

経済統計学 - 宿題 No.1

蛭川雅之

提出期限：2025 年 10 月 24 日（金）11:00（厳守）

注意事項

- 学生相互間で相談をしても構いませんが、答案は必ず個人で作成してください。
- 答案は手書きでもタイプ打ちでも構いませんが、最初のページに必ず学籍番号と氏名を明記してください。
- 答案を作成する際には、忘れずに問題番号を記入してください。具体的な答案の作成要領は次の通りです。
 1. 第 1 部の各問題については、計算根拠などをできるだけ具体的に説明してください。必要に応じて途中計算を書いても構いません。なお、説明がなくただ結果だけが書いてある場合、その問題の点数は零点とします。また、数値は電卓を用いて計算してください。
 2. 第 2 部の各問題については、Excel ワークシートの該当箇所のプリントアウトもしくはスクリーンショットを添付してください。コメントが必要な場合は、ワークシートに直接タイプする、あるいは、プリントアウト・スクリーンショットにメモ書きをする等の方法で対処してください。
- 答案は第 1 部・第 2 部をまとめて左肩にホチキス止めをした上で提出してください。
- 提出期限を厳守してください。いかなる理由があろうと、期限後に提出された答案は受け取りません。
- ルールに違反して答案を書いた（例：本講義に履修登録していない学生に助力を求める、他の学生の答案もしくは過去問の解答例を丸写しする）場合、内容の如何を問わず減点の対象となりますので注意してください。

第 1 部

問題 1-1

新生児 60 人の体重データを基に次の度数分布表を作成した。階級の読み方は、階級“2000～2350”を例にとると「2000g 以上 2350g 未満」を意味する。また、階級値は各階級の中央値である。この度数分布表の空欄 (1)-(8) に該当する数値を求めよ。なお、空欄 (5)(6) はそれぞれ小数第 4 位まで求めること。空欄以外の数値の一部も意図的に空白にしてあるので注意せよ。

階 級	階級値	度 数	相対度数	累積度数
2000 ～ 2350		(3)		(7)
2350 ～ 2700	(1)	8		10
2700 ～ 3050		(4)		26
3050 ～ 3400		18	(5)	
3400 ～ 3750	(2)	7		
3750 ～ 4100		4	(6)	
4100 ～ 4450		3		
4450 ～ 4800		2		(8)

問題 1-2

まず、 n 個の生データ $(X_1, Y_1), \dots, (X_n, Y_n)$ を降順に並べ替え、スピアマン順位相関係数を計算する。次に、同じ生データを昇順に並べ替え、改めてスピアマン順位相関係数を計算する。このとき、これら 2 つの値は等しいことを示せ。なお、単純化のため、 (X, Y) いずれにも同順位はないと仮定せよ。

問題 1-3

2 つの離散確率変数 X と Y について、 X は $x_1, x_2, \dots, x_n$ のいずれか、 Y は $y_1, y_2, \dots, y_m$ のいずれかをとるものとする。このとき、 X と Y の共分散 $Cov(X, Y)$ について、

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

が成り立つことを示せ。さらに、この結果を利用して、 X と Y が独立である場合、これらは無相関であることを示せ。

第 2 部

問題 2-1

Excel ファイル hw1_q2-1.xls には女子大学生 80 名分の身長が記録されている。このデータについて、(1) 度数分布表、ヒストグラム、箱ひげ図を作成し、(2) 平均値、中央値、分散 $\hat{\sigma}^2$ 、標準偏差 $\hat{\sigma}$ を計算せよ。なお、度数分布表は以下の手順で作成するものとする。

1. 度数分布表の体裁は問題 1-1 と同様とする。累積比率は求めなくてよい。
2. 階級は上から順に

143 ~ 146 146 ~ 149 149 ~ 152 152 ~ 155 155 ~ 158
158 ~ 161 161 ~ 164 164 ~ 167 167 ~ 170

の 9 つとする。なお、階級の読み方は、例えば“143 ~ 146”の場合、「143cm 以上 146cm 未満」を意味する。

3. 各階級値はその階級の中央値とする。

問題 2-2

Excel ファイル hw1_q2-2.xls には男子大生 50 名分の身長と体重が記録されている。このデータから (1) 身長を横軸、体重を縦軸に取った散布図を作成し、(2) 身長・体重それぞれの平均値、中央値、分散 $\hat{\sigma}^2$ 、標準偏差 $\hat{\sigma}$ を計算し、さらに (3) 身長と体重についてのピアソン積率相関係数およびスピアマン順位相関係数を求めよ。

問題 2-3

2 つの離散確率変数 X と Y について、 (X, Y) の同時確率 $p_{ij} = p_{X,Y}(x_i, y_j) = \Pr(X = x_i, Y = y_j)$ が下表に与えられている。このとき、以下の問いに答えよ。

$X \backslash Y$	0	1
1	3/20	3/20
2	1/10	3/10
3	3/20	3/20

- (1) X の周辺確率分布を求めよ。

- (2) Y の周辺確率分布を求めよ。
- (3) X と Y は独立か。理由を付して説明せよ。
- (4) X の期待値 $E(X)$ および分散 $Var(X)$ を求めよ。
- (5) Y の期待値 $E(Y)$ および分散 $Var(Y)$ を求めよ。
- (6) X と Y の共分散 $Cov(X, Y)$ および相関係数 $Corr(X, Y)$ を求め、結果についてコメントせよ。
- (7) $X = 1, 2, 3$ それぞれの場合について Y の条件付き確率分布を求めよ。
- (8) $X = 1, 2, 3$ それぞれの場合について Y の条件付き期待値を求め、期待値繰り返しの法則が成り立つことを確認せよ。
- (9) $X = 1, 2, 3$ それぞれの場合について Y の条件付き分散を求め、全分散の法則が成り立つことを確認せよ。

問題 2-4

ある保険会社は、企業向けに詐欺被害に対応する損害保険を開発し、2500 社に販売した。詐欺被害に遭う企業の割合は 1000 社あたり 1 社であると仮定して、この保険会社に対し実際に保険金請求を行う企業数が (1) 皆無である確率、(2) ちょうど 1 社である確率、(3) ちょうど 2 社である確率、(4) 3 社以上である確率をそれぞれ小数第 4 位まで求めよ。なお、各確率とも二項分布、ポアソン分布による近似の 2 通りで計算し、結果を比較すること。

問題 2-5

高名な料理人 K 氏が鯛のあら炊き 1 皿を調理するのに要する時間は平均 22 分、標準偏差 5 分の正規分布とみなして差し支えないとする。このとき、以下の問いに答えよ。

- (1) K 氏が鯛のあら炊き 1 皿を調理するのに要する時間が 20 分から 25 分の間である確率を小数第 3 位まで求めよ。

(2) K 氏が鯛のあら炊き 1 皿を調理するのに要する時間がある長さを超える確率は 0.1 であるという。この時間の長さ（単位：分）を小数第 1 位まで求めよ。

(3) あるテレビ局は K 氏をスタジオに招き、30 分間の生放送情報番組内で鯛のあら炊き 1 皿を調理してもらおうと計画している。この番組の終了時点で K 氏の調理が完了していない確率を小数第 3 位まで求めよ。なお、K 氏は番組冒頭から調理を開始し、番組の前後に一切 CM は入らない（＝番組は正味 30 分間である）とする。また、番組出演者の試食時間を考慮に入れる必要はない。