検定の基礎
（教科書第8章）

北九州市立大学経済学部

齋藤 朗宏

今日の話題

- 背理法
- 統計的仮説検定

背理法

問題

- 日本人成人男性の平均身長が165cmではないことを示せるか？

背理法

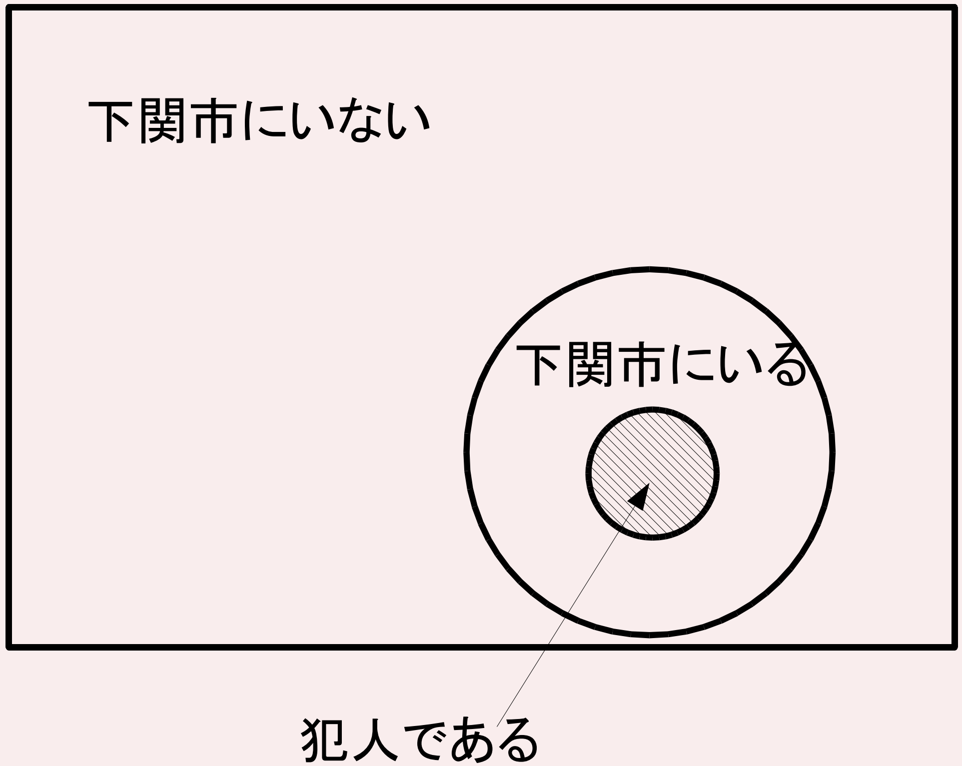
- 証明したい命題（統計学においては集合と考えても問題ない）A について，それを敢えて否定するA^cを提示し，そのA^c が成立し得ないという事実を示すことで, A の正しさを示す方法.

- 証明したいこと
  - 私は6月14日19時半に下関で起こった殺人事件の犯人ではない（A）.
- 証明方法
  - もし私が犯人であった（A^c）とする．そうであれば, 私はその時間下関にいる筈である．しかし，私はその時間，心理データ解析Ｉの講義をしていた．これは矛盾である．よって，私は犯人ではありえない.

背理法

- 背理法で証明できない場合でも，否定したかったA^cが真実であったことの証明にはならない．
  - 仮に私が講義を休講にしてその時間下関にいたとする．この場合，犯人ならばその時間下関にいる筈であるという命題は否定できない．しかし，だからと言って私が犯人だという証明にはならない．そもそも，下関には30万の人がいる．犯人はそのうちのたった一人であり，それ以外は犯人ではないので，「下関にいなければ犯人ではない」は事実でも，「下関にいたら犯人である」には無理がある．
  - 魔女ではないことを証明できないからと言って，魔女であるという証明にはならない．

