長崎大学大学院総合生産科学研究科（博士前期課程）総合生産科学専攻
共生システム科学コース（情報データ科学分野）

2024年度入学試験問題　情報データ科学科目　　　　　　　受験番号

1.　8桁(8bit)の符号付き2進数について答えよ．ただし，負数は2の補数で表現するものとする．

(1)　表すことのできる最大値を求め10進法で答えよ．またその数値を表すビット列を示せ．さらに，そのビット列を16進法で表記せよ．

解答欄

10進法：

ビット列：

16進法：

(2)　表すことのできる最小値を求め10進法で答えよ．またその数値を表すビット列を示せ．さらに，そのビット列を16進法で表記せよ．

解答欄

10進法：

ビット列：

16進法：

(3)　10進数の100を表すビット列を示せ．

解答欄

(4)　10進数の −100 を表すビット列を示せ．

解答欄

2.　100 円以下の与えられた金額ちょうどになる最小の硬貨枚数を計算するプログラムを python で作成する．空欄を埋めてプログラムを完成させよ．ただし，新たな変数は追加しないものとする．また，使用できる硬貨は，100 円玉，50 円玉，10 円玉，5 円玉，1 円玉の 5 種類とする．

例：46 円の場合，10 円玉×4 個　＋　5 円玉×1 個　＋　1 円玉×1 個　で，最小の硬貨枚数は 6 枚

[アルゴリズムの概要]

- 高額の硬貨から使うように考えて枚数と残金を計算していく．
- 金額に対して，ある額の硬貨が何枚まで使えて，残金がいくらになるかを計算する．
  例：46 円に対して 10 円玉が何枚使えるかは 46÷10 の「商」，いくら残るかは 46÷10 の「余り」で求められる．

[プログラム(python)]

```
# 与えられた金額に対して
# 金額ちょうどになる最小の硬貨枚数を計算する
# 関数(num_of_coins)を定義する.
# 整数除算には // を使用すること.
 (1)  num_of_coins(total):
    # 使用できる硬貨の種類を coins にリストとして定義
    coins = (2)
    num = 0
    remain = total
    #ある額の硬貨が何枚まで使えて，残金がいくらになるかを計算する
    for coin in coins:
        num = (3)
        remain = (4)
    # 硬貨枚数を返す
    return num

# メインプログラム部
kingaku = 46 # 金額
maisuu = num_of_coins(kingaku)
print(maisuu) # 枚数を出力
```

解答欄

(1)	
(2)	
(3)	
(4)	

3.　決定性有限オートマトン $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ がある．状態の集合は $Q = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6\}$，入力アルファベットは $\Sigma = \{a, b\}$，初期状態は q_0，受理状態の集合は $F = \{q_0, q_1, q_2, q_3, q_4, q_6\}$であり，動作関数は以下のように与えられている．

$\delta(q_0, a) = q_1$,　$\delta(q_1, a) = q_5$,　$\delta(q_2, a) = q_4$,　$\delta(q_3, a) = q_4$,　$\delta(q_4, a) = q_4$,　$\delta(q_5, a) = q_5$,　$\delta(q_6, a) = q_2$,

$\delta(q_0, b) = q_3$,　$\delta(q_1, b) = q_2$,　$\delta(q_2, b) = q_3$,　$\delta(q_3, b) = q_3$,　$\delta(q_4, b) = q_3$,　$\delta(q_5, b) = q_5$,　$\delta(q_6, b) = q_4$

以下の設問に答えよ．

(1)　次の記号列 w_1〜w_5 のうち，M が受理するものをすべて選び，○で囲め．

解答欄

$w_1 = ababaabb$　　$w_2 = baaabbbb$　　$w_3 = babaabbb$　　$w_4 = aabbbaab$　　$w_5 = bbaabaab$

(2)　M が受理しない記号列 $w \notin L(M)$ のうち，長さ $|w|$ が最小の記号列 $w \in \Sigma^*$をすべて示せ．

解答欄

(3)　M の状態 q_4 と等価な状態をすべて示せ．ただし，初期状態から到達不可能な状態は除外すること．

解答欄

(4)　受理する言語が M と同じで，状態の数が最小の決定性有限オートマトンを求め，その状態遷移図を示せ．

解答欄

4.　空記号列を ε で表すものとする．非終端記号の集合 $N = \{ S \}$，終端記号の集合 $\Sigma = \{0, 1\}$，書き換え規則の集合 $P = \{ S \rightarrow 0S1S,\ S \rightarrow 1S0S,\ S \rightarrow \varepsilon \}$，開始記号 S で定義される文法 $G = \langle N, \Sigma, P, S \rangle$ について，以下の設問に答えよ．

