

第3講

問1

(1) 平均：79，中央値：85.5，最頻値：87

最も好ましいもの：中央値

(平均は最小値に引きずられていて中心から低くなっており，最頻値は2回しかない87を選んでいて適切とはいえない。)

(2) 分散：422.6，標準偏差：20.56，変動係数：0.26 (26%)

(3) 最大値：100，最小値：23，範囲（最大値-最小値）：77，四分位点範囲：15

(4) Zスコアによる方法：外れ値なし（どのZスコアも3より小さい）

箱ひげ図による方法：外れ値あり（最小値である4番目の観測値23）

（箱ひげ図の場合，ひげの長さを分位点偏差の1.5倍とすると，分位点偏差が15なのでその1.5倍の22.5の長さを第一四分位の下側と第三四分位の上側にとることで，下限が51.75で上限が111.75となる箱ひげ図を書くことができる。最小値23は下限より小さいため。）

問2

(1) 平均：1.18，中央値：1，最頻値：1

(2) 分散：1.148，標準偏差：1.071，変動係数：0.908 (90.8%)

(3) 四分位点範囲：2

問3

(1) 共分散：28.14

(2) 相関係数：0.91

最高気温とアイスクリームの売上個数との間には強い正の相関が存在する。

第4講

問1

(1) $A \cap B = \{E_3, E_6\}$

(2) $A \cup B = \{E_1, E_3, E_4, E_6, E_9, E_{10}\}$

(3) $A^c \cap B = \{E_4, E_{10}\}$

(4) $(A \cap B) \cup (A^c \cap B) = \{E_3, E_4, E_6, E_{10}\}$

(5) $A \cup (A^c \cap B) = \{E_1, E_3, E_4, E_6, E_9, E_{10}\}$

問 2

- (1) ①
(2) ①

問 3

(1) 確率の公理 3 を用いると, $A \cap A^c = \emptyset$ から $P(A \cup A^c) = P(A) + P(A^c)$

確率の公理 2 を用いると, $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$

よって, $P(A^c) = 1 - P(A)$

(2) $(A \cap B) \cap (A^c \cap B) = \emptyset$ であるから, 確率の公理 3 を用いると

$$P((A \cap B) \cup (A^c \cap B)) = P(A \cap B) + P(A^c \cap B)$$

また, 集合の性質から

$$B = (A \cap B) \cup (A^c \cap B)$$

よって

$$P(B) = P(A \cap B) + P(A^c \cap B)$$

(3) (2) の結果を用いると

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

$$P(B) = P(A \cap B) + P(A^c \cap B)$$

よって

$$P(A) + P(B) = P(A \cap B) + P(A \cap B) + P(A \cap B^c) + P(A^c \cap B)$$

また $(A \cap B)$, $(A \cap B^c)$, $(A^c \cap B)$ はそれぞれ排反なので, 上式の右辺の最後の 3 項は

$$P(A \cap B) + P(A \cap B^c) + P(A^c \cap B) = P(A \cup B)$$

つまり

$$P(A) + P(B) = P(A \cap B) + P(A \cup B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

となる。

問 4

$$(1) P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.4} = 0.75$$

$$(2) P(A \cap B) = P(A|B)P(B) = 0.60 \times 0.80 = 0.48$$

問 5

$$(1) A = \{1, 3, 5\}, B = \{1, 2, 3, 4\}, A \cap B = \{1, 3\}$$

$$P(A \cap B) = 1/3$$

$$P(A)P(B) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$$

より, $P(A \cap B) = P(A)P(B)$ なので, 事象 A と事象 B は独立である。

$$(2) A = \{1, 3, 5\}, B = \{1, 2, 3\}, A \cap B = \{1, 3\}$$

$$P(A \cap B) = 1/3$$

$$P(A)P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

より, $P(A \cap B) \neq P(A)P(B)$ なので, 事象 A と事象 B は独立でない。

(3) $A = \{3,4,5,6\}$, $B = \{1,2\}$, $A \cap B = \emptyset$

$$P(A \cap B) = 0$$

$$P(A)P(B) = \frac{2}{3} \times \frac{1}{3} = \frac{2}{9}$$

より, $P(A \cap B) \neq P(A)P(B)$ なので, 事象 A と事象 B は独立でない。

問 6

(1) ベイズの定理より

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.6 \times 0.4}{0.6 \times 0.4 + 0.7 \times 0.6} = \frac{4}{11} \end{aligned}$$

(2) ベイズの定理より

$$P(A|B^c) = \frac{P(B^c|A)P(A)}{P(B^c|A)P(A) + P(B^c|A^c)P(A^c)}$$

ここで, $P(A) = 0.5$ から $P(A^c) = 0.5$ 。さらに

$$P(B^c|A) = 1 - P(B|A) = 0.60$$

$$P(B^c|A^c) = 1 - P(B|A^c) = 0.40$$

だから

$$P(A|B^c) = \frac{0.60 \times 0.5}{0.60 \times 0.5 + 0.40 \times 0.5} = 0.6$$

問 7

「資格試験に合格」をA, 「資格試験に不合格」を A^c とする。「模擬試験で合格判定」をB, 「模擬試験で不合格判定」を B^c とすると, 題意より $P(A) = 0.3$, $P(B|A) = 0.6$, $P(B|A^c) = 0.1$ である。ベイズの定理を用いると, 求める確率は

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.6 \times 0.3}{0.6 \times 0.3 + 0.1 \times 0.7} = 0.72 \end{aligned}$$

第 5 講

問 1

- (1) 離散確率変数
- (2) 離散確率変数
- (3) 離散確率変数
- (4) 連続確率変数
- (5) 連続確率変数

問 2

- (1) 略
- (2) 略
- (3) $E(X) = 1$
- (4) $V(X) = 0.5$

問 3

- (1) $P(X = 0) = 0.55, P(X = 1) = 0.45$
 $P(Y = 0) = 0.50, P(Y = 1) = 0.50$
- (2) $E(Y) = 0.5, V(Y) = 0.25, Sd(Y) = 0.5$
- (3) $P(Y = 0|X = 1) = \frac{P(X=1, Y=0)}{P(X=1)} = \frac{4}{9}$
 $P(Y = 1|X = 1) = \frac{P(X=1, Y=1)}{P(X=1)} = \frac{5}{9}$
- (4) 確率変数 X と Y は独立ではない。

