

2.1 確率変数

例えば、均等に作られたさいころを投げたときに出る目の数を表す変数を X とおくと、変数 X がとり得る値は $1, 2, 3, 4, 5, 6$ のいずれかであり、 $X = 1$ となる確率は $\frac{1}{6}$ 、 $X = 2$ となる確率は $\frac{1}{6}$ 、 $X = 3$ となる確率は $\frac{1}{6}$ 、 $X = 4$ となる確率は $\frac{1}{6}$ 、 $X = 5$ となる確率は $\frac{1}{6}$ 、 $X = 6$ となる確率は $\frac{1}{6}$ 、というように、変数 X の各々の値について X がその値をとる確率が決まっています。このような変数を **確率変数 (random variable)** といいます。数学的には次のように定義します。

定義 実数を表す変数 X が確率変数であるとは、任意の実数 a に対して $X \leq a$ となる確率が定まることである。

証明は省略しますが次の定理が成り立ちます。

定理 2.1.1 確率変数 X 及び任意の実数 a, b に対して、 $X \geq a$ となる確率、 $X < a$ となる確率、 $X > a$ となる確率、 $a \leq X \leq b$ となる確率、などの確率が定まる。

確率変数 X 及び実数 a, b に対して、
 $X \leq a$ となる確率を $P(X \leq a)$ と、
 $X \geq a$ となる確率を $P(X \geq a)$ と、
 $a \leq X \leq b$ となる確率を $P(a \leq X \leq b)$ と、
 $X = a$ となる確率を $P(X = a)$ と、
 いうように書き表します。

例 均等に作られたさいころを投げて出る目の数を表す変数を X とおきます。この変数 X の値は $1, 2, 3, 4, 5, 6$ のどれかであり、
 $P(X=1) = P(X=2) = P(X=3) = P(X=4) = P(X=5) = P(X=6) = \frac{1}{6}$ 。
 例えば、実数 $\frac{8}{3}$ に対して、
 $X \leq \frac{8}{3} \iff X = 1 \text{ または } X = 2$ 、
 よって
 $P(X \leq \frac{8}{3}) = P(X=1) + P(X=2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ 。
 例えば、実数 $\sqrt{11}$ に対して、
 $X \leq \sqrt{11} \iff X = 1 \text{ または } X = 2 \text{ または } X = 3$ 、
 よって
 $P(X \leq \sqrt{11}) = P(X=1) + P(X=2) + P(X=3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$ 。
 例えば、実数 6.2 に対して、
 $P(X \leq 6.2) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$ 。
 例えば、実数 -2 に対して、 $X \leq -2$ となることはないので、
 $P(X \leq -2) = 0$ 。

このように、任意の実数 a について次のように確率 $P(X \leq a)$ が決まります⁶⁾：

$$P(X \leq a) = \begin{cases} 0 & (a < 1 \text{ のとき}) \\ \frac{[a]}{6} & (1 \leq a < 6 \text{ のとき}) \\ 1 & (a \geq 6 \text{ のとき}) \end{cases}$$

従って変数 X は確率変数です。 [終]

確率変数の値を実際に得るためには試行が必要です。例えば、あるさいころを投げて出る目の数を表す確率変数を考えるとき、このさいころを投げて出る目の数を調べるのが試行です。例えば、あるものを何回も測定するときの測定値の分布が決まっていると考えると、測定量を表す変数は確率変数であり、測定が試行です。試行によって実際に得られた確率変数の値を**実現値 (sample value)**といいます⁷⁾。例えば、あるさいころを投げて出る目の数を表す確率変数の実現値とは、実際にそのさいころを投げて出た目の数のことです。例えば、あるものの測定量を表す確率変数の実現値とは、実際に測定を行って得られた測定値のことです。確率変数の実現値は試行を行うごとに得られます。同じ試行でも1回めと2回めとは同じ確率変数の実現値が異なったりします。

変数 X が確率変数であるとき、任意の実数 x に対して確率 $P(X \leq x)$ が決まります。よって各実数 x に確率 $P(X \leq x)$ を対応させる関数ができます。この関数を確率変数 X の**累積分布関数 (cumulative distribution function)** といいます。つまり、確率変数 X の累積分布関数とは次のような関数 $F(x)$ のことです：任意の実数 x に対して

