今日の内容：確率変数 (教科書の 5 章)
確率変数の定義

標本空間 Ω と、Ω 上の確率 P が与えられているとする。Ω 上の確率変数 X とは、Ω を定義域とし実数を値域とする関数である。$\omega \in \Omega$ における関数 X の値を $X(\omega)$ と書く。言い換えれば、確率変数 X は標本点 ω を実数 $X(\omega)$ に対応させる関数である。
確率変数自体はただの関数なので、それだけでは確率変数がある値を取る確率を考えることができない。Ω 上の確率 P と一緒になってはじめて確率変数がある値を取る確率を考えられるようになる。
例えば $X(\omega) = x$ となる標本点 ω の集合を事象 A であらわすと、確率変数 X が値 x になる確率は $P(A)$ で表される。
また、X, Y が確率変数であるとき $X + Y$, X^2 なども確率変数である。また f を実関数として $f(X)$ も確率変数である。(上記の定義は教科書とは違う。教科書の定義は数学的に曖昧なので使用しなかった)

確率変数の例

前回の、赤玉と黄玉が１つずつ入っているつぼから玉を取り出して戻す操作を２回繰り返す例では、標本空間は

$$\Omega = \{(\text{赤}, \text{赤}), (\text{赤}, \text{黄}), (\text{黄}, \text{赤}), (\text{黄}, \text{黄})\}$$

であった。確率変数 X を $X((\text{赤}, \text{赤}))=1$, $X((\text{赤}, \text{黄}))=1$, $X((\text{黄}, \text{赤}))=0$, $X((\text{黄}, \text{黄}))=0$ とすると、X は１回目に赤玉を取り出したとき 1, それ以外のときに 0 を取る確率変数になる。
同様に確率変数 Y を $Y((\text{赤}, \text{赤}))=1$, $Y((\text{赤}, \text{黄}))=0$, $Y((\text{黄}, \text{赤}))=1$, $Y((\text{黄}, \text{黄}))=0$ とすると、X は２回目に赤玉を取り出したとき 1, それ以外のときに 0 を取る確率変数になる。確率変数 $X+Y$ は２回玉を取り出したときに赤玉が出た回数の合計になる。
玉を２回取り出す操作を行うと確率変数 X, Y の値が決まる。一般に標本点を１つ選ぶと確率変数の値が決まるが、このように標本点を１つ選んで決まった確率変数の値を**実現値**と呼ぶ。確率変数はアルファベット大文字、実現値は小文字で表わすことが多い。

累積分布関数

ある確率変数 X の実現値が x 以下になる標本点の集合は

$$\{\omega \in \Omega \mid X(\omega) \leq x\}$$

と書ける。したがって確率変数 X の実現値が x 以下になる確率は

$$P(\{\omega \in \Omega \mid X(\omega) \leq x\})$$

で表わされ、これは実数 x の関数になる。これを

$$P(X \leq x)$$

と省略して表記し、累積分布関数と呼ぶ。$P(X \leq x)$ を $F_X(x)$ と書くこともある。
例：前のページの赤玉と黄玉を取り出す例で確率変数 Z を 2 回の取り出しで赤玉が出た合計回数とする。このとき

$$\{\omega \in \Omega \mid Z(\omega) \leq 1.5\} = \{(\text{赤}, \text{黄}), (\text{黄}, \text{赤}), (\text{黄}, \text{黄})\}$$

であり、$F_Z(1.5) = P(Z \leq 1.5) = 0.75$ である。（黒板に F_Z を図示する）
累積分布関数 F_X は以下の性質を持つ。そのことを標本空間と確率の公理から証明できる。

1. F_X は右連続

2. $x_1 < x_2$ ならば $F_X(x_1) \leq F_X(x_2)$

3. $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to +\infty} F_X(x) = 1$.

また、$a \leq b$ のときに、X が a より大きく b 以下の値を取る確率は $F_X(b) - F_X(a)$ で与えられる。

離散型の確率変数

確率変数 X の取り得る値の集合が有限集合または可算無限集合のとき X を離散型の確率変数と呼ぶ。今まで出てきた例はすべて離散型の確率変数である。X が実現値 x をとる確率を $P(X=x)$ または $P_X(x)$ で表わし、確率関数 (または確率質量関数) と呼ぶ。これは x の関数である。

例：赤玉と黄玉を取り出す例における確率変数 Z の確率分布は $P_Z(0)=P(Z=0)=0.25$, $P_Z(1)=P(Z=1)=0.5$, $P_Z(2)=P(Z=2)=0.25$, $z\neq 0,1,2$ のとき $P_Z(z)=P(Z=z)=0$ となる。

標本空間の省略

標本空間を考える必要がなく、確率変数を考えることが重要である場合、背後にある標本空間 Ω と Ω 上の確率 P を省略して、Ω と P の影響を受ける確率変数 X とその累積分布関数だけ考えることがある。（教科書の確率変数の定義は実はそうなっている）ときどき確率変数の背後にある標本空間を持ち出さないと議論ができなくなることがある。

