Econometrics I, Estimation

Department of Economics
Stanford University

September, 2008

Part I
Confidence intervals

• An random interval $L(X), U(X)$ such that for prespecified probability $\alpha$:

$$P(\theta_0 \in (L(X), U(X))) = \alpha.$$  

• Confidence intervals with exact coverage are difficult to find.

• Approximate (asymptotic) confidence intervals:

$$\lim_{n \to \infty} P_n(\theta_0 \in (L_n(X_n), U_n(X_n))) = \alpha.$$  

• Ideally, want short intervals with precise coverage.

• Many ways to construct confidence intervals.

• Equal tail intervals, shortest intervals, one-sided intervals.

• We will focus on estimator-based intervals.
• Suppose $\sqrt{n} \left( \hat{\theta} - \theta_0 \right) \xrightarrow{d} N(0, \sigma^2)$.

• Usually $\sigma^2$ unknown but replaced by consistent estimate $\hat{\sigma}^2$.

$$\frac{\sqrt{n} \left( \hat{\theta} - \theta_0 \right)}{\hat{\sigma}} \xrightarrow{d} N(0, 1).$$

• Symmetric confidence interval: $P \left( Z > z_{\alpha/2} \right) = \alpha/2$.

$$\left( L(X) = \hat{\theta} - z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}}, U(X) = \hat{\theta} + z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}} \right).$$

• Asymptotically correct coverage:

$$P \left( \theta_0 \in (L(X), U(X)) \right) = P \left( |\hat{\theta} - \theta_0| \leq z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}} \right)$$

$$= P \left( \frac{\sqrt{n}|\hat{\theta} - \theta_0|}{\hat{\sigma}} \leq z_{\alpha/2} \right) \xrightarrow{n \to \infty} \alpha.$$
• Prior $\pi(\theta)$. likelihood $f(x|\theta)$.

• Posterior density

$$p(\theta|x) = \frac{f(x|\theta) \pi(\theta)}{\int f(x|\theta) \pi(\theta) \, d\theta}.$$

• In general, computing $p(\theta|x)$ is difficult.

• Exception: conjugate family. Let $\mathcal{F}$ denote the class of likelihoods $f(x|\theta)$. A class $\Pi$ of prior distributions is a conjugate family for $\mathcal{F}$ if the posterior distribution is in the class $\Pi$ for all $f \in \mathcal{F}$, all priors in $\Pi$, and all $x \in \mathcal{X}$.

• The conjugate family for the normal mean when variances are known is normal.
• \{X_t\}, t = 1, \ldots, n \text{ i.i.d. } X_t \sim N(\mu, \sigma^2). \sigma^2 \text{ known.}

• Prior \pi(\mu) \sim N(\mu_0, \lambda_0), \mu_0, \lambda_0 \text{ known.}

• Posterior distribution

\[ p(\mu|X) \sim N\left(\frac{\lambda^2 \bar{x} + \frac{\sigma^2}{n} \mu_0}{\lambda^2 + \frac{\sigma^2}{n}}, \frac{\frac{\sigma^2}{n} \lambda^2}{\lambda^2 + \frac{\sigma^2}{n}}\right). \]

• Write \( t_0 = 1/\lambda^2, \bar{t} = n/\sigma^2 \): precision parameters.

\[ p(\mu|X) \sim N\left(\frac{t_0 \mu_0 + \bar{t} \bar{x}}{t_0 + \bar{t}}, \frac{1}{t_0 + \bar{t}}\right) \]

• prior mean and sample mean are weighted by their precisions.

• Posterior precision sum of prior and data precisions.
• Bayesian point estimator.

• minimizes posterior expected loss functions:

\[
\hat{\theta} = \min_{\theta \in \Theta} \int \rho (\theta - \tilde{\theta}) p (\tilde{\theta}|x) \ d\tilde{\theta}.
\]

• If \( \rho(x) = x^2 \), square loss:

\[
\hat{\theta} = \int \tilde{\theta} p (\tilde{\theta}|x) \ d\tilde{\theta} \quad \text{posterior mean}.
\]

• In the normal example:

\[
\hat{\mu} = \frac{t_0 \mu_0 + \bar{t} \bar{x}}{t_0 + \bar{t}}.
\]

• Other posterior locations, or loss functions, can be used.

• Posteriors interval: region under \( p (\theta|x) \) with a given area.