統計学II

早稲田大学政治経済学術院
西郷 浩

本日の目標

- 統計的仮説検定の仕組み
- 帰無仮説と対立仮説
- 棄却域
- 片側検定

統計的仮説検定の仕組み(1)

- たとえ話
  - いつも慎重な友人が、今日はミスを繰り返している。
  - それを見て、「何かあったのかな？」と思う。

統計的仮説検定の仕組み(2)

- どのようにしてそのような結論に到達したか。
  - 作業仮説の設定
    - 「その友人に何も起きていない（いつもどおり）」と仮定する。
  - 作業仮説の下での可能性（確率）の見積もり
    - いつもどおりなら、その友人がミスを繰り返すことはほとんどない。
  - 最終判断（結論）
    - 作業仮説を否定する。
      - 滅多におきないことが起きたとは考えない。
    - 作業仮説を否定しない
      - 滅多に起きないことが実際に発生している。

統計的仮説検定の仕組み(3)

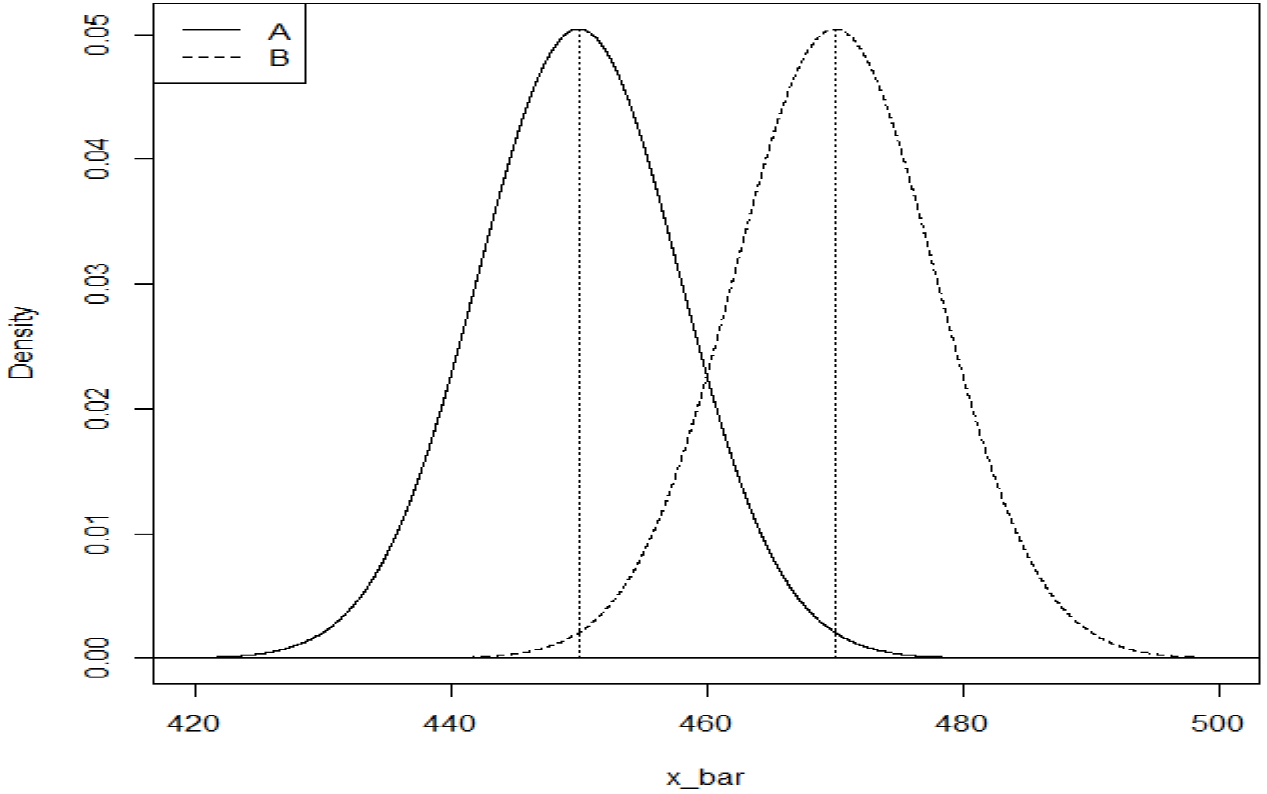
- 2種類の果樹
  - 果樹Aの果実の重さ: $N(450, 50^2)$ にしたがって分布する。
  - 果樹Bの果実の重さ: $N(470, 50^2)$ にしたがって分布する。
    - 果樹Aから取れた果実の評判が良い。
- ラベルを付け忘れた箱詰めの果実40個
  - 全部果樹 A から取れたか、全部果樹 B から取れたか。
  - 平均重量を測ったら 464g であった。
  - どちらの果樹から取れたと判断すべきか。
    - 果樹Bの果実の方が、平均重量が高めに出やすい。
    - けれども、果樹Aの果実にも大小があり、450gよりも平均重量が大きくなることはある。
    - 平均重量が470gに近いとうだけで、「果樹Bから取れた果実である」と結論をくだすわけにはいかない。

