

2.1 確率変数

例えば、均等に作られた一つのさいころを投げたときに出る目の数を表す変数を X とおくとき、変数 X がとり得る値は $1, 2, 3, 4, 5, 6$ のいずれかであり、 $X = 1$ である確率は $\frac{1}{6}$ 、 $X = 2$ である確率は $\frac{1}{6}$ 、 $X = 3$ である確率は $\frac{1}{6}$ 、 $X = 4$ である確率は $\frac{1}{6}$ 、 $X = 5$ である確率は $\frac{1}{6}$ 、 $X = 6$ である確率は $\frac{1}{6}$ 、である。変数 X がとり得る値の各々について X がその値をとる確率が決まる。このような確率が定まる変数を**確率変数**(random variable)という。

数学的には次のように定義する。

[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 $X \leq a$ である確率が定まることである。

証明は省くが次の定理が成り立つ。

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

確率変数 X 及び実数 a, b に対して、

$$\begin{aligned} X \leq a \text{ である確率を } & \mathrm{P}(X \leq a) \text{ と、} \\ X \geq a \text{ である確率を } & \mathrm{P}(X \geq a) \text{ と、} \\ a \leq X \leq b \text{ である確率を } & \mathrm{P}(a \leq X \leq b) \text{ と、} \\ X = a \text{ である確率を } & \mathrm{P}(X = a) \text{ と、} \end{aligned}$$

いうように書き表す。

[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 とおく。この変数 X の値は $1, 2, 3, 4, 5, 6$ のどれかであり、

$$\begin{aligned} \mathrm{P}(X = 1) &= \mathrm{P}(X = 2) = \mathrm{P}(X = 3) = \mathrm{P}(X = 4) = \mathrm{P}(X = 5) = \mathrm{P}(X = 6) \\ &= \frac{1}{6} . \end{aligned}$$

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

$$\mathrm{P}\left(X \leq \frac{8}{3}\right) = \mathrm{P}(X = 1) + \mathrm{P}(X = 2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} .$$

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

$$\mathrm{P}(X \leq \sqrt{11}) = \mathrm{P}(X = 1) + \mathrm{P}(X = 2) + \mathrm{P}(X = 3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} .$$

実数 6.2 に対して、

$$\begin{aligned} \mathrm{P}(X \leq 6.2) &= \mathrm{P}(X = 1) + \mathrm{P}(X = 2) + \mathrm{P}(X = 3) + \mathrm{P}(X = 4) + \mathrm{P}(X = 5) + \mathrm{P}(X = 6) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= 1 . \end{aligned}$$

例えば、実数 -2 に対して、 $X \leq -2$ であることはないので、

$$\mathrm{P}(X \leq -2) = 0 .$$

このように、任意の実数 a について次のように確率 $\mathrm{P}(X \leq a)$ が決まる：

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

従って変数 X は確率変数である。

確率変数の値を実際に得るためには試行が必要になる。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ことが試行である。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定まっているとき、測定量を表す変数は確率変数であり、測定が試行である。

試行によって実際に得られた確率変数の値を**実現値**(realized value)あるいは観測値(observed value)という⁵⁾。という。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ある。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ある。確率変数の実現値は試行を行うごとに得られる。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する。

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 $\mathrm{P}(X \leq x)$ が唯一つ決まる。よって各実数 x に確率 $\mathrm{P}(X \leq x)$ を対応させる関数ができる。この関数を確率変数 X の**(累積)分布関数**((cumulative) distribution function)という。つまり、確率変数 X の累積分布関数とは次のような関数 $F(x)$ のことである：任意の実数 x に対して

$$F(x) = \mathrm{P}(X \leq x) .$$

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る。

確率変数 X がと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)という。つまり、離散型確率変数 X のとり得る値は、0 番の値 x_0 、1 番の値 x_1 、2 番の値 x_2 、3 番の値 x_3 、4 番の値 x_4 、 $\cdots$ のどれかである。なので、番号 i と j について $i \neq j$ のとき $x_i \neq x_j$ であるとき、確率 $\mathrm{P}(X = x_0)$ 、 $\mathrm{P}(X = x_1)$ 、 $\mathrm{P}(X = x_2)$ 、 $\mathrm{P}(X = x_3)$ 、 $\mathrm{P}(X = x_4)$ 、 $\cdots$ の総和は 1 である。

