3. 連続型の確率変数と確率密度に慣れよう！

確率変数には飛び飛びの値をとる“**離散型**（りさんがた）**のもの**”と，“**連続型**（れんぞくがた）**のもの**”とがあるんだよ。前章まではすべて離散型の確率変数ばかりを勉強してきたけれど，今回は**連続型の確率変数**について勉強していこう！

● 確率密度の意味をマスターしよう！

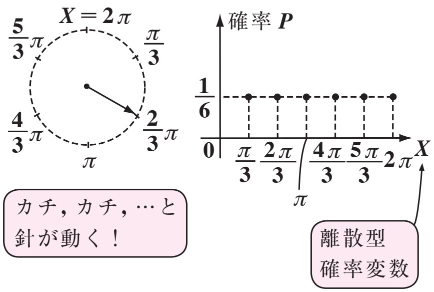
まず，離散型と連続型の確率分布の違いを例で示すよ。

(Ⅰ) 離散型の確率分布

図1(ⅰ)に示すように，円周上に6つの目盛り $\frac{\pi}{3}, \frac{2}{3}\pi, \cdots\cdots, 2\pi$ がつけてあり，この6点を同様に確からしく針がカチカチ…と指す場合を考える。これらの目盛りを確率変数 X とおくと $X=\frac{\pi}{3}, \frac{2}{3}\pi, \cdots\cdots, 2\pi$ となる確率は，図1(ⅱ)のようにすべて $\frac{1}{6}$ となる。これはいいね？

図1 (Ⅰ) 離散型確率分布の例

(ⅰ)　　　　(ⅱ) 確率分布

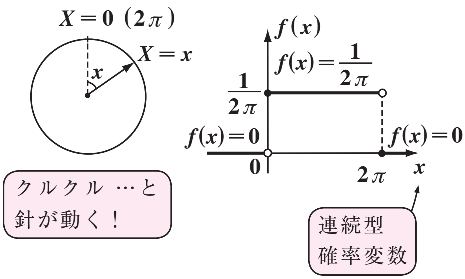


(Ⅱ) 連続型の確率分布

次，図2(ⅰ)に示すように，同じ円周に対して針が自由にクルクルと回って，無作為にある1点に止まる場合を考える。このとき，針が $X=x$ $(0 \leqq x < 2\pi)$ の点を指す確率を計算できる？　そう。今回は円周上には連続的に無限に点が並んでいるので，$X=x$ となる確率は x の値に関わらず常に 0 $\left(=\frac{1}{\infty}\right)$ となるんだね。

図2 (Ⅱ) 連続型確率分布の例

(ⅰ)　　　　(ⅱ) 確率密度



でも，x が $0 \leqq x \leqq \frac{\pi}{3}$ の範囲に入る確率は，$\frac{60^\circ}{360^\circ}=\frac{1}{6}$ とすぐに求まるだろう？ 一般に連続型の確率計算では，確率変数 X が $a \leqq X \leqq b$ の範囲に入る

確率 $P(a \leqq X \leqq b)$ を，$P(a \leqq X \leqq b) = \int_a^b f(x)dx$ の定積分の形で表すんだよ。

この被積分関数 $f(x)$ のことを“**確率密度関数**(かくりつみつどかんすう)” または “**確率密度**(かくりつみつど)” と呼ぶ。

図 2 の例では，確率変数 X が，$0 \leqq X < 2\pi$ の範囲に入る確率が全確率 1 で

ここに等号をつけてもかまわない。どうせ $X = 2\pi$ となる確率は 0 だからね。

あり，また，この範囲内のどの点に対しても針は同様に確からしく指すはずだから，この確率密度 $f(x)$ は $f(x) = c$ (定数) と定数関数になるはずだ。よって，

$$\int_0^{2\pi} \overset{c}{f(x)} dx = c[x]_0^{2\pi} = \boxed{2\pi c = 1}$$ (全確率) となるね。

$\therefore c = \dfrac{1}{2\pi}$ より，この確率密度 $f(x)$ は $f(x) = \dfrac{1}{2\pi}$ $(0 \leqq x < 2\pi)$ となる。当然，$x < 0,\ 2\pi \leqq x$ のとき $f(x) = 0$ だね。(図 2 (ⅱ) を参照してくれ。)

それでは，連続型確率変数と確率密度について，まとめておこう。

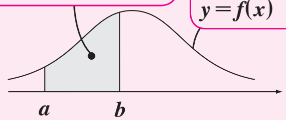
連続型確率変数 X と確率密度 $f(x)$

連続型確率変数 X が $a \leqq X \leqq b$ となる確率 $P(a \leqq X \leqq b)$ は次式で表される。

$$\boxed{P(a \leqq X \leqq b) = \int_a^b f(x)dx} \quad (a < b)$$

このような関数 $f(x)$ が存在するとき，$f(x)$ を“**確率密度**”と呼び，確率変数 X は確率密度 $f(x)$ の連続型確率分布に従うという。

また，$y = f(x)$ のグラフを X の**分布曲線**(ぶんぷきょくせん)と呼ぶ。



注意 連続型確率分布では，$X = x$ のように表す場合がよくある。この場合，「確率変数 X が，ある値 x である」というように考えるといいよ。ただし，確率密度 $f(x)$ では，x は変数として扱われる。このような独特の表現法にも慣れていくことだね。