統計的仮説検定の仕組み(4)

- 問題の整理
  - 標本抽出:
    - $X_i \sim_{iid} N(\mu, 50^2), i = 1,\ 2, \dots, 40.$
  - 仮説の設定
    - 仮説A:果樹Aから取れた果実である（$\mu = 450$）。
    - 仮説B:果樹Bから取れた果実である（$\mu = 470$）。
      - 両者のうちどちらを取るかの意思決定の問題とみなせる。
  - 意思決定のための判断材料＝標本
    - 標本から計算した標本平均 $\bar{X} = \frac{1}{40}\sum_{i=1}^{40} X_i$
    - 標本平均 $\bar{X}$ の値によって、仮説Aと仮説Bのどちらかを取る。

統計的仮説検定の仕組み(5)

図1：仮説A（仮説B）から取れた40個の果実の平均重量の標本分布



帰無仮説と対立仮説(1)

- 意思決定（受け入れる仮説）と真の状態との関係

表1: 下される結論と真の状態との関係

下される結論	真の状態	
	仮説Aが真	仮説Bが真
仮説Aを取る	正解	過誤
仮説Bを取る	過誤	正解

  – 2つの過誤を同時に0にすることはできない。
    - $\bar{X}$ が多少大きくても仮説Aを取る。→ 右上の過誤が大きくなる。
    - $\bar{X}$ が少しでも大きければ仮説Bを取る。→ 左下の過誤が大きくなる。
  – 妥協案を模索する必要がある。

帰無仮説と対立仮説(2)

- 仮説に役割分担をあたえる。
  - 仮説Aを帰無仮説にする。$H_0 : \mu = 450$
  - 仮説Bを対立仮説にする。$H_1 : \mu = 470$
    - 帰無仮説は作業仮説に当たる。
    - 対立仮説は、帰無仮説が棄却(否定)されたときに採択される仮説である。

