6

無作為抽出と標本分布

6-1　無作為抽出

母集団と標本　調査や実験で観測の対象となる同種の事物の集まりを母集団という．母集団にはそれを構成する要素の数が有限か無限かによって，**有限母集団**と**無限母集団**がある．母集団に関する情報を得るため，それからとり出された母集団の一部分を**標本**という．

無作為抽出　標本から母集団に関する統計的推測を行うためには，標本は母集団の縮図になるようなものでなければならない．そのような標本は無作為抽出によって得られる．**無作為抽出**とは，母集団を構成するすべての要素が等確率で標本のなかに選ばれるような標本抽出の方法である．無作為抽出によって得られた標本を**無作為標本**(または単に**標本**)という．実際の無作為抽出では**乱数表**がよく使われる．

復元抽出と非復元抽出　母集団から標本を抽出するとき，一度とり出したものを元に戻し，次のものをとり出す方法を**復元抽出**といい，とり出したものを元に戻さず，次のものをとり出す方法を**非復元抽出**という．

母数と統計量　母集団におけるある変量 X の確率分布をその変量の**母集団分布**といい，この分布の特性値である平均 μ，分散 σ^2 などを**母数**という．一般に，母集団からとられた大きさ n の無作為標本を表す確率変数 $X_1, X_2, \cdots, X_n$ を**標本変量**という．標本変量 $X_1, X_2, \cdots, X_n$ は互いに独立な n 個の確率変数で，その確率分布はすべて母集団分布と同じである．標本の個数 n を**標本の大きさ**といい，

$$\text{標本平均}\quad \overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$$

標本分散　$s^2=\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$

のような標本変量 $X_1, X_2, \cdots, X_n$ の関数を**統計量**という.

6-2　標本平均の分布

標本平均 $\overline{X}$ の分布に関して, 以下の定理が成り立つ.

$\overline{X}$ の平均と分散

定理 1　平均 μ, 分散 σ^2 の無限母集団からとられた大きさ n の標本の平均を $\overline{X}$ とすれば

$$E(\overline{X})=\mu$$
$$V(\overline{X})=\frac{\sigma^2}{n}$$

定理 2　平均 μ, 分散 σ^2 の有限母集団からとられた大きさ n の標本の平均を $\overline{X}$, 母集団の大きさを N とすれば

$$E(\overline{X})=\mu$$
$$V(\overline{X})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}$$

$\overline{X}$ の標本分布

定理 3　平均 μ, 分散 σ^2 の正規母集団からとられた大きさ n の標本の平均 $\overline{X}$ の分布は, 平均 μ, 分散 $\dfrac{\sigma^2}{n}$ の正規分布に従う.

定理 4（中心極限定理）　平均 μ, 分散 σ^2 のある母集団からとられた大きさ n の標本の平均 $\overline{X}$ の分布は, n が十分大きいならば, 近似的に平均 μ, 分散 $\dfrac{\sigma^2}{n}$ の正規分布に従う.

6-3　χ^2 分布, t 分布, F 分布

χ^2 分布, t 分布, F 分布はいずれも正規母集団からの標本抽出に関連して導出された標本分布である.

χ^2 分布　確率密度関数が

$$f_\nu(x)=ce^{-\frac{x}{2}}x^{\frac{\nu}{2}-1}\qquad(x>0)$$

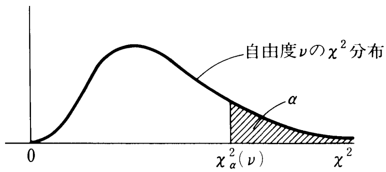
で与えられる分布を**自由度 ν の χ^2 分布**という. ここで, 定数 c は $f_\nu(x)$ が密度関数であるという条件から決まる.

χ^2 分布表　確率変数 X が自由度 ν の χ^2 分布をするとき, 確率 α に対し

て

$$P(X > \chi_\alpha^2(\nu)) = \alpha$$

を満たす $\chi_\alpha^2(\nu)$ の値の表を χ^2 分布表という（付表 5 参照）．



定理 5　平均 μ，分散 σ^2 の正規母集団からとられた大きさ n の標本の分散を s^2 とするとき，

$$\chi^2 = \frac{ns^2}{\sigma^2}$$

は自由度 $\nu = n-1$ の χ^2 分布に従う．

***t* 分布**　確率密度関数が

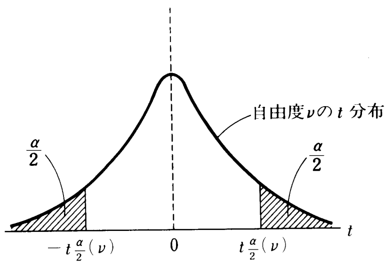
$$g_\nu(x) = c\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \qquad (-\infty < x < \infty)$$

