統計学
離散型確率変数

担当：　長倉　大輔
（ながくらだいすけ）

離散型確率変数

- 確率変数とは？

ある現象が確率的に起こったときにその現象に対応する数値の事を**確率変数**という。

株価収益率などは数値そのものだが、お天気などの場合にも、晴れの時には「1」、曇りの時には「2」などのようにそれぞれの現象に数値を対応させる事ができる。

そのような変数を考えれば、それは確率変数である。

離散型確率変数

- 確率変数の種類

(離散型確率変数)
サイコロの目や、晴れの時には「1」、曇りの時には「2」のように離散的に値をとる確率変数を**離散型確率変数**という。

(連続型確率変数)
株価収益率、身長、体重などのように連続的に値をとる確率変数を**連続型確率変数**という。

離散型確率変数

■ 確率関数

離散型確率変数 X は m 個の実数 x_i , $i=1,\ldots,m$ を実現値として取り、それぞれの実現値の確率は p_i , $i=1,\ldots,m$ であるとしよう。この時、この離散型確率変数 X の**確率関数** $p_X(.)$ は

$$p_X(x_i) = p_i, \qquad 0 < p_i < 1,$$

を満たし、また $y \neq x_i$ に対しては

$$p_X(y) = 0$$

を満たす関数である(慣例として大文字 X で確率変数、小文字 x で確率変数 X が取りうる値 を表す)。

離散型確率変数

つまり**確率関数**とは離散型確率変数の取りうる値にその**確率**を対応させる**関数**の事である。

p_i , i =1,.., m は確率であるから

$$\sum_{i=1}^{m} p_i = p_1 + \cdots + p_m = 1$$

を満たす。

離散型確率変数

(例1)　コイン投げ

確率変数 X はコインを投げて表が出たら $X = 1$、裏が出たら $X = 0$ という実現値を取るとしよう。

表が出る確率も裏が出る確率も 1/2 とする。この時、X の確率関数 $p_X(x)$ は

$$p_X(1) = 1/2, \quad p_X(0) = 1/2, \quad p_X(y) = 0 \text{ for } y \neq 1, 0$$

となる。

離散型確率変数

(例2) サイコロの目
サイコロの目を離散型確率変数とみなし、それを X としよう。X の取りうる値は $X = 1, 2, 3, 4, 5, 6$ である。

全ての目は同じ確率で出るとすると、この X の確率関数 $p_X(x)$ は

$$p_X(1) = 1/6,\quad p_X(2) = 1/6,\quad p_X(3) = 1/6,$$
$$p_X(4) = 1/6,\quad p_X(5) = 1/6,\quad p_X(6) = 1/6,$$
$$p_X(x) = 0 \quad \text{for } x \neq 1, 2, 3, 4, 5, 6$$

となる。

離散型確率変数

- 累積分布関数

ある実数 x に対して確率変数 X が x 以下の値である確率 $\Pr(X \leq x)$を対応させる関数を**累積分布関数**という。

この定義は連続型確率変数に対してもあてはまる。

離散型確率変数

離散型確率変数の場合、具体的には以下のようになる。

ある離散型確率変数 X が m 個の実数: $x_1 < x_2 < \ldots < x_m$ を取るとしよう。この時、X の**累積分布関数** $F_X(x)$ は

