

2027年度 大学入学共通テスト 予想問題κ

試験開始の指示があるまで、この問題冊子の中を見てはいけません。

情 報 [『情報Ⅰ』] (100 点 60 分)

I 注 意 事 項

- 1 この問題冊子は、35 ページあります。
- 2 解答用紙に、正しく記入・マークされていない場合は、採点できないことがあります。
- 3 試験中に問題冊子の印刷不鮮明、ページの落丁・乱丁及び解答用紙の汚れ等に気付いた場合は、手を高く挙げて監督者に知らせなさい。
- 4 問題冊子の余白等は適宜利用してよいが、どのページも切り離してはいけません。
- 5 試験終了後、問題冊子は持ち帰りなさい。

II 解 答 上 の 注 意

解答上の注意は、裏表紙に記載してあります。問題冊子を裏返して必ず読みなさい。

(下 書 き 用 紙)

情報 I の試験問題は次に続く。

(下 書 き 用 紙)

情報 I の試験問題は次に続く。

情 報 I

(全 問 必 答)

第 1 問 次の問い(問 1 ～ 4)に答えよ。(配点 20)

問 1 次の文章を読み、空欄

ア	イ	ウ
---	---	---

，

エ	オ	カ
---	---	---

，

キ	ク	ケ
---	---	---

に当てはまる数字をマークせよ。また、空欄

コ

に入れるのに最も適当なものを、後の解答群のうちから一つ選べ。

通信の誤りを検出する方法の一つに **CRC**(巡回冗長検査)がある。CRC では、桁上がりや桁借りを行わない 2 を法とする計算でビット列の割り算を行う。この計算では、各桁の足し算と引き算はどちらも排他的論理和(XOR)になる。送信側は、データの後ろに検査用のビットを付け加え、全体が決められたビット列(**生成多項式**に対応するビット列)で割り切れるようにして送信する。

ここでは、生成多項式に対応するビット列を 1011 とし、次の手順で検査用の 3 ビットを求める。

- 手順 1 データの後ろに 0 を 3 個付け加えたビット列を作る。

手順 2 手順 1 のビット列を 1011 で割ったときの余り(3 ビット)を求める。割り算は、残っているビット列の先頭の 1 の位置に 1011 の先頭をそろえて排他的論理和をとる操作を、そろえる位置がなくなるまで繰り返せばよい。

手順 3 データの後ろに、手順 2 で求めた余りを付け加えたものを送信する。

例えば，データが 1110 の場合は図 1 のようになり，送信するビット列は 1110100 となる。

$$\begin{array}{rcl}
 \text{割られる数} & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \leftarrow \text{データ1110の後ろに0を3個付けたもの} \\
 \text{XOR} & 1 & 0 & 1 & 1 & & & & \\
 \hline
 & & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
 \text{XOR} & & 1 & 0 & 1 & 1 & & & \\
 \hline
 & & & 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 & \leftarrow \text{余り (100)}
 \end{array}$$

図 1 データ 1110 の場合の割り算

データが 1100101 のとき，送信するビット列の末尾の 3 ビットは **ア****イ****ウ** となる。受信側では，受信したビット列全体を 1011 で割り，余りが 000 でなければ誤りがあると判断する。このデータを送信したところ，受信したビット列は 1100001010 であった。これを 1011 で割った余りは **エ****オ****カ** となるので，誤りがあることがわかる。また，受信したビット列が 1100101110 であった場合は，余りは **キ****ク****ケ** となる。この余りは，送信したビット列の先頭の 1 ビットだけが反転した場合の余りと同じであり，余りだけから誤りの位置を一つに決めることはできない。

ただし，この方法でも検出できない誤りがある。例えば，送信した 10 ビットのうち **コ** 場合は，誤りがあっても余りが 000 となる。

コ の解答群

- ① 1 ビットだけが反転した
- ② 隣り合う 2 ビットが反転した
- ③ 先頭から 1 ビット目と 8 ビット目の 2 ビットが反転した
- ④ 連続する 3 ビットがすべて反転した

問 2 次の文章を読み、空欄

サ	シ	ス
---	---	---

、

セ	ソ	タ
---	---	---

、

チ	ツ
---	---

 に当てはまる数字をマークせよ。また、空欄

テ

 に入れるのに最も適当なものを、後の解答群のうちから一つ選べ。

公開鍵暗号方式の一つである RSA 暗号では、次の手順で鍵を作る。

- 手順 1 二つの異なる素数 p , q を選び, $n = p \times q$ とする。

手順 2 $L = (p - 1) \times (q - 1)$ とし, L との最大公約数が 1 である整数 e を選ぶ。

手順 3 $e \times d$ を L で割った余りが 1 となる正の整数 d ($d < L$) を求める。

手順 4 e と n の組を公開鍵, d を秘密鍵とする。

平文 m ($0 \leq m < n$ の整数) は, m^e を n で割った余り c として暗号化する。 c^d を n で割った余りを求めると, m に戻る。

$p = 11$, $q = 13$, $e = 7$ とすると, $L =$

サ	シ	ス
---	---	---

 であり, $d =$

セ	ソ	タ
---	---	---

 となる。また, 平文 $m = 9$ を暗号化すると, $c =$

チ	ツ
---	---

 となる。なお, 9^7 のような大きな数を直接計算する必要はない。 9^2 , 9^4 を順に求めながら, そのたびに n で割った余りに置き換えて計算すればよい。