で与えられる分布を自由度 ν の ***t* 分布**という．ここで，定数 c は $g_\nu(t)$ が密度関数であるという条件から決まる．

***t* 分布表**　確率変数 T が自由度 ν の t 分布をするとき，確率 α に対して

$$P(|T| > t_{\frac{\alpha}{2}}(\nu)) = \alpha$$

を満たす $t_{\frac{\alpha}{2}}(\nu)$ の値の表を t 分布表という（付表 4 参照）．



定理 6　正規母集団からとられた大きさ n の標本の平均と分散を，$\bar{x}$，s^2 とするとき，

$$t = \frac{\sqrt{n-1}(\bar{x}-\mu)}{s}$$

の分布は，自由度 $\nu = n-1$ の t 分布に従う．

定理 7　2 つの正規母集団 $N(\mu_1, \sigma^2)$，$N(\mu_2, \sigma^2)$ から独立にとられた大きさ n_1，n_2 の標本の平均と分散を $\bar{x}_1$，$\bar{x}_2$ および $s_1{}^2$，$s_2{}^2$ とする．このとき，

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$$

ここで，

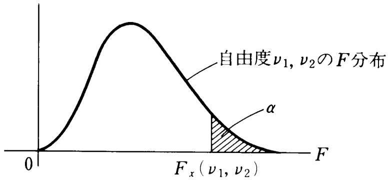
$$s^2 = \frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2}$$

は，自由度 $\nu = n_1 + n_2 - 2$ の t 分布に従う．

F 分布　確率密度関数が

$$h_{\nu_1,\nu_2}(x)=cx^{\frac{\nu_1}{2}-1}\left(1+\frac{\nu_1}{\nu_2}x\right)^{-\frac{\nu_1+\nu_2}{2}}\qquad(x>0)$$

で与えられる分布を自由度 ν_1, ν_2 の F 分布という．ここで定数 c は $h_{\nu_1,\nu_2}(x)$ が密度関数であるという条件から決まる．



F 分布表　確率変数 F が自由度 $(\nu_1,\ \nu_2)$ の F 分布をするとき，確率 α に対して，

$$P(F>F_\alpha(\nu_1,\ \nu_2))=\alpha$$

を満たす $F_\alpha(\nu_1,\ \nu_2)$ の値の表を F 分布表という．本書では，$\alpha=0.05$, 0.025 に対する表を与える(付表 6, 付表 7 参照)．

定理 8　2 つの正規母集団 $N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$ から独立にとられた n_1 個と n_2 個の標本の不偏分散を $u_1{}^2$, $u_2{}^2$ とするとき，

$$F=\frac{u_1{}^2}{u_2{}^2}$$

の分布は，自由度 $(n_1-1,\ n_2-1)$ の F 分布に従う．

例　題

例題 1　(非復元抽出による $\overline{X}$ の標本分布)

5 個の数，$\{1, 2, 3, 4, 5\}$ からなる母集団がある．(a) この母集団の平均と分散を求めよ．(b) この母集団から大きさ 2 の標本を非復元抽出でとり出すとき，すべての可能な標本を列挙せよ．(c) 各標本の平均を求め，標本平均 $\overline{X}$ の標本分布を導け．(d) この分布の平均と分散を求めて，これらの値が公式から求めた値と一致することを示せ．(e) 母集団分布と $\overline{X}$ の標本分布をそれぞれ図示せよ．

解　(a)　母平均：　$\mu=\dfrac{1+2+3+4+5}{5}=\mathbf{3}$

　　母分散：　$\sigma^2=\dfrac{1^2+2^2+3^2+4^2+5^2}{5}-9=\mathbf{2}$

(b)　可能な標本は次に示す ${}_5C_2=10$ 通りである．

標本	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(2, 3)	(2, 4)	(2, 5)	(3, 4)	(3, 5)	(4, 5)
$\bar{x}$	1.5	2	2.5	3	2.5	3	3.5	3.5	4	4.5

（c）　よって，$\overline{X}$ の標本分布は

$\bar{x}$	1.5	2	2.5	3	3.5	4	4.5
$P(\overline{X}=\bar{x})$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

