

臨床研究における統計学

放射線影響研究所統計部
三角 宗近

Munechika Misumi, Ph.D.
Department of Statistics,
Radiation Effects Research Foundation





日本糖尿病学会 COI 開示

発表者名：三角 宗近

演題発表に関連し、開示すべきCOI関係にある
企業などはありません。

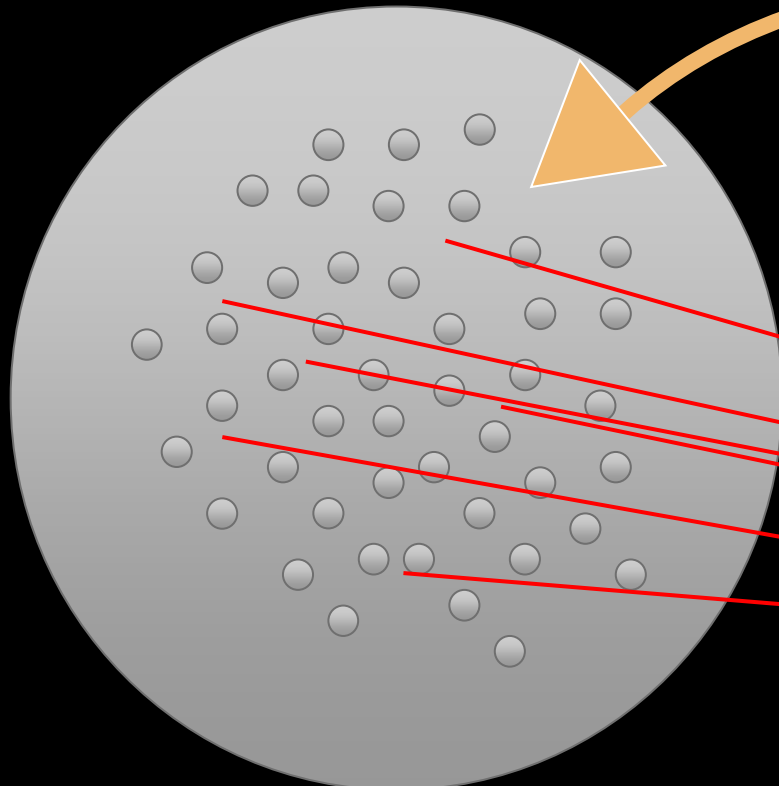
な ぜ

統 計 学 ？

母集団と標本

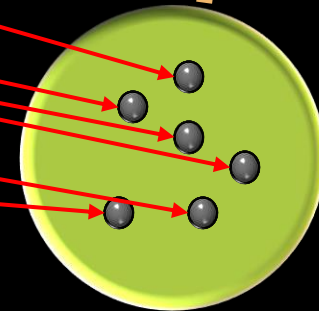
- 母集団(population)
- 調べたいことの対象となる集団

母集団



統計学的推測

標本(データ)



有限または無限

大きさ N

無作為抽出

統計的推測：母集団についての推測

- 同じ母集団について再度同じことを行えば、正しい推測ができていれば同じ結果が得られるはず（再現性）
- 現実では、有限のデータを再度集めて、同じように推定を行っても必ずしも全く同じ結果になるとは限らない

統計的推測：母集団についての推測

- 同じ母集団について再度同じことを行えば、正しい推測ができていれば同じ結果が得られるはず（再現性）
- 現実では、有限のデータを再度集めて、同じように推定を行っても必ずしも全く同じ結果になるとは限らない

