定理 1 の証明

新納浩幸

2003/8/1

$\bar{p}_n(x) \equiv E_{X^n}\{p_n(x)\}$ と定義。

ここで $p_n(x)$ は $\frac{1}{h^M}p(\frac{x-X_i}{h})$ の標本平均であることに注意すると、$p_n(x)$ の平均は $\frac{1}{h^M}p(\frac{x-X_i}{h})$ の平均に等しいので、

$$\bar{p}_n(x) = \int \frac{1}{h^M}p(\frac{x-y}{h})q(y)dy$$

が成立する。さらに、$z = (x-y)/h$ とおくと、以下が成立する ($dy = -hdz$ だが、積分の範囲が逆転するので、マイナスが消えることに注意)。

$$\bar{p}_n(x) = \int \frac{1}{h^M}p(\frac{x-y}{h})q(y)dy = \int p(z)q(x-hz)dz \quad (3.10)$$

ここから $\bar{p}_n(x)$ は、n に依存しないことがわかる (当たり前)。次に、

$$\begin{aligned}
E(n,h) &\equiv E_{X^n}\{\int (q(x)-p_n(x))^2dx\} \\
&= E_{X^n}\{\int (q(x)-\bar{p}_n(x)+\bar{p}_n(x)-p_n(x))^2dx\} \\
&= E_{X^n}\{\int (q(x)-\bar{p}_n(x))^2dx + 2\int (q(x)-\bar{p}_n(x))(\bar{p}_n(x)-p_n(x))dx + \int (\bar{p}_n(x)-p_n(x))^2dx\} \\
&= E_{X^n}\{\int (q(x)-\bar{p}_n(x))^2dx\} + E_{X^n}\{2\int (q(x)-\bar{p}_n(x))(\bar{p}_n(x)-p_n(x))dx\} + E_{X^n}\{\int (\bar{p}_n(x)-p_n(x))^2dx\}
\end{aligned}$$

左辺の E_{X^n} の中の第 1 項は n に依存していないので、

$$E_{X^n}\{\int (q(x)-\bar{p}_n(x))^2dx\} = \int (q(x)-\bar{p}_n(x))^2dx$$

次に第 2 項も、$q(x)-\bar{p}_n(x)$ が n に依存しないことに注意すると、

$$E_{X^n}\{2\int (q(x)-\bar{p}_n(x))(\bar{p}_n(x)-p_n(x))dx\} = 2\int (q(x)-\bar{p}_n(x))E_{X^n}\{\bar{p}_n(x)-p_n(x)\}dx = 0$$

以上より、

$$\begin{aligned}
E(n,h) &= \int (q(x)-\bar{p}_n(x))^2dx + E_{X^n}\{\int (\bar{p}_n(x)-p_n(x))^2dx\} \\
&= \int (q(x)-\bar{p}_n(x))^2dx + \int E_{X^n}\{(\bar{p}_n(x)-p_n(x))^2\}dx \\
&= \int B(x)^2dx + \int C(x)dx
\end{aligned}$$

ただしここで、

$$B(x) = q(x) - \bar{p}_n(x)$$

$$C(x) = E_{X^n}\{(\bar{p}_n(x) - p_n(x))^2\}$$

次に $B(x)$ と $C(x)$ のそれぞれについて考える。
まず $B(x)$ について、(3.10) 式を使うと、

$$\begin{aligned} B(x) &= q(x) - \bar{p}_n(x) \\ &= q(x) - \int p(z)q(x - hz)dz \\ &= \int p(z)q(x)dz - \int p(z)q(x - hz)dz \\ &= \int p(z)(q(x) - q(x - hz))dz \end{aligned}$$

ここでテイラー 展開[1]を利用すると、ある θ が存在して、

$$B(x) = \int p(z) \left(\sum_{i=1}^{M} hz_i \frac{\partial}{\partial x_i} q(x) - \frac{h^2}{2} \sum_{i,j=1}^{M} z_i z_j \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} q(x + \theta) \right) dz$$

今、仮定から $\int xp(x)dx = 0$ なので第 1 項は 0 となる。そして $n \to \infty$ とするとき、$h \to 0$、そして θ は $x < x + \theta < x - hz$ の範囲[2]でとられているので $\theta \to 0$ となる。

以上より、$B(x)$ は以下のように近似できる。

$$B(x) = -\frac{h^2}{2} \sum_{i,j=1}^{M} \int z_i z_j p(z)dz \frac{\partial^2 q(x)}{\partial x_i \partial x_j} = -\frac{h^2}{2} \sum_{i,j=1}^{M} P_{ij} \frac{\partial^2 q(x)}{\partial x_i \partial x_j}$$