RSA 暗号の安全性に関する記述として最も適当なものは

テ

 である。

テ の解答群

- ① e が公開されているので、 d は誰でも簡単に計算できるが、平文 m が秘密であるので安全である
- ② n が大きくても、 L は公開鍵から簡単に求めることができる
- ③ 同じ平文を同じ公開鍵で暗号化すると、毎回異なる暗号文が得られるので、平文を推測されることはない
- ④ n を p と q に素因数分解できれば秘密鍵 d を求められるが、 n が十分大きければ、素因数分解を現実的な時間で行うことができない

問 3 次の文章を読み、空欄ト～ネに当てはまる数字をマークせよ。

ある会社のデータベースには、表1の「部署」表と「社員」表がある。社員表の部署コードが空欄(NULL)である社員は、まだどの部署にも配属されていない。

表1 部署表(左)と社員表(右)

部署コード	部署名	社員番号	氏名	部署コード	給与
D1	営業	E01	青木	D1	30
D2	開発	E02	井上	D1	26
D3	総務	E03	上田	D2	40
D4	広報	E04	遠藤	D2	34
		E05	大野	D2	28
		E06	加藤	D3	25
		E07	木村		27

SQL では、NULL との比較(=, <>, >など)の結果は、真でも偽でもない「不明」となる。WHERE 句では、条件が真となる行だけが残る。また、集約関数 COUNT(*)は行の数を数え、COUNT(列名)はその列が NULL でない行の数を数える。AVG(列名)は NULL を除いて平均を求める。

```
SQL1  SELECT B. 部署名, COUNT(*) AS 人数
        FROM 部署 B LEFT OUTER JOIN 社員 S
        ON B. 部署コード = S. 部署コード
        GROUP BY B. 部署名

SQL2  SELECT COUNT(*) FROM 社員 WHERE 部署コード <> 'D1'

SQL3  SELECT COUNT(*) FROM 部署 B
        WHERE NOT EXISTS (SELECT * FROM 社員 S
                           WHERE S. 部署コード = B. 部署コード)

SQL4  SELECT COUNT(*) FROM 部署
        WHERE 部署コード NOT IN (SELECT 部署コード FROM 社員)

SQL5  SELECT COUNT(*) FROM 社員 S
        WHERE S. 給与 > (SELECT AVG(T. 給与) FROM 社員 T
                           WHERE T. 部署コード = S. 部署コード)
```

SQL1 の結果のうち、部署名が広報である行の人数の値は となる。また、SQL2 の結果は 、SQL3 の結果は 、SQL4 の結果は 、SQL5 の結果は となる。

問 4 次の文章を読み、空欄 , , , , , に当てはまる数字をマークせよ。また、空欄 に入れるのに最も適当なものを、後の解答群のうちから一つ選べ。

あるシステムでは、四つのプロセス P1 ～ P4 が 2 種類の資源 R1, R2 を共有している。システム全体の資源の数は、R1 が 10 個、R2 が 7 個である。各プロセスは、実行中に必要とする資源の最大数(最大要求数)をあらかじめ申告し、資源を少しずつ要求する。現在の割当て数と最大要求数は表 2 のとおりである。各プロセスは、最大要求数まで資源を割り当てられれば処理を完了し、割り当てられていた資源をすべて返却する。

表 2 各プロセスの割当て数と最大要求数

プロセス	P1	P2	P3	P4
割当て数(R1, R2)	2, 1	3, 2	2, 1	1, 1
最大要求数(R1, R2)	5, 3	4, 4	6, 2	3, 3

OS は、資源の要求に応じても、すべてのプロセスが何らかの順序で処理を完了できる状態(安全な状態)が保たれる場合にだけ要求に応じ、そうでなければ要求したプロセスを待たせる(銀行家のアルゴリズム)。こうすることで、プロセスどうしが互いに資源の返却を待ち続けるデッドロックを避けることができる。

現在、どのプロセスにも割り当てられていない資源は、R1 が 個、R2 が 個である。P3 が処理を完了するまでにさらに必要な R1 の数は 個である。また、現在の状態から、必要な資源をすべて直ちに割り当てて処理を完了させることができるプロセスは 個ある。例えば、P2 に必要な資源を割り当て、P2 が処理を完了して資源をすべて返却すると、割り当てられていない資源は R1 が 個、R2 が 個となる。

現在の状態で次の要求があったとき，応じても安全な状態が保たれるものは

マである。

マの解答群

- ① P1 が R2 を 1 個要求する
- ② P3 が R1 と R2 を 1 個ずつ要求する
- ③ P4 が R1 を 2 個要求する
- ④ P1 が R1 を 2 個と R2 を 1 個要求する

第2問 次の問い(A・B)に答えよ。(配点 30)

A 次の文章を読み、後の問い(問1～5)に答えよ。

X社の社内LANでは、図1のように5台のスイッチSW1～SW5を接続している。線上の数値は、そのリンクのコストを表す。故障に備えて経路を冗長にしているが、このままではフレームがループ状に転送され続けるおそれがある。そこで、**スパニングツリープロトコル(STP)**を用いて、一部のポートをフレームの転送に使わない状態(**ブロッキング**)にし、ループのない木構造の経路を作っている。

