# 2014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】【74】**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．ある村に存在する6戸の農家（Hさん～Mさん）について，農場（田んぼ）の大きさとそこで生産される1ha（ﾍｸﾀｰﾙ）あたりの米の生産量（土地生産性）の調査を行い，下の表1のような結果を得た．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　農家の農場規模と1haあたりの米の生産量

| | 農場規模 | 土地生産性 |
|---|---|---|
| H | 3.0 | 4.2 |
| I | 1.2 | 5.2 |
| J | 0.8 | 5.4 |
| K | 2.5 | 4.6 |
| L | 1.3 | 5.1 |
| M | 2.0 | 4.3 |

単位）農場規模：ha

土地生産性：t（トン）

表2　各種記述統計量

| 統計量 | 農場規模 | 土地生産性 |
|---|---|---|
| 平　均 | 1.800 | 4.800 |
| 分　散 | 〔1〕．〔2〕〔3〕〔4〕 | 0.252 |
| 標準偏差 | 〔5〕．〔6〕〔7〕〔8〕 | 0.502 |
| 変動係数 | 0.47 | 〔9〕．〔10〕〔11〕〔12〕 |
| 共分散 | －0.390 | |
| 相関係数 | **【13】**〔14〕．〔15〕〔16〕〔17〕 | |

（1）表2中の，農場規模の標本分散と標本標準偏差，土地生産性の変動係数，および両変数間の相関係数を求め，適切な数値および符号（符号は**【13】**）をマークせよ．

（２）表 2 で導いた相関係数から読み取ることができることとして最も適切なものを 1 つ選び，その番号をマークせよ． …………………………解答番号〔18〕

①農場規模と土地生産性との間には，強い正の相関がある．
②農場規模と土地生産性との間には，弱い正の相関がある．
③農場規模と土地生産性とは無相関である．
④農場規模と土地生産性との間には，弱い負の相関がある．
⑤農場規模と土地生産性との間には，強い負の相関がある．
⑥農場規模と土地生産性との間には，相関があるかわからない．

Ⅱ．次の（1）（2）の〔19〕～〔30〕に適切な数値をマークせよ．また，（3）の〔31〕～〔36〕には最も適切なものを指定された【語群】から 1 つ選び，その番号をマークせよ．さらに，（4）では，各語句説明文が表す内容として最も適切なものを下の【語群】から 1 つ選び，その番号をマークせよ．ただし，語句説明文の前にある〔37〕～〔39〕は解答番号である．

（1）ゆがみのないコインを 2 枚同時に投げるゲームを 40 回繰り返して行った場合，2 枚とも表が出る回数の平均は〔19〕〔20〕，分散は〔21〕．〔22〕〔23〕〔24〕である．

（2）サッカーワールドカップのブラジル代表チームは，3 分間に平均して 0.1 本のシュートを決めるという．このチームが 1 試合（90 分間）で平均的に決めるシュートの数は〔25〕本であり，1 試合で 2 本以上シュートを決める確率は〔26〕〔27〕．〔28〕〔29〕〔30〕**%**である．

（3）平均が $\mu$ ，分散が $\sigma^2$ 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 $n$ 個の標本から作られる標本平均 $\bar{X}$ の平均は〔31〕，分散は〔32〕となる．この標本平均 $\bar{X}$ を母集団平均（ $\mu$ ）の推定量として用いた場合，標本平均 $\bar{X}$ は $n$ の極限（ $n \to \infty$ ）において真の平均 $\mu$ に収束する（真の平均 $\mu$ から外れる確率が 0 となる）ことから〔33〕推定量であり，その期待値（ $E(\bar{X})$ ）が真の平均 $\mu$ に等しくなることから〔34〕推定量でもある．さらに，この標本平均 $\bar{X}$ は,〔34〕推定量のうち〔35〕が最も小さいことから最も効率的であるといい，最小〔35〕〔34〕推定量，もしくは〔36〕推定量でもある．

