大数の法則と中心極限定理

芝浦工業大学　数理科学研究会
BV19083　久保田　静希

2020 年 6 月 12 日

目 次

1 確率変数 **1**
1.1 確率変数と確率分布 1
1.2 累積分布関数 1
1.3 確率変数の期待値と分散 2

2 モーメント **3**
2.1 モーメント 3
2.2 モーメント母関数 4

3 正規分布 **5**
3.1 正規分布 5
3.2 正規分布のモーメント母関数 5

4 大数の法則 **6**
4.1 チェビシェフの不等式 6
4.2 大数の法則 6

5 中心極限定理 **7**

6 今後の課題 **8**

研究背景

去年の大宮祭でベイズ統計学を学んだがベイズ統計学だけではなく，ほかの統計学の考え方についても興味が出てきたため大数の法則と中心極限定理を学ぼうと思った．この二つの定理を証明しより深く理解することを目標にしたいと思う．

1　確率変数

その各変数に対しそれぞれ確率が与えられている変数のことを確率変数といい一般に X などの大文字で表すことが多い．またその確率を $P(X)$ と表す．以下 X を任意の確率分布に従う確率変数．$P(X)$ を X の起こる確率とする．

1.1　確率変数と確率分布

一般に加算集合 $\{x_1, x_2, \dots\}$ の中の値をとる確率変数 X は離散型といわる．このとき．

$$P(X = x_k) = f(x_k)\ (k = 1, 2, \dots)$$

を満たす関数 f を X の離散型の確率分布という．また $f(x_k)$ は確率であるので，

$$f(x_k) \geq 0\ (k = 1, 2, \dots)\ \text{かつ}\ \sum_{k=1}^{\infty} f(x_k) = 1$$

と表す．x の添え字は省略されることも多い．確率分布 f は確率の重みの様子を表している．このような重みの考えから連続型の確率変数を考える．非加算集合の中の値をとる確率変数 X を連続型の確率変数という．今，X が $a \leq X \leq b$ の範囲の値をとる確率を

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

としたとき関数 f を確率密度関数という．このとき集合 $(-\infty, \infty)$ の任意の要素 x に対し

$$f(x) \geq 0,\ \text{かつ}\ \int_{-\infty}^{\infty} f(x) = 1$$

を満たす.

1.2　累積分布関数

確率変数 X に対し x を実数とするとき，x 以下の値をとる確率を累積分布関数 $F(x)$ という．

定義 1.1. 累積分布関数

$$F(x) = P(X \geq x)$$

離散型の場合 F は確率密度関数 f の定積分

$$F(x) = \int_{-\infty}^{x} f(u)du$$

もしくは微分積分学の基本定理より

$$F'(x) = f(x)$$

で定義される．離散型の場合

$$F(x) = \sum_{u \leq x} f(u)$$

で定義される．

1.3 確率変数の期待値と分散

確率変数 X に対する期待値を $E(X)$ と表す.

定義 1.2. 確率変数の期待値

$$E(X) = \sum_x xf(x) \text{（離散型）}$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \text{ (連続型)}$$

但し，無限和，積分であるので存在しないこともある．X の関数 $\phi(X)$ についても同様に

$$E(\phi(X)) = \sum_x \phi(X)f(x) \text{（離散型）}$$

$$E(\phi(X)) = \int_{-\infty}^{\infty} \phi(X)f(x)dx \text{ (連続型)}$$

但し $f\{\phi(X)\} = f(x)$ である.

定理 1.3. 期待値の性質

$$E(c) = c \ldots (a)$$

$$E(X+c) = E(X) + c \ldots (b)$$

$$E(cX) = cE(X) \ldots (c)$$

$$E(X+Y) = E(X) + E(Y) \ldots (d)$$

但し c は定数，Y は確率変数.
確率変数 X に対する分散を $V(X)$ と表す.

定義 1.4. 確率変数の分散

$$V(X) = \sum_x (x-\mu)^2 f(x) \text{（離散型）}$$

$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \text{ (連続型)}$$

しかしこの定義のまま計算するよりも定理 1.3 を用いて式変形した形

$$\begin{aligned} V(X) &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - \{E(X)\}^2 \end{aligned}$$

を用いたほうが良い.

