

保健統計学特講

2024.5.14

- 交絡と回帰モデル
- 症例数とパワー計算: 保健学共通特講 IV のサンプルサイズの設計の話の裏返し
 - 一般にサンプルサイズが大きいほど検出力 (Statistical power) は大きくなる
 - 検出力は 1 から第二種の過誤 (本当は差があるのにそれを見逃す確率) を引いた値
- 多重検定
 - 3群以上の比較をするとき, 2群間の比較を繰り返すと全体としての第一種の過誤 (本当は差がないのに誤って差があると判定してしまう確率) が大きくなってしまいうので, 個々の検定の有意水準あるいは有意確率を調整する (これを検定の多重性の調整という) 必要がある。多重比較ともいう
 - ボンフェローニの不等式に基づくボンフェローニの方法が最も単純。ホルムの方法や FDR が標準的。

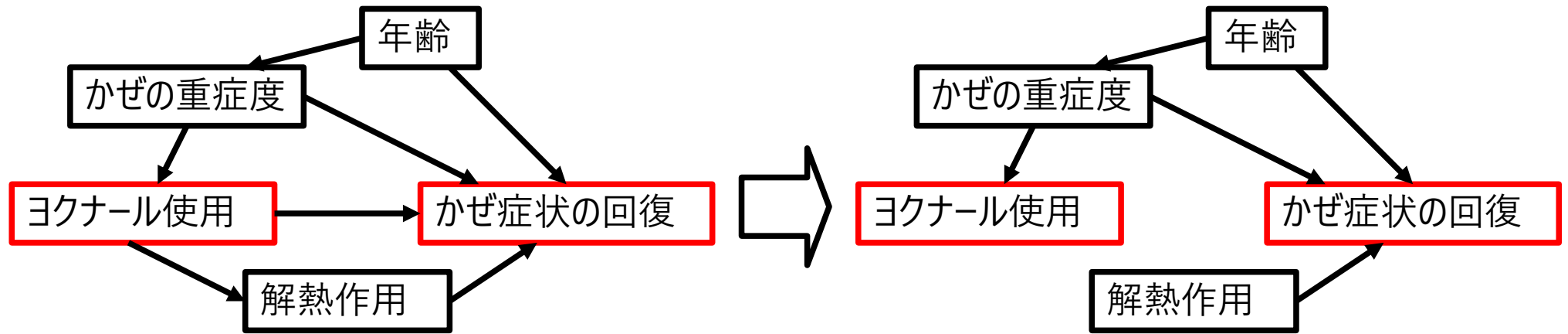
交絡 (confounding) とは

- 要因曝露→結果という因果関係を考えているとき、要因とも結果とも関連があつて、しかも中間媒介因子ではない因子があるとき、これを交絡因子と呼ぶ。交絡因子の影響によって、注目している因果関係は歪められる。この歪みを交絡と呼ぶ→交絡を制御して解析しないと、要因曝露と結果の関係を正しく評価できない
- 交絡要因の必要条件(例:肥満⇒高血圧に対する年齢)
 - 結果(高血圧)に影響している(年齢が高いほど高血圧は増える)
 - 要因曝露に偏りがある(肥満群と非肥満群で年齢分布が異なっている)
 - 中間媒介因子ではない(肥満が高齢をもたらすのではない):中間媒介因子を介する影響を取り除いてしまうと、要因曝露がもたらす結果への効果を過小評価することになるので、中間媒介因子は交絡扱いできない
- 要因曝露に偏りがあることが交絡因子の必要条件だからといって、検定によって分布に有意差があるという理由だけで、変数を交絡として制御するのは誤り
 - RCTでは割り付けた変数以外の変数の分布には差がないようにランダム割り付けを行っているので、差がないことは既知。いくらp値が小さくても、それは差があることを意味しない。第一種の過誤そのもの。
 - 観察研究でも、検定によって要因曝露の有無間で分布に有意差がある変数があったとしても、その変数が結果に影響しているかどうかを考慮していないので交絡の必要条件を満たしていない。

DAG の利用

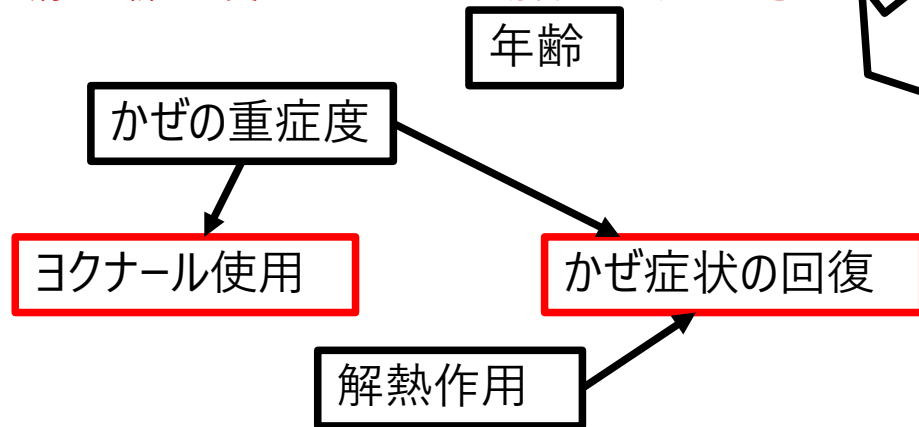
- DAG (Directed Acyclic Graph = 有向非巡回グラフ) は、因果推論をする際に、制御すべき交絡因子を見つけるのに役立つ
- 参考文献としてお薦めしたい本
 - 佐藤俊哉『宇宙怪人しまりす 統計よりも大事なことを学ぶ』朝倉書店、2024
 - 林岳彦『はじめての統計的因果推論』岩波書店、2024
- 観測された変数を四角で囲み、それらの間を矢印の元が原因となる変数、矢印の先が結果となる変数であるように矢印でつなぎ、矢印の向きを辿っていっても元の変数に戻ってこないようなグラフを DAG と呼ぶ
- 複数の観測された変数について想定される因果関係に基づいた DAG を描くと、「裏口テスト」と呼ばれる手法を使って、制御すべき交絡変数を見出すことができる
- (注) その研究では観測されていない変数や、原理的に観測不可能な要因も、結果となる変数に影響を与えている可能性も忘れてはならない。

裏口テストの例 (佐藤 2024、pp.57-63)



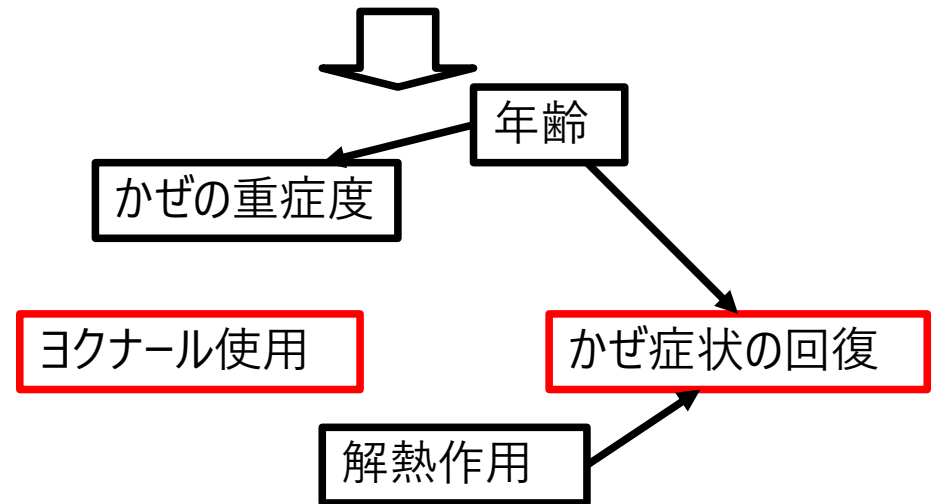
想定される因果関係の図 (DAG になっている)