〔31〕〔32〕の【語群】

| | | | | |
|---|---|---|---|---|
| ① $n$ | ② $n-1$ | ③ $\lambda$ | ④ $\frac{1}{\lambda^2}$ | ⑤ $\frac{1}{k^2}$ |
| ⑥ $\frac{1}{n}$ | ⑦ $\frac{\sigma^2}{n}$ | ⑧ $p$ | ⑨ $\mu$ | |

〔33〕～〔36〕の【語群】

| | | | | |
|---|---|---|---|---|
| ①不変 | ②不偏 | ③普遍 | ④一致 | ⑤最良 |
| ⑥平均 | ⑦分散 | ⑧中央値 | ⑨線形 | |

**（裏面につづく　Ⅱの問題は（4）まである）**

（４）〔37〕上の（３）の文章中の下線部が表しているもの

〔38〕上の（３）の標本平均 $\bar{X}$ において，標本数 $n$ が大のとき，これを標準化した確率変数 $Z$ が標準正規分布に近似できるという性質

〔39〕事象 A と事象 B という 2 つの事象のもとで成立する $P(\mathrm{A} \cup \mathrm{B}) = \mathrm{P(A)} + \mathrm{P(B)} - \mathrm{P}(\mathrm{A} \cap \mathrm{B})$

【語群】

| | | |
|---|---|---|
| ①チェビシェフの不等式 | ②確率の加法定理 | ③確率の乗法定理 |
| ④互いに背反な事象 | ⑤互いに独立な事象 | ⑥中心極限定理 |
| ⑦大数の法則 | ⑧根元事象 | ⑨標本空間 |

Ⅲ．ベイズの定理は，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〔40〕

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〔41〕

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

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

（３）サッカーワールドカップで，1 次リーグを突破するという事象を $A$ とおき，1 次リーグ初戦で負ける事象を $B$ とする． $A$ の割合は，4 チームのうち 2 チームが決勝トーナメント進出することから 50%と考える．今，これまでのデータから，決勝トーナメント進出した（1 次リーグを突破した）チームのうち，初戦で負けたチームの比率が 6%であることが分かっている．また，1 次リーグで敗退したチームのうち，初戦で負けたチームの比率は 56%であることもわかっている．

以上の確率をもとにして，「初戦で負けたチーム」（ $B$ ）が，「1 次リーグを突破する」（ $A$ ）確率 $P(A|B)$ を求めると，

〔42〕．〔43〕〔44〕〔45〕

となる．（この問題は，百分率（%）で答えるのではない．）

Ⅳ．以下の（１）～（４）の各解答番号に適切な数値をマークせよ．ただし，**【74】**には正負の符号を冒頭の説明に従ってマークせよ．また，〔83〕には検定結果として最も適切なものを下の**【選択肢】**の中から１つ選び，その番号をマークせよ．

※なお，（２）（３）を解答する際には（４）の情報は用いないこととする．

（１）アルゼンチン代表選手であるメッシのシュートスピードは，母平均（母集団平均）が 80km/h，母集団標準偏差が 16km/h の正規分布に従うことが知られている．このとき，メッシのシュートが 88km/h 以上となる確率は，〔46〕〔47〕．〔48〕〔49〕〔50〕**%**となる．

（２）ブラジル代表選手であるネイマールのシュートスピードは，正規分布に従うことがわかっている．ここで，ネイマールが蹴ったシュートを 16 本調べたところ，その平均は 90km/h，標準偏差は 20km/h であった．ネイマールのシュートのスピードについて信頼係数 99%でその母平均（母集団平均）に関する区間推定を行うと，信頼区間は

〔51〕〔52〕．〔53〕〔54〕〔55〕以上，〔56〕〔57〕〔58〕．〔59〕〔60〕〔61〕以下

となる．

（３）ネイマールのシュートのスピードについて，その母分散（母集団分散）の区間推定を信頼係数 90%で行うと，信頼区間は

〔62〕〔63〕〔64〕．〔65〕〔66〕〔67〕以上，〔68〕〔69〕〔70〕．〔71〕〔72〕〔73〕以下

となる．

（４）今回のワールドカップ出場選手全員のシュートの平均速度は 84km/h である．ネイマールのシュートのスピードの母集団標準偏差が 16km/h であることがわかっているとした場合，「ネイマールのシュートはワールドカップ出場選手の平均値より速い」と言えるかについて有意水準 5%で検定する．

