2012 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．ある大学の学生 5 人（H さん〜L さん）の GPA を 2 年次と 4 年次で調査したところ，下の表のようになった．この表をもとに，以下の問いに答えよ．

	H さん	I さん	J さん	K さん	L さん
2 年次	2.0	3.6	2.4	2.2	0.8
4 年次	1.8	3.2	2.2	1.8	1.0

（1）2 年次の成績の中央値は〔1〕.〔2〕, 4 年次の成績の最頻値は〔3〕.〔4〕である．

（2）上の表から計算される 2 年次の標本標準偏差は 1.00，4 年次の標本標準偏差は 0.80 である．したがって，2 年次の GPA の変動係数は〔5〕.〔6〕〔7〕〔8〕である．

（3）2 年次と 4 年次の成績の共分散は〔9〕.〔10〕〔11〕〔12〕であり，
相関係数は〔13〕.〔14〕〔15〕〔16〕であるから,〔17〕.

〔17〕の【選択肢】

①2 年次の成績より 4 年次の成績の方が高くなる．

②2 年次の成績と 4 年次の成績との散布図は右上がりになるか否かわからない．

③2 年次の成績と 4 年次の成績とには正の相関がある．

④2 年次の成績と 4 年次の成績とには負の相関がある．

⑤2 年次の成績と 4 年次の成績には相関がない．

Ⅱ．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〔18〕

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

④ $P(A|B)=\dfrac{P(A\cup B)}{P(B)}$　　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〔19〕

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

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

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

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

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

（３）癌（がん）の検診にあたって，検診を受ける人は「癌を患っていると感じている」人（A）と，「癌を患っているとは感じていないが一応診てもらおう」という人（A^C）に二分できるとしよう．A の割合は 20%，A^C の割合は〔20〕〔21〕．〔22〕〔23〕〔24〕%で，さらに従来の検診の結果，A のうち 6%が「実際に癌にかかっている」ことがわかっており，A^C のうちでも 1%が「実際に癌にかかっている」ことがわかっているとしよう．

以上の確率をもとにして，「実際に癌にかかっている」人（B）が、自分でも「癌を患っていると感じている」（A）確率として $P(A|B)$ を求めると，

〔25〕〔26〕．〔27〕〔28〕〔29〕**%**

となる．

（裏面に続く）

Ⅲ．次の文章について，最も適切な語句を指定された【選択肢】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，〔33〕と〔34〕〔35〕には求めた値をマークせよ．

確率分布では，パラメータを用いて平均や分散を表すことができる．たとえば，パラメータ λ のポアソン分布の平均は〔30〕，分散は〔31〕である．また，5 本中 2 本の当たりくじが入っているくじ引きを 15 回連続して引いたときに（毎回引いたくじは元に戻すこととする）当たりが出る回数 x は〔32〕分布に従うことになるが，その平均は〔33〕，分散は〔34〕.〔35〕となる．

次に，たとえば平均＝μ，分散＝σ^2 の母集団から無作為抽出（無限母集団からの復元抽出）された n 個の標本による標本平均 $\overline{X}$ について考える．標本数 n が大きいときには母集団の分布がいかなる場合でも標本平均 $\overline{X}$ は〔36〕分布に近似され，それを標準化した確率変数 Z は〔37〕分布に近似できる．これを〔38〕という．そして，標本数 n が無限大のとき，〔39〕により標本平均 $\overline{X}$ は母集団平均 μ に収束することになる．このように，n の極限において推定量（この場合は標本平均 $\overline{X}$）が真の母集団パラメータ（この場合は母集団平均）に収束する性質を，推定における〔40〕性という．

なお，この場合の標本平均 $\overline{X}$ の〔41〕は常に真の平均（母集団平均）μ になることが証明されており，標本平均 $\overline{X}$ は母集団平均 μ の〔42〕推定量でもある．

〔30〕〔31〕の【選択肢】

①n　②$n-1$　③np　④$np(1-p)$　⑤λ　⑥$\dfrac{1}{\lambda^2}$

〔32〕〔36〕〔37〕の【選択肢】

①二項　②一様　③正規　④標準正規
⑤自由度 $n-1$ の t　⑥自由度 n の χ^2　⑦自由度 $n-1$ の χ^2

〔38〕～〔42〕の【選択肢】

①大数の法則　②中心極限定理　③ベイズの定理
④加法定理　⑤乗法定理　⑥不偏
⑦一致　⑧臨界値　⑨期待値

（右ページに続く）

Ⅳ. ある高速道路の料金所を通過する車の台数は平均的に 1 時間 90 台であり，ポアソン分布に従っている．この料金所でたまたま選んだ午後 4 時 30 分から 4 時 33 分の 3 分間に 3 台以上の車が通過する確率は，

