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。

母集団としては現実に目の前にある数値の集まりだけでなく、想像上の話として考える事が出来るようなデータも扱います。例えばある工場で生産される全ての製品の重量など、いちいち全て量るわけにはいきませんが、仮定の話としてそう云うデータを空想する事は出来ます。このような場合、母集団は未知のデータと言えるでしょう。

そして未知の母集団の様子を窺うために、この母集団から n 個のサンプルを無作為に取り出して調査する事を考えます。具体的には n 回復元抽出をするわけですが、どんな n 個の数値の組が出て来るかは復元抽出をする度にランダムに変わるわけで、その全体を考えればこれは n 個の（あるいは n 次元の）確率変数であると考えられます。そうして得られた n 個の確率変数は独立であり、1 つ 1 つは同じくじ引きの結果ですから全て母集団と同じ分布に従います。この様に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。

この講義では、一貫して復元抽出の結果の総体を表す確率変数を『標本』と呼び、実際の抽出によって得られる具体的な数値の事を『サンプル』と呼んで区別する事にします。具体的なサンプルは標本と云う『さいころ』の 1 つの『出目』です。

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

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

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。

確率変数である標本平均の 1 つの実現値は、具体的な n 個のサンプルの平均値になっています。逆に言えば、実際に n 個の数値をサンプルとして取り出した時にどんな数値が出るかはランダムでしたからその平均値もランダムであって、その可能性の総体を確率変数として考えたものがこの標本平均であると言えます。

母集団が平均 m・分散 v をもつ場合、多次元確率変数についての知識から

$$E\left[\bar{X}\right] = m, \quad Var\left[\bar{X}\right] = \frac{v}{n}$$

である事を既に知っています。従って標本の大きさが非常に大きければ $\bar{X}$ の分散は非常に小さい事になり、平均値である母平均のまわりに密集している事が分かります。つまり、非常に大きなサンプルをとれば、その平均値は母平均からそう離れた値にはならないだろうと云う事になります。

8.2　母集団が正規分布であった場合

正規分布 $N(0,t)$ の密度を $N_t(x)$ として $N_1 * N_1$ を計算してみると、

$$\begin{aligned}(N_1 * N_1)(x) =& \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-y)^2}{2}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}} dy \\ =& \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\left(y^2 - xy + \frac{x^2}{2}\right)} dy \\ =& \frac{1}{2\pi} \mathrm{e}^{-\frac{x^2}{4}} \int_{-\infty}^{\infty} \mathrm{e}^{-\left(y - \frac{x}{2}\right)^2} dy\end{aligned}$$

となります。更に $y - \frac{x}{2} = z$ と云う風に変数変換すれば、Gauß積分が出て来て

$$(N_1 * N_1)(x) = \frac{1}{2\pi} \mathrm{e}^{-\frac{x^2}{4}} \int_{-\infty}^{\infty} \mathrm{e}^{-z^2} dz = \frac{1}{\sqrt{2\pi \cdot 2}} \mathrm{e}^{-\frac{x^2}{2 \cdot 2}} = N_2(x)$$

が得られますから、丁度分散が $1+1=2$ の正規分布になっている事が分かります。

同様に計算すれば、$N_s * N_t = N_{s+t}$ となる事が分かります。

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 $N\left(m+n, s+t\right)$ に従います。

この結果から、平均が 0 でない n 個の独立な正規分布 $N(m,t)$ の和は平均 nm、分散 nt の正規分布に従う事が分かり、さらに $\frac{1}{n}$ 倍すれば $\frac{X_1 + \cdots + X_n}{n}$ は正規分布 $N(m,t)$ に従う母集団からとった大きさ n の標本平均であり、次が分かります：

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 $N\left(m,\frac{t}{n}\right)$ に従います。

8.3　母集団が一様分布であった場合

$[-1,1]$ 上の一様分布の密度関数 $f_1(x)$ は

$$f_1(x)=\begin{cases}1 & -1\le x\le 1\\ 0 & \text{otherwise}\end{cases}$$

であり、平均は 0、分散は $\frac{1}{3}$ です。

ここでたたみ込み $f_1*f_1=f_2$ を計算すると

$$f_2(x)=\begin{cases}0 & 2\le x\\ \frac{1}{4}(2-x) & 0\le x\le 2\\ \frac{1}{4}(2+x) & -2\le x\le 0\\ 0 & x\le -2\end{cases}$$

