg))	í i	VHERE IMPW_0006.喫煙	= ″ A″ AND IMPW_0006.喫煙 = ″ C″	<u> </u>
	高血圧(mmHg)(最 Emp⊑(mmHg)(長	高血圧()			_
	3.血)至(mmHe)(最 糖(尿糖)		フィルタの組み合せ(F)		
	蛋白(尿蛋白)				
脉?	倉町 (水)塩町)	「思報	可テータにフィルタを適用―		

- (k)「フィルタの組み合せ」というダイアログが新たに開くので、2つの漏斗の絵の間にある「AND ▼」をクリックし、その下に表示されたメニューから「OR に変更」をク
 - リックする。



- (1)「AND ▼」が「OR ▼」に変ったことを確認した後、最下部の「OK」ボタンをクリッ クする。
- (m) 最初のダイアログに戻るので「実行」ボタンをクリックする。
- (n) 画面上に、グループ分けに使用する変数と、比較する変数のみの表が、新たに表示 されたことを確認する。
- 3) 上部メニュー内の [分析]をクリックし、その下に表示されるメニュー内の [分散分析] をクリックし、更にその右に表示されたメニュー内の「ノンパラメトリックな一元配置分 散分析」をクリックする。

ų	SAS Enterprise Guide - Sample-Load								
	ファイル(E) 編集(E) 表示(⊻) コード(C)	データ(<u>D</u>)	記述統計(<u>S</u>)	グラフ(<u>G</u>) [分析	F(A) アドイン(1)	OLAP(<u>O</u>)	<u>_</u>	ル(工) ウィンドウ(W) ヘルプ(円)
1	1 1 • 6 • 6 • 6 × 1 4	a Q	ູ 🍇 ສູງດະນ	ェクトデザイナ		分散分析(<u>A</u>)		Ы	t 検定(<u>1</u>)
l	QUERY FOR SASUSER IMPW 0006 -					回帰分析(<u>R</u>)		颅	一元配置分散分析(0)
I			*			多変量解析(<u>M</u>)		渔	ノンパラメトリックな一元配置分散分析(N)
l	70919119270-7	80870	ジェクトデザイナ	SASL		生存時間分析(S)		st.	線形モデル(L)
I	-Load		喫煙 🖌	GOTI				1.22	
I	セスフロー					上程能刀分析(目))	<u>×.</u>	旭日モノハ(四)…
I	Samplexis (Sample\$)		B			管理図(0)			
I	🖉 データのインポート	2 C			-				
I		3 C			lion	バレート図(円)			
1	<u>المعامم المعامم المعام</u>	4 C				時系列分析(工)			
l	SASUSERIMPW 0006	5 C				-17(2-17)-6(C <u>T</u>)			
	SASUSERIMPW_0006	6 A	·····		ę.	モデルのスコアリング	ΰ(<u>L</u>)		
18	H		······		_			_	

4)新しいダイアログが開くので、「変数リスト」からグループ分けに用いる変数をクリックした後、その右の右向き矢印のボタンをクリックし、新たに表示されたメニューから「独立変数」をクリックし、「タスクの役割」の下の「独立変数」の下に、その変数が表示されたことを確認する。

QUERY_FOR_SASUSER_IMPW_0006_0005 の ノンパラメトリックな一元配置分散分析					
タスクの役割	<u>タスクの役割</u>				
万1/1 正確な p 値	変数リスト(<u>A</u>)	タスクの役割(<u>E</u>):			
結果	名前	🞯 従属変数(分析変数)			
97.FJU					
		◎ 法立案数 (進択の上陸 1)			
		🞯 グループ分析			
		 ・ ・ ・			
		[
		- グループ分析			

5)「変数リスト」から比較する変数をクリックした後、その右の右向き矢印のボタンをクリッ クし、新たに表示されたメニューから「従属変数」をクリックし、「タスクの役割」の下 の「従属変数」の下に、その変数名が表示されたことを確認する。

A QUERY_FOR_S	ASUSER_IMPW_0006_0005 の ノンパラ>	トリックな一元配置分散分析
タスクの役割	タスクの役割	
一分析	変数リスト(<u>A</u>)	タスクの役害!(E):
結果	名前	💿 従属変数(分析変数)
タイトル	🗌 💩 喫煙	
	🔞 GOT(IU/I)	🛞 独立変数 (選択の上限:1)
	l ·	
		⑧ グループ分析
		 ・ ・ ・
		◎ 従属変数(分析変数)
		· · · · · · · · · · · · · · · · · · ·
		- 齋 ヴループ分析

6) 左端のリストから「分析」をクリックし、その右に新たに表示された「検定に用いるス コア」の下の「Wilcoxon」の前のボックスにチェックを入れ、最下部の「実行」ボタンを クリックする。

CODERY_FOR_SASUSER_IMPW_0006	0005のノンパラメトリックな一元配置分散分析
QUERY_FOR_SASUSER_IMPW_0000 タスクの役割 予研 正確な p 値 結果 タイトル ダイトル ダイトル マーWilcoxon(W) □ Median(M) □ Savaee(G) □ Van der Waer (V) □ Ansari-Bradit □ Kutz の正規元 □ Mood(Q)	20005 の ノンパラメトリックな一元配置分散分析 ✓ 経験分布関数統計量 (EDF) を計算する(E) 「 欠損値を分類水準として含める() 「 欠損値を分類水準として含める() 「 速続性の補正を表示しない(U) (A) コア(K)
☐ Noong ☐ Siegel-Tukey ☐ 生データ(D)	D 注: [正確な p 値] ページで "正確な p 値"を使用可能に し、 磁果」 ページで "統計量"を使用可能にするためには [検定に用いるスコア] チェックボックスをオンにしてください。
	<u> </u>
<u> </u>	実行(B) 上書き保存(S) キャンセル ヘルプ

