
統計研修受講前問題集

統計研究研修所では、統計の理論と演習による幅広い研修を行なっています。当所での統計の学習のためには数学の知識が必要となりますが、統計分析で用いる数学は必ずしも中学校や高等学校で学んだ数学をそのまま使うわけではありません。

近年、いろいろな事情から、研修期間が短くなる傾向にあります。そのため、当所の講義時間だけでは、統計分析に必要な基礎知識を習得するのに必要かつ十分な時間を確保することが難しい状態です。このような現状を踏まえ、統計研修を受講する前に統計分析の基礎知識の一端を知ることが、講義内容の理解に大きく役立つものと考え、本問題集を作成しました。

本問題集の解説は簡単なものであり、多くの重要な説明は割愛しております。受講生の皆様におかれては、本問題集が講義を理解するための最初のステップと理解され、本問題集に書かれた項目などを統計関係書籍やインターネットなどで調べ、一層の統計分析のための知識をもって受講されることを切望します。

なお、本問題集の一部の事項は、課程によっては講義されないことがあることをご了承ください。

目 次

1. 統計の基礎	1
(1) Σ (総和) 記号	1
(2) 増加率	2
(3) 平均値	3
ア 算術平均	3
イ 加重算術平均 (加重平均)	4
(4) 対数	6
(5) 標準偏差と分散	7
2. 関数のグラフ	8
3. 場合の数	10
4. 行列	12

1. 統計の基礎

(1) Σ (総和) 記号

数学や統計で用いられる記号は、数式を簡単に表すために用いられます。

Σ (「シグマ」という。) は、数列のすべての数を合計 (総和の計算) するという意味で使用される記号です。

数列 “ $a_1, a_2, a_3, \dots, a_n$ ” の総和 “ $a_1 + a_2 + a_3 + \dots + a_n$ ” を求める式を、 Σ 記号を用い

て表すと、“ $\sum_{k=1}^n a_k$ ” になります。 ($a_1 + a_2 + a_3 + \dots + a_n = \sum_{k=1}^n a_k$)

(注1) Σ 記号の下の “ $k=1$ ” は、「 k 」が Σ 記号の右の「 a 」の順番を表し、「 $=1$ 」は1番目の「 a 」から順に加えるという意味です。

(注2) Σ 記号の上の “ n ” は、 Σ 記号の下の文字 (この場合 k) について n 番目まで加えるという意味です。

(注3) Σ 記号の右の “ a_k ” は加える対象の「 a 」と順番を表す「 k 」です。

【例】 Σ 記号を用いた算式の例です。 Σ 記号が意味するところを考えてください。

$$\sum_{k=3}^8 a_k = a_3 + a_4 + a_5 + a_6 + a_7 + a_8 \qquad \sum_{k=2}^4 \frac{2k}{n} = \frac{2 \cdot 2}{n} + \frac{2 \cdot 3}{n} + \frac{2 \cdot 4}{n}$$

$$\sum_{k=1}^n (a_k + b_k) = \sum_{k=1}^n a_k + \sum_{k=1}^n b_k \qquad \sum_{k=1}^n c \cdot a_k = c \sum_{k=1}^n a_k$$

$$\sum_{k=1}^n c = c + c + \dots + c = c \cdot n \quad (c \text{ は定数})$$

$$\sum_{k=1}^n k = 1 + 2 + 3 + \dots + n = \frac{1}{2} n (n+1) \quad (\text{等差数列の和})$$

$$\sum_{k=1}^n k^2 = 1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{1}{6} n (n+1)(2n+1)$$

$$\sum_{k=1}^n r^{k-1} = 1 + r + r^2 + \dots + r^{n-1} = \frac{r^n - 1}{r - 1} \quad (\text{等比数列の和})$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_{n-1} + x_n}{n} = \frac{1}{n} \sum_{k=1}^n x_k \quad (\text{平均を求める式})$$

【問題 1】

1. 次の式を、 Σ 記号を使わないで記述しなさい。

$$(1) \sum_{k=1}^5 a_{3k}$$

$$(2) \sum_{k=1}^4 (a_k + b_k)$$

2. 「表 1」の 10 人の身長合計と平均を求める式を、 Σ 記号を用いて記述しなさい。

3. 「表 1」の 10 人の身長合計値と平均値を計算しなさい。

表 1 身長

番号	1	2	3	4	5
身長 (cm)	158	167	150	169	160
番号	6	7	8	9	10
身長 (cm)	159	152	162	156	164

