2015年度「統計」期末試験問題

付表3から付表6を参照し，正しい答えを選択しなさい．ただし，数値は四捨五入により小数点以下3桁までとする（計算過程においても小数点以下4桁を四捨五入し，小数点以下3桁までを利用する）．

選択肢を選ぶ問題の場合には適切な選択肢の番号をマークし，数値を答える問題の場合には1桁ごとに数値をマークすること．

問題内の，〔1〕，〔2〕……は，解答番号を表している．

ただし，解答番号【7】【13】は正負の符号を表すものとし，解答が正（＋）である場合は「9」に，負（−）である場合は「0」にマークせよ．

注意：解答が小数点以下の位において0となる部分には0をマークすること．

例）解答が23.28の場合で，〔1〕〔2〕.〔3〕〔4〕〔5〕に解答する場合には，

〔1〕2

〔2〕3

〔3〕2

〔4〕8

〔5〕0

をマークする．

Ⅰ．A社とB社が「垂直的分業関係」にある場合，両社の生産額には正の相関が見られる．また，A社とB社とが「競争関係」にある場合には両社の生産額には負の相関が見られ，この2社が「水平的分業関係」にある場合，両社の生産額は無相関になる．

ここで，近年のA社とB社の生産額を調べると，下の表1の通りとなった．また，その結果から導かれる記述統計量は表2の通りである．

表1　A社およびB社の生産額

年	A社	B社
2008	35	19
2009	28	22
2010	25	30
2011	28	25
2012	24	30
2013	19	30

（単位）億円

表2　各種記述統計量

統計量	A社	B社
標本平均	26.5	〔1〕〔2〕
標本分散	28.3	22.8
標本標準偏差	5.320	〔3〕.〔4〕〔5〕〔6〕
標本共分散	【7】〔8〕〔9〕.〔10〕〔11〕〔12〕	
相関係数	【13】〔14〕.〔15〕〔16〕〔17〕	

（1）表 2 中の，B 社の生産額の標本平均，標本標準偏差と，両社の生産額の標本共分散，相関係数を求め，適切な数値および符号（符号は【7】【13】）をマークせよ．

（2）A 社の生産額の中央値は〔18〕〔19〕.〔20〕，最頻値は〔21〕〔22〕である．

（3）以上の結果から導かれることとして，最も適切なものを 1 つ選び，その番号を解答欄にマークせよ．　…………………解答番号〔23〕

①A 社の生産量が増大した結果，B 社の生産量が増大した．
②B 社の生産量が増大した結果，A 社の生産量が増大した．
③A 社と B 社は，「垂直的分業関係」にある．
④A 社と B 社は，「水平的分業関係」にある．
⑤A 社と B 社は，「競争関係」にある．
⑥A 社と B 社の関係はわからない．

Ⅱ．次の（1）の空欄に，適切な数値をマークせよ．
また，（2）については，下の文章の内容が正しい場合には「1」を，誤りである場合には「0」をマークせよ．なお，問題文の前の〔番号〕は解答番号である．

（1）4 つの選択肢の中から 1 つの正解を選ぶ問題が 20 問用意されている．なお，各問題の配点はすべて 5 点である．

J 君は，どの問題もわからなかったので，すべての問題にでたらめに答えた．このとき，J 君の正解数の平均は〔24〕問であり，J 君の点数の期待値は〔25〕〔26〕点である．

（2）平均が μ，分散が σ^2 の無限母集団から n 個の無作為標本を復元抽出したとして，その n 個の標本から作られる標本平均 $\bar{X}$ を母集団平均（母平均：μ）の推定量として用いる場合を考える．

〔27〕この標本平均 $\bar{X}$ の平均は μ，分散は σ^2 である．
〔28〕この標本平均 $\bar{X}$ の期待値（$E(\bar{X})$）は，真の平均（母集団平均（μ））に等しくなるから，一致推定量である．
〔29〕この標本平均 $\bar{X}$ は，標本数 n の極限（$n \to \infty$）において真の平均（母集団平均（μ））に収束することがチェビシェフの不等式によって証明される．これを大数の法則という．
〔30〕この標本平均 $\bar{X}$ は，標本数 n の極限（$n \to \infty$）において真の平均（母集団平均（μ））に収束するから，不偏推定量である．

