2016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【10】【17】【53】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．日本の都道府県のうちの6つの府県について加入電話[1]の加入率（契約数／世帯総数）と携帯電話の加入率（契約数／県（府）民人口数）を調べたところ，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値は表2の通りである．

表1　加入電話の加入率　と
携帯電話の加入率

都道府県	加入電話	携帯電話
A県	31	95
B県	33	103
C県	29	117
D県	28	102
E府	36	99
F県	42	96

表2　各種記述統計量

統計量	加入電話	携帯電話
標本平均	33.167	102.000
標本分散	26.940	〔1〕〔2〕.〔3〕〔4〕〔5〕
標本標準偏差	5.190	〔6〕.〔7〕〔8〕〔9〕
標本共分散	－21.841	
相関係数	【10】〔11〕.〔12〕〔13〕〔14〕	

（1）表2中の，携帯電話の加入率の標本分散，標本標準偏差と，加入電話加入率と携帯電話加入率に関する相関係数を求め，適切な数値および符号（符号は【10】）をマークせよ．

[1] 固定電話のこと

（２）加入電話加入率の中央値は〔15〕〔16〕％である．

（３）データの標準化の式を用いると，B 県の携帯電話加入率の標準化された値は

【17】〔18〕.〔19〕〔20〕〔21〕

となる．(符号は【17】)

Ⅱ．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〔22〕

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

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

⑤ $P(A|B)=\dfrac{P(A\cup B)}{P(A)}$

（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〔23〕

① $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(B)+P(B|A^C)P(A^C)}$　　② $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^C)P(A^C)}$

③ $P(A|B)=\dfrac{P(B|A)P(B)}{P(B|A)P(A)+P(B|A^C)P(A^C)}$　　④ $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(A)+P(A|B^C)P(B^C)}$

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

（３）ある企業における入社試験の合否結果に，男女差別が存在しているかどうかが問題となった．「合格すること」を事象 A,「男性であること」を事象 B としたとき，

全体の合格率 $P(A)=0.756$

男性比率 $P(B)=0.844$

合格者に占める男性の比率 $P(B|A)=0.888$

不合格者に占める男性の比率 $P(B|A^C)=0.707$

であることが分かっている．このとき，男性の合格率 $P(A|B)$ は〔24〕.〔25〕〔26〕〔27〕であり，女性の合格率 0.545 よりも高い．

（裏面に続く）

Ⅲ．以下の（１）～（３）の空欄に入る最も適切なものを下の指定された【選択肢】から選び，その番号をマークせよ．

（１）成功確率 p のベルヌーイ試行（成功＝1，失敗＝0 とする）を n 回繰り返したとき，x 回成功する確率は〔28〕分布に従い，その平均は〔29〕，分散は〔30〕である．

（２）母集団の分布がいかなるものであっても，その母集団の平均が μ ，分散が σ^2 であることがわかっている場合，そこから n 個の標本を無作為抽出[2]し，その無作為標本を用いて導かれる標本平均（ $\bar{X}$ ）の平均は〔31〕，分散は〔32〕となることがわかっている．

（３）上の（２）と同様に，平均 μ ，分散 σ^2 の分布を持つ母集団から n 個の標本を無作為抽出し，その無作為標本から標本平均（ $\bar{X}$ ）を導出する．このとき，標本数が無限大に近づくにつれて（$n \to \infty$），標本平均 $\bar{X}$ が母集団平均（母平均） μ に収束することが証明される．これを〔33〕といい，この性質から標本平均 $\bar{X}$ は母集団平均 μ の〔34〕であることがわかる.〔33〕は, 正確には n の極限（$n \to \infty$）において標本平均 $\bar{X}$ が母集団平均 μ から外れる確率が 0 となることを示す〔35〕を用いて証明するのであるが，直感的には標本平均 $\bar{X}$ の分散が〔32〕であることから導かれる．

