数理統計学・期末試験問題 (2017.07.19)

- **[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) 2 つの事象 E, F は独立であって, $P(E) = a$, $P(F) = b$ がわかっている. (5 点 ×2)

(1) $P(E \cup F)$ を a, b で表せ.

(2) 条件付確率 $P(E \cap F^c | E \cup F)$ を a, b で表せ. ただし, F^c は F の余事象である.

[2] (必答) 長さ L の棒をランダムに折って長いほうの断片の長さを X とする. (5 点 ×4)

(1) X の密度関数を求めよ.

(2) (1) を用いて, 長いほうの断片の長さが短いほうの 2 倍以上になる確率を求めよ.

(3) X の平均値を求めよ.

(4) X の分散を求めよ.

[3] (必答) 大規模な選抜試験が実施され, 上位から 33% までが合格となる. 試験の結果, 平均点は 58 点, 標準偏差は 12 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整数で求めよ. (10 点)

[4] (必答) ある国では, 病気 A の感染者は 500 人に 2 人の割合であるという. 検査 B は, 感染者の 95%に陽性反応を示すが, 非感染者の $100p$ % にも陽性反応が出てしまう $(0 < p < 1)$. この検査を受けて陽性反応が出たとき, その人が感染者である確率を $f(p)$ とする. (5 点 ×2)

(1) $f(p)$ を求め, $f(p)$ のグラフの特徴を述べよ.

(2) $f(p) \geq 0.5$ となるのは p がどのような範囲にあるときか?

[5] (必答) 100 万世帯から 600 世帯を無作為抽出して, 番組 A の視聴率調査を行い, 信頼係数 95% の信頼区間 $22.1 \pm 3.3\%$ が得られたという. ここでいう信頼区間とはどのようにして導かれ, どのような意味をもつか? 500 文字を目安に (数式を除く. この問題文の 1 行が 50 文字程度である) 説明せよ. (15 点)

[6] (必答) ある食品の製造ラインでは, 1 袋の重量が 102g になるように調整している. 一方, 工程の特性によって, 1 袋の重量は標準偏差 0.36g の正規分布であることが経験的に知られている. あるロットから選んだ 9 個の標本の平均重量は 101.74g であった. 製造ラインに狂いが生じているかどうかを仮説検定で判定せよ. 標準偏差 0.36g があらかじめ知られていないときは, どのような検定をすればよいか答えよ. (15 点)

[7] (選択) コインを 256 回投げたところ, 表が 116 回出た. このコインが公平かどうかを有意水準 5% の仮説検定によって判定したい. (10 点 ×2)

(1) 「第 2 種誤り確率」の定義を述べて, その基本的な性質を上の問題に即して説明せよ.

(2) 第 2 種誤り確率が 10%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.

[8] (選択) 確率変数 X はパラメータ $\lambda > 0$ の指数分布に従うものとする．すなわち, $0 < a < b$ に対して,

$$P(a \leq X \leq b) = \int_a^b \lambda e^{-\lambda x} dx$$

が成り立つ．次の問いに答えよ．(10 点 ×2)

(1) X の平均値 $\mathbf{E}[X]$ と分散 $\mathbf{V}[X]$ を求めよ．

(2) 確率変数 Y を

$$Y = \begin{cases} [X], & X \geq 0, \\ 0, & X < 0 \end{cases}$$

と定義する．ただし, 実数 x に対して, $[x]$ は x を超えない最大の整数を表す．確率変数 Y の平均値 $\mathbf{E}[Y]$ を計算せよ．

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・期末試験解説 (2017.07.19)

[1] (1) 独立性から $P(E\cap F)=P(E)P(F)$ となることを用いる. 排反 $(P(E\cup F)=P(E)+P(F))$ と勘違いしないこと.

