2018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．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．ただし，問題文中に指示されている場合および解答欄（解答番号）が特定の桁までで区切られている場合を除き，数値は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桁までを利用すること（最終的な解答も小数点以下第3位までを解答する）．

標準正規分布を用いる際に分布表から直接導出できない場合は，最も近い2つの値の中点とする．

問題内の，〔1〕，〔2〕……は，解答番号を表している．

ただし，**解答番号【14】【60】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．

【学生ID欄】左詰めで記入
2016年度以前生（学生ID8桁）は，必ずマーク部分の下2桁の「*」にマークすること．

I．ある6つの村における農業補助金額と米の生産量は下の表1の通りである．以下の問いに答えよ．**ただし，この問題の（1）については，小数点以下第2位を四捨五入し，小数点以下第1位までを解答せよ．（2）については整数値のみを解答せよ．また，（3）（4）については，表2中の数値を用いて計算せよ．**

表1　6カ村の農業補助金額と米の生産量

国	補助金額	米生産量
A村	28	42
B村	3	16
C村	16	38
D村	24	37
E村	9	18
F村	16	35

（単位）補助金額：100万円，米生産量：トン

表2　記述統計量

記述統計量	補助金額	米生産量
平均	16	31
分散	〔1〕〔2〕.〔3〕	123.2
標準偏差	〔4〕.〔5〕	11.1

（１）農業補助金額の分散と標準偏差を求めよ（解答番号は表２中にあるとおり）.

（２）米の生産量の中央値は〔6〕〔7〕であり，農業補助金額の最頻値は〔8〕〔9〕である.

（３）農業補助金額の変動係数は〔10〕.〔11〕〔12〕〔13〕，米の生産量の変動係数は 0.358 である.

（４）農業補助金と米の生産量との間の共分散は 93.2 であり，
この２変数間の相関係数は【14】〔15〕.〔16〕〔17〕〔18〕である.

（５）上の（３）（４）の結果からわかることの組み合わせとして適切なものを下の【選択肢】から選び，解答欄にマークせよ. ……………………解答番号〔19〕

（あ）変動係数で見た場合，米の生産量の散らばり具合の方が農業補助金額の散らばり具合よりも大きい.
（い）変動係数で見た場合，農業補助金額と米の生産量の単位が異なるので，その散らばり具合を比較することはできない.
（う）農業補助金額と米の生産量の散布図では，右上がりのプロットが描かれる.
（え）農業補助金額と米の生産量との間には正の相関があるので，農業補助金額を増やすことによって米の生産量が増大する.
（お）農業補助金額と米の生産量との間には正の相関があるので，米の生産量がより多い村により多くの農業補助金額が投入されているといえる.
（か）農業補助金額と米の生産量との間には負の相関があるので，米の生産量を増大させるためには農業補助金を減らせばよい.

【選択肢】
①（あ）と（う）が正しい　②（い）と（う）が正しい　③（あ）と（え）が正しい
④（い）と（か）が正しい　⑤（う）と（え）が正しい　⑥（う）と（お）が正しい
⑦（え）と（お）が正しい　⑧（あ）と（う）と（お）が正しい
⑨（い）と（え）と（お）が正しい

（裏面に続く）

Ⅱ．ベイズの定理（ベイズの公式）は，条件と結果を入れ替えた（因果関係を逆にした）確率を求める方法である．すなわち，$P(B|A)$ を使って $P(A|B)$ を求める方法である．ベイズの定理の導出は以下のステップで行う．

（１）条件付き確率の公式として最も適切なものを以下から選択せよ．………………………解答番号〔20〕

① $P(A|B)=\dfrac{P(A\cap B)}{P(B)}$　② $P(A|B)=\dfrac{P(A\cap B)}{P(A)}$　③ $P(A|B)=\dfrac{P(A\cup B)}{P(A)}$

④ $P(A|B)=\dfrac{P(A\cup B)}{P(B)}$　⑤ $P(A|B)=\dfrac{P(B)}{P(A\cap B)}$　⑥ $P(A|B)=\dfrac{P(A)}{P(A\cap B)}$

（２）一般に事象 A と事象 B が生じる場合について，$P(B)=P(A\cap B)+P(A^C\cap B)$ が成立する．これを応用すればベイズの定理が得られる．ベイズの定理として正しいものを以下から選択せよ．

……………解答番号〔21〕

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

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

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

（３）あるスポーツのワールドカップでは，過去に予選開始 3 ヶ月以内に監督が交代して予選通過をした確率は低いと言われている．予選突破を事象 A，予選敗退を事象 A^C とする．予選突破の確率は 0.5 とする．予選開始 3 ヶ月以内に監督が交代する事象を B とする．また，予選突破をしたチームの中で予選開始 3 ヶ月以内に監督が交代していた確率は 0.05 で，予選敗退したチームの中で予選開始 3 ヶ月以内に監督が交代していた確率は 0.25 であった．以上の確率をもとにして，予選開始 3 ヶ月以内に監督が交代して予選通過をする確率 $P(A|B)$ を求めると，

〔22〕.〔23〕〔24〕〔25〕

となる（この問題は，百分率（%）で答えるのではない）．

Ⅲ. 以下の（１）（２）について，（１）の〔26〕～〔30〕については最も適切なものを指定された【選択肢】から選び，その番号をマークせよ（同じ選択肢を複数回用いてもよい）.