(2) 増加率

2 時点間の増減を求めるときに、増加率（「増減率、上昇率など」ともいう。）を計算します。率で表すのは、比較する時点が変わっても同じものさしで評価できるためです。

$$r = \left(\frac{x_i}{x_0} - 1 \right) \times 100$$

年率計算の式は、次のとおりです。

$$r = \sqrt[N]{x_1 \times x_2 \times \cdots \times x_{N-1} \times x_N}$$

【問題 2】

「表 2」から、5 年間の人口増加率を計算しなさい。

表 2 人口数

年	2000年	2005年
人口 (人)	121,338,000	125,437,000

(3) 平均値

統計分析とは、バラバラなデータの集まり（データの分布）から、手法を用いて特性などを読み取ることです。データの全体を代表する値としては、一般に、平均値が使われます。“計算による平均”には、算術平均、加重算術平均、幾何平均、調和平均があります。また、平均値以外にデータ全体を代表する値として、中央値と最頻値があります。

ア 算術平均

算術平均は、平均値の計算方法の中で最も多く使われる計算方法で、データをすべて足した総和をデータの総数で割って計算します。

また、平均値は、“算術平均の値からデータの値を引いた値の総和は「0」になる”という性格を持っています。

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_{n-1} + x_n}{n} = \frac{1}{n} \sum_{k=1}^n x_k$$

【問題 3】

「表 3」の 10 人の身長と体重の、それぞれの合計値と平均値を計算しなさい。

表 3 身長と体重

番号	1	2	3	4	5
身長(cm)(X)	158	167	150	169	160
体重(kg)(Y)	55	57	45	64	50
番号	6	7	8	9	10
身長(cm)(X)	159	152	162	156	164
体重(kg)(Y)	60	53	49	44	57

(参考 1) 度数とヒストグラム

度数とは、データをそれぞれの階級に振り分けたときの階級内のデータの数をいう。

ヒストグラムとは、度数を棒の面積で表したものです。横軸に変数（階級幅別）、縦軸に度数（目盛り）をとって描きます。

イ 加重算術平均 (加重平均)

加重算術平均は、データに重み (ウェイト) を付けて計算する算術平均です。

実際には、消費者物価指数、企業物価指数、賃金指数計算などに使われている計算方法です。

$$\bar{x} = \frac{w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + \cdots + w_{n-1} \times x_{n-1} + w_n \times x_n}{w_1 + w_2 + \cdots + w_{n-1} + w_n} = \frac{\sum_{i=1}^n w_i \times x_i}{\sum_{i=1}^n w_i}$$

【問題 4】

「表 4」(A) の国語の平均点及び (B) の牛肉 100 グラム当たりの平均価格を計算しなさい。

表 4 (A) 国語の試験結果

番号	1	2	3	4	5	6	7	8	9	10
国語(点)	42	48	49	50	52	55	57	68	80	88
人数(人)	1	1	1	2	3	8	3	4	2	1

表 4 (B) 牛肉価格と購入量

店番号	1	2	3	4	5
牛肉価格(円/100g)	385	415	305	515	990
購入量(g)	300	250	650	700	250
店番号	6	7	8	9	10
牛肉価格(円/100g)	684	680	840	755	485
購入量(g)	560	600	350	400	350

(参考 2) 幾何平均 → 「(4) 対数」参照

(参考 3) 調和平均

調和平均は、データを逆数変換して、計算する平均値です。

$$\frac{1}{H} = \frac{1}{N} \left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_N} \right)$$

また、次の式に置き換えられます。

$$H = \frac{N}{\sum \frac{1}{x}}$$

(参考4) 移動平均

移動平均は、時系列分析に用いられる平均値です。

時系列分析では、時間の経過に伴うデータの変化から将来の傾向を予測するため、分析に使用するデータから月々の違いや季節による違いなどを取り除く必要があります。この季節変動を除去するために、一定期間を定めて算術平均し、期間をずらして季節変動をならした平均値を求める計算方法が、移動平均です。

例えば、12か月移動平均の場合、1月～12月までの算術平均を計算し、次に2月～翌年1月までの算術平均、・・・、12月～翌年11月までの算術平均、翌年1月～12月までの算術平均というように計算する方法です。

(参考5) 中央値 (中位数, メジアン)

中央値は、データを一定の順に並べた (例えば、小から大へと) ときに、全データを真ん中で二つに分ける境界点の値です。したがって、中央値は、度数分布の面積を2等分する垂直線の座標ともいえます。データの数が奇数個のときは

