

8 標本平均と中心極限定理

8.1 標本平均

統計処理／分析の対象となるデータ・確率変数を母集団と呼びます。

独立で同じ分布に従う確率変数のファミリーは今後略して i.i.d. (independent and identically distributed) であると言います。

定義 8.1.1 母集団と同じ分布に従う i.i.d. な n 個の確率変数のファミリーを、この母集団からとった大きさ n の標本と呼びます。

定義 8.1.2 母集団 X からとった大きさ n の標本 $X_1, \dots, X_n$ に対して、

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

で定まる確率変数 $\bar{X}$ を、この母集団からとった大きさ n の標本平均と言います。

$$E[\bar{X}] = m, \quad \text{Var}[\bar{X}] = \frac{v}{n} \quad (\text{母集団の平均 } m \cdot \text{分散 } v)$$

8.2 母集団が正規分布であった場合

定理 8.2.1 それぞれ $N(m, s), N(n, t)$ に従う 2 つの独立な正規分布の和は正規分布 $N(m+n, s+t)$ に従います。

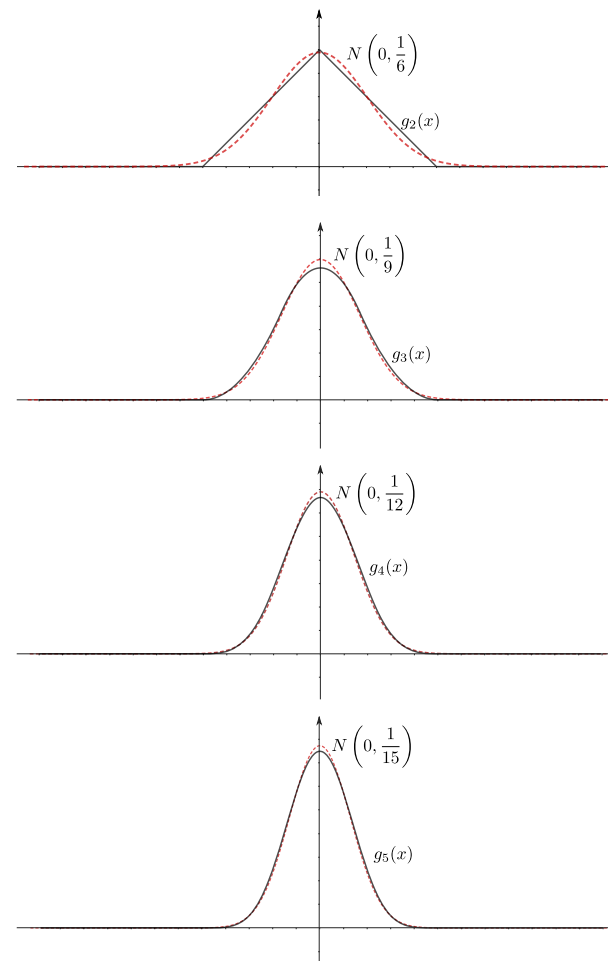
定理 8.2.2 $N(m, t)$ に従う母集団 X からとった大きさ n の標本平均 $\bar{X}$ は正規分布 $N(m, \frac{t}{n})$ に従います。

8.3 母集団が一様分布であった場合

$[-1, 1]$ 上の一様分布 (平均 0、分散 $\frac{1}{3}$) の密度関数 $f_1(x)$ を n 個たたみ込んだものを $f_n(x)$ とします。

$[-1, 1]$ 上の一様分布に従う母集団からとった大きさ n の標本平均 $\bar{X}_n$ の密度 $g_n(x)$ は $nf_n(nx)$ であり、これと正規分布 $N(0, \frac{1}{3n})$ の密度関数を比較すると以下の通りになっ

ています。 n が大きくなるにつれて正規分布に近づいてゆくのが分かるでしょうか？：



8.4 the central limit theorem

定理 8.4.1 [the central limit theorem] 平均 m 、分散 $v \neq 0$ である母集団 X からとった大きさ n の標本平均 $\bar{X}$ は n が十分大きい時 (50 以上) には正規分布 $N(m, \frac{v}{n})$ で近似されます。

8.5 中心極限定理による近似

問題 8.5.1 [教 例題 16.1 改] 正規分布 $N(169.2, 3.87^2)$ に従う母集団から大きさ 5 の標本をとるとき、標本平均 $\bar{X}$ について確率 $P[|\bar{X} - 169.2| \leq 4]$ を求めて下さい。

