2013 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．ある大学で 1 年生から 4 年生共通の英語の到達度試験を行った結果，36 名が受験し，その結果は下の表の通りとなった．

受験番号	点数（1 年）	受験番号	点数（2 年）	受験番号	点数（3 年）	受験番号	点数（4 年）
101	68	201	90	301	70	401	56
102	72	202	30	302	58	402	42
103	45	203	68	303	42	403	52
104	30	204	16	304	68	404	79
105	40	205	88	305	52	405	68
106	52	206	68	306	60	406	23
107	68			307	98	407	15
108	69			308	32	408	47
				309	70	409	81
				310	78	410	60
				311	78		
				312	68		

（右ページに続く）

また，全体（36 名）と各学年の主な記述統計量は下の表の通りである．

	人数	標本平均	標本分散
全体	36	58.361	426.365
1 年	8	55.500	254.286
2 年	6	60.000	〔1〕〔2〕〔3〕.〔4〕〔5〕〔6〕
3 年	12	64.500	304.818
4 年	10	52.300	471.122

（1）上の表の「2 年次の点数の標本分散」を計算し，各桁をマークせよ．

（2）4 年生の中央値は〔7〕〔8〕.〔9〕〔10〕〔11〕であり，全体（36 名）における最頻値は〔12〕〔13〕である．

（3）1 年生の点数の標本分散＝254.286 より標本標準偏差＝〔14〕〔15〕.〔16〕〔17〕〔18〕が導かれる．そこで，受験番号 101 番の学生の点数を「1 年生の中だけで」標準化すると，標準化された点数は〔19〕.〔20〕〔21〕〔22〕となり，その偏差値は〔23〕〔24〕.〔25〕〔26〕〔27〕となる．

Ⅱ．次の文章について，最も適切な語句を指定された指定された【語群】から選択し，その番号をマークせよ．

２つのデータ系列（$x_1, x_2 \cdots, x_n$）と（$y_1, y_2 \cdots, y_n$）があるとき，その共分散は〔28〕の式で表され，相関係数は

$$\frac{x\text{と}y\text{の共分散}}{x\text{の}(\quad〔29〕\quad)\times y\text{の}(\quad〔29〕\quad)}$$

で表される．

いま，２つのデータ系列の相関係数が－0.869 であるとき，この２つの変数の間には〔30〕といい，〔31〕ということがわかる．

〔28〕の【語群】

① $\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2$　　② $\frac{1}{n-1}\sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2$

③ $\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\left(y_i-\overline{y}\right)^2$　　④ $\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)$

⑤ $\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\left(y_i-\overline{y}\right)^2}$

（裏面に続く）

〔29〕〔30〕の【語群】

①標本平均　②標本分散　③標本標準偏差
④変動係数　⑤四分位範囲　⑥正の相関がある
⑦負の相関がある　⑧無相関である

〔31〕の【語群】

①「xがより大きいとき，yもより大きくなっている」
②「xがより小さいとき，yもより小さくなっている」
③「xがより大きいとき，yはより小さくなっている」
④「xがより大きくなったので，yはより大きくなった」
⑤「xがより大きくなったので，yはより小さくなった」

Ⅲ．次の（1）（2）と（3）の〔45〕〔46〕については適切な数値を，（3）の〔44〕については最も適切な語句を【語群】から選択してその番号をマークせよ．

（1）ゆがみのないコインを 30 回連続で投げたとき，表が出る回数の平均は〔32〕〔33〕，分散は〔34〕.〔35〕〔36〕〔37〕である．

（2）ある営業マンは，勤務時間の 8 時間に平均して 14 本の電話を独立に受けるという．この営業マンがある日の勤務時間中の午後 1 時から 5 時までの 4 時間で受ける電話の本数の平均は〔38〕本である．また，この 4 時間で 3 本以上電話を受ける確率は〔39〕〔40〕.〔41〕〔42〕〔43〕**%**である．

（3）自由度が無限大のt分布は〔44〕分布と重なり合う．しがって，〔44〕分布であっても自由度が無限大のt分布であっても，右すそに〔45〕.〔46〕**%**を取る位置は 1.96 となっている．

【語群】

①二項　②ポアソン　③正規　④標準正規　⑤χ^2

（右ページに続く）

Ⅳ．ベイズの定理は，条件と結果の入れ替え，因果関係を逆にした確率を求める方法である．すなわち，$P(A|B)$ を使って $P(B|A)$ を求める方法である．ベイズの定理を説明した（１）（２）の文章について，最も適切な式を下の選択肢から選び，その番号をマークせよ．また，（３）については適切な数値をマークせよ．

（１）一般に事象 A と事象 B が生じる場合について，事象 A が生じる確率（周辺確率）は〔47〕で表される．

① $P(A)=P(A)+P(B)$
② $P(A)=P(A\cap B)+P(A^C\cap B^C)$
③ $P(A)=P(A\cap B)+P(A\cap B^C)$
④ $P(A)=P(A\cup B)+P(A^C\cup B^C)$
⑤ $P(A)=P(A\cup B)+P(A\cup B^C)$

