演習問題 2.08

同時確率が $p(\ x\ ,\ y\ )$ であるような 2 つの変数 x と y を考える。これについて、次の 2 つの結果を証明せよ。

$$\mathrm{E}[\ x\ ] = \mathrm{E}_y[\ \mathrm{E}_x[\ x \mid y\ ]\ ] \quad \cdots (\ 2.270\ )$$

$$\mathrm{var}[\ x\ ] = \mathrm{E}_y[\ \mathrm{var}_x[\ x \mid y\ ]\ ] + \mathrm{var}_y[\ \mathrm{E}_x[\ x \mid y\ ]\ ] \quad \cdots (\ 2.271\ )$$

ただし、$\mathrm{E}_x[\ x \mid y\ ]$ は、条件付き分布 $p(\ x \mid y\ )$ の下での x の期待値である。条件付き分散についても同様である。

[期待値]

ある関数 $f\,(\,x\,)$ の確率分布 $p\,(\,x\,)$ の下での平均値のこと。

$$\mathrm{E}[\,f\,] = \int p\,(\,x\,)\,f\,(\,x\,)\,dx \quad \cdots (\ 1.34\ )$$

[分散]

$f\,(\,x\,)$ がその平均値 $\mathrm{E}[\,f\,(\,x\,)\,]$ の周りでどれぐらいバラつくかを表す尺度のこと。

$$\mathrm{var}[\,f\,] = \mathrm{E}[\ (\,f\,(\,x\,) - \mathrm{E}[\,f\,(\,x\,)\,]\,)^2\ ] \quad \cdots (\ 1.38\ )$$

$$= \mathrm{E}[\,f\,(\,x\,)^2\,] - \mathrm{E}[\,f\,(\,x\,)\,]^2 \quad \cdots (\ 1.39\ )$$

[解]

まず、式 (2.270) を証明する。式 (2.270) の右辺は、式 (1.34) を用いて、

$$\mathrm{E}_y[\ \mathrm{E}_x[\ x \mid y\ ]\ ] = \int p(\,y\,)\,\mathrm{E}_x[\ x \mid y\ ]\,dy$$

$$= \int p(\,y\,) \int x\,p(\ x \mid y\ )\,dx\,dy$$

と展開できる。ここで、ベイズの定理と確率の乗法定理より、

$$p(\ x \mid y\ ) = \frac{p(\ y \mid x\ )\,p(\,x\,)}{p(\,y\,)} = \frac{p(\ x\ ,\ y\ )}{p(\,y\,)}$$

であるから、上記の式は、

$$= \int p(\,y\,) \int x\,\frac{p(\ x\ ,\ y\ )}{p(\,y\,)}\,dx\,dy$$

$$= \int\!\!\int x\,p(\ x\ ,\ y\ )\,dx\,dy$$

$$= \int x \int p(\,x,\ y\,)\,dy\,dx$$

となり、確率の加法定理より、さらに上記の式は、

$$= \int x\,p(\,x\,)dx$$

$$= \mathrm{E}[\,x\,]$$

となる。よって、式 (2.270) が示された。

[確率の加法定理]

$$p(\,x\,) = \int p(\,x,\ y\,)\,dy = \int p(\,x \mid y\,)\,p(\,y\,)\,dy$$

確率の乗法定

[確率の乗法定理]

$$p(\,x,\ y\,) = p(\,x \mid y\,)\,p(\,y\,) = p(\,y \mid x\,)\,p(\,x\,)$$

[ベイズの定理]

$$p(\,y \mid x\,) = \frac{p(\,x \mid y\,)\,p(\,y\,)}{p(\,x\,)}$$

↓ 確率の乗法定理

$$= \frac{p(\,x,\ y\,)}{p(\,x\,)}$$

↓ 確率の加法定理

$$= \frac{p(\,x,\ y\,)}{\int p(\,x,\ y\,)\,dy}$$

↓ 確率の乗法定理

$$= \frac{p(\,x,\ y\,)}{\int p(\,y \mid x\,)\,p(\,x\,)\,dy}$$

次に、式 (2.271) を証明する。式 (2.271) の右辺は、

$$\mathrm{E}_y[\,\mathrm{var}_x[\,x \mid y\,]\,] + \mathrm{var}_y[\,\mathrm{E}_x[\,x \mid y\,]\,] \qquad \cdots ※$$

であり、上記の式の第一項は、式 (1.39) を用いて、

$$\mathrm{E}_y[\,\mathrm{var}_x[\,x \mid y\,]\,] = \mathrm{E}_y[\,\mathrm{E}_x[\,(\,x \mid y\,)^2\,] - \mathrm{E}_x[\,x \mid y\,]^2\,]$$

$$= \mathrm{E}_y[\,\mathrm{E}_x[\,(\,x \mid y\,)^2\,]\,] - \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]^2\,]$$

と展開でき、また第二項は、式 (1.39) を用いて、

$$\mathrm{var}_y[\,\mathrm{E}_x[\,x \mid y\,]\,] = \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]^2\,] - \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]\,]^2$$

と展開できる。よって、式 ※ は、

$$= \mathrm{E}_y[\,\mathrm{E}_x[\,(\,x \mid y\,)^2\,]\,] - \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]^2\,] + \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]^2\,] - \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]\,]^2$$

$$= \mathrm{E}_y[\,\mathrm{E}_x[\,(\,x \mid y\,)^2\,]\,] - \mathrm{E}_y[\,\mathrm{E}_x[\,x \mid y\,]\,]^2$$

$\cdots$ ※※

と整理できる。ここで、上記の式の第一項は、式 (1.34) より、

$$\mathrm{E}_y[\ \mathrm{E}_x[\ (\ x\ |\ y\ )^2\ ]\ ] = \int\int x^2\, p(\ x\ |\ y\ )\, dx\, dy$$

$$= \int x^2 \int p(\ x\ |\ y\ )\, dy\, dx$$

$$= \int x^2\, p(\ x\ )\, dx$$

$$= \mathrm{E}_x[\ x^2\ ]$$

となり、第二項は、

$$\mathrm{E}_y[\ \mathrm{E}_x[\ x\ |\ y\ ]\ ]^2 = \mathrm{E}_x[\ x\ ]^2$$

となるので、最終的に式 ※※ は、

$$= \mathrm{E}_x[\ x^2\ ] + \mathrm{E}_x[\ x\ ]^2$$

となり、これは式 (1.39) を用いて、

$$= \mathrm{var}[\ x\ ]$$

と変形できる。よって、式 (2.271) が示された。