確率・統計：第１回講義 (4月7日(月))

このノートを読んで，演習問題を解くことで，2025年度前期(月曜３限)「確率・統計」(担当：松本)の第１回講義とします．

何を学習するにも教科書を持っているべきだと思います．この講義であれば，指定した

「確率・統計の基礎」(松本裕行著，学術図書)

を参照してほしいと思います．

「定理」とあるのは，高校で学習した平均値の定理のように，重要な数学的な事実です．そこから公式が生まれることが多いです．定理に関しては，このノートを見るだけではなく，自筆のノートを作成して，例題や応用例などを合わせてメモしてください．その際，単に写すだけではなく，自分が理解しているかどうかを考えてください．このときに，教科書があると役に立つと思います．

私から特にお願いしたいのは，

> いろいろな事実を丸覚えするのではなく，理由をつけて頭に入れてほしいということです．演習問題を解く際は，公式や例題を見ながら解いてください．そのときも，理由付けを忘れないようにしてください．

演習問題，問題は必ず，解くこと．上にも書いたように，教科書，このノートを見ながらで全くかまいません．

重要な公式

$$(x+y)^n = \sum_{r=0}^{n} {}_nC_r x^r y^{n-r} \qquad \text{二項定理}$$

$$e^x = \sum_{r=0}^{\infty} \frac{x^r}{r!} \qquad \text{指数関数のマクローリン展開}$$

$$\sum_{r=0}^{N} cx^r = \frac{c(1-x^{N+1})}{1-x} \qquad \text{等比級数の有限和}$$

$$\sum_{r=0}^{\infty} cx^r = \frac{c}{1-x} \qquad (\text{ただし，}|x|<1) \qquad \text{等比級数の無限和}$$

第1章 確率変数，確率分布

1　離散分布

有限集合または正の整数のような可算集合に値をとる確率変数を考える．例から始める．

1.1　二項分布

サイコロを n 回ふるとき 6 の目が r 回出る確率は,

$$ {}_nC_r\left(\frac{1}{6}\right)^r\left(1-\frac{1}{6}\right)^{n-r} $$

である．このとき，6 の回数を X と書くと，X はサイコロを振る前は値が分からず，$X=r$ の確率が上のように分かっているだけである．

このような変数を**確率変数**という．

サイコロを振るという試行を一般化することにより，確率・統計で最も基本的な二項分布が定義される．

定義 1. p を $0<p<1$ を満たす実数とし，n を自然数とする．確率変数 X が，$\{0,1,2,...,n\}$ に値をもち，$X=r$ の確率 $P(X=r)$ が

$$ P(X=r)={}_nC_rp^r(1-p)^{n-r} \qquad (r=0,1,2,...,n) $$

によって与えられるとき，X は二項分布 $B(n,p)$ に従うという．

同じ試行を独立に n 回繰り返すとき，確率 p の事象が起きる回数は二項分布に従う．統計において，世論調査やテレビ番組の視聴率などを統計的に考える際の基本である．

なお，二項定理

$$ (x+y)^n=\sum_{r=0}^{n}{}_nC_rx^ry^{n-r} $$

を用いると，$\sum_{r=0}^{n}{}_nC_rp^r(1-p)^{n-r}=1$ が分かる．

1.2　幾何分布

サイコロを何回も独立に振るとき，始めて 6 が出るまでに 6 以外の目が出る回数を X とすると，$X=r$ の確率 $P(X=r)$ は次で与えられる：

$$ P(X=r)=\left(1-\frac{1}{6}\right)^r\frac{1}{6} \qquad (r=0,1,2,...). $$

このように，同じ試行を独立に繰り返すとき，確率 p の事象が始めて起きるまでに何回試行を行ったかを考える．

定義 2. p を $0 < p < 1$ を満たす実数とする．確率変数 X が，0 以上の整数に値をもち，$X = r$ の確率 $P(X = r)$ が

$$P(X = r) = (1-p)^r p \qquad (r = 0, 1, 2, ...)$$

によって与えられるとき，X はパラメータ p の**幾何分布**に従うという．

等比数列の和の公式

$$\sum_{r=0}^{\infty} cx^r = \frac{c}{1-x} \qquad (|x| < 1)$$

を用いると，$\sum_{r=0}^{\infty}(1-p)^r p = 1$ が分かる．

上で少しもどかしい言い方をしたのは，次の命題を簡単に述べるためである．

命題 1.1. m, r を 1 以上の整数とし，X が幾何分布に従うとすると，$X = m$ の下での $X = r + m$ の条件付き確率について次が成り立つ：

$$P(X \geqq r + m \mid X \geqq m) = P(X \geqq r).$$

$X \geqq r$ は r 回までに，考えている事象が起きなかったということを意味する．上の命題は，m 回試行を行って起きなかったときにさらに r 回試行を行っても起きないという事象の確率は m には依らず，最初から考えて r 回までに起きない確率と等しいことを意味する．

