

## 9. 統計学的推定と検定

高木 廣文  
東邦大学看護学部  
国際保健看護学研究室

1

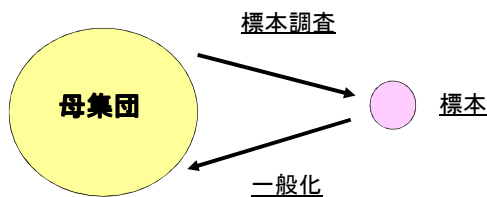
### 統計学的な推論とは？

- 1) 母集団から標本を抽出→調査, データ収集
- 2) データ ⇒ 母集団の特性値について推論
  - ・ 推定 estimation
  - ・ 検定 test

- 推定：母集団の特性値の大きさを推測
- 検定：母集団に関する**仮説の成立・不成立**を検証

2

### 推定・検定の考え方



3

### 統計学的推定とは？

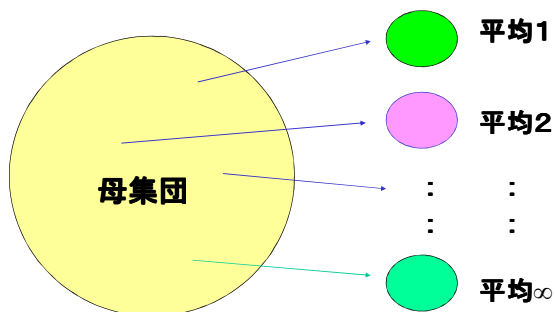
■ 標本のデータから, 母集団の特性値を推測

- 期待値：  
標本の取り方は極めて多数あり, 個々のデータから得られ割合や平均値は異なる。

あらゆる標本で求められた割合や平均値の「平均」を期待値と呼ぶ

4

### 期待値とは統計量の平均



5

### 標本の抽出方法は無数？

100人から10人を無作為に抽出：

$$\text{組合せの個数} = {}_{100}C_{10} = \frac{100!}{10! \times 90!}$$

$$= 17,310,309,456,440 \text{ 通り}$$

6

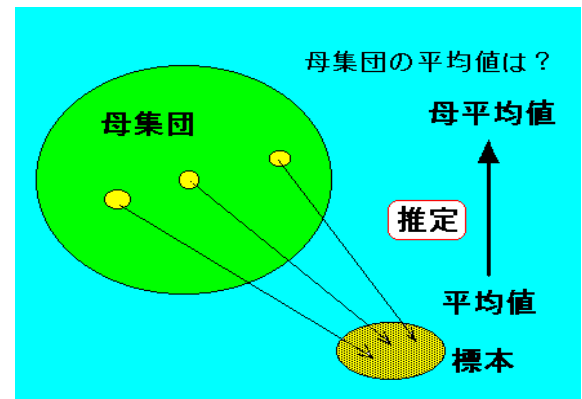
## 期待値の考え方

取り得る全ての平均値の平均：

$$\lim_{n \rightarrow \infty} \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_n}{n}$$

7

## 推定の考え方



## 点推定について

標本データから得られた割合や平均値などの数値で、母集団の特性値を推測する。

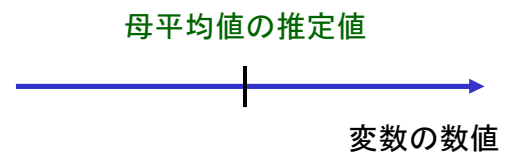
標本平均値の期待値 = 母平均値  
標本割合の期待値 = 母割合 (比率)

