ＥＺＲによる傾向スコア分析

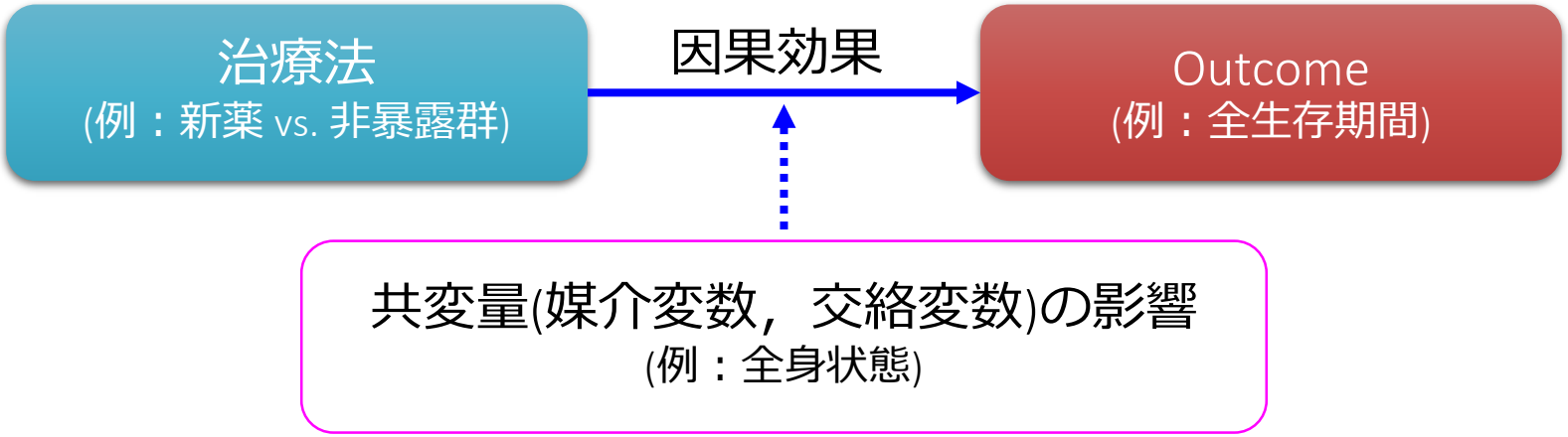
下川敏雄

和歌山県立医科大学 医学系研究科 医療データサイエンス講座

和歌山県立医科大学附属病院 臨床研究センター

傾向スコアの動機

□ 治療法の評価 = ある医学的介入による影響(因果効果)の検証



共変量が厄介な理由：コントロールできないものが多い

検証的な側面から共変量の影響を最小限にして因果効果を評価する

- **無作為化比較試験**(RCT:Randomized Controlled Trial)

観察的な側面(観察研究)において(治療選択に対する)共変量の影響を最小限にして因果効果を評価する

- **統計学的な調整法 (共変量とOutcomeの線形性などの制約)**
- **傾向スコア(Propensity score) の利用**

統計的方法による共変量の調整の方法とその限界

観察研究における共変量を考慮した研究及び解析(星野・岡田，2006)

研究デザイン

(1) 均衡化：共変量の値が同じになるペアを作ることで2つの群の被験者を構成
- 完全に一致するペアを作ることは事実上不可能
- 多数の共変量を含めることができない
- ペアは観測者が決めるため，主観的で恣意性を排除できない.

(2) 恒常化・限定：共変量のある被験者のみに限定して解析を行う
- 一部の被験者に限定するため，研究の一般可能性が低くなる.
- 多数の共変量を含めることができない

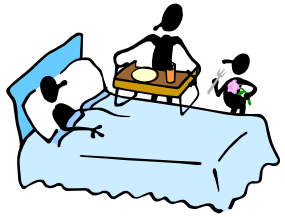
統計

(3) 統計的調整(多変量解析等)
- 誤った統計的モデルを用いると誤った結果を導く可能性がある.

複数の共変量を一つの変数(ものさし)におきかえ，そのもとで均衡化や層別化などを行い，誤った統計モデルを用いても「頑健」な結果をえることができる方法

傾向スコア分析(propensity score)

因果効果の定義：Neyman-Rubinの反事実的モデル



被験者i

既存薬を投与したときの効果 $y_{i|0}$

新薬を投与したときの効果 $y_{i|1}$

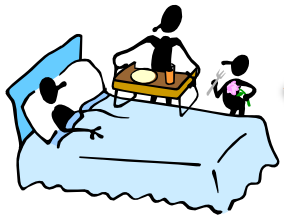
本来の治療効果

$\mathrm{E}[y_{i|1}] - \mathrm{E}[y_{i|0}]$

全ての患者にそれぞれの治療を実施した場合の平均的な効果の差

個体処理効果ITEという

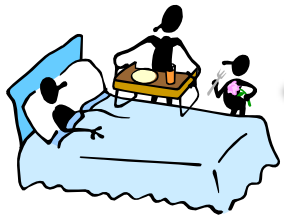
しかしながら，現実には…



被験者i
→ 新薬投与

$y_{i|0}$ 欠測

$y_{i|1}$



被験者i'
→ 既存薬投与

$y_{i|0}$

$y_{i|1}$ 欠測

となり，いずれか一方は欠測であると考えられる．

因果効果の定義：Neyman-Rubinの反事実的モデル

□ 無作為化比較試験の枠組みで考える

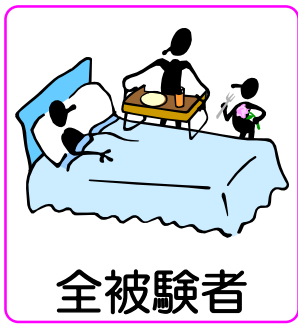
既存薬投与群$z=0$，新薬投与群$z=1$としたときの平均的な効果(zは割付変数と呼ばれる)

	新薬投与	既存薬投与
新薬投与群$z=1$	$E[y_{i\|1} \mid z=1]$	$E[y_{i\|0} \mid z=1]$ 欠測値
既存薬投与群$z=0$	$E[y_{i\|1} \mid z=0]$ 欠測値	$E[y_{i\|0} \mid z=0]$
	$E[y_{i\|1}]$	$E[y_{i\|0}]$

共変量

本来の因果効果
$E[y_{i|1}] - E[y_{i|0}]$

無作為化比較試験による群間の均一化



全被験者

2群間の共変量の影響を均一化

R

	新薬投与	既存薬投与
新薬投与群$z=1$	$E[y_{i\|1} \mid z=1]$ ⋮ $E[y_{i\|1}]$	$E[y_{i\|0} \mid z=0]$ ⋮ $E[y_{i\|0}]$
既存薬投与群$z=0$		

共変量

となることから，欠測と捉えられていた部分を無作為化によって補うことで，本来の因果効果を評価することができる.

傾向スコアで期待されること

傾向スコアの計算に用いた共変量は事後の調整因子等に用いない.

共変量 → PS

	暴露	非暴露群
暴露群 $z=1$	$E[y_{i\|1} \mid z=1]$ ⋮ $E[y_{i\|1}]$	$E[y_{i\|0} \mid z=0]$ ⋮ $E[y_{i\|0}]$
非暴露群 $z=0$		

