

応用確率統計 教科書章末問題解答補足

担当教官：荒木純道

東京工業大学 工学部 電気電子工学科

1 第2章

1.1 問2

1. 確率の定義より

$$\int p(x)dx = 1 \quad (1)$$

が成立するような A を求めれば良い。

$$\begin{aligned} 1 &= A \int_{-\infty}^{\infty} \frac{1}{e^x + e^{-x}} dx \\ &= A \int_0^{\infty} \frac{dt}{t^2 + 1} \quad (e^x = t) \\ &= A \left[\tan^{-1}(t) \right]_{t=0}^{\infty} \\ &= A \frac{\pi}{2} \\ A &= \frac{2}{\pi} \end{aligned}$$

2. 累積分布関数の定義は以下の通り

$$P(a) = \int_{-\infty}^a p(x)dx \quad (2)$$

定義と先ほどの問題と同様の変数変換により

$$\begin{aligned} P(a) &= \frac{2}{\pi} \int_{-\infty}^a \frac{1}{e^x + e^{-x}} dx \\ &= \frac{2}{\pi} \left[\tan^{-1}(t) \right]_{t=0}^{e^a} \\ &= \frac{2}{\pi} \tan^{-1}(e^a) \end{aligned}$$

3. 本問題は教科書の最後の方の解答が誤りである。区間 $[0, 1]$ の連続一様分布に従う確率変数を X とし、本問題で与えられている確率密度関数に従う確率変数を Y とすると、以下のような変数変換で求めることが可能

$$X = \frac{2}{\pi} \tan^{-1}(e^Y) \quad (3)$$

一変数での変数変換であれば以下のように求めることができる。

$$\begin{aligned} p_Y(y) &= p_X(x) \left| \frac{dx}{dy} \right| \\ &= \frac{2}{\pi} \frac{1}{1 + (e^y)^2} e^y \\ &= \frac{2}{\pi} \frac{1}{e^y + e^{-y}} \end{aligned}$$

ただし、途中の式変形では区間 $[0, 1]$ の連続一様分布は $p_X(x) = 1 (0 < x < 1)$ を利用し、また $(\tan^{-1})' = (1 + x^2)^{-1}$ を利用した。

1.2 問3

本問題はボックス・ミュラー法と呼ばれる一様分布から標準正規分布を生成する方法である。元々の形が非常に煩雑なため、ある程度簡単化するために以下のような変数変換を行い、それぞれの確率変数に関して確率密度関数を算出する。

$$S = \sqrt{-2 \ln X} \quad (4)$$

$$T = \cos(2\pi Y) \quad (5)$$

この式を逆に X, Y について解き直すと

$$X = e^{-s^2/2} \quad (6)$$

$$Y = \frac{1}{2\pi} \cos^{-1}(T) \quad (7)$$

となる。1変数の場合の変数変換は上述のように求めることができるため

$$\begin{aligned} p_S(s) &= p_X(x) \left| \frac{dx}{ds} \right| \\ &= s e^{-s^2/2} \\ p_T(t) &= p_Y(y) \left| \frac{dy}{dt} \right| \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{1-t^2}} \end{aligned}$$

最後に Z の確率密度関数を求める。2変数から1変数の確率変数を算出するには以下の式を用いる必要がある。

$$\begin{aligned} p_Z(z) &= \int_0^{\infty} \int_{-1}^1 \delta(z - st) p_S(s) p_T(t) ds dt \\ &= \int_{|s| > z} \frac{p_S(s) p_T(z/s)}{|s|} ds \\ &= \int_{|s| > z} \frac{s e^{-s^2/2}}{|s|} \frac{|s|}{2\pi \sqrt{s^2 - z^2}} ds \\ &= \frac{1}{2\pi} \int_{|s| > z} \frac{s e^{-s^2/2}}{\sqrt{s^2 - z^2}} ds \\ &= \frac{1}{\pi} \int_z^{\infty} \frac{s e^{-s^2/2}}{\sqrt{s^2 - z^2}} ds \end{aligned}$$

続いて、何度かの変数変換を行うと算出できる

$$\begin{aligned}
 &= \frac{1}{\pi} \int_0^\infty \frac{e^{-(a+z^2)/2}}{\sqrt{a}} da/2 \quad (a = s^2 - z^2) \\
 &= \frac{e^{-z^2/2}}{2\pi} \int_0^\infty \frac{e^{-a/2}}{\sqrt{a}} da \\
 &= \frac{e^{-z^2/2}}{2\pi} \int_0^\infty \frac{e^{-b^2/2}}{b} 2b db \quad (b = \sqrt{a}) \\
 &= \frac{e^{-z^2/2}}{\pi} \int_0^\infty e^{-b^2/2} db \\
 &= \frac{e^{-z^2/2}}{\pi} \sqrt{\pi/2} \\
 &= \frac{e^{-z^2/2}}{\sqrt{2\pi}}
 \end{aligned}$$