$\frac{N+1}{2}$ 番目が、偶数個のときは中央の2つのデータの間値、 $\frac{N}{2}$ と $\frac{N}{2}+1$

の間値が中央値になります。

(参考6) 四分位数, 十分位数

中央値の仲間として、四分位数や十分位数などがあります。

四分位数は、データを一定の順に並べ (例えば、小から大へと)、4分の1ずつに区切ったときの境界点です。25%点を第1・四分位数、50%点 (中央値) を第2・四分位数、75%点を第3・四分位数といいます。また、十分位数は、10%で区切ったときの境界点をいい、第1・十分位数 (10%点)、第2・十分位数 (20%点)、・・・、第9・十分位数 (90%点) といいます。

(参考7) 最頻値 (並み数, モード)

最頻値は、度数分布でいえば最大度数となるデータ又はその階級をいいます。

最頻値は、データの観察においてそのデータやデータが含まれる階級が最大であるという分かりやすさがある反面、変数の数が少なく度数分布の形が整わない場合は説明が難しいため、使用されません。

(4) 対数

統計分析では、数値を対数に変換してから分析を行うことがよくあります。対数変換とは、例えば、 $10,000=10^4$ の場合、 $10,000$ を 4 に変換させるもので、 $4 = \log_{10}10,000$ と表します。 10 の平方根 $\sqrt{10} = 10^{1/2} = 3.162\dots$ の対数は $1/2$ で、 $1/2 = \log_{10}10^{1/2}$ と表します。ここで、 \log_{10} の 10 を対数の底 (てい) といい、 10 を底とする対数を常用対数といいます。このほか $e = 2.71828\dots$ を底とする自然対数もよく使われます。底の表記はしばしば省略されますが、常用対数と自然対数のどちらを用いているか注意する必要があります。(Excel では、それぞれ LOG, LN という関数が対応しています。)

対数は、掛け算を足し算にする性質があります。

$$\log(X \times Y) = \log X + \log Y$$

この性質を利用して年率計算や幾何平均の算出に用いられます。

$$G = \sqrt[N]{x_1 \times x_2 \times \dots \times x_{N-1} \times x_N}$$

上記の N 個の数値 $x_1, x_2, \dots, x_{N-1}, x_N$ の幾何平均について、対数をとると、以下のよう
に $x_1, x_2, \dots, x_{N-1}, x_N$ それぞれの対数の算術平均になります。

$$\log G = \frac{1}{N} (\log x_1 + \log x_2 + \dots + \log x_{N-1} + \log x_N) = \frac{1}{N} \sum_{k=1}^N \log x_k$$

【例】 A市の地価は4年間で毎年12.1%、10.9%、8.6%、7.2%値上がりしました。4年間の平均上昇率を計算しなさい。

$$(答え) \quad r = \sqrt[4]{1.121 \times 1.109 \times 1.086 \times 1.072} = \sqrt[4]{1.4473} = 1.0968$$

A市の地価の4年間の平均上昇率 9.7%

(注) 計算方法は、解答集を参照してください。

【問題5】

「表5」から、5年間の人口増加率の年率を計算しなさい。

表5 人口数

年	2000年	2005年
人口(人)	121,338,000	125,437,000

(5) 標準偏差と分散

データの分布の状態は、データの分散度（ばらつきの大きさ）を測定することで分かります。データの分散度を測定する方法としては、最大値から最小値を引いた値を用いる「レンジ」、平均値に対するデータの偏差（偏り）を測定する「平均偏差」、「標準偏差」、「分散」がありますが、一般には、「分散」と「標準偏差」がよく使われます。

標準偏差は、平均値に対する偏差の2乗を平均し、これをデータと同じ次元の単位にするために平方に開いたもので、“ s ”で表わします。

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{N}} \quad (\text{標準偏差の計算式})$$

これを平方に開く前のもの（標準偏差の2乗値に相当）を分散といい“ s^2 ”で表します。

$$s^2 = \frac{\sum (x - \bar{x})^2}{N} \quad (\text{分散の計算式})$$

なお、統計数理では、ばらつきの測定は標準偏差よりも分散を使う場合が多いのですが、分散は標準偏差の2乗値であるため、分散の値にデータと同じ単位（例えば、円、kgなど）をつけることはできません。

【問題6 標準偏差, 分散】

次の数列の、平均値、分散、標準偏差を計算しなさい。

5 0 5 1 5 5 5 6 5 8

(参考8) 標準偏差が使用される理由

標準偏差は、数学的には最もすぐれたばらつきの測定方法といえます。それは、変数の平均値に対する偏差の大きさを、2乗（正負の符号をなくすため）して平均し、これを平方に開いて元の単位で示しているため、結果の意味が容易に理解できるからです。