連続型確率分布の 4 つの性質を下に示すよ。

連続型確率分布の性質

(ⅰ) $P(X=a)=0$　　(ⅱ) $f(x)\geqq 0$　　(ⅲ) $\int_{-\infty}^{\infty} f(x)\,dx=1$ (全確率)

$x=a$ となる確率は $\mathbf{0}$

(ⅳ) $\int_a^b f(x)\,dx=P(a\leqq X\leqq b)=P(a<X\leqq b)$

$=P(a\leqq X<b)=P(a<X<b)$

$X=a$, $X=b$ となる確率は $\mathbf{0}$ なので，等号はあってもなくても同じになる。

● 連続型確率分布の期待値と分散を押さえよう！

確率密度 $f(x)$ に従う確率変数 X の**期待値**(**平均**)$m=E(X)$ と**分散** $\sigma^2=V(X)$ と**標準偏差** $\sigma=D(X)$ の定義式と計算式を次に示す。

“シグマの 2 乗”と読む　　“シグマ”と読む

X の期待値・分散・標準偏差

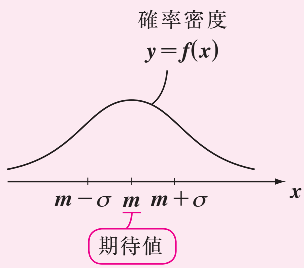
確率密度 $f(x)$ に従う連続型確率変数 X の期待値，分散，標準偏差は次のようになる。

(1) 期待値 $m=E(X)=\int_{-\infty}^{\infty} xf(x)\,dx$

(2) 分散 $\sigma^2=V(X)=\int_{-\infty}^{\infty}(x-m)^2f(x)\,dx$

$=E(X^2)-\{E(X)\}^2$

(3) 標準偏差 $\sigma=D(X)=\sqrt{V(X)}$



離散型確率変数 X の場合のもの，すなわち，

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

(3) 標準偏差 $\sigma=D(X)=\sqrt{V(X)}$　と対比して覚えよう！

期待値 m，分散σ^2，標準偏差 σ を求める場合，離散型のものの Σ 計算が，連続型のものでは積分計算に変わっていることに注意しよう。

離散型のときと同様に連続型確率変数においても，$E(X)=\int_{-\infty}^{\infty}xf(x)dx$ と定義すると，$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)dx$ などとなるのは大丈夫だね。これから (2) の分散 σ^2 も次のように変形できるんだね。

$$(2)\ \sigma^2=V(X)=\int_{-\infty}^{\infty}(x-m)^2f(x)dx=\int_{-\infty}^{\infty}(x^2-2mx+m^2)f(x)dx$$

$$=\underbrace{\int_{-\infty}^{\infty}x^2f(x)dx}_{E(X^2)}-2m\underbrace{\int_{-\infty}^{\infty}xf(x)dx}_{m=E(X)}+m^2\underbrace{\int_{-\infty}^{\infty}f(x)dx}_{1\text{（性質（iii）より）}}$$

$$=E(X^2)-2m^2+m^2=E(X^2)-\underbrace{m^2}_{\{E(X)\}^2}$$

$$=E(X^2)-\{E(X)\}^2$$ が導けた！←離散型の σ^2 と同じ式だね。

さらに，確率変数 X を使って，新たな確率変数 Y を $Y=aX+b$ （a, b：実数定数）と定義したとき，Y の期待値 $E(Y)$，分散 $V(Y)$ と標準偏差 $D(Y)$ は次のようになる。この結果も離散型のときのものと同様だから覚えやすいはずだ。

Y の期待値・分散・標準偏差

$Y=aX+b$ （a, b：実数定数）により，Y を新たに定義すると，

(1) 期待値 $E(Y)=E(aX+b)=aE(X)+b$

(2) 分散 $V(Y)=V(aX+b)=a^2V(X)$

(3) 標準偏差 $D(Y)=\sqrt{V(Y)}=\sqrt{a^2V(X)}=|a|\sqrt{V(X)}=|a|D(X)$

$$(1)\ E(Y)=E(aX+b)=\int_{-\infty}^{\infty}(ax+b)f(x)dx=a\underbrace{\int_{-\infty}^{\infty}xf(x)dx}_{E(X)}+b\underbrace{\int_{-\infty}^{\infty}f(x)dx}_{1\text{（全確率）}}$$

