Bayesiánská analýza

I. Základní principy a pojmy bayesiánské ekonometrie

Obsah tématu

Bayesiánský přístup – motivace

- Příklad pro ilustrace základních principů:

  Eddy, Sean R. (2004) – What is Bayesian statistics?

Obsah tématu

Úvod

- Cíl: obecné principy bayesiánské ekonometrie.
- Koop (2003) – kapitola 1.
- Ekonometrie:
  - odhad parametrů modelu (regresní koeficienty);
  - porovnání různých modelů (testování hypotéz);
  - předpověď.
- Bayesiánská ekonometrie – využití několika jednoduchých pravidel pravděpodobnosti = univerzálnost.

Základní principy I

- Nechť A a B označují dva jevy, $p(B|A)$ je podmíněná pravděpodobnost jevu B za podmínky realizace jevu A; souhrn toho co víme o B známe-li A.
- Bayesiánský přístup – $A =$ něco známého (např. data); $B =$ něco neznámého (např. koeficienty v modelu).
- Nechť y jsou data, y^* nepozorovaná data (tj. předpověď), M_i pro $i = 1, \ldots, m$ je množina modelů za předpokladu, že každý závisí na parametrech θ^i.
    - Znalosti o parametrech v daném modelu – posteriorní hustota $p(\theta^i|M_i, y)$.
    - Porovnání modelů – posteriorní pravděpodobnost modelu $p(M_i|y)$.
    - Předpověď – předpovědní hustota $p(y^*|y)$.

Základní principy II

- Podmíněná pravděpodobnost A za podmínky B, označována $Pr(A|B)$ je pravděpodobnost realizace jevu A za podmínky, že nastal jev B.
- Pravidla pro podmíněnou pravděpodobnost:

$$Pr(A|B) = \frac{Pr(A,B)}{Pr(B)},$$

$$Pr(B|A) = \frac{Pr(A,B)}{Pr(A)}.$$

Bayesův teorém

$$Pr(A|B) = \frac{Pr(B|A)Pr(A)}{Pr(B)}$$

- Lze interpretovat nejen pro dva jevy ale i pro dvě náhodné veličiny.
- Nahrazení $Pr()$ pravděpodobnostní funkcí či hustotou pravděpodobnosti.

Odhad I

- Jeden model závisející na parametrech θ (např. regresní model s vysvětlujícími proměnnými).
- Zajímá nás posteriorní hustota a její vlastnosti $p(\theta|y)$.
- Bayesovo pravidlo:
$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}$$
- Pro náš model:
$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

Odhad II

- $p(\theta|y)$ – otázka „Při daných datech, co víme o θ?".
- Bayesiánský přístup – θ je náhodná veličina.
- Klasický přístup – θ není náhodná veličina.
- V rámci odhadu lze ignorovat $p(y)$, neboť neobsahuje θ:

$$p(\theta|y) \propto p(y|\theta)p(\theta).$$

Popis výrazů

- $p(\theta|y)$ – posteriorní hustota pravděpodobnosti.
- $p(y|\theta)$ – věrohodnostní funkce.
- $p(\theta)$ – apriorní hustota pravděpodobnosti.
- $\propto$ – „je proporcionální vzhledem“.
- $p(\theta)$ – nezávisí na datech, obsahuje nedatovou informaci o θ.

Pojetí pravděpodobnosti a nejistoty

- „Frekvencistický“ (klasický) přístup – pravděpodobnost nastání nějaké události je limitou její relativní četnosti (četnosti, se kterou událost nastává pokud se neomezeně zvyšuje množství dostupných dat).
- Problém v praxi při studiu zřídka se opakujících jevů (potřeba využití teoretických výsledků).
- Bayesovský přístup – subjektivní chápání pravděpodobnosti vyjadřující stupeň víry (přesvědčení), který je aktualizován po získání dodatečných dat.
- „Frekvencistická nejistota“ – zdrojem nejistoty je náhodnost spjatá s realizací náhodné veličiny (pravděpodobnostní rozdělení proměnných není předmětem nejistoty).
- „Bayesovská nejistota“ – pravděpodobnostní rozdělení samo o sobě nejisté (jeho možné změny a modifikace po získání nových informací) $\rightarrow$ nejistota je implicitně spojená s aktualizací pravděpodobnosti (apriorní pravděpodobnost $\rightarrow$ posteriorní pravděpodobnost).

Apriorní informace

- Kontraverzní – nevědeckost?
- Máme-li apriorní informaci, měli bychom ji využít.
- Lze využít i neinformativní priory.
- Empirické bayesiánské metody – prior z dat.
- Citlivostní analýza pro priory.

