

治験・倫理審査委員会委員研修 (2023/12/8)

医学統計の基本

93

PAGES

三橋利晴@新医療研究開発センター

mitsuh-t@cc.okayama-u.ac.jp

sankyoh@gmail.com

1. 研究において統計解析は下流工程
2. 統計解析手法よりも**上流工程**が重要
3. 研究目的・研究デザインにより解析方法は決まる。

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

実例

要約する

見やすくする

因果推論

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

実例

要約する

見やすくする

因果推論

愛媛では、各家庭にポンジューズ
が出る蛇口があるらしい。



愛媛県庁HPより引用

ジュースを作らないと。



株式会社えひめ飲料HPより引用

みかんも作らないと。



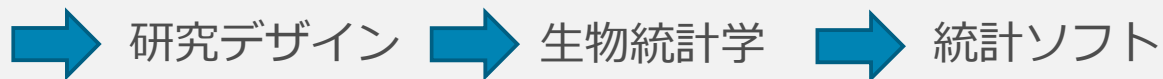
研究に当てはめると。

研究デザイン → 生物統計学 → 統計ソフト



研究に当てはめると。

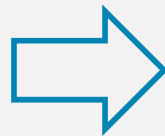
研究仮説
研究目的



適切な統計解析のためには

研究仮説・研究目的の明確化が重要

例えば



PICO/PECOの活用

PICO/PECOを使うと...

曖昧な臨床上や公衆衛生学上の疑問を

研究するのに適した形にするために、

シンプルに定式化・構造化することができる。

- **P=Patient, Population**
誰を対象者とするのか
- **I=Intervention, E=Exposure**
どのような介入・曝露を考えるのか
- **C=Control, Comparison**
何と比較するのか
- **O=Outcome**
主要なアウトカム

統計について考える前に...

研究の目的を把握することが重要

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

実例

要約する

見やすくする

因果推論

なぜか…？



誤った手法で誤った結論にしてしまうことを避けたい。

審査において誤った手法をきちんと指摘したい。

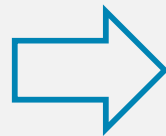
All models are wrong,
but some are useful.

George E. Box

正しい解析手法があるわけではない。

研究仮説・研究目的を満足できれば良い。

ただし、致命的なエラーは避けたい。



ご相談下さい。

- 1 要約する
- 2 見やすくする
- 3 分類する
- 4 比較する
- 5 予測する
- 6 因果推論

1 要約する

2 見やすくする

3 分類する

4 比較する

5 予測する

6 因果推論

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

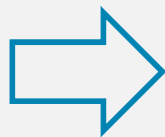
実例

要約する

見やすくする

因果推論

研究対象となっている症例などについて、
様子を提示するために、データを要約する。



記述統計量の算出

平均値・標準偏差

中央値・四分位範囲

度数・割合

など

Table 1 Baseline characteristics of the sample

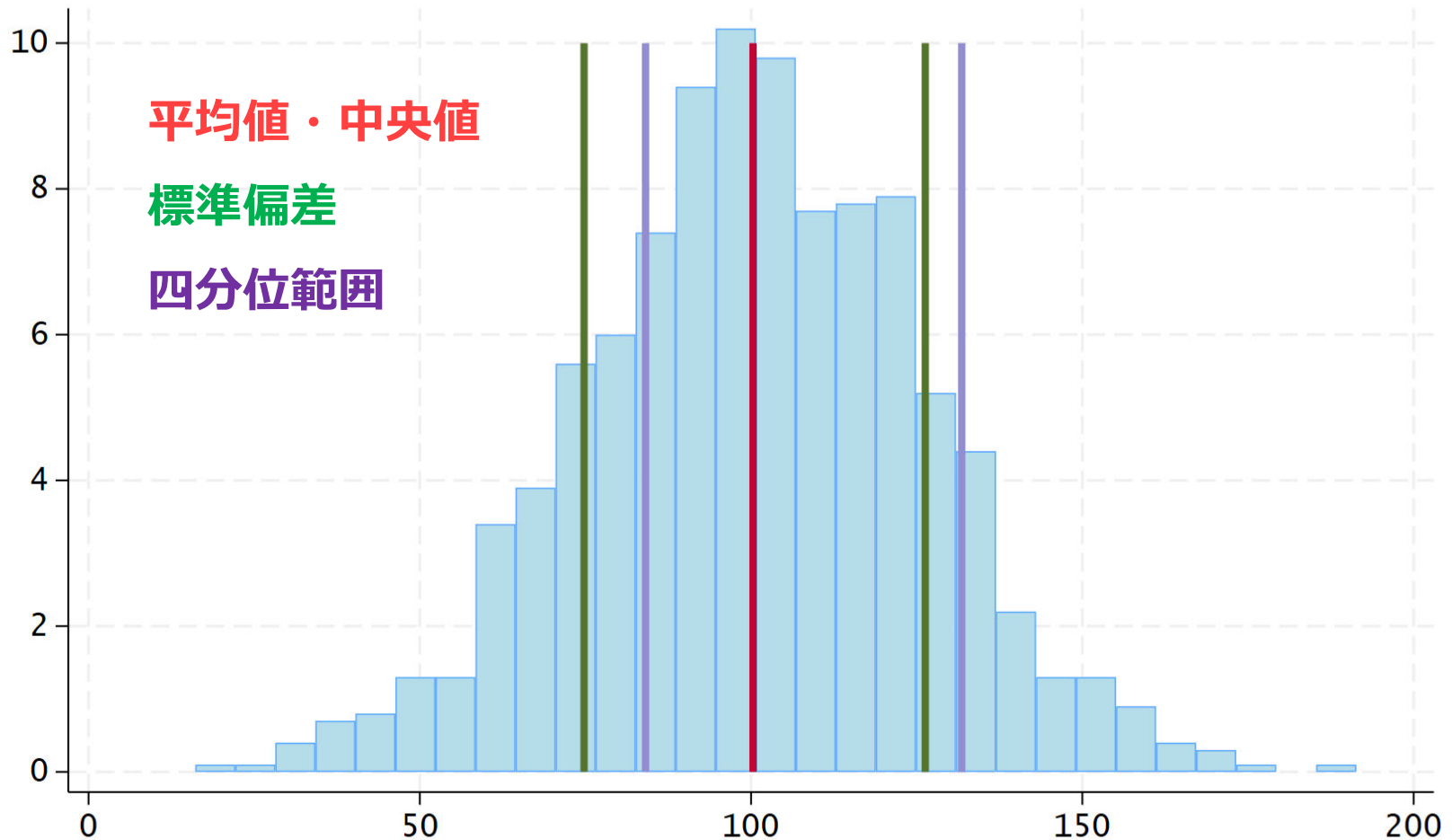
Variable	Entire sample (n = 301)	Intervention group (n = 148)	Control group (n = 153)
Female (n (%))	150 (49.8)	74 (50.0)	76 (49.7)
Age in years (mean (SD))	40.2 (10.1)	40.2 (9.9)	40.2 (10.2)
Educational attainment of university or more (n (%))	178 (59.1)	81 (54.7)	97 (63.4)
Parental status (n (%))	133 (44.2)	67 (45.3)	66 (43.1)
Household income/year in JPY millions (n (%))			
Less than 2	19 (6.3)	10 (6.8)	9 (5.9)
2 or more and less than 4	56 (18.6)	30 (20.3)	26 (17.0)
4 or more and less than 6	78 (25.9)	35 (23.6)	43 (28.1)
6 or more and less than 8	44 (14.6)	25 (16.9)	19 (12.4)
8 or more and less than 10	22 (7.3)	12 (8.1)	10 (6.5)
10 or more and less than 12	18 (6.0)	7 (4.7)	11 (7.2)
12 or more and less than 15	3 (1.0)	1 (0.7)	2 (1.3)
15 or more and less than 20	7 (2.3)	4 (2.7)	3 (2.0)
20 or more	3 (1.0)	1 (0.7)	2 (1.3)
Unknown	16 (5.3)	9 (6.1)	7 (4.6)
Missing	35 (11.6)	14 (9.5)	21 (13.7)
Marital status (n (%))			
Married	161 (53.5)	82 (55.4)	79 (51.6)
Never married	120 (39.9)	55 (37.2)	65 (42.5)
Divorced/widowed	20 (6.6)	11 (7.4)	9 (5.9)
Employment status (n (%))			
Full-time	163 (54.2)	73 (49.3)	90 (58.8)
Part-time	46 (15.3)	29 (19.6)	17 (11.1)
Self-employed	24 (8.0)	15 (10.1)	9 (5.9)
Other	4 (1.3)	2 (1.4)	2 (1.3)
None	64 (21.3)	29 (19.6)	35 (22.9)
Self-rated health (n (%))	250 (83.1)	118 (79.7)	132 (86.3)
Internet search engine use <once/day (n (%))	43 (14.3)	19 (12.8)	24 (15.7)

