

付録. 日々もよくよくみれば仮説検定 (2)

- 他にも例えば
 - 冬だったらほとんど桜が咲かないはただけ
桜が咲いたからもう春が来たね (＝冬は終わったね)
 - 眼が良かったらほとんど眼鏡をかけていないはただけ
眼鏡をかけているから眼が悪いね
 - 正常だったらほとんどCK-MBが高値にならないはただけ
CK-MBが高値だから心筋梗塞だね
 - 地毛だったらほとんどズシないはただけ
ズシしていることがあるからツラだね (＝地毛じゃないね)

69

片側p値と両側p値



- 「等しいかより極端な結果」をどう決めるか
- 両側検定：どちらの方向に偏っても問題としたい (興味がある)
 - 帰無仮説「表と裏の出やすさが等しい」
 - 対立仮説「表と裏の出やすさが等しくない」
- 片側検定：一方の方向に偏った場合に問題としたい
 - 帰無仮説「表と裏の出やすさが等しい」
 - 対立仮説「表が出やすい」 (裏が出やすくても興味ない)
或いは「裏が～」 (表が～) のどちらか
- 片側か両側かは事前に (できれば研究開始前に) 決めるべき
- 中田くんにとってはどちらも問題だから**両側検定**



両側p値を求める状況 70

生物統計学2
 a) 交絡とその調整
 b) 因果推論とランダム化

ICR2006 第4回 (2006.8.23)
 国立がんセンター情報研究部
 吉村 健一


前回 (ICR2006 第3回) の復習

先々週の放送

- 視聴者4人に毎日眠る前、ミニトマト3つを摂取してもらう



- 一週間後、4人中3人も血液がさらさらに！
- 各物回会者にご機嫌かつ得意げにミニトマトをオススメ
 - 夕方、スーパーでミニトマトが飛びように売れる



前回のまとめ  さらさら制ら

3/4 = 75% に対して効果

- 4人中3人ぐらいは偶然でも起こるかもしれない αエラーの問題
- 4人の結果では少なすぎで信用できないかもしれない βエラーの問題
サンプルサイズ、検出力、βエラーの問題

αエラーとβエラーを考慮してプロトコルを事前作成し、実験後に仮説検定を行えば良い

前回を踏まえて

- 「ミニトマトを食べると血液さらさらになるか」を調べたい
- 仮説検定をしよう！
 - 帰無仮説「血液さらさら割合10%以下」
 - 10%以下だったら、、、 食べる意味がない（効果なし）
 - 対立仮説「血液さらさら割合50%以上」
 - 半分もさらさらになるなら  食べる意味がある！（効果あり）
- 場面に応じた臨界的最小有意差

5

α エラーと検出力、サンプルサイズ

- $\alpha = 5\%$ として検定
 - さらさら割合が10%以下であるのに「食べる意味がある」と誤って判断してしまう確率を5%に設定
- さらさら割合50%に対する検出力90%
 - さらさら割合が50%である場合に「食べる意味がある」と正しく判断できる確率を90%に設定
- これより、必要なサンプルサイズ **10人**
 - 仮説検定を行う前提で統計的に導出

6

ここまで、前回の復習




7

他にも「かもしれない」

3/4 = 75% に対して効果

- 4人中3人ぐらいは偶然でも起こるかもしれない✓
 α エラーの問題
- 4人の結果では少なすぎて信用できないかもしれない✓
サンプルサイズ、検出力、 β エラーの問題
- ミニトマトを食べなくても、さらさらになったかもしれない
- 他にも（血液さらさら効果があると云われている）
黒豆とか、納豆とか、青汁とか、、、摂取したかもしれない

 を食べなかつた人と比較しよう！

8

今週の放送「比較をしよう！」

- 視聴者50人に毎日眠る前、ミニトマト3つ摂取してもらって一週間後に血液さらさらか調べる
- それとは別に、毎日眠る前にミニトマトを摂取しなかった視聴者50人を募集して、一週間後に血液さらさらか調べる