標本がランダムに選ばれていれば、平均値や割合はそのまま母集団値の推定値になる。

9

## 点推定 point estimation

- 点推定は数直線上の1つの点で母数を推定



10

## 区間推定について

1つの数値ではなく、ある程度の幅で母集団の特性値の推測を行なう。

【問】 幅の大きさはどのように決めるのか？

【答】 一定の信頼性を付与した区間を考える

・ 信頼区間 confidence interval

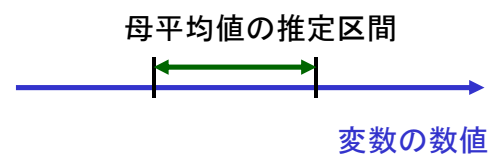
— 95%信頼区間, 99%信頼区間

同一の調査を繰り返した場合、95%、99%は区間内に母平均値等が存在

11

## 区間推定 interval estimation

- 区間推定は数直線上の区間で母数を推定



12

## 点か区間か？

例) 内閣の支持・不支持を調べたところ支持するとした割合は35%であった。

- 【問】日本全体での内閣支持率は、  
点推定 : 35% (?)  
区間推定 : 95%信頼区間は32~38%  
⇒ わかりやすさか、信頼性か？

13

## 統計学的検定について

推定と同様に、標本についてではなく、母集団の各種の特性について言及する方法

- ・ **帰無仮説 null hypothesis** :  
母集団についての仮説を設定。  
その仮説が成立つかどうかを、データを用いて調べる。  
- 研究上の「**作業仮説**」とは異なる！

14

## 作業仮説と帰無仮説

(例) 喫煙と心疾患の関係を調べる。

- **研究上の作業仮説** :  
「喫煙は心疾患の原因である」  
どうやって調べれば、作業仮説が証明できるのか？  
- 喫煙者と非喫煙者で心疾患の発生割合などを比較

15

## 作業仮説の直接証明は困難！

- 2つのグループしかなければ、標本データでは必ずどちらかの割合が大きくなる。  
- 喫煙者で心疾患の発生割合が高くても、偶然なのか必然なのか、直接的に証明するのは困難。

【問題】 どうすれば、2つのグループで差があることを言えるのか？

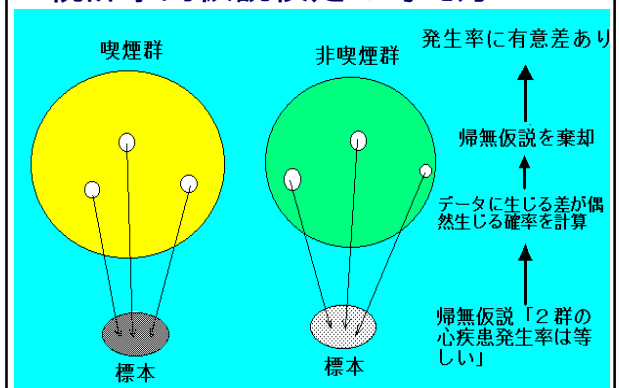
16

## 統計学的仮説検定の考え方

- 1) 帰無仮説を立てる。  
「喫煙グループと非喫煙グループの心疾患発生率は等しい」
- 2) 仮説が正しいとした場合、実際のデータで観察される差が偶然生じる確率を求める。
- 3) 確率が小さい場合、その差が偶然生じるとは考えにくい。
- 4) 仮説を否定し、差があるとする (**有意差**) 。

17

## 統計学的仮説検定の考え方



## 様々な検定方法

- 確率を計算するのに便利のように、仮説が正しいと仮定した場合、ある統計量を計算  
⇒ 統計量によって、**z検定**、**t検定**、 **$\chi^2$ 検定**、**F検定**などと呼ばれる。
- 統計学的検定では、仮説は**棄却reject**されないと意味がない。
- 確率が大きい場合 ⇒ 「**保留**」となる。

19

## 検定における2つの誤り

- 1) 有意水準 **significance level ( $\alpha$ )**  
帰無仮説を棄却するか否かという、検定の目安となる確率の大きさ。

仮説が正しい場合、誤って仮説を棄却する誤り、の判断の基準となる確率

- ・ **第1種の誤り (過誤) type I error**  
通常、5%、1%が用いられる。

20

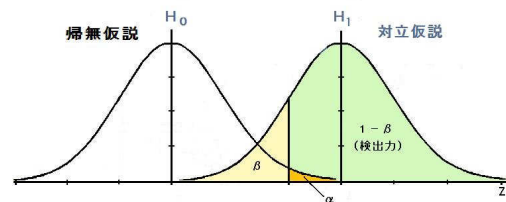
## 第2種の誤りと検出力

- 2) 第2種の誤り (過誤) **type II error ( $\beta$ )**  
仮説が誤っているのに、仮説を採択してしまう誤り (**仮説を棄却できない誤り**)

