統計検定
Japan Statistical Society Certificate

2 級

2019 年 6 月 16 日

【注意事項】

1 試験開始の合図があるまで，この問題冊子の中を見てはいけません。

2 この問題冊子は，28 ページあります。

3 試験時間は 90 分です。

4 試験中に問題冊子の印刷不鮮明，ページの落丁・乱丁およびマークシートの汚れ等に気付いた場合は，手を挙げて監督者に知らせなさい。

5 マークシートの **A** 面には次の項目があるので，それぞれの指示に従い記入あるいは確認しなさい。項目の内容に誤りがある場合は，手を挙げて監督者に知らせなさい。

① 氏名
氏名を記入しなさい。

② 検定種別
受験する検定種別を確認しなさい。

③ 受験番号
受験番号を確認しなさい。

④ **Web** 合格発表
Web 合格発表について，希望の有無をマークしなさい。

6 解答は，マークシートの B 面の解答にマークしなさい。例えば，$\boxed{\textbf{10}}$ と表示のある問に対して ③ と解答する場合は，次の（例）のように解答番号 **10** の解答の ③ にマークしなさい。

（例）

解答番号	解 答
10	① ② ● ④ ⑤

7 解答番号は，**35** まであります。

8 23 ページ以降に付表を掲載しています。必要に応じて利用しなさい。

9 問題冊子の余白等は適宜利用してよいが，どのページも切り離してはいけません。

10 試験終了後，問題冊子は持ち帰りなさい。

問 1　次の表は，2008 年および 2015 年の，2 人以上の勤労者世帯における，貯蓄額の階級別相対度数分布表である。

	階級	2008 年 相対度数 (%)	2015 年 相対度数 (%)
(A)	100 万円未満	（ア）	13.2
(B)	100 万円以上 200 万円未満	7.1	7.2
(C)	200 万円以上 300 万円未満	6.9	7.0
(D)	300 万円以上 400 万円未満	6.3	6.1
(E)	400 万円以上 500 万円未満	5.5	5.6
(F)	500 万円以上 600 万円未満	5.7	5.5
(G)	600 万円以上 700 万円未満	5.2	4.5
(H)	700 万円以上 800 万円未満	3.9	4.2
(I)	800 万円以上 900 万円未満	3.5	3.3
(J)	900 万円以上 1000 万円未満	3.4	3.2
(K)	1000 万円以上 1200 万円未満	5.8	6.0
(L)	1200 万円以上 1400 万円未満	4.7	4.6
(M)	1400 万円以上 1600 万円未満	4.3	4.2
(N)	1600 万円以上 1800 万円未満	2.8	3.0
(O)	1800 万円以上 2000 万円未満	2.8	2.5
(P)	2000 万円以上 2500 万円未満	5.3	5.3
(Q)	2500 万円以上 3000 万円未満	3.8	3.2
(R)	3000 万円以上 4000 万円未満	4.7	4.2
(S)	4000 万円以上	（イ）	7.2

資料：総務省「家計調査」

〔1〕　2008 年における貯蓄額が 2000 万円以上の世帯は，全体の 19.6%であった。（イ）に入る数値はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **1**

① 1.2　② 3.0　③ 5.8　④ 8.2　⑤ 11.1

〔2〕　2015 年における貯蓄額の中央値が含まれる階級はどれか。次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **2**

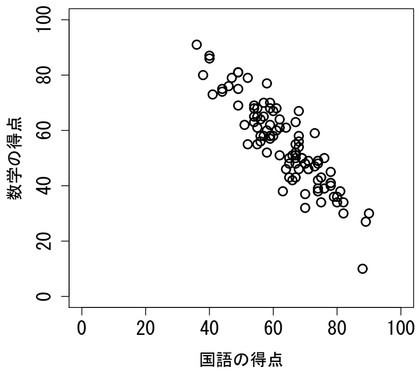
① (H)　② (I)　③ (J)　④ (K)　⑤ (L)

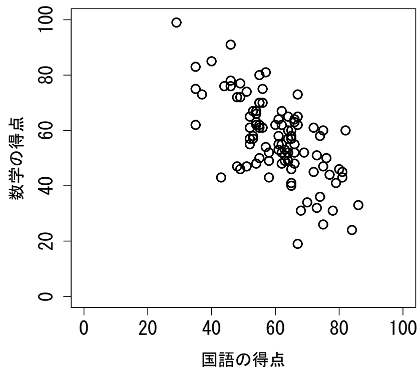
〔3〕　2015 年における貯蓄額の平均値は 1309 万円であった。2015 年における貯蓄額が平均未満の世帯の割合を x% とする。x の 1 の位を四捨五入した値はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **3**