第2回の付録のデルタ関数の性質を利用し、またガウス積分により最後の積分を終えた。ガウス積分を行わなくとも正規分布の性質を利用して求めることも可能である。

2 第3章

2.1 問1

こちらは問題文の分散の式に誤りがあり、以下の式が正しい。

$$V(X+Y) = V(X) + V(Y) + 2E(XY) - 2E(X)E(Y) \quad (8)$$

2.2 問3

1. マルコフ不等式を $f(x) = |x|^p$ ($p > 0$) として利用。

$$\begin{aligned}
 P(|X|^p > tE(|X|^p)) &\leq 1/t \\
 P(|X| > (tE(|X|^p))^{1/p}) &\leq 1/t
 \end{aligned}$$

ここで $\epsilon = (tE(|X|^p))^{1/p}$ とし、更に t について解いて右辺の t に代入すれば

$$\begin{aligned}
 t &= \epsilon^p / E(|X|^p) \\
 P(|X| > \epsilon) &\leq E(|X|^p) / \epsilon^p
 \end{aligned}$$

2. $X \rightarrow X_n - X$ と置換すれば $n \rightarrow \infty$ としたとき右辺は p 次平均収束するため0に収束する。このため、前問の不等式より確率収束することがわかる。

3 第4章

3.1 問1

1. 与えられている式はテイラー展開により導かれる式である。またポアソン分布は離散分布であるため、積分ではなく、 n に関しての和の演算を行うことで特性関数を求めることができる。

$$\begin{aligned}
 \phi(t) &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} e^{itn} \\
 &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(e^{it}\lambda)^n}{n!} \\
 &= e^{-\lambda} e^{\lambda e^{it}} \\
 &= e^{\lambda(e^{it}-1)}
 \end{aligned}$$

となる。途中で与えられた式を用いた。次の問題は特性関数とモーメントの関係をを使えば容易に求められる。

3.2 問3

確率変数 X, Y が独立であるため、 $E(e^{t(X+Y)}) = E(e^{tX})E(e^{tY})$ が成立する。このため、

$$\begin{aligned}
 k_n(X+Y) &= \frac{d^n}{dt^n} \log(E(e^{t(X+Y)})) \\
 &= \frac{d^n}{dt^n} \log(E(e^{tX})E(e^{tY})) \\
 &= \frac{d^n}{dt^n} \log(E(e^{tX})) + \frac{d^n}{dt^n} \log(E(e^{tY})) \\
 &= k_n(X) + k_n(Y)
 \end{aligned}$$

となる。

4 第5章

4.1 問2

δ 関数を用いると離散の確率分布を連続確率分布へと変形することができる。この場合、分布は連続になっているため、積分で平均や分散を求めることができる。ベイズの定理より、 $p(x, y) = p(y|x)p(x)$ となるため、 $p(y|x) = p(x, y)/p(x)$ となる。問1の結果を用いることで

$$\begin{aligned}
 r(x) &= \int y \frac{p(x, y)}{p(x)} dy \\
 &= \frac{1}{p(x)} \int y [a\delta(y)p_0(x) + (1-a)\delta(y-1)p_1(x)] dy \\
 &= \frac{1}{p(x)} [(1-a)p_1(x)] \\
 &= p(n=1|x)
 \end{aligned}$$

となる。

5 第6章

5.1 問2

以下のように確率変数 X を定義する。

$$X = \begin{cases} 1 & (p_X(x=1) = p) \\ 0 & (p_X(x=0) = 1-p) \end{cases} \quad (9)$$

また、独立で X と同一な分布に従う確率変数 $X_1, \dots, X_n$ を定義し、新しい確率変数 Y_n を以下のように定義する

$$Y_n = \frac{X_1 + \dots + X_n}{n} \quad (10)$$

中心極限定理より、 Y_n は平均 $E(Y_n) = p$ 、分散 $V(Y_n) = (p-p^2)/n$ の正規分布に従う。

$$p_{Y_n}(y_n) = \frac{1}{\sqrt{2\pi(p-p^2)/n}} \exp\left(-\frac{(y_n-p)^2}{2(p-p^2)/n}\right) \quad (11)$$