なぜばらつくのか？ 個体差？環境に起因する違い？
人では特にばらつく要因は多数

不確実性の中で再現性を保証・主張する（外的妥当性、一般化）

➔ 統計学

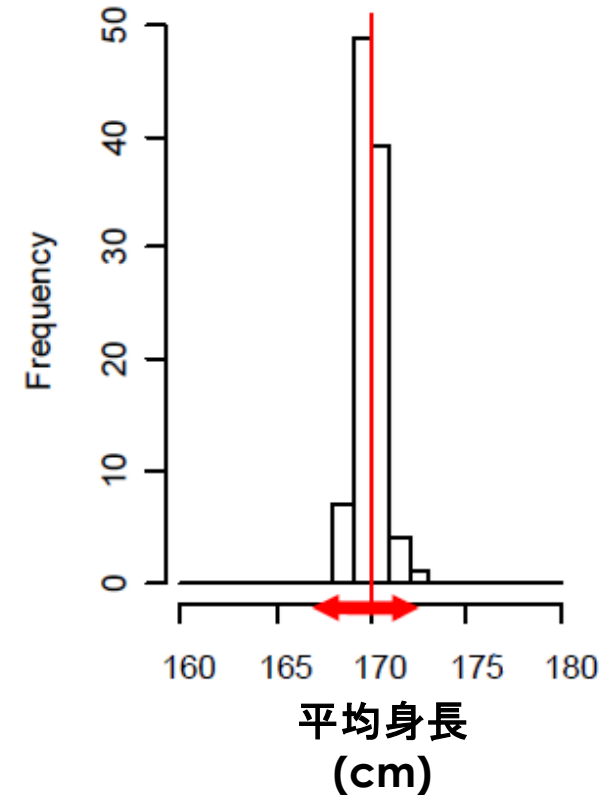
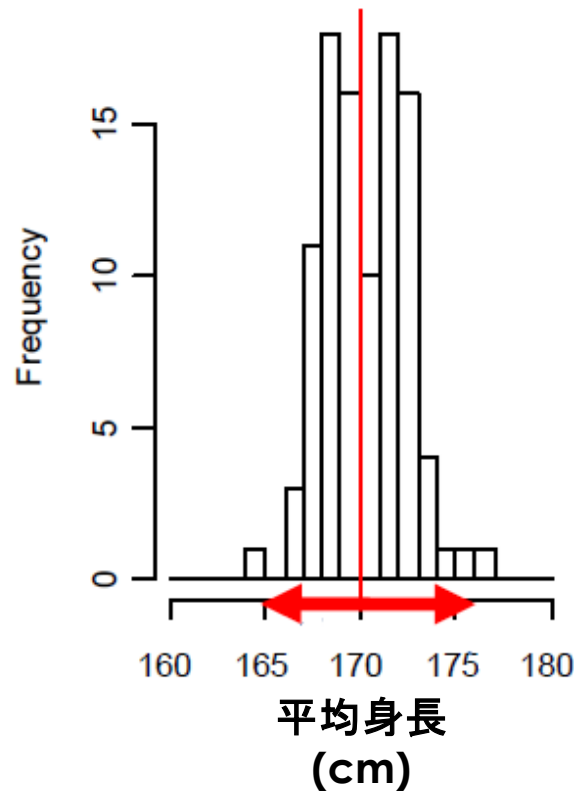
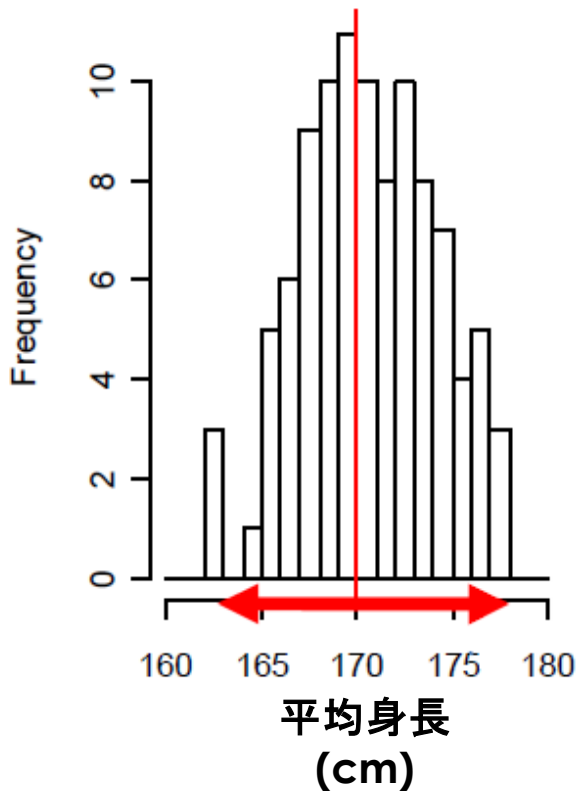
標本サイズと誤差

