統計学入門
～統計的仮説検定の問題設定～

2025年度1学期：　　月曜2限

担当教員：　　　　　石垣　司

統計的仮説検定って？

母集団の分布の母数に関する仮説の妥当性を標本から検証する方法

- **例**1: サイコロを600回投げたとき1の目が120回も出た。この回数は偶然の範囲内か？
- **例**2: 2020年8月に実施されたTOEFL ITPテストを自主的に受検した東北大学生のスコアの平均値は500点であった。ここで，未受検の学生30名を無作為抽出して同じテストを受検させたときのスコアの平均値が480点であったとする。このスコアの違いは偶然の範囲内といえるか？
- **例**3: 2024年度の東北大学の入学者の割合（入学者数 / 志願者数）は宮城県出身者は26.6%，北海道出身者は42.4%であった。この入学者の割合の違いは偶然の範囲内か？

数学的内容は既学習+α。 問題設定を理解することが重要

仮説検定のロジック（帰無仮説 vs 対立仮説）

帰無仮説：“仮説は正しくない”を表現する仮説

対立仮説：“仮説は正しい”を表現する仮説

- 例2：自ら受検した学生と未受検の学生の母集団のTOEFL ITPの平均スコアに差はあるか？

  帰無仮説：“平均スコアに差はない”

  対立仮説：“平均スコアに差はある”

仮説検定のロジック

- 仮説に対する標本が得られていて，帰無仮説が正しいという設定の下でその標本が生じる確率を計算すると，その確率は有意水準よりも小さい
  ⇒ ゆえに，対立仮説を支持する

検定統計量

仮説検定で使用する統計量

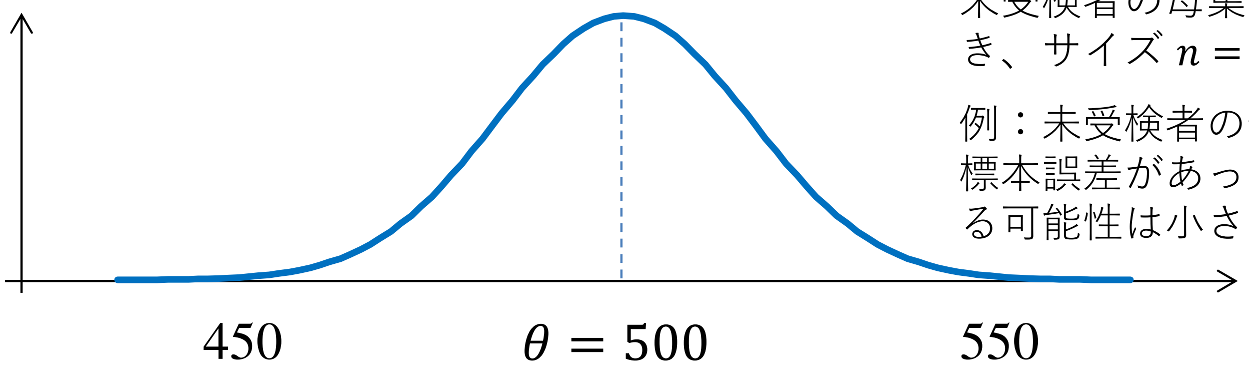
- 帰無仮説を表現する検定統計量の分布と実際の観測値とのズレの大きさを測るために利用

- 例2：TOEFL ITPの平均スコアは異なるか？
  ⇒ 検定統計量に標本平均 $\bar{X}$ を用いる

  帰無仮説：“平均スコアに差はない”

  対立仮説：“平均スコアに差はある”