[定義] 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ある：

(1) X がとり得る値の総てを互いに異なる実数より成る有限数列 $\{x_n\}_{0 \leq n \leq N}$ (N は自然数)で枚挙できて、

$$\begin{aligned} \sum_{n=0}^N \mathrm{P}(X = x_n) &= \mathrm{P}(X = x_0) + \mathrm{P}(X = x_1) + \mathrm{P}(X = x_2) + \cdots \\ &\quad + \mathrm{P}(X = x_{N-1}) + \mathrm{P}(X = x_N) \\ &= 1 . \end{aligned}$$

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

$$\begin{aligned} \sum_{n=0}^{\infty} \mathrm{P}(X = x_n) &= \mathrm{P}(X = x_0) + \mathrm{P}(X = x_1) + \mathrm{P}(X = x_2) + \cdots \\ &\quad + \mathrm{P}(X = x_n) + \mathrm{P}(X = x_{n+1}) + \cdots \\ &= 1 . \end{aligned}$$

[例] 確率変数 X がとり得る値は $-\frac{7}{3}$ と $-\sqrt{2}$ と $\frac{11}{5}$ と π とだけであり、

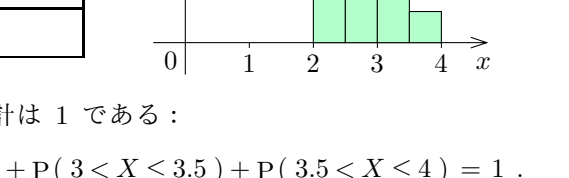
$$\begin{aligned} \mathrm{P}\left(X = -\frac{7}{3}\right) &= \frac{1}{4} , & \mathrm{P}(X = -\sqrt{2}) &= \frac{1}{6} , \\ \mathrm{P}\left(X = \frac{11}{5}\right) &= \frac{1}{4} , & \mathrm{P}(X = \pi) &= \frac{1}{3} \end{aligned}$$

とする。このとき、

$$\mathrm{P}\left(X = -\frac{7}{3}\right) + \mathrm{P}(X = -\sqrt{2}) + \mathrm{P}\left(X = \frac{11}{5}\right) + \mathrm{P}(X = \pi) = \frac{1}{4} + \frac{1}{6} + \frac{1}{4} + \frac{1}{3} = 1 .$$

従ってこの確率変数 X は離散型確率変数である。この離散型確率変数 X に関する確率は次のように表とがグラフなどでも表せる。

x の値	確率 $\mathrm{P}(X = x)$
$-\frac{7}{3}$	$\frac{1}{4}$
$-\sqrt{2}$	$\frac{1}{6}$
$\frac{11}{5}$	$\frac{1}{4}$
π	$\frac{1}{3}$
上にある値以外	0



[例] 確率変数 X がとり得る値は $-\frac{11}{3}$ と -2 と $\sqrt{3}$ と $\frac{17}{4}$ とだけであり、

$$\begin{aligned} \mathrm{P}\left(X = -\frac{11}{3}\right) &= \frac{3}{10} , & \mathrm{P}(X = -2) &= \frac{1}{4} , \\ \mathrm{P}(X = \sqrt{3}) &= \frac{1}{20} , & \mathrm{P}\left(X = \frac{17}{4}\right) &= \frac{2}{5} \end{aligned}$$

とする。この確率変数 X が離散型確率変数であることを示す。

$$\begin{aligned} \mathrm{P}\left(X = -\frac{11}{3}\right) + \mathrm{P}(X = -2) + \mathrm{P}(X = \sqrt{3}) + \mathrm{P}\left(X = \frac{17}{4}\right) \\ = \frac{3}{10} + \frac{1}{4} + \frac{1}{20} + \frac{2}{5} = \frac{6+5+1+8}{20} = 1 . \end{aligned}$$

従って変数 X は離散型確率変数である。 X に関する確率 $\mathrm{P}\left(X \leq \frac{3}{2}\right)$ を求める。

$$\mathrm{P}\left(X \leq \frac{3}{2}\right) = \mathrm{P}\left(X = -\frac{11}{3}\right) + \mathrm{P}(X = -2) = \frac{3}{10} + \frac{1}{4} = \frac{11}{20} .$$