(補足;「Wilcoxon」の下にある「Median」の前のボックスにチェックを入れると、「中 央値検定(メディアン検定)」が行われる)

8) 分析結果が表示されるので、下図のような「Wilcoxon の順位和検定(2標本)」と書かれた表を見る。その表中から「正規近似」の3つ下にある「両側 Pr>|Z|」の右横の数値を確認する。この値が、「Mann-WhitneyのU検定(Wilcoxon 順位和検定)の危険率」(2つのグループの間で比較した変数に差が無い」ことが起こりうる確率の値)なので、この値が<u>5%(0.05)未満</u>の場合は、帰無仮説を棄却できる(否定できる)ため、5%の有意水準で「<u>比</u>較した変数は2つのグループの間で有意な差がある」とみなすことができる。

これに対し、この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できな い)ため、2つのグループの間に差があるかどうかの判定は「<u>保留</u>」となり、「<u>2つのグ</u> ループの間で比較した変数に有意な差があるとはいえない」あるいは「<u>2つのグループの</u> 間で比較した変数に有意な差は見られなかった」という結論になる。決して「2つのグルー プ間で比較した変数に<u>差が無い」という結論は出せない</u>。これは、危険率が「帰無仮説が 正しいのに、間違って棄却してしまう確率」であるためである。また現実的にも、サンプ ル数を増加させた場合や他の検定方法を使用した場合、有意差が検出される可能性がある ためである。

Wilcoxonの順位和検定(2標本)				
統計量	99725.0000			
正規近似				
Z	-0.5769			
片側 Pr < Z	0.2820			
両側 Pr > Z	0.5640			
t分布で近似				
片側 Pr < Z	0.2821			
両側 Pr > Z	0.5642			
Zには 0.5の連続性の補正が含まれています。				

- (5) 記号の順位(順序)尺度データを数字に変換する
 - 1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。
 - 2) 上部メニュー内の [データ]をクリックし、その下に表示されるメニュー内の [フィルタ とクエリ]をクリックする。



3)新しいウィンドウが開くので、左側の変数のリストからグループ分けに用いる変数を選び、ダブルクリックし、更に、数字に変換する変数を選び、ダブルクリックし、右側の枠内にその2つの変数が表示されていることを確認する。

📸 Samplexis (Sample\$) のクエリー クエリビルダ 📃 🔀						
クエリ名(Q): Samplexis (Sample\$) のク	エリ 出力名(0):	SASUSER.QUERY_FOR	SAMPLE_XLS_SAI	変更(<u>C</u>)		
⊞ 計算列(M) 📑 パラメータ(P) 😼 検証(V)	• 🗒 ブレビュー(E) 🖺 オコ	プション(<u>N</u>) ・				
🔠 追加(I) 🗙 削除(D) 🖽 結合(J)	データの選択 フィルタデータ	データの並べ替え		1		
🖃 🎥 Sample xis Sample 🔺	列名	入力	概要			
F1 (F1)	💩 喫煙 (喫煙)	Sample_xls_Sample喫				
──▲ 漢字氏名(漢字氏名)	🔊 '尿糖' (尿糖)	Sample_xls_Sample_尿				
				÷		
				-\$-		
- (1) 身長(cm) (身長(cm)) - ○ (ナモ(ル-) (ナモ(ル-))						

4) 左上にある「計算列」をクリックする。

	SASUSER IMP	₩_0006 のクエリ・	- クエリビルダ	
	クエリ・タ(の):	SASUSERI	MPW_0006 のク	IU
(⊞計算列(M)	▶ パラメータ(P)) 🔊 検証(⊻)	- EZ 7
		¥ 削除(D)	==a 結合(J)	データの選
		006		別々

5)「計算列」というダイアログが開くので、右上の「新規作成」ボタンをクリックし、その 下に新たに表示されたメニュー内の「列の再コード化」をクリックする。



6)「項目の選択」というダイアログが開くので、「項目の選択」の下のリストから数字に変換する変数をクリックし、最下部の「継続」ボタンをクリックする。

I	11 項目の選択	×I
	項目の選択	
ł	Sample_xis_Sample_	
	· · · · · · · · · · · · · · · · · · ·	
	□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □	
1	(1) 年令階級 (1) 中令階級	
	(1) 本致(本/日) (1) 身長(cm)	
	最低血圧(mmHg)	
	▲ 尿潜血 → → 赤血球(万個/mm3)	1 [
		1
,	維続(C) キャンセル	
¢		//

7)「列の再コード化」というダイアログが開くので、もし、新しい変数名を付ける場合は、 「新しい列名」の右横のボックス内の名前を変更する。次に、その下の「追加」ボタンを クリックする。

III 列の再コード化- Sample♪	xls_Sample_尿_糖 🛛 🔀
新しい列名(<u>N</u>):	Recode_尿
置換(<u>E</u>):	追加(<u>A</u>) 肖耶余(<u>R</u>)
置換	置換後
他の値	
すべての他の値を次の値でお	置き換える:
● 現在の値(C)	
○ 欠損値(<u>M</u>)	
○ 値(丁):	
	,
「新しい列の種類」	
④ 文字(H)	C 数値(U)
OK	***/7/1 / /1/7 /
01	

8)「置換の指定」というダイアログが開くので、中央右にある、「値の取得」ボタンをクリッ クする。

ş	置換の指定
(値の置換 範囲の置換
l	
	値の取得(⊻)▼
-	この値で置き換える(<u>W</u>):
Γ	
ł	
	ヒンド 欠損値は、空日 ご表示されます。
	ビンド、欠損値は、空日 じ表示されます。
	ビンド、欠損値は、空日 ご表示されます。