以下のように仮説を変更して比較

帰無仮説「さらさら割合の差無し（差が0%）」
 対立仮説「さらさら割合の差30%」
 $\alpha=5\%$ 、検出力約85%で必要サンプルサイズ計100人

9

今週の放送

■ **実験デザイン**



摂取した50人



摂取しなかった50人

コントロール (対照)

□ 一週間後、2つのグループの結果
 (血液がさらさらになった視聴者の割合) を比較

10

実験の結果が得られた！

	さらさら	効果なし	
食べた	35人	15人	50人
食べず	25人	25人	50人

11

割合の差 (リスク差) は？

	さらさら	効果なし	
食べた	35人	15人	50人
食べず	25人	25人	50人

$70\% - 50\% = 20\%$ の差

を食べたグループで20%さらさら割合が高い

12

前回と同様に p 値 を求める！

- 帰無仮説「さらさら割合が等しい」が正しい場合
結果が偶然でどの程度ばらつくか

p値 = 4% < α = 5%

Fisherの正確な検定
に對照

0% (割合が等しい)

実験結果は **20%の差**

13

この結果を見て

- ミニトマトの効果はある！ 統計的有意差！
 - 食べなかったときと食べたときのさらさら割合の差は**20%**もある ($p=4% < \alpha$)
- 名物会者は「リコピン・パワー！でさらさら」と得意げに生物学的機構を説明しながら再びオススメ
 - 夕方のスーパーマーケットでは、再び完売
- みなさんはこの結果から を買いますか？

今日から買って食べます or これだけやっても信じません

14

ミニトマトを食べなくても、さらさらになったかもしれない

「食べた場合」と「食べなかった場合」との比較を行う実験
(コントロール)

15

交絡と交絡因子

16

ところで

- 違う曜日のライバルTV健康番組が
 - ブロッコリの太陽がんに対する予防効果をうたっていた
- もしかして血液さらさら効果もあるかも（**新しい仮説**）
 - 生物学的機構：イソチオシアネートの効果？
 - 先ほどのミニトマト実験の100人について
 - を食べていたかどうかとも聞いてみよう

17

ブロッコリの効果

- ブロッコリを食べた人と食べなかった人とで比べよう

	さらさら	効果なし
食べた	60	0
食べず	0	40

割合の差は $100 - 0 = 100\%$ 、**めっちゃめっちゃ効果あり**

18

ブロッコリも効果ありということ？

- ブロッコリのさらさら効果は**絶大**
- そうすると、どういう解釈ができるか・・・
 -  と  共にさらさら効果あり
 -  のみにさらさら効果あり
 -  のみにさらさら効果あり
- 個別の結果だけではどれが正しいか判断できない！
 - 本当はどちらに効果があるのかわからない

19

お互いに影響を取り除いてみる

	食べた	食べた
	食べない	食べた
	食べた	食べない
	食べない	食べない

の効果をみる

20

お互いに影響を取り除いてみる

トマト	食べた	食べた
ブロッコリー	食べた	食べた

を食べた者同士で比較

食べた	食べた
食べない	食べない

を食べない者同士で比較

21

お互いに影響を取り除いてみる

トマト	食べた	食べた
ブロッコリー	食べた	食べた

を食べた者同士で比較

食べた	食べた
食べない	食べない

を食べない者同士で比較

22

お互いに影響を取り除いてみる

トマト	食べた	食べた
ブロッコリー	食べた	食べた

23

お互いに影響を取り除いてみる

トマト	食べた	食べた
ブロッコリー	食べた	食べた

を食べた者同士で比較

食べた	食べた
食べない	食べない

24

お互いに影響を取り除いてみる

25

サブグループ解析 (層別解析) の考え方

26

サブグループごとに比較すれば互いの影響を取り除くこと (調整すること) ができる

ミニトマトで分けて の効果をみると

	さらさら	効果なし	
食べた	35	0	35
食べず	0	15	15

	さらさら	効果なし	
食べた	25	0	25
食べず	0	25	25

を食べた人でも食べなかった人でも のさらさら効果は絶大!