「標本を無作為に選んで平均値を計算」を100回繰り返す

標本数 5

標本数 10

標本数 100

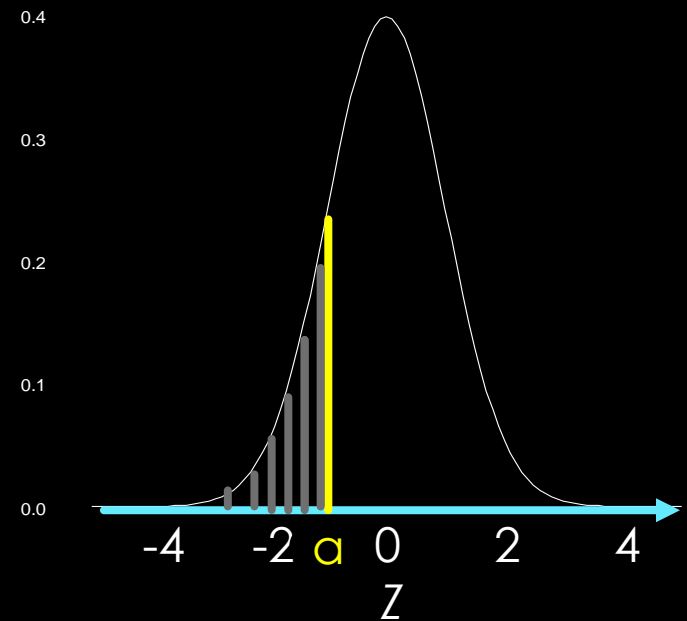


正規分布 NORMAL DISTRIBUTION

$Z < \alpha$ となる確率が $-\infty$ から α までの曲線下の面積で得られる

- 統計解析で最もよく出てくる確率分布
- 特徴：
 - 全ての実数値の確率を考えられる
 - 平均を中心に左右対称な釣り鐘型の分布
 - 平均 μ と分散 σ^2 によって記述される：
実数 X を確率変数とすると、 $X \sim N(\mu, \sigma^2)$
- どんな分布の母集団から取ったデータでも、 N が十分に大きいとき、その標本平均は近似的に正規分布に従う（中心極限定理）

$N(0, 1)$ を標準正規分布と呼ぶ

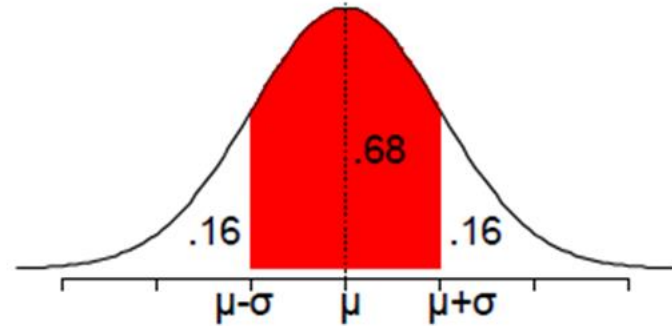


$\bar{X} = (X_1 + X_2 + \dots + X_n)/n$ とすると

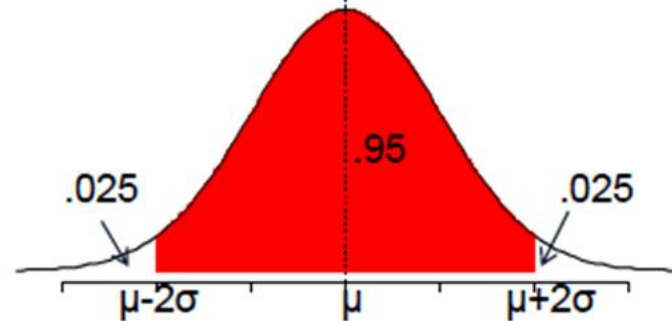
$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{\sqrt{n}}\right)$$

正規分布の性質

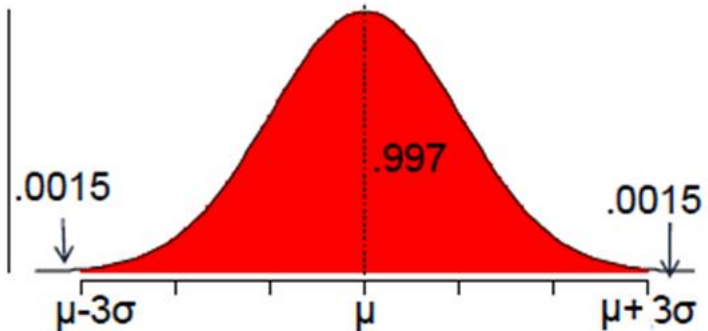
- 平均から1標準偏差内に約70%



- 平均から2標準偏差内に約95%



- 平均から3標準偏差内に約99%



例 正規分布に従う母集団の平均値の検定

- 血清総蛋白量 (TP) は母標準偏差 $\sigma = 0.4$ の正規分布に従うことが解っているとす。無作為標本 $N = 10$ 人の TP を測定し、標本平均 = 6.5 となった。母集団平均 μ が 7.0 と異なるかどうか検討せよ。