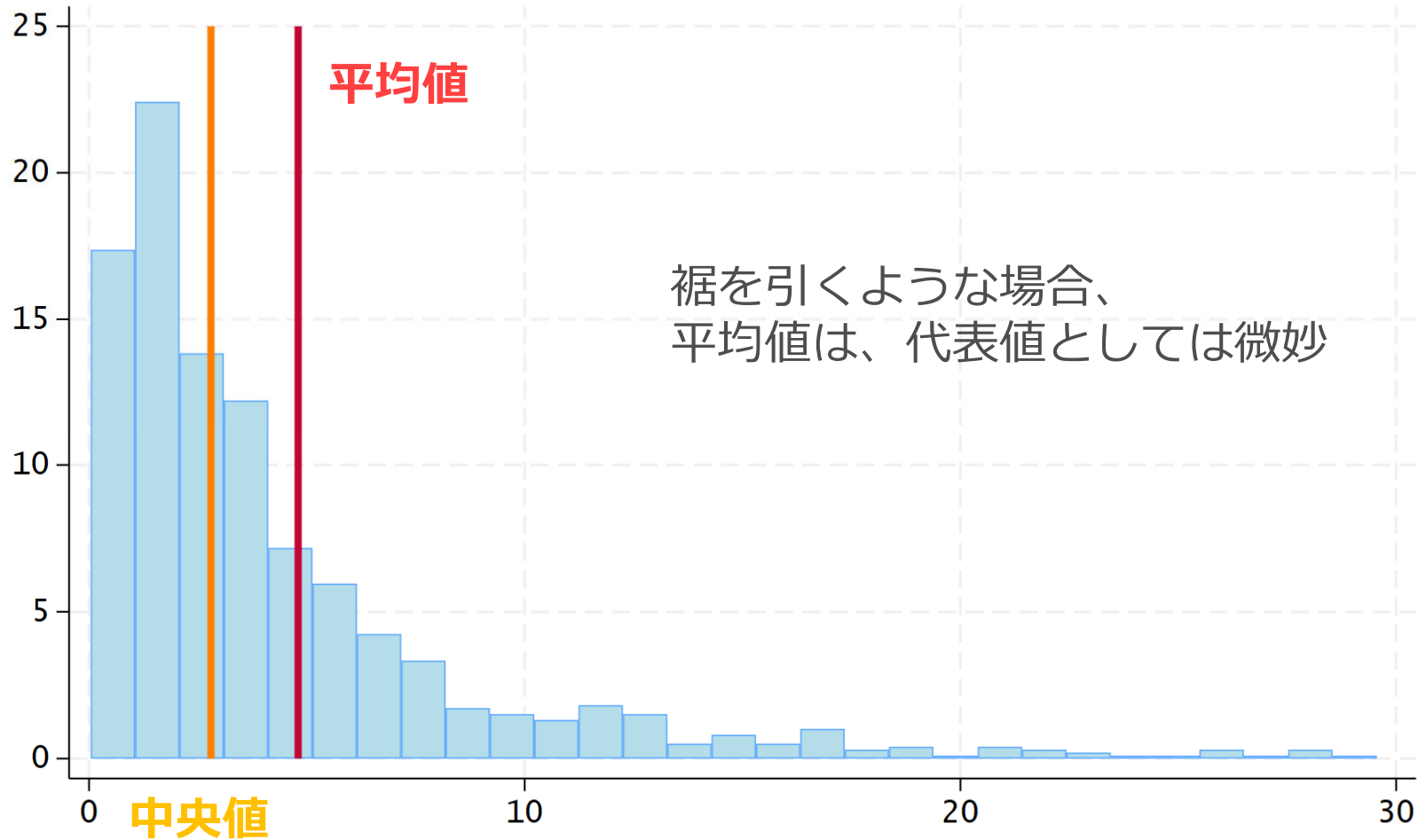
- データの代表値（中心的な値）の一種
- データ内の全ての数値を合計し、その合計を数値の個数で割ったもの。

- データの散らばり具合を示す。
- 平均値との差を用いて算出する。

- データの代表値（中心的な値）の一種
- データ内の数値を小さい順に並べた時に、真ん中に来る数値。

- データの**散らばり具合**を示す。
- データ内の数値を小さい順に並べた時に、25%のところと75%の所に来る数値の差。





記述統計量を示す時には、

検定しない。

CONSORTやSTROBEでも

注意されている。

Table 1 Baseline characteristics of the sample

Variable	Entire sample (n = 301)	Intervention group (n = 148)	Control group (n = 153)
Female (n (%))	150 (49.8)	74 (50.0)	76 (49.7)
Age in years (mean (SD))	40.2 (10.1)	40.2 (9.9)	40.2 (10.2)
Educational attainment of university or more (n (%))	178 (59.1)	81 (54.7)	97 (63.4)
Parental status (n (%))	133 (44.2)	67 (45.3)	66 (43.1)
Household income/year in JPY millions (n (%))			
Less than 2	19 (6.3)	10 (6.8)	9 (5.9)
2 or more and less than 4	56 (18.6)	30 (20.3)	26 (17.0)
4 or more and less than 6	78 (25.9)	35 (23.6)	43 (28.1)
6 or more and less than 8	44 (14.6)	25 (16.9)	19 (12.4)
8 or more and less than 10	22 (7.3)	12 (8.1)	10 (6.5)
10 or more and less than 12	18 (6.0)	7 (4.7)	11 (7.2)
12 or more and less than 15	3 (1.0)	1 (0.7)	2 (1.3)
15 or more and less than 20	7 (2.3)	4 (2.7)	3 (2.0)
20 or more	3 (1.0)	1 (0.7)	2 (1.3)
Unknown	16 (5.3)	9 (6.1)	7 (4.6)
Missing	35 (11.6)	14 (9.5)	21 (13.7)
Marital status (n (%))			
Married	161 (53.5)	82 (55.4)	79 (51.6)
Never married	120 (39.9)	55 (37.2)	65 (42.5)
Divorced/widowed	20 (6.6)	11 (7.4)	9 (5.9)
Employment status (n (%))			
Full-time	163 (54.2)	73 (49.3)	90 (58.8)
Part-time	46 (15.3)	29 (19.6)	17 (11.1)
Self-employed	24 (8.0)	15 (10.1)	9 (5.9)
Other	4 (1.3)	2 (1.4)	2 (1.3)
None	64 (21.3)	29 (19.6)	35 (22.9)
Self-rated health (n (%))	250 (83.1)	118 (79.7)	132 (86.3)
Internet search engine use <once/day (n (%))	43 (14.3)	19 (12.8)	24 (15.7)