※ 裏口道がすべてなくなるまで、通り道にある変数を除去する。
 ただし2つの変数から1つの変数に矢印が入る合流点があると、
 そこで道は塞がるので裏口道ではなくなることに、除去によって合
 流が消えて新たな裏口道が生まれる場合がある点に注意



年齢を調整する→裏口道が残るので調整不十分 (→この場合、この調整は不要)

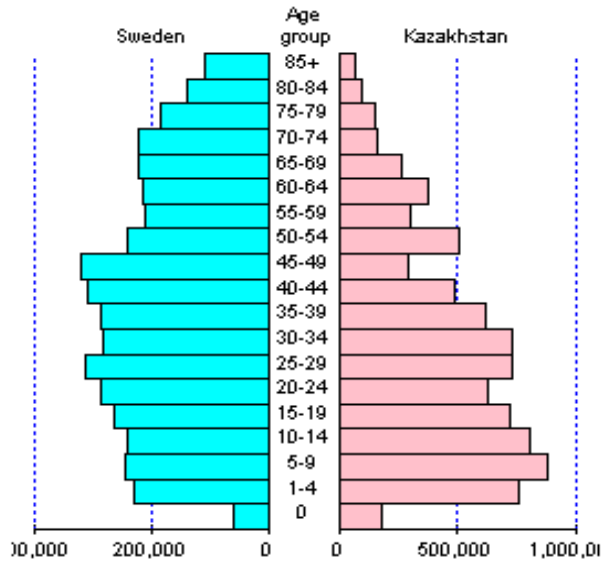
ヨクナールの効果を取り除いてもヨクナール使用と咳症状の回復に関連がある (= 裏口道がある) かを調べる→2つの裏口道がある



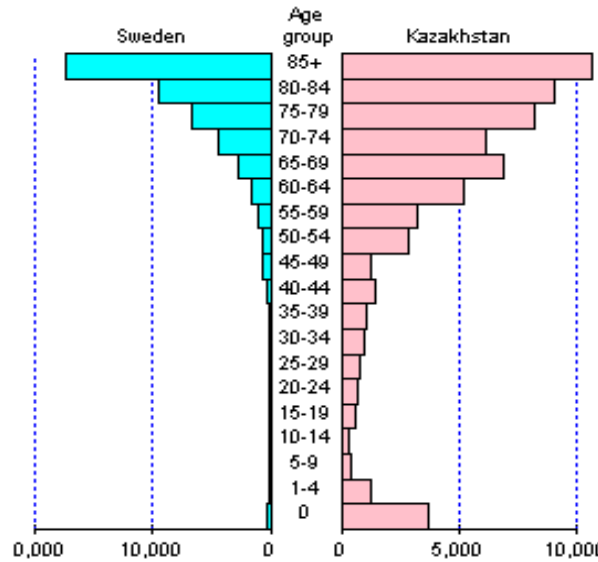
咳の重症度を調整する→裏口道が消える = 交絡として咳の重症度を調整すればヨクナール使用の効果調べることができる

交絡の例 (1)

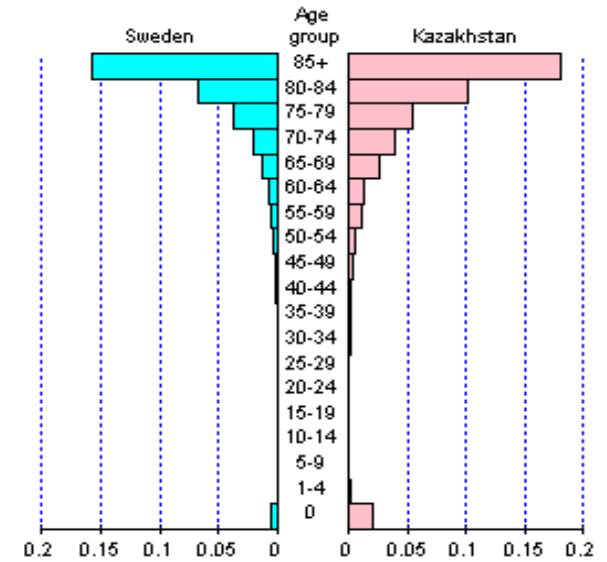
Population structure of Sweden and Kazakhstan females in 1992



Number of death by age in Sweden and Kazakhstan females in 1992



Age specific death rates of Sweden and Kazakhstan females in 1992



- スウェーデンの女性とカザフスタンの女性の 1992 年の死亡率
- スウェーデンの方がカザフスタンより衛生水準・医療水準・生活水準が高い
→ 年齢別死亡率は全年齢層でスウェーデン < カザフスタン
→ 当然, 平均寿命はスウェーデン > カザフスタン
→ が, 粗死亡率はスウェーデン (10.5) > カザフスタン (6.3)
- 理由は以下 2 点
 - どちらの国でも高齢者の死亡率 >> 若者や成人の死亡率
 - 高齢者の割合はスウェーデン > カザフスタン
- 年齢構造が交絡因子となっている

交絡の例 (2)

- 年齢が交絡因子となっているために、喫煙者の方が非喫煙者よりも死亡率が低くなることもある(出典:『ロスマンの疫学』)

表 1-1 英国ウィッカムの女性における調査開始時の喫煙状況による 20 年間の死亡リスク*

生存状況	喫煙者	非喫煙者	合計
死亡	139	230	369
生存	443	502	945
総数	582	732	1,314
リスク (死亡/総数)	0.24	0.31	0.28

* Appleton ら³⁾ のデータによる

表 1-2 英国ウィッカムの女性における調査開始時の年齢別喫煙状況による 20 年間の死亡リスク

年齢 (歳)	生存状況	喫煙者	非喫煙者	総数
18-24	死亡	2	1	3
	生存	53	61	114
	死亡リスク	0.04	0.02	0.03
25-34	死亡	3	5	8
	生存	121	152	273
	死亡リスク	0.02	0.03	0.03
35-44	死亡	14	7	21
	生存	95	114	209
	死亡リスク	0.13	0.06	0.09
45-54	死亡	27	12	39
	生存	103	66	169
	死亡リスク	0.21	0.15	0.19
55-64	死亡	51	40	91
	生存	64	81	145
	死亡リスク	0.44	0.33	0.39
65-74	死亡	29	101	130
	生存	7	28	35
	死亡リスク	0.81	0.78	0.79
75+	死亡	13	64	77
	生存	0	0	0
	死亡リスク	1.00	1.00	1.00

交絡の例 (3)

- ダウン症ハイリスクなのは出生順位が後の子か？ それとも母の年齢が高い子か？ (出典『ロスマンの疫学』)