（d）　$\overline{X}$ の標本分布より

$E(\overline{X})=\mathbf{3}$　　（分布の対称性より）

$$V(\overline{X})=1.5^2\times\frac{1}{10}+2^2\times\frac{1}{10}+\cdots+4.5^2\times\frac{1}{10}-3^2=\mathbf{0.75}$$

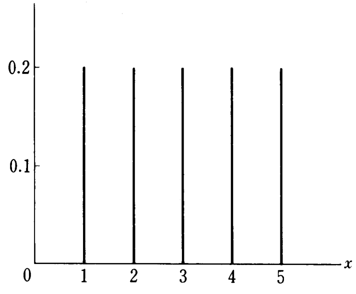
一方，公式からは

$$E(\overline{X})=\mu=3$$

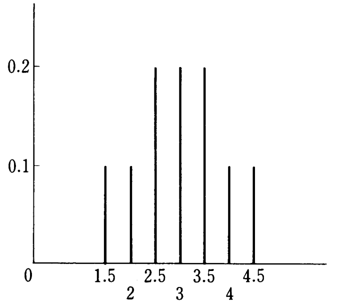
$$V(\overline{X})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}=\frac{5-2}{5-1}\cdot\frac{2}{2}=0.75$$

これらの値は $\overline{X}$ の標本分布から求めた値に一致している．

（e）



母集団分布



$\overline{X}$ の標本分布

例題 2　（平均とメジアンの標本分布）

1, 2, 3 の目がそれぞれ 2 個ずつ記入されたサイコロを 1 回投げるとき，出る目の平均 μ と分散 σ^2 を求めよ．

（a）このサイコロを 3 回投げるとき，（i）出る目の平均，（ii）出る目のメジアン，の標本分布を導け．

（b）　これら標本分布の平均はいずれも μ に等しいことを示し，どちらの分散が小さいかを述べよ．

解　（a）　このサイコロの 1 回の投げの結果を X とすると，X の確率分布は

x	1	2	3
$P(X=x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

よって，$\mu=1\times\frac{1}{3}+2\times\frac{1}{3}+3\times\frac{1}{3}=2$

$$\sigma^2=1^2\times\frac{1}{3}+2^2\times\frac{1}{3}+3^2\times\frac{1}{3}-2^2=\frac{2}{3}$$

サイコロの3回の投げの結果を X_1, X_2, X_3，その標本平均を $\overline{X}$，標本メジアンを M とし，$\overline{X}$ と M の標本値をそれぞれ $\bar{x}$ と m とする．

(X_1, X_2, X_3) のすべての可能な結果と，それに対する $\bar{x}$ の値，および m の値を次の表に示す．

可能な結果 (x_1, x_2, x_3)	$\bar{x}$	m	可能な結果 (x_1, x_2, x_3)	$\bar{x}$	m
(1, 1, 1)	1	1	(3, 3, 1)	$\frac{7}{3}$	3
(1, 1, 2)	$\frac{4}{3}$	1	(3, 1, 3)	$\frac{7}{3}$	3
(1, 2, 1)	$\frac{4}{3}$	1	(1, 3, 3)	$\frac{7}{3}$	3
(2, 1, 1)	$\frac{4}{3}$	1	(3, 3, 2)	$\frac{8}{3}$	3
(1, 1, 3)	$\frac{5}{3}$	1	(3, 2, 3)	$\frac{8}{3}$	3
(1, 3, 1)	$\frac{5}{3}$	1	(2, 3, 3)	$\frac{8}{3}$	3
(3, 1, 1)	$\frac{5}{3}$	1	(1, 2, 3)	2	2
(2, 2, 1)	$\frac{5}{3}$	2	(1, 3, 2)	2	2
(2, 1, 2)	$\frac{5}{3}$	2	(2, 1, 3)	2	2
(1, 2, 2)	$\frac{5}{3}$	2	(2, 3, 1)	2	2
(2, 2, 2)	2	2	(3, 1, 2)	2	2
(2, 2, 3)	$\frac{7}{3}$	2	(3, 2, 1)	2	2
(2, 3, 2)	$\frac{7}{3}$	2	(3, 3, 3)	3	3
(3, 2, 2)	$\frac{7}{3}$	2			

これら27通りの結果の得られる確率はいずれも $\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{27}$ である．

よって，この表より $\overline{X}$ の標本分布と M の標本分布は

（i）

$\bar{x}$	1	$\frac{4}{3}$	$\frac{5}{3}$	2	$\frac{7}{3}$	$\frac{8}{3}$	3
$P(\bar{X}=\bar{x})$	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{7}{27}$	$\frac{6}{27}$	$\frac{3}{27}$	$\frac{1}{27}$

