2017 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【13】【18】【75】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（下 2 桁は空欄）．マーク部分の下 2 桁には必ず「＊」にマークすること．

Ｉ．６つの地域の高齢化率（全人口に占める 65 歳以上の人口の割合）と，ある政党の支持率を調べたところ表 1 のようになった．各種記述統計量は表 2 のとおりである．なお，（２）（３）については，表中に示されている数値を用いて導出すること．

表 1　高齢化率と政党支持率

地区	高齢化率（%）	政党支持率（%）
A	30	28
B	18	43
C	25	〔1〕〔2〕
D	38	21
E	23	40
F	20	45

表 2　各種記述統計量

統　計　量	高齢化率	政党支持率
標 本 平 均	25.667	34
標 本 分 散	53.867	〔3〕〔4〕.〔5〕
標本標準偏差	7.339	〔6〕.〔7〕〔8〕〔9〕
変 動 係 数	0.286	0.〔10〕〔11〕〔12〕
標本共分散	− 65.8	
相 関 係 数	【13】〔14〕.〔15〕〔16〕〔17〕	

（1）表 1 中の地区 C の数値を求め，マークせよ．

（2）表 2 中の政党支持率の標本分散，標本標準偏差，変動係数，および高齢化率と政党支持率の相関係数を求め，適切な数値および符号をマークせよ．（符号は【13】）

（3）地区 E の高齢化率を標準化すると【18】〔19〕．〔20〕〔21〕〔22〕となる．（符号は【18】）

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〔23〕

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

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

（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〔24〕

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

③ $P(A|B)=\dfrac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^C)P(A^C)}$　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）ある電子製品には，純正品と偽物が存在している。純正品を A，偽物を A^C とする。純正品の市場シェアは 0.8，偽物のシェアは 0.2 である。外見は純正品と偽物は同じであるが，不良品（B）である比率が異なっている。不良品である比率は，純正品で 0.02，偽物で 0.30 である。

以上の確率をもとにして，購入した製品が不良品 B であった時に，それが純正品 A である確率 $P(A|B)$ を求めると，

〔25〕．〔26〕〔27〕〔28〕

となる．（この問題は，百分率（%）で答えるのではない）

（裏面に続く）

Ⅲ. 以下の（1）～（3）について，求めた数値，確率をマークせよ．ただし，〔33〕については求めた金額を下の【選択肢】から最も適切なものを選び，その番号をマークせよ．

（1）ある店の客単価の平均（1 人あたりの平均支払額）が 2400 円であるとする．この店では，会計の際に 20 本中 3 本の当たりが入ったくじを引いてもらい，当たりが出た場合にはその支払いを無料にするというキャンペーンを行おうとしている．なお，くじは引くたびに元に戻すことになっている．25 人に連続してこのくじ引きを行った場合，無料になる人の平均は〔29〕.〔30〕〔31〕〔32〕人であるから，この店は〔33〕円をこのキャンペーンのコストとして見込んでおかなければならない．

〔33〕の【選択肢】

①5,000　　②7,000　　③9,000　　④12,000

（2）ある池では 1 時間に平均 2 匹魚が釣れる．ある釣り人が 3 時間に 5 匹以上の魚を釣ることができる確率は〔34〕.〔35〕〔36〕〔37〕である．（この問題は百分率（%）で答えるのではない）

（3）プロ野球の M 選手の打球の飛距離は，平均 78.35m，標準偏差 7.5m の正規分布に従うことがわかっている．ただし，M 選手は無心にバットを振っており，飛距離を全くコントロールしていない．また，ゴロになった打球については飛距離の分布には含まれていない．95m 以上の飛距離であればホームランとなるとき，M 選手がホームランを打つ確率は〔38〕.〔39〕〔40〕〔41〕%である．

Ⅳ. 以下の【A】【B】の状況のもとで，それぞれの語句説明文が示す最も適切な語句を下の【選択肢】から選び，その番号をマークせよ．

【A】の状況

標本空間中に事象 A と事象 B があると考える．

〔42〕事象 A と事象 B に共通部分がない（事象 A と事象 B の積事象が空事象である）場合，事象 A か事象 B のどちらかが生じる確率は，事象 A が生じる確率と事象 B が生じる確率を足せばよいという性質（確率は足すことができるという性質）．

〔43〕事象 A と事象 B に共通部分がある（事象 A と事象 B の積事象が空事象でない）場合，事象 A か事象 B のどちらかが生じる確率が $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ で表されるということ．