$$P(E\cup F)=P(E)+P(F)-P(E\cap F)=P(E)+P(F)-P(E)P(F)=a+b-ab.$$

(2) まず, $(E\cap F^c)\cap(E\cup F)=E\cap F^c$ を確認する (集合算で計算するか, またはベン図でも描く). そして,

$$P(E\cap F^c)=P(E)-P(E\cap F)=a-ab.$$

そうすれば,

$$P(E\cap F^c|E\cup F)=\frac{P((E\cap F^c)\cap(E\cup F))}{P(E\cup F)}=\frac{a-ab}{a+b-ab}\,.$$

[2] (1) 題意から X は区間 $[L/2,L]$ に値をとる連続型確率変数になる. したがって, 分布関数 $F(x)=P(X\le x)$ について, $x<L/2$ ならば $F(x)=0$, $x>L$ ならば $F(x)=1$ は明らか. $L/2\le x\le L$ としよう. $X\le x$ となるのは, ランダム点が図内の線分から選ばれた時で, その確率は,

$$P(X\le x)=\frac{L-2(L-x)}{L}=\frac{2}{L}\,x-1$$

で与えられる. したがって,

$$F(x)=\begin{cases}0, & x<L/2,\\ \dfrac{2}{L}\,x-1, & L/2\le x\le L,\\ 1, & x>L.\end{cases}$$

微分して, 密度関数が求められる.

$$f(x)=\begin{cases}\dfrac{2}{L} & L/2\le x\le L,\\ 0, & \text{その他}.\end{cases}$$

これは, $[L/2,L]$ 上の一様分布である.

(2) 長いほうの断片の長さが短いほうの 2 倍以上になるのは,

$$X\ge 2(L-X)\quad\Longleftrightarrow\quad X\ge\frac{2}{3}\,L$$

の時である. したがって, 求める確率は

$$P\Big(X\ge\frac{2}{3}\,L\Big)=\int_{\frac{2}{3}L}^{+\infty}f(x)dx=\int_{\frac{2}{3}L}^{L}\frac{2}{L}\,dx=\frac{2}{L}\times\frac{L}{3}=\frac{2}{3}\,.$$

(3)

$$\mathbf{E}[X]=\int_{-\infty}^{+\infty}xf(x)=\int_{L/2}^{L}x\,\frac{2}{L}\,dx=\frac{3}{4}\,L.$$

(4)

$$\mathbf{E}[X^2]=\int_{-\infty}^{+\infty}x^2f(x)=\int_{L/2}^{L}x^2\,\frac{2}{L}\,dx=\frac{7}{12}\,L^2.$$

したがって,

$$\mathbf{V}[X]=\mathbf{E}[X^2]-\mathbf{E}[X]^2=\frac{7}{12}\,L^2-\left(\frac{3}{4}\,L\right)^2=\frac{L^2}{48}\,.$$

[3] 受験者の得点を X, 合格の最低点を a とすれば, $P(X \geq a) = 0.33$ が成り立つ. $X \sim N(58, 12^2)$ であるから,

$$P(X \geq a) = P\left(\frac{X-58}{12} \geq \frac{a-58}{12}\right) = P\left(Z \geq \frac{a-58}{12}\right) = 0.33.$$

ここで, $Z \sim N(0,1)$ である. 標準正規分布表によって, $P(Z \geq 0.44) = 0.33$ であるから, 求める a は,

$$\frac{a-58}{12} = 0.44 \quad \Longleftrightarrow \quad a = 63.28$$

したがって, 64 点が合格するための最低点である.

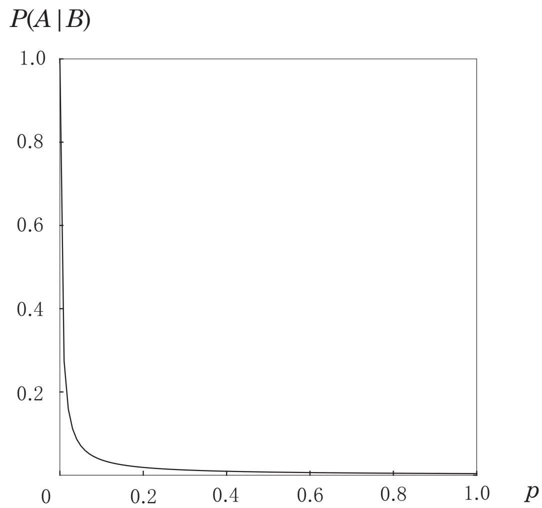
[4] 感染している事象を A, 陽性反応が出る事象を B とすると,

$$P(A) = \frac{2}{500}, \qquad P(A^c) = \frac{498}{500}, \qquad P(B|A) = 0.95, \qquad P(B|A^c) = p.$$