背理法のイメージ



- 「犯人である」は「下関市にいる」の部分集合なので，「下関市にいる」を否定すれば「犯人である」も否定される．ただし，下関市にいるからといって犯人ではない．

統計的仮説検定

問題

- 日本人100人に対して身長を調べたところ，その平均は172.7cmであった．ここから日本人の平均身長は165cmではないと言えるか．

回答

- 実際には完全に証明することは極めて難しいが，真の身長165cmのとき，100人から得られた身長が172.7cmであるという事実がどの程度起こりにくいのかを示すことで，確率的に導くことができる．→統計的仮説検定
  - 教科書p.195参照

統計的仮説検定

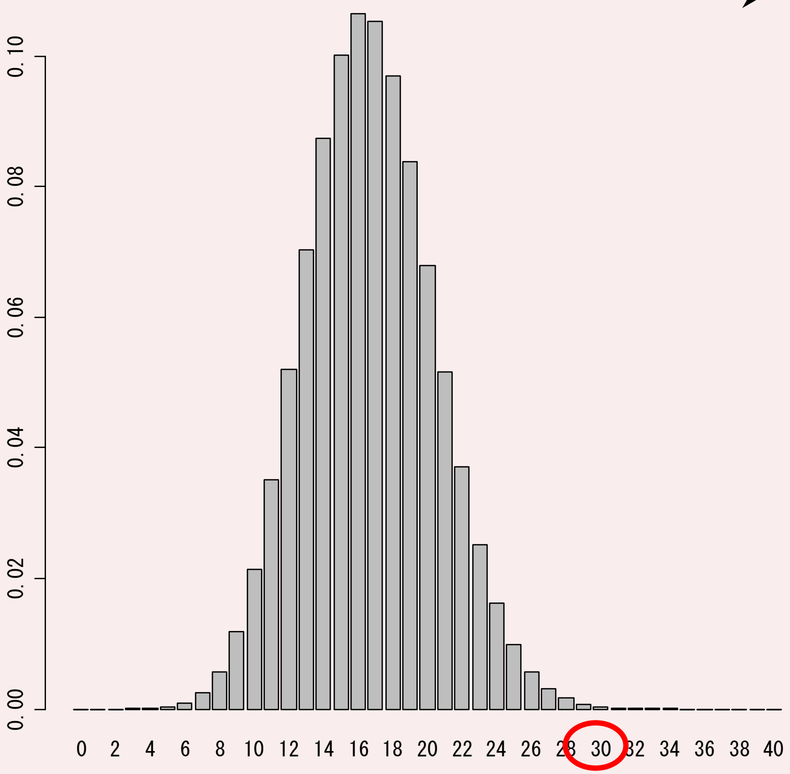
- 統計学では，母数がある値ではないことを証明したい場合が多々ある.
  - 6面さいころの，各々の目が出る確率は1/6である. しかし，イカサマ用のさいころだったら1/6にはならない. たとえば，6の目が出やすいさいころであることを証明するには？
  - 日本人と韓国人は同じ東アジア人であり，体格も比較的似ている. 日本人と韓国人の平均身長が違うことを証明するには？

統計的仮説検定

- 6面さいころの，各々の目が出る確率は1/6である．しかし，イカサマ用のさいころだったら1/6にはならない．たとえば，6の目が出やすいさいころであることを証明するには？

- さいころの目が6となることを成功，6以外となることを失敗と置くと，出目が6か否か，その回数は2項分布と呼ばれるに従うことがわかっている．従って，$p = 1/6$ の2項分布ではまず起こりえない程6の目の出る回数が多ければ，イカサマ用のさいころと見なすことができる．

統計的仮説検定



- さいころを100回振ったとする．もし6の目が出る確率が1/6なら，成功回数は以下のような2項分布に従う．ここで，もし6の目が30回出たとすれば，これは6の目が出る確率が1/6なのだけどたまたま起こりにくいことが起こったと考えるより，6が出る確率は1/6ではないと考える方が自然である．