となります。さらに計算すると

$$f_3(x)=(f_1*f_1*f_1)(x)=\begin{cases}0 & 3\le x\\ \frac{1}{16}(3-x)^2 & 1\le x\le 3\\ \frac{1}{8}(3-x^2) & -1\le x\le 1\\ \frac{1}{16}(3+x)^2 & -3\le x\le -1\\ 0 & x\le -3\end{cases}$$

$$f_4(x)=(f_1*f_1*f_1*f_1)(x)=\begin{cases}0 & 4\le x\\ \frac{1}{96}(4-x)^3 & 2\le x\le 4\\ \frac{1}{96}(32-12x^2+3x^3) & 0\le x\le 2\\ \frac{1}{96}(32-12x^2-3x^3) & -2\le x\le 0\\ \frac{1}{96}(4+x)^3 & -4\le x\le -2\\ 0 & x\le -4\end{cases}$$

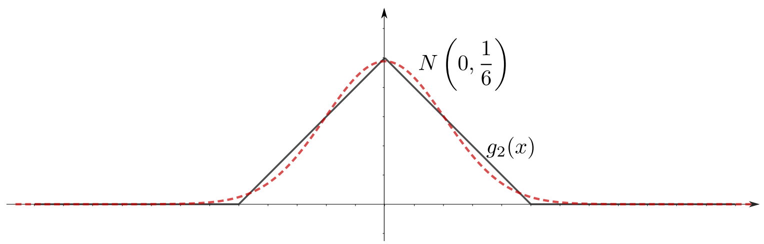
$$f_5(x)=\underbrace{(f_1*\cdots*f_1)}_{5\text{ コ}}(x)=\begin{cases}0 & 5\le x\\ \frac{1}{768}(5-x)^4 & 3\le x\le 5\\ \frac{1}{192}(55+10x-30x^2+10x^3-x^4) & 1\le x\le 3\\ \frac{1}{768}(230-60x^2+6x^2) & -1\le x\le 1\\ \frac{1}{192}(55-10x-30x^2-10x^3-x^4) & -3\le x\le -1\\ \frac{1}{768}(5+x)^4 & -5\le x\le -3\\ 0 & x\le -5\end{cases}$$

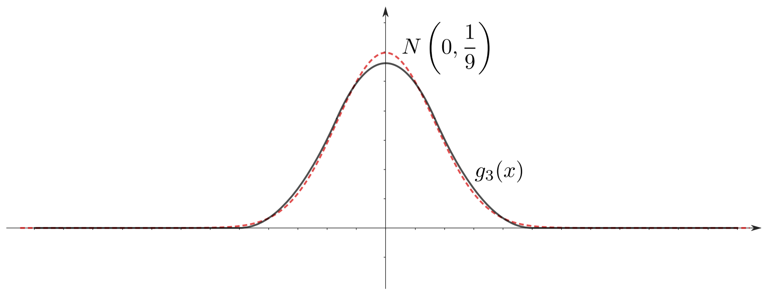
などとなっています。

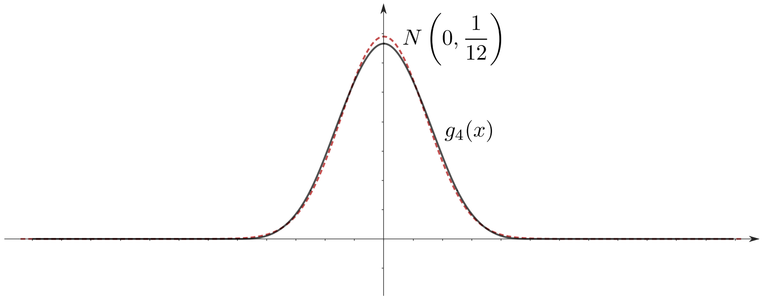
一般に S_n の密度が $f_n(x)$ であるとき、$\frac{S_n}{n}$ の密度は