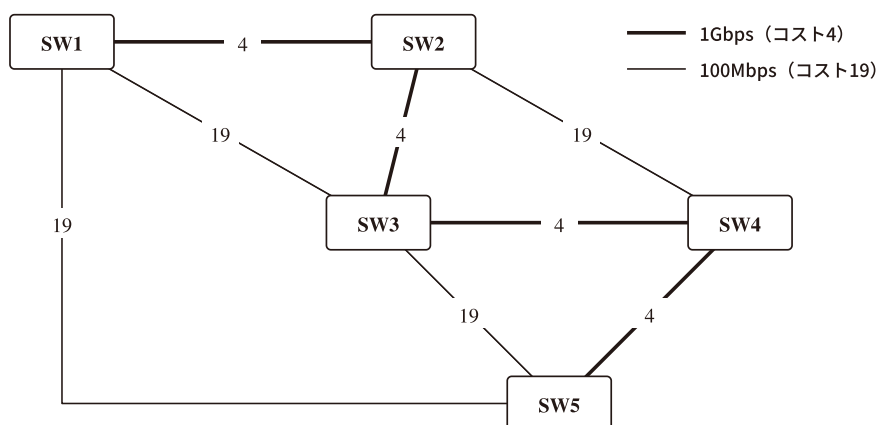


図1 スイッチの接続

表1 各スイッチの優先度と MAC アドレス

スイッチ	優先度	MAC アドレス
SW1	32768	00-00-5E-00-53-21
SW2	4096	00-00-5E-00-53-35
SW3	32768	00-00-5E-00-53-12
SW4	32768	00-00-5E-00-53-30
SW5	8192	00-00-5E-00-53-44

各スイッチは、優先度と MAC アドレスを組み合わせた**ブリッジ ID**をもつ。二つのブリッジ ID は、まず優先度を比べ、優先度が等しい場合は MAC アドレスを比べて、値が小さい方を小さいとする。STP は、次の規則で各ポートの役割を決める。

- 規則1 ブリッジ ID が最小のスイッチを**ルートブリッジ**とする。
- 規則2 各スイッチについて、ルートブリッジまでの経路上のリンクのコストの合計の最小値を、そのスイッチの**ルートパスコスト**とする(ルートブリッジでは0とする)。
- 規則3 ルートブリッジ以外の各スイッチでは、ルートブリッジまでのコストの合計が最小となる経路の最初のリンクにつながるポートを**ルートポート**とする。そのような経路が複数ある場合は、最初のリンクの接続先のスイッチのブリッジ ID が小さい方を選ぶ。
- 規則4 各リンクについて、両端のスイッチのうちルートパスコストが小さい方(等しい場合はブリッジ ID が小さい方)のポートを**指定ポート**とする。
- 規則5 ルートポートでも指定ポートでもないポートを**ブロッキング**にする。

問 1 ルートブリッジとなるスイッチとして最も適当なものを，次の①～④のうちから一つ選べ。 ア

- ① SW1 ② SW2 ③ SW3 ④ SW4 ⑤ SW5

問 2 次の文章中の空欄 イ，ウエ に当てはまる数字をマークせよ。

SW4 のルートパスコストは イ，SW5 のルートパスコストは ウエ である。

問 3 次の文章を読み，空欄 オ に当てはまる数字をマークせよ。また，空欄 カ に入れるのに最も適当なものを，後の解答群のうちから一つ選べ。

ブロッキングになるポートは全部で オ 個ある。また，SW1 と SW3 を結ぶリンクについての記述として最も適当なものは カ である。

カ の解答群

- ① SW1 側のポートがブロッキングになる
- ② SW3 側のポートがブロッキングになる
- ③ 両端のポートがともにブロッキングになる
- ④ 両端のポートがともにフレームを転送する

問 4 次の文章を読み、空欄 **キ****ク**、**ケ****コ**、**シ** に当てはまる数字をマークせよ。また、空欄 **サ** に入れるのに最も適当なものを、後の解答群のうちから一つ選べ。

SW3 と SW4 を結ぶリンクが故障した。STP によって各ポートの役割が決め直された後、SW4 のルートパスコストは **キ****ク**、SW5 のルートパスコストは **ケ****コ** となり、SW5 のルートポートは **サ** となる。また、ブロッキングになるポートは全部で **シ** 個となる。

サ の解答群

- ① SW1 とのリンクにつながるポート
- ② SW3 とのリンクにつながるポート
- ③ SW4 とのリンクにつながるポート
- ④ なし(ルートポートは決まらない)

問 5 STP に関する記述として最も適当なものを、次の①～④のうちから一つ選べ。 **ス**

- ① ブロッキングのポートがあるので、冗長な経路を用意していても、リンクの故障時に通信を継続することはできない。
- ② STP を用いると、すべてのリンクを同時にフレームの転送に使うので、通信速度が経路の数に比例して向上する。
- ③ ポートの役割は、ルートブリッジからのコストには関係なく、MAC アドレスの大小だけで決まる。
- ④ ループがあると、ブロードキャストフレームが転送され続けて増え、ネットワークが使えなくなることがあるので、STP でループを防ぐ。

