

統計学の基礎理論①

仮説検定の考え方

東京科学大学 ILA国府台
中林 潤

JH人材育成課 バイオインフォマティクソン育成講座 ②

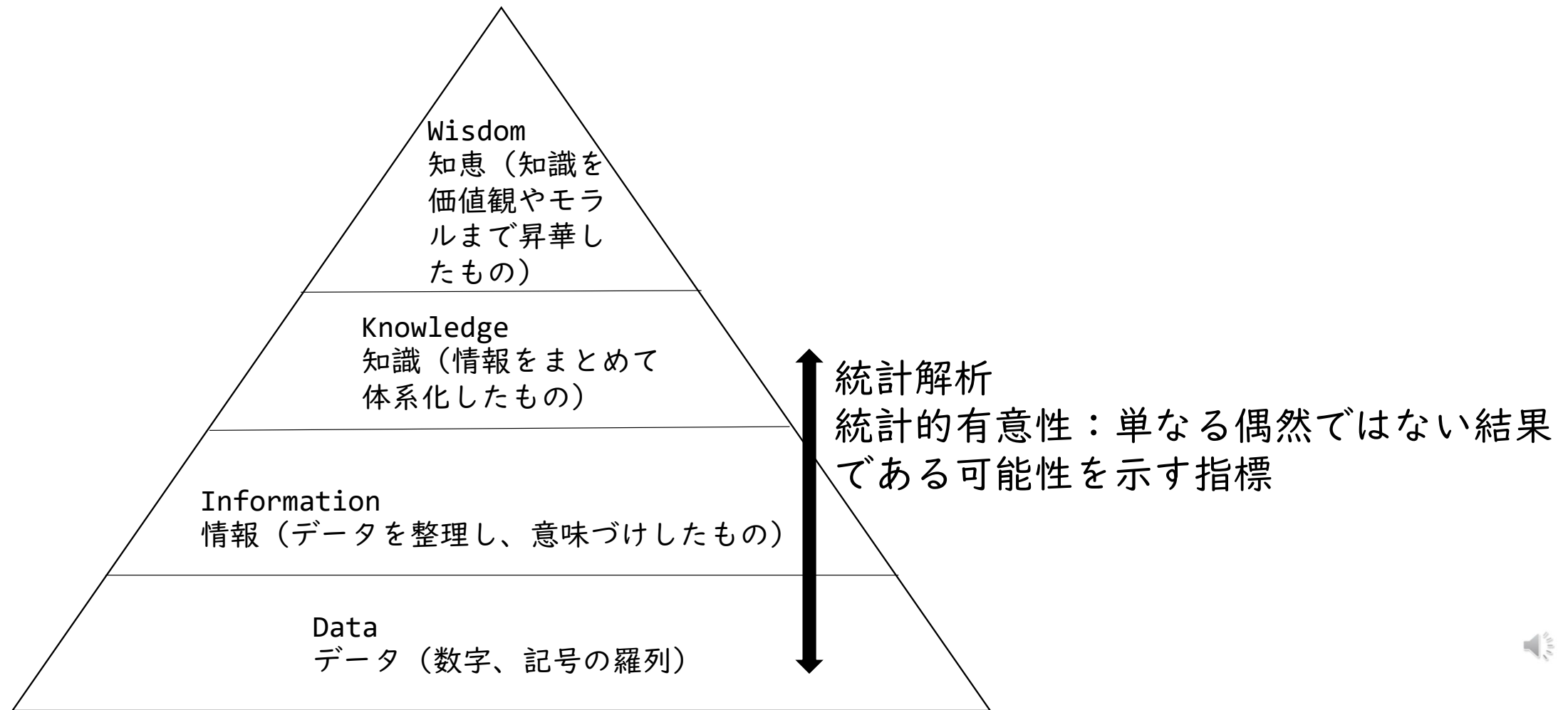


内容

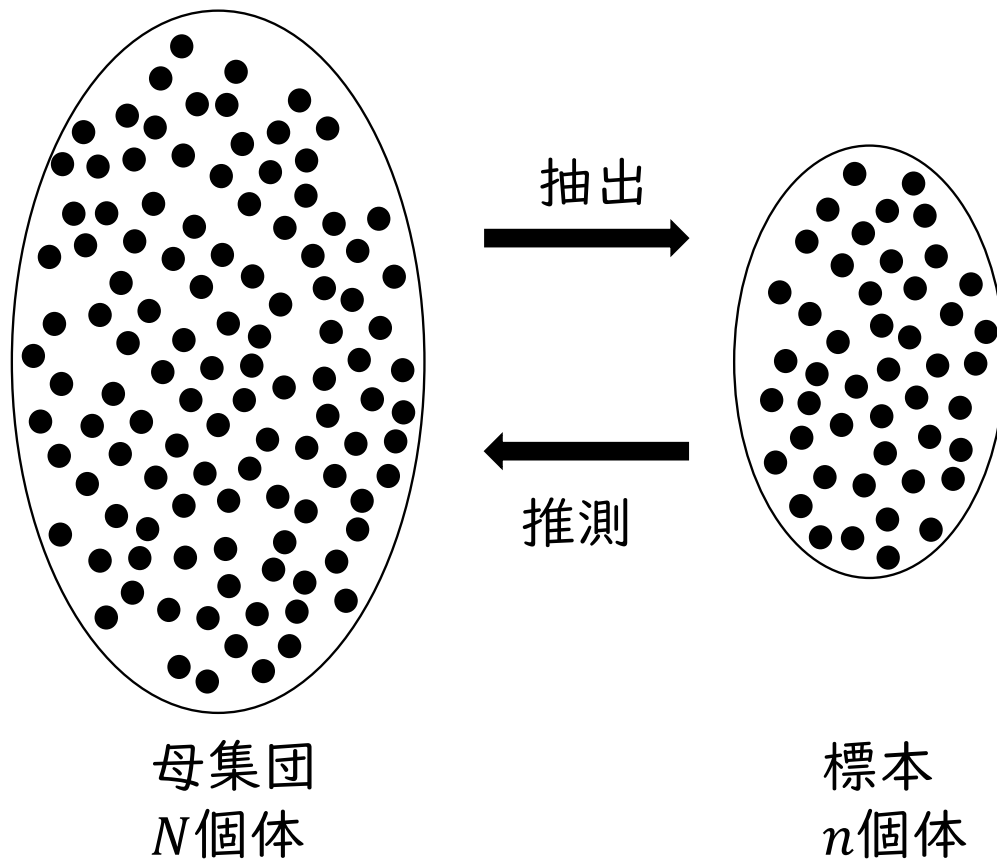
- 有意水準仮説検定の考え方
- 有意水準仮説検定の問題点



データ処理のDIKWモデル



母集団と標本



- 母集団について知りたい。
- 標本から母集団について推測する。
- 母集団と標本が同じ特性を持つことが重要である。
- 母集団をいくつかの抽出単位に分け、全ての抽出単位が抽出される確率を等しくする。