- 検定の検出力 **power of a test**

第2種の誤り  $\beta$  に対して、 $1 - \beta$  は「帰無仮説が誤っている場合に、帰無仮説を棄却する」確率

21



有意水準と検出力の関係

22

## 片側検定と両側検定

- 帰無仮説と対立仮説  
帰無仮説  $H_0$ : 「 $A = B$ 」の場合
  - ・ 帰無仮説の否定  
対立仮説  $H_A$ : 「 $A \neq B$ 」
  - ・ 2つの仮説に分けられる。  
対立仮説  $H_1$ : 「 $A > B$ 」  
対立仮説  $H_2$ : 「 $A < B$ 」

23

## 両側検定 both sided test, two tailed test

- 有意確率の計算を実際に行う場合、帰無仮説のもとでの検定統計量の分布を考え、その分布の両側の確率を求める。
  - ・ 「**両側検定**」と呼ばれる。  
有意水準を  $\alpha$  とした場合、上側と下側の有意水準はそれぞれ  $\alpha/2$  と等しくする。

24

## 両側検定の考え方

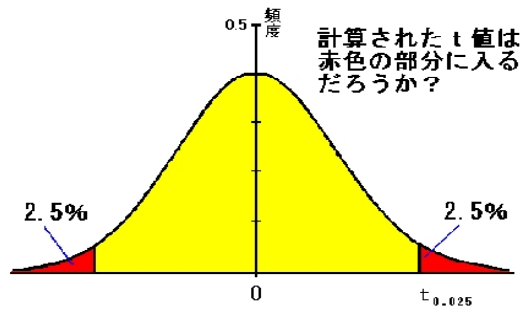


図 両側検定の考え方

## 片側検定 one sided test, one tailed test

- 有意確率の計算では、検定統計量の分布の上側の部分のみ、または下側の部分のみの確率のみを考える。  
⇒ 「片側検定」と呼ばれる。

(注) 仮説についての事前情報がない場合、両側検定が用いられる。  
データを見てから片側検定にしてはならない。

26

## 片側検定の考え方

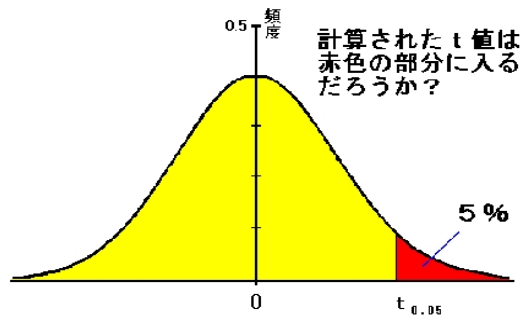


図 片側検定の考え方

## 検定方法の分類

- **パラメトリック検定:**

変数の母集団分布として、特定の分布 (通常は正規分布) を仮定する方法

- **ノンパラメトリック検定:**

母集団の分布を仮定しない方法。  
主に順位に基づく検定方法が多い。

パラメトリック検定より若干効率が悪い。

28

## よく用いられる検定方法

- **分布の位置について:**
  - P: 2群の母平均値の差の検定 (t検定)
  - 一元配置分散分析 (F検定)
  - N: 順位和検定 (U検定),
  - クラスカル・ウォリス検定 ( $\chi^2$ 検定)
- **2変数の関係について:**
  - P: 無相関の検定 (t検定)
  - N: 独立性の検定 ( $\chi^2$ 検定), 直接確率

29

## 10. 因果推論と研究デザイン

高木廣文

東邦大学看護学部  
国際保健看護学研究室

●タバコは肺がんの原因？

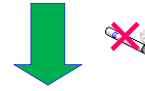
- ・ Hiroさんはタバコを吸っていた。
- ・ Hiroさんは肺がんになった。