27

ブロッコリーで分けて の効果をみると

	さらさら	効果なし	
食べた	35	0	35
食べず	25	0	25

	さらさら	効果なし	
食べた	0	15	15
食べず	0	25	25

を食べた人でも食べなかった人でも のさらさら効果はゼロ!

28


あれっ??!

- の影響を取り除くと の効果はなし?
- の効果はウソだったのか・・・?
- この結果から、どういう解釈が適切か
 - ① には効果が有ることには変わらない
 - ② と一緒に食べた場合に は効果あり
 - ③ ではなく に効果あり

29

当然 に効果無しと解釈するのが適切

		食べた人のみ		食べなかった人のみ	
		さらさら	効果なし	さらさら	効果なし
食べた	さらさら	35	0	0	15
	食べた	25	0	0	25


 を食べた人でも食べなかった人でも
 のさらさら効果は完全にゼロ!

30

どうしてこのようなことになったのか

人数	35人		V	25人	
	食べた	食べない		食べた	食べない
人数	食べた	食べない	食べた	食べない	
	食べた	食べない	食べた	食べない	

- 人数をよーく見ると
 - を食べた人に を食べた人が多い
 - サラダと一緒に摂取? の影響により見かけ上 の効果があるようにみえてしまった

31

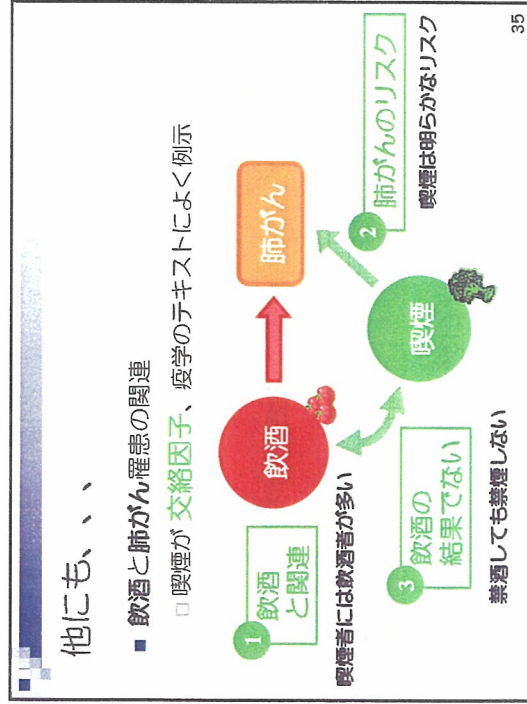
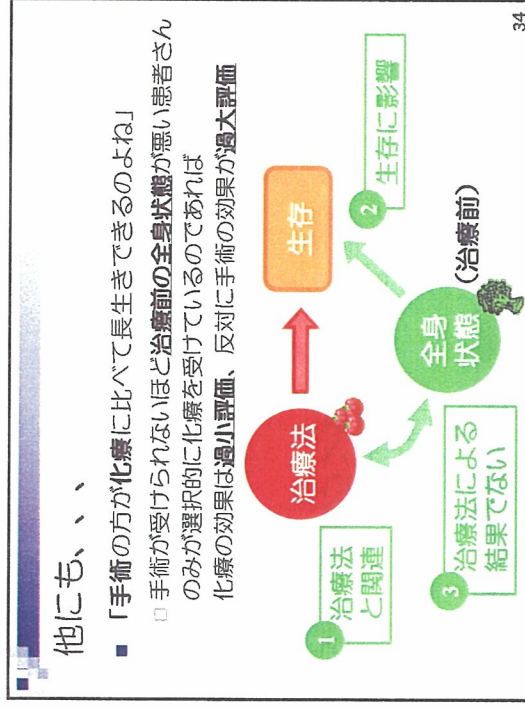
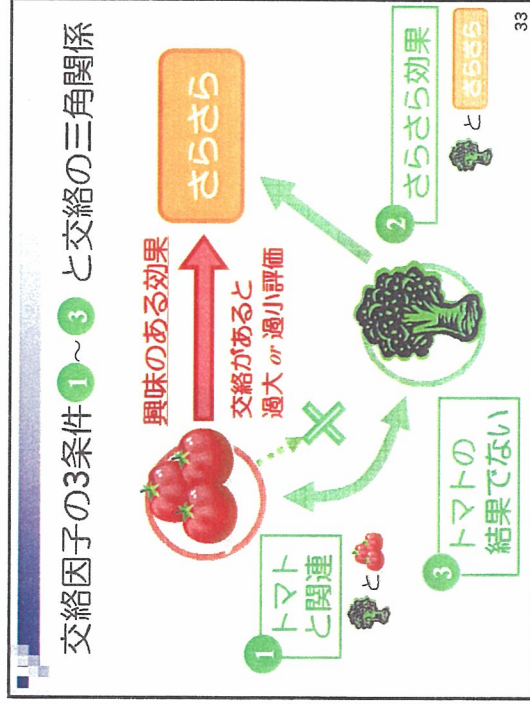
この現象が交絡 confounding


