

数理統計学・期末試験問題 (2013.07.24)

- [1]–[6] は必答. [7]–[8] から 1 題だけを選択解答せよ. (100 点満点)
- 電卓などの計算機の使用禁止.
- 提出する解答用紙には学籍番号と氏名を記入せよ.
- 判読不能な文字 (薄い, 小さい, 汚いなど) や論理不明瞭な文章は読みません.
- 試験終了後, 問題の解説を担当者のホームページに掲載するので参考にされたい.

[1] (必答) 10 本中あたりが 2 本含まれているくじがある (1 回引いたくじは元に戻さない). このくじを 3 人が順に引くとき, 1 番目, 2 番目, 3 番目に引く人の当たる確率は同じであることを示せ. (10 点)

[2] (必答) 中心を O とする半径 R の円の内部にランダムに 1 点を選び, その点と中心 O との距離を X とする. X の分布関数 $F(x) = P(X \leq x)$ を求めよ. 次に, それを用いて, X の確率密度関数, 平均, 分散を求めよ. (20 点)

[3] (必答) 標準正規分布表を用いて次の問いに答えよ. (10 点)

- (1) X が正規分布 $N(4, 5^2)$ に従う確率変数であるとき $P(X \leq 5.6)$ を求めよ.
- (2) 大規模な選抜試験が実施され, 上位 10% が合格となる. 試験の結果, 平均点は 65 点, 標準偏差は 8 点であった. 受験者全体の得点分布は正規分布であると仮定して, 合格するための最低点を求めよ.

[4] (必答) 「100 万世帯に対して番組 A の視聴率調査を行った. 600 世帯を無作為抽出して視聴率 22.1% が得られたので, 標準的な計算によって, 信頼係数 95% の信頼区間 $22.1 \pm 3.3\%$ が導かれた」このことに対して以下の問いに答えよ. (20 点)

- (1) 信頼区間とは何か? 信頼区間の求め方を確率論的な根拠も合わせて説明せよ.
- (2) 信頼係数 95% を保ったまま, 信頼区間の幅をより狭く $22.1\% \pm 0.33\%$ のように精度を高めるためには, どうすればよいか? 理由も合わせて述べよ.

[5] (必答) ある国では, 病気 A の感染者は 100 人に 4 人の割合であるという. 検査 B は, 感染者の 90% に陽性反応を示すが, 非感染者の 5% にも陽性反応が出てしまう. ある人がこの検査を受けて陽性反応が出た. この人が感染者である確率 (% で表わし, 小数第 1 位を四捨五入せよ) を求めよ. (10 点)

[6] (必答) コインを 400 回投げたところ表が 222 回出た. このコインは公平であるといえるだろうか? 有意水準 5% で仮説検定せよ. 有意水準 1% ではどうか? これら 2 つの結論を比較検討せよ. (10 点)

[7] (選択) サイコロを 2 個投げて出た目の大きい方を X , 小さい方を Y とする. ただし, 同じ目が出たときは, $X = Y$ とする. (20 点)

- (1) 確率 $P(X = 4)$ を求めよ.
- (2) 条件付き確率 $P(Y \leq 2 | X \geq 5)$ を求めよ.
- (3) X の平均値 $\mathbf{E}[X]$ を計算せよ.
- (4) X, Y の共分散 $\sigma_{XY} = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])]$ を計算せよ.

