2010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．

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．次の文章の（1）～（3）について，最も適切な語句を指定された【語群】から選択し，その番号をマークせよ．ただし，同じ選択肢を複数回用いてもよい．また，〔6〕〔7〕〔16〕については適切な数値をマークせよ．

小問（4）に関しては，最も適切な選択肢の番号をマークせよ．

（1）母集団から無作為抽出された標本から計算される標本平均（$\overline{X}$）や標本分散（S^2）などが，真のパラメータ（母集団のパラメータ，母数）の推定に用いられる場合，これを〔1〕という．〔1〕の望ましさの指標としては，標本数nの極限（$n \to \infty$）において〔1〕が真のパラメータに収束する場合の〔2〕性や，〔1〕の〔3〕が真のパラメータの値に等しくなる場合の〔4〕性などが重要である．

【語群】

①P値　②F値　③絶対値　④期待値　⑤推定量
⑥統計量　⑦一致　⑧不偏　⑨効率（有効）

（2）平均＝μ，分散＝σ^2の確率変数Xの標準化は〔5〕という式で与えられる．標準化された確率変数Zの平均は〔6〕，分散は〔7〕である．この確率変数Xが正規分布である場合には，標準化変数Zは〔8〕となる．

〔5〕の【語群】

① $Z=\sqrt{\dfrac{X-\mu}{\sigma^2}}$　② $Z=\sqrt{\dfrac{(X-\mu)^2}{\sigma}}$　③ $Z=\sqrt{\dfrac{X-\mu}{\sigma}}$　④ $Z=\dfrac{X-\mu}{\sigma^2}$

⑤ $Z=\dfrac{(X-\mu)^2}{\sigma}$　⑥ $Z=\dfrac{X-\mu}{\sigma}$　⑦ $Z=E(X)$　⑧ $Z=E(X-\mu)$

⑨ $Z=E\left[(X-\mu)^2\right]$

〔8〕の【語群】

①ベルヌーイ確率分布　②二項分布　③ポアソン分布
④一様分布　⑤正規分布　⑥標準正規分布
⑦ χ^2 分布　⑧ t 分布

（3）2 つの事象（事象 A と事象 B）において，これらに共通な事象が存在するとき，この和事象が生じる確率は〔9〕で計算される．

また，事象 A が生じたもとで事象 B が生じる確率を〔10〕といい，その計算式は〔11〕である．これを用いれば，〔12〕を求める式〔13〕が導出され，これを「確率の〔14〕」という．

さらに，これら 2 つの事象に共通な事象がないとき，〔12〕は〔15〕となり，その確率は〔16〕である．

〔9〕〔11〕〔13〕の【語群】

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

③ $P(A\cup B)=P(A)+P(B)+P(A\cap B)$　④ $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

⑤ $P(A\cup B)=P(A)+P(B)$　⑥ $P(A\cap B)=P(A)+P(B)$

⑦ $P(A\cap B)=P(B\,|\,A)P(B)$　⑧ $P(A\cap B)=P(B\,|\,A)P(A)$

〔10〕の【語群】

①スタージェスの公式　②等確率　③周辺確率　④条件付き確率
⑤制約された標本空間　⑥同時確率　⑦確率関数　⑧密度関数
⑨分布関数

〔12〕〔14〕〔15〕の【語群】

①加法性　②加法定理　③乗法定理
④ベイズの定理　⑤積事象　⑥空事象
⑦余事象　⑧補事象　⑨根元事象

（4）中心極限定理の説明として最も適切なものを選び，その番号をマークせよ．

母平均（母集団平均）＝μ，母分散（母集団分散）＝σ^2のとき，（無限母集団からの復元）無作為抽出によるn個の無作為標本による標本平均（$\overline{X}$）について考える．

①中心極限定理とは，標本数nが極限（$n \to \infty$）のときに標本平均（$\overline{X}$）が母集団平均（μ）から外れる確率が 0 となることを表す定理である．

②中心極限定理は，標本平均（$\overline{X}$）の分散が$\dfrac{\sigma^2}{n}$であることによって証明される．

③中心極限定理は，標本平均（$\overline{X}$）の分散がnpであることによって証明される．

④中心極限定理によって，母集団分布がいかなる場合でも標本数nが大のとき，標本平均（$\overline{X}$）の標準化変数は標準正規分布に近似できる．

解答番号〔17〕

Ⅱ．次の表は，20 世紀後半の 40 年間におけるアメリカと日本の失業率（年平均）をまとめたものである．このデータをもとにして，以下の問に答えよ．　　単位）％

失業率／年	1960-1967	1968-1973	1974-1979	1980-1990	1990-2000
アメリカ	5.0	4.6	6.7	7.0	5.6
日　本	1.3	1.2	1.9	2.5	3.2

出所）春田素夫・鈴木直次（2005）『アメリカの経済』（第 2 版）p.145，表 5-1.