B 次の文章を読み、後の問い(問 1 ～ 5)に答えよ。

Web サービスを運営する W 社では、不正なアクセスに使われた端末の ID (8 ビット) を記録したリストを持っている。アクセスのたびにリストを検索すると時間がかかるので、あらかじめ**ブルームフィルタ**と呼ばれる仕組みで判定し、「リストに含まれる可能性がある」と判定された場合だけリストを検索することにした。

ここで用いるブルームフィルタは、位置 0 ～ 15 の 16 ビットからなるビット列 F と、二つのハッシュ関数 h_1 , h_2 で構成される。ID を v とするとき、 $h_1(v)$ は v の下位 4 ビットを、 $h_2(v)$ は v の上位 4 ビットを、それぞれ 0 ～ 15 の整数とみなした値である。

登録 F の初期値はすべて 0 とする。ID v を登録するときは、 F の位置 $h_1(v)$ と位置 $h_2(v)$ のビットを 1 にする。

判定 ID v を判定するときは、 F の位置 $h_1(v)$ と位置 $h_2(v)$ のビットを調べ、両方とも 1 であれば「含まれる可能性がある」、そうでなければ「含まれない」と判定する。

以下では、ID を 16 進数で表す。例えば、ID 3A では、 $h_1 = 10$ (A は 10 を表す)、 $h_2 = 3$ である。また、 F は、位置 15 を最上位、位置 0 を最下位とする 16 ビットの数とみなす。

問 1 次の文章中の空欄 ，， に当てはまる数字または文字をマークせよ。

ID 3A，5C，A7，93 の四つを登録した。このとき， F のうち 1 であるビットは 個であり， F を 16 進数で表すと となる。また，仮にこの四つに加えて ID 6E も登録したとすると， F を 16 進数で表すと となる。ただし，16 進数の各桁は，解答欄の 0 ～ 9 および a ～ f のうちから一つずつマークせよ。

問 2 次の文章を読み，空欄 に当てはまる数字をマークせよ。また，空欄 に入れるのに最も適当なものを，後の解答群のうちから一つ選べ。

ID 3A，5C，A7，93 の四つだけを登録した状態で，00 から FF までの 256 個の ID をすべて判定すると，「含まれる可能性がある」と判定される ID は，登録した四つを含めて全部で 個ある。また，登録していない ID 5A，4C，97，2B のうち，「含まれる可能性がある」と判定されるものは である。

- の解答群
- | | |
|-----------|--------------------|
| ① 5A のみ | ④ 5A と 97 |
| ② 4C と 2B | ③ 5A，4C，97，2B のすべて |

問 3 ID v について，位置 $h1(v)$ と位置 $h2(v)$ のビットだけが 1 である 16 ビット
 の数を M とする。判定の条件「両方とも 1」を，ビットごとの論理積
 (AND)，論理和(OR)，排他的論理和(XOR)を用いて表したものととして最も
 適当なものを，次の①～④のうちから一つ選べ。ただし，「 $=$ 」は両辺の値が
 等しいことを表す。 ハ

① $(F \text{ OR } M) = M$

② $(F \text{ AND } M) \neq 0$

③ $(F \text{ XOR } M) = 0$

④ $(F \text{ AND } M) = M$

問 4 次の文章を読み，空欄 ヒフ，ヘ，ホマ に当てはまる数字をマ
 クせよ。

ブルームフィルタでは，登録していない ID が「含まれる可能性がある」と
 判定されること(偽陽性)があるが，登録した ID が「含まれない」と判定され
 ることはない。ビット列のビット数を m ，ハッシュ関数の数を k ，登録した
 ID の数を n とし，ハッシュ関数の値が一様にばらつくとみなすと，偽陽性
 の確率はおおよそ $(1 - e^{-kn/m})^k$ となる。ここで， e は自然対数の底であり，必
 要であれば表 1 の値を用いよ。

表 1 e^{-x} の値

x	0.25	0.5	0.75	1	1.25
e^{-x}	0.779	0.607	0.472	0.368	0.287

$m=16$, $k=2$, $n=4$ のとき, 偽陽性の確率はおよそ $0.\boxed{\text{ヒ}}\boxed{\text{フ}}$ (小数第 3 位を四捨五入) である。 $m=16$, $n=4$ のまま k を 1 から 5 までの整数で変えたとき, 偽陽性の確率が最も小さくなるのは $k=\boxed{\text{へ}}$ のときである。

また, 100 万個の ID を登録し, k を最適に選んで偽陽性の確率を 1 % にするには, およそ $m = -n \times \ln 0.01 \div (\ln 2)^2$ ビットが必要であることが知られている。 $\ln 0.01 \approx -4.61$, $\ln 2 \approx 0.693$ とすると, 必要なビット列の大きさは約 $\boxed{\text{ホ}}.\boxed{\text{マ}}$ M バイト (1M バイト = 10^6 バイト, 小数第 2 位を四捨五入) となる。

