確率への招待　14回目

確率変数と確率分布②

共分散と相関係数
確率変数の例

１．独立な確率変数の期待値と分散

定理）２つの確率変数XとYが独立ならば、

平均　E（XY）＝E（X）E（Y）

分散　V（X＋Y）＝V（X）＋V（Y）となる。

（なお、E（X＋Y）は、XとYが独立でなくても、E（X）＋E（Y））

いちおう証明をつけておく。

（証明を覚える必要はないが、定理そのものを覚えることと、
計算に慣れることは必要）

XとYが独立とは、確率分布の表で、$p_{ij}=p_iq_j$ということだった。
あとは定義にしたがって、E（XY）、V（X＋Y）を計算する。

$$\mathrm{E(XY)} = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j p_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j p_i q_j$$
$$= \sum_{i=1}^{m} x_i p_i \times \sum_{j=1}^{n} y_j q_j = E(X)E(Y)$$

$$
\begin{aligned}
&V(X+Y) \\
&= E\left((X+Y)^2\right) - \left(E(X+Y)\right)^2 \\
&= E(X^2 + 2XY + Y^2) - \left(E(X) + E(Y)\right)^2 \\
&= E(X^2) + 2E(\mathrm{X}Y) + E(Y^2) - \left(E(X)\right)^2 - 2E(X)E(Y) - \left(E(Y)\right)^2 \\
&= E(X^2) - \left(E(X)\right)^2 + E(Y^2) - \left(E(Y)\right)^2 + 2[E(XY) - E(X)E(Y)] \\
&= V(X) + V(Y) + 2[E(XY) - E(X)E(Y)]
\end{aligned}
$$

ここで、XとYが独立ならE（XY）＝E（X）E（Y）なので、

（上式）＝V（X）＋V（Y）

２．共分散と相関係数

<u>XとYが必ずしも独立でない場合、V（X＋Y）はどうなるだろうか。</u>

確率変数の分散の定義にならって２つの確率変数X,Yの共分散Cov（X,Y）を次のように定義する。

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

すると、V（X）＝E（X^2）－{E（X）}^{2}を導いたときと同様の計算で、

$$\begin{aligned} Cov(X,Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY - XE(Y) - E(X)Y + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

とくに、X とY が独立ならば $Cov(X, Y) = 0$

以上の準備の下に、V(X＋Y)を計算する。

$$V(X+Y)=V(X)+V(Y)+2\{E(XY)-E(X)E(Y)\}$$

これと前ページの式を合わせると、
XとYが必ずしも独立ではない場合には、
V(X＋Y)＝V(X)＋V(Y)＋2Cov(X, Y)
　　これも覚えておくべき式
とくに、XとYが独立ならばV(X＋Y)＝V(X)＋V(Y)
　　←XとYが独立ならばCov(X, Y)＝0だから

共分散については、次の不等式が成り立つ。

$$\{Cov(X,Y)\}^2 \leq V(X)V(Y) \qquad \text{(シュワルツの不等式)}$$

証明）高校数学でも、シュワルツの不等式をやったことと思う。
証明も大体同じである。

任意の実数 t に対し、
$\{(X-E[X])t+(Y-E[Y])\}^2$ は常に0以上なので、期待値をとった
$E[\{(X-E[X])t+(Y-E[Y])\}^2]$ も常に0以上。

これを展開して、
$E((X-E(X))^2)t^2+2E[(X-E[X])(Y-E[Y])]t+E((Y-E(Y))^2)$
$=V(X)t^2+2Cov(X,Y)t+V(Y)$ は常に0以上。

これを t の2次式と見ると、判別式≦0でなければならないから、
$\{Cov(X,Y)\}^2-V(X)V(Y)\leq 0$

２つの確率変数X，Yの相関係数 (correlation) を、次の式で定義する。

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}$$

シュワルツの不等式より、－１≦$\rho(X,Y)$≦１

おおざっぱに言って、
「Xが増えるときにYも増える傾向があるとき」はρ＞０
「Xが増えるときにYが減る傾向があるとき」 はρ＜０

３．チェビシェフの不等式、大数の法則

チェビシェフの不等式は、「平均から離れたと値を取る確率は低い」ことを主張するものであり、次のように定式化される。

定理）Xを確率変数とし、その期待値をm、分散をσ^2とする。

任意の正の実数aに対し、$P(|X-m|>a\sigma)\leq\dfrac{1}{a^2}$

証明）V（X）の計算において、Xの範囲を、|X－m|＞aσ の部分と|X－m|≦aσ の部分に分けると、

・|X－m|＞aσ の部分では、（X－m）^{2}＞$a^2\sigma^2$ だから、
これに確率を掛け算して足し合わせると
（この部分）＞$a^2\sigma^2$P（|X－m|＞aσ）

・|X－m|≦aσの部分では、かなり甘い評価だが、
（X－m）^{2}≧0だから、
（この部分）≧0

足し算して、V（X）≧$a^2\sigma^2$P（|X－m|＞aσ）
V（X）＝σ^2だから、両辺を$a^2\sigma^2$で割ると求める式になる。

チェビシェフの不等式から、有名な「大数の法則」が導かれる。

大数の法則

X_1、X_2、・・・、X_n が独立で同一の分布に従う確率変数とし、その期待値m、分散σ^2が存在するとする。

このとき、X_1、X_2、・・・、X_n の平均（X_1＋X_2＋・・・＋X_n）／nは、n が十分に大きいとき、mに近づく。

証明）（X_1＋X_2＋・・・＋X_n）／nは確率変数だが、その平均はm、分散はσ^2/nとなる。

ここでnが十分に大きいと、分散は0に近づくので、チェビシェフの不等式から、題意が示される。

※　実は、世の中には、「平均や分散が存在しない（無限大に発散する）」確率変数も存在する。そういった確率変数については、大数の法則は必ずしも成立しない。

４．確率変数の例

（１）二項分布

①二項分布の定義

　反復試行のところで、
ある試行を１回行って事象Aが起こる確率をpとするとき、
この独立な試行をn回行ってAがちょうどr回起こる確率は

$$ {}_nC_r p^r (1-p)^{n-r} $$

となることを学んだ。

　rを確率変数と考えるとき、この分布を二項分布といい、
記号Ｂ（n,p）で表す。（binomial distribution）

　Ｘ～Ｂ（n,p）

Ｘ	０	１	・・・	r	・・・	n	計
Ｐ	$(1-p)^n$	$np(1-p)^{n-1}$	・・・	${}_nC_r p^r (1-p)^{n-r}$		p^n	1