表2: 下される結論と真の状態との関係

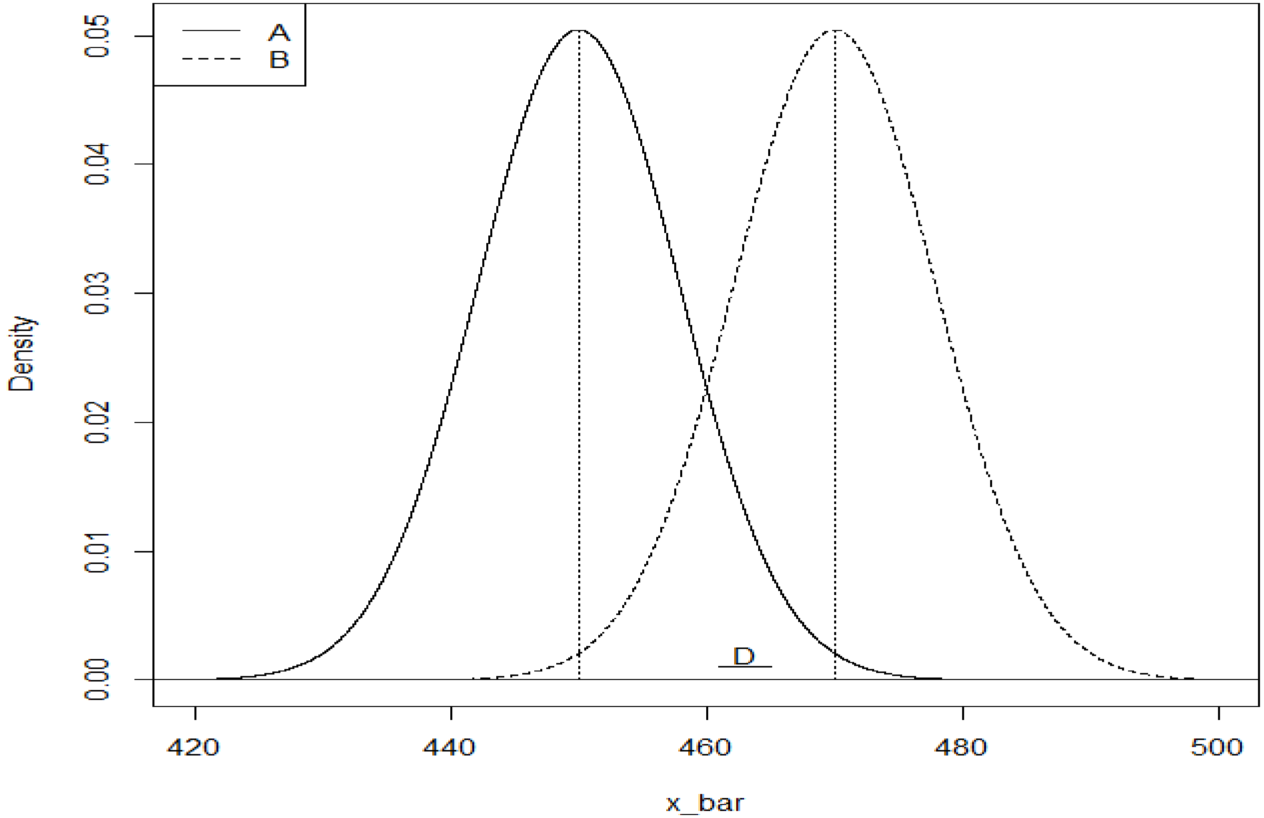
下される結論	真の状態	
	H_0が真	H_1が真
H_0を棄却しない	正解	第2種の過誤
H_0を棄却する	第1種の過誤	正解

棄却域の決定(1)

- 棄却域
  - 標本情報は標本平均 $\bar{X}$ に集約されている。
  - $\bar{X}$ の値によって、H_0を棄却するか否かを決める。
    - 「数直線上に、ある領域 D を定め、観察された標本平均が領域 D に入るか否かで結論を下す」と表現しても、内容は同じである。
      - もし、$\bar{X} \in D$ であれば、H_0 を棄却する。
      - もし、$\bar{X} \notin D$ であれば、 H_0 を棄却しない。
        - » そのような領域 D を、（H_0の）棄却域とよぶ。
  - 棄却域と第1種の過誤の確率αとおよび第2種の過誤の確率 β との関係