$=aE(X)+b$ となる。また，

$$(2)\ V(Y)=V(aX+b)=\int_{-\infty}^{\infty}\{\underbrace{(ax+\cancel{b})}_{Y}-\underbrace{(am+\cancel{b})}_{Y\text{の期待値}E(Y)}\}^2f(x)dx$$ ←分散の定義式

$$=a^2\int_{-\infty}^{\infty}(x-m)^2f(x)dx=a^2V(X)$$ となるのも大丈夫？

連続型確率分布の期待値・分散・標準偏差（Ⅰ）

絶対暗記問題 56	難易度 ★★	CHECK1	CHECK2	CHECK3

確率密度$f(x)$が，$f(x)=\begin{cases} ax^2 & (0 \leqq x \leqq \sqrt{3},\ a\text{は正の定数}) \\ 0 & (x<0,\ \sqrt{3}<x) \end{cases}$ で定義されている。

このとき，次の各問いに答えよ。

(1) a の値を求めよ。

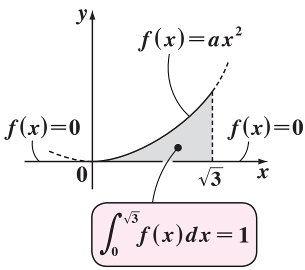
(2) 確率密度 $f(x)$ に従う確率変数 X の期待値 $m=E(X)$，
　分散 $\sigma^2=V(X)$，標準偏差 $\sigma=D(X)$ を求めよ。

(3) X を使って新たな確率変数 Y を $Y=4X-\sqrt{3}$ で定義するとき，
　Y の期待値 $E(Y)$，分散 $V(Y)$，標準偏差 $D(Y)$ を求めよ。

ヒント！ 確率密度 $f(x)$ は，$x<0$ または $\sqrt{3}<x$ のとき 0 より，$0 \leqq x \leqq \sqrt{3}$ でのみ定義されていると考えればいい。よって，(1)では，$\int_0^{\sqrt{3}} f(x)dx=1$（全確率）から，$a$ の値を求めよう。(2)では，定義式に従って，$E(X)$，$V(X)$，$D(X)$ を求めればいい。(3)では，$Y=4X-\sqrt{3}$ より，$E(Y)=4E(X)-\sqrt{3}$，$V(Y)=4^2V(X)$，$D(Y)=\sqrt{V(Y)}$ となるんだね。正確に迅速に計算できるように練習しよう。

解答＆解説

確率密度 $f(x)=\begin{cases} ax^2 & (0 \leqq x \leqq \sqrt{3}) \\ 0 & (x<0,\ \sqrt{3}<x) \end{cases}$



(1) 確率密度の性質より，

$$\underbrace{\int_{-\infty}^{\infty} f(x)dx}_{\cancel{\int_{-\infty}^{0} f(x)dx}+\int_0^{\sqrt{3}} f(x)dx+\cancel{\int_{\sqrt{3}}^{\infty} f(x)dx}} = 1 \quad (\text{全確率})\quad \text{よって，}$$

$$\underbrace{\int_{-\infty}^{0} f(x)dx}_{0},\quad \underbrace{\int_{\sqrt{3}}^{\infty} f(x)dx}_{0}$$

$$\int_0^{\sqrt{3}} ax^2dx = a\left[\frac{1}{3}x^3\right]_0^{\sqrt{3}} = \frac{a}{3}(3\sqrt{3}-0) = \boxed{\sqrt{3}a=1}\ (\text{全確率})$$

$\therefore a=\dfrac{1}{\sqrt{3}}$ となる。……………………………………………………(答)

(2) $f(x)$ に従う確率変数 X の期待値 $E(X)$，分散 $V(X)$，標準偏差 $D(X)$ は，

$$m = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_0^{\sqrt{3}} x \cdot \frac{1}{\sqrt{3}}x^2 dx = \frac{1}{\sqrt{3}}\int_0^{\sqrt{3}} x^3 dx$$

$$= \frac{1}{\sqrt{3}}\left[\frac{1}{4}x^4\right]_0^{\sqrt{3}} = \frac{1}{4\sqrt{3}}\left\{(\sqrt{3})^4 \cancel{-0}\right\} = \frac{9}{4\sqrt{3}} = \frac{3\sqrt{3}}{4} \quad \cdots\cdots\cdots\cdots\text{(答)}$$

$$\sigma^2 = V(X) = E(X^2) - \{\underbrace{E(X)}_{\frac{3\sqrt{3}}{4}}\}^2 = \int_0^{\sqrt{3}} x^2 \cdot \frac{1}{\sqrt{3}}x^2 dx - \left(\frac{3\sqrt{3}}{4}\right)^2$$