（１）の〔31〕～〔35〕と（２）については適切な数値をマークせよ.

また（３）については，下線部（あ）～（き）に含まれる誤りを適切に修正するものを選択肢から選び，その番号をマークせよ．選択肢の前の番号は解答番号である.

（１）生じる結果（根元事象）が「ある事象が起こる（成功)」と「ある事象が起こらない（失敗)」の 2 種類しかない試行を〔26〕という．ここで，ある事象が起こる（成功）の確率が p であるとすると，ある事象が起こらない確率は〔27〕となる.

このような試行（実験）を n 回繰り返したときに x 回成功する確率は〔28〕によって表され，その平均は〔29〕，分散は〔30〕である.

したがって，赤玉 5 個，白玉 45 個（合計 50 個）の中から 1 つを引いて赤玉が出た場合を当たり（成功）とし，その場合 100 円の賞金がもらえるというゲームを 24 回続けて行ったとする（ただし，引いた玉は毎回元に戻す)．このとき，当たりとなる回数の平均は〔31〕.〔32〕回であり，期待される賞金額は〔33〕〔34〕〔35〕円である.

〔27〕〔29〕〔30〕【選択肢】

①p　②$(1-p)$　③np　④$np(1-p)$　⑤$p(1-p)$

⑥n　⑦μ　⑧σ^2　⑨$\dfrac{\sigma^2}{n}$

〔26〕〔28〕【選択肢】

①ベルヌーイ試行　②ポアソン分布　③二項分布　④自由度 n の t

⑤標準正規分布　⑥正規分布　⑦自由度 n の χ^2 分布　⑧自由度 $n-1$ の t

（２）ある自動車販売店では，30 日間で 15 台車が売れると言われている．この販売店において，6 日間で 3 台以上の車を売ることができる確率は〔36〕.〔37〕〔38〕〔39〕である（この問題は百分率（%）で答えるのではない).

（Ⅲの問題は 2 枚目に続いている）

（3）母集団の分布がいかなるものであっても，平均が μ ，分散が σ^2 の母集団から n 個の無作為標本を取り出したとき（無限母集団の復元抽出とする），その n 個の無作為標本から作られる標本平均 $\bar{X}$ の平均は

(あ) $\underline{\mu}$ ，分散は (い) $\underline{\sigma^2}$ となる．

この標本平均 $\bar{X}$ を母集団平均（ μ ）の推定量として用いた場合，標本平均 $\bar{X}$ が n の極限（ $n \to \infty$ ）において真の平均 μ に収束する（真の平均 μ から外れる確率が 0 となる）ことから (う) 不偏推定量である．そして，この性質を (え) 大数の法則という．また，その期待値（ $E(\bar{X})$ ）は真の平均 μ に等しくなることが証明されており，標本平均 $\bar{X}$ は母集団平均（ μ ）の (お) 不偏推定量でもある．なお，不偏推定量のうち，より (か) 分散が小さい方が効率的（有効）であるといい，この標本平均 $\bar{X}$ は最も効率的であることから (き) 最小推定量でもある．

〔40〕（あ）（い）について

① （あ）を $\underline{n\mu}$ に修正　　② （あ）を $\underline{p(1-p)}$ に修正

③ （い）を $\underline{\dfrac{\sqrt{n}(\bar{X}-\mu)^2}{\sigma}}$ に修正　　④ （い）を $\underline{\dfrac{\sigma^2}{n}}$ に修正

〔41〕（う）（え）（お）について

① （う）を普遍推定量に修正　　② （う）を一致推定量に修正

③ （え）を中心極限定理に修正　　④ （お）を普遍推定量に修正

〔42〕（か）（き）について

① （か）を確率に修正　　② （か）を標本成功率に修正

③ （き）を有意推定量に修正　　④ （き）最小分散不偏推定量に修正

（次ページに続く）

Ⅳ. 年間 25 試合に登板する二刀流メジャーリーガーO選手のピッチングの球速は, 平均 135km/h, 分散 2500 の正規分布に従うことがわかっている. O選手のピッチングの球速が 90km/h 以上 160km/h 以下となる確率は〔43〕〔44〕.〔45〕〔46〕〔47〕%である（この問題は，計算途中では四捨五入しないこと）.

Ⅴ. スペインのプロサッカーチームから移籍してきたⅠ選手のシュートのスピードを無作為に 16 回調べたところ，標本平均が 135km/h，標本分散が 2500 という結果であった. Ⅰ選手のシュートのスピード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 95%で区間推定を行うと，信頼区間は,
〔48〕〔49〕〔50〕.〔51〕〔52〕〔53〕以上,〔54〕〔55〕〔56〕.〔57〕〔58〕〔59〕以下
となる.

（2）「Ⅰ選手の平均シュートスピードが，日本人プロサッカー選手H選手のシュートのスピードの平均 110km/h より速いといえるか」について有意水準 5%で検定を行った場合，検定統計量は,
【60】〔61〕.〔62〕〔63〕〔64〕
であり，帰無仮説は〔65〕される.

