2025/02/21

藪友良

条件付き期待値と繰り返し期待値の法則

ここでは条件付き確率、条件付き期待値と分散、繰り返し期待値の法則を紹介します。以下では、Xのとりうる値を$\{x_1,x_2,\cdots,x_m\}$、Yのとりうる値を$\{y_1,y_2,\cdots,y_s\}$として議論をします。

本稿は、藪友良『入門　実践する統計学』(2012 年、東洋経済新報社)の補足資料です。同時確率と周辺確率に馴染みがない読者は、『入門　実践する統計学』の 5.4.1 節を参照してください。

確率変数の条件付き確率

事象の条件付き確率と同様(4.4.1 節参照)、確率変数の条件付き確率は$P\{Y=y_j|X=x_i\}$と表記し、次のように定義されます。

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}$$

これはXがx_iという値をとったという条件のもとで、Yがy_jという値になる条件付き確率です。

条件付き確率が、確率の公理を満たすことは簡単に確認できます(確率の公理は 4.2.4 節参照)。まず、周辺確率$P\{X=x_i\}$、同時確率$P\{X=x_i,Y=y_j\}$はともに 0 以上の値ですから、条件付き確率$P\{Y=y_j|X=x_i\}$も 0 以上です。次に、jに関して和をとると、条件付き確率の和は次のように 1 になります。

$$\sum_{j=1}^{s}P\{Y=y_j|X=x_i\}=\frac{\sum_{j=1}^{s}P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=1$$

式展開では、周辺確率の定義$P\{X=x_i\}=\sum_{j=1}^{s}P\{X=x_i,Y=y_j\}$を用いました[1]。

例 1（天気と通勤時間①）　5.4.1 節の例を思い出してください。Xは天気を表す変数であり、X=0 なら雨、X=1 なら晴れとします。また、Yは通勤時間を表す変数であり、Y=0 なら短時間、Y=1 なら中時間、Y=2 なら長時間で目的地に着く

[1] 公理 3 も同様に証明できますが、自明なので証明を省きました。

とします。下表の同時確率分布表は、同時確率と周辺確率をまとめたものです。たとえば、天気が雨で通勤時間が長時間である確率 $P\{X=0,Y=2\}$は 0.15 とわかります。また、天気が雨の確率は $P\{X=0\}$は 0.3 です。

		Y			
		0	**1**	**2**	
X	**0**	0.05	0.10	0.15	0.30
	1	0.50	0.15	0.05	0.70
		0.55	0.25	0.20	1.00

この表から、雨が降ったという条件のもとで(X=0)、短時間(Y=0)で目的地に着く確率は次のように計算されます。

$$P\{Y=0|X=0\}=\frac{P\{X=0,Y=0\}}{P\{X=0\}}=\frac{0.05}{0.3}=0.167$$

また、雨のとき(X=0)、中時間(Y=1)となる確率は$P\{Y=1|X=0\}=\frac{0.1}{0.3}=0.333$、雨のとき($X$=0)、長時間($Y$=2)となる確率は$P\{Y=2|X=0\}=\frac{0.15}{0.3}=0.5$です。当然ですが、これらの和は 1 です。同様にして、天気が晴のとき(X=1)、Yの確率はそれぞれ$P\{Y=0|X=1\}=0.714$、$P\{Y=1|X=1\}=0.214$、$P\{Y=2|X=1\}=0.072$です。以上から、雨のとき通勤時間は長く、晴れのとき通勤時間が短いことがわかります。

事象の独立と同様(4.4.3 節参照)、確率変数 X、Y が独立なら同時確率は$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$となります。このとき、次式が成立します。

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i\}P\{Y=y_j\}}{P\{X=x_i\}}=P\{Y=y_j\}$$

この結果から、確率変数 X、Y が独立なら、通常の確率と条件付き確率が同じになることがわかります。これは X の情報が、Yの確率を考えるうえで何の意味もないためです。条件付き確率では、「独立な変数は条件から除くことができます」。

<u>条件付き期待値</u>

条件付き期待値$E[Y|X=x_i]$は、X が x_i という値をとったという条件のもとでの Y の期待値です。何らかの条件のもとでの期待値であり、たとえば、女性の所得の期待値、前日に株価が上昇したときの株価変化率の期待値などです。