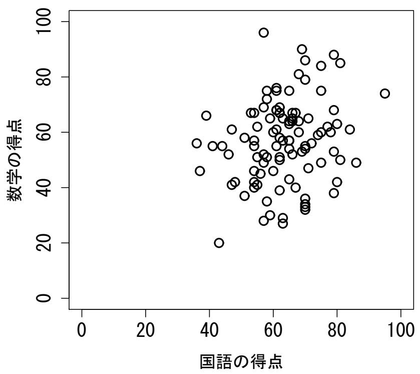
① 30　② 40　③ 50　④ 60　⑤ 70

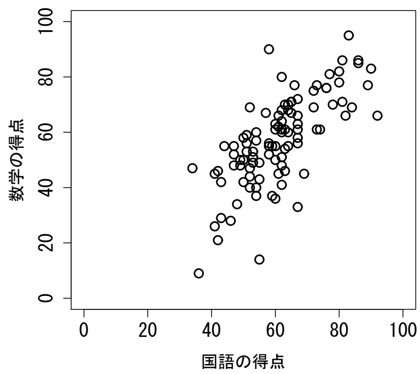
問 2　ある中学校の生徒 100 人が，国語と数学のテストを受けた。いずれも 100 点満点である。この結果，国語の得点の標準偏差は 12.5，数学の得点の標準偏差は 16.4，国語と数学の得点の相関係数は 0.72 であった。

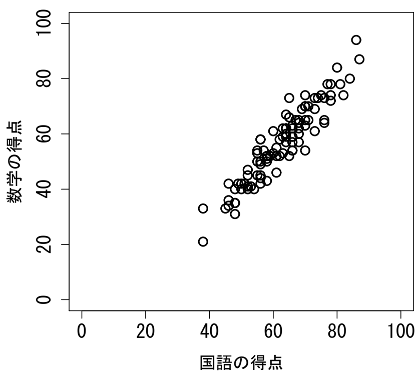
〔1〕　国語と数学の得点の散布図として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\textbf{4}}$

①


②


③


④


⑤


〔2〕　国語と数学の得点の共分散はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **5**

① 112.5　② 147.6　③ 184.7　④ 193.7　⑤ 205.0

〔3〕　次の記述は，数学の得点のみ 2 倍にしたときの，変動係数と共分散の変化に関するものである。

> 全ての生徒について数学の得点のみ 2 倍にすると，数学の得点の変動係数は (A)。また，国語と数学の得点の共分散は (B)。

(A) と (B) にあてはまるものの組合せとして，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **6**

① (A)　変わらない　(B)　変わらない
② (A)　変わらない　(B)　2 倍になる
③ (A)　2 倍になる　(B)　変わらない
④ (A)　2 倍になる　(B)　2 倍になる
⑤ (A)　2 倍になる　(B)　4 倍になる

問 3　気温を測る単位として，日本では摂氏が用いられている。一方で，アメリカにおいては，華氏を用いるのが一般的であり，摂氏 (C) から華氏 (F) への変換公式は $F = 1.8C + 32$ となる。次の表は，2018 年 12 月 9 日のアメリカの 17 の主要都市における最低気温のデータを摂氏と華氏，双方の単位で記載したものである。

No.	主要都市	摂氏	華氏	No.	主要都市	摂氏	華氏
1	アトランタ	1	33.8	10	ニューヨーク	−1	30.2
2	アンカレジ	−6	21.2	11	ヒューストン	4	39.2
3	サンフランシスコ	6	42.8	12	ボストン	−5	23.0
4	シアトル	4	39.2	13	ポートランド	6	42.8
5	シカゴ	−6	21.2	14	マイアミ	22	71.6
6	デトロイト	−4	24.8	15	ラスベガス	7	44.6
7	デンバー	−1	30.2	16	ロサンゼルス	10	50.0
8	ニューオーリンズ	4	39.2	17	ワシントン D.C.	0	32.0
9	メンフィス	−1	30.2				

資料：日本気象協会

〔1〕　上記の摂氏で表されたデータを標準化得点に変換したものを $z_1, \ldots, z_{17}$ とし，華氏で表されたデータを標準化得点に変換したものを $w_1, \ldots, w_{17}$ とする。ただし，下付きの添え字はこれらのデータの No. に対応している。また，標準化得点の計算に用いる標準偏差は不偏分散の正の平方根とし，摂氏で表されたデータの平均は 2.4，標準偏差は 7.0 であった。次の記述 I ～ III は，上のデータの標準化得点に関する説明である。

I.　$\dfrac{1}{17}\displaystyle\sum_{i=1}^{17} z_i = 0$ であり，かつ $\dfrac{1}{16}\displaystyle\sum_{i=1}^{17} z_i^2 = 1$ である。

II.　標準化得点 $z_1, \ldots, z_{17}$ のどの値も 2.5 より小さい値をとる。

III.　すべての $i = 1, \ldots, 17$ に対して，$z_i = w_i$ となる。

記述 I ～ III に関して，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\textbf{7}}$

① I のみ正しい
② II のみ正しい
③ I と II のみ正しい
④ I と III のみ正しい
⑤ I と II と III はすべて正しい

〔2〕　華氏で表されたデータの平均を $\overline{F}$，標準偏差（不偏分散の正の平方根）を s_F とおく。このとき，$\overline{F}$ と s_F の値の組合せとして，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\mathbf{8}}$

① $\overline{F} = 4.2,\ s_F = 12.6$

② $\overline{F} = 4.2,\ s_F = 44.6$

③ $\overline{F} = 36.3,\ s_F = 7.0$

④ $\overline{F} = 36.3,\ s_F = 12.6$

⑤ $\overline{F} = 36.3,\ s_F = 44.6$

問 4　世帯人員と持家率の関係を調べたい。次の表は，2017 年の 2 人以上の勤労者世帯について，世帯人員別に持家率と勤め先収入をまとめたものである。

世帯人員別の持家率と 1 世帯当たり 1 か月間の収入（2 人以上の勤労者世帯）

世帯人員（人）	持家率（%）	勤め先収入（万円）
2	75.1	41.3
3	77.3	49.0
4	83.7	54.0
5	82.9	55.6
6 以上	84.8	52.1