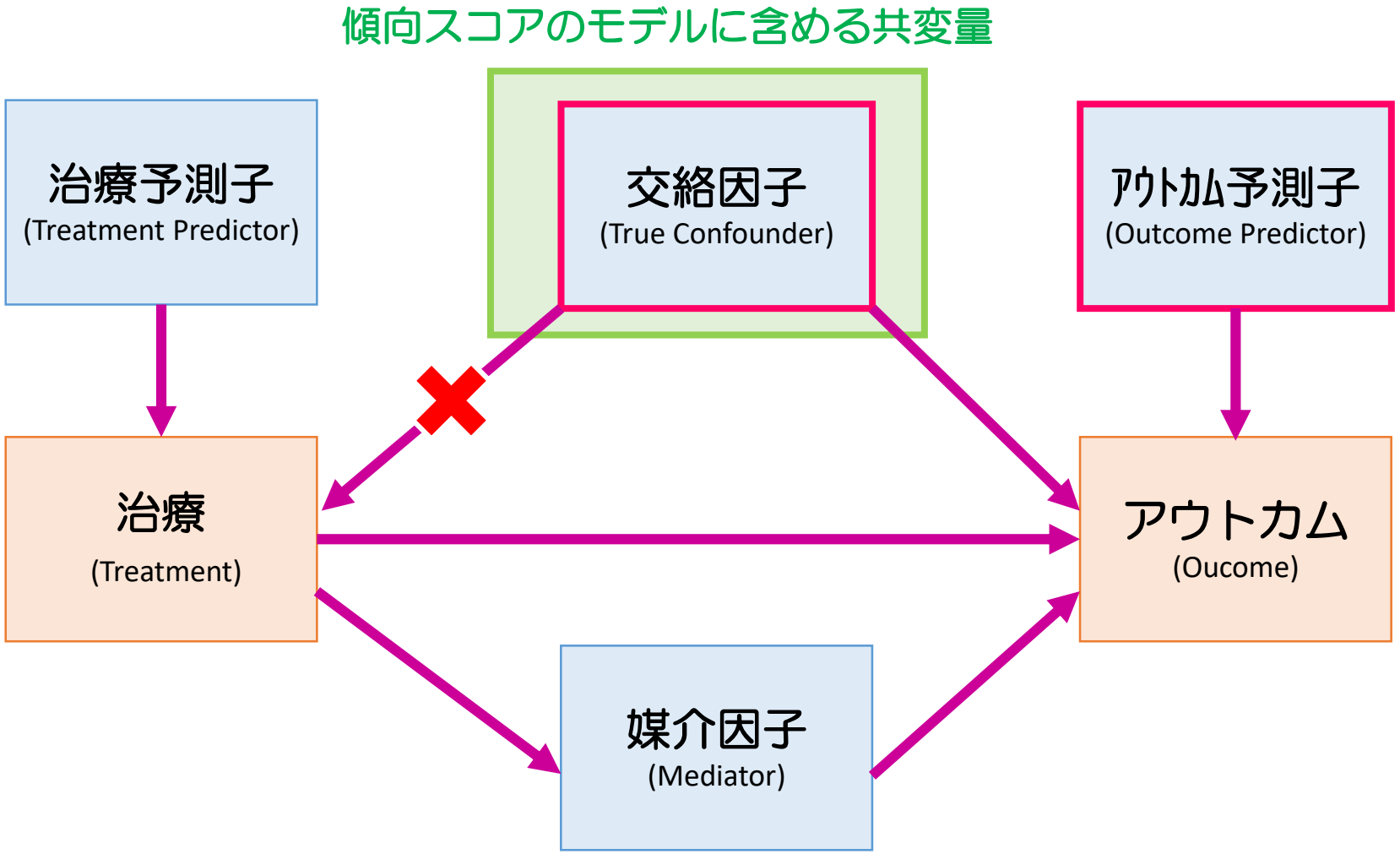
その他共変量

傾向スコアを用いることで複数の共変量から1つの変数に集約できる (傾向スコアモデルに関する適切性の評価をc統計量，疑似決定係数で評価すること(評価していない論文が多いことが指摘されている(Weizen et al., 2004))).

共変量による投与群の選択性を揃えることで，処理の割付については偶発性のみに依存するようにする．これにより，群間の被験者層を揃えることができる.

すなわち，傾向スコアとは，観察研究でのインバランスな群間のバランシングを行う「ものさし」となる.

因果関係に関する模式図(Leite, 2017)



それぞれの因子の意味(Brookhart et al., 2006)

交絡因子：処置効果の偏りとバラツキの削減に繋がる.

アウトカム予測子：処置効果のバラツキの削減に繋がる(偏りは関係ない).

観察研究における効果量の定義

	暴露	非暴露
暴露群$z=1$	$E[y_{i\|1} \mid z=1]$	$E[y_{i\|0} \mid z=1]$ 欠測値
非暴露群$z=0$	$E[y_{i\|1} \mid z=0]$ 欠測値	$E[y_{i\|0} \mid z=0]$
	$E[y_{i\|1}]$	$E[y_{i\|0}]$

ATT

ATE

- 平均処理効果(ATE:Average Treatment Effect)・平均因果効果(ACE:Average Causal Effect)

母集団のすべての個体での暴露でのアウトカムの期待値と非暴露でのアウトカムの期待値の差

$$ATE=E[y_{i|1}]-E[y_{i|0}]$$

- 暴露群の平均処理効果(ATT: Average Treatment effect for Treated)

母集団のうち暴露群での暴露のアウトカムの期待値と非暴露でのアウトカムの期待値の差

$$ATT=E[y_{i|1} \mid Z=1]-E[y_{i|0} \mid Z=1]$$

傾向スコア解析の手順(Leite, 2017)

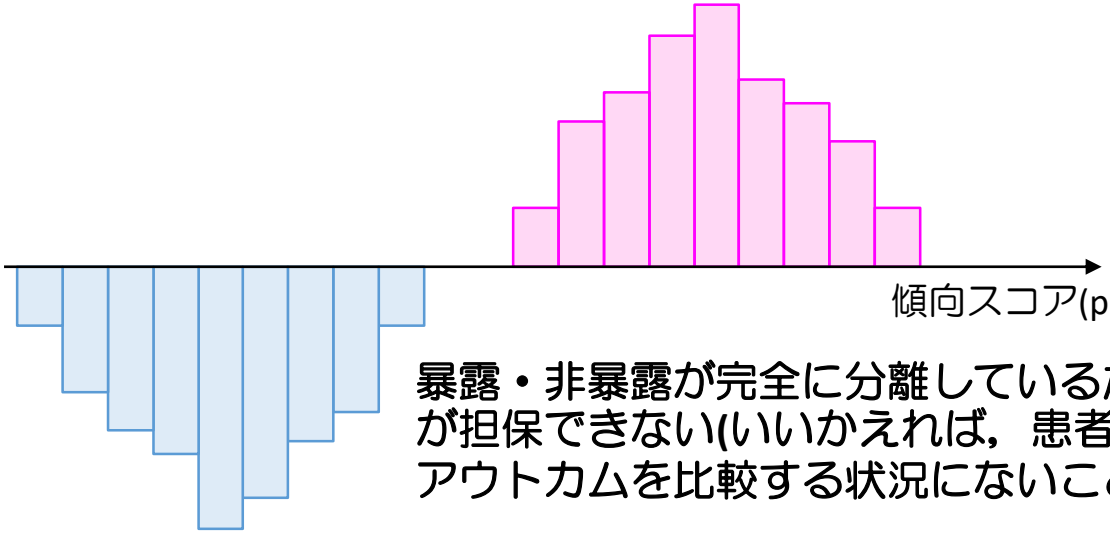
Step	目標	手順の例
1. 準備	解析可能な「完全な」データ集合を得る	• 共変量選択 • 欠測値に対する処理
2. 傾向スコアの推定	暴露群と非暴露群に対して傾向スコアを推定する.	• ロジスティック回帰 • Random Forests • 一般化ブースティング樹木
3. 傾向スコア手順の実施	暴露群と非暴露群の共変量の分布のバランスをとるための戦略を実施する.	• 傾向スコア・マッチング • 傾向スコア層化 • IPW(傾向スコア重み付け)
4. 共変量のバランスの評価	暴露群と非暴露群のあいだの共分散の分布のバランスが達成された程度を評価する.	• 標準化された平均距離の計算 • 分散比の計算
5. 処置効果推定	処置効果とその標準誤差を推定する.	• 重み付け平均距離 • 一般化線形モデル
6. 感度分析	除外された共変量が処置効果の有意性検定をどの程度変化させるかについて評価する.	• Rosenbaum(2002)の方法 • Carnegie et al.(2016)の方法

傾向スコアによる解析の留意点

傾向スコアとは，研究対象因子に暴露する確率で表される(つまり，傾向スコアは0～1の範囲をとる).

暴露群

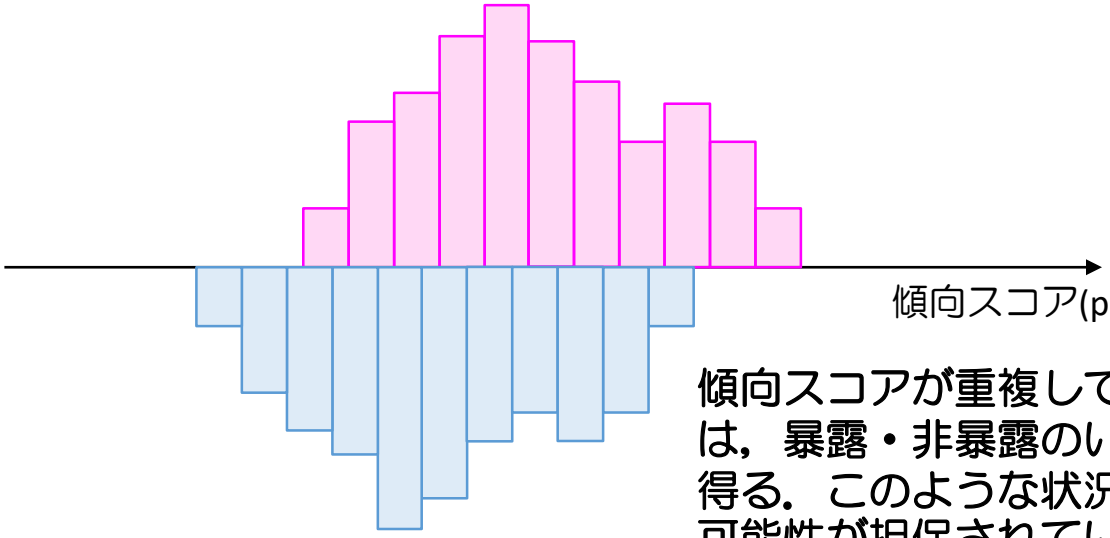
非暴露群



暴露・非暴露が完全に分離しているため，比較可能性が担保できない(いいかえれば，患者の違い明確であり，アウトカムを比較する状況にないことを示している)

暴露群

非暴露群



傾向スコアが重複している部分については，暴露・非暴露のいずれも候補になり得る．このような状況においては，比較可能性が担保されている.

傾向スコア手順の実施

- **マッチング**

  傾向スコアの一致した(あるいは極めて近い)個体同士を選択する.

- **層別**

  傾向スコアを用いて層別(5層程度)したうえで層別解析を行う.