帰無仮説が成立すると仮定したときの検定統計量 $\bar{X}$ の分布の例



未受検者の母集団の母平均も 500 であると仮定したとき、サイズ $n = 30$ の標本が従う標本平均 $\bar{X}$ の分布

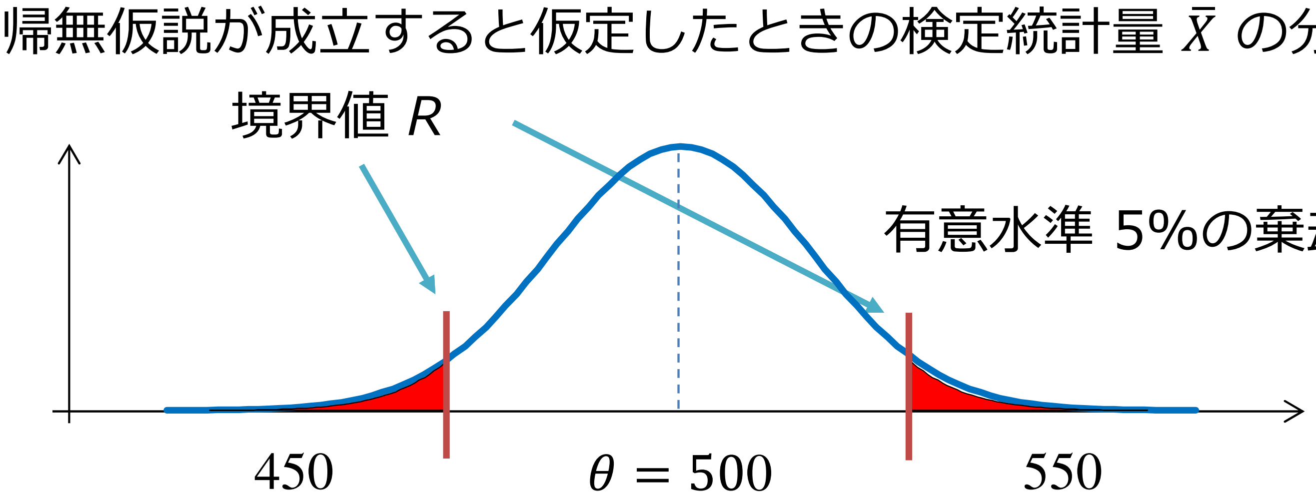
例：未受検者の母集団の母平均も 500 であるならば、標本誤差があったとしても平均スコアが400や600となる可能性は小さい