Hiroさんの肺がんの原因はタバコである？

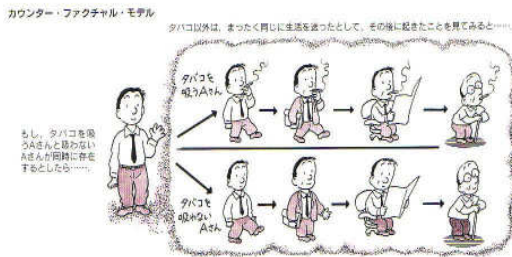
●タバコを吸わないのに肺がんになった

- ・ Hiroさんはタバコを吸っていない。
- ・ Hiroさんは肺がんになった。

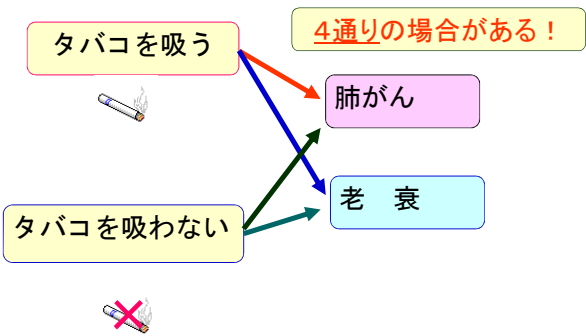


Hiroさんの肺がんの原因はタバコではない！

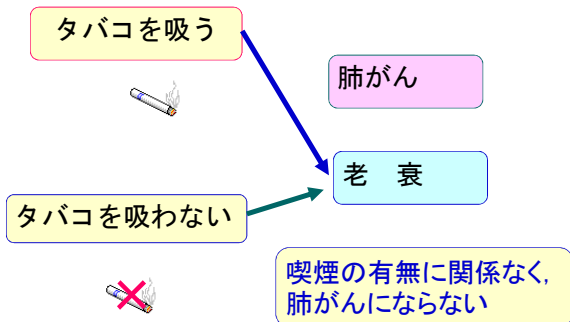
タバコを吸う人生と吸わない人生



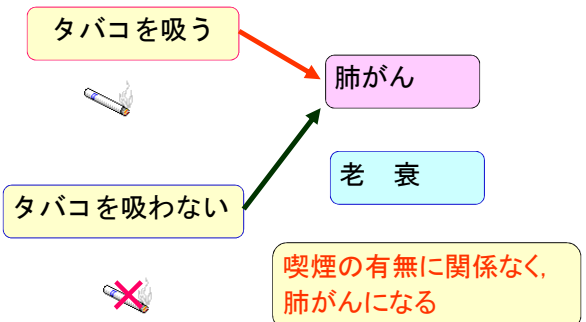
正しい因果推論は？



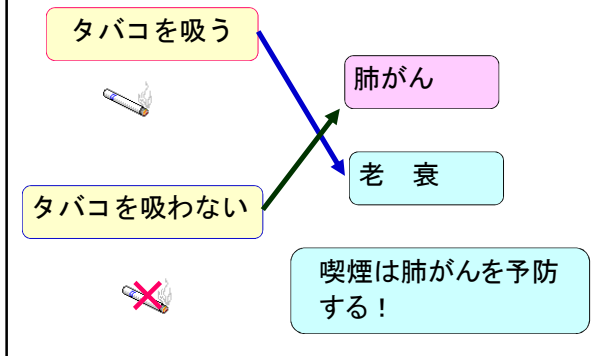
1) タバコは無関係：健康



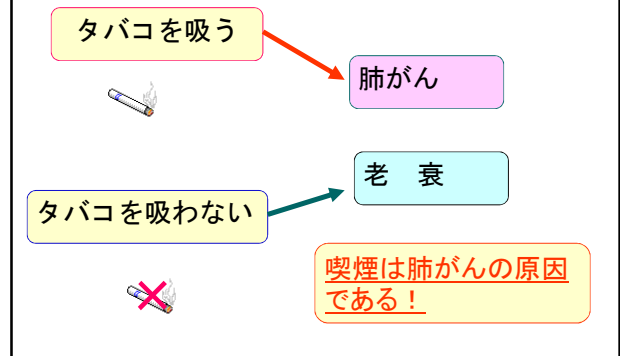
2) タバコは無関係：肺がん



### 3) タバコは健康のもと？



### 4) タバコは肺がんの原因



### タバコと肺がんの関係は？

違いは喫煙の有無のみなので、Aさんに起きたことの原因としてタバコの影響だけを考えることができるはず。



### ●正しい因果推論のためには

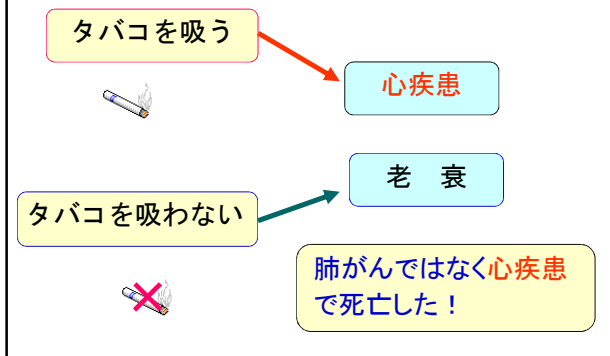
Hiroさんへのタバコの影響を調べるため

- ・タバコを吸うHiroさんを追跡調査： A
- ・タバコを吸わないHiroさんを追跡調査： B

正しい因果推論： AとBの結果を比較

Counter-Factual Model (半現実的モデル)

### ●Aさんへのタバコの影響は？



### ●集団へのタバコの影響を調べるには

多くの対象について

- ・タバコを吸う場合を追跡調査： A
- ・タバコを吸わない場合を追跡調査： B

正しい因果推論： AとBの結果を比較

肺がん、心疾患等、個別にリスクを推定する。

現実にはCounter-Factual Modelを適用する？

タイムマシンか多重世界を観測する装置が必要になる！