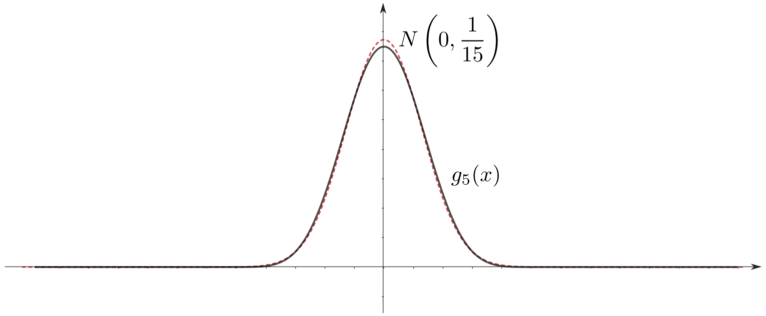
$$P\left[\frac{S_n}{n}\le t\right]=P[S_n\le nt]=\int_{-\infty}^{nt}f_n(x)dx=\int_{-\infty}^{t}f_n(nz)n\,dz$$

となりますから、$[-1,1]$ 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 $N\left(0,\frac{1}{3n}\right)$ 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 モーメントとその母関数

確率変数 X の分散は X^2 の平均値に関連していましたが、一般に 3 次以上の X^n の平均値も扱う場合があり、これらは存在するならばモーメント（積率）と呼ばれます。

例えば区間 $(-1, 1)$ 上の一様分布の平均値は 0、分散は $\frac{1}{3}$ ですから、1・2 次のモーメントは正規分布 $N\left(0, \frac{1}{3}\right)$ と同じですが、これらは全く別の分布をしています。しかし、これが 3 次、4 次と続いた場合、つまり、任意の次数のモーメントが等しいような場合には実は 2 つの確率変数は同じ分布に従うことが知られています。

指数関数 e^x の Taylor 展開を使えば

$$E[\mathrm{e}^{tX}] = E\left[1 + tX + \frac{1}{2!}t^2X^2 + \cdots\right] = 1 + tE[X] + \frac{1}{2!}t^2E[X^2] + \cdots \tag{8.1}$$

ですから、展開式の係数が丁度モーメントになっている（階乗は除く）事が分かります。ただしモーメントそのものや期待値 $E[\mathrm{e}^{tX}]$ が存在しない事もあります。

この $E\left[\mathrm{e}^{tX}\right]$ の事を（存在する場合に限る）X の積率母関数（moment generating function）と呼び、この講義では多くの場合記号 $M_X(t)$ で表します。

2 つの確率変数があって moment generating function が等しければ全てのモーメントも等しい事が分かりますから、その場合は確率変数同士が同じ分布に従う事が言えます。

例えば標準正規分布に従う確率変数 X の moment generating function は、簡単な計算により $M_X(t) = \mathrm{e}^{\frac{1}{2}t^2}$ となる事が分かります。

独立な確率変数 X, Y の和の moment generating function を計算してみると、e^{tX} と e^{tY} も独立ですから

$$M_{X+Y}(t) = E\left[\mathrm{e}^{t(X+Y)}\right] = E\left[\mathrm{e}^{tX}\mathrm{e}^{tY}\right] = E\left[\mathrm{e}^{tX}\right]E\left[\mathrm{e}^{tY}\right] = M_X(t)M_Y(t)$$

となってそれぞれの moment generating function の積になります。

また、W の moment generating function が $M_W(t)$ であるとき、その定数倍 aW の moment generating function$M_{aW}(t)$ は

$$M_{aW}(t) = E[\mathrm{e}^{t(aW)}] = E[\mathrm{e}^{(at)W}] = M_W(at)$$

となる事も注意しておきます。

8.5 the central limit theorem

平均 m・分散 $v > 0$ である母集団 X からとった大きさ n の標本平均 $\bar{X} = \frac{X_1+\cdots+X_n}{n}$ の分布はどうなっているでしょうか？　標準化した Z_n について調べてみましょう。

$$\bar{X} - m = \frac{X_1 + \cdots + X_n - nm}{n} = \frac{(X_1 - m) + \cdots + (X_n - m)}{n}$$

$$Z_n = \frac{\bar{X} - m}{\sqrt{\frac{v}{n}}} = \frac{(X_1 - m) + \cdots + (X_n - m)}{\sqrt{nv}} = \frac{\frac{X_1 - m}{\sqrt{v}} + \cdots + \frac{X_n - m}{\sqrt{v}}}{\sqrt{n}}$$