ここで, ベイズの公式によって,

$$f(p) = P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)} = \frac{\frac{2}{500} \times 0.95}{\frac{2}{500} \times 0.95 + \frac{498}{500} \times p} = \frac{1.9}{1.9 + 498p}$$

ここで, p はこの検査薬の一種のエラーを表すといえる. $p = 0$ ならエラーなしなので, $P(A|B) = 1$ となり, 陽性反応が出れば感染者と断言できる. 実際上, $p = 0$ とはできないが, できるだけ 0 に近いほうがよい. また, $P(A|B)$ は p とともに単調減少することも明らかである. しかし, この程度の観察では不十分. 重要なことは, $P(A|B)$ が $p = 0$ から少しでも $p > 0$ となるだけで劇的に減少することである.



(2) $f(p) \geq 0.5$ は,

$$\frac{1.9}{1.9 + 498p} \geq 0.5$$

を意味する.

$$p \leq \frac{1.9}{498} = 0.0038.$$

[5] 省略

[6] $H_0 : m = 102, \quad H_1 : m \neq 102$ とおく. 大きさ 9 の標本平均は $\bar{X} \sim N\left(102, \frac{0.36^2}{9}\right) = N(102, 0.12^2)$. 規準化して

$$Z = \frac{\bar{X} - 102}{0.12} \sim N(0,1).$$

実現値 $\bar{x} = 101.74$ を代入して,

$$z = \frac{101.74 - 102}{0.12} = -\frac{0.26}{0.12} = -2.17$$

H_1 から両側検定となる．有意水準 $\alpha = 0.05$ (5%) に対する棄却域は $|z| \geq 1.96$. 実現値 $z = -2.17$ は棄却域に落ちる．したがって, 有意水準 5% で H_0 は棄却され, 製造ラインに狂いが生じていると判定される．

有意水準 $\alpha = 0.01$ とすると, 両側検定の 5%棄却域は $|z| \geq 2.58$. 実現値 $z = -2.17$ は棄却域に落ちない．したがって, 有意水準 1% では H_0 は棄却されない．（高度に有意ではない）

母分散が未知の時は t-検定を行う．詳しくは教科書を参照.

[7] (1)

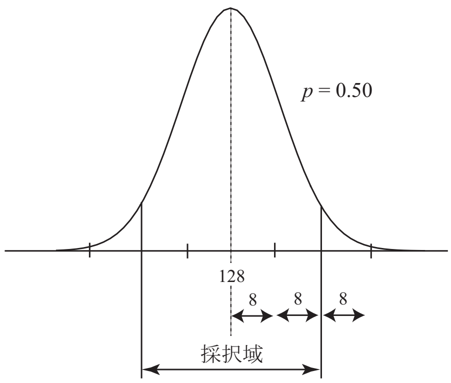
$$H_0 : p = \frac{1}{2} \qquad H_1 : p \neq \frac{1}{2} \qquad \alpha = 0.05$$

によって検定を行う． コインを 256 回投げるときの表の回数を X とすると, $X \sim B(256, 1/2) \approx N(128, 8^2)$. H_1 から両側検定であり, 棄却域は