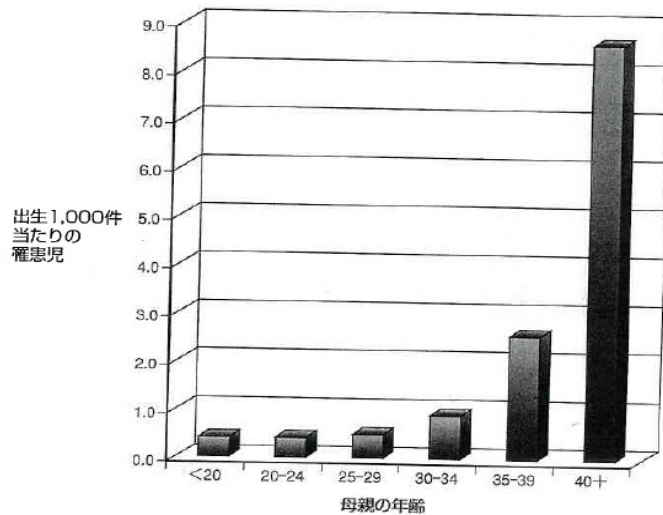
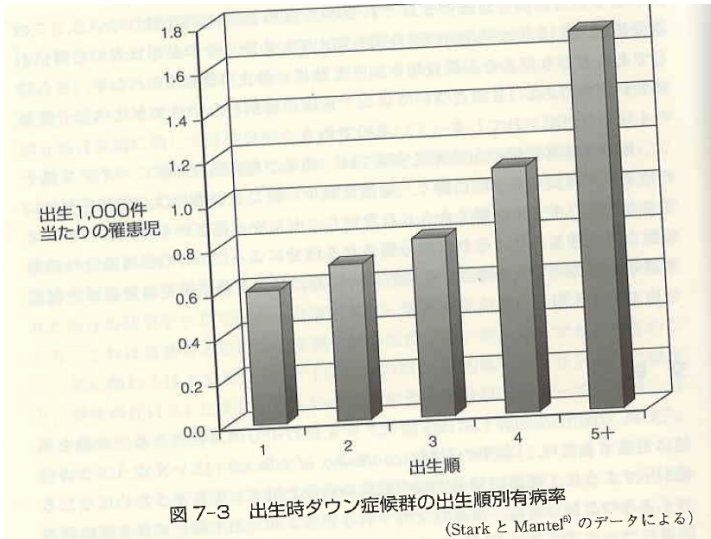


図 7-4 出生時ダウン症候群の母親の年齢別有病率

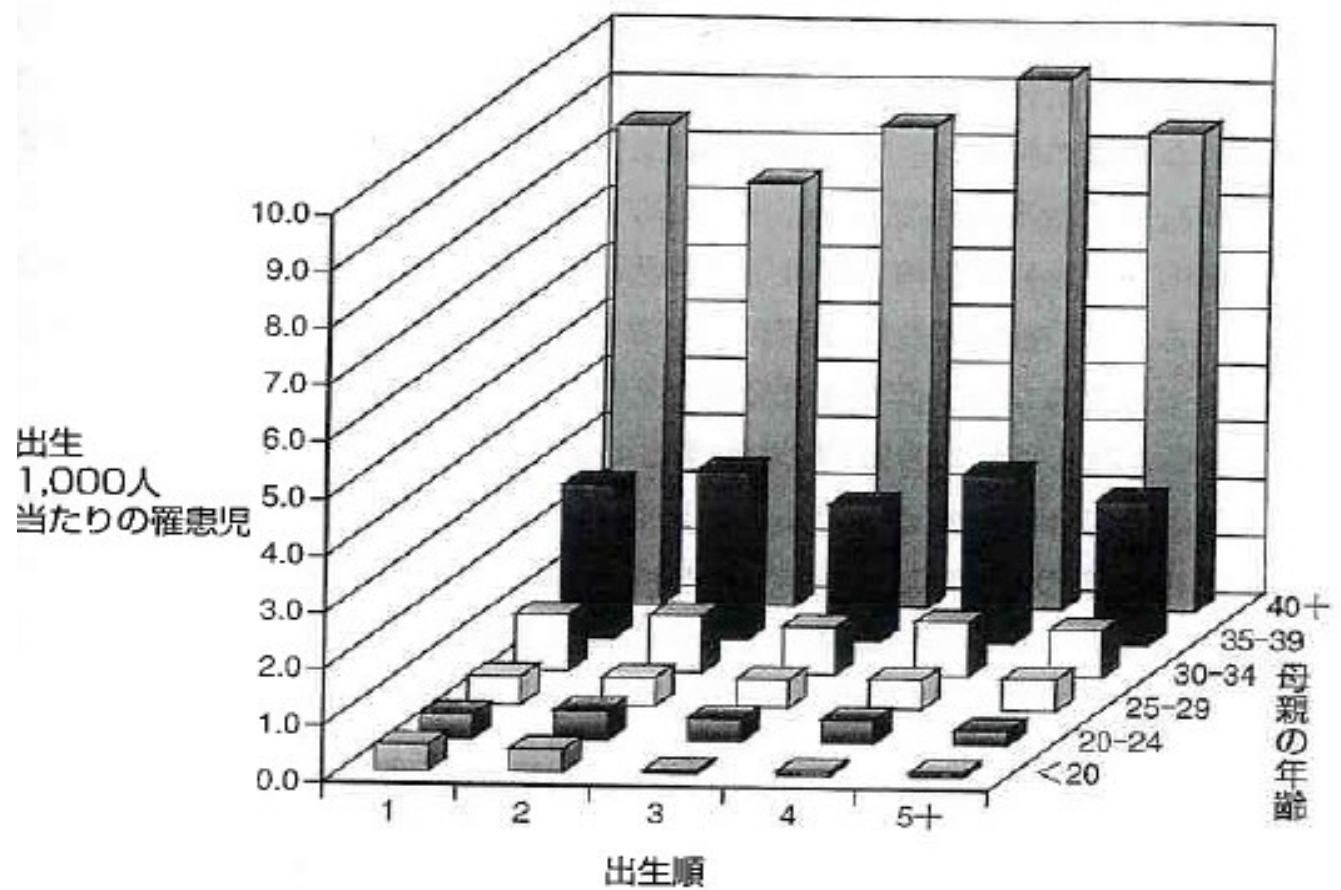


図 7-5 出生時ダウン症候群の出生順および母親の年齢別有病率

交絡の制御

- デザインによる制御
 - 性別や年齢による交絡がありそうな場合は、それらで層別した層別無作為抽出をする、コホート研究なら曝露群と非曝露群を設定するときに年齢や性別をマッチングするか、特定の年齢や性別に限定化して研究すると、それらの影響がデータに表れない
- 解析方法の工夫による制御
 - 層別解析：層別に比較する
 - 標準化：年齢調整死亡率とか
 - プール化：どの層でも同じ関連や違いがあると仮定して要約統計量を計算する（マンテルヘンツェルの要約オッズ比など）
 - 多変量解析：交絡因子も独立変数に含めたロジスティック回帰分析などの一般化線型モデル、傾向スコアを用いた共変量の調整、DID (Difference in Difference) 法など
 - Rothman は、独立変数に含めることによって、他の独立変数の係数が大きく変わるかどうか、交絡因子かどうかを判定する目安になるとしている

ロジスティック回帰分析の例

- エコポイントチェックデータ
<https://minato.sip21c.org/advanced-statistics/ecopx.txt>
- <https://minato.sip21c.org/humeco/ecopoint.html>
 - 高月紘(編著)「自分の暮らしがわかるエコロジー・テスト:環境問題は生活のエコ度チェックから」(講談社ブルーバックス)
 - CGI 化→表にチェックして,「テスト実行」
 - エコポイント総合点が 99.8 点満点(丸め誤差で≠ 100)
 - 温暖化問題, 廃棄物問題, 水質汚染問題, 大気汚染問題, 有害化学物質問題のエコポイントがそれぞれ 100 点満点
 - 若い層を中心とした対象者 356 人の平均が 42.6
環境問題の講演を聞きに来た人たち 182 人の平均が 48.1

お名前 (本名でなくて構いません)					
年齢	10代 ▼				
性別	男性 ▼				
同居人の人数 (本人を含む)	1人 ▼				
以下の取り組みについて	いつも	大体	時々	たまに(A)	皆無(B)
1. 新聞・雑誌をリサイクルに出している	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. 古紙100%のトイレトペーパーを使用している	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. 飲料容器やトレーをリサイクルに出している	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. 買い物袋を持参している	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. 着る服で調節して, 冷暖房をできるだけ控えている	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

