以下の問題を解答するにあたっては，必要に応じて 4 ページの正規分布表を用いてもよい．また，ここでの**晴れ**の定義については，気象庁の天気概況の「快晴」または「晴」とする．

(1) 太郎さんは，自分が住んでいる地域において，日曜日に**晴れ**となる確率を考えている．

晴れの場合は 1 ，**晴れ**以外の場合は 0 の値をとる確率変数を X と定義する．また，$X=1$ である確率を p とすると，その確率分布は表 1 のようになる．

表　1

X	0	1	計
確　率	$1-p$	p	1

この確率変数 X の平均 (期待値) を m とすると

$$m = \boxed{\text{ア}}$$

となる．

太郎さんは，ある期間における連続した n 週の日曜日の天気を，表 1 の確率分布をもつ母集団から無作為に抽出した大きさ n の標本とみなし，それらの X を確率変数 X_1，X_2，$\cdots$，X_n で表すことにした．そして，その標本平均 $\overline{X}$ を利用して，母平均 m を推定しようと考えた．実際に $n=300$ として**晴れ**の日数を調べたところ，表 2 のようになった．

表　2

天　気	日　数
晴れ	75
晴れ以外	225
計	300

母標準偏差を σ とすると，$n=300$ は十分に大きいので，標本平均 $\overline{X}$ は近似的に正規分布 $N\left(m, \boxed{\text{イ}}\right)$ に従う．

一般に，母標準偏差 σ がわからないとき，標本の大きさ n が大きければ，σ の代わりに標本の標準偏差 S を用いてもよいことが知られている．S は

$$\begin{aligned} S &= \sqrt{\frac{1}{n}\{(X_1-\overline{X})^2+(X_2-\overline{X})^2+\cdots+(X_n-\overline{X})^2\}} \\ &= \sqrt{\frac{1}{n}({X_1}^2+{X_2}^2+\cdots+{X_n}^2)-\boxed{\text{ウ}}} \end{aligned}$$

で計算できる．ここで，${X_1}^2=X_1$，${X_2}^2=X_2$，$\cdots$，${X_n}^2=X_n$ であることに着目し，右辺を整理すると，$S=\sqrt{\boxed{\text{エ}}}$ と表されることがわかる．

よって，表 2 より，大きさ $n=300$ の標本から求められる母平均 m に対する信頼度 95% の信頼区間は $\boxed{\text{オ}}$ となる．

$\boxed{\text{ア}}$ の解答群

⓪ p	① p^2	② $1-p$	③ $(1-p)^2$

$\boxed{\text{イ}}$ の解答群

⓪ σ	① σ^2	② $\dfrac{\sigma}{n}$	③ $\dfrac{\sigma^2}{n}$	④ $\dfrac{\sigma}{\sqrt{n}}$

$\boxed{\text{ウ}}$，$\boxed{\text{エ}}$ の解答群 (同じものを繰り返し選んでもよい．)

⓪ $\overline{X}$	① $(\overline{X})^2$	② $\overline{X}(1-\overline{X})$	③ $1-\overline{X}$

$\boxed{\text{オ}}$ については，最も適当なものを，次の⓪～⑤のうちから一つ選べ．

⓪ $0.201 \leqq m \leqq 0.299$	① $0.209 \leqq m \leqq 0.291$
② $0.225 \leqq m \leqq 0.250$	③ $0.225 \leqq m \leqq 0.275$
④ $0.247 \leqq m \leqq 0.253$	⑤ $0.250 \leqq m \leqq 0.275$

(2) ある期間において，「ちょうど 3 週続けて日曜日の天気が晴れになること」がどのくらいの頻度で起こり得るのかを考察しよう．以下では，連続する k 週の日曜日の天気について，(1) の太郎さんが考えた確率変数のうち X_1，X_2，$\cdots$，X_k を用いて調べる．ただし，k は 3 以上 300 以下の自然数とする．

X_1，X_2，$\cdots$，X_k の値を順に並べたときの 0 と 1 からなる列において，「ちょうど三つ続けて 1 が現れる部分」を A とし，A の個数を確率変数 U_k で表す．例えば，$k=20$ とし，X_1，X_2，$\cdots$，X_{20} の値を順に並べたとき

$$1,\ 1,\ 1,\ 1,\ 0,\ \underbrace{\underline{1,\ 1,\ 1}}_{\text{A}},\ 0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 1,\ 0,\ 0,\ \underbrace{\underline{1,\ 1,\ 1}}_{\text{A}}$$