$$F_X(x) = \Pr(X \le x) = \sum_{i=1}^{j} p_X(x_i)$$

と定義される。ここで $p_X(x)$ は X の確率関数、x_j は $x_1, \ldots, x_m$ のうち **x 以下で最も大きい数**である。

離散型確率変数

(例3)　サイコロの目

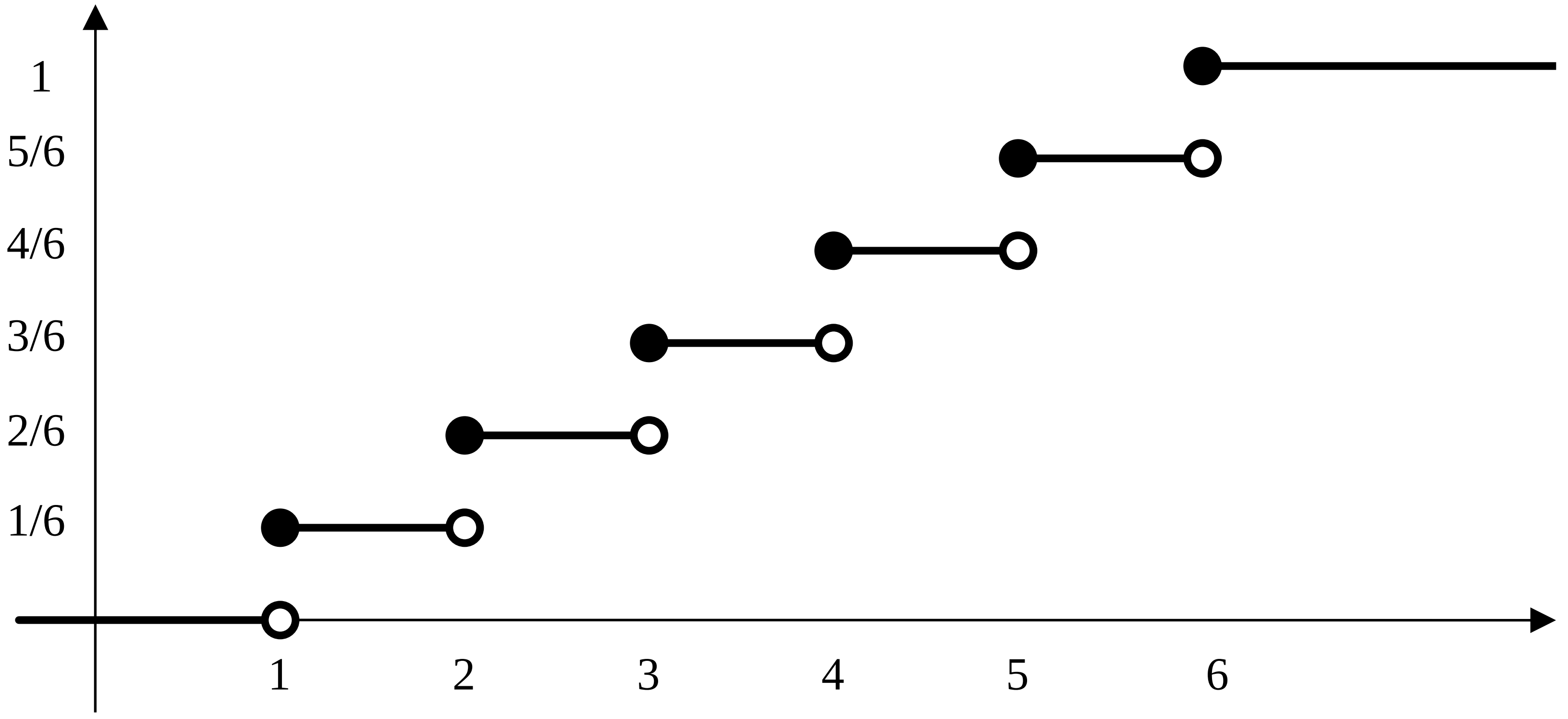
先ほどのサイコロの例では

$$F(x) = \Pr(X \le x) = \begin{cases} 0, & x < 1 \\ 1/6, & 1 \le x < 2 \\ 2/6, & 2 \le x < 3 \\ 3/6, & 3 \le x < 4 \\ 4/6, & 4 \le x < 5 \\ 5/6, & 5 \le x < 6 \\ 1, & 6 \le x \end{cases}$$

となる。

離散型確率変数

(サイコロの目の分布関数のグラフ)