定理 1.5. 分散の性質

$$V(c) = 0 \ldots (a)$$

$$V(X+c) = V(X) \ldots (b)$$

$$V(cX) = c^2 V(X) \ldots (c)$$

また，$D(X) = \sqrt{V(X)}$ を X の標準偏差という.

定理 1.6. 標準化
期待値の性質と分散の性質を用いると

$$Z = \frac{X - E(X)}{\sqrt{V(X)}}$$

という変数を考えたとき，

$$E(Z) = 0, V(X) = 1$$

となるこの変換を X の標準化といい Z を標準化変数という．

2　モーメント

2.1　モーメント

モーメント（積率）とは確率分布の形を決める指標であり μ_r と表す．

定義 2.1. モーメント

$$\mu_r = E(X^r)$$

を X の（原点周りの）r 次のモーメント（積率）といい

$$\mu_r' = E(X - \mu)^r$$

を X の期待値のまわりの r 次のモーメントという．また，

$$\alpha_r = E(\frac{X - \mu}{\sigma})^r$$

を X の r 次の標準化モーメントという．期待値，分散はモーメントの基礎的なもので

$$\mu_1 = E(X) \quad \mu_2' = V(X)$$

である．

$$\alpha_3 = E(\frac{X - \mu}{\sigma})^3$$

を歪度といい確率密度関数 $f(x)$ がどの程度非対象で，右側に偏っているか，左に偏っているかがわかる．$\alpha > 0$ だと右に偏っており，$\alpha < 0$ だと左に偏っている．$|\alpha|$ の大きさでその程度を表すことができる．歪度は定義通り計算することは少なく期待値の性質を用いて

$$E(\frac{X - \mu}{\sigma})^3 = E(X^3) - 3\mu E(X^2) - 2\mu^2$$

とすることが多い．

$$\alpha_4 = E(\frac{X - \mu}{\sigma})^4$$

は中心の周囲のとがり具合を表す．正規分布は $\alpha_4 = 3$ なので $\alpha_4 - 3$ を尖度といい $\alpha_4 - 3 > 0$ なら正規分布より尖っており，$\alpha_4 - 3 < 0$ なら丸く鈍った形をしている．
期待値，分散，歪度，尖度などの値を指定すると，それに該当する確率分布の候補は制限されてくる．全ての次数のモーメントを指定できればそれにより一つの確率分布が決定されるはずである．よってモーメント母関数 $M_X(t)$ というものを定義する．

2.2 モーメント母関数

定義 2.2. モーメント母関数
$$M_X(t) = E(e^{tX})$$
計算方法は
$$M_X(t) = \sum_x e^{tx} f(x) \text{ (離散型)}$$
$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx \text{ (連続型)}$$
となる．ただし無限和，積分が存在しないこともある．定義 2.2 よりモーメント母関数を繰り返し微分して 0 とおいた導関数は

$$M'_X(0) = \mu_1,\ M''_X(0) = \mu_2,\ M'''_X(0) = \mu_3\ \ \ldots$$

などとなる．一般に
$$M_X^{(r)}(0) = \mu_r$$
のように，各次数のモーメントがわかる．証明は e^x の展開式
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \ldots$$
に $x = tX$ を代入すると
$$e^{tX} = 1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} \ldots$$
ここで両辺の期待値をとる（すなわち $E(X)$ に代入する）と定理 1.3 と定義 2.2 より
$$\begin{aligned} M_X(t) &= 1 + tE(X) + \frac{t^2 E(X^2)}{2!} + \frac{t^3 E(X^3)}{3!} \ldots \\ &= 1 + t\mu_1 + \frac{t^2\mu_2}{2!} + \frac{t^3\mu_3}{3!} \ldots \end{aligned}$$
よってモーメント母関数の導関数からモーメントがわかることは自明である．

定理 2.3. モーメント母関数の性質モーメント母関数には以下の性質が成り立つ。
$$M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t) \ldots (a)$$
$$M_{aX}(t) = M_X(at) \ldots (b)$$
但し，X_1, X_2 は互いに独立な確率変数とし，a は定数とする．

証明. (a) X_1, X_2 を離散型確率変数とする．また、f, g, h をそれぞれ $X_1 + X_2, X_1, X_2$ の確率密度関数とすると，X_1, X_2 は互いに独立であることから $f(X_1 + X_2) = g(X_1)h(X_2)$ となるので
$$\begin{aligned} M_{X_1+X_2}(t) &= E(e^{t(X_1+X_2)}) \\ &= \sum_{x_1+x_2} e^{t(x_1+x_2)} f(x_1 + x_2) \\ &= \sum_{x_1} e^{tx_1} g(x_1) \sum_{x_2} e^{tx_2} h(x_2) \\ &= M_{X_1}(t) M_{X_2}(t) \end{aligned}$$
連続型も同様である．
(b) 定義より自明.