$$= \frac{1}{\sqrt{3}}\left[\frac{1}{5}x^5\right]_0^{\sqrt{3}} - \frac{27}{16} = \frac{\overbrace{(\sqrt{3})^5}^{9\sqrt{3}}}{5\sqrt{3}} - \frac{27}{16} = \frac{9}{5} - \frac{27}{16}$$

$$= \frac{9\times16 - 27\times5}{80} = \frac{144-135}{80} = \frac{9}{80} \quad \cdots\cdots\cdots\cdots\text{(答)}$$

$$\sigma = D(X) = \sqrt{V(X)} = \sqrt{\frac{9}{80}} = \frac{\sqrt{3^2}}{\sqrt{4^2\times5}} = \frac{3}{4\sqrt{5}} = \frac{3\sqrt{5}}{20} \quad \cdots\cdots\cdots\cdots\text{(答)}$$

(3) X を使って定義された確率変数 $Y = 4X - \sqrt{3}$ の期待値 $E(Y)$，分散 $V(Y)$，標準偏差 $D(Y)$ を求めると，

$$E(Y) = E(4X - \sqrt{3}) = 4\underbrace{E(X)}_{\frac{3\sqrt{3}}{4}} - \sqrt{3}$$

公式：
$E(aX+b) = aE(X)+b$

$$= \cancel{4} \times \frac{3\sqrt{3}}{\cancel{4}} - \sqrt{3} = 3\sqrt{3} - \sqrt{3} = 2\sqrt{3} \quad \cdots\cdots\cdots\cdots\text{(答)}$$

$$V(Y) = V(4X - \sqrt{3}) = 4^2\underbrace{V(X)}_{\frac{9}{80}}$$

公式：
$V(aX+b) = a^2V(X)$

$$= 16 \times \frac{9}{80} = \frac{9}{5} \quad \cdots\cdots\cdots\cdots\text{(答)}$$

$$D(Y) = \sqrt{V(Y)} = \sqrt{\frac{9}{5}} = \frac{3}{\sqrt{5}} = \frac{3\sqrt{5}}{5} \quad \cdots\cdots\cdots\cdots\text{(答)}$$

連続型確率分布の期待値・分散・標準偏差（Ⅱ）

絶対暗記問題 57	難易度 ★★	CHECK1	CHECK2	CHECK3

確率密度 $f(x)$ が，$f(x)=\begin{cases} ax(2-x) & (0 \leqq x \leqq 2) \\ 0 & (x<0,\ 2<x) \end{cases}$ で定義されている。

このとき，次の問いに答えよ。（ただし，a は正の定数である。）

(1) a の値を求めよ。

(2) 確率密度 $f(x)$ に従う確率変数 X の期待値 $m=E(X)$，分散 $\sigma^2=V(X)$，標準偏差 $\sigma=D(X)$ を求めよ。

(3) X を使って新たな確率変数 Y を $Y=pX+q$ (p, q：定数，$p>0$) で定義する。Y の期待値 $E(Y)$ と標準偏差 $D(Y)$ がそれぞれ $E(Y)=0$，$D(Y)=1$ となるように p，q の値を定めよ。　（東京都立大＊）

ヒント！ 確率密度 $f(x)$ は $x<0,\ 2<x$ のとき 0 より，$0 \leqq x \leqq 2$ でのみ定義されていると考えればいいんだよ。よって，(1) $\int_0^2 f(x)dx=1$（全確率）から，a の値が求まる。(2) $E(X)$, $V(X)$, $D(X)$ の定義式，計算式に従って計算しよう。(3) 公式 $E(Y)=pE(X)+q$, $D(Y)=|p|D(X)$ を使って解けばいい。

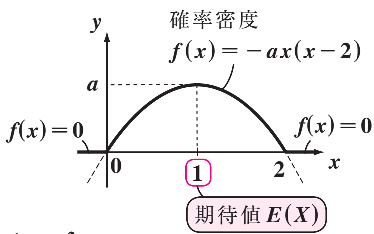
解答＆解説

確率密度 $f(x)=\begin{cases} a(2x-x^2) & (0 \leqq x \leqq 2) \\ 0 & (x<0,\ 2<x) \end{cases}$

(1) 確率密度の性質より，

$$\int_{-\infty}^{\infty} f(x)dx = 1 \text{（全確率）}\quad \text{よって，}$$

$$\int_{-\infty}^{\infty} f(x)dx = \underbrace{\cancel{\int_{-\infty}^{0} f(x)dx}}_{0} + \int_0^2 f(x)dx + \underbrace{\cancel{\int_2^{\infty} f(x)dx}}_{0}$$



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

$$=a\left(4-\frac{8}{3}\right)=\boxed{\frac{4}{3}a=1}\text{（全確率）}\qquad \therefore\ a=\frac{3}{4}\ \cdots\cdots\cdots\text{（答）}$$