棄却域と有意水準

棄却域：帰無仮説から離れたある領域（範囲）

有意水準：ある事象が起こる確率が偶然とは考えにくい（有意である）と判断する基準となる確率（デジタル大辞泉）

– 検定統計量の分布（確率密度関数）の棄却域での面積
– 通常は 0.05 (5%) や 0.01 (1%)を設定
– **境界値**：棄却域の境界線の値



帰無仮説の棄却と解釈

帰無仮説の「棄却」

- “帰無仮説は統計的に認められない”という結果
  対立仮説が採択される, ともいう
- 例:「5%有意水準で棄却」,「5%水準で有意差あり」

帰無仮説が棄却された場合

⇒ 対立仮説が統計的に認められる

帰無仮説が棄却できない場合

⇒ 対立仮説に関して何もわからない

※「対立仮説が棄却される」,「帰無仮説が採択される」ではないことに注意

両側検定と片側検定 #1

両側検定：棄却域を分布の両端に定める

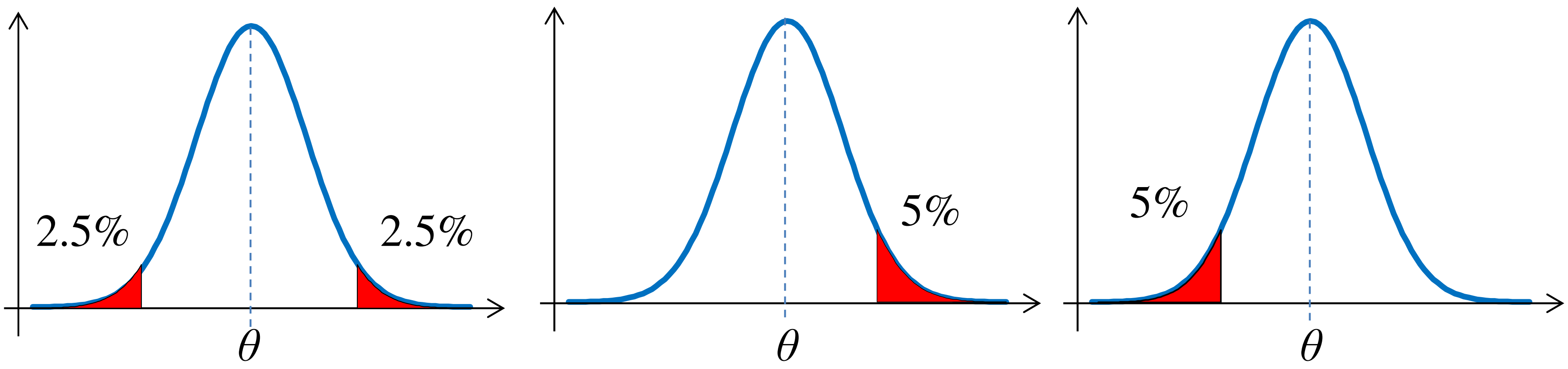
– 例：平均スコアは異なるか？

片側検定：棄却域を分布の片端に定める

– 例：未受検者の平均スコアは高いか？（右側検定）

– 例：未受検者の平均スコアは低いか？（左側検定）

有意水準5%



両側検定　　片側検定（右側検定）　　片側検定（左側検定）

両側検定と片側検定 #2

帰無仮説 H_0 と対立仮説 H_1 の立て方

θ は母集団の母数の値（例 / 未受検者の本当の平均スコア）

θ_0 は帰無仮説での母数の値（例：$\theta_0 = 500$）

- 両側検定　$H_0 : \theta = \theta_0,\quad H_1 : \theta \neq \theta_0$
- 右側検定　$H_0 : \theta = \theta_0 \text{ or } \theta \leq \theta_0,\quad H_1 : \theta > \theta_0$
- 左側検定　$H_0 : \theta = \theta_0 \text{ or } \theta \geq \theta_0,\quad H_1 : \theta < \theta_0$

- 帰無仮説は適切な検定統計量を定義できるように定める

  例：“$H_0 : \theta \neq \theta_0$” では, 平均の検定統計量はうまく定義できない

- 帰無仮説に不等号を用いる場合でも, 実際には等号が成り立つ下での検定統計量を用いる

検定の手順

1. 仮説を立てる
2. 検定統計量を定める
3. 有意水準 α（0.05や0.01など）を定める
4. 標本から検定統計量を計算
5. 検定統計量の分布から棄却域を定める
6. 帰無仮説を棄却できるかを検定

検定の例 #1

サイコロAを600回投げたとき1の目が120回も出た。これは1の目が多く出やすいイカサマサイコロか？

1. **仮説を立てる**
   - **帰無仮説:「イカサマサイコロではない」** $H_0 : \theta = \theta_0$
   - **対立仮説:「イカサマサイコロである」** $H_1 : \theta > \theta_0$

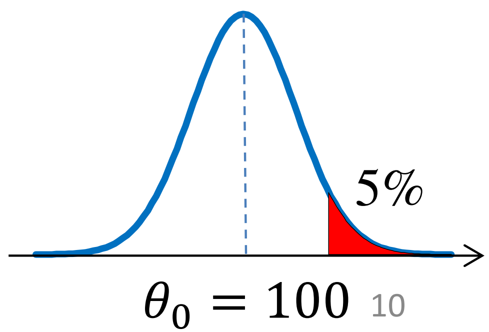
     1の目が出る回数が多すぎるという検定

     θ_0: **普通のサイコロで1の目が出る回数の理論値**

     θ: **サイコロAの真の母数から計算される1の目が出る回数**

2. **検定統計量 X を定める**
   - **1の目の出る回数の確率変数** $X \sim \text{Binominal}(n, 1/6)$
   - n **が大きいので中心極限定理より正規分布で近似**