3　正規分布

正規分布 (ガウス分布) とは代表的な連続型確率変数の確率分布である．誤差の測定や自然界など，とても広い現象に対して当てはまり統計学の理論上も応用上も非常に重要な確率分布である．

3.1　正規分布

定義 3.1. 正規分布の確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ -\infty < x < \infty$$

ただし $\mu = E(X), \sigma^2 = V(X)$ である．ここで定数 $\frac{1}{\sqrt{2\pi}\sigma}$ は

$$\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$$

から来ており $\int_{-\infty}^{\infty} f(x)dx = 1$ とするための規格定数である．また，平均 μ, 分散 σ^2 の正規分布を $N(\mu, \sigma^2)$ と表記する．変数を標準化した正規分布 $N(0,1)$ を標準正規分布という．

3.2　正規分布のモーメント母関数

定理 3.2. 正規分布のモーメント母関数
確率変数 X が正規分布 $N(\mu, \sigma)$ に従うとする．このとき，X のモーメント母関数 $M_X(t)$ は

$$M_X(t) = e^{\mu t + \frac{\sigma^2}{2}t^2}$$

証明.

$$\begin{aligned}
M_X(t) &= E\{e^{t(X)}\} \\
&= \int_{-\infty}^{\infty} e^{tx} f(x) dx \\
&= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2} + tx} dx
\end{aligned}$$

ここで被積分関数の指数部を平方完成すると，

$$-\frac{(x-\mu)^2}{2\sigma^2} + tx = -\frac{1}{2\sigma^2}\{(x-(\mu+\sigma^2 t))^2\} + \mu t + \frac{\sigma^2}{2}t^2$$

従って

$$\begin{aligned}
M_X(t) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\{(x-(\mu+\sigma^2 t))^2\} + \mu t + \frac{\sigma^2}{2}t^2} dx \\
&= e^{\mu t + \frac{\sigma^2}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{\{x-(\mu+\sigma^2 t)\}^2}{2\sigma^2}} dx
\end{aligned}$$

ここで被積分関数 $\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{\{x-(\mu+\sigma^2 t)\}^2}{2\sigma^2}}$ は正規分布 $N(\mu+\sigma^2 t, 2\sigma^2)$ の確率密度関数となるので

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{\{x-(\mu+\sigma^2 t)\}^2}{2\sigma^2}} dx = 1$$

よって

$$M_X(t) = e^{\mu t + \sigma^2 t^2}$$

4　大数の法則

大数の法則とは標本の数が十分に大きいとき，その標本平均を母集団の平均としてもよいという常識を数学的に厳密に証明したものである．

4.1　チェビシェフの不等式

チェビシェフの不等式とは具体的な確率密度関数が分かっていない場合でも期待値と分散が分かっていれば，確率を不等式で表せるという定理である．

定理 4.1. チェビシェフの不等式
$\mu = E(X), \sigma^2 = V(X)$ とする．いかなる確率変数 X に対しても

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}\ (k:\text{定数})$$

が成り立つ．

証明. $I = \{x : |x-\mu| \geq k\sigma\}$, f を確率密度関数とすると

$$\begin{aligned}\sigma^2 &= V(X)\\ &= \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx\\ &\geq \int_I (x-\mu)^2 f(x)dx\end{aligned}$$

ここで $(x-\mu)^2 \geq k\sigma$ より

$$\begin{aligned}\sigma^2 &\geq (k\sigma)^2\int_I f(x)dx\\ &= (k\sigma)^2 P(|X-\mu| \geq k\sigma)\end{aligned}$$