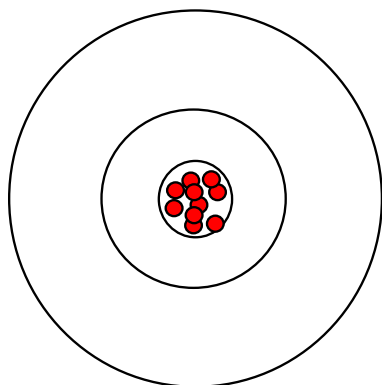
↓
無作為抽出



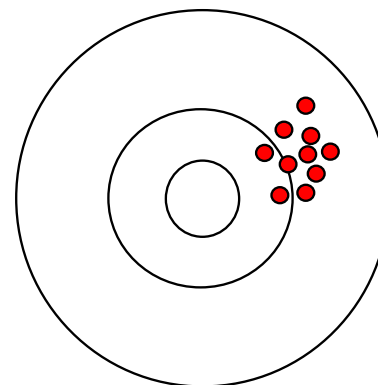
分散とバイアス

分散：ばらつき
バイアス：偏り

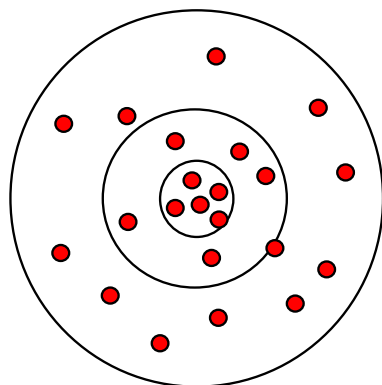
分散：小
バイアス：小



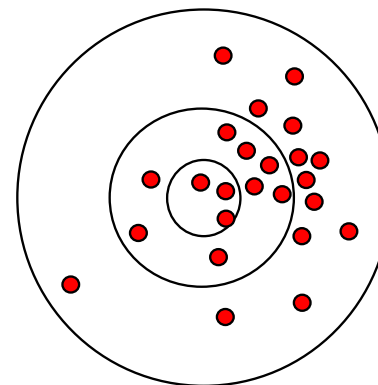
分散：小
バイアス：大



分散：大
バイアス：小



分散：大
バイアス：大



フィッシャーの3原則

- 反復 (replication)
同じ条件で2回以上実験（評価）を行う
偶然誤差（たまたま起こる誤差）の大きさを評価できる
- 無作為化 (randomization)
実験の場所や時間、処理方法などをランダムに割り振る
結果に影響を与える要因の偏りを小さくする
- 局所管理 (local control)
バックグラウンドが均一になるよう管理する
系統誤差（処理による誤差）を小さくする



基本統計量

- 平均値
- 最大値：データ値の中で最も大きい値
- 最小値：データ値の中で最も小さい値
- 中央値：データの中央の値、データ数が偶数の場合、2つの値の平均となる
- 最頻値：データ値の中で最も頻繁に出てくる値