- ラテン語で“一緒に混ざる” confundere が語源
- 結果に影響を与える他の要因との関連により、ある要因の結果への影響が見かけ上歪められてしまうこと



今回の場合

- 結果に影響を与える の関連により、の結果への影響が見かけ上歪められてしまうこと
 - を交絡因子という

32




交絡因子  二調整しなければならぬ変数



- 1 興味のある曝露因子 () と関連する
- 2 結果 (さらさら) に影響する
- 3 興味のある曝露因子 () の結果でない


- 3条件全てを満たす場合に交絡因子、交絡を生む
- 交絡があると、結果を適切に解釈できない
 - 過大評価 (効果がないのに効果あり) にも、過小評価 (あるのに なし) にもつながる
 - 適切に調整しないと解釈できない



37



適切に交絡調整して初めて解釈できる

 食べた人のみ

	さらさら	効果なし	
	35	0	35
	25	0	25

 食べなかった人のみ

	さらさら	効果なし	
	0	15	15
	0	25	25

 を食べた人でも食べなかった人でも  のさらさら効果は完全にゼロ！

38

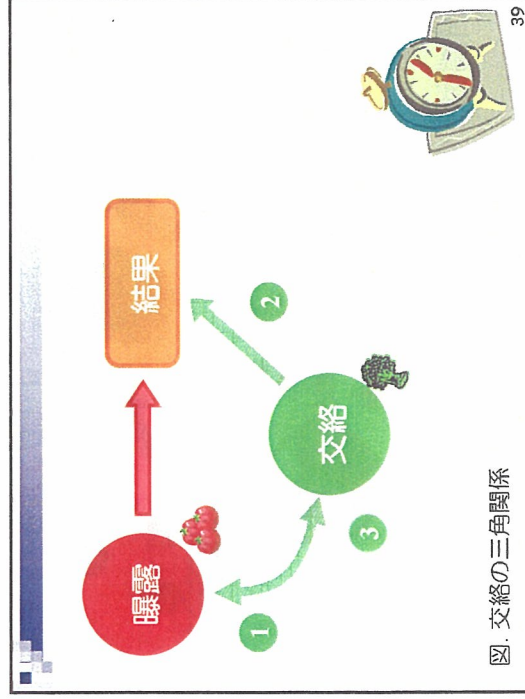





図. 交絡の三角関係

因果推論とランダム化

40


残念ながらさらさら効果なし

-  は  を調整すると効果なし
-  は効果がないようだ・・・
- ところで、、、そもそも「効果あり」とか「効果なし」とはどういうことか・・・
 - みなさん、常識のように使っているかもしれませんが 人にうまく説明できますか？
 - 基本的な部分より整理しながら、ゆっくり考えてみよう！