であったとする．この例では，下線部分は A を示しており，1 が四つ以上続く部分は A とはみなさないので，$U_{20}=2$ となる．

$k=4$ のとき，X_1，X_2，X_3，X_4 のとり得る値と，それに対応した U_4 の値を書き出すと，表 3 のようになる．

表　3

X_1	X_2	X_3	X_4	U_4
0	0	0	0	0
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
1	1	0	0	0
1	0	1	0	0
1	0	0	1	0
0	1	1	0	0
0	1	0	1	0
0	0	1	1	0
1	1	1	0	1
1	1	0	1	0
1	0	1	1	0
0	1	1	1	1
1	1	1	1	0

ここで，U_k の期待値を求めてみよう．(1) における p の値を $p = \dfrac{1}{4}$ とする．$k = 4$ のとき，U_4 の期待値は

$$E(U_4) = \frac{\boxed{\text{カ}}}{128}$$

となる．$k = 5$ のとき，U_5 の期待値は

$$E(U_5) = \frac{\boxed{\text{キク}}}{1024}$$

となる．

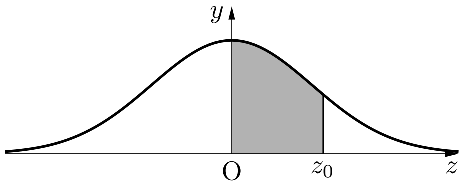
4 以上の k について，k と $E(U_k)$ の関係を詳しく調べると，座標平面上の点 $(4,\ E(U_4))$，$(5,\ E(U_5))$，$\cdots$，$(300,\ E(U_{300}))$ は一つの直線上にあることがわかる．この事実によって

$$E(U_{300}) = \frac{\boxed{\text{ケコ}}}{\boxed{\text{サ}}}$$

となる．

正　規　分　布　表

次の表は，標準正規分布の分布曲線における右図の灰色部分の面積の値をまとめたものである．



z_0	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

(24 共通テスト本試験 IIB 3)

【答】

ア	イ	ウ	エ	オ	カ	キク	ケコ	サ
0	3	1	2	0	3	33	21	8

【解答】

(1) 確率変数 X のとりうる値は 0, 1 であり，それぞれの確率は

$$P(X=0)=1-p,\quad P(X=1)=p$$

であるから，X の平均 (期待値)m は

$$m=0\cdot(1-p)+1\cdot p=\boldsymbol{p}\quad (0) \qquad \cdots\cdots(答)$$

となる．

母平均 m，母標準偏差 σ の母集団からの標本平均 $\overline{X}$ は近似的に

$$\text{正規分布 } N\left(m,\ \boldsymbol{\frac{\sigma^2}{n}}\right)\quad (3) \qquad \cdots\cdots(答)$$

に従う．

一般に，母標準偏差 σ がわからないとき，標本の大きさ n が大きければ，σ の代わりに標本の標準偏差 S を用いてもよいことが知られている．S は

$$S=\sqrt{\frac{1}{n}\{(X_1-\overline{X})^2+(X_2-\overline{X})^2+\cdots+(X_n-\overline{X})^2\}}\quad (\because\ 定義)$$

$$=\sqrt{\frac{1}{n}({X_1}^2+{X_2}^2+\cdots+{X_n}^2)-\boldsymbol{\overline{X}^2}}\quad (\because\ 公式)\quad (1) \qquad \cdots\cdots(答)$$

で計算できる．ここで，${X_1}^2=X_1$，${X_2}^2=X_2$，$\cdots$，${X_n}^2=X_n$ であることに着目すると

$$S=\sqrt{\frac{1}{n}(X_1+X_2+\cdots+X_n)-\overline{X}^2}$$

$$=\sqrt{\overline{X}-\overline{X}^2}$$

$$=\sqrt{\boldsymbol{\overline{X}(1-\overline{X})}}\quad (2) \qquad \cdots\cdots(答)$$

と表されることがわかる．

表 2 より

$$\overline{X}=\frac{75}{300}=\frac{1}{4},\quad S=\sqrt{\frac{1}{4}\left(1-\frac{1}{4}\right)}=\frac{\sqrt{3}}{4}$$

標本の大きさ $n=300$ は十分大きいから，$\overline{X}$ は正規分布 $N\left(m,\ \dfrac{S^2}{n}\right)$ に従い，確率変数 $Z=\dfrac{\overline{X}-m}{\dfrac{S}{\sqrt{n}}}$ は標準正規分布 $N(0,\ 1)$ に従う．正規分布表から

$$P(|Z|\leqq 1.96)\fallingdotseq 0.95$$

であり，$|Z|\leqq 1.96$ を書き換えると

$$-1.96\leqq\frac{m-\dfrac{1}{4}}{\dfrac{\sqrt{3}}{4\sqrt{300}}}\leqq 1.96$$

$$\frac{1}{4}-1.96\cdot\frac{1}{40}\leqq m\leqq\frac{1}{4}+1.96\cdot\frac{1}{40}$$

$$0.25-0.049\leqq m\leqq 0.25+0.049$$

よって，母平均 m に対する 95%の信頼区間は

$$\mathbf{0.201} \leqq \boldsymbol{m} \leqq \mathbf{0.299} \quad (0) \qquad \cdots\cdots(答)$$

