統計的な推測【連続した値をとる確率変数】　p.74～75

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り（確認テスト・相互採点・リフレクションの記入）

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

【内容目標】正規分布の考え方を理解しよう

　これまでは、さいころの目や硬貨の枚数などのように、その値がとびとびで有限個であるような確率変数を考えてきた。ここでは、長さや時間のように、連続した値をとる確率変数について考えよう。

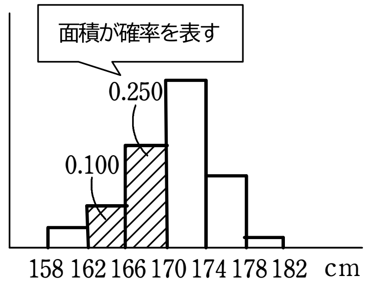
□連続した値をとる確率変数

　右の表は，あるクラス 40 人の身長を調べた結果で，度数分布表において，各階級の相対度数をまとめたものである。この 40 人の中から 1 人を選び，その生徒の身長を X とする。

　たとえば，X が 174 cm 以上 178 cm 未満の階級に属する確率は相対度数 0.175 であると考えてよい。このように考えると，X が各階級に属する確率は，各階級の相対度数に一致する。したがって，X は各階級に対して属する確率の定まる確率変数と考えられる。

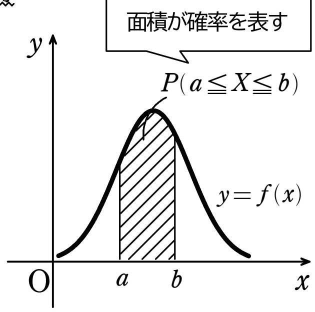
身長の階級（cm）	度数	相対度数
158 以上 162 未満	2	0.050
162　～ 166	4	0.100
166　～ 170	10	0.250
170　～ 174	16	0.400
174　～ 178	7	0.175
178　～ 182	1	0.025
計	40	1.000

　右の図は，上の表で示した各階級の相対度数を，長方形の面積で表したヒストグラムである。たとえば，図の斜線部分の面積が $162 \leqq X < 170$ となる確率を表す。



　データの大きさを増やし階級の幅を細かく分けて同様なヒストグラムを作っていくと，ヒストグラムの形はある曲線に近づいていくと考えられる。

　一般に，連続した値をとる確率変数 X を **連続型確率変数** という。連続型確率変数 X の確率分布を考える場合は，X に 1 つの曲線 $y=f(x)$ を対応させ，$a \leqq X \leqq b$ となる確率 $P(a \leqq X \leqq b)$ が，右の図の斜線部分の面積で表されるようにする。



　この曲線 $y=f(x)$ を，X の **分布曲線** といい，

　関数 $f(x)$ を **確率密度関数** という。

確率密度関数になるかどうかは，一般に関数 $y=f(x)$ において
x が確率変数 X の実現値として扱うことが前提になっていれば
$f(a) \geqq 0, \int_{-\infty}^{\infty} f(x)\,dx=1$　が示されればよい

統計的な推測【連続した値をとる確率変数】　p.74～75

確率密度関数 $f(x)$ は，次のような性質をもつ。

確率密度関数の性質

1　常に　$f(x) \geqq 0$

2　確率 $P(a \leqq X \leqq b)$ は，曲線 $y=f(x)$ と x 軸，2 直線 $x=a$，$x=b$ で囲まれた部分の面積に等しい。

すなわち　　$P(a \leqq X \leqq b)=\int_a^b f(x)\,dx$

> 1点の確率は0であるから
> $P(a \leqq X \leqq b)=P(a<X<b)$

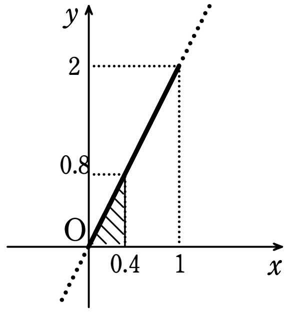
3　X のとる値の範囲が $\alpha \leqq X \leqq \beta$ のとき　$\int_\alpha^\beta f(x)\,dx=1$

> 端から端までの
> 面積（確率の和）は1

例１６） 確率変数 X のとる値の範囲が $0 \leqq X \leqq 1$ で，
その確率密度関数が $f(x)=2x$ $(0 \leqq x \leqq 1)$ であるとき

$$P(0 \leqq X \leqq 0.4)=\frac{1}{2}\times 0.4\times 0.8$$
$$=0.16$$

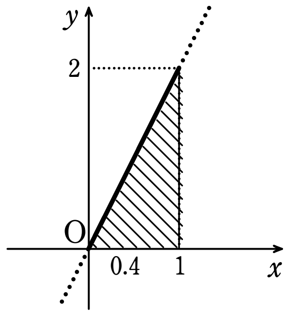
> 三角形の面積の公式



$$P(0 \leqq X \leqq 1)=\frac{1}{2}\times 1\times 2$$
$$=1$$

> 三角形の面積の公式

終



＜補足＞　定積分の計算を用いると，次のように求められる。

$$P(0 \leqq X \leqq 0.4)=\int_0^{0.4} 2x\,dx=\Big[x^2\Big]_0^{0.4}=0.4^2-0^2=0.16$$

$$P(0 \leqq X \leqq 1)=\int_0^{1} 2x\,dx=\Big[x^2\Big]_0^{1}=1^2-0^2=1$$

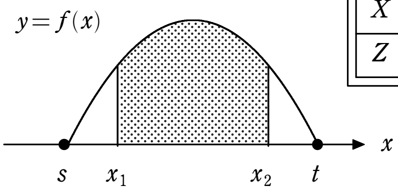
> 定積分の公式

統計的な推測【連続した値をとる確率変数】　p.74～75

補足　確率変数 X の確率密度関数を $y=f(a)$ とする。
この確率変数 X を，$Z=aX+b$ として，確率変数 Z に変換するとき，
得られる確率密度関数 $y=g(z)$
このとき，$y=f(x)$ と $y=g(z)$ との間には，どのような関係があるか？