- **逆確率重み付け(IPW(E): Inverse Probability Weighting (Estimator))**

  傾向スコアを用いてアウトカムに重み付けを行う方法

- **共分散分析**

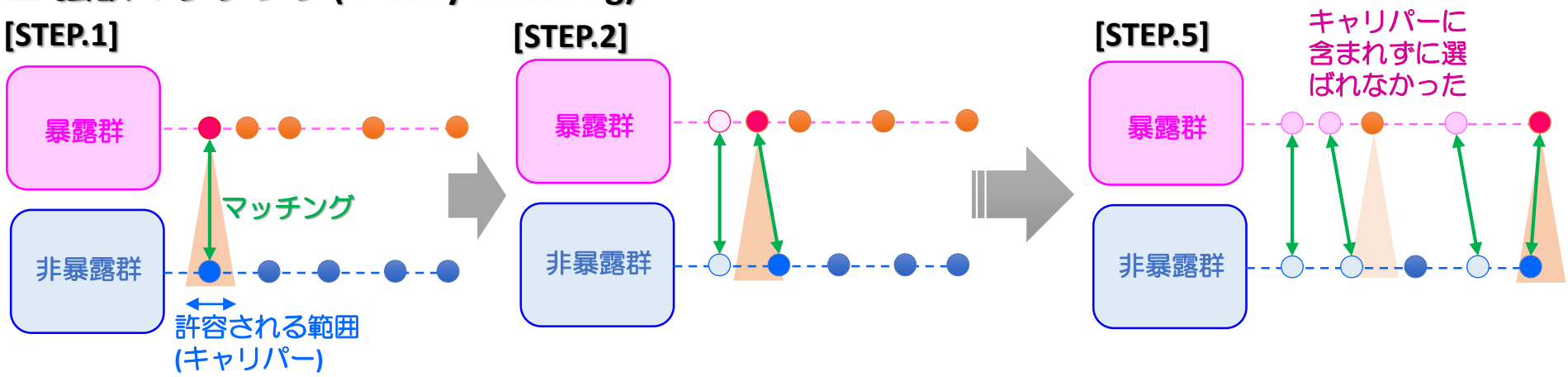
  傾向スコアを共変量にとった共分散分析を実施する.

これらの手法はいずれかを使うということではなく，組み合わせて用いることもできる．例えば，マッチング後に共分散分析を行うことが考えられる.

実際の応用分野では，マッチングあるいは逆確率重み付けを用いることが多いようである.

マッチング

■ 強欲マッチング(Greedy matching)



暴露群の任意の個体で，傾向スコアが最も近い非暴露群の個体を逐次に探索する方法である．強欲マッチングには様々なバリエーションがあり，以下を組み合わせて実行する．

- 復元マッチング vs. 非復元マッチング
- 1:1マッチング vs. 固定比マッチング(1:kマッチング) vs. 変動比マッチング
- 最近傍法 vs. キャリパーを伴う最近傍法

例題：テキスト5.2節

新薬(A)を投与した71例と既存薬B(C)を投与した101例の背景因子(性別，年齢，喫煙の有無，BMI，重症度スコア)と治療効果を調査した後ろ向き研究のデータである．

Sex：性別(M：男性，F：女性),
Smoke：喫煙歴(1：有，0：無),
Score：重症度スコア,
Outcome：アウトカム(1：改善，0：非改善)

Age：年齢,
BMI：Body Mass Index,
group：薬剤(A:新薬，C：既存薬)

[STEP.1] 傾向スコアを計算する．

応答変数： group(薬剤)
説明変数： Sex(性別)， Age(年齢)， Smoke(喫煙歴)， BMI， Score(重症度スコア)

＊連続変数の場合には，2値化をしてはいけない．

[STEP.2] 傾向スコアを用いてマッチングする．

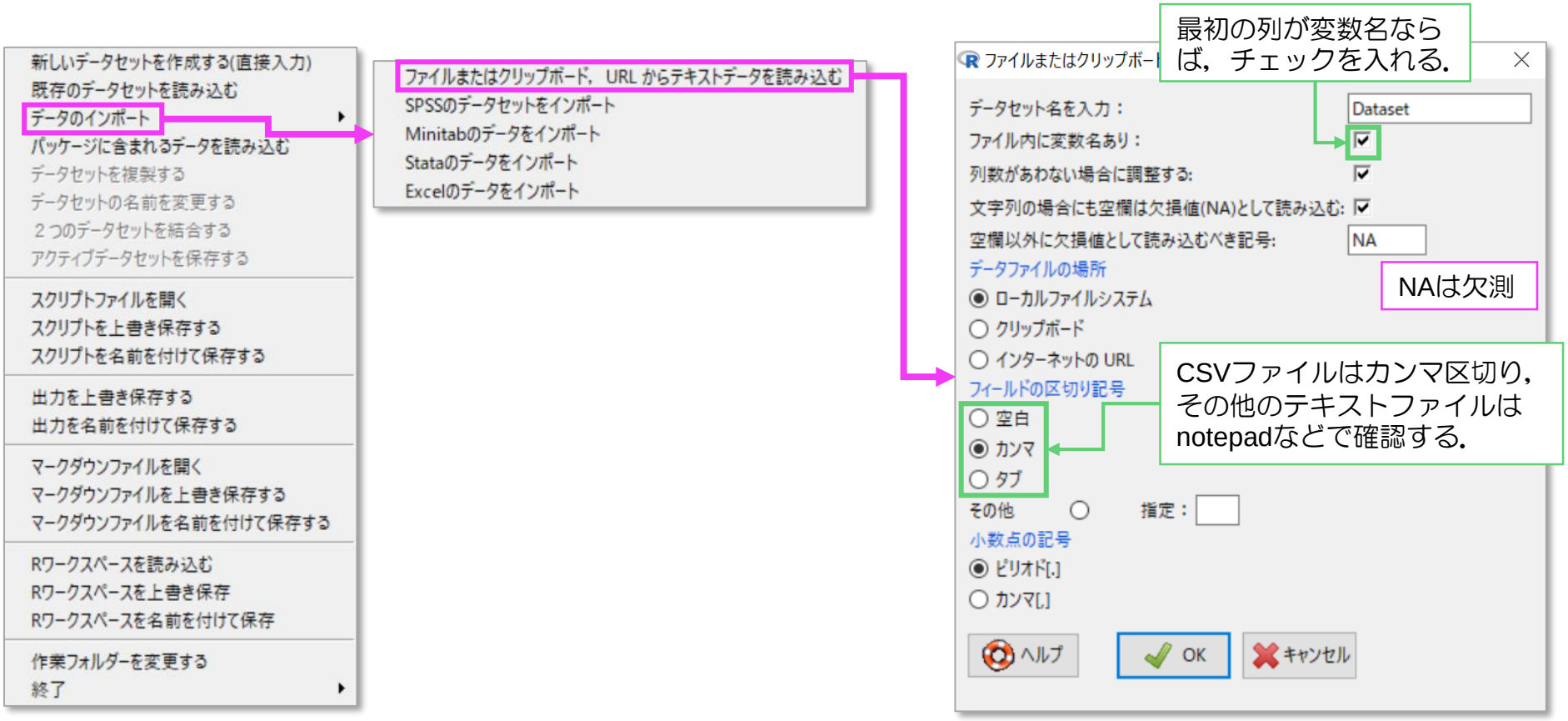
[STEP.3] マッチングしたデータを用いて解析を行う．

CSVファイルの読み込み

PSexample.csv

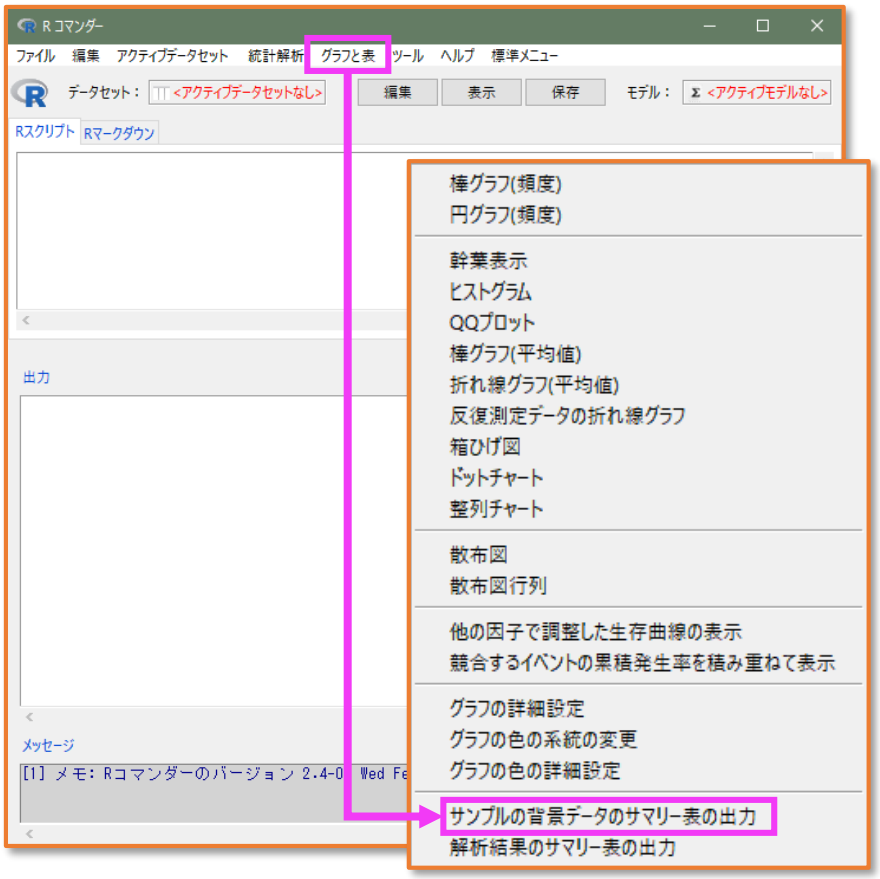
最初の列は変数名

	A	B	C	D	E	F	G
1	Sex	Age	Smoke	BMI	Score	group	Outcome
2	M	50	1	29.5	10	A	1
3	F	67	0	24	8	C	0
4	F	50	0	20.2	8	C	0



あとは，「PSexample.csv」を読み込めばよい．

データの概要



グループ変数： group(薬剤)
連続変数：年齢(Age)， BMI， Score(重症度スコア)
カテゴリカル変数： Sex(性別)， Smoke(喫煙歴)

結果

出力

実行

```
> FinalTable <- FinalTable[which(rownames(FinalTable)!="n"),]
> FinalTable <- rbind(n=row1, FinalTable)
> FinalTable <- rbind(row0, FinalTable)
> row0 <- rep("", length(colnames(FinalTable)))
> row0[3] <- "group"
> FinalTable <- rbind(row0, FinalTable)
> finaltable.dataframe.print(FinalTable)
```

