

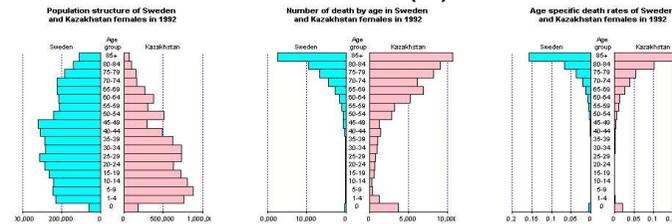
保健統計学特講 2019.5.21

- 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ますので、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

- 原因→結果という因果関係を考えているとき、原因とも結果とも関連があって、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
- 交絡要因の3条件 (例: 肥満⇒高血圧に対する年齢)
 - 注目している要因 (肥満) ではない
 - 注目している要因 (肥満) と結果である健康影響 (高血圧) の両方と因果的に関連
 - 要因曝露の結果ではない (肥満が高齢をもたらすのではない)

交絡の例 (1)



- 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)

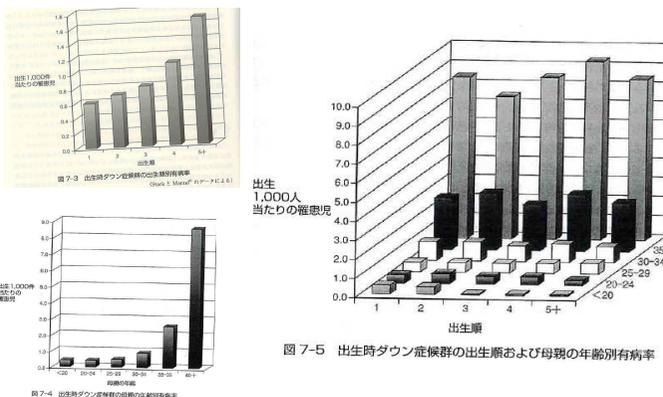
- 年齢が交絡因子となっているために、喫煙者の方が非喫煙者よりも死亡率が低くなることもある (出典: 『ロスマンの疫学』)

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

年齢 (歳)	生存状況	喫煙者	非喫煙者	総数
15-24	死亡	2	1	3
	生存	53	61	114
25-34	死亡リスク	0.04	0.02	0.03
	生存	121	152	273
35-44	死亡	3	5	8
	死亡リスク	0.02	0.03	0.03
45-54	死亡	14	7	21
	死亡リスク	0.13	0.06	0.09
55-64	死亡	27	12	39
	死亡リスク	0.21	0.15	0.19
65-74	死亡	51	40	91
	死亡リスク	0.44	0.33	0.39
75+	死亡	29	103	130
	死亡リスク	0.81	0.78	0.79
総数	死亡	13	64	77
	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』)



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、特定の年齢や性別に限定して研究すると、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等

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
<http://minato.sip21c.org/advanced-statistics/ecopx.txt>

<http://minato.sip21c.org/humeco/ecopoint.html>

エコポイント・チェック (Eco-Point Check)

Last updated on January 9, 2008 (08:01:14:48) (リンク先変更)

環境問題に真剣に取り組む必要はありますが、21世紀の生活の場において、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。環境問題に取り組むためには、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。環境問題に取り組むためには、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。

エコポイント・チェック (Eco-Point Check) は、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。環境問題に取り組むためには、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。

エコポイント・チェック (Eco-Point Check) は、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。環境問題に取り組むためには、環境問題に関心する方々の努力によって環境問題が解決されることは、必ずしも期待できません。また、意識だけでなく、行動にも取り組む必要があります。

項目	スコア
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=70-79
性別	0=M 1=F
FAMILYSIZE	自分を含む同居人数
Q01	新聞雑誌リサイクル
Q02	古紙100%トレイットペーパー使用
Q03	飲料容器トレーリサイクル
Q04	買い物袋持参
Q05	冷蔵庫より着る服で調節
Q06	食材を期限切れで捨てない
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)
Q08	風呂水を洗濯等に利用
Q09	車のアイドリングストップ (車を停めている人はいつても)
Q10	マイカーを避けて公共交通を利用
Q11	太陽熱温水器を利用
Q12	家電製品は省エネ型以外は買わない
Q13	米のとぎ汁は流さず有効利用
Q14	油をふき取ってから洗い
Q15	塩ビ系プラスチックを買わない
Q16	洗剤は合成洗剤でなく石鹸利用
Q17	洗剤 (石鹸を含む) をはかって適量使用
Q18	除草剤や殺虫剤を使わない
Q19	車のバッテリーや電池を適正処理
Q20	トイレや風呂場の強力洗浄剤を利用しない
Q21	有機溶剤を利用しない
Q22	有機農産物を選ぶ
Q23	地域の農産物を選ぶ
Q24	朝食早起
Q25	煙草を吸わない

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=70-79
SEX	性別	0=M 1=F
FAMILYSIZE	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%トレイットペーパー使用	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	冷蔵庫より着る服で調節	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 (車を停めている人はいつても)	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	塩ビ系プラスチックを買わない	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鹸利用	0=いつも 1=大体 2=時々 3=たまに 4=皆無
Q17	洗剤 (石鹸を含む) 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	地域の農産物を選ぶ	0=いつも 1=大体 2=時々 3=たまに 4=皆無
Q24	朝食早起	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 (公共交通便利がたまに / 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。 $\beta_3 \sim \beta_3$ 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("http://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 R^2 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", Nagelkerke2(res)$R2))
cat("\end{document}\n")
sink()
出力結果
(Intercept) -0.5479 0.3232 -1.695 0.09006
Q10c 1.3036 0.2886 4.517 6.28e-06 ← 6.28x10^-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
(Intercept) Q10c AGE SEX
0.5781719 3.6826760 0.8299727 2.1308305
2.5 % 97.5 %
(Intercept) 0.3032334 1.081392
Q10c 2.1093893 6.560746
AGE 0.6667677 1.0309811
SEX 1.2550376 3.700273
$R2 [1] 0.1333421 ← モデルの説明力は 13% 程度
  
```

	Dependent variable:
	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
Alkalai Inf. Crit.	391.206

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

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)
  
```

症例数とパワー計算 (1)

- 「保健学共通特講 IV」テキスト第 4 章を参照 <http://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, α エラー, β 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, α エラーがわかれば β 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
- 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いで良い

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、偶然でないと思われ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
- 参考文献としては、大久保恒亜, 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 がオススメ

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、有名なのは、ボンフェローニ (Bonferroni) の方法, ホルム (Holm) の方法, シェフェ (Scheffe) 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よい
- 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 `userfriendlyscience` 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、`userfriendlyscience` 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択、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 → 中央の縦線が「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))` `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