　とはいえ，実際に標本数 n の極限（$n \to \infty$）を取ることは不可能である．そこで，標本数 n が大きいとき，母集団がいかなる分布を有していても，標本平均 $\bar{X}$ が平均＝〔31〕，分散＝〔32〕の〔36〕分布となり，それを $Z=$〔37〕で標準化した場合，この標準化された確率変数 Z が〔38〕分布となることがわかっている．これを〔39〕という．

〔29〕～〔32〕の【選択肢】

①p　②np　③$p(1-p)$　④$np(1-p)$　⑤μ

⑥$\dfrac{p(1-p)}{n}$　⑦$\dfrac{\mu}{n}$　⑧$\dfrac{\sigma}{n}$　⑨$\dfrac{\sigma^2}{n}$

〔28〕〔36〕〔38〕の【選択肢】

①ベルヌーイ確率　②二項　③ポアソン　④正規

⑤標準正規　⑥自由度 n の t　⑦自由度 $n-1$ の t

〔33〕〔34〕〔35〕〔37〕〔39〕の【選択肢】

①$\dfrac{\bar{X}-\mu}{\sigma^2}$　②$\dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$　③不偏推定量　④一致推定量

⑤チェビシェフの不等式　⑥大数の法則　⑦加法定理　⑧乗法定理

⑨中心極限定理

[2] ここでの無作為抽出は，無限母集団からの無作為抽出もしくは有限母集団からの復元無作為抽出を考えている．

Ⅳ．ある駐車場に到着する車の台数はポアソン分布に従っており，1 時間に平均して 6 台の車が到着することがわかっている．この駐車場において，ある 20 分間に車が 1 台も到着しない確率は,

〔40〕〔41〕.〔42〕〔43〕〔44〕％

である．

Ⅴ．近畿地方にある D 大学の 2 回生 16 人の 1 か月の支出金額を調べたところ，標本平均が 8.2 万円，標本標準偏差が 2.2 万円という結果であった．この大学の 2 回生の 1 か月の支出金額が正規分布に従っているとして，以下の問いに答えよ．ただし，標本数 16 は小標本であるとする．

（1）母平均（母集団平均：μ）について信頼係数 99％で区間推定を行うと，信頼区間は,

〔45〕.〔46〕〔47〕〔48〕以上,〔49〕.〔50〕〔51〕〔52〕以下

となる．

（2）「D 大学 2 回生の 1 か月の支出金額が，関東圏の大学 2 回生の 1 か月の支出金額の平均 8.95 万円より低いといえるか」について有意水準 1％で検定を行った場合，検定統計量は,

【53】〔54〕.〔55〕〔56〕〔57〕

であり，検定統計量が臨界値よりも〔58〕にあるので，帰無仮説は〔59〕される．

〔58〕の【選択肢】

①外側（0 から遠い側）　　②内側（0 に近い側）

〔59〕の【選択肢】

①棄却　　②採択

※【53】は符号（正（＋）の場合「9」，負（－）の場合「0」）

※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．

（3）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,

〔60〕.〔61〕〔62〕〔63〕以上,〔64〕〔65〕.〔66〕〔67〕〔68〕以下

となる．

（4）この問題文にあるとおりの調査結果のもとで，母分散（母集団分散）が 3.24 である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,

〔69〕.〔70〕〔71〕〔72〕以上,〔73〕.〔74〕〔75〕〔76〕以下

となる．

（2 枚目に続く）

Ⅵ. 真の（母集団の）D 大学 2 回生の 1 か月の支出金額が平均（μ）＝8.0 万円，標準偏差（σ）＝1.2 万円の正規分布に従っていることがわかっている場合，〔77〕.〔78〕〔79〕〔80〕万円以上支出していれば，支出金額が上位 5%に入っているといえる．（分布表から正確な値が導出できない場合は，最も近い値の中点を取るとする）．

Ⅶ. 大卒以上の学歴を持つ男性有業者を対象に，理系学部および文系学部出身別に平均所得を調査した結果，表 3 のような結果を得た．これに基づいて「理系学部出身者の平均所得が文系学部出身者の平均所得よりも高いと言えるか」についての検定を行う．なお，2 つの母集団の分散は異なると仮定する．また，標本数が大きいため，t 分布は標準正規分布と一致していると考えてよい．