質問項目

AGE	年齢(10歳階級)	0=10-19	1=20-29	2=30-39	3=40-49	4=50-59	5=60-69	6=70-
SEX	性別	0=M	1=F					
FAMSIZE	自分を含む同居人数							
Q01	新聞雑誌リサイクル	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q02	古紙100%トイレトペーパー使用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q03	飲料容器トレーリサイクル	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q04	買い物袋持参	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q05	冷暖房より着る服で調節	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q06	食材を期限切れで捨てない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q07	風呂は家族で続けて入る(一人暮らしの場合はお湯を少なくする)	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q08	風呂水を洗濯等に利用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q09	車のアイドリングストップ(車を持っていない人は「いつも」になる)	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q10	マイカーを避けて公共交通を利用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q11	太陽熱温水器を利用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q12	家電製品は省エネ型以外は買わない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q13	米のとぎ汁は流さず有効利用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q14	油をふき取ってから皿洗い	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q15	塩ビ系プラ製品を買わない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q16	洗剤は合成洗剤でなく石鹼利用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q17	洗剤(石鹼を含む)をはかって適量使用	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q18	除草剤や殺虫剤を使わない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q19	車のバッテリーや電池を適正処理	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q20	トイレや風呂場の強力洗浄剤を利用しない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q21	有機溶剤を利用しない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q22	有機農産物を選ぶ	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q23	地場の農産物を選ぶ	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q24	早寝早起き	0=いつも	1=大体	2=時々	3=たまに	4=皆無		
Q25	煙草を吸わない	0=いつも	1=大体	2=時々	3=たまに	4=皆無		

ロジスティック回帰 (1)

- Q4 と Q10 の関係には年齢と性別の影響が両方ある(年齢と性別は Q4 と Q10 の関係を交絡している)
- 年齢と性別と Q10 を独立変数, Q4 を従属変数としてロジスティック回帰分析
- 年齢と性別の影響を調整して,
 - Q10 が TRUE の人(いつも / 大体 / 時々公共交通利用 = 公共交通利用者)は
 - Q10 が FALSE の人(公共交通利用がたまに / 皆無 = マイカー族)に比べて,
 - どれくらい Q4 が TRUE である(いつも / 大体 / 時々買い物袋持参 = エコバッグ派である)割合が高いかわかる

ロジスティック回帰 (2)

- Q4 が TRUE である割合を p とすると,
 $\log(p/(1-p)) = \beta_0 + \beta_1 \text{AGE} + \beta_2 \text{SEX} + \beta_3 \text{Q10}$
TRUE を 1, FALSE を 0 とすると, 他の変数を調整した上での Q10 の効果は,
 - Q10 が TRUE の場合の Q4 が TRUE である割合を p_1 とすれば,
 $\log(p_1/(1-p_1)) = \beta_0 + \beta_1 \text{AGE} + \beta_2 \text{SEX} + \beta_3$
 - Q10 が FALSE の場合の Q4 が TRUE である割合を p_0 とすれば,
 $\log(p_0/(1-p_0)) = \beta_0 + \beta_1 \text{AGE} + \beta_2 \text{SEX}$
 - 辺々引けば $\log(p_1/(1-p_1)) - \log(p_0/(1-p_0)) = \beta_3$
 - 結局この左辺は, $\log[(p_1/(1-p_1))/(p_0/(1-p_0))]$ となる。これは対数オッズ比である。つまり, AGE と SEX の影響を調整した上で Q10 が TRUE である人は FALSE の人に比べて, $\exp(\beta_3)$ 倍, Q4 が TRUE になりやすい
 - 「性別と年齢の影響を調整した上で, 公共交通利用者はマイカー族に比べて, エコバッグ派である可能性が $\exp(\beta_3)$ 倍大きい」ということである。 $\beta_0 \sim \beta_3$ は, 一般化線型モデルの係数として推定される。

ロジスティック回帰 (3)

- R のコード

```
x <- read.delim("https://minato.sip21c.org/advanced-statistics/ecopx.txt") # 読み込み
x$Q04c <- ifelse(x$Q04>=2.5, 1, 0) # スコアなので 1/0 に変換
x$Q10c <- ifelse(x$Q10>=4.8, 1, 0) # スコアなので 1/0 に変換
res <- glm(Q04c ~ Q10c+AGE+SEX, family=binomial(logit), data=x) # ロジスティック回帰
summary(res) # 回帰分析の結果を出力
exp(coef(res)) # 対数オッズ比なので係数の exp() をとってオッズ比を得る
exp(confint(res)) # 信頼区間の exp() をとってオッズ比の 95% 信頼区間を得る
library(fmsb) # fmsb パッケージを読み込む
NagelkerkeR2(res) # モデルの説明力を示す Nagelkerke の R2 の計算
library(stargazer) # きれいに出力するための stargazer パッケージを読み込む
sink("logistic-res.tex") # このファイルを TeXLive でコンパイルすると右の表生成
cat("\\documentclass[a4paper,12pt]{article}\\n\\begin{document}\\n")
stargazer(res, ci=TRUE, apply.coef=exp, type="latex",
  notes=sprintf("Nagelkerke's R-squared: %4.2f", NagelkerkeR2(res)$R2))
cat("\\end{document}\\n")
sink()
```

- 出力結果

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.5479	0.3232	-1.695	0.09006
Q10c	1.3036	0.2886	4.517	6.28e-06 ← 6.28×10 ⁻⁶ の意味
AGE	-0.1864	0.1106	-1.685	0.09202
SEX	0.7565	0.2751	2.750	0.00596

AIC: 391.21

	Q10c	AGE	SEX
(Intercept)	0.5781719	0.8299727	2.1308305
	2.5 %	97.5 %	
(Intercept)	0.3032334	1.081392	
Q10c	2.1093893	6.560746	
AGE	0.6667677	1.030981	
SEX	1.2550376	3.700273	

\$R2 [1] 0.1333421 ← モデルの説明力は 13% 程度

Table 1:

<i>Dependent variable:</i>	
	Q04c
Q10c	3.683*** (3.117, 4.248)
AGE	0.830*** (0.613, 1.047)
SEX	2.131*** (1.592, 2.670)
Constant	0.578* (-0.055, 1.212)
Observations	308
Log Likelihood	-191.603
Akaike Inf. Crit.	391.206

Note: *p<0.1; **p<0.05; ***p<0.01
Nagelkerke's R-squared: 0.13

jamovi でやると

jamovi - Untitled

Variables Data Analyses Edit

Exploration T-Tests ANOVA Regression Frequencies Factor rosetta SEM Modules

Binomial Logistic Regression

Collinearity statistics

Model Fit

Fit Measures

Deviance
 AIC
 BIC
 Overall model test

Pseudo R²

McFadden's R²
 Cox & Snell's R²
 Nagelkerke's R²
 Tjur's R²

Omnibus Tests

Likelihood ratio tests

Estimate (Log Odds Ratio)

Confidence interval
Interval 95 %

Odds ratio
 Confidence interval
Interval 95 %

Estimated Marginal Means
Prediction
Save

Binomial Logistic Regression

Model Fit Measures

Model	Deviance	AIC	R ² _{McF}	R ² _N
1	383	391	0.0771	0.133

Model Coefficients - Q04C

Predictor	Estimate	SE	Z	p	Odds ratio	95% Confidence Interval	
						Lower	Upper
Intercept	-0.548	0.323	-1.70	0.090	0.578	0.307	1.09
Q10C:							
1 - 0	1.304	0.289	4.52	< .001	3.683	2.092	6.48
AGE	-0.186	0.111	-1.68	0.092	0.830	0.668	1.03
SEX	0.757	0.275	2.75	0.006	2.131	1.243	3.65

Note. Estimates represent the log odds of "Q04C = 1" vs. "Q04C = 0"