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

多くの確率変数は以下に述べる離散型か連続型かのどちらかに分類されます。

確率変数 X が取り得る値の総てに番号が付けられるとき、 X を**離散型確率変数 (discrete random variable)** といいます。つまり、離散型確率変数 X の取り得る値は1番の値 x_1 、2番の値 x_2 、3番の値 x_3 、4番の値 x_4 、 \dots のどれかでなければなりません。ですから、番号 i と j に対して $i \neq j$ のとき $x_i \neq x_j$ であるとき、確率 $P(X=x_1)$ 、 $P(X=x_2)$ 、 $P(X=x_3)$ 、 $P(X=x_4)$ 、 \dots の総和は1でなければなりません。

正確には、確率変数 X が離散型確率変数であるとは以下の2条件のどちらかを満たすことです：

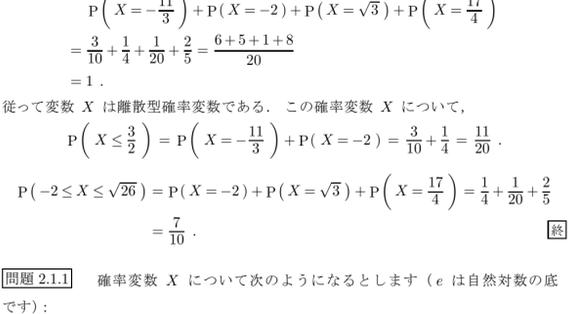
(1) X が取り得る値の総てを互いに異なる実数より成る有限数列 $\{x_n\}_{n \leq N}$ (N は自然数) で枚挙でき、

$$\sum_{n=0}^N P(X=x_n) = P(X=x_0) + P(X=x_1) + P(X=x_2) + \dots + P(X=x_{N-1}) + P(X=x_N) = 1$$

(2) X が取り得る値の総てを互いに異なる実数より成る無限数列 $\{x_n\}_{n \geq 0}$ で枚挙でき、

$$\sum_{n=0}^{\infty} P(X=x_n) = P(X=x_0) + P(X=x_1) + P(X=x_2) + \dots + P(X=x_n) + P(X=x_{n+1}) + \dots = 1$$

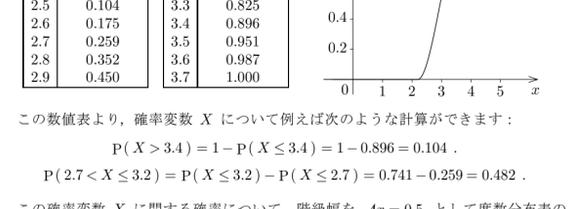
例 確率変数 X について次のようになります：
 $P(X = -\frac{7}{3}) = \frac{1}{4}$ 、 $P(X = -\sqrt{2}) = \frac{1}{6}$ 、
 $P(X = \frac{11}{5}) = \frac{1}{4}$ 、 $P(X = \pi) = \frac{1}{3}$ ；
 $\frac{7}{3}, -\sqrt{2}, \frac{11}{5}, \pi$ 以外の実数 x に対して $P(X=x) = 0$ 。このとき、
 $P(X = -\frac{7}{3}) + P(X = -\sqrt{2}) + P(X = \frac{11}{5}) + P(X = \pi) = \frac{1}{4} + \frac{1}{6} + \frac{1}{4} + \frac{1}{3} = 1$ 。
 従ってこの確率変数 X は離散型確率変数です。この離散型確率変数 X に関する確率は次のように表とかグラフなどでも表せます。



例題 2.1.1 確率変数 X について次のようになります：
 $P(X = -\frac{11}{3}) = \frac{3}{10}$ 、 $P(X = -2) = \frac{1}{4}$ 、
 $P(X = \sqrt{3}) = \frac{1}{20}$ 、 $P(X = \frac{17}{4}) = \frac{2}{5}$ ；
 $-\frac{11}{3}, -2, \sqrt{3}, \frac{17}{4}$ 以外の実数 x に対して $P(X=x) = 0$ 。この確率変数 X が離散型確率変数であることを示し、確率 $P(X \leq \frac{3}{2})$ 、 $P(-2 \leq X \leq \sqrt{26})$ を求める。
 $-\frac{11}{3}, -2, \sqrt{3}, \frac{17}{4}$ 以外の実数 x に対して $P(X=x) = 0$ であり、
 $P(X = -\frac{11}{3}) + P(X = -2) + P(X = \sqrt{3}) + P(X = \frac{17}{4}) = \frac{3}{10} + \frac{1}{4} + \frac{1}{20} + \frac{2}{5} = \frac{6+5+1+8}{20} = 1$ 。
 従って変数 X は離散型確率変数である。この確率変数 X について、
 $P(X \leq \frac{3}{2}) = P(X = -\frac{11}{3}) + P(X = -2) = \frac{3}{10} + \frac{1}{4} = \frac{11}{20}$ 。
 $P(-2 \leq X \leq \sqrt{26}) = P(X = -2) + P(X = \sqrt{3}) + P(X = \frac{17}{4}) = \frac{1}{4} + \frac{1}{20} + \frac{2}{5} = \frac{7}{10}$ 。 [終]

