

## 2009 年度 統計 中間試験 解説

問題(A)、問題(B)は、ほぼ同一の問題について出題方法を変えたものであり、難易度に大差はないことを述べておく。

### 問題(A) 解答例

Q1:

【解説】記述統計として、平均、分散(標準偏差)など標本分布の特性を表す指標の理解ができていないかを問うた問題である。問題文中にヒントを記したように、

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

の関係を使うと、計算が遥かに簡単になる。

【解答】

(1) テスト成績の(分布の)平均を求めなさい。[3点]

$$\bar{Y} = \frac{1}{10} \sum_{i=1}^{10} Y_i = \frac{350}{10} = 35$$

(2) テスト成績の(分布の)標準偏差を求めなさい。[4点]

記述統計であるから、分散を求めるときに偏差平方和を標本の大きさ  $n$  で割ることに注意。

$$\text{Var}(Y) = \frac{1}{10} \sum_{i=1}^{10} (Y_i - \bar{Y})^2 = \frac{1}{10} \left( \sum_{i=1}^{10} Y_i^2 - 10\bar{Y}^2 \right) = \frac{1}{10} (13060 - 10 \times 35^2) = \frac{810}{10} = 81$$

ゆえに、標準偏差は  $\sqrt{\text{Var}(Y)} = \sqrt{81} = 9$

(3) 練習時間の(分布の)レンジを求めなさい。[4点]

レンジは(最大値 - 最小値)として定義されるから、 $200 - 75 = 125$

(4) 練習時間とテスト成績の相関係数を求めなさい [4点]

相関係数は  $\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \times \sqrt{\text{Var}(Y)}}$  として定義されるから、

$$\begin{aligned} \rho &= \frac{\sum_{i=1}^{10} (X_i - \bar{X})(Y_i - \bar{Y}) / 10}{\sqrt{\sum_{i=1}^{10} (X_i - \bar{X})^2 / 10} \sqrt{\sum_{i=1}^{10} (Y_i - \bar{Y})^2 / 10}} = \frac{\left( \sum_{i=1}^{10} X_i Y_i - 10\bar{X}\bar{Y} \right) / 10}{\sqrt{\left( \sum_{i=1}^{10} X_i^2 - 10\bar{X}^2 \right) / 10} \sqrt{\left( \sum_{i=1}^{10} Y_i^2 - 10\bar{Y}^2 \right) / 10}} \\ &= \frac{(44180 - 10 \times 120 \times 35) / 10}{\sqrt{(160000 - 10 \times 120^2) / 10} \sqrt{810} / 10} = \frac{218}{\sqrt{1600} \sqrt{81}} = \frac{218}{40 \times 9} = 0.60\dot{5} = 0.606 \end{aligned}$$

Q2:

【解説】二項分布にしたがう確率変数  $X$  について、確率関数、分布関数、(確率分布の)平均、分散を求める問題である。また、 $X$  の一次変換によって分散がどのように変化するかについても問われている。

【解答】

(5)  $E(X)$  を求めなさい。[ 3 点 ]

実験回数を  $n$ 、成功確率を  $p$  とすると、 $E(X) = np$ 。よって、 $E(X) = 6 \times 0.4 = 2.4$

(6)  $Z$  の確率分布の標準偏差  $\sqrt{\text{Var}(Z)}$  を求めなさい。[ 4 点 ]

まず、 $\text{Var}(X) = np(1-p)$  であるから、 $\text{Var}(X) = 6 \times 0.4 \times (1-0.4) = 1.44$ 。

$Z = 2X + 1$  より、 $\text{Var}(Z) = 2^2 \text{Var}(X) = 4 \times 1.44 = 5.76$ 。

よって  $\sqrt{\text{Var}(Z)} = \sqrt{5.76} = 2.4$

(7)  $P(Z = 3)$  を求めなさい。[ 4 点 ]

$X$	0	1	2	3	4	5	6
$Z$	1	3	5	7	9	11	13

という関係があるから、

$$P(Z = 3) = P(X = 1) = {}_6C_1 \times 0.4 \times (1-0.4)^{6-1} = 6 \times \frac{2}{5} \times \left(\frac{3}{5}\right)^5 = \frac{2916}{15625} = 0.186624.$$