Assumption Checks

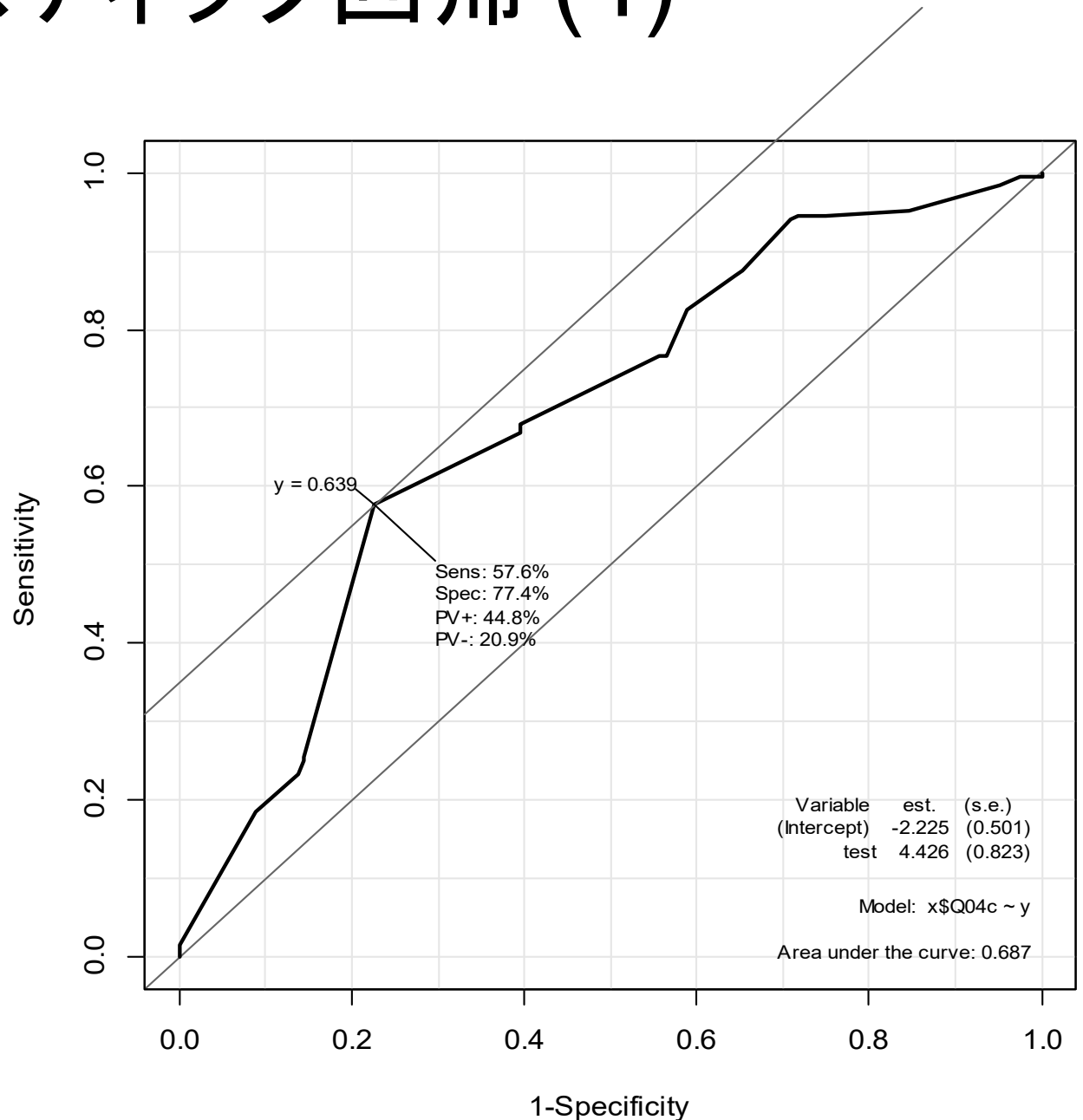
Collinearity Statistics

	VIF	Tolerance
Q10C	1.09	0.916
AGE	1.04	0.957
SEX	1.08	0.929

[3]

ロジスティック回帰 (4)

- 得られたロジスティック回帰式を使って Q04c を予測するための ROC 分析
- ```
y <- predict(res,
 data.frame(Q10c=x$Q10c,
 AGE=x$AGE, SEX=x$SEX),
 "response")
```
- `library(Epi)`
- `ROC(y, x$Q04c)`
- 最適カットオフ 0.639,  
 感度 0.576, 特異度 0.774
- AUC は 0.687  
 (あまり良くない)





# 混合効果モデル

- アウトカムとなる健康事象が個人レベルで、個人レベルの要因と集団レベルの(サブグループ)要因があり、個人レベルでの回帰モデルの切片と傾きの片方または両方がサブグループによって違う場合、混合効果モデルを当てはめて交絡を調整する
- 例えば、Demographic and Health Surveys の 2017 年インドネシアのデータ (IDIR71FL.SAV という SPSS 形式のデータ)には V012 (年齢), V025 (都市部/農村部), V106 (教育歴), V201 (既往出生児数), V361 (避妊利用)が含まれている。既往出生児数を他の変数で説明するモデルを考える。
  - `library(foreign)`
  - `x <- read.spss("IDIR71FL.SAV")`
  - `lm(V201 ~ V012+V025+V106+V361, data=x) # 重回帰`
  - `glm(V201 ~ V012+V025+V106+V361, data=x, family="poisson") # ポアソン回帰`
  - `library(lmerTest)`
  - `lmer(V201 ~ V012 + V106 + V361 + (1 | V025), data=x, REML=FALSE) # 混合`
- この結果はどれも大差ない。係数の信頼区間を得るには `confint(coef())` を使う。年齢の係数は正, 教育歴なしに比べどの教育歴も負, 避妊利用は現在利用中に比べたのカテゴリは負, 都市部に比べ農村部は正の係数となっている
- 混合効果モデルにはグループレベルの変数により切片だけが異なるランダム切片モデル(上のコードはそれを仮定している)と、傾きも異なると仮定するモデルがある

# 結果の比較 (重回帰とポアソン回帰)

Call:  
lm(formula = V201 ~ V012 + V025 + V106 + V361, data = x)

Residuals:  
Min 1Q Median 3Q Max  
-4.2150 -0.6136 -0.0643 0.3418 10.9900

Coefficients:  
(Intercept) 0.6291598 0.0482432 13.041 < 2e-16 \*\*\*  
V012 0.0762938 0.0006318 120.759 < 2e-16 \*\*\*  
V025Rural 0.1844247 0.0105574 17.469 < 2e-16 \*\*\*  
V106Primary -0.6379498 0.0394769 -16.160 < 2e-16 \*\*\*  
V106Secondary -0.9298056 0.0392179 -23.709 < 2e-16 \*\*\*  
V106Higher -1.2269346 0.0403892 -30.378 < 2e-16 \*\*\*  
V361Used since last birth -0.3381545 0.0173691 -19.469 < 2e-16 \*\*\*  
V361Used before last birth -0.1027063 0.0188352 -5.453 4.98e-08 \*\*\*  
V361Never used -1.1855481 0.0133170 -89.025 < 2e-16 \*\*\*

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

Residual standard error: 1.138 on 49618 degrees of freedom  
Multiple R-squared: 0.5461, Adjusted R-squared: 0.546  
F-statistic: 7461 on 8 and 49618 DF, p-value: < 2.2e-16

glm(formula = V201 ~ V012 + V025 + V106 + V361, family = "poisson", data = x)

Deviance Residuals:  
Min 1Q Median 3Q Max  
-3.4452 -0.7446 -0.3526 0.3175 6.5011

Coefficients:  
(Intercept) -0.6740534 0.0281197 -23.97 <2e-16 \*\*\*  
V012 0.0522313 0.0004569 114.33 <2e-16 \*\*\*  
V025Rural 0.1268545 0.0070699 17.94 <2e-16 \*\*\*  
V106Primary -0.2814560 0.0194853 -14.45 <2e-16 \*\*\*  
V106Secondary -0.4031578 0.0196109 -20.56 <2e-16 \*\*\*  
V106Higher -0.5929119 0.0216679 -27.36 <2e-16 \*\*\*  
V361Used since last birth -0.1458802 0.0102629 -14.21 <2e-16 \*\*\*  
V361Used before last birth -0.1405757 0.0100127 -14.04 <2e-16 \*\*\*  
V361Never used -1.2205031 0.0122418 -99.70 <2e-16 \*\*\*