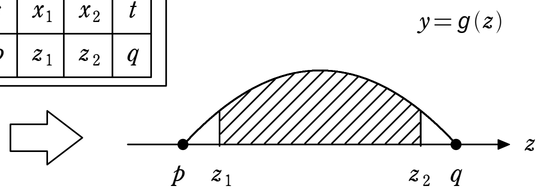
確率変数 X に対して，確率変数 Z を　$Z=aX+b\ (a>0)$ とする。
確率変数 X の値の範囲が区間 $[s,\ t]$，これに対して確率変数 Z の値の範囲が区間 $[p,\ q]$ とする。
ただし，$p=as+b$，$q=at+b$ である。

任意の x_1，x_2 に対して，$z_1=ax_1+b$，$z_2=ax_2+b$ のとき，以下の対応をする。

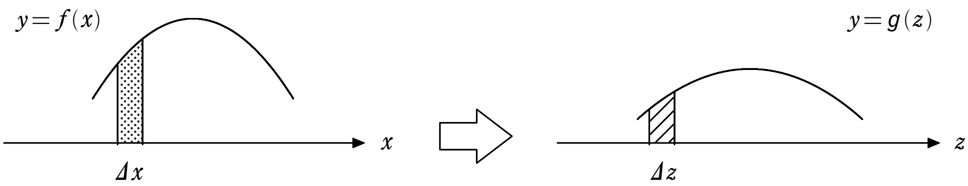


$z=ax+b$

X	s	x_1	x_2	t
Z	p	z_1	z_2	q



ここで確率密度関数 $y=g(z)$ は，$P(x_1 \leqq X \leqq x_2)=P(z_1 \leqq Z \leqq z_2)$ を満たすので
x，z の増分 $\varDelta x$，$\varDelta z$ とすると，$P(x \leqq X \leqq x+\varDelta x)=P(z \leqq Z \leqq z+\varDelta z)$ である。



$\varDelta x$，$\varDelta z$ が微小なら　$f(x)\varDelta x=g(z)\varDelta z$　　$\therefore$　$f(x)=g(z)\dfrac{\varDelta z}{\varDelta x}$ が成り立つ

$\varDelta x \longrightarrow 0$ のとき，$f(x)=g(z)\dfrac{dz}{dx}=ag(z)$　　よって　$g(z)=\dfrac{1}{a}f(x)$

【ポイント】
$y=g(z)$ のグラフは，$y=f(x)$ を横軸（ x 軸）方向に a 倍して，$+b$ だけ平行移動しただけではなく，確率密度関数として成立するために，縦軸（ y 軸）方向に $\dfrac{1}{a}$ 倍している

統計的な推測【連続した値をとる確率変数】　p.74～75

補足　得られた関数 $y=g(z)$ が確率密度関数の性質をもつか確認する

$$g(z)=g(ax+b)=\frac{1}{a}f(x)$$ より

(1)　$g(z)=\dfrac{1}{a}f(x)\geqq 0$

(2)　$P(z_1\leqq z\leqq z_2)=P(x_1\leqq x\leqq x_2)=\displaystyle\int_{x_1}^{x_2}f(x)dx=\int_{z_1}^{z_2}ag(z)\cdot\frac{1}{a}dz=\int_{z_1}^{z_2}g(z)dz$

$P(z_1\leqq z\leqq z_2)=\displaystyle\int_{z_1}^{z_2}g(z)dz$

$\displaystyle\int_s^t f(x)dx=1$

(3)　$\displaystyle\int_p^q g(z)dz=\int_s^t g(ax+b)adx=\int_s^t\frac{1}{a}f(x)\cdot adx=\int_s^t f(x)dx=1$

以上より　$y=g(z)$ は確率密度関数の性質をもつことが確認できる。

補足　$Z=aX+b$ で変換された Z の平均と分散を調べる。

ここで連続型確率変数 X の平均 $E(X)$ および分散 $V(X)$ は

平均　$E(X)=\displaystyle\int_s^t xf(x)dx$　　　分散　$V(X)=\displaystyle\int_s^t(x-m)^2f(x)dx$

ただし，$m=E(X)$ として定義さている。

$$E(X)=\int_p^q zg(z)dz=\int_s^t(ax+b)g(ax+b)adx$$

$\displaystyle\int_s^t f(x)dx=1$

$$=\int_s^t(ax+b)f(x)dx=a\int_s^t xf(x)dx+b\int_s^t f(x)dx=am+b=\boldsymbol{aE(X)+b}$$

$$V(Z)=\int_p^q(z-E(Z))^2g(z)dz=\int_p^q\{z-(am+b)\}^2g(z)dz$$

$$=\int_s^t(ax+b-am-b)^2g(ax+b)\cdot adx$$

$$=\int_s^t a^2(x-m)^2f(x)dx=a^2\int_s^t(x-m)^2f(x)dx=\boldsymbol{a^2V(X)}$$

【ポイント】

X の１次式 $Z=aX+b$ による確率分布の平均および分散は，連続型も離散型同様に算出できる

$$E(Z)=E(aX+b)=aE(X)+b$$

$$V(Z)=V(aX+b)=a^2V(X)$$