（裏面に続く）

Ⅲ．次の（1）（2）の〔31〕～〔37〕について，適切な数値をマークせよ．

（1）ある国ではM伝染病がはやっており，1日に平均1人感染するという．5日間に5人以上感染する確率は〔31〕〔32〕%である（%（百分率）にしたときに小数点以下が出た場合は四捨五入すること）．なお，M伝染病の感染者数はポアッソン分布（ポアソン分布）に従うものとする．

（2）ある大学では，2年次終了時のGPAが1.00未満である場合，3年次に進級できないことになっている．ここで，ある年の2年次終了時の全員の成績を調査したところ，平均＝2.22，標準偏差＝1.525の正規分布に従っていた．したがって，この学年では〔33〕〔34〕．〔35〕〔36〕〔37〕%の学生が3年次に進級できないことになる（ただし，この問題では計算途中で四捨五入しないこと）．

Ⅳ．以下では，ベイズの定理（ベイズの公式）を導出し，ベイズの定理を用いた確率計算を行う．

（1）まず，ベイズの定理を導出する第1ステップとして，条件付き確率を確認する．$P(A|B)$として最も適当なものを1つ選びなさい．　　…………………………………解答番号〔38〕

① $P(A|B)=\dfrac{P(A\cup B)}{P(B)}$　　② $P(A|B)=\dfrac{P(A\cup B)}{P(A)}$

③ $P(A|B)=\dfrac{P(A\cap B)}{P(B)}$　　④ $P(A|B)=\dfrac{P(A\cap B)}{P(A)}$

⑤ $P(A|B)=\dfrac{P(A)}{P(A\cup B)}$

（2）上記の条件付き確率と$P(B)=P(B\cap A)+P(B\cap A^C)$であることを用いると，ベイズの定理が導出される．ベイズの定理として最も適当なものを1つ選びなさい．　　…………………解答番号〔39〕

① $P(A|B)=\dfrac{P(B|A)P(B)}{P(B|A)P(A)+P(A|B^C)P(B^C)}$

② $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(A)+P(A|B^C)P(B^C)}$

③ $P(A|B)=\dfrac{P(B|A)P(B)}{P(B|A)P(A)+P(B|A^C)P(A^C)}$

④ $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^C)P(A^C)}$

⑤ $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(B)+P(B|A^C)P(A^C)}$

（3）ある都市でシートベルト着用率と事故死の状況を調査したところ，事故死になる（A）人のうち2割はシートベルトを着用（B）していた．他方，事故死にならなかった（$A^C(\bar{A})$）人のうち7割はシートベルトを着用していることもわかった．ちなみに，事故死になる人は全体の 5%だとわかっている．このような情報をもとにして計算される，シートベルトをしながら（B）事故死になる人（A）の確率は〔40〕．〔41〕〔42〕〔43〕%である．

V．プロ野球選手のＦ投手の球速について９球をしらべたところ，標本平均は 140km/h，標本標準偏差は 15km/h となった．Ｆ投手のプロ野球時代の球速が正規分布に従うとして，以下の問いに答えよ．

（１）母分散（母集団分散：σ^2）について信頼係数 95％で区間推定を行うと，信頼区間は，
〔44〕〔45〕〔46〕．〔47〕〔48〕〔49〕以上，〔50〕〔51〕〔52〕．〔53〕〔54〕〔55〕以下となる．

（２）母平均（母集団平均：μ）について信頼係数 95％で区間推定を行うと，信頼区間は，
〔56〕〔57〕〔58〕．〔59〕〔60〕〔61〕以上，〔62〕〔63〕〔64〕．〔65〕〔66〕〔67〕以下となる．

（３）母分散（母集団分散）が 144 とわかっている場合，母平均（母集団平均：μ）について信頼係数 95％で区間推定を行うと，信頼区間は，
〔68〕〔69〕〔70〕．〔71〕〔72〕〔73〕以上，〔74〕〔75〕〔76〕．〔77〕〔78〕〔79〕以下となる．

（４）母分散（母集団分散）が 144 とわかっている場合，「Ｆ投手のプロ野球時代の平均球速が，Ｆ投手の高校時代の最速スピード 132km/h より速いと言えるか」について有意水準 1％で検定を行った場合，P 値は〔80〕．〔81〕〔82〕〔83〕であり，帰無仮説は〔84〕される．
※〔84〕に入る語が，「棄却」の場合には「1」を，「採択」の場合には「0」をマークせよ．
※帰無仮説が「採択」されるとは，帰無仮説が「棄却されない」ということである．

(2 枚目に続く)