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

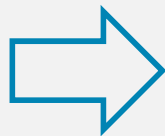
実例

要約する

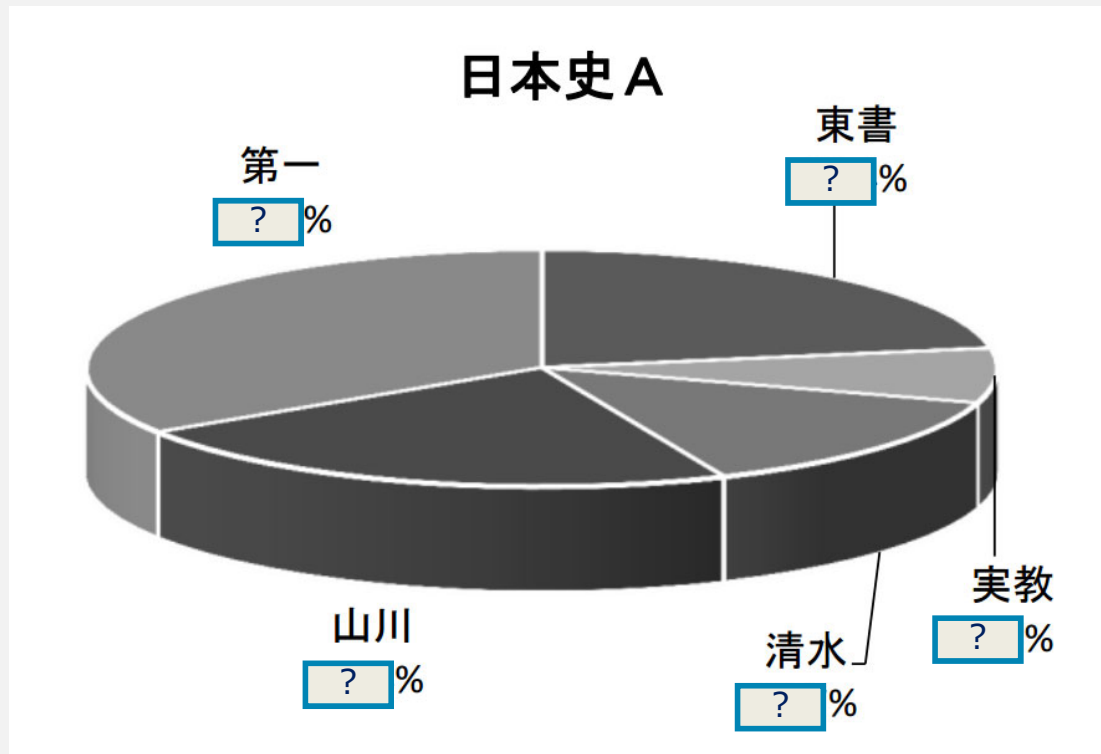
見やすくする

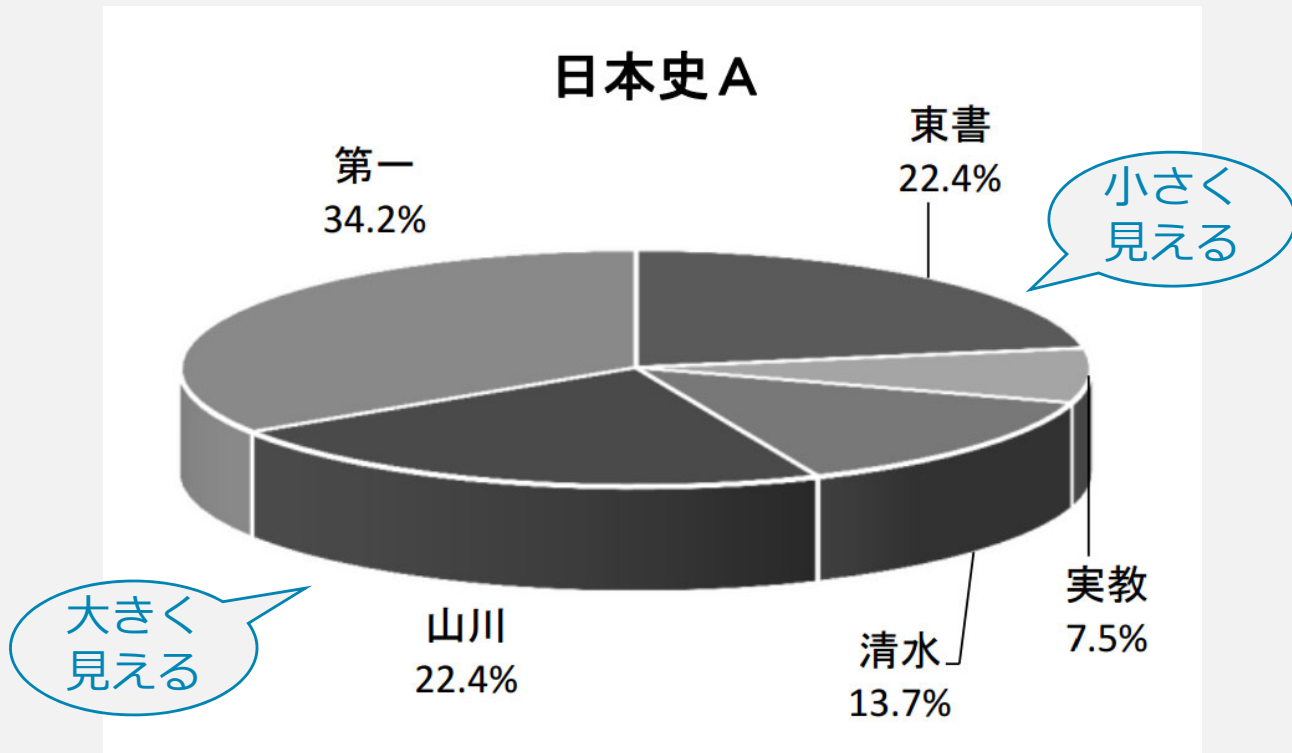
因果推論

直感的には分かりにくい数値を
ひとめ見て分かるように図示する。
要約統計量だけでは、実情を掴めない。

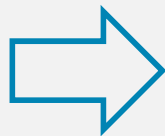


適切なグラフ作成





直感的には分かりにくい数値を
ひとめ見て分かるように図示する。
要約統計量だけでは、実情を掴めない。



適切なグラフ作成

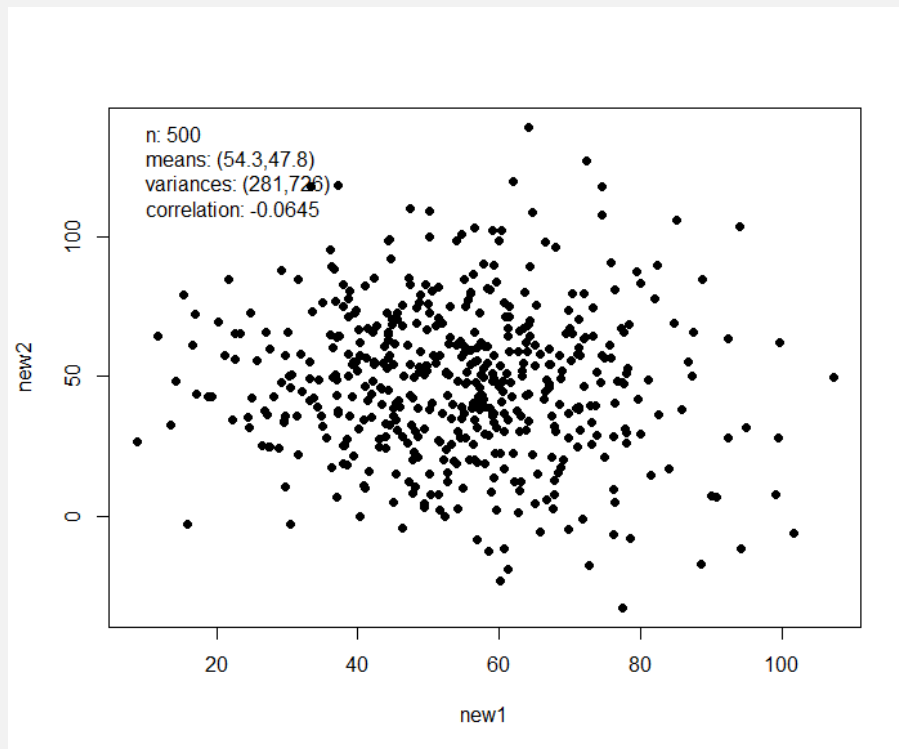
データセットA

- 500症例
- データX
 - 平均=54.3
 - 分散=281 (SD=16.8)
- データY
 - 平均=47.8
 - 分散=726 (SD=26.9)
- 相関係数=-0.0645

データセットB

- 142症例
- データX
 - 平均=54.3
 - 分散=281 (SD=16.8)
- データY
 - 平均=47.8
 - 分散=726 (SD=26.9)
- 相関係数=-0.0645



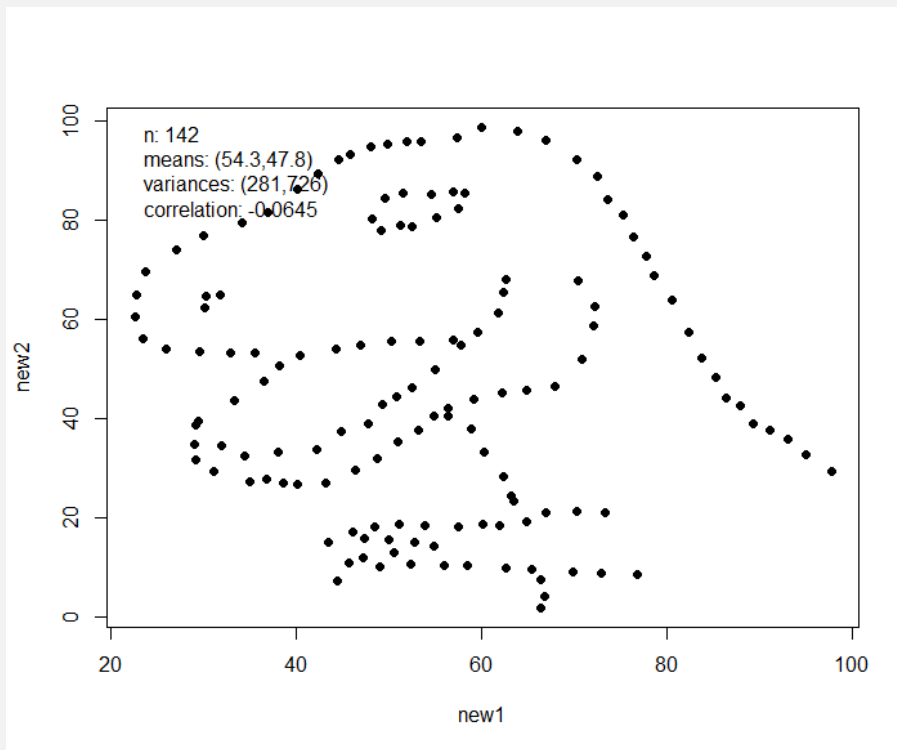


データセットA

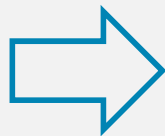
- 500症例
- データX
 - 平均=54.3
 - 分散=281 (SD=16.8)
- データY
 - 平均=47.8
 - 分散=726 (SD=26.9)
- 相関係数=-0.0645