〔44〕事象 A と事象 B の積事象が $P(B|A)P(A)$ （もしくは $P(A|B)P(B)$ ）で表されるということ．

（【選択肢】および問題の続きは右ページ）

【B】の状況

母集団平均（母平均）が μ，母集団分散（母分散）が σ^2 の母集団から n 個の標本を抽出（無限母集団からの無作為復元抽出）し，その無作為標本を用いて標本平均（$\bar{X}$）を求め，母集団平均（μ）の推定量とする．

〔45〕標本数 n が大きいとき，母集団の分布（母分布）がいかなる場合であっても，標本平均（$\bar{X}$）が正規分布に従う（もしくは，標本平均を標準化した確率変数が標準正規分布に従う）こと．

〔46〕母集団の分布（母分布）がいかなる場合であっても，標本平均（$\bar{X}$）が，標本数 n の極限（$n \to \infty$）において母集団平均（μ）から外れる確率が 0 となること．

〔47〕母集団の分布（母分布）がいかなる場合であっても，標本平均（$\bar{X}$）が，標本数 n の極限（$n \to \infty$）において母集団平均（μ）に収束することによって示される推定量としての望ましい性質．

〔48〕母集団の分布（母分布）がいかなる場合であっても，標本平均（$\bar{X}$）の期待値（$E(\bar{X})$）が，母集団平均（μ）に等しくなるという推定量としての望ましい性質．

〔42〕〔47〕〔48〕の【選択肢】

①不偏性　②加法性　③効率性（有効性）　④乗法性
⑤一致性　⑥独立性　⑦従属性　⑧連続性　⑨再生性

〔43〕〔44〕〔45〕〔46〕の【選択肢】

①中心極限定理　②チェビシェフの不等式　③乗法定理　④加法定理
⑤大数の法則　⑥同時確率　⑦周辺確率　⑧条件付き確率

（2 枚目に続く）

Ｖ．スマートフォンのある機種の連続待ち受け時間は 570 時間と公表されている．なお，この機種の連続待ち受け時間の分布は正規分布に従うことがわかっている．この機種の工場において 9 個の製品を無作為に抽出し，連続待ち受け時間を調べたところ以下の結果を得た（単位：時間）.

568,　572,　580,　571,　573,　588,　569,　573,　572

以下の（１）～（５）について，求めた数値，符号，適した選択肢の番号をマークせよ．
ただし，（２）～（４）については，（５）の情報を用いないこと．

（１）この結果から導かれる連続待ち受け時間の標本平均は〔49〕〔50〕〔51〕時間，標本分散は 39（標本標準偏差は 6.245）である．

（２）この機種の連続待ち受け時間の母集団平均（母平均）μ を信頼係数 90%で区間推定を行うと，信頼区間は

〔52〕〔53〕〔54〕.〔55〕〔56〕〔57〕～〔58〕〔59〕〔60〕.〔61〕〔62〕〔63〕時間

である．

（３）この機種の連続待ち受け時間の母集団分散（母分散）σ^2 を信頼係数 95%で区間推定を行うと，信頼区間は

〔64〕〔65〕.〔66〕〔67〕〔68〕～〔69〕〔70〕〔71〕.〔72〕〔73〕〔74〕

である．

（４）この調査をもとに，この機種の連続待ち受け時間が公表値である 570 時間を超えていると言えるかについて有意水準 1%で検定を行うと，（標準化された）検定統計量は【75】〔76〕.〔77〕〔78〕〔79〕であり，〔80〕.

〔80〕の【選択肢】

①臨界値は 1.96 なので帰無仮説を棄却する．したがって，連続待ち受け時間は公表値 570 時間を超えていると言える

②臨界値は 2.23（2.225）なので帰無仮説を棄却する．したがって，連続待ち受け時間は公表値 570 時間を超えているとは言えない

③臨界値は 2.58（2.575）なので帰無仮説を棄却しない．したがって，連続待ち受け時間は公表値 570 時間を超えていると言える

④臨界値は 2.897 なので帰無仮説を棄却しない．したがって，連続待ち受け時間は公表値 570 時間を超えているとは言えない

⑤臨界値は 3.355 なので帰無仮説を棄却する．したがって，連続待ち受け時間は公表値 570 時間を超えていると言える

（Ｖの問題は右ページに続く）

（5）仮にこの機種の連続待ち受け時間の母集団分散（σ^2）が 56.25 であることがわかっている場合，この調査をもとに，この機種の連続待ち受け時間が公表値である 570 時間を超えていると言えるかについて有意水準 5%で検定を行うとする．このとき，片側検定での P 値は〔81〕.〔82〕〔83〕〔84〕であり，〔85〕.