Factor	Group	group A	group C	p.value
n		71	110	
Sex (%)	F	19 (26.8)	41 (37.3)	0.150
	M	52 (73.2)	69 (62.7)	
Smoke (%)	0	37 (52.1)	80 (72.7)	0.007
	1	34 (47.9)	30 (27.3)	
Age		54.11 (8.62)	53.45 (8.32)	0.604
BMI		25.03 (3.77)	24.10 (2.92)	0.063
Score		8.85 (1.68)	7.30 (1.94)	<0.001

```
> write.table(FinalTable, "clipboard", sep=" ", row.names=FALSE, col.names=FALSE)
```

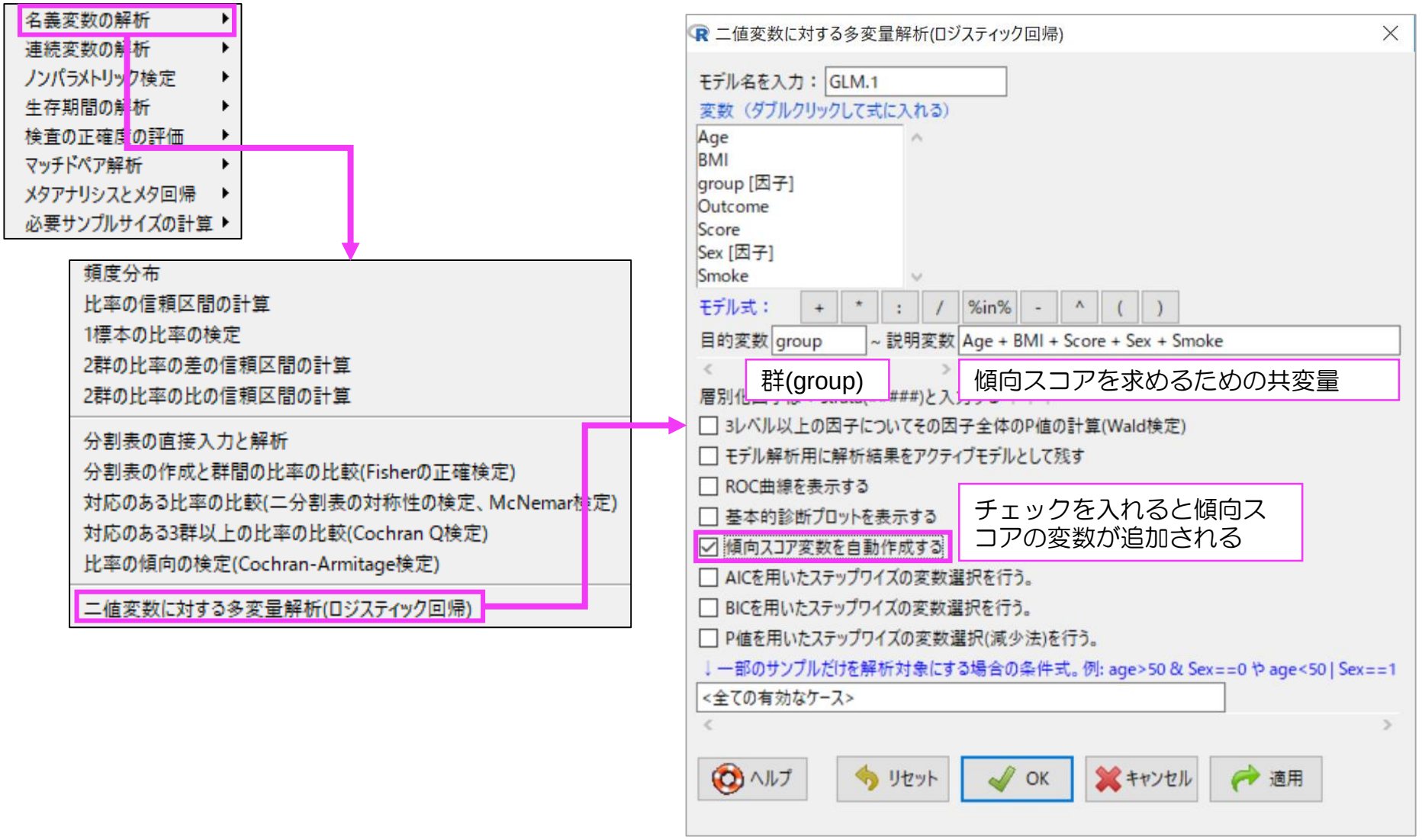
クリップボードのデータからExcelで作成

		group A	group C	p.value
		71	110	
Sex (%)	F	19 (26.8)	41 (37.3)	0.150
	M	52 (73.2)	69 (62.7)	
Smoke (%)	0	37 (52.1)	80 (72.7)	0.007
	1	34 (47.9)	30 (27.3)	
Age		54.11 (8.62)	53.45 (8.32)	0.604
BMI		25.03 (3.77)	24.10 (2.92)	0.063
Score		8.85 (1.68)	7.30 (1.94)	<0.001

喫煙歴，重症度スコアが有意である．したがって，この研究の被験者層において，これらの共変量に対して違いが認められる．

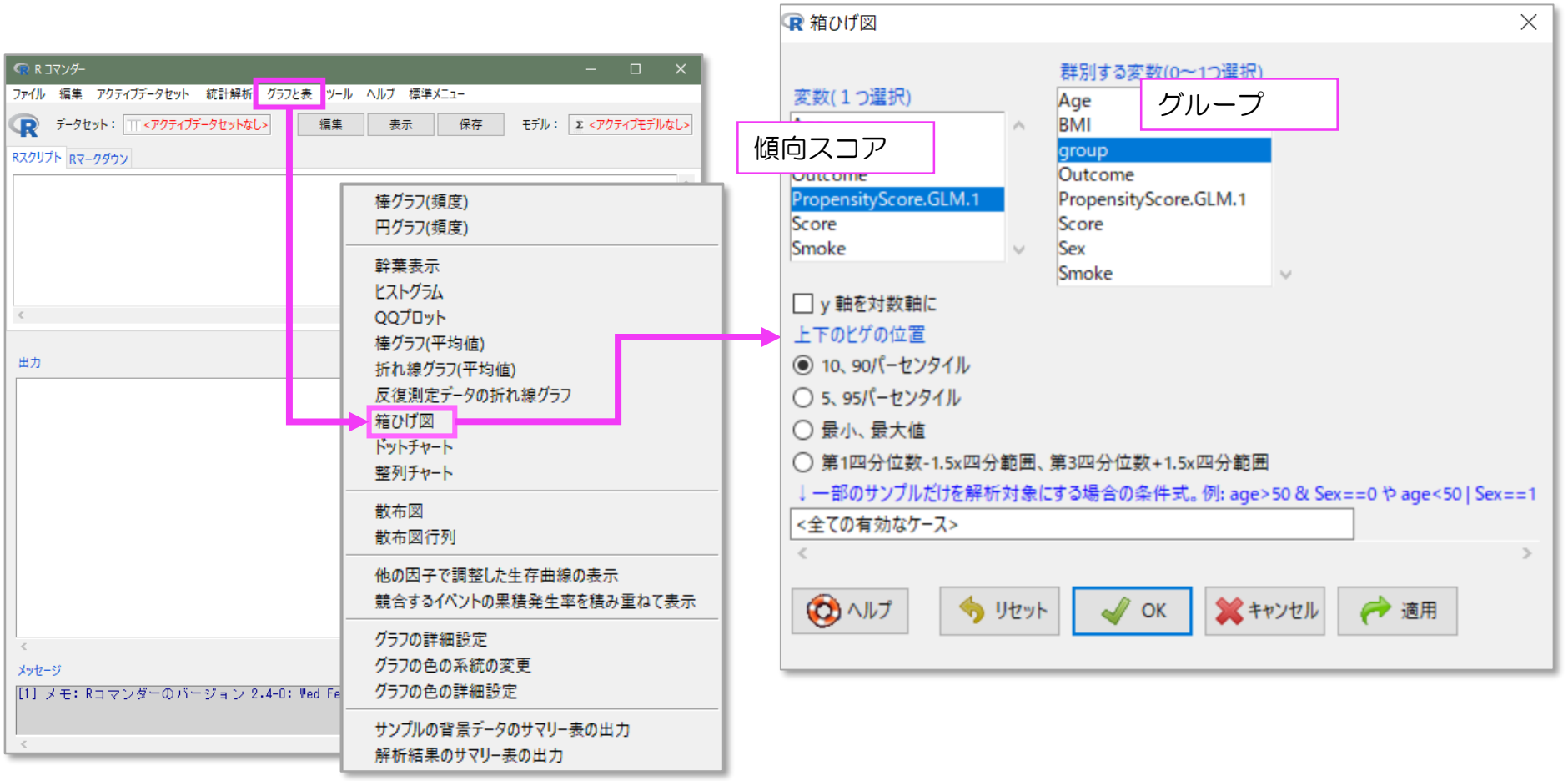
[STEP.1] 傾向スコアの計算

傾向スコアは，通常のロジスティック回帰によって求めることができる.