9)新しいダイアログが開くので、データ値のリストからどれかの値をクリックし、「OK」 ボタンをクリックする。_____

		×
データ値		
値	出力形式が適用された値	I
-	-	
+	+	
+-	+-	
++++	+++	
1		
値の追加取得()) C OK キャンセル	1
		-//

10)「置換の指定」ダイアログに戻るので、「この値で置き換える」の下のボックスに変換 する数字の値を入力し、最下部の「0K」ボタンをクリックする。

IN THE A LARGE AND A L	
鬱 置換の指定	×
値の置換 範囲の置換	
-	
	値の取得(⊻)▼
この値で置き換える(W):	
0	
ヒント: 欠損値は、空白 で表示されます。	
OK キャンセノ	

11) 7)~10)を繰り返して、再分類前の値と再分類後の値を全て設定する。

例)例えば、尿糖の「-」「+-」「+」「++」「+++」を数字に変換する場合は、 最初に9)で「-」を選んだ後、10)で「0」を入力、 次に9)で「+-」を選んだ後、10)で「1」を入力、 更に9)で「+」を選んだ後、10)で「2」を入力、 その次に9)で「++」を選んだ後、10)で「3」を入力、 最後に9)で「+++」を選んだ後、10)で「4」を入力する

12) 全ての設定が終了したら、「列の再コード化」ダイアログの下部にある「新しい列 の種類」の所で、「数値」の前の丸をクリックしてマークを付けた後、最下部の「OK」 ボタンをクリックする。

鬱 列の再コード化- Sample_xis	:_Sample_尿塘	x
新しい列名(<u>N</u>):	Recode_尿糖	
置換(<u>E</u>):	追加(<u>A</u>)	削除(<u>R</u>)
置換	置換後	
-	0	
+-	1	
++	3	
++++	4	
J		
- 他の値		
 すべての他の値を次の値で置	き換える:	
	C170C0-	
C 現在(別値(G)		
⊙ 欠損値(M)		
(二)(市(工))		
	I	
オイトリップリップが学売		
「新作しに19月の時里発見」		
○ 文字(H)	○ 新値(1))	
	MALE COL	
	-	
OK	キャンセル	ヘルプ

- 13)「計算列」ダイアログに戻るので、最下部の「閉じる」ボタンをクリックする。
- 14) 最初のダイアログに戻るので、最下部の「実行」ボタンをクリックする。
- 15) グループ分けに用いる変数と、数字変換前のデータと数字変換後のデータの表が表示されるので、正しく数字変換されていることを確認する。
- 16) 上記の表を用いて、適切な統計手法で分析を行う。

(6) χ²検定 (カイ二乗検定)を行う

1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。

- 2) χ²検定(カイ二乗検定)の場合、比較するグループが3つ以上でも特に問題は無いが、分析結果を判断する際に、グループが2つである方が分かりやすいため、以下の操作で、グループ分けに用いる変数の値またはカテゴリーを、2つの値またはカテゴリーであるデータのみに限定した方が良い(再分類によって、2つのグループにまとめ直す方法でも良い)。 また、比較するデータ(変数)についても、分析後の判断を分かりやすくするためには、カテゴリーを限定するか、再分類することで、少なくすることが望ましい(グループ分けする変数も、比較する変数も、カテゴリーが2つである場合が、最も分かりやすいので、最良と言える)。
 - (a) 上部メニュー内の [データ]をクリックし、その下に表示されるメニュー内の [フィ ルタとクエリ]をクリックする。

8 CAC Futurning Oxida - Samala Load		
SAS Enterprise Guide - Sample-Load		
ファイル(E) 編集(E) 表示(V) コード(C)	データ(<u>D</u>) 記述統計(<u>S</u>)	グラフ(<u>G</u>) 分析(<u>A</u>)
1 1 • 6 • 6 4 • 6 6 × 0	✓ 読み取り専用(E)	
SASUSERIMPW_0006(データのインポ 🔹 🖕	フィルタとクエリ(E)	
プロジェクトエクスプローラ 🛛 🖡 🗙	行(<u>R</u>)	PV
💽 Sample-Load	Ⅲ 列(L)	12
	移動(<u>G</u>)	Ctrl+G

(b) 下図のようなダイアログが開くので、左側の変数リストから、グループ分けに用い る変数と、比較する変数を、右側の「データの選択」の下のボックスにドラッグアン ドドロップし、これらの変数名が表示されていることを確認する。

SASUSERIMPW_0006 のクエリー クエリビルタ				×
クエリ名(Q): SASUSERIMPW_0006の	ウエリ 出力名(0):	SASUSERQUERY_FO	R_SASUSER_IMPW_0	変更(<u>C</u>)
🌐 計算列(M) 🛛 🧱 パラメータ(P) 💋 検証(V) → 🗟 ブレビュー(E) 😫 オブ	€∋)(<u>N</u>) ▼		
	データの選択 フィルタデータ	データの並べ替え		1
□ IMPW_0006 ④ F1 (F1) ▲ 漢字氏名(漢字氏名) ● 性別(性別) ● 年齢(才)(年齢(オ)) ● 年齢(オ)(年齢(オ)) ● 年齢(オ)(年齢(オ)) ● 年勤(年)(年勤(本)) ● ● ●	<u>列名</u> ▲ '喫煙' (喫煙) (2) '年令階級' (年令階級)	入力 IMPW_0006.喫煙 IMPW_0006.年令階級	概要	