よって両辺 $(k\sigma)^2$ で割ると

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

4.2　大数の法則

定理 4.2. 大数の法則 ϵ を任意の正の定数，もとの確率分布の平均を μ，その分布から n 個とられた観測値の平均を $\overline{X_n}$ としたとき

$$P(|\overline{X_n}-\mu| < \epsilon) \to 1\ (n \to \infty)$$

証明. 一般に，確率変数 $X_1, X_2 \ldots X_n$ は独立で，同一確率分布に従うとする．このとき，各 $X_i(i = 1, 2, \ldots n)$ の分散，期待値は同一なので $E(X_i) = \mu, V(X_i) = \sigma^2$ とする．$\overline{X} = X_1 + X_2 + \cdots + X_n$ とすると，定理 1.3 より

$$E(\overline{X}) = \mu, V(\overline{X}) = \frac{\sigma^2}{n}$$

ここで定理 4.1 チェビシェフの不等式

$$P(|\overline{X_n} - \mu| > k\sigma_n) \leq \frac{1}{k^2} \ (k : 定数)$$

より $k = \frac{\epsilon}{\sigma_n}$ とすると，$\sigma_n^2 = V(\overline{X_n})$ より $k = \frac{\sqrt{n}\epsilon}{\sigma}$ なので

$$P(|\overline{X_n} - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \to 0 \ (n \to \infty)$$

すなわち

$$P(|\overline{X_n} - \mu| < \epsilon) \to 1 \ (n \to \infty)$$

5　中心極限定理

定理 5.1. 中心極限定理中心極限定理とは，X が平均 μ，標準偏差 σ のある分布に従うならば，大きさ n の無作為標本基づく標本平均 $\overline{X_n}$ の分布は正規分布 $N(\mu, \frac{\sigma}{\sqrt{n}})$ に近づくという定理である．標準化したものを数式で表すと

$$P(a \leq \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} \leq b) \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \ (n \to \infty)$$

証明. $\frac{X_n - \mu}{\sqrt{n}\sigma} = Y_n$ とすると

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} = \frac{Y_1 + Y_2 + \cdots + Y_n}{\sqrt{n}}$$

となる．また，$X_1, X_2, \ldots, X_n$ の確率分布は同一なので $Y_1, Y_2, \ldots, Y_n$ の確率分布も同一である．ここで，$Y_1, Y_2, \ldots, Y_n$ のなかから一つとったものを Y とし，そのモーメント母関数を $M_Y(t)$ とすると Y は $X_1, X_2, \ldots, X_n$ を標準化した確率変数なので $E(Y) = 0,\ E(Y^2) = V(Y) = 1$ となる．よって

$$M_Y(t) = 1 + \frac{t^2}{2} + \frac{E(Y^3)}{3!}t^3 + \ldots$$

となる．このとき，$Y_1 + Y_2 + \cdots + Y_n = T$ とすると T のモーメント母関数 $M_T(t)$ は定理 $2.3 \ldots (a)$ より

$$\begin{aligned} M_T(t) &= \{M_Y(t)\}^n \\ &= \{1 + \frac{t^2}{2} + \frac{E(Y^3)}{3!}t^3 + \ldots\}^n \end{aligned}$$

となる．よって確率変数 $\frac{T}{\sqrt{n}}$ のモーメント母関数 $M_{\frac{T}{\sqrt{n}}}(t)$ は定理 $2.3 \ldots (b)$ より

$$\begin{aligned} M_{\frac{T}{\sqrt{n}}}(t) &= M_T(\frac{t}{\sqrt{n}}) \\ &= \{1 + \frac{t^2}{2n} + \frac{E(Y^3)}{3!n\sqrt{n}}t^3 + \ldots\}^n \\ &\to e^{\frac{t^2}{2}} \ (n \to \infty) \end{aligned}$$

定理 3.2 より，これは標準正規分布のモーメント母関数なので中心極限定理は示された．

6 今後の課題

今回は大数の法則と中心極限定理を学んだが次は区間推定法などについても学んでいきたいと思った.

参考文献

[1] 松原望, 縄田和満, 中井検裕, 統計学入門，東京大学出版会, 1991.

[2] 分析ノート，正規分布のモーメント母関数を導出する，2020 /05 /18, https://analytics-note.xyz/statistics/normal-moment/