離散型確率変数

例題 1

コイン投げの例に出てきた確率変数 X の累積分布関数を求め、そのグラフを書け。

離散型確率変数

- 平均と分散

データを要約する際に標本平均と標本分散を計算した。

同様に、**確率分布を要約する**ために確率分布に対しても**平均**と**分散**を定義する。

離散型確率変数の平均と分散は以下のように定義される。

確率変数の平均は**期待値**とも呼ばれる。今後はこの2つの単語を併用する。

離散型確率変数

離散型確率変数 X が m 個の実数 $x_1, \ldots, x_m$ を取り、その確率関数を $p_X(x_i)$ としよう。この時、X の平均は

$$E(X) = \mu = x_1 p_X(x_1) + \ldots + x_m p_X(x_m)$$
$$= \sum_{i=1}^{m} x_i p_X(x_i)$$

分散は

$$\mathrm{var}(X) = \sum_{i=1}^{m} (x_i - \mu)^2 p_X(x_i)$$

と定義される。$E(X)$ と $\mathrm{var}(X)$ はそれぞれ平均と分散を表す(英語の Expectation (期待値) と Variance (分散)の略)。

離散型確率変数

標本数 n の標本平均はそれぞれの観測値に **$1/n$ という重みを付けた加重平均**であるのに対して、確率分布の平均は取りえる値のそれぞれをその値の**確率で重みをつけた加重平均**である。

確率分布の平均と分散についても、平均は**分布の中心**、分散(およびその平方根をとった標準偏差)は**分布のばらつき具合**を表すという解釈は標本平均、標本分散の場合と同じである。

離散型確率変数

- 平均と分散の性質

標本平均や標本分散で成り立つ性質は確率分布の平均や分散でもそのまま成り立つ。以下は代表的なものである。

(平均)

(1)　定数 c について $E(c) = c$,

(2)　$E(X - \mu) = 0$,

(3)　$E(aX + b) = a\mu + b$,

(4)　$E(X + Y) = E(X) + E(Y)$

(分散)

(1)　$\mathrm{var}(X) = E(X^2) - \mu^2$,

(2)　$\mathrm{var}(aX + b) = a^2\mathrm{var}(X)$,

離散型確率変数

- ベルヌーイ分布

コインを投げて表か裏が出るような、

(1) 起こりうる出来事(根元事象)の数が 2 つしかなく、
(2) 表が出れば $X = 1$、裏が出れば $X = 0$ のようにそれぞれの起こりうる出来事に 0 と 1 が対応し、
(3) それぞれの確率が

$$\Pr(X = 1) = p,\ \Pr(X = 0) = 1 - p,\ 0 < p < 1$$

で与えられる確率変数 X の分布を**ベルヌーイ分布**という。

離散型確率変数

例題 2 (ベルヌーイ分布の期待値と分散)

ベルヌーイ分布の期待値と分散が

$$E(X) = p,\ \ \mathrm{var}(X) = (1 - p)\,p$$

となる事を示しなさい。

離散型確率変数

- 2 項分布 (Binomial distribution)

n 個の独立なベルヌーイ確率変数 $X_1, X_2, \ldots, X_n$を考える X_i $(i = 1, \ldots, n)$ は全て同じベルヌーイ分布 $\Pr(X_i = 1) = p$, $\Pr(X_i = 0) = 1 - p$ に従うとする。この時

