

第 10 回統計の考え方レポート

- 1 1 と書いたカードを 1 枚、2 と書いたカード 2 枚、...、 n と書いたカード n 枚 がある。この中から 1 枚のカードを取りだすとき、カードの示す数 X を確率変数とする。

- (1) $X = k$ である確率 p_k を求めよ。
- (2) X の平均値 m を求めよ。
- (3) X の標準偏差 σ を求めよ。

- 2 白球 4 個、赤球 5 個が入っている袋の中から 5 個取りだすとき、赤球の含まれる個数の確率分布を求めよ。

- 3 ある体格検査で次のような身長の数値分布表が得られた。これから平均の身長を求めよ。

身長 (cm)	人数
150 - 155	3
155 - 160	6
160 - 165	32
165 - 170	17
170 - 175	10
175 - 180	2

- 4 3 面に数字 1 を刻み、他の 3 面にそれぞれ 2, 3, 4 の数字を刻んだサイコロがある。このようなサイコロ 2 個を同時に振って出た目の和の期待値すなわち和の平均値を求めよ。ただし、このサイコロのどの面の現われる見込みも同等であるとする。

- 5 サイコロを 10 回振るとき、1 の目が出る回数の平均値を求めよ。

- 6 50 歳の夫と 48 歳の妻が 20 年後まで生存する確率は、夫が 0.15、妻が 0.2 である。現在夫が 50 歳、妻が 48 歳である 10 組の夫婦が 20 年後に、夫婦のうち少なくとも 1 方が生存している組の数を X とする。 X が二項分布に従う確率変数であるとして、 X の平均値を求めよ。

- 7 袋の中に 20 枚のカードが入っている。それぞれ、1, 2, 3, $\dots$, 20 の数字が記入されている。最初に抽出した 1 枚の数値を X_1 とし、残りから抽出した数値を X_2 とする。 $X_1 + X_2$ の期待値 (平均値) を求めよ。

- 8 X は 1, 2, $\dots$, n を等しい確率で取る確率変数であって、分散は 10 である。このとき n を求めよ。

- 9 あるバレーボールの試合では、どちらかのチームが 3 セット先に勝つと試合が終了する。今、A, B 両チームがこのバレーボールの試合を行うとき、1 セットの勝負で A が B に勝つ確率を p 、負ける確率を q ($p + q = 1$) とするとき、 p はこの試合を通して一定で、各セットの勝負は独立であるとする。このとき、次の問に答えよ。

- (1) この試合が 4 セットで終了する確率
- (2) この試合が終了するまで要するセット数を X とするときの X の確率分布
- (3) この試合で $p = \frac{1}{2}$ とするとき (2) の X の平均値と標準偏差

- 10 数字 0, 1, 2 を 1 つずつ記入した 3 枚のカードの中から任意に 1 枚を抜き取ってカードに記入された数字 X を読む。次にそのカードをもとに戻して、さらに 1 枚を抜き取ってカードに記入された数字 Y を読む。このとき $Z = XY$ として、

- (1) Z のとるそれぞれの値に対する確率を求めよ。
- (2) それぞれの確率変数 Z の平均値と分散を求めよ。

- 21 次の調査をするには、全数調査と標本調査のどちらがよいか、理由をつけて答えよ。

- (1) 新薬の薬効
(2) 缶詰の不良品検査
(3) 選挙での結果の予想

22 ある中学校の 3 年生全員について、国語の成績の分布のようすを調べるために、1 組から 5 組まで 5 クラス 200 人の中から標本を選んで調査することにした。このとき、標本の選びかたとして適切なものを次の (a) ~ (e) から選べ。ただし、答えは 1 つとは限らない。

- (a) 1 番から 5 番までのくじをつくり、1 つのクラスを選んでそのクラスの生徒全員を標本にする。
(b) 各クラスで投票をして、8 人ずつ選ぶ。
(c) 全員でくじ引きをして、40 人選ぶ。
(d) 男子だけ調べる。
(e) 出席番号を 5 で割って、3 余る生徒を選ぶ。