（標準化された）「検定統計量」は**【74】**〔75〕．〔76〕〔77〕〔78〕となり，$P$ 値は〔79〕．〔80〕〔81〕〔82〕**%**であるから，〔83〕．

※ここでの解答には，上記（２）のネイマールのシュートスピードに関する情報を利用する．

〔83〕の**【選択肢】**

①$P$ 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，ネイマールのシュートスピードはワールドカップ出場選手全体よりも速いといえる

②$P$ 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，ネイマールのシュートスピードはワールドカップ出場選手全体よりも速いとはいえない

③$P$ 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，ネイマールのシュートスピードはワールドカップ出場選手全体よりも速いといえる

④$P$ 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，ネイマールのシュートスピードはワールドカップ出場選手全体よりも速いとはいえない

（2 枚に目につづく）

V．2013 年に行われたある国際比較調査結果によると，被雇用者の勤続年数に関して次表のような結果を得ている．

| 国 | 標本数 | 勤続年数の標本平均（単位：年） | 勤続年数の標本標準偏差（単位：年） |
|---|---|---|---|
| 日本 | 100 | 12.5 | 11.0 |
| ドイツ | 100 | 11.0 | 10.0 |

これら 2 国の母集団標準偏差が異なるとして，日本とドイツの平均勤続年数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〔84〕．〔85〕〔86〕〔87〕となる．

（2）対立仮説を「日本とドイツの平均勤続年数に差がある」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．　………解答番号〔88〕

①検定統計量が棄却域に入るので，日本とドイツの平均勤続年数に差がない．

②検定統計量が棄却域に入らないので，日本とドイツの平均勤続年数に差がない．

③検定統計量が棄却域に入るので，日本とドイツの平均勤続年数が等しいとは言えない．

④検定統計量が棄却域に入らないので日本とドイツの平均勤続年数が等しいとは言えない．

⑤日本とドイツの平均勤続年数に差があるかどうかは判らない．

（右ページにつづく）

VI. 回帰分析に関する以下の文章の各解答番号部分について，下の解答番号につづく選択肢の中から最も適切なものを 1 つ選び，その番号をマークせよ．

変数 $X$ と $Y$ の関係を，線形回帰モデル $Y_i = a + bX_i + u_i$ （$i = 1,\cdots,n$， $u_i$ は誤差項）を用いて検定する場合を考える．このとき，$X$ を〔89〕，$Y$ を〔90〕という．さて，推定されるパラメータ $b$ に関する帰無仮説は〔91〕であり，（$a$ が有意であり，）この帰無仮説が〔92〕場合にこの 2 変数間には $Y = a + bX$ という関係があるとみなされる．そして，この線形回帰モデルの説明力を表す指標は〔93〕である．

このときの検定統計量 $\dfrac{\hat{b}}{s_{\hat{b}}}$ （$s_{\hat{b}}$ は標準誤差）は〔94〕に従うとして検定するが，この検定を行うためには，誤差項（$u_i$）が通常の誤差項が満たすべき条件（標準的線形回帰モデルの仮定）を満たしていなければならない．この標準的線形回帰モデルの仮定では，誤差項は独立であり，平均が〔95〕，分散が均一な〔96〕に従うとされ，さらに〔89〕である $X$ は〔97〕とされる．そして，この標準的線形回帰モデルの仮定を満たしている場合，パラメータ $b$ の推定量（$\hat{b}$）は〔98〕となる．

〔89〕①独立変数　②従属変数
〔90〕①独立変数　②従属変数
〔91〕① $b > 0$　② $b < 0$　③ $b = 0$　④ $b \neq 0$
〔92〕①棄却される　②棄却されない
〔93〕①ジニ係数　②決定係数　③変動係数
〔94〕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
〔95〕①0　②1　③ $\mu$　④ $\sigma^2$
〔96〕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
〔97〕①確率変数である　②確率変数ではない
〔98〕①LED　②RED　③BLUES　④BLUE

（以下余白）