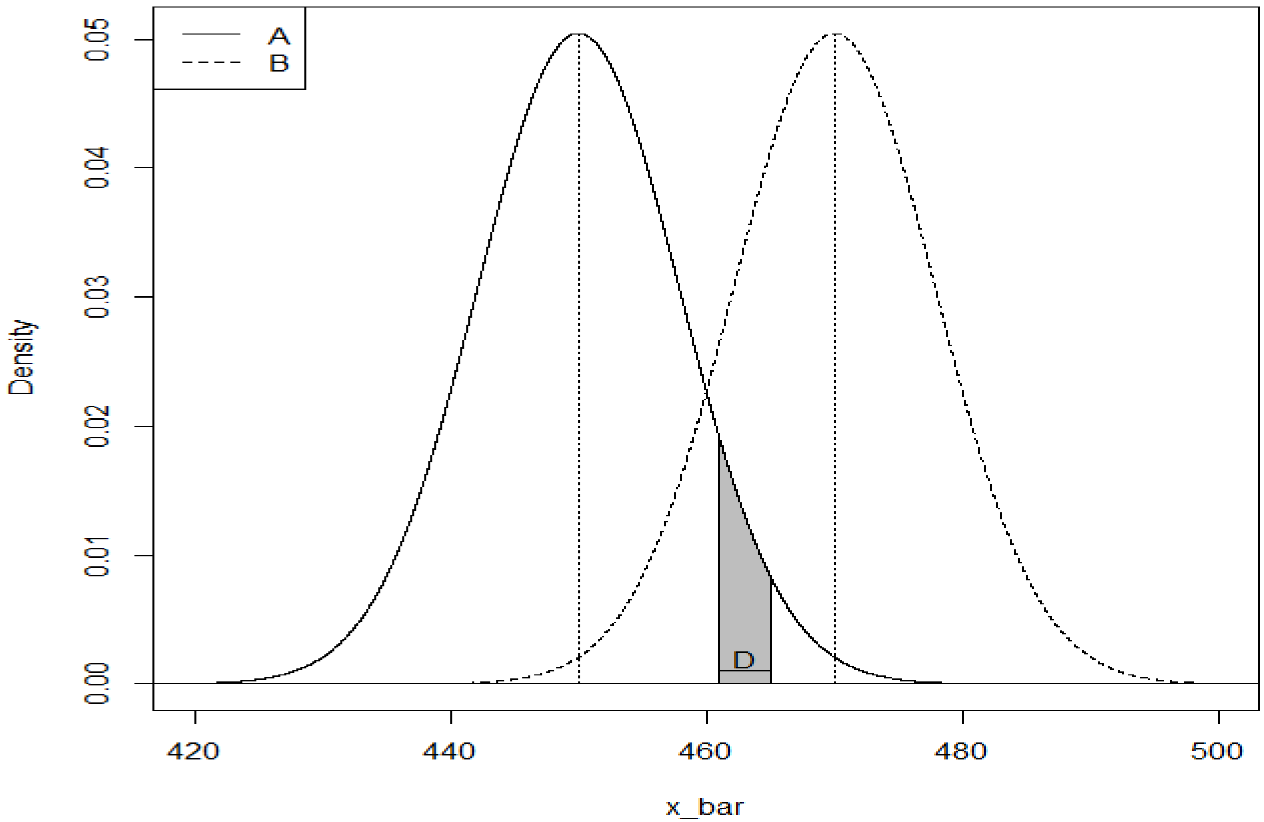
棄却域の決定(2)