表 3　文理別平均所得（単位：万円）

	文理	標本数	平均値	標準偏差	平均値の標準誤差
課税前労働所得	理系	900	515	300	10.0
	文系	1600	490	300	7.5

（1）この検定を行うための（標準化された）検定統計量は，〔81〕.〔82〕〔83〕〔84〕である．

（2）この検定において，有意水準 5%と 1%の双方について検討する．この検定における P 値と有意水準との関係として最も適当なものを 1 つ選びなさい．　…………解答番号〔85〕

	有意水準 5%	有意水準 1%
①	P 値＞有意水準	P 値＞有意水準
②	P 値＞有意水準	P 値＜有意水準
③	P 値＜有意水準	P 値＞有意水準
④	P 値＜有意水準	P 値＜有意水準

（3）検定結果として最も適切なものを選びなさい．　…………解答番号〔86〕

①有意水準 5%では理系学部出身者の平均所得が文系学部出身者の平均所得より高いと言えるが，有意水準 1%では理系学部出身者の平均所得が文系学部出身者の平均所得より高いとは言えない．

②有意水準 5%では理系学部出身者の平均所得が文系学部出身者の平均所得より高いとは言えないが，有意水準 1%では理系学部出身者の平均所得が文系学部出身者の平均所得より高いといえる．

③有意水準 5%および有意水準 1%の双方において，理系学部出身者の平均所得は文系学部出身者の平均所得よりも高いと言える．

④有意水準 5%および有意水準 1%の双方において，理系学部出身者の平均所得は文系学部出身者の平均所得よりも高いとは言えない．

Ⅷ．回帰分析に関する以下の問いに答えよ．

（1）回帰分析に関する以下の文章〔87〕～〔93〕について，正しいものには「1」を，誤っているものには「0」をマークせよ．なお，問題文前の〔数字〕は解答番号である．

〔87〕標準的線形回帰モデルの仮定のもとでの誤差項（u_i）の平均は 0 である．

〔88〕標準的線形回帰モデルの仮定のもとでの誤差項（u_i）の分散は 1 である．

〔89〕標準的線形回帰モデルの仮定のもとでの誤差項（u_i）は独立である（自己相関していない）．

〔90〕標準的線形回帰モデルの仮定のもとでの誤差項（u_i）は標準正規分布に従う．

〔91〕標準的線形回帰モデルの仮定のもとでの誤差項（u_i）と独立変数とは独立であるから，独立変数は確率変数である．

〔92〕残差（誤差）とは，回帰直線と各データの間の最短の距離（データから回帰直線に対して垂直に引かれた線分の距離）のことである．

〔93〕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．

（2）回帰モデル$Y_i = a + bX_i + u_i$について，最小自乗法（最小二乗法）でパラメータbを推定する場合を考える．ここで、Y_iは従属変数，X_iは独立変数，u_iは標準的線形回帰モデルの仮定を満たす誤差項とする．以下の〔94〕および〔95〕の文章のうち**誤っているもの**を 1 つ選び，その番号をマークせよ．なお，〔数字〕は解答番号である．

〔94〕
①パラメータbの推定量$\hat{b}$は，不偏推定量である．
②パラメータbの推定量$\hat{b}$は，一致推定量である．
③パラメータbの推定量$\hat{b}$は，最良線形不偏推定量（BLUE）である．
④パラメータbの推定量$\hat{b}$は，決定係数といわれる．

〔95〕
①パラメータbの有意性を検定する際の帰無仮説は，$H_0 : b = 0$である．
②パラメータbの有意性を上側の片側検定で検定する際の対立仮説は，$H_1 : b > 0$である．
③パラメータbの推定量を$\hat{b}$としたとき，その標準化された検定統計量$\dfrac{\hat{b}}{s_{\hat{b}}}$（$s_{\hat{b}}$は標準誤差）は，自由度$n-2$の$t$分布に従う．
④上記の帰無仮説が棄却された場合，この 2 つの変数XとYとの間には有意な関係がないと判断される．

（以下，余白）