統計学II

早稲田大学政治経済学術院
西郷 浩

本日の講義の目的

- 推定
  - 点推定と区間推定
- 点推定
  - 推定量の評価の基準
    - 平均2乗誤差；不偏性；一致性
- 点推定量の求め方
  - 積率法と最尤法

推定

- 点推定
  - 母数のなるべく近くに出現する統計量を見つけること。
    - そのような統計量を、とくに「推定量」とよぶ。
      - 例：母平均 μ の点推定量：$\hat{\mu} = \bar{X}$
- 区間推定
  - 所与の値以上の確率で母数をふくむような区間を標本情報から見つけること。
    - そのような区間を、信頼区間または区間推定値とよぶ。
      - 例：母平均 μ の信頼区間：$[L, U]$ ただし $P(L \leq \mu \leq U) \geq 0.95.$

点推定

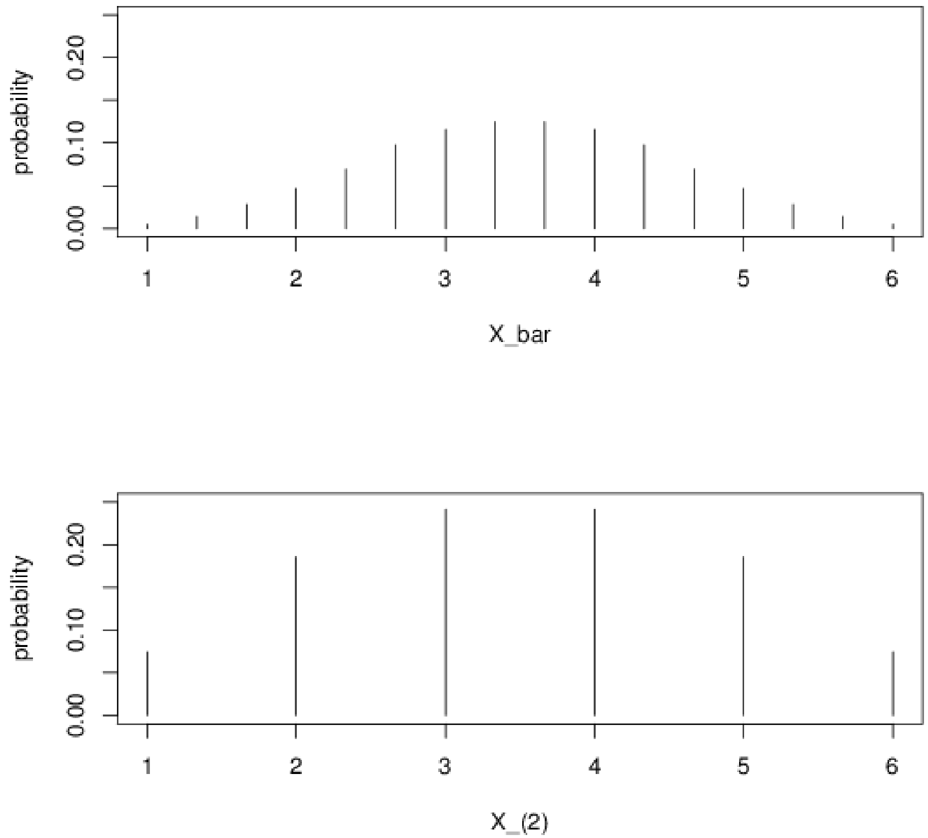
- 推定量
  - 標本:$X_i \sim_{iid} (\mu, \sigma^2)$ $(i = 1, 2, \dots, n)$
  - 推定量:母数のなるべく近くに出現する統計量
    - 一般の母数 θ の点推定量:$\hat{\theta} = f(X_1, X_2, \dots, X_n)$
    - 例:
      - $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$
      - $\tilde{\mu}$ =(標本中央値)
      - $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
      - $\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$

平均2乗誤差(1)

- 平均2乗誤差
  - 母平均 μ の推定量 $\hat{\mu}$ の平均2乗誤差
    - $MSE_{\hat{\mu}}(P) = E_P\{(\hat{\mu} - \mu)^2\} = \{E_P(\hat{\mu}) - \mu\}^2 + V_P(\hat{\mu})$
    - 平均2乗誤差が小さい推定量の方が、平均的な誤差の2乗が小さいという意味で、母数に近いといえる。
- 例:
  - 1から6の目までが一様に出やすいサイコロを3回投げるとき:
    - $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$
    - $\tilde{\mu} = X_{(2)}$=(標本中央値)

平均2乗誤差(2)