$H_0 : \mu = 7.0 (= \mu_0)$ vs. $H_1 : \mu \neq 7.0$ の両側検定を行う

統計量

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

データから Z を計算すると $Z \sim N(0,1)$
これを用いて、P値を計算

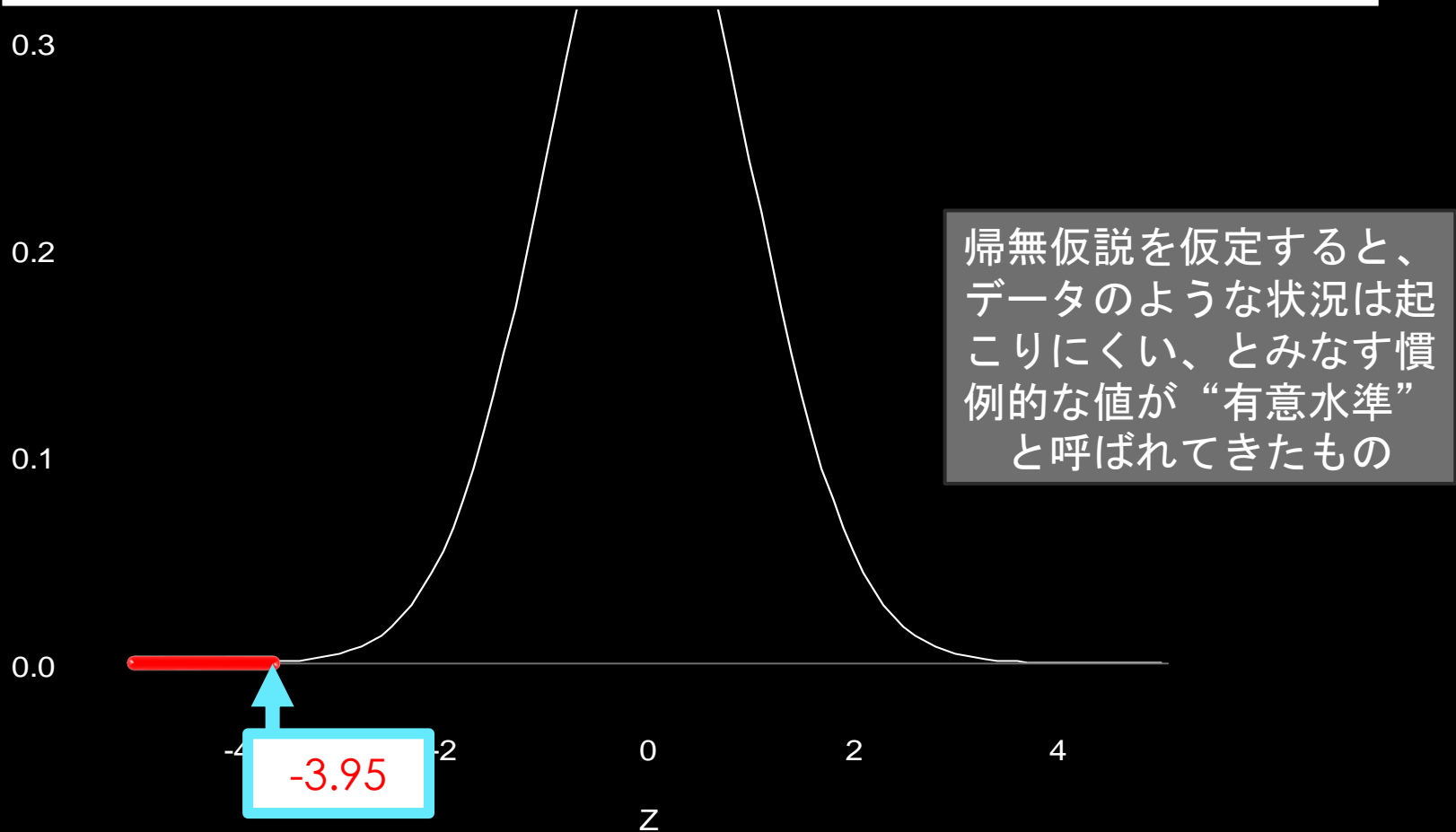
P値