平均値と分散

- 平均値

全てのデータ値の合計をサンプル数で割った値

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$$

- 標本分散

データ値のばらつきを表す指標

$$s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2$$

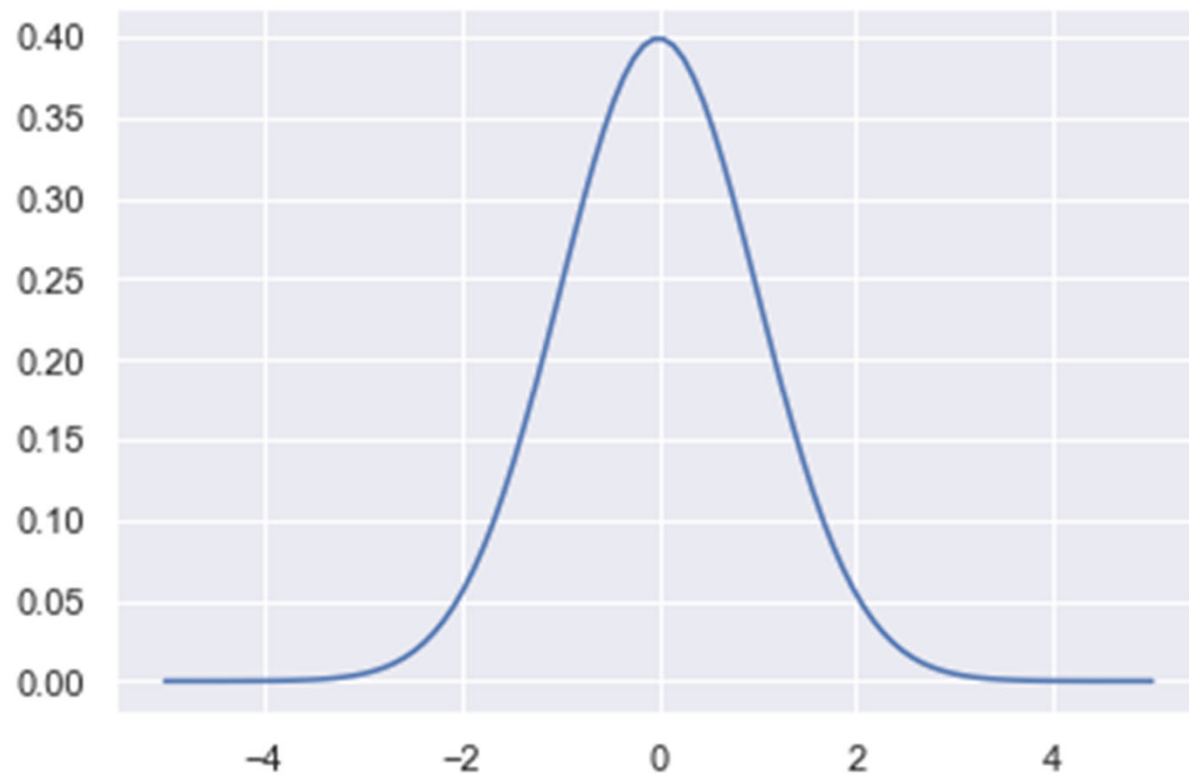
- 不偏分散

母集団の分散を推定した値

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$$