（1）日本における失業率の変動係数は〔18〕．〔19〕〔20〕〔21〕である．

（2）両国の失業率の共分散は〔22〕．〔23〕〔24〕〔25〕である．

（3）両国の失業率の相関係数を求める方法として最も適切なものを選択し，その番号をマークせよ．

①（2）で求めた共分散をアメリカの標本分散で割る

②（2）で求めた共分散を日本の標本分散で割る

③（2）で求めた共分散を両国の標本分散を足したもので割る

④（2）で求めた共分散を両国の標本分散をかけたもので割る

⑤（2）で求めた共分散を両国の標本標準偏差を足したもので割る

⑥（2）で求めた共分散を両国の標本標準偏差をかけたもので割る

解答番号〔26〕

（４）上表のデータから相関係数を計算すると 0.510 という結果となった．この結果からわかる内容として最も適切なものを選び，その番号をマークせよ．

①アメリカと日本の失業率には強い正の相関があり，アメリカの失業率が日本の失業率に影響を与えている．

②アメリカと日本の失業率には正の相関があるが，日本の失業率がアメリカの失業率に与える影響は小さい．

③アメリカと日本の失業率には負の相関があり，アメリカの失業率が上昇することによって日本の失業率は低下する．

④アメリカと日本の失業率には負の相関があり，日本の失業率悪化はアメリカの景気悪化が原因である．

⑤アメリカと日本の失業率には正の相関があるが，両者の因果関係はわからない．

解答番号〔27〕

Ⅲ．あるサッカーの予選リーグでは，勝者のチームは 1 ゲーム（90 分の試合）で平均 3 点，敗者のチームは平均 1 点得点することが知られている．どちらも得点はポアッソン分布（ポアソン分布）に従っているとして，以下の問いに答えよ．

（１）試合中のある 30 分間に，勝者のチームが少なくとも 1 点得点する確率は，

〔28〕〔29〕.〔30〕〔31〕〔32〕**%**である．

（２）試合中のある 45 分間に，敗者のチームが 1 点も入れられない確率は，

〔33〕〔34〕.〔35〕〔36〕〔37〕**%**である．

Ⅳ．正規母集団（確率変数は正規分布に従っている母集団）から無作為抽出した 11 個の観測値が，

9，7，8，4，8，7，5，7，8，6，8

であったとする．これについて以下の問に答えよ．

なお，この観測値から計算した偏差平方和は 22 である．

※偏差平方和とは，平均からの差の二乗和：$\sum_{i=1}^{n}(X_i - \overline{X})^2$ である．

（１）標本平均は〔38〕，中央値は〔39〕，最頻値は〔40〕である．

（２）標本標準偏差は〔41〕.〔42〕〔43〕〔44〕である．

（３）母平均（母集団平均：μ）について，信頼係数 95%で区間推定をした場合，その信頼区間は，

〔45〕.〔46〕〔47〕〔48〕～〔49〕.〔50〕〔51〕〔52〕

である．

（4）この問題で，もし母標準偏差（母集団標準偏差：σ）が 1.4 であるとわかっているとした場合，母平均（母集団平均：μ）についての信頼係数 95%で区間推定をすると，その信頼区間は，

〔53〕.〔54〕〔55〕〔56〕～〔57〕.〔58〕〔59〕〔60〕

である.

（5）この問題で，母平均（母集団平均：μ）が未知である場合，母分散（母集団分散：σ^2）について信頼係数 90%で区間推定をすると，その信頼区間は，

〔61〕.〔62〕〔63〕〔64〕～〔65〕.〔66〕〔67〕〔68〕

である.

V．制限速度が 60km/h のある高速道路のある区間では，車のスピードは平均 75km/h，標準偏差は 30km/h の正規分布に従っていることがわかっている．その区間で実際に 36 台の車を調べたところ，その平均は 76km/h，標準偏差は 32km/h であった．これを受けて，制限速度を 76km/h まで緩和することになった．そこで再び 64 台の車のスピードを調べたところ，その平均は 85km/h，標準偏差は 40km/h であった．なお，制限速度が 76km/h の時の車のスピードの母集団は，もはや未知となっている．以上をもとにして，以下の問に答えよ.

（1）制限速度が 60km/h の場合，〔69〕〔70〕.〔71〕〔72〕〔73〕**%**の車がスピード違反をしていることになる.

（2）制限速度が 76km/h まで緩和することによって，車の平均速度は制限速度が 60km/h の時より速くなったと言えるであろうか．帰無仮説を「車の平均速度は制限速度が 60km/h の時より速くなっていない」として有意水準 1%で検定した場合，

P 値は〔74〕.〔75〕〔76〕〔77〕

となる.

（3）（2）の検定結果として最も適切なものを次から選び，その番号をマークせよ.

① *P* 値が有意水準より小さくなるので帰無仮説が採択され，平均速度は以前より速くなったとは言えない.

② *P* 値が有意水準より小さくなるので帰無仮説が棄却され，平均速度は以前より速くなった.

③ *P* 値が有意水準より大きくなるので帰無仮説が採択され，平均速度は以前より速くなったとは言えない.

