

## cs/csi - リスクデータの分析 【評価版】

epitab (tables for epidemiologists) 系コマンドの1つである cs コマンドは ir コマンドの場合と同様、コホート型の研究で記録された発症データを分析するものですが、対象とするのは累積罹患率 (cumulative incidence) データです。この場合、被験者に対する観察時間は一律であるため、データは人時ではなく単なる人の形で記録されることになります。

- |             |           |
|-------------|-----------|
| 1. 累積罹患率データ |           |
| 2. 基本的な用例   | Example 6 |
|             | Example 7 |
| 3. 層化データ    | Example 8 |
| 4. 標準化      | Example 9 |

## 1. 累積罹患率データ

累積罹患率 (cumulative incidence) データとは ir/iri コマンドが対象とする罹患率データ (incidence-rate data) と類似の性格を有するデータですが、分母が人時ではなく単なる人数となる点が異なります。この場合も非症例の被験者群からスタートして、それぞれの被験者が症例に移行したかどうかをある一定の期間モニタします。その際、ある危険因子に曝露されたグループと曝露されていないグループの2つの集団が監視の対象となります。データは次の形式で記録されます。

	曝露	非曝露	計
症例	$a$	$b$	$a + b$
非症例	$c$	$d$	$c + d$
計	$a + c$	$b + d$	$a + b + c + d$

この形式のデータの場合、観測時間はすべての被験者について一律となる点に注意してください。情報は通常リスク比 (risk ratio)

$$\frac{a/(a+c)}{b/(b+d)}$$

の形で集約されます。リスク比が2ということは曝露群に属する被験者の方が非曝露群に属する被験者より症例となる可能性が2倍大きいことを意味します。逆にリスク比が0.5であれば症例となる可能性が半分ということになります。従って効果がなかったことに対応する“null”値は1ということになります。

## 2. 基本的な用例

### ▷ Example 6: csi

評価版では割愛しています。

### ▷ Example 7: cs

Example 6 ではデータを直接ダイアログ上に入力したわけですが、今度は同じデータがファイルとして与えられたときの扱いを示します。使用する Example データセットは `csxmpl.dta` です。

```
. use http://www.stata-press.com/data/r15/csxmpl.dta, clear *1
```

データセットの内容を `list` コマンドにより確認しておきます。

```
. list *2
```

	case	exp	pop
1.	1	1	7
2.	1	0	12
3.	0	1	9
4.	0	0	2

このデータセットの場合には、 $2 \times 2$  分割表の 4 つのセルに対応する人数が `pop` という変数で記録されています。このように集約された形のデータの場合には度数重み (frequency weights) の指定が必要となる点に注意してください。

\*1 メニュー操作 : File ▷ Example Datasets ▷ Stata 15 manual datasets と操作、Base Reference Manual [R] の `epitab` の項よりダウンロードする

\*2 メニュー操作 : Data ▷ Describe data ▷ List data

- Statistics ▷ Epidemiology and related ▷ Tables for epidemiologists ▷ Cohort study risk-ratio etc. と操作
- Main タブ: Case variable: case  
Exposed variable: exp

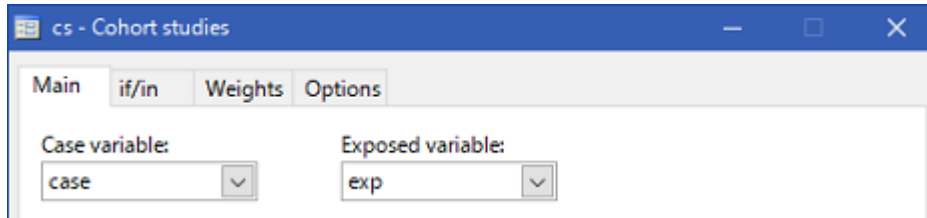


図 2 cs ダイアログ – Main タブ

- Weights タブ: Frequency weight: pop
- Options タブ: Fisher's exact p: ✓

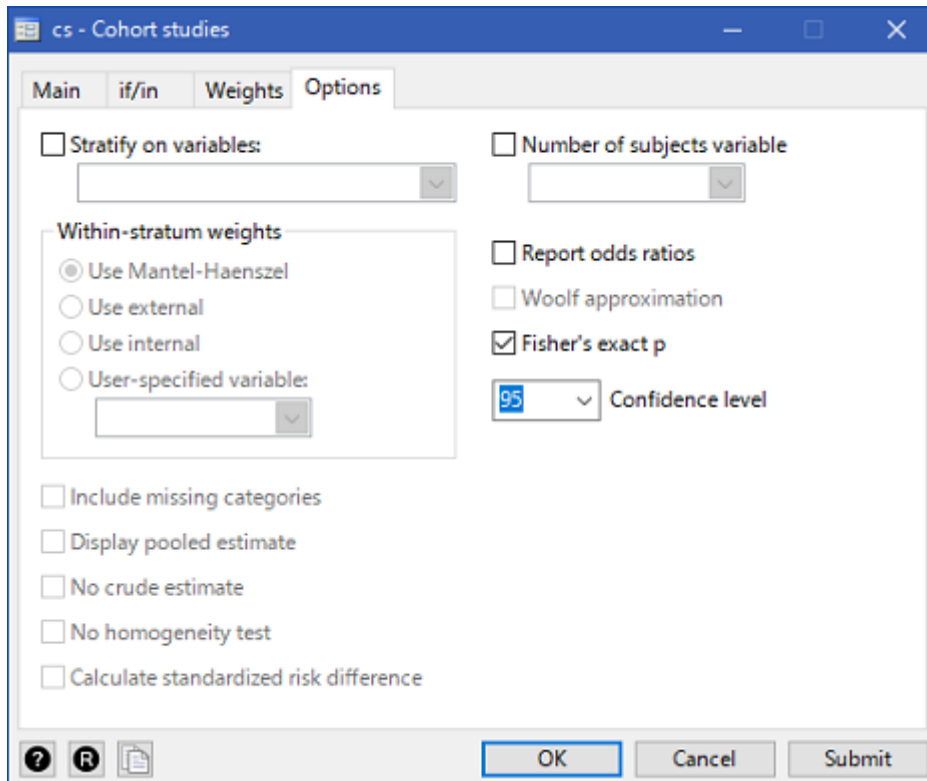


図 3 cs ダイアログ – Options タブ

```
. cs case exp [fweight = pop], exact
```

	exp		Total
	Exposed	Unexposed	
Cases	7	12	19
Noncases	9	2	11
Total	16	14	30
Risk	.4375	.8571429	.6333333
	Point estimate		[95% Conf. Interval]
Risk difference	-.4196429		-.7240828 -.1152029
Risk ratio	.5104167		.2814332 .9257086
Prev. frac. ex.	.4895833		.0742914 .7185668
Prev. frac. pop	.2611111		

1-sided Fisher's exact P = 0.0212  
2-sided Fisher's exact P = 0.0259

csi のときと同じ結果であることが確認できます。

一方、データセット上には個人ごとの情報が記録されている場合もあります。この例で言えば乳児 30 人の観測結果を 30 件のデータとして記録するわけです。cs に対してはそのような個人レベルのデータも入力できません。その場合、データは cs によって自動的に集計され、上記と同じ結果が出力されることとなります。 ◀

### 3. 層化データ

#### ▷ Example 8: 層化データ

評価版では割愛しています。

### 4. 標準化

評価版では割愛しています。

#### ▷ Example 9: 標準化

評価版では割愛しています。

◀

■