$X \sim N(np, np(1-p))$



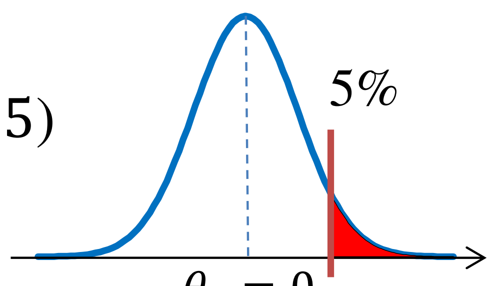
3. **有意水準を定める**
   - **5%有意水準($\alpha = 0.05$)の片側検定（右側検定）**

検定の例 #2

サイコロAを600回投げたとき1の目が120回も出た。これは1の目が多く出やすいイカサマサイコロか？

4. **標本から検定統計量を計算**
   - $E[X] = np = 100, V[X] = np(1-p) \fallingdotseq 83.3$
   - **帰無仮説の下では** 1**の目が出る回数は** $X \sim N(100, 83.3)$
   - $X = 120$ **のときの標準化統計量** $Z = (X - E[X])/\sqrt{V[X]} \fallingdotseq 2.19$

5. **検定統計量の分布から棄却域を求める**
   - $Z_{0.05} = 1.65$ **より棄却域は** $Z > 1.65$ (**境界値** $R = 1.65$)



6. **帰無仮説を棄却できるかの検定**
   - $Z > R$ **となり** Z **は棄却域の中にあるため帰無仮説は棄却される**
   - **結論　このサイコロAは有意水準**5%**でイカサマサイコロである**

検定の誤り

第1種の過誤 (Type I error, False positive)

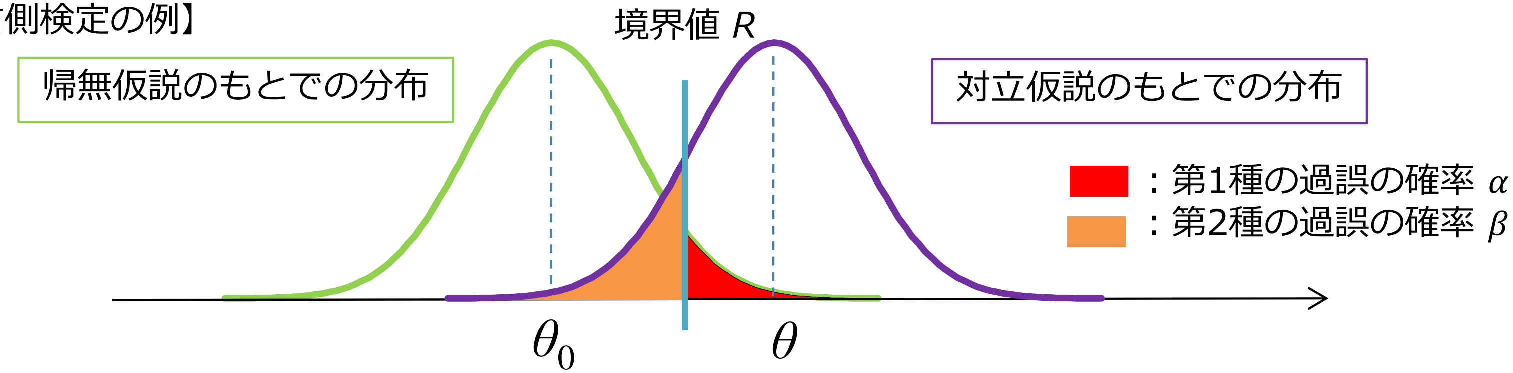
- 本当は帰無仮説を棄却できないのに, 検定では棄却してしまう間違え
- P(帰無仮説を棄却する | 帰無仮説が正しい) = α　(α は有意水準)

第2種の過誤(Type II error, False negative)