〔43〕〔44〕．〔45〕〔46〕〔47〕**%**

である．

また，車が料金所を通過する時のスピードが平均 15km/h，標準偏差 5km/h の正規分布に従っている場合，通過するスピードが 10km/h 以上 25km/h 以下となる確率は，

〔48〕〔49〕．〔50〕〔51〕〔52〕**%**

である．<u>（※この問題では計算途中の四捨五入はしないこと）</u>

Ｖ. T 社で生産されている LED 電球から無作為抽出した 9 個の LED 電球の寿命時間（単位：万時間）が，

2，2，3，4，4，4，5，6，6

であったとする．この寿命時間が正規分布に従っているとして，以下の問いに答えよ．

（１）標本平均は〔53〕．〔54〕〔55〕〔56〕，標本分散は 2.25 である．

（２）母平均（母集団平均：μ）について，信頼係数 99％で区間推定をした場合，その信頼区間は，

〔57〕．〔58〕〔59〕〔60〕～〔61〕．〔62〕〔63〕〔64〕

である．

（３）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 95％で区間推定をすると，その信頼区間は，

〔65〕．〔66〕〔67〕〔68〕～〔69〕．〔70〕〔71〕〔72〕

である．

（2 枚目に続く）

（4）ライバルのP社で生産されているLED電球の平均寿命時間は，3時間であることが知られている．T社の寿命時間の方がP社の平均寿命時間より長いと言えるかについて有意水準5%で検定する場合，（標準化された）「検定統計量」は絶対値で，

〔73〕.〔74〕〔75〕〔76〕

となる．また，この検定結果として最も適切なものを下の選択肢から選ぶと，〔77〕となる．

〔77〕の【選択肢】

①臨界値は2.306で，検定統計量が棄却域に入るので帰無仮説は棄却される．
②臨界値は2.306で，検定統計量が棄却域に入らないので帰無仮説は採択される．
③臨界値は1.860で，検定統計量が棄却域に入るので帰無仮説は棄却される．
④臨界値は1.860で，検定統計量が棄却域に入らないので帰無仮説は採択される．
⑤臨界値は1.860で，検定統計量が棄却域に入るので帰無仮説は採択される．

（5）もしT社のLED電球の寿命時間について，その母標準偏差（母集団標準偏差：σ）が1.2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，

〔78〕.〔79〕〔80〕〔81〕～〔82〕.〔83〕〔84〕〔85〕

である．

VI. 2011年に40歳以上の既婚者の金融資産額について，親と「同居している世帯」と「同居していない世帯」とに分けて調査が行われ，その結果は以下の通りであった．

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）
同居している世帯	400	14.7	17
同居していない世帯	2500	13.0	18

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つのグループの金融資産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」（の絶対値）として最も適切なもの（最も近いもの）を選択せよ．（金融資産額の単位は100万円）　解答番号〔86〕

①1.371　②1.515　③1.840　④2.107
⑤3.696　⑥5.782

（右ページに続く）

（2）対立仮説を「平均金融資産額は，同居している世帯の方が同居していない世帯より大きい」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〔87〕

① P 値が有意水準よりも大きいので帰無仮説を棄却し，同居している世帯の方が同居していない世帯より平均金融資産額が大きいと言える．
② P 値が有意水準よりも大きいので帰無仮説を採択し，同居している世帯の方が同居していない世帯より平均金融資産額が大きいと言える．
③ P 値が有意水準よりも大きいので帰無仮説を採択し，同居している世帯の方が同居していない世帯より平均金融資産額が大きいとは言えない．
④ P 値が有意水準よりも小さいので帰無仮説を棄却し，同居している世帯の方が同居していない世帯より平均金融資産額が大きいと言える．
⑤ P 値が有意水準よりも小さいので帰無仮説を採択し，同居している世帯の方が同居していない世帯より平均金融資産額が大きいとは言えない．

（3）対立仮説が「2 つのグループに差がある」となった場合，有意水準 5%で検定した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，両グループの平均金融資産額に差がないと言える．
②検定統計量が棄却域に入らないので，両グループの平均金融資産額に差がないと言える．
③検定統計量が棄却域に入るので，両グループの平均金融資産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，両グループの平均金融資産額が等しいとは言えない．
⑤両グループの平均金融資産額に差があるかどうかはわからない．

（以下余白）