統計的仮説検定の手順

1. 検証したい事柄を確認する（このさいころはイカサマか？）．
   - イカサマであるなら個々の出目の確率は等しくない（ここでは簡単のために6の目だけ出易いとする）が，イカサマでなければ個々の出目の確率は等しい．
2. 背理法に従い，検証したい事柄を否定する仮説を立てる．これを帰無仮説と呼ぶ（各々の目が出る確率は1/6で等しい）．
   - もしイカサマでないのだとすれば，さいころを100回振った時，6の目の出る回数は$p = 1/6, n = 100$の2項分布に従う．

統計的仮説検定の手順

3. データを取り，その結果が帰無仮説の下で十分に起こりにくいことであるか確認する．

- どのくらいの確率で起こることだったら「起こりにくい」と考えるか，その基準を定める必要がある．この基準となる確率を有意水準と呼び，αで表す．一般的には，$\alpha = 0.05, 0.01, 0.001$などが基準としてよく用いられる．
- 2項分布の累積分布関数から，$p = 1/6, n = 100$であれば，10回から24回の間に全体の95%が入る．よって，有意水準を5%と設定した場合，9回以下あるいは25回以上であったならば，起こりにくいことが起きた．10回から24回の間なら起こりにくいことは起きてないと考える．

統計的仮説検定の手順

4. 起こりにくいことが起こった場合には，帰無仮説を棄却して対立仮説を採択する．
   - 26回6の目が出たので，このさいころはイカサマではないという帰無仮説は棄却され，このさいころはイカサマであり，各々の目が出る確率は1/6ではないという対立仮説が採択される．また，このとき有意差（「差」は仮説との差）があったという言い方をする．

➢ 帰無仮説を棄却出来なかった場合，帰無仮説が正しいと証明された訳ではない点に注意．
   - つまり，23回6の目が出た場合，帰無仮説，「6の目が出る確率は1/6である」は棄却できないが，確率1/6であることが証明された訳ではない（6の目が出る確率は23/100の信頼区間と被っている）．

統計的仮説検定

- 日本人男性100人に対して身長を調べたところ，その平均は172.7cm，標準偏差8.7であった．ここから日本人男性の平均身長は165cmではないと言えるか．

- 身長は正規分布に従っていると考えられる．中心極限定理から，標本平均は平均が母平均 μ_x ，標準偏差は $\sigma_x/\sqrt{100}$ の正規分布に従っていると言える．得られた標本平均172.7cmが，母平均165cmであるときにどれだけ起こりにくいことかを考え，それが起こりにくいことであるならば，165cmではないと考える．

統計的仮説検定の手順

1. 検証したい事柄を確認する（日本人の平均身長（母平均）は165cmか？）．
2. 背理法に従い，検証したい事柄を否定する仮説を立てる．これを帰無仮説と呼ぶ（日本人の身長の母平均は165cmである）．
   - 165cmであるならば，標本平均は平均165cm，標準偏差 $\sigma_x/\sqrt{100}$ の正規分布に従う．

統計的仮説検定の手順

3. データを取り，その結果が帰無仮説の下で十分に起こりにくいことであるか確認する．

- どのくらいの確率で起こることだったら「起こりにくい」と考えるか，その基準を定める必要がある．この基準となる確率を**有意水準**と呼び，αで表す．一般的には, $\alpha = 0.05, 0.01, 0.001$などが基準としてよく用いられる．
- $\sigma_x/\sqrt{100}$について，母標準偏差が未知であるため，標本の標準偏差で代用し，$8.7/\sqrt{100}$を考える．このとき，帰無仮説が正しいのであれば，以下の式で得られるtは自由度99のｔ分布に従う．

$$t = \frac{\bar{x} - \mu_x}{s/\sqrt{n}}, \ t = \frac{172.7 - 165}{8.7/\sqrt{100}}$$

統計的仮説検定の手順

4. 起こりにくいことが起こった場合には，帰無仮説を棄却して対立仮説を採択する．
   - $t = 8.85$ である．自由度99のt分布の場合（即ち，帰無仮説が正しいのであれば）こうして得られたｔの95%が±1.98の範囲に入る筈である．しかし，その範囲から大きく外れている．これは，たまたま残り5%が発生したのではなく，帰無仮説が間違っていた，つまり母平均は165cmではないと考える．