問 5 ブルームフィルタに関する記述として最も適当なものを, 次の①～③のうちから一つ選べ。 $\boxed{\text{ミ}}$

- ① 登録した ID を削除するには, その ID の位置 $h1(v)$ と位置 $h2(v)$ のビットを 0 に戻せばよい。
- ① 登録する ID の数が増えるほど, 偽陽性の確率は小さくなる。
- ② 同じビット数と同じハッシュ関数を用いる二つのブルームフィルタは, ビットごとの論理和をとることで, 両方の ID を登録したブルームフィルタと同じものにできる。
- ③ ビット列 F から, 登録された ID の一覧を一通りに復元することができる。

第3問 次の文章を読み、後の問い(問1～6)に答えよ。(配点 25)

Mさんの学校では、七つの拠点0～6を光ファイバで結ぶ計画を立てている。図1は、配線できる拠点の組と、その配線の費用(単位は十万円)を表したものである。Mさんは、すべての拠点どうしが(他の拠点を経由してもよいので)通信できるようにしつつ、費用の合計が最小となる配線の組合せ(最小全域木)を求めたい。

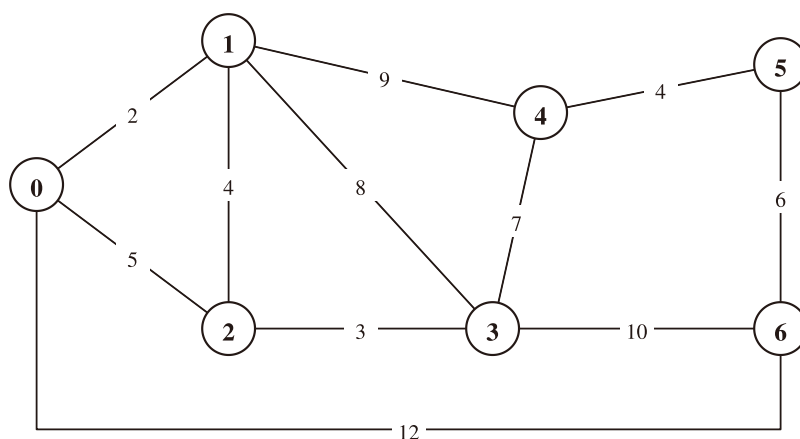


図1 配線の候補と費用

Mさんは、次の方法(クラスカル法)を用いることにした。

- 手順1 配線の候補を費用の小さい順に並べる(費用が等しい場合は表1の順とする)。
- 手順2 候補を順に一つずつ調べ、その候補を採用しても、採用した配線だけで閉じた経路(閉路)ができなければ採用し、閉路ができるならば採用しない。
- 手順3 すべての候補を調べ終わったら終了する。

表 1 費用の小さい順に並べた配線の候補

j	0	1	2	3	4	5	6	7	8	9	10
u[j]	0	2	1	4	0	5	3	1	1	3	0
v[j]	1	3	2	5	2	6	4	3	4	6	6
w[j]	2	3	4	4	5	6	7	8	9	10	12

問 1 次の文章を読み、空欄 **ア** **イ**，**エ**，**オ** **カ** に当てはまる数字をマークせよ。また、空欄 **ウ** に入れるのに最も適当なものを、後の解答群のうちから一つ選べ。

この方法で求めた配線の費用の合計は **ア** **イ** (十万円)である。また、最初に「採用しない」と判断される候補は **ウ** であり、採用されない候補は全部で **エ** 個ある。なお、拠点 3 と拠点 4 を結ぶ候補が工事の都合で使えなくなった場合、同じ方法で求めた費用の合計は **オ** **カ** (十万円)となる。

ウ の解答群

- | | |
|--------------|--------------|
| ① 拠点 0 と拠点 2 | ① 拠点 5 と拠点 6 |
| ② 拠点 1 と拠点 3 | ③ 拠点 3 と拠点 4 |

問 2 次の文章を読み、空欄 **キ** ～ **ケ** に入れるのに最も適当なものを、後の解答群のうちから一つずつ選べ。

Mさんは、閉路ができるかどうかを効率よく判定するため、拠点をグループに分けて管理することにした。最初は、各拠点がそれぞれ一つのグループである。候補の両端の拠点が同じグループに属していれば、その候補を採用すると閉路ができる。採用した場合は、二つのグループを一つにまとめる。

各グループは、グループ内の一つの拠点(代表)を根とする木で表す。配列 **Oya** の要素 **Oya[i]** には拠点 i の親を格納し、代表の拠点は自分自身を親とする。親を順にたどれば代表が見つかる。また、配列 **Ookisa** の要素 **Ookisa[i]** には、拠点 i が代表であるグループの拠点の数を格納する。二つのグループをまとめるときは、拠点の数が少ない方の代表の親を、多い方の代表にする(等しい場合は、**v[j]** の側の代表の親を **u[j]** の側の代表にする)。図 2 はこの方法で費用の合計を求めるプログラムであり、配列 **u**, **v**, **w** には表 1 の値が格納されている。

```

(01) Oya = [0, 1, 2, 3, 4, 5, 6]
(02) Ookisa = [1, 1, 1, 1, 1, 1, 1]
(03) goukei = 0
(04) j を 0 から 要素数 (W) - 1 まで 1 ずつ増やしながら繰り返す:
(05) | a = U[j]
(06) | Oya[a] != a の間, 繰り返す:
(07) | | a = キ
(08) | b = V[j]
(09) | Oya[b] != b の間, 繰り返す:
(10) | | b = Oya[b]
(11) | | もし ク ならば:
(12) | | | もし Ookisa[a] < Ookisa[b] ならば:
(13) | | | | Oya[a] = b
(14) | | | | Ookisa[b] = Ookisa[b] + Ookisa[a]
(15) | | | そうでなければ:
(16) | | | | Oya[b] = a
(17) | | | | Ookisa[a] = ケ
(18) | | goukei = goukei + W[j]
(19) 表示する (goukei)