41


被験者の中田くんでごえてみよう！

実験開始日

- 激戦続きでお疲れの「中田くん」
- 就寝前に  を食べた

1週間後

- 血液がさらさらになった 
- はさらさら効果があった




でも、もしかしたら

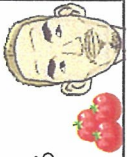
反事実 counterfactual を考える

実験開始日

- 激戦続きでお疲れの「中田くん」
- 就寝前に何も食べなかった (場合を考えてみて、、、)


1週間後

- 血液が (同じく) さらさらになった
- 食べなくてもさらさらになったのだから当然  はさらさらと関係ないことになる




つまり、1つの事実だけでは判断できない

実験開始日

- 激戦続きでお疲れの「中田くん」
- 就寝前に  を食べた


1週間後

- 血液がさらさらになった 
- はさらさら効果があった

三「た」論法

くすりを、

- ・ 飲んだら
- ・ 治った
- ・ だから効いた

はだめ！ 

つまり、効果の有無は

 を食べると	さらさら	どろどろ	効果あり
	さらさら	どろどろ	なし
さらさら	さらさら	どろどろ	効果あり
どろどろ	さらさら	どろどろ	なし

45

効果があることを言いたい場合には

 を食べると	さらさら	どろどろ	効果あり
	さらさら	どろどろ	なし
さらさら	さらさら	どろどろ	効果あり
どろどろ	さらさら	どろどろ	なし

46

2つの状況を組み合わせないと分からない

 を食べると	さらさら	どろどろ	効果あり
	さらさら	どろどろ	なし
さらさら	さらさら	どろどろ	効果あり
どろどろ	さらさら	どろどろ	なし

47

中田くん個人に関して

因果推論

-  と  の因果関係を調べたい (原因と結果の関係)
- 実験開始日に  を食べた中田くん (理想的なコントロール)
-  を食べなかった中田くん


という食べたか/食べなかったかだけ異なる  が必要

□ この両方の状況の結果を比較しないと、効果の有り無し (=因果関係) は分からない

48


小休止

- 「反事実と比較しないと因果関係は分からない」という因果推論の基礎について





49

でも

- 現実の世界では、どちらか一方は絶対に調べられない
- どうしたら良いか
 - 個人に興味があるが、調べられないので、**仕方なくグループについて**
 - のさらさら効果を調べられるか考えてみよう
 - 特徴の似たコントロールが必要
 - でも、中田くんのような人はそういさそうなので、 の仲間内、サッカー日本代表メンバーの中で比べる (内部比較をする) 50

どのように調べるか

- サッカー日本代表を2つに分けて互いに似たような特徴をもつグループを作る
 -  を食べるグループ
 -  を食べないグループ (コントロール)
- 一週間後、グループ間でさらさら割合 (エンドポイント) を比較する
- さらに、交絡があったら**結果が適切に解釈できなくなる**ので、交絡因子を考慮したい
 - 交絡因子としてプロックリと年齢 (さらさらに影響しそう) 51

プロックリと年齢 (交絡因子の候補) を調べる



52

プロックリと年齢が等しい同士で比べる

大黒くん 1980生

宮本くん 1977生

稲本くん 1979生

小野くん 1979生

中田くん 1977生

玉田くん 1980生

マッチング
同じ特徴の人
同士で比べる

53

プロックリと年齢が等しい同士で比べる

大黒くん 1980生

宮本くん 1977生

稲本くん 1979生

小野くん 1979生

中田くん 1977生

玉田くん 1980生

マッチング
同じ特徴の人
同士で比べる

54

でも、プロックリと年齢だけでは

- プロックリと年齢だけで、ヒトの特徴はとも表せない
単純でない
他にも交絡因子の候補として
 - 体格 (身長、体重、)
 - 運動能力
 - 居住地域
 - 血圧
 - もっと前に食べたプロックリ
 - ...
- 特徴がたくさんありすぎて、
本当は全部マッチしたいができない