(参考9) 平均偏差

度数分布のばらつきをデータの平均値に対する偏差の和で表すと「0」になってしまうため、ばらつきを考える場合は、偏差の方向を無視して偏差の大きさを測ります。そのときに、偏差の絶対値を平均する方法が平均偏差です。

$$M.D = \frac{1}{N} \sum |x - \bar{x}|$$

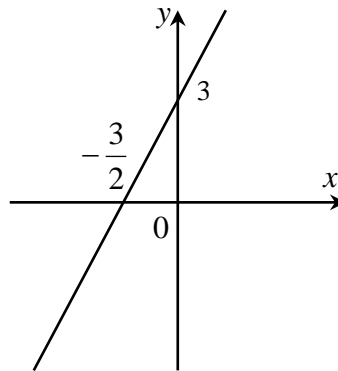
2. 関数のグラフ

関数の変化の様子や方程式を解く場合は、グラフにするとわかりやすいことが多くあります。

【例1：直線のグラフ】

$y = 2x + 3$ のグラフ
(傾き：2, y 切片：3,
 x 切片： $-\frac{3}{2}$)

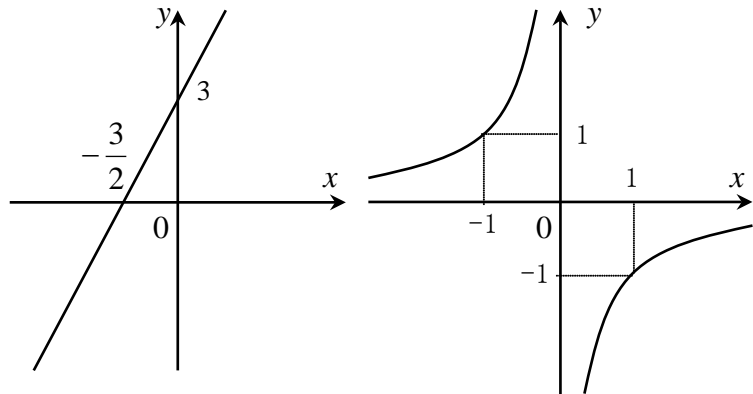
例1：直線のグラフ



例2：双曲線のグラフ

【例2：双曲線のグラフ】

$y = -\frac{1}{x}$ のグラフ
(漸近線： $x = 0$, $y = 0$)

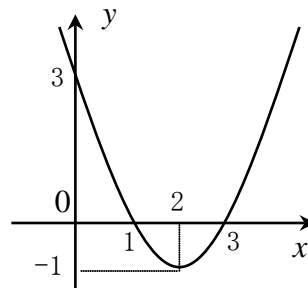


【例3：放物線のグラフ】

$y = x^2 - 4x + 3 = (x - 2)^2 - 1$ のグラフ

例3：放物線のグラフ

(軸： $x = 2$, 頂点： $(2, -1)$)



【問題7 関数のグラフ】

次の関数のグラフを描きなさい。

1. $y = -3x - 6$

2. $y = \frac{1}{2}x + 3$

(参考) 関数のグラフと方程式の関係

方程式 $f(x)=0$ の解は、関数 $y=f(x)$ と x 軸の交点の x 座標になります。

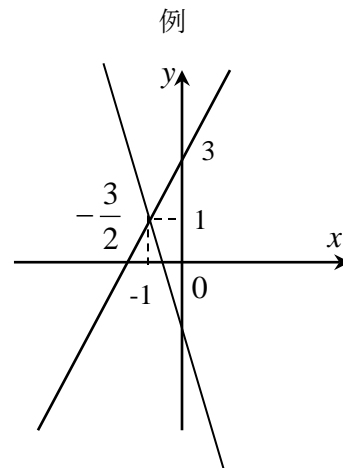
連立方程式 $\begin{cases} y=f(x) \\ y=g(x) \end{cases}$ の解は、関数 $y=f(x)$ と関数 $y=g(x)$ の交点の座標

(すなわち、両方の関数を同時に満たす x, y の値の組) になります。

【例】

連立方程式 $\begin{cases} y=2x+3 \\ y=-3x-2 \end{cases}$ の解 $\begin{cases} x=-1 \\ y=1 \end{cases}$ は、

一次関数 $y=2x+3$ と $y=-3x-2$ の交点の座標 $(-1,1)$ になります。



3. 場合の数

『和の法則』により「AとBは同時には起こらないとすると、AまたはBの起こる場合の数は $m+n$ 通り」に、また『積の法則』により「AとBの起こり方が互いに影響しない」とすると、AとBがともに起こる場合の数は $m \times n$ 通り」となります。

