

(2019年11月5日版)

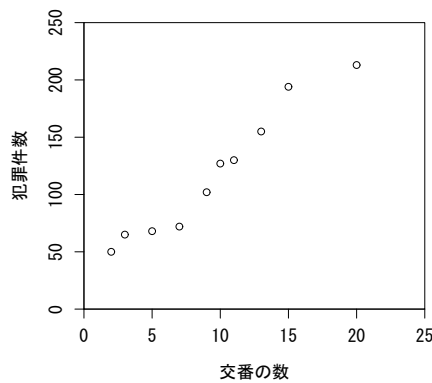
統計検定3級新出題範囲例題集（問題および略解）

1. これは2020年4月から変更される出題範囲に対応した統計検定3級の例題集です。
2. 本例題集は3級の全範囲を網羅しているのではなく、新出題範囲「相関と回帰」、「確率分布」、「統計的な推測」に関する問題のみ掲載しています。
3. 付表「標準正規分布の上側確率」を10ページに掲載しています。
4. 本例題集は必要に応じて改訂されます。

問1 相関関係および因果関係に関する記述について、次の①～⑤のうちから適切でないものを一つ選べ。

- ① 因果関係とは、2つの事象について一方の事象がもう一方の事象の直接的な原因となっている関係のことである。
- ② 2つの事象の間に擬相関があるとき、これら2つの事象の両方と相関のある事象が存在する。
- ③ ある2つの事象の因果関係を調べるには、実験研究を行うとよい。
- ④ あるコンビニエンスストアチェーンの各店舗で、苦情の数と売上高の相関を調べたところ正の相関があった。このとき、苦情の数と売上高の間には因果関係があると言える。
- ⑤ ある店舗において、気温とアイスクリームの販売数には正の相関があり、気温と炭酸飲料の販売数にも正の相関があった。また、アイスクリームと炭酸飲料の販売数にも正の相関があったが、これは擬相関と考えられる。

問2 次の散布図は、いくつかの町について交番の数と犯罪件数についてまとめたものである。



この散布図からわかることとして、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 交番の多い町では犯罪件数が多い傾向がある。
- ② 交番の数が増えると警察官が増えるため、犯罪件数は減る傾向にある。
- ③ 交番の数を増やせば、犯罪件数が増える。
- ④ 犯罪件数が増えることにより、交番の数が増える。
- ⑤ 交番の数と犯罪件数の間に因果関係がある。

問3 あるコンビニエンスストアで売られている商品 A の 1 日あたりの売上数 (個) とその日の最高気温 (°C) について調べた。最高気温を x , 売上げ数を y とし, x が y を説明する回帰直線を求めたところ,

$$y = 3.73 + 2.33x$$

という式が得られた。最高気温が 25°C のときの商品 A の販売数の予測値はいくらか。次の ① ~ ⑤ のうちから最も適切なものを一つ選べ。 3

① 9

② 25

③ 58

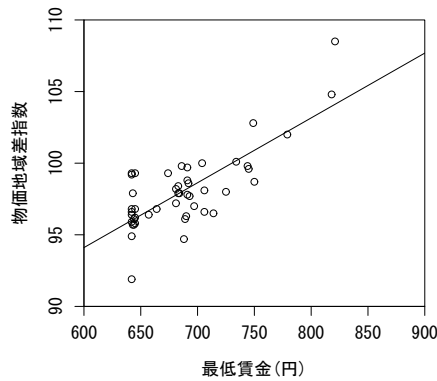
④ 62

⑤ 91

問4 次の図は、各都道府県の最低賃金（単位：円）と全国物価地域差指数（全国平均＝100）の散布図および回帰直線である。この回帰直線の式は

$$\text{全国物価地域差指数} = 66.95 + 0.045 \times \text{最低賃金}$$

である。



資料：総務省「平成19年全国物価統計調査」、
厚生労働省「地域別最低賃金改定状況（平成22年）」

この散布図および回帰直線の式から読み取れることとして、次のI～IIIの記述を考えた。

- I. 最低賃金を2000円にすれば、全国物価地域差指数は平均的に156.95となる。
- II. 最低賃金が700円であれば、全国物価地域差指数は平均的に98.45である。
- III. 全国物価地域差指数が98.45であれば、最低賃金は平均的に700円である。

この記述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。

4

- 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 IIとIIIのみ正しい

問5 次の表は、2018年の年末に販売された第771回全国自治宝くじの年末ジャンボ宝くじと年末ジャンボミニ宝くじの当選金（単位：円）と当選確率を表したものである。