ここで分子の $Y_j = \frac{X_j - m}{\sqrt{v}}$（$X_j$ の標準化）は平均 0、分散 1 ですから

$$M_{Y_j}(t) = 1 + \frac{1}{2}t^2 + \text{(3 次以上の項)}$$

とべき級数展開されます。従ってそれらの独立和である分子 $Y_1 + \cdots + Y_n$ の moment generating function $M(t)$ は

$$M(t) = \left\{1 + \frac{1}{2}t^2 + \text{(3 次以上の項)}\right\}^n$$

になりますから、先に見た定数倍と moment generating function の関係から、

$$M_{Z_n}(t) = \left\{ M\left(\frac{t}{\sqrt{n}}\right) \right\}^n = \left\{ 1 + \frac{1}{2n}t^2 + \left(\frac{t}{\sqrt{n}} \text{ の 3 次以上の項}\right) \right\}^n$$

と書ける事が分かります。

ここで定数項以外の部分を R と書くことにして対数をとれば、n が十分大きいとき n を分母に含む R の部分は $|R| < 1$ ですから $\log(1+R)$ の Taylor 展開式から

$$\begin{aligned} \log M_{Z_n}(t) &= n \log(1+R) \\ &= n\left(R - \frac{1}{2}R^2 + \frac{1}{3}R^3 - \cdots\right) \\ &= nR - \frac{1}{2}nR^2 + \frac{1}{3}nR^3 - \cdots \end{aligned}$$

となります。するとまず

$$nR = n\left\{ \frac{1}{2n}t^2 + \left(\frac{t}{\sqrt{n}} \text{ の 3 次以上の項}\right) \right\} = \frac{1}{2}t^2 + \left(\frac{1}{\sqrt{n}} \text{ の 1 次以上の項}\right)$$

ですからここで $n \to \infty$ の極限をとれば

$$\lim nR = \frac{1}{2}t^2$$

となります。また、

$$nR^2 = n\left\{ \frac{1}{2n}t^2 + \left(\frac{t}{\sqrt{n}} \text{ の 3 次以上の項}\right) \right\}^2 = \left(\frac{1}{n} \text{ の 1 次以上の項}\right)$$

なのでこちらは極限をとると 0 である事が分かります。nR^3 等、これ以上のべきは全て同様に 0 に収束しますから、結局のところ

$$\lim_{n\to\infty} \log M_{Z_n}(t) = \frac{1}{2}t^2$$

であり、従って

$$\lim_{n\to\infty} M_{Z_n}(t) = \mathrm{e}^{\frac{1}{2}t^2}$$

が得られることになります。これは標準正規分布の moment generating function に一致しており、従って Z_n は n が大きい時には標準正規分布で近似される事が分かります。

これは $\bar{X}$ を標準化したものでしたから、$\sqrt{\frac{v}{n}}$ 倍して m を加えれば元に戻って $\bar{X}$ は平均 m、分散 $\frac{v}{n}$ の正規分布で近似されます。

定理 8.5.1 [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 が十分大きい時には正規分布 $N\left(m, \frac{v}{n}\right)$ で近似されます。

これを今日初めに見た母集団が正規分布に従う場合と比べてみると、標本平均と云うものは標本数が十分大きければ、母集団がどんな分布であったとしても母集団が正規分布であった場合と（ほぼ）同じ分布をしていることが分かるはずです。統計学では正規分布が重要である、あるいは、至る所に正規分布が現れると言われる所以です。

多くの場合標本のサイズが 50 以上であればこの中心極限定理を適用して差し支えないとされています。

8.6 中心極限定理による近似

問題 8.6.1 [教科書　例題 16.1 改題] 正規分布 $N\left(169.2, 3.87^2\right)$ 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 $P\left[\left|\bar{X} - 169.2\right| \leq 4\right]$ を求めて下さい。