正しい因果推論のためには次善の策が必要！

●現実世界ではコントロールが必要

対象との比較のためのコントロール



調べたい要因以外は全て同一であることが望ましい

- ・一卵性双生児：遺伝的に同一
- ・兄弟姉妹：遺伝的に近い，同一家庭環境等
- ・似たものを集める

●マッチング matching

- ・対象にあわせたコントロールを選ぶ：性，年齢，職業などをあわせる。



前向き研究：曝露－非曝露研究  
 後向き研究：ケース・コントロール研究  
 ・1例ずつマッチさせるのは結構大変！

無作為割付 random allocation

研究対象者を多数集める

乱数で決める

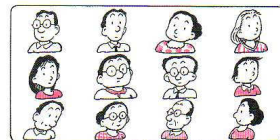
Case (新薬治療群)

Control (標準薬群)

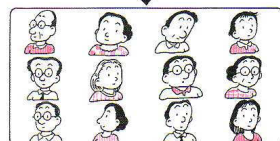
無作為割付の効果は？

無作為化

Case Group



Control Group



一人ひとり異なるが，集団で見るとかなり似たものになる。

無作為割付の特徴

- ・割り当てたグループでの各種特性は，平均的には同一になると期待される。
- ・一回の割り当てでは，必ずしも同一特性をもった集団に分割されるとは限らない。
- ・対象者数が多いほど，集団特性が均一化されるものと期待できる。
- ・各種の推測統計学の手法が適用可能。



## EBMで求められる研究デザイン

- I a 複数のランダム化比較試験のメタ分析
- I b 少なくとも1つのランダム化比較試験
- II a 少なくとも1つのよくデザインされた非ランダム化比較試験
- II b 少なくとも1つの他のタイプがよくデザインされた準実験的研究
- III 比較試験や相関研究, 症例対照研究など, よくデザインされた非実験的記述的研究
- IV 専門家委員会の報告や意見, あるいは権威者の臨床経験

(US Agency for Health Care Policy and Research)