2009-2-6

連続型確率変数

確率変数 X とその累積分布関数 F_X があるとする。ある関数 f_X が有って

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

と表わせるとき、確率変数 X は連続型であると呼び、f_X を X の確率密度関数と呼ぶ。f_X は F_X の導関数だと思ってよいが、F_X に微分不可能な点がある連続型確率変数が存在する（次の例を参照）。
F_X はいかなる場合も右連続であるが、X が連続型であるためには F_X が連続であることがさらに必要である。離散型確率変数では、$P(X = x)$ が非零になる x で F_X が連続ではなくなるから、離散型確率変数は連続型ではない。
離散型でも連続型でもない確率変数もあるが、取り扱いが難しいのでこの講義では取り上げない。

例：一様確率変数

0 以上 1 以下のすべての値に一様に（つまり等しく）実現値をとる確率変数 U を考える。$0 \leq a \leq b \leq 1$ のとき a より大きく b 以下の値を U が取る確率は $b-a$ で与えられるべきである。

$$F_U(u) = \begin{cases} 0 & \text{if } u < 0, \\ u & \text{if } 0 \leq u \leq 1, \\ 1 & \text{if } 1 < u \end{cases}$$

とおくと、a より大きく b 以下の値を U が取る確率は $F_U(b) - F_U(a) = b - a$ となる。
F_U のグラフを以下に書く。

例：一様確率変数の確率密度関数

F_U のグラフが 0 と 1 で微分できないことに注意した後に、f_U のグラフを以下に書く。

例：一様確率変数の確率質量関数

一様確率変数 U はすべての実数 u について確率質量関数 $P_U(u) = 0$ となる。なぜならば、もしある $0 < u < 1$ の u で $P_U(u) = a > 0$ となったとすると、U は 0 以上 1 以下のどの値も等しい確率で取るからある $u_1 \neq u$ においても $P_U(u_1) = a > 0$ である。お互いに異なる $u_1, \ldots, u_n$ を十分に多くとると $P_U(u_1) + \cdots + P_U(u_n) = na > 1$ とすることができて矛盾が生じる。したがって $P_U(u) > 0$ となる u は存在しない。

2009-2-10

離散型確率変数の期待値

離散型確率変数 X が 有限種類の値 $x_1, \ldots, x_n$ だけを取るとする。このとき X の期待値 (または平均) を

$$E[X] = \sum_{i=1}^{n} P_X(x_i) x_i$$

で定義する。
無限和は、和を取る順番によって無限和の値が変わることがある (大学１年の微分積分の講義内容を参照)。したがって離散確率型確率変数 Y が無限種類の値 y_1, $y_2, \ldots$ を取るときに、期待値を

$$E[Y] = \sum_{i=1}^{\infty} P_Y(y_i) y_i$$

と定義すると、y_i に添字 i を付ける順番が変わると式の値が変わってしまい、うまく定義できないことがある。添字 i を付ける順番が変わると右辺の値が変わる場合は Y の期待値は定義されない。

2009-2-11

連続型確率変数の期待値

連続型確率変数 X の期待値を定義したい。確率変数 X の期待値を

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

で定義する。
例：先程の一様確率変数の例では

$$\begin{aligned} E[U] &= \int_{-\infty}^{\infty} u f_U(u) du \\ &= \int_{0}^{1} u f_U(u) du \\ &= \int_{0}^{1} u du \\ &= 0.5 \end{aligned}$$

となり、直観と一致している。

期待値が存在しない連続型確率変数の例

確率密度関数が

$$f_X(x) = \frac{1}{\pi(1+x^2)}$$

で与えられる確率変数 X を考える。この関数は $x=0$ に対して左右対称であるから期待値が 0 になりそうに見えるが、そうではない。$xf_X(x) = x/(\pi(1+x^2))$ の不定積分は $\ln(1+x^2)/2\pi$ である。したがって

$$\begin{aligned}\int_{-\infty}^{0} x\frac{1}{\pi(1+x^2)}dx &= -\infty \\ \int_{0}^{\infty} x\frac{1}{\pi(1+x^2)}dx &= \infty\end{aligned}$$

である。したがって

$$\int_{-\infty}^{\infty} x\frac{1}{\pi(1+x^2)}dx = \lim_{\substack{r\to\infty \\ s\to\infty}} \int_{-r}^{s} x\frac{1}{\pi(1+x^2)}dx$$