正規母集団からとった標本平均はまた正規分布に従います。

この場合 $\bar{X}$ は $N\left(169.2, \frac{3.87^2}{5}\right)$ に従っていますから、

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

ですから正規分布表から

$$P\left[\left|\bar{X}-169.2\right|\leq 4\right]\approx 2\cdot 0.4896=0.9792$$

が分かります。 □

問題 8.6.2 [教科書　例題 16.2 改題] 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 $P\left[19.4\leq\bar{X}\leq 20.0\right]$ 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 $N\left(19.7,\frac{0.8^2}{50}\right)$ で近似されます。従って

$$\begin{aligned}
P\left[19.4\leq\bar{X}\leq 20.0\right]&=P\left[19.4\leq N\left(19.7,\frac{0.8^2}{50}\right)\leq 20.0\right]\\
&=P\left[19.4-19.7\leq N\left(19.7,\frac{0.8^2}{50}\right)-19.7\leq 20.0-19.7\right]\\
&=P\left[-0.3\leq N\left(0,\frac{0.8^2}{50}\right)\leq 0.3\right]\\
&=2P\left[0\leq N\left(0,\frac{0.8^2}{50}\right)\leq 0.3\right]\\
&=2P\left[0\leq N(0,1)\leq\frac{0.3}{\frac{0.8}{\sqrt{50}}}\right]\\
&=2P\left[0\leq N(0,1)\leq\frac{1.5\sqrt{2}}{0.8}\right]\\
&\approx 2P\left[0\leq N(0,1)\leq 2.651\right]\\
&\approx 2\cdot 0.496\\
&=0.992
\end{aligned}$$

です。 □

8.7 問題演習

基本演習 8.1 [教科書　問題 16.4] $N\left(10,2^2\right)$ 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) $P\left[\bar{X}>10.5\right]$　(2) $P\left[\bar{X}<9.8\right]$　(3) $P\left[9.6\leq\bar{X}\leq 10.4\right]$

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

基本演習 8.3 $N\left(7,2.3^2\right)$ 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、$P\left[7.35\leq\bar{X}\right]$ 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 $E[\mathrm{e}^{tX}]$ が $\mathrm{e}^{\frac{1}{2}t^2}$ 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：

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

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：

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

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 $h(x)=\frac{1}{2}\mathrm{e}^{-|x|}$ 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。

演習問題解答例

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

【解答例】 $\bar{X}$ は正規分布 $N\left(10, \frac{2^2}{25}\right)$ に従います。

(1)

$$
\begin{aligned}
P\left[\bar{X} > 10.5\right] &= P\left[N\left(10, \frac{2^2}{25}\right) > 10.5\right] \\
&= P\left[N\left(10, \frac{2^2}{25}\right) - 10 > 10.5 - 10\right] \\
&= P\left[N\left(0, \frac{2^2}{25}\right) > 0.5\right] \\
&= P\left[\frac{N\left(0, \frac{2^2}{25}\right)}{\frac{2}{5}} > \frac{0.5}{\frac{2}{5}}\right] \\
&= P\left[N\left(0, 1\right) > 1.25\right] \\
&= 0.5 - P[0 \leq N(0,1) \leq 1.25] \\
&\approx 0.5 - 0.3944 \\
&\approx 0.1056
\end{aligned}
$$

(2)

$$
\begin{aligned}
P\left[\bar{X} < 9.8\right] &= P\left[N\left(10, \frac{2^2}{25}\right) < 9.8\right] \\
&= P\left[\frac{N\left(0, \frac{2^2}{25}\right)}{\frac{2}{5}} < -\frac{0.2}{\frac{2}{5}}\right] \\
&= P\left[N\left(0, 1\right) < -0.5\right] \\
&= 0.5 - P[0 \leq N(0,1) \leq 0.5] \\
&\approx 0.5 - 0.1915 \\
&\approx 0.3085
\end{aligned}
$$

(3)

$$
\begin{aligned}
P\left[9.6 < \bar{X} < 10.4\right] &= P\left[9.6 < N\left(10, \frac{2^2}{25}\right) < 10.4\right] \\
&= P\left[-\frac{0.4}{\frac{2}{5}} < \frac{N\left(0, \frac{2^2}{25}\right)}{\frac{2}{5}} < \frac{0.4}{\frac{2}{5}}\right] \\
&= 2P[0 \leq N(0,1) \leq 1] \\
&\approx 2 \cdot 0.3413 \\
&\approx 0.6826
\end{aligned}
$$