[8] (選択) コインを 400 回投げたとき, 表が 215 回出た. 有意水準 5% の仮説検定によってコインは公平といえるかどうかを判定するときに生ずる第 2 種誤り確率とは何か説明せよ. 次に, 第 2 種誤り確率が 5% 以下になるような状況はどのような場合であるか答えよ. (20 点)

付録：標準正規分布表 $P = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-x^2/2} dx$

z	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	0.4773	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	0.4983	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

数理統計学 (2013.07.24 実施) 期末試験解説

[1] 1 番目の人が当たる事象を A , 2 番目の人が当たる事象を B , 3 番目の人が当たる事象を C とする. 樹形図を用いてもよいし, 実質的に同じことであるが, 条件付確率を繰り返し用いてもよい. 題意から

$$P(A) = \frac{2}{10} \quad P(A^c) = \frac{8}{10}.$$

は明らか. 1 番目が引き終わった状況から,

$$P(B|A) = \frac{1}{9}, \quad P(B|A^c) = \frac{2}{9}$$

であるから,

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) = \frac{1}{9} \cdot \frac{2}{10} + \frac{2}{9} \cdot \frac{8}{10} = \frac{18}{90} = \frac{2}{10} = P(A).$$

同様に, 2 番目が引き終わった状況から,

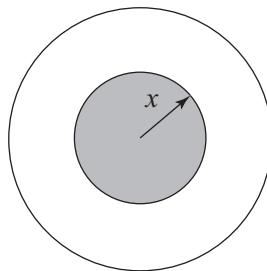
$$P(C|A \cap B) = 0, \quad P(C|A \cap B^c) = P(C|A^c \cap B) = \frac{1}{8}, \quad P(C|A^c \cap B^c) = \frac{2}{8}$$

したがって,

$$\begin{aligned} P(C) &= P(C|A \cap B)P(A \cap B) + P(C|A \cap B^c)P(A \cap B^c) \\ &\quad + P(C|A^c \cap B)P(A^c \cap B) + P(C|A^c \cap B^c)P(A^c \cap B^c) \\ &= \frac{1}{8} \cdot \frac{2}{10} \cdot \frac{8}{9} + \frac{1}{8} \cdot \frac{8}{10} \cdot \frac{2}{9} + \frac{2}{8} \cdot \frac{8}{10} \cdot \frac{7}{9} \\ &= \frac{144}{720} = \frac{2}{10} = P(A) \end{aligned}$$

上記の条件付き確率による説明は, 樹系図によるものと同じである.

[2]



題意から X は $0 \leq X \leq R$ の範囲の値をとるから, $x < 0$ では $F(x) = 0$, $x \geq R$ では $F(x) = 1$ である. そこで, $0 \leq x \leq R$ とする. $X \leq x$ はランダムに選んだ 1 点と中心 O との距離が x 以下となることを意味するが, それはランダム点が O を中心とする半径 x の円板から選ばれたことと同値である. したがって,

$$F(x) = P(X \leq x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}.$$

以上まとめて,

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \frac{x^2}{R^2}, & 0 \leq x \leq R, \\ 1, & x \geq R. \end{cases}$$

これを微分したものが確率密度関数である:

$$f(x) = \begin{cases} \frac{2x}{R^2} \cdot & 0 \leq x \leq R, \\ 0, & \text{その他.} \end{cases}$$

その平均値は

$$m = \int_0^R x f(x) dx = \frac{2}{3} R$$

分散は,

$$\sigma^2 = \int_0^R x^2 f(x) dx - m^2 = \int_0^R \frac{2x^3}{R^2} dx - m^2 = \frac{R^2}{2} - \left(\frac{2}{3} R\right)^2 = \frac{1}{18} R^2$$

[3] (1) $X \sim N(4, 5^2)$ から $Z = \frac{X-4}{5} \sim N(0, 1)$ に注意して,

$$P(X \leq 5.6) = P\left(\frac{X-4}{5} \leq \frac{5.6-4}{5}\right) = P(Z \leq 0.32) = 0.5 + 0.1255 = 0.6255$$

(2) $X \sim N(65, 8^2)$ に対して, $P(X \geq a) = 0.1$ となるような a を求めればよい. 標準正規分布表により, $Z \sim N(0, 1)$ なら, $P(Z \geq 1.28) = 0.1$. したがって,

$$Z = \frac{X-65}{8} \geq 1.28 \quad \Leftrightarrow \quad X \geq 75.2$$

したがって, 76 点以上とればよい.