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

(Dispersion parameter for poisson family taken to be 1)  
Null deviance: 90239 on 49626 degrees of freedom  
Residual deviance: 37663 on 49618 degrees of freedom  
AIC: 129952

Number of Fisher Scoring iterations: 5

Table 1:

|                            | Dependent variable:          |                      |
|----------------------------|------------------------------|----------------------|
|                            | OLS<br>(1)                   | Poisson<br>(2)       |
| V012                       | 0.076***<br>(0.001)          | 0.052***<br>(0.0005) |
| V025Rural                  | 0.184***<br>(0.011)          | 0.127***<br>(0.007)  |
| V106Primary                | -0.638***<br>(0.039)         | -0.281***<br>(0.019) |
| V106Secondary              | -0.930***<br>(0.039)         | -0.403***<br>(0.020) |
| V106Higher                 | -1.227***<br>(0.040)         | -0.593***<br>(0.022) |
| V361Used since last birth  | -0.338***<br>(0.017)         | -0.146***<br>(0.010) |
| V361Used before last birth | -0.103***<br>(0.019)         | -0.141***<br>(0.010) |
| V361Never used             | -1.186***<br>(0.013)         | -1.221***<br>(0.012) |
| Constant                   | 0.629***<br>(0.048)          | -0.674***<br>(0.028) |
| Observations               | 49,627                       | 49,627               |
| R <sup>2</sup>             | 0.546                        |                      |
| Adjusted R <sup>2</sup>    | 0.546                        |                      |
| Log Likelihood             |                              | -64,967.080          |
| Akaike Inf. Crit.          |                              | 129,952.200          |
| Residual Std. Error        | 1.138 (df = 49618)           |                      |
| F Statistic                | 7,461.021*** (df = 8; 49618) |                      |

Note: 1 \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

# 線形混合モデルの結果

Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's method  
[LmerModLmerTest]

Formula: V201 ~ V012 + V106 + V361 + (1 | V025)

| AIC      | BIC      | logLik   | deviance | df.resid |
|----------|----------|----------|----------|----------|
| 153677.8 | 153765.9 | -76828.9 | 153657.8 | 49617    |

Scaled residuals:

| Min     | 1Q      | Median  | 3Q     | Max    |
|---------|---------|---------|--------|--------|
| -3.7054 | -0.5386 | -0.0568 | 0.2999 | 9.6585 |

Random effects:

| Groups | Name        | Variance | Std.Dev. |
|--------|-------------|----------|----------|
| V025   | (Intercept) | 0.008444 | 0.09189  |
|        | Residual    | 1.294560 | 1.13779  |

Number of obs: 49627, groups: V025, 2

Fixed effects:

|                            | Estimate   | Std. Error | df        | t value | Pr(> t )     |
|----------------------------|------------|------------|-----------|---------|--------------|
| (Intercept)                | 7.220e-01  | 8.041e-02  | 4.630e+00 | 8.978   | 0.000423 *** |
| V012                       | 7.629e-02  | 6.317e-04  | 4.963e+04 | 120.761 | < 2e-16 ***  |
| V106Primary                | -6.382e-01 | 3.947e-02  | 4.963e+04 | -16.166 | < 2e-16 ***  |
| V106Secondary              | -9.303e-01 | 3.921e-02  | 4.963e+04 | -23.724 | < 2e-16 ***  |
| V106Higher                 | -1.228e+00 | 4.038e-02  | 4.962e+04 | -30.396 | < 2e-16 ***  |
| V361Used since last birth  | -3.382e-01 | 1.737e-02  | 4.963e+04 | -19.471 | < 2e-16 ***  |
| V361Used before last birth | -1.027e-01 | 1.883e-02  | 4.963e+04 | -5.455  | 4.92e-08 *** |
| V361Never used             | -1.186e+00 | 1.332e-02  | 4.963e+04 | -89.038 | < 2e-16 ***  |

---

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

Correlation of Fixed Effects:

|             | (Intr) | V012   | V106Pr | V106Sc | V106Hg | V36slb | V36blb      |
|-------------|--------|--------|--------|--------|--------|--------|-------------|
| V012        |        | -0.340 |        |        |        |        |             |
| V106Primary |        | -0.485 | 0.077  |        |        |        |             |
| V106Secndry |        | -0.517 | 0.165  | 0.943  |        |        |             |
| V106Higher  |        | -0.491 | 0.133  | 0.912  | 0.939  |        |             |
| V361Usdsnlb |        | -0.059 | 0.030  | 0.013  | 0.013  | 0.013  |             |
| V361Usdbflb |        | -0.014 | -0.137 | 0.022  | 0.023  | 0.026  | 0.184       |
| V361Nevrusd |        | -0.229 | 0.465  | 0.101  | 0.094  | 0.044  | 0.281 0.176 |

# jamovi の場合

## 線形回帰

モデル適合度指標

モデル

|   | R     | R <sup>2</sup> |
|---|-------|----------------|
| 1 | 0.739 | 0.546          |

モデル係数 - V201

予測変数

|                                                  | 推定値     | 標準誤差     | t      | p      |
|--------------------------------------------------|---------|----------|--------|--------|
| 切片 <sup>a</sup>                                  | 0.6292  | 0.0482   | 13.04  | < .001 |
| V012                                             | 0.0763  | 6.32E-04 | 120.76 | < .001 |
| V025:<br>Rural – Urban                           | 0.1844  | 0.0106   | 17.47  | < .001 |
| V106:<br>Primary – No education                  | -0.6379 | 0.0395   | -16.16 | < .001 |
| Secondary – No education                         | -0.9298 | 0.0392   | -23.71 | < .001 |
| Higher – No education                            | -1.2269 | 0.0404   | -30.38 | < .001 |
| V361:<br>Used since last birth – Currently using | -0.3382 | 0.0174   | -19.47 | < .001 |
| Used before last birth – Currently using         | -0.1027 | 0.0188   | -5.45  | < .001 |
| Never used – Currently using                     | -1.1855 | 0.0133   | -89.03 | < .001 |

### 線形回帰

モデル適合度指標

| モデル | R     | R <sup>2</sup> |
|-----|-------|----------------|
| 1   | 0.739 | 0.546          |

モデル係数 - V201

| 予測変数                                             | 推定値     | 標準誤差    | t      | p      |
|--------------------------------------------------|---------|---------|--------|--------|
| 切片 <sup>a</sup>                                  | 0.6292  | 0.0482  | 13.04  | < .001 |
| V012                                             | 0.0763  | 6.32e-4 | 120.76 | < .001 |
| V025:<br>Rural – Urban                           | 0.1844  | 0.0106  | 17.47  | < .001 |
| V106:<br>Primary – No education                  | -0.6379 | 0.0395  | -16.16 | < .001 |
| Secondary – No education                         | -0.9298 | 0.0392  | -23.71 | < .001 |
| Higher – No education                            | -1.2269 | 0.0404  | -30.38 | < .001 |
| V361:<br>Used since last birth – Currently using | -0.3382 | 0.0174  | -19.47 | < .001 |
| Used before last birth – Currently using         | -0.1027 | 0.0188  | -5.45  | < .001 |
| Never used – Currently using                     | -1.1855 | 0.0133  | -89.03 | < .001 |

<sup>a</sup> 基準レベルの値を示しています

### Generalized Linear Models

Negative binomial     Custom  
 Distribution: Gaussian  
 Link Function: Identity