問 4

- (1) $E(a + bX) = \sum_{i=1}^m (a + bx_i)P(x_i) = a \sum_{i=1}^m P(x_i) + b \sum_{i=1}^m x_i P(x_i) = a + bE(X)$
- (2) $V(a + bX) = E[\{(a + bX) - E(a + bX)\}^2]$
 $= E[\{bX - bE(X)\}^2]$
 $= b^2 E[\{X - E(X)\}^2]$
 $= b^2 V(X)$
- (3) $V(X) = E[\{X - E(X)\}^2] = E[X^2 - 2XE(X) + E(X)^2]$
 $= E(X^2) - 2E(X)E(X) + E(X)^2$
 $= E(X^2) - E(X)^2$
- (4) $E(a + bX + cY) = a \sum_{i=1}^m P(x_i) + b \sum_{i=1}^m x_i P(x_i) + c \sum_{j=1}^n y_j P(y_j) = a + bE(X) + cE(Y)$
- (5) $V(a + bX + cY) = E[\{(a + bX + cY) - E(a + bX + cY)\}^2]$
 $= E[\{b(X - E(X)) + c(Y - E(Y))\}^2]$
 $= E[b^2\{X - E(X)\}^2 + c^2\{Y - E(Y)\}^2 + 2bc\{X - E(X)\}\{Y - E(Y)\}]$
 $= b^2 E[\{X - E(X)\}^2] + c^2 E[\{Y - E(Y)\}^2] + 2bc E[\{X - E(X)\}\{Y - E(Y)\}]$
 $= b^2 V(X) + c^2 V(Y) + 2bc Cov(X, Y)$

問 5

- (1) 14.4 万円
- (2) 84.24 万円

問 6

- (1) $P(X = 7) = 0.062$
- (2) $P(X \leq 7) = 0.093$
- (3) $P(\geq 8) = 1 - P(X \leq 7) = 0.907$

問 7

- (1) $P(X = 0) + P(X = 1) = 0.074 + 0.193 = 0.267$
- (2) $P(X > 10) = 1 - P(X \leq 10) = 1 - 0.533 = 0.467$

第 6 講

問 1

※この問題を解くためには，次の基本的な積分を用います。

$$\int_a^b c dx = [cx]_a^b = cb - ca$$
$$\int_a^b x dx = \left[\frac{1}{2}x^2\right]_a^b = \frac{1}{2}b^2 - \frac{1}{2}a^2$$

- (1) $P(1.4 \leq X \leq 1.8) = 0.5 \times (1.8 - 1.4) = 0.2$
- (2) $P(X \leq 1.4) = 0.5 \times 1.4 = 0.7$
- (3) $E(X) = \int_0^2 0.5x dx = \left[\frac{1}{4}x^2\right]_0^2 = 1 - 0 = 1$
- (4) $V(X) = \int_0^2 0.5(x - 1)^2 dx = 0.5 \times \left[\frac{1}{3}x^2 - x^2 + x\right]_0^2 = \frac{1}{3}$

問 2

ある図書館で利用者がサービスカウンターを利用する時間間隔は指数分布に従うと考えることが出来る。問題文より，時間間隔を 5 分とすると $\lambda = 1$ で，10 分以上利用者がいない確率は， $1 -$ ”10 分以内に利用者がある確率” で求められる。つまり

$$1 - F(2) = \exp(-2) = 0.135$$

問 3

小数点第 4 位以下は四捨五入する。

- (1) 0.885
- (2) 0.841
- (3) 0.726
- (4) 0.524
- (5) 0.842
- (6) 1.282

問 4

- (1) 0.106
- (2) 0.903
- (3) 0.734
- (4) 56.733
- (5) 10.252

問 5

大学の入学者数を確率変数 X とすると、2 項分布の平均と分散は

$$E(X) = 1100 \times 0.7 = 770$$

$$V(X) = 1100 \times 0.7 \times 0.3 = 231$$

よって、正規分布近似を用いると

$$P(X > 800) = P\left(\frac{X - 770}{\sqrt{231}} > \frac{800 - 770}{\sqrt{231}}\right) = P(Z > 1.973) = 0.024$$

問 6

- (1) 0.871
- (2) 0.701
- (3) 0.542
- (4) 1.372

問 7

- (1) 0.647
- (2) 0.535
- (3) 15.507
- (4) 2.733
- (5) 自由度が 20 の場合 (3) の解答は $x = 31.41$ 、(4) の解答は $x = 10.85$ と、より大きな値をとる。

問 8

- (1) 0.182
- (2) 0.445
- (3) 2.913
- (4) 4.640

第 7 講

問 1

- (1) $E(\bar{X}) = \mu = 100$, $V(\bar{X}) = \frac{\sigma^2}{n} = \frac{9}{4}$
- (2) $P(\bar{X} \geq 102) = P\left(Z > \frac{4}{3}\right) = 0.0912$
- (3) $P(98 \leq \bar{X} \leq 101) = P\left(-\frac{4}{3} \leq Z \leq \frac{2}{3}\right) = 0.656$
- (4) $P(\bar{X} \leq 101.5) = P(Z \leq 1) = 0.8413$

問 2

$$(1) E(\bar{X}) = \mu = 100, V(\bar{X}) = \frac{\sigma^2}{n} = 1$$

$$(2) P(\bar{X} \geq 102) = P(Z \geq 2) = 0.02275$$

$$(3) P(98 \leq \bar{X} \leq 101) = P(-2 \leq Z \leq 1) = 0.8186$$

$$(4) P(\bar{X} \leq 101.5) = P(Z \leq 1.5) = 0.933$$

$n = 36$ のときと比べ、 $n = 81$ の方が抽出分布の分散が小さくなり、母平均周辺の値をとりやすくなっている。

問 3

$$(1) \bar{X} \sim N(500, \frac{125^2}{25}) \text{だから}$$

$$\begin{aligned} P(\bar{X} > 550) &= P\left(Z > \frac{550 - 500}{\frac{125}{5}}\right) \\ &= P(Z > 2.0) = 0.023 \end{aligned}$$