資料：総務省「家計調査」

世帯人員と持家率の相関係数は 0.91，勤め先収入の影響を除去した世帯人員と持家率の偏相関係数は 0.79 と計算された。ここで,「6 以上」という世帯人員については，平均値として与えられている 6.36 を用いた。

〔1〕 次の記述 I 〜 III は，この相関係数と偏相関係数に関するものである。

I.　相関係数が 0.91 ということから，世帯人員と持家率に，近似的に傾きが正の直線の関係があると考えられる。

II.　偏相関係数は，非線形関係（直線でない関係）を捉えるものである。偏相関係数が 0.79 ということは，世帯人員と持家率に非線形関係が存在する可能性を示唆する。

III.　一般的に，相関係数が正なら偏相関係数は負になるという法則性がある。相関係数も偏相関係数も正という今回の計算結果から，世帯人員と持家率には全く関係がないことがわかる。

記述 I 〜 III に関して，次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。 **9**

① I のみ正しい
② II のみ正しい
③ III のみ正しい
④ I と III のみ正しい
⑤ I と II と III はすべて正しい

〔2〕 次の記述 I 〜 III は，この相関係数と偏相関係数を比較したときの解釈に関するものである。

I. 相関係数が 0.91 で偏相関係数が 0.79 ということは，収入の水準が上昇すると，世帯人員と持家率の相関が 0.79 から 0.91 に増加することを示している。世帯人員と持家率の相関は高収入の世帯ほど高いと考えられる。

II. 相関係数が 0.91 で偏相関係数が 0.79 ということは，収入の水準が変動すると，世帯人員と持家率の相関が 0.79 から 0.91 の間で変動することを示している。世帯人員と持家率の相関はやや不安定だと考えられる。

III. 相関係数が 0.91 で偏相関係数が 0.79 ということは，収入の影響を取り除くと，世帯人員と持家率の相関が 0.91 から 0.79 に減少することを示している。世帯人員と持家率の相関には，収入を共通の要因とする見かけ上の相関（擬相関）による部分が含まれていると考えられる。

記述 I 〜 III に関して，次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。

10

① I のみ正しい
② II のみ正しい
③ III のみ正しい
④ I と II と III はすべて正しい
⑤ I と II と III はすべて誤り

問 5 実験計画における「フィッシャーの 3 原則」とは,「無作為化」,「繰り返し」,「局所管理」である。次の記述 I 〜 III は，この 3 原則に関するものである。

I. 「無作為化」により，制御できない要因の影響を偶然誤差に転化できる。

II. 「繰り返し」とは，同一の被験者から繰り返しデータを得ることである。同一の実験条件に複数の被験者を割り当てても「繰り返し」を行ったことにはならない。

III. 「局所管理」とは，実験全体をいくつかのブロックに分割し，実験を監督・監視する人を各ブロックに無作為に割り付けることを意味する。

記述 I 〜 III に関して，次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。

11

① I のみ正しい
② II のみ正しい
③ III のみ正しい
④ I と II のみ正しい
⑤ I と II と III はすべて誤り

問 6　標本抽出法に関する記述として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\mathbf{12}}$

① 多段抽出では，段数を増やせば増やすほど高い精度を得ることができる。

② 系統抽出は，似た傾向をもつように母集団を系統的にグループ分けし，すべてのグループから少数の個体を無作為に抽出し，標本とする方法である。

③ 回答率の低い調査であっても，無作為抽出で，有効回答数が十分にあれば，高い精度を達成できる。

④ 系統抽出した標本による調査結果の方が，単純無作為抽出した標本による調査結果よりもいつでも高い精度であるといえる。

⑤ クラスター（集落）抽出は，母集団を網羅的に分割し小集団（クラスター）を構成した上で，その中から抽出されたいくつかのクラスター内の個体すべてを調査する方法である。

問 7　2 つの事象 A, B に関して，次が成り立つとする。

$$P(A) = 0.4, \quad P(B) = 0.35, \quad P(A \cup B) = 0.61$$

これらから読み取れることとして，次の ① ～ ⑤ のうちから適切なものを一つ選べ。
$\boxed{\mathbf{13}}$

① 事象 A と B は独立であり，かつ，排反でもある。

② 事象 A と B は独立であるが，排反ではない。

③ 事象 A と B は排反であるが，独立ではない。

④ 事象 A と B は排反でも，独立でもない。

⑤ 事象 A と B は排反ではなく，また，独立であるかどうかはわからない。

問 8　袋 A には赤玉が 2 個，白玉が 3 個入っており，袋 B には赤玉が 1 個，白玉が 4 個入っている。1 から 6 の目が等しい確率で出るサイコロを 1 回投げて 2 以下の目が出たら袋 A から 2 回玉を取り出し，3 以上の目が出たら袋 B から 2 回玉を取り出すこととする。玉を取り出す際はその度に元に戻すものとする。