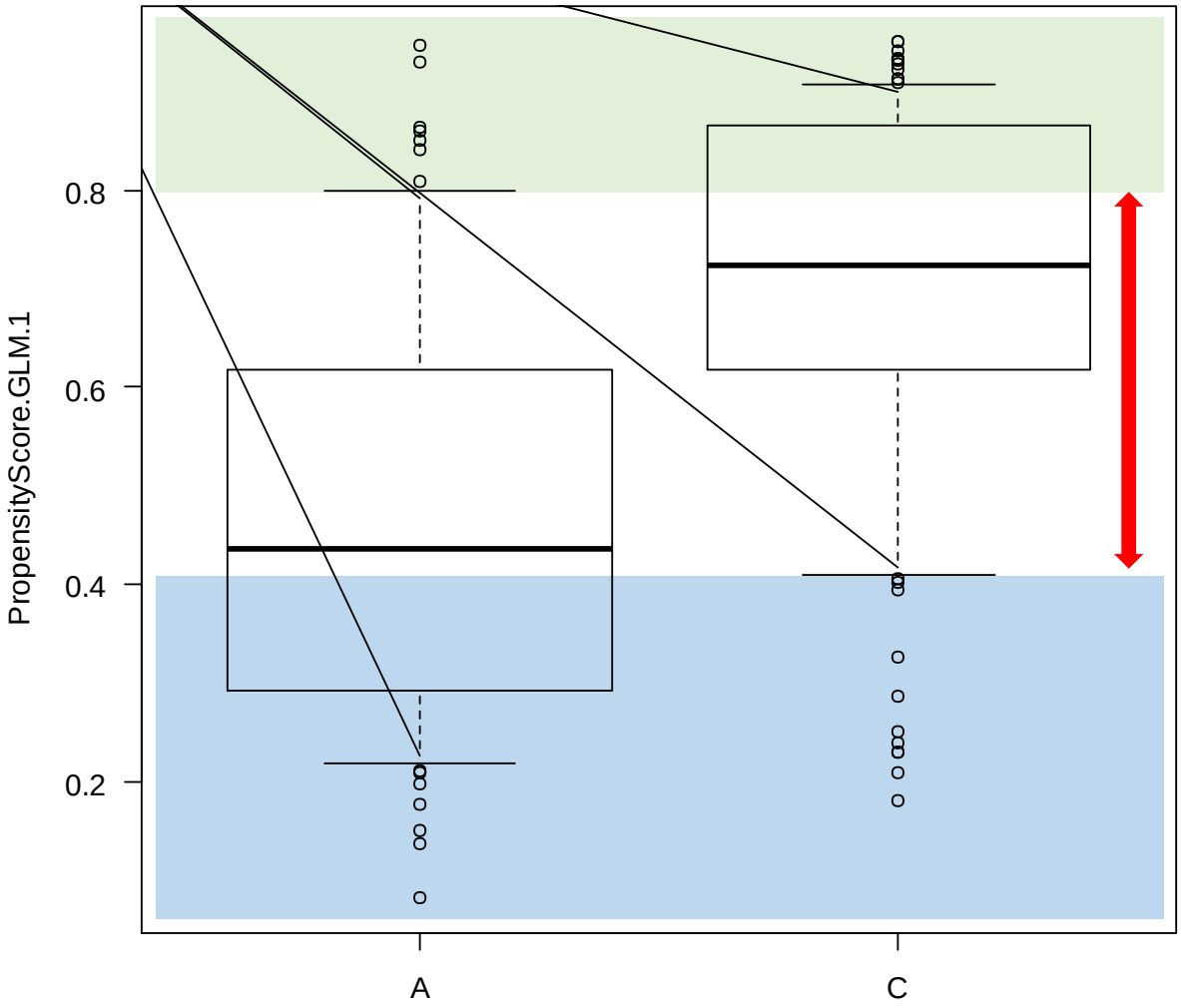


新たな変数(傾向スコア)が追加される.

(余談) 傾向スコアをグラフ表示する



箱ひげ図



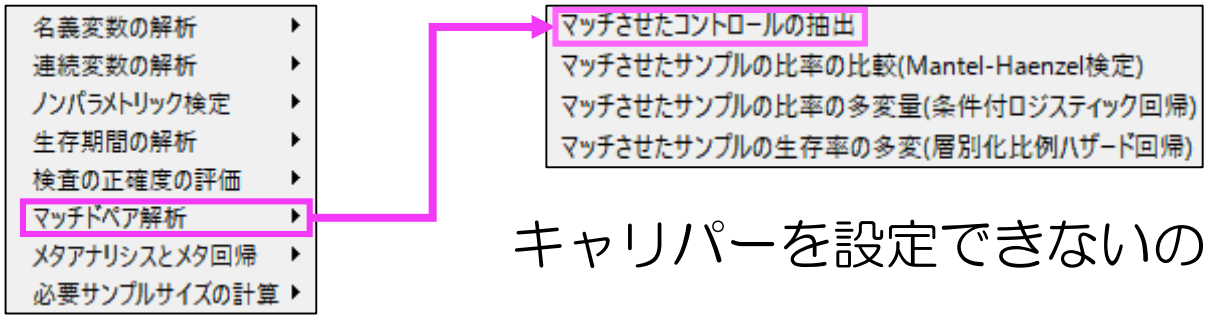
傾向スコアが小さい被験者は対照薬(C) を投与されているが新薬(A)を投与されている被験者は少ない.

新薬(A)と対照薬(C)のいずれも投与される可能性がある.

傾向スコアが小さい被験者は新薬(A)を投与されているが対照薬(C)を投与されている被験者は少ない.

マッチングの方法

卑近なマッチング



キャリパーを設定できないので推奨されない.

Rを応用する(追加パッケージmatchingを利用する)

matchingパッケージではグループ変数が0(コントロール群), 1(アクティブ群)で定義しなければならない(グループ変数groupはAとCになっている).

「アクティブデータセット」→「変数の操作」→「ダミー変数を作成する」

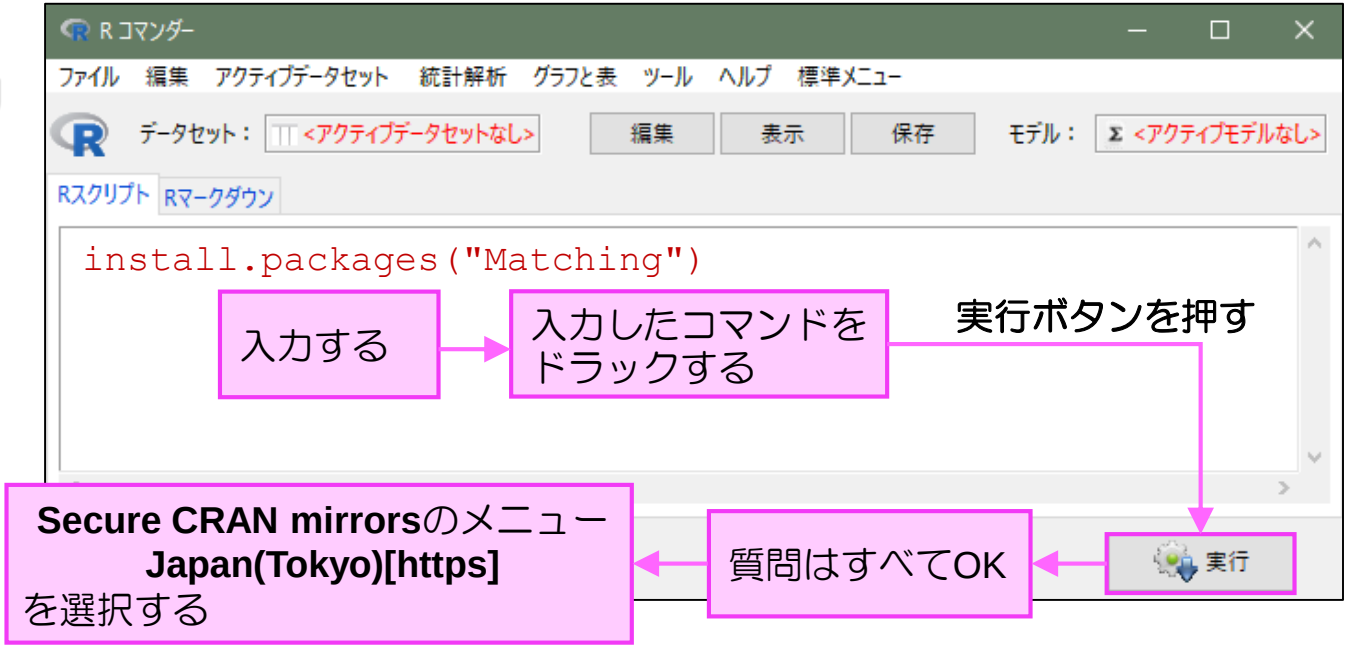
新たに生成される変数

groupDummyGroupA
　処理群(A)を1, 対照群(C)を0とした場合
groupDummy.GroupC
　処理群(A)を0, 対照群(C)を1とした場合