図1：標本平均と標本中央値の標本分布



注：公正なサイコロを3回投げるとき。

野口・西郷(2014)『基本 統計学』培風館 p. 139

平均2乗誤差(3)

- 平均2乗誤差による比較
  - $MSE_{\bar{X}}(P) \approx 0.97$
  - $MSE_{X_{(2)}}(P) \approx 1.88$
    - この場合の P は「1から6までの目が一様に出やすい」という母集団分布をあらわす。
  - 標本平均 $\bar{X}$ の方が標本中央値 $X_{(2)}$ よりも母平均 μ からの平均的な2乗誤差が小さい。
- 平均2乗誤差を利用する場面
  - 複数の推定量の精度を比較する。

不偏性(1)

- 不偏性：
  - 母平均 μ の推定量 $\hat{\mu}$ が以下の性質を満たすとき、$\hat{\mu}$ は μ について不偏であるという。
    - $E_P(\hat{\mu}) = \mu$ （すべてのμについて）
      - 推定量 $\hat{\mu}$ は、母平均 μ を過大推定することも過小推定することもあるが、過大推定・過小推定のどちらか一方に偏ることはない。
      - 不偏性をもつ推定量を不偏推定量と呼ぶ。
      - 不偏性をもたない推定量を偏りのある推定量と呼ぶ。

不偏性(2)

- 例:
  - 必ずしも公正でないサイコロを3回投げる。
    - $P: P(X_i = j) = p_j$ ただし、$\sum_{j=1}^{6} p_j = 1.$
  - $E_P(\bar{X}) = \mu$
  - $E_P\left(X_{(2)}\right) \neq \mu$ (一般に)
  - 例えば、
    - $p_1 = p_2 = \frac{3}{12}, p_3 = p_4 = \frac{2}{12}, p_5 = p_6 = \frac{1}{12}$ のとき。
  - $\mu \approx 2.8, E(\bar{X}) = \mu, E\left(X_{(2)}\right) \approx 2.7 < \mu$

一致性(1)

- 一致性
  - 母平均 μ の推定量 $\hat{\mu}$ が以下の性質を満たすとき、$\hat{\mu}$ は μ について一致性を持つという。
    - 任意の正数 ε に対して、$P(|\hat{\mu}-\mu| \leq \varepsilon) \rightarrow 1$ as $n \rightarrow \infty$.
      - 標本の大きさ n が大きくなるほど、母平均 μ の近辺に高い確率で推定量 $\hat{\mu}$ が出現する。
    - 一致性を持つ推定量を一致推定量と呼ぶ。

一致性(2)

- 例
  - $X_i \sim_{iid} (\mu, \sigma^2)$ $(i = 1, 2, \ldots, n)$
  - $P(|\bar{X} - \mu| \leq \varepsilon) \to 1$ as $n \to \infty$.
    - 大数の法則による主張。
  - 標本平均 $\bar{X}$ は母平均 μ の一致推定量である。

点推定量の求め方：積率法(1)

- 積率
  - 母集団における k 次の積率：　$\mu_k = E\left(X^k\right)$
  - 標本における k 次の積率：　$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$
- 積率法
  - 母数を、母集団における積率の式で表現する。
  - そこにおいて、母集団における積率を標本における積率で置き換える。

点推定量の求め方:積率法(2)

- 例
  - $X_i \sim_{iid} (\mu, \sigma^2)$ $(i = 1, 2, \dots, n)$
  - $\mu = E(X_i) = \mu_1$, $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$.
  - $\sigma^2 = E(X^2) - \{E(X)\}^2 = \mu_2 - \mu_1^2$,