X に関する確率 $\mathrm{P}(-2 \leq X \leq \sqrt{26})$ を求める。

$$\begin{aligned} \mathrm{P}(-2 \leq X \leq \sqrt{26}) &= \mathrm{P}(X = -2) + \mathrm{P}(X = \sqrt{3}) + \mathrm{P}\left(X = \frac{17}{4}\right) = \frac{1}{4} + \frac{1}{20} + \frac{2}{5} \\ &= \frac{7}{10} . \end{aligned}$$

問 2.1.1 確率変数 X がとり得る値は $-\frac{9}{4}$ と $-\sqrt{5}$ と e と $\frac{11}{2}$ とだけであり、

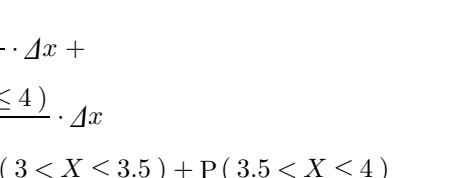
$$\begin{aligned} \mathrm{P}\left(X = -\frac{9}{4}\right) &= \frac{1}{5} , & \mathrm{P}(X = -\sqrt{5}) &= \frac{1}{6} , \\ \mathrm{P}(X = e) &= \frac{1}{3} , & \mathrm{P}\left(X = \frac{11}{2}\right) &= \frac{3}{10} ; \end{aligned}$$

とする。(e は自然対数の底です)。この確率変数 X が離散型確率変数であることを示し、 X に関する確率 $\mathrm{P}\left(X \leq \frac{11}{4}\right)$ 、 $\mathrm{P}(-2.2 \leq X \leq 5.5)$ を求めよ。

連続型確率変数を説明する。

[例] 人が長さを測らずに目分量で紐を $3\mathrm{m}$ 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 とおく。仮に多数の様々な人達に目分量で $3\mathrm{m}$ であ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 $3.5\mathrm{m}$ 以下である割合が近似的に確率 $\mathrm{P}(X \leq 3.5)$ であると考えられる。このように、任意の実数 a について確率 $\mathrm{P}(X \leq a)$ が近似的に決まるとする。よって、多数の実験によって、各実数 x に対して確率 $\mathrm{P}(X \leq x)$ を近似的に定めることができる。従って、そのような実験によって、変数 X は確率変数と考えることができる。この確率変数のとり得る値は 3 の近くに“べったりと”(隙間がない実数の集合として)あるので、変数 X がとり得る値の総てに番号を付けることができない⁶⁾。よって X は離散型確率変数ではない。しかし、多数の実験によって実数 x に対する累積分布関数の値 $\mathrm{P}(X \leq x)$ を決めることができる。この累積分布関数の数値表およびグラフが例えば次のようになるとする。

x	$\mathrm{P}(X \leq x)$	x	$\mathrm{P}(X \leq x)$
2.2	0.000	3.1	0.550
2.3	0.013	3.1	0.648
2.4	0.049	3.2	0.741
2.5	0.104	3.3	0.825
2.6	0.175	3.4	0.896
2.7	0.259	3.5	0.951
2.8	0.352	3.6	0.987
2.9	0.450	3.7	1.000



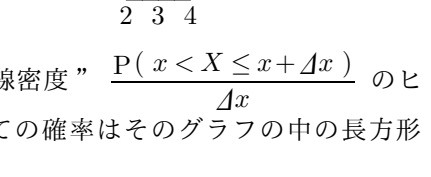
この数値表より、確率変数 X について例えば次のような計算ができる：

$$\mathrm{P}(X > 3.4) = 1 - \mathrm{P}(X \leq 3.4) = 1 - 0.896 = 0.104 .$$

$$\mathrm{P}(2.7 < X \leq 3.2) = \mathrm{P}(X \leq 3.2) - \mathrm{P}(X \leq 2.7) = 0.741 - 0.259 = 0.482 .$$