(2) $f(x)$ に従う確率変数 X の期待値 $E(X)$，分散 $V(X)$，標準偏差 $D(X)$ は，

$$m=E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_0^2 x\underbrace{f(x)}_{\frac{3}{4}(2x-x^2)}dx=\frac{3}{4}\int_0^2(2x^2-x^3)dx$$

分布の対称性から，当然の結果だね。

$$=\frac{3}{4}\left[\frac{2}{3}x^3-\frac{1}{4}x^4\right]_0^2=\frac{3}{4}\left(\frac{16}{3}-4\right)=\frac{3}{4}\cdot\frac{4}{3}=1 \quad\cdots\cdots\cdots\cdots\text{(答)}$$

$$\sigma^2=V(X)=E(X^2)-\{\underbrace{E(X)}_{1}\}^2=\int_0^2 x^2\underbrace{f(x)}_{\frac{3}{4}(2x-x^2)}dx-1^2$$

公式：
$\int x^n dx=\frac{1}{n+1}x^{n+1}+C$
を使った。

$$=\frac{3}{4}\int_0^2(2x^3-x^4)dx-1=\frac{3}{4}\left[\frac{1}{2}x^4-\frac{1}{5}x^5\right]_0^2-1$$

$$=\frac{3}{4}\left(8-\frac{32}{5}\right)-1=\frac{3}{4}\cdot\frac{8}{5}-1=\frac{6}{5}-1=\frac{1}{5} \quad\cdots\cdots\cdots\cdots\text{(答)}$$

$$\sigma=D(X)=\sqrt{V(X)}=\sqrt{\frac{1}{5}}=\frac{1}{\sqrt{5}}=\frac{\sqrt{5}}{5} \quad\cdots\cdots\cdots\cdots\text{(答)}$$

(3) X を使って，新たな確率変数 Y を $Y=pX+q$ (p，q：定数，$p>0$) で定義するとき，期待値 $E(Y)=0$，標準偏差 $D(Y)=1$ となるので，

$$\begin{cases} E(Y)=E(pX+q)=p\underbrace{E(X)}_{1}+q=\boxed{p+q=0} \\ D(Y)=D(pX+q)=|p|D(X)=\boxed{p\cdot\frac{1}{\sqrt{5}}=1} \quad (\because p>0) \end{cases}$$

以上より，$p+q=0$，　$p=\sqrt{5}$　より

$p=\sqrt{5}$，　$q=-\sqrt{5}$ ……………(答)

$Y=\frac{X-m}{\sigma}=\frac{X-1}{\frac{1}{\sqrt{5}}}$
$=\sqrt{5}(X-1)=\underbrace{\sqrt{5}}_{p}X\underbrace{-\sqrt{5}}_{q}$ のこと

参考

確率変数 X の期待値を $m=E(X)$，標準偏差を $\sigma=D(X)$ とする。

このとき，新たな確率変数 Y を $Y=\frac{X-m}{\sigma}$ で定義すると，

このように変数を変換することを変数の“**標準化**”という

Y の期待値 $E(Y)=0$，標準偏差 $D(Y)=1$ となることも覚えておこう。

なぜなら

$$\begin{cases} E(Y)=E\left(\frac{1}{\sigma}X-\frac{m}{\sigma}\right)=\frac{1}{\sigma}\cdot\underbrace{E(X)}_{m}-\frac{m}{\sigma}=0 \\ D(Y)=D\left(\frac{1}{\sigma}X-\frac{m}{\sigma}\right)=\left|\frac{1}{\sigma}\right|\cdot\underbrace{D(X)}_{\sigma}=\frac{1}{\sigma}\cdot\sigma=1 \quad (\because \sigma>0) \end{cases}$$ となるからだ。

4. 正規分布と標準正規分布をマスターしよう！

離散型の二項分布 $B(n, p)$ の n を大きくしていくと，近似的に連続型の**正規分布** $N(m, \sigma^2)$ になることがわかっている。今回は，連続型の確率分布の中でも最も重要なこの“**正規分布**”と，その確率変数を標準化して得られる“**標準正規分布**”について，教えるつもりだ。

● 正規分布の確率密度 $f_N(x)$ をマスターしよう！

二項分布 $B(n, p)$ を表す確率分布の関数を $P_B(x)$ とおくと，

$P_B(x) = {}_n\mathrm{C}_x p^x q^{n-x}$ $(x = 0, 1, 2, \cdots, n)$ になるのはいいね。

期待値 $m = E(X) = np$，分散 $\sigma^2 = V(X) = npq$ だったね！

この $P_B(x)$ は，$x = 0, 1, 2, \cdots, n$ と離散的な確率変数の確率分布なんだけれど，ここで n を $50, 100, \cdots$ と十分に大きくとり，そして x を連続的な確率変数とみなすことにより，次のような“**正規分布**”と呼ばれる確率分布になることがわかっている。この正規分布は，その期待値 (平均) m と分散 σ^2 を使って $N(m, \sigma^2)$ と表され，その確率密度を $f_N(x)$ とおくと，$f_N(x)$ は次のように表される。