正規分布



- 平均値の頻度が最も高く、外れるに従って頻度が減少する。
- 平均値、最頻値、中央値が一致している。
- 左右対称の釣り鐘型。
- ガウスの公理を満たす。

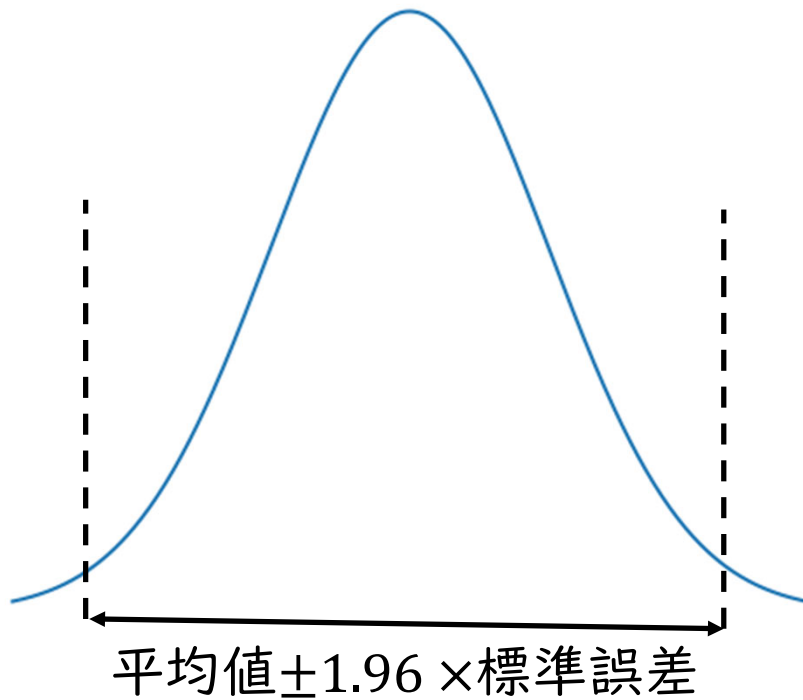


ガウスの公理

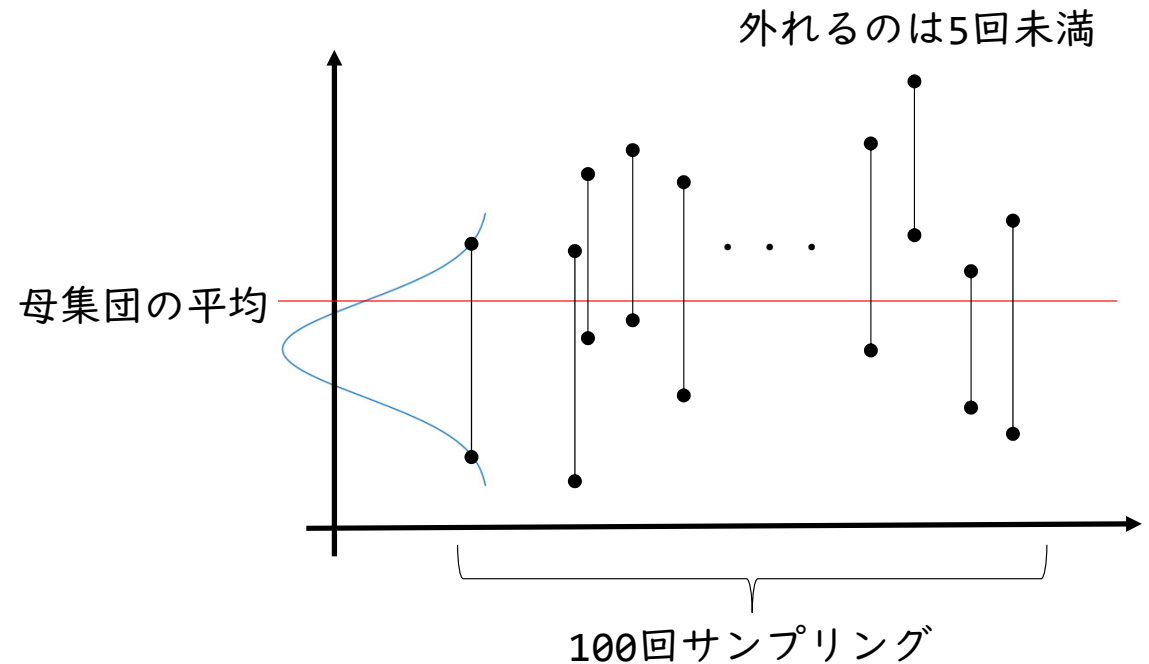
- 大きさの等しい正と負の誤差は同じ確率で起こる。
左右対称
- 小さな誤差は大きな誤差よりも起こりやすい。
平均値の近傍の値が起こりやすい
- ある限界値より大きい誤差は実際上起こらない。
平均から大きく外れた値は非常に起こりにくい