データセットB

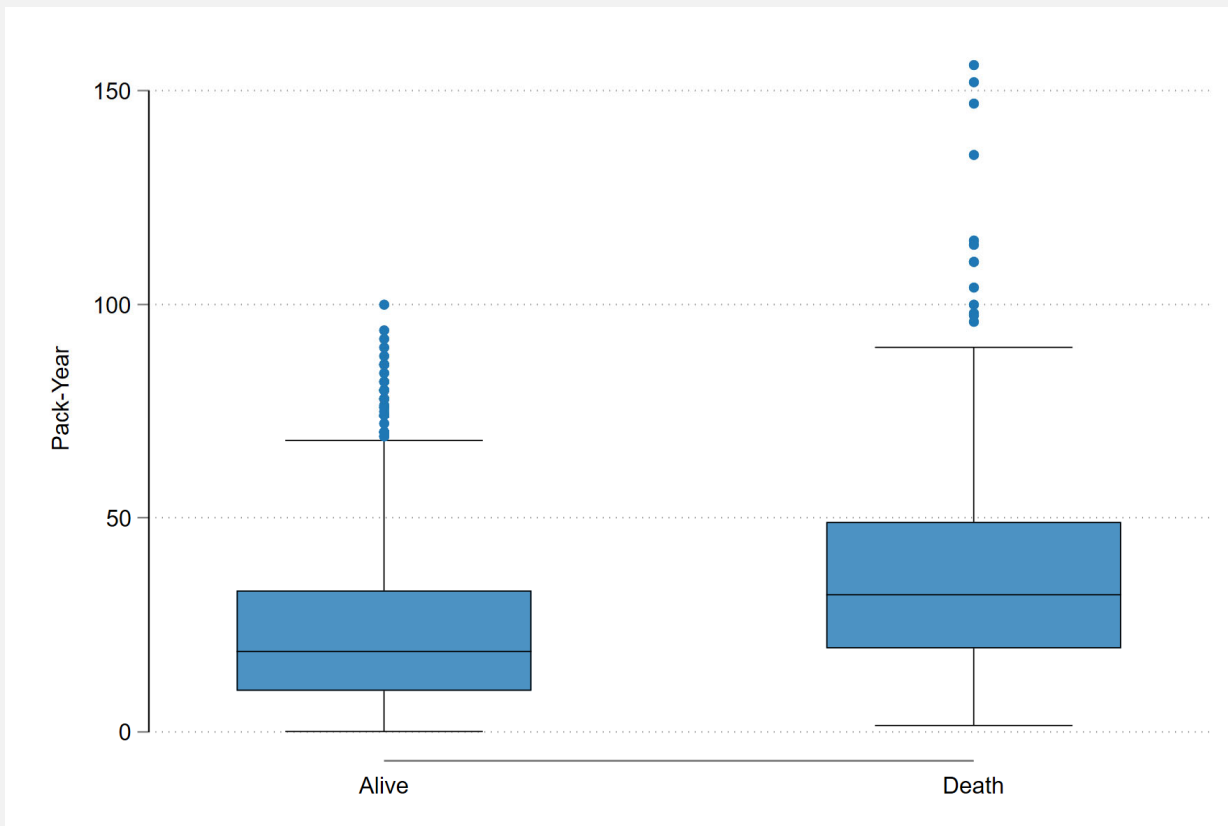
- 142症例
- データX
 - 平均=54.3
 - 分散=281 (SD=16.8)
- データY
 - 平均=47.8
 - 分散=726 (SD=26.9)
- 相関係数=-0.0645