(c)「データの選択」の右隣の「フィルタデータ」のタグをクリックし、その下に表示

されたボックスに、グループ分けに用いる変数をドラッグアンドドロップする。

SHOUSERIMPW	10006 000Th - 0ThEM8				X
クエリ名(Q):	SASUSERIMPW_0006 D	עדע	出力名(0):	SASUSERQUERY_FOR_SASUSER_IMPW_0	変更(<u>C</u>)
⊞計算列(M)	🍃 パラメータ(P) 💋 検証(V)	• 🗔	プレビュー(E) 🖹 オ	プション(<u>N</u>) ▼	
□ 追加(丁) □ 認 IMPW.000 □ ③ F1 (1 □ ④ F1 (1 □ ④ 年前(1) □ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓	★ 削除(D) 型:結合(J) 36 (1) 5(名(漢字氏名) (1性別) 才)(年齢(才)) 皆級(年令階級) (喫煙) 年)(年数(年)) 本/日)(本数(本/日)) cm)(身長(cm)) Kg)(体重(Kg))	データの浅 一未加工:	R フィルタデータ デーション ロカナ データ (オブション) データ	タの並べ替え にフィルタを適用するには、ここに列をドロップしてください。	· R X

(d)「フィルタの編集」というダイアログが開くので、「値(V)」の行の右端の黒い逆三 角形ボタンをクリックする。

👿 フィルタの編集	
列:	IMPW_0006.年令階級
演算子(0):	次の値に等しい
値(⊻):	
	U
┌ フィルタ定義 ――	
[IMPW/0006年会	ji 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
IMPW_0006.年令	皆級 =
IMPW_0006.年令	皆級 =
IMPW_0006.年令	皆級 = 目む(<u>E</u>)

(e)新しいダイアログが更に表示されるので、その下部にある「値の取得」ボタンをク リックする。

一一一个 一个	 パラメータ		×
値	1022 2	- 出力形式が適用され	1.た値 🔤
値の取得(<u>G</u>)			in.

(f) グループ分けに用いる変数の値やカテゴリーのリストが表示されるので、まず、2 つのグループの一方の値またはカテゴリーをクリックする。

	×
データ値 列 パラメータ	
値	出力形式が適用された値
20	20
30	30
50	40 50
000	
<u> </u>	
(中の)自由市田(日)(人	
10())101/14以(守(型)	

(g)「フィルタの編集」のダイアログに戻るので、「値」の右に、選択した値またはカ テゴリーが表示されていることを確認し、下の「OK」ボタンをクリックする。

鬱 フィルタの編集		×
矛 リ:	IMPW_0006.年令階級	
演算子(<u>0</u>):	次の値に等しい	
値(以):	30	
─ フィルタ定義 IMPW_0006年令階級	§ = 30	1
□ 値を引用符で囲む	5(E)	
止力形式が適用された。	5れた日付を使用する(山)	
	OK ++>211 /113	

(h) 最初のダイアログに戻り、選択した値またはカテゴリーが表示されていることを確 認する

	A COLORADO DE	1.6.4	· · · · · · · · · · · · · · · · · · ·	E 1 40 10		
📓 SASUSERIMI	PW_0006 のクエリー	クエリビルダ				x
クエリ名(<u>Q</u>):	SASUSERIM	PW_0006 のクエリ	出力名(0):	SASUSER.QUERY_F	DR_SASUSER_IMPW_0	変更(<u>C</u>)
⊞計算列(<u>M</u>)	🕞 パラメータ(<u>P</u>)	🔊 検証(⊻) 🔹 🔹	🗟 プレビュー(E) 🔮 オプジ	/∃ン(<u>N</u>) ▼		
田道加(① 日間 IMPW_(★ 削除(D) 目 0006 (F1) 字氏名(漢字氏名)		タの選択 フィルタデータ う DIエデータにフィルタを適用 - IMPW_0006 年令階級 = 30	"~タの並べ替え】		

- (i) (c)~(h)の操作を繰り返し、2つのグループのもう一方の値またはカテゴリーを設 定する。
- (j) もし、比較する変数もカテゴリーを限定する場合は、(c)~(h)の操作を繰り返し、 分析に用いるカテゴリーを設定する。
- (k) 2つのグループの値またはカテゴリー、および、比較する変数も限定した場合は、 それらの値やカテゴリーが選択されていることを確認した後、その下の「フィルタの 組み合せ」ボタンをクリックする。

SASUSERIMPW_0000 ())	19 - 9196769			2
. クエリ名(Q): SASUS	ERIMPW_0006 のクエリ	出力名(0):	SASUSERQUERY_FOR_SASUSER_IMPW_0	j 変更(<u>C</u>)
🏢 計算列(M) 📑 パラメー	タ(P) 🔊 検証(V) 🔹	蔵 ブレビュー(E) 🔮 オブシ	• (<u>N</u>) →	
世 追加(1))	200選択 フィルタデータ ロエデータにフィルタを適用 MPW 0006 年令階級 = 30 MPW 0006 年令階級 = 30 MPW 0006 単厚 = "A" MPW 0006 喫煙 = "A" MPW 0006 喫煙 = "A" AHERE IMPW 0006 年令階額 PW 0006 喫煙 = "A" AND フィルタの組み 合世(E)…	- タの並べ替え 皮 = 30 AND IMPW_0006年令階級 = 50 AND IMPW_0006喫煙 = "C"	X X X

(1)「フィルタの組み合せ」というダイアログが新たに開くので、同じ変数の2つの漏
 斗の絵の間にある「AND ▼」をクリックし、その下に表示されたメニューから「OR に
 変更」をクリックし、「AND ▼」が「OR ▼」に変ったことを確認する。



(m)「Ctr1」(コントロール)キーを押しながら、同じ変数の漏斗をクリックして色を反転させた後、左端上の「グループ」をクリックし、同じ変数の漏斗を括弧でまとめた後、最下部の「OK」ボタンを押す。



- (n) 最初のダイアログに戻るので「実行」ボタンをクリックする。
- (o) 画面上に、グループ分けに使用する変数と、比較する変数のみの表が、新たに表示 されたことを確認する。
- 3) 上部メニュー内の [記述統計]をクリックし、その下に表示されるメニュー内の [分割表 分析]をクリックする。