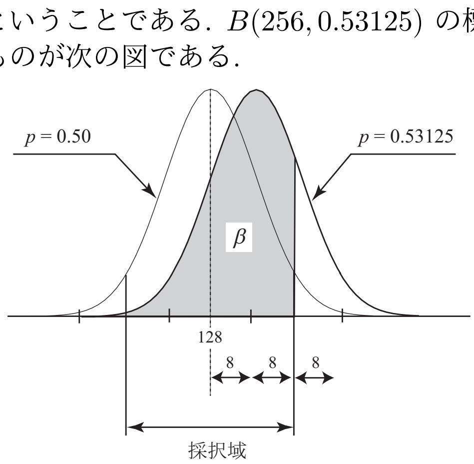
$$W : \left| \frac{x - 128}{8} \right| \geq 1.96 \iff x \leq 128 - 15.68 \text{ または } 128 + 15.68 \leq x$$

実現値 $x = 116$ 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, コインは公平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, コインは公平ではないのにもかかわらず公平であると判定してしまう確率が第 2 種誤り確率である．コインが公平ではない場合, 可能な p は無限にあり, 第 2 種誤り確率を簡単に評価することはできない.

採択域を図示したものが次の図である．$1.96 \approx 2$ として説明すれば十分である.

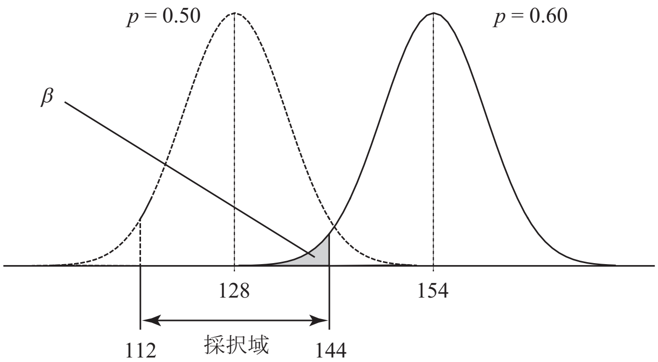


イカサマコインの例示として, 平均値が $136 = 128 + 8$ であるコインを投げていたとする．逆算して $p = 0.53125$ のイカサマコインということである．$B(256, 0.53125)$ の標準偏差も約 8 として見積もってよい．その分布を重ねて書いたものが次の図である.



採択域に実現値が現れる確率は, 網掛け部分の面積であり, これが第 2 種誤り確率 β である. 明らかに, $\beta > 0.5$ となるほどに大きい. つまり, 投げているコインがイカサマであっても, 公平なコインに近い ($p = 0.53125$) 場合は, それを区別することは難しく, 仮説検定では「公平なコインである」と判定されがちである. つまり, 第 2 種の誤りを犯しやすい.

(2) 第 2 種誤り確率が小さい状況とは, 投げているコインのイカサマ度が大きく, その分布が $p = 1/2$ の分布とかけ離れている場合である. たとえば, $p = 0.6$ の分布を書いて, 採択域に実現値が出る確率を見積もろう. $p = 0.6$ のとき, 表の回数は概ね $N(154, 8^2)$ に従うとしてよい. そうすると, β は標準正規分布の $z = (154 - 144)/8 = 1.25$ より上側に対応する確率である. 標準正規分布表によって $\beta = 0.5 - 0.3944 \approx 0.1$ である. したがって, 表の出る確率が $p \leq 0.4$ あるいは $p \geq 0.6$ のように公平なコインから相当にずれているときは, $\beta \leq 0.1$ といえる.



[8] (1)

$$\mathbf{E}[X] = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda},$$
$$\mathbf{V}[X] = \int_0^\infty (x - \mathbf{E}[X])^2 \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2}.$$

(2) Y は $\{0, 1, 2, \dots\}$ に値をとる離散型確率変数になる.

$$P(Y = n) = \int_n^{n+1} \lambda e^{-\lambda x} dx = e^{-\lambda n} - e^{-\lambda(n+1)} = e^{-\lambda n}(1 - e^{-\lambda}), \qquad n = 0, 1, 2, \dots.$$

したがって,

$$\mathbf{E}[Y] = \sum_{n=0}^\infty nP(Y = n) = \sum_{n=0}^\infty ne^{-\lambda n}(1 - e^{-\lambda}) = \frac{e^{-\lambda}}{1 - e^{-\lambda}}$$