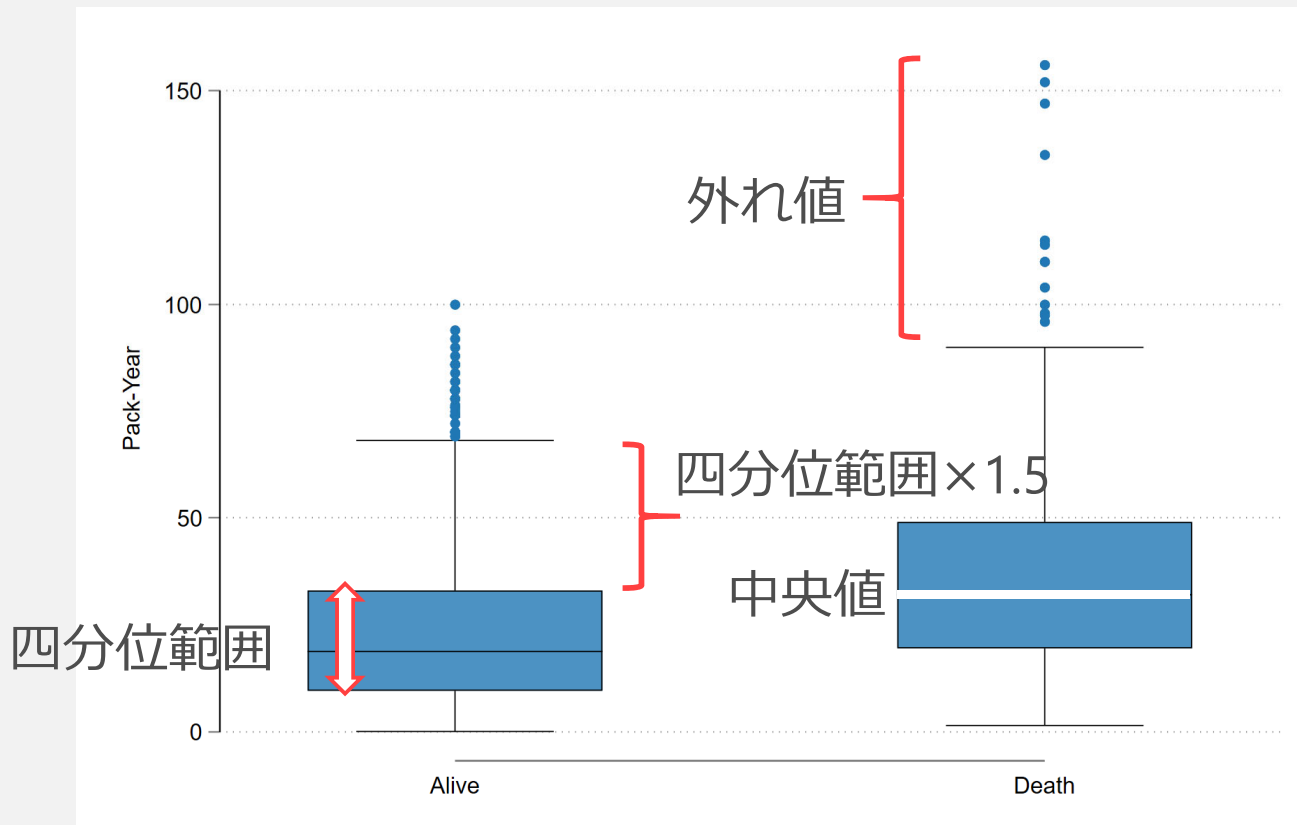


直感的には分かりにくい数値を
ひとめ見て分かるように図示する。
要約統計量だけでは、実情を掴めない。

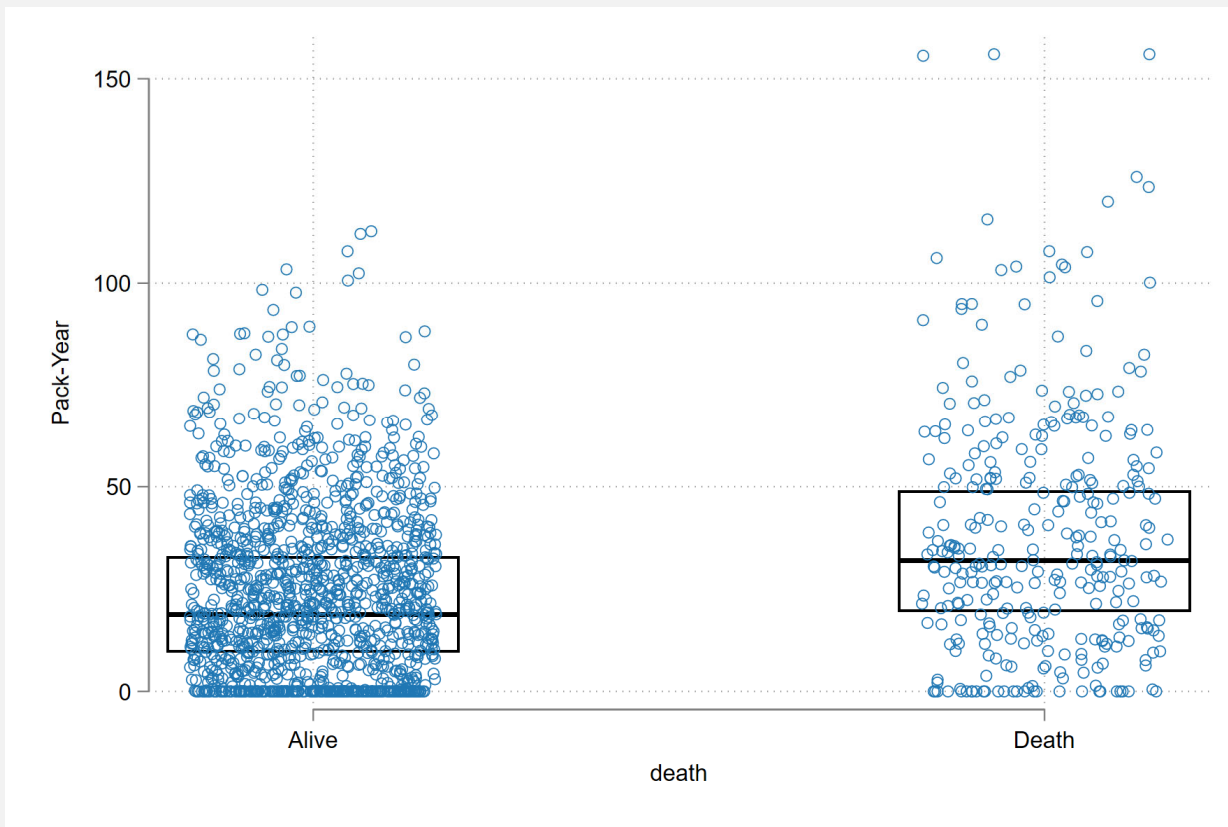


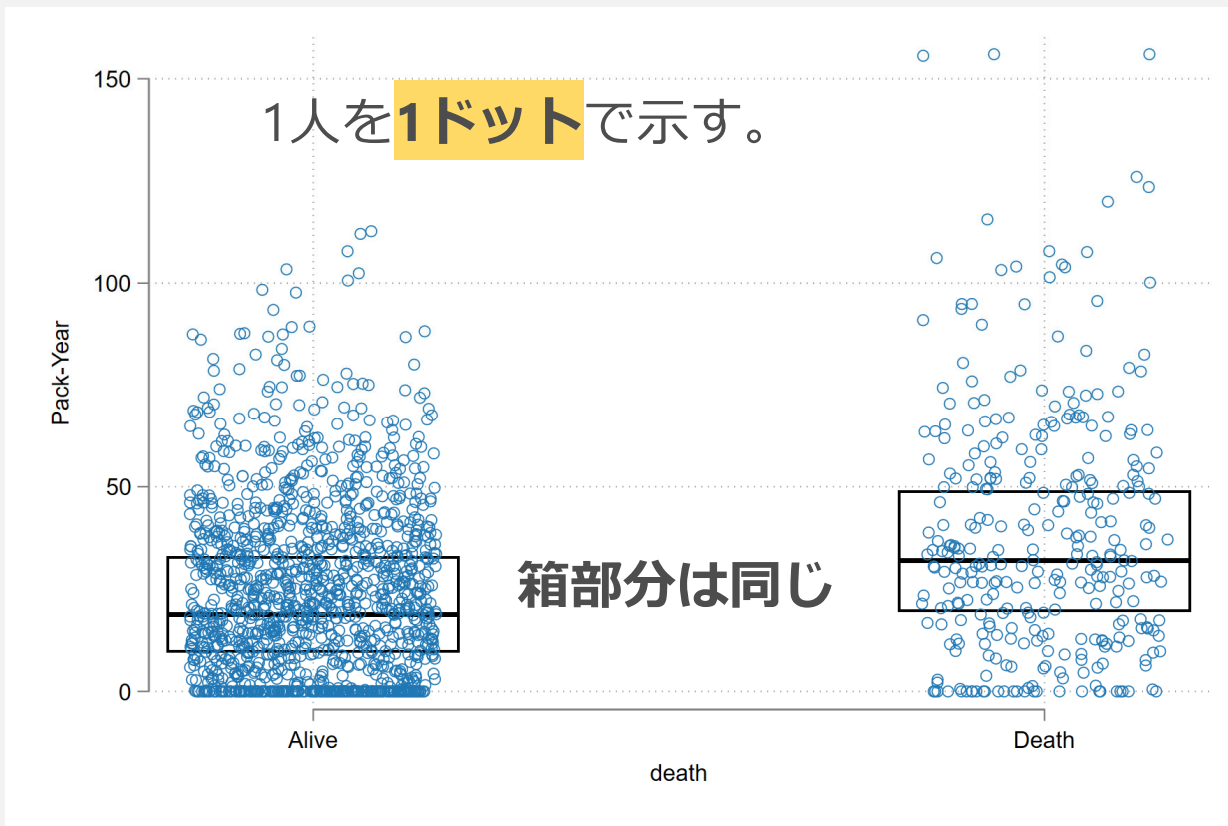
適切なグラフ作成

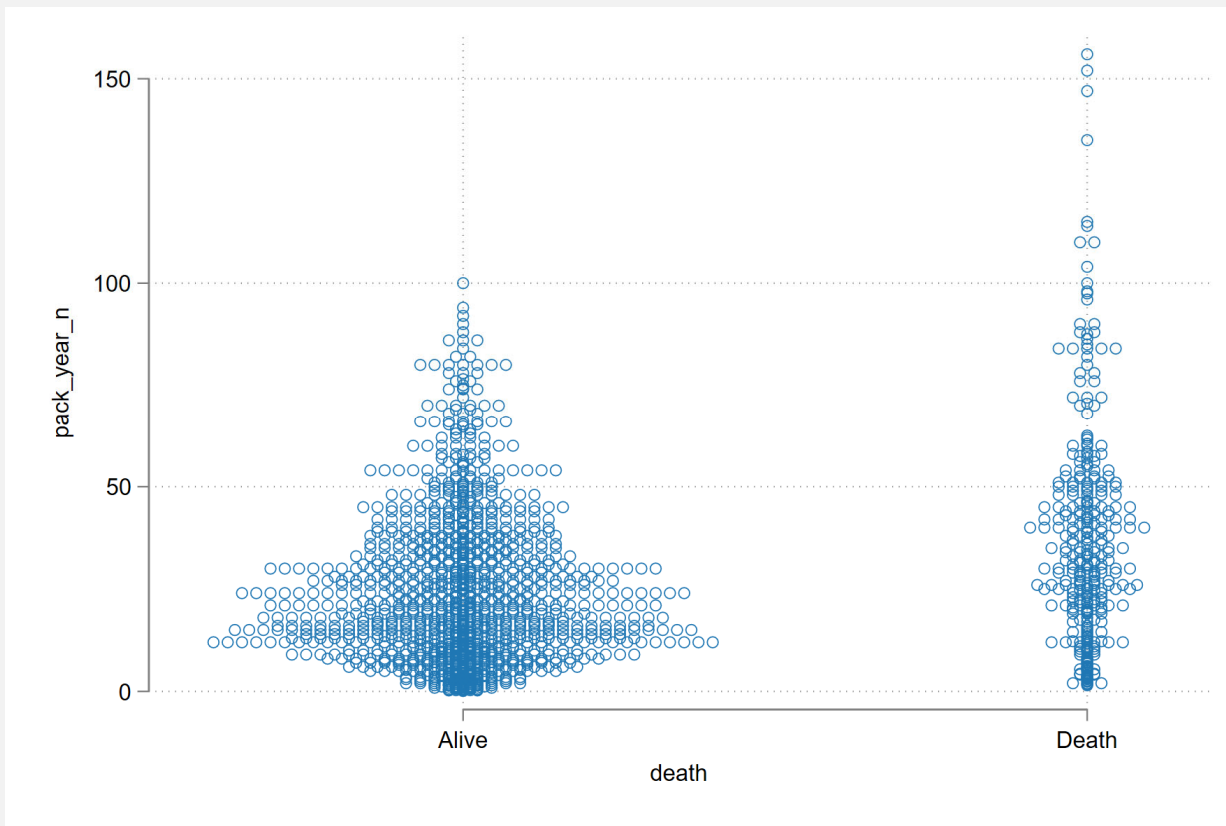


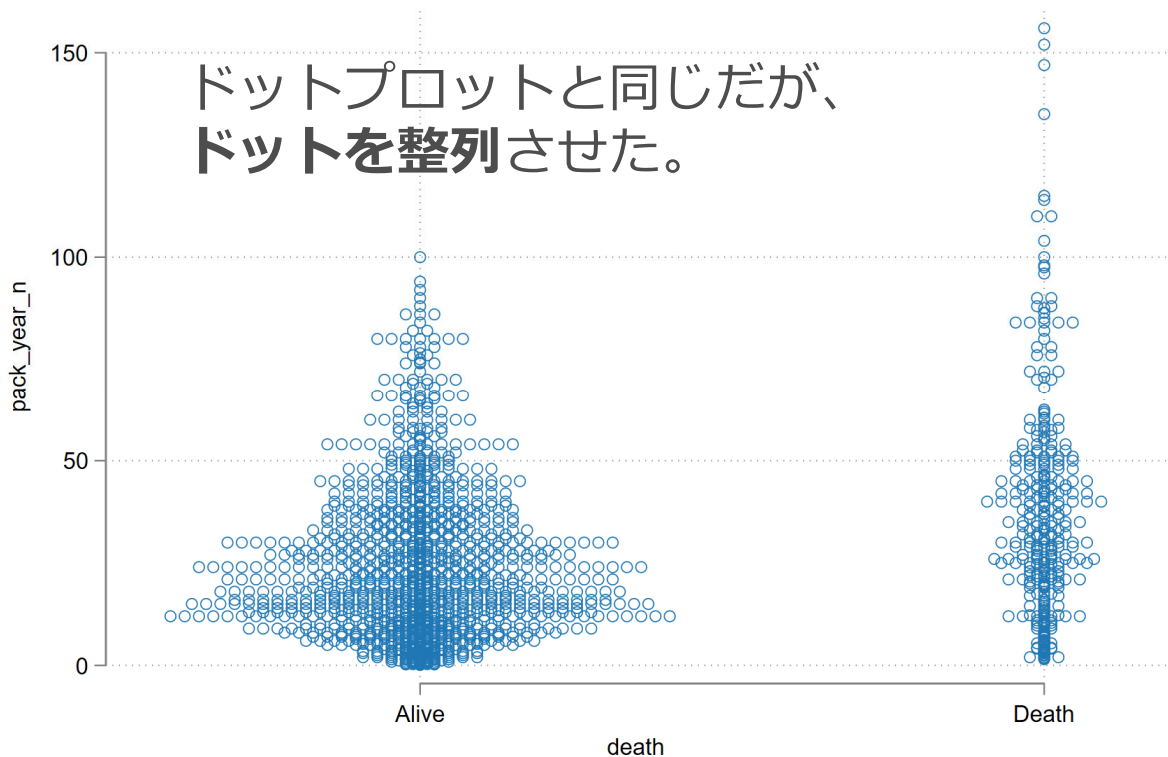


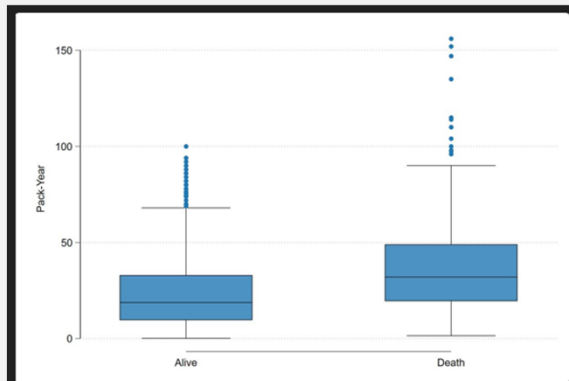
- 図示することで
分かりやすくなっているが、**情報が落ちている。**
- 生存者と死亡者はどちらが多い？



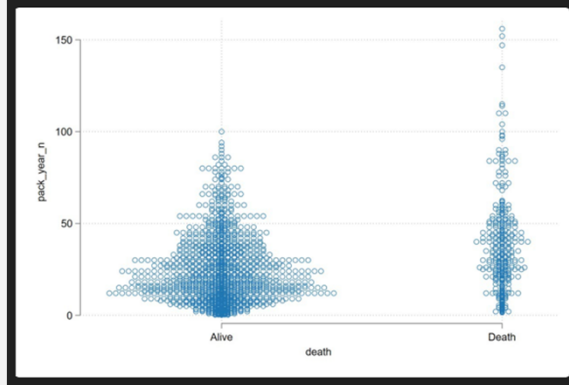






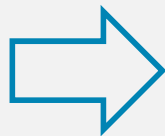


同じ分布だが、死亡群でやや喫煙本数が多い
(少し上にシフトした印象)



分布が全然違う。

直感的には分かりにくい数値を
ひとめ見て分かるように図示する。
要約統計量だけでは、実情を掴めない。



適切なグラフ作成

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

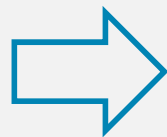
実例

要約する

見やすくする

因果推論

介入／曝露がアウトカムに与える影響を推定する。



各種統計モデル

因果
推論

検定の話はしないのか？

線形回帰分析

重回帰分析

バイアス

因果
推論

検定の話はしないのか？

線形回帰分析

重回帰分析

バイアス

統計学の講義では、Student's t-testを教わったが...？

今日はその話は無いのか？

ない

検定は、データから

「有意である」

「有意とは言えない」

この**いずれかの情報しか**

引き出せない。



推定を行い、

「どのくらいの差があるのか？」

「どのくらい確からしいのか？」

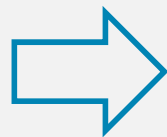
という**豊富な情報を引き出す**べき。



推定は検定よりも計算量が多い。
コンピュータ能力が低い頃は
難しく、検定が行われていた。



介入／曝露がアウトカムに与える影響を推定する。



各種統計モデル

アウトカム変数	よく使われる統計モデル
連続量	重回帰分析
二値 (あり or なし)	ロジスティック回帰分析
非負整数 (1, 2, 3, ...)	ポアソン回帰分析
生存時間	Cox比例ハザードモデル