55

そこで、

- ランダム化 randomization
 - ランダム化によってグループ分け (割り付け) を決める
 - ミニトマトを食べるかコントロールとなるかを
(ここではプロックリや年齢は全く考慮せずに)
確率に基づいてランダムに決める

ミニトマト・グループ

コントロール・グループ

56

ランダム化によって

- 食べない場合の結果（さらさら割合）が平均的に等しくなるグループをつくれる
 - つまり、個々の特徴に関しても、平均的に似た特徴をもつグループ同士となることを期待できる

平均年齢も平均身長も平均体重もプロックコリ採取割合も、、、何でもグループ間で、人数が増えるにつれて似てきそう

57

ランダム化によって交絡を断ち切る

と交絡因子の関連がなくなる（平均的に似る）

興味のある効果

さらさら

と交絡

トマトの結果でない!

58

ランダム だから既知でない交絡も断ち切る

- 実験前には交絡因子がどうか分からなかったもの、納豆でも、黒豆でも、青汁でも、それ以外でも、グループ間で平均的に似るので平均的に交絡が無くなる

さらさら

交絡因子

既知のものでも 未知のものでも

59

最強のマッチングとも比較しよう！

- マッチングで因果推論できるか？
- 臨床研究で最強のマッチングは、、、
 - 世界中を探して自分と特徴の似た人を見つけ出す？
 - 究極は・・・双子研究（一卵性双生児のみで臨床研究）
 - 双子がいなくても、未来にはクローン技術が発展し、
 - 遺伝子が等しい、もうひとりの自分をつくって比べれば効果の有無が判断できる？
- わざわざランダム化までして他人と比べなくても双子やクローンと比較する方が良いのかもしれない！

60

ところが、クローンでも一卵性双生児でも

- 等しくできるのは、遺伝子という**既知の交絡因子のみ**
- 遺伝子以外の環境要因や未知の要因は等しくできない
 - 運動習慣とか食習慣とか、際限なく一緒にできない
 - まだよく分かっていない未知の交絡因子に関しては（分からないのだから当然）一緒にできない
- やっぱり、ランダム化は最良の手段！
 - マッチングは既知のみ、ランダム化は未知も
 - 可能であればランダム化が好ましい

61



因果推論をする際の交絡への対抗手段

- 可能であれば、最良の手段であるランダム化
 - 交絡が明らかな場合、マッチングしてからランダム化すること（層別ランダム化）もある
- 不可能であれば、交絡因子となりうるもの全てを測定しかつ適切に調整する
 - 層別解析（サブグループ解析）、モデル（ロジスティック回帰、Cox回帰、OΔ回帰、、、）による調整
 - 交絡因子全てを測定できているか
 - 本当に適切に調整できているか は神のみぞ知るところ
 - 非ランダム化研究の統計解析は統計家でも難しい解析
 - ランダム化研究よりエビデンスとしての質は**劣る**

62

 **スタディのプロトコールを書こう！**

因果推論とランダム化を踏まえて

- 帰無仮説「 にさらさら効果がない」
- 対立仮説「 にさらさら効果がある」
 - 臨床的に意味のある最小の差（臨床的最小有意差）は30%
- 視聴者100人をランダム化によって
 - 毎日眠る前にミニマト3つ摂取するグループ（50人）
 - ミニマトを摂取しないコントロールグループ（50人）
 のどちらかに割り付ける
- 一週間後、さらさらかどうかを調べる
- グループ間でさらさら割合を比較する
 - $\alpha=5\%$ で検定、30%の差に対する検出力約**85%**

63

Take Home Messages（まとめ）

- 交絡があると結果が適切に解釈できない
- 交絡因子となる三条件
 - ①興味のある曝露因子と関連、②結果に影響
 - ③興味のある曝露因子の結果でない
- 反事実を考えないと因果は推論できない
- ランダム化は因果を推論する上で最良の方法
 - 既知・未知問わず、交絡因子との関連を断ち切ることが可能
 - 非ランダム化の場合、交絡が生じていないことは誰も保証できない、これより適切に結果を解釈できているかどうかも不確か
 - 結果の解釈には注意を要する

