

問題 1 次の2次元データ (A, B) について、各成分の期待値と分散 $(E[A], Var[A], E[B], Var[B])$ を求め、共分散 $Cov[A, B]$ も求めて下さい。

データ (A, B)
$(8, 3), (6, 6), (5, 4), (7, 3), (7, 5), (7, 6), (11, 5), (7, 9), (4, 2), (8, 5)$

ただし、四捨五入等の近似は一切せず 有限小数の形で解答して下さい。

配点：25点 シラバス達成度目標：ア、イ

【解答例】

$$E[A] = \frac{8 + 6 + 5 + 7 + 7 + 7 + 11 + 7 + 4 + 8}{10} = 7$$

$$Var[A] = E[(A - E[A])^2] = \frac{1^2 + 1^2 + 2^2 + 4^2 + 3^2 + 1^2}{10} = \frac{32}{10} = 3.2$$

$$E[B] = \frac{3 + 6 + 4 + 3 + 5 + 6 + 5 + 9 + 2 + 5}{10} = \frac{48}{10} = 4.8$$

$$\begin{aligned} Var[B] &= E[B^2] - E[B]^2 \\ &= \frac{9 + 36 + 16 + 9 + 25 + 36 + 25 + 81 + 4 + 25}{10} - \left(\frac{48}{10}\right)^2 \\ &= \frac{266}{10} - \frac{24^2}{5^2} \\ &= \frac{266 \cdot 5 - 24^2 \cdot 2}{50} \\ &= \frac{1330 - 1152}{50} \\ &= \frac{178}{50} \\ &= 3.56 \end{aligned}$$

$$\begin{aligned} Cov[A, B] &= E[AB] - E[A]E[B] \\ &= \frac{24 + 36 + 20 + 21 + 35 + 42 + 55 + 63 + 8 + 40}{10} - 7 \cdot \frac{48}{10} \\ &= \frac{344 - 7 \cdot 48}{10} \\ &= \frac{8}{10} \\ &= 0.8 \end{aligned}$$

- $E[A]$ 、 $E[B]$ 、 $Var[A]$ 、 $Var[B]$ 、 $Cov[A, B]$ それぞれ5点
- 計算ミスだけのもの 3点
- 定義が間違っているもの 2点
- 書き写しミスなどケアレスミスのみのもの 4点

問題 2 独立な成分をもつ 2 次元の確率変数 (X, Y) において、 X, Y の密度関数がそれぞれ次の $f(x), g(x)$ であるとき、 $X + Y$ の密度関数、すなわちこれらのたたみ込みを求めて下さい。

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}, \quad g(x) = \begin{cases} x & 0 \leq x \leq \sqrt{2} \\ 0 & \text{otherwise} \end{cases}.$$

配点：10 点 | シラバス達成度目標：イ、オ

【解答例】 まずたたみ込みの定義と $g(x)$ の定義から

$$\begin{aligned} (f * g)(x) &= \int_{-\infty}^{\infty} f(x-y)g(y)dy \\ &= \int_0^{\sqrt{2}} f(x-y)ydy \end{aligned}$$

です。ここで $f(x-y)$ は $0 \leq x-y \leq 1$ 、すなわち $x-1 \leq y \leq x$ のところ以外では 0 ですからこの区間 $[x-1, x]$ と積分区間の $[0, \sqrt{2}]$ の重なり具合によって値が変わって来ます。

【 $x \leq 0$ のとき】

このとき 2 つの区間は内部に共通点を持ちませんから $(f * g)(x) = 0$ です。

【 $0 \leq x \leq 1$ のとき】

このとき、2 つの区間の重なりは $[0, x]$ ですから、

$$(f * g)(x) = \int_0^x ydy = \frac{1}{2}x^2$$

です。

【 $1 \leq x \leq \sqrt{2}$ のとき】

このときは区間の重なりは $[x-1, x]$ ですから、

$$(f * g)(x) = \int_{x-1}^x ydy = \frac{1}{2}\{x^2 - (x-1)^2\} = x - \frac{1}{2}$$

となります。

【 $\sqrt{2} \leq x \leq \sqrt{2} + 1$ のとき】

このとき 2 つの区間の重なりは $[x-1, \sqrt{2}]$ ですから、

$$(f * g)(x) = \int_{x-1}^{\sqrt{2}} ydy = \frac{1}{2}\{2 - (x-1)^2\} = \frac{1}{2}(1 + 2x - x^2)$$