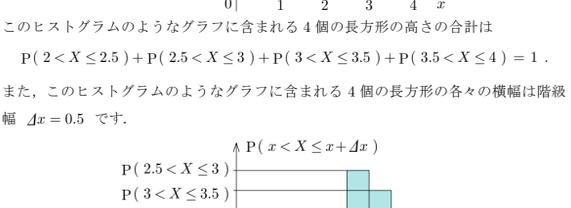
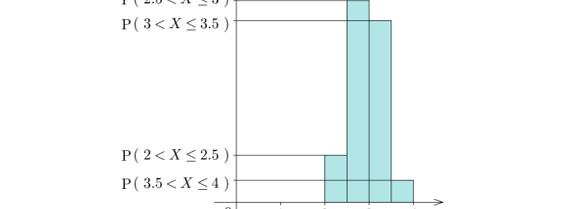
問題 2.1.1 確率変数 X について次のようになります (e は自然対数の底です)：
 $P(X = -\frac{9}{4}) = \frac{1}{5}$ 、 $P(X = -\sqrt{5}) = \frac{1}{6}$ 、
 $P(X = e) = \frac{1}{3}$ 、 $P(X = \frac{11}{2}) = \frac{3}{10}$ ；
 $-\frac{9}{4}, -\sqrt{5}, e, \frac{11}{2}$ 以外の実数 x に対して $P(X=x) = 0$ 。この確率変数 X が離散型確率変数であることを示し、確率 $P(X \leq \frac{11}{4})$ 、 $P(-2.2 \leq X \leq 5.5)$ を求めなさい。

例として、人が身長を測らずに目分量で紐を3mになるべく近い長さで切り出すとき、切り出した紐の実際の長さ(単位はm)を表す変数を X とおきます。仮に多数の様々な人達に目分量で3mになるべく近い長さで紐を切り出してもらって実際の長さを調べると、例えば紐の長さが3.5m以下になる割合が近似的に確率 $P(X \leq 3.5)$ になると考えられます。このように、任意の実数 a に対して確率 $P(X \leq a)$ が近似的に決まります。ですから、実験的に、各実数 x に対して確率 $P(X \leq x)$ を近似的に定めることができます。従って、そのような実験を行ったとして、変数 X は確率変数と考えることができます。この確率変数の取り得る値は3の近くに“べったりと”(ひと続きの実数の集合として)ありますから、離散型確率変数のように確率の値を表とかグラフの値で正確に記述することができません。しかし、実数 x に対する累積分布関数 $P(X \leq x)$ は決まります。この累積分布関数の数値表およびグラフは例えば次のようになります。

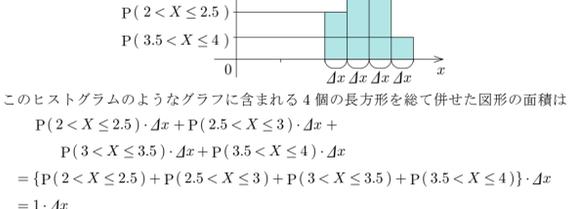


この数値表より、確率変数 X について例えば次のような計算ができます：
 $P(X > 3.4) = 1 - P(X \leq 3.4) = 1 - 0.896 = 0.104$ 。
 $P(2.7 < X \leq 3.2) = P(X \leq 3.2) - P(X \leq 2.7) = 0.741 - 0.259 = 0.482$ 。