平均値の95%信頼区間

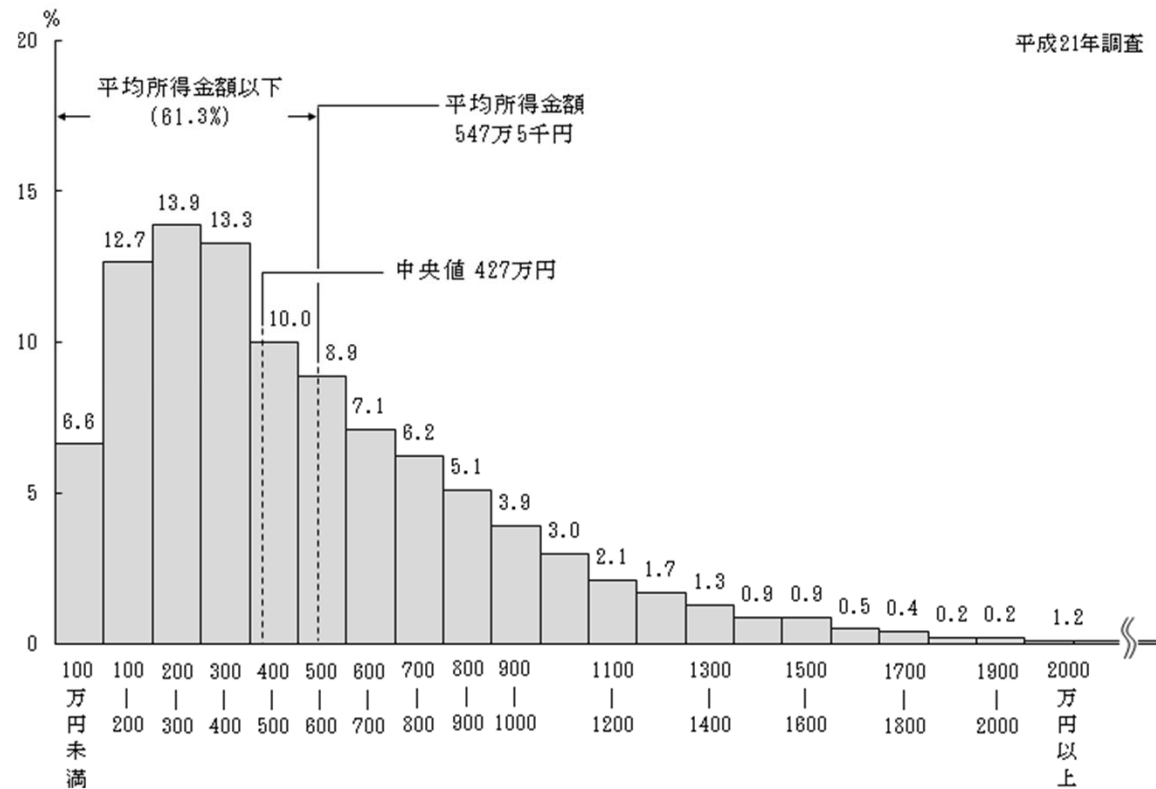


$$\text{標準誤差} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2}}{\sqrt{N}} = \text{標準偏差} / \sqrt{N}$$



