応用システム工学
第一回　確率統計の復習

平成25年04月19日

測定の不確かさ

- 科学・工学の発展を支えてきたのは測定技術
  - 測定の信頼性とトレーサビリティ
- 測定の信頼性はなぜ必要？
  - 製品性能の測定（契約等）
  - 製品の規格・規制への適合（法律・基準等）
  - 世界最高性能！等を公表する場合
- トレーサビリティ
  - 標準器による測定機器の定期校正
  - 標準器がいつ・どこで・どのような標準器で校正されたかを，さかのぼって国家標準・国際標準機関まで追跡できること

不確かさ

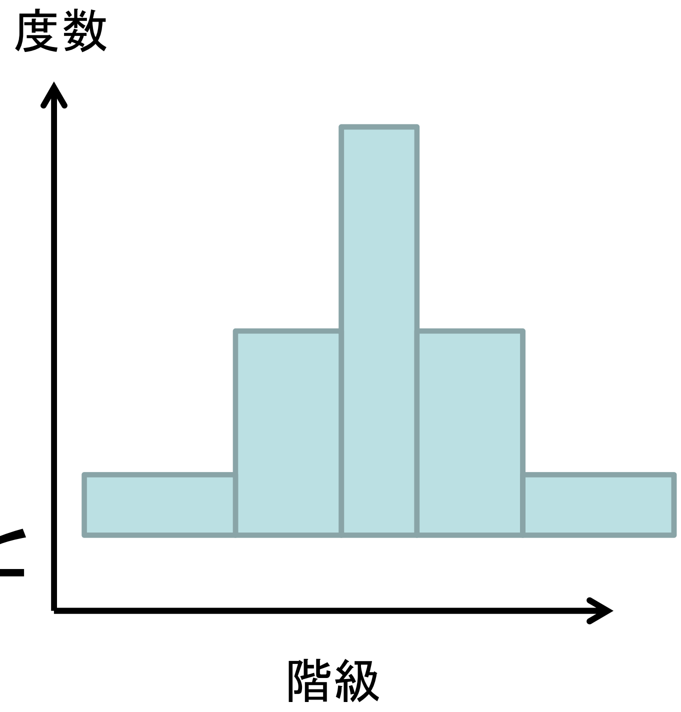
- 定量的な測定値の信頼性を表すために定義されたもの
- 定義
  - 測定の結果に付随した，合理的に結び付けられ得る値のばらつきを特徴づけるパラメータ
    - ISO Guide to the Expression of Uncertainty in Measurement (GUM) 1993年
    - 測定の結果に付随するパラメータであり，「測定値」からどの程度のばらつきの範囲内に「真の値」があるかを示すもの
- 統計解析により不確かさを評価する方法と，統計的解析以外の方法により評価する方法がある

統計用語

- 統計的推測　→　標本調査で(データの一部を調べて)全体を推測する
- 分布　→　標本調査したデータのバラつき
- 度数　→　データの現れる頻度を表す量
  - 相対度数　→　%で表したもの
- 度数分布　→　度数で表した分布
- 階級　→　データの値が入る区間
  - データが整数であらわされない場合等
  - 区間の大きさを階級幅
  - 階級の上限・下限の中間値を階級値
- 代表値　→　分布を一つの数字で表す

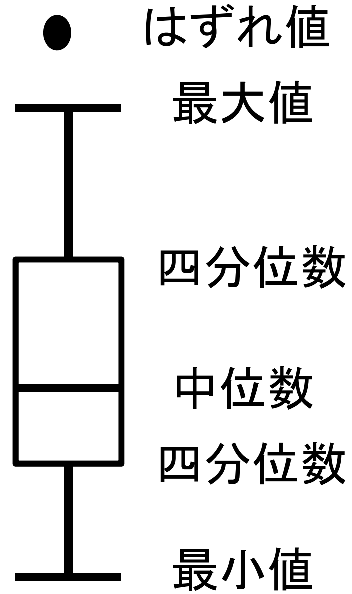
ヒストグラム

- 単なる棒グラフとは異なる
  - 棒の間隔を開けない
- 横軸　階級
- 縦軸　度数(相対度数)
- 階級幅を底辺・度数を面積とする長方形で，各階級の度数を表す
  - 面積で度数を表すことで，階級の取り方の自由度が高くなる



箱ひげ図(ボックスプロット・四分位図)

- ヒストグラムの発展系
- 下記の値のデータで表す
  - 最小値(小さい方から0%)
  - 第一四分位数(小さい方から25%)
  - 中位数(小さい方から50%)
    - 中央値，メディアン
  - 第三四分位数(小さい方から75%)
  - 最大値(小さい方から100%)



飛び離れた値を外れ値として，最大値・最小値から除外することもあり

分布を表す量
代表値の例

- 算術平均(相加平均)=$\dfrac{\text{データの合計}}{\text{データの総数}}$
  - $\bar{x} = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$
- 中央値(中位数，メジアン)=データを大きさ順に並べた時の中央の順位の値
  - データ数が偶数の場合は，中央の二つの算術平均をとる
- モード(最頻値) → 度数が最も多い階級値

分布を表す量

- 偏差 → データ値と平均との差
  - 偏差の平均は0
- 分散 → 偏差の二乗平均
  - $\sigma^2 = V(x) = \frac{1}{n}\{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2\}$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

  - 標準偏差 → 分散の平方根 (もとの単位にもどる)

分布を表す量
モーメント

- モーメント(積率) → 分散の発展形
- 変数Xの度数分布を考える
  - 階級値xの階級の相対度数f(x)
    - 平均$\mu = E(X) = \sum_x xf(x)$
    - 分散$\sigma^2 = V(X) = E\{(x-\mu)^2\} = \sum_x (x-\mu)^2 f(x)$
    - 一般形 変数Xの関数g(x)の平均

$$E\left(g(x)\right) = \sum_x g(x)f(x)$$

  - 関数g(x)がk次 → 変数Xのk次のモーメント

分布を表す量
モーメント

- 原点の周りのk次のモーメント　$\mu'_k = E\left(X^k\right)$
  - 平均μは原点の周りの1次のモーメントに相当
- 平均の周りのk次のモーメント　$\mu_k = E\left((X-\mu)^k\right)$
  - 分散σ^2は平均の周りの2次のモーメントに相当
- 高次のモーメントを用いた分布の特徴量
  - 3次のモーメント　歪度　$\alpha_3 = \frac{\mu_3}{\sigma_3}$
    - ヒストグラムの正または負への偏り具合
  - 4次のモーメント　尖度　$\alpha_4 - 3 = \frac{\mu_4}{\sigma_4} - 3$
    - 正規分布の尖度を0とする（正規分布の値が3）
    - 平均値付近での分布の集中の度合い