（ii）

m	1	2	3
$P(M=m)$	$\frac{7}{27}$	$\frac{13}{27}$	$\frac{7}{27}$

（b）　これら 2 つの分布はいずれも 2 を中心として対称であるから

$$E(\overline{X})=E(M)=2=\mu$$

$\overline{X}$ と M の分散は，それぞれ

$$V(\overline{X})=1^2\times\frac{7}{27}+\left(\frac{4}{3}\right)^2\times\frac{3}{27}+\cdots+3^2\times\frac{1}{27}-4=\mathbf{\frac{2}{9}}$$

$$V(M)=1^2\times\frac{7}{27}+2^2\times\frac{13}{27}+3^2\times\frac{7}{27}-4=\mathbf{\frac{14}{27}}$$

よって，

$$\boldsymbol{V(\overline{X})<V(M)}.$$

例題 3　（**$\overline{X}$ の標本分布**）

ある工場で生産される電球の寿命は平均 1180 時間，標準偏差 20 時間の正規分布に従う．このとき，次を求めよ．

（a）　25 個の電球の無作為標本の寿命の平均が 1170 時間を超える確率．

（b）　無作為に選んだ n 個の電球の寿命の平均が，少なくとも 0.9 の確率で 1175 時間を超えるといえる n の値．

解　電球の寿命を X とすると

$$X \sim N(1180, 20^2)$$

25 個の電球の寿命を $X_1, X_2, \cdots, X_{25}$，その平均を

$$\overline{X}=\frac{1}{25}\sum_{i=1}^{25} X_i$$

とすれば，

$$E(\overline{X})=1180, \qquad V(\overline{X})=\frac{20^2}{25}=16$$

よって，$\overline{X} \sim N(1180, 4^2)$ であるから

（a） $P(\overline{X}>1170)=P\left(Z>\dfrac{1170-1180}{4}\right)$

$$=1-\Phi(-2.5)=\Phi(2.5)=\mathbf{0.9938}$$

（b） n 個の電球の寿命を $X_1, X_2, \cdots, X_n$，その平均を

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n} X_i$$

とすれば，

$$\overline{X}_n \sim N\left(1180, \frac{20^2}{n}\right)$$

与えられた情報より，

$$P(\overline{X}_n>1175)\geqq 0.90$$
$$P(\overline{X}_n<1175)<0.10$$
$$P\left(Z<\frac{1175-1180}{20/\sqrt{n}}\right)\leqq 0.10$$
$$\Phi\left(-\frac{\sqrt{n}}{4}\right)\leqq 0.10, \qquad \Phi\left(\frac{\sqrt{n}}{4}\right)\geqq 0.90$$

正規分布表より

$$\frac{\sqrt{n}}{4}\geqq 1.28 \quad \Rightarrow \quad n\geqq 26.2$$

よって，**27 個以上**.

例題 4 （**正規変量の 1 結合**）

X_1 と X_2 は独立な確率変数で，

$$X_1 \sim N(5, 2), \qquad X_2 \sim N(3, 1)$$

のとき，$Y=2X_1-X_2$ に対して

（a） $E(Y)$,

（b） $V(Y)$,

（c） $P(Y>8.5)$

を求めよ.

解 （a） $E(Y)=E(2X_1-X_2)=2E(X_1)-E(X_2)=2\times 5-3=\mathbf{7}$

（b） $V(Y)=V(2X_1-X_2)$

$$= 4V(X_1) + V(X_2) \qquad (X_1 \text{ と } X_2 \text{ は独立であるから})$$
$$= 4 \times 2 + 1 = \mathbf{9}$$

（c）正規分布に従う独立な確率変数 X_1, X_2 の一次結合 $a_1X_1 + a_2X_2$ は正規分布に従うから

$$Y \sim N(7, 9)$$

よって，

$$\begin{aligned} P(Y > 8.5) &= 1 - P(Y < 8.5) \\ &= 1 - P\left(Z < \frac{8.5-7}{3}\right) \\ &= 1 - P(Z < 0.5) \\ &= 1 - \Phi(0.5) = 1 - 0.6915 = \mathbf{0.3085} \end{aligned}$$

例題 5　(2 つの平均の差の分布)

確率変数 X と Y は独立で，$X \sim N(2, 1)$，$Y \sim N(3, 2)$ とする．$\overline{X}$ と $\overline{Y}$ は X と Y に関するそれぞれ 5 個の観測値の平均を表す．このとき，次を求めよ．