□

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

【解答例】 標本のサイズが大きいので中心極限定理により $\bar{X}$ は $N\left(1800, \frac{100^2}{100}\right)$ で近似されます。

(1)

$$
\begin{aligned}
P\left[\bar{X} > 1830\right] &\approx P[N(1800, 100) > 1830] \\
&= P[N(0, 100) > 30] \\
&= P[N(0,1) > 3] \\
&= 0.5 - P[0 \leq N(0,1) < 3] \\
&\approx 0.5 - 0.4987 \\
&\approx 0.0013
\end{aligned}
$$

(2)

$$
\begin{aligned}
P\left[\bar{X} < 1790\right] &\approx P[N(1800, 100) < 1790] \\
&= P[N(0,1) < -1] \\
&= 0.5 - P[0 \leq N(0,1) < 1] \\
&\approx 0.5 - 0.3413 \\
&\approx 0.1587
\end{aligned}
$$

(3)

$$\begin{aligned} P\left[1790 \leq \bar{X} \leq 1810\right] &\approx P[1790 \leq N(1800, 100) \leq 1810] \\ &= P[-1 \leq N(0,1) \leq 1] \\ &= 2P[0 \leq N(0,1) \leq 1] \\ &\approx 2 \cdot 0.3413 \\ &\approx 0.6826 \end{aligned}$$

□

基本演習 8.3 $N\left(7, 2.3^2\right)$ 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、$P\left[7.35 \leq \bar{X}\right]$ を求めてください。

$\bar{X}$ は $N\left(7, \frac{2.3^2}{100}\right)$ に従いますから、

$$\begin{aligned} P\left[7.35 \leq \bar{X}\right] &= P\left[7.35 \leq N\left(7, \frac{2.3^2}{100}\right)\right] \\ &= P\left[0.35 \leq N\left(0, \frac{2.3^2}{100}\right)\right] \\ &= P\left[\frac{0.35}{\frac{2.3}{10}} \leq N(0,1)\right] \\ &= P[1.52 \leq N(0,1)] \\ &= 0.5 - P[0 \leq N(0,1) \leq 1.52] \\ &\approx 0.5 - 0.4357 \\ &= 0.0643 \end{aligned}$$

です。 □

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

$$\begin{aligned} M_X(t) &= E[\mathrm{e}^{tX}] \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{tx} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2 + tx} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\{(x-t)^2 - t^2\}} dx \\ &= \mathrm{e}^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}(x-t)^2} dx \\ &= \mathrm{e}^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}y^2} dy \\ &= \mathrm{e}^{\frac{1}{2}t^2} \end{aligned}$$

□

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：

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

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。

【解答例】

$$E[X^{2n+1}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} x^{2n+1} \mathrm{e}^{-\frac{x^2}{2}} dx = \lim_{\substack{L \to -\infty \\ M \to \infty}} \int_L^M \frac{1}{\sqrt{2\pi}} x^{2n+1} \mathrm{e}^{-\frac{x^2}{2}} dx$$

ですが、

$$\begin{aligned} \int_L^0 \frac{1}{\sqrt{2\pi}} x^{2n+1} \mathrm{e}^{-\frac{x^2}{2}} dx &= \int_{-L}^0 \frac{1}{\sqrt{2\pi}} (-y)^{2n+1} \mathrm{e}^{-\frac{(-y)^2}{2}} (-1) dy \\ &= -\int_0^{-L} \frac{1}{\sqrt{2\pi}} y^{2n+1} \mathrm{e}^{-\frac{y^2}{2}} y dy \end{aligned}$$

となるので、この積分が $L \to -\infty$ のとき $\lim_{M\to\infty} \int_0^M$ と同じ有限値に収束する事から明らかに奇数次のモーメントは 0 です。

偶数次では

$$\begin{aligned}
E[X^{2n}] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} x^{2n} \mathrm{e}^{-\frac{x^2}{2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} (-x^{2n-1})(-x) \mathrm{e}^{-\frac{x^2}{2}} dx \\
&= \left[\frac{1}{\sqrt{2\pi}} (-x^{2n-1}) \mathrm{e}^{-\frac{x^2}{2}} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} -(2n-1) x^{2n-2} \mathrm{e}^{-\frac{x^2}{2}} dx \\
&= (2n-1) E[X^{2(n-1)}]
\end{aligned}$$