統計的仮説検定の手順

- 帰無仮説を棄却出来なかった場合，帰無仮説が正しいと証明された訳ではない点に注意．
  - 帰無仮説が母平均171cmであるかどうか．であったとするならば，$t = 1.93$であり，±1.98の範囲内なので帰無仮説は棄却できない．しかし，母平均171cmが証明された訳ではない．
  - 母平均が171cmである可能性もあるが，母平均が172cmや173cmである確率の方が高そうに見える実際，このデータは母平均173cmの集団からの標本である．

実習

- お年玉のデータに関して，お年玉の真の平均値は20000円ではないと言うことが出来るだろうか？
- $\alpha = 0.05$ とする.

$$t = \frac{\bar{x} - \mu_x}{s/\sqrt{n}}$$

追加資料

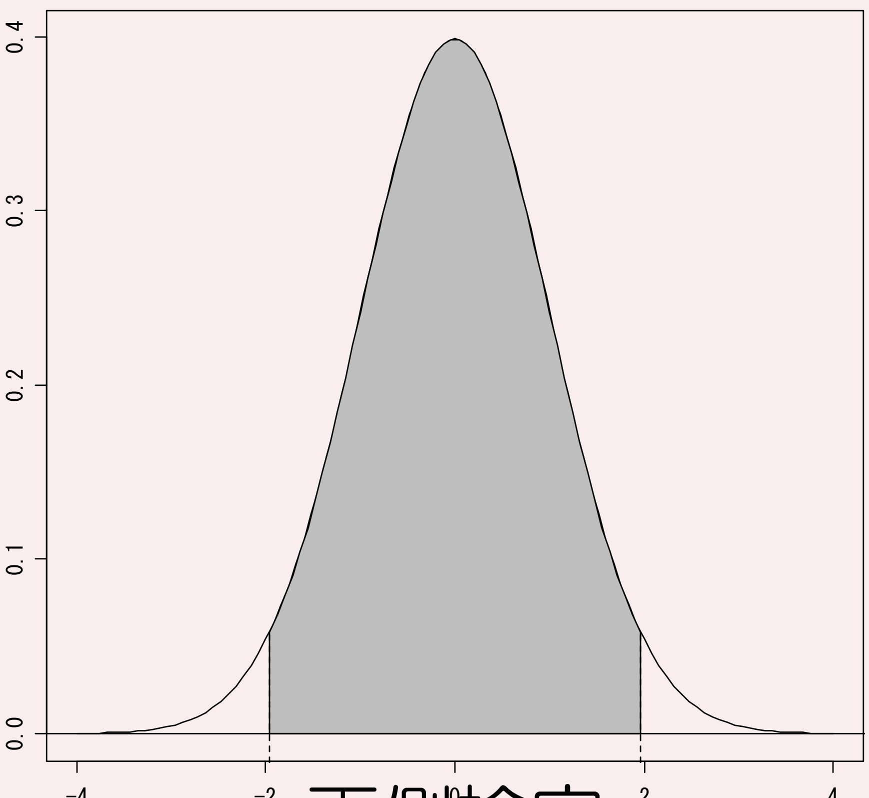
t 統計量と z 統計量

- 母標準偏差が既知であるとみなせる場合には，左の式によりz統計量が得られる．このとき，帰無仮説が真であれば，ｚは標準正規分布に従う．
- 母標準偏差が未知である場合には，右の式によりｔ統計量が得られる．このとき，帰無仮説が真であれば，ｔはｔ分布に従う．
- これらの値の絶対値が標準正規分布，当該自由度におけるｔ分布の95%範囲より大きいならば，有意水準0.05で帰無仮説は棄却できる．