(8)  $F_Z(4) = P(Z \leq 4)$  を求めなさい。[ 4 点 ]

$$P(Z \leq 4) = P(X \leq 1.5) = P(X = 0) + P(X = 1) = 1 \times \left(\frac{3}{5}\right)^6 + \frac{2916}{15625} = \frac{729 + 2916}{15625} = 0.23328$$

Q3:

【解説】同時確率分布、周辺確率分布の関係と条件付確率および条件付期待値の基本を問うている。

【解答】

(9) 数値が入っていないセル(a) ~ (f)の値を求めなさい。[ 各 2 点 ]

セル(b)は  $P(Y = 10 \cap X = 5) = P(Y = 10 | X = 5)P(X = 5) = 0.25 \times 0.4 = 0.1$ 。あとは順次求めることができる。(a) 0.3, (b) 0.1, (c) 0.2, (d) 0.1, (e) 0.4, (f) 0.3。

(10) 条件付確率  $P(X = 5 | Y = 25)$  を求めなさい。[ 4 点 ]

$$\text{セル(d), (f)より } P(X = 5 | Y = 25) = \frac{P(X = 5 \cap Y = 25)}{P(Y = 25)} = \frac{0.1}{0.3} = \frac{1}{3} = 0.33.$$

(11)  $X=3$  を条件としたときの  $Y$  の条件付期待値  $E(Y | X = 3)$  を求めなさい。[ 4 点 ]

$$\begin{aligned} E(Y | X = 3) &= 10 \times P(Y = 10 | X = 3) + 20 \times P(Y = 20 | X = 3) + 25 \times P(Y = 25 | X = 3) \\ &= 10 \times \frac{0.3}{0.6} + 20 \times \frac{0.1}{0.6} + 25 \times \frac{0.2}{0.6} = \frac{10}{0.6} = \frac{50}{3} = 16.67 \end{aligned}$$

## 問題 (B) 解答例

Q1:

【解説】記述統計として、平均、分散(標準偏差)など標本分布の特性を表す指標の理解ができて  
いるかを問うた問題である。問題文中にヒントを記したように、

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

の関係を使うと、計算が遥かに簡単になる。

【解答】

(1) 練習時間の(分布の)平均を求めなさい。[3点]

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i = \frac{1200}{10} = 120$$

(2) 練習時間の(分布の)標準偏差を求めなさい。[4点]

記述統計であるから、分散を求めるときに偏差平方和を標本の大きさ  $n$  で割ることに注意。

$$\text{Var}(X) = \frac{1}{10} \sum_{i=1}^{10} (X_i - \bar{X})^2 = \frac{1}{10} \left( \sum_{i=1}^{10} X_i^2 - 10\bar{X}^2 \right) = \frac{1}{10} (160000 - 10 \times 120^2) = \frac{16000}{10} = 1600$$

ゆえに、標準偏差は  $\sqrt{\text{Var}(X)} = \sqrt{1600} = 40$

(3) テスト成績のメディアン(中央値)を求めなさい。[4点]

標本の大きさ  $n$  が 10 で偶数であるので、メディアンは小さい順にならべて 5 番目と 6 番目のデ

ータの算術平均として定義されるから、 $\frac{32+35}{2} = 33.5$ 。

(4) 練習時間とテスト成績の共分散を求めなさい [4点]

$$\text{共分散は } \text{Cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n} \quad \text{として定義されるから、}$$

$$\frac{\sum_{i=1}^{10} (X_i - \bar{X})(Y_i - \bar{Y})}{10} = \frac{\sum_{i=1}^{10} X_i Y_i - 10\bar{X}\bar{Y}}{10} = \frac{44180 - 10 \times 120 \times 35}{10} = 218.$$

Q2:

【解説】二項分布にしたがう確率変数  $X$  について、確率関数、分布関数、(確率分布の)平均、分散を求める問題である。また、 $X$  の一次変換によって分散がどのように変化するかについても問うている。

【解答】

(5)  $E(X)$  を求めなさい。[3点]