4)分割表分析のダイアログが開くので、「変数リスト」からグループ分けに用いる変数と 比較する変数をそれぞれクリックした後、それぞれで、その右の右向き矢印のボタンをク リックし、新たに表示されたメニューから「表変数」をクリックし、「タスクの役割」の 下の「表変数」の下に、グループ分けに用いる変数と比較する変数の両方が表示されたこ とを確認する。



5) ダイアログの左端のリストから「テーブル」をクリックし、新たに右に表示された「テー ブルで許可されている変数」の下のリストから、最初にグループ分けに用いる変数をその 右の「プレビュー」の下の領域にドラッグアンドドロップし、続いて、比較する変数を同 じ領域にドラッグアンドドロップする。その結果、「プレビュー」の下の表で、上にグルー プ分けに用いる変数、横に比較する変数が表示されていることを確認する。更に、その下 の「生成されるテーブル」の下に「(比較する変数名) by (グループ分けに用いる変数名)」 と記載されていることも確認する。

QUERY_FOR_SASUSE	R_IMPW_0006_0007 の 分割表分析	<u>.</u>	×
タスクの役割	テーブル		
テーブル テーブル セル統計量 裏統計量 関連 一致 順序付けされた差 傾向検定 計算オブション 結果 セル統計量の結果 表統計量の結果 タイトル	テーブルで許可されている変数(<u>V</u>): ② 喫煙 ② 年令階級	フレヒュー: < ここに変数をドラッグしてください>	*
	生成されるテーブル(T):		
	\checkmark		
QUERY_FOR_SASUSEF	R_IMPW_0006_0007 の 分割表分析		×
タスクの役割	テーブル		
テーノル セル統計量	テーブルで許可されている変数(⊻):	プレビュー:	
- 表統計量 - 関連		"年令階級"	1
一致 順序付けされた差			
傾向検定			
計算オノンヨン 結果			
セル統計量の結果 表統計量の結果			
27 F/V			
	王カxc4loテーノルL/ 『嗖煙" by "年令階級"		-
-	· 가하게 가는 · · · · · · · · · · · · · · · · · ·	テノザチリット	

6)ダイアログの左端のリストから「セル統計量」をクリックし、新たに右に表示された「使用できる統計値」の下で、既にチェックの入っているものはそのままにして、「期待セル度数」の前のボックスにチェックを入れる。

🏢 QUERY_FOR_SASUSER_IMPW_0006_0007 の 分割表分析			
タスクの役割	セル統計量		
テーブル セル統計量 表統計量 関連 一致 順序付けされた差 傾向検定 計算オプション 結果 セル統計量の結果 表統計量の結果 タイトル	使用できる統計値 「列の累積パーセント(M) 「行のパーセント(M) マ列のパーセント(U) マセルの度数(E) 「セルのパーセント(P) 「欠損値の度数(V) 「セルごとの x2 乗続計量(L) 「セルごとの x2 乗続計量(L) 「セルごとの x2 乗続計量(L) 「セルごとの x2 乗続計量(L) 「セルごとの x2 乗続計量(L) 「日本の基本の期待値からの偏差(D) 「割将ゼル度数(E) 「日本の表示のパーセント(T) 「出力データセット(こパーセントで含める(N)		

7) ダイアログの左端のリストから「表統計量」の下の「関連」をクリックし、新たに右に 表示された「独立性の検定」の下で、「χ2 乗検定(Pearson、.....」の前のボックスと、 「r×c 表に対するフィッシャーの直接法」の前のボックスにチェックを入れ、最下部の「実 行」ボタンをクリックする。

QUERY_FOR_SASUSE	R_IMPW_0006_0007 の 分割表分析		×
タスクの役割	表統計量>関連		
 テーブル セル統計量 表統計量 関連 一致 順序付けされた差 傾向検定 計算オプション 結果 セル統計量の結果 タイトル タイトル 	独立性の検定 ×2 乗検定(1) 「Pearson、尤度比、 Cochran-Mantel-Haenszel の ×2 乗 検定、2x2 表に対する Fisher の直接注 「正確な p 値(E) 「▼ x c 表に対するフィッシャーの直接注(F) Cochran-Mantel-Haenszel 統計量 「CMH 統計量(L) スコアの種類:表(B) 2×2 より大きい表に対する、Fisher の正確な p 値 も呼ばれます。	 独立性の統計量 統計量(M) (2 × 2 表に対する相対危険度およびオッズ比を含む) オッズ比に対する正確な p 値および信頼区間(P) 統計量が 0 に等しいという検定(D) 2 × 2 表に対するリスクの差(K) 2 × 2 表に対する相対危険度(L) 次の場合は [計算オブション] ページを使用してくださ(い。 正確な p 値の計算時に制限時間を設ける。 CMH 統計量のスコアを選択する。 	
, 	, 	上書き保存(S) キャンセル ヘルプ	

- 8) 分析結果が表示されるので、まず、下図のような「FREQ プロシジャ」と書かれた下の表 を見る。その表の各セルで、真ん中の「期待度数」を全て確認し、下記の条件に合うもの があれば、その指示に従う。
 - (a) 2×2表(分割表の行も列も2個のもの)で、どれかのセルの期待度数が5未満の 場合、10)に進み、そこに記載された方法で結果を判定する。

(b) <u>2×2表以外の場合で、セルの期待度数が以下のどちらかの条件に当てはまる場</u>合

(1) セルのうちで 20%の期待度数が 5 未満

(2) セルの中に、1つでも期待度数が0のものがある

どちらかの条件に当てはまる場合は、11)に進み、そこに記載された方法で分析 する。

上記の(a)(b)どちらの条件にも当てはまらない場合は、9)に進み、そこに記載された 方法で結果を判定する。

度数	表:喫煙*年令階級						
列のバーセント		年令階級(年令階級)				
	喫煙(喫煙)	30	50	合計			
	A	259 233.33	<u>192</u> 217.67	451			
		82.75	65.75				
	с	54	100	154			
		17.25	34.25				
	合計	313	292	605			