64

健康関連TVショーの正しい視聴法

- **ICR-Biostatistics版子エックリス上**
 - 実験の仮説（帰無仮説・対立仮説）は明確か
 - α エラーを考慮しているか
 - P値を使って仮説検定しているか
 - 検出力（または β エラー）を考慮して必要なサンプルサイズを事前に定めているか
 - ランダム化がなされているか
 - 非ランダム化の場合、全ての交絡因子を測定し適切に考慮されているか
 - 事前にプロトコールで実験方法を全て規定しているか

65

補足. 区間推定（信頼区間）とは

66

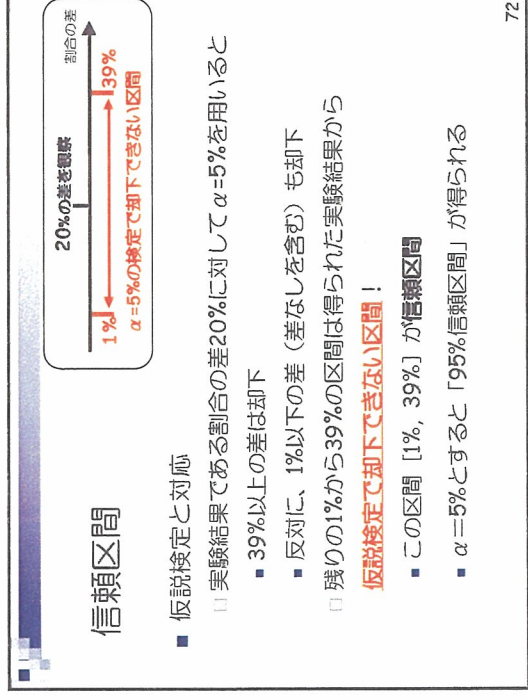
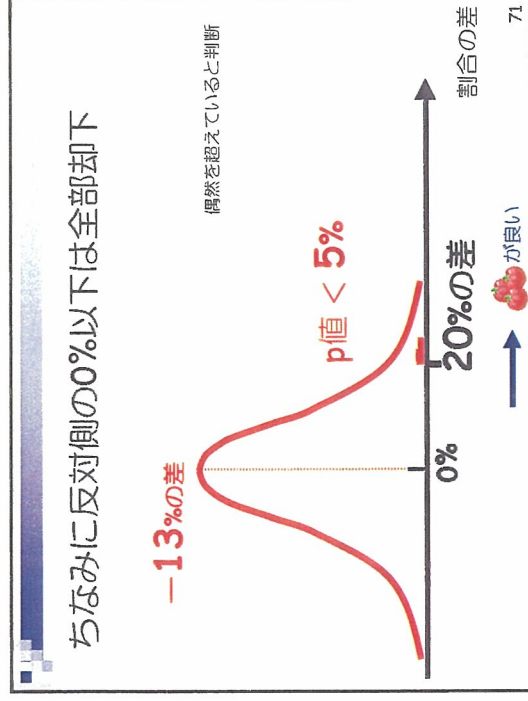
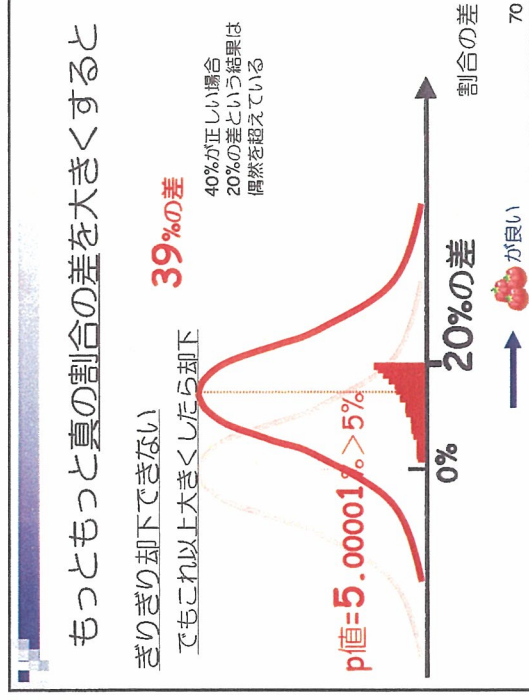
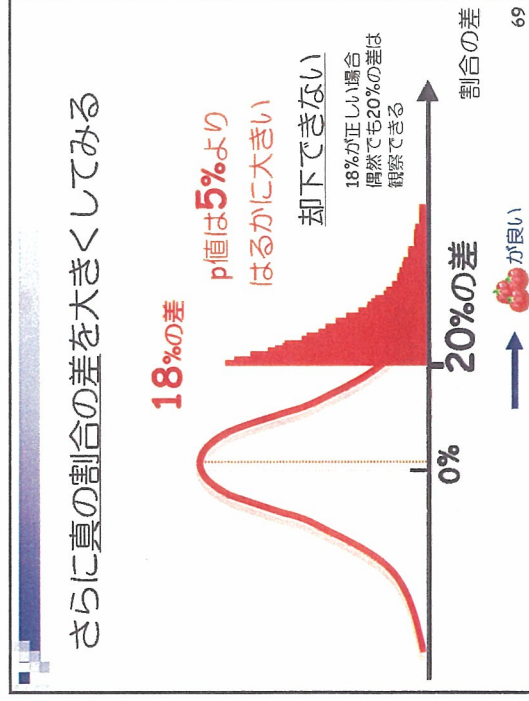
帰無仮説が真の場合の観察される割合の分布