です。

【 $\sqrt{2} + 1 \leq x$ のとき】

このときは 2 つの区間に交わりはありません（端点のみ）から、 $(f * g)(x) = 0$ です。

以上から

$$(f * g)(x) = \begin{cases} 0 & \sqrt{2} + 1 \leq x \\ \frac{1}{2}(1 + 2x - x^2) & \sqrt{2} \leq x \leq \sqrt{2} + 1 \\ x - \frac{1}{2} & 1 \leq x \leq \sqrt{2} \\ \frac{1}{2}x^2 & 0 \leq x \leq 1 \\ 0 & x \leq 0 \end{cases}$$

です。

- ミスの多いもの 3 点
- ミスの少ないもの 7 点
- 最初の定義式程度のもの 1 点

問題 3 密度関数が次の $h(x)$ で与えられる確率変数 W に対して確率 $P[-2.2 \leq W \leq 0.3]$ と期待値 $E[W]$ を求めて下さい。

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

配点：15 点 シラバス達成度目標：ウ

【解答例】

$$\begin{aligned} P[-2.2 \leq W \leq 0.3] &= \int_{-2.2}^{0.3} h(x) dx \\ &= \int_{-1}^{0.3} \frac{3}{4}(1-x^2) dx \\ &= \left[\frac{3}{4}x - \frac{1}{4}x^3 \right]_{-1}^{0.3} \\ &= \frac{0.9 - 0.027}{4} + \frac{3}{4} - \frac{1}{4} \\ &= \frac{2.873}{4} \\ &= 0.71825 \end{aligned}$$

● 確率で 8 点
計算ミス少しのもの 6 点
ミスが多いもの、積分範囲にミスがあるもの 3 点

● 分布表からの読み取りのミス - 3 点
● 標準偏差でなく分散で割るミス - 4 点
● 計算のミス - 3 点
● ミスが多い、あるいは重大なもの - 5 点

$$\begin{aligned} E[W] &= \int_{-\infty}^{\infty} xh(x) dx \\ &= \int_{-1}^1 \frac{3}{4}x(1-x^2) dx \\ &= 0 \end{aligned}$$

● 期待値で 7 点
計算ミスが少しのもの 5 点
ミスが多いもの、積分範囲にミスがあるもの 3 点

問題 4 標準正規分布表を使って、正規分布 $N(5, 2^2)$ に従う確率変数 Q に対して確率 $P[3 \leq Q \leq 8]$ を求めて下さい。

配点：10 点 シラバス達成度目標：エ

【解答例】

$$\begin{aligned} P[3 \leq Q \leq 8] &= P[3-5 \leq Q-5 \leq 8-5] \\ &= P[-2 \leq N(0, 2^2) \leq 3] \\ &= P\left[-\frac{2}{2} \leq \frac{N(0, 2^2)}{2} \leq \frac{3}{2}\right] \\ &= P[-1 \leq N(0, 1) \leq 1.5] \\ &= P[0 \leq N(0, 1) \leq 1] + P[0 \leq N(0, 1) \leq 1.5] \\ &= 0.3413 + 0.4332 \\ &= 0.7745 \end{aligned}$$

問題 5 次の各文章が正しいか間違っているかを○/×のみで答えて下さい。

(1) 標本分散の平均値は母分散に一致します。

(2) 平均 13、分散 2 の母集団からとった大きさ 100 の標本平均は正規分布 $N(13, \frac{2}{100})$ に従います。

(3) 大きさ 7 の標本分散は自由度 6 のカイ自乗分布に従います。

配点：各 5 点 シラバス達成度目標：カ、キ

【解答例】 (1) ×、(2) ×、(3) ×。

● 部分点なし

問題 6 ある工場で生産している製品 A の重さは、通常は平均値が 80g、標準偏差が 2g の正規分布をしています。ある日の製品の中から 50 個のサンプルをとって調べたところ、重さの平均値が 80.6g でした。この日の製品の重さの平均値は通常と比べて重いと言えるでしょうか。