(2)

$$\begin{aligned} P(475 \leq \bar{X} \leq 540) &= P\left(\frac{475 - 500}{25} \leq Z \leq \frac{540 - 500}{25}\right) \\ &= P(-1.0 \leq Z \leq 1.6) = 0.787 \end{aligned}$$

問 4

(1) $np(1-p) = 100 \times 0.4 \times (1-0.4) = 24 > 9$ なので、標本割合が正規分布で近似できるほど標本数が大きいといえる。

(2) 標本割合の抽出分布は

$$\hat{p} \sim N(0.4, 0.0024)$$

なので

$$P(\hat{p} > 0.45) = P\left(Z > \frac{0.45 - 0.4}{\sqrt{0.0024}}\right) = P(Z > 1.02) = 0.1537$$

(3)

$$P(\hat{p} < 0.3) = P\left(Z < \frac{0.3 - 0.4}{\sqrt{0.0024}}\right) = P(Z < -2.04) = 0.02$$

(4) $P(\hat{p} > 0.45) \leq 0.05$ となる最小の n を求める。

$$P\left(Z > \frac{0.45 - 0.4}{\sqrt{\frac{0.4 \times 0.6}{n}}} = 1.645\right) = 0.05$$

を解いて $n = 259.73$ なので、 n は最低でも 260 にする必要がある。

問 5

(1) $np(1-p) = 300 \times 0.28 \times 0.72 = 60.48 > 9$ より、標本割合は正規分布で近似できるため

$$\hat{p} \sim N(0.28, 0.000672)$$

$$(2) P(\hat{p} > 0.31) = P\left(Z > \frac{0.31-0.28}{\sqrt{\frac{0.28 \times 0.72}{300}}}\right) = P(Z > 1.157) = 0.1236$$

問 6

$$(1) P\left(\frac{14s^2}{\sigma^2} > \frac{14 \times 30^2}{25^2}\right) = P(\chi_{14}^2 > 20.16) = 0.125$$

$$(2) P\left(\frac{29s^2}{\sigma^2} < \frac{29 \times 20^2}{25^2}\right) = P(\chi_{29}^2 < 18.56) = 0.068$$

問 7

$$(1) E(\bar{X}) = \frac{1}{3}E(X_1) + \frac{1}{3}E(X_2) + \frac{1}{3}E(X_3) = \mu \text{ となるので不偏性を満たす}$$

$$E(\tilde{X}) = \frac{1}{2}\mu + \frac{1}{4}\mu + \frac{1}{8}\mu = \frac{7}{8}\mu \text{ となるので不偏性を満たさない。}$$

$$E(\ddot{X}) = \frac{1}{2}\mu + \frac{1}{2}\mu = \mu \text{ となるので不偏性を満たす。}$$

$$(2) V(\bar{X}) = \frac{1}{9}\sigma^2 = \frac{1}{3}$$

$$V(\tilde{X}) = \frac{1}{2}\sigma^2 = \frac{1}{2}$$

より、不偏推定量の中で最も効率的な推定量は $\bar{X}$ である。

(3) 平均平方誤差を MSE と書くと

$$MSE(\bar{X}) = \frac{1}{3}$$

$$MSE(\tilde{X}) = \text{偏り}^2 + \text{分散} = \frac{85}{256}$$

$$MSE(\ddot{X}) = \frac{1}{2}$$

$MSE(\tilde{X}) < MSE(\bar{X}) < MSE(\ddot{X})$ より、平均平方誤差の意味で最も望ましい推定量は $\tilde{X}$ である。

第 8 講

問 1

$$(1) \bar{X} \pm z_{0.05} \frac{\sigma}{\sqrt{n}} = 50 \pm 1.64 \frac{4}{\sqrt{64}} = [49.18, 50.82]$$

$$(2) \bar{X} \pm z_{0.025} \frac{\sigma}{\sqrt{n}} = 50 \pm 1.96 \frac{4}{\sqrt{64}} = [49.02, 50.98]$$

$$(3) \bar{X} \pm z_{0.025} \frac{\sigma}{\sqrt{n}} = 50 \pm 1.96 \frac{4}{\sqrt{225}} = [49.48, 50.52]$$

問 2

$$(1) \bar{X} \pm t_{15,0.025} \frac{s}{\sqrt{n}} = 50 \pm 2.13 \frac{4}{\sqrt{16}} = [47.87, 52.13]$$

$$(2) \bar{X} \pm t_{63,0.025} \frac{s}{\sqrt{n}} = 50 \pm 2.00 \frac{4}{\sqrt{64}} = [49.00, 51.00]$$

$$(3) \bar{X} \pm t_{63,0.025} \frac{\sigma}{\sqrt{n}} = 50 \pm 2.00 \frac{8}{\sqrt{64}} = [48.00, 52.00]$$

(4) 標本数 $n \geq 30$ を中心極限定理が近似できる目安とすると, (2) と (3) は母集団の分布が正規分布でなくても概ね結果は正確であるといえる。

問 3

$$(1) 686.375$$

(2) $\bar{X} = 686.375$, $s = 166.40$ より, 95%信頼区間は

$$\bar{X} \pm t_{n-1,\alpha/2} \frac{s}{\sqrt{n}} = 686.375 \pm 2.13 \frac{166.40}{4} = [597.71, 775.04]$$

となり, 598 人以上 775 人以下である。

問 4

$$(1) \bar{X} \pm t_{35,0.05} \frac{s}{\sqrt{n}} = 46 \pm 2.03 \frac{20}{\sqrt{36}} = [39.23, 52.77]$$

となり, 40 個以上 52 個以下である。

(2) 誤差範囲は $t_{63,0.05} \frac{s}{\sqrt{n}} = 2.00 \frac{20}{\sqrt{64}} < 2.03 \frac{20}{\sqrt{36}}$ なので, (1) の区間よりも小さくなる。

問 5

$$(1) \hat{p} \pm z_{0.025} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3 \pm 1.96 \sqrt{\frac{0.3 \times 0.7}{250}} = [0.243, 0.357]$$

$$(2) \hat{p} \pm z_{0.05} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.45 \pm 1.64 \sqrt{\frac{0.45 \times 0.55}{175}} = [0.388, 0.512]$$

$$(3) \hat{p} \pm z_{0.005} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05 \pm 2.58 \sqrt{\frac{0.05 \times 0.95}{400}} = [0.022, 0.078]$$

$$(4) \hat{p} \pm z_{0.0125} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.9 \pm 2.24 \sqrt{\frac{0.9 \times 0.1}{1000}} = [0.879, 0.921]$$