（a）$E[(X-3)^2]$

（b）$E[(X+2Y)^2]$

（c）$\overline{X} - \overline{Y}$ の分布

（d）$P(\overline{X} > \overline{Y})$

解　（a）

$$\begin{aligned} E[(X-3)^2] &= E[(X-2-1)^2] \\ &= E[(X-2)^2 - 2(X-2) + 1] \\ &= E[(X-2)^2] - 2E(X-2) + 1 \\ &= 1 - 0 + 1 = \mathbf{2} \end{aligned}$$

（b）

$$\begin{aligned} E[(X+2Y)^2] &= E[\{X-2+2(Y-3)+8\}^2] \\ &= E[(X-2)^2] + 4E[(Y-3)^2] + 64 \\ &\quad + 4E[(X-2)(Y-3)] + 16E(X-2) + 32E(Y-3) \end{aligned}$$

仮定より

$$E[(X-2)^2] = 1, \quad E[(Y-3)^2] = 2,$$
$$E(X-2) = 0, \quad E(Y-3) = 0,$$
$$\begin{aligned} E[(X-2)(Y-3)] &= E(X-2)E(Y-3) \qquad (X-2 \text{ と } Y-3 \text{ は独立だから}) \\ &= 0 \end{aligned}$$

であるから,

$$E[(X+2Y)^2]=1+4\times 2+64=\mathbf{73}$$

(c) X, Y は正規変数であるから, $\overline{X}-\overline{Y}$ も正規変数である.

$$E(\overline{X}-\overline{Y})=E(\overline{X})-E(\overline{Y})=2-3=-1$$

$$V(\overline{X}-\overline{Y})=V(\overline{X})+V(\overline{Y})$$

$$=\frac{1}{5}+\frac{2}{5}=\frac{3}{5}$$

よって,

$$\overline{X}-\overline{Y}\sim N\left(\mathbf{-1},\ \boldsymbol{\frac{3}{5}}\right)$$

(d) $P(\overline{X}>\overline{Y})=P(\overline{X}-\overline{Y}>0)$

$$=P\left(Z>\frac{0+1}{\sqrt{3/5}}\right)=1-\Phi(1.291)=1-0.9017=\mathbf{0.0983}$$

例題 6 (正規変量の和)

大型のりんご1個の重さは平均 330 g, 標準偏差 15 g の正規分布に従い, 並型のりんご1個の重さは平均 280 g, 標準偏差 10 g の正規分布に従う. そのとき, 次の確率を求めよ.

(a) 無作為に選んだ大型のりんご3個の重さの合計が 1000 g を超える.

(b) 大型のりんご1個と並型のりんご1個を無作為に選ぶとき, 並型の重さが大型の重さを超える.

(c) 無作為に選んだ並型のりんご5個の重さの合計が, 無作為に選んだ大型のりんご4個の重さの合計を超える.

解 大型のりんご1個の重さを X, 並型のりんご1個の重さを Y とすると

$$X\sim N(330,15^2),\quad Y\sim N(280,10^2)$$

(a) 大型のりんご3個の重さを X_1, X_2, X_3 とすれば, ,$T=X_1+X_2+X_3$ の分布は

$$E(T)=E(X_1+X_2+X_3)$$
$$=E(X_1)+E(X_2)+E(X_3)$$
$$=330+330+330=990$$

$$V(T)=V(X_1+X_2+X_3)$$
$$=V(X_1)+V(X_2)+V(X_3)\qquad (X_1, X_2, X_3 \text{ は独立だから})$$

$$=15^2+15^2+15^2=675$$

より，

$$T\sim N(990,675)$$

よって，

$$P(T>1000)=P\left(Z>\frac{1000-990}{\sqrt{675}}\right)$$

$$\fallingdotseq 1-\Phi(0.385)=1-0.6499=\mathbf{0.3501}$$

（b）$X-Y$ の分布は

$$E(X-Y)=E(X)-E(Y)$$
$$=330-280=50$$
$$V(X-Y)=V(X)+V(Y)$$
$$=15^2+10^2=325$$

より，

$$X-Y\sim N(50,325)$$

よって，

$$P(Y>X)=P(X-Y<0)=P\left(Z<\frac{0-50}{\sqrt{325}}\right)$$
$$=1-\Phi(2.774)=1-0.9972=\mathbf{0.0028}$$