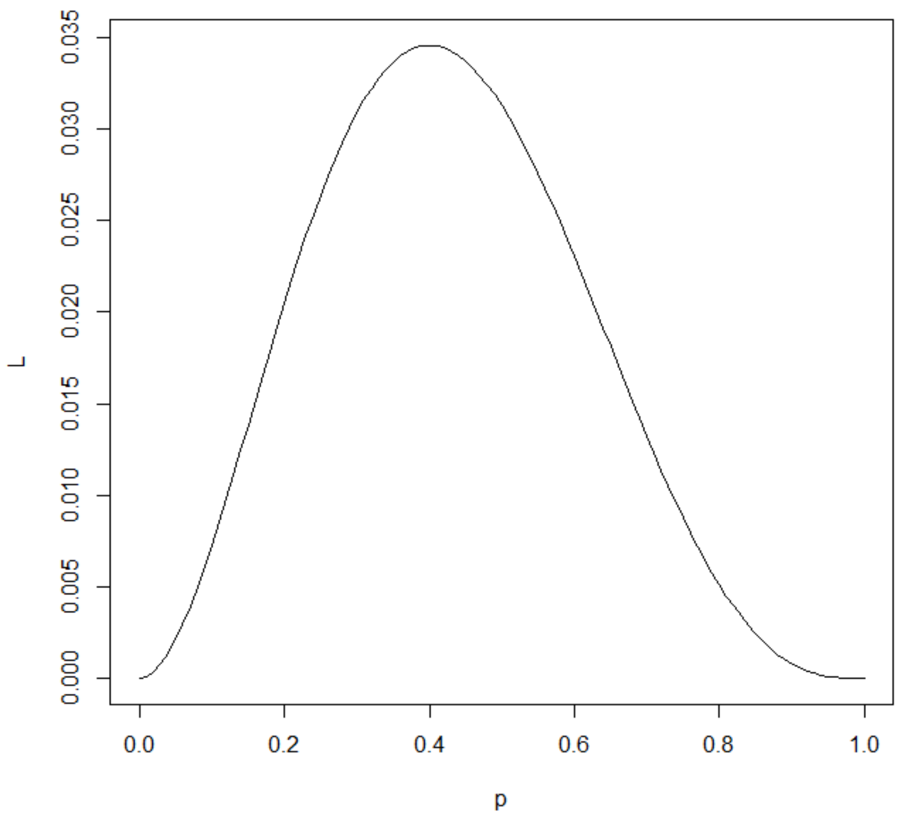
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

点推定量の求め方：最尤法(1)

- 尤度（関数）
  - 同時確率（密度）関数を、母数についての関数とみなしたもの。
    - 例： サイコロをn回投げて6が出る確率p を推定する。
      - $X_i \sim_{iid} Bernoulli(p)\ \ (i = 1, 2, \dots, n)$
      - $p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = p^{\sum x_i}(1-p)^{n-\sum x_i} = L(p;\ x_1, x_2, \dots, x_n)$
      - 数値例
        - » $n = 5$ として、$(x_1, x_2, x_3, x_4, x_5) = (1,0,1,0,0)$ のとき、
          - $L(p; 1,0,1,0,0) = p^2(1-p)^3.$

点推定量の求め方：最尤法(2)

図2：尤度関数

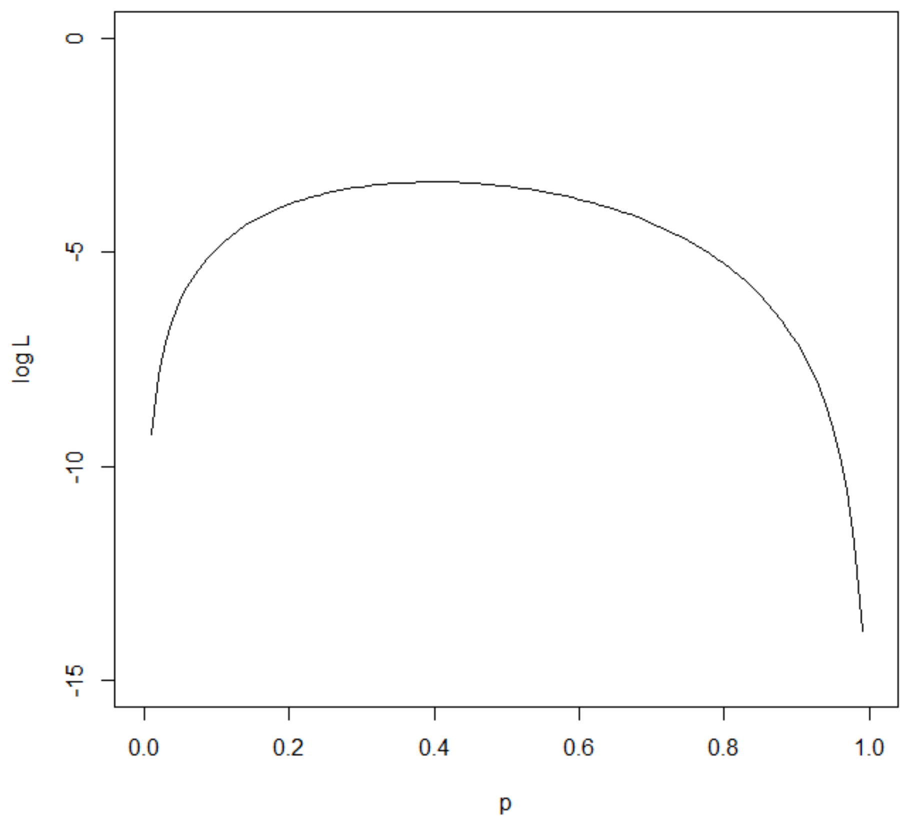


点推定量の求め方：最尤法(3)