正規分布 $N(m, \sigma^2)$

正規分布 $N(m, \sigma^2)$ の確率密度 $f_N(x)$ は

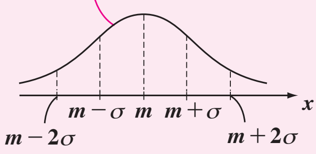
$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad \cdots (*)$$ であり，

(x：連続型の確率変数，$-\infty < x < \infty$)

その期待値と分散は，

$E(X) = m$，$V(X) = \sigma^2$ である。

正規分布の
確率密度
$f_N(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$



初めて，正規分布 $N(m, \sigma^2)$ の確率密度 $f_N(x)$ を見ると，ほとんどの人が「ヒェ～!」ってことになると思う。この $(*)$ の右辺の e は**ネイピア数**と呼ばれる定数で，微分積分では重要な定数なんだけれど，この意味については，数学Ⅲの講義で解説しよう。今は，e は，$e \fallingdotseq 2.72$ の定数であることだけ頭に入れておいてくれたらいいんだよ。

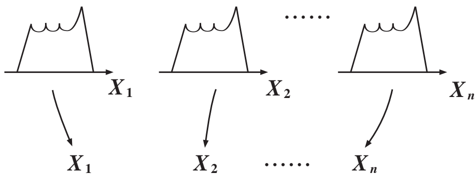
でも，この正規分布 $f_N(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$ のグラフは，平均 m に関して左右対称なキレイなすり鉢(ばち)型になっており，その分散は σ^2（標準偏差は σ）の確率密度関数なんだね。たとえば，たく山の学生がある数学のテストを受けたとき，その得点分布が，この正規分布に近い形になることも，経験的によく知られているんだね。

さらに，理論的な証明は難しいんだけれど，平均 m，分散 σ^2 の同一の確率分布から取り出された n 個の変数 $X_1, X_2, \cdots, X_n$ の相加平均を $\overline{X}=\dfrac{X_1+X_2+\ \cdots\ +X_n}{n}$ とおくと，n が十分大きいとき，この $\overline{X}$ は正規分布 $N\Big(\underbrace{m}_{\text{平均}},\ \underbrace{\dfrac{\sigma^2}{n}}_{\text{分散}}\Big)$ に従うことがわかっている。これを "**中心極限定理**(ちゅうしんきょくげんていり)" という。

本格的に勉強したい人は，「**統計学キャンパス・ゼミ**」（**マセマ**）を読むといいよ。もちろん，大学に入ってからだね。

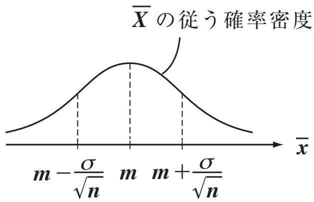
図1　中心極限定理のイメージ

平均 m，分散 σ^2 の n 個の同一の分布



$\overline{X}=\dfrac{X_1+X_2+\ \cdots\cdots\ +X_n}{n}$ とおくと，

$\overline{X}$ は正規分布 $N\left(m, \dfrac{\sigma^2}{n}\right)$ に従う。



このように正規分布 $N(m, \sigma^2)$ は，連続型の確率分布として，非常に重要なものなので，その確率密度 $f_N(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$ を何回も自分で書いて，頭にたたき込んでおくことを勧める。

具体的に，正規分布 $(10, 25)$ の確率密度 $f_N(x)$ は，$m=10$，$\sigma^2=25\ (\sigma=5)$ なので，$f_N(x)=\dfrac{1}{5\sqrt{2\pi}}e^{-\frac{(x-10)^2}{50}}$ となるんだね。これで，少しは慣れたかな？

● 標準正規分布 $N(0,\,1)$ は，数表で示されている！

正規分布 $N(m,\,\sigma^2)$ に従う確率変数 X の平均は $E(X)=m$，標準偏差は $D(X)=\sigma$ となるので，この X を使って，新たな確率変数 Z を $Z=\dfrac{X-m}{\sigma}$ と定義すると，Z は平均が 0，分散が 1（標準偏差が 1）の正規分布に従うことは大丈夫？ つまり，

$$E(Z)=E\left(\frac{1}{\sigma}X-\frac{m}{\sigma}\right)=\frac{1}{\sigma}\overbrace{E(X)}^{m}-\frac{m}{\sigma}=\frac{m}{\sigma}-\frac{m}{\sigma}=0$$

$$V(Z)=V\left(\frac{1}{\sigma}X-\frac{m}{\sigma}\right)=\frac{1}{\sigma^2}\overbrace{V(X)}^{\sigma^2}=\frac{\sigma^2}{\sigma^2}=1$$

