ECE 7251: Signal Detection and Estimation

Spring 2002
Prof. Aaron Lanterman
Georgia Institute of Technology

Lecture 8, 1/23/02:
Nonrandom Parameter Estimation

What Can We Do?

- No prior! q is not a random variable.
- How about "conditional" MSE?

$$MSE_q = E_q[(\hat{q}(\underbrace{Y}_{}) - q)^2]$$

- Trivial estimator: $Y \sim p(y;q)$

$$\hat{q}(Y) = q_0$$

- Does really well if $q = q_0$, but might suck otherwise

Consistency

- Consistency: for all q, for any $e>0$

$$\lim_{n\to\infty} \Pr_q(|\underbrace{\hat{q}(\overbrace{Y}^{Y \sim p(y;q)})}_{\text{func. of } n} - q| > e) = 0$$

"convergence in probability"

- Useful result: *If*
  - $a = g(q)$, where g is a continuous func., and
  - $\hat{q}$ is consistent for q,

  then

  $\hat{a} = g(\hat{q})$ is consistent for a

More Definitions

- Estimator Bias: $b_q(\hat{q}) = E_q[\hat{q}(Y)] - q$
- Unbiasedness: for all q,

$$b_q(\hat{q}) = 0, \text{ i.e. } E_q[\hat{q}(Y)] = q$$

- Estimator Variance:

$$\mathrm{var}_q(\hat{q}) = E_q[(\hat{q}(Y) - E_q[\hat{q}(Y)])^2]$$

Decomposition of MSE

$$MSE_q(\hat{q}) = E_q[(\hat{q}(Y) - q)^2]$$

$$= E_q[\{(\hat{q}(Y) - E_q[\hat{q}(Y)]) + (E_q[\hat{q}(Y)] - q)\}^2]$$

$$= \overbrace{E_q[(\hat{q}(Y) - E_q[\hat{q}(Y)])^2]}^{\mathrm{var}_q(\hat{q})} + \overbrace{(E_q[\hat{q}(Y)] - q)^2}^{b_q(\hat{q})}$$

$$+2E_q[\hat{q}(Y) - E_q[\hat{q}(Y)]](E_q[\hat{q}(Y)] - q)$$ (struck out)

$$= \mathrm{var}_q(\hat{q}) + b_q^2(\hat{q})$$

Often have a "bias-variance tradeoff"

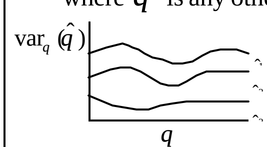
MVU Estimators

- Minimum variance unbiased estimator:

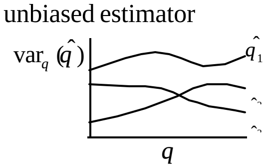
$$E_q[\hat{q}(Y)] = q$$

$$\mathrm{var}_q(\hat{q}) \le \mathrm{var}_q(\breve{q}) \text{ for all } q$$

where $\breve{q}$ is any other unbiased estimator



$\hat{q}_3$ is the MVUE.



There is no MVUE!

Stuff about MVU Estimators

- No easy way to know if they exist
- No easy way to find them if they do exist
- Closest thing to a turn-the-crank procedure:
  1. Find some (not necessarily so great) unbiased estimator $\breve{q}$
  2. Find a <u>complete</u> sufficient statistic *T(Y)*
  3. Apply <u>Rao-Blackwell-Lehmann-Scheffe</u> theorem:

$$\hat{q} = E[\breve{q} \mid T(Y)]$$

Best (?) explanation: Kay, Vol. I, Ch. 5

Useful Recipes

Method of Moments

Karl Pearson

Maximum Likelihood

Sir R.A. Fisher

Pics from MacTutor History of Mathematics Archive (http://www-groups.dcs.st-andrews.ac.uk/~history)

Maximum Likelihood Estimators

- Likelihood: $p(y;q)$
- Loglikelihood: $l(q) = \ln p(y;q)$

  (often drop terms which are constant with respect to q)
- ML estimator:

$$\hat{q}_{ML}(y) = \underset{q}{\operatorname{argmax}}\, p(y;q) = \underset{q}{\operatorname{argmax}}\, l(q)$$

- Sometimes cannot be found in closed form
- Most commonly used nonrandom parameter estimator in engineering literature

Some Properties of ML Estimators

- Consistent
- Asymptotically unbiased: As $n \to \infty$,

$$E_q[\hat{q}(Y)] \to q$$

- Asymptotically MVUE
- Invariant to functional transformations:

  If $a = g(q)$,

  then $\hat{a}_{ML} = g(\hat{q}_{ML})$
- Asymptotically <u>efficient</u> (to be discussed later)

Method of Moments

- Quick, cheap & dirty
- Define moments:

$$m_k(q) = E[Y^k] = \int y^k p(y;q) dq$$

- Find K moments such that

$$g(q) = [m_1(q), \ldots, m_k(q)]$$

  can be inverted for q

$$q = g^{-1}([m_1(q), \ldots, m_k(q)])$$

Method of Moments Con't.

- Just replace moments with empirical estimates:

$$\hat{q}_{MoM}(y) = g^{-1}([\hat{m}_1(q), \ldots, \hat{m}_k(q)])$$

$$\hat{m}_k(q) = \frac{1}{n} \sum_{i=1}^{n} y_i^k$$

- If moments are smooth functions of q, MoM estimator is
  - Asymptotically unbiased
  - Consistent
  - No other grand wonderful properties in general

Comments on Method of Moments

- If regular moments don't exist (i.e. Cauchy, other heavy-tailed distributions) can use "fractional" moments
- MoM often used to get initial value for iterative likelihood maximization techniques
- MoM doesn't always give unique estimator (different choices of moments gives different answers!)
- More details: Kay, Vol. I, Ch. 9