問 6

$$(1) \frac{32}{200} = 0.16$$

$$(2) \text{割合の 95\%信頼区間は } 0.16 \pm 1.96 \sqrt{\frac{0.16 \times 0.84}{200}} = [0.109, 0.211]$$

(3) 売上利益の 95%信頼区間は, $[0.109 \times 200 \times 50, 0.211 \times 200 \times 50] = [1091.9, 2108.1]$ となるので, 1092 円以上 2108 円以下。

問 7

$$(1) \left[\frac{(12-1)s^2}{\chi_{12-1,0.025}^2}, \frac{(12-1)s^2}{\chi_{12-1,0.975}^2} \right] = \left[\frac{11 \times 16}{21.92}, \frac{11 \times 16}{3.82} \right] = [8.029, 46.073]$$

$$(2) \left[\frac{(25-1)s^2}{\chi_{25-1,0.05}^2}, \frac{(25-1)s^2}{\chi_{25-1,0.95}^2} \right] = \left[\frac{24 \times 0.8}{36.42}, \frac{24 \times 0.8}{13.85} \right] = [0.53, 1.386]$$

$$(3) \left[\frac{(44-1)s^2}{\chi_{44-1,0.005}^2}, \frac{(44-1)s^2}{\chi_{44-1,0.995}^2} \right] = \left[\frac{43 \times 49}{70.62}, \frac{43 \times 49}{22.86} \right] = [29.84, 97.17]$$

問 8

$$(1) P(\sigma^2 \geq 16) = P\left(\frac{(17-1)s^2}{\sigma^2} \leq \frac{(17-1)s^2}{16}\right) = P\left(\chi_{17-1}^2 \leq \frac{16 \times 7.96}{16}\right) = P(\chi_{16}^2 \leq 7.96) = 0.05$$

$$(2) \left[\frac{16 \times 7.96}{28.85}, \frac{16 \times 7.96}{6.908} \right] = [4.42, 18.44]$$

第 9 講

問 1

$$(1) (\bar{X} - \bar{Y}) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} = 70 - 60 \pm 1.96 \sqrt{\frac{25}{36} + \frac{9}{100}} = [8.26, 11.74]$$

$$(2) (\bar{X} - \bar{Y}) \pm t_{n_X+n_Y-2, \frac{\alpha}{2}} \sqrt{\frac{s^2}{n_X} + \frac{s^2}{n_Y}} = 70 - 60 \pm 1.978 \sqrt{\frac{8.96}{36} + \frac{8.96}{100}} = [8.85, 11.15]$$

$$(3) (\bar{X} - \bar{Y}) \pm t_{\nu, \frac{\alpha}{2}} \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}} = 70 - 60 \pm 1.99 \sqrt{\frac{6}{36} + \frac{10}{100}} = [8.97, 11.03]$$

問 2

$$(\bar{X} - \bar{Y}) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} = 130 - 120 \pm 1.96 \sqrt{\frac{8.4}{40} + \frac{11.3}{36}} = [8.973, 11.027]$$

問 3

$$9.82 - 3.22 \pm 1.96 \sqrt{\frac{12.33}{9} + \frac{18.01}{12}} = [0.973, 12.227]$$

問 4

$$(1) (\hat{p}_X - \hat{p}_Y) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}}$$

$$= 0.37 - 0.40 \pm 1.96 \sqrt{\frac{0.37 \times 0.63}{300} + \frac{0.4 \times 0.6}{380}} = [-0.103, 0.044]$$

$$\begin{aligned}
 (2) (\hat{p}_X - \hat{p}_Y) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}} \\
 = 0.49 - 0.52 \pm 1.96 \sqrt{\frac{0.49 \times 0.51}{200} + \frac{0.52 \times 0.48}{220}} = [-0.1257, 0.0657]
 \end{aligned}$$

$$\begin{aligned}
 (3) (\hat{p}_X - \hat{p}_Y) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}} \\
 = 0.05 - 0.08 \pm 1.96 \sqrt{\frac{0.05 \times 0.95}{622} + \frac{0.08 \times 0.92}{488}} = [-0.05954, -0.00046]
 \end{aligned}$$

問 5

日本の大学生を母集団X, 韓国の大学生を母集団Yとすると,

$$\hat{p}_X = \frac{218}{360} = 0.61$$

$$\hat{p}_Y = \frac{98}{220} = 0.45$$

95%信頼区間は

$$\begin{aligned}
 (\hat{p}_X - \hat{p}_Y) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}} \\
 = (0.61 - 0.45) \pm 1.96 \sqrt{\frac{0.61 \times 0.39}{360} + \frac{0.45 \times 0.55}{220}} \\
 = [0.077, 0.243]
 \end{aligned}$$

第 10 講

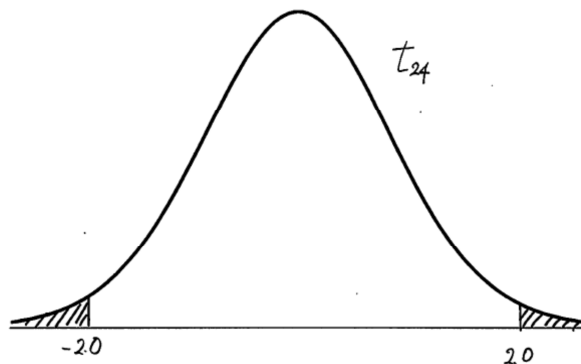
問 1

$$(1) t \leq -t_{24,0.025} = -2.06, t \geq t_{24,0.025} = 2.06$$

(2) 検定統計量 $t = \frac{106-100}{15/\sqrt{25}} = \frac{6}{3} = 2.0$ は棄却域に含まれないため, 有意水準 5% で帰無仮説は

棄却されない。

(3) p 値は下図の斜線部分で $P(t_{24} \geq 2.0) \times 2 = 0.057$ となる。



問 2

(1) $t \geq t_{24,0.05} = 1.71$

(2) 検定統計量 $t = \frac{106-100}{15/\sqrt{25}} = \frac{6}{3} = 2.0$ は棄却域に含まれるため、有意水準 5% で帰無仮説は棄却される。