- 本当は帰無仮説を棄却できるのに, 検定では棄却できない間違え
- P(帰無仮説を棄却できない | 対立仮説が正しい) = β

		検定結果	
		帰無仮説を棄却	棄却できない
真の構造	対立仮説が正しい	正しい検定	第2種の過誤
	帰無仮説が正しい	第1種の過誤	正しい検定

【有意水準5%の右側検定の例】



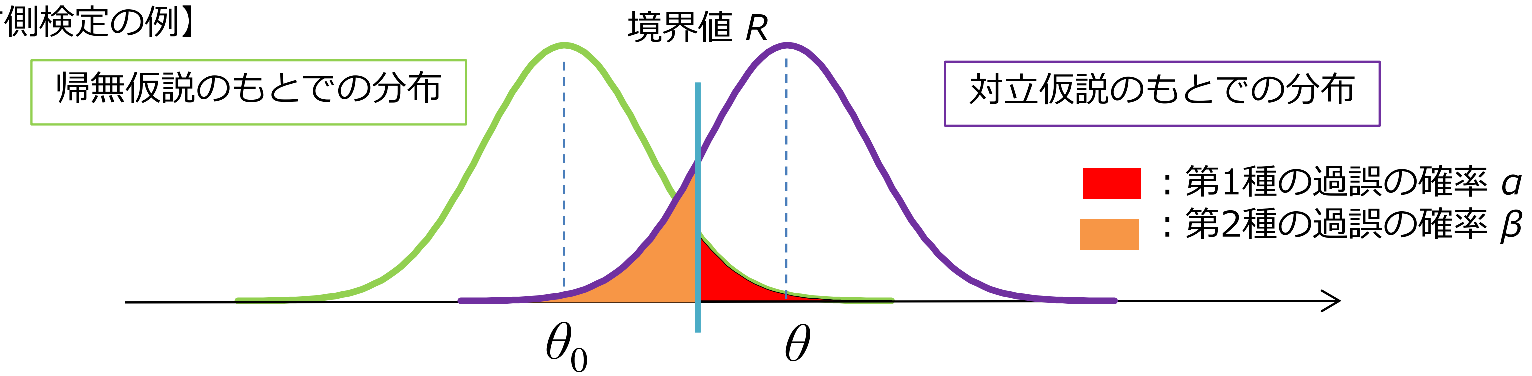
補足：検出力

対立仮説が正しいときに帰無仮説を棄却できる確率

- 検出力 = P(帰無仮説を棄却 | 対立仮説が正しい) $= 1 - \beta$
- αが小⇒βが大, αが大⇒βが小, のトレードオフ
- αを固定(例：有意水準5%など)したうえで
  βが小さな検定（検出力が大きな検定）が望ましい

		検定結果	
		帰無仮説を棄却	棄却できない
真の構造	対立仮説が正しい	正しい検定	第2種の過誤
	帰無仮説が正しい	第1種の過誤	正しい検定

【有意水準5%の右側検定の例】



P 値

観測された有意水準

- P 値が事前に設定された有意水準 α よりも小さければ, 帰無仮説は棄却できる

- 右側検定での例

  サイコロを600回投げたとき1の目が120回も出た。これは1の目が多く出やすいイカサマサイコロか?