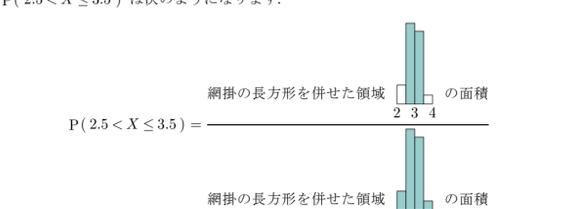
この確率変数 X に関する確率について、階級幅を $\Delta x = 0.5$ として度数分布表のようなものを考え、更にヒストグラムのようなグラフを描きます。



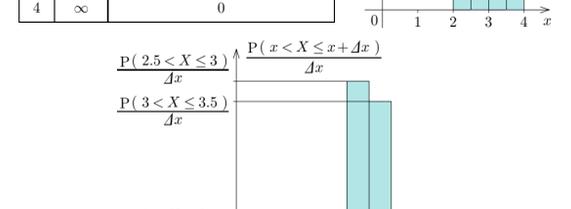
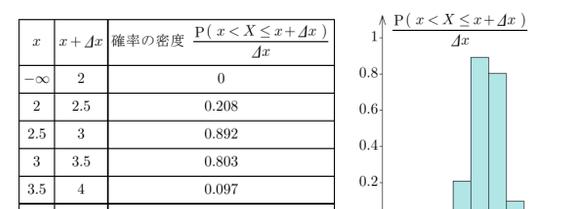
このヒストグラムのようなグラフに含まれる4個の長方形の高さの合計は
 $P(2 < X \leq 2.5) + P(2.5 < X \leq 3) + P(3 < X \leq 3.5) + P(3.5 < X \leq 4) = 1$ 。
 また、このヒストグラムのようなグラフに含まれる4個の長方形の各々の横幅は階級幅 $\Delta x = 0.5$ です。



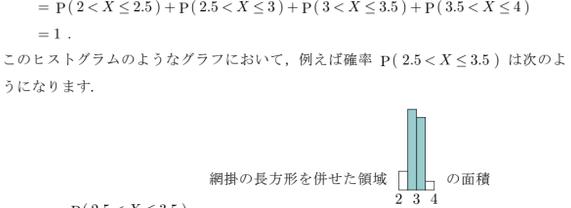
このヒストグラムのようなグラフに含まれる4個の長方形を総て併せた図形の面積は
 $P(2 < X \leq 2.5) \cdot \Delta x + P(2.5 < X \leq 3) \cdot \Delta x + P(3 < X \leq 3.5) \cdot \Delta x + P(3.5 < X \leq 4) \cdot \Delta x = \{P(2 < X \leq 2.5) + P(2.5 < X \leq 3) + P(3 < X \leq 3.5) + P(3.5 < X \leq 4)\} \cdot \Delta x = 1 \cdot \Delta x = \Delta x$ 。
 このヒストグラムのようなグラフに含まれる4個の長方形を総て併せた図形の面積は階級幅 Δx です。このヒストグラムのようなグラフにおいて、例えば確率 $P(2.5 < X \leq 3.5)$ は次のようになります。



確率 $P(x < X \leq x + \Delta x)$ を Δx で割った確率の“密度” $\frac{P(x < X \leq x + \Delta x)}{\Delta x}$ を考えます。

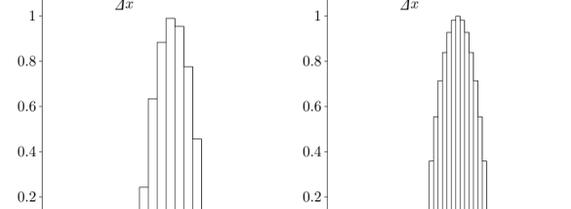


このヒストグラムのようなグラフに含まれる4個の長方形を総て併せた図形の面積は
 $\frac{P(2 < X \leq 2.5)}{\Delta x} \cdot \Delta x + \frac{P(2.5 < X \leq 3)}{\Delta x} \cdot \Delta x + \frac{P(3 < X \leq 3.5)}{\Delta x} \cdot \Delta x + \frac{P(3.5 < X \leq 4)}{\Delta x} \cdot \Delta x = P(2 < X \leq 2.5) + P(2.5 < X \leq 3) + P(3 < X \leq 3.5) + P(3.5 < X \leq 4) = 1$ 。
 このヒストグラムのようなグラフにおいて、例えば確率 $P(2.5 < X \leq 3.5)$ は次のようになります。