証明. 等比数列の和の公式より，自然数 n に対して

$$P(X \geqq n) = \sum_{k=n}^{\infty}(1-p)^k p = \frac{(1-p)^n p}{1-(1-p)} = (1-p)^n$$

である．したがって，

$$\begin{aligned} P(X \geqq r + m \mid X \geqq m) &= \frac{P(X \geqq r + m \text{ かつ } X \geqq m)}{P(X \geqq m)} = \frac{P(X \geqq r + m)}{P(X \geqq m)} \\ &= \frac{(1-p)^{r+m}}{(1-p)^m} = (1-p)^r = P(X \geqq r) \end{aligned}$$

となる． □

1.3　一般の離散分布，平均，分散

二項分布や幾何分布に従う確率変数のように，取り得る値が実数の離散的な部分集合である確率変数を**離散型確率変数**という．

X が離散型確率変数のとき，取り得る値の集合を $\{a_1, a_2, ..., a_N\}$ と書き，その確率を

$$P(X = a_i) = p_i \qquad (i = 1, 2, ...)$$

と書く．幾何分布のように，$N = \infty$ となる場合もある．

取り得る値とその確率のペア $\{(a_i, p_i)\}_{i=1}^N$ によって定まる値のばらつきを，**離散分布**という．試験の結果は，p_i を結果が a_i である受験生の割合とすると，離散分布の例を与える．

離散分布の特徴を表す量が，次の平均と分散である．

定義 3. (1) 確率変数 X が $\{(a_i, p_i)\}_{i=1}^N$ によって定まる離散分布に従うとき，

$$\sum_{i=1}^N a_i p_i$$

を X の**期待値**または**平均**と呼び，$E[X]$ または m によって表す．

(2) $X - m$ を X の**偏差**と呼び，$(X - m)^2$ の期待値

$$E[(X - m)^2] = \sum_{i=1}^N (a_i - m)^2 p_i$$

を X の分散と呼び，$V[X]$ または σ^2 $(\sigma > 0)$ と表す．

分散は散らばりの広さを表す基本的な量である．分散がゼロであるのは，確率変数の取り得る値が１つだけで確率的でない場合である．分散の平方根 $\sigma = \sqrt{V[X]}$ を**標準偏差**と呼ぶが，この講義ではあまり使わない．

$\{(a_i, p_i)\}_{i=1}^N$ が試験の結果の場合，平均については容易に理解されると思う．そして，

$$\frac{a_i - m}{\sigma} \times 10 + 50$$

が，得点 a_i の偏差値である．

命題 1.2. α, β を定数とすると，

$$E[\alpha X + \beta] = \alpha E[X] + \beta, \qquad V[\alpha X + \beta] = \alpha^2 V[X]$$

が成り立つ．とくに，$T = \dfrac{X - m}{\sigma}$ とおくと，T の平均は 0，分散は 1 である．

確率変数 T を X の**正規化**と呼ぶ．

証明. 平均については直感的に明らかだし，分散についても容易に分かるが，丁寧に示す．

確率変数 X が $\{(a_i, p_i)\}_{i=1}^N$ によって定まる離散分布に従うとする．平均については，

$$E[\alpha X+\beta]=\sum_{i=1}^N(\alpha a_i+\beta)p_i=\alpha\sum_{i=1}^N a_ip_i+\beta\sum_{i=1}^N p_i=\alpha E[X]+\beta$$

となる．

分散についても，$E[X]=m$ と書くと，

$$V[\alpha X+\beta]=E[\{\alpha X+\beta-(\alpha m+\beta)\}^2]=E[\alpha^2(X-m)^2]$$

であり，平均の性質から $\alpha^2E[(X-m)^2]$ に等しい．または，さらに，

$$E[\alpha^2(X-m)^2]=\sum_{i=1}^N\alpha^2(a_i-m)^2p_i=\alpha^2\sum_{i=1}^N(a_i-m)^2p_i=\alpha^2V[X]$$

となり，結論を得る． □

注意 1.3. $V[X]=E[X^2]-(E[X])^2$ が成り立ち，分散の計算に便利である．実際，確率変数 X が $\{(a_i, p_i)\}_{i=1}^N$ によって定まる離散分布に従うとき，

$$\begin{aligned}V[X]&=\sum_{i=1}^N(a_i-m)^2p_i=\sum_{i=1}^N(a_i^2-2ma_i+m^2)p_i\\&=\sum_{i=1}^n a_i^2p_i-2m\sum_{i=1}^N a_ip_i+m^2\sum_{i=1}^N p_i\end{aligned}$$