- 対数尤度（関数）
  - 尤度関数の対数変換
    - 例：
      - $X_i \sim_{iid} Bernoulli(p) \ \ (i = 1, 2, \dots, n)$
      - $\log L(p;\ x_1, x_2, \dots, x_n)$
        $= (\sum_{i=1}^{n} x_i) \log p + (n - \sum_{i=1}^{n} x_i) \log(1 - p).$
      - 数値例
        - » $n = 5$ として、$(x_1, x_2, x_3, x_4, x_5) = (1,0,1,0,0)$ のとき、
          - $\log L(p; 1,0,1,0,0) = 2 \log p + 3 \log(1 - p).$

点推定量の求め方：最尤法(4)

図3：対数尤度関数



点推定量の求め方：最尤法(5)

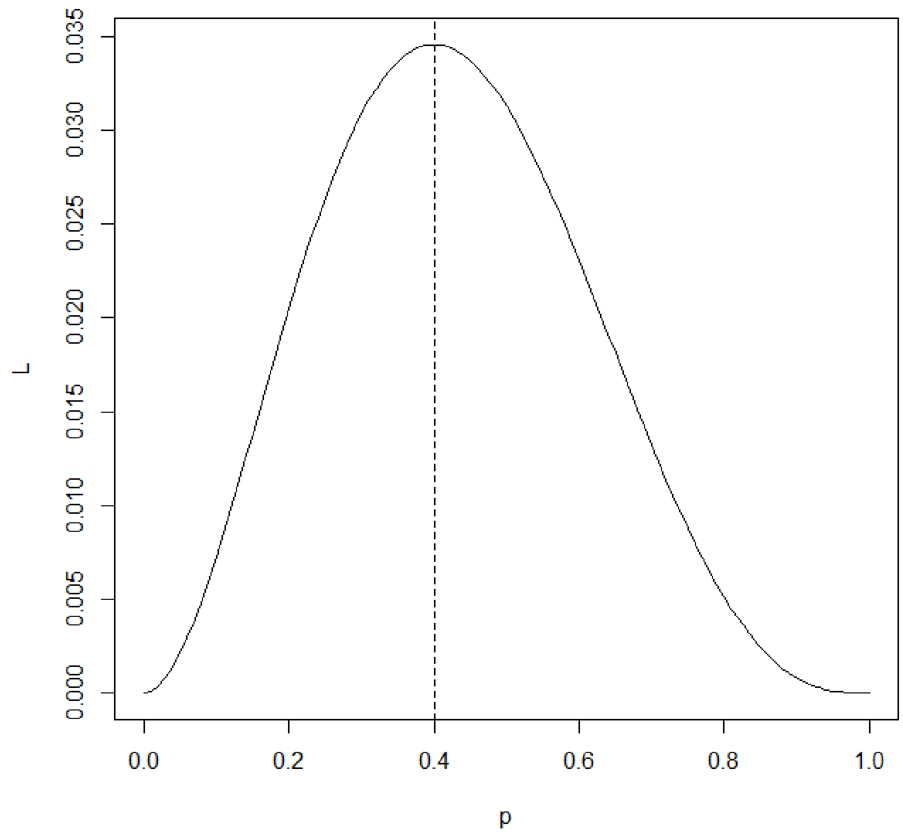
- 最尤法
  - 尤度関数を最大にする母数の値を母数の推定量とする推定方法。
    - 対数尤度関数を最大にする母数としても答えは同じ。
  - 最尤法による推定量を最尤推定量と呼ぶ。
    - 漸近的な（標本の大きさ n が大きいとき成り立つ）性質に優れている場合が多いことが知られている。
      - 漸近的に不偏である。
      - 一致性を持つ。
      - 漸近的に有効である。

点推定量の求め方：最尤法(6)

– 例
  - $X_i \sim_{iid} Bernoulli(p)\ \ (i = 1, 2, \dots, n)$
  - 最尤推定量
    – $\hat{p}$ such that $L(\hat{p};\ x_1, x_2, \dots, x_n) = max_p\ L(p;\ x_1, x_2, \dots, x_n)$.
      » $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$=（6の目が出た割合）
        - $\frac{d\,L}{d\,p} = (\sum x_i)p^{\sum x_i - 1} - (n - \sum x_i)(1 - p)^{n - \sum x_i - 1} = 0$
          - これを解くと $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
        - $\frac{d\log L}{d\,p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0$
          - これを解くと $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

点推定量の求め方：最尤法(7)

図4：尤度関数と最尤推定値



点推定量の求め方：最尤法(8)

図5：対数尤度関数と最尤推定値