因果
推論

検定の話はしないのか？

線形回帰分析

重回帰分析

バイアス

研究仮説（目的）

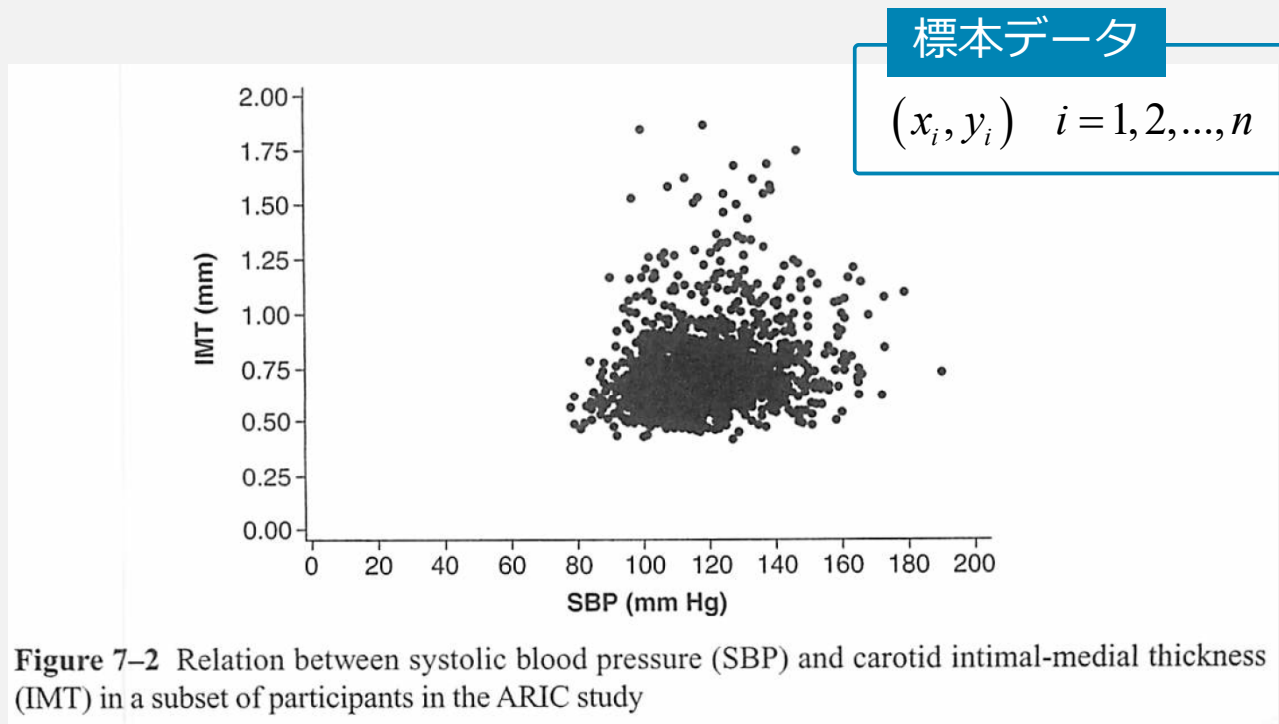
収縮期血圧と頸動脈の内膜中膜複合体厚の関連評価

研究デザイン：横断研究

曝露：収縮期血圧（systolic blood pressure: [SBP](#)）

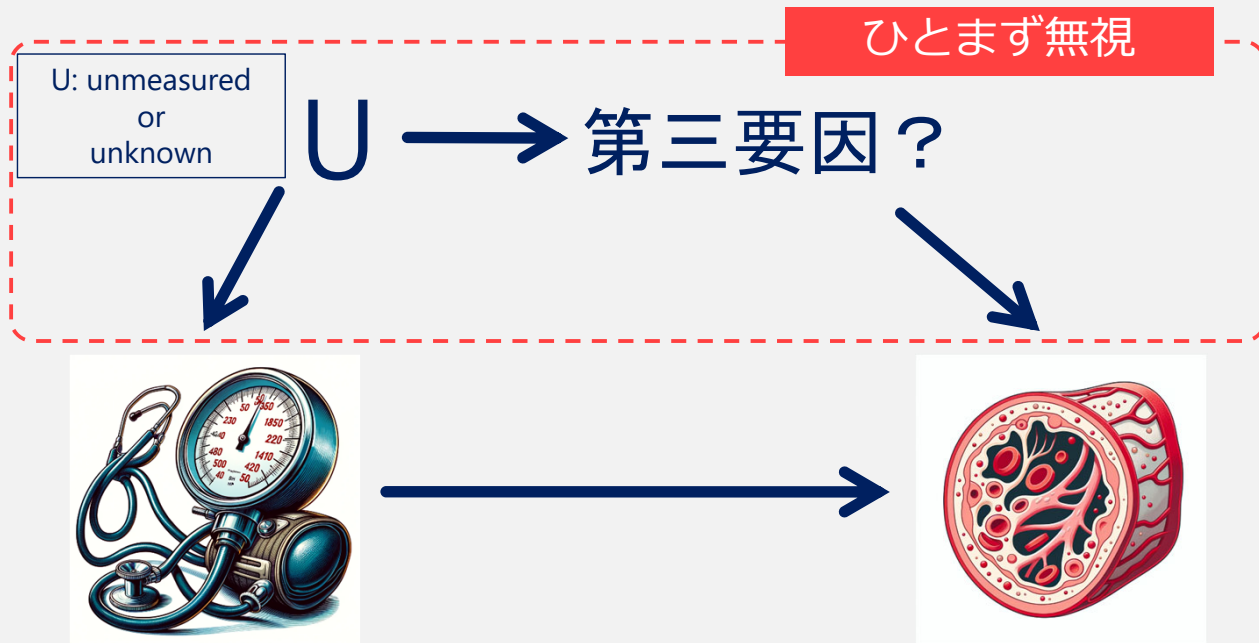
アウトカム：頸動脈の内膜中膜複合体厚（[IMT](#)）

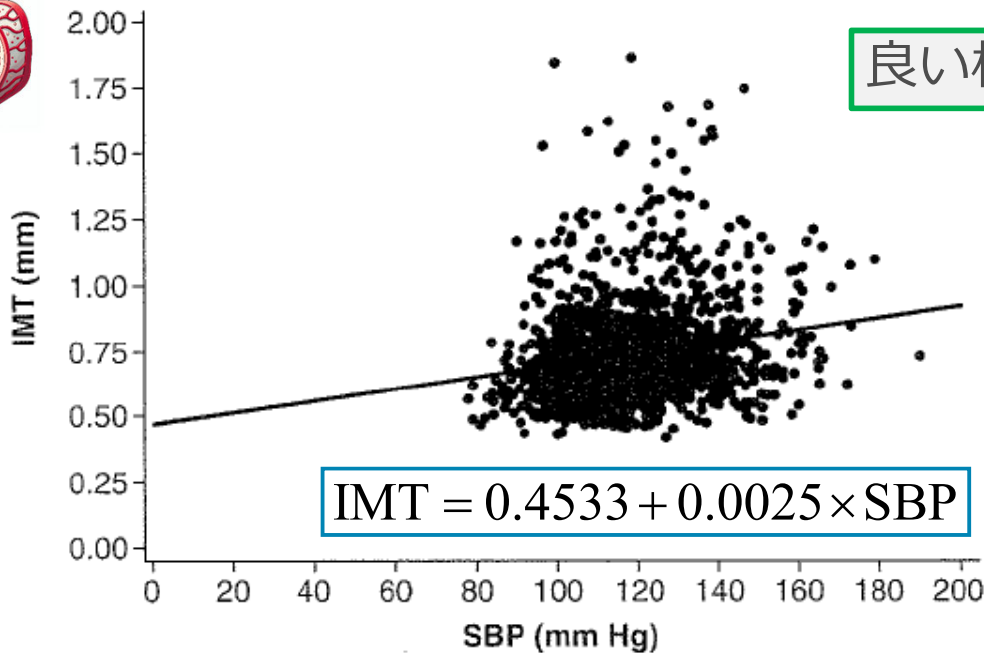
対象者：ARIC Study (1987–1989) に参加した1,410人



- 収縮期血圧が高くなればなるほど、
頸動脈の内膜中膜複合体厚 (IMT) は厚くなるのだろうか？

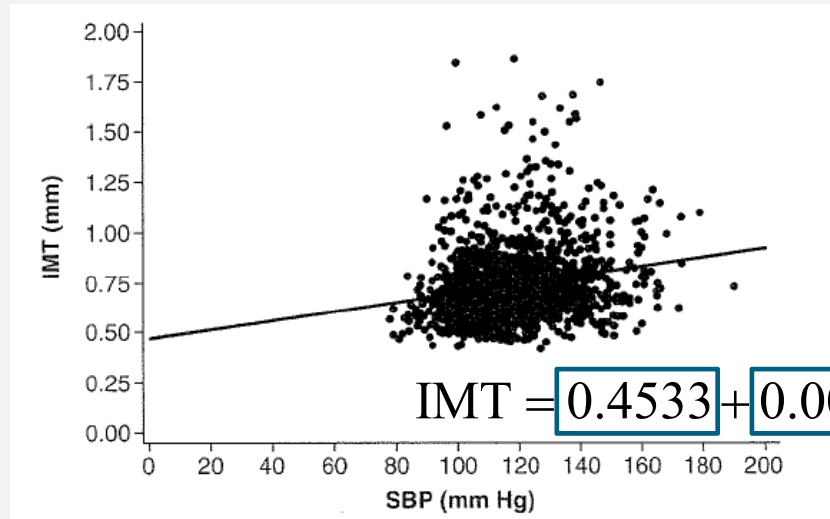
- もしそうだとしたら、
その**関連の強さはどの程度なのだろうか？**



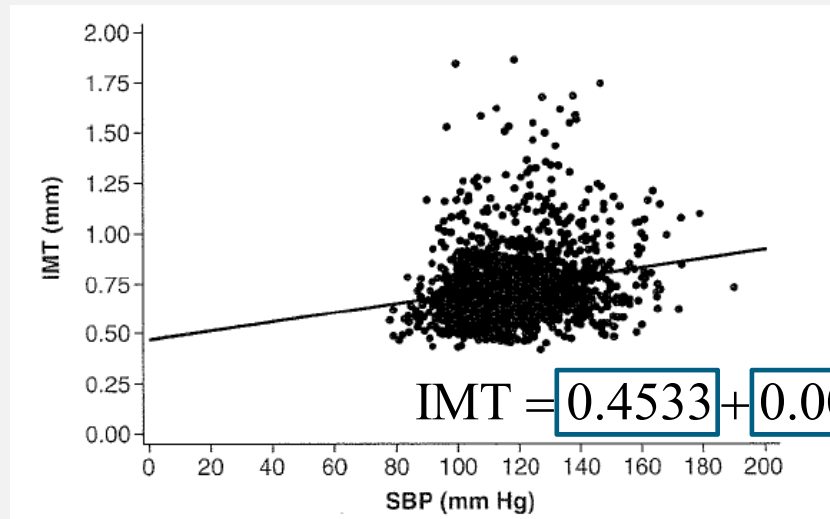


良い様に線を引いた



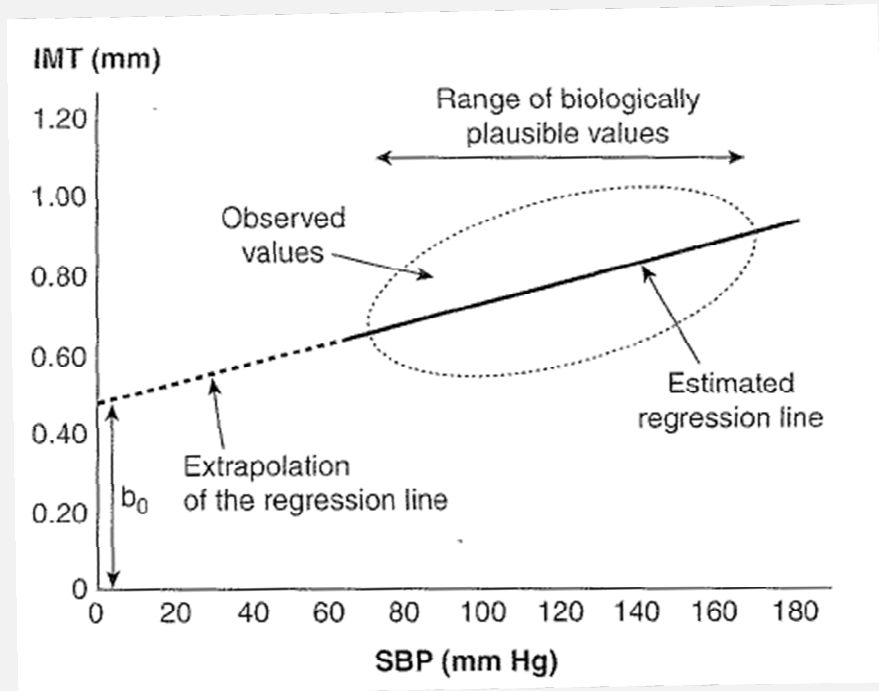


- 傾き slope である **0.0025** はどのように解釈できるか？
 - 収縮期血圧 SBP が1単位 上昇するごとに、
“平均して”内膜中膜複合体厚 IMT が0.0025 mm厚くなることを意味する
 - 分析の単位が重要である（説明変数間の単純な比較は要注意）



● 切片 intercept である **0.4533** はどのように解釈できるか？

- 収縮期血圧 SBP が 0 mm Hg である時に期待される内膜中膜複合体厚 IMT が 0.4533 mm であることを意味する



- この場合、切片の推定値に意味のある解釈はできない
 - 単に、外挿 extrapolation することによって推定されたに過ぎない

$$E(Y_i) = \beta_0 + \beta_1 x_i \quad (i = 1, 2, \dots, n)$$

アウトカム
outcome

切片
intercept

傾き
slope

曝露
exposure

パラメータ

回帰モデル regression model

$$Y_i = \beta_0 + \beta_1 x_i + E_i \quad (i = 1, 2, \dots, n)$$

系統的成分
systematic part

ランダム成分
random part

- E_i の誤差モデル
 - 期待値0 $E(E_i) = 0$
 - 分散一定 $\text{Var}(E_i) = \sigma^2$
 - 独立 (無相関) $E_i \perp\!\!\!\perp E_j \quad (i \neq j)$
 - 正規性 $E_i \sim N(0, \sigma^2)$

良い様に線を引いた

左の条件を満たすように
 β_0 と β_1 の値を決めた

- 収縮期血圧が高くなればなるほど、
頸動脈の内膜中膜複合体厚は厚くなるのだろうか？
- ◆厚くなる

□もしそうだとしたら、

その関連の強さは**どの程度なのだろうか？**

◆収縮期血圧が**1 mm Hg上昇**するごとに、
内膜中膜複合体厚が**0.0025 mm厚く**なる

因果
推論

検定の話はしないのか？

線形回帰分析

重回帰分析

バイアス



- 単純な線形回帰モデルの**説明変数を複数に拡張した**モデル
 - 目的変数：連続変数
 - 説明変数：連続変数、カテゴリー変数、順序変数など、変数の型は問わない
- 層別分析と比べて、以下のメリットがある
 - 容易に複数の交絡因子を同時に調整できる
 - 連続変数の説明変数を扱うことができる

$$E(Y_i) = \beta_0 + \beta_1 x_{1i} + \sum_{l=2}^k \beta_l x_{li} \quad (i = 1, 2, \dots, n)$$

アウトカム
outcome

切片
intercept

傾き
slope

曝露
exposure

($k-1$ 個の)
交絡因子

- 研究仮説（目的）：

収縮期血圧と頸動脈の「内膜中膜複合体厚」の関連評価

- 曝露：収縮期血圧 (SBP)

- アウトカム：頸動脈の内膜中膜複合体厚 (IMT)

- **交絡因子（共変量 covariates）：年齢、性別、BMI**

Y_i (IMT) の平均値モデル

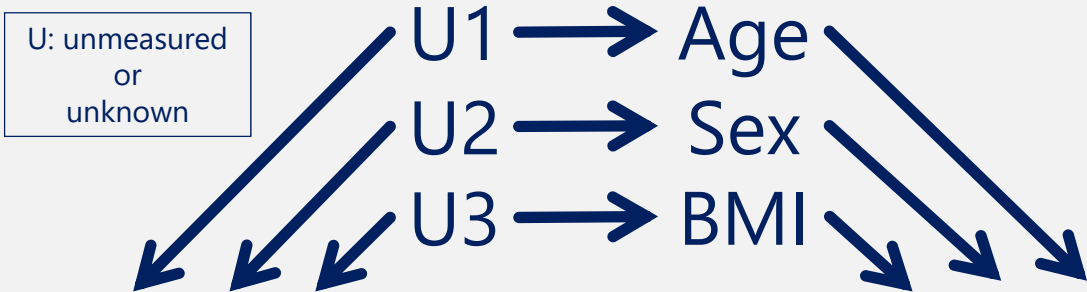
