

1 確率変数列の収束について

確率変数列 X_n の収束には概収束、平均収束、確率収束、分布収束がある。その中でも特に、大数の法則や中心極限定理で出てく確率収束と分布収束についてまとめる。

1.1 確率収束

1.2 定義

確率変数列 X_n (たとえば標本平均 $X_n = \frac{1}{n} \sum_i X_i$) が確率変数 X (定数でもよい) に確率収束するとは、

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} P(|X_n - X| \geq \epsilon) = 0 \quad (1)$$

となることを表し、 $X_n \xrightarrow{P} X$ と記す。

X の周りの区間にどれだけ狭い区間 $(X - \epsilon, X + \epsilon)$ をとっても、 n を大きくすることで X_n がその区間外になる確率をいくらでも小さくできる。

1.3 例

サイコロを振って 1 が出る確率を測定するとする。試行回数を増やすと出る確率は $1/6$ 収束するはずであり、このときに確率収束するといえる。 i 回目サイコロを振って 1 が出た場合に 1、それ以外の場合にゼロを取る確率変数 X_i を考える。 n 回試行したときの確率変数の平均は

$$X_n = \frac{1}{n} \sum_{i=1}^n X_i$$

と表され、 $X_n \xrightarrow{P} \frac{1}{6}$ ($n \rightarrow \infty$) である。この収束を理論的に裏付けるのが大数の法則である。

1.4 分布収束 (法則収束、弱収束)

確率変数 Z_n の分布 F_n が特定の連続分布 F に分布収束 (弱収束、法則収束) するとは、

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \quad (2)$$

となることを表し、 $F_n \xrightarrow{d} F$ と記す。この場合 F を Z_n の漸近分布 (極限分布) といい、 Z_n は漸近的に分布 F に従うという。

2 大数の法則

大数の法則

X_1, \dots, X_n が互いに独立に分布関数に従い、 $E[X_i] = \mu, V[X_i] = \sigma^2$ が存在するとする。 $n \rightarrow \infty$ のときに 標本平均 \bar{X}_n が母平均 μ に収束すること を主張する。

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mu \quad (n \rightarrow \infty) \quad (3)$$

- 標本数 n を大きくすると標本平均 \bar{X}_n は、個々の確率変数が従う母平均 $E[X] = \mu$ に確率収束する。

2.1 適当な感覚

$E[X] = \mu, V[X] = \sigma^2$ が存在するとしているので、標本平均についての平均と分散はそれぞれ

$$E\left[\bar{X} = \frac{1}{n} \sum X_i\right] = \mu \quad (4)$$

$$V\left[\bar{X} = \frac{1}{n} \sum X_i\right] = \sigma^2/n \quad (5)$$

である。 $n \rightarrow \infty$ のとき分散がゼロになるので、標本平均は μ の周りに分散ゼロで分布することになり、母平均に収束しそうな気がする。

3 中心極限定理

中心極限定理

X_1, \dots, X_n が互いに独立に分布関数に従い、標本平均 \bar{X} について $E[\bar{X}_n] = \mu, V[\bar{X}_n] = \sigma^2/n$ が存在するとする。標準化した標本平均の累積分布関数が標準正規分布の累積分布関数に収束する。

$$P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x\right) \rightarrow \Phi(x) \quad (n \rightarrow \infty) \quad (6)$$