〔1〕　サイコロを 1 回投げるとき，袋 B から赤玉が 1 回だけ取り出される確率はいくらか。次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **14**

① $\frac{2}{75}$　② $\frac{4}{75}$　③ $\frac{8}{75}$　④ $\frac{4}{25}$　⑤ $\frac{16}{75}$

〔2〕　サイコロを 1 回投げるとき，赤玉が取り出される回数を X とする。X の期待値として，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **15**

① $\frac{4}{75}$　② $\frac{8}{75}$　③ $\frac{16}{75}$　④ $\frac{4}{15}$　⑤ $\frac{8}{15}$

問 9　2 つの確率変数 X と Y に関して，期待値 $E[X], E[Y]$ および X と Y の積の期待値 $E[XY]$ が以下のようになっている。

$$E[X] = 1, \quad E[Y] = 2, \quad E[XY] = 4$$

いま，$Z = X + Y, W = 2X - Y$ としたとき，分散 $V[Z], V[W]$ が

$$V[Z] = V[W] = 24$$

であった。

〔1〕　X と Y の共分散 $\mathrm{Cov}[X, Y]$ と，X, Y の 2 乗の期待値 $E[X^2], E[Y^2]$ の値の組合せとして，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **16**

① $\mathrm{Cov}[X,Y] = 2, \quad E[X^2] = 4, \quad E[Y^2] = 16$
② $\mathrm{Cov}[X,Y] = 2, \quad E[X^2] = 4, \quad E[Y^2] = 21$
③ $\mathrm{Cov}[X,Y] = 2, \quad E[X^2] = 5, \quad E[Y^2] = 20$
④ $\mathrm{Cov}[X,Y] = 6, \quad E[X^2] = 4, \quad E[Y^2] = 16$
⑤ $\mathrm{Cov}[X,Y] = 6, \quad E[X^2] = 5, \quad E[Y^2] = 20$

〔2〕　X と Y の相関係数はいくらか。次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **17**

① -0.75　② -0.25　③ 0　④ 0.25　⑤ 0.75

問 10　ある調査員が個人を対象とした訪問調査を行う。ある時間帯に調査対象者が在宅している確率が 0.2 であるとし，各訪問で調査対象者が在宅か否かは独立とする。

〔1〕　3 軒目の訪問で初めて調査対象者が在宅している確率はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **18**

① 0.13　② 0.24　③ 0.48　④ 0.67　⑤ 0.78

〔2〕　初めて調査対象者が在宅しているまでに訪問する軒数の確率分布として，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **19**

① 期待値 4，分散 15 の二項分布
② 期待値 5，分散 20 の幾何分布
③ 期待値 5，分散 20 の正規分布
④ 期待値 6，分散 25 の二項分布
⑤ 期待値 6，分散 25 の幾何分布

問 11　確率変数 X は期待値 2，分散 9 の正規分布に従うとする。このとき，確率 $P(-1 < X \leq 4)$ はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **20**

① 0.16　② 0.22　③ 0.34　④ 0.41　⑤ 0.59

問 12　$X_1, ..., X_9$ は母平均 μ，母分散 σ^2 の正規母集団からの大きさ 9 の無作為標本とする。また $\overline{X}$ を $X_1, ..., X_9$ の標本平均とし，S^2 を不偏分散とする。このとき，確率

$$P\left(\overline{X} \geq \mu + 0.62S\right)$$

はいくらか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **21**

① 0.0250　② 0.0314　③ 0.0479　④ 0.0500　⑤ 0.2676

問 13　既知の母集団 $\{2, 4, 6, 8\}$ を考える。この母集団から大きさ 2 の標本 X_1, X_2 を無作為復元抽出する。この標本に対する標本平均を $\overline{X} = \dfrac{X_1 + X_2}{2}$ とし，$p_k = P(\overline{X} = k)$ とおく。ただし，k は自然数とする。

〔1〕　(p_3, p_6) の組合せとして，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **22**

① $(1/16, 1/16)$　② $(1/16, 3/16)$　③ $(1/8, 1/4)$
④ $(1/8, 3/16)$　⑤ $(3/16, 1/8)$

〔2〕　$\overline{X}$ の (中央値，最頻値) の組合せとして，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **23**

① $(4.0, 5.0)$　② $(4.5, 6.0)$　③ $(5.0, 5.0)$
④ $(5.0, 6.0)$　⑤ $(6.0, 6.0)$

〔3〕　$\overline{X}$ の期待値 $E[\overline{X}]$ に関する説明として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **24**

① $E[\overline{X}]$ の値を計算するためには，$p_1, \ldots, p_8$ をすべて計算する必要がある。
② $E[\overline{X}]$ の厳密な値を知るのは不可能である。
③ $E[\overline{X}]$ は母平均の不偏推定量であるから，4 か 6 のいずれかである。
④ $E[\overline{X}]$ を知るには実際に (X_1, X_2) を抽出したデータが必要である。
⑤ $\overline{X}$ は標本抽出のたびに異なる値をとり得るが，$E[\overline{X}]$ の値は定数である。

問 14　ある池には総数 N 匹の魚がいる。この池から 300 匹の魚を捕獲し，目印を付けて池に戻す。十分時間が経過してから，再び 200 匹を捕獲して調べたところ，目印のついている魚が 20 匹いた。N が十分大きいとしたときの，目印のついている魚の比率の 95% 信頼区間として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。