$$E(Y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} \quad (i = 1, 2, \dots, n)$$

X_1 : SBP (1 mm Hg)

X_2 : age (1 year)

X_3 : sex (men vs. women)

X_4 : body mass index (1 kg/m²)



- 以下のように、調整する交絡因子を徐々に増やした

- Model 1: 何も調整しない（粗解析）

$$\text{IMT} = \beta_0 + \beta_1 \times \text{SBP}$$

- Model 2: 年齢だけ調整する

$$\text{IMT} = \beta_0 + \beta_1 \times \text{SBP} + \beta_2 \times \text{age}$$

- Model 3: 年齢と性別を調整する

$$\text{IMT} = \beta_0 + \beta_1 \times \text{SBP} + \beta_2 \times \text{age} + \beta_3 \times \text{sex}$$

- Model 4: 年齢、性別、BMIを調整する

$$\text{IMT} = \beta_0 + \beta_1 \times \text{SBP} + \beta_2 \times \text{age} + \beta_3 \times \text{sex} + \beta_4 \times \text{BMI}$$

調整することにより、SBPの係数 β_1 はどう変化したか？

	Model 1	Model 2	Model 3	Model 4
Intercept	0.4533	-0.0080	0.0107	-0.0680
SBP (1 mmHg)	0.0025	0.0016	0.0014	0.0012
Age (1 yr)	Not included	0.0104	0.0096	0.0099
Sex (men = 1, women = 0)	Not included	Not included	0.0970	0.0981
BMI (1 kg/m ²)	Not included	Not included	Not included	0.0033

Interpret the estimated coefficient of SBP in each model.



回帰モデル regression model

$$Y_i = \beta_0 + \sum_k \beta_k x_{ki} + E_i \quad (i = 1, 2, \dots, n)$$

系統的成分
systematic part

ランダム成分
random part

- E_i の誤差モデル
 - 期待値0 $E(E_i) = 0$
 - 分散一定 $\text{Var}(E_i) = \sigma^2$
 - 独立 (無相関) $E_i \perp E_j \quad (i \neq j)$
 - 正規性 $E_i \sim N(0, \sigma^2)$

因果
推論

検定の話はしないのか？

線形回帰分析

重回帰分析

バイアス

- 因果推論では、各種バイアスによって因果関係が歪められることに注意して、統計解析を行う。

- 測定バイアス
 - データ取得の際によって生じる。
- 選択バイアス
 - 対象者選択などによって生じる。
- 交絡バイアス
 - 共通原因などによって生じる。

- 研究デザインで対応することが重要だが、統計解析でも何とかならないこともない。
- Directed Acyclic Graphなどを活用する。

	研究デザイン	統計解析
交絡バイアス	○	○
選択バイアス	○	×
測定バイアス	○	×

従来からよく言われていること

	研究デザイン	統計解析
交絡バイアス	○	○
選択バイアス	○	○
測定バイアス	○	○

理論的には何とかできる。

	研究デザイン	統計解析
交絡バイアス	△	△
選択バイアス	△	△
測定バイアス	△	△

理論的には可能だが、現実はこのもの。
△を○にするためには、地道な努力が必要

- 傾向スコアを用いる方法
 - IPTW (Inverse Probability of Treatment Weighting)
 - 傾向スコアマッチング
 - 傾向スコアによる層別
 - g-formula
 - Doubly robust推定法
 - 操作変数法
- その他百花繚乱

基本

どのレイヤーが重要なのか？

データをどうしたいのか？

実例

要約する

見やすくする

因果推論

1. 研究において統計解析は下流工程
2. 統計解析手法よりも**上流工程**が重要
3. 研究目的・研究デザインにより解析方法は決まる。