年末ジャンボ宝くじ		年末ジャンボミニ宝くじ	
当選金額（円）	当選確率	当選金額（円）	当選確率
7億	2000万分の1	3000万	200万分の1
1億5000万	1000万分の1	1000万	50万分の1
1000万	2000万分の3	100万	10万分の1
100万	20万分の1	10万	3333分の1
10万	4763分の1	2万	5000分の1
1万	1000分の1	1万	1000分の1
3000	100分の1	3000	100分の1
300	10分の1	300	10分の1

[1] 年末ジャンボミニ宝くじの1枚あたりの当選金額の平均は149円であった。年末ジャンボ宝くじと年末ジャンボミニ宝くじの当選金額の平均の比較について、次の①～④のうちから最も適切なものを一つ選べ。 5

- ① 年末ジャンボ宝くじの当選金額の平均の方が高い
- ② 2つの宝くじの当選金額の平均は等しい
- ③ 年末ジャンボミニ宝くじの当選金額の平均の方が高い
- ④ 購入するたびに、どちらの平均が高いかわ変わる

[2] 年末ジャンボミニ宝くじが3億枚売れたとする。また、年末ジャンボミニ宝くじは1枚300円で販売されている。このとき、販売額と総当選金額の差はおよそいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。 6

- ① 300億円
- ② 447億円
- ③ 453億円
- ④ 514億円
- ⑤ 900億円

問6 箱の中にある製品が入っていて、その中の不良品の割合は5%である。この箱の中から100個の製品を無作為に取り出し、不良品か否かを確認する。100個のうち不良品の数を確率変数 X とし、その標本比率を $\hat{p} = X/100$ とする。また、この箱の中の製品の数は十分多いものとする。

[1] 標本比率 \hat{p} の平均 μ 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 0 ② 0.0005 ③ 0.05 ④ 0.95 ⑤ 1

[2] 標本比率 \hat{p} の標準偏差 σ 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 0.0005 ② 0.022 ③ 0.048 ④ 0.22 ⑤ 4.8

[3] $(\hat{p} - \mu)/\sigma$ が 1.96 以上の値を取る確率はいくらか。次の ① ~ ⑤ のうちから最も適切なものを一つ選べ。ここで、標本比率を標準化した $(\hat{p} - \mu)/\sigma$ が標準正規分布 $N(0, 1)$ に近似的に従うことを用いてよい。

- ① 0.01 ② 0.025 ③ 0.05 ④ 0.95 ⑤ 0.975

問7 野球において打者を評価する指標の一つに打率がある。ここでは打率を、

$$\text{打率} = \frac{\text{ヒットを打った回数}}{\text{打席に立った回数}}$$

と定義する*。一般的にはこの数値が高い選手ほどよい選手とされる。A選手が毎回の打席でヒットを打つ確率 p は一定とし、互いの打席は独立とする。

*実際の定義は分母が打数（打席に立った回数から四死球，犠打，犠飛，打撃妨害，走塁妨害の数を除いた回数）である。

[1] A選手が100打席経過した段階でヒットを32本打っていた。ヒットを打つ確率 p に対する信頼度95%の信頼区間として、次の①～⑤のうちから最も適切なものを一つ選べ。 10

- ① $-0.59 \leq p \leq 1.23$ ② $0.23 \leq p \leq 0.41$ ③ $0.27 \leq p \leq 0.37$
 ④ $0.31 \leq p \leq 0.33$ ⑤ $0.63 \leq p \leq 1.27$

[2] A選手が200打席経過した段階でヒットを64本打っていた。200打席経過した結果から、ヒットを打つ確率 p に対する信頼度95%の信頼区間を求めた場合、その幅は[1]の結果に対する信頼区間の幅の何倍となるか。次の①～⑤のうちから最も適切なものを一つ選べ。 11

- ① 1/4倍 ② 1/2倍 ③ $1/\sqrt{2}$ 倍 ④ 2倍 ⑤ 変わらない

問8 ある製品の重量(単位:g)の母平均を μ とする。 μ の95%信頼区間は $110 \leq \mu \leq 120$ であった。次のI～IIIの記述は、この信頼区間について述べたものである。

- I. 標本の95%が含まれる区間が $[110, 120]$ である。
 II. 標本平均が区間 $[110, 120]$ に入っている確率は95%である。
 III. 無作為抽出による同じ大きさを持つ標本を多数用意し、それぞれの標本を用いて95%信頼区間を求める手続きを行うと、 μ はこれらの信頼区間のうち約95%の区間に含まれる。

この記述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。

12

- 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IIIのみ正しい

問9 ある店舗で扱っている商品 A の目標販売数は日平均 500 個である。店長は目標より売れていると主張しており、それを調べるために商品 A について 30 日間の販売数を調べ、仮説検定を行うことにした。