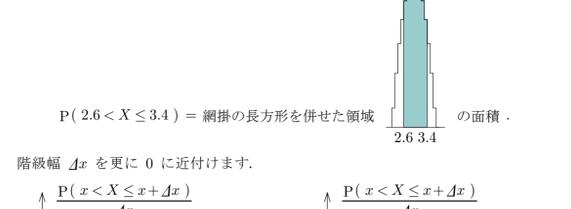


このように、確率変数 X についての確率の“密度” $\frac{P(x < X \leq x + \Delta x)}{\Delta x}$ のヒストグラムのようなグラフにおいて、 X についての確率はそのグラフの中の長方形を併せた領域の面積になります。

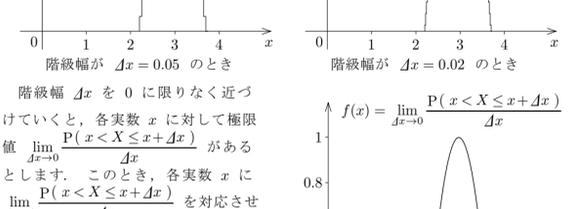
階級幅 $\Delta x = 0.2$ のとき及び $\Delta x = 0.05$ のとき、確率変数 X について確率の“密度” $\frac{P(x < X \leq x + \Delta x)}{\Delta x}$ のヒストグラムのようなグラフは次のようになります。



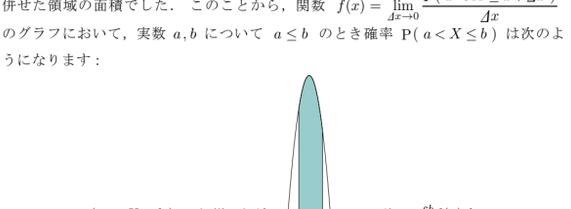
確率の“密度” $\frac{P(x < X \leq x + \Delta x)}{\Delta x}$ のヒストグラムのようなグラフにおいて、 X についての確率はそのグラフの中の長方形を併せた領域の面積ですから、例えば確率 $P(2.6 < X \leq 3.4)$ は次のようになります：



階級幅 Δx を更に0に近づけます。



階級幅 Δx を0に限りなく近づけていくと、各実数 x に対して極限值 $\lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x}$ があるとき、このとき、各実数 x に $\lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x}$ を対応させる関数ができます。この関数を $f(x)$ とおきます：各実数 x に対して $f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x}$ 。 $P(x < X \leq x + \Delta x)$ のヒストグラム



このように、関数 $f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x}$ を考えると、 X に関する確率 $P(a < X \leq b)$ は $f(x)$ の定積分 $\int_a^b f(x) dx$ になります。このことは次のようにも分かります。確率変数 X の累積分布関数を $F(x)$ とおきます：各実数 x に対して $F(x) = P(X \leq x)$ 。このとき、
 $P(a < X \leq x + \Delta x) = P(X \leq x + \Delta x) - P(X \leq x) = F(x + \Delta x) - F(x)$ 。
 従って、
 $f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x)$ 。
 このように、関数 $f(x)$ は関数 $F(x)$ の導関数です。つまり関数 $F(x)$ は関数 $f(x)$ の原始関数です。従って、関数 $f(x)$ が積分可能であれば、微分積分の基本定理より $\int_a^b f(x) dx = F(b) - F(a) = P(X \leq b) - P(X \leq a) = P(a < X \leq b)$ 。
 このように各実数 a, b ($a \leq b$) について $\int_a^b f(x) dx = P(a < X \leq b)$ となる関数 $f(x)$ を**確率密度関数 (probability density function)** といいます。正確には次のように定義します。

定義 確率変数 X の確率密度関数とは、実数全体を定義域とする関数 $f(x)$ が次の3条件を満たすものである：
 (1) 広義積分 $\int_{-\infty}^{\infty} f(x) dx$ の値があつて、 $\int_{-\infty}^{\infty} f(x) dx = 1$ ；
 (2) 任意の実数 x について $f(x) \geq 0$ ；
 (3) 任意の実数 a について $P(X \leq a) = \int_{-\infty}^a f(x) dx$ 。
 確率変数 X の確率密度関数が存在するとき、 X は**連続型確率変数 (continuous random variable)** であるといいます。連続型確率変数について次の定理が成り立ちます。