次に $C(x)$ について考える。
$p_n(x)$ は $\frac{1}{h^M} p(\frac{x - X_i}{h})$ の標本平均であることに注意すると、中心極限定理から、$\sqrt{n}(\bar{p}_n(x) - p_n(x))$ は、平均が 0 分散が以下の式で表せる $V(X)$ に法則収束する。

$$V(x) \equiv E_{X^n}\{\left(\frac{1}{h^M} p(\frac{x - X_i}{h}) - \bar{p}_n(x) \right)^2\}$$

これは以下のように変形できる。

$$\begin{aligned} V(x) &\equiv E_{X^n}\{\left(\frac{1}{h^M} p(\frac{x - X_i}{h}) - \bar{p}_n(x) \right)^2\} \\ &= \frac{1}{h^M} \int p(z)^2 q(x - hz)dz - \{\int p(z)q(x - hz)dz\}^2 \end{aligned}$$

この部分は以下のように示せる。

[1] テキストでは平均値の定理だけどテイラー 展開の方が直接的。正則性から 2 回部分可能であることも注意。
[2] テイラー 展開は $a < b$ の関係が逆転しても成立することに注意。

$$
\begin{aligned}
V(x) &\equiv E_{X^n}\{\left(\frac{1}{h^M}p(\frac{x-X_i}{h}) - \bar{p}_n(x)\right)^2\} \\
&= \int \left(\frac{1}{h^M}p(\frac{x-y}{h}) - \bar{p}_n(x)\right)^2 q(y)dy
\end{aligned}
$$

今、$\bar{p}_n(x)$ は n に無関係なので、簡単化のために μ とおく。$z = (x-y)/h$ で変数変換すると、

$$
\begin{aligned}
V(x) &= \int \left(\frac{1}{h^M}p(\frac{x-y}{h}) - \bar{p}_n(x)\right)^2 q(y)dy \\
&= \int \left(\frac{1}{h^M}p(z) - \mu\right)^2 h^M q(x-hz)dz \\
&= \frac{1}{h^M}\int (p(z) - h^M\mu)^2 q(z)dz \\
&= \frac{1}{h^M}\int (p(z)^2 - 2h^M\mu p(z) + h^{2M}\mu^2) q(x-hz)dz \\
&= \frac{1}{h^M}\left(\int p(z)^2 q(x-hz)dz - 2h^M\mu\int p(z)q(x-hz)dz + h^{2M}\mu^2\int q(x-hz)dz\right)
\end{aligned}
$$

ここで第 2 項は、$\int p(z)q(x-hz)dz = \mu$ なので、$-2h^M\mu^2$。また第 3 項は、$y = x - hz$ で変数変換すると、

$$
\begin{aligned}
h^{2M}\mu^2\int q(x-hz)dz &= h^{2M}\mu^2\int \frac{1}{h^M}q(y)dy \\
&= h^M\mu^2\int q(y)dy \\
&= h^M\mu^2
\end{aligned}
$$

以上より、

$$
\begin{aligned}
V(x) &= \frac{1}{h^M}\int p(z)^2 q(x-hz)dz - \mu^2 \\
&= \frac{1}{h^M}\int p(z)^2 q(x-hz)dz - \{\int p(z)q(x-hz)dz\}^2
\end{aligned}
$$

$n \to \infty$、$h \to 0$ の極限を考えると、上の式の第 1 項は第 2 項よりもずっと大きい。また $C(x)$ は $p_n(x)$ の平均が $\bar{p}_n(x)$ なので、分散公式を利用すると、$V(\bar{p}_n(x) - p_n(x)) + E(\bar{p}_n(x) - p_n(x))^2$ に等しく、

$$
V(\bar{p}_n(x) - p_n(x)) = \frac{1}{n}V(\sqrt{n}(\bar{p}_n(x) - p_n(x))) = \frac{C(x)}{n}
$$

$$
E(\bar{p}_n(x) - p_n(x)) = 0
$$

なので、結局、$C(x) = \frac{V(x)}{n}$ が成立する。つまり、$C(x)$ は以下で近似できる。

$$
C(x) = \frac{q(x)}{nh^M}P_2
$$

以上をまとめると、$n \to \infty$、$h \to 0$ のとき、$E(n,h)$ は以下で近似できる。

$$E(n,h) = \frac{h^4}{4} \sum_{i,j,k,l=1}^{M} P_{ij} P_{kl} Q_{ij} Q_{kl} + \frac{1}{nh^M} P_2$$

n を固定して $E(n,h)$ を最小にする h を求める。左辺を h で偏微分して、極値になる点をもとめると、

$$h = \left(\frac{M P_2}{n \sum_{i,j,k,l=1}^{M} P_{ij} P_{kl} Q_{ij} Q_{kl}} \right)^{\frac{1}{M+4}}$$

を得る。