$$Y = X_1 + X_2 + \ldots + X_n$$

を **2 項確率変数**と呼び、その確率分布を **2 項分布**と呼ぶ。

確率変数 Y が 上のような 2 項分布に従っているとき、

$$\boldsymbol{Y \sim B(n, p)}$$

と書く。Bは Binomial の頭文字、～はその右側の分布に従っているという意味である。

離散型確率変数

- 2 項分布の意味

2 項分布は成功確率 p の試行を n 回行った時の**成功した回数**と見なす事ができる。

「成功」したら $X_i = 1$、「失敗」したら $X_i = 0$ とすると

試行	X_1,	X_2,	X_2,	X_3,	……	,	X_n
結果	0,	1,	1,	0,	……	,	1

$\rightarrow\ Y = X_1 + X_2 + \ldots + X_n$ は成功した回数と等しい。

離散型確率変数

- 2 項分布の確率、期待値、分散

2項分布において、$Y = k$ となる確率は

$$\Pr(Y = k) = {}_nC_k\, p^k\, (1 - p)^{n-k}$$

で与えられる。ここで${}_nC_k$ は

$${}_nC_k = \frac{n!}{k!(n-k)!}$$

で与えられる。$n!$ は n の**階乗**と呼ばれ、

$$n! = n \times (n-1) \times (n-2) \ldots 3 \times 2 \times 1$$

と定義される。特に n が 0 の場合は **$0! = 1$** と定義される。1 の階乗も 1 である事に注意 ($1\,! = 1$)。

離散型確率変数

- 2 項分布の確率、期待値、分散

2項分布の期待値と分散は

$$E(Y) = pn, \quad \mathrm{var}(Y) = np\,(1-p)$$

で与えられる。これは宿題で確認する。

離散型確率変数

例題3 (2項確率の計算)

ある子供が5匹の金魚をすくったとする。個々の金魚が1週間以内に死ぬ確率は0.9だとしよう。また個々の金魚の生死は独立だとしよう。一週間後

(1) 少なくとも一匹生き残る確率を求めなさい。

(2) 多くとも一匹しか生き残れない確率を求めなさい

(3) 何匹生き残る事が期待できるか？
(期待値を求める)

演習問題

1. サイコロを10回振った時の1の目が出る回数の期待値を求めなさい。

2. あるクイズ問題は合計で10問あるとする。この1問につき正解すると4点、失敗すると –1点になるとしよう。あなたの正答率を Pr(正答する) =1/3とすると、合計点が 0 点以上となる確率はいくらになるか。

3. 上の問題において得点の期待値を求めなさい。

離散型確率変数

- 同時確率と周辺確率

表が出る確率も裏が出る確率も共に1/2 の 2 枚のコインをそれぞれ コイン1, コイン2 とし、X_i をコイン i (i =1, 2) の目に対応する確率変数とする。

X_i はコインi を投げて、表が出たとき 1 を取り、裏が出たときに 0 を取るとする。この時、X_1, X_2 が共に 1 を取るような確率、すなわち

$$\Pr(X_1 = 1, X_2 = 1)$$

のような確率を X_1 と X_2 の**同時確率**という。

離散型確率変数

同時確率に対して

$$\Pr(X_i = 1) = 1/2\ ,\quad \Pr(X_i = 0) = 1/2,\quad i = 1, 2$$

を**周辺確率**とよぶ。

離散型確率変数

- 独立

2 つの(離散型)確率変数 X_1 と X_2 に対して、その同時確率がそれぞれの周辺確率の積と等しいとき、すなわち

$$\Pr(X_1 = k, X_2 = j) = \Pr(X_1 = k) \Pr(X_2 = j),$$

(k, j は X_1, X_2 が取りうる値) が成り立つとき、この 2 つの(離散型)確率変数は**独立**であるという。

離散型確率変数

- 同時確率と周辺確率の表

先ほどの**独立な** 2 つのコインの同時確率と周辺確率を表にすると以下のようになる(これをみれば、なぜ周辺確率がそう呼ばれるかがイメージできるでしょう)。

コイン1＼コイン2	表	裏	コイン1の周辺確率
表	1/4	1/4	1/2
裏	1/4	1/4	1/2
コイン2の周辺確率	1/2	1/2	1

ここで表と表がクロスする部分は 2 枚とも表が出る確率、等とする。

離散型確率変数

- 同時確率と周辺確率の関係

さきほどのコインの例で同時確率と周辺確率には

$$\Pr(X_1{=}1) = \Pr(X_1{=}1, X_2{=}1) + \Pr(X_1{=}1, X_2{=}0)$$

のような関係がある事がわかる。これは、根元事象が

{コイン1, コイン2} = {表表}, {表裏}, {裏表}, {裏裏}

の4つであり、これらは互いに排反であるので、

$$\begin{aligned}\Pr(X_1{=}1) &= \Pr(\{\text{表表}, \text{表裏}\}) \\ &= \Pr(\{\text{表表}\}) + \Pr(\{\text{表裏}\}) \\ &= \Pr(X_1{=}1, X_2{=}1) + \Pr(X_1{=}1, X_2{=}0),\end{aligned}$$