[4] (1) 内容は教科書を参照のこと.

(2) 95%信頼区間は

$$\bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

で与えられるので, 信頼区間の幅を 1/10 にするためには標本数は 100 倍必要になる.

※ 試験問題として出題しているのだから, 数理的な解答が必要。「精度を上げるためには標本数を増やせばよい」などという情緒的な解答の人は, 何のための数理統計学だったのか, 反省せよ.

[5] 病気 A に感染している確率と感染していない確率は

$$P(A) = \frac{4}{100}, \quad P(A^c) = \frac{96}{100}.$$

検査 B に陽性反応を示す確率は, 条件付確率であって,

$$P(B|A) = 0.9 \quad P(B|A^c) = 0.05$$

ベイズの公式によって,

$$\begin{aligned} P(A|B) &= \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)} \\ &= \frac{\frac{4}{100} \times 0.9}{\frac{4}{100} \times 0.9 + \frac{96}{100} \times 0.05} = \frac{3.6}{3.6 + 4.8} = 0.428 \approx 43\%. \end{aligned}$$

[6] $H_0 : p = \frac{1}{2}$, $H_1 : p \neq \frac{1}{2}$, $\alpha = 0.05$ によって検定を行う. コインを 400 回投げるときの表の回数を X とすると, $X \sim B(400, 1/2) \approx N(200, 10^2)$. H_1 から両側検定の棄却域は

$$W : \left| \frac{x - 200}{10} \right| \geq 1.96$$

実現値 $\bar{x} = 222$ は W に落ちる. したがって, 有意水準 5% で H_0 は棄却され, コインは公正ではないと判定される.

有意水準 $\alpha = 0.01$ とすると, 両側検定の棄却域は

$$W : \left| \frac{x - 200}{10} \right| \geq 2.58$$

であるから, 実現値 $\bar{x} = 222$ は W に落ちない. したがって, 有意水準 1% で H_0 は棄却されず, 高度に有意とは判定できない.

[7] サイコロを 2 個振って出た目の大きいほうを X , 小さいほうを Y とする. サイコロの 1 個目の目を i , 2 個目の目を j とすると, 標本空間と確率は,

$$\Omega = \{\omega = (i, j); i, j \in \{1, 2, 3, 4, 5, 6\}\}, \quad P(\{\omega\}) = P(\{(i, j)\}) = \frac{1}{36}$$

で与えられる. X, Y は Ω で定義された確率変数であり,

$$X : \Omega \ni \omega = (i, j) \mapsto \max\{i, j\}, \quad Y : \Omega \ni \omega = (i, j) \mapsto \min\{i, j\},$$

で与えられることになる. X, Y の値について表を作ると, 次のようになる.

$X(i, j)$							$Y(i, j)$						
$i \backslash j$	1	2	3	4	5	6	$i \backslash j$	1	1	1	1	1	1
1	1	2	3	4	5	6	1	1	1	1	1	1	1
2	2	2	3	4	5	6	2	1	2	2	2	2	2
3	3	3	3	4	5	6	3	1	2	3	3	3	3
4	4	4	4	4	5	6	4	1	2	3	4	4	4
5	5	5	5	5	5	6	5	1	2	3	4	5	5
6	6	6	6	6	6	6	6	1	2	3	4	5	6

(1) 事象 $\{X = 4\}$ は 7 個の根元事象から成るから,

$$P(X = 4) = \frac{7}{36}$$

(2) まず,

$$\{X \geq 5\} = \left\{ \begin{array}{cccccc} & & & & (1, 5) & (1, 6) \\ & & & & (2, 5) & (2, 6) \\ & & & & (3, 5) & (3, 6) \\ & & & & (4, 5) & (4, 6) \\ (5, 1) & (5, 2) & (5, 3) & (5, 4) & (5, 5) & (5, 6) \\ (6, 1) & (6, 2) & (6, 3) & (6, 4) & (6, 5) & (6, 6) \end{array} \right\},$$

$$\{X \geq 5, Y \leq 2\} = \left\{ \begin{array}{cc} & (1, 5) \quad (1, 6) \\ & (2, 5) \quad (2, 6) \\ (5, 1) & (5, 2) \\ (6, 1) & (6, 2) \end{array} \right\}$$

