条件付き確率とベイズの定理

荒木 純道

2012/11/1

条件付き確率とベイズの定理

- 同時確率：2つの確率変数, $X, Y \rightarrow p(x,y)$
- 周辺確率： XもしくはYのみの確率 $\rightarrow p(x), p(y)$
- 条件付き確率

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

＊X,Yの周辺確率密度関数

$$p(x) = \int_{-\infty}^{\infty} p(x,y)dy$$

$$p(y) = \int_{-\infty}^{\infty} p(x,y)dx$$

定義から、ベイズの定理は

$$p(x,y) = p(y|x)p(x) = p(x|y)p(y)$$

特にXとYが独立なときには

$$p(y|x) = p(y), \quad p(x|y) = p(x)$$

条件付き確率とベイズの定理

- あたりくじがn 枚中r 枚入っている箱がある。
  この箱から１枚ずつくじを引く。（くじは箱に戻さない。）

１人目のA が当たる確率： $p(A) = \frac{r}{n}$

１人目のA がはずれる確率： $p(A^c) = \frac{n-r}{n}$

A があたったあとに、２人目のB が当たる確率： $p(B|A) = \frac{r-1}{n-1}$

A がはずれたあとに、２人目のB が当たる確率： $p(B|A^c) = \frac{r}{n-1}$

結局B が当たる確率は $p(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$

$$= \frac{r-1}{n-1}\cdot\frac{r}{n} + \frac{r}{n-1}\cdot\frac{n-r}{n} = \frac{r}{n} = P(A)$$

A,B のあたる確率はくじを引く順番によらない！

条件付き確率とベイズの定理

事象A が生じたという条件のもとで、事象B が生じる確率：

$$p(B|A) = \frac{P(A\cap B)}{P(A)}$$

A とB を入れ替えると、

$$p(A|B) = \frac{P(A\cap B)}{P(B)}$$

以上より、ベイズの定理は

$$p(A\cap B) = P(A|B)P(B) = P(B|A)P(A)$$

これを変形すると、

$$p(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

$$= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

A：原因、B：結果なら結果から原因を推測する式。

くじ引きの問題で、A,B の順に引いたが、A はまだ結果を見ていない。Bの結果はあたりであった。
A があたっている確率は？

$$p(A|B) = \frac{\frac{r-1}{n-1}\frac{r}{n}}{\frac{r-1}{n-1}\frac{r}{n} + \frac{r}{n-1}\frac{n-r}{n}} = \frac{r-1}{n-1}$$

条件付き確率とベイズの定理

$$\Omega = \bigcup_i A_i, \quad A_i \cap A_{i'} = \phi \quad (i \neq i')$$

$$\Omega = \bigcup_i B_j, \quad B_j \cap B_{j'} = \phi \quad (j \neq j')$$

$P(A_i \cap B_j) = p(A_i, B_j)$ ならば、

$$P(A_i|B_j) = \frac{p(A_i, B_j)}{p(A_1, B_j) + p(A_2, B_j) + \cdots}$$

$$= \frac{p(A_i, B_j)}{\sum_i p(A_i, B_j)}$$

条件付き確率とベイズの定理

- 確率密度関数による表現

確率変数 $X(\omega), Y(\omega) \in R$

事象A,B

$$A = \{\omega; x < X(\omega) < x + \Delta x\}$$

$$B = \{\omega; y < Y(\omega) < y + \Delta y\}$$

確率 $P(A) = p(x)\Delta x$

$$P(B) = p(y)\Delta y$$

$$P(A\cap B) = p(x,y)\Delta x\Delta y$$

条件付き確率 $P(B|A)$ を考えると、

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{p(x,y)}{p(x)}\Delta y$$

確率密度関数でのベイズの定理

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

多次元の場合 $p(x_1, x_2 | x_3, x_4, x_5) = \frac{p(x_1, x_2, x_3, x_4, x_5)}{\iint p(x_1, x_2, x_3, x_4, x_5)dx_1dx_2}$,

例）アイスクリーム売上げ vs. 温度

同時確率分布

Y

		15~25℃	25~30℃	30℃以上	
X	0~100個	0.02	0.1	0.2	0.32
	100~200個	0.02	0.3	0.1	0.42
	200~300個	0.06	0.1	0.1	0.26
		0.1	0.5	0.4	

$p(x), p(y)$

200～300 売り上げ → 温度が25～30℃ となる確率

$$p\left(\underset{y2}{Y} \mid \underset{x3}{X}\right) = \frac{p\left(\overset{x3}{X}, \overset{y2}{Y}\right)}{p\left(\underset{x3}{X}\right)} = \frac{0.1}{0.26} = \frac{10}{26} \approx 0.38$$

回帰曲線

- 確率変数X, Yの同時確率密度関数が $p(x, y)$ であるとき、条件X=xのもとでのYの平均値：

$$E[Y \mid x] = \int_{-\infty}^{\infty} y\, p(y \mid x) dy \quad = \int_{-\infty}^{\infty} y \frac{p(x, y)}{\int_{-\infty}^{\infty} p(x, y) dy} dy$$

- 関数$r(x) = E(Y \mid x)$を回帰曲線 or 回帰関数という。

回帰関数 $r(x) = E(Y \mid x)$ は、平均２乗誤差

$$\underset{\text{汎関数}}{L[f]} = E\left[(y - f(x))^2\right] = \iint (y - f(x))^2 \underbrace{p(x, y)}_{\text{既知}} dxdy$$

を最小化する関数である。なお、汎関数　functional: $f(\cdot) \to L[f]$

最小推定分散

Yの推定値 $\hat{y}$

$$v(x) = \int y^2 p(y \mid x) dx \qquad r(x) = E[Y \mid x] \text{ とおく}$$

$$L[f] = \int p(x) \int p(y \mid x) \cdot \left(y^2 - 2yf(x) + f^2(x)\right) dydx$$

$$= \int p(x) \left(v(x) - 2r(x) f(x) + f^2(x)\right) dx$$

$$= \int p(x) \left[\underbrace{\{f(x) - r(x)\}}_{\text{最適解：} f(x) = r(x)}^2 + v(x) - r^2(x)\right] dx$$

回帰曲線

$\hat{y} = r(x) = E[Y \mid x]$ の推定分散は最小（誤差）$= \int p(x)\left(v(x) - r^2(x)\right) dx$ になる。

また、不偏推定 $E(Y - \hat{y}) = 0$ である。