25

① 0.100 ± 0.017　② 0.100 ± 0.021　③ 0.100 ± 0.034
④ 0.100 ± 0.042　⑤ 0.100 ± 0.131

問 15　次の表は，2017 年 1 月から 2018 年 12 月までの，Amazon.com の株価の月次変化率 (単位：%) の基本統計量をまとめたものである。

	標本サイズ	標本平均	不偏分散
Amazon.com	24	3.23	8.72^2

資料：Yahoo! Finance (https://finance.yahoo.com)

Amazon.com の株価の月次変化率は，互いに独立に平均 μ，分散 σ^2 の正規分布に従うと仮定する。

〔1〕　μ の 95% 信頼区間として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **26**

① 3.23 ± 2.93　② 3.23 ± 3.05　③ 3.23 ± 3.68
④ 3.23 ± 4.86　⑤ 3.23 ± 6.56

〔2〕　帰無仮説 $\mu = 0$，対立仮説 $\mu > 0$ の検定結果として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **27**

① 有意水準 1% で棄却できるが，0.1% では棄却できない。
② 有意水準 2.5% で棄却できるが，1% では棄却できない。
③ 有意水準 5% で棄却できるが，2.5% では棄却できない。
④ 有意水準 10% で棄却できるが，5% では棄却できない。
⑤ 有意水準 10% では棄却できない。

問 16　X を平均 θ，分散 1 の正規分布に従う確率変数とし，帰無仮説 H_0，対立仮説 H_1 をそれぞれ

$$H_0 : \theta = 0, \qquad H_1 : \theta = 1$$

と想定した仮説検定を考える。X の観測結果 x に対して，棄却域を

$$x \geq 0.8$$

と定めると，第 1 種過誤の確率は（ア）であり，第 2 種過誤の確率は（イ）である。

次に，棄却域を

$$x \geq x_0$$

としたときの第 1 種過誤の確率を $\alpha(x_0)$，第 2 種過誤の確率を $\beta(x_0)$ とする。座標平面上に $(\beta(x_0), 1-\alpha(x_0))$ で与えられる点を P とし，x_0 を 0 から 1 まで動かしたときの点 P の軌跡を表したグラフの概形は（ウ）のようになる。

このグラフを参考にすると，第 1 種過誤の確率と第 2 種過誤の確率の和

$$\alpha(x_0) + \beta(x_0)$$

を最小にする x_0 は（エ）であることがわかる。

〔1〕　文中の（ア），（イ）に当てはまる数値の組合せとして，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **28**

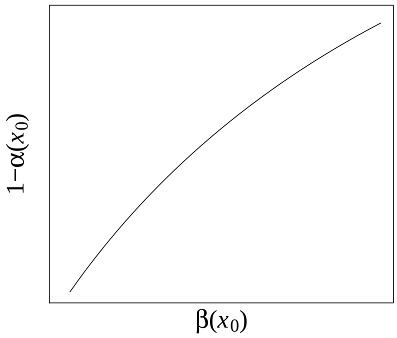
① （ア）0.212 （イ）0.212　　② （ア）0.212 （イ）0.421

③ （ア）0.421 （イ）0.212　　④ （ア）0.421 （イ）0.421

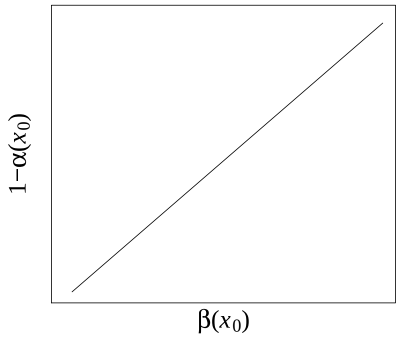
⑤ （ア）0.421 （イ）0.655

〔2〕　文中の（ウ）に当てはまるグラフの概形として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **29**