H_0 の下でデータもしくはそれよりもっと極端なデータが得られる確率

両側検定： H_0 の下で Z は標準正規分布 $N(0,1)$

帰無仮説の下で、 p 値は非常に小さく ($p < 0.0001$)、
血清TPは7.0と異なる。

さらに、データは7.0よりも小さいことを示唆 $\rightarrow \mu$ は7.0よりも小さい

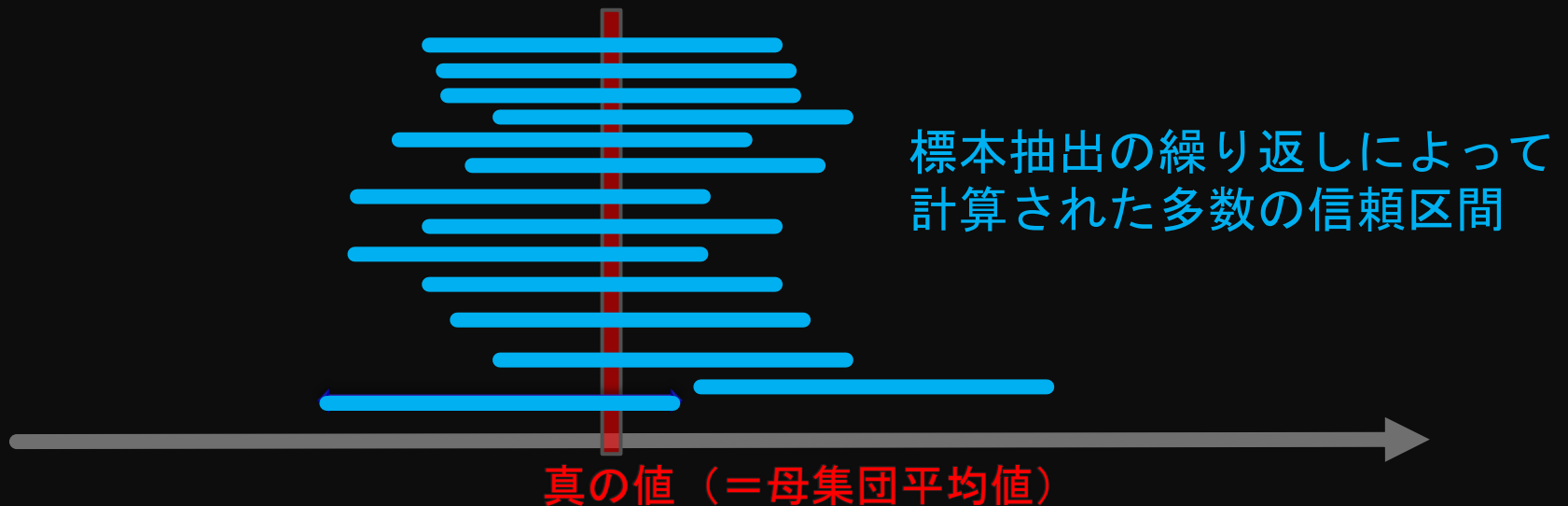


例の95%信頼区間： [6.252, 6.748] (<7)

