

1 次の 2 次元のデータ (A, B) に対して、 A と B の共分散 $Cov[A, B]$ を求めて下さい。

$$(A, B) : (3, 5), (2, 8), (7, 6), (1, 5), (1, 8), (2, 8)$$

ただし、結果は一切の近似をせずに完全に正しい値を、有限小数か出来るだけ約分された分数の形で答えて下さい。

配点：20 点

シラバス到達目標：ア、イ

【解答例】

$$\begin{aligned} E[AB] &= \frac{15 + 16 + 42 + 5 + 8 + 16}{6} = \frac{102}{6} = 17 \\ E[A] &= \frac{3 + 2 + 7 + 1 + 1 + 2}{6} = \frac{16}{6} = \frac{8}{3} \\ E[B] &= \frac{5 + 8 + 6 + 5 + 8 + 8}{6} = \frac{40}{6} = \frac{20}{3} \end{aligned}$$

なので、

$$\begin{aligned} Cov[A, B] &= E[AB] - E[A]E[B] \\ &= 17 - \frac{8 \cdot 20}{3^2} \\ &= \frac{153 - 160}{9} \\ &= -\frac{7}{9} \end{aligned}$$

です。

□

2 密度関数が

$$h(x) = \begin{cases} 2x & 0 \leq x \leq \frac{1}{2} \\ -2x + 2 & \frac{1}{2} \leq x \leq 1 \\ 2x - 2 & 1 \leq x \leq \frac{3}{2} \\ -2x + 4 & \frac{3}{2} \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

で与えられる確率変数 X に対して確率 $P\left[\frac{1}{2} \leq X \leq \frac{9}{4}\right]$ を求めて下さい。

配点：15 点

シラバス到達目標：ア、ウ

【解答例】

$$\begin{aligned} P\left[\frac{1}{2} \leq X \leq \frac{9}{4}\right] &= 1 - P\left[0 \leq X \leq \frac{1}{2}\right] \\ &= 1 - \int_0^{\frac{1}{2}} 2x dx \\ &= 1 - \left[x^2\right]_0^{\frac{1}{2}} \\ &= \frac{3}{4} \end{aligned}$$

□

3 次の中から正しいものを 1 つ選んでください。

- (1) 正規母集団からとった大きさ 3 の標本分散は自由度 2 のカイ自乗分布に従います。
- (2) $Var[X] = 7, Var[Y] = 1$ で X, Y が独立であるとき、 $Var[X - 2Y] = 11$ です。
- (3) 母集団が正規分布に従っていれば、標本平均の分散は母集団の分散に一致します。
- (4) 他の条件を変えずに有意水準を 1 % から 5 % にすると棄却域は小さくなります。
- (5) 2 次元の確率変数 (A, B) があって A, B それぞれの密度関数が $f(x), g(y)$ であれば、 (A, B) の密度関数は $f(x)g(y)$ です。

配点：5 点 シラバス到達目標：ア、エ、オ、キ

【解答例】 (2)。

□

4 不良品が 0.8 % 混ざっている部品の山から 250 個とったとき、不良品が 3 個以上入っている確率を求めてください。

ただし、部品は沢山あるので復元抽出と見なせるとします。

配点：10 点 シラバス到達目標：エ

【解答例】 250 個とったときの不良品の数を X とすると X は 2 項分布 $B(250, 0.008)$ に従います。参考資料の事実 2 によれば、 X はパラメータ 2 のポアソン分布 $P_O(2)$ で近似されます。

$$\begin{aligned} P[X \geq 3] &\approx P[P_O(2) \geq 3] \\ &= 1 - P[P_O(2) \leq 2] \\ &= 1 - (1 + 2 + 2)e^{-2} \\ &\approx 1 - 5 \cdot 0.135 \\ &= 0.325 \end{aligned}$$

中心極限定理によって近似した場合は

$$\begin{aligned} P[X \geq 3] &= P[B(250, 0.008) \geq 3] \\ &\approx P[N(2, 2 \cdot 0.992) \geq 3] \\ &= P[N(0, 1.984) \geq 1] \\ &= P\left[N(0, 1) \geq \frac{1}{\sqrt{1.984}}\right] \\ &\approx P\left[N(0, 1) \geq \frac{1}{1.41}\right] \\ &\approx P[N(0, 1) \geq 0.709] \\ &= 0.5 - P[0 \leq N(0, 1) \leq 0.709] \\ &\approx 0.5 - 0.2611 \\ &= 0.2389 \end{aligned}$$