〔85〕の【選択肢】

①P 値が有意水準よりも大きいので, 帰無仮説を棄却する．したがって, 連続待ち受け時間は 570 時間を超えていると言える.

②P 値が有意水準よりも大きいので, 帰無仮説を棄却する．したがって, 連続待ち受け時間は 570 時間を超えているとは言えない.

③P 値が有意水準よりも大きいので，帰無仮説を棄却しない．したがって，連続待ち受け時間は 570 時間を超えているとは言えない.

④P 値が有意水準よりも小さいので, 帰無仮説を棄却する．したがって, 連続待ち受け時間は 570 時間を超えていると言える.

⑤P 値が有意水準よりも小さいので, 帰無仮説を棄却する．したがって, 連続待ち受け時間は 570 時間を超えているはと言えない.

⑥P 値が有意水準よりも小さいので，帰無仮説を棄却しない．したがって，連続待ち受け時間は 570 時間を超えているとは言えない.

（裏面に続く）

VI. 2017 年に行われたある調査結果によると，22 歳までに海外旅行の経験が無いグループと有るグループ別に，現在の年収に関して次表のような結果を得ている．

海外旅行経験の有無	標本数	個人年収の標本平均（単位：万円）	個人年収の標本標準偏差（単位：万円）
無し	900	375	250
有り	400	410	300

これら 2 つのグループの母集団標準偏差が異なるとして，海外旅行経験の有無にかかわらず，個人年収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〕となる．

（2）対立仮説を「海外旅行経験の有無によって，個人年収に差がある」として，有意水準 5%で両側検定したとする．この検定結果として最も適切なものを選択せよ．ただし，臨界値が分布表から直接導出できない場合には，最も近い 2 つの値の中点とする．　…………………解答番号〔90〕

①検定統計量が棄却域に入るので，海外旅行経験の有無にかかわらず，個人年収に差がない．
②検定統計量が棄却域に入らないので，海外旅行経験の有無にかかわらず，個人年収に差がない．
③検定統計量が棄却域に入るので，海外旅行経験の有無によって，個人年収に差がある．
④検定統計量が棄却域に入らないので，海外旅行経験の有無によって，個人年収に差がある．
⑤海外旅行経験の有無と個人年収に差との関係については，統計的に何も言えない．

（右ページにも問題は続いている）

Ⅶ. 回帰分析に関する以下の文章について，空欄（〔解答番号〕）に入る最も適切なものを指定された【選択肢】から１つ選び，その番号をマークせよ．ただし，〔98〕については適切な数値をマークせよ．

回帰モデル$Y_i = a + bX_i + u_i$を最小二乗法(最小自乗法)で推定する場合を考える．ここでX_iは〔91〕，Y_iは〔92〕といわれる．また，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．

最小二乗法でbについて検定する場合，帰無仮説をH_0:〔93〕として，標準化された検定統計量$\frac{\hat{b}}{s_{\hat{b}}}$（$s_{\hat{b}}$は標準誤差）が〔94〕に従うことを用いる．なお，たとえば回帰係数bが正であると言えるかを検定する場合の対立仮説はH_1:〔95〕である．検定によって帰無仮説が〔96〕場合，2つの変数X_iとY_iには有意な関係があるとされる．

さて，誤差項（u_i）に関する標準的線形回帰モデルの仮定には，誤差項は〔97〕に従っており，その平均は〔98〕，分散が〔99〕であることなどが含まれる．そして，この仮定のもとでの回帰係数bの推定量$\hat{b}$は〔100〕．

〔91〕〔92〕〔99〕の【選択肢】

①従属変数　②独立変数　③均一　④一致

〔93〕〔95〕の【選択肢】

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

〔94〕〔97〕の【選択肢】

①正規分布　②標準正規分布　③自由度$n-1$のt分布　④自由度$n-2$のt分布　⑤自由度$n-1$のχ^2分布　⑥自由度$n-2$のχ^2分布　⑦ベルヌーイ確率分布

〔96〕の【選択肢】

①棄却される　②棄却されない

〔100〕の【選択肢】

①不偏推定量であるが最良推定量ではない
②不偏推定量であり，最良推定量である
③不偏推定量ではない

（問題終了）