見やすくなるので曲線におきかえます

67

却下できなくなる（p値>5%となる）まで、真の割合の差を大きくしてみる

68



信頼区間の正しい解釈

- 「信頼区間の真ん中の値ほど真の値に近く、反対に値が区間の端になればなるほど真の値から離れている」と解釈されてしまいがちであるが **誤り**
- 信頼区間の内側は、どの値も平等・等価値
 - 今回の結果からは、1%でも10%でも20%でも39%でも真の値としての確からしさは全て同じ
 - [1%, 39%]の内側はどの値も、検定で却下できないという点で平等
 - 却下できない区間で信頼区間を構成する方法を用いれば結果的に、100回に95回は真の値を含むことができる

73

付録

参考文献

<コントロール・ランダム化・因果推論について>

- 佐藤俊哉, "Pコントロール". 橋広計 他(編), "これからの臨床試験 医薬品の科学的評価—原理と方法", 第2章, pp.21-34, 朝倉書店, 1999.

74

付録

比較可能性について

- 理想的なコントロールは自分自身
= 同じ予後をもつコントロール・グループ
 - 「このコントロールは比較可能性がある」という
- 比較可能性のあるコントロール・グループ
 - 自分自身コントロールは事実上反し不可能
 - ランダム化のみによって無条件で達成可能
- 比較可能性がない三交絡がある と定義すると
 - 交絡因子とは比較可能性をくずす因子
 - 交絡因子の3条件は比較可能性をなくすための必要条件として導くことができる
 - 両群で分布が異なる
 - 予後に影響をする
 - 因果推論の一般論から中間変数は調整すべきでない

75

付録

きやーるが鳴くんで雨ずらよ!

- カエルが鳴くから、雨が降るでしょう
 - カエルが鳴き声をあげると、昔の人は田植え、稲刈りの予定変更
- この表現は正しいか?
 - 数年前の統計関係のメーリングリスト*で心理統計の専門家によってまじめに議論
 - 「カエル → 雨」は因果関係と違ってよいか否か

* 心理学メーリングリスト ipr より