- $X = 120$ のときの標準化統計量 $Z = (X - E[X])/\sqrt{V[X]} \fallingdotseq 2.19$

$$p \text{ value} = P(Z > 2.19 | \theta = \theta_0)$$

- 標準正規分布表から P 値 $= 0.014$
- 5%水準では有意差あり, 1%水準では有意差なし

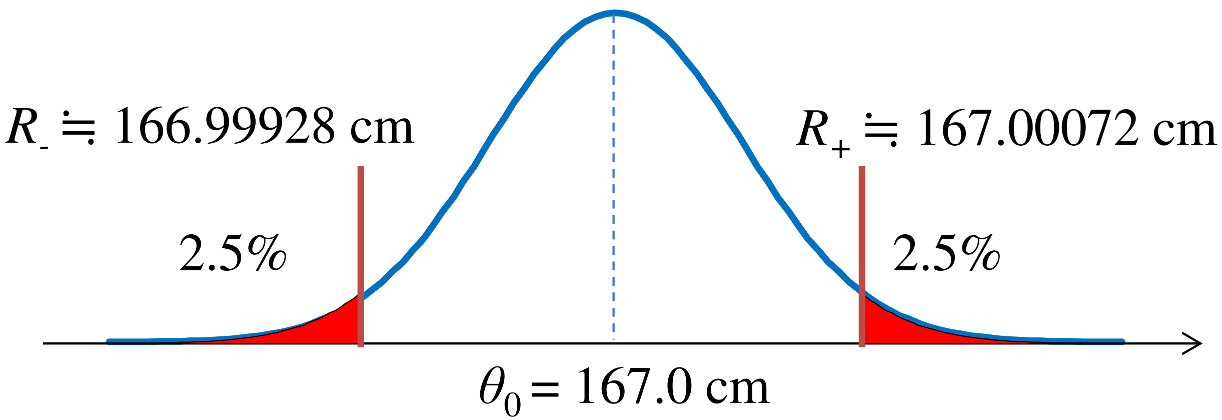
仮説検定の実用上の注意 #1

n を大きくすれば，ほとんどの帰無仮説が棄却される

- **検定の結果として有意差があった場合でも，現実の問題としての意味を考察する必要あり**
- 例：日本人の成人男性の平均身長が167.0cmと異なるかを $n = 100{,}000$ 人の標本から検定

  $X \sim N(167, 6^2)$ と仮定すると $V[\bar{X}] = 0.00036$ cm

  棄却域(±0.00072cm)に現実的な意味があるだろうか？



5%水準での両側検定

仮説検定の実用上の注意 #2

有意水準で測るのは仮説の統計的な確度のみ
⇒　仮説の効果の大きさは測っていない

誤用例

- 既存薬と比較し新薬Aは5%有意水準で効果の差あり
- 既存薬と比較し新薬Bは1%有意水準で効果の差あり
- よって, 新薬Bの方が新薬Aよりも効果が高い

誤用の可能性がある例

- $n = 3{,}000$ 人の調査結果から, 既存薬と比較して新薬Cは1%有意水準で効果の差あり
- よって, 新薬Cは既存薬よりも効果が高い

仮説検定の実用上の注意 #3

“なぜ有意差が生じたのか？” の情報はない

– **例**1：サイコロを600回投げたとき1の目が120回も出た。この回数は偶然の範囲内か？

⇒ **サイコロに何らの偏りがあるかどうかを調べることで比較的容易に原因を特定できる可能性がある**

– **例**3：2024年度の東北大学の入学者の割合（入学者数 / 志願者数）は宮城県出身者は26.6%，北海道出身者は42.4%であった。この入学者の割合の違いは偶然の範囲内か？

⇒ **この事象が生じる要因はいくつも考え得るし，原因は複合的である可能性もある。加えて，ある仮説を検証・実証できるかは，また別の問題**

「北海道から仙台へ越境するには何か強い動機があるはずだ。だから入学率が高いのでは？」という仮説を立てたときに，それを実証するのは容易ではない

\# 例3では、2021年度の大学入試の志願者と入学者は全数調査のため標本調査ではない。しかし，各受験者の入学・非入学はベルヌーイ分布から発生する確率変数と見なすことができる。その上で、宮城県出身者の合否のベルヌーイ分布の平均母数 p_M と北海道出身者の合否のベルヌーイ分布の平均母数 p_H が異なるかどうかを検定（帰無仮説 $p_M = p_H$）するための検定統計量を適切に定義できる

演習問題

2021年12月に東北大学全学教育内で実施されたTOEFL ITPテストの東北大学文学部生と経済学部生の受検者のスコアの平均値は, それぞれ, 517点と507点であった。このスコアの平均値の差は偶然であるかどうかを調べたい

1. 仮説検定のための帰無仮説と対立仮説を立てなさい
2. 帰無仮説が棄却されたとする。そのときの仮説検定の結論を述べなさい
3. 帰無仮説が棄却できなかったとする。そのときの仮説検定の結論を述べなさい