図2-1: 棄却域と第1種の過誤、第2種の過誤



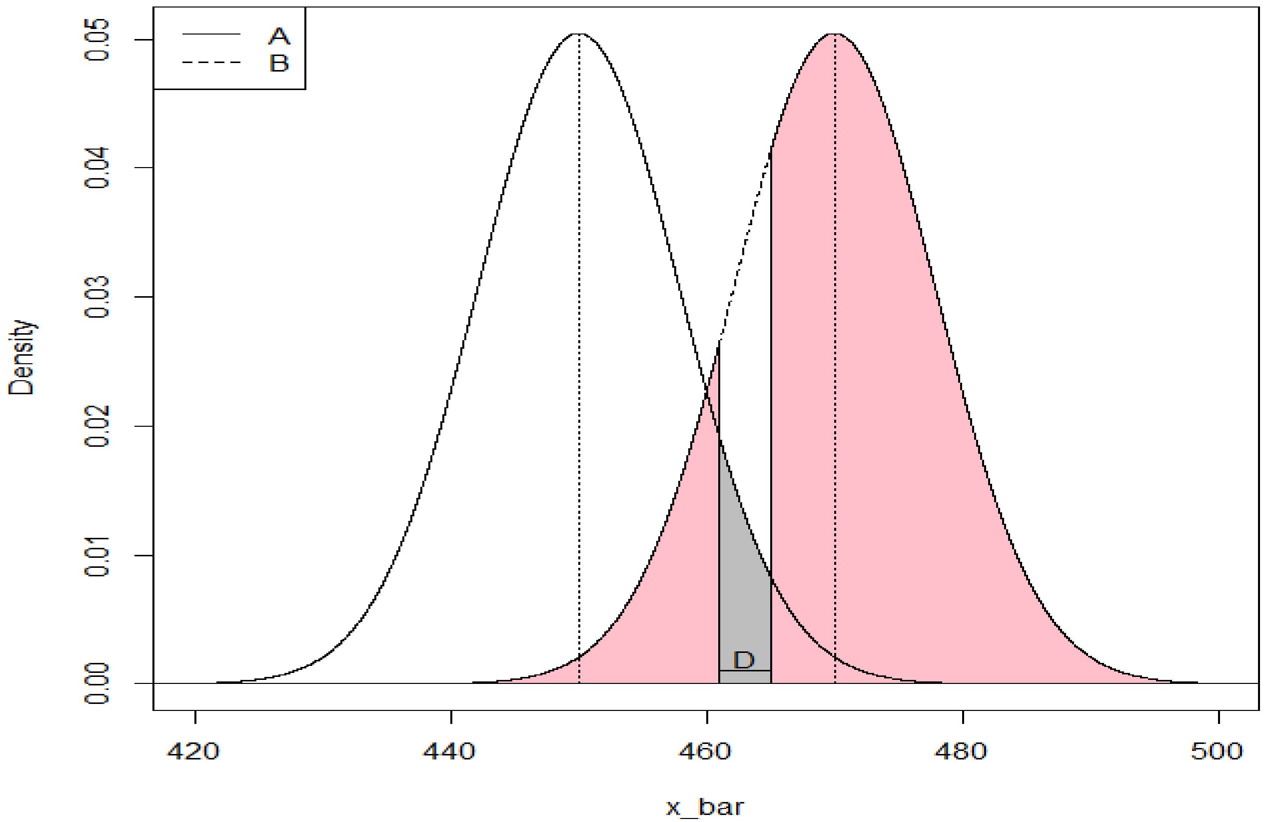
棄却域の決定(3)

図2-2: 棄却域と第1種の過誤、第2種の過誤



棄却域の決定(4)

図2-3: 棄却域と第1種の過誤、第2種の過誤



検定手続きの構成(1)