(3) p 値は下図の斜線部分で $P(t_{24} \geq 2.0) = 0.028$ となる。

問 3

仮説は

$$H_0: \mu \leq 4$$

$$H_1: \mu > 4$$

検定統計量は

$$t = \frac{4.3 - 4.0}{0.7/\sqrt{120}} = 4.69$$

棄却域は $t \geq t_{119,\alpha} = 1.66$ なので、帰無仮説は有意水準 5% で棄却され、満足度は 4.0 を上回る。なお、 $n \geq 30$ を中心極限定理が近似できる目安とすると、標本数 $n = 120$ は帰無仮説の下での検定統計量を近似するのに十分に大きいため、結果は信頼できる。

問 4

(1) 検定統計量は $t = \frac{55-60}{20/\sqrt{20}} = -1.12$ となり、 p 値は

$$P(t_{19} \leq -1.12) = 0.139$$

(2) 検定統計量は $t = \frac{55-60}{20/\sqrt{50}} = -1.77$ となり、 p 値は

$$P(t_{49} \leq -1.12) = 0.047$$

(3) 検定統計量は $t = \frac{55-60}{20/\sqrt{100}} = -2.50$ となり、 p 値は

$$P(t_{99} \leq -2.50) = 0.011$$

問 5

(1) 検定統計量は $t = \frac{55-60}{50/6} = -0.6$ となり、 p 値は

$$P(t_{35} \leq -0.6) \times 2 = 0.555$$

(2) 検定統計量は $t = \frac{55-60}{20/6} = -1.5$ となり、 p 値は

$$P(t_{35} \leq -1.5) \times 2 = 0.150$$

(3) 検定統計量は $t = \frac{55-60}{12/6} = -2.50$ となり、 p 値は

$$P(t_{35} \leq -2.50) \times 2 = 0.022$$

問 6

仮説は

$$H_0: \mu = 12$$

$$H_1: \mu \neq 12$$

標本平均および標本標準偏差を計算すると、 $\bar{X} = 12.244$ および $s = 0.705$ なので、 t 検定統計量は

$$t = \frac{12.244 - 12}{0.705/\sqrt{9}} = 1.039$$

棄却域は $t \leq -t_{8,0.025} = -2.306$, $t \geq t_{8,0.025} = 2.306$ なので、有意水準 5%で帰無仮説は棄却されず、バッテリーの平均消耗時間は 12 時間と異なる。

第 11 講

問 1

(1) $z = \frac{\hat{p} - 0.25}{\sqrt{\frac{0.25 \times 0.75}{100}}} < -z_{0.05} = -1.64$ を、 $\hat{p}$ について解くと

$$\hat{p} < -1.64 \times \sqrt{\frac{0.25 \times 0.75}{100}} + 0.25 = 0.179$$

(2) 同様に

$$\hat{p} < -1.64 \times \sqrt{\frac{0.25 \times 0.75}{400}} + 0.25 = 0.214$$

(3) 同様に

$$\hat{p} < -1.64 \times \sqrt{\frac{0.25 \times 0.75}{800}} + 0.25 = 0.225$$

問 2

(1) 棄却域は

$$\chi^2 \leq \chi_{15,0.975}^2 = 6.26, \quad \chi^2 \geq \chi_{15,0.025}^2 = 27.49$$

(2) 検定統計量は

$$\chi^2 = \frac{15 \times 180}{100} = 27.0$$

なので、有意水準 5%で帰無仮説は棄却されず、母分散は 100 と異なる。

(3) p 値は下図の斜線領域で、 $P(\chi_{15}^2 \geq 27.0) \times 2 = 0.057$

(4) 棄却域は

$$\chi^2 \geq \chi_{15,0.05}^2 = 24.995$$

(5) 検定統計量は(2)と同じなので、有意水準 5%で帰無仮説は棄却され、母分散は 100 と異なる。

(6) p 値は下図の斜線領域で、 $P(\chi_{15}^2 \geq 27.0) = 0.029$

問 3

(1) 検定統計量は

$$t = \frac{0.92 - 0.88}{\sqrt{\frac{0.36^2}{600} + \frac{0.36^2}{562}}} = 1.89$$

p 値は $P(t_{600+562-2} > 1.89) \times 2 = 0.059$ なので、有意水準 5%で帰無仮説は棄却されず、両国の平均転職回数は異なる。

(2) 検定統計量は

$$t = \frac{0.92 - 0.88}{\sqrt{\frac{0.4^2}{600} + \frac{0.3^2}{562}}} = 1.94$$

自由度は

$$v = \frac{\left(\frac{0.4^2}{600} + \frac{0.3^2}{562}\right)^2}{\frac{\left(\frac{0.4^2}{600}\right)^2}{599} + \frac{\left(\frac{0.3^2}{562}\right)^2}{561}} = 1107.86$$

p 値は $P(t_v > 1.94) \times 2 = 0.053$ なので、有意水準 5%で帰無仮説は棄却されず、両国の平均転職回数は異なる。

(3) $n \geq 30$ を中心極限定理が近似できる目安とすると、正確な結果を得るために十分な標本数があると考えられる。

問 4

(1) 検定統計量は

$$z = \frac{0.68 - 0.58}{\sqrt{\frac{0.68 \times 0.32}{200} + \frac{0.58 \times 0.42}{300}}} = 2.29$$

p 値は $P(Z > 2.29) \times 2 = 0.022 < 0.05$ となり、有意水準 5%で帰無仮説は棄却される。

(2) 検定統計量は

$$z = \frac{0.22 - 0.20}{\sqrt{\frac{0.22 \times 0.78}{500} + \frac{0.20 \times 0.80}{800}}} = 0.858$$

p 値は $P(Z > 0.858) \times 2 = 0.39$ となり、有意水準 5%で帰無仮説は棄却されない。

問 5

(1) 検定統計量は

$$F = \frac{s_X^2}{s_Y^2} = \frac{66}{40} = 1.65$$

p 値は $P(F_{23,35} > 1.65) \times 2 = 0.177$ となり、有意水準 5%で帰無仮説は棄却されない。

(2) 検定統計量は

$$F = \frac{s_X^2}{s_Y^2} = \frac{32}{12} = 2.67$$

p 値は $P(F_{7,119} > 2.67) \times 2 = 0.026$ となり、有意水準 5%で帰無仮説は棄却される。

第 12 講

問 1

(1) 検定統計量は

$$F = \frac{488/6}{226/43} = 15.47$$

p 値は $P(F_{6,43} \geq 15.47) = 0.000$ だから、有意水準 5%で帰無仮説は棄却され、グループ母平均は全て等しいわけではないと結論される。