④ *P* 値が有意水準より大きくなるので帰無仮説が棄却され，平均速度は以前より速くなった.

⑤ *P* 値が有意水準より大きくなるので対立仮説が採択され，平均速度は以前より速くなった.

解答番号〔78〕

Ⅵ．東大志望者と阪大志望者の英語の成績は，正規分布になることが知られている．ある予備校で大学受験模擬試験を実施した結果，英語科目の成績は，東大志望者のクラス 1（受験者は 100 人）では，平均点（$\overline{X}_1$）が 73 点，標準偏差（S_1）が 15 点であった．また，阪大志望者のクラス 2（受験者は 100 人）では，平均点（$\overline{X}_2$）が 70 点，標準偏差（S_2）が 13 点であった．ここで，標本平均の差を $\overline{X} = \overline{X}_1 - \overline{X}_2$ で与えるとして，以下の問に答えよ．

ただし，両クラスの母集団分散（σ_1^2, σ_2^2）はそれぞれ未知で異なっているものとする．

（1）$\overline{X}$ の分散（の推定量）は〔79〕．〔80〕〔81〕〔82〕である．

（2）帰無仮説を「東大志望者と阪大志望者の英語成績の平均に差はない」として平均値の差の検定を行う場合の（標準化された）検定統計量は，

〔83〕．〔84〕〔85〕〔86〕

である．

（3）対立仮説を「東大志望者と阪大志望者の英語成績の平均は異なる」として有意水準 5% で検定を行う場合，検定結果として最も適切なものを次から選び，その番号をマークせよ．

①東大志望者と阪大志望者の英語成績の平均に統計的に有意な差はない．

②東大志望者と阪大志望者の英語成績の平均に有意な差がある．

③東大志望者と阪大志望者の英語成績の平均に統計的に有意な差があるか否かは判断できない．

解答番号〔87〕

Ⅶ．回帰分析に関して，以下の問に答えよ．

（1）誤差項が満たすべき条件（標準的線形回帰モデルの仮定）として通常想定されるもののうち**正しくないもの**を選び，その番号をマークせよ．

①誤差項の平均は 0 である．

②誤差項は自己相関していない（誤差項は独立である）．

③誤差項の分散は均一である．

④誤差項と説明変数は独立であるので，説明変数は確率変数となっている．

⑤誤差項は正規分布である．

解答番号〔88〕

（裏面に続く）

（２）最小自乗法（最小二乗法）に関する以下の説明について，正しい場合には①に，誤っている場合には②にマークせよ．問題文前の〔89〕～〔99〕は解答番号である．

ただし，説明変数を X_i，被説明変数を Y_i，誤差項を u_i とし，$Y_i = a + bX_i + u_i$ を推定するものとする．また，パラメータの推定値はそれぞれ $\hat{a}, \hat{b}$ と標記する．

〔89〕 最小自乗法における残差（e_i）は，回帰直線と各データとの垂直方向の距離である．

〔90〕最小自乗法とは，残差の和を最小にするように直線を引くルールである．

〔91〕最小自乗法では，残差の和は 0 となる．

〔92〕最小自乗法におけるモデルの説明力は，全変動に対する説明される変動の割合によって表される．

〔93〕決定係数が 1 もしくは－1 に近づくほど，最小自乗法におけるモデルの説明力は高くなる．

〔94〕最小自乗法では，通常，説明変数 X_i が被説明変数 Y_i に影響を及ぼすという因果関係を前提としてモデルを設定する．

〔95〕最小自乗法においてパラメータ b の有意性を検定する場合，帰無仮説は $H_0 : b = 0$ とする．

〔96〕最小自乗法においてパラメータ b の有意性に関する帰無仮説が棄却された場合，説明変数と被説明変数には有意な関係がないと判断する．

〔97〕最小自乗法においてパラメータ b を標準化した検定統計量＝$\dfrac{\hat{b}}{S_{\hat{b}}}$ は，自由度 $n-2$ の t 分布に従っている．

〔98〕誤差項が通常満たすべき条件（標準的線形回帰モデルの仮定）を有している場合，最小自乗法におけるパラメータ b の推定量 $\hat{b}$ は一致推定量である．

〔99〕誤差項が通常満たすべき条件（標準的線形回帰モデルの仮定）を有している場合，最小自乗法におけるパラメータ b の推定量 $\hat{b}$ は BLUE である．

《コメント》

・〔94〕は，新関先生のコメントにあったように，通常，因果関係が明らかな場合を想定しているという表現に変更してみました．

・Ⅱ（４）【赤で示しています】は，本来であれば相関係数の有意性を検定しなければ正の相関があると明言できないように思います（もちろん検定方法は教えていません）．自分でも検定していないので有意性があるかどうかは確認していないのですが，自由度が小さいので微妙かと…．母集団のことを考えていないと認識しますか？それとも，いっそ，外してしまいましょうか?

・角井問題の解答の間違い，ご指摘ありがとうございました．