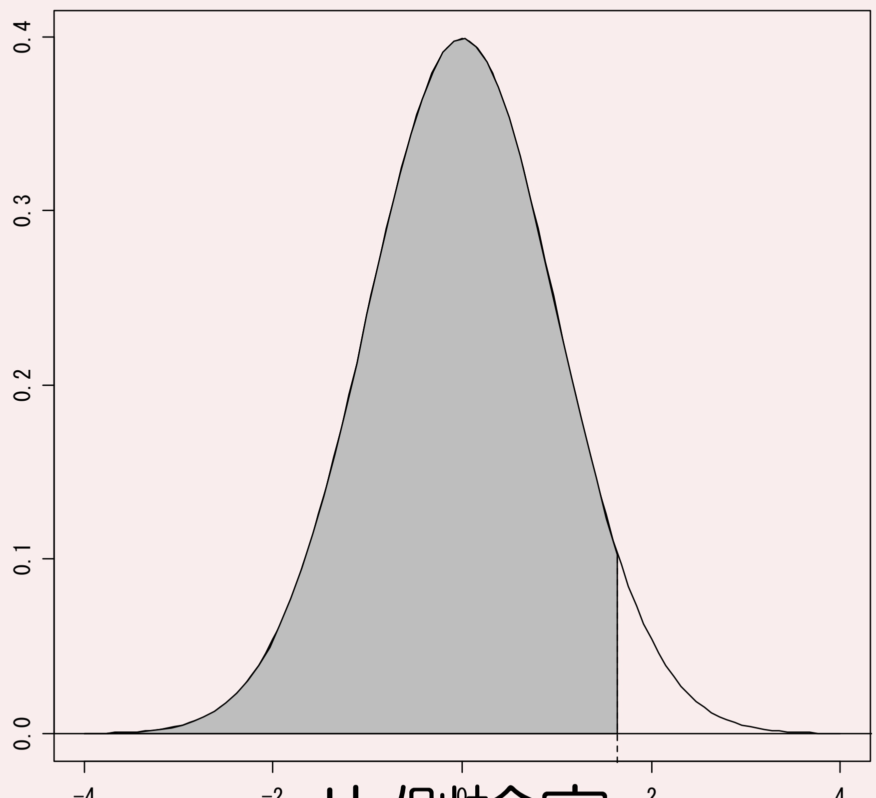
$$z = \frac{\bar{x} - \mu_x}{\sigma_x/\sqrt{n}}, \quad t = \frac{\bar{x} - \mu_x}{s/\sqrt{n}}$$

片側検定と両側検定

- 「起こりにくい」をどう定義するのかによって範囲の設定の仕方が変わる.
- どちらも塗りつぶした部分の面積は95%.



