

Lecture XII: Likelihood Principle

Charles B. Moss

May 28, 2013

- I. We start by assuming $\omega = g(\theta_A)$ – say we are interested in the mean of the posterior distribution.

- A. The posterior of the ω can then be expressed as

$$\begin{aligned} E[\omega | y^0, A] &= E[g(\theta_A) | y^0, A] \\ &= \int_{\Theta_A} g(\theta_A) p(\theta_A | y^0, A) d\nu(\theta_A) \\ &= \frac{\int_{\Theta_A} g(\theta_A) L(\theta_A; y^0, A) p(\theta_A | A) d\nu(\theta_A)}{\int_{\Theta_A} L(\theta_A; y^0, A) p(\theta_A | A) d\nu(\theta_A)} \end{aligned} \tag{1}$$

1. In forming the posterior distribution we only have information on the data beyond the $L(\theta_A; y^0, A)$.
 2. All the information about $g(\theta_A)$ is contained in the likelihood function.
- B. "The likelihood principle ... states that if two data-generating mechanisms lead to the same likelihood function, then they convey the same evidence about the unknown parameters (Geweke)."
1. **Definition 3.5.1** An *experiment* is characterized by the triplet

$$E = \{ y, \theta_A, p(y | \theta_A, A) \} . \tag{2}$$

2. **Definition 3.5.2** The *evidence* about θ_A arising from E and the realization $y = y^0$, denoted $E\nu(E, y^0)$, is any inference, conclusion, or report concerning θ_A based on E and $y = y^0$.

3. **Definition 3.5.3 Weak Sufficiency Principle** Suppose that $s(y; A)$ is a sufficient statistic in the experiment $E = \{y, \theta_A, p(y|\theta_A, A)\}$. Let two runs of the experiment result in realizations y_1^0 and y_2^0 , respectively, and suppose $s(y_1^0; A) = s(y_2^0; A)$. Then $E\nu(E, y_1^0) = E\nu(E, y_2^0)$.
4. **Definition 3.5.3** states that two actions with the same sufficient statistics are equally valid.
5. **Definition 3.5.4** Given two experiments

$$E_j = \{y_j, \theta_A, p(y_j|\theta_A, A_j)\} \quad j = 1, 2 \quad (3)$$

involving the same parameter vector θ_A , a *mixed experiment*, based on E_1 and E_2 , is $E_* = [y_*, \theta_A, p(j, y_j|\theta_A, A)]$ where the random vector $y_* = (j, y_j)$, $A = (A_1, A_2)$, and $p(j, y_j|\theta_A, A) = 0.5p(y_j|\theta_A, A_j)$.

6. **Definition 3.5.5 Weak Conditionality Principle** Consider two experiments $E_j = \{y_j, \theta_A, p(y_j|\theta_A, A_j)\}$ ($j = 1, 2$), as well as the mixed experiment $E_* = [y_*, \theta_A, p(j, y_j|\theta_A, A)]$. Then $E\nu[E_*, (j, y_j)] = E\nu(E_j, y_j)$.
- C. **Definition 3.5.6 Likelihood Principle** Consider two experiments $E_j = [y_j, \theta_A, p(y_j|\theta_A, A_j)]$ ($j = 1, 2$). Suppose that for the particular realizations y_1^0 and y_2^0 , the respective likelihood functions satisfy

$$L(\theta_A; y_1^0, A_1) \propto L(\theta_A; y_2^0, A_2). \quad (4)$$

Then $E\nu(E_1, y_1^0) = E\nu(E_2, y_2^0)$.

- D. **Theorem 3.5.1 Equivalence of the Likelihood Principle, and the Weak Sufficiency and Weak Conditionality Principles** The weak sufficiency principle and the weak conditionality principle are together equivalent to the likelihood principle.
1. **Definition 3.5.7** A *sequential experiment* is an experiment in which the *stopping time* T is random.
 2. **Definition 3.5.8** A *stopping rule* in a sequential experiment is a sequence of probabilities $\tau = \{\tau_t\}_{t=0}^{\infty}$ in which τ_0 is constant and

$$\tau_m = P(T = m | Y_m, \theta_A, A) = P(T = m | Y_m, A) = \tau_m(Y_m; A). \quad (5)$$

The stopping probability may depend on the observables Y_m , but not the unobservables (i.e., the parameters θ_A).

3. Corollary 3.5.1 Stopping Rule Corollary of the Likelihood Principle In a sequential experiment $E \nu [E, (T, Y_T)]$ depends on (T, Y_T) only through $L(\theta_A; Y_T, A)$. The likelihood principle implies that the stopping rule τ is irrelevant.

II. Stopping rules will become important in simulating the posterior distribution – how many observations to draw is actually a question of when do you stop.