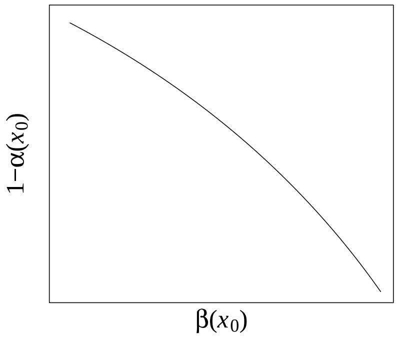
①



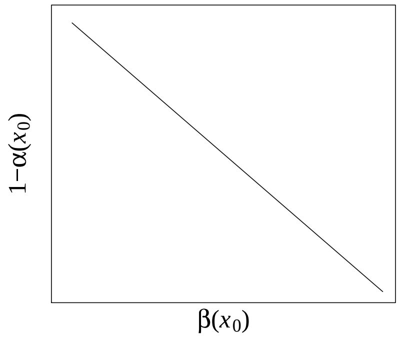
②



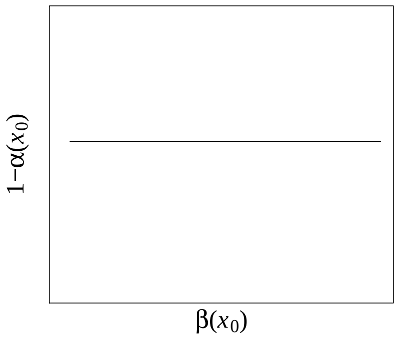
③



④



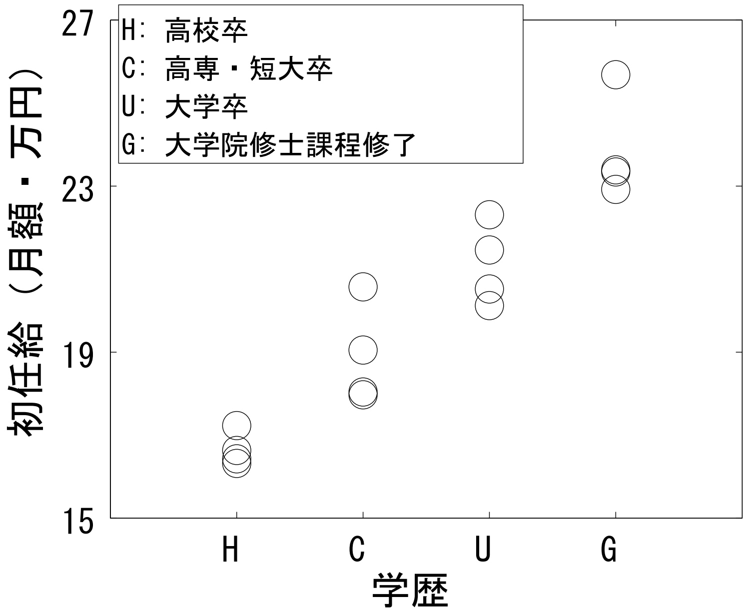
⑤



〔3〕　文中の（エ）に当てはまるものとして，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **30**

① 0のみ　　② 0.5のみ　　③ 1のみ
④ 0と1のみ　　⑤ 0から1の実数すべて

問 17　新卒者の初任給と最終学歴（以下，学歴）の関係を，4 つの業種（鉱業等，建設業，製造業，電気業等）において調べたい。次の図は，学歴別に，4 つの業種における 2018 年の新卒者の平均初任給をプロットしたものである。



資料：厚生労働省「平成 30 年賃金構造基本統計調査（新規学卒者の初任給の推移）」

〔1〕　高専・短大卒ダミー変数 C を，高専・短大卒なら 1，それ以外なら 0 をとる変数とする。同様に大学卒ダミー変数 U と大学院修士課程修了ダミー変数 G を作成する。初任給 y を被説明変数，3 つの学歴ダミー変数 C, U, G を説明変数，互いに独立に正規分布 $N(0, \sigma^2)$ に従う誤差項を u とする重回帰モデル

$$y = \beta_1 + \beta_2 C + \beta_3 U + \beta_4 G + u$$

を最小二乗法で推定したところ次の表のようになった。ここで，$\hat{\sigma}$ は，σ^2 の不偏推定値の正の平方根である。

	回帰係数	標準誤差	t-値	P-値
切片	16.653	0.510	32.652	4.31×10^{-13}
C	2.255	0.721	3.127	8.75×10^{-3}
U	4.450	0.721	6.170	4.80×10^{-5}
G	7.180	0.721	9.955	3.76×10^{-7}

観測数	16	$\hat{\sigma}$	1.020
決定係数	0.900	自由度調整済み決定係数	0.876

次の記述 I ～ III は，この推定結果に関するものである。

I. この推定結果からは，高校卒の学歴と初任給の関係がわからない。高校卒ダミー変数 H を用いて $y = \gamma_1 + \gamma_2 H + \gamma_3 C + \gamma_4 U + \gamma_5 G + v$ を最小二乗法で推定すべきである。

II. 大学院修士課程修了の初任給は，大学卒の初任給よりも 2.73 万円高い傾向がある。

III. P-値は，自由度 13 の t 分布を用いて計算されている。

記述 I ～ III に関して，次の ①～⑤ のうちから最も適切なものを一つ選べ。 **31**

① I のみ正しい
② II のみ正しい
③ III のみ正しい
④ I と II のみ正しい
⑤ II と III のみ正しい

〔2〕 教育年数 x を，高校卒は 12 年，高専・短大卒は 14 年，大学卒は 16 年，大学院修士課程修了は 18 年として作成する。初任給 y を被説明変数，教育年数 x を説明変数，u を互いに独立に正規分布 $N(0, \sigma^2)$ に従う誤差項とする単回帰モデル

$$y = \alpha + \beta x + u$$

を最小二乗法で推定したところ次の表のようになった。ここで，$\hat{\sigma}$ は，σ^2 の不偏推定値の正の平方根である。

	回帰係数	標準誤差	t-値	P-値
切片	2.323	1.620	1.434	0.174
x	1.187	0.107	11.109	2.5×10^{-8}

観測数	16	$\hat{\sigma}$	0.955
決定係数	0.898	自由度調整済み決定係数	0.891

次の記述 I ～ III は，この推定結果に関するものである。

I. 教育年数が 1 年増えると初任給は 1.187 万円上がる傾向がある。

II. 自由度調整済み決定係数とは，重回帰モデルにおいて説明変数の数に応じて決定係数を調整したものである。よって，単回帰モデルでは決定係数と自由度調整済み決定係数は等しい。今回の推定結果では 0.898 と 0.891 のように異なっているが，これは計算の丸め誤差のためである。

