2004 年 7 月 6 日（今野）

数理・自然–6（統計学入門）レポート問題

解答上の注意

提出先と締め切り

- 締め切り： 7 月 28 日 (水) 16:30
- 提出先：百年館 10 階数学研究室前廊下にあるレポート回収ポスト
- レポート用紙は A4（必ずホチキスで綴じること）とし，表紙には科目名，所属学科名，学籍番号，および名前（ふりがな）を記載のこと．
- 少なくとも 3 題解答すること．

以下の点に留意してレポートを作成すること．

(1) 解答の途中過程は丁寧に記述すること．

(2) 講義で学んだ知識を利用する場合には，なにをどのように用いたかをできるだけ明示すること．

(3) 各自が理解していることを採点者にわかるように解答を作成することを心がけること．

問題 1 下のデータは 82 人の試験の得点である．このデータについてつぎの問いに答えよ．

(1) 階級数や階級のつくり方についてどのように選択をしたかについてその理由を述べた上で，下のデータの度数分布表を作成せよ．

(2) 作成された度数分布表から観察できることを述べよ．

(3) このデータの算術平均と中央値の大小関係を度数分布表から推論せよ．さらに，推論の根拠を述べよ．

57	76	57	39	52	49	64	51	64	53	51	67	52	38	63
61	57	46	49	51	48	51	56	48	70	46	39	48	62	47
65	64	45	70	54	49	47	58	51	43	58	73	58	54	65
72	50	66	52	37	38	74	68	64	67	81	56	83	72	49
51	33	51	46	84	50	41	52	53	57	47	54	45	58	50
54	63	40	77	97	42	47								

$\boxed{\text{問題 2}}$　下のデータは 1991 年から１年ごとのある地区の年間のゴミの排出量を調べた結果である．回帰方程式を求めるために下記の数表を参考にしてよい．ただし，$x =$ 西暦 $- 1990$，y をその年一年間に排出されたゴミの量とする．

x	1	2	3	4	5	6	7
y	49	44	42	33	27	12	20

(1) 上記のデータについて散布図を作図せよ．

(2) 回帰方程式を求め，(1) で作成した散布図に書き込め．

(3) 求めた回帰直線を用いて，2000 年のゴミの排出量の予測値を求め，その値を予測値として利用することの妥当性について議論せよ．

変量								合計
x	1	2	3	4	5	6	7	28
y	49	44	42	33	27	12	20	227
x^2	1	4	9	16	25	36	49	150
$x \cdot y$	49	88	126	132	135	72	140	742
y^2	2401	1936	1764	1089	729	144	400	8463

$\boxed{\text{問題 3}}$　コインを 4 回投げる実験．4 回中に表の出た回数に対応する確率変数 X の確率分布表は

X の値	0	1	2	3	4	合計
確率	1 / 16	4 / 16	6 / 16	4 / 16	1/16	1

であった．以下の問いに答えよ．ただし，解答だけでなく途中経過（途中の計算式など）もできるだけ丁寧に記述すること．

(1) X の期待値 $\mathbb{E}[X]$ を求めよ．

(2) X の分散 $\mathbb{VAR}[X]$ を求めよ．

(3) $Y = \dfrac{X - \mathbb{E}[X]}{\sqrt{\mathbb{VAR}[X]}}$ とおく．Y の確率分布表を求めよ．

(4) 確率 $\mathbb{P}(-1 \leq Y \leq 1)$ を求めよ．

$\boxed{\text{問題 4}}$ 下の表は，30 日間の（自動車の）運転免許の停止をはじめて受けたのものを年齢（30 未満/30 以上) と 30 日間の運転免許の停止後の 1 年間に再度免許停止となったものとそうでないもの 4 つのカテゴリーに分類された比率をまとめたものである．

年齢	停止後の 1 年間の状況		計
	再度免停	免停を受けず	
30 未満	0.27	0.33	0.60
30 以上	0.10	0.30	0.40
計	0.37	0.63	1.00

事象 A と B を以下のように定義する．

A：30 日間の（自動車の）運転免許の停止をはじめて受けたのものの年齢が 30 以上．
B： 30 日間の（自動車の）運転免許の停止をはじめて受けたのものが停止後の 1 年間に再度免許停止になる．