【例1】二つのさいころを同時に投げたとき、上になった目の数の和が5の倍数になるのは、いくつありますか。

(答え) 7通り

目の数の和が5の倍数になるのは、「5」と「10」の2通りがあります。和が「5」になる場合は4通り、「10」になる場合は3通りですが、これらは同時に起こらないので、『和の法則』により $4+3=7$ になります。

【例2】電車でA駅からB駅まで行くのに2とおり、徒歩でB駅から目的地まで行くのに3通りの方法がありますが、A駅から目的地まで行くのにいくつの方法がありますか。

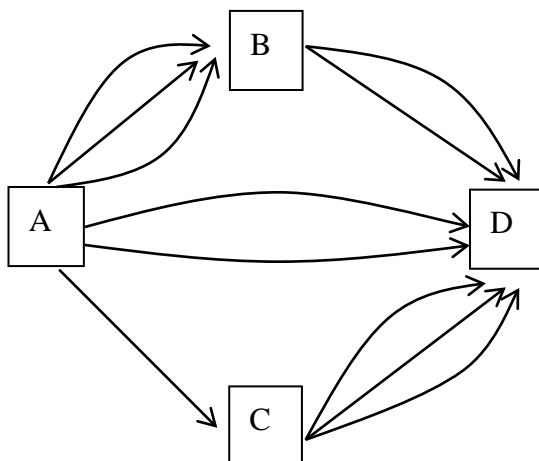
(答え) 6通り

『積の法則』により $2 \times 3 = 6$ になります。

【問題8】

下図のとおり、何本かの道でつながるA、B、C、Dの4つの町があります。同じ町を2回通らないでA町からD町へ行く方法はいくつありますか。

図 A町からD町へ行く道路



(参考 1 1) 順列

異なる n 個のものを一列に並べる方法は、

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \quad \text{とおりです (} n \text{ の階乗)}。$$

異なる n 個のものから r 個を選んで一列に並べる方法は、

$${}_n P_r = \underbrace{n \times (n-1) \times (n-2) \times \cdots \times (n-r+2) \times (n-r+1)}_{r \text{ 個の積}} = \frac{n!}{(n-r)!} \text{ 通りです。}$$

ここで、 $0! = 1$ と定義すると、異なる n 個のものから n 個を選んで一列に並べる方法の数は ${}_n P_n = n!$ となります。

エクセル関数 “PERMUT” を使うと計算できます。

(参考 1 2) 組合せ

異なる n 個のものから r 個を選ぶ方法は、

$${}_n C_r = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r(r-1)(r-2)\cdots 1} = \frac{n!}{r!(n-r)!} \text{ 通りです。}$$

ここで、 $0! = 1$ と定義すると、 ${}_n C_0 = {}_n C_n = 1$ となります。

エクセル関数 “COMBIN” を使うと計算できます。

4. 行列

行列 A とスカラー k に対して、スカラー倍 kA が定義できます。

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ のとき, } kA = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}$$

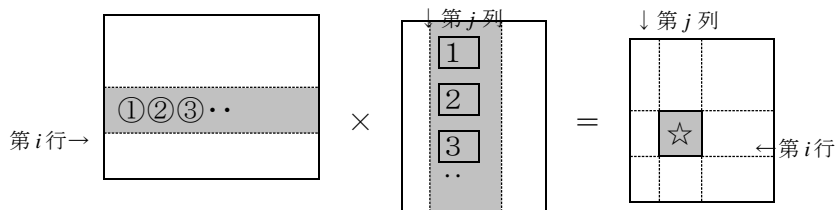
また、同じ型の行列 A, B に対して、和 $A+B$ が定義できます。

$$A = \text{上記}, B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \text{ のとき, } A+B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

積 AB は、 A の列数と B の行数が等しいときに初めて定義できます。このときの AB の行数 A と、列数 B は同じ数でなければなりません。

(l 行 m 列の行列) \times (m 行 n 列の行列) = (l 行 n 列の行列)

計算方法： $\star = \textcircled{1} \times \boxed{1} + \textcircled{2} \times \boxed{2} + \textcircled{3} \times \boxed{3} + \cdots$



【問題 9 行列】

1. 次の計算をしなさい。

$$2 \times \begin{bmatrix} 3 & -2 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix}$$

2. 次の計算をしなさい。

$$\begin{bmatrix} 6 & 4 & 5 \\ 5 & 3 & 0 \end{bmatrix} \times \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

3. 行列 $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ の逆行列を求めなさい。

(ヒント)

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{かつ} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ となる行列 } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ を}$$

求めます。