条件付き期待値(conditional expectation)は、厳密には、Yのとりうる値y_jに、条件付き確率$P\{Y = y_j | X = x_i\}$を掛けて加重平均したものです。

$$E[Y|X = x_i] = \sum_{j=1}^{s} y_j P\{Y = y_j | X = x_i\}$$

例を通じて、条件付き期待値の理解を深めていきましょう。

例 2（天気と通勤時間②）　例 1 の数値例を用いて、雨のとき(X=0)、通勤時間 Yの期待値を求めます。

$$E[Y|X = 0] = 0 \times P\{Y = 0|X = 0\} + 1 \times P\{Y = 1|X = 0\} + 2 \times P\{Y = 2|X = 0\}$$

$$= 0 \times 0.167 + 1 \times 0.333 + 2 \times 0.5 = 1.333$$

同様に、晴れのとき(X=1)、通勤時間 Yの期待値は次のようになります。

$$E[Y|X = 1] = 0 \times P\{Y = 0|X = 1\} + 1 \times P\{Y = 1|X = 1\} + 2 \times P\{Y = 2|X = 1\}$$

$$= 0 \times 0.714 + 1 \times 0.214 + 2 \times 0.072 = 0.358$$

この結果からも、晴れのとき通勤時間が短くなることがわかります。

<u>条件付き期待値の性質</u>

ここで、$g(X)$をXの任意の関数としたとき、$g(X)Y$の条件付き期待値は次のようになります。

$$E[g(X)Y|X = x_i] = g(x_i)E[Y|X = x_i]$$

Xは定数x_iであるため、任意の関数$g(x_i)$も定数となり、期待値の外に出せます。たとえば、$g(X) = a + bX$とし、a、bは定数とすると次のようになります。

$$E[(a + bX)Y|X = x_i] = (a + bx_i)\,E[Y|X = x_i]$$

確率変数 X、Yが独立なら、$P\{Y = y_j | X = x_i\} = P\{Y = y_j\}$となりました。この結果から、条件付き期待値は次のようになります。

$$E[Y|X = x_i] = E[Y]$$

つまり、条件付き期待値では、独立な変数は条件から除くことができます。

条件付き期待値を$E[Y|X]$と表記し、条件付き期待値$E[Y|X]$を確率変数と見なすことがあります。ここで、Xの値はまだ決まっていないため、条件付き期待値$E[Y|X]$もまた確率変数です。これに対して、$E[Y|X = x_i]$は、Xの値がx_iと既に決まっているため、定数であり確率変数ではありません。たとえば、$Y = a + bX$のと

き、$E[Y|X] = a + bX$は確率変数である一方、$E[Y|X = x_i] = a + bx_i$は定数です。

繰り返し期待値の法則

条件付き期待値を用いた有名な定理として、繰り返し期待値の法則があります。**繰り返し期待値の法則**(law of iterated expectation)は、「Yの期待値$E[Y]$は条件付き期待値$E[Y|X]$の期待値に等しい」とします(証明は補足参照)。

繰り返し期待値の法則
$$E[Y] = E[E[Y|X]]$$

$E[Y|X]$の期待値とは、条件付き期待値の取りうる値$E[Y|X = x_i]$に、確率$P\{X = x_i\}$を掛けて加重平均したものです。つまり、繰り返し期待値の法則は、

$$E[Y] = \sum_{i=1}^{m} E[Y|X = x_i]P\{X = x_i\}$$

と表現できます。次の例を通じて理解を深めていきましょう。

例 3(男女の所得)　Yは所得、Xは性別を表す変数($X = 1$なら男性、$X = 0$なら女性)とします。繰り返し期待値の法則は、全体の期待所得$E[Y]$が、男性の期待所得$E[Y|X = 1]$と女性の期待所得$E[Y|X = 0]$に、それぞれの確率を掛けた加重平均になるとしています。

$$E[Y] = E[Y|X = 1]P\{X = 1\} + E[Y|X = 0]P\{X = 0\}$$

たとえば、20 代男性の平均収入は 377 万円、女性の平均収入は 330 万円、男女の割合は各 50%とすると、全体の平均年収は 353 万円$(= 377 \times 0.5 + 330 \times 0.5)$になります。

条件付き分散

条件付き分散$V(Y|X = x_i)$は、X が x_i という値をとったという条件のもとでの Y の分散になります。**条件付き分散**(conditional variance)は、次のように定義されます。

$$V(Y|X = x_i) = E[(Y - E[Y|X = x_i])^2|X = x_i]$$

これは、Y が条件付き期待値$E[Y|X = x_i]$から乖離している程度を$(Y - E[Y|X = x_i])^2$とし、条件付き期待値をとったものです。

条件付き分散の性質をいくつか紹介します。まず、XとYが独立なら、条件付き期待値の条件から独立な変数Xを外せるので、次のようになります。

$$V(Y|X=x_i)=E[(Y-E[Y])^2]=V(Y)$$

つまり、条件付き分散の計算でも、独立な変数は条件から除くことができます。

条件付き分散を$V(Y|X)$と表記し、確率変数とみなすことがあります。Xの値はまだ決まっていないため、条件付き分散$V(Y|X)$もまた確率変数です。これに対して、$V(Y|X=x_i)$は、Xの値がx_iと既に決まっているため、定数であり確率変数ではありません。