という広義積分の値は存在しない。
なお、上の分布は教科書 128 ページで紹介されているコーシー分布の特別な例である。

2009-2-13

期待値の性質

期待値をとる操作 $E[\cdot]$ について以下の性質が成り立つ。ただし c は定数 (確率 1 で値 c を取る確率変数) とする。

$E[c] = c$

$E[X + c] = E[X] + c$

$E[X + Y] = E[X] + E[Y]$

$E[cX] = cE[X]$

2009-2-14

$Y = g(X)$ の期待値

確率変数 X と関数 g を用いて新しい確率変数 $Y = g(X)$ を定義したとする。$E[Y]$ を計算したい。

まず X, Y が両方離散型確率変数の場合を検討する。X の取り得る値を $x_1, x_2, \ldots,$ とし $y_i = g(x_i)$ とする。もし $x_i \neq x_j$ ならば必ず $g(x_i) \neq g(x_j)$ が成り立つとすると $P_Y(y_i) = P_X(x_i)$ が成立するので

$$\begin{aligned} E[Y] &= \sum_i P_Y(y_i) y_i & (1) \\ &= \sum_i P_X(x_i) g(x_i) & (2) \end{aligned}$$

$g(x_i) = g(x_j)$ かつ $x_i \neq x_j$ のときは $y_i = y_j$ なので等式 (1) が成立しないが、よく考えると

$$E[Y] = \sum_i P_X(x_i) g(x_i)$$

の成立がわかる。

X, Y の両方が連続型確率変数の場合は

$$E[Y] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \tag{3}$$

が成立する。この式が成立することの証明はかなり難しいので説明しない。式 (3) を用いると $E[Y] = \int y f_Y(y) dy$ の中の確率密度関数 $f_Y(y)$ の計算を省略して直接 $E[Y]$ を計算できるため、計算の手間が減る。

分散

確率変数の値のばらつきについて考えてみたい。Y を ± 1 をそれぞれ確率 0.5 でとる確率変数、Z を $\pm 1, \pm 2$ をそれぞれ確率 0.25 で取る確率変数とする。Y と Z は両方とも期待値は 0 だが Z のほうがばらつきが大きい。このばらつきを定量的に捉える概念を紹介する。確率変数 X の期待値を m とする。$(X-m)^2$ は X が期待値から遠く離れた値を取るほど大きくなる。この期待値 $E[(X-m)^2]$ を分散と呼び、$V[X]$ で表わし確率変数のばらつきを表わす。上記 Y の分散は 1 で Z の分散は 2.5 である。分散の平方根を標準偏差と呼ぶ。

2009-2-16

分散の性質

分散を実際に手で計算するときには以下の公式が便利である。

$$
\begin{aligned}
V[X] &= E[(X-m)^2] \\
&= E[X^2 - 2mX + m^2] \\
&= E[X^2] - 2mE[X] + m^2 \\
&= E[X^2] - 2m \cdot m + m^2 \\
&= E[X^2] - m^2
\end{aligned}
$$

$E[X^2]$ の計算には式 (2) または式 (3) を用いると手間が減る。

c を定数として、分散には以下の性質がある：

$V[c] = 0$

$V[X+c] = V[X]$

$V[cX] = c^2 V[X]$

2009-2-17

確率変数の標準化

ある確率変数 X が与えられたときに新しい確率変数

$$Z = \frac{X - E[X]}{\sqrt{V[X]}}$$

を考える。このとき $E[Z] = 0, V[Z] = 1$ である。このような Z を得る操作を標準化と呼び、Z を標準化変数と呼ぶ。

中央値 (メディアン) と最頻値 (モード)

確率変数 X の中央値とは $F_X(x) = 0.5$ となる値 x である。つまり、確率変数が中央値より小さい値をとる確率と大きい値を取る確率が丁度等しくなっている。最頻値とは、確率変数 X が最も頻繁に取る値で、X が連続型の場合確率密度関数 f_X を最大にする x、離散型の場合確率質量関数 P_X を最大にする x である。次回は、多次元の確率分布 (教科書の 7 章) を講義予定。E/O クラス分けを間違えている学生に注意をする。

2009-2-18

演習問題

原則として授業終了までに答案を提出して下さい。終らない場合は授業が行われた日の夜までに南３号館３１１号室前の箱に演習答案を提出して下さい。
以下のような累積分布関数を持つ確率変数 X を考える。

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \frac{1}{2}x^2 & \text{if } 0 \leq x \leq 1, \\ -\frac{1}{2}(x-2)^2 + 1 & \text{if } 1 \leq x \leq 2, \\ 1 & \text{if } 2 \leq x \end{cases}$$

11. X は離散型確率変数であるか否か、理由と共に答えよ。
12. X の確率密度関数を図示せよ。
13. X が 0.5 より大きく 1 以下の値をとる確率を求めよ。
14. X の期待値、最頻値、中央値、分散、標準偏差を求めよ。
15. X の標準化変数 Z の累積分布関数を求めよ。(ヒント：Z の累積分布関数 F_Z を X の累積分布関数 F_X を用いて $F_Z(z) = F_X($「z の関数」$)$ という形式で表わしたあと、$F_Z(z)$ の具体的な式の形を求めるのが標準的な方法でしょう)

2009-2-19

16. Z の確率密度関数を図示せよ。

17. 授業でわかりにくかった所や要望を書いて下さい。また前回の答案で×になったところで、理由がわからないところは申し出て下さい。