となりますから、

$$E[X^{2n}] = (2n-1)(2n-3)\cdots 1 = (2n-1)!!$$

が分かります。 □

発展演習 8.6 関数：

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

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。

【解答例】部分積分を繰り返せば

$$E[X^n] = \int_0^{\infty} x^n \mathrm{e}^{-x} dx = n!$$

です。 □

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。

【解答例】密度関数は偶関数であり、モーメントの計算の積分は原点対象な有界区間での積分になりますから、奇関数の性質によって奇数次のモーメントは明らかに 0 です。偶数次の場合も簡単に積分されて結局 n 次のモーメントは、

$$\begin{cases} \frac{1}{n+1} & n : \text{even} \\ 0 & n : \text{odd.} \end{cases}$$

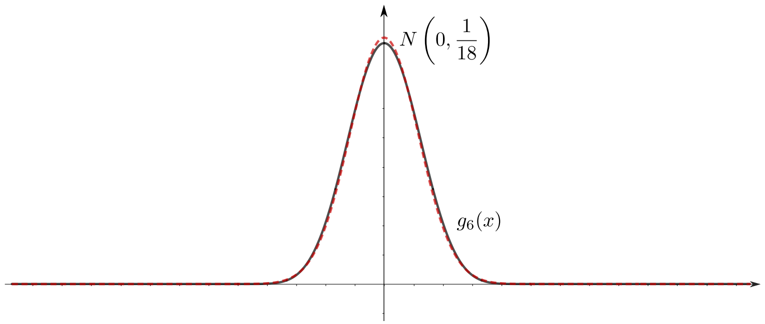
です。 □

発展演習 8.8 密度関数が $h(x) = \frac{1}{2}\mathrm{e}^{-|x|}$ 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。

【解答例】4 番の問題と同様に計算して（奇数次のモーメントは 0 になる事に注意）モーメントが求まりますから母関数を級数で書けばそれは $\dfrac{1}{1-t^2}$ である事が分かる筈です。 □

一様分布の畳み込みと正規分布

$$f_6(x) = \begin{cases} 0 & 6 \leq x \\ \frac{1}{7680}(6-x)^5 & 4 \leq x \leq 6 \\ \frac{1}{7680}(5x^5 - 90x^4 + 600x^3 - 1680x^2 + 1200x + 1632) & 2 \leq x \leq 4 \\ \frac{1}{7680}(-10x^5 + 60x^4 - 480x^2 + 2112) & 0 \leq x \leq 2 \\ \frac{1}{7680}(10x^5 + 60x^4 - 480x^2 + 2112) & -2 \leq x \leq 0 \\ \frac{1}{7680}(-5x^5 - 90x^4 - 600x^3 - 1680x^2 - 1200x + 1632) & -4 \leq x \leq -2 \\ \frac{1}{7680}(6+x)^5 & -6 \leq x \leq -4 \\ 0 & x \leq -6 \end{cases}$$



$$f_7(x) = \begin{cases} 0 & 7 \leq x \\ \frac{1}{92160}(7-x)^6 & 5 \leq x \leq 7 \\ \frac{1}{92160}(-6x^6 + 168x^5 - 1890x^4 + 10640x^3 - 29610x^2 + 30408x + 8274) & 3 \leq x \leq 5 \\ \frac{1}{92160}(15x^6 - 210x^5 + 945x^4 - 700x^3 - 4095x^2 - 210x + 23583) & 1 \leq x \leq 3 \\ \frac{1}{92160}(-20x^6 + 420x^4 - 4620x^2 + 23548) & -1 \leq x \leq 1 \\ \frac{1}{92160}(15x^6 + 210x^5 + 945x^4 + 700x^3 - 4095x^2 + 210x + 23583) & -3 \leq x \leq -1 \\ \frac{1}{92160}(-6x^6 - 168x^5 - 1890x^4 - 10640x^3 - 29610x^2 - 30408x + 8274) & -5 \leq x \leq -3 \\ \frac{1}{92160}(7+x)^6 & -7 \leq x \leq -5 \\ 0 & x \leq -7 \end{cases}$$