23 袋の中に白と黒の碁石がたくさん入っている。この袋をよくかき混ぜてから 10 個の碁石を取りだし、白と黒の碁石の数を調べる。この作業をくり返したところ、次のような結果になった。以下の問に答えよ。

回数	1	2	3	4	5	6	7	8	9	10
白の碁石	6	6	7	5	8	7	4	5	6	6
黒の碁石	4	4	3	5	2	3	6	5	4	4

- (1) この実験で取りだされる白の碁石の平均個数を求めよ。
(2) 黒の碁石の全体に対する割合は、だいたいどのくらいであると推定できるか。
(3) 袋の中に碁石が 130 個入っていたとすると、黒の碁石は何個入っていると推定できるか。

24 ある山で、サル 30 匹を捕獲し、印をつけて放した。後日、またサルを 30 匹捕獲したところ、印のついているサルは 4 匹いた。この山にはおよそ何匹のサルが生息していると考えられるか。

25 64 枚のカードに右の表のように数字を書いた。それらのカードをよく切って、10 枚取りだし、それらの数字の平均を記録することを 100 回くり返し、その結果を度数分布表にした。度数分布表によって、標本平均の平均値を求めよ。ただし、小数第 3 位を四捨五入して答えること。

数字	枚数
1	1
2	6
3	15
4	20
5	15
6	6
7	1

標本平均以上 未満	回数
3.0 ~ 3.3	1
3.3 ~ 3.6	11
3.6 ~ 3.9	23
3.9 ~ 4.2	34
4.2 ~ 4.5	21
4.5 ~ 4.8	7
4.8 ~ 5.1	3
計	100

26 袋の中に、区別がつかない白のピンポン球とオレンジのピンポン球が 420 個入っている。この中から 20 個のピンポン球を取りだし、白とオレンジの割合を 10 回くり返して調べたところ、標本平均の平均がおよそ 3 : 7 になった。このとき、次の問に答えよ。

- (1) この袋にオレンジ色のピンポン球は何個入っていると考えられるか。
(2) この袋の中から 70 個のピンポン球を取り出したとき、白のピンポン球は何個含まれていると考えられるか。

第 9 回統計の考え方解答

- 4 与式の展開式における一般項が ${}_{10}C_r 1^{10-5} (2x)^r = 2^r {}_{10}C_r x^r$ で与えられる。となりあう係数の差が初めて負になったとき、すなわち $f(r+1) - f(r) < 0$ となるとき、その r が最大係数の値となる。 $\frac{f(r+1)}{f(r)} < 1$ かどうかを調べる。

$$\frac{2^{r+1} {}_{10}C_{r+1}}{2^r {}_{10}C_r} = 2 \cdot \frac{\frac{10 \cdot 9 \cdots (11-r) \cdot (10-r)}{(r+1) \cdot r \cdots 2 \cdot 1}}{\frac{10 \cdot 9 \cdots (11-r)}{r \cdots 2 \cdot 1}} = 2 \frac{10-r}{r+1} < 1$$

これを解いて、 $\frac{19}{3} < r$ 数直線から $r = 7$ のとき。

- 6 (1) $x + y = 1$ より、 $(n+1)yx^n = (n+1)(1-x)x^n$ となる。 $f(x) = (n+1)(1-x)x^n$ を考えると、 $x > 0$ より、 $f'(x) = 0$ なる x で最大最小を取る。 $f'(x) = (n+1)x^{n-1}\{n - (n+1)x\}$ である。