```

図2 費用の合計を求めるプログラム

キ

の解答群

- ① U[j] ② Oya[b] ③ a + 1 ④ Oya[a]

ク

の解答群

- ① a == b ② a != b
 ③ Oya[a] == Oya[b] ④ U[j] != V[j]

ケ の解答群

- | | |
|-------------------------|-------------------------|
| ① Ookisa[b] | ① Ookisa[a] + 1 |
| ② Ookisa[a] + Ookisa[b] | ③ Ookisa[a] - Ookisa[b] |

問 3 次の文章中の空欄 **コ**，**サ**，**シ**，**ス**，**セ****ソ** に当てはまる数字をマークせよ。

図 2 のプログラムを実行すると， j が 5 の処理を終えた時点で，**Oya**[3] の値は **コ**，**Oya**[6] の値は **サ** であり， j が 6 の処理を終えた時点で，**Oya**[4] の値は **シ** である。すべての処理を終えた時点で **Ookisa**[0] の値は **ス** であり，(07) 行目と (10) 行目が実行される回数の合計は **セ****ソ** である。

問 4 図 2 のプログラムで，拠点の数が少ない方のグループを多い方のグループの下にまとめる理由として最も適当なものを，次の①～③のうちから一つ選べ。

タ

- ① 拠点の数が多い方のグループを少ない方の下にまとめると，木の高さ（根からの段数）が低くなり，代表を見つける処理が速くなるから。
- ② まとめ方によって，求められる最小全域木の費用の合計が変わるので，正しい答えを得るためにはこのまとめ方が必要だから。
- ③ このまとめ方をすると，どの木の高さも，拠点の数を N として $\log_2 N$ 程度以下に抑えられるので，代表を見つける処理が速くなるから。
- ④ まとめ方を工夫しても処理の速さは変わらないが，配列 **Ookisa** を使わないと閉路を正しく判定できないから。

問 5 次の文章を読み、空欄 ・ に入れるのに最も適当なものを、後の解答群のうちから一つずつ選べ。また、空欄 に当てはまる数字をマークせよ。

代表を見つけた後に、たどった拠点の親をすべて代表に付け替えておく（経路圧縮）、次からはより速く代表を見つけられる。そこで、図 2 の (07) 行目の後に、(06) 行目と同じ字下げで次の処理を追加する。

```
x = U[j]
x != a の間、繰り返す：
    | tsugi = Oya[x]
    | Oya[x] = 
    | x = 
```

・ の解答群

- | | | |
|---------|----------|--------------|
| ① a | ① b | ② x |
| ③ tsugi | ④ Oya[a] | ⑤ Oya[tsugi] |

例えば、拠点 0 ～ 4 について **Oya**=[0, 0, 1, 2, 3]であるとき、拠点 4 の代表を求めて経路圧縮を行うと、**Oya** の値が変化する拠点は 個である。

問 6 次の文章中の空欄 ， に当てはまる数字をマークせよ。

最小全域木の考え方は，データを似たもの同士のグループに分けること(クラスタリング)にも使える。図 2 のプログラムで，グループの数が 3 になった時点で処理を止めると，拠点は三つのグループに分かれる。このとき，最も多くの拠点を含むグループの拠点数は であり，異なるグループに属する拠点どうしを直接結ぶ候補のうち，費用が最小のものの費用は (十万円)である。

第4問 次の文章を読み、後の問い(問1～6)に答えよ。(配点 25)

ECサイトを運営するV社では、購入手続きの画面を改良した。新しい画面Bが現在の画面Aより購入率(訪問者のうち購入した人の割合)を高めるかを調べるため、訪問者を無作為に二つに分け、一方には画面A、もう一方には画面Bを表示する実験(A/Bテスト)を14日間行った。表1はその結果である。

表1 A/Bテストの結果

画面	訪問者数	購入者数	購入率
A	4000	200	5.0%
B	4000	248	6.2%

Vさんは、この差が偶然によるものかどうかを**仮説検定**で調べることにした。「画面AとBの真の購入率は等しい」という仮説(**帰無仮説**)のもとで、次の値 Z を計算する。

$$Z = (\text{Bの購入率} - \text{Aの購入率}) \div \text{SE}$$

ここで、SEは、二つの画面を合わせた購入率を p として、 $\sqrt{\{p(1-p) \times (1 \div 4000 + 1 \div 4000)\}}$ で求められ、この実験ではおよそ0.00514である。帰無仮説が正しければ、 Z はおよそ標準正規分布に従う。表2は、標準正規分布に従う変数の絶対値が各値以上となる確率である。