となる．

(2) $P(X=1)=p=\dfrac{1}{4}$ である．

U_4 の値は $U_4=1,\ 0$ のいずれかであり，$U_4=1$ となるのは

「$1,1,1,0$」,「$0,1,1,1$」

の 2 通りがある．それ以外は $U_4=0$ である．

$$P(U_4=1)=\left(\frac{1}{4}\right)^3\cdot\frac{3}{4}+\frac{3}{4}\cdot\left(\frac{1}{4}\right)^3=\frac{6}{4^4}=\frac{3}{128}$$

U_4 の期待値は

$$E(U_4)=1\cdot\frac{3}{128}+0\cdot P(U_4=0)=\frac{\mathbf{3}}{\mathbf{128}} \qquad \cdots\cdots(答)$$

となる．

U_5 の値は $U_5=1,\ 0$ のいずれかであり，$U_5=1$ となるのは

「$1,1,1,0,\square$」,「$0,1,1,1,0$」,「$\square,\ 0,1,1,1$」 (ここで, □は 1 でも 0 でもよい)

の 3 つのタイプがある．それ以外は $U_5=0$ である．

$$\begin{aligned}P(U_5=1)&=\left(\frac{1}{4}\right)^3\cdot\left(\frac{3}{4}\right)\cdot 1+\frac{3}{4}\cdot\left(\frac{1}{4}\right)^3\cdot\frac{3}{4}+1\cdot\left(\frac{3}{4}\right)\cdot\left(\frac{1}{4}\right)^3\\&=\frac{12+9+12}{4^5}=\frac{33}{1024}\end{aligned}$$

U_5 の期待値は

$$E(U_5)=1\cdot\frac{33}{1024}+0\cdot P(U_5=0)=\frac{\mathbf{33}}{\mathbf{1024}} \qquad \cdots\cdots(答)$$

となる．

- $P(U_5=1)$ は U_4 をもとに計算することもできる．

(i) $U_4=1$ かつ $X_5=0$ のとき，$U_5=1$ となる．このときの確率は

$$P(U_4=1)\times\frac{3}{4}=\frac{6}{4^4}\times\frac{3}{4}=\frac{18}{4^5}$$

(ii) $U_4=1$ かつ $X_5=1$ のとき，$U_5=1$ となるのは「1,1,1,0,1」のときだけであり，このときの確率は

$$\left(\frac{1}{4}\right)^3\cdot\frac{3}{4}\times\frac{1}{4}=\frac{3}{4^5}$$

(iii) $U_4=0$ かつ $X_5=0$ のとき，$U_5=1$ となるのはことはない．

(iv) $U_4=0$ かつ $X_5=1$ のとき，$U_5=1$ となるのは「□,0,1,1,1」(ここで, □は 1 でも 0 でもよい) のときであり，このときの確率は

$$1\cdot\frac{3}{4}\cdot\left(\frac{1}{4}\right)^2\times\frac{1}{4}=\frac{3}{4^4}=\frac{12}{4^5}$$

以上から

$$P(U_5=1)=\frac{18}{4^5}+\frac{3}{4^5}+0+\frac{12}{4^5}=\frac{33}{1024}$$

である．

座標平面上の点 $(4,\ E(U_4))$，$(5,\ E(U_5))$，$\cdots$，$(300,\ E(U_{300}))$ は一つの直線 $y=px+q$ 上にあることを認めると，$\left(4,\ \dfrac{3}{128}\right)$，$\left(5,\ \dfrac{33}{1024}\right)$ はこの直線 $y=px+q$ 上の点であるから

$$\begin{cases} \dfrac{3}{128}=4p+q \\ \dfrac{33}{1024}=5p+q \end{cases}$$

$$\therefore\quad p=\frac{33}{1024}-\frac{3}{128}=\frac{33-3\cdot 8}{1024}=\frac{9}{1024}$$

$$q=\frac{33}{1024}-5\cdot\frac{9}{1024}=-\frac{12}{1024}$$

点 $(300,\ E(U_{300}))$ も直線 $y=\dfrac{9}{1024}x-\dfrac{12}{1024}$ 上の点であるから

$$E(U_{300})=\frac{9}{1024}\cdot 300-\frac{12}{1024}=\frac{2688}{1024}=\boldsymbol{\frac{21}{8}}$$ ……(答)

となる.