Dependent Variable: V201  
 Factors: V025, V106, V361  
 Covariates: V012

**Effect Size**  
 Odd Ratios (expB)     For exp(B)  
 Relative Risk (RR)     For estimates

**Confidence Intervals**  
 For exp(B)     For estimates  
 Interval: 95 %

### Model Results

Loglikelihood ratio tests

|      | X <sup>2</sup> | df | p      |
|------|----------------|----|--------|
| V012 | 13467          | 1  | < .001 |
| V025 | 322            | 1  | < .001 |
| V106 | 1040           | 3  | < .001 |
| V361 | 12810          | 3  | < .001 |

Parameter Estimates

| Names       | Effect                                   | Estimate | SE      | exp(B) | 95% Exp(B) Confidence Interval |       | z     | p      |
|-------------|------------------------------------------|----------|---------|--------|--------------------------------|-------|-------|--------|
|             |                                          |          |         |        | Lower                          | Upper |       |        |
| (Intercept) | (Intercept)                              | 0.3430   | 0.00683 | 1.409  | 1.390                          | 1.428 | 50.2  | < .001 |
| V012        | V012                                     | 0.0522   | 4.57e-4 | 1.054  | 1.053                          | 1.055 | 114.3 | < .001 |
| V0251       | Rural - Urban                            | 0.1269   | 0.00707 | 1.135  | 1.120                          | 1.151 | 17.9  | < .001 |
| V1061       | Primary - No education                   | -0.2815  | 0.01949 | 0.755  | 0.727                          | 0.784 | -14.4 | < .001 |
| V1062       | Secondary - No education                 | -0.4032  | 0.01961 | 0.668  | 0.643                          | 0.695 | -20.6 | < .001 |
| V1063       | Higher - No education                    | -0.5929  | 0.02167 | 0.553  | 0.530                          | 0.577 | -27.4 | < .001 |
| V3611       | Used since last birth - Currently using  | -0.1459  | 0.01026 | 0.864  | 0.847                          | 0.882 | -14.2 | < .001 |
| V3612       | Used before last birth - Currently using | -0.1406  | 0.01001 | 0.869  | 0.852                          | 0.886 | -14.0 | < .001 |
| V3613       | Never used - Currently using             | -1.2205  | 0.01224 | 0.295  | 0.288                          | 0.302 | -99.7 | < .001 |

# jamovi によるポアソン混合効果モデル

Generalized Mixed Models →

Logistic       Poisson  
 Probit       Negative Binomial

**Custom Model**

Custom

Distribution: Gaussian

Link Function: Identity

Dependent Variable: V201

Factors: V106, V361

Covariates: V012

Cluster variables: V025

Effect Size:  Odd Ratios (expB)

Confidence Intervals:  For exp(B)

## Model Results

### Fixed Effect Omnibus tests

|      | $\chi^2$ | df   | p      |
|------|----------|------|--------|
| V361 | 10045    | 4.00 | < .001 |
| V012 | 13068    | 1.00 | < .001 |
| V106 | 1049     | 3.00 | < .001 |

### Fixed Effects Parameter Estimates

| Names                      | Effect                     | Estimate | SE      | exp(B) | 95% Exp(B) Confidence Interval |       | z     | p      |
|----------------------------|----------------------------|----------|---------|--------|--------------------------------|-------|-------|--------|
|                            |                            |          |         |        | Lower                          | Upper |       |        |
| V361Currently using        | V361Currently using        | 0.7198   | 0.0452  | 2.054  | 1.880                          | 2.244 | 15.9  | < .001 |
| V361Used since last birth  | V361Used since last birth  | 0.5739   | 0.0459  | 1.775  | 1.622                          | 1.942 | 12.5  | < .001 |
| V361Used before last birth | V361Used before last birth | 0.5792   | 0.0460  | 1.785  | 1.631                          | 1.953 | 12.6  | < .001 |
| V361Never used             | V361Never used             | -0.5008  | 0.0461  | 0.606  | 0.554                          | 0.663 | -10.9 | < .001 |
| V012                       | V012                       | 0.0522   | 4.57e-4 | 1.054  | 1.053                          | 1.055 | 114.3 | < .001 |
| V1061                      | Primary - No education     | -0.2816  | 0.0195  | 0.755  | 0.726                          | 0.784 | -14.5 | < .001 |
| V1062                      | Secondary - No education   | -0.4034  | 0.0196  | 0.668  | 0.643                          | 0.694 | -20.6 | < .001 |
| V1063                      | Higher - No education      | -0.5933  | 0.0217  | 0.553  | 0.530                          | 0.576 | -27.4 | < .001 |

### Random Components

| Groups    | Name        | SD     | Variance | ICC |
|-----------|-------------|--------|----------|-----|
| V025      | (Intercept) | 0.0632 | 0.00400  |     |
| Residuals |             | 1.0000 | 1.00000  |     |

注: Number of Obs: 49627 , groups: V025 2

# DID（時間経過が交絡である場合）

- Snow J (1855) が分析したロンドンのコレラのデータ
- 1854 年の感染地図と自然実験（右の 2 社の両方から給水を受けている地域で、水の塩分を調べて、どちらの会社からの給水か特定し、S&V 社からの給水を受けている群で死亡リスクが数倍あることを示した）は有名
- 1849 年には両社ともテムズ川下流から採水していたが、1852 年に L 社は上流に採水地点を移した

| 水道会社 | 1849 年コレラ死者数 | 1854 年コレラ死者数 |
|------|--------------|--------------|
| S&V  | 2261         | 2458         |
| L    | 162          | 37           |
| 両方から | 3905         | 2547         |

- 採水を変えなかった（非介入群）S&V 社から給水を受けている群は死者やや増、採水を変えた（介入群）L 社から給水を受けている群は死者減。ただしその減少が採水点を変えた効果かどうかは、時間経過があるので不明
- 差の差を取ると 43 % 減  
 $(37 - 162) / 162 - (2458 - 2261) / 2261 = -0.43$

# 症例数とパワー計算 (1)

- 「保健学研究共通特講 IV」テキスト第 4 章を参照  
<https://minato.sip21c.org/ebhc-text.pdf>
- 例えば独立 2 群の平均値の差の t 検定をする場合
  - 標準偏差, 意味のある差,  $\alpha$  エラー,  $\beta$  エラーを決めればサンプルサイズを計算できる

$$n = 2(z_{\alpha/2} - z_{1-\beta})^2 SD^2 / d^2 + z_{\alpha/2}^2 / 4$$

⇔ サンプルサイズ, 標準偏差, 実際の差,  $\alpha$  エラーがわかれば  
 $\beta$  エラーがわかる →  $1 - \beta$  エラーとして検出力も

$$z_{1-\beta} = z_{\alpha/2} - \sqrt{\frac{(n - z_{\alpha/2}^2 / 4) \cdot d^2}{2 \cdot SD^2}}$$

- パワー計算とは検出力を求めること。
- 他の条件が同じなら, サンプルサイズが大きいほど検出力は上がる



# 症例数とパワー計算 (2) 計算例

- 独立2群間の両側t検定をして、サンプルサイズが各群100人、平均値の差が1.5、標準偏差が4で、有意水準が0.05の場合、
- PSによれば(右図)検出力は0.752。G\*Powerは次スライド
- Rでは  
`power.t.test(n=100, delta=1.5, sd=4, sig.level=0.05, type="two.sample", alternative="two.sided")`  
と打てば、`power=0.7513...`
- EZRでは0.755となる
- 小数点以下3桁目が微妙に違うが気にしなくて良い

Power and Sample Size Program: Main Window

File Edit Log Help

Survival t-test Regression 1 Regression 2 Dichotomous Mantel-Haenszel Log

Output [Studies that are analyzed by t-tests](#)

[What do you want to know?](#) Power

[Power](#) .752

Design

[Paired or independent?](#) Independent

Input

$\alpha$  0.05  $\delta$  1.5 Calculate

$n$  100  $\sigma$  4 Graphs

$m$  1

Description

We are planning a study with 100 experimental subjects and 100 control subjects. In a previous study the response within each subject group was normally distributed with standard deviation 4. If the true difference in the experimental and control means is 1.5, we will be able to reject the null hypothesis that the population means of the experimental and control groups are equal with probability (power) .752. The Type I error probability associated with this test of this null hypothesis is 0.05.