(2) 検定統計量は

$$F = \frac{27.5/3}{18.8/16} = 7.80$$

p 値は $P(F_{3,16} \geq 7.80) = 0.002$ だから、有意水準 5%で帰無仮説は棄却され、グループ母平均は全て等しいわけではないと結論される。

問 2

(1) 全変動は 1831.2、グループ間変動は 594、グループ内変動は 1237.2 である。

(2)

	変動	自由度
グループ間	594.0	3
グループ内	1237.2	16
合計	1831.2	19

(3) 検定統計量は

$$F = \frac{594.0/3}{1237.2/16} = 2.56$$

p 値は $P(F_{3,16} \geq 2.56) = 0.09$ だから、有意水準 5%で帰無仮説は棄却されず、曜日による平均客数の違いはないと結論される。

問 3

(1) 全変動は 204.93, グループ間変動は 34.15, グループ内変動は 170.78 である。

(2)

	変動	自由度
グループ間	34.15	2
グループ内	170.78	12
合計	204.93	14

問 4

(1) グループ標本平均を推定量とする。 $\hat{\mu}_1 = 4.25$, $\hat{\mu}_2 = 7.83$, $\hat{\mu}_3 = 7.4$ 。

(2) 4 人の「誤差項の実現値」は,

$$6 - 4.25 = 1.75$$

$$0 - 4.25 = -4.25$$

$$4 - 4.25 = -0.25$$

$$7 - 4.25 = 2.75$$

(3) 検定統計量

$$F = \frac{34.15/2}{170.78/12} = 1.20$$

p 値は $P(F_{2,12} \geq 1.20) = 0.33$ だから, 有意水準 5% で帰無仮説は棄却されず, 職種による平均有給休暇取得日数の違いはないと結論される。

第 13 講

問1

(1) 9

(2) -3.6

(3) 2.3

(4) 902.3

問 2

(1)

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{60}{300} = 0.2$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 10 - 0.2 \times 30 = 4$$

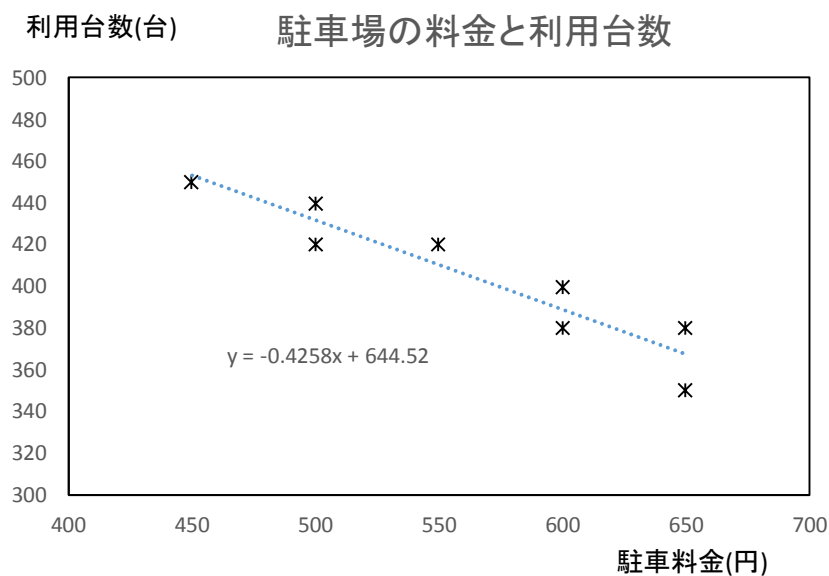
(2) $\hat{\beta}_1$: 客数 1 人の増加によるワイン注文本数増加分の期待値。

$\hat{\beta}_0$: 客数が 0 人の日の平均ワイン注文本数。なお, 客数が 0 人であればワインの注文本数は 0 本であるから, 切片は意味を持たない。

(3) $\hat{Y}_i = 4 + 0.2 \times 50 = 5$ (本)

問 3

(1)



(2) 駐車料金を 100 円上げると、利用台数は期待値で 43 台減少する。

問 4

(1) $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ を用いると

$$\begin{aligned}\hat{U}_i &= Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \\ &= Y_i - (\bar{Y} - \hat{\beta}_1 \bar{X}) - \hat{\beta}_1 X_i \\ &= (Y_i - \bar{Y}) - \hat{\beta}_1 (X_i - \bar{X})\end{aligned}$$

(2)

$$\sum_{i=1}^n \hat{U}_i = \sum_{i=1}^n (Y_i - \bar{Y}) - \hat{\beta}_1 \sum_{i=1}^n (X_i - \bar{X})$$

ここで

$$\begin{aligned}\sum_{i=1}^n (Y_i - \bar{Y}) &= \sum_{i=1}^n Y_i - \sum_{i=1}^n \bar{Y} \\ &= n\bar{Y} - n\bar{Y} = 0\end{aligned}$$

同様に

$$\sum_{i=1}^n (X_i - \bar{X}) = \sum_{i=1}^n X_i - \sum_{i=1}^n \bar{X} = 0$$

よって

$$\sum_{i=1}^n \hat{U}_i = 0$$

(3) (1)の結果を用いると

$$\hat{U}_i^2 = (Y_i - \bar{Y})^2 + \hat{\beta}_1^2 (X_i - \bar{X})^2 - 2\hat{\beta}_1 (X_i - \bar{X})(Y_i - \bar{Y})$$

総和をとると

$$\sum_{i=1}^n \hat{U}_i^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2 + \hat{\beta}_1^2 \sum_{i=1}^n (X_i - \bar{X})^2 - 2\hat{\beta}_1 \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