1. 今、試行回数 10000 回で表が 3500 回出ている。このとき、99% 信頼区間を求めると以下の値の範囲となる。

$$|Y_n - p| < 3\sqrt{\frac{p - p^2}{10000}} \quad (12)$$

今回の試行により、得られた Y_n の実現値は 3500/10000 であるため、以上の不等式に代入して p に関して変形すると

$$|p - 0.35| < 3\sqrt{\frac{p - p^2}{10000}} \quad (13)$$

となり、 $p \approx 0.35$ より、この値を右辺に代入すること

$$0.3357 < p < 0.3643 \quad (14)$$

となり、確率 p の推定を行うことができる。

2. 以上の議論を一般化する。 N 回の試行で $0.35N$ 回表が出たものとする。99% 信頼区間は

$$|Y_N - p| < 3\sqrt{\frac{p - p^2}{N}} \quad (15)$$

であるため、 p に関して整理すれば

$$|p - Y_N| < 3\sqrt{\frac{p - p^2}{N}} \quad (16)$$

となる。このとき、 Y_N の実現値は 0.35 であり、 $p \approx 0.35$ より、この値を右辺に代入することで

$$|p - 0.35| < 3\sqrt{\frac{0.35 \cdot 0.65}{N}} \quad (17)$$

精度を 2 倍にするため、精度の低い試行回数を N_1 、精度の高い試行回数を N_2 とする。ただし、 $N_1 < N_2$ である。このとき、互いのばらつきが以下の不等式を満たせば良い

$$3\sqrt{\frac{0.35 \cdot 0.65}{N_1}} / 3\sqrt{\frac{0.35 \cdot 0.65}{N_2}} \geq 2 \quad (18)$$

この式を変形することで

$$\begin{aligned} \sqrt{\frac{N_2}{N_1}} &\geq 2 \\ \frac{N_2}{N_1} &\geq 4 \\ \therefore N_2 &\geq 4N_1 \end{aligned}$$

となるため、 $N_1 = 10000$ とすれば、 $N_2 \geq 40000$ となる。

6 第7章

6.1 問1

2 項分布は実は、第 7 回の演習の問 1(b) で定義された Y_n が本来従う分布である。このため、2 項分布の平均は

$$\sum_{r=0}^n r_n C_r a^r (1-a)^{n-r} = na \quad (19)$$

これを本問題では利用する。離散のカルバック情報量の定義式に基づいて式変形していく

$$\begin{aligned} D(p||q) &= \sum_{r=0}^n n C_r a^r (1-a)^{n-r} \log \frac{a^r (1-a)^{n-r}}{b^r (1-b)^{n-r}} \\ &= \sum_{r=0}^n n C_r a^r (1-a)^{n-r} \left[r \log \left(\frac{a}{b} \right) \right. \\ &\quad \left. + (n-r) \log \left(\frac{1-a}{1-b} \right) \right] \\ &= \log \left(\frac{a}{b} \right) \sum_{r=0}^n r_n C_r a^r (1-a)^{n-r} \\ &\quad + \log \left(\frac{1-a}{1-b} \right) \left[n \sum_{r=0}^n n C_r a^r (1-a)^{n-r} \right. \\ &\quad \left. - \sum_{r=0}^n r_n C_r a^r (1-a)^{n-r} \right] \\ &= \log \left(\frac{a}{b} \right) na + \log \left(\frac{1-a}{1-b} \right) (n - na) \\ &= n \left[a \log \left(\frac{a}{b} \right) + (1-a) \log \left(\frac{1-a}{1-b} \right) \right] \end{aligned}$$

となる。よって $f(a, b)$ は以下の通りである

$$\therefore f(a, b) = a \log \left(\frac{a}{b} \right) + (1-a) \log \left(\frac{1-a}{1-b} \right) \quad (20)$$

6.2 問2

1. 統計分野では最尤推定と呼ばれる未知のパラメタを推定する方法があり、その方法と同じことを行うのが本問題である。ここで $q_a(x)$ は

$$q_a(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-a)^2/2} \quad (21)$$

であり、この関数の対数をとると

$$\log(q_a(x)) = -(x-a)^2/2 - \log(2\pi)/2 \quad (22)$$

本問題で与えられている $L(a)$ も同様に対数をとれば

$$\begin{aligned} \log(L(a)) &= \log\left(\prod_{i=1}^n q_a(X_i)\right) \\ &= \sum_{i=1}^n \log(q_a(X_i)) \\ &= \sum_{i=1}^n \left[-(X_i - a)^2/2 - \log(2\pi)/2 \right] \\ &= \sum_{i=1}^n \left[-(X_i - a)^2/2 \right] - n \log(2\pi)/2 \end{aligned}$$

となり、右辺の第 1 項のみ a に依存することがわかる。ここで、両辺 a について微分すると

$$\begin{aligned}\frac{d \log(L(a))}{da} &= \sum_{i=1}^n (X_i - a) \\ &= \sum_{i=1}^n X_i - an = 0 \\ \therefore a^* &= \frac{1}{n} \sum_{i=1}^n X_i\end{aligned}$$

となり、算術平均が最適な a の値を決めることがわかる。

2. 第 6 回の講義資料より明らかなように平均 0, 分散 $1/n$ となる。