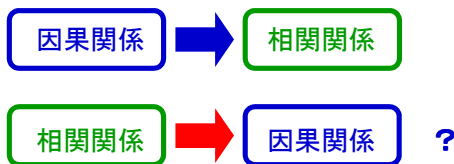
## 11. 因果推論とHillの9条件

高木廣文  
東邦大学看護学部  
国際保健看護学研究室

50

## 相関関係と因果関係

### ●相関correlationと因果causality



- 相関関係が変数間に存在しても, 必ずしも因果関係の存在を意味しない。

51

## 多様な因果関係

- ・ 単一因子 - 単一結果
- ・ 単一因子 - 複数結果
- ・ 複数因子 - 単一結果
- ・ 複数因子 - 複数結果

■ 第3因子の存在 - 見かけ上の関係

■ 因果連鎖

52

## 因果関係のパスモデル

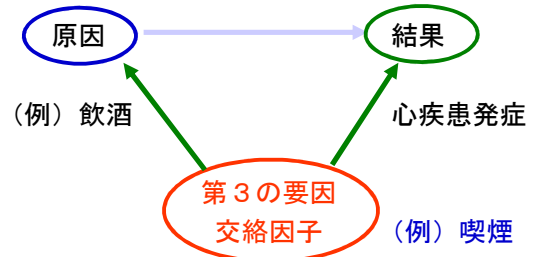


(例) 喫煙 肺癌発症

- 原因と結果を矢印で結び, 因果関係があることを示す (パス図, パスダイアグラム)。

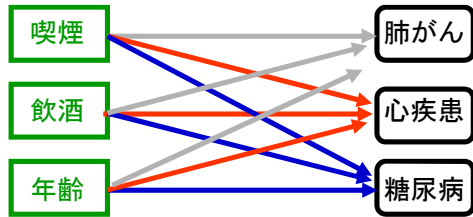
53

## 交絡のある因果関係



54

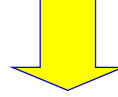
## 多要因一多結果



55

## 相関から因果を推論するには？

- 相関関係は因果関係を示すとは限らない



- 因果推論のためのマニュアルが必要

56

## コッホの原則 (1882)

1. 病原体は、それが引き起こすと考えられる疾患のどの症例にも存在することが純粋な培地内での分離によって示されねばならないこと
2. その病原体は他の疾患では見出されないこと
3. 分離された病原体は実験動物内で、その疾患を再び起こすことができること
4. 病原体はその実験動物から再び分離されること

⇒ 細菌学的方法論として有用  
非感染性の疾患については別の論理が必要

57

## Hillの9条件 (1965)

1. Strength (強さ) \*
2. Consistency (一致性) \*
3. Specificity (特異性) \*
4. Temporality (時間的先行性) \*
5. Biological gradient (生物学的傾向性)
6. Plausibility (尤もらしさ)
7. Coherence (整合性) \*
8. Experiment (実験)
9. Analogy (類似性)

\* "Smoking and Health" (米国公衆衛生局) (1964)

58

### 1. Strength (強さ) \*

相関係数, リスク比, オッズ比などの数値  
自体が大きいこと。

統計学的に有意だけでは十分ではない。

59

### 2. Consistency (一致性)

Has it been repeatedly observed by different  
persons, in different places, circumstances and  
times?

対象者, 地域, 環境, 調査時期などを変えて  
も, 同一の結果が得られるか?

60

### 3. Specificity (特異性)

If, as here, the association is limited to specific workers and to particular sites and types of disease and there is no association between the work and other modes of dying, then clearly that is a strong argument in favour of causation.

原因-結果には1対1対応が望ましい。  
しかし、1対多対応も多い。

61

### 4. Temporality (時間的先行性)

which is the cart and which the horse?

原因は結果よりも時間的に先にある。

62

### 5. Biological gradient (生物学的傾向性)

量反応関係 (doseresponse relationship) が認められること

望ましい条件であるが、必ずしも量-反応関係が認められるとは限らない。

63

### 6. Plausibility (尤もらしさ)

現在の医学的、生物学的常識などからみて尤もらしいこと

新しい知見は、常識に矛盾するかもしれない。

What is biologically plausible depends upon the biological knowledge of the day.

Sherlock Holmes advised Dr Watson, 'when you have eliminated the impossible, whatever remains, *however improbable*, must be the truth.'

64

### 7. Coherence (整合性)

すでに分かっている疾病の自然史や、生物学的事実と矛盾しないこと

過去の知見と矛盾しないこと。

65

### 8. Experiment (実験)

(動物) 実験で確認できれば、因果関係の強力な証拠になる

66

## 9. Analogy (類似性)

同様な因果関係がすでに認められていれば、観察した関係が因果であると考えやすくなる

67