この日も製品の重さの標準偏差は 2g であるとしたうえで、

帰無仮説： この日の製品の重さの平均値は通常通り 80g である。

対立仮説： この日の製品の重さの平均値は通常と比べて重い。

について有意水準 5 % で仮説検定して下さい。

配点： 15 点 シラバス達成度目標：ケ

【解答例】 帰無仮説が正しいと仮定します。このとき母集団は平均 80、分散 2^2 ですが、サンプル数が多いので中心極限定理により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 $N\left(80, \frac{2^2}{50}\right)$ で近似されます。

問題になっているのは通常と比べて重いと言えるかどうかですから、有意水準 5 % の右片側棄却域を求めます。

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

ですから、標準正規分布表から $\frac{d-80}{\sqrt{\frac{2^2}{50}}} \sim 1.645$ 、すなわち

$$d \sim 80 + 1.645 \cdot \frac{\sqrt{2}}{5} \sim 80.465$$

が得られます。従って求める棄却域は $[80.465, \infty)$ であり、今回のサンプル平均はここに入っていますから帰無仮説は棄却され、この日の製品の重さの平均値は通常と比べて重いと判断されます。

標準正規分布表の読み取りで $\frac{d-80}{\sqrt{\frac{2^2}{50}}} \sim 1.645$ としたところは、1.64 もしくは 1.65 でも構いません。それぞれの場合について解答を示しておくと、

【1.64 の場合】

$$d \sim 80 + 1.64 \cdot \frac{\sqrt{2}}{5} \sim 80.464$$

が得られ、棄却域は $[80.464, \infty)$ です。

【1.65 の場合】

$$d \sim 80 + 1.65 \cdot \frac{\sqrt{2}}{5} \sim 80.467$$

が得られ、棄却域は $[80.467, \infty)$ です。

いずれの場合もサンプル値は棄却域に入っており、結論は同じです。

また、以上の計算において、 $\sqrt{2} \sim 1.414$ として計算していますが、これも 1.41 あるいは 1.4 でも構いません。どの場合もサンプル値は棄却域に入る事になります。

●最初に帰無仮説が正しいと仮定していないもの - 1 点

●少しのみで途中のもの 2 ~ 5 点

●片側検定できちんとやっているもの

ミスの少ないもの 12 点

ミスの多いもの 10 点

●両側検定でやっているもの

大体OKなら 10 点

ミスが多いもの 6 点

問題 7 分散 3.5^2 の正規分布に従う母集団から大きさ 25 のサンプルをとってその平均値を調べたところ 209.7 でした。母平均 m の信頼度 95 % の信頼区間を求めて下さい。

- 標本平均の記号の定義のないものなど、最初の説明が不十分なもの - 1 点

配点：10 点 シラバス達成度目標：ク

【解答例】 母集団 $N(m, 3.5^2)$ からとった大きさ 25 の標本平均 $\bar{X}$ は正規分布 $N\left(m, \frac{3.5^2}{25}\right)$ に従いますので、まずは

$$P[|\bar{X} - m| \leq d] = 0.95$$

となるような正数 d を求めます。

標準化変形を行えば

$$\begin{aligned} 0.95 &= P[|\bar{X} - m| \leq d] \\ &= P\left[\left|N\left(0, \frac{3.5^2}{25}\right)\right| \leq d\right] \\ &= P\left[\left|N(0, 1)\right| \leq \frac{d}{\frac{3.5}{5}}\right] \\ 0.475 &= P\left[0 \leq N(0, 1) \leq \frac{5d}{3.5}\right] \end{aligned}$$

- 少しのミス - 2 点
- ミスが多い、あるいは重大なもの - 4 点

なので、標準正規分布表から $\frac{5d}{3.5} \sim 1.96$ 、すなわち、

$$d \sim \frac{1.96 \cdot 3.5}{5} \sim 1.372$$

が得られます。以上から

$$P[|\bar{X} - m| \leq 1.372] \sim 0.95$$

ですから、今回のサンプル平均 209.7 について、信頼度 95 % で

$$|209.7 - m| \leq 1.372$$

$$209.7 - 1.372 \leq m \leq 209.7 + 1.372$$

$$208.328 \leq m \leq 211.072$$

- この不等式に対して信頼度等の説明がないもの - 1 点

と言えます。従って求める信頼区間は $[208.328, 211.072]$ です。