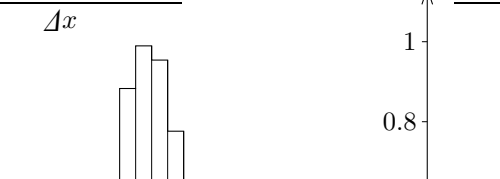
この確率変数 X に関する確率について、階級幅を $\Delta x = 0.5$ として度数分布表のようなものを考え、更にヒストグラムのようなグラフを描く。

x	$x + \Delta x$	確率 $\mathrm{P}(x < X \leq x + \Delta x)$
$-\infty$	2	0
2	2.5	0.104
2.5	3	0.446
3	3.5	0.401
3.5	4	0.049
4	∞	0



X の値が各階級幅に属す確率の合計は 1 である：

$$\mathrm{P}(2 < X \leq 2.5) + \mathrm{P}(2.5 < X \leq 3) + \mathrm{P}(3 < X \leq 3.5) + \mathrm{P}(3.5 < X \leq 4) = 1 .$$



このヒストグラムのようなグラフに含まれる各長方形の幅は階級幅の $\Delta x = 0.5$ である。このグラフに含まれる長方形を総て併せた図形の面積は

$$\begin{aligned} \mathrm{P}(2 < X \leq 2.5) \cdot \Delta x + \mathrm{P}(2.5 < X \leq 3) \cdot \Delta x \\ + \mathrm{P}(3 < X \leq 3.5) \cdot \Delta x + \mathrm{P}(3.5 < X \leq 4) \cdot \Delta x \\ = \{\mathrm{P}(2 < X \leq 2.5) + \mathrm{P}(2.5 < X \leq 3) + \mathrm{P}(3 < X \leq 3.5) + \mathrm{P}(3.5 < X \leq 4)\} \Delta x \\ = 1 \times \Delta x \\ = \Delta x . \end{aligned}$$

このグラフに含まれる長方形を総て併せた図形の面積は階級幅 Δx である。

このヒストグラムのようなグラフにおいて、例えば確率 $\mathrm{P}(2.5 < X \leq 3.5)$ は次のようになる。

$$\mathrm{P}(2.5 < X \leq 3.5) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

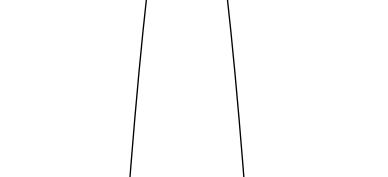
$$\mathrm{P}(2.5 < X \leq 3.5) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

この等式の右辺の分数の分母は階級幅 Δx である。

$$\mathrm{P}(2.5 < X \leq 3.5) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

そこで、確率 $\mathrm{P}(x < X \leq x + \Delta x)$ の替わりに $\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ を Δx で割った確率の“線密度” $\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ を考える。

x	$x + \Delta x$	確率の線密度 $\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$
$-\infty$	2	0
2	2.5	0.208
2.5	3	0.892
3	3.5	0.803
3.5	4	0.097
4	∞	0



$$\frac{\mathrm{P}(2.5 < X \leq 3)}{\Delta x} = \frac{\text{明緑色の長方形を併せた領域の面積}}{\Delta x}$$

$$\frac{\mathrm{P}(3 < X \leq 3.5)}{\Delta x} = \frac{\text{明緑色の長方形を併せた領域の面積}}{\Delta x}$$

$$\frac{\mathrm{P}(2 < X \leq 2.5)}{\Delta x} = \frac{\text{明緑色の長方形を併せた領域の面積}}{\Delta x}$$

$$\frac{\mathrm{P}(3.5 < X \leq 4)}{\Delta x} = \frac{\text{明緑色の長方形を併せた領域の面積}}{\Delta x}$$

ヒストグラムのようなグラフに含まれる長方形を総て併せた図形の面積は

$$\begin{aligned} \frac{\mathrm{P}(2 < X \leq 2.5)}{\Delta x} \cdot \Delta x + \frac{\mathrm{P}(2.5 < X \leq 3)}{\Delta x} \cdot \Delta x + \\ \frac{\mathrm{P}(3 < X \leq 3.5)}{\Delta x} \cdot \Delta x + \frac{\mathrm{P}(3.5 < X \leq 4)}{\Delta x} \cdot \Delta x \\ = \mathrm{P}(2 < X \leq 2.5) + \mathrm{P}(2.5 < X \leq 3) + \mathrm{P}(3 < X \leq 3.5) + \mathrm{P}(3.5 < X \leq 4) \\ = 1 . \end{aligned}$$