（c）大型りんご 4 個の重さを X_1, X_2, X_3, X_4, その合計を $T_1=\sum_{i=1}^{4}X_i$,

並型りんご 5 個の重さを Y_1, Y_2, Y_3, Y_4, Y_5 その合計を $T_2=\sum_{i=1}^{5}Y_i$

とすれば，仮定より

$$T_1\sim N(1320, 4\times 15^2)$$
$$T_2\sim N(1400, 5\times 10^2)$$

よって，

$$T_2-T_1\sim N(80, 5\times 10^2+4\times 15^2)=N(80, 1400)$$

ゆえに，

$$P(T_2>T_1)=P(T_2-T_1>0)$$
$$=P\left(Z>\frac{0-80}{\sqrt{1400}}\right)$$
$$=1-\Phi(2.138)=1-0.9837=\mathbf{0.0163}$$

例題 7 （$\overline{X}$ の平均と分散）

$X_1, X_2, \cdots, X_n$ を平均 μ，分散 σ^2 の母集団からの大きさ n の無作為標本とし，その平均を $\overline{X}$ とするとき

$$E(\overline{X})=\mu, \qquad V(\overline{X})=\frac{\sigma^2}{n}$$

を示せ．

解 $X_1, X_2, \cdots, X_n$ は同一母集団からの無作為標本であるから，互いに独立で，かつ

$$E(X_1)=E(X_2)=\cdots=E(X_n)=\mu$$
$$V(X_1)=V(X_2)=\cdots=V(X_n)=\sigma^2$$

よって，

$$\begin{aligned}
E(\overline{X})&=E\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\\
&=\frac{1}{n}E(X_1+X_2+\cdots+X_n)\\
&=\frac{1}{n}\{E(X_1)+E(X_2)+\cdots+E(X_n)\}\\
&=\frac{1}{n}(\mu+\mu+\cdots+\mu)\\
&=\mu
\end{aligned}$$

$$\begin{aligned}
V(\overline{X})&=V\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\\
&=\frac{1}{n^2}V(X_1+X_2+\cdots+X_n) \qquad (X_1, X_2, \cdots, X_n \text{ の独立性より})\\
&=\frac{1}{n^2}\{V(X_1)+V(X_2)+\cdots+V(X_n)\}\\
&=\frac{1}{n^2}(\sigma^2+\sigma^2+\cdots+\sigma^2)\\
&=\frac{\sigma^2}{n}
\end{aligned}$$

例題 8（独立な正規変数の和と差）

ある大学の男子学生の体重は平均 60 kg，標準偏差 8 kg の正規分布に従い，女子学生の体重は平均 52 kg，標準偏差 5 kg の正規分布に従うことが知られている．いま，1 人の男子学生と 1 人の女子学生を無作為に選ぶとき，次の確率を求めよ．

（a） 2 人の体重の合計が 100 kg を超える．

（b） 女子学生の方が男子学生よりも重い．

（c） 女子学生の体重が男子学生の体重の少なくとも $\frac{3}{4}$ はある．

解　男子学生の体重を X，女子学生の体重を Y とすると，

$$X \sim N(60, 8^2), \qquad Y \sim N(52, 5^2)$$

（a） $S=X+Y$ とおくと，

$$E(S)=E(X)+E(Y)=60+52=112$$
$$V(S)=V(X)+V(Y)=8^2+5^2=89$$

よって，

$$S \sim N(112, 89)$$

ゆえに

$$\begin{aligned}P(S>100)&=1-P(S<100)\\&=1-P\left(Z<\frac{100-112}{\sqrt{89}}\right)\\&=1-\Phi(-1.272)=\Phi(1.272)=\mathbf{0.8983}\end{aligned}$$

（b） $D=X-Y$ とおくと，

$$E(D)=E(X)-E(Y)=60-52=8$$
$$V(D)=V(X)+V(Y)=89$$

よって，

$$D \sim N(8, 89)$$

ゆえに

$$\begin{aligned}P(Y>X)&=P(X-Y<0)\\&=P\left(Z<\frac{0-8}{\sqrt{89}}\right)\\&=\Phi(-0.848)\\&=1-\Phi(0.848)=1-0.8017=\mathbf{0.1983}\end{aligned}$$

（c）　$W=Y-\frac{3}{4}X$ とおくと

$$E(W)=E(Y)-\frac{3}{4}E(X)=52-\frac{3}{4}\times 60=7$$

$$V(W)=V(Y)+\left(\frac{3}{4}\right)^2 V(X)=25+\frac{9}{16}\times 64=61$$

よって，

$$P\left(Y>\frac{3}{4}X\right)=P(W>0)$$
$$=P\left(Z>\frac{0-7}{\sqrt{61}}\right)$$
$$=\Phi(0.896)=\mathbf{0.8149}$$