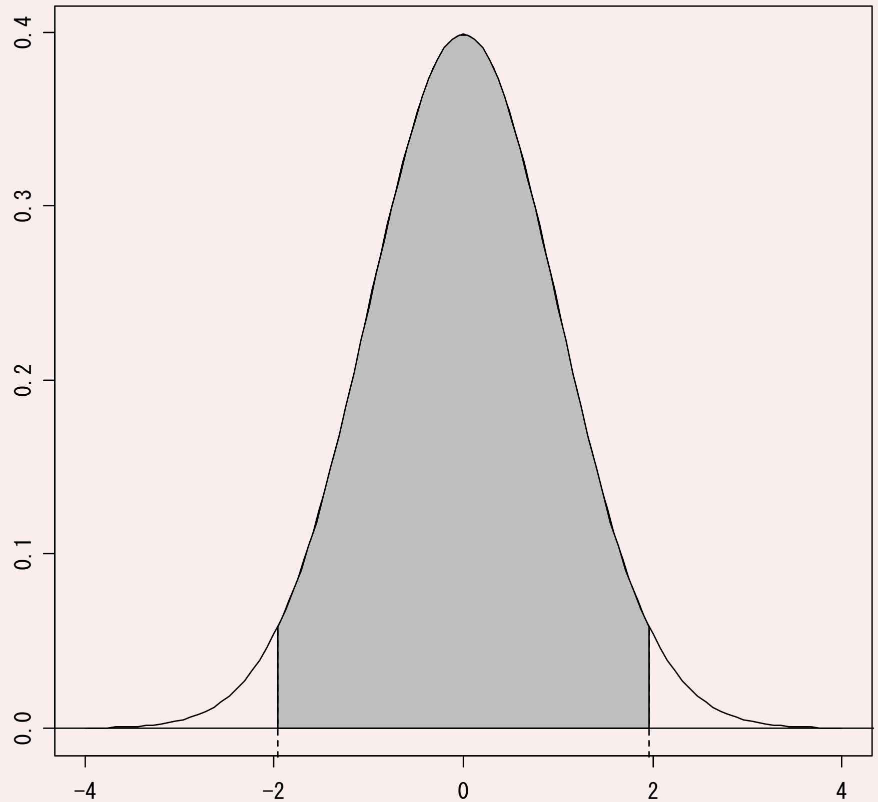
両側検定



片側検定

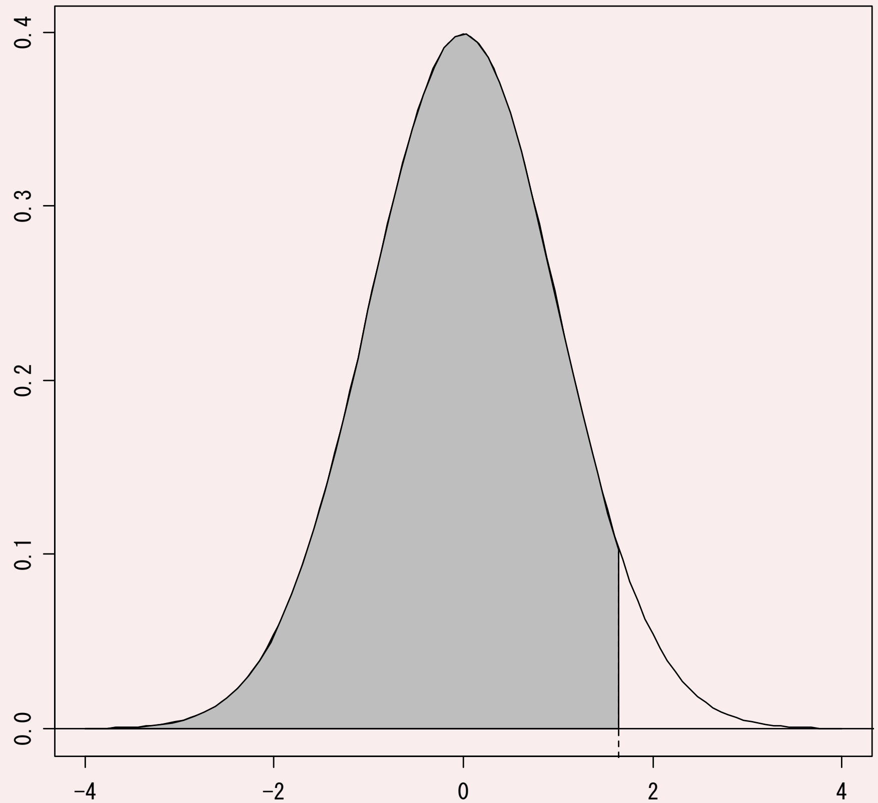
片側検定と両側検定

- 平均値が〇〇であるという帰無仮説を検定したいなら，両側検定を用いるのが一般的．



片側検定と両側検定

- 平均値がある値以上（以下）であることが明らかな状態で検定したい場合には，片側検定を用いることができる．迷ったら両側検定を使うのが確実．



片側検定を行うことが出来る例

- 1年生と6年生ではどちらが体重が重いか？
- 勉強前と後では学力向上が見られるか？

- これらの例では，興味の対象は「向上しているか否か」であって，「高いか低いか」ではない．

統計的仮説検定の問題点

- あくまでも「確率」での評価であるため，間違える可能性がある．
- 帰無仮説が正しいとき棄却してしまう場合と，帰無仮説が間違いなのに棄却できない場合は間違いである．
- 帰無仮説が正しい時棄却しないのは正しい行動だが，結果としては何も言えない．帰無仮説が間違えている時に棄却するのが検定の目的．

帰無仮説	正しい	間違い
棄却しない	正しい	第二種の誤り β
棄却する	第一種の誤り α	検定成功