このヒストグラムのようなグラフにおいて、例えば確率 $\mathrm{P}(2.5 < X \leq 3.5)$ は次のようになる：

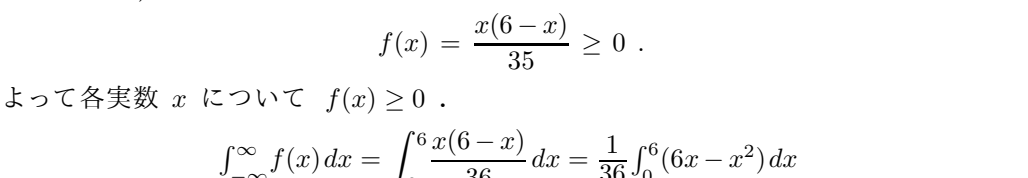
$$\mathrm{P}(2.5 < X \leq 3.5) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

$$\mathrm{P}(2.5 < X \leq 3.5) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

$$\mathrm{P}(2.5 < X \leq 3.5) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

このように、確率変数 X についての確率の“線密度” $\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ 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ある。

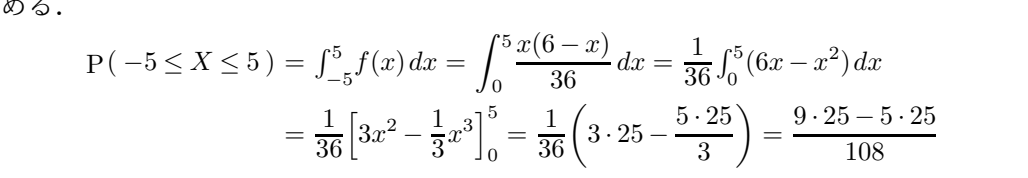
階級幅が $\Delta x = 0.2$ のとき及び $\Delta x = 0.05$ のとき、確率変数 X について確率の“線密度” $\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ のヒストグラムのようなグラフは次のようになる。



確率の“線密度” $\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ 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なので、例えば確率 $\mathrm{P}(2.6 < X \leq 3.4)$ は次のようになる：

$$\mathrm{P}(2.6 < X \leq 3.4) = \frac{\text{明緑色の長方形を併せた領域の面積}}{\text{階級幅 } \Delta x}$$

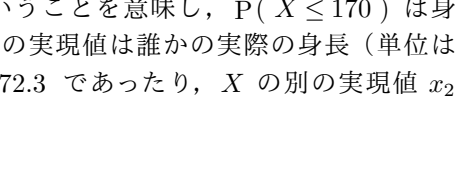
階級幅 Δx を更に 0 に近づける。



階級幅 Δx を 0 に限りなく近づけていく。各実数 x に対して極限値 $\lim_{\Delta x \rightarrow 0} \frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ があるとする。このとき、各実数 x に $\lim_{\Delta x \rightarrow 0} \frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ を対応させる関数ができる。この関数を $f(x)$ とおく：各実数 x に対して

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x} .$$

$\frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ 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あった。このことから、関数 $f(x) = \lim_{\Delta x \rightarrow 0} \frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ のグラフにおいて、実数 a, b について $a \leq b$ のとき確率 $\mathrm{P}(a < X \leq b)$ は次のようになる：

$$\mathrm{P}(a < X \leq b) = \frac{\text{明緑色の領域の面積}}{\text{階級幅 } \Delta x} = \int_a^b f(x) dx .$$

このように、関数 $f(x) = \lim_{\Delta x \rightarrow 0} \frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x}$ を考えると、 X に関する確率 $\mathrm{P}(a < X \leq b)$ は $f(x)$ の定積分 $\int_a^b f(x) dx$ である。

このことは次のようにも分かる。確率変数 X の累積分布関数を $F(x)$ とおく：各実数 x に対して $F(x) = \mathrm{P}(X \leq x)$ 。このとき、

$$\mathrm{P}(x < X \leq x + \Delta x) = \mathrm{P}(X \leq x + \Delta x) - \mathrm{P}(X \leq x) = F(x + \Delta x) - F(x) .$$

従って、

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{\mathrm{P}(x < X \leq x + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta$$