学会発表の抄録等では必ず報告！

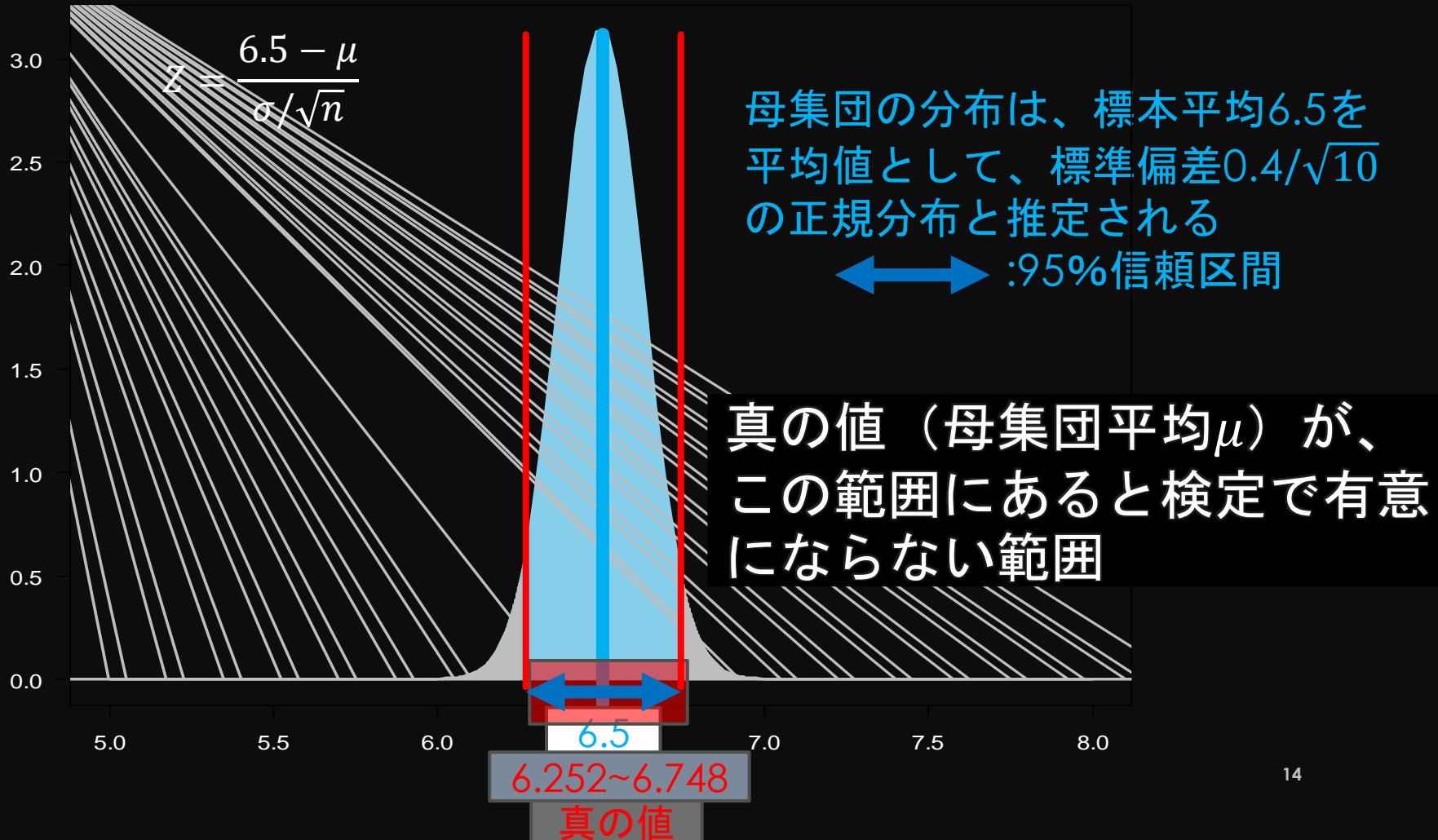
95%信頼区間とは？

-真の値があったとして、「標本抽出→区間推定」を繰り返し
行い、多数の信頼区間を計算した場合に、そのうち95%の
信頼区間が真の値を含むような区間



95%信頼区間とは？

5%有意水準で帰無仮説とデータが矛盾しない範囲



5%有意水準での検定で $P < 0.05$



95%信頼区間が対象とする値を含まない

95%信頼区間とは？

- 真の値(母集団平均)がこの区間に含まれていれば、5%有意水準の検定で有意にならない真の値の範囲

すなわち、5%の有意水準で、帰無仮説とデータが矛盾しない真の値の区間

- 真の値の推定値1つ、ではなく研究データの精度を反映した区間推定値として報告する

母集団についての推測

- 成立するためには、偏りなく標本が取られていることが大前提
- 臨床研究では、現実にデータ収集では母集団が条件付きになることが多い（個人の印象です）
- 実際、実験データと異なり、**観察研究データ**◆では、解釈が条件付きになることも多く、「どこまでその研究の結果が一般化できるか」をlimitationとして認めた上で発表することも多い
- ◆ ある介入の効果（または曝露の影響）を評価したいが、倫理的な制約等から研究対象者のランダム化（無作為に対象者を選ぶこと）ができない研究データ

事例) データの解釈に絡んで

同じ病院の2人の医師A,Bの手術の成績を検索し、表1を得た。ともに100例の手術を行ったが、良好な予後を示したのはそれぞれ48人、52人であった。

表1	良好	悪化	計
A	48	52	100
B	52	48	100
計	100	100	200

良好な予後を示した患者の割合は
Aでは48%、Bでは52%
Bの方が優れているといえる?