例題 9　（独立な正規変数の和）

毎日定刻に自宅を出て会社に通勤しているあるサラリーマンの通勤所要時間（分）は次の3段階からなる．

X：自宅から乗車駅までの歩行時間と下車駅から会社までの歩行時間の合計．

Y：乗車駅での電車の待ち時間．

Z：電車に乗っている時間．

X，Y，Z の平均と分散が

	平均	標準偏差
X	20	1
Y	5	3
Z	40	2

で与えられるとき，次を求めよ．

（a）　このサラリーマンの通勤所要時間の平均と標準偏差．

（b）　X，Y，Z はいずれも正規分布に従うとして，このサラリーマンが自宅を出て1時間以内に会社に着く確率．

解　通勤所要時間を W とすると

$$W=X+Y+Z$$

ゆえに，

（a）$E(W)=E(X+Y+Z)$

$=E(X)+E(Y)+E(Z)$

$=20+5+40=\mathbf{65}$（分）

$V(W)=V(X+Y+Z)$

$=V(X)+V(Y)+V(Z)$　（X, Y, Z は互いに独立とみなして）

$=1^2+3^2+2^2=14$

よって,

$$\sqrt{V(W)}=\sqrt{14}=\mathbf{3.74}\ (\text{分})$$

（b）W は平均 65, 標準偏差 3.74 の正規分布に従うから

$$P(W<60)=P\left(Z<\frac{60-65}{\sqrt{14}}\right)=1-\Phi(1.336)\fallingdotseq\mathbf{0.0908}$$

例題 10（中心極限定理）

ある種の繊維 1 本の破断強度は平均 1.2 kg, 標準偏差 0.1 kg の正規分布に従う．この繊維 100 本の束に 122.5 kg の荷重をかけるとき，束が破壊される確率を求めよ．

解　繊維 100 本の破断強度を $X_1, X_2, \cdots, X_{100}$ とすると

$$X_i \sim N(1.2, 0.1^2) \qquad (i=1, 2, \cdots, 100)$$

100 本の束の破断強度を Y とすると

$$Y=X_1+X_2+\cdots+X_{100}$$

よって，Y は

$$E(Y)=E(X_1)+E(X_2)+\cdots+E(X_{100})=100\times1.2=120$$

$$V(Y)=V(X_1)+V(X_2)+\cdots+V(X_{100})=100\times0.1^2=1$$

の正規分布に従うから，

$$P(Y>122.5)=P\left(Z>\frac{122.5-120}{1}\right)=1-\Phi(2.5)=\mathbf{0.006}$$

［注］　この問題では，$n=100$ より中心極限定理が使えるので，繊維の破断強度が**正規分布**に従うとの仮定がなくても解ける．

例題 11（中心極限定理）

乱数表から 1 桁の乱数を 50 個とるとき，その平均が 4 と 5 の間にある確率を求めよ．

解 乱数の分布は離散型一様分布

x	0	1	2	3	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

で，この分布の平均と分散は

$$\mu=0\times\frac{1}{10}+1\times\frac{1}{10}+\cdots+9\times\frac{1}{10}=4.5$$

$$\sigma^2=0^2\times\frac{1}{10}+1^2\times\frac{1}{10}+\cdots+9^2\times\frac{1}{10}-4.5^2=8.25$$

である．とられた50個の乱数はこの分布をもつ母集団からの大きさ50の無作為標本とみなされるから，50個の標本の平均を $\overline{X}$ とすれば

$$E(\overline{X})=\mu=4.5$$

$$V(\overline{X})=\frac{\sigma^2}{n}=\frac{8.25}{50}=0.165$$

$n=50$ は大きいから，中心極限定理によって，$\overline{X}$ は近似的に $N(4.5, 0.165)$ に従う．よって，

$$P(4<\overline{X}<5)=P\left(\frac{4-4.5}{\sqrt{0.165}}<Z<\frac{5-4.5}{\sqrt{0.165}}\right)$$

$$=2\Phi(1.231)-1=2\times0.8909-1=\mathbf{0.7818}$$

6章の問題

6.1 サイコロを2回投げるとき，出た目の和の標本分布を求め，この分布の平均と分散を求めよ．

6.2 1の目を3つ，2の目を2つ，3の目を1つもつサイコロを1回投げるとき，(a) 出る目 X の確率分布を示し，(b) その平均と分散を求めよ．(c) このサイコロを7回投げるとき，出る目の平均 $\overline{X}$ の平均と分散を求めよ．