表2 標準正規分布に従う変数の絶対値が z 以上となる確率

z	1.64	1.96	2.33	2.58	3.29
確率	0.10	0.05	0.02	0.01	0.001

問 1 次の文章を読み，空欄 ， に当てはまる数字をマークせよ。また，空欄 に入れるのに最も適当なものを，後の解答群のうちから一つ選べ。

二つの画面を合わせた購入率 p は $0.\text{アイウ}$ であり， Z の値は (小数第 3 位を四捨五入) となる。この結果の解釈として最も適当なものは である。

の解答群

- ① 画面 B の真の購入率が画面 A より高い確率は，およそ 98% である
- ② 帰無仮説が正しいとすると，これ以上の大きさの差が偶然に生じる確率はおよそ 2% であり，有意水準 5% では有意な差があるといえるが，有意水準 1% では有意な差があるとはいえない
- ③ 有意水準 1% でも，有意な差があるといえる
- ④ 帰無仮説が正しい確率は，およそ 2% である

問 2 次の文章を読み、空欄 、 に当てはまる数字をマークせよ。
また、空欄 に入れるのに最も適当なものを、後の解答群のうちから一つ選べ。

購入率の差 (B-A) の 95 % 信頼区間は、(B の購入率 - A の購入率) $\pm 1.96 \times \text{SE}$ で求められる。SE としておよそ 0.00514 を用いると、95%信頼区間は約 ポイントから約 ポイントまでとなる(小数第 2 位を四捨五入)。この信頼区間に関する記述として最も適当なものは である。

の解答群

- ④ 真の差は、必ず 0.2 ポイント以上である
- ① 信頼区間の幅は、訪問者数をそれぞれ 4 倍にすると 4 分の 1 になる
- ② 信頼区間が 0 を含まないので、画面 B の効果は実務上も十分に大きいといえる
- ③ 信頼区間が 0 を含まないことは、有意水準 5 % の検定で有意な差があるという結論と矛盾しない

問 3 次の文章を読み，空欄 **ス** **セ** **ソ** **タ** に当てはまる数字をマークせよ。また，空欄 **チ** に入れるのに最も適当なものを，後の解答群のうちから一つ選べ。

Vさんは，次の実験に必要な訪問者数を見積もることにした。購入率が p_1 から p_2 に変わる効果を，有意水準 5 % (両側検定)，検出力 80 % で見つけるために必要な各画面の訪問者数 n は，次の式で見積もることができる。

$$n = (1.96 + 0.84)^2 \times \{p_1(1 - p_1) + p_2(1 - p_2)\} \div (p_1 - p_2)^2$$

$p_1 = 0.05$ ， $p_2 = 0.06$ とすると， n は約 **ス** **セ** **ソ** **タ** 人 (小数第 1 位を切り上げ) となる。また， p_2 を 0.055 に変えて，検出したい差を 0.5 ポイントにすると，必要な訪問者数は **チ** 。

チ の解答群

- | | |
|------------|-------------|
| ① 約半分になる | ① ほとんど変わらない |
| ② 約 2 倍になる | ③ 約 4 倍になる |

問 4 次の文章を読み、空欄 , に当てはまる数字をマークせよ。また、空欄 ・ に入れるのに最も適当なものを、後の解答群のうちから一つずつ選べ。

V 社では別の施策として、二つのクーポン C、D の効果を比べた。この施策では、訪問者を無作為に分けずに、PC とスマートフォン(SP)の利用者に、それぞれ担当者が選んだクーポンを配布した。表 3 はその結果である。

表 3 クーポンの配布数と購入者数

端末	クーポン	配布数	購入者数
PC	C	200	30
PC	D	800	104
SP	C	800	40
SP	D	200	8

PC での購入率は C が 15%、D が 13% であり、SP での購入率は C が 5%、D が 4% である。一方、PC と SP を合わせた全体の購入率は、C が %、D が % となる。このように、どちらの端末でも C の購入率の方が高いのに、全体では D の購入率の方が高くなっている。この理由として最も適当なものは である。また、二つのクーポンの効果を正しく比べる方法として最も適当なものは である。

の解答群

- ① 購入率の高い PC の利用者に、D が多く配布されていたこと
- ② D の方が、クーポンとしての効果が大きいこと
- ③ 配布数の合計が、C と D で異なること
- ④ SP の購入者数が少なく、端末ごとの購入率が信頼できないこと

ネ の解答群

- ① 全体の購入率だけを比べる
- ② 購入者数の合計が多い方を，効果が大きいと判断する
- ③ 訪問者を無作為に C と D に割り当てる実験を行うか，端末ごとに分けて比べる
- ④ 配布数が多い方のクーポンを優先して配布し，その結果を比べる

問 5 次の文章を読み，空欄 に当てはまる数字をマークせよ。また，空欄 に入れるのに最も適当なものを，後の解答群のうちから一つ選べ。

Vさんは，画面 A と B の比較で，購入率以外にも滞在時間や再訪問率など 20 種類の指標について，それぞれ有意水準 5 % で検定を行おうと考えた。どの指標にも本当は差がない場合でも，20 回の検定のうち少なくとも一つで有意な差があると判定される確率は，各検定が独立であるとする $1 - 0.95^{20}$ であり， $0.95^{20} \approx 0.358$ を用いると約 0. となる。このような問題への対処として最も適当なものは である。