実験回数を  $n$ 、成功確率を  $p$  とすると、 $E(X) = np$ 。よって、 $E(X) = 5 \times 0.6 = 3$

(6)  $\text{Var}(Z)$  を求めなさい。[ 4 点 ]

まず、 $\text{Var}(X) = np(1-p)$  であるから、 $\text{Var}(X) = 5 \times 0.6 \times (1-0.6) = 1.2$ 。  $Z = 3X - 2$  より、 $\text{Var}(Z) = 3^2 \text{Var}(X) = 9 \times 1.2 = 10.8$ 。

(7)  $P(Z = 3)$  を求めなさい。[ 4 点 ]

$X$	0	1	2	3	4	5	6
$Z$	-2	1	4	7	10	13	16

という関係があるから、 $P(Z = 3) = P(X = 5/3) = 0$ 。

(8)  $P(Z > 3) = 1 - F_Z(3)$  を求めなさい。なお、 $F_Z(3) = P(Z \leq 3)$  である。[ 4 点 ]

$$P(Z \leq 3) = P(X = 0) + P(X = 1) = 1 \times \left(\frac{2}{5}\right)^5 + 5 \times \left(\frac{3}{5}\right) \times \left(\frac{2}{5}\right)^4 = \frac{32 + 240}{3125} = 0.08704$$

よって、 $P(Z > 3) = 1 - F_Z(3) = 1 - 0.08704 = 0.91296$ 。

Q3:

【解説】同時確率分布、周辺確率分布の関係と条件付確率および条件付期待値の基本を問うている。

【解答】

(9) 数値が入っていないセル(a) ~ (f)の値を求めなさい。[ 各 2 点 ]

セル(c)は  $P(Y = 20 \cap X = 5) = P(Y = 20 | X = 5)P(X = 5) = 0.5 \times 0.4 = 0.2$ 。あとは順次求めることができる。(a) 0.2, (b) 0.1, (c) 0.2, (d) 0.1, (e) 0.3, (f) 0.3。

(10) 条件付確率  $P(X = 3 | Y = 25)$  を求めなさい。[ 4 点 ]

$$\text{セル(a), (f)より } P(X = 3 | Y = 25) = \frac{P(X = 3 \cap Y = 25)}{P(Y = 25)} = \frac{0.2}{0.3} = \frac{2}{3} = 0.67。$$

(11)  $X=3$  を条件としたときの  $Y$  の条件付期待値  $E(Y | X = 5)$  を求めなさい。[ 4 点 ]

$$\begin{aligned} E(Y | X = 5) &= 10 \times P(Y = 10 | X = 5) + 20 \times P(Y = 20 | X = 5) + 25 \times P(Y = 25 | X = 5) \\ &= 10 \times \frac{0.1}{0.4} + 20 \times \frac{0.2}{0.4} + 25 \times \frac{0.1}{0.4} = \frac{7.5}{0.4} = 18.75 \end{aligned}$$

## 成績分布など

	問題 (A)	問題 (B)
受験者数	146 名	121 名
( 受験率 )	97.3%	81.8%
平均点	32.62	31.66
標準偏差	9.10	8.63
問題別正答率		
(1)	97.9%	100.0%
(2)	68.8%	64.5%
(3)	89.0%	78.5%
(4)	29.8%	35.1%
(5)	81.5%	81.8%
(6)	14.0%	26.9%
(7)	52.7%	27.3%
(8)	25.0%	9.1%
(9)	(a)	100.0%
	(b)	82.9%
	(c)	100.0%
	(d)	83.6%
	(e)	83.6%
	(f)	83.6%
(10)	73.6%	75.6%
(11)	57.2%	61.2%

問題 (B) は、学年によって成績に差があり、

	2 年生	3 年生	4 年生その他
平均点	32.62	28.42	34.03
( 標準偏差 )	(8.63)	(8.43)	(7.66)

となっている。

### 【講評】

最近の学生の傾向なのかもしれないが、問題文をよく読まない学生が散見される。Q1 の(1)～(2)と(3)では求めている変数が異なるにもかかわらず、気がつかない者が少なくなかった。また問題 (A) の Q3 の(9)では、求めているのが確率であるにもかかわらず、負の値を答えている者がいた。( 確率は定義より[0,1]の実数値をとる )

【成績分布】