$\bar{X}$ は $N(169.2, \frac{3.87^2}{5})$ に従っています。

$$\begin{aligned} P[|\bar{X} - 169.2| \leq 4] &= P\left[\left|N\left(169.2, \frac{3.87^2}{5}\right) - 169.2\right| \leq 4\right] \\ &= P\left[\left|N(0, 1)\right| \leq \frac{4}{\frac{3.87}{\sqrt{5}}}\right] \\ &\approx 2P[0 \leq N(0, 1) \leq 2.31] \\ &\approx 2 \cdot 0.4896 \\ &= 0.9792 \end{aligned}$$

□

問題 8.5.2 [教 例題 16.2 改] 平均値が 19.7、分散が 0.8^2 である母集団からとった大きさ 50 の標本平均 $\bar{X}$ について確率 $P[19.4 \leq \bar{X} \leq 20.0]$ を求めて下さい。

中心極限定理によれば、この母集団からとった大きさ 50 の標本平均 $\bar{X}$ は正規分布 $N(19.7, \frac{0.8^2}{50})$ で近似されます。従って

$$\begin{aligned} P[19.4 \leq \bar{X} \leq 20.0] &= P\left[19.4 \leq N\left(19.7, \frac{0.8^2}{50}\right) \leq 20.0\right] \\ &= P\left[19.4 - 19.7 \leq N\left(19.7, \frac{0.8^2}{50}\right) - 19.7 \leq 20.0 - 19.7\right] \\ &= P\left[-0.3 \leq N\left(0, \frac{0.8^2}{50}\right) \leq 0.3\right] \\ &= 2P\left[0 \leq N\left(0, \frac{0.8^2}{50}\right) \leq 0.3\right] \\ &= 2P\left[0 \leq N(0, 1) \leq \frac{0.3}{\frac{0.8}{\sqrt{50}}}\right] \end{aligned}$$

$$\begin{aligned} &= 2P\left[0 \leq N(0, 1) \leq \frac{1.5\sqrt{2}}{0.8}\right] \\ &\approx 2P[0 \leq N(0, 1) \leq 2.651] \\ &\approx 2 \cdot 0.496 \\ &= 0.992 \end{aligned}$$

です。

□

8.6 問題演習

基本演習 8.1 [教科書 問題 16.4] $N(10, 2^2)$ に従う母集団からとった大きさ 25 の標本平均 $\bar{X}$ はどんな分布に従いますか。また次の確率は幾らになりますか。

$$(1) P[\bar{X} > 10.5] \quad (2) P[\bar{X} < 9.8] \quad (3) P[9.6 \leq \bar{X} \leq 10.4]$$

基本演習 8.2 [教科書 問題 16.5 改題] 平均 1800、標準偏差 100 の母集団からとった大きさ 100 の標本平均 $\bar{X}$ に対して次の確率を求めて下さい。

$$(1) P[\bar{X} > 1830] \quad (2) P[\bar{X} < 1790] \quad (3) P[1790 \leq \bar{X} \leq 1810]$$

基本演習 8.3 $N(7, 2.3^2)$ に従う母集団からとった大きさ 100 の標本平均 $\bar{X}$ に対して、 $P[7.35 \leq \bar{X}]$ を求めてください。

発展演習 8.4 標準正規分布に従う確率変数 X の moment generating function $E[e^{tX}]$ が $e^{\frac{1}{2}t^2}$ となることを示してください。

発展演習 8.5 標準正規分布に従う確率変数 X のモーメントを、定義に従った計算：

$$E[X^n] = \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

によって計算して下さい。まず奇数次のモーメントが 0 である事を示し、偶数次については漸化式を求めると良いでしょう。

発展演習 8.6 関数：

$$f(x) = \begin{cases} e^{-x} & 0 \leq x \\ 0 & \text{otherwise} \end{cases}$$

を密度関数とする確率変数 X のモーメントを計算して下さい。

発展演習 8.7 区間 $[-1, 1]$ 上の一様分布のモーメントを求めて下さい。

発展演習 8.8 密度関数が $h(x) = \frac{1}{2}e^{-|x|}$ であるような確率変数 X に対して moment generating function を求めて下さい。