統計的な推測【連続型確率変数】p.78～80

青チャートB: 例題 78～79

【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】 連続型確率変数について理解できる。

○正規分布

これまでに学んだ確率変数は，さいころの目や硬貨の枚数などのように、とびとびの値しかとらなかったが，長さや時間のように、確率変数のとる値を連続的なものとして扱うことが必要であったり，便利であったりすることがある。

ここでは連続的な値をとる確率変数の確率分布について考えてみよう。

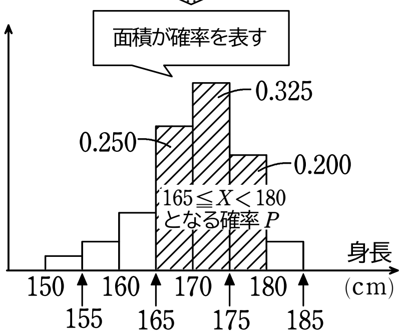
□連続型確率変数とその分布

右の表は，あるクラスの生徒の身長の度数分布表である。この 40 人の生徒から 1 人を選んだとき，その生徒の身長 X の属する階級の階級値が 157.5 である確率は，相対度数 0.050 であると考えてよい。

つまり，X は階級値の値をとる確率変数と考えられ，その確率分布は相対度数の分布と一致する。

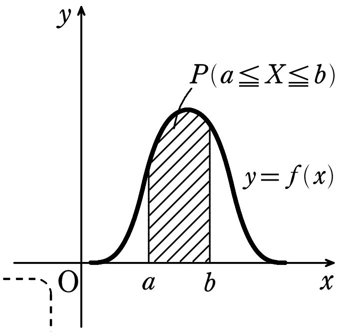
階級（cm）	階級値	度数	相対度数
150 以上 ～ 155 未満	152.5	1	0.025
155 ～ 160	157.5	2	0.050
160 ～ 165	162.5	4	0.100
165 ～ 170	167.5	10	0.250
170 ～ 175	172.5	13	0.325
175 ～ 180	177.5	8	0.200
180 ～ 185	182.5	2	0.050
計		40	1.000

この確率分布を図示するには，右の図のように，各階級の上の長方形の面積が，その階級の相対度数を表すようなヒストグラムをかけばよい。
このとき，例えば，$165 \leqq X < 180$ となる**確率 P は**，右の図の**斜線部分の面積で表される**。



データの大きさを増し，階級の幅も狭くしていくと，ヒストグラムの形は次第にある 1 つの曲線に近づいていく。

そこで，一般に，連続的な値をとる確率変数 X の確率分布を考える場合には，X に 1 つの曲線を対応させ，**$a \leqq X \leqq b$ となる確率が図の斜線を施した部分の面積で表される**ようにする。このような曲線を X の **分布曲線** という。



参考 (1) ヒストグラムを作成する際に縦軸は各階級の度数 f_x ではなく，

$\dfrac{f_x}{nh}$ （n は総度数，h は階級の幅）

(2) 各柱の底辺は h （階級の幅）であるから，各階級に対応する柱状形の面積は

$\dfrac{f_x}{nh} \cdot h = \dfrac{f_x}{n}$ であり，ヒストグラムの面積は 1 になる。

→ n を大きくして，h を小さくするとヒストグラムはなめらかな曲線に近づく

統計的な推測【連続型確率変数】 p.78～80

青チャートB: 例題 78～79

横軸を x 軸に，縦軸を y 軸にとるとき，X の分布曲線の方程式を $y=f(x)$ とすると，関数 $f(x)$ は次の性質をもつ。

[1] 常に $f(x)\geqq 0$

[2] 確率 $P(a\leqq X\leqq b)$ は，$y=f(x)$ のグラフと x 軸，および 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

【注意】$f(x)$ の定義域は，実数全体のことも，その一部分のこともある。

連続的な値をとる確率変数を **連続型確率変数** といい，$f(x)$ を X の **確率密度関数** という。
これに対し，今まで扱ってきたような，とびとびの値をとる確率変数を **離散型確率変数** という。

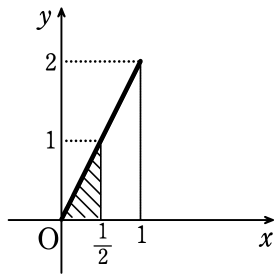
確率密度関数になるかどうかは，一般に関数 $y=f(x)$ において
x が確率変数 X の実現値として扱うことが前提になっていれば
$f(a)\geqq 0$，　$\int_{-\infty}^{\infty} f(x)\,dx=1$　が示されればよい

ここでは，特に断りがない場合，確率変数 X は連続型であるとする。

例９） 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

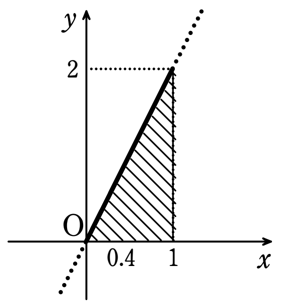
$$P\left(0\leqq X\leqq \frac{1}{2}\right)=\frac{1}{2}\cdot\frac{1}{2}\cdot 1=\frac{1}{4}$$

三角形の面積の公式



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

三角形の面積の公式



補足　定積分の計算を用いると，次のように求められる。

$$P\left(0\leqq X\leqq \frac{1}{2}\right)=\int_0^{\frac{1}{2}} 2x\,dx=\Big[x^2\Big]_0^{\frac{1}{2}}=\left(\frac{1}{2}\right)^2-0^2=\frac{1}{4}$$