PS version 3.1.2 Copy to Log Exit

# G\*Power の場合

The screenshot displays the G\*Power 3.1.9.7 interface. At the top, the title bar reads "G\*Power 3.1.9.7". The menu bar includes "File", "Edit", "View", "Tests", "Calculator", and "Help". The main window is titled "Central and noncentral distributions" and "Protocol of power analyses".

The central graph shows two normal distributions: a solid red curve centered at 0 and a dashed blue curve centered at approximately 2.65. Two vertical green lines are drawn at  $t = -2$  and  $t = 1.97202$ . The area under the red curve to the right of  $t = 1.97202$  is shaded red and labeled  $\frac{\alpha}{2}$ . The area under the blue curve to the left of  $t = 1.97202$  is shaded blue and labeled  $\beta$ . The text "critical t = 1.97202" is positioned above the right vertical line.

The "Test family" is set to "t tests" and the "Statistical test" is "Means: Difference between two independent means (two groups)". The "Type of power analysis" is "Post hoc: Compute achieved power - given  $\alpha$ , sample size, and effect size".

**Input Parameters:**

- Tail(s): Two
- Determine =>
- Effect size d: 0.3750000
- $\alpha$  err prob: 0.05
- Sample size group 1: 100
- Sample size group 2: 100

**Output Parameters:**

- Noncentrality parameter  $\delta$ : 2.6516504
- Critical t: 1.9720175
- Df: 198
- Power ( $1 - \beta$  err prob): 0.7513714 (circled in red)

On the right side, there are two radio button options for sample sizes:

- $n1 \neq n2$ : Mean group 1 (0), Mean group 2 (1), SD  $\sigma$  within each group (0.5)
- $n1 = n2$ : Mean group 1 (0), Mean group 2 (1.5), SD  $\sigma$  group 1 (4), SD  $\sigma$  group 2 (4)

At the bottom right, there are buttons for "Calculate", "Calculate and transfer to main window", and "Close". The "Effect size d" is set to 0.375.

At the bottom center, there are buttons for "X-Y plot for a range of values" and "Calculate".

# 効果量について

- 統計的仮説検定では、「差がない」とか「関連がない」という帰無仮説の下で、解析対象のデータよりも「差がある」「関連がある」側に外れたデータが偶然得られる確率に相当する  $p$  値が「有意水準」より低いとき、偶然でないとなしして帰無仮説を棄却する
  - サンプルサイズを大きくすれば、「偶然でない」差や関連を検出できる可能性(～検出力)は上がるが、差や関連が「ゼロではない」ことの確信度があがるだけ
  - $p$  値からでは、差や関連の大きさは不明
- しかし、研究者が本当に知りたいのは、処置やグループの違いが結果にどれくらい大きな差をもたらすのか、あるいは二つの変数の関連はどれくらい大きいのか、という点。その指標が**効果量**
- 最近では信頼区間に加えて効果量も示すべきとする Journal が多い
  - G\*Power では効果量も計算できるし、必要な効果量からサンプルサイズを設計することもできる
- 参考文献としては、大久保街亜、岡田謙介『伝えるための心理統計：効果量・信頼区間・検定力』勁草書房、ISBN978-4-326-25072-1 がお薦め

## 2 系統の効果量

- 効果量とは、帰無仮説が正しくない程度を量的に表す指標
- データに適用する統計モデルにより 2 系統に分かれる
  - 独立変数と従属変数の区別を含む場合 (独立 2 群間の平均値の差の検定も、群分け変数を独立変数、量の変数を従属変数と考える線型モデルで表せるので、この枠組みに含まれる)
    - 独立変数が従属変数に対して影響を及ぼす程度が効果量
    - d 族の効果量 (Cohen の d や Hedges の g など)
      - \* R では `effsize` パッケージの `cohen.d()` 関数で計算できる
  - 独立変数と従属変数の区別がない場合
    - 変数間の関連の大きさを表すのが効果量
    - r 族の効果量 (ピアソンの相関係数, 決定係数, イータ 2 乗, オメガ 2 乗など)
      - \* R では `lsr` パッケージの `etaSquared()` 関数でイータ 2 乗を得る
- 一般に効果量は標準化された値なので元データの単位によらない
- 帰無仮説が完全に正しいとき効果量は 0

# 多重検定 (1)

- 「保健学共通特講 IV」テキスト第9章を参照
- 多重比較の方法は多数あるが、有名なものは、ボンフェローニ (Bonferroni) の方法, ホルム (Holm) の方法, シェフェ (Scheffé) の方法, テューキー (Tukey) の HSD, ダネット (Dunnett) の方法, ウィリアムズ (Williams) の方法である。最近では, FDR(False Discovery Rate) 法も良く使われるようになった。
- ボンフェローニの方法とシェフェの方法は検出力が悪いので, 特別な場合を除いては使わない方が良い
- 平均値の差の多重比較ならテューキーの HSD ( R では TukeyHSD() 関数を使う), その他の場合はホルムの方法か FDR 法が薦められる
- R の pairwise.t.test() や pairwise.prop.test() や fmsb パッケージの pairwise.fisher.test() などでは, p.adjust.method="fdr" のようなオプションとして検定の多重性の調整方法を指定する
- 等分散を仮定しない方法として Games-Howell 法がある。R では rosetta パッケージの posthocTGH() 関数で可能である。
  - 組み込みデータ chickwts で, 餌の違いによる鶏の体重を多重比較するには, rosetta パッケージを呼び出した後,  
posthocTGH(chickwts\$weight, chickwts\$feed, method="games-howell")

# 多重検定 (2) EZR での実行例

- Rの組み込みデータ（6週間異なる餌で飼育した後のヒヨコ体重 chickwts : 変数 feed が餌, 変数 weight が体重）を分析
- 一元配置分散分析で feed の weight への効果は有意
- EZR ではファイル>パッケージに含まれるデータを読み込む> datasets の chickwts を選択し, 統計解析>連続変数の解析>3群以上の間の平均値の比較と選び, 右上画面でオプション選択
- Tukey の場合右下の図が表示される  
→中央の縦棒が「差がゼロ」の線。95% 同時信頼区間がそこに掛かっていない組合せが5%水準で「有意差あり」
- # RではTukeyは  

```
res <- TukeyHSD(aov(weight~feed, data=chickwts))
plot(res)
print(res) # 多重性調整済 p 値表示
他は (調整法は "" 内に指定)
pairwise.t.test(chickwts$weight,
chickwts$feed, p.adjust.method="fdr")
行列形式で2つずつの餌の組合せについて
検定の多重性調整済 p 値が表示される
```

R 3群以上の間の平均値の比較(一元配置分散分析one-way ANOVA)

モデル名を入力: AnovaModel.1  
複数の選択はCtrlキーを押しながらクリック。↓

目的変数(1つ選択) weight  
比較する群(1つ以上選択) feed, weight

グラフ 等分散と考えますか?  
 箱ひげ  はい(一元配置分散分析)  
 棒  いいえ(Welch検定)  
 折れ線

↑ Welch検定の場合は多重比較やモデルの保存は行われません。  
↓ 2組ずつの比較(post-hoc検定)は比較する群が1つの場合のみ実施される。

2組ずつの比較(Bonferroniの多重比較)  
 2組ずつの比較(Holmの多重比較)  
 2組ずつの比較(Tukeyの多重比較)  
 2組ずつの比較(Dunnettの多重比較)

↑ アルファベット順で一番若い群が対照群となる。変更したい場合は因子水準を再順序化する。  
 モデル解析用に解析結果をアクティブモデルとして残す

↓ 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