ここで $\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$ を上の式に代入すると、右辺の第3項目は

$$-2\hat{\beta}_1^2 \sum_{i=1}^n (X_i - \bar{X})^2$$

となるので

$$\sum_{i=1}^n \hat{U}_i^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2 - \hat{\beta}_1^2 \sum_{i=1}^n (X_i - \bar{X})^2$$

(4) $\hat{Y}_i$ の定義から

$$\hat{Y}_i - \bar{Y} = Y_i - \hat{U}_i - \bar{Y}$$

これに(1)の結果を代入して

$$\begin{aligned} \hat{Y}_i - \bar{Y} &= Y_i - (Y_i - \bar{Y}) + \hat{\beta}_1 (X_i - \bar{X}) - \bar{Y} \\ &= \hat{\beta}_1 (X_i - \bar{X}) \end{aligned}$$

(5) (4)の結果を(3)の結果に代入すると

$$\sum_{i=1}^n \hat{U}_i^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2 - \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

つまり

$$\text{残差二乗和} = \text{全変動} - \text{回帰変動}$$

よって

$$\text{全変動} = \text{回帰変動} + \text{残差二乗和}$$

が成立する。

(6) (1)の結果を用いると

$$\hat{U}_i (X_i - \bar{X}) = (X_i - \bar{X})(Y_i - \bar{Y}) - \hat{\beta}_1 (X_i - \bar{X})^2$$

両辺の総和をとると

$$\sum_{i=1}^n \hat{U}_i (X_i - \bar{X}) = \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) - \hat{\beta}_1 \sum_{i=1}^n (X_i - \bar{X})^2$$

ここで $\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$ を上の式に代入すると、

$$\sum_{i=1}^n \hat{U}_i (X_i - \bar{X}) = 0$$

(7) (4)の両辺を2乗して総和をとると

$$\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = \hat{\beta}_1^2 \sum_{i=1}^n (X_i - \bar{X})^2$$

ここで

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = \hat{\beta}_1^2 \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

(8) (7)の結果を用いると

$$\begin{aligned} R^2 &= \hat{\beta}_1^2 \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{[\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})]^2}{[\sum_{i=1}^n (X_i - \bar{X})^2][\sum_{i=1}^n (Y_i - \bar{Y})^2]} \times \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\ &= \frac{[\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})]^2}{[\sum_{i=1}^n (X_i - \bar{X})^2][\sum_{i=1}^n (Y_i - \bar{Y})^2]} \end{aligned}$$

$$= \left(\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \right)^2 = \rho_{XY}^2$$

第 14 講

問 1

$$(1) \hat{\beta}_1 \pm t_{48-2, 0.025} s_{\hat{\beta}_1} = 6.9 \pm 2.01 \times 4.6 = [-2.36, 16.16]$$

$$(2) \hat{\beta}_1 \pm t_{30-2, 0.025} s_{\hat{\beta}_1} = 12.8 \pm 2.05 \times 4.2 = [4.21, 21.39]$$

$$(3) \hat{\beta}_1 \pm t_{120-2, 0.025} s_{\hat{\beta}_1} = 0.03 \pm 1.98 \times 0.01 = [0.01, 0.05]$$

問 2

係数推定量は正規分布に従うと考えられるため，信頼区間は小さくなる。具体的には問 1 での信頼区間の大きさは (1) 18.52，(2) 17.18，(3) 0.0396。誤差項の分散が既知であれば，信頼区間の大きさはそれぞれ

$$(1) 2 \times (1.96 \times 4.6) = 18.03$$

$$(2) 2 \times (1.96 \times 4.2) = 16.46$$

$$(3) 2 \times (1.96 \times 0.01) = 0.0392$$

となる。

問 3

(1) 検定統計量は

$$t = \frac{6.9}{4.6} = 1.50$$

p 値は

$$2 \times P(t_{46} \leq -|1.50|) = 0.14 > 0.05$$

となるので，有意水準 5%で帰無仮説は棄却されない。

(2) 検定統計量は

$$t = \frac{12.8}{4.2} = 3.05$$

p 値は

$$2 \times P(t_{28} \leq -|3.05|) = 0.005 < 0.05$$

となるので，有意水準 5%で帰無仮説は棄却される。

(3) 検定統計量は

$$t = \frac{0.03}{0.01} = 3.00$$

p 値は

$$2 \times P(t_{118} \leq -|3.00|) = 0.003 < 0.05$$

となるので，有意水準 5%で帰無仮説は棄却される。

問 4

(1) $s = 2.904$

(2) $[-0.49, 1.19]$

(3) t 検定統計量は 1.32。棄却域は

$$t \leq -3.18 \quad \text{および} \quad t \geq 3.18$$

なので、帰無仮説は棄却されない。また p 値は 0.278 なので、帰無仮説は棄却されない。

問 5

(1) 傾き係数の信頼区間は

$$547.0 \pm 2.06 \times 118.1 = [303.77, 790.23]$$

打率の上昇分は 0.1 だから $[30.377, 79.023]$ (百万円) となる。

(2) 打率の低下分は 0.05 だから $[-39.512, -15.188]$ (百万円) となる。

(3) t 検定統計量は、

$$t = \frac{547.0}{118.1} = 4.63$$

p 値は $2 \times P(t_{25} \leq -|4.63|) = 0.000 < 0.05$ となるので、打率の上昇が翌年の年棒に影響しないという帰無仮説は 5% 有意水準で棄却される。

(4) 切片 β_0 につき

$$H_0: \beta_0 = 0$$

$$H_1: \beta_0 \neq 0$$

を有意水準 5% で行う。 t 検定統計量は

$$t = \frac{-31.7}{27.0} = -1.17$$

p 値は $2 \times P(t_{25} \geq |-1.17|) = 0.251 > 0.05$ となるので、打率がゼロの打者は翌年の年棒はゼロであるという帰無仮説は有意水準 5% で棄却されない。

問 6

$\hat{\beta}_1$ の確率的表現を用いると

$$V(\hat{\beta}_1) = E \left[\left(\frac{\sum_{i=1}^n (X_i - \bar{X}) u_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \right)^2 \right] = \frac{E[(\sum_{i=1}^n (X_i - \bar{X}) u_i)^2]}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2}$$