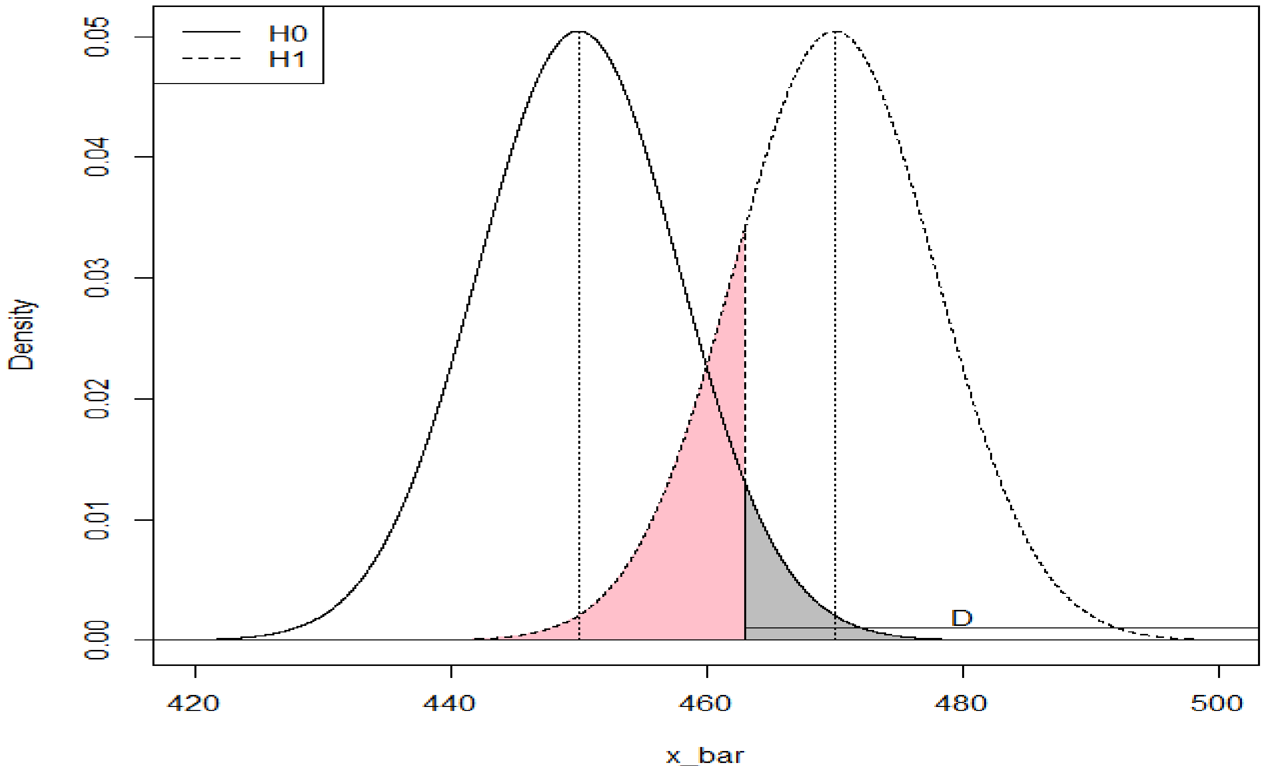
- トレードオフの解決方法
  - 第1段階:
    - 第1種の過誤が発生する確率αを、ある値q以下にする。
      - 値 q を有意水準とよぶ。
      - この講義では $q = 0.05$ (5%) とすることが多い。
  - 第2段階
    - 第1段階の条件を満たしたうえで、第2種の過誤が発生する確率βを最低にする。
      - 検出力 $1 - \beta$ を最高にする、と言い換えても同じになる。
  - この接近法により検定方法が一意に定まる。

検定手続きの構成(2)

- 2段階法による棄却域の決定
  - $D = \{\bar{x}\text{: } \bar{x} > c\}$
    - ただし、c は、H_0 が正しいときに、$P(\bar{X} \in D) = q(= 0.05)$ となるように選ぶ。
    - より正確には、棄却域は標本空間に定められる。
      - $D = \left\{(x_1, x_2, \dots, x_n)\text{: } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i > c\right\}$
- 教科書の例
  - $c = 450 + 1.64\sqrt{50^2/40} \approx 463$
  - 標本から計算された標本平均は464gで c よりも大きい。→ H_0 を棄却する。
    - つまり「箱のなかの40個の果実は、果樹Aではなく果樹Bから取れた」と結論する。

検定手続きの構成(3)

図3: 最適な棄却域と第1種の過誤、第2種の過誤



検定手続きの構成(4)

- 棄却域の別表現
  - Z検定
    - $Z=\frac{\bar{X}-450}{\sqrt{50^2/40}}$
      - 分子：標本平均とH_0が想定する母平均との差
      - 分母：標本平均の標準誤差（散らばり）
    - もし、$Z>1.64$ であれば、H_0 を棄却する。
      - 標本平均に含まれる散らばりに比べて、標本平均とH_0が想定する母平均との差があまりにも大きければ、H_0 を棄却する。

これまでのまとめ:片側検定(右側)