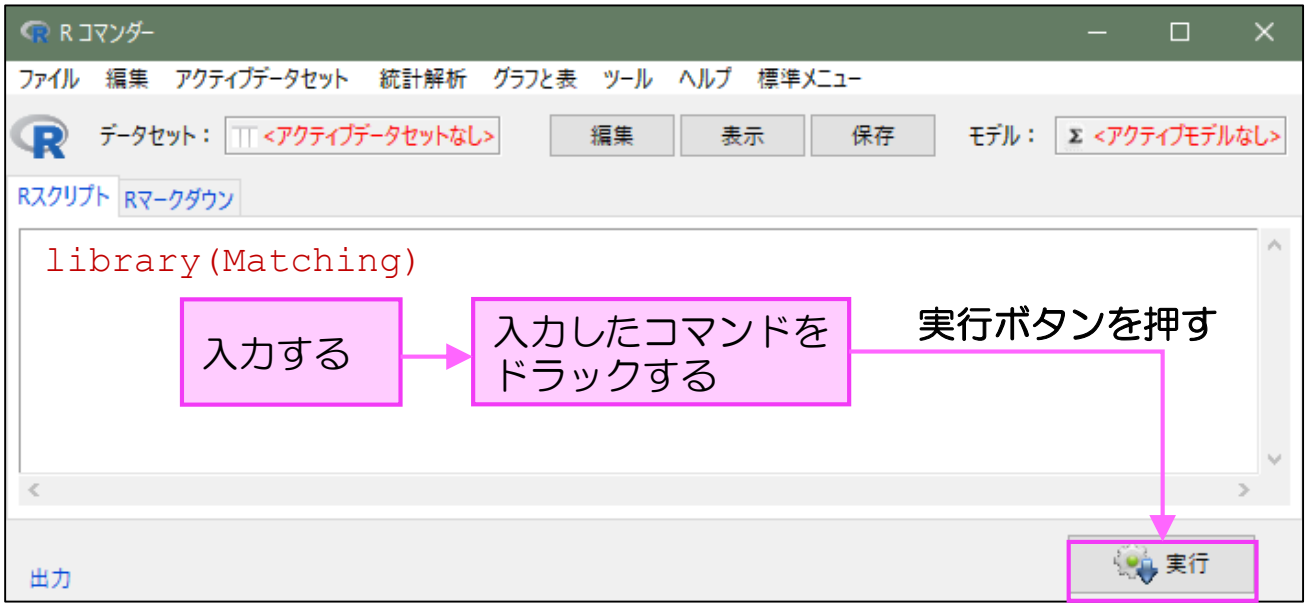
Rを用いた解析

[STEP.1]
パッケージmatching
をインストールする

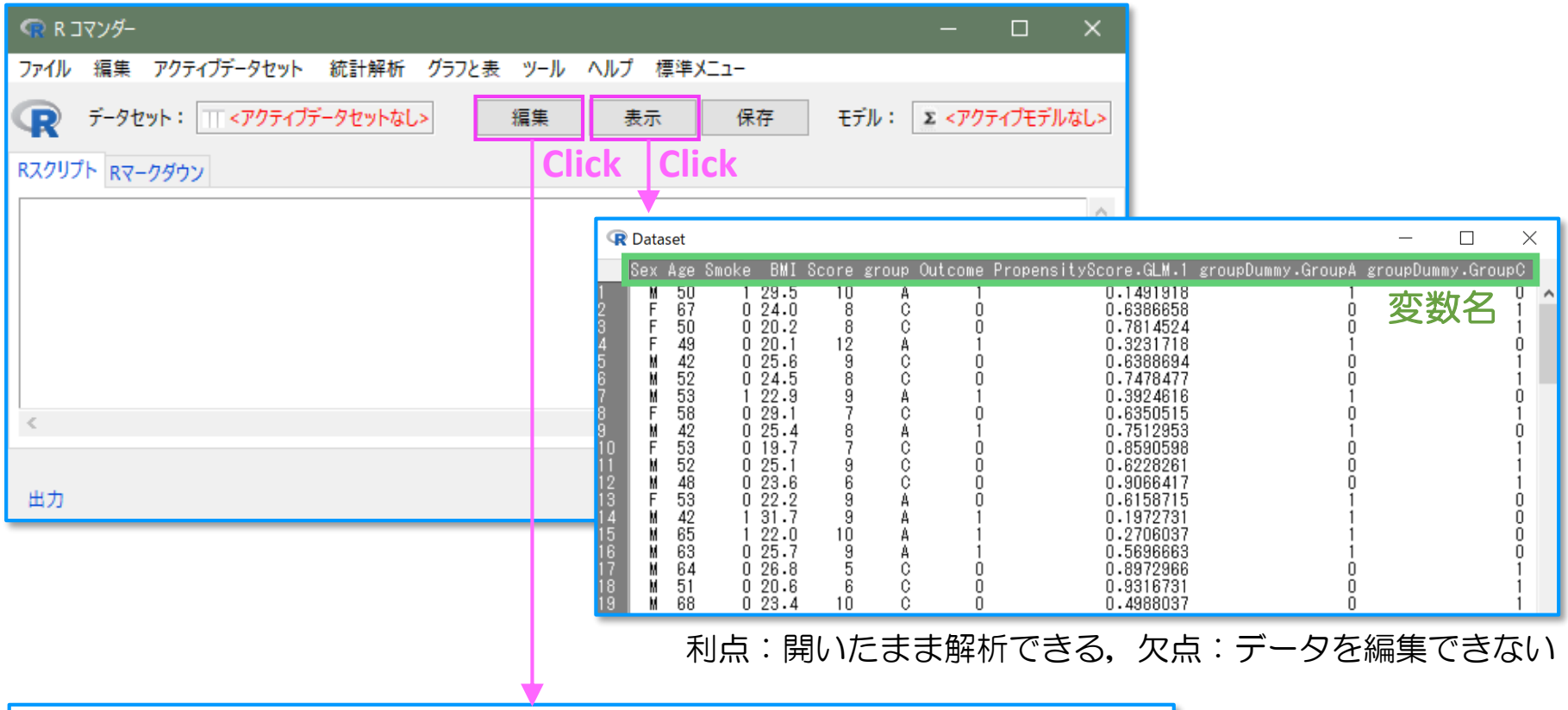
これは1回やればよい
(毎回実施しなくてもいい)



[STEP.2]
パッケージmatching
を読み込む

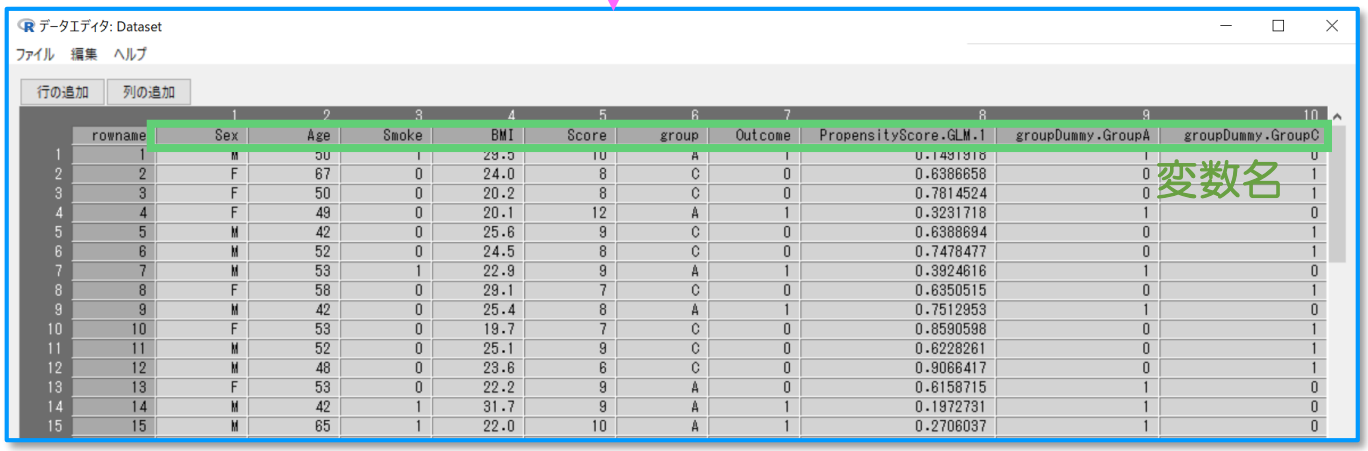


(余談) 変数名を確認するには？



	Sex	Age	Smoke	BMI	Score	group	Outcome	PropensityScore.GLM.1	groupDummy.GroupA	groupDummy.GroupC
1	M	50	1	29.5	10	A	1	0.1491918	1	0
2	F	67	0	24.0	8	C	0	0.6386658	0	1
3	F	50	0	20.2	8	C	0	0.7814524	0	1
4	F	49	0	20.1	12	A	1	0.3231718	1	0
5	M	42	0	25.6	9	C	0	0.6388694	0	1
6	M	52	0	24.5	8	C	0	0.7478477	0	1
7	M	53	1	22.9	9	A	1	0.3924616	1	0
8	F	58	0	29.1	7	C	0	0.6350515	0	1
9	M	42	0	25.4	8	A	1	0.7512953	1	0
10	F	53	0	19.7	7	C	0	0.8590598	0	1
11	M	52	0	25.1	9	C	0	0.6228261	0	1
12	M	48	0	23.6	6	C	0	0.9066417	0	1
13	F	53	0	22.2	9	A	0	0.6158715	1	0
14	M	42	1	31.7	9	A	1	0.1972731	1	0
15	M	65	1	22.0	10	A	1	0.2706037	1	0
16	M	63	0	25.7	9	A	1	0.5696663	1	0
17	M	64	0	26.8	5	C	0	0.8972966	0	1
18	M	51	0	20.6	6	C	0	0.9316731	0	1
19	M	68	0	23.4	10	C	0	0.4988037	0	1

利点：開いたまま解析できる，欠点：データを編集できない



	rowname	Sex	Age	Smoke	BMI	Score	group	Outcome	PropensityScore.GLM.1	groupDummy.GroupA	groupDummy.GroupC
1	1	M	50	1	29.5	10	A	1	0.1491918	1	0
2	2	F	67	0	24.0	8	C	0	0.6386658	0	1
3	3	F	50	0	20.2	8	C	0	0.7814524	0	1
4	4	F	49	0	20.1	12	A	1	0.3231718	1	0
5	5	M	42	0	25.6	9	C	0	0.6388694	0	1
6	6	M	52	0	24.5	8	C	0	0.7478477	0	1
7	7	M	53	1	22.9	9	A	1	0.3924616	1	0
8	8	F	58	0	29.1	7	C	0	0.6350515	0	1
9	9	M	42	0	25.4	8	A	1	0.7512953	1	0
10	10	F	53	0	19.7	7	C	0	0.8590598	0	1
11	11	M	52	0	25.1	9	C	0	0.6228261	0	1
12	12	M	48	0	23.6	6	C	0	0.9066417	0	1
13	13	F	53	0	22.2	9	A	0	0.6158715	1	0
14	14	M	42	1	31.7	9	A	1	0.1972731	1	0
15	15	M	65	1	22.0	10	A	1	0.2706037	1	0