(1)　次の記号列 w_1～w_5 のうち，文法 G で導出できるものをすべて選び，○で囲め．

解答欄

$w_1 = 01010011$　　$w_2 = 10001111$　　$w_3 = 10100111$　　$w_4 = 00111001$　　$w_5 = 11001001$

(2)　記号列 0101 が文法 G で導出できることを次のように示すことができる．

$S \Rightarrow 0S1S \Rightarrow 01S \Rightarrow 010S1S \Rightarrow 0101S \Rightarrow 0101$

これと同様に，書き換え規則をひとつずつ適用しながら記号列 111000 の導出を示せ．

解答欄

(3)　文法 G は曖昧であることを示せ．

解答欄

(4)　$L(M) = L(G)$となるようなプッシュダウンオートマトンを $M = \langle Q, \Sigma, \Gamma, \delta, q_0, Z_0, F \rangle$ とする．
状態の集合は $Q = \{q_0\}$，入力アルファベットは $\Sigma = \{0,1\}$，スタック記号は $\Gamma = \{0, 1, S, Z_0\}$である．
δ は動作関数，q_0 と Z_0 はそれぞれ初期状態と初期スタック記号であり，受理状態の集合は $F = \{q_0\}$である．
M の状態遷移図を示せ．なお，入力記号列を読み終えたときにスタックが空かつ受理状態であることのみを受理条件とする．

解答欄

5.　2 値 1，0 をそれぞれ確率 p，確率 $1-p$ でとるベルヌーイ分布に従う確率変数 X がある．ここで，p は $0<p<1$ を満たす実数とする．無作為標本 $X_1, X_2, \dots, X_n$ を，X と同じ分布に従う n 個の独立な確率変数とする．ただし，n は 1 より大きい自然数である．$X_1, X_2, \dots, X_n$ の標本比率を $\hat{p} = n^{-1}\sum_{i=1}^{n} X_i$ とするとき，以下の問いに答えよ．

(1) X の平均 $E(X)$，分散 $V(X)$ がそれぞれ p，$p(1-p)$ と等しいことを示せ．

解答欄

(2) $\hat{p}$ の平均 $E(\hat{p})$，分散 $V(\hat{p})$ がそれぞれ p，$p(1-p)/n$ と等しいことを示せ．

解答欄

(3) $n = 1600$，$\hat{p} = 0.2$ であるとき，p に対する信頼度 90%の信頼区間および信頼度 95%の信頼区間を小数点第 3 位以下を四捨五入してそれぞれ求めよ．ただし，$n = 1600$ は十分大きいことから，$\hat{p}$ が従う分布が正規分布で近似できるとし，事象 $|Z| \le 1.64$，事象 $|Z| \le 1.96$ の確率がそれぞれ 0.9，0.95 であるとして解答せよ．ここで，Z は標準正規分布に従う確率変数とする．

解答欄

6．　大学生の 1 日の摂取カロリーを説明するため，次の単回帰モデルを考える．

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

被説明変数は 1 日の摂取カロリー（y，単位：kcal）であり，説明変数は身長（x_1，単位：cm）であり，誤差項（ε）は互いに独立に正規分布$N(0, \sigma_\varepsilon^2)$に従うとする．統計環境 R を使って，単回帰モデル（モデル 1）を通常の最小二乗法（ordinary least squares）で推定したところ，以下の出力結果が得られた．

モデル 1（被説明変数：1 日の摂取カロリー）

	推定値	標準誤差	t 値	p 値
切片	−123.5936	84.2080	−1.468	0.143
x_1	15.9139	0.4997	(a)	< 2e−16

(1) 身長が 150cm の人と比較して, 身長が 160cm の人は, 平均して 1 日の摂取カロリーが何 kcal 多いといえるか. 小数点以下を四捨五入して答えよ.