FREQ プロシジャ

9) 8)の分割表のすぐ下にある下図のような表を見て、「カイ2乗値」の行の「p値」を確認 し、以下の方法で判定する。

統計量	自由度	値	p値
カイ2乗値	1	22.9925	<.0001
尤度比カイ2乗値	1	23.2135	<.0001
連続性補正カイ2乗値	1	22.1056	<.0001
Mantel-Haenszelのカイ2乗値	1	22.9545	<.0001
ファイ係数		0.1949	
一致係数		0.1913	
Cramerの V 統計量		0.1949	

この値は「2つのグループ間で、比較したデータの割合に差が無い」ことが起こりう る確率の値である。従って、この値が<u>5%(0.05)未満</u>の場合は、帰無仮説を棄却できる (否定できる)ため、5%の有意水準で「<u>比較したデータ(の割合)は2つのグループの間</u> で有意な差がある」とみなすことができる。

この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できない)ため、2 つのグループの間で比較したデータの割合に差があるかどうかの判定は「<u>保留</u>」となり、 「<u>2つのグループの間で比較したデータ(の割合)に有意な差があるとはいえない</u>」ある いは「<u>2つのグループの間で比較したデータ(の割合)に有意な差は見られなかった</u>」と いう結論になる。決して「2つのグループ間で比較したデータ(の割合)に差が無い」と **いう結論は出せない**。これは、危険率が「帰無仮説が正しいのに、間違って棄却してし まう確率」であるためである。また現実的にも、サンプル数を増加させた場合や他の検 定方法を使用した場合、有意差が検出される可能性があるためである。

 2×2表の分析で、4つのセルのうち1つでも期待度数が5未満のものがある場合は、 Fisher の直接確率法(正確確率法, Fisher's exact probability test)</u>を使用する必要が ある。このため、9)のカイ2乗検定のすぐ下にある下図のような「Fisher の正確検定」の 表の「両側 Pr<=P」右横の値を確認する。

この値は「2つのグループ間で、比較したデータの割合に差が無い」ことが起こりう る確率の値である。従って、この値が<u>5%(0.05)未満</u>の場合は、帰無仮説を棄却できる (否定できる)ため、5%の有意水準で「<u>比較したデータ(の割合)は2つのグループの間</u> で有意な差がある」とみなすことができる。

この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できない)ため、2 つのグループの間で比較したデータの割合に差があるかどうかの判定は「<u>保留</u>」となり、 「<u>2つのグループの間で比較したデータ(の割合)に有意な差があるとはいえない</u>」ある いは「<u>2つのグループの間で比較したデータ(の割合)に有意な差は見られなかった</u>」と いう結論になる。決して「2つのグループ間で比較したデータ(の割合)に<u>差が無い」と</u> いう結論は出せない。これは、危険率が「帰無仮説が正しいのに、間違って棄却してし まう確率」であるためである。また現実的にも、サンプル数を増加させた場合や他の検 定方法を使用した場合、有意差が検出される可能性があるためである。

Fisherの正確検定				
セル (1,1) 度数 (F)	259			
左側 Pr <= F	1.0000			
右側 Pr >= F	1.167E-06			
表の確率 (P)	7.165E-07			
両側 Pr <= P	1.720E-06			

(注; 上記の表の危険率のように「1.720E-06」という数値は、「1.720×10⁻⁶」 を意味している。すなわち、「○○E-X」という数値は、「○○×10^{-x}」 のことである。

 2×2表以外の分析では、下記のような場合、χ²検定(カイ二乗検定)はできない可能性 がある。

(1) セルのうち 20%の期待度数が 5 未満

(2) セルの中に、1つでも期待度数が0のものがある

このような場合は、下記のような操作を行う。

- 1) データの再分類を行い、期待度数の小さなセルを無いようにして、再度、χ²検 定(カイニ乗検定)を行う。
- 2) データの再分類ができない場合は、χ²検定(カイ二乗検定)の結果の解釈や採用 するかどうかを慎重に検討(統計の専門家に相談するのも良い)し、場合によって は分析を中止する。

- (7) 対応のあるデータの差が正規分布しているかどうかを判定する
 - 1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。
 - 2) 上部メニュー内の [データ]をクリックし、その下に表示されるメニュー内の [フィルタ とクエリ]をクリックする。

/]] / / / / /	G 0						
🕵 SAS Enterprise	e Guide – Sample	-Load					
- ファイル(E) 編	集(E) 表示(⊻)	⊐-ド(<u>C</u>) [デー	<u> タ(D)</u>	記述統計(<u>S</u>)	グラフ(<u>G</u>)	分析(<u>A</u>)
1 11 - 63 - 15	8 % Pa 6	SX 10	×	読み取	マリ専用(<u>E</u>)]
SASUSERIMPV	V_0006(プロセスフ	□-) • _•		フィルタ	1とクエリ(<u>F</u>)		
プロジェクトエクスプロ	1-5	+ ×		行(<u>R</u>)			P
💽 Sample-Load				列(L)			L. L.
白 🎭 プロセスフロ]-			移動((<u>a</u>)	Ctr	1+G
Sampl	lexis (Sample\$) ⊨≂∧_∞_'		8790.	- 	l /n;etn/ D)		