したがって,

$$P(Y \leq 2 | X \geq 5) = \frac{P(Y \leq 2, X \geq 5)}{P(X \geq 5)} = \frac{8/36}{20/36} = \frac{2}{5}$$

(3) X の確率分布がわかる.

k	1	2	3	4	5	6
$P(X = k)$	1/36	3/36	5/36	7/36	9/36	11/36

平均値は公式にあてはめて計算する.

$$\begin{aligned} \mathbf{E}(X) &= \sum_{k=1}^6 kP(X = k) \\ &= 1 \times \frac{1}{36} + 2 \times \frac{3}{36} + 3 \times \frac{5}{36} + 4 \times \frac{7}{36} + 5 \times \frac{9}{36} + 6 \times \frac{11}{36} = \frac{161}{36} \end{aligned}$$

(4) $\sigma_{XY} = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$ を用いるのが便利である. まず, $XY = ij$ で i, j が独立であることから,

$$\mathbf{E}[XY] = \mathbf{E}[ij] = \mathbf{E}[i]\mathbf{E}[j] = \frac{7}{2} \cdot \frac{7}{2} = \frac{49}{4}$$

(3) と同様にして,

$$\mathbf{E}(Y) = \frac{91}{36}$$

したがって,

$$\sigma_{XY} = \frac{49}{4} - \frac{161}{36} \cdot \frac{91}{36} = \frac{49 \cdot 9 \cdot 36 - 161 \cdot 91}{36^2} = \frac{15876 - 14651}{36^2} = \frac{15876 - 14651}{36^2} = \frac{1225}{36^2}$$

[8] $H_0 : p = \frac{1}{2}, H_1 : p \neq \frac{1}{2}, \alpha = 0.05$ によって検定を行う. コインを 400 回投げるときの表の回数を X とすると, $X \sim B(400, 1/2) \approx N(200, 10^2)$. H_1 から両側検定の棄却域は

$$W : \left| \frac{x - 200}{10} \right| \geq 1.96$$

実現値 $\bar{x} = 215$ は W に落ちない. したがって, 有意水準 5% で H_0 は棄却されず, コインは公正であると判定される. この判定が間違っている確率, すなわち, コインは公正ではないにもかかわらず公正であると判定してしまう確率が第 2 種誤り確率である. コインが公正ではない場合, 可能な p は無限にあり, 第 2 種誤り確率を評価することができない.

たとえば, $p = 0.525$ (左図) と $p = 0.6$ (右図) の場合を考えよう. $p = 0.525$ の場合は, $p = 0.5$ の分布と近いため, 採択域の範囲にかぶさってくる範囲がかなり大きく, 図からわかるように $\beta \geq 0.5$ である. つまり, 公正なコインに近い偽物に対しては, 公正であるという判定をしがちであり, その間違い確率 (第 2 種誤り確率) は極めて大きい.

$p = 0.6$ の場合は, 2 つの分布がかなり離れており, $\beta \leq 0.025$ となることが見て取れる. したがって, $\beta = 0.05$ となるのは p が 0.6 より少し小さい時である. このとき, 分散はほぼ 10^2 と考えてよいので, $219.6 + 1.64 \times 10 = 236$ から $p = 0.59$ がわかる. つまり, 調べているコインが公平 ($p = 0.5$) であるか, あるいは表が出やすく $p \geq 0.59$ であるか, いずれかであれば, 第 2 種誤り確率が 0.05 に抑え込まれる.