②二項分布の再生性

Xを二項分布B（n,p）に従う二項分布、
Yを二項分布B（m,p）に従う二項分布とし、pは同じとする。

XとYが独立であれば、二項分布の意味を考えると、
Xは確率pの独立試行をn回行ったときに事象が起こる回数、
Yは確率pの独立試行をm回行ったときに事象が起こる回数
なので、
X＋Yは、確率pの独立試行をn＋m回行ったときに事象が起こる回数、すなわち、二項分布B（n＋m,p）となる。

これを繰り返し使うと、$X_1, X_2, \cdots, X_n$が互いに独立な確率変数で、それぞれが二項分布B（1,p）に従うならば（pは共通）、

$$X_1 + X_2 + \cdots + X_n \sim B(n, p)$$

これらを、「二項分布の再生性」という。

③二項分布の期待値、分散

二項分布B（n,p）の期待値、分散を求めるために、まずは、n＝1の場合、二項分布B（1,p）の期待値、分散を求めよう。

B（1，p）は、

X	0　　1	計
P	1－p　　p	1

よって、$E(X)=0\cdot(1-p)+1\cdot p=p$

$E(X^2)=0^2\cdot(1-p)+1^2\cdot p=p$

$V(X)=E(X^2)-\{E(X)\}^2=p-p^2=p(1-p)$

次に、B（n，p）の平均、分散を考えよう。

$X_1, X_2, \cdots, X_n$が互いに独立な確率変数で、それぞれが二項分布B（1,p）に従うならば$X_1+X_2+\cdots+X_n \sim B(n, p)$

$E(X_1+X_2+\cdots+X_n)=E(X_1)+\cdots+E(X_n)=np$

$V(X_1+X_2+\cdots+X_n)=V(X_1)+\cdots+V(X_n)=np(1-p)$

（2）連続型の確率変数

①定義

これまでは、確率変数Xは、とびとびの値をとるものとして計算を進めてきた。しかし、世の中には、連続的な値をとるものも多くある。（時間、モノの長さなど）

これらに対しても、これまでの議論が使えるようにしよう。

例）[０，１] 間の一様分布

Xが０から１の間の値を同じ確率でとるものとしよう。

　　０から１の間の乱数

　　０から１の数直線上にエンピツを落としてみる、等。

Xが特定の値（例えば0.1とか0.5とかπ/5とか）をとる確率は、ゼロになる（０から１の間には数は無限に多くあるから）。

したがって、P（X＝a）を考えてもうまくいかない。

幅を持たせて、P（a≦X≦b）を考えるとうまくいく。

０≦a＜b≦１なるa，bに対して、P（a≦X≦b）＝b－aとなる。

（全体の区間 [0, 1] のうち [a, b] の割合として確率を定義。）

一般に、確率変数Xが連続的な値をとるとき**連続型の確率変数**という（これに対し、Xがとびとびの値をとるときは**離散型の確率変数**という）。

離散型の確率変数で度数分布表を描いたのと同様、連続型の確率変数についても「分布曲線」を描くことができ、その関数を「確率密度関数」という。

確率密度関数f(x)は次の性質を持つ。

①常にf(x)≧0

②Xのとる値の範囲がα≦X≦βのとき　$\int_{\alpha}^{\beta} f(x)dx = 1$

確率P(a≦X≦b)は、y＝f(x)のグラフとx軸、および2直線x＝a、x＝bで囲まれた部分の面積に等しい。すなわち、

$$P(a \le X \le b) = \int_{a}^{b} f(x)dx$$

例1）一様分布の確率密度関数

a，bをa＜bを満たす任意の実数とするとき

$$f(x)=\frac{1}{b-a}\qquad (\text{ただし}\,a\leq x\leq b)$$

は確率密度関数になる。

これを区間（a，b）上の一様分布という。

例2）f（x）＝2x　（0≦x≦1）は確率密度関数になる。

実際、0≦x≦1においてf（x）≧0であり、

$$\int_0^1 2x\,dx=[x^2]_0^1=1$$

確率変数・例題

サイコロを１回振ったときに出た目の数をXとするとき、

（１）Xの確率分布を求めよ。

（２）期待値Ｅ（X）、分散Ｖ（X）を求めよ。

（３）2^Xの期待値Ｅ（2^X）を求めよ。

（答え）

(1)

X	１	２	３	４	５	６	計
Ｐ	1/6	1/6	1/6	1/6	1/6	1/6	1

(2)$\mathrm{E(X)} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2}$

$\mathrm{E(X^2)} = \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$なので

$\mathrm{V(X)} = \mathrm{E(X^2)} - (E(X))^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$

(3)$\mathrm{E}\left(2^{\mathrm{X}}\right) = \frac{2+4+8+16+32+64}{6} = \frac{126}{6} = 21$

問題のバリエーションとして、「コインを３枚投げたときに表の出る枚数」とか、E(2X+3)やV(3X+1)を求めよ、とか。。。。