利点：データを編集できる，欠点：開いたまま解析できない

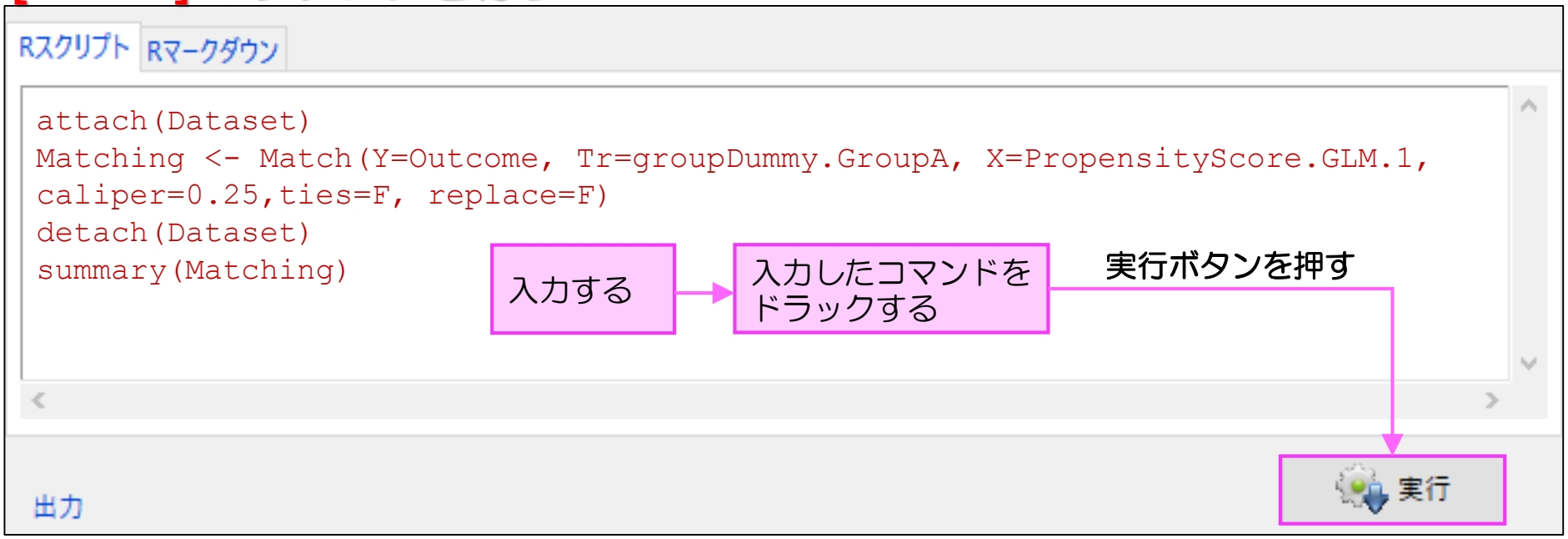
Rによる傾向スコアの計算（続き）

パッケージmatchingにおけるマッチングの関数

```
Match(Y=応答, Tr=群, X=傾向スコア, caliper=キャリパー, ties=F, replace=F)
```

- 応答 → **Outcome**
- 群 → **groupDummyGroupA**
  - 試験薬(A)が1，コントロール群(C)が0のダミー変数
- 傾向スコア → **PropensityScore.GLM.1**
  - 傾向スコアの値
- キャリパー
  - 0.20～0.25の間で自由に設定

[STEP.3] マッチングを行う



マッチング結果

出力

実行

```
> attach(Dataset)
> Matching <- Match(Y=Outcome, Tr=groupDummy.GroupA, X=PropensityScore.GLM.1, caliper=0.25,ties=F, replace=F)
> Detach(Dataset)
> attach(Dataset)
> Matching <- Match(Y=Outcome, Tr=groupDummy.GroupA, X=PropensityScore.GLM.1, caliper=0.25,ties=F, replace=F)
> detach(Dataset)
> summary(Matching)

Estimate...  0.2381
SE.........  0.11071
T-stat.....  2.1507
p.val......  0.031499

Original number of observations..............  181
Original number of treated obs...............  71
Matched number of observations...............  42
Matched number of observations  (unweighted).  42

Caliper (SDs)........................................  0.25
Number of obs dropped by 'exact' or 'caliper'  29
```

被験者総数

試験群の例数

マッチング後の方群当たりの例数

Rによる傾向スコアの計算（続き）

[STEP.4] マッチング後のデータを抽出する

Rスクリプト　Rマークダウン

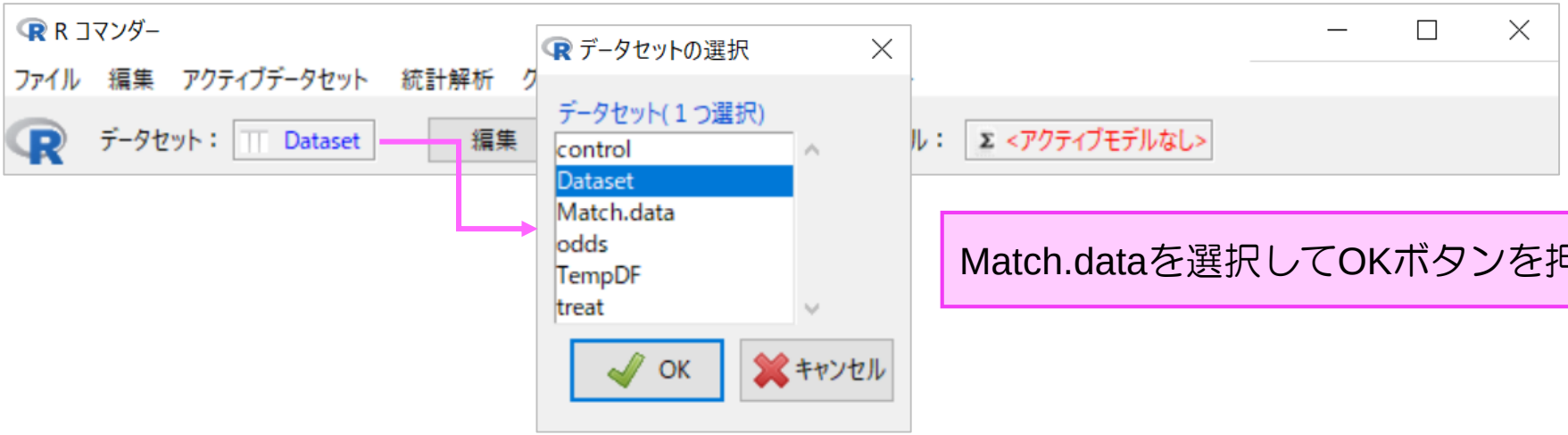
```
treat <- Dataset[Matching$index.treated,]
control <- Dataset[Matching$index.control,]
Match.data <- rbind(treat,control)
```