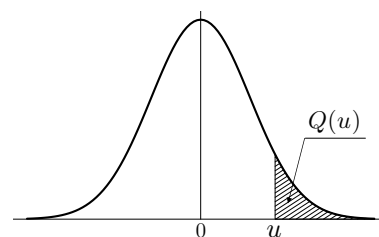
[1] 販売数が目標より多いことを調べたいときの帰無仮説と対立仮説について、次の①～④のうちから最も適切なものを一つ選べ。 13

- ① 帰無仮説：販売数の日平均は 500 個より多い
対立仮説：販売数の日平均は 500 個である
- ② 帰無仮説：販売数の日平均は 500 個である
対立仮説：販売数の日平均は 500 個より多い
- ③ 帰無仮説：販売数の日平均は 500 個より多い
対立仮説：販売数の日平均は 500 個より多いか少ないかである
- ④ 帰無仮説：販売数の日平均は 500 個である
対立仮説：販売数の日平均は 500 個より多いか少ないかである

[2] 商品 A の 30 日間の販売数を調べ、日平均値を求めたところその値は a となり、有意水準 1 % で有意であった。この結果からわかることとして、次の①～④のうちから最も適切なものを一つ選べ。 14

- ① もし販売数の日平均が 500 個より大きいなら、 a 以下となる確率は 1 % 以下である。
- ② もし販売数の日平均が 500 個より大きいなら、 a 以上となる確率は 1 % 以下である。
- ③ もし販売数の日平均が 500 個なら、 a 以下となる確率は 1 % 以下である。
- ④ もし販売数の日平均が 500 個なら、 a 以上となる確率は 1 % 以下である。

付表 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。

略解

問 1 ④

- ① 因果関係とは、2つの事象について、一方の事象がもう一方の事象の直接的な原因となっている関係のことであるので、正しい。
- ② 擬相関とは2つの事象の背後に、これらの事象に影響を与える（相関のある）事象があり、そのために現れる相関のことであるので、正しい。
- ③ 因果関係の有無を調べるには、ある種の介入を対象者にする実験研究を行うとよいので、正しい。
- ④ 相関（関係）があるだけでは因果関係があるとは言えないので、誤り。
- ⑤ 気温が、アイスクリームの販売数と炭酸飲料の販売数に影響を与え、そのために現れる擬相関の例であるので、正しい。

問 2 ①

- ① 正の相関が強く、交番の多い町では犯罪件数が多い傾向にあるので、正しい。
- ② 交番の数が増えると、犯罪件数が増える傾向にあるので、誤り。また、警察官については判断するデータは示されていない。
- ③ 交番の数と犯罪件数の間には強い相関があるが、因果関係があるとは言えず、交番の数を増やすと犯罪件数が増えるとは言えないので、誤り。
- ④ ③と同様、犯罪件数と交番の数の間には因果関係があるとは言えず、犯罪件数が増えると交番の数が増えるとは言えないので、誤り。
- ⑤ 相関があるからと言って、因果関係があるとは言えないので、誤り。

問 3 ④

回帰直線の式から、最高気温が 25℃のときの商品 A の販売数の予測値は、 $3.73 + 2.33 \times 25 = 61.98$ (個) である。

問 4 ②

- I. 誤り。回帰直線において、説明変数の取りうる値から大幅にずれた点についての予測を行うことは適切ではない。このデータでは、最低賃金はおよそ 640 円から 830 円である。2000 円を回帰直線の式に代入して値を求めることは適切ではないので、誤り。
- II. 正しい。回帰直線の式の最低賃金に 700 円を代入すると、 $66.95 + 0.045 \times 700 = 98.45$ と予測されるので、正しい。
- III. 誤り。回帰直線において、目的変数から説明変数を予測することは好ましくないなので、誤り。

以上から、正しい記述はIIのみなので、正解は②である。

問5 [1] : ③

年末ジャンボ宝くじの当選金額の平均は、

$$\begin{aligned} & 7 \text{億} \times \frac{1}{2000 \text{万}} + 1 \text{億} 5000 \text{万} \times \frac{1}{1000 \text{万}} + 1000 \text{万} \times \frac{3}{2000 \text{万}} + 100 \text{万} \times \frac{1}{20 \text{万}} \\ & + 10 \text{万} \times \frac{1}{4763} + 1 \text{万} \times \frac{1}{1000} + 3000 \times \frac{1}{100} + 300 \times \frac{1}{10} \\ & \approx 35 + 15 + 1.5 + 5 + 21 + 10 + 30 + 30 = 147.5 \text{ (円)} \end{aligned}$$

となるので、年末ジャンボミニ宝くじの当選金額の平均 (149 円) の方が高い。

[2] : ③

販売額は3 (億枚) \times 300 (円) = 900 (億円) であり、総当選金額はおよそ 300 (億枚) \times 149 (円) = 447 (億円) である。その差は $900 - 447 = 453$ (億円) となる。