となり，

$$V[X]=E[X^2]-2mE[X]+m^2=E[X^2]-m^2$$

となる．記号としてはもっと簡単に，次のように計算しても同じ結果が得られる：

$$V[X]=E[(X-m)^2]=E[X^2-2mX+m^2]=E[X^2]-2mE[X]+m^2=E[X^2]-m^2.$$

演習 1.1. ある試験の結果が

点	40	60	80	100
人数	20	40	20	20

であったとする．結果の平均と分散を求めよ．

解答. 平均 68，分散 416

二項分布，幾何分布の平均と分散を計算する．

二項分布の平均，分散

命題 1.4. X が二項分布 $B(n,p)$ に従う確率変数のとき，平均と分散は次で与えられる：

$$E[X] = np, \qquad V[X] = np(1-p).$$

証明. 教科書とは違う証明を与える．示すべきことは

$$E[X] = \sum_{r=0}^{n} r\,{}_nC_r p^r (1-p)^{n-r} = np$$

である．

まず，二項定理を思い出す：

$$(x+y)^n = \sum_{r=0}^{n} {}_nC_r x^r y^{n-r}.$$

上の示すべきことと見比べると，x で微分することを思いつくので実行すると，

$$n(x+y)^{n-1} = \sum_{r=0}^{n} r\,{}_nC_r x^{r-1} y^{n-r}$$

となり，x をかけると

$$nx(x+y)^{n-1} = \sum_{r=0}^{n} r\,{}_nC_r x^r y^{n-r} \tag{\#}$$

となる．

よって，$x=p, y=1-p$ を代入すると，

$$\sum_{r=0}^{n} r\,{}_nC_r p^r (1-p)^{n-r} = np(p+(1-p)) = np$$

となり，$E[X]=np$ を得る．

次に，

$$E[X^2] = \sum_{r=0}^{n} r^2\,{}_nC_r p^r (1-p)^{n-r}$$

を計算する．

このためには，(#) の両辺を x で微分して，両辺に x を掛けると

$$n\{(x+y)^{n-1} + (n-1)x(x+y)^{n-2}\} = \sum_{r=0}^{n} r^2\,{}_nC_r x^{r-1} y^{n-r},$$

$$n\{(x+y)^{n-1} + (n-1)x(x+y)^{n-2}\}x = \sum_{r=0}^{n} r^2\,{}_nC_r x^r y^{n-r}$$

を得る．ここで，$x = p, y = 1 - p$ を代入すると，

$$E[X^2] = \sum_{r=0}^{n} r^2 {}_nC_r p^r (1-p)^{n-r} = n(1 + (n-1)p)p = np(1 - p + np^2)$$

となる．したがって，次のように結論を得る：

$$V[X] = E[X^2] - (E[X])^2 = np(1 - p + np^2) - (np)^2 = np(1-p).$$

□

幾何分布の平均，分散

命題 1.5. X がパラメータ p の幾何分布に従う確率変数のとき，平均と分散は次で与えられる：

$$E[X] = \frac{1-p}{p}, \qquad V[X] = \frac{1-p}{p^2}.$$

証明. 平均について，計算べきものは

$$E[X] = \sum_{r=0}^{\infty} r(1-p)^r p = (1-p)p + 2(1-p)^2 p + 3(1-p)^3 p + \cdots + r(1-p)^r p + \cdots$$

である．ここで，両辺に $1-p$ を掛けると

$$(1-p)E[X] = \sum_{r=0}^{\infty} r(1-p)^{r+1} p = (1-p)p^2 + 2(1-p)^3 p + 3(1-p)^4 p + \cdots + r(1-p)^{r+1} p + \cdots$$

だから，辺々引くと

$$pE[X] = (1-p)p + (1-p)^2 p + \cdots + (1-p)^r p + \cdots = \frac{(1-p)p}{1-(1-p)} = 1 - p$$

となり，$E[X] = \dfrac{1-p}{p}$ を得る．

同様に，$E[X^2]$ と $(1-p)E[X^2]$ を無限和で表して辺々を引くと

$$\begin{aligned}(1-(1-p))E[X^2] &= (1-p)p + \sum_{r=2}^{\infty} (r^2 - (r-1)^2)(1-p)^r p \\ &= (1-p)p + \sum_{r=2}^{\infty} (2r-1)(1-p)^r p = \sum_{r=1}^{\infty} (2r-1)(1-p)^r p\end{aligned}$$

となる．したがって，

$$pE[X^2] = 2E[X] - (1-p) = \frac{2(1-p)}{p} - (1-p) = \frac{p^2 - 3p + 2}{p}$$

となり (なぜか？考えてみてほしい),

$$V[X] = E[X^2] - (E[X])^2 = \frac{p^2 - 3p + 2}{p^2} - \left(\frac{1-p}{p}\right)^2 = \frac{1-p}{p^2}$$

となり，結論を得る．(少し計算を省略した．演習で別の方法を与える．) □