（２）また，乗法定理：$P(A\cap B)=P(A|B)P(B)$ を応用すれば，〔48〕のようなベイズの定理が得られる．

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

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

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

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

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

（３）ある都市銀行で住宅ローンの返済不履行の状況を調査したところ，返済不能になった人のうち 3 割は審査を優良でパスしていた．他方，返済上の問題がなかった人のうち 7 割は審査を優良でパスしていることもわかった．ちなみに，返済不能になる人は全体の 8%であるとわかっている．「返済不能」を事象 B，「優良で審査をパス」を事象 A とすると，審査を優良でパスしながら返済不能に陥る人の確率は，上記のベイズの定理を用いると，

〔49〕．〔50〕〔51〕〔52〕

となる．（この問題は百分率（%）で答えるのではない）

（2 枚目に続く）

Ｖ．アイドルグループ「ＹＮＴ」において，チームのセンター・ポジションを決定する人気投票（選抜総選挙）が行われた．１位になったＳ子さんを応援する九州のファンを無作為に 9 人選び，投票券の入った CD を何枚買ったか聞いたところ，

2，5，8，15，20，20，50，60，90

であった. 九州のファン全体の購入枚数が正規分布にしたがっているとして, 以下の問いに答えよ.（ただし,（２）（３）においては（１）に示された標本標準偏差＝30.037 を利用し，下の（４）の情報は用いないこと）

（１）上記標本の標本平均は〔53〕〔54〕．〔55〕〔56〕〔57〕，標本標準偏差は 30.037 である.

（２）S 子さんに投票した九州のファンの CD 購入枚数について，信頼係数 95%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

〔58〕.〔59〕〔60〕〔61〕以上,〔62〕〔63〕.〔64〕〔65〕〔66〕以下

となる.

（３）S 子さんに投票した九州のファンの CD 購入枚数について，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

〔67〕〔68〕〔69〕.〔70〕〔71〕〔72〕以上,〔73〕〔74〕〔75〕〔76〕.〔77〕〔78〕〔79〕以下

となる.

（４）S 子さんに投票した全国のファンの CD 購入枚数の平均（期待値）と九州のファン全体の CD 購入枚数の標準偏差は，ともに 15 枚であることが知られている．この場合,「九州のファンの方が全国のファンより多く CD を購入している」と言えるか．有意水準 1%で検定する．(標準化された)「検定統計量」は，絶対値で〔80〕.〔81〕〔82〕〔83〕となり，検定結果は〔84〕となる.

〔84〕については，当てはまる最も適切なものを下の選択肢から選べ.

〔84〕の【選択肢】

①臨界値は 2.33 で P 値は有意水準の 1%より大きく,「九州のファンの方が全国のファンより多く CD を購入している」と言える.

②臨界値は 2.33 で P 値は有意水準の 1%より小さく,「九州のファンの方が全国のファンより多く CD を購入している」と言える.

③臨界値は 2.58 で P 値は有意水準の 1%より大きく,「九州のファンの方が全国のファンより多く CD を購入している」と言える.

④臨界値は 2.58 で P 値は有意水準の 1%より小さく,「九州のファンの方が全国のファンより多く CD を購入している」と言える.

⑤臨界値は 2.897 で P 値は有意水準の 1%より大きく,「九州のファンの方が全国のファンより多く CD を購入している」と言える.

⑥臨界値は 2.897 で P 値は有意水準の 1%より小さく,「九州のファンの方が全国のファンより多く CD を購入している」と言える.

（右ページに続く）

Ⅵ. 平成 24 年に行われたある調査結果によると，子供の頃に親からモラル教育として「ルールは必ず守りなさい」と教えられた者と，そのようなモラル教育を受けていない者とでは，平均年収は以下のようになっている．

グループ	標本世帯数	標本平均（単位：百万円）	標本標準偏差（単位：百万円）
教育を受けていない世帯	9500	3.3	3.2
教育を受けている世帯	4000	3.5	3.4

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，２つのグループの年収金額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〔85〕.〔86〕〔87〕〔88〕となる．

（２）対立仮説を「2 つのグループに差がある」として，有意水準 5%で検定したとする．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〔89〕

①検定統計量が棄却域に入るので，両グループの平均年収金額に差がない．
②検定統計量が棄却域に入らないので，両グループの平均年収金額に差がない．
③検定統計量が棄却域に入るので，両グループの平均年収金額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，両グループの平均年収金額が等しいとは言えない．
⑤両グループの平均年収金額に差があるかどうかは判らない．

Ⅶ. 回帰分析に関する以下の文章のうち，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．

〔90〕残差（誤差）は回帰直線（回帰式）と各データの最短の距離で表される．

〔91〕最小二乗法（最小自乗法）とは，残差（誤差）の合計を最小にする回帰直線（回帰式）の引き方である．

〔92〕通常の誤差項（u_i）が満たすべき条件（標準的線形回帰モデルの仮定）では，誤差項（u_i）は独立で分散が均一な正規分布であるとされる．

〔93〕説明変数を X_i，被説明変数を Y_i，誤差項を u_i とし，$Y_i = a + bX_i + u_i$ の回帰直線（回帰式）を推定する場合，パラメータ b を推定するための帰無仮説は $H_0 : X_i = 0$ である．

〔94〕説明変数を X_i，被説明変数を Y_i，誤差項を u_i とし，$Y_i = a + bX_i + u_i$ を最小二乗法（最小自乗法）によって推定した場合，パラメータ b の推定量 $\hat{b}$ は BLUE となる．

〔95〕最小二乗法（最小自乗法）における決定係数とは，推定されたパラメータ $\hat{b}$ の有意性を検証するために用いられる統計量である．

（以下余白）