III. 両側検定 $H_0 : \alpha = 0$，$H_1 : \alpha \neq 0$ を行っても，片側検定 $H_0 : \alpha = 0$，$H_1 : \alpha > 0$ を行っても，P-値は 0.174 で同じである。

記述 I ～ III に関して，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **32**

① I のみ正しい
② II のみ正しい
③ III のみ正しい
④ I と II のみ正しい
⑤ I と II と III はすべて誤り

〔3〕 次の記述 I ～ III は，学歴ダミー変数を使った重回帰モデルと教育年数を使った単回帰モデルの比較に関するものである。

I. 学歴ダミー変数を使った重回帰モデルの決定係数は，教育年数を使った単回帰モデルのそれよりも 0.002 高い。したがって，重回帰モデルの方を選択すべきである。

II. 学歴ダミー変数を使った重回帰モデルでは，学歴が高専・短大卒から大学卒に変わった時の初任給の変化と，大学卒から大学院修士課程修了に変わった時の初任給の変化は異なる。一方，教育年数を使った単回帰モデルでは，両者の初任給の変化は同じである。

III. 学歴ダミー変数を使った重回帰モデルでは，中学卒という学歴の初任給を予測することはできない。一方，教育年数を使った単回帰モデルでは，推定された単回帰モデルに $x = 9$ を代入すれば形式的には予測できるが，外挿には注意する必要がある。

記述 I ～ III に関して，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **33**

① I のみ正しい
② II のみ正しい
③ III のみ正しい
④ I と II のみ正しい
⑤ II と III のみ正しい

問 18　都道府県別の 1 人当たり小売店舗事業所数を説明するため，以下の重回帰モデルを推定した。

$$(1\text{人当たり小売店舗事業所数}) = \\ \alpha + \beta_1 \times (1\text{人当たり乗用車数}) + \beta_2 \times (1\text{人当たり貨物車数}) + u$$

ここで，u は互いに独立に正規分布 $N(0, \sigma^2)$ に従う誤差項とする。

1 人当たり小売店舗事業所数，1 人当たり乗用車数，1 人当たり貨物車数にそれぞれ対応する変数を retail(単位：事業所/人)，car(単位：台/人)，truck(単位：台/人) として上の重回帰モデルを最小二乗法で推定したところ，次のような出力結果が得られた。なお出力結果の一部を削除している。また出力結果の (Intercept) は定数項 α を表している。

出力結果

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0084345  0.0004812  17.528  < 2e-16
car         -0.0077833  0.0022781  -3.417  0.00138
truck        0.0310015  0.0066308   4.675 2.79e-05
---