従って、関数の極値を調べると、

x		$\frac{n}{n+1}$	
$f'(x)$		0	
$f'(x)$	$\nearrow$	最大	$\searrow$

ただし最大は $f(\frac{n}{n+1}) = \left(\frac{n}{n+1}\right)^n$

(2)

$$\left(1 + \frac{1}{n}\right)^n = 1 + {}_nC_1 \frac{1}{n} + {}_nC_2 \left(\frac{1}{n}\right)^2 + \cdots + {}_nC_{n-1} \left(\frac{1}{n}\right)^{n-1} + \left(\frac{1}{n}\right)^n \geq 1 + {}_nC_1 \frac{1}{n} = 1 + n \times \frac{1}{n} = 2$$

展開した第 2 項以降は、必ず正の数の和となるため。等号は $n = 1$ のとき。

- (3) (2) より、 $\left(\frac{n+1}{n}\right) \geq 2$ であるので、 $\left(\frac{n}{n+1}\right) \leq \frac{1}{2}$ ところが、(1) からは、 $(n+1)yx^n \leq \left(\frac{n}{n+1}\right)$ となる。これらの結果を合わせて、 $(n+1)yx^n \leq \frac{1}{2}$ となる。等号は $n = 1$ のとき。

- 7 どの目の出る確率も $\frac{1}{6}$

- 8 当たりくじの数で求めるので $p_3 = 0$ $p_2 = \frac{{}_2C_2 \cdot {}_8C_1}{{}_{10}C_3} = \frac{1}{15}$, $p_1 = \frac{{}_2C_1 \cdot {}_8C_2}{{}_{10}C_3} = \frac{7}{15}$, $p_0 = \frac{{}_8C_3}{{}_{10}C_3} = \frac{7}{15}$ となる。確率分布表は

X	0	1	2	3	計
p_X	$\frac{7}{15}$	$\frac{7}{15}$	$\frac{1}{15}$	0	1

- 9 X は 0, 1, 2 の値をとり、その確率は $p_0 = \frac{1}{4}$, $p_1 = \frac{2}{4} = \frac{1}{2}$ となる。確率分布表は

X	0	1	2	計
p_x	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

- 10 2 つのさいころの目の和は 2, 3, $\dots$, 12 の値をとる。場合の数を数えると以下ようになる。

目の和	2	3	4	5	6	7	8	9	10	11	12	計
場合の数	1	2	3	4	5	6	5	4	3	2	1	36

確率分布表は場合の数を全体の組み合わせ $6^2 = 36$ で割ることによって得られる。

目の和 X	2	3	4	5	6	7	8	9	10	11	12	計
$p(X)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

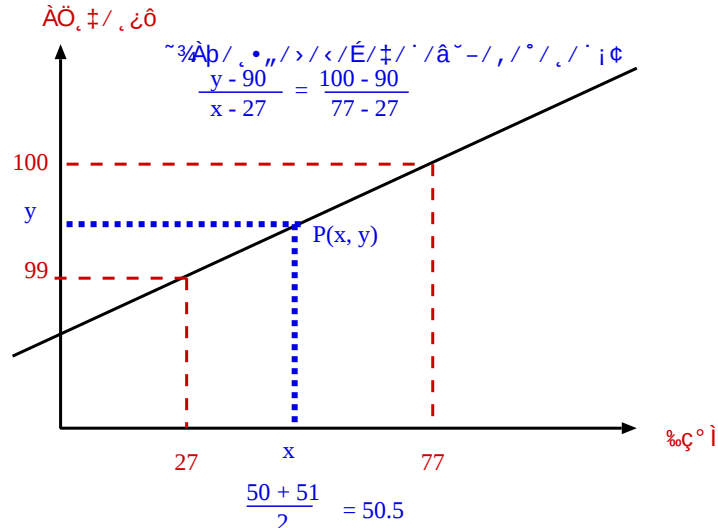
約分を実行してあっても正解とする。数の数え方をみせるため、ここでは約分していない。