です。

(本来は使えないが) 2 項分布を正規分布で近似する事実 1 によって半整数補正を加えて近似した場合は

$$\begin{aligned}P[X \geq 3] &= P[B(250, 0.008) \geq 3] \\&\approx P[N(2, 2 \cdot 0.992) \geq 2.5] \\&= P[N(0, 1.984) \geq 0.5] \\&= P\left[N(0, 1) \geq \frac{0.5}{\sqrt{1.984}}\right] \\&\approx P\left[N(0, 1) \geq \frac{0.5}{1.41}\right] \\&\approx P[N(0, 1) \geq 0.3546] \\&= 0.5 - P[0 \leq N(0, 1) \leq 0.3546] \\&\approx 0.5 - 0.138 \\&= 0.362\end{aligned}$$

です。

□

5 2 つの独立な確率変数 X, Y があって、 X は正規分布 $N(1, 4)$ 、 Y は正規分布 $N(0, 9)$ に従うとき、確率 $P[2X + Y \geq 7.65]$ を求めてください。

配点：10 点	シラバス到達目標：ウ、エ、オ
---------	----------------

【解答例】 $2X + Y$ は正規分布 $N(2, 25)$ に従いますから

$$\begin{aligned}P[2X + Y \geq 7.65] &= P[N(2, 25) \geq 7.65] \\&= P[N(0, 25) \geq 5.65] \\&= P[N(0, 1) \geq 1.13] \\&= 0.5 - P[0 \leq N(0, 1) \leq 1.13] \\&\approx 0.5 - 0.3708 \\&= 0.1292\end{aligned}$$

です。

□

6 コインを 3 回投げて表の出た回数を X 、表の出た回数と裏の出た回数の差の絶対値を Y とします。

- (1) $P[X = 0, Y = 1]$ を計算して下さい。
- (2) X, Y が独立であるかどうか理由をつけて述べて下さい。
- (3) $X + Y$ の期待値と分散を求めて下さい。

配点：(1)(2) 合わせて 5 点、(3)5 点	シラバス到達目標：ア、イ、エ
---------------------------	----------------

【解答例】 (1) $Y = |X - (3 - X)| = |2X - 3|$ ですから、 $X = 0$ のとき $Y = 3$ であって、

$$P[X = 0, Y = 1] = 0$$

です。

(2) X は 2 項分布 $B(3, 0.5)$ に従いますから

$$P[X = 0] = \frac{1}{2^3}, \quad P[Y = 1] = P[|2X - 3| = 1] = P[X = 1, 2] = \frac{3}{2^3} + \frac{3}{2^3} = \frac{6}{2^3}$$

であって、

$$P[X = 0, Y = 1] \neq P[X = 0]P[Y = 1]$$

ですから、独立ではありません。

(3)

$$\begin{aligned} E[X + Y] &= E[X + |2X - 3|] \\ &= \sum_{j=0}^3 (j + |2j - 3|) \binom{3}{j} \left(\frac{1}{2}\right)^3 \\ &= (3 + 6 + 9 + 6) \frac{1}{8} \\ &= 3 \end{aligned}$$

$$\begin{aligned} Var[X + Y] &= E[(X + Y)^2] - E[X + Y]^2 \\ &= \sum_{j=0}^3 (j + |2j - 3|)^2 \binom{3}{j} \frac{1}{2^3} - 9 \\ &= (9 + 12 + 27 + 36) \frac{1}{8} - 9 \\ &= 1.5 \end{aligned}$$

□

7 ある工場で生産している製品の重さは、通常は平均値が 80g、標準偏差が 4g の正規分布に従っていますが、最近『重くなっているのでは?』という疑義が生じています。そこで 50 個の標本をとって有意水準 5 % で『平常と比べて重い』と言えるかどうか仮説検定を行うことになりました。ただし、標準偏差の値は疑われておらず、4g であると考えます。

ある日の製品の中から 50 個の標本を抽出して測定したところ、重さの平均値が 80.8g でした。その日の製品は平常と比べて重いと言えるでしょうか。有意水準 5 % で検定して下さい。

配点：15 点	シラバス到達目標：ケ
---------	------------

【解答例】

帰無仮説 H_0 ：重さの平均値は 80g である。

対立仮説 H_1 ：重さの平均値は 80g より重い。

帰無仮説が正しいと仮定します。すると母集団は $N(80, 4^2)$ に従いますから、ここからとった大きさ 50 の標本平均 $\bar{X}$ は $N\left(80, \frac{4^2}{50}\right)$ に従います。