表1	良好	悪化	計
A	48	52	100
B	52	48	100
計	100	100	200

- 表1の例を患者の年齢によって2つの表に分けた。

表2 (年齢 > 60)	良好	悪化	計
A	32 (40%)	48	80
B	4 (20%)	16	20
計	36	64	100

表3 (年齢 ≤ 60)	良好	悪化	計
A	16 (80%)	4	20
B	48 (60%)	32	80
計	64	36	100

表2でも表3でもAの方が20%も良好な人が多く、よい成績なのに、それを合わせた表1ではBの方が成績がよいように見える。

内部妥当性：比較が適切かどうか

例) 放影研の集団調査：放射線被曝によってがん等の疾患や死亡が増えているか

- 被曝した人、と被曝していない人で比較

比較は妥当か？

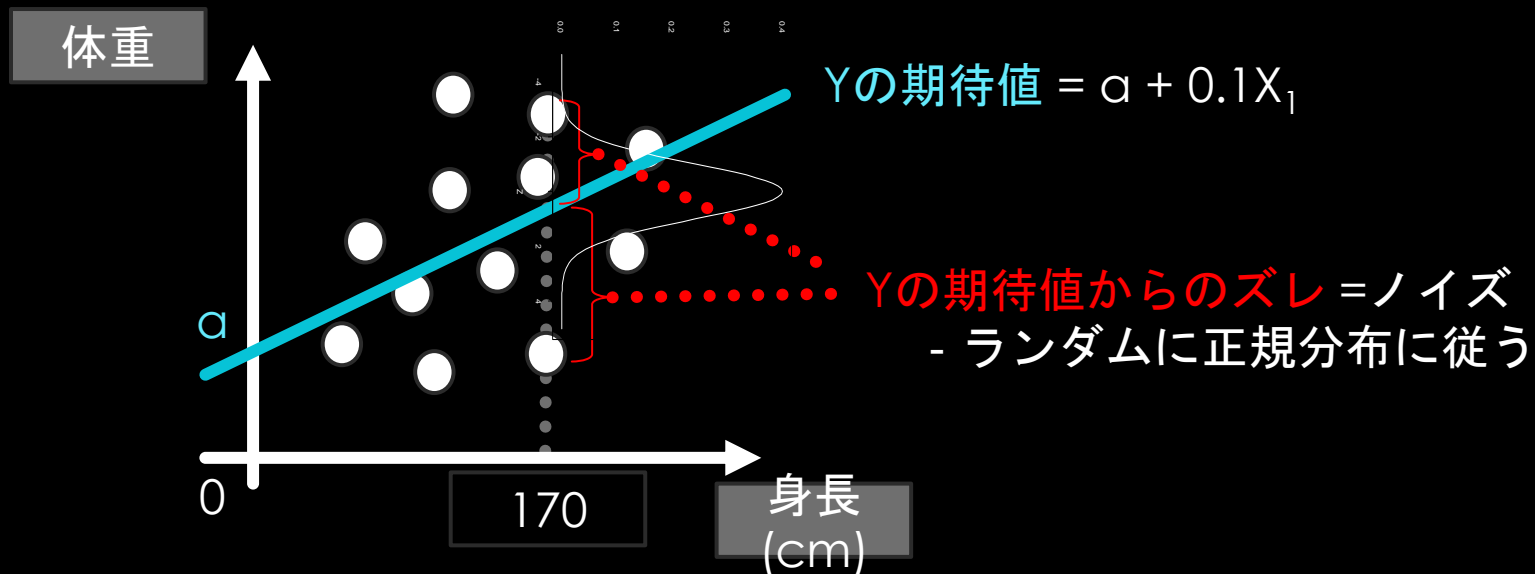
被曝時の条件が偏っていないか
性、年齢など

→ できるだけ多くの情報

統計学的方法（多変量解析など）により、数学的に対象としている**曝露以外の要因の影響をそろえて、**曝露（介入）の影響を推定

統計モデル

- 生物統計で用いるモデルは、多くが線形モデル
- 対象者のYにXが関連するか調べたい場合、
$$Y = a + cX_1 + \text{ノイズ}$$
という式を考え、集団におけるYの平均値のXとの関連を評価
- 例えば、Y=体重、X=身長とすると、