となる事よりわかる。

離散型確率変数

- 同時確率関数と周辺確率関数

先ほどの例において、例えば X_1 と X_2 がそれぞれ k と j になる確率を与える関数

$$p_{12}(k, j) = \Pr(X_1 = k, X_2 = j)$$

をこの2つの確率変数の**同時確率関数**という。

またこの時、X_1 の確率関数 $p_1(x)$は**周辺確率関数**と呼ばれる場合がある。

X_1 と X_2 が独立であれば、$p_{12}(k, j) = p_1(k)\, p_2(j)$ となる。

離散型確率変数

- 一般の場合

2 つの確率変数 X と Y があり、X は $x_1, x_2, \dots, x_m$ の m 個の値を正の確率で取り、Y は $y_1, y_2, \dots, y_n$ の n 個の値を正の確率で取るとする。この時、X と Y の同時確率関数は

$$p_{XY}(x_i, y_j) = \Pr(X = x_i, Y = y_j),\ i = 1, 2, \dots, m,\ \ j = 1, 2, \dots n,$$

であり、その周辺確率関数は同時確率関数の和と等しく

$$p_X(x_i) = \Pr(X = x_i) = \sum_{j=1}^{n} p_{XY}(x_i, y_j), \quad i = 1, 2, \dots, m$$

$$p_Y(y_j) = \Pr(Y = y_j) = \sum_{i=1}^{m} p_{XY}(x_i, y_j), \quad j = 1, 2, \dots, n$$

を満たす。

離散型確率変数

同時確率関数は、定義より

$$\sum_{i=1}^{m}\sum_{j=1}^{n} p_{XY}(x_i, y_j) = 1$$

を満たす

離散型確率変数

- 条件付確率関数

条件付確率に対しては**条件付確率関数**が定義される。

条件付確率関数の値を求めるには同時確率関数の値を周辺確率関数の値でわってやればよい。先ほどのコインの例では、例えば $X_1 = 1$ という条件付の $X_2 = 0$ の条件付確率関数 $p_{2|1}(0|1)$ の値は

$$p_{2|1}(0 \mid 1) = \Pr(X_2 = 0 \mid X_1 = 1)$$

$$= \frac{\Pr(X_1 = 1,\ X_2 = 0)}{\Pr(X_1 = 1)} = \frac{p_{12}(1,\ 0)}{p_1(1)} = \frac{1/4}{1/2} = \frac{1}{2}$$

となる。X_1と X_2 は独立なのでこれは $p_2(0)$に等しい。

離散型確率変数

- 条件付平均と条件付分散

離散型確率変数 X は $x_1,\ldots,x_m$ の m 個の値を取り $X=x_i$ の $Y=y_j$ という条件付確率関数の値は $p_{X|Y}(x_i\,|\,y_j)=p_{i\,|\,j}$ で与えられているとする。この時 X の**条件付平均**と**条件付分散**は

(条件付平均)

$$E(X\,|\,Y=y_j)=\mu_{X|Y=y_j}=\sum_{i=1}^{m}x_i\,p_{i\,|\,j}$$

(条件付分散)

$$\mathrm{var}(X\,|\,Y=y_j)=\sum_{i=1}^{m}(x_i-\mu_{X|Y=y_j})^2\,p_{i\,|\,j}$$

のように定義される。

離散型確率変数

- 共分散

離散型確率変数 X と Y の**共分散** cov(X, Y) は

$$\operatorname{cov}(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij}(x_i - \mu_X)(y_j - \mu_Y)$$

によって定義される(cov は英語の covariance (共分散)の略)。ここで μ_X と μ_Y はそれぞれ X と Y の平均であり、p_{ij} は $p_{ij} = p_{XY}(x_i, y_i) = \Pr(X = x_i, Y = y_j)$ である。