（よって，$D(Z)=\sqrt{V(Z)}=1$）　　　　となるんだね。

このように，変数 X から，変数 $Z=\dfrac{X-m}{\sigma}$ に変換することを，確率変数を“**標準化**（ひょうじゅんか）”するといい，この標準化した変数 Z が従う正規分布 $N(0,\,1)$ のことを特に“**標準正規分布**（ひょうじゅんせいきぶんぷ）”と呼ぶ。

これから，標準正規分布 $N(0,\,1)$ の確率密度 $f_S(z)$ は

$$\boxed{f_S(z)=\frac{1}{\sqrt{2\pi}}\,e^{-\frac{z^2}{2}}}$$

となるんだね。

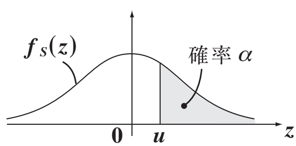
> $f_S(z)=\dfrac{1}{\sqrt{2\pi}\cdot\underbrace{\boxed{1}}_{\sigma}}\,e^{-\frac{(z-\overbrace{\boxed{0}}^{m})^2}{2\cdot\underbrace{\boxed{1}}_{\sigma^2}}}$ と書き換えると，
>
> $m=0,\ \sigma^2=1$ の $N(0,\,1)$ になっていることがわかるね。

どんな正規分布 $N(m,\,\sigma^2)$ に従う変数 X も，$Z=\dfrac{X-m}{\sigma}$ と標準化することにより，標準正規分布 $N(0,\,1)$ $\left(\text{確率密度 } f_S(z)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\right)$ にもち込むことができる。ここで，u をある 0 以上の定数とおくよ。そして，$z\geqq u$ となる確率 $P(z\geqq u)$ を α とおくと，図 2 に示すように，

$$\alpha=P(z\geqq u)=\int_u^{\infty}f_S(z)dz=\frac{1}{\sqrt{2\pi}}\int_u^{\infty}e^{-\frac{z^2}{2}}\,dz$$

となる。だけど，残念ながら，この積分は高校

図 2　標準正規分布における確率 $\alpha=P(z\geqq u)$



数学の範囲では解けないんだね。

でも心配は要らない。試験では **0** 以上の各定数 u の値に対して，確率 $\alpha = P(z \geqq u)$ の値が表 **1** に示すように与えられるので，この表を利用して，実際に問題を解いていけばいいんだよ。

では，例題で練習してみよう！

表 **1**　$\alpha = \int_u^{\infty} f_S(z)dz$ の表

u	$\alpha = P(z \geqq u)$
0.0	**0.5000**
0.1	**0.4602**
0.2	**0.4207**
0.3	**0.3821**
0.4	**0.3446**
0.5	**0.3085**
⋮	⋮

◆ 例題 **13** ◆

確率変数 X が正規分布 $N(2, 16)$ に従うとき，確率 $P(X \geqq 4)$ を求めよ。
(ただし，上記の表 **1** を用いてよいものとする。)

確率変数 X は，正規分布 $N(\overset{m}{2}, \overset{\sigma^2}{16})$ に従うので，これを標準化すると，

$Z = \dfrac{X-2}{4}$ ← $\boxed{Z = \dfrac{X-m}{\sigma}}$ となる。

$X \geqq 4$ となる確率 $P(X \geqq 4)$ は，$X-2 \geqq 4-2$，　$\underbrace{\dfrac{X-2}{4}}_{z} \geqq \dfrac{4-2}{4}$ $\left(\dfrac{1}{2} = u\right)$ より，

$P(z \geqq 0.5)$ に等しい。

よって，表 **1** より，$P(z \geqq 0.5) = 0.3085$ となる。

$P(X \geqq 4) = P(z \geqq 0.5) = 0.3085$ ……………………………………(答)

参考

$N(m, \sigma^2)$ に従う変数 X を

$Y = \dfrac{10(X-m)}{\sigma} + 50$ と変換すると，

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

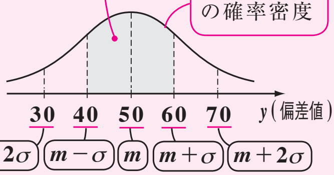
$E(Y) = E\left(\overset{a}{\dfrac{10}{\sigma}}X \overset{b}{- \dfrac{10m}{\sigma} + 50}\right) = \dfrac{10}{\cancel{\sigma}}\overset{m}{\cancel{E(X)}} - \cancel{\dfrac{10m}{\sigma}} + 50 = 50$

$V(Y) = V\left(\dfrac{10}{\sigma}X - \dfrac{10m}{\sigma} + 50\right) = \dfrac{100}{\sigma^2}\overset{\sigma^2}{V(X)} = 100$