分散の公式を紹介します(証明は練習問題)。第 1 の公式は、5.3.1 節で紹介した分散の公式$V(X)=E[X^2]-E[X]^2$に対応した次式です。

$$V(Y|X)=E[Y^2|X]-E[Y|X]^2$$

第 2 の公式は、分散$V(Y)$を 2 つの項に分解するものです。

$$V(Y)=E[V(Y|X)]+V(E[Y|X])$$

第 1 項$E[V(Y|X)]$はグループ内分散、第 2 項$V(E[Y|X])$はグループ間分散と呼ばれます。たとえば、Yを所得、Xを性別($X=1$なら男性、$X=0$なら女性)としたとき、所得分散$V(Y)$は、男女別の所得分散の期待値(各グループ内でのばらつき)、男女別の所得の期待値の分散(グループ間でのばらつき)の和となります。

補足：証明

繰り返し期待値の法則の証明

Xのとりうる値を$\{x_1,x_2,\cdots,x_m\}$、Yのとりうる値を$\{y_1,y_2,\cdots,y_s\}$とします。このとき、$E[Y|X]=\sum_{j=1}^{s}y_jP\{Y=y_j|X=x_i\}$に注意すると、次のように展開できます。

$$\begin{aligned}E\big[E[Y|X]\big]&=\sum_{i=1}^{m}\left(\sum_{j=1}^{s}y_jP\{Y=y_j|X=x_i\}\right)P\{X=x_i\}\\&=\sum_{j=1}^{s}y_j\left(\sum_{i=1}^{m}\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}P\{X=x_i\}\right)\\&=\sum_{j=1}^{s}y_j\left(\sum_{i=1}^{m}P\{X=x_i,Y=y_j\}\right)\\&=\sum_{j=1}^{s}y_jP\{Y=y_j\}=E[Y]\end{aligned}$$

4行目の式展開は、周辺確率の定義$P\{Y = y_j\} = \sum_{i=1}^{m} P\{X = x_i, Y = y_j\}$を用いました。

練習問題

1. $V(Y|X) = E[Y^2|X] - E[Y|X]^2$を示せ。
2. $V(Y) = E[V(Y|X)] + V(E[Y|X])$を示せ。

練習問題1の答え

条件付き分散$V(Y|X = x_i)$の定義式から、次のようになる。

$$E[(Y - E[Y|X = x_i])^2|X = x_i] = E[(Y^2 + E[Y|X = x_i]^2 - 2E[Y|X = x_i]Y)|\, X = x_i]$$
$$= E[Y^2|\, X = x_i] + E[Y|X = x_i]^2 - 2E[Y|X = x_i]^2$$
$$= E[Y^2|\, X = x_i] - E[Y|X = x_i]^2$$

式展開では、$X = x_i$の条件のもとで$E[Y|X = x_i]$は定数とみなせるため、以下の関係式が成立するとした。

$$E[E[Y|X = x_i]^2|\, X = x_i] = E[Y|X = x_i]^2$$
$$E[E[Y|X = x_i]Y|\, X = x_i] = E[Y|X = x_i]E[Y|\, X = x_i]$$

練習問題2の答え

$V(Y|X) = E[Y^2|X] - E[Y|X]^2$となるため、次のように表せる。

$$E[V(Y|X)] = E[\, E[Y^2|X] - E[Y|X]^2]$$
$$= E\big[\, E[Y^2|X]\big] - E[\, E[Y|X]^2]$$

第1項は繰り返し期待値の法則から$E\big[\, E[Y^2|X]\big] = E[Y^2]$となる。第2項を考えよう。$E[Y|X]$は確率変数であり、$V(E[Y|X])$は分散の公式から

$$V(E[Y|X]) = E[E[Y|X]^2] - E\big[E[Y|X]\big]^2 = E[E[Y|X]^2] - E[Y]^2$$

となる(式展開では、繰り返し期待値の法則から$E\big[E[Y|X]\big] = E[Y]$となることを用いた)。上式を書き換えると、$E[E[Y|X]^2] = V(E[Y|X]) + E[Y]^2$となる。

$E[V(Y|X)] = E\big[\, E[Y^2|X]\big] - E[\, E[Y|X]^2]$の式に、$E\big[\, E[Y^2|X]\big] = E[Y^2]$と$E[E[Y|X]^2] = V(E[Y|X]) + E[Y]^2$を代入すると、

$$E[V(Y|X)] = E[Y^2] - (V(E[Y|X]) + E[Y]^2)$$
$$= (E[Y^2] - E[Y]^2) - V(E[Y|X])$$

$V(Y) = E[Y^2] - E[Y]^2$から、上式は$E[V(Y|X)] = V(Y) - V(E[Y|X])$であり、書き換えると$V(Y) = E[V(Y|X)] + V(E[Y|X])$となる。