3)新しいウィンドウが開くので、左側の変数のリストから対応のあるデータそれぞれをダ ブルクリックし、右側の枠内にその変数が表示されていることを確認する。

I	SASUSERIMF	יע_0008 שי	- クエリビルダ				×
	クエリ名(<u>Q</u>):	SASUSERI	IMPW_0008 のク	IU 出力名	GO: SASUSER	RQuery_for_SASUSER_IMPW_00	OE 変更(<u>C</u>)
•	⊞計算列(<u>M</u>)	🕞 パラメータ(<u>P</u>) 🔊 検証(V)	 ・ 🗟 ブレビュー(E) 	E) 🖺 オブション(N) 🔹		
į	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	★ 削除(<u>D</u>)	≞ 結合(」)	データの選択 フィル	,タデータ データの並べ替	ž]	
	□ <u>課</u> , IMPW_0 	008 号(番号) 笔前(授業前) 笔 後(授業後)		列名 (12) '授業前' (授業 (12) '授業後' (授業	入力 (前) IMPW_0008 (後) IMPW_0008 兆 (後) IMPW_0008 (約)	概要 受業前 受業後	

4) 左上にある「計算列」をクリックする。



5)「計算列」というダイアログが開くので、右上の「新規作成」ボタンをクリックし、その 下に新たに表示されたメニュー内の「式の作成」をクリックする。

鬱 計算列	×
	新規作成(№)▼
	列の再コード化(R)…
	<u>I\()</u> [F\$%[<u>B</u>]
	名前の変更(<u>R</u>)
	閉じる
✓ クエリに新しい計算列を追加する(A)	
	11

6)「拡張式エディタ」というダイアログが開くので、左中央の「使用できる変数」の下のリストから、対応のあるデータの一方をクリックした後、右端の「式へ追加」ボタンをクリックして、上の「式テキスト」の下のボックス内にその変数の名前が表示されることを確認する。次に、その下のボタンの中で「一」(マイナス)ボタンをクリックし、「式テキスト」に表示されることを確認する。最後に、対応のあるデータのもう一方をクリックした後、右端の「式へ追加」ボタンをクリックして、上の「式テキスト」の下のボックス内にその変数の名前が表示されることを確認する。式が間違っていないことを確認した後、最下部の「OK」ボタンをクリックする。

「式テキスト」には、「(対応のあるデータの一方)-(対応のあるデータのもう-方)」という式を作成する。

電拡張式エディタ		<u>×</u> .
式テキスト(工):		
IMPW_0008.'授業前'n - IMPW_0008.'授業後'n		X
- / ** AND OR NOT =	$\Leftrightarrow < <= \rightarrow \rightarrow= =^{\times}$	N X I I () .
	式のクリア(<u>C</u>)	式へ追加(D)
データ 閲数		
使用できる変数(<u>A</u>):	変数値(<u>R</u>):	
 □ 提 IMPW_0008 ○ 御号 ○ 授業前 ○ 授業後 	_値 フォーマットされた値 5 5 6 6 7 7 8 8 9 9	
		複数の値を取得(<u>G</u>)
	OK	++>tll 11/7

7)「拡張式エディタ」が閉じて、「計算式」ダイアログに戻るので、左のボックス内に表示 されている「Calculation1」をクリックした後、「名前の変更」ボタンをクリックする。

鬱 計算列	×
Calculation1	新規作成(<u>N</u>)▼
	編集(E)
	削除(<u>D</u>)
	名前の変更(<u>R</u>)
	閉じる
☑ クエリに新しい計算列を追加する(A)	,
	111

8)変換後の変数の名前を適当に入力した後、「閉じる」ボタンをクリックする。

鬱 計算列	×
Difference	新規作成(N)▼
	編集(<u>E</u>)
	肖耶余(<u>D</u>)
	名前の変更(B)
	閉じる
✓ クエリに新しい計算列を追加する(A)	

- 9)新しい変数名が、グループ分けに用いる変数の下に表示されていることを確認した後、 「実行」ボタンをクリックする。
- 10) 対応のある2つのデータとその差の表が表示されるので、この表を用いて、配布資料「SAS (SAS Enterprise Guide) 使用方法(1因子(1種類のデータ)の概要を見る)」の「(1) ヒ ストグラムおよび、累積度数分布の確率プロットの表示」の操作を行い、対応のある2つ のデータの差の値が正規分布しているかどうかを判定する。

もし、正規分布していた場合は、この資料の「(8)対応のある t 検定 (Paired t test) を行う」の操作を行い、2 つのグループ間で比較する変数に差があるかどうかを判定する。 もし、正規分布していない場合は、配布資料「統計学基礎(3) 2 つのグループの1因子 (1種類のデータ)の差を検定する」の(20)以降を参照する。

- (8) 対応のある t 検定 (Paired t test)を行う
 - 1) SAS Enterprise Guide を起動し、使用するデータの表が表示されている状態まで進める。
 - 2) 上部メニュー内の [分析]をクリックし、その下に表示されるメニュー内の [分散分析]
 - をクリックし、更にその右に表示されたメニュー内の「t 検定」をクリックする。

	SAS Enterprise Guide					
	ファイル(E) 編集(E) 表示(⊻) コード(<u>C</u>)	データ(<u>D</u>) 記述統計(<u>S</u>) グラフ(<u>G</u>)	分标	析(<u>A)</u> アドイン(J) OLAP(<u>O</u>)	<u>س</u> –	ル(① ウィンドウ(W) ヘルプ(出)
	19-6-61 8 × 6 8 × 8	🤊 🔊 🔍 🚬 🍡 🎭 プロジェクトデザイナ		分散分析(<u>A</u>)	M	t 検定(<u>1</u>)
				回帰分析(<u>R</u>)	藨	一元配置分散分析(0)
				多変量解析(<u>M</u>)	≙	ノンパラメトリックなー元配置分散分析(<u>N</u>)
		Beg プロジェクトデザイナ 📋 SASI		生存時間分析(S)	15	線形モデル(L)
		123 番号 123 授業 123 123 123 123 123 123 123 123 123 123			1	混合モデル(<u>M</u>)
	白 🎥 授業とテスト×Is (授業とテスト\$)			管理図(0)		
	□ 1 データのインボート	3 3	lín	パレート図(円)		
	「「「「「「「」」 取り返し ジンミット ひと」 一下	4 4				
	SASUSERIMPW 0008	5 5				
	SASUSERIMPW_0008	6 6	ę.	モデルのスコアリング(<u>L</u>)		
_						