重くなっているかどうかの問題ですから、右片側棄却域を求めます。

$$\begin{aligned}
 0.05 &= P[d \leq \bar{X}] \\
 &= P\left[d \leq N\left(80, \frac{4^2}{50}\right)\right] \\
 &= P\left[d - 80 \leq N\left(0, \frac{4^2}{50}\right)\right] \\
 &= P\left[\frac{d - 80}{\sqrt{\frac{4^2}{50}}} \leq N(0, 1)\right] \\
 0.45 &= P\left[0 \leq N(0, 1) \leq \frac{d - 80}{\sqrt{\frac{4^2}{50}}}\right]
 \end{aligned}$$

正規分布表によれば

$$\frac{d - 80}{\sqrt{\frac{4^2}{50}}} \approx 1.645 \quad \text{従って} \quad d \approx 80 + 1.645 \cdot \frac{4}{\sqrt{50}} \approx 80.933$$

ですから、

$$0.05 \approx P[80.933 \leq \bar{X}]$$

が分かります。

従って有意水準 0.05 での右片側棄却域は $[80.933, +\infty)$ となります。

今回のサンプル値 80.8 は棄却域に入っていないので、帰無仮説を棄却するに足る合理的な理由はなく、重さの平均値は 80g より重かったとは言えないと判断されます。□

8.1 [選択問題] ある選挙での候補者 A の得票率を p とします。400 人に出口調査をした結果、80 人が候補者 A に投票したと言っています。400 人のランダムサンプルにおける候補者 A に投票した人数を X として以下の問いに答えて下さい。ただし、有権者数が非常に大きいのでサンプルの抽出は復元抽出とみなせるものとします。

(1) X はどんな分布に従いますか。 p を使って表現してください。

(2) X を正規分布 $N(400p, 400 \cdot 0.2 \cdot 0.8)$ で近似して (その際、半整数補正はしないでください) 候補者 A の得票率 p の信頼度 0.95 の信頼区間を求めて下さい。

配点：(1)5 点、(2)10 点 シラバス到達目標：ク

【解答例】 (1) 2 項分布 $B(400, p)$ 。

(2) X を正規分布 $N(400p, 64)$ で近似して

$$0.95 = P[|X - 400p| \leq d]$$

となるような d を求めます。

$$0.95 = P[|X - 400p| \leq d]$$

$$\approx P[|N(400p, 64) - 400p| \leq d]$$

$$= P\left[|N(0, 1)| \leq \frac{d}{8}\right]$$

$$0.475 = P\left[0 \leq N(0, 1) \leq \frac{d}{8}\right]$$

ですから、正規分布表により $\frac{d}{8} \approx 1.96$ 、すなわち $d \approx 15.68$ が分かります。

従って

$$0.95 = P[|X - 400p| \leq 15.68]$$

ですから、今回のサンプル値 80 について、信頼度 95 % で

$$|80 - 400p| \leq 15.68$$

$$80 - 15.68 \leq 400p \leq 80 + 15.68$$

$$64.32 \leq 400p \leq 95.68$$

$$0.1608 \leq p \leq 0.2392$$

が成り立ちますから、求める得票率 p の信頼区間は $[0.1608, 0.2392]$ です。

□

8.2 [選択問題] ある母集団から大きさ 50 の無作為サンプルを抽出して、そのサンプルの平均値が 54.3、分散が 24.5 でした。母平均の信頼度 99 % の信頼区間を求めて下さい。

配点：15 点 シラバス到達目標：ク

【解答例】 サンプルサイズが大きいので、サンプルの分散で母分散の代用とします。また、中心極限定理によれば、母平均を m とするとこの母集団からとった大きさ 50 の標本平均 $\bar{X}$ は正規分布 $N\left(m, \frac{24.5}{50}\right)$ で近似されます。

まず

$$0.99 = P[|\bar{X} - m| \leq d]$$

となるような $d > 0$ を求めます。

$$0.99 = P[|\bar{X} - m| \leq d]$$

$$\approx P\left[\left|N\left(m, \frac{24.5}{50}\right) - m\right| \leq d\right]$$

$$= P[|N(0, 0.49)| \leq d]$$

$$= P\left[|N(0, 1)| \leq \frac{d}{0.7}\right]$$

$$0.495 = P\left[0 \leq N(0, 1) \leq \frac{d}{0.7}\right]$$

すると正規分布表より

$$2.575 \approx \frac{d}{0.7} \quad \text{従って} \quad d \approx 1.8025$$

が得られますので、

$$0.99 \approx P[|\bar{X} - m| \leq 1.80]$$

が成り立ちます。従って今回のサンプル値 54.3 について、信頼度 99 % で

$$|54.3 - m| \leq 1.80$$

$$54.3 - 1.80 \leq m \leq 54.3 + 1.80$$

$$52.50 \leq m \leq 56.10$$

が成り立ちますから、求める信頼区間は $[52.5, 56.1]$ です。

□