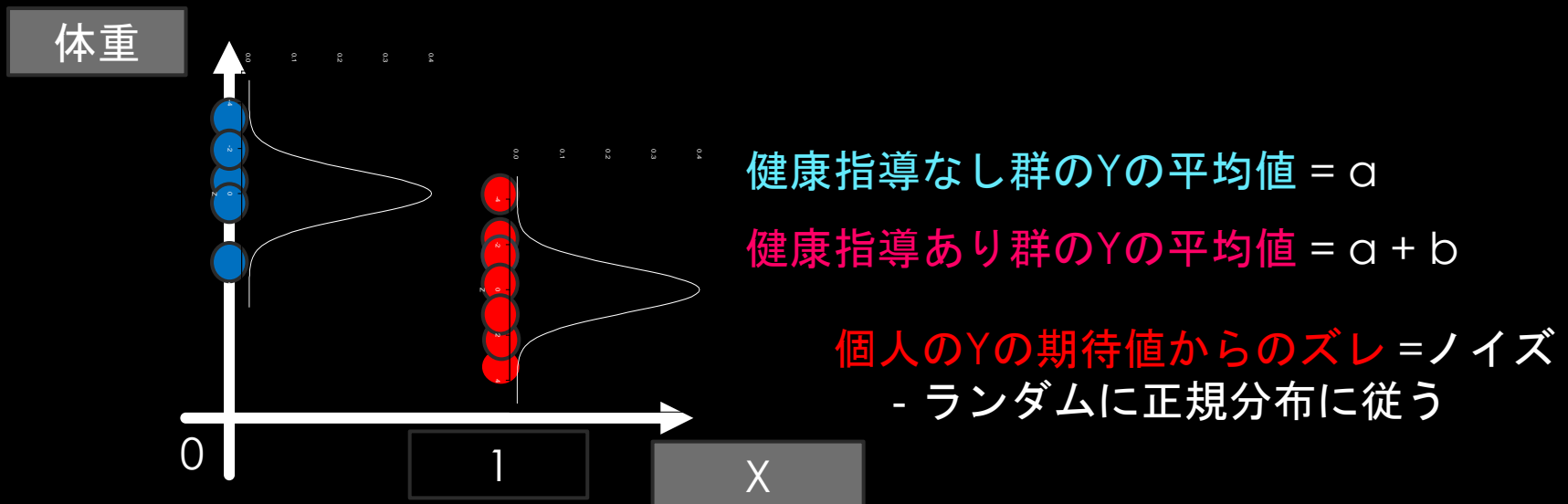
統計モデル（2）

- Xとして連続量ではなく、対照群にX=0、介入群にX=1という値を付けた場合

$$Y = a + bX_0 + \text{ノイズ}$$

として、集団におけるYの平均値のXとの関連を評価

- 例えば、Y=3か月後の体重、X=健康指導なし or ありとする

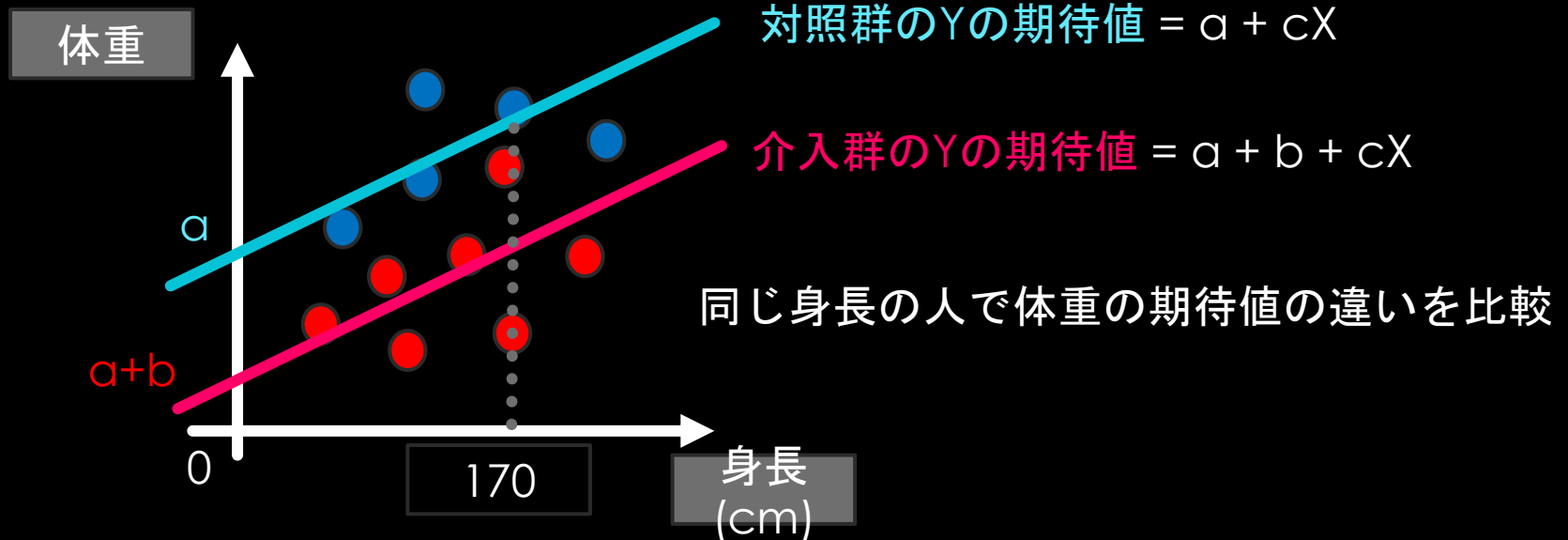


統計モデル (3)

- 介入の効果を身長の影響を調整して評価
 $X_0=0$ を介入なし、 $X_0=1$ を介入あり、 X_1 を身長として

$$Y = a + bX_0 + cX_1 + \text{ノイズ}$$

- 例えば、 $Y=3$ か月後の体重、 $X_0=0$ を健康指導なし、 $X_0=1$ を健康指導あり、 X_1 と身長すると、



まとめ

- 統計学は、背景にある母集団についてものを言うため、すなわち、不確実性を考慮したうえで、再現性を伴った科学として結果をしめすためのもの
- 観察研究データにおいては、研究対象者に偏りがあると介入の効果を誤って判断してしまう状況が起こり得る
- 学会発表は、最新の研究成果等の情報を得たり、データ分析において考えるべき要因についてコメントや提案など助言をもらい、研究内容の改善の機会にもなる
- 観察研究を想定して話をしたが、研究計画の段階から生物統計学者に相談していただくことが研究効率向上や統計解析の単純化につながる