仮説検定の概念

仮説検定による検証

- 仮説検定
  - ある仮説をたて，これを確率の概念を用いて検討する方法
  - 仮説の例

> 20歳の女性13人について，入浴を済ませた2分後の
> 最高血圧と安静時の最高血圧を測定した。
> これらの血圧値の間に差があるかを検証したい

- データ
  - (入浴を済ませた2分後の最高血圧値)－(安静時の最高血圧値)
  - 測定値

20, 4, 10, 2, 10, -10, 4, 24, 10, -6, 14, 10, 16

データを眺める

- データ：入浴後－安静時

$$\underbrace{20,\ 4,\ 10,\ 2,\ 10}_{\text{正}},\ -10,\ \underbrace{4,\ 24,\ 10}_{\text{正}},\ -6,\ \underbrace{14,\ 10,\ 16}_{\text{正}}$$

負　負

- 考察
  - 正の値が多く，入浴後の血圧値の方が高いと考えられる
- 問題
  - この差は<u>本質的な差</u>から生じたものか？
    それとも<u>偶然に生じたもの</u>か？

仮説検定

- 仮説検定の概念
  - 「血圧値に差はない」という仮説（帰無仮説）をたて，
    この仮説のもとで，今回のデータが生じる確率を考える
  - この確率が著しく小さい場合
    - 帰無仮説のもとでは
      「めったに起こらないことが起きた」と考える
    - 帰無仮説を捨てる（棄却する）ことで
      「血圧値には差がある」ことを主張する
  - この確率が著しく小さいと言えない場合
    - 結論を保留する

仮説検定の手順

- 手順１：帰無仮説をたてる
- 手順２：標本(データ)を無作為抽出する
- 手順３：帰無仮説を真としたときに，そのような標本が出現する確率を調べる
- 手順４：その確率がきわめて小さいときには帰無仮説を棄却する。確率が小さいとはいえないときは判定を保留する

仮説検定の例

安静時と入浴後の血圧値の例で考える

- 帰無仮説 H_0
  - 入浴をすませた2分後の最高血圧と安静時の最高血圧との間に差はない
- 真の差(未知) μ で表すと, 帰無仮説は
  - $H_0 : \mu = 0$
- 「差がある」という仮説（対立仮説）は
  - $H_1 : \mu \neq 0$

正しい判定と誤った判定

- 正しい判定
  - 帰無仮説が真のとき，帰無仮説を棄却しない
    - 差があるときに，差がないと判定しない
  - 対立仮説が真のとき，帰無仮説を棄却する
    - 差があるときに，差があると判定する
- 誤った判定
  - 帰無仮説が真であるのに，帰無仮説を棄却する（第1種の過誤）
    - 差がないときに，差があると判定してしまう誤り
  - 対立仮説が真であるのに，帰無仮説を棄却しない（第2種の過誤）
    - 差があるときに，差があるとは判定しない誤り

母集団の状態と判定との関係

判定 / 母集団の状態	帰無仮説を棄却しない	帰無仮説を棄却する
帰無仮説が真	正しい判定	誤った判定 （第1種の過誤）
対立仮説が真	誤った判定 （第2種の過誤）	正しい判定

- 有意水準：第1種の過誤をおかす確率
  - 例：有意水準5%で仮説検定する

検出力について

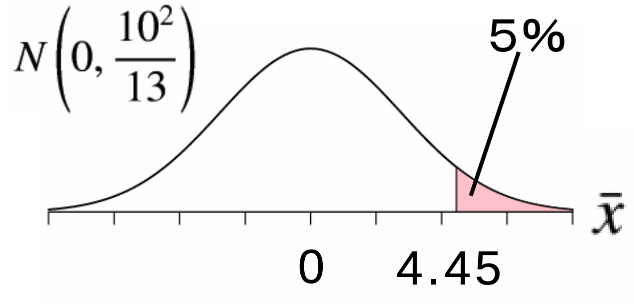
検出力

- 検出力の定義
  - 検出力 ＝ 1－（第2種の過誤の確率）
  - 対立が真のとき，帰無仮説を正しく棄却する確率
- 第2種の過誤
  - 対立仮説が真のとき，帰無仮説を棄却しない誤り

例：検出力の求め方

- 標本数13，標準偏差10の正規分布で考える
  - 仮説の設定

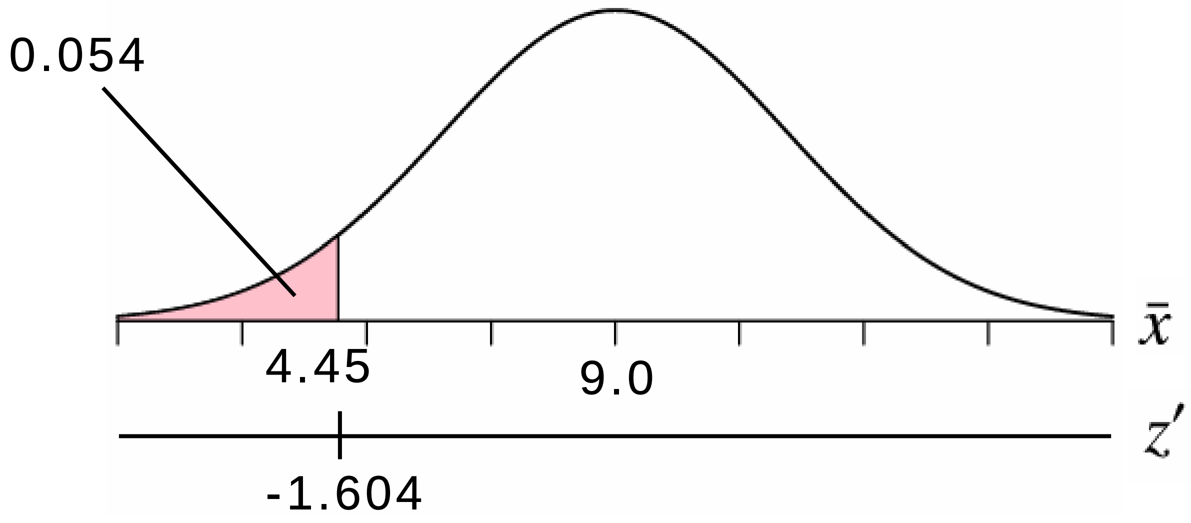
$$H_0 : \mu = 0, \quad H_1 : \mu = 9.0$$