解答欄

(2) モデル 1 の出力結果の(a)に入る数値はいくらか. 計算には電卓を用いてよいが, t 値の計算式を明示せよ. 小数点第 4 位以下を四捨五入して答えよ.

解答欄

7．　大学生の 1 日の摂取カロリーを説明するため，次の重回帰モデルを考える．

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$$

被説明変数は 1 日の摂取カロリー（y，単位：kcal），説明変数は身長（x_1，単位：cm），体重（x_2，単位：kg），年齢（x_3，単位：歳），性別（x_4，0 = 男，1 = 女）であり，誤差項（ε）は互いに独立に正規分布$N(0, \sigma_\varepsilon^2)$に従うとする．統計環境 R を使って，すべての説明変数を用いて重回帰モデル（モデル 2）を通常の最小二乗法（ordinary least squares）で推定したところ，以下の出力結果が得られた．

モデル 2（被説明変数：1 日の摂取カロリー）

	推定値	標準誤差	t 値	p 値
切片	1118.7129	75.3204	14.853	< 2e−16
x_1	5.2555	0.4538	11.581	< 2e−16
x_2	12.0593	0.5343	22.572	< 2e−16
x_3	−6.8891	1.6905	−4.075	5.36e−05
x_4	−94.7315	8.5195	−11.119	< 2e−16
R^2	0.8884			
$\tilde{R}^2$	0.8875			
n	500			

注：R^2は決定係数を表す．$\tilde{R}^2$は自由度調整済み決定係数を表す．nは標本サイズを表す．

(1) 帰無仮説を$\beta_1 = 0$とする．身長（x_1）の偏回帰係数は，5%の有意水準で統計的に有意か．その根拠も述べよ．
解答欄

(2) 推定された偏回帰係数を用いて，1 日の摂取カロリー（y）の値を予測するための回帰式を作成せよ．
解答欄

(3) 身長（x_1）の値が165cm，体重（x_2）の値が65kg，年齢（x_3）が20歳，性別（x_4）が男のとき，1日の摂取カロリーを予測せよ．計算には電卓を用いてよいが，計算式を明示せよ．小数点第4位以下を四捨五入して答えよ．

解答欄

(4) 推定された回帰式の決定係数は0.8884であり，自由度調整済み決定係数は0.8875である．説明変数という用語を使って，これらの数値が異なる理由を述べよ．

解答欄

8.　2023 年の Times Higher Education による世界大学ランキング上位 200 校のデータに対して，統計環境 R を使って，主成分分析を実行したところ，以下の出力結果が得られた．使用した変数は，teaching（教育），research（研究），citations（被引用数），industry（産業界からの収入），international（国際性）のスコアである．summary(df1)は，これら 5 個の変数の要約統計量を表示している．

モデル 3（主成分分析）

```
> summary(df1)
    teaching          research          citations          industry          international
 Min.    :29.40   Min.    :27.60   Min.    : 9.30   Min.    : 37.30   Min.    :25.50
 1st Qu.:42.35   1st Qu.:45.85   1st Qu.:79.08   1st Qu.: 44.80   1st Qu.:60.27
 Median :50.30   Median :55.65   Median :87.50   Median : 58.25   Median :75.30
 Mean    :54.06   Mean    :59.26   Mean    :84.84   Mean    : 63.41   Mean    :73.58
 3rd Qu.:62.23   3rd Qu.:72.90   3rd Qu.:93.60   3rd Qu.: 79.42   3rd Qu.:91.83
 Max.    :94.80   Max.    :99.70   Max.    :99.80   Max.    :100.00   Max.    :99.60

> model3<-princomp(df1, cor=TRUE)
> summary(model3)
Importance of components:
                          Comp.1    Comp.2    Comp.3    Comp.4     Comp.5
Standard deviation     1.4378177 1.2162992 0.9083189 0.7430716 0.27585872
Proportion of Variance 0.4134639 0.2958768 0.1650086 0.1104311 0.01521961
Cumulative Proportion  0.4134639 0.7093407 0.8743493 0.9847804 1.00000000
```

出典：Times Higher Education（2023）World University Rankings

(1) 第 1 主成分の寄与率はいくらか．

解答欄

(2) 第 1 主成分と第 2 主成分の 2 個の主成分によって，全体の分散の何%を説明できるといえるか．その根拠も説明せよ．

解答欄