ヒ の解答群

- ① 有意水準を検定の数で割った値(この場合は 0.0025)を，各検定の有意水準として用いる
- ② 有意な差が出た指標だけを報告し，それ以外の指標は報告しない
- ③ 有意な差が出るまで，検定する指標を増やし続ける
- ④ 指標の数が増えるほど，各検定の有意水準を大きくする

問 6 次の文章を読み、空欄 フ ヘ , ホ マ に当てはまる数字をマークせよ。
 また、空欄 ミ ・ ム に入れるのに最も適当なものを、後の解答群のうちから一つずつ選べ。

図 1 は、A/B テストの期間中の日ごとの購入率である。ただし、各日の訪問者数は、画面 A、B ともに同じであった。

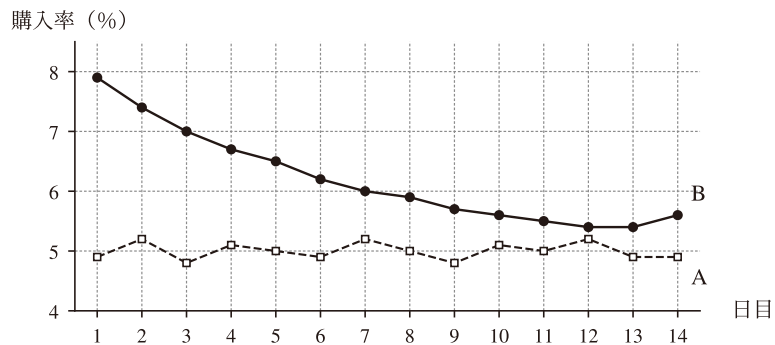


図 1 日ごとの購入率

表 4 日ごとの購入率(%)

日目	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	4.9	5.2	4.8	5.1	5.0	4.9	5.2	5.0	4.8	5.1	5.0	5.2	4.9	4.9
B	7.9	7.4	7.0	6.7	6.5	6.2	6.0	5.9	5.7	5.6	5.5	5.4	5.4	5.6

表 4 から、前半の 7 日間(1 ～ 7 日目)の購入率の平均の差(B－A)は約 フ ヘ ポイント、後半の 7 日間(8 ～ 14 日目)の差は約 ホ マ ポイントである(小数第 2 位を四捨五入)。図 1 と表 4 から読み取れることとして最も適当なものは ミ である。

ミの解答群

- ⑩ 画面 B の購入率は実験の後半ほど高くなっており，長く使うほど効果が大きくなる
- ⑪ 画面 B の購入率は初めのころが高く，次第に画面 A に近づいている。新しい画面の目新しさによる一時的な効果が含まれている可能性があるので，実験期間を延ばして効果が続くかを確かめるべきである
- ⑫ 毎日の購入率の差はほぼ一定なので，14 日間の平均の差 1.2 ポイントが今後も続くと考えられる
- ⑬ 画面 A の購入率は日によって大きく変動しているので，画面 A の結果は信頼できない

また，V さんの同僚は，実験の途中で毎日検定を行い，初めて有意な差が出た時点で実験を終了すればよいと提案した。この提案に関する記述として最も適当なものは **ム** である。

ムの解答群

- ⑭ 毎日検定を行うことで偽陽性の確率が下がるので，望ましい方法である
- ⑮ 実験を早く終了できれば，検出力は必ず上がる
- ⑯ 途中で何度も検定を行い，有意になった時点で止めると，本当は差がなくても有意と判定される確率が 5 % より大きくなる
- ⑰ 検定の回数は結果に影響しないので，どちらの方法でもよい

(下 書 き 用 紙)

(下 書 き 用 紙)

II 解答上の注意

- 1 解答は、解答用紙の問題番号に対応した解答欄にマークしなさい。例えば、第2問の **ア** と表示のある問いに対して③と解答する場合は、次の例のように問題番号 **2** の解答記号 **ア** の解答欄の③にマークしなさい。

例1

2	解 答 欄															
	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
ア	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	Ⓕ

- 2 この試験では、選択肢から選んで答える問題と、解答欄の数字(0～9)又は文字(a～f)から選んで答える問題があります。各問題の指示に従って解答しなさい。
- 3 問題の文中の **イ**，**ウ****エ** などの **□** に、数字(0～9)又は文字(a～f)を入れるよう指示された場合、**イ**，**ウ**，**エ**…の一つ一つは、これらのいずれか一つに対応します。それらを解答用紙の**イ**，**ウ**，**エ**…で示された解答欄にマークして答えなさい。

例2 **ウ****エ** に 38 と答えたいとき

ウ	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	Ⓕ
エ	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	Ⓕ

- 4 同一の問題文中に **ア**，**ウ****エ** などが2度以上現れる場合、原則として、2度目以降は、**ア**，**ウ****エ** のように細字で表記します。

※この問題冊子は、大学入学共通テスト「情報Ⅰ」の過去の出題傾向をもとに作成した予想問題であり、独立行政法人大学入試センターが作成・公表したものではありません。また、問題中のデータは架空のものです。