Predikce

- Prediktivní hustota pravděpodobnosti – $p(y^*|y)$.
- Integrací sdružené hustoty pravděpodobnosti – marginální hustota:

$$p(y^*|y) = \int p(y^*, \theta|y) d\theta$$

$$p(y^*|y) = \int p(y^*|y, \theta) p(\theta|y) d\theta$$

Porovnání modelů

- Modely M_i pro $i = 1, \ldots, m$ závisí na parametrech θ^i.
- Posteriorní pravděpodobnost modelů $p(M_i|y)$.
- Bayesovo pravidlo – $B = M_i$ a $A = y$:

$$p(M_i|y) = \frac{p(y|M_i)p(M_i)}{p(y)}$$

- $p(M_i)$ – apriorní pravděpodobnost modelu.
- $p(y|M_i)$ – marginální věrohodnost.

Výpočet marginální věrohodnosti

- Posteriorní hustotu pravděpodobnosti lze zapsat jako:

$$p(\theta^i|y, M_i) = \frac{p(y|\theta^i, M_i)p(\theta^i|M_i)}{p(y|M_i)}$$

- Integrací obou stran dle θ^i lze získat:

$$p(y|M_i) = \int p(y|\theta^i, M_i)p(\theta^i|M_i)d\theta^i$$

- Marginální pravděpodobnost modelu – prior a věrohodnostní funkce.

Posteriorní podíl šancí

- Posteriorní podíl šancí:

$$PO_{ij} = \frac{p(M_i|y)}{p(M_j|y)} = \frac{p(y|M_i)p(M_i)}{p(y|M_j)p(M_j)}$$

- Posteriorní podíl šancí a marginální pravděpodobnost modelu.
- Bayesův faktor:

$$BF_{ij} = \frac{p(y|M_i)}{p(y|M_j)}$$

Obsah tématu

Prezentace výsledků

- Analogicky ke klasickému přístupu – bodové odhady a konfidenční intervaly.
- Obvyklý bodový odhad – posteriorní střední hodnota.
- Nechť θ je k-prvkový vektor, $\theta = (\theta_1, \ldots, \theta_k)'$. Posteriorní střední hodnota každého prvku vektoru je dána jako:

$$E(\theta_i|y) = \int \theta_i p(\theta|y) d\theta$$

Očekávaná hodnota I

- Nechť $g()$ je funkce, pak očekávaná (střední) hodnota $E[g(X)]$ je definována pro diskrétní náhodnou veličinu X jako:

$$E[g(X)] = \sum_{i=1}^{N} g(x_i)p(x_i)$$

- Pro spojitou náhodnou veličinu X (za podmínky $E[g(X)] < \infty$):

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)p(x)dx$$

Očekávaná hodnota II

- Měřítkem disperze je posteriorní standardní odchylka, tedy odmocnina posteriorního rozptylu počítaného jako:

$$var(\theta_i|y) = E(\theta_i^2|y) - E(\theta_i|y)^2$$

- Vyžaduje vyhodnotit posteriorní střední hodnotu a rovněž také:

$$E(\theta_i^2|y) = \int \theta_i^2 p(\theta|y) d\theta$$

- Pravděpodobnost, že koeficient je kladný:

$$p(\theta_i \geq 0|y) = \int_0^\infty p(\theta|y)$$

Charakteristiky posteriorní hustoty

- Většina charakteristik má formu:

$$E[g(\theta)|y] = \int g(\theta)p(\theta|y)d\theta$$

- $g(\theta)$ – funkce, která nás zajímá.
- Všechny vlastnosti – forma integrálu, stejně jako marginální věrohodnost a prediktivní hustota.
- Až na výjimky - nemožnost vyhodnocení integrálu analyticky.

Posteriorní simulace

- Integrály v Bayesiánské analýze vyhodnocovány simulačními metodami.
- Zákon velkých čísel a centrální limitní věta.
- Zákon velkých čísel – předpokládejme náhodný výběr s posteriorní hustoty $\theta^{(1)}, \ldots, \theta^{(S)}$. Pro S jdoucí k nekonečnu konverguje výběrový průměr ke střední hodnotě $E[\theta|y]$.
- Bayesiánský přístup – využívá asymptotické teorie, ovšem asymptotické v S (pod kontrolou výzkumníka).

Monte Carlo integrace

- Nechť $\theta^{(s)}$ pro $s = 1, \ldots, S$ je náhodný výběr z $p(\theta|y)$ a definujme:

$$\hat{g}_S = \frac{1}{S} \sum_{s=1}^{S} g(\theta^{(s)}).$$

- $\hat{g}_S$ konverguje k $E[g(\theta)|y]$ pro S jdoucí k nekonečnu.
- MC integrace – aproximace $E[g(\theta)|y]$, jen pro $S = \infty$ jde chyba aproximace k nule.
- Volba S na nás – výpočetní nároky.
- Pro odhad chyby aproximace - centrální limitní věta $\rightarrow$ numerická standardní chyba (NSE).

Počítačový software

- Většinou vlastní programy v prostředí Matlab, Gauss či R.
- Balíčky BUGS (Bayesian Analysis Using Gibbs Sampling) – omezené typy modelů.
- Nástroje pro speciální typ modelů (makroekonomické modely) – Dynare toolbox a IRIS toolbox pro Matlab (případně Octave).