母集団の平均値がこの範囲を外れるのは5%未満

正規分布ではない分布



厚生労働省

平成21年国民生活基礎調査

<https://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa09/2-2.html>

仮説検定の考え方

- 母集団について仮定された命題を、標本に基づいて検証する。
- 帰無仮説：棄却されることが期待される仮説
- 対立仮説：帰無仮説と対立する仮説
- 帰無仮説を棄却し、対立仮説を採用する。



有意水準仮説検定の手順

1. 帰無仮説を立てる。
2. 統計検定量からp値を求める。
3. p値が有意水準よりも低ければ、帰無仮説を棄却し対立仮説を採用する。



有意性検定

- 検定統計量：標本の値から計算される分布が既知の値
- p値：帰無仮説が正しいという仮定の下で観測値が得られる確率
- 有意水準：帰無仮説を棄却するかどうか判断する基準。1%か5%を使うことが多い。
- p値を求め、有意水準よりも低ければ帰無仮説を棄却し対立仮説を支持する。



第1種、第2種の過誤

		真実	
		帰無仮説が正しい	帰無仮説が誤り
検定	帰無仮説を棄却しない	正しい	第2種の過誤
	帰無仮説を棄却する	第1種の過誤	正しい

第1種の過誤：差がないものを、差があると判定してしまう。
第2種の過誤：差があるものを、差がないと判定してしまう。



再現性とQRPs

- 再現性の問題

ある研究者によって統計的に有意の差があるとされた結果が、他の研究者には同程度の有意の差と判定されないこと。

- QRPs (Questionable Research Practice)

正しい結論にたどり着く確率を下げるような研究手法



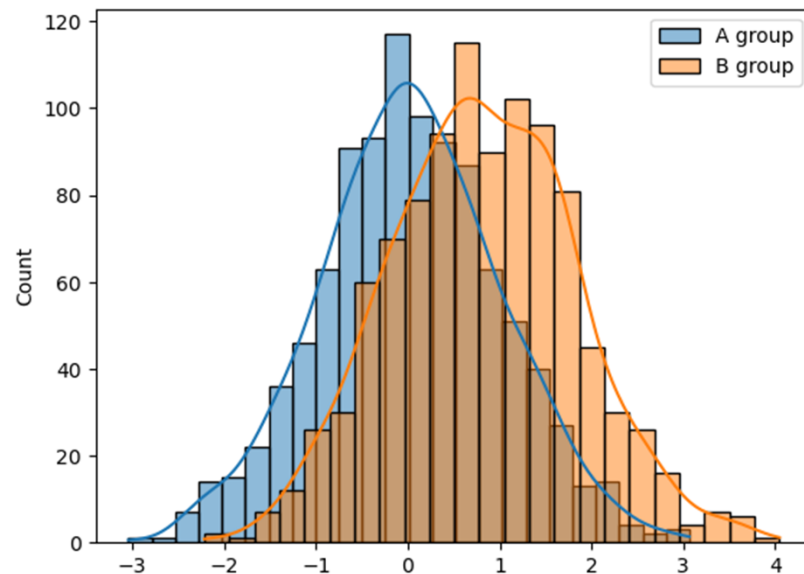
Statement on statistical significance and p-value

1. P-values can indicate how incompatible the data are with a specified statistical model.
2. P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
6. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.



ビッグデータと帰無仮説

- データの精度が上がる→帰無仮説が棄却される確率は上がる。
- 些末な対立仮説が支持される。



帰無仮説：2グループは平均の等しい母集団から抽出された。
対立仮説：2グループの平均は異なる。

バイオインフォマティクスで扱うデータの特徴

- 大容量である。
- 分散が大きい。
- 離散的な値を含む。0を多く含む。
- 測定を反復することが困難な場合がある。

適切な統計解析方法を選択することが重要



仮説検定の手順

1. 帰無仮説と対立仮説を立てる。
2. 検定統計量を計算する。
3. 有意水準と比較し、帰無仮説を棄却すべきかどうか確認する。
4. 帰無仮説を棄却→対立仮説を採用する。