このとき，以下の事象が起こる確率を求めよ．

(a) $P(A)$ （事象 A が起こる確率） (b) $P(B)$
(c) $P(A \cap B)$ （A と B の積事象） (d) $P(A \cup B)$ （A と B の和事象）
(e) $P(A^c)$ （A の補事象） (f) $P\{(A \cup B)^c\}$
(g) $P\{(A \cap B)^c\}$
(h) $P(B|A)$ （A が与えられたときの B の起こる確率）
(i) $P(A|B)$ （B が与えられたときの A の起こる確率）
(j) $P(A^c|B)$

$\boxed{\text{問題 5}}$ ビアガーデンでひとりあたりの客が注文するビールの中ジョッキの数の確率分布が以下になることが調査で知られているとする．

個数	1	2	3	4	5	6	計
確率	0.05	0.20	0.40	0.20	0.10	0.05	1

(1) この確率分布の平均を求めよ．

(2) この確率分布の分散と標準偏差を求めよ．計算をするときには，小数第 3 位以下を四捨五入してよい．(ヒント：$\sqrt{139} = 11.7$)

(3) 100 人の客が来る予約が入っているとき，ビールが足りなくなるような事態が稀にしか起きないようするためには，すくなくとも何杯分のビールを準備するのが妥当かを答えよ．また，その理由についても述べよ．ただし，予約された客が注文するビールの数は独立同一に上の確率分布に従うと仮定する．

ヒント：確率変数列 $X_1, X_2, \ldots, X_n$ が平均 μ，分散 σ^2 の分布に独立同一に従うとき，中心極限定理から

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \quad \text{もしくは} \quad \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

の分布は標準正規分布で近似できることを使うとよい．ここで，$\bar{X}_n = (1/n)(X_1 + X_2 + \cdots + X_n)$ と $S_n = X_1 + X_2 + \cdots + X_n$ とした．

$\boxed{\text{問題 6}}$　下の表は，オスカー賞における最優秀男優賞と最優秀女優賞の受賞者の年齢を過去 30 年にわたり調査した結果である．それぞれの受賞年齢について度数分布表と図を作成したうえ，男女の受賞時の年齢の分布について比較し，その特徴や性別間での相違（そのようなものがあれば）を述べよ．

男優	32	51	33	61	35	45	55	39	76	37
	42	40	32	60	38	56	48	48	40	43
	62	43	42	44	41	56	39	46	31	47
女優	80	26	41	21	61	38	49	33	74	30
	33	41	31	35	41	42	37	26	34	34
	35	26	61	60	34	24	30	37	31	27

$\boxed{\text{問題 7}}$　次の貯蓄高のデータに対し平均値を求め、最頻値（モード）との関係について論じよ
.

$$100, 110, 120, 120, 120, 130, 130, 140, 150\,(\text{万円})$$

(1) 母集団、標本、標本抽出という用語を用いて、母平均と標本平均の違いを説明せよ．

$\boxed{\text{問題 8}}$　1 年生の統計学の試験（100 満点で得点は整数とする）の成績の平均が 60 点で標準偏差が 15 点であった．ただし，得点の分布は正規分布に従うものと仮定する．

(1) 上位 20 % 以内に「A」以上（「A または A^+」）の評価をつけるとすれば，何点以上の得点で評価が「A」以上になるか．

(2) 上位 5 % 以内に 「A^+」の評価ををつけるとすれば，何点以上の得点で評価が「A^+」になるか．

(3) 下位 20 % 以内に「不合格」の評価をつけるとすれば，何点以下の得点で評価が「不合格」になるか．

$\boxed{\text{問題 9}}$　コインを投げたときの表を H，裏を T で表す．1 枚のコインを 2 回独立に投げたときの標本空間を簡単に $\Omega = \{HH, HT, TH, TT\}$ と書く．ここで，HH は 1 回目は表で 2 回目も表がでた事象を表し，HT は 1 回目は表で 2 回目は裏が出た事象を表す．他についても同様な表記とする．つぎの事象を定義する：

$$A = \{HH, HT\}, \quad B = \{HH, TH\}, \quad C = \{HT, TH\}$$

(1) A, B, C はどんな事象かをその性質をうまく捕らえる言葉で説明せよ．

(2) つぎの関係が成り立つことを確かめよ．

$$(\mathrm{a})\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \quad (\mathrm{b})\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C) \quad (\mathrm{c})\mathbb{P}(C \cap A) = \mathbb{P}(C)\mathbb{P}(A)$$

(3) 確率 $\mathbb{P}(A \cap B \cap C)$ を求めよ．