ここで分子は

$$\begin{aligned} \sum_{i=1}^n (X_i - \bar{X})^2 E(u_i^2) + \sum_{i=1}^n \sum_{j=1(j \neq i)}^n (X_i - \bar{X})(X_j - \bar{X}) E(u_i u_j) \\ = \sum_{i=1}^n (X_i - \bar{X})^2 \sigma^2 \end{aligned}$$

となることから結果が得られる。

また

$$\begin{aligned} \hat{\beta}_0 - \beta_0 &= \bar{Y} - \hat{\beta}_1 \bar{X} - \beta_0 \\ &= \beta_0 + \beta_1 \bar{X} + \bar{u} - \hat{\beta}_1 \bar{X} - \beta_0 \\ &= \bar{u} - \bar{X}(\hat{\beta}_1 - \beta_1) \end{aligned}$$

よって

$$V(\hat{\beta}_1) = E[(\hat{\beta}_0 - \beta_0)^2] = E(\bar{u}^2) + \bar{X}^2 E[(\hat{\beta}_1 - \beta_1)^2] + 2\bar{X} E[(\hat{\beta}_1 - \beta_1)\bar{u}]$$

右辺の第1項は標本平均の分散だから $\frac{\sigma^2}{n}$ 、第2項は $\bar{X}^2 V(\hat{\beta}_1) = \frac{\sigma^2 \bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$ 、第3項は以下の理由からゼロになる。

$$\begin{aligned} E[(\hat{\beta}_1 - \beta_1)\bar{u}] &= E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})u_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \times \frac{1}{n} \sum_{j=1}^n u_j\right] \\ &= \frac{1}{n \sum_{i=1}^n (X_i - \bar{X})^2} E\left[\sum_{i=1}^n \sum_{j=1}^n (X_i - \bar{X})u_i u_j\right] \\ &= \frac{\sigma^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \left[\sum_{i=1}^n (X_i - \bar{X})\right] \end{aligned}$$

ところが $\sum_{i=1}^n (X_i - \bar{X}) = n\bar{X} - n\bar{X} = 0$ 。

よって、

$$V(\hat{\beta}_1) = \frac{\sigma^2}{n} + \frac{\sigma^2 \bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

をまとめると

$$V(\hat{\beta}_1) = \sigma^2 \left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2 + n\bar{X}^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \right]$$

となり、括弧内の分子は

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

となるので結果が得られる。

第 15 講

問 1

- (1) X_1 の値が1単位増加し、 X_2 の値が変わらなかった時の Y の増加分の期待値の推定値。
- (2) このモデルでは X_2 をコントロールしていないため。

問 2

- (1) 平均気温が一度上昇し、湿度が変わらなければ、貸し自転車の利用者数は平均的に172.2人増加すると推定される。また、平均湿度が1%上昇し、平均気温が変わらなければ、貸し自転車の利用者数は平均的に9.1人減少すると推定される。
- (2) 平均気温が0度で平均湿度が0%の日の貸し自転車の利用者数の期待値。
- (3) $\hat{Y}_i = 453.0 + 172.2 \times 20 - 9.1 \times 10 = 3806$ (人)
- (4) 定数項は、平均気温が0度の日の貸し自転車の利用者数の期待値と解釈される。また、説明変数の係数推定値は、平均気温が1度上昇すると貸し自転車の利用者数は平均的に102.2人増加すると解釈される。いずれも、湿度をコントロールしていないので、平均気温に伴い平均湿度が変化し、平均湿度が貸し自転車の利用者数に与える効果も含んでいる。

問 3

(1) 回帰変動 $= 15.51 - 8.85 = 6.66$

(2) 決定係数 $= \frac{6.66}{15.51} = 0.4294$

(3) $s^2 = \frac{8.85}{15-3} = 0.7375$

(4) 自由度修正済決定係数 $= 1 - \frac{8.85/(15-3)}{15.51/(15-1)} = 0.3343$

問 4

(1) 95%信頼区間は以下のようになる

$$\beta_0: 13.2 \pm t_{14,0.025} \times 4.3 = [3.98, 22.42]$$

$$\beta_1: 4.2 \pm t_{14,0.025} \times 2.2 = [-0.52, 8.92]$$

$$\beta_2: -0.8 \pm t_{14,0.025} \times 0.2 = [-1.23, -0.37]$$

$$\beta_3: 3.8 \pm t_{14,0.025} \times 0.8 = [2.08, 5.52]$$

(2) t 検定統計量はそれぞれ

$$\beta_0: t = \frac{13.2}{4.3} = 3.07$$

$$\beta_1: t = \frac{4.2}{2.2} = 1.91$$

$$\beta_2: t = \frac{-0.8}{0.2} = -4.0$$

$$\beta_3: t = \frac{3.8}{0.8} = 4.75$$

両側検定の p 値は

$$\beta_0: 2 \times P(t_{14} \geq 3.07) = 0.008$$

$$\beta_1: 2 \times P(t_{14} \geq 1.91) = 0.077$$

$$\beta_2: 2 \times P(t_{14} \geq 4.0) = 0.0013$$

$$\beta_3: 2 \times P(t_{14} \geq 4.75) = 0.00031$$

より、個別係数が0である帰無仮説は β_0 , β_2 , β_3 については棄却され、 β_1 については棄却される。

問 5

(1) t 検定統計量は

$$t = \frac{-0.2671}{0.1023} = -2.61$$

となるので p 値は

$$2 \times P(t_{96} \geq 2.61) = 0.0105$$

となり、有意水準 5%で帰無仮説 $\beta_1 = 0$ は棄却され、酸濃度はワインの品質に影響を与える
と結論される。

(2) t 検定統計量は

$$t = \frac{0.3629}{0.0750} = 4.84$$

となるので p 値は

$$2 \times P(t_{96} \geq 4.84) = 0.000$$

となり，有意水準 5%で帰無仮説 $\beta_3 = 0$ は棄却され，アルコール濃度はワインの品質に影響を与えると結論される。

(3) F 検定統計量は

$$F = \frac{(67.64 - 47.18)/3}{47.18/(100 - 4)} = 13.877$$

となるので p 値は

$$P(F_{3,96} \geq 13.877) = 0.000$$

となり，有意水準 5%で帰無仮説 $\beta_1 = \beta_2 = \beta_3 = 0$ は棄却される。