- 標本抽出
  - $X_i \sim_{iid} N(\mu, \sigma^2)\ i = 1, 2, \ldots, n.$
- 前提
  - 母分散 σ^2 が既知である。
    - それが未知の場合は次回にあつかう。
- 帰無仮説と対立仮説
  - $\begin{cases} H_0\colon \mu = \mu_0 \\ H_1\colon \mu = \mu_1 \end{cases}$ ただし、$\mu_0 < \mu_1$ であるとする。
    - 単純帰無仮説 vs 単純対立仮説
- 有意水準0.05の棄却域
  - $D = \left\{\bar{x}\colon \bar{x} > \mu_0 + 1.64\sqrt{\sigma^2/n}\right\}$
    - または $Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}}$ として、$D = \{z\colon z > 1.64\}$

これまでのまとめ：片側検定（左側）

- 標本抽出と前提
  - 前に同じ。
- 帰無仮説と対立仮説
  - $\begin{cases} H_0\colon\ \mu = \mu_0 \\ H_1\colon\ \mu = \mu_1 \end{cases}$ ただし、$\mu_0 > \mu_1$ であるとする。
- 有意水準0.05の棄却域
  - $D = \left\{\bar{x}\colon\ \bar{x} < \mu_0 - 1.64\sqrt{\sigma^2/n}\right\}$
    - または $Z = \frac{\bar{X}-\mu_0}{\sqrt{\sigma^2/n}}$ として、$D = \{z\colon z < -1.64\}$

練習問題1(1)

- W大学政治経済学部
  - 入学時に英語能力試験を実施している。
  - 過去の実績（得点の分布）
    - 平均点500点、標準偏差100点の正規分布。
  - ある教員の主張:
    - 今年の入学者900人の実力は例年より高く、平均点で見積もって510点である。
  - 入学者900人に対して実施した試験の実際の平均点が505点だった。
  - 主張は正しいといえるか。

練習問題1(2)

- 解答
  - 仮説の設定
    - データ発生の仕組み：$X_i \sim N(\mu, 100^2)$
    - 帰無仮説：H_0: $\mu = 500$
    - 対立仮説：H_1: $\mu = 510$
    - 有意水準を0.05とする。
  - 棄却域
    - $\bar{X} > 500 + 1.64\sqrt{100^2/900} = 505.5$
      - または、$Z = \frac{\bar{X}-500}{\sqrt{100^2/900}} > 1.64$

練習問題1(3)

- 結論
  - $\bar{X}_{obs} = 505$ なので棄却域に入らない。
    → H_0は棄却できない。
    - または、$Z_{obs} = \frac{505-500}{\sqrt{100^2/900}} = 1.5$ なので H_0は棄却できない。
  - つまり:
    - 今年度の入学者の平均点で測った実力が例年の500点よりも高いという証拠は得られなかった。

練習問題2(1)

- あるメーカー
  - 携帯端末Aの使用可能時間（寿命）
    - 平均8760時間（＝24時間×365日）、標準偏差250時間の正規分布
  - 部品を取り換えて軽量化した。
  - 「平均的な使用時間が100時間短くなって、8660時間になった」という懸念あり。
  - 軽量化した製品25個の寿命を測定したところ、平均的な寿命が8670時間であった。
  - 懸念は正しいといえるか。

練習問題2(2)

- 解答
  - 仮説の設定
    - データ発生の仕組み: $X_i \sim N(\mu, 250^2)$
    - 帰無仮説:H_0: $\mu = 8760$
    - 対立仮説:H_1: $\mu = 8660$
    - 有意水準を0.05とする。
  - 棄却域
    - $\bar{X} < 8760 - 1.64\sqrt{250^2/25} = 8678$
      - または、$Z = \frac{\bar{X}-8760}{\sqrt{250^2/25}} < -1.64$

練習問題2(3)

– 結論
  - $\bar{X}_{obs} = 8670$ なので棄却域に入る。
    → H_0は棄却される。
    – または、$Z_{obs} = \frac{8670-8760}{\sqrt{250^2/25}} = -1.8$ なので H_0は棄却される。
  - つまり:
    – 「平均的な使用時間が8760時間である」という仮説が疑われる証拠が得られた。