Residual standard error: 0.0009454 on 44 degrees of freedom
Multiple R-squared:  0.4285,Adjusted R-squared:  0.4026
F-statistic:  16.5 on 2 and 44 DF,  p-value: 4.507e-06
```

資料： 総務省「平成 28 年 経済センサス－活動調査」
総務省「平成 28 年 住民基本台帳人口・世帯数」
総務省「一般社団法人 自動車検査登録情報協会」

〔1〕　有意水準 5% で有意な係数（定数項を含む）の組合せについて，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **34**

① α　② β_1　③ α, β_2
④ β_1, β_2　⑤ α, β_1, β_2

〔2〕 出力結果から読み取れる情報として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\textbf{35}}$

① Adjusted R-squared の値が Multiple R-squared の値よりも小さいことから，正規性の仮定を疑うべきである。

② t value の値は対応する変数の説明力を表している。1 人当たり乗用車数のように t value の値がマイナスである変数は，説明力が非常に低いと判断される。

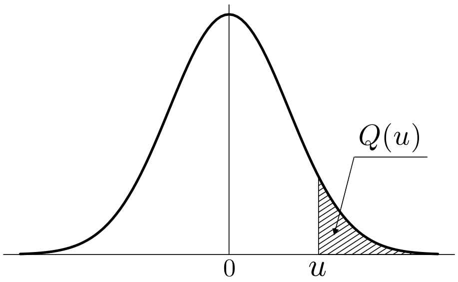
③ 他の変数が同じ値である場合，1 人当たり乗用車数が多い都道府県では，1 人当たり小売店舗事業所数は少ない傾向がある。

④ F-statistic の値は，定数項を含むすべての係数が 0 であるという帰無仮説の検定に用いられる。

⑤ 重回帰分析で変数選択を行うときには，Multiple R-squared が大きいモデルを選択すればよい。

付　表

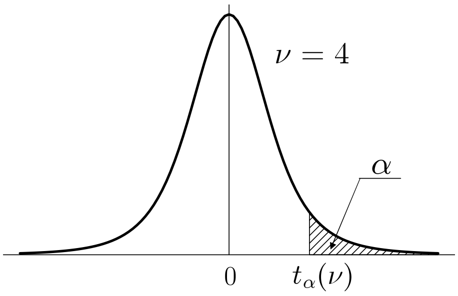
付表 1. 標準正規分布の上側確率



u	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

$u = 0.00 \sim 3.99$ に対する，正規分布の上側確率 $Q(u)$ を与える。
例：$u = 1.96$ に対しては，左の見出し 1.9 と上の見出し .06 との交差点で，$Q(u) = 0.0250$ と読む。表にない u に対しては適宜補間すること。

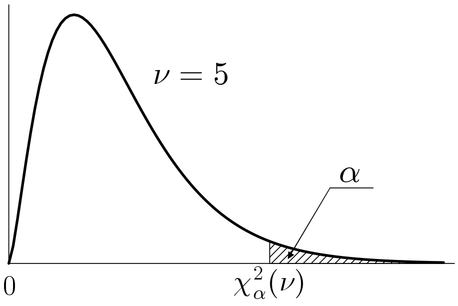
付表 2. t 分布のパーセント点



	α				
ν	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.656
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
240	1.285	1.651	1.970	2.342	2.596
∞	1.282	1.645	1.960	2.326	2.576

自由度 ν の t 分布の上側確率 α に対する t の値を $t_\alpha(\nu)$ で表す。
例：自由度 $\nu = 20$ の上側 5% 点 ($\alpha = 0.05$) は，$t_{0.05}(20) = 1.725$ である。
表にない自由度に対しては適宜補間すること。

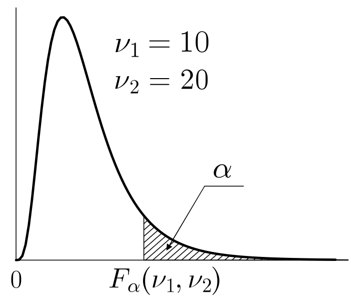
付表 3. カイ二乗分布のパーセント点



	α							
ν	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
1	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63
2	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21
3	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34
4	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28
5	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09
6	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81
7	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48
8	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09
9	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67
10	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21
11	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72
12	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22
13	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69
14	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14
15	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58
16	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00
17	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41
18	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81
19	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19
20	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89
35	18.51	20.57	22.47	24.80	46.06	49.80	53.20	57.34
40	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15
60	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38
70	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.43
80	53.54	57.15	60.39	64.28	96.58	101.88	106.63	112.33
90	61.75	65.65	69.13	73.29	107.57	113.15	118.14	124.12
100	70.06	74.22	77.93	82.36	118.50	124.34	129.56	135.81
120	86.92	91.57	95.70	100.62	140.23	146.57	152.21	158.95
140	104.03	109.14	113.66	119.03	161.83	168.61	174.65	181.84
160	121.35	126.87	131.76	137.55	183.31	190.52	196.92	204.53
180	138.82	144.74	149.97	156.15	204.70	212.30	219.04	227.06
200	156.43	162.73	168.28	174.84	226.02	233.99	241.06	249.45
240	191.99	198.98	205.14	212.39	268.47	277.14	284.80	293.89

自由度 ν のカイ二乗分布の上側確率 α に対する χ^2 の値を $\chi^2_\alpha(\nu)$ で表す。
例：自由度 $\nu = 20$ の上側 5% 点 ($\alpha = 0.05$) は，$\chi^2_{0.05}(20) = 31.41$ である。
表にない自由度に対しては適宜補間すること。

付表 4. F 分布のパーセント点



$\alpha = 0.05$

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	15	20	40	60	120	∞
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.619	4.558	4.464	4.431	4.398	4.365
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.845	2.774	2.661	2.621	2.580	2.538
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.403	2.328	2.204	2.160	2.114	2.066
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.203	2.124	1.994	1.946	1.896	1.843
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.089	2.007	1.872	1.822	1.768	1.711
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.015	1.932	1.792	1.740	1.683	1.622
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	1.924	1.839	1.693	1.637	1.577	1.509
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.836	1.748	1.594	1.534	1.467	1.389
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.750	1.659	1.495	1.429	1.352	1.254

$\alpha = 0.025$

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	15	20	40	60	120	∞
5	10.007	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619	6.428	6.329	6.175	6.123	6.069	6.015
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.522	3.419	3.255	3.198	3.140	3.080
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	2.862	2.756	2.585	2.524	2.461	2.395
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774	2.573	2.464	2.287	2.223	2.156	2.085
25	5.686	4.291	3.694	3.353	3.129	2.969	2.848	2.753	2.677	2.613	2.411	2.300	2.118	2.052	1.981	1.906
30	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575	2.511	2.307	2.195	2.009	1.940	1.866	1.787
40	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452	2.388	2.182	2.068	1.875	1.803	1.724	1.637
60	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334	2.270	2.061	1.944	1.744	1.667	1.581	1.482
120	5.152	3.805	3.227	2.894	2.674	2.515	2.395	2.299	2.222	2.157	1.945	1.825	1.614	1.530	1.433	1.310

自由度 (ν_1, ν_2) の F 分布の上側確率 α に対する F の値を $F_\alpha(\nu_1, \nu_2)$ で表す。
例：自由度 $\nu_1 = 5$, $\nu_2 = 20$ の上側 5% 点 $(\alpha = 0.05)$ は，$F_{0.05}(5, 20) = 2.711$ である。
表にない自由度に対しては適宜補間すること。

一般財団法人　統計質保証推進協会

統計検定センター

〒101-0051 東京都千代田区神田神保町３丁目６番

URL **http://www.toukei-kentei.jp**

2019.6