定義より $\operatorname{cov}(X, X) = \operatorname{var}(X)$ となることに注意。

離散型確率変数

- 共分散の性質

(1) $\mathrm{cov}(X, Y) = E(XY) - \mu_X \mu_Y$

ここで $\mu_X = E(X)$, $\mu_Y = E(Y)$, および

$$E(XY) = x_1 y_1 p_{11} + x_1 y_2 p_{12} + \ldots + x_m y_n p_{mn}$$
$$= \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j p_{ij}$$

である。

(2) 2つの確率変数 X と Y が独立であれば、

$$\mathrm{cov}(X, Y) = 0,$$

(3) $\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\mathrm{cov}(X, Y)$

(4) $\mathrm{cov}(aX, bY) = ab\,\mathrm{cov}(X, Y)$

離散型確率変数

- 相関係数

確率変数 X と確率変数 Y の**相関係数**は

$$\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Y)}}$$

によって定義される。標本相関係数の場合と同様、

$$-1 \le \rho_{XY} \le 1$$

であり、単位の変化に影響されず**線形関係の強さを測る**ものである。

離散型確率変数

相関係数の定義より $\rho_{XY} = 0$ となるのは $\mathrm{cov}(X, Y) = 0$ となる時、およびその時に限る。

相関係数は X と Y の間の線形関係の強さを測るものである。よってもし X と Y が独立 (つまり何の関係もない)であれば $\rho_{XY} = 0$ (つまり $\mathrm{cov}(X, Y) = 0$) である。しかしながら、$\mathrm{cov}(X, Y) = 0$ であっても ***X* と *Y* は独立とは言えない**事に注意が必要である($Y = X^2$ のようは非線形の関係があるかもしれない)。

宿題 (提出する必要はありません)

1. 離散型確率変数, X, Y, Z について

   (1) $\mathrm{var}(X) = \mathrm{cov}(X, X)$

   (2) $\mathrm{cov}(X + Y, Z) = \mathrm{cov}(X, Z) + \mathrm{cov}(Y, Z)$

を示しなさい。

2. m 個の独立な確率変数 $X_1, \ldots, X_m$ に関して、

$$\mathrm{var}(X_1 + X_2 + \ldots + X_m) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + \ldots + \mathrm{var}(X_m)$$

となる事を共分散の性質(2)と上記の問題 1の(1)と(2)を用いて示しなさい。

宿題 (提出する必要はありません)

3. 2 項分布の期待値と分散が

$$E(Y) = pn, \ \ \mathrm{var}(Y) = np\,(1- p)$$

となる事を、期待値については平均の性質(4)を用いることによって、また分散については Y が独立な確率変数の和である事に注意して、上記問題2の結果を用いて示せ。

演習問題

4. 確率変数 X と Y の同時確率の表が

$X \diagdown Y$	2	0	Xの周辺確率
1	1/6	1/4	5/12
0	1/2	a	7/12
Yの周辺確率	2/3	1/3	1

で与えられているとする。以下の問いに答えよ。

(a) $\Pr(X=1 \mid Y=2)$を求めよ。

(b) a の値を求めよ

(c) X と Y は独立か?

(d) X の $Y=2$ という条件付平均を求めよ。

宿題 (提出する必要はありません)

4. スライド「確率論の基礎」の宿題に出てきた「3囚人問題1」について、以下の同時確率表を完成させなさい。

釈放＼看守の答え	B	C	釈放の周辺確率
A			1/3
B			1/3
C			1/3
看守の答えの周辺確率			1

宿題 (提出する必要はありません)

5. 問題4と同様に、今度は「3囚人問題2」、「3囚人問題3」、「3囚人問題4」について同時確率表を完成させなさい。

(3囚人問題2)

釈放＼看守の答え	B	C	釈放の周辺確率
A			1/4
B			1/4
C			1/2
看守の答えの周辺確率			1

宿題 (提出する必要はありません)

(3囚人問題3)

釈放＼看守の答え	処刑される	処刑されない	釈放の周辺確率
A			1/3
B			1/3
C			1/3
看守の答えの周辺確率			1

(3囚人問題4)

釈放＼看守の答え	B	C	釈放の周辺確率
A			1/3
B			1/3
C			1/3
看守の答えの周辺確率			1