定理 2.1.2 関数 $f(x)$ は連続型確率変数 X の確率密度関数であるとする。任意の実数 a について、
 $P(X \leq a) = P(X < a) = \int_{-\infty}^a f(x) dx$ 。
 $P(X \geq a) = P(X > a) = \int_a^{\infty} f(x) dx$ 。
 任意の実数 a, b について、 $a \leq b$ のとき、
 $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = \int_a^b f(x) dx$ 。

例題 2.1.2 実数全体を定義域とする関数 $f(x)$ を次のように定める：

$$f(x) = \begin{cases} \frac{x(6-x)}{36} & (0 \leq x \leq 6 \text{ のとき}) \\ 0 & (x < 0 \text{ または } x > 6 \text{ のとき}) \end{cases}$$

 この関数 $f(x)$ が確率密度関数になることを示す。また、この関数 $f(x)$ を確率密度関数とする確率変数 X について、確率 $P(X \geq 4)$ 、 $P(-5 \leq X \leq 5)$ を求める。
 任意の実数 x について $f(x) \geq 0$ 。
 $\int_{-\infty}^{\infty} f(x) dx = \int_0^6 \frac{x(6-x)}{36} dx = \frac{1}{36} \int_0^6 (6x - x^2) dx = \frac{1}{36} [3x^2 - \frac{1}{3}x^3]_0^6 = \frac{1}{36} (3 \cdot 6^2 - \frac{1}{3} \cdot 6^3) = \frac{1}{36} (3 \cdot 36 - 2 \cdot 36) = \frac{1}{36} (36) = 1$ 。
 従って関数 $f(x)$ は確率変数の確率密度関数になる。
 この関数 $f(x)$ を確率密度関数とする確率変数 X について、
 $P(X \geq 4) = \int_4^6 f(x) dx = \int_4^6 \frac{x(6-x)}{36} dx = \frac{1}{36} \int_4^6 (6x - x^2) dx = \frac{1}{36} [3x^2 - \frac{1}{3}x^3]_4^6 = \frac{1}{36} (3 \cdot 6^2 - \frac{1}{3} \cdot 6^3 - (3 \cdot 4^2 - \frac{1}{3} \cdot 4^3)) = \frac{1}{36} (36 - 24 - (48 - \frac{64}{3})) = \frac{1}{36} (12 - \frac{104}{3}) = \frac{1}{36} (\frac{36 - 104}{3}) = \frac{1}{36} (-\frac{68}{3}) = \frac{17}{27}$ 。
 $P(-5 \leq X \leq 5) = \int_{-5}^5 f(x) dx = \int_0^5 \frac{x(6-x)}{36} dx = \frac{1}{36} \int_0^5 (6x - x^2) dx = \frac{1}{36} [3x^2 - \frac{1}{3}x^3]_0^5 = \frac{1}{36} (3 \cdot 5^2 - \frac{1}{3} \cdot 5^3) = \frac{1}{36} (75 - \frac{125}{3}) = \frac{1}{36} (\frac{225 - 125}{3}) = \frac{100}{108} = \frac{25}{27}$ 。 [終]

問題 2.1.2 実数全体を定義域とする関数 $f(x)$ を次のように定める：

$$f(x) = \begin{cases} \frac{3}{32}(2x+3-x^2) & (-1 \leq x \leq 3 \text{ のとき}) \\ 0 & (x < -1 \text{ または } x > 3 \text{ のとき}) \end{cases}$$

 この関数 $f(x)$ が確率密度関数になることを示さない。また、この関数 $f(x)$ を確率密度関数とする確率変数 X について、確率 $P(X \leq 0)$ 、 $P(1 \leq X \leq 7)$ を求めなさい。

⁶⁾ 実数 x に対して $[x]$ は x 以下の最大の整数を表します。
⁷⁾ 現実的には、確率変数の実現値はもの長さとかもの重さなどの量そのものを表す。例えば、確率変数 X を人の身長とすると、つまり“ $X = l$ の身長 (単位は cm)”とすると、不等式 $X \leq 170$ は“身長が170cm以下”ということの意味し、 $P(X \leq 170)$ は身長が170cm以下になる確率を表します；また、 X の実現値は誰かの実際の身長 (単位は cm) であり、 X の実現値 x_1 について $x_1 = 172.3$ となったり、 X の別の実現値 x_2 について $x_2 = 168.7$ となったりします。