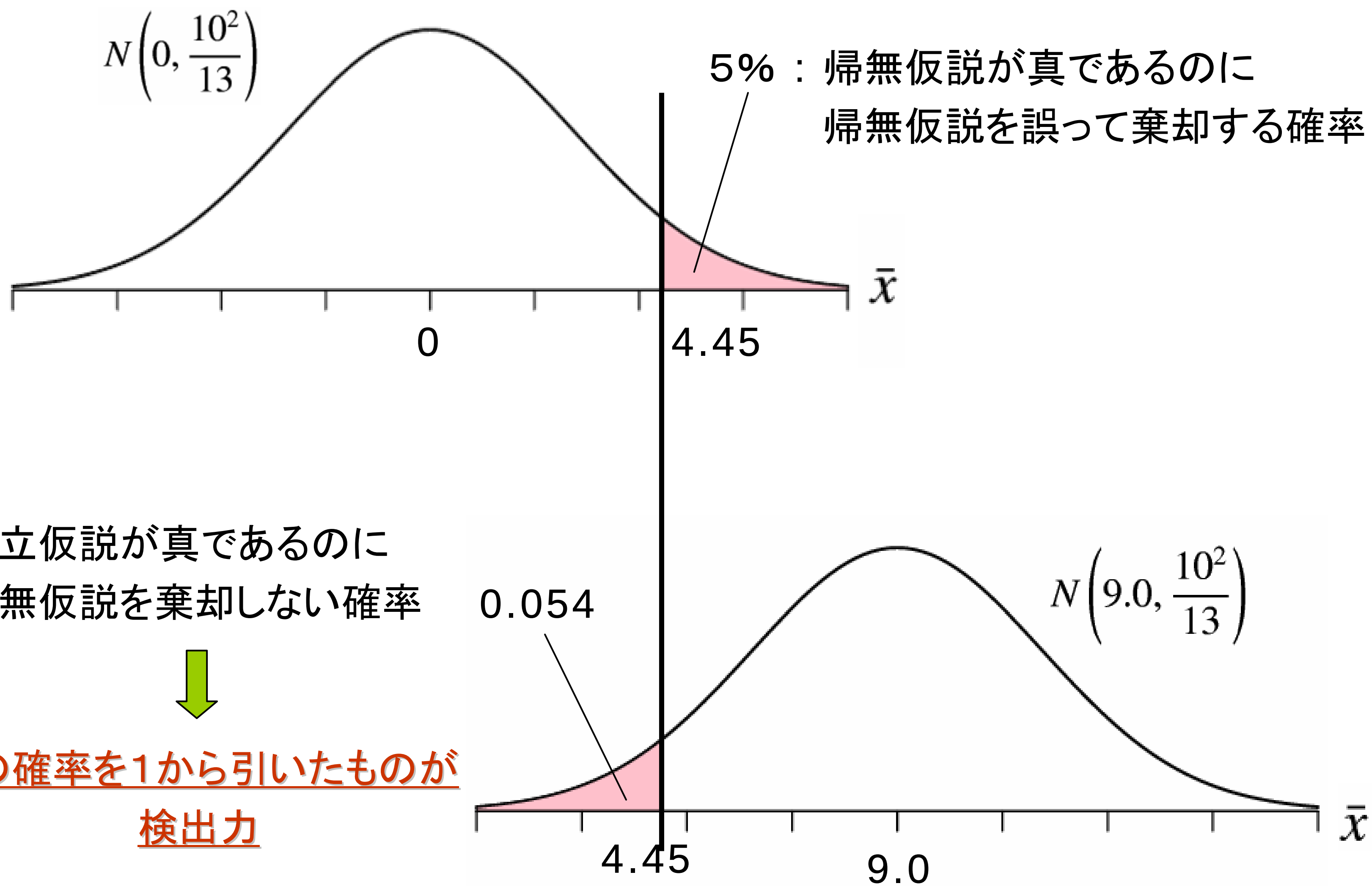
- まず帰無仮説のもとで考える
  - 標準正規分布の上側5%点 ＝ 1.64
  - $\bar{x} = 1.64 \times \dfrac{10}{\sqrt{13}} = 4.55$
  - 標本平均が4.55よりも大きいとき，帰無仮説を棄却する

例：検出力の求め方

- 次に，対立仮説のもとで考える
  - 対立仮説 $H_1 : \mu = 9.0$
  - $\bar{x} = 4.55$ のとき，$z' = \dfrac{4.55 - 9.0}{\dfrac{10}{\sqrt{13}}} = -1.604$



検出力



対立仮説が真であるのに
帰無仮説を棄却しない確率

<u>この確率を1から引いたものが</u>
<u>検出力</u>

検出力

- 今回の例では
  - 検出力 = 1 – 0.054 = 0.946 となる
- 検出力について
  - 母平均が帰無仮説の値から離れるほど検出力は高くなる
  - 標本数が大きいほど，標本平均の分散は小さくなる
    ゆえに，検出力は高くなる