問6 不良品の割合が p の製品を n 個無作為に取り出す。このとき、取り出された不良品の数を確率変数 X とすると、 X は二項分布 $B(n, p)$ に従う。その標本比率を $\hat{p} = X/n$ とすると、取り出す数 n が十分多いとき、 \hat{p} は正規分布 $N(p, p(1-p)/n)$ に近似的に従う。さらに、標本比率を標準化した $(\hat{p} - \mu)/\sigma = (\hat{p} - p)/\sqrt{p(1-p)/n}$ が標準正規分布 $N(0, 1)$ に近似的に従う。これらのことを用いて考えるとよい。

[1] : ③

標本比率の平均は箱の中の不良品の割合 0.05 と一致する。

[2] : ②

箱の中の不良品の割合 $p = 0.05$ に対し、標本比率の標準偏差は、 $\sqrt{p(1-p)/n} = \sqrt{0.05 \times 0.95/100} = 0.022$ である。

[3] : ②

標本比率を標準化した $(\hat{p} - \mu)/\sigma$ は標準正規分布 $N(0, 1)$ に近似的に従うので、 $(\hat{p} - \mu)/\sigma$ が 1.96 以上となる確率は 0.025 である。

問7 [1] : ②

ヒットを打つ確率 p に対する標本比率は $32/100 = 0.32$ となり、信頼度 95 % の信頼区間は、

$$0.32 \pm 1.96 \times \sqrt{\frac{0.32 \times (1 - 0.32)}{100}} = 0.32 \pm 0.09$$

なので、 $0.23 \leq p \leq 0.41$ となる。

[2] : ③

標本サイズが 2 倍なので、信頼区間の幅は $1/\sqrt{2}$ 倍となる。実際、標本比率は [1] と同様に $64/200 = 0.32$ となり、このときの信頼区間の幅は、

$$1.96 \times \sqrt{\frac{0.32 \times (1 - 0.32)}{200}} = 1.96 \times \frac{1}{\sqrt{2}} \times \sqrt{\frac{0.32 \times (1 - 0.32)}{100}}$$

であるので、信頼区間の幅は $1/\sqrt{2}$ 倍となる。

問 8 ③

信頼区間とは、標本の大きさと信頼度（信頼係数）を固定して、同じ手順で信頼区間を数多く作成したとき、「作成された信頼区間が母平均 μ を含む確率が信頼度である」と解釈する。

- I. 誤り。区間 $[110, 120]$ に標本の 95 % が含まれるのではないので、誤り。
- II. 誤り。区間 $[110, 120]$ に標本平均がある確率が 95 % ではないので、誤り。
- III. 正しい。上で説明した事柄を示すので、正しい。たとえば、同じ大きさの標本を用いて 95 % 信頼区間を求める手続きにより 100 個の μ に対する信頼区間を作成すると、これら 100 個の信頼区間のうち約 95 個の信頼区間が μ を含む。

以上から、正しい記述は III のみなので、正解は ③ である。

問 9 [1] : ②

目標販売数より売れているという主張を調べたいので、販売数の日平均が目標と同じである（500 個である）ことを帰無仮説とし、目標より多いこと（500 個より多い）を対立仮説とすることが適切である。

[2] : ④

有意水準 1 % で有意であるとは、「帰無仮説が真のとき、調査により得られた値以上の値が得られる確率が 1 % 以下である」ということである。この問題で帰無仮説が真とは「販売数の日平均が 500 個」であり、得られた値が a なので、④が適切である。

問 10 [1] : ⑤

一郎くんの方が強い場合、 X は 5 以上になると考えられる。また、帰無仮説が真の場合（一郎くんと次郎くんの強さが同じとする場合）、 $X \geq 8$ となる確率は 0.004、 $X \geq 7$ となる確率は $0.004 + 0.031 = 0.035$ 、 $X \geq 6$ となる確率は $0.004 + 0.031 + 0.109 = 0.144$ である。よって、有意水準 0.05 のときの棄却域は 0.05 を超えない $X \geq 7$ である。

[2] : ④

X の分布が正規分布 $N(4, 2)$ で近似できるので、 $(X - 4)/\sqrt{2}$ の分布が標準正規分布 $N(0, 1)$ となる。また、有意水準 0.05 なので、標準正規分布表より、 $(X - 4)/\sqrt{2} \geq 1.645$ が棄却域となる。つまり、棄却域は $X \geq 6.33$ である。

[3] : ①

有意水準 0.05 のとき、 $X = 6$ は棄却域には入らない。よって、帰無仮説は棄却されず、一郎くんの方が（次郎くんより）強いとは言えない。帰無仮説が棄却されないとき、「一郎くんと次郎くんの強さが同じ」とも言えないことに注意されたい。