- 1 小さい順にならべて、中央の値を取りだす。奇数なら真ん中、偶数ならとなりあう数字の平均値を取る。

$$M_e = \frac{1}{2}(169.8 + 170.5) = 170.15$$

- 2 真ん中の来そうな階級を調べる。1 カプセルが 91 - 100 粒のところに 50 位と 51 位は表れるだろう。全部で 100 カプセルある。50.5 位の粒を調べればよい。27 位が 90 粒あり、77 位が 100 粒であった

と考える。傾きが一緒であることから、直線の方程式を求め、 $x = 50.5$ を代入して、 $y = 94.7$ を得る。



- 3 度数のもっとも多いのは 7 である。よってその階級値が Mo であり、 $Mo = 14$
- 4 間隔の平均 $\bar{m}$ を求めると、 $\bar{m} = 100.454545454545 \dots$ となるので、有効数字は年の最小単位が 1 年であるのでその $\frac{1}{10}$ 、すなわち小数第 1 位まで。よって、小数第 2 位を四捨五入して、100.5 年と求められる。
- 5 表を作って計算しよう。 $\bar{x} = \frac{1799}{40} = 44.975$ ここで、有効数字を考えると最小単位が 1 kg より、小数第 1 位までとなる。したがって、小数第 2 位を四捨五入して 45.0 と求められる。
- 6 ワークシートのように仮平均を書くと分かりやすい。 $m = 113.75$ と求められるが、得点の最小単位が 1 点なので、小数第 1 位まで求めることができればよい。従って、小数第 2 位を四捨五入して、113.8 と求められる。
- 7 練習だと思ってワークシートのように表を作成して計算してみよう。 $m = 3$, $\sigma^2 = 2$ と求められる。
- 8 分散の公式が与えられたので、使ってみよう。 $\bar{x} = 9$, $\sum_{k=1}^{50} x_k^2 = 4900$, $N = 50$ である。度数は 1 ずつなので、 $f_k = 1$ ($k = 1, 2, \dots, 50$) より、 $\sigma^2 = \frac{1}{50} \sum_{k=1}^{50} x_k^2 \cdot 1 - \bar{x}^2 = \frac{1}{50} \times 4900 - 9^2 = 17$
- 9 階級は階級の中央の値を取る。ワークシートを作ってみよう。各値は、座高の最小単位が 1cm だから、小数第 1 位までとなる。よって小数第 2 位を四捨五入して、 $m = \frac{3870}{44} = 87.954545454 = 88.0$ となる。 $\sigma^2 = \frac{1}{44} \times 340703 - 87.95^2 = 8.0475 = 8.0$ となる。標準偏差は $\sqrt{8.0475} = 2.83681159050085665007 = 2.8$ 講義で学んだような仮平均を使ってもよい。そのほうが計算が楽である。同じ計算が楽にできることを確かめよう。

10 平均値を m とすると、

$$m^2 = \left(\frac{1}{5} \sum_{i=1}^5 x_i^2 \right)^2 = \frac{1}{25} \left(\sum_{i=1}^5 x_i^2 + 2 \sum_{i=1}^5 \sum_{j=1}^5 x_i x_j \right) = 25$$

ここで、 $x_i > 0$ より $m > 0$ であるので、 $m = 5$ 。分散 σ^2 は

$$\sigma^2 = \frac{1}{5} \sum_{i=1}^5 x_i^2 - m^2 = \frac{1}{5} \times 151 - 5^2 = 5.2$$

11 n 人中賛成の割合が p なので、賛成の人数は np 人である。反対者は $n - np = (1 - p)n$ 人。

(1) 度数分布表は

(2) 平均値 m は上の表から計算して $m = \{0 \times n(1 - p) + 1 \times np\} = p$ 。分散 σ^2 は、

$$\sigma^2 = \frac{1}{N} \sum (x^2 - m^2) = \frac{1}{n} \{0^2 \times n(1 - p) + 1^2 \times np\} - p^2 = p(1 - p)$$

従って、 $\sigma = \sqrt{p(1 - p)}$

階級値 x	度数 f
0	$(1 - p)n$
1	np
計	n