VI. ある調査で，子供の頃に「ルールを守る」という躾（しつけ）を受けたと記憶している者（以下，「躾を受けた者」）と，記憶していない者（以下，「躾を受けていない者」）について，現在の所得がどのように異なるかを調べた．その結果，「躾を受けた者」の回答者数は 1600 人で，このグループの平均所得は 680 万円，標準偏差は 200 万円であった．また，「躾を受けていない者」の回答者数は 2500 人であり，このグループの平均所得は 665 万円，標準偏差は 180 万円であった．この調査結果から，「子供の頃にルールを守るという『躾を受けた者』の平均所得が，『躾を受けていない者』の平均所得よりも高いと言えるか」について仮説検定しなさい．なお，2 つのグループの母分散は異なると仮定する．

（1）この検定を行うための（標準化された）検定統計量は，〔85〕.〔86〕〔87〕〔88〕である．

（2）この問題の検定結果として最も適当なものを 1 つ選びなさい．　　　　…………解答番号〔89〕

①有意水準 5%の臨界値は 1.96 であり，「躾を受けた者」の平均所得は「躾を受けていない者」の平均所得よりも有意には高くない．

②有意水準 5%の臨界値は 1.645（1.65）であり，「躾を受けた者」の平均所得は，「躾を受けていない者」の平均所得よりも有意に高くない．

③有意水準 1%の臨界値は 2.33 であり，「躾を受けた者」の平均所得は，「躾を受けていない者」の平均所得よりも有意に高くない．

④有意水準 1%の臨界値は 2.33 であり，「躾を受けた者」の平均所得は，「躾を受けていない者」の平均所得よりも有意に高い．

⑤有意水準 1%の臨界値は 2.575（2.58）であり，「躾を受けた者」の平均所得は，「躾を受けていない者」の平均所得よりも有意に高い．

Ⅶ．回帰分析に関する以下の文章の（１）～（３）について，（１）の〔91〕～〔93〕，および（３）の〔96〕～〔100〕に最も適切な語句を指定された【語群】から選び，その番号をマークせよ．なお，同じ選択肢を複数回用いても良い．

また，（１）の〔90〕，（２）の〔94〕〔95〕には，適切な数値をマークせよ．

（１）標準的線形回帰モデルの仮定のもとでは，誤差項（u_i）は平均が〔90〕，分散は独立で〔91〕な分散σ^2をもつ〔92〕分布となっている．また，〔93〕変数は確率変数ではなく，誤差項とは〔93〕である．

（２）最小自乗法（最小二乗法）（特に単回帰）においてモデルの説明力を表す決定係数は，最小で〔94〕，最大で〔95〕の値を取り，〔95〕に近いほどモデルの説明力が高い．

（３）最小自乗法（最小二乗法）では，残差（誤差）の２乗の和を最小にするように回帰直線を引く．残差（誤差）とは回帰直線と各データの間の距離のことであるが，この距離は〔96〕に対して垂直に取っている．

回帰モデル$Y_i = a + bX_i + u_i$（ここで，u_iは標準的線形回帰モデルの仮定を満たす誤差項）について最小自乗法でパラメータbを推定すると，この推定量は〔97〕となる．この推定量をもとにパラメータbの有意性を検定する場合，帰無仮説を〔98〕として，検定統計量$\frac{\hat{b}}{s_{\hat{b}}}$（$s_{\hat{b}}$は標準誤差）が〔99〕分布に従うことを用いて行う．

検定結果は，上記の帰無仮説が〔100〕場合，この２つの変数XとYとの間には$Y = a + bX$となる有意な関係があるとされる．

〔91〕と〔93〕の【語群】

①被説明　②従属　③独立　④周辺　⑤均一

〔92〕と〔99〕の【語群】

①正規　②標準正規　③自由度$n-1$のχ^2　④自由度$n-2$のχ^2
⑤自由度$n-1$のt　⑥自由度$n-2$のt　⑦自由度n_1+n_2-2のt

〔96〕〔97〕と〔100〕の【語群】

①NABY　②BIUS　③BLUE　④X軸（横軸）　⑤Y軸（縦軸）
⑥回帰直線　⑦棄却される　⑧棄却されない　⑨無相関である

〔98〕の【語群】

①$H_0: b = 0$　②$H_0: b \neq 0$　③$H_0: b > 0$　④$H_0: b < 0$　⑤$H_0: b \geq 0$
⑥$H_0: b \leq 0$

（以下，余白）