入力する → 入力したコマンドをドラックする → 実行ボタンを押す

出力

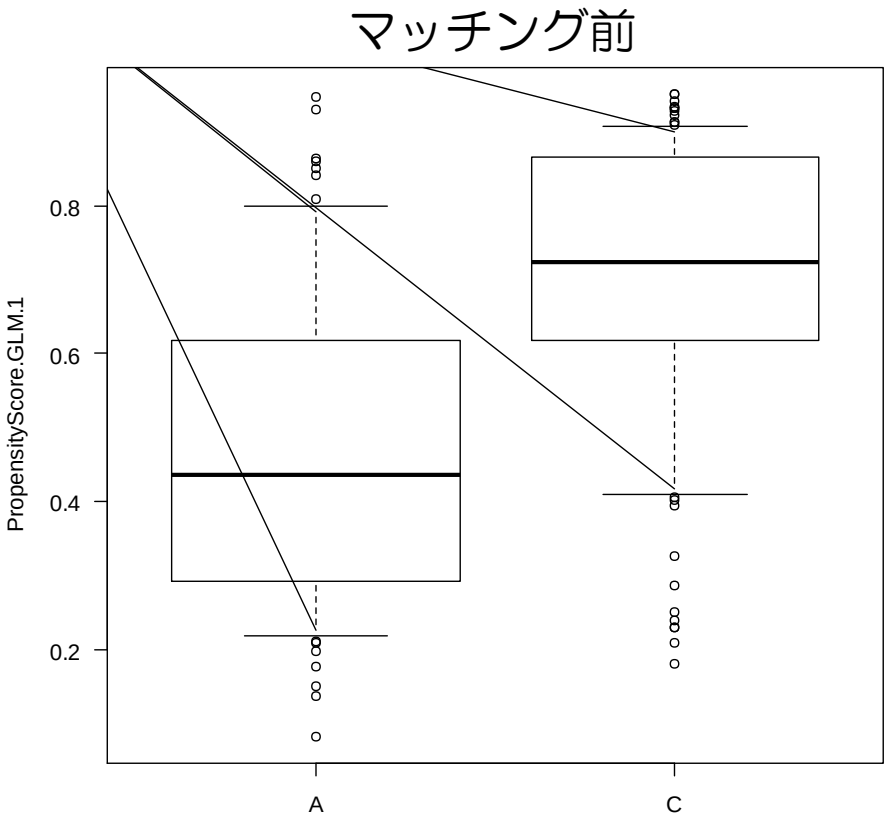
実行

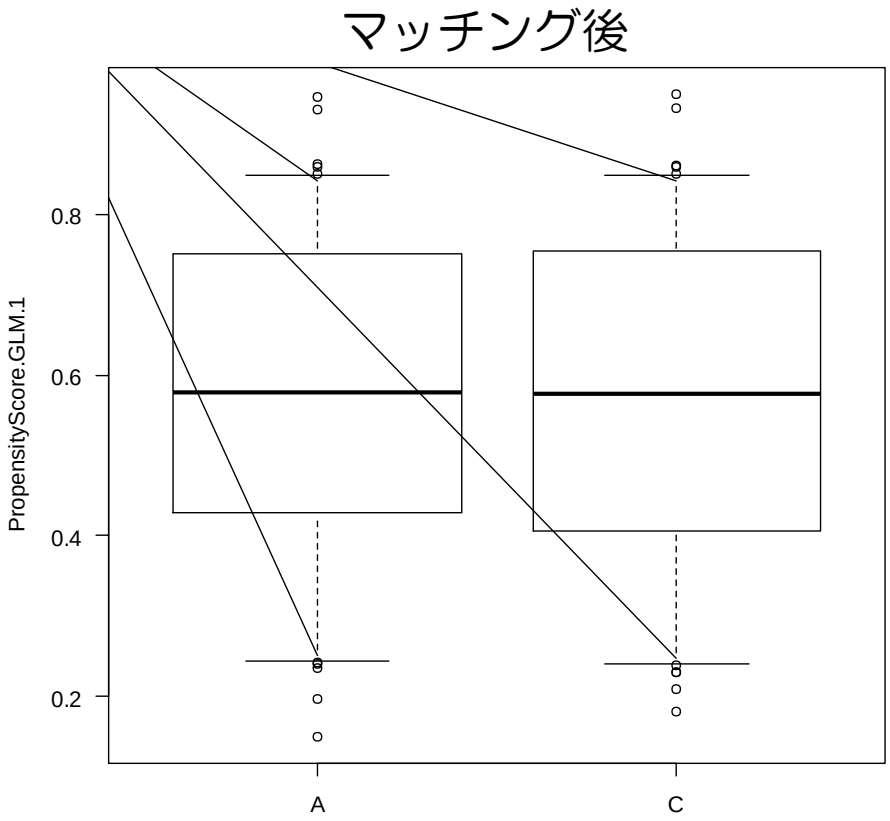
マッチングされたデータが新たなデータ集合Match.dataに保存される.



Match.dataを選択してOKボタンを押す

マッチング後の傾向スコアの分布を箱ひげ図で確認する





明らかにマッチング後では分布の違いが認められなくなった.

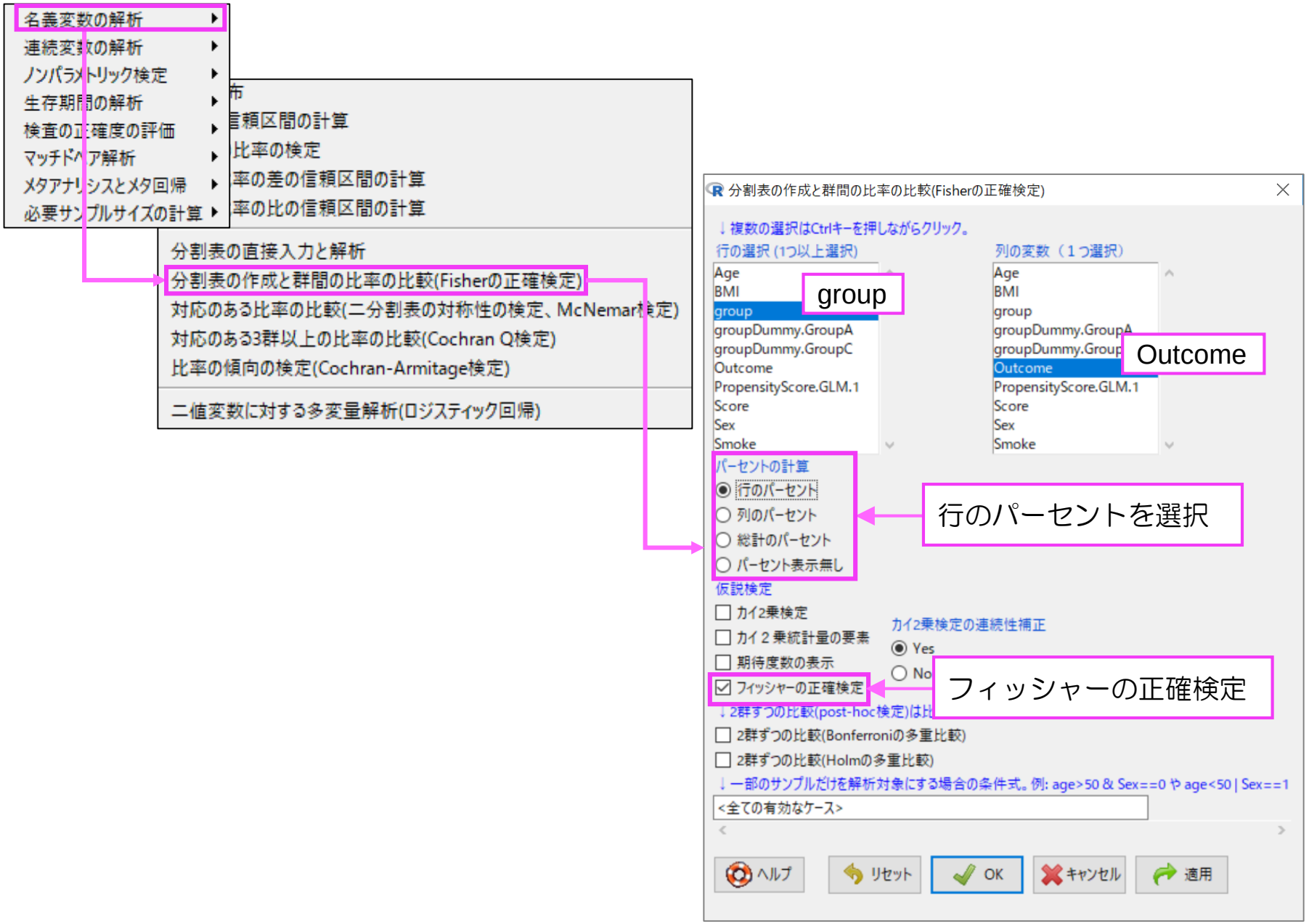
共変量毎に確認する.

		group		p.value
		A	C	
		71	110	
Sex (%)	F	19 (26.8)	41 (37.3)	0.150
	M	52 (73.2)	69 (62.7)	
Smoke (%)	0	37 (52.1)	80 (72.7)	0.007
	1	34 (47.9)	30 (27.3)	
Age		54.11 (8.62)	53.45 (8.32)	0.604
BMI		25.03 (3.77)	24.10 (2.92)	0.063
Score		8.85 (1.68)	7.30 (1.94)	<0.001

		group		p.value
		A	C	
		71	110	
Sex (%)	F	14 (33.3)	16 (38.1)	0.820
	M	28 (66.7)	26 (61.9)	
Smoke (%)	0	27 (64.3)	26 (61.9)	1.000
	1	15 (35.7)	16 (38.1)	
Age		54.83 (9.38)	53.52 (7.84)	0.490
BMI		24.61 (2.96)	24.14 (3.16)	0.482
Score		8.26 (1.65)	8.31 (1.94)	0.904

分布間の差が認められなくなった

主要評価項目の評価


名義変数の解析
連続変数の解析
ノンパラメトリック検定
生存期間の解析
検査の正確度の評価
マッチドペア解析
メタアナリシスとメタ回帰
必要サンプルサイズの計算
分割表の直接入力と解析
分割表の作成と群間の比率の比較(Fisherの正確検定)
対応のある比率の比較(二分割表の対称性の検定、McNemar検定)
対応のある3群以上の比率の比較(Cochran Q検定)
比率の傾向の検定(Cochran-Armitage検定)
二値変数に対する多変量解析(ロジスティック回帰)
分割表の作成と群間の比率の比較(Fisherの正確検定)
↓複数の選択はCtrlキーを押しながらクリック。
行の選択 (1つ以上選択)
列の変数 (1つ選択)
Age
BMI
group
groupDummy.GroupA
groupDummy.GroupC
Outcome
PropensityScore.GLM.1
Score
Sex
Smoke
group
Outcome
パーセントの計算
行のパーセント
列のパーセント
総計のパーセント
パーセント表示無し
行のパーセントを選択
仮説検定
カイ2乗検定
カイ２乗統計量の要素
期待度数の表示
フィッシャーの正確検定
カイ2乗検定の連続性補正
Yes
フィッシャーの正確検定
2群ずつの比較(Bonferroniの多重比較)
2群ずつの比較(Holmの多重比較)
↓一部のサンプルだけを解析対象にする場合の条件式。例: age>50 & Sex==0 や age<50 | Sex==1
<全ての有効なケース>
ヘルプ
リセット
OK
キャンセル
適用

アウトプット

アウトプット

```
> rowPercents(.Table) # 行のパーセント表示
     Outcome
group     0     1 Total Count
    A 50.0 50.0   100    42
    C 73.8 26.2   100    42
```

有効割合

```
> fisher.test(.Table)

     Fisher's Exact Test for Count Data

data:  .Table
p-value = 0.04238
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.1272557 0.9693956
sample estimates:
odds ratio
 0.3594044
```

C/Aのオッズ比
(A/Cのオッズ比は逆数を取ればよい)

```
> res <- NULL
> res <- fisher.test(.Table)
> Fisher.summary.table <- rbind(Fisher.summary.table, summary.table.twoway(table=.Table, res=res))
> colnames(Fisher.summary.table)[length(Fisher.summary.table)] <-  gettextRcmdr(
+   colnames(Fisher.summary.table)[length(Fisher.summary.table)])

> Fisher.summary.table
        Outcome=0 Outcome=1 Fisher検定のP値
group=A        21        21          0.0424
group=C        31        11
```

新薬と既存薬で有効割合に有意な違いが認められた.