6.3 X と Y は独立な確率変数で，

$$X\sim N(5,9),\qquad Y\sim N(7,16)$$

のとき，

(a) $X-Y$,

（b）　$3X+Y$,

（c）　$\overline{X}$,

（d）　$\overline{X}-\overline{Y}$

の分布を求めよ．ただし，$\overline{X}$ は $N(5, 9)$ からとった大きさ 25 個の無作為標本の平均で，$\overline{Y}$ は $N(7, 16)$ からとった大きさ 16 個の無作為標本の平均である．

6.4　ある中学校の男子 3 年生の身長の分布は近似的に平均 163.0 cm，標準偏差 8 cm の正規分布に従い，女子 3 年生の身長の分布は近似的に平均 155.5 cm，標準偏差 6 cm の正規分布に従う．そのとき，次の確率を求めよ．

（a）　男子生徒 2 人を無作為に選ぶとき，2 人の身長の差が 5 cm を超える．

（b）　女子生徒 2 人を無作為に選ぶとき，2 人の身長の差が 5 cm を超える．

（c）　男子生徒 1 人と女子生徒 1 人を無作為に通ぶとき，2 人の身長の差が 10 cm を超える．

6.5　高校 3 年生の身長は次のような正規分布に従うことが知られている．

男子：　169.3 cm，標準偏差 4.6 cm

女子：　156.6 cm，標準偏差 4.2 cm

これら 2 つの正規母集団からとった大きさ 64 と 36 の標本平均をそれぞれ $\overline{X}$，$\overline{Y}$ とするとき，次の値を求めよ．

（a）　$\overline{X}$，$\overline{Y}$ の平均と標準偏差．

（b）　$\overline{X}-\overline{Y}$ の平均と標準偏差．

（c）　$\overline{X}$ が $\overline{Y}$ を少なくとも 12 cm 超える確率．

6.6　桃の缶詰の中味の重さは平均 300 g，標準偏差 1.2 g の正規分布に従い，空き缶の重さは平均 20 g，標準偏差 0.5 g の正規分布に従う．このとき，次の値を求めよ．

（a）　桃の缶詰 1 個の重さの平均と標準偏差．

（b）　無作為に選んだ桃の缶詰 1 個の重さが 323 g を超える確率．

6.7　鋼板にあけた穴にボルトをはめこむ．ボルトの直径は平均 2.80 cm，標準偏差 0.03 cm の正規分布に従い，穴の直径は平均 2.90 cm，標準偏差 0.04 cm の正規分布に従うとき，次を求めよ．

（a） ボルトと穴をそれぞれ無作為に選んではめるとき，ボルトが穴にはまらない確率.

（b） 無作為にとった5個のボルトがすべて穴にはまる確率.

6.8 正規母集団 $N(\mu, \sigma^2)$ からとられた大きさ n の無作為標本を $X_1, X_2, \cdots, X_n$ とするとき，

$$Y = X_1 + 2X_2 + 3X_3 + \cdots + nX_n$$

はどんな分布に従うか.

6.9 （a） 平均 μ，分散 σ^2 の正規母集団からの大きさ4の無作為標本の平均を $\overline{X}$ とするとき，確率

$$P(|\overline{X} - \mu| < \sigma)$$

の値を求めよ.

（b） 平均80，標準偏差10の母集団から大きさ64の無作為標本をとるとき，標本平均 $\overline{X}$ が82を超える確率を求めよ.

6.10 密度関数

$$f(x) = 6x(1-x) \qquad (0 \leqq x \leqq 1)$$

を分布にもつ母集団から大きさ45の標本がとられた. 標本平均を $\overline{X}$ とするとき，次を求めよ.

（a） $\overline{X}$ の標本分布.

（b） $P(\overline{X} < 0.55)$.

（c） $P(\overline{X} > 0.46)$.

6.11 小数点以下を四捨五入することで，測定値をそれに最も近い整数値に丸めるとき，丸めの誤差 X は区間 $\left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$ で一様分布に従う確率変数とみなされる. n 個の測定値をそれぞれそれに最も近い整数値に丸めるとき，これら丸め誤差の平均 $\overline{X}$ の平均と分散を求めよ. n は十分大きいと仮定して，$\overline{X}$ の絶対値が $\dfrac{1}{2\sqrt{n}}$ を超えない確率を求めよ.