〔65〕の【選択肢】
①棄却　　　　②採択

※【60】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 90%で区間推定を行うと，信頼区間の整数部分は,
〔66〕〔67〕〔68〕〔69〕以上,〔70〕〔71〕〔72〕〔73〕以下となる
（この問題では，計算途中は小数点以下第 4 位を四捨五入して小数点以下第 3 位までを利用し，最終的な解答では小数点以下第 1 位を四捨五入して整数値のみを解答すること）.

（4）この問題文にあるとおりの調査結果のもとで，母集団の標準偏差が 50km/h 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 99%で区間推定を行うと，信頼区間は,
〔74〕〔75〕〔76〕.〔77〕〔78〕〔79〕以上,〔80〕〔81〕〔82〕.〔83〕〔84〕〔85〕以下となる.

（裏面に続く）

VI. 全国の中学 3 年生を対象とした数学に関する学力調査結果によると，都市部に住むグループと非都市部に住むグループでは，得点に関して次表のような結果を得ている．

都市部と非都市部	標本数	数学の平均得点	数学得点の標本標準偏差
都市部	400	68	22
非都市部	400	65	20

これら 2 つのグループの母集団分散（標準偏差）が異なるとして，都市部と非都市部の数学得点の平均値が等しいと言えるかについて検定を行うことにする．両グループの標本数が多いことから正規分布を用いた検定を行う．

（1）この場合の（標準化された）「検定統計量」は，絶対値で〔86〕**.**〔87〕〔88〕〔89〕となり，

P 値は〔90〕**.**〔91〕〔92〕〔93〕となる

（P 値は，求めた検定統計量の小数点以下第 3 位を四捨五入して，小数点以下第 2 までを利用して求めること．ただし，求めた P 値は小数点以下第 4 位を四捨五入して小数点以下第 3 位までを解答すること（百分率（%）で答えるのではない））.

（2）対立仮説を「数学得点の平均は都市部の方が非都市部より高い」として，有意水準 5%で片側検定したとする．この検定結果として最も適切なものを選択せよ．　…………………解答番号〔94〕

① （標準化された）「検定統計量」が臨界値よりも小さいので，数学得点の平均は都市部の方が高いと言える．
② （標準化された）「検定統計量」が臨界値よりも大きいので，数学得点の平均は都市部の方が高いとは言えない．
③ P 値が有意水準よりも小さいので，数学得点の平均は都市部の方が高いとは言えない．
④ P 値が有意水準よりも小さいので，数学得点の平均は都市部の方が高いと言える．
⑤ P 値が有意水準よりも大きいので，数学得点の平均は都市部の方が高いと言える．
⑥都市部と非都市部にかかわらず，数学得点の平均値の差については，統計的に何も言えない．

（次ページに続く）

Ⅶ. X_i を独立変数（説明変数），Y_i を従属変数（被説明変数）とする線形回帰モデル $Y_i = a + bX_i + u_i$ を最小二乗法（最小自乗法）で推定（検定）する場合を考える．なお，u_i は標準的線形回帰モデルの仮定を満たす誤差項である．

以下の文章〔95〕～〔98〕について，内容が正しいものの組み合わせを以下の選択肢から選び，その番号をマークせよ．

解答番号〔95〕

（あ）標準的線形回帰モデルの仮定において誤差項は，平均が 0，分散が 1 の正規分布に従う

（い）標準的線形回帰モデルの仮定において誤差項は，独立変数と独立である．

（う）標準的線形回帰モデルの仮定において誤差項の分散は均一である．

解答番号〔96〕

（あ）残差とは，各データと回帰直線との最短の距離（データから回帰直線への垂線の距離）のことである．

（い）残差とは，X 軸へ垂直の方向（縦方向）の，各データと回帰直線との距離のことである．

（う）最小二乗法とは，残差の和を最小にする回帰直線を引く方法である．

解答番号〔97〕

（あ）最小二乗法で係数 b について検定する場合，帰無仮説は $H_0 : b = 0$ とする．

（い）最小二乗法で係数 b が正であるといえるかについて検定する場合，対立仮説は $H_0 : b > 0$ とする．

（う）最小二乗法で係数 b が正であるといえるかについて検定する場合，両側検定を行う．

解答番号〔98〕

（あ）最小二乗法で係数 b が正であるといえるかについて検定する場合，検定統計量（t 値）が臨界値よりも大きければ（外側にあれば）帰無仮説が棄却される．

（い）最小二乗法で係数 b が正であるといえるかについて検定する場合，帰無仮説が棄却されれば，2 変数（独立変数と従属変数）との間に有意な関係があるといえる．

（う）最小二乗法における決定係数は，回帰モデルの説明力を表している．

【選択肢】

①（あ）のみが正しい ②（い）のみが正しい ③（う）のみが正しい

④（あ）と（い）が正しい ⑤（あ）と（う）が正しい ⑥（い）と（う）が正しい

⑦（あ）（い）（う）とも正しい ⑧（あ）（い）（う）とも誤り

（以下余白）