定積分の公式

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

統計的な推測【連続型確率変数】 p.78～80

青チャートB: 例題 78～79

確率変数 X のとる値の範囲が $\alpha \leqq X \leqq \beta$ で，その確率密度関数が $f(x)$ のとき，X の期待値 $E(X)$ と分散 $V(X)$ は，$E(X)=m$ とすると，次のように定義される。

連続型確率変数の期待値と分散

$$E(X)=\int_\alpha^\beta xf(x)dx,$$

$$V(X)=\int_\alpha^\beta (x-m)^2 f(x)dx$$

離散的な確率変数の期待値，分散は

期待値　$m=E(X)=\sum_{k=1}^{n} x_k p_k$

分散　$V(X)=\sum_{k=1}^{n} (x_k-m)^2 p_k$

離散的な確率変数から連続型の確率変数への
変形には数学Ⅲで学ぶ区分求積法の考えを用いる

また，X の標準偏差 $\sigma(X)$ を $\sqrt{V(X)}$ で定める。

注意 関数 x^n の定積分の計算では，次のことを用いる。

n が 0 または正の整数のとき　$\displaystyle\int_a^b x^n dx=\left[\frac{1}{n+1}x^{n+1}\right]_a^b$

問4） 例 9 同様　確率変数 X のとる値 x の範囲が $0\leqq x\leqq 1$ で，その確率密度関数 $f(x)$ が

$$f(x)=2x \ (0\leqq x\leqq 1)$$

で表されるとき，の期待値，分散，標準偏差を求めよ。

解答

$$E(X)=\int_0^1 2x^2 dx=\left[\frac{2}{3}x^3\right]_0^1=\frac{2}{3}$$

$$V(X)=\int_0^1 2x\left(x-\frac{2}{3}\right)^2 dx$$

$(x-m)^2 f(x)=\left(x-\frac{2}{3}\right)^2\cdot 2x$

$$=\int_0^1\left(2x^3-\frac{8}{3}x^2+\frac{8}{9}x\right)dx$$

$$=\left[\frac{1}{2}x^4-\frac{8}{9}x^3+\frac{4}{9}x^2\right]_0^1=\frac{1}{18}$$

$$\sigma(X)=\sqrt{\frac{1}{18}}=\frac{\sqrt{2}}{6}$$

分散については
(2乗の平均)－(平均の2乗) から
$V(X)=\int_\alpha^\beta x^2 f(x)dx-\{E(X)\}^2$
を用いる方が計算しやすい。
($V(X)=\int_\alpha^\beta (x-m)^2 f(x)dx$ を
展開して整理すると証明できる)

補足 連続型確率変数 X の確率密度関数が $f(x)$ であり，X のとり得る値の範囲が $\alpha\leqq X\leqq\beta$ のとき，
領域 D「$0\leqq y\leqq f(x)$ かつ $\alpha\leqq X\leqq\beta$」の重心 G の x 座標が $E(X)$ である。
したがって $E(X)$ を求める検算用のテクニックとして次のようなものがある。

1　D が三角形のとき 頂点の座標を $(x_k,\ y_k)$ $(k=1,\ 2,\ 3)$ と表すと G の x 座標 $E(X)$ は　$E(X)=\dfrac{x_1+x_2+x_3}{3}$

2　D が直線 $x=a$ について対称なとき G は $x=a$ 上にあるから，G の x 座標 $E(X)$ は　$E(X)=a$

統計的な推測【連続型確率変数】 p.78～80

青チャートB: 例題 78～79

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 $y=g(z)$

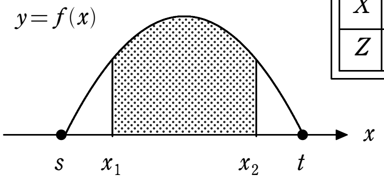
このとき， $y=f(x)$ と $y=g(z)$ との間には，どのような関係があるか？

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。

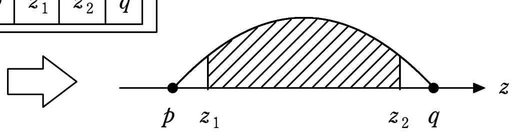
ただし， $p=as+b$， $q=at+b$ である。

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。



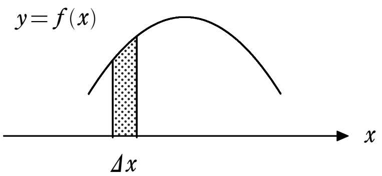
$z=ax+b$

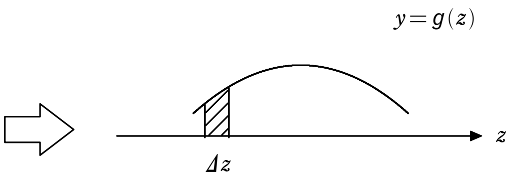
X	s	x_1	x_2	t
Z	p	z_1	z_2	q



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 Δx， Δz とすると， $P(x \leqq X \leqq x+\Delta x)=P(z \leqq Z \leqq z+\Delta z)$ である。





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

$\Delta x \to 0$ 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

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)　$g(z)=\frac{1}{a}f(x)\geqq 0$

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

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

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

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

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

ただし，$m=E(X)$ として定義さている。

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

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

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

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

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

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

【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

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

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