$D(Y) = \sqrt{V(Y)} = 10$ となる。得点 X を上記のように Y に変換すると，Y は平均 **50**，標準偏差 **10** となって，俗にいう**偏差値**になる。

偏差値 **40** ～ **60** に含まれる確率は，約 **68.3**% だ。

$N(50, 10^2)$ の確率密度



絶対暗記問題 58	難易度 ★	CHECK1	CHECK2	CHECK3

確率変数 X が，正規分布 $N(1,\ 4)$ に従うとき，確率 $P(2 \leqq X \leqq 4)$ を求めよ。ただし，右の標準正規分布表を使ってよい。

標準正規分布表 $\alpha = \int_u^\infty f_S(z)dz$

u	α
0.5	0.3085
1.5	0.0668

ヒント！ 確率変数 X を標準化して $Z = \dfrac{X-m}{\sigma}$ とおくと，$P(2 \leqq X \leqq 4)$ は，$P(0.5 \leqq z \leqq 1.5)$ で求められることがわかるはずだ。

解答＆解説

確率変数 X は，正規分布 $N(\overset{m}{1},\ \overset{\sigma^2}{4})$，すなわち平均 $m=1$，標準偏差 $\sigma=2$ の正規分布に従うことがわかる。X を標準化して Z で表すと，

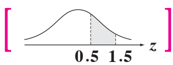
$Z = \dfrac{X-1}{2}$ ← $\boxed{Z = \dfrac{X-m}{\sigma}}$ となる。

ここで，$2 \leqq X \leqq 4$ となる確率 $P(2 \leqq X \leqq 4)$ は，

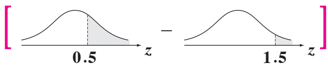
$2-1 \leqq X-1 \leqq 4-1$，$\dfrac{2-1}{2} \leqq \overset{z}{\boxed{\dfrac{X-1}{2}}} \leqq \dfrac{4-1}{2}$ より，

$0.5 \leqq z \leqq 1.5$ となる確率 $P(0.5 \leqq z \leqq 1.5)$ に等しい。

$\therefore P(2 \leqq X \leqq 4) = P(0.5 \leqq z \leqq 1.5)$



$= P(z \geqq 0.5) - P(z \geqq 1.5)$



$= 0.3085 - 0.0668$

$= 0.2417$ ……………………………………………………(答)

標準正規分布における確率（Ⅱ）

絶対暗記問題 59	難易度 ★	CHECK1	CHECK2	CHECK3

確率変数 X が，正規分布 $N(3,\ 100)$ に従うとき，確率 $P(0 \leqq X \leqq 10)$ を求めよ。ただし，右の標準正規分布表を使ってよい。

標準正規分布表 $\alpha = \int_u^{\infty} f_S(z)dz$

u	α
0.3	0.3821
0.7	0.2420

ヒント！ 確率変数 X を標準化した変数 Z を用いると，$P(0 \leqq X \leqq 10) = P(-0.3 \leqq z \leqq 0.7)$ となる。今回は，標準正規分布表を使うとき，標準正規分布のグラフの対称性を利用することがポイントとなるんだよ。

解答＆解説

確率変数 X は，正規分布 $N(\underset{m}{3},\ \underset{\sigma^2}{100})$，すなわち平均 $m = 3$，標準偏差 $\sigma = 10$ の正規分布に従うことがわかる。X を標準化して Z で表すと，

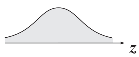
$Z = \dfrac{X-3}{10}$ ← $\boxed{Z = \dfrac{X-m}{\sigma}}$ となる。

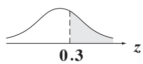
ここで，$0 \leqq X \leqq 10$ となる確率 $P(0 \leqq X \leqq 10)$ は，

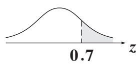
$0-3 \leqq X-3 \leqq 10-3$，$\dfrac{0-3}{10} \leqq \underbrace{\dfrac{X-3}{10}}_{z} \leqq \dfrac{10-3}{10}$ より，

$-0.3 \leqq z \leqq 0.7$ となる確率 $P(-0.3 \leqq z \leqq 0.7)$ に等しい。

$\therefore\ P(0 \leqq X \leqq 10) = P(-0.3 \leqq z \leqq 0.7)$ ［
］

$= 1 - P(z \geqq 0.3) - P(z \geqq 0.7)$ ［

－ 

－ 
］

（注意 表で，u は 0 以上の値に対してしか，確率 α の値を与えてはいないが，標準正規分布は，直線 $z = 0$ に関して左右対称なグラフになっているので，上記のような計算ができるんだね。工夫が大切だ！）

$= 1 - 0.3821 - 0.2420 = 0.3759$ ……………………………（答）