2) t 検定のダイアログが開くので、「t 検定の種類を選択してください」の下の「対応のあ る検定」の前の丸をクリックしてマークを付ける。



3) 左端のリストで「タスクの役割」をクリックし、右側に新たに表示された「変数リスト」 から、対応のあるデータの一方をクリックした後、その右にある右向き矢印のボタンをク リックし、新たに表示されたメニューから「対応のある変数」をクリックする。更に、対 応のあるデータのもう一方をクリックした後、その右にある右向き矢印のボタンをクリッ クし、新たに表示されたメニューから「対応のある変数」をクリックして、2つの対応の あるデータが、「タスクの役割」の下の「対応のある変数」の下に表示されていることを 確認して、最後に最下部の「実行」ボタンをクリックする。

IMPW_0008 の t	D t 検定	×
t 検定の種類		
分析 分析 グラフ タイトル	変数リスト(A) タスクの役割(E): 名前 が応のある変数(選択の上限: 2) し、 〈変数が必要です〉 〈変数が必要です〉 〈変数が必要です〉 〈変数が必要です〉	
	 ・	teary tear of the second sec
	▲ ● 対応のある変数 ③ グループ分析	<u>+ +</u>
	度数カウント 通 重み	
		 ▼
 	ビュー(<u>O</u>) 実行(<u>B</u>) 上書き保存(<u>S</u>) キャンセ	
"対応のある変数"	如 役割には少なくとも 2 変数を割り当てる必要があります。	1.

4)分析結果が表示されるので、表示されている表の中から、下図のような「t検定」と書かれた表を見る。その表中から、「Pr>|t|」の下の数値を確認する。この値が、「対応のあるt検定の危険率」(2つのグループの間で比較した変数に差が無い」ことが起こりうる確率の値)なので、この値が<u>5%(0.05)未満</u>の場合は、帰無仮説を棄却できる(否定できる)ため、5%の有意水準で「<u>比較した変数は2つのグループの間で有意な差がある</u>」とみなすことができる。

これに対し、この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できな い)ため、2つのグループの間に差があるかどうかの判定は「<u>保留</u>」となり、「<u>2つのグ</u> <u>ループの間で比較した変数に有意な差があるとはいえない</u>」あるいは「<u>2つのグループの</u> <u>間で比較した変数に有意な差は見られなかった</u>」という結論になる。決して「2つのグルー プ間で比較した変数に<u>差が無い」という結論は出せない</u>。これは、危険率が「帰無仮説が 正しいのに、間違って棄却してしまう確率」であるためである。また現実的にも、サンプ ル数を増加させた場合や他の検定方法を使用した場合、有意差が検出される可能性がある ためである。

t検定						
差	自由度	t値	Pr > t			
授業前 - 授業後	9	-11.76	<.0001			

- (9) Wilcoxon 符号順位和検定(Wilcoxon signed-ranks test)を行う
 - この資料の「(7)対応のあるデータの差が正規分布しているかどうかを判定する」を参照して、「対応のある2つのデータの差」の変数を作成し、その値が表示されている表を 画面に出す。
 - 2) 上部メニュー内の [記述統計]をクリックし、その下に表示されるメニュー内の [分布] をクリックする。



3)新しいウィンドウが開くので、「対応のある2つのデータの差」を「変数リスト」の下のリストから選んでクリックする。そして、その右隣りにある右向き矢印のボタンをクリックし、新たに表示されたメニューから「分析変数」をクリックし、右の「タスクの役割」の下の「分析変数」の下に、選択した変数名が表示されていることを確認した後、最下部の「実行」ボタンをクリックする。

4)分析結果が表示されるので、表示されている表の中から、下図のような「位置の検定」 と書かれた表を見る。その表中から、「符号付順位検定」の行の「p値」の下の数値を確認 する。この値が、「Wilcoxon 符号順位和検定の危険率」(2つのグループの間で比較した変 数に差が無い」ことが起こりうる確率の値)なので、この値が<u>5%(0.05)未満</u>の場合は、帰 無仮説を棄却できる(否定できる)ため、5%の有意水準で「<u>比較した変数は2つのグルー</u> プの間で有意な差がある」とみなすことができる。

これに対し、この値が<u>5%(0.05)以上</u>の場合は、帰無仮説を棄却できない(否定できない)ため、2つのグループの間に差があるかどうかの判定は「<u>保留</u>」となり、「<u>2つのグ</u> ループの間で比較した変数に有意な差があるとはいえない」あるいは「<u>2つのグループの</u> 間で比較した変数に有意な差は見られなかった」という結論になる。決して「2つのグルー プ間で比較した変数に<u>差が無い」という結論は出せない</u>。これは、危険率が「帰無仮説が 正しいのに、間違って棄却してしまう確率」であるためである。また現実的にも、サンプ ル数を増加させた場合や他の検定方法を使用した場合、有意差が検出される可能性がある ためである。

位置の検定 H0: Mu0=0						
検定	統計量		p値			
Studentの t 検定	t	-11.7589	Pr > t	<.0001		
符号検定	М	-5	$\Pr \ge M $	0.0020		
符号付順